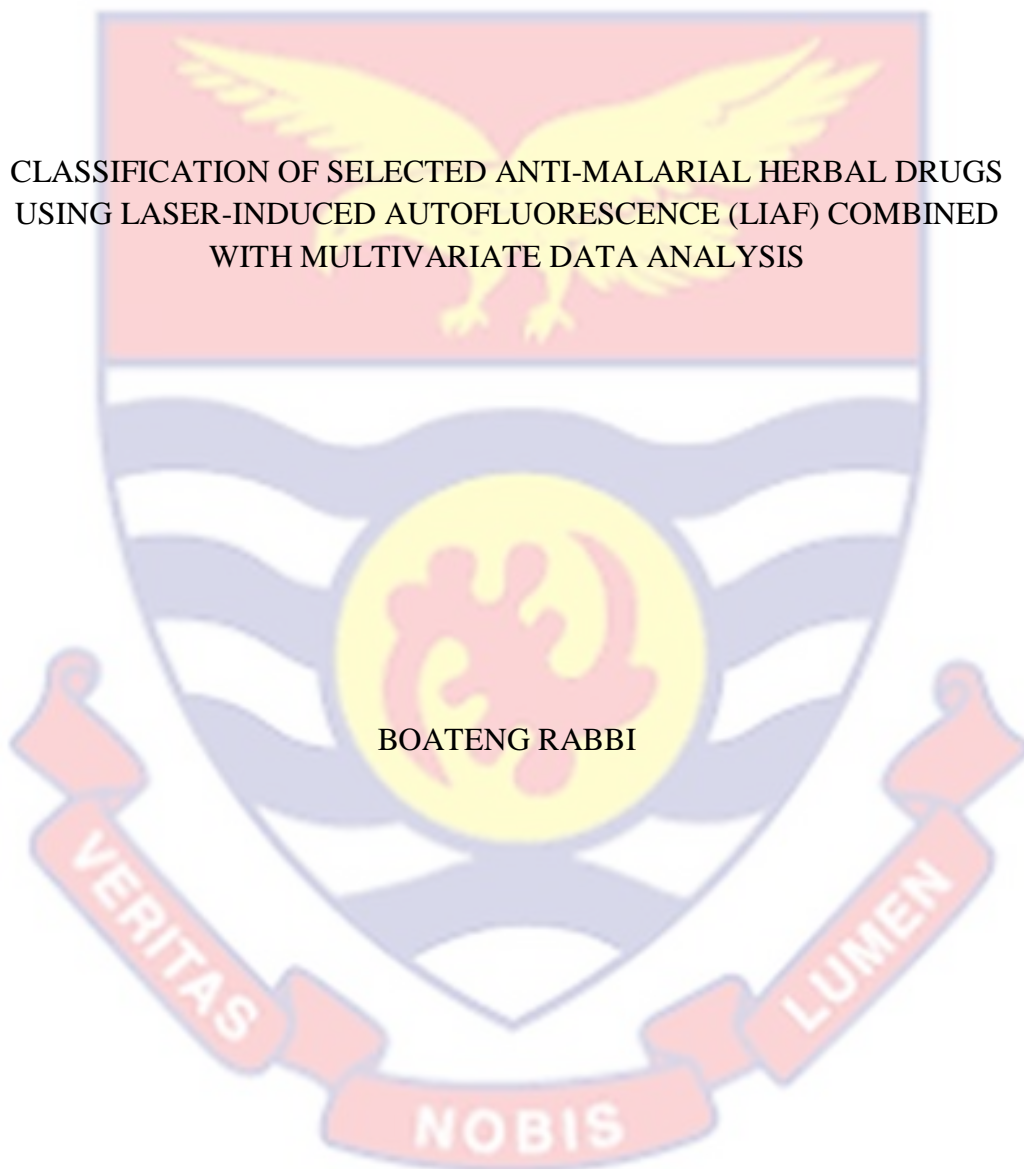


UNIVERSITY OF CAPE COAST



CLASSIFICATION OF SELECTED ANTI-MALARIAL HERBAL DRUGS  
USING LASER-INDUCED AUTOFLUORESCENCE (LIAF) COMBINED  
WITH MULTIVARIATE DATA ANALYSIS

BOATENG RABBI

2021

UNIVERSITY OF CAPE COAST

CLASSIFICATION OF SELECTED ANTI-MALARIAL HERBAL DRUGS  
USING LASER-INDUCED AUTOFLUORESCENCE (LIAF) COMBINED  
WITH MULTIVARIATE DATA ANALYSIS

BY

BOATENG RABBI

This thesis submitted to the Department of Physics of the School of Physical Sciences,  
College of Agriculture and Natural Sciences, University of Cape Coast, in partial  
fulfilment of the requirement for the award of Master of Philosophy degree in  
Physics

JULY 2021

## DECLARATION

### Candidate's Declaration

I hereby declare that this thesis is the result of my own original work and that no part of it has been presented for another degree in this University or elsewhere.

Candidate's Signature: ..... Date: .....

Name: Boateng Rabbi

### Supervisors' Declaration

We hereby declare that the preparation and presentation of the thesis were supervised in accordance with the guidelines on supervision of thesis laid down by the University of Cape Coast.

Principal Supervisor's Signature: ..... Date: .....

Name: Prof. Moses Jojo Eghan

Co-Supervisor's Signature: ..... Date: .....

Name: Dr. Jerry Opoku Ansah

## ABSTRACT

Laser-induced autofluorescence (LIAF) was used to characterize and classify some selected commercial anti-malarial herbal drugs (AMHDs) to avoid sample destruction. Results from deconvoluted peaks of the LIAF spectra of AMHDs showed secondary metabolites belonging to derivatives of alkaloids and classes of phenolic compounds (tannins, lignin, flavonoid, and coumarins) present in all the AMHDs samples. Analyses with unsupervised methods (Principal component analysis (PCA), K-means clustering and Hierarchical Clustering Analysis (HCA)) classified the AMHDs that were similar in physicochemical properties. Based on two PCs, supervised pattern recognition methods such as Supervised Vector Machine (SVM), K-Nearest Neighbor (KNN), Quadratic Discriminant Analysis (QDA), and Linear Discriminant Analysis (LDA) models generated were used to identify and classify unknown AMHDs with 100.00, 100.00, 99.52 and 99.04 % accuracy respectively. LIAF technique in combination with multivariate analysis may offer non-destructive characterization and classification of AMHDs.

## KEY WORDS

Active pharmaceutical ingredient

Anti-Malarial Herbal Medicines

Laser Induced Autofluorescence

Multivariate data analysis

Optical Spectroscopic Techniques

Secondary metabolites

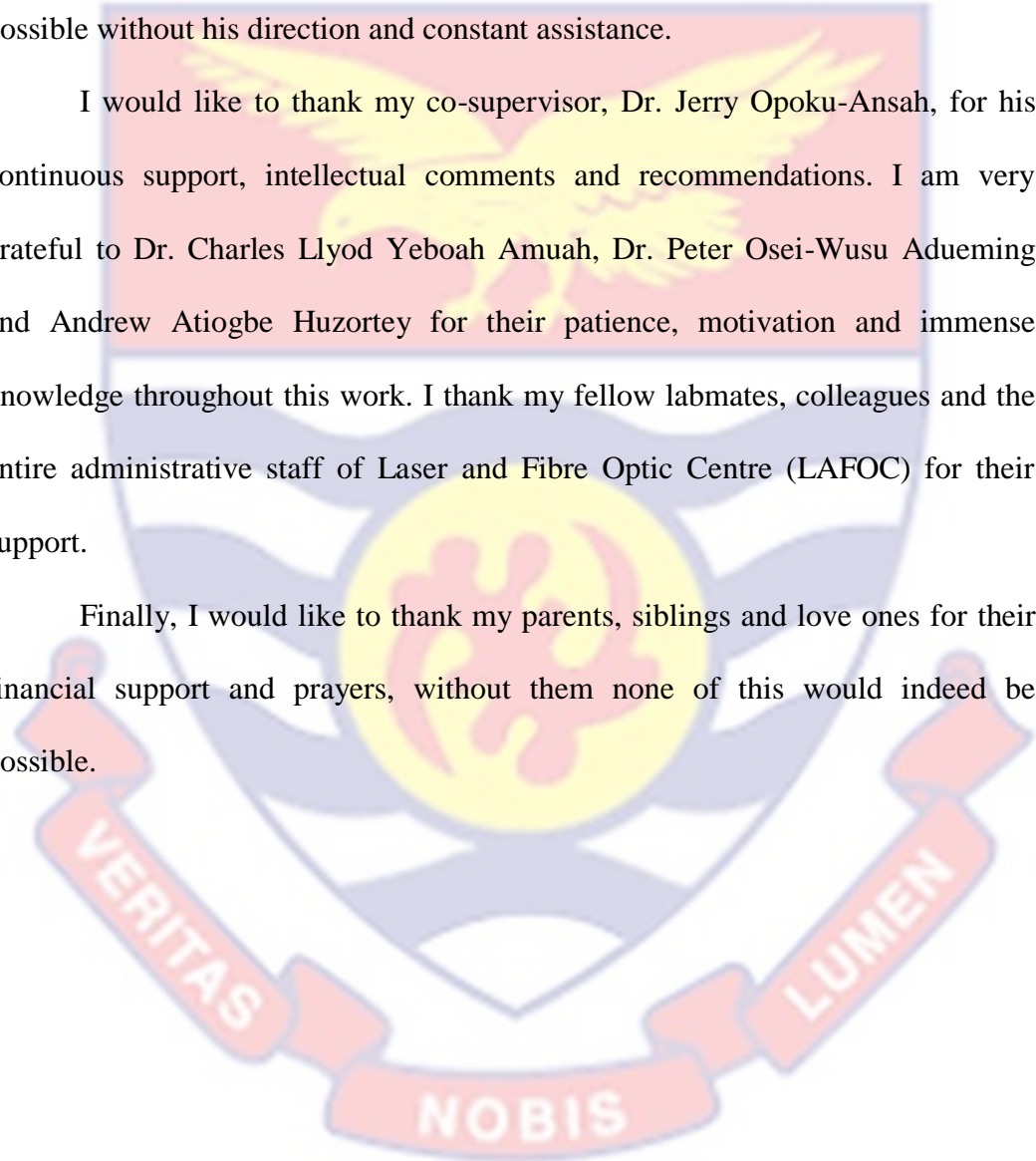


## ACKNOWLEDGEMENTS

I would like to express my gratitude to my principal supervisor, Professor Moses Jojo Eghan, for his support and guidance during my research. In regards to instructing and mentoring me during my study. This thesis would not have been possible without his direction and constant assistance.

I would like to thank my co-supervisor, Dr. Jerry Opoku-Ansah, for his continuous support, intellectual comments and recommendations. I am very grateful to Dr. Charles Llyod Yeboah Amuah, Dr. Peter Osei-Wusu Adueming and Andrew Atiogbe Huzortey for their patience, motivation and immense knowledge throughout this work. I thank my fellow labmates, colleagues and the entire administrative staff of Laser and Fibre Optic Centre (LAFOC) for their support.

Finally, I would like to thank my parents, siblings and love ones for their financial support and prayers, without them none of this would indeed be possible.



## DEDICATION

To my father Rev. Boateng Atta Simon.



## TABLE OF CONTENTS

	<b>Page</b>
DECLARATION	ii
ABSTRACT	iii
KEY WORDS	iv
ACKNOWLEDGEMENTS	v
DEDICATION	vi
LIST OF TABLES	x
LIST OF FIGURES	xi
LIST OF ABBREVIATIONS	xiv
CHAPTER ONE: INTRODUCTION	1
Background to the Study	1
Scope of Study	9
Statement of the Problem	9
Research Objectives	10
Significance of the Study	11
Delimitation	11
Limitation	11
Organization of Study	11
Chapter Summary	11
CHAPTER TWO: LITERATURE REVIEW	13
Introduction	13
Light Interaction with Matter	13
Fluorescence Spectroscopy	16



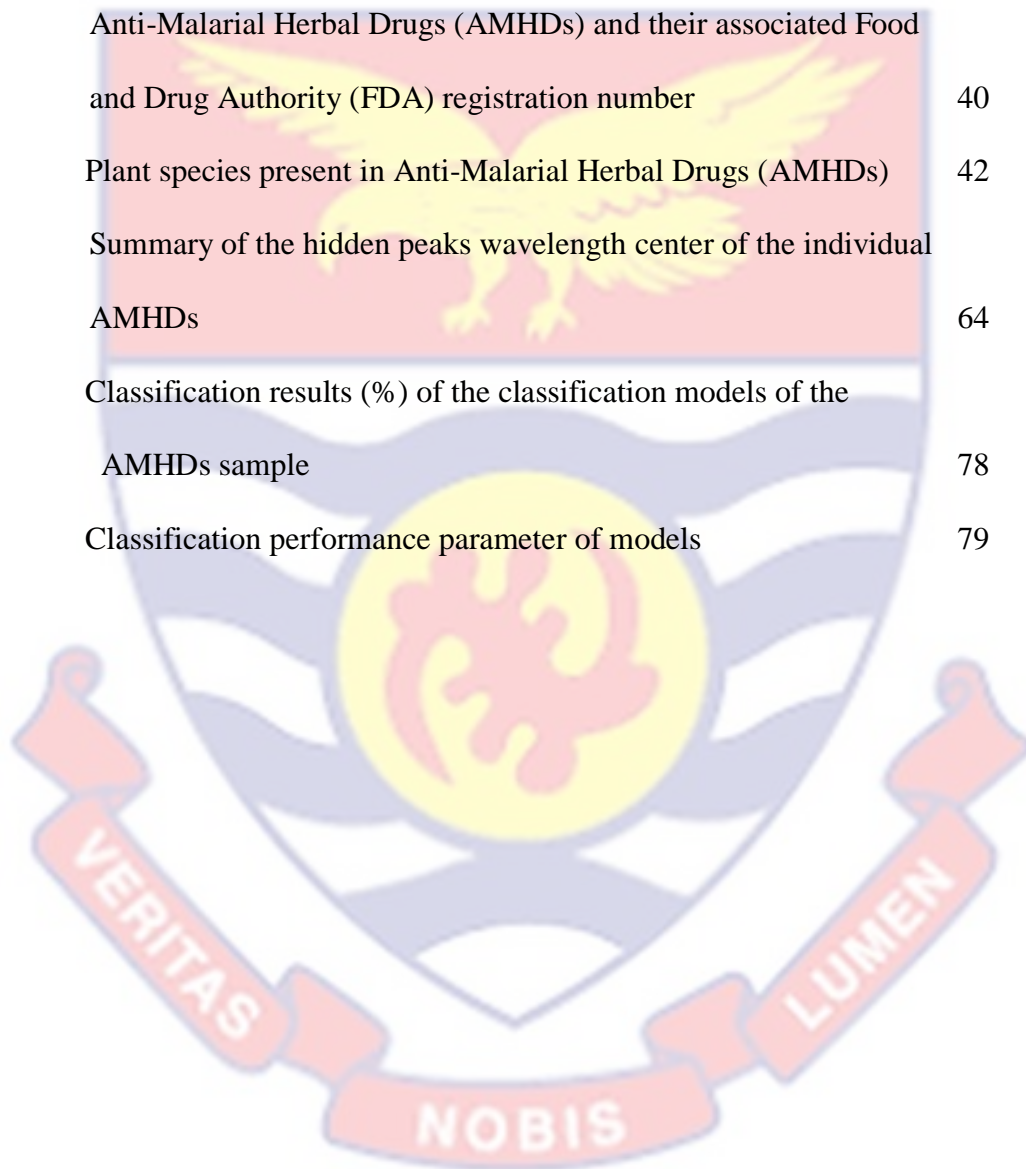
Laser-Induced Autofluorescence (LIAF)	19
Secondary Metabolites	21
Multivariate Analysis	26
Chapter Summary	38
CHAPTER THREE: RESEARCH METHODS	40
Introduction	39
Anti-Malarial Herbal Drugs (AMHDs) Samples Collection	39
Laser-Induced Autofluorescence Experimental Set-up	45
Laser-Induced Autofluorescence Spectra Measurement	47
Pre-processing of Laser-Induced Autofluorescence Spectra Data	48
Chapter Summary	50
CHAPTER FOUR: RESULTS AND DISCUSSION	51
Introduction	51
LIAF Spectra of AMHDs	51
Chemometrics of Anti-malarial Herbal Drugs (AMHDs)	65
Chapter Summary	81
CHAPTER FIVE:	
SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS	85
Summary	82
Conclusions	82
Recommendations	85
REFERENCES	87
APPENDICES	118

APPENDIX A: LIAF SPECTRA OF AMHDS SAMPLE	117
APPENDIX B: DECONVOLUTED PEAKS OF AMHDS SAMPLE	119



## LIST OF TABLES

		Page
1	The confusion matrix of a classification model arranged in two classes as a $2 \times 2$ matrix	36
2	Anti-Malarial Herbal Drugs (AMHDs) and their associated Food and Drug Authority (FDA) registration number	40
3	Plant species present in Anti-Malarial Herbal Drugs (AMHDs)	42
4	Summary of the hidden peaks wavelength center of the individual AMHDs	64
5	Classification results (%) of the classification models of the AMHDs sample	78
6	Classification performance parameter of models	79



## LIST OF FIGURES

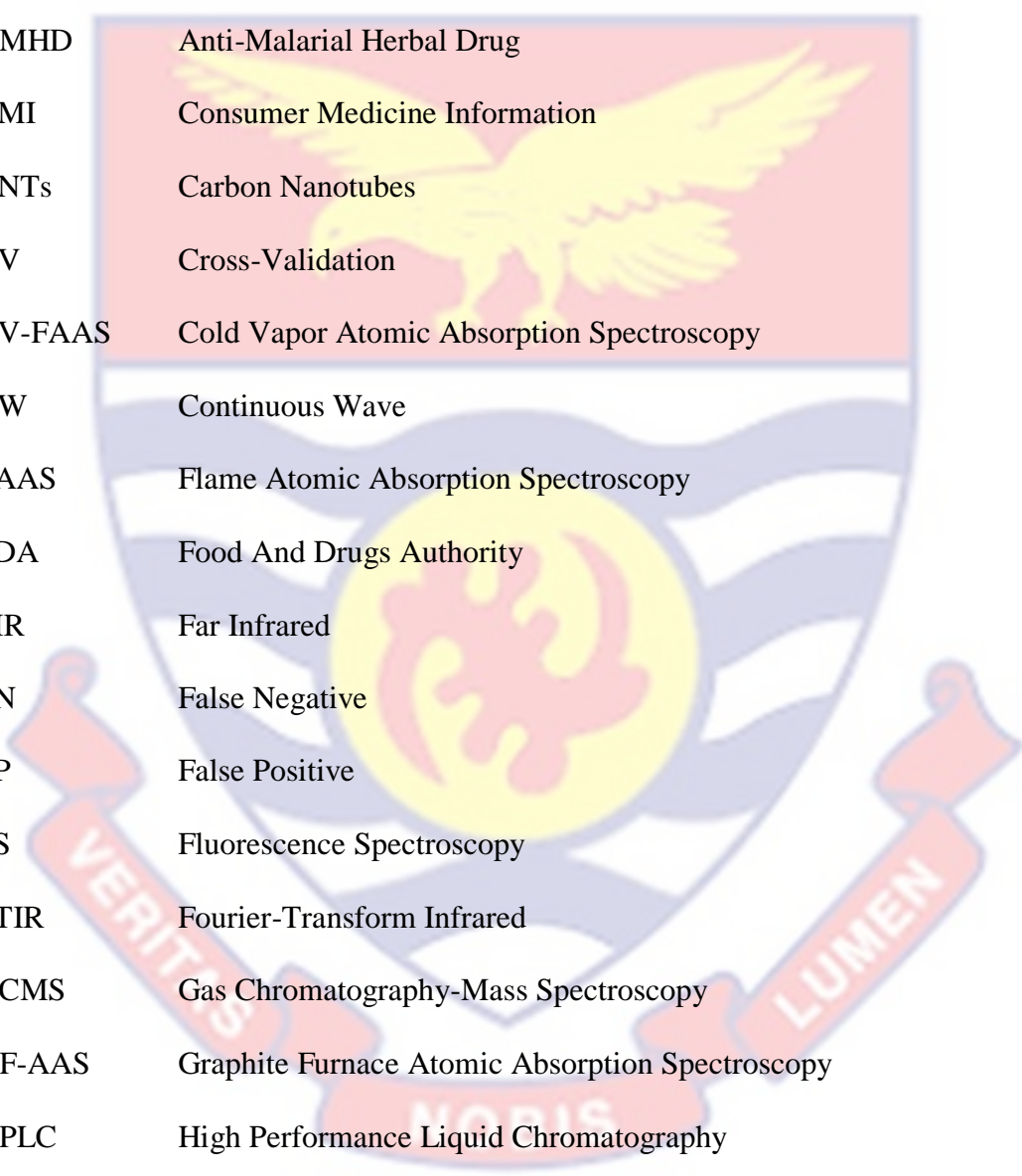
	Page
1	13
2	17
3	46
4	48
5	50
6	53
7	54
8	55
9	57

10	Deconvoluted LIAF spectra of AMHD_K.	60
11	Deconvoluted LIAF spectra of AMHD_A, AMHD_B, AMHD_C and AMHD_G	62
12	Scree plot showing percentage variance of the first ten (10) principal components of AMHDs	66
13	Loadings plot the first two PCs obtained from PCA of the LIAF spectra of AMHDs samples	67
14	Score plot showing clustering of AMHDs	69
15	A PCA biplot showing secondary metabolites and principal component scores of AMHDs	71
16	A visual curve display of the performance of K-cluster and the corresponding WCSS values	73
17	K-means clustering indicating three clusters regions (1, 2 and 3) derived from the fluorescence spectra of the AMHDs	74
18	Dendrogram indicating three clusters (1, 2 and 3) derived from the fluorescence spectra of the AMHDs	75
19	Flow Chart showing the sequence of spectra processing used for classification	76
20	Confusion Matrix for the Prediction set of the AMHDs for the KNN model	80
21	Confusion Matrix for the Prediction set of the AMHDs for the SVM model	80
22	LIAF spectra of AMHD_D	117

23	LIAF spectra of AMHD_E	117
24	LIAF spectra of AMHD_F, AMHD_H, AMHD_I and AMHD_J.	118
25	Deconvoluted LIAF spectra of AMHD_F, AMHD_H, AMHD_I and AMHD_J	119
26	Deconvoluted LIAF spectra of AMHD_D, and AMHD_E	120



## LIST OF ABBREVIATIONS



AAS	Atomic Absorption Spectroscopy
ACT	Artemisinin-Based Combination Therapy
ATR	Attenuated Total Reflection
AMHD	Anti-Malarial Herbal Drug
CMI	Consumer Medicine Information
CNTs	Carbon Nanotubes
CV	Cross-Validation
CV-FAAS	Cold Vapor Atomic Absorption Spectroscopy
CW	Continuous Wave
FAAS	Flame Atomic Absorption Spectroscopy
FDA	Food And Drugs Authority
FIR	Far Infrared
FN	False Negative
FP	False Positive
FS	Fluorescence Spectroscopy
FTIR	Fourier-Transform Infrared
GCMS	Gas Chromatography-Mass Spectroscopy
GF-AAS	Graphite Furnace Atomic Absorption Spectroscopy
HPLC	High Performance Liquid Chromatography
KNN	K-Nearest Neighbor
LDA	Linear Discriminant Analysis
LIAF	Laser-Induced Autofluorescence

LIF	Laser-Induced Fluorescence
LIP	Laser-Induced Phosphorescence
NIR	Near Infrared
PCA	Principal Component Analysis
PET	Polyethylene Terephthalate
QDA	Quadratic Discriminant Analysis
ROC	Receiver Operating Curves
SMA	Subminiature Version A
SVM	Support Vector Machine
TLC	Thin Layer Chromatography
TN	True Negative
TP	True Positive
TTL	Transistor-Transistor Logic





## CHAPTER ONE

### INTRODUCTION

#### Background to the Study

Over the last decade, applications of optical spectroscopic techniques have increased significantly. It is possible to use different optical phenomena as well as different light properties to get information from samples of interest (Titus et al., 2019; Vanlanduit & Guillaume, 2009). Generally, optical phenomena deals with electromagnetic radiation being in contact with a medium. The optical phenomena may be linear or nonlinear when considering the amplitude of the light as the key factor. The interaction of the electromagnetic radiation with a medium results in the absorption, dispersion, fluorescence, transmission, reflection or scattering of the electromagnetic radiation and it can be responsive to slight changes in the medium's molecular/atomic structure or soundings (Gruber, 1956; Parson, 2015; Sirohi, 2009; Vanlanduit & Guillaume, 2009).

Optical spectroscopic methods propose to solve some limitations by detecting molecules/atoms simultaneously and non-destructively (Strachan et al., 2007; Watanabe et al., 2019). Optical spectroscopic techniques are used in a wide range of applications, from semiconductors to medical sciences, art conservation to defense, physics to pharmaceutical, biology to geology, due to their exceptional sensitivity and speed. Optical spectroscopic techniques are used in laboratories for the examination of medical and industrial samples due to their non-invasiveness. These techniques could also be used in robust measuring systems and routine inspection of solid, liquid, and gaseous samples (Allen et al., 1971; Gruber, 1956;

Mayinger, 2001; Parson, 2015; Titus et al., 2019; Wartewig, 2004; Watanabe et al., 2019).

Fiber optic probes can be used in a variety of spectroscopic measurement, such as detecting fluorescence from a sample or elucidate the nature of a body by monitoring fluorescent traces in tract tracing. Optical spectroscopic techniques have low cost, long life, and resistance to thermal and mechanical disruptions make them ideal for use in harsh environments (Haruna et al., 2019; Lodder, 2002).

The choice of optical spectroscopic technique is based on the sample under study. There are several types of optical spectroscopy techniques namely: Raman spectroscopy, Infrared (Near and Far) spectroscopy, Atomic Absorption spectroscopy, Laser induced breakdown Spectroscopy, Ultraviolet and Visible spectroscopy, and their applications vary depending on the study's objective. Optical spectroscopy, such as Raman spectroscopy, is a good example used in atomic or molecular studies. Raman spectroscopy is a technique that deals with the vibrational, and rotational modes of a medium by using Raman scattering of monochromatic laser light (Edwards, 2005; Strachan et al., 2007). The interaction of the phonons with the laser light causes an energy change, which provides information about the phonon modes in the system. Raman spectroscopy has a number of advantages, including the fact that there is little or no sample preparation and non-destructive, water bands are typically small and easily subtracted making the strength of the bands in a Raman intensity spectrum be dependent on the concentration of the corresponding molecules, allowing for

quantitative study (Edwards, 2005; Huzortey et al., 2017; Strachan et al., 2007; Wartewig, 2004).

Raman spectroscopy could be used in the pharmaceutical industry to provide information on various aspects such as medicinal uniformity, chemical identification, and the physical state of medicine. It has helped to improve new areas in pharmaceutical study (John & George, 2017; Nasdala et al., 2012). In the study of polymers, the information obtained from a Raman spectrum is useful not only for determining the relationship between chemical and morphological structure, but also for determining the layered structure of polymer coatings without need destroying the layers (Koenig, 1971; Seyrek & Decher, 2012; Speight, 2017; Weselucha-birczy et al., 2014). Raman spectroscopy is also very useful in nano-technological applications for separating single walled Carbon nanotubes (CNTs) from multi-walled CNTs, as well as traces in cells (John & George, 2017; Weselucha-birczy et al., 2014).

Another method of optical spectroscopy is Infrared (IR) spectroscopy. Infrared spectroscopy is a method that uses the vibrations of atoms to determine its properties. The most popular method for obtaining an infrared spectrum is to transmit infrared radiation through a sample and determine what fraction of the incident radiation is absorbed at each energy level (Lodder, 2002; Stuart, 2005). The frequency of a vibration of a part of a sample molecule corresponds to the energy at which any peak in an absorption spectrum appears. The information found in the IR spectrum pertains to any element of the molecular structure. This is because molecules have different isotopic distributions, constitutions,

structures, conditions or conformations. The IR spectrum may be used to distinguish two distinct molecules (Schrader, 1996). The introduction of low-cost, high-performance computers, combined with the use of multivariate data techniques, has resulted in an explosion of IR spectroscopy applications in a variety of fields, ranging from medical to pharmaceutical to traditional food analysis (Hussain & Keçili, 2020a; Schrader, 1996; Stuart, 2005; Sun, 2015; Wulandari, Retnaningtyas, Nuri, & Lukman, 2016). Infrared spectroscopy has also advanced significantly as a result of constant advancements in computer technology. The invention of Fourier-transform spectrometers, on the other hand, has resulted in the most important developments in infrared spectroscopy. The principle behind Fourier-transform infrared (FTIR) spectroscopy is unique, in that two beams of light interact to produce an interferogram. An interferometer is used in this type of instrument, which uses the well-known mathematical method of Fourier-transformation. The mathematical method of Fourier-transformation can convert the two domain of space and time. Infrared spectroscopy using Fourier-transform infrared (FTIR) spectroscopy has greatly improved the accuracy of infrared spectra while reducing the amount of time it takes to acquire results (Stuart, 2005). This is important in remote sensing of air pollution from factory and power plant chimneys, as well as vehicular (Vanlanduit & Guillaume, 2009). IR spectroscopy can be Near Infrared (NIR) or Far Infrared (FIR) spectroscopy depending on the electromagnetic radiation techniques.

NIR spectroscopy is a sub branch of the IR spectroscopy that has evolved into the following acquisition modes: transmittance, transreflectance, reflectance,

interactance, and transmittance through scattering medium, which refer to the various geometric arrangements of the detected radiation. Different reflectance methods, such as attenuated total reflection (ATR), specular reflectance, and diffuse reflectance, are now widely practiced in infrared spectroscopy technique (Pasquini, 2003; B. H. Stuart, 2005). Because it is fast and requires no sample preparation, ATR is one of the most popular sampling methods used by infrared spectroscopists. It is a contact sampling technique that takes advantage of the spectral information available from reflection phenomena. Unlike many sample techniques the ATR is able to measure the IR spectra of a medium that is too dense or absorbing to be measured (Larkin, 2018). Another common spectroscopic technique usually used in the studies of atoms or molecules of a medium is the Atomic Absorption Spectroscopy (AAS).

AAS is a quantitative analytical technique that measures the optical radiation absorbed by an element to determine and analyze its presence in a complex sample. (Biswas, Karn, Paresh, & Balasubramanian, 2017; Hussain & Keçili, 2020). AAS is a highly sensitive method of elemental analysis that allows for the detection of metals in a variety of samples down to the picogram stage, as well as determining the concentration of a single metal constituent in a sample. (Biswas, et al., 2017; Butcher, 2005). The technique is often used to assess the toxicology of complex samples such as food, biological, industrial and environmental samples. In the medical sector, clinical samples are often analyzed to determine the presence of metals and metalloids in fluids and biological tissues for therapeutic purposes (Calatayud & Icardo, 2004). It has a wide range of

applications and is relatively interference-free when measuring a variety of elements (Hanrahan et al., 2004; Hill & Fisher, 2017). AAS has been used to evaluate heavy metals in medicinal plants. (Akram et al., 2015).

Depending on the measurement range and concentration of the element of interest, there are other optical spectroscopic techniques such as flame atomic absorption spectroscopy (FAAS), cold vapor atomic absorption spectroscopy (CV-AAS), hydride-generating atomic absorption spectroscopy (HG-AAS), and graphite furnace atomic absorption spectroscopy (GF-AAS) systems (Michalke & Nischwitz, 2013). FAAS has been used in biological fluid studies as well as in biomedical applications to analyze solutions derived from biological tissues and to assess metal cation concentrations in filtered water sources (Campbell & Ingram, 2014; Taylor, 1999). Though some systems have low detection limits, their sensitivity for derivatized species has been improved (Michalke & Nischwitz, 2013).

Another sensitive spectroscopic technique is the Fluorescence Spectroscopy (FS). FS became widely used in several fields of science after the groundbreaking work of Weber, Stokes, Vavilov, Perrin and Jaboski (Wolfbeis, 2004). FS is an absorption spectroscopy complementary technique that analyzes the fluorescence from a sample. FS instruments use short wavelength light in the visible or ultraviolet wavelength ranges to illuminate samples, resulting in a longer wavelength (lower energy) of photon (fluorescence) (Braeuer, 2015; Da Silva & Utkin, 2018; Eyring & Martin, 2013). It is a responsive and adaptable analytical tool that can be used for both quantitative and qualitative analysis.

Fluorescence is used in basic analytical assays in environmental science and clinical chemistry, cell recognition and sorting in flow cytometry, and single-cell imaging in medicine, for example. FS is a generally recognized and effective technique that's used in a wide range of applications, including DNA sequencing, genetic analysis, forensics, medical diagnostics and biotechnology, environmental and manufacturing sites (Chirayil et al., 2017). Since water is not an interferent in FS measurement, it makes the technology extremely sensitive to fluorescent target molecules. As a result, its applications in dilute aqueous media are feasible (Kessler & Kessler, 2020). Fluorescence spectroscopy has been used to classify and measure difficult-to-quantify products such as morphine extracted from body fluids (Eyring & Martin, 2013; Tranter, 1999). The techniques of Laser-Induced Phosphorescence (LIP) and Laser-Induced Fluorescence (LIF) are both fluorescence processes that occur after light has been absorbed.

The LIF and LIP phenomena are based on the ability of a visible/ultraviolet laser beam to selectively excite fluorescent atom/molecules. In order to detect LIF/LIP, a specific wavelength of light source is used. When an excited molecule returns to its ground state, a photon with a slightly longer wavelength is released (lower energy) (Karlinsey, 2012; Kumar et al., 2018). Due to its higher sensitivity, LIF is favoured over IR, Raman and UV spectroscopy for detecting compounds in low concentrations (Brinck et al., 2003; Karlinsey, 2012). In a thirty-one component multivitamin powder mix, LIF was used to analyze the concentrations of five different constituents in real-time. LIF was also used to monitor the mixing of pharmaceutical components in a V-blender with the aid of

a non-contact probe (Nadeem & Heindel, 2018). LIF's nondestructive nature has been used to identify various anti-malarial plants from different locations (Amuah, Eghan, Anderson, Adueming, & Opoku-Ansah, 2017). When induced by an UV/visible laser beam, the fluorescence emission from innate molecule fluorophores is referred to as Laser-Induced Autofluorescence (LIAF).

LIAF is a non-invasive procedure that does not require the use of special histochemical dyes (García-plazaola et al., 2015; Rost, 2016). The autofluorescence of a molecule can be described by a number of factors, the most important of which are the emission spectrum, quantum yield, lifetime, and polarization, all of which can be measured in a steady-state or time-resolved manner. They provide details about the fluorescing molecule's photophysical properties as well as the chemical and physical nature of its microenvironment (Rost, 2016; Wolfbeis, 2004). Several new LIAF-based applications have been developed in recent years, promoting autofluorescence spectroscopy from a strictly scientific to a more routine tool. Despite the fact that the number of regular fluorescence applications is rapidly increasing, the concepts remain the same. In disease diagnosis, LIAF has been used for the identification of infected as well as uninfected RBCs and the estimation of parasite density of *Plasmodium falciparum* (Opoku-Ansah, Eghan, Anderson, & Boampong, 2014; Opoku-Ansah, Eghan, Anderson, Boampong, & Buah-Bassuah, 2016). In geology, LIAF is characterized by the presence of small quantities of activator compounds, which are often referred to as "impurities" and are widely used in coal research (Jaffrennou et al., 2007; Rost, 2016). In the studies of secretory cells, many



researchers have employed the LIAF technique. Since secretory cells of leaves, stems, flowers, and roots fluoresce in all colors of the visible spectrum. Secretory cells in plants contain a variety of fluorescent products, primary and secondary metabolites, as has recently been discovered (Roshchina, 2018; Roshchina et al., 2017). Secondary substances such as phenols, flavins, quinones, alkaloids, polyacetylenes, coumarins, terpenoids, and others are found in secretory cells and can fluorescence when exposed to ultraviolet radiation (Roshchina, 2018). These techniques has been employed in many fields of study (Biswas, Karn, Balasubramanian, & Kale, 2017; Chen, Zheng, & Liu, 2011; Giridhar, Manepalli, & Apparao, 2017; Lamirel, 2014; Liu & Brezinski, 2014; McCreery, 2005; Weselucha-birczy et al., 2014).

### **Scope of Study**

This work seeks to apply LIAF to characterize, and classify eleven (11) commercial anti-malarial herbal drugs purchased from licensed herbal shops within Cape Coast Metropolis.

### **Statement of the Problem**

The wide use of AMHDs has drawn the attention of many researchers to its study. The use of plant-based multi-herbal preparations for malaria treatment is very popular due to the high cost and restricted availability of artemisinin-based combination therapy (ACT) (Adusei-Mensah et al., 2020; Frank Adusei-Mensah et al., 2019; WHO, 2017a, 2018). Plant-based medicines are used more often because of their year-round availability, low cost, and ease of access as opposed to pharmaceutical medicines (Febir et al., 2016; WHO, 2017b).

Researchers studied the phytochemical constituents of AMHDs using methods such as gas chromatography-mass spectroscopy (GCMS), high-performance liquid chromatography (HPLC), thin-layer chromatography (TLC) and others. Wilmot in 2017 assessed the chemo-suppressive activity of AMHDs against *P. berghei* in vivo (Wilmot et al., 2017). Köhler in 2002 also assessed the physiochemical properties of AMHDs using HPLC and TLC (Köhler et al., 2002). These processes are expensive, destructive and may require enough time which makes them not feasible enough.

In the assessment of phytochemical analysis of medicinal plants, LIAF has been used widely (Donaldson, 2020; Roshchina et al., 2016). LIAF has unique characteristics such as rapidity, non-destructiveness, non-invasiveness, less expensive, and sensitivity. LIAF has been proven to be an effective and objective tool for characterizing, discriminating, and identifying medicinal plants (Amuah et al., 2017; Donaldson, 2018; García-plazaola et al., 2015; Roshchina, 2012). Despite LIAF's numerous advantages, commercial AMHDs have not been explored with this technique.

### **Research Objectives**

This work aims to determine the spectral fingerprint of some selected AMHDs samples using LIAF. This is to infer secondary metabolites from literature to the AMHD spectral fingerprint. This work also seeks to combine the LIAF spectra fingerprint and multivariate analysis Principal Component Analysis (PCA), K-Means Clustering, Hierarchical Clustering Analysis (HCA), Linear Discriminant Analysis (LDA), Quadratic Linear Analysis (QDA), K-Nearest

Neighbor (KNN) and Support Vector Machine (SVM)) to characterize and classify the AMHDs of common physiochemical properties.

### **Significance of the Study**

The work would be important because the spectral fingerprints of AMHDs would provide faster and cheaper alternative for detecting any AMHDs. It is a non-destructive method and can be an objective confirmation of any AMHDs.

### **Organization of Study**

This thesis consists of five main chapters. Chapter one gives an overview of optical spectroscopic techniques and discusses fluorescence spectroscopy applications to medicinal plants. The scope of the work and the organization of the thesis are included within this chapter. The second chapter reviews the literature on the concepts of light-matter interaction, principles of optical spectroscopy, as well as phytochemical constituents present in medicinal plants. Chapter three looks at the experimental methods and procedure for the research work. The LIAF spectral fingerprints results, analysis, and discussions are presented in Chapter four. In the fifth chapter, conclusions are drawn and relevant recommendations are given to assist in further research.

### **Chapter Summary**

This chapter has discussed the background to various optical spectroscopic techniques and their associated applications, and also discussed fluorescence spectroscopy applications on medicinal plants (Anti-malarial herbal medicines).

The scope of the work and the organization of the thesis were included in this chapter.



## CHAPTER TWO

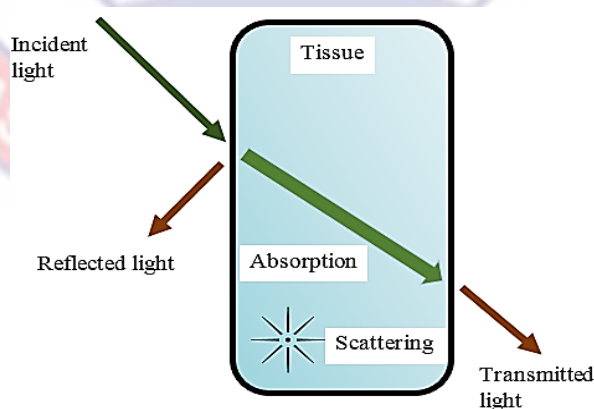
### LITERATURE REVIEW

#### Introduction

This chapter reviews the literature on light-matter interaction with emphasis on optical spectroscopic techniques used for the characterization and classification of samples. The secondary metabolites found in medicinal plants are outlined and finally the theories involved in multivariate data analysis discussed.

#### Light Interaction with Matter

Light interacts with matter in a variety of ways. Fundamentally, the wavelengths, and intensity of light determines the result of their interaction with matter. The optical properties of matter, such as the depth, refractive index, surface contour, type of targeted medium and temperature of the medium, influence light interaction with matter (Bordin-Aykroyd et al., 2019). Due to the different optical properties of matter, light can be strongly absorbed, transmitted, reflected towards the source, refracted, or scattered by another as showed in figure 1 (Bordin-Aykroyd et al., 2019; Keiser, 2016; Weiner, 2003).



*Figure 1:* Some probable light-tissue interactions (reflection, transmission, absorption, and scattering)

When a light beam strikes a surface separated by two dielectric materials, a portion of the light beam is reflected by the first media. Once it reaches the second material, the remaining light is refracted. The type of reflecting surface has a significant impact on the amount of light reflected by a material in a certain direction. Specular reflection is the reflection of light off of smooth surfaces. Diffuse reflection occurs as light is reflected off of microscopically rough surfaces (Keiser, 2016). One aspect of the law of reflection is shown in equation 1.

$$\theta_i = \theta_r \quad (1)$$

where  $\theta_i$  is the angle of incidence and  $\theta_r$  is the angle of reflection. Thus, the angle of incidence is equal to the angle of reflection.

Following the entry of light into a medium, the molecules of the medium and interactions between the electric fields of the incoming photons allow some of the light energy to be absorbed. Absorption therefore takes place when a portion of the light energy is converted into thermal motion or an increase in molecular motions in the material. That is, photons can give up their energy in order to shift a molecule's vibrational state to a higher vibrational state. The electronic structure of the atoms and molecules in the material, the wavelength of the light, the thickness of the absorbing layer, and the temperature all influence the degree of absorption.

The entry of light may also scatter. Light scattering is described as the deflection of a light ray from a straight path due to irregularities in the propagation medium, particles, or the interface between two media with different

refractive indices than the surrounding material (Gomes & Rocha-Santos, 2019). The wavelength of light, as well as the refractive indices, structures, and sizes of the material components liable for the scattering, all influence light scattering in tissue. Scattering is a mechanism that can be elastic or inelastic. In elastic scattering, the energy of incident and scattering photons is the same. The majority of light-matter interactions are defined by this form of scattering. The transfer of energy between a scattering molecule and the photon is known as inelastic scattering. During inelastic scattering, energy can be transferred from the photon to the molecule or the photon can absorb energy from the molecule (Keiser, 2016).

Transmittance ( $T$ ) process occurs when photons are passed on to neighboring atoms through the bulk of the material and re-emitted on the opposite side of the medium. The light propagation through a medium is a repeated process of scattering and rescattering (Keiser, 2016). The transmittance is defined by the expression in equation 2.

$$T = \frac{I}{I_o} \quad (2)$$

where  $I_o$  is the incident light intensity,  $I$  the transmitted light intensity upon the detecting area of the detector. Various techniques based on the principle of light-matter interaction have been developed. Raman spectroscopy, infrared spectroscopy, atomic absorption spectroscopy, and fluorescence spectroscopy are some of the techniques that employ this phenomena.

## Fluorescence Spectroscopy

Luminescence is the emission of ultraviolet, visible, or infrared photons by an electronically excited molecule. Luminescence in activity includes fluorescence and phosphorescence. The mode of excitation is the absorption of a photon, which causes the absorbing species to become electronically excited. The emission of photons that occurs as a result of de-excitation is known as photoluminescence which can be in the form of phosphorescence, fluorescence, and delayed fluorescence, and it is one of the physical phenomena that can occur when light interacts with matter (Valeur, 2001; Virk, 2014).

The spontaneous emission of light by the transitions of light from excited singlet states to different vibrational levels of the electronic ground state is known as fluorescence. Interactions in the excited state with other molecules, such as energy transfer, proton transfer, electron transfer, can be termed de-excitation (Braeuer, 2015; Valeur, 2001; Zhang, 2010). These de-excitation pathways can compete with fluorescence emission if they occur on a time scale comparable to the lifetime of the molecules in the excited state. Several parameters can be used to describe a molecule's fluorescence. The intensity of fluorescence at a specific wavelength,  $F(\lambda)$ , the emission spectrum (emission intensity dependence on emission wavelength), quantum yield ( $\Phi$ ), lifetime ( $\tau$ ), and polarization ( $P$ ) are significant parameters for the characterization of fluorescence of a molecule (Lakowicz, 1999; Sauer, Hofkens, & Enderlein, 2013; Wolfbeis, 2004). These variables can be tracked in both steady-state and time-resolved modes. They



provide details about the fluorescing molecule's photophysical properties as well as the physical and chemical nature of its microenvironment.

Fluorophores are the molecules that cause fluorescence. The fluorophore, as well as the fluorophore's chemical environment, determine the wavelength of the emitted energy. The fluorophore normally resides in one of the several vibrational levels of an excited singlet state after being excited with light of appropriate wavelength (on a time scale of  $\leq 10^{-15}$  seconds), as shown in the Jablonski diagram in figure 2 (Lakowicz, 1999; Sauer et al., 2013; Valeur, 2001; Wolfbeis, 2004).

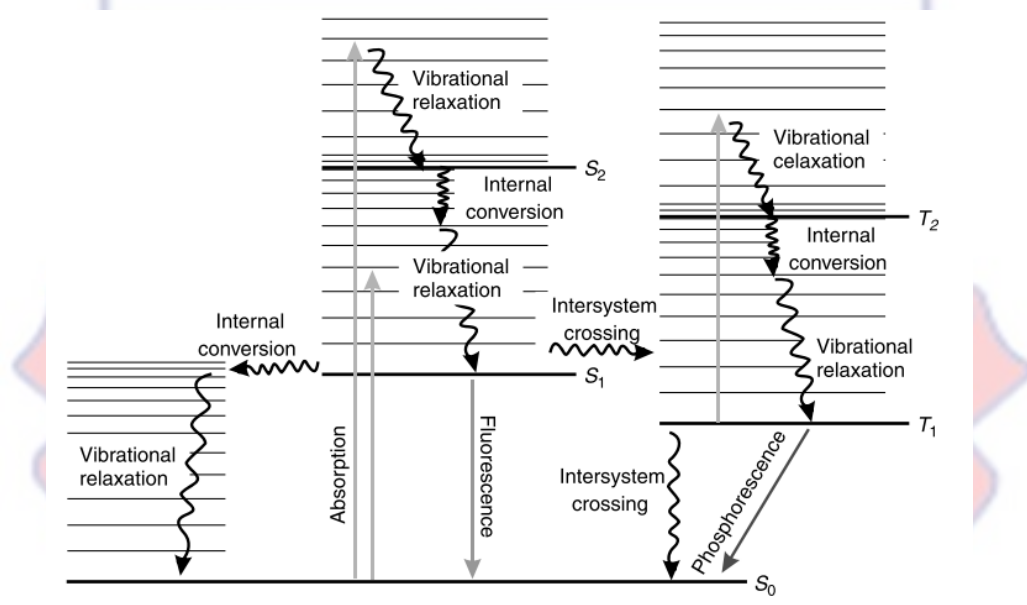


Figure 2: Jablonski diagram describing the electronic energy levels (Sauer et al., 2013)

The likelihood of discovering the fluorophore in one of the potential excited singlet states,  $S_0, S_1 \dots S_n$ , is determined by the excitation wavelength and the transition probabilities (Lakowicz, 1999; Markus et al., 2013; Virk, 2014). Within  $10^{-11} - 10^{-14}$  seconds after being excited to higher excited singlet states,  $S_n$ , fluorophore undergoes internal conversion to highest vibrational levels

of the excited singlet state. Vibrational relaxation causes molecules in highest vibrational levels ( $10^{-10} - 10^{-12}$  s) to rapidly fall to a lower vibrational level of the same excitation state by losing energy.

The intersystem crossing can also reverse the spin of an excited electron, leaving the molecule in the first excited triplet state,  $T_1$  (Markus et al., 2013). Though excited singlet state has higher electronic energy than the triplet state, should in case the vibrational level of the singlet and triplet states overlap, the chances of intersystem crossing increase. Overall, the intersystem crossing efficiency is highly dependent on the fluorophore's nature and transition probability, and it is rarely predictive. The presence of heavy atoms, on the other hand, is known to significantly increase the intersystem crossing rate constant (Markus et al., 2013; Valeur, 2001).

Triplet state lifetime can range from milliseconds to seconds, minutes, or even longer, compared to an excited singlet state's average lifetime of  $10^{-7} - 10^{-9}$  seconds. The intersystem crossing rate, the lifetime and the extinction coefficient at the excitation wavelength of the triplet state all have a strong influence on  $T_1 \rightarrow T_n$  absorption performance (Lakowicz, 1999; Markus et al., 2013; Virk, 2014).

The spontaneous emission of a fluorescence photon provides details about the fluorophore's environment and interactions. The fluorescence emission spectrum, for example, reveals the solvent's polarity, while the fluorescence lifetime and quantum yield directly reveals the fluorophore's quenching interactions with its microenvironment. A fluorophore's fluorescence quantum

yield and lifetime are two of their most significant properties. The mechanism is controlled by the fluorophore's emission rate ( $\Gamma$ ) and its non-radiative decay rate to the lowest singlet state,  $S_0(k_f)$  both depopulate the excited state (Lakowicz, 1999). The quantum yield is determined by the fraction of fluorophores that decay by emission. The quantum yield is expressed in equation 3.

$$Q = \frac{\Gamma}{\Gamma + k_f} \quad (3)$$

where  $\Gamma$  is the fluorophore's emission rate and  $k_f$  is the non-radiative decay rate to  $S_0$ . The brightest emission is seen in substances with the highest quantum yields when it approaches unity. The lifetime of a fluorophore is also significant since it specifies the amount of time it has to communicate with or diffuse in its environment, and therefore the amount of information it can provide from its emission. Many other factors, such as pH, polarity, hydrogen bonding, viscosity, and the existence of quenchers, can affect quantum yields and lifetimes in the condensed phase (Valeur, 2001).

### **Laser-Induced Autofluorescence (LIAF)**

The fluorescence emission from atoms or molecules that have been excited to a higher energy level by absorption of laser light is known as LIAF. This molecule absorbs laser photons, resulting in an electronically excited state without the addition of any histochemical dyes (García-plazaola et al., 2015; Rost, 2016). A filter or a monochromator is positioned for the fluorescent emission to pass through, which is then detected by a photomultiplier (Crosley, 1982; Telle & Donovan, 2007).

If the laser source resonantly stimulates a particle, a photon with the energy,  $h\nu_{12}$  ( $h = \text{Planck's constant}$  and  $\nu = \text{frequency of photon}$ ), is absorbed, raising the particle to the excited state. In general, the selective effect on tissues occurs when the light source is monochromatic in nature, such as a diode laser. Because biological tissue has light receptors (chromophores) that are highly selective to the wavelength it absorbs, the monochromaticity property of laser light, which is related to its specificity of wavelength, becomes a determinant factor in the laser-biological tissue interaction. Due to its collimation property, which allows non-divergent emission, the laser can deposit a large amount of energy in a very small area (spot size).

Diode or semiconductor lasers, which are made up of a p-n junction, are the smallest lasers. They work in milliamps of current and a few volts of voltage. The laser kit as a whole is very thin, and it may be put onto an integrated circuit board if necessary (Bolshov et al., 2015; Selden, 1973). Heterostructure lasers, a type of diode laser, have multiple layers of different materials with same electronic configuration. The materials, such as phosphorous, indium, and aluminum, are configured to help confine the electronic current to the junction region and reduce heat dissipation requirements and current. Semiconductor lasers have wavelengths ranging from 0.4 to 1.8  $\mu\text{m}$  and can produce up to 100 mW of continuous output power (Bolshov et al., 2015; Halina, 2005; Selden, 1973). Induced fluorescence was first observed and debated by R. W. Wood in 1905, several decades before the invention of the laser (Telle & Donovan, 2007).

The popularity of LIAF in many fields, including the investigation of chemical processes, can be attributed to its high sensitivity. For decades, LIAF has been one of the most common laser spectroscopic techniques for studying unimolecular and bimolecular chemical reactions (Rigler & Widengren, 2018; Telle & Donovan, 2007).

### **Secondary Metabolites**

Secretory cells have compartments that contain secondary metabolites that are concentrated in secretory structures like vacuoles and intracellular and extracellular secretion (Roshchina, 2018; Roshchina et al., 2016). Secondary metabolites are specialized compounds that provide the plant with additional surviving and competitive advantages. Natural products and phytochemicals are terms used to describe secondary metabolites (Neilson et al., 2017a). The bioactivities of various secondary metabolites are responsible for their chemopreventive properties, thus, antioxidant, antiparasitic, anticarcinogenic, antiplasmodial and anti-inflammatory effects. Phenols, terpenoids, alkaloids, and their derivatives are the main groups of secondary metabolites that have medicinal properties in humans and animals (Neilson et al., 2017b; Roshchina, 2018). The actions of medicinal herbs and many modern medicines are dependent on secondary metabolites.

### **Polyphenols**

Polyphenols are a structural variety of molecules with numerous phenol functional groups in their structure. Polyphenols are studied for their antioxidant and anti-radical action, natural chelation agents, enzymatic activity,

antiatherosclerotic, anti-inflammatory, anticarcinogenic, hepatoprotective, spasmolytic, antiviral, oestrogenic activity, anticancer properties, antimicrobial, and again useful health effects (Agatonovic-Kustrin & Morton, 2018; Khalaf et al., 2019; Moreno & Peinado, 2012; Pietta et al., 2003; Simpson & Amos, 2017; Tamokou et al., 2017; Teoh, 2016). Phenolic acids, Flavonoids, tannins, lignans, stilbenes, and anthocyanins are polyphenols bearing one or more hydroxyl groups (Khalaf et al., 2019; Pandey & Rizvi, 2009).

### **Phenolic Acid**

Phenolic molecules are widely dispersed in plant tissues, and they play a very important role in fruit color, flavor, and astringency. The content of phenolic compounds in plant tissues can range from 0.5 to 5.0 g per 100 g dry weight. Phenolic molecules are secondary metabolites of plant metabolism that have little impact on the plant's biological or environmental functions (Khalaf et al., 2019). Phenolic acids are the most basic polyphenols based on their chemical structure. Phenolic acids include carboxylic acids with benzoic or cinnamic acid skeletons. (Neilson et al., 2017b).

### **Flavonoids**

Flavonoid chemicals are plant-derived chemicals that can be found in various areas of the plant in nature. They are one of the most distinctive types of chemicals found in higher plants. This compound is found in plants' seeds, leaves, flowers, and bark (Hernández-Rodríguez et al., 2018). Flavones, flavonols, flavanones, flavanonols, flavanols or catechins, anthocyanins, and chalcones are

the subgroups (Hernández-Rodríguez et al., 2018; Panche et al., 2016; Pietta et al., 2003).

Plants, animals, and microorganisms all use flavonoids for a range of biological functions. Flavonoids have antioxidant and biochemical properties that have been associated to diseases such as cancer, atherosclerosis, Alzheimer's disease (AD), and others. Flavonoids are an essential component in a number of nutraceutical, pharmacological, medical, and cosmetic uses because they have a wide range of health-promoting benefits. This is due to their ability to control important cellular enzyme functions as well as their anti-carcinogenic, antioxidative, anti-mutagenic, and anti-inflammatory capabilities (Panche et al., 2016).

### **Tannins**

Tannins are polyphenols with molecular weights ranging between 500 to 3000 u from a chemical standpoint. Tannic acid's chemical structure is determined by the plant species that produces it. All tannins share some characteristics that allow them to be classified as either hydrolysable or condensed. Gallotannines (GTs), ellagitannines (ETs), and complex tannins are three forms of hydrolysable tannins that are based on mainly gallic acid, glucose, and ellagic derivatives. Proanthocyanidins (PAs) are nonhydrolysable tannins made up of flavonoid monomers, quinine units, or flavone derivatives. Condensed tannins are more resistant to microbial destruction than hydrolysable tannins, and they have better antiviral, antibacterial, and antifungal properties (Krzyzowska et al., 2017; Mena et al., 2015; Tamokou et al., 2017).

Tannins have the potential to operate as a biological antioxidant, as well as a complexing or precipitating agent, based on their concentration. Tannins' biological activity is determined by their chemical structure (Krzyzowska et al., 2017; Pietta et al., 2003).

### **Lignans**

Lignans are natural compounds that are dimeric and have the same monomers as lignins. The structure of lignans is 2,3-dibenzylbutane. Many plants include such compounds as minor ingredients, where they serve as the building blocks for the production of lignin in the plant cell wall (Simpson & Amos, 2017). Anticancer, antihypertensive, antioxidant, estrogenic, antiviral, and insecticidal qualities have all been attributed to lignans. Pharmacological properties of its derivatives, podophyllotoxin is the most well-known lignan (Rodríguez-García et al., 2019; Simpson & Amos, 2017; Zitterman, 2003).

### **Alkaloids**

Alkaloids are secondary metabolites that were originally characterized as pharmacologically active nitrogen-based chemicals. There are around 20000 alkaloids known, the majority of which have been isolated from plants (Kukula-Koch & Widelski, 2017; O'Connor, 2010; Verpoorte, 2004).

Alkaloids are important in human medicine as well as in an organism's natural defense. Alkaloids account for about 20% of all secondary metabolites discovered in plants which protect plants from predators while also controlling their growth. Alkaloids are used as cardioprotective agents, anesthetics, and anti-inflammatory drugs in medicine. Morphine, ephedrine, strychnine, quinine,



cocaine, and nicotine are some of the most well-known alkaloids utilized in clinical settings. (Heinrich et al., 2021; Qiu et al., 2014). A variety of alkaloids have commercial potential as medications or as research tools in pharmacology (Verpoorte, 2004).

### **Terpenoids**

Terpenoids, or isoprenoids, are a varied group of naturally occurring chemicals made up of five carbon isoprene units. The basic skeleton and functional groupings of terpenoids distinguish them from one another. Terpenoids are found in practically all kinds of living beings because they are so common. Terpenoid structures have been identified in over 60,000 natural sources, making them one of the most diverse classes of natural compounds ever discovered. They influence the flavor, fragrance, color of the leaves, flowers, and fruits of plants. Terpenoids are also necessary for plant metabolism (Reyes et al., 2017).

Terpenoids are found in many plants and are utilized in traditional medicine for anti-inflammatory and pain-relieving qualities. Many members in the Asteraceae family have traditionally been used to alleviate inflammatory diseases. Their medicinal action is assumed to be due to the sesquiterpene lactones they generate (Agatonovic-Kustrin & Morton, 2018; Cox-Georgian et al., 2019). Artemisinin, which is used to treat malaria, and vincristine and taxol, which are used to treat cancer, are examples of pharmaceuticals created from naturally generated terpenoids (Reyes et al., 2017; Thoppil & Bishayee, 2011).

## **Multivariate Analysis**

Multivariate analysis is a set of techniques for analyzing large datasets containing multiple variables in which the relationships between multivariate measurements and underlying structure are significant (Goodman et al., 1979; Grimnes & Martinsen, 2015). The variables can be evaluated simultaneously employing multivariate analysis in order to reach the key elements of the process that produced them. The multivariate approach allows for both (1) examination of the variables' joint performance and (2) determination of each variable's effect in the presence of the others.

Both descriptive and inferential techniques are proposed in multivariate analysis. The essence of the system becomes the major emphasis with multivariate descriptive approaches and hypothesis tests that include whatever intercorrelations the variables have are known as multivariate inferential techniques (Rencher, 2002).

## **Data Pre-treatment Methods**

Multivariate data processing is used in many procedures, and it is common to perform some pre-processing on the data before the methods may be properly used. This pre-processing stage of data analysis has three basic goals. The initial goal is to limit the amount of data by removing data that isn't related to the research study. Pre-processing data also entails preserving or enhancing enough information within the data to meet the desired aim, as well as extracting or transforming the data into a format appropriate for future analysis (Adams, 2004).

For data processing, the most important spectroscopic data treatment are in terms of noise elimination, dimension reduction, baseline correction, and data standardization. In comparison to untreated data samples, the treated data samples do not differ considerably (Liu, Kyne, Wang, & Yu, 2020; Roy, Sharma, Nath, Bhattacharyya, & Kalita, 2018).

### Noise Elimination

To reduce noise, a variety of procedures are used, including signal averaging and smoothing. The noise associated with a recorded signal has a significant impact on analysis, and the signal-to-noise ratio, or S/N, is one figure of merit used to describe the measurement's quality (Adams, 2004). S/N is defined as,

$$S/N = \frac{\text{average signal magnitude}}{\text{rms noise}} \quad (4)$$

where the rms (root mean square) noise is  $\text{rms noise} = \sqrt{\frac{\sum(\bar{x}-x_i)^2}{n-1}}$  where  $x_i$  and  $\bar{x}$  are the signal and mean noise values, respectively, of n variables (Adams, 2004).

### Signal Smoothing

Smoothing spectral data can be done using a variety of mathematical manipulation approaches. Boxcar averaging is one simple smoothing method. The spectral data is divided into a sequence of distinct, evenly spaced bands, with each band being replaced by a centroid average value. The degree of smoothing increases as the number of points averaged increases. (Adams, 2004; Y.-J. Liu et al., 2020; Press & Teukolsky, 1990). Savitzky and Golay developed a method for smoothing spectral data employing convolution filter vectors obtained from least-

squares-fit polynomial function coefficients. The simple moving average technique can be described in equation 5 as

$$S_i = \sum_{j=-n}^n x_{i+j} w_j \quad (5)$$

where  $x$ , and  $S_i$  are the original and smoothed data vectors, respectively, and  $w_j$  is the smoothing window's weighting factors. When performing a Savitzky-Golay smoothing operation on spectral data, it's important to consider the filtering function (quadratic or quartic), the span frequency (the smoothing function's width), and the number of times the filter will be applied to the data sequentially (Adams, 2004; Press & Teukolsky, 1990).

### **Baseline Correction**

Fluorescence, phosphorescence, and black body radiation all cause unequal amplitude shifts across distinct wavenumbers, resulting in a slowly shifting curve known as baseline. Before continuing with the analysis, these amplitude shifts must be corrected (Qian et al., 2017). Baseline correction can be done in a variety of ways.

The baseline correction can be done either physically or mathematically. Sample purification, gating, and pulsing are all examples of physical approaches that involve instrument adjustments. The mathematical approach includes the polynomial fitting, frequency-domain filtering, and derivative calculation (Guo et al., 2016). The basic baseline correction is the 'baseline offset correction,' in which all variables that constitute the spectrum are subtracted from a constant number (Andrade-Garda, 2009).

## Standardization

Standardization is one of the most used methods for pre-processing spectral data. The purpose of standardization procedures is to make the data values of the measured samples comparable so that statistical analysis may be performed afterwards. Because of differences in volume, samples cannot always be directly compared. This is known as the size effect. The  $j$ -th variable's mean,  $\bar{x}_j$ , and standard deviation,  $s_j$ , are denoted by  $\frac{1}{n} \sum_{i=1}^n x_{ij}$  and  $\sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}$  respectively. The data after standardization is represented by  $\tilde{x}_{ij}$ . Some of the standardization methods are:

### Mean-centering

Mean-centering is useful when the variances of the variables are similar, and the converted variables have mean values of zero as shown in equation 6 (Adams, 2004).

$$\tilde{x}_{ij} = x_{ij} - \bar{x}_j \quad (6)$$

### Autoscaling

The method's goal is to standardize the variables so that each observation has a variance of one and a mean of zero. It uses the mean and standard deviation of variables to alter their values (Berg et al., 2006; Roy et al., 2018). When the variables being measured are recorded in different units, such as concentration, pH, particle size, conductivity, and so on, autoscaling is almost always required (Adams, 2004; Roy et al., 2018). Each property of a variable can be changed using equation 7.

$$\tilde{x}_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j} \tag{7}$$

### Principal Component Analysis

Principal component analysis (PCA) is a well-known multivariate analytic method that employs the spectral decomposition of a correlation coefficient or covariance matrix. Its goal is to keep as much statistical information as feasible while reducing the dimensionality of a dataset (Hotelling, 1933; Salkind, 2012). Considering the values of  $p$  variables  $x_1, x_2, \dots, x_p$  are acquired from each of  $n$  individuals (1, 2, ... ..  $n$ ) in a sample. The mean of the  $j$ th variable is  $\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij}$ . The SP (corrected sum of products), is the corresponding quantity for covariance, and it is defined by equation 8.

$$\frac{1}{n} \sum_{i=1}^n (x_{ij} - \bar{x}_j) (x_{ik} - \bar{x}_k) \tag{8}$$

where  $x_{ij}$  is the  $i$ th variable  $j$  measure and  $x_{ik}$  is the  $i$ th variable  $k$  measure. The purpose of principal components analysis is to find the linear combination of variables that best summarizes the  $n$ -dimensional distribution of data. For each linear combination of variables, a new variable  $X$  is defined as in equation 9.

$$X = a_1x_1 + a_2x_2 + \dots + a_nx_n \tag{9}$$

where  $a$  is a vector of  $a_1, a_2, \dots, a_p$  constants. The new variable's variance,  $s_x^2$ , can be obtained using equation 10.

$$s_x^2 = \sum_{j=1}^n \sum_{k=1}^n a_j a_k Cov_{jk} \tag{10}$$

which, from the definition of covariance, can be rewritten as shown in equation 11.

$$s_x^2 = \sum_{j=i}^n a_j^2 s_j^2 + \sum_{j=i}^n \sum_{k=1}^n a_j s_j a_k s_k r_{jk} \quad (11)$$

where ,  $r_{jk}$ , denotes the linear correlation coefficient between the two variables j and k, as  $r_{jk} = \text{Covariance} / x_j . x_k$ .

With standardized coefficients given to the variables used in the linear combinations, principal components analysis seeks the linear combination with the greatest variance thus the first principal axis, or first principal component. After then, the search for a second linear combination that has the majority of the remaining variance which is uncorrelated to the first principal component continues. This process is repeated till all the basic components are estimated, which is frequently the case. For interpretation and further analysis, the components are then used (Adams, 2004; Hancock & Smyth, 2009; Hotelling, 1933; Jolliffe & Cadima, 2016; Salkind, 2012; Stuart, 1982).

### **Hierarchical Clustering Techniques**

When using hierarchical clustering techniques, the raw data is divided into a few broad categories, each of which is further subdivided into smaller groups until only the individual objects remain. These techniques can be either agglomerative or divisive (Adams, 2004; Holak, 1974; Johnson & Wichern, 2007; Yeturu, 2020). Small groups, starting with individual samples, are fused to form larger groups via agglomerative clustering. Divisive clustering, on the other hand, starts with a single cluster comprising all samples, which is divided into smaller divisions (Adams, 2004). Hierarchical approaches are popular, not least because

their use results in the creation of a dendrogram, which may show both the clustering process and the final result in a two-dimensional visual representation. The generic formula of agglomerative hierarchical clustering that could be used to define the various between-group distance metrics is defined as show in equation 12.

$$d_{k(ij)} = \alpha_i d_{ki} + \alpha_j d_{kj} + \beta d_{ij} + \gamma |d_{ki} - d_{kj}| \quad (12)$$

The distance between objects  $i$  and  $j$  is  $d_{i,j}$ , while the distance between group  $k$  and a new group  $(i, j)$  formed by the fusion of groups  $i$  and  $j$  is  $d_{k(i,j)}$ . The values of the coefficients  $\alpha_i, \alpha_j, \beta$ , and  $\gamma$  are chosen to determine which between-group metric will be employed (Adams, 2004).

### **K-Means Clustering**

K-means clustering main goal is to divide the  $m$  objects, each with  $n$  variables, into  $K$  clusters with the square of the within-cluster sum of distances as small as possible. The elements  $x_{i,j} (1 \leq i \leq m, 1 \leq j \leq n)$  define the  $X$  data matrix, where  $m$  and  $n$  are the number of observations and the number of variables used to define the observations respectively. Cluster analysis looks for  $K$  partitions or clusters, with each item belonging to only one of them (Adams, 2004; Subasi, 2020; Yeturu, 2020).  $B_{L,j} (1 \leq L \leq K)$  is the mean value for each variable  $j$  for all objects in cluster  $L$ . The Euclidean metric gives the distance  $D_{i,L}$  between the  $i$ th object and the center of each cluster as shown in equation 13.



$$D_{i,L} = \left[ (x_{ij} - B_{Lj})^2 \right]^{\frac{1}{2}} \quad (13)$$

The algorithm works by shifting an object from one cluster to the next in order to reduce the error, and it stops when no more movement can be made (Adams, 2004).

### Linear Discriminate Analysis (LDA)

LDA is a classifier that finds a linear combination of features that distinguishes two or more data classes. LDA, like PCA, can be used for dimension reduction as well as classification. The classes in LDA are supposed to be regularly distributed. The a priori probabilities for class 1 and class 2 in a two-class dataset are  $p_1$  and  $p_2$ , respectively; the class means  $\mu_1$  and  $\mu_2$ , overall mean,  $\mu$ , and the class variances ( $cov_1$  and  $cov_2$ ) can be related as shown in equation 14.

$$\mu = p_1 \times \mu_1 + p_2 \times \mu_2 \quad (14)$$

Between-class and within-class scatters are then used to represent the required criteria for class separability. For a multiclass setting, the scatter measures are determined in equation 15 and 16 as follows:

$$S_w = \sum_{j=1}^c p_j cov_j \quad (15)$$

where  $c$  denotes the number of classes and

$$cov_j = (x_j - \mu_j)(x_j - \mu_j)^T \quad (16)$$

The spread between classes is calculated as shown in equation 17.

$$S_b = \frac{1}{C} \sum_{j=1}^c (\mu_j - \mu)(\mu_j - \mu)^T \quad (17)$$

The goal is to find a discriminant plane that maximizes the within-class to between-class scattering ratio (variances) which is defined in equation 18 as:

$$J_{LDA} = \frac{wS_b w^T}{wS_w w^T} \quad (18)$$

In practice, the class covariance and means are not always known, but the training set could be used to estimate them. In the aforementioned equations, the maximum likelihood estimate might be utilized instead of the precise value (Adams, 2004; Hotelling, 1933; Neath & Johnson, 2010; Subasi, 2020; Yeturu, 2020).

### **Quadratic Discriminant Analysis (QDA)**

This method is a probabilistic parametric classification algorithm that emerged from the LDA algorithm for nonlinear class separations. QDA, like LDA, assumes multivariate normal probability density distributions, but the dispersion is not uniform across all categories. As a result, the categories differ in terms of their centroid position as well as the variance–covariance (Oliveri et al., 2010). QDA is a quadratic decision limit discriminant function that could be used to categorize data sets having two or more classes. QDA offers a higher prediction power than LDA, however it requires the estimation of each class's covariance matrix (Canizo et al., 2019).

### **K-Nearest Neighbor (KNN)**

K-nearest-neighbor technique is a nonparametric and relatively a straightforward classification method. An observed data sets that are closest are

identified and given a point,  $x_o$ , that needs to be classified into one of the  $K$  groups. The population with the maximum data sets observed among the  $k$ -nearest neighbors receives  $x_o$  as a categorization criteria. Points where there is no majority would either be assigned to one of the largest categories at random or left unclassified. The simplicity of  $k$ -nearest neighbor classification is an advantage. A user just has to choose between two options: (1) the number of neighbors and (2) the distance measure to utilize. Mahalanobis distance, Euclidean distance, and city-block distance are all popular distance metrics. The Euclidean distance between two pattern vectors is the most commonly used distance measure. The Mahalanobis distance would take into consideration the data's intrinsic correlations. In testing or cross-validation (CV) the classifier's quality, test data set are frequently used to determine the number of neighbors (Adams, 2004; Johnson & Wichern, 2007; Kessler & Kessler, 2020; Neath & Johnson, 2010).

### **Support Vector Machine (SVM)**

The support vector machine (SVM) technique is comparable to linear discriminant analysis in terms of how it operates. Similar to LDA, a set of hyperplanes is formed to split the feature vectors into various classes, this then selects the hyperplane that is the furthest away from the nearest training samples. According to Cover's theorem, SVM determines the hyperplane with the largest margin by translating input data into high-dimensional space. When a complex classification problem is represented in high-dimensional nonlinear space rather than low-dimensional nonlinear space, it is more likely to be linearly separable. Regularization is also used by SVM to avoid artifacts. Nonlinear SVM can also

exist with a nonlinear decision boundary attributable to the Kernel function, giving it more flexibility (Jothilakshmi & Gudivada, 2016; Neath & Johnson, 2010; Vaibhaw et al., 2020).

### Classification Performance

One of the most important aspects in classification is the evaluation of the classified models. A classifier can be judged on its sensitivity, specificity, relevance, precision recall, accuracy, and error. Receiver operating curves (ROC), learning curves, and confusion matrix are standard methods for assessing the outcomes of classification algorithms (Kotu & Deshpande, 2019; MathWorks, 2014; Oprea & Ti, 2014). The confusion matrix is the best way to describe classification performance.

Confusion matrix shows the number of right and wrong predictions made by the model in comparison to the actual classifications in the test data (MathWorks, 2019; Oprea & Ti, 2014). Table 1 shows the confusion matrix for a classifier with two classes: True and False.

**Table 1: The confusion matrix of a classification model arranged in two classes as a  $2 \times 2$  matrix**

		Actual Class (Observation)	
		True class	False Class
Predicted class (expectation)	True class	True Positive (TP)	False Positive (FP)
	False Class	True Negative (TN)	False Negative (FN)

Source: (Kotu & Deshpande, 2019; Oprea & Ti, 2014)

### Sensitivity

Sensitivity refers to a classifier's ability to select all of the examples that must be selected. Sensitivity is defined as a percentage (or ratio) using the formula as shown in equation 19.

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (19)$$

### Specificity

A classifier's specificity refers to its capacity to reject all of the cases that need to be rejected. The ratio (or percentage) of specificity is defined in equation 20 as:

$$\text{Specificity} = \frac{TN}{TN + FN} \quad (20)$$

### Precision

The proportion of cases found that were genuinely relevant is defined as precision. Precision is therefore defined in equation 21 as:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (21)$$

### Accuracy

Classifiers ability to choose all cases that need to be selected and reject all cases that need to be rejected is defined as accuracy. This would indicate that FP = FN = 0 for a classification model with 100 % accuracy. The accuracy formula is shown in equation 22 as:

$$\text{Accuracy} = \frac{TN + TP}{TP + FP + TN + FN} \quad (22)$$

### Chapter Summary

This chapter reviewed the literature on some basic principles of light interaction with matter, with emphasis on fluorescence. The secondary metabolites present in medicinal plants were described, followed by a discussion of the theories involved in multivariate data analysis.



## CHAPTER THREE

### RESEARCH METHODS

#### Introduction

The experimental set-up, method and procedure for acquiring data of LIAF from AMHDs samples are presented in this chapter. The first part describes the collection of AMHDs samples, the second part focuses on experimental set-up and method for the acquisition of LIAF spectra data while the last part of this chapter describes the various methods used for the pre-processing of LIAF spectra data.

#### Anti-Malarial Herbal Drugs (AMHDs) Samples Collection

This research work utilized eleven (11) commercial anti-malaria herbal drugs (AMHDs), prepared as decoction. One of the AMHDs samples purchased from Centre for Scientific Research into Plant Medicine (CSRPM) at Mampong-Akwapim was used as the standard. This AMHD is trademarked as Mampong Herbal Medicine and has been used extensively in the treatment of malaria at CSRPM's herbal clinic in Mampong-Akwapim, Ghana. This AMHD was prepared from *Cryptolepis Sanguinolenta*, also known as nibima, which is of a particular interest and has been shown to be clinically potent in the treatment of malaria.

The remaining ten (10) AMHDs samples were purchased from licensed herbal shops distributors within the Cape Coast Metropolis. These AMHDs were selected based on a study carried out by Wilmot et al. (2017), which laid emphasis

on the number of mostly used AMHDs in the coastal regions of Ghana (Wilmot et al., 2017).

For the purpose of anonymity and confidentiality, the eleven (11) AMHDs were re-labelled as; AMHD\_A, AMHD\_B, AMHD\_C, AMHD\_D, AMHD\_E, AMHD\_F, AMHD\_G, AMHD\_H, AMHD\_I, AMHD\_J and AMHD\_K. All the AMHDs samples bear the registration number of the Food and Drug Authority (FDA) as shown in Table 2.

**Table 2: Anti-malaria herbal drugs (AMHDs) and their associated Food and Drug Authority (FDA) registration number.**

AMHDs	FDA registration number
AMHD_A	FDB/HD 07-716X
AMHD_B	FDB/HD 15-123X
AMHD_C	FDB/HD 09-510X
AMHD_D	FDB/HD 05-907X
AMHD_E	FDA/HD 17-103X
AMHD_F	FDA/HD 18-041X
AMHD_G	FDB/HD 05-806X
AMHD_H	FDB/HD 10-509X
AMHD_I	FDB/HD 17-030X
AMHD_J	FDA/HD 16-041X
AMHD_K	FDA/HD1 16-020X

The AMHDs samples on market were in 500 ml volumes in amber polyethylene terephthalate (PET) plastic bottles. Consumer Medicine Information (CMI) of all the AMHDs samples were either written on the sample bottle or box, showing the medication and constituents making up the decoction. From the CMI of all AMHDs samples, only AMHD\_G and AMHD\_K were prepared from a



single plant constituent with the rest having as many as two (2) to six (6) plants as constituents. The plants used in the preparation of each AMHDs are shown in Table 3.



**Table 3: Plant species present in Anti-Malaria Herbal Drugs (AMHDs)**

Plant specie	Family	Local name	Part used	AMHD present in	Reference
<i>Alchornea cordifolia</i>	Euphorbiaceae	Ogyama	L	E	(Appiah, Oppong, Mardani, Omari, Kpabitey, Amoatey, Onwona-agyeman, Oikawa, Katsura & Fujii, 2018)
<i>Alstonia boonei</i>	Apocynaceae	Onyame dua	F/R	E	(Oppong, Agyare, Duah, Mbeah, Asase, Sarkodie, Nettey, Adu, Boatema & Agyarkwa, Amoateng, 2019)
<i>Anthocleista nobilis</i>	Longaniaceae/ Gentisnaceae	Wudini k&E	L/Sb	B	(Ngwoke, 2018); (Oppong et al., 2019); (Kumlaga, Agyare, Dickson, Mensah, Annan, Loiseau & Champy, 2015)
<i>Azadirachta indica</i>	Meliaceae	Neem tree	Sb	A/F/I	(Appiah et al., 2018)
<i>Bindens</i>	Asteraceae	Gyinantwi	L	C	(Kumlaga et al., 2015)
<i>Bombax buonopozense</i>	Malveceae/ Bombonaceae	Akata	Sb	D/J	(Oppong et al., 2019); (Kumlaga et al., 2015)
<i>Citrus aurantifolia</i>	Rutaceae	Ankaa twade&	L/F	C	(Appiah et al., 2018); (Oppong et al., 2019)
<i>Cola gigantea</i>	Malveceae	Watapuo	L	J/D	(Kumlaga et al., 2015)

Table 3, continued

<i>Cryptolepis Sanguinolenta</i>	Periplocaceae	Nibima	R	G/I	(Komlaga et al., 2015)
<i>Cymbopogon citratus</i>	Poaceae	Lemon grass	L	I	(Komlaga et al., 2015)
<i>Khaya senegalensis</i>	Meliaceae	Kuntunkuri	Sb	E/F	(Appiah et al., 2018); (Adeniyi et al., 2018)
<i>Morinda lucida</i>	Rubiaceae	Konkrɔma	L/Sb	H	(Appiah et al., 2018); (Oppong et al., 2019)
<i>Moringa oleifera</i>	Moringaceae	Moringa	L/S	I	(Appiah et al., 2018)
<i>Nauclea latifolia</i>	Rubiaceae		R	H	(Osei-Djarbeng Agyekum- Attobra, Nkansah, Solaga, Osei-Asante & Owusu-Dapaah, 2015); (Asanga et al., 2017)
<i>Ocimum viride</i>	Lamiaceae	Nunum	L	I	(Osei-Djarbeng et al., 2015); (Komlaga et al., 2015)
<i>Paullinia pinnata</i>	Sapindaceae	Toa-ntini	L	C/E	(Appiah et al., 2018)
<i>Polyathia longifolia</i>	Annonaceae	Weeping willow	F	H	(Komlaga et al., 2015)
<i>Psidium guajava</i>	Myrtaceae	Guave	L	C	(Komlaga et al., 2015); (Adeniyi et al., 2018)
<i>Pycnanthus angolensis</i>	Myristcaceae	Otie	L	E	(Appiah et al., 2018)

Table 3, continued

<i>Rauwolfia vomitoria</i>	Apocynaceae	Kakapenpen	F	E	(Oppong et al., 2019); (Kumlaga et al., 2015)
<i>Senna siamea</i>	Fabaceae	Acacia	L/Sb	H	(Kumlaga et al., 2016)
<i>Solanum torvum</i>	Solanaceae	Nsusuwa	L/F	D/J/ H	(Adeniyi et al., 2018); (Appiah et al., 2018)
<i>Spathodea campanulata</i>	Bignoniaceae	Akuakuoninsuo	L/Sb	D/J	(Kumlaga et al., 2015);
<i>Synedrella nodiflora</i>	Asteraceae		L	H	(Gaertn, 2017)
<i>Theobroma cacao</i>	Malvaceae	kookoo	L	I	(Souza & Moreira, 2018)
<i>Vitex grandifolia</i>	Lamiaceae	Samanibir	L	B	(Oppong et al., 2019); (Bello, Zaki, Khan, Fasinu, Ali, Khan, Usman, & Oguntoye, 2017)
<i>Vernonia amygdalina</i>	Asteraceae	AwOnwono	L/S	D/J	(Appiah et al., 2018)
<i>Xylopla aethiopica</i>	Annonaceae	Hwentia	Sd	C	(Oppong et al., 2019); (Kumlaga et al., 2015)

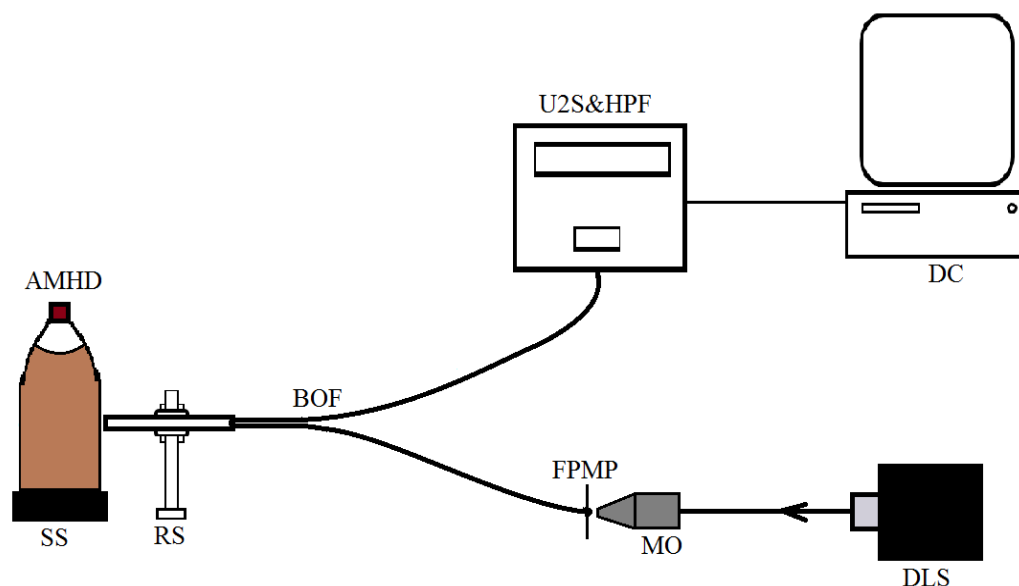
L – Leaves    F – Fruit    Sb – Stem bark    Sd – Seed    S – Stem    R – Root

### **Laser-Induced Autofluorescence Experimental Set-up**

The laser-induced autofluorescence (LIAF) set-up (Figure 3) consists of a light source, optical components and a detecting system. The light source is a 405 nm continuous wave (CW) diode laser (Laser module, China). It has a DC input voltage of 12 V and current of 2 A, with a maximum output power of 100 mW. The diode laser is controlled by a Transistor-Transistor Logic (TTL) modulator laser switch between 2.5 and 5 V.

The optical components includes a microscope objective (040AS016 MELLES GRIOT, Thorlabs, USA) of a 40× magnification and a numerical aperture of 0.65, a fibre port micropositioner (PAF-SMA-7, Thorlabs, USA) with a SubMiniature version A (SMA) 905 connector and a visible/near-infrared bifurcated optical fibre (R400-7-VIS-NIR, Ocean Optics) that has a common end of 600 µm probe with a diameter of 3.175 mm and a length of 76.2 mm. The bifurcated optical fibre has two split arms from an optical fiber probe, with one arm connected to the light source and the other arm connected to the spectrometer.

The detecting system of the LIAF set-up is made up of a USB2000 spectrometer (USB2000, Ocean Optics, USA) which operates on a power of 100 mW at 5 V and a high pass absorptive edge filter (GG445;  $\lambda > 455$  nm) of diameter 12.5 mm. Other components in the LIAF experimental set-up are the sample stage, retort stand and a desktop computer (Hewlett-Packard Company HP Compaq dx2390 Microtower, Intel (R) Pentium (R) Dual CPPU E2180 @2.00 GHz, 0.99 GB RAM).



*Figure 3:* Schematic diagram of LIAF experimental set-up for anti-malaria herbal drugs (AMHDs) data acquisition. DLS- Diode Laser Source, MO- microscope objective, FPMP- fibre port micro-positioner, RS- retort stand, SS- sample stage, AMHD- Anti-malarial herbal drug, BOFP- bifurcated optical fibre, U2S&HPF- USB2000 spectrometer and high pass filter, and DC- desktop computer

The values from the output of the detector were read by OOIBase32 spectrometer operating software (Ocean Optics Inc. Version 1.0, USA) on a Microsoft Windows XP Desktop Computer. The Operating Software provided a real time display of spectra and helps a variety of process functions such as integration time, signal averaging and boxcar to be altered to prevent peaks from being off-scale. Before performing a measurement with the LIAF set up, the spectrometer was calibrated at 0% transmittance by unchecking the check box of the “Strobe/Lamp Enable” and also subtracting the stored dark spectrum with the light source blocked.

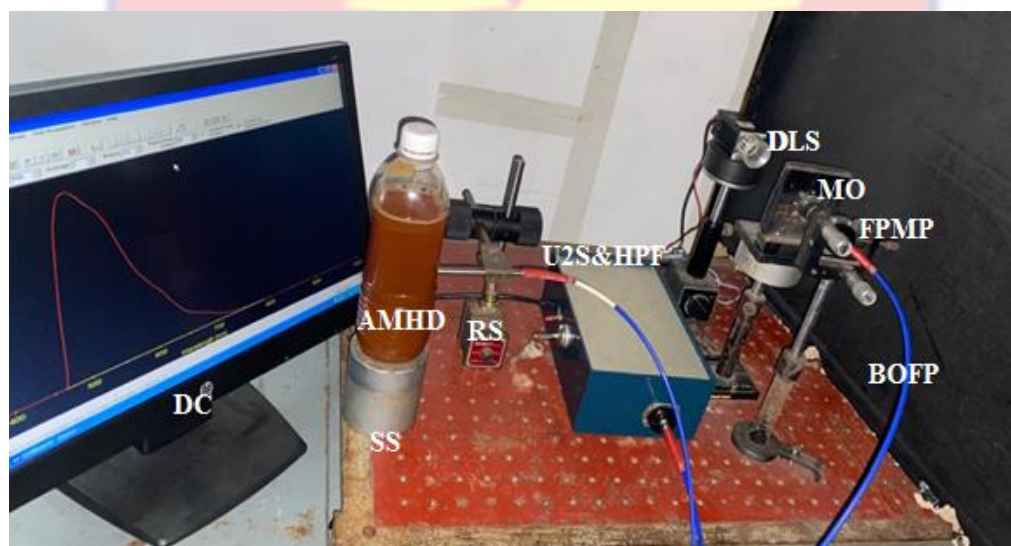
### LIAF Spectra Measurement

All measurements were taken in a Polyethylene terephthalate (PET) plastic bottle since it exhibits minimal absorption in the ultraviolet wavelength ranges. A 500 ml AMHD\_A was transferred to a clear 500 ml PET plastic bottle and capped. The AMHD\_A sample was shaken vigorously and positioned onto a sample stage. A continuous wave (CW) diode laser with a maximum power output of 100 mW was focused onto the AMHD sample by a bifurcated fibre optical probe through a microscope objective and a fibre port micropositioner. Fluorescence from the AMHD\_A sample was collected by the same optical probe and transmitted through the output arm of the bifurcated optical fibre to the ocean optics spectrometer through the long pass edge filter which cut off the excitation source.

The fluorescence spectrum of AMHD\_A was monitored in real time with the aid of the OOIBase32 software. The acquisition parameters such as the *integration time*, *average* and *boxcar* of the OOI Base 32 software were altered for an anticipated AMHD\_A spectrum. The *integration time* was adjusted to 500 ms with the *average* adjusted to 2 and *boxcar* to 10, which respectively set the signal-to-noise ratio and smoothing width of AMHD\_A spectrum. With an *initial delay* of 2 ms a total of 57 spectra were acquired.

With the same acquisition parameters, LIAF measurements were repeated three times on different spots of the AMHD\_A sample. The process was repeated for all the remaining AMHDs samples and the fluorescence emission spectra were recorded in each case. Since the spectrometer operating software does not allow

vigorous analysis of fluorescence emission spectra, all the individual spectra data were saved and exported into an excel matrix format for pre-processing and further analysis using a commercial third party software package MATLAB (Mathworks, R2020a) and OriginLab (OriginPro, 2019b). A photograph showing arrangement of instruments are shown in figure 4.



*Figure 4:* Components of LIAF set-up used for AMHDs data acquisition. DLS-Diode Laser Source, MO-microscope objective, FPMP-fibre port micro-positioner, RS-retort stand, SS-sample stage, AMHD-Antimalarial herbal drug, BOFP-bifurcated optical fibre, U2S&HPF-USB20000 spectrometer and high pass filter, and DC-desktop computer

### **Pre-processing of LIAF Spectra Data**

In the preparation of LIAF spectra data for analysis, two preprocessing techniques were employed thus Savitzky-Golay smoothing filter and auto-scaling technique. Prior to the preprocessing techniques, the spectra data were extracted into a matrix format in MAT-files version 7 by a self-written MATLAB algorithm and their intensities were averagely computed. Since medicinal plants lighten all



colors in the visible spectrum due to the presence of different fluorescent secondary metabolites, a wavelength range of 450-800 nm was subjected to the preprocessing methods. A structured chart of the pre-processing methods for the LIAF spectra is presented in figure 5.

Savitzky-Golay smoothing filter was first applied to remove noisy signals from each LIAF spectra data. The method smoothens the observations individually at a width of the frequency band (span frequency) to be 0.1 in a second degree polynomial. The linear least square regression would be calculated at every frequency band of 0.1. Savitzky-Golay filter ideally minimizes the least error of a fitting polynomial to each noisy data frame.

So to significantly better the classification between the AMHDs samples, standardization technique was used in modifying the variables of spectra data within appropriate range. A comparative analysis of the LIAF spectra was done using an auto-scaling technique which was classified as a standardization technique, it was applied to the smoothen the individual spectra intensities data to adjust each variance of variables allowing making further analysis to be done on the basis of correlation rather than variances. The returns would be used in measuring the similarity and relationship between the variables, allowing characterization of each spectral intensity on comparable scale across all AMHDs samples. The skewness and kurtosis of the individual LIAF spectra remains the same as the original LIAF spectra.

A *standardize* syntax in Matlab was used to auto-scale the smoothen data set. The syntax calculated the mean centering of each column of the spectra data,

thus, subtracting the mean of variables from each value of the spectra data. The standard deviation of the spectra data at each column was subtracted from the mean centering. The mean spectra data output from the Savitzky-Golay smoothing filter and autoscaling technique were plotted separately for further analysis. The new sets of the fluorescence data were further analyzed using a different multivariate analysis to classify and compare the secondary metabolites that may be present in the AMHDs.

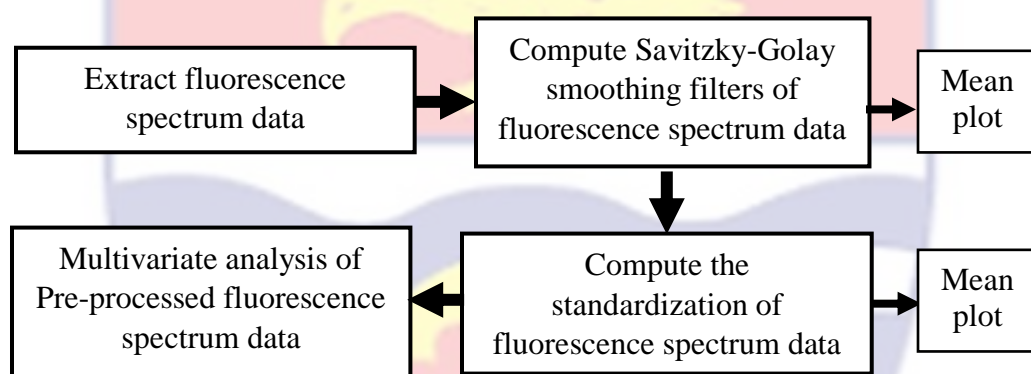


Figure 5: A structured chart for pre-processing LIAF spectra data of AMHDs

### Chapter Summary

The first part of this chapter described the collection of AMHDs samples. LIAF measurements that was employed in the acquisition of AMHD samples were further described in the second part of this chapter. The various pre-processing techniques used in the smoothing and standardization of the fluorescence spectral data of the individual AMHD sample were described in the final part of the chapter.

## CHAPTER FOUR

### RESULTS AND DISCUSSION

#### Introduction

This chapter presents results, analysis and discussions on LIAF spectra for the identification and classification of AMHDs samples. The first part of this chapter discusses the spectral fingerprints of AMHDs and the identification of possible secondary metabolites that may be present. The last part discusses the use of chemometrics in the classification of the AMHDs.

#### LIAF Spectra of AMHDs

The LIAF spectra obtained from each AMHDs were the extract of water soluble secondary metabolites from either parts of a specified plant species tissues, such as the leaves, flowers, seeds, barks, roots, or the shrubs. Examples of these water soluble secondary metabolites are phenols and phenolic acid, flavonoid, stilbenes, lignin, etc. and are responsible for their anti-malarial potencies (Harborne, 1973; Oladeji et al., 2020; Roshchina, 2018; Tay et al., 2011).

Visual inspection of the AMHDs showed that AMHD\_H had a very light color and almost transparent, compared to the least transparent AMHD (AMHD\_D), which had a charcoal gray color. Though none of the AMHDs was the same when visually compared and the variations can also not be certified by a mere look and labeling. As a result, a more rapid sensitive and non-destructive technique was employed in evaluating some selected AMHDs using their fluorescence emission spectra at an excitation wavelength of 405 nm.

A plot of the smoothed LIAF spectra for the eleven (11) AMHDs: AMHD\_A, AMHD\_B, AMHD\_C, AMHD\_D, AMHD\_E, AMHD\_F, AMHD\_G, AMHD\_H, AMHD\_I, AMHD\_J, and AMHD\_K displayed a unique spectra signature. The LIAF spectra data of each AMHD was smoothed to remove noisy signals from each fluorescence spectra.

A plot showing the LIAF spectra signature of AMHD\_K is shown in figure 6. The plot displayed a broad emission band within the spectra range of 460-800 nm with a major emission at 500-650 nm which was characterized in the green to orange fluorescence. AMHD\_G exhibited a likely spectra nature of AMHD\_K, with the same major broad band emission spectra range. The LIAF spectra of AMHD\_A, AMHD\_B and AMHD\_C also exhibited one major broad emission band which ranges from a wavelength of 460 to 620 nm and can be characterized in the blue (455-492 nm) to orange (590-620 nm) fluorescence as described in figure 7.

The LIAF spectra of AMHD\_D, AMHD\_E, AMHD\_F, AMHD\_H, AMHD\_I and AMHD\_J exhibited two emission bands. The major emission band was within 460-650 nm and a minor emission band was within 650-750 nm. The major emission band was characterized by the blue to orange fluorescence while the minor emission band was characterized by the red to far-red fluorescence. The spectra plot of AMHD\_D, AMHD\_E, AMHD\_F, AMHD\_H, AMHD\_I and AMHD\_J are shown in appendix A.

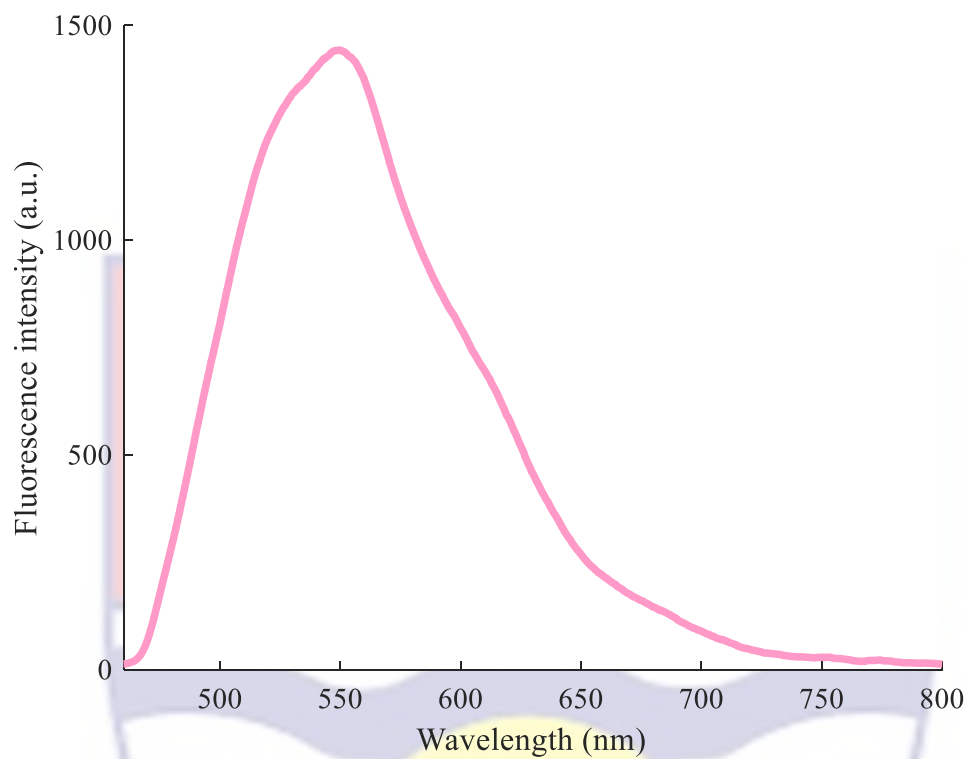
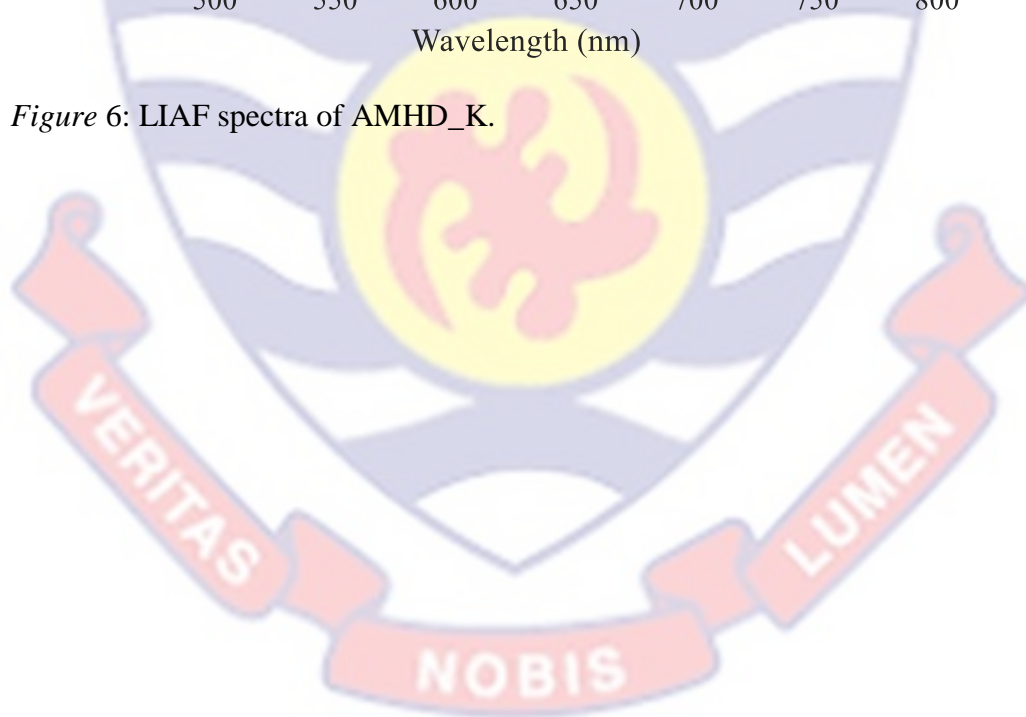


Figure 6: LIAF spectra of AMHD\_K.



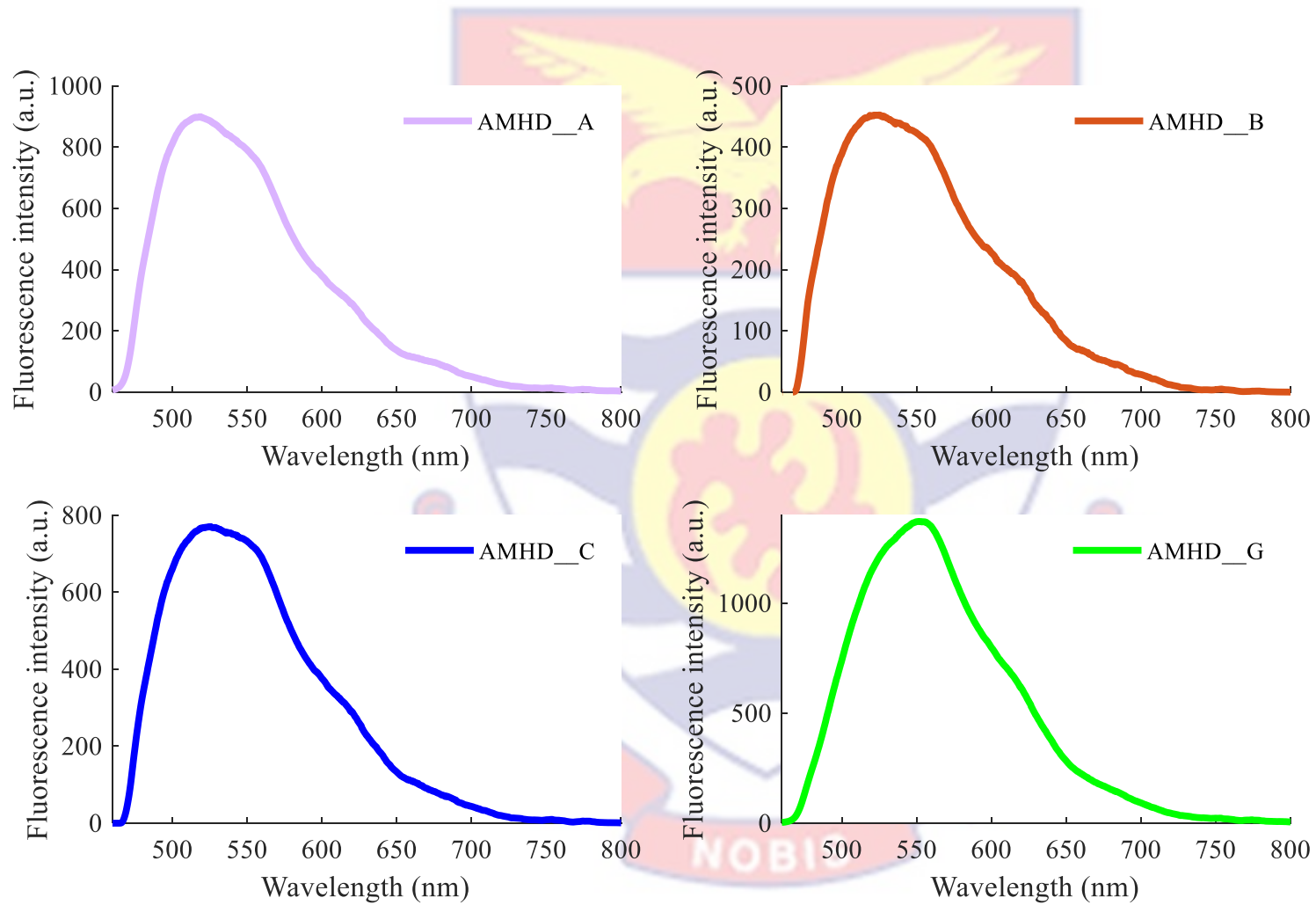


Figure 7: LIAF spectra of AMHD\_A, AMHD\_B, AMHD\_C and AMHD\_G

Intensity-wise, all fluorescence spectra differ depending on the AMHD. The results showed in figure 8 indicate that the fluorescence spectra intensity of AMHD\_H was higher and easier to differentiate as compared to the other ten (10) AMHDs. Among the ten (10) AMHDs, AMHD\_E, AMHD\_K, and AMHD\_G showed fluorescence intensity within a range of 1250-1500 a.u. On the contrary, AMHD\_A, AMHD\_C, AMHD\_J, and AMHD\_I showed a fluorescence intensity range within 500-1000 a.u. Followed by AMHD\_B and AMHD\_F then AMHD\_D that has the least fluorescence intensity among all the AMHDs. The spectra signature exhibited among the AMHDs were as a result of their different biophysical properties among them. The AMHDs could therefore be grouped according to their fluorescence intensities that were comparably the same.

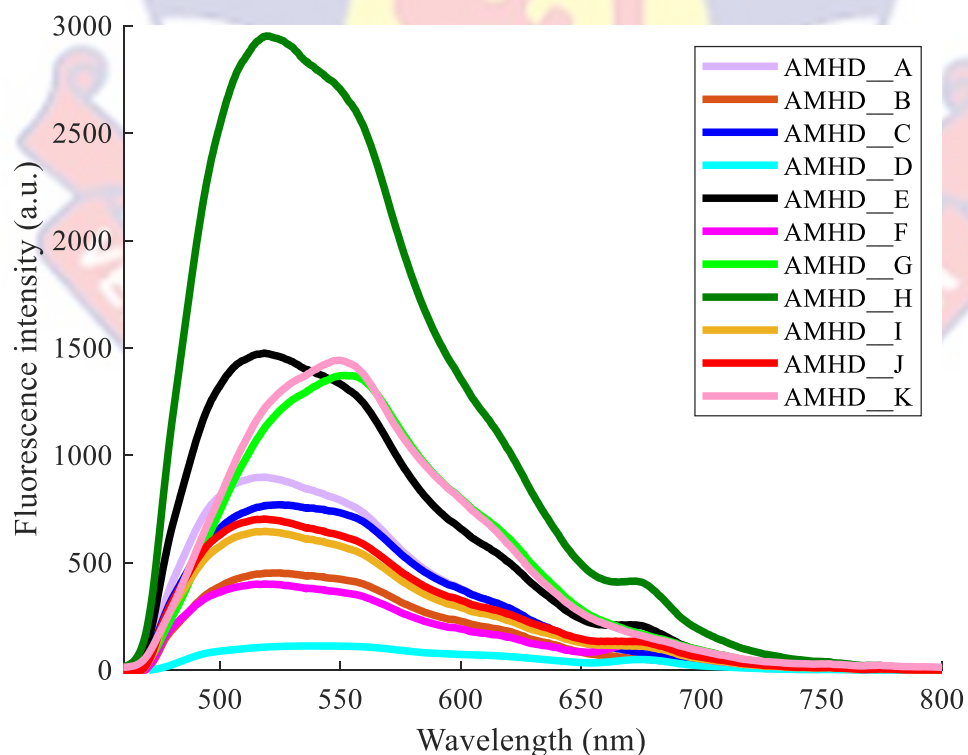


Figure 8: LIAF spectra showing disparities in AMHDs

The plot of standardized fluorescence spectra of all AMHDs is shown in figure 9. In this case after standardization (autoscaling), the structure of spectral signature of AMHDs became the main aim of study thus, the plot of all the AMHDs spectra together was evaluated based on correlation instead of covariance. Therefore making the secondary metabolites equally important for the study. In the correlation study of the AMHDs, two emission bands could be observed when all the LIAF spectra of the AMHD were plotted together. The first fluorescence emission band was the major band with wavelength range of 460-620 nm. Eight of the AMHDs (AMHD\_A, AMHD\_B, AMHD\_C, AMHD\_E, AMHD\_F, AMHD\_H, AMHD\_I and AMHD\_J) showed high fluorescence within a wavelength range of 495-560 nm and could be characterized by the green (490-575 nm) fluorescence.

Three of the AMHDs: AMHD\_D, AMHD\_G and AMHD\_K showed high fluorescence within a wavelength range of 510–580 nm and could be characterized by the fluorescence of green (492-575 nm) and yellow (575-597 nm). The second emission band was within a wavelength range of 650-750 nm. This emission band among the fluorescence spectra of AMHDs (AMHD\_D, AMHD\_E, AMHD\_F, AMHD\_H, AMHD\_I and AMHD\_J) was characterized in the red to far-red (620-800 nm) fluorescence. The fluorescence emission band in this range was higher for AMHD\_D followed by AMHD\_F, AMHD\_J, AMHD\_I and AMHD\_H in that order as compared to the other AMHDs. This may be the reason for the high concentration of secondary metabolites that were characterized in the red fluorescence (620-700 nm) as compared to the other AMHDs.



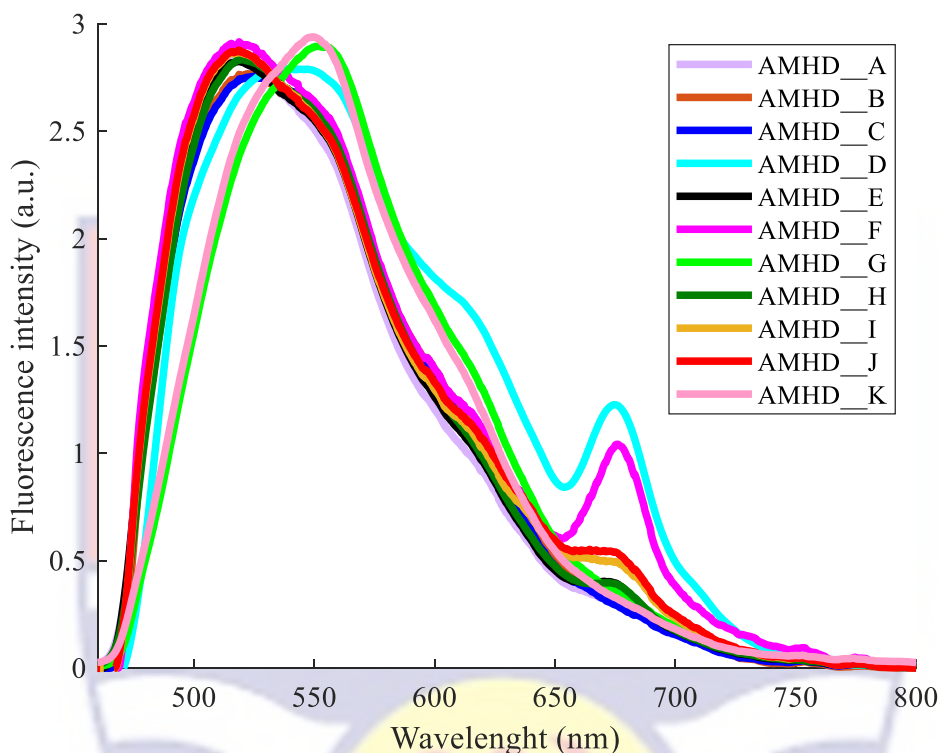


Figure 9: Standardized LIAF spectra showing differences in AMHDs

Secondary metabolites found in the secretory cells of plant could fluorescence in all color of the visible spectrum when exposed to ultraviolet radiation (Roshchina, 2018; Roshchina et al., 2017). The hidden peaks of the individual fluorescence spectra could therefore be used to predict the type of secondary metabolites that may be present in each AMHDs. The LIAF spectra for all the AMHDs were therefore deconvolved, where the spectra signature were decomposed to extract information on each hidden peak. For each spectra signature, broad and overlapping hidden peaks with their respective different peak wavelength centers are presented. Each hidden peak was characterized by a secondary metabolite that may be present in each AMHD sample. The

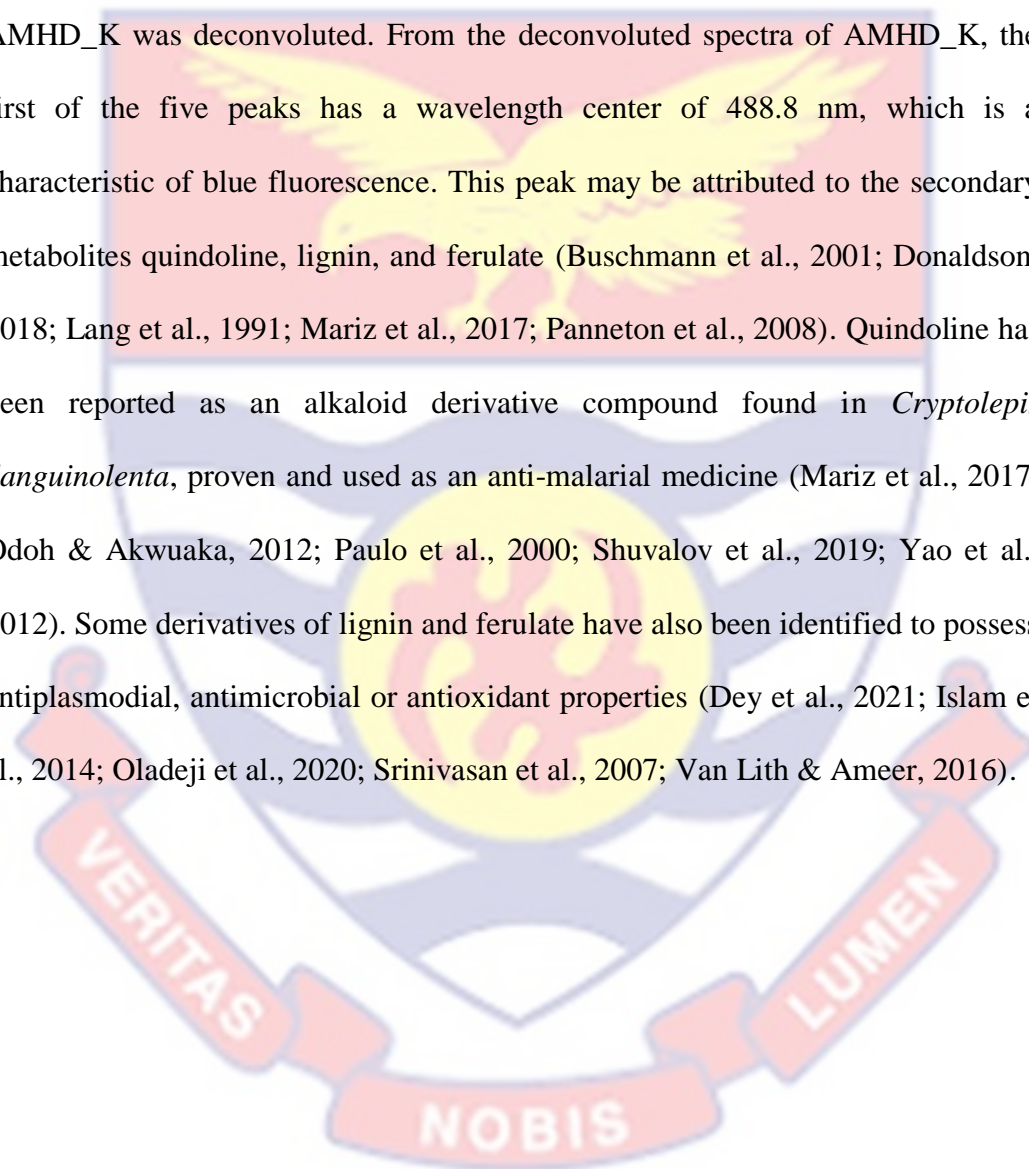
fluorescence intensities height of the deconvoluted peaks from the individual standardized LIAF spectra depends on the base, wavelength center, width and area of their respective hidden peak. Each deconvoluted spectra contains the spectra signature superimposed with their corresponding fitted spectra, the deconvoluted peaks, their respective wavelength centers and the coefficient of determination.

The degree of the interrelation of variables between the fitted spectra and their respective AMHDs spectra was estimated by the coefficient of determination,  $R^2$ . The fitted spectra of all the AMHDs fluorescence spectra indicated an average of  $0.9981 \pm 0.0018$  goodness-of-fit. The coefficient of determination,  $R^2$ , showed a highly positive interdependence across all the averaged LIAF spectra, indicating similarity and association between the fitted spectra variables and their respective LIAF spectra variables. Five hidden peaks were identified for each deconvoluted AMHDs spectra.

In this analysis, AMHD\_K was used as a standard against which all AMHDs were assessed. This AMHD was prepared from *Cryptolepis Sanguinolenta*, also known as nibima, which is of a particular interest and has been shown to be clinically potent in the treatment of malaria. The LIAF spectra of AMHD\_K showed only one predominant broad band without any shoulder. The five hidden peaks identified for the AMHD\_K spectra could be characterized by blue-red fluorescence as shown in figure 10. According to the manufacturers, AMHD\_K was prepared from *Cryptolepis Sanguinolenta*. Variety of secondary metabolites such as terpenes, phenolic compounds and derivatives of alkaloids

have been reported to be present in *Cryptolepis Sanguinolenta* when extracted through decoction (Cimanga et al., 1997; Osafo et al., 2017; Tay et al., 2011; Wright et al., 2001).

Five hidden peaks were observed when the LIAF spectra signature of AMHD\_K was deconvoluted. From the deconvoluted spectra of AMHD\_K, the first of the five peaks has a wavelength center of 488.8 nm, which is a characteristic of blue fluorescence. This peak may be attributed to the secondary metabolites quindoline, lignin, and ferulate (Buschmann et al., 2001; Donaldson, 2018; Lang et al., 1991; Mariz et al., 2017; Panneton et al., 2008). Quindoline has been reported as an alkaloid derivative compound found in *Cryptolepis Sanguinolenta*, proven and used as an anti-malarial medicine (Mariz et al., 2017; Odoh & Akwuaka, 2012; Paulo et al., 2000; Shuvalov et al., 2019; Yao et al., 2012). Some derivatives of lignin and ferulate have also been identified to possess antiplasmodial, antimicrobial or antioxidant properties (Dey et al., 2021; Islam et al., 2014; Oladeji et al., 2020; Srinivasan et al., 2007; Van Lith & Ameer, 2016).



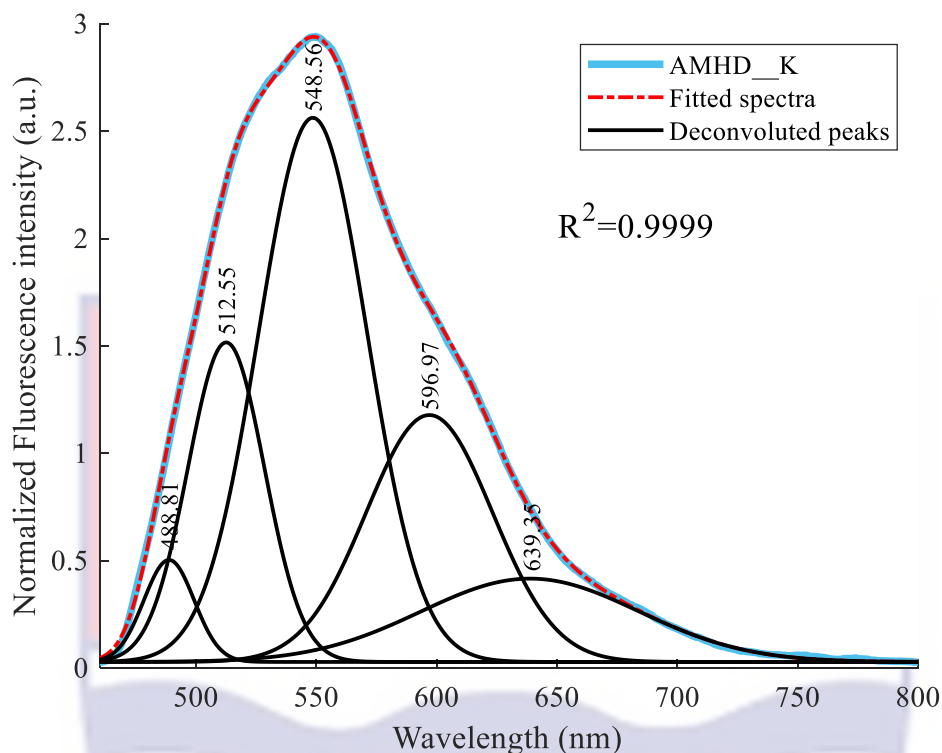


Figure 10: Deconvoluted LIAF spectra of AMHD\_K

Three peaks were observed in the green–yellow region which had wavelength centers at 512.6, 548.57 and 596.98 nm respectively. The wavelength centers may be attributed to the presence of cryptolepine, flavonoids such as anthocyanins, dihydrokaempferol and tannins (Buschmann et al., 2001; Donaldson, 2018, 2020; Harborne, 1973; Lang et al., 1991; Mariz et al., 2017; Roshchina, 2018; Roshchina et al., 2016, 2017). Cryptolepine is a vital derivative of alkaloid and a major extract after the decoction of *Cryptolepis Sanguinolenta* which is proven to be most effective in the treatment of uncomplicated malaria issues (Bugyei et al., 2011; Forkuo et al., 2016; Kirby et al., 1995a; Mariz et al., 2017; Osafo et al., 2017; Osei-Djarbeng et al., 2015; Yao et al., 2012). Flavonoids and tannins also form major constituents after the extraction of *Cryptolepis*

*Sanguinolenta* and their derivatives have been reported to possess a significant antiplasmodial activity (Bankole et al., 2015; Forkuo et al., 2016, 2017; Okokon et al., 2017; Wright et al., 2001).

The last peak was observed in the orange region which had a wavelength center at 638.81 nm. Certain derivatives of the secondary metabolites such as alkaloids, anthocyanins and tannins fluorescence at the yellow–red region, the last peak could therefore be attributed to these secondary metabolites (Donaldson, 2018, 2020; Roshchina et al., 2017). The last peak may also be attributed to kaempferol and sesquiterpenoid (azulene) (Donaldson, 2020; García-plazaola et al., 2015).

Figure 11 show deconvoluted LIAF spectra of AMHD\_A, AMHD\_B, AMHD\_C and AMHD\_G. With the exception of AMHD\_D, all the other AMHDs (AMHD\_A, AMHD\_B, AMHD\_C, AMHD\_E, MHD\_F, AMHD\_G, AMHD\_H, AMHD\_I, and AMHD\_J) had a peak at  $(489.20 \pm 1.85 \text{ nm})$  at the blue region which may be attributed to certain phenolic compounds, terpenoids and alkaloids (Donaldson, 2018, 2020; García-plazaola et al., 2015; Lang et al., 1991; Roshchina, 2018; Roshchina et al., 2016, 2017). The deconvoluted LIAF spectra of AMHD\_D, AMHD\_E, AMHD\_F, AMHD\_H, AMHD\_I and AMHD\_J are shown in appendix B.

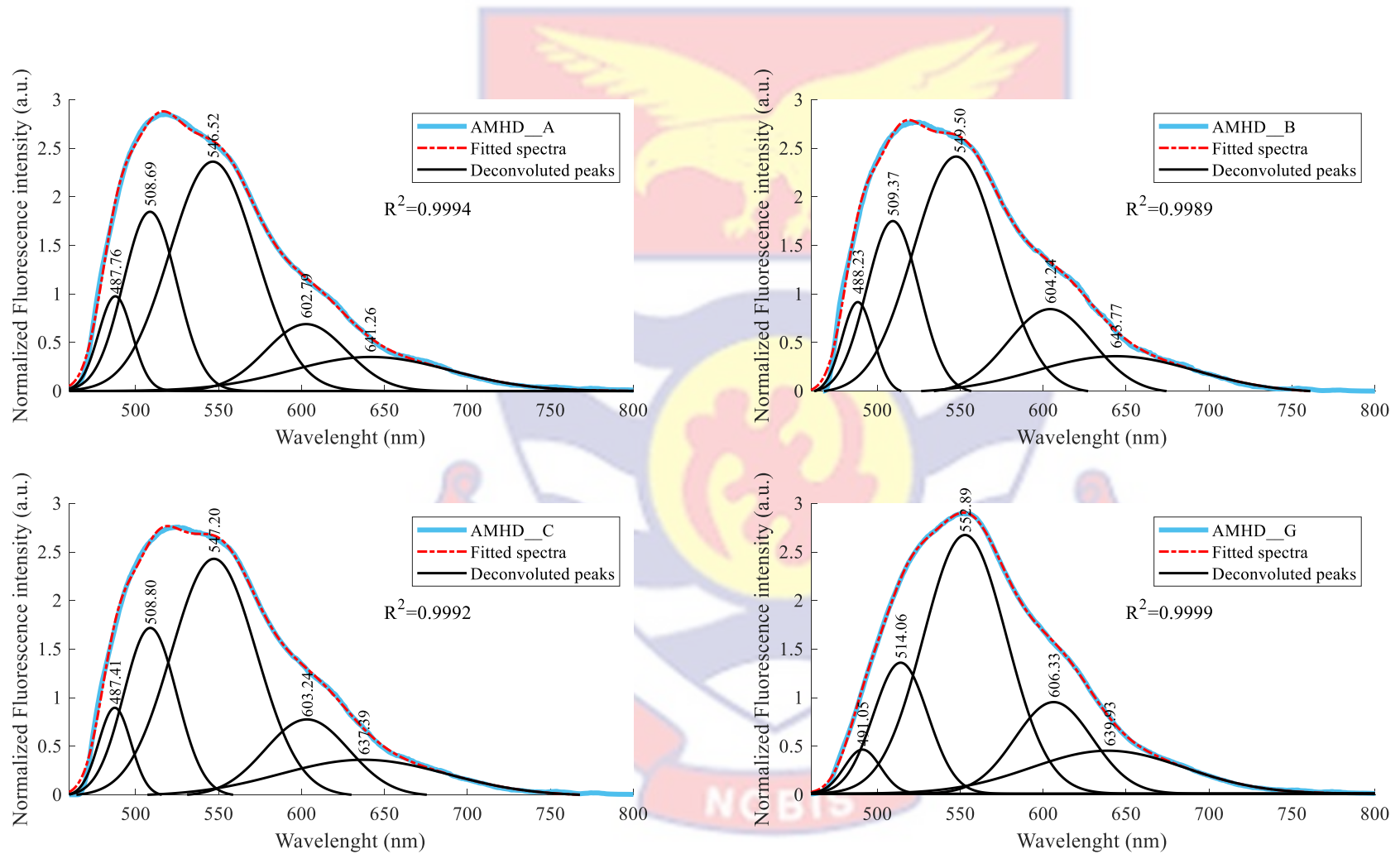


Figure 11: Deconvoluted LIAF spectra of AMHD\_A, AMHD\_B, AMHD\_C and AMHD\_G

Two deconvoluted peaks from the AMHDs: AMHD\_A, AMHD\_B, AMHD\_C, AMHD\_E, AMHD\_F, AMHD\_G, AMHD\_H, AMHD\_J, and AMHD\_I were observed in the green-yellow region and at wavelength center  $511.22 \pm 2.86$  and  $549.14 \pm 3.74$  nm. However, AMHD\_D displayed three deconvoluted peaks at wavelength centers 493.50, 515.04 and 553.65 nm which are in the green-yellow fluorescence region. Deconvoluted fluorescence peaks at the green-yellow region may be attributed to several derivatives of alkaloids, numerous classes of phenolic compounds (tannins, lignin, flavonoid, and coumarins) and carotenoids (Buschmann et al., 2001; Donaldson, 2018, 2020; García-plazaola et al., 2015; Harborne, 1973; Roshchina, 2012, 2018; Roshchina et al., 2016, 2017).

Fluorescence peaks at orange-red region for all AMHDs occurred at  $604.56 \pm 1.77$  and  $657.91 \pm 20.45$  nm respectively. Secondary metabolites such as azulene, anthocyanins and derivatives of tannins, alkaloids and quinones fluorescence at the orange-red region (Donaldson, 2018, 2020; Harborne, 1973; Roshchina, 2018; Roshchina et al., 2016). Generally, the red fluorescence is attributed to the chlorophyll pigments in plants (Buschmann et al., 2001; Donaldson, 2018, 2020; Lopes et al., 2020; Ombinda-lemboumba, 2006). Intensity wise, AMHD\_F and AMHD\_D peaked prominently at red fluorescence indicating a high concentration of the chlorophyll pigments and derivatives of alkaloids and tannins (Ahmad & Saleem, 2018; Donaldson, 2020). A summary of all wavelength centers and their corresponding standard error of the hidden peaks obtained from each AMHDs is shown in Table 4.

**Table 4: Summary of the hidden peaks wavelength center of the individual AMHDs**

AMHDs	Peak 1/nm	Peak 2/ nm	Peak 3/nm	Peak 4/nm	Peak 5/nm
AMHD_K (standard)	488.81 ±0.331	512.56 ±0.236	548.57 ±0.314	596.98 ±1.252	638.81 ±4.990
AMHD_A	487.76 ±0.350	508.68 ±0.631	546.53 ±1.222	602.79 ±3.745	641.82 ±10.592
AMHD_B	488.23 ±0.338	509.37 ±0.603	547.50 ±1.017	604.23 ±3.949	643.77 ±15.859
AMHD_C	487.41 ±0.336	508.80 ±0.570	547.21 ±0.936	603.28 ±3.982	637.46 ±17.185
AMHD_D	493.50 ±0.485	515.04 ±0.932	553.65 ±1.540	613.75 ±2.413	677.14 ±0.556
AMHD_E	487.71 ±0.379	508.87 ±0.654	547.46 ±0.980	605.86 ±2.971	670.24 ±2.799
AMHD_F	487.34 ±0.402	508.36 ±0.775	545.39 ±1.336	605.27 ±5.302	678.36 ±0.423
AMHD_G	491.05 ±0.491	514.07 ±0.352	552.88 ±0.466	606.33 ±1.362	639.02 ±4.796
AMHD_H	488.80 ±0.385	510.10 ±0.592	548.01 ±0.874	605.55 ±2.923	670.93 ±2.513
AMHD_I	487.65 ±0.370	508.57 ±0.642	545.79 ±1.056	603.46 ±3.987	669.65 ±3.767
AMHD_J	487.71 ±0.377	508.78 ±0.658	546.39 ±1.091	605.08 ±3.889	671.20 ±3.325



## Chemometrics of AMHDs

### Unsupervised methods

Chemometric technique called PCA was employed in determining the differences and similarities among the fluorescence spectra data of the AMHDs. The PCA of the AMHDs data set was principally focused on two objectives: first, it was concerned with interpreting the variance-covariance structure of the set of variables, and thus, the intensity variation of the individual AMHDs to be characterized. The second objective was for PCA to disclose the set of variables that set out some inherent spectra structure among the individual AMHDs. These objectives explain the physicochemical properties among the AMHDs. The PCA provided a method and technique that reduces the dimension of the parameter space, and thus, with the same spectra data information on the AMHDs, the large set of variables were transformed into a smaller set of variables. The PCA algorithm using MATLAB returned the outputs: the latent, coefficients, scores, t-squared, explained variables and estimated means of variables.

The algorithm returns the latent in a single column vector which was the calculated eigenvalues of the covariance of the AMHDs in percentage. The resulting percentage eigenvalues of the total variance for the first ten (10) PCs are shown in the scree plot in figure 12. The first two PCs were the most significant in the separation and clustering of AMHDs accounting for more than 97.93 % of the cumulative variability which could be used to classify the AMHDs.

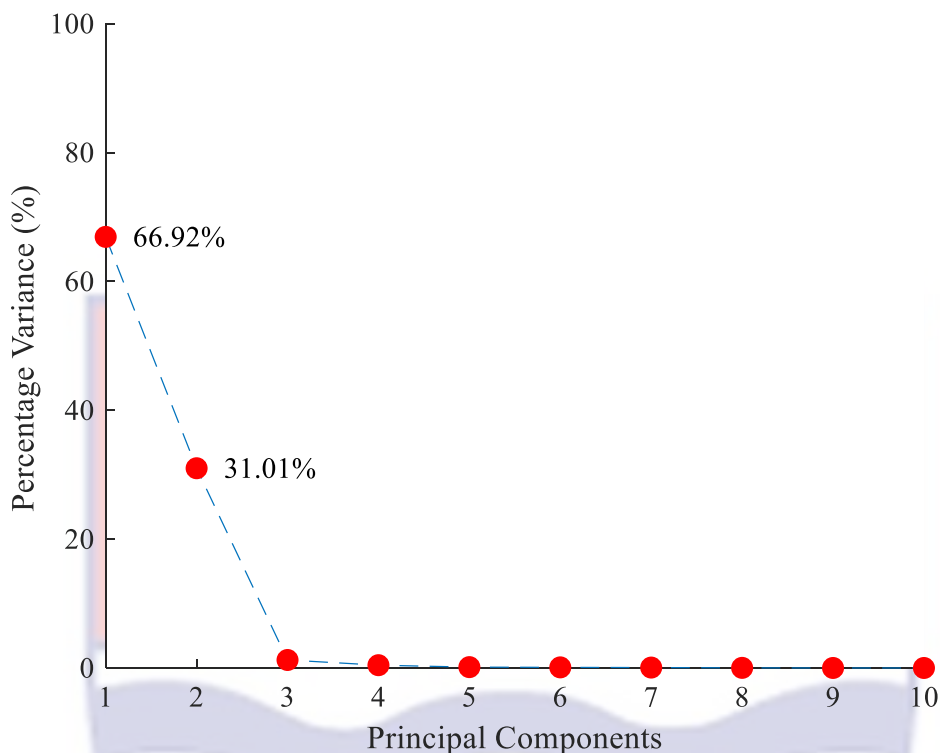


Figure 12: Scree plot showing percentage variance of the first ten (10) PCs of AMHDs

The PC coefficients (PCCs) also known as the PC loadings represent the eigenvectors of the AMHDs data matrix. The return of the PC loadings were arranged in descending order of component variance in a single column. The first two eigenvectors namely Principal Component 1 and 2 (PC1 and PC2) were plotted as a function of wavelength as shown in figure 13. Variables within a wavelength range 460-800 nm were significant in determining the distribution of the AMHDs. Together, the PC1 and PC2 revealed separation and clear distinct difference between the AMHDs. From the PC1, the wavelength center of 493 and 585 nm which were loaded on the positive and negative axis respectively indicate the most significant separation of the AMHDs samples.

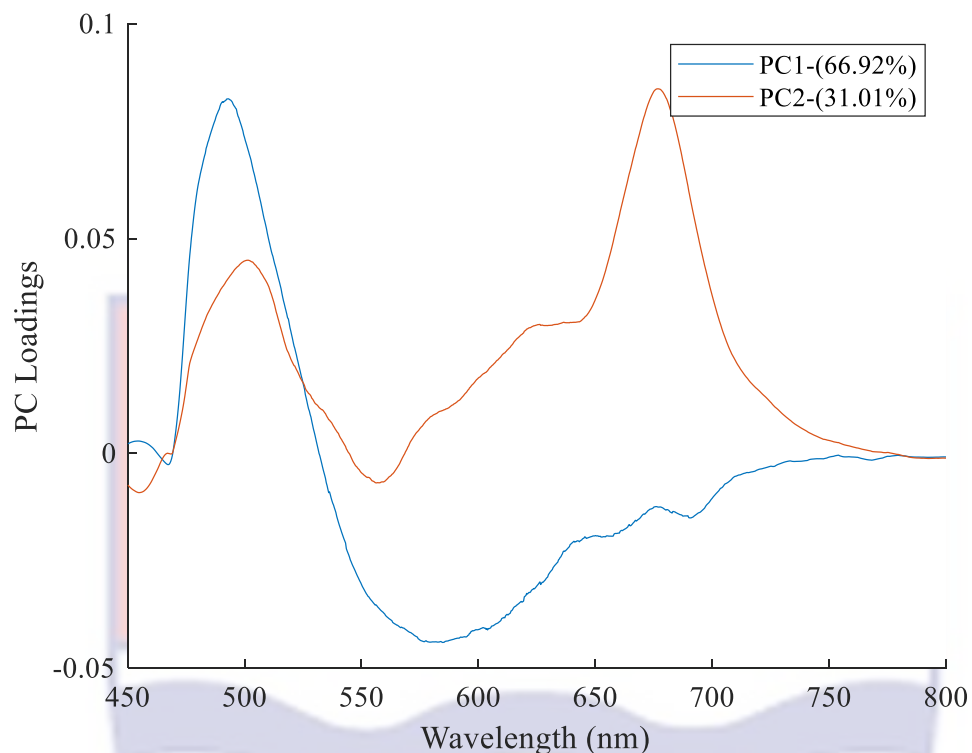


Figure 13: Loadings plot the first two PCs obtained from PCA of the LIAF spectra of AMHDs samples

Similarly, from the PC2, the wavelength center of 501 and 677 nm on the positive axis shown the most significant separation of AMHDs. The variation in chlorophyll content and derivatives of alkaloids and tannins concentration among the AMHDs may be the reason for the significant separation at emission band 677 nm on PC2 axis.

Different plant species have been used in the preparations of the individual AMHDs making them contain substantially different secondary metabolites. To compare the AMHDs according to their physio-chemical properties, a PC score plot was performed by loading the PC scores into the two PCs (PC1 = 66.92% and PC2 = 31.01%) which shows the pattern in the individual AMHDs represented in PC space as shown in figure 14. On the negative side of the PC1

axis, eight (8) AMHDs (AMHD\_A, AMHD\_B, AMHD\_C, AMHD\_D, AMHD\_E, AMHD\_F, AMHD\_I, and AMHD\_J) were closely clustered. However, among these AMHDs, the trend of clustering was within bounds. AMHD\_G, AMHD\_H and AMHD\_K, on the other hand, can be located on the positive side of the PC1 axis with AMHD\_K and AMHD\_G clustering very closely. The clear distinction of AMHD\_H from the rest of the samples was evident. The separation on the PC1 axis was as a result of the spectra signature of the AMHDs (AMHD\_G, AMHD\_H, and AMHD\_K) which tended to depict a high fluorescence in the green-yellow region as compared to the other eight AMHDs depicting a high fluorescence in the blue-green region.

On the PC2 axis, all AMHDs cluster towards the positive side of the PC1 axis. The difference in chlorophyll and derivatives of alkaloids and tannins concentrations may be the reason for their separation along the PC2 axis, thus, AMHD\_D having the highest concentration and AMHD\_K having the least. Comparing all the AMHDs on the score plot, only AMHD\_G clustered best with AMHD\_K (standard AMHD) indicating the common chemical composition among them. Since AMHD G and AMHD\_K (standard AMHD) were prepared from common plant species, they would have a similar chemical composition making them cluster more closely, as shown in the score plot of figure 14.

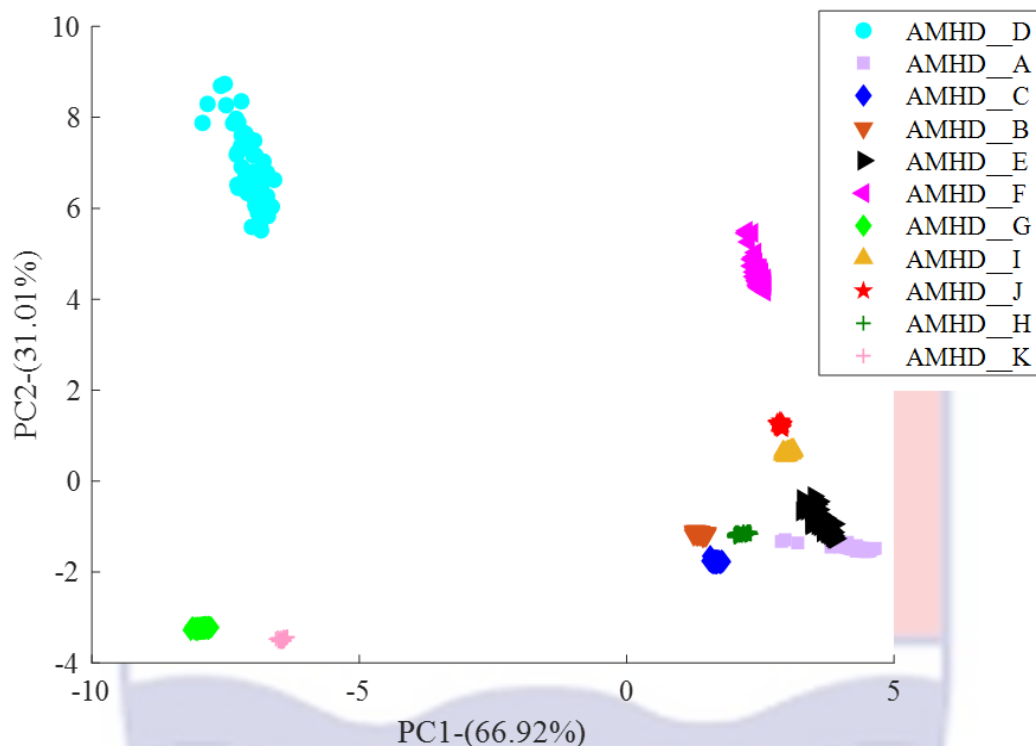


Figure 14: Score plot showing clustering of AMHDs

In order to compare secondary metabolites that may be present among the AMHDs, some secondary metabolites reported by Donaldson (2020), Lang et al. (1991), Mariz et al. (2017), and Roshchina et al. (2016) to be found in medicinal plants and individual AMHDs were analyzed using a PCA Biplot (Donaldson, 2020; Mariz et al., 2017; Roshchina et al., 2016). PCA Biplot is a graphical representation of the PC scores and PC coefficient that is visualized in a low dimensional graph. The PCA Biplot then presents the fluorescence emission wavelength centers of certain secondary metabolites (phenolic compounds, derivatives of alkaloids, riboflavin and anthraquinones) reported to have antiplasmodial activities and the PC score of AMHDs on the same axes to observe their interrelationship.

The biplot (figure 15) was represented in quadrants of PC score with the various secondary metabolites unevenly distributed by a line joining the origin. A phenolic compound (coumarins) and riboflavin could be observed in the first quadrant of the biplot, indicating their high contribution in the AMHD\_F, AMHD\_I and AMHD\_J than in the other AMHDs. Another phenolic compound (flavonoid) and a derivative of quinone (anthraquinones) appeared to be in the second quadrant which correlated with the AMHD\_D. This shows the high contribution of flavonoid and anthraquinones in AMHD\_D than in the other AMHDs.

In the third quadrant, phenolic compound (tannin) and two major derivatives of alkaloids (cryptolepine and quindoline) correlated with AMHD\_G and AMHD\_K. This shows the high contribution of tannin, cryptolepine and quindoline in AMHD\_G and AMHD\_K than in the other AMHDs. Cryptolepine and quindoline in the third quadrant are two important secondary metabolites which have been proved to treat uncomplicated malaria issues. The decoction of *Cryptolepis Sanguinolenta* has been reported to contain cryptolepine and quindoline and some other secondary metabolites (Cimanga et al., 1997; Mariz et al., 2017; Osafo et al., 2017; Shuvalov et al., 2019; Yao et al., 2012). Only one phenolic compound appeared in the last quadrant thus lignin, which correlated with five AMHDs (AMHD\_A, AMHD\_B, AMHD\_C, AMHD\_H and AMHD\_E). This also indicates the high contribution of the lignin in AMHD\_A, AMHD\_B, AMHD\_C, AMHD\_H and AMHD\_E than in the other AMHDs.

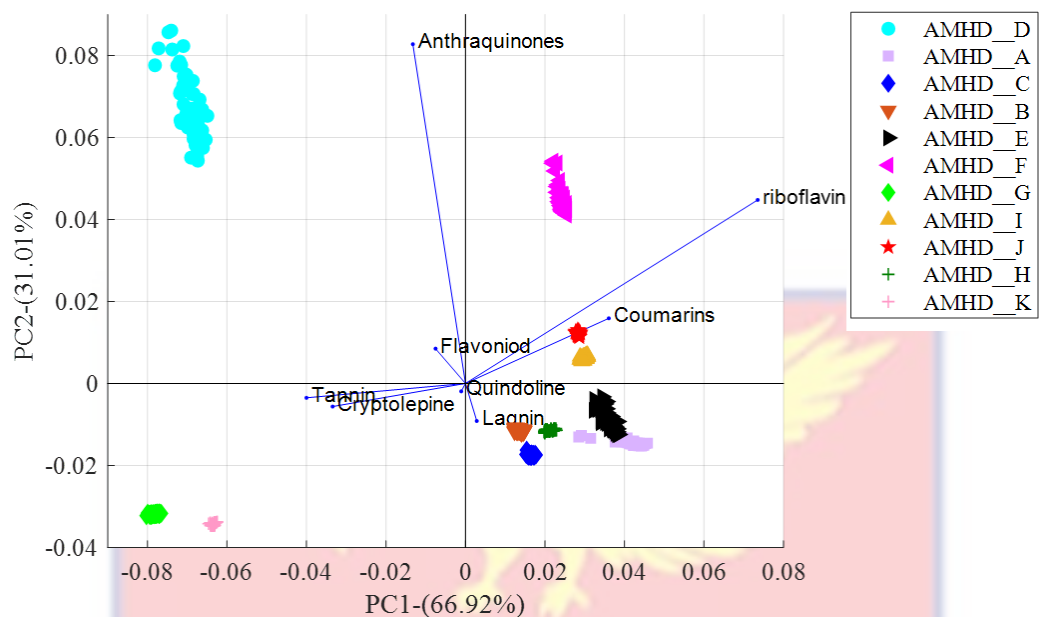


Figure 15: A PCA biplot showing secondary metabolites and PC scores of AMHDs

Secondary metabolites that appeared in the third quadrant of the biplot could be derived from *Cryptolepis Sanguinolenta*. AMHD\_K and AMHD\_G were prepared from only one medicinal plant (*Cryptolepis Sanguinolenta*) which could be confirmed by the biplot as its extracts contributed highly to the samples preparations. Other secondary metabolites may be present but contributed very less in the AMHDs (AMHD\_K and AMHD\_G) decoction. The rest of the AMHDs were prepared from more than two plant species, for which their secondary metabolites are said to be mostly phenolic compounds (flavonoids, coumarins, lignin and tannins) (Neilson et al., 2017b; Okokon et al., 2017; Tay et al., 2011).

A nonhierarchical clustering technique: K-means method was used to group AMHDs into a set of K clusters (number of clusters). In the process of locating suitable centroids for clusters of AMHD data points, K-means clustering

minimizes the data dimension. The individual AMHDs within a cluster could be characterized by the cluster's correlating centroid. As a result, the clustering algorithm divides the data into clusters that share comparable characteristics, such as geometric closeness in the 2D feature space. In the clustering analysis of the AMHDs dataset, the Elbow method was used to estimate the optimal cluster number. It uses the same AMHD dataset with a given continuous cluster number range ( $k = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10$ ) to perform K-means. Then, for each user-specified cluster number  $k$ , the Within Cluster Sum of Squares (WCSS) is computed, and a plot of WCSS values against each cluster number  $k$  was done. WCSS computes the sum squared distance of each point belonging to a certain cluster and the corresponding centroid in the cluster.

The optimal cluster  $k = 3$  was chosen by the fact that from the initial clustering numbers ( $k = 1, 2$ ), there were high values of WCSS but as the clustering number increases the values of WCSS fall off significantly. Between  $k = 1$  and  $k = 2$ , as well as  $k = 2$  and  $k = 3$ , there was a significant change, which was followed by a linear change as the clustering numbers increased making the elbow point to be  $k = 3$ , as indicated in the figure 16.



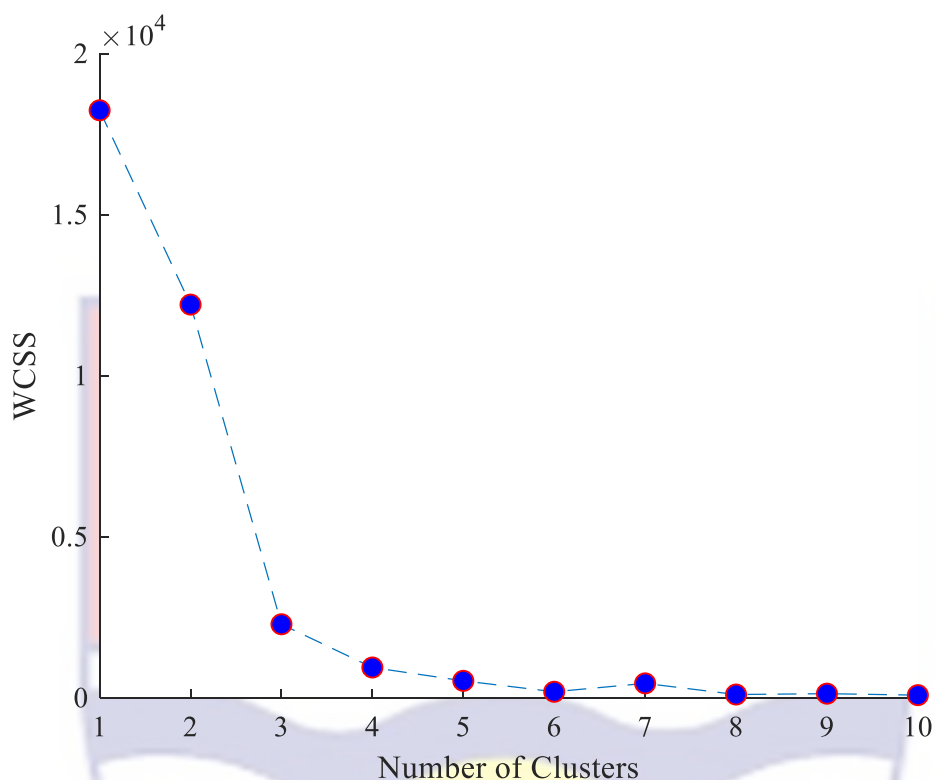


Figure 16: A visual curve display of the performance of K-cluster and the corresponding WCSS values

The k-means++ algorithm was used as a heuristic to find the centroid for k-means clustering in order to implement the 2-means method. The k-means++ algorithm improves the speed of the K-means algorithm while also improving the consistency of the final solution. The Euclidean distance of each AMHD from the centroids point was then computed using the K-means algorithm, and each was reassigned to the nearest group on a grid. Figure 17 shows a plot of three cluster regions: Cluster 1 (AMHD\_A, AMHD\_B, AMHD\_C, AMHD\_E, AMHD\_F, AMHD\_H, AMHD\_I and AMHD\_J), cluster 2 (AMHD\_K and AMHD\_G) and cluster 3 (AMHD\_D). The K-means clustering confirmed the clustering done in the PC score plot indicating the common secondary metabolites that may be

present in AMHD\_G and AMHD\_K. Whereas AMHD\_J, AMHD\_I, AMHD\_E, AMHD\_A, AMHD\_B, AMHD\_C, AMHD\_H and AMHD\_F may also have common secondary metabolites.

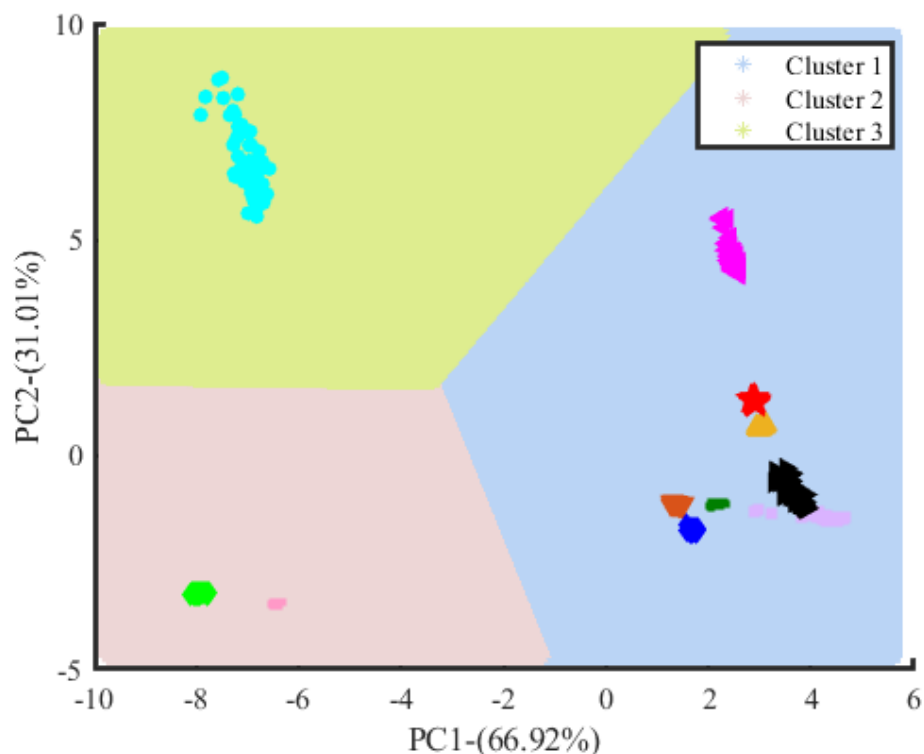


Figure 17: K-means clustering indicating three clusters regions (1, 2 and 3) derived from the fluorescence spectra of the AMHDs

In order to ascertain further the variability in the AMHDs, a hierarchical clustering analysis (HCA) was performed. HCA is an unsupervised pattern technique, which is complementary to the findings obtained from the PCA. As shown in figure 18, the hierarchical clustering results for the AMHDs were graphically represented on the dendrogram. The value of the modified dissimilarity between the combined clusters, which could be viewed as a measure of the separation of paired AMHDs, was labeled by the vertical line along with

the clustering dendrogram. Selecting a distance threshold of eight (8) showed that there were three clusters (1, 2 and 3). In figure 18, cluster 1 consist of AMHD\_J, AMHD\_I, AMHD\_E, AMHD\_A, AMHD\_B, AMHD\_C, AMHD\_H and AMHD\_F, cluster 2 consists of AMHD\_K and AMHD\_G and cluster 3 consist of AMHD\_D. The AMHDs showed a good fit from the figure indicating a significant cophenetic correlation coefficient of 0.9225. The cophenetic correlation coefficient gives a measures of how accurately the dendrogram represents the dissimilarities among the AMHDs. The HCA confirms the clustering done by the K-means clustering indicating the common secondary metabolites in AMHD\_K and ANHD\_G.

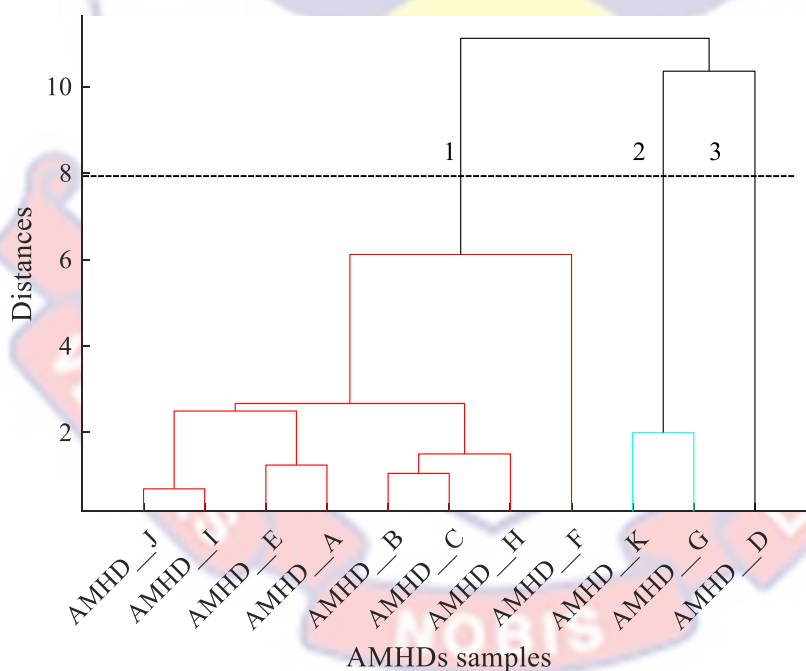


Figure 18: Dendrogram indicating three cluster regions (1, 2 and 3) derived from the fluorescence spectra of the AMHDs

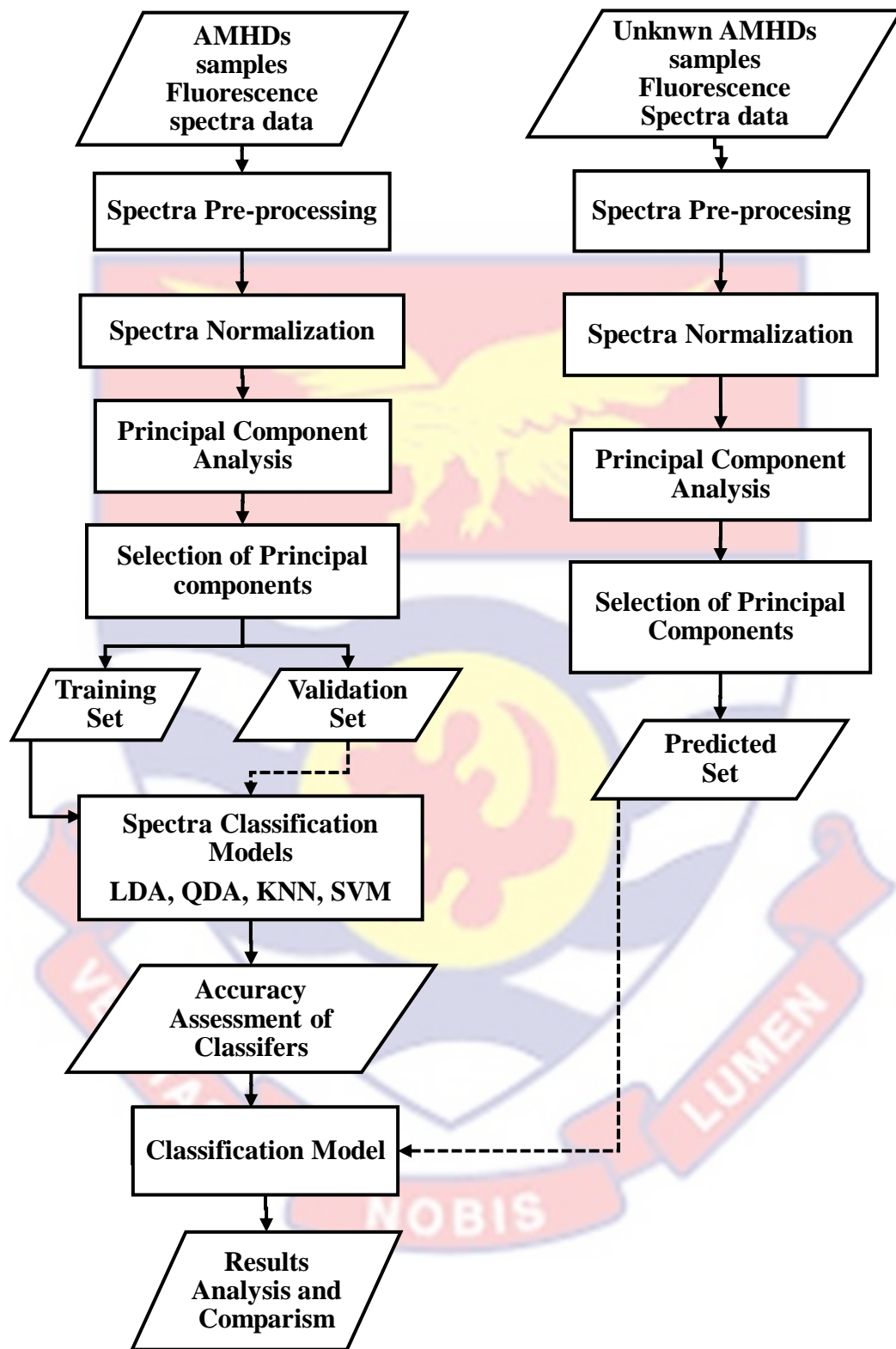


Figure 19: Flow Chart showing the sequence of spectra processing used for classification

### Supervised methods

The results from PCA were able to clearly differentiate the AMHDs, so a model could therefore be created in groups where its data could be used to identify and classify unknown AMHDs as belonging to one of the groups (Adams, 2004; Amuah et al., 2017; Pamp et al., 2020). Therefore, supervised methods of pattern recognition were used: Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Quadratic Discriminant Analysis (QDA) and Linear Discriminant Analysis (LDA), to help create a model in the identification and classification of unknown AMHDs. In figure 19, the complete steps for classifying unknown AMHDs spectra data is summarized by means of a flow diagram.

PCA of the LIAF spectra were used as input data for the model to be generated. Based on the data provided in figure 12, using the two PCs, the prediction set was derived which suggested how unknown AMHDs could be classified. The first two PCs explained more than 97.93 % of the total variance in the LIAF spectra comprising a total of 625 samples. Based on the data provided, using the two PCs, the data set was subdivided into two sets; training set and test set for the classification model. The training set was used in generating a model for the individual classifier and the test set was used in testing the predictability of the model obtained in the training set. MATLAB was used to systematically divide the data set into training and test sets for each sub-group using the Random Sampling algorithm. The Random Sampling algorithm divides the observations into  $k$  disjoint sub-groups (or  $K$ -folds) that are selected at random but are

approximately equal in size. The performance of classifiers based on different K values was measured using the accuracy rate of K-fold cross-validation testing sets (CV accuracy rates) (MacArthur et al., 2020; Morais et al., 2019; Xu et al., 2019). The algorithm randomly chose  $k = 3$ , where KNN, SVM, LDA and QDA models were obtained. Therefore, parameter  $K = 5$  was used in all of the models in the PC data set partition.

The classification models were constructed based on the random parameter. The classification models (KNN, SVM, QDA and LDA) were generated and the stability of the models was measured using the mean classification accuracy rate of 10-fold cross validation sets, and the generalization of the models was evaluated using the classification accuracy rate of the testing set. Table 5 summarizes the classification outcomes for the training and testing sets.

**Table 5: Classification results (%) of the classification models of the AMHDs sample**

Classification model	Training set (502) (%)	Testing set (125) (%)
SVM	100.00	100.00
KNN	100.00	100.00
LDA	99.04	99.04
QDA	99.52	99.52

In both the training and test sets, the results showed that KNN and SVM performed better, with a CV accuracy rate of 100.00 %. Sensitivity, specificity, and precision were also found to be one (1) as shown in Table 6. The model produced was accurate and stable. The good performance in classification model output offers valuable concrete proof on how perfect KNN and SVM could be used to classify and identify unknown potential AMHDs datasets.

**Table 6: Classification performance parameter of models**

Classification model	Sensitivity	Specificity	Precision
SVM	1.00	1.00	1.00
KNN	1.00	1.00	1.00
LDA	1.00	1.00	0.99
QDA	1.00	1.00	0.99

To detail the accuracy of KNN and SVM, a confusion matrix of the dataset were graphically displayed making use of the confusion chart function in MATLAB. Figures 20 and 21 shows the confusion chart for the two classifiers: KNN and SVM respectively of the AMHDs dataset. Each row of the confusion chart indicates the true class (the correct label of the individual AMHD samples created using the training dataset) and each column indicates the predicted class (incorrect label of the individual AMHD samples created using the testing dataset). Out of 20 % of all the AMHD dataset, both KNN and SVM correctly classified the dataset with no misclassified sample.

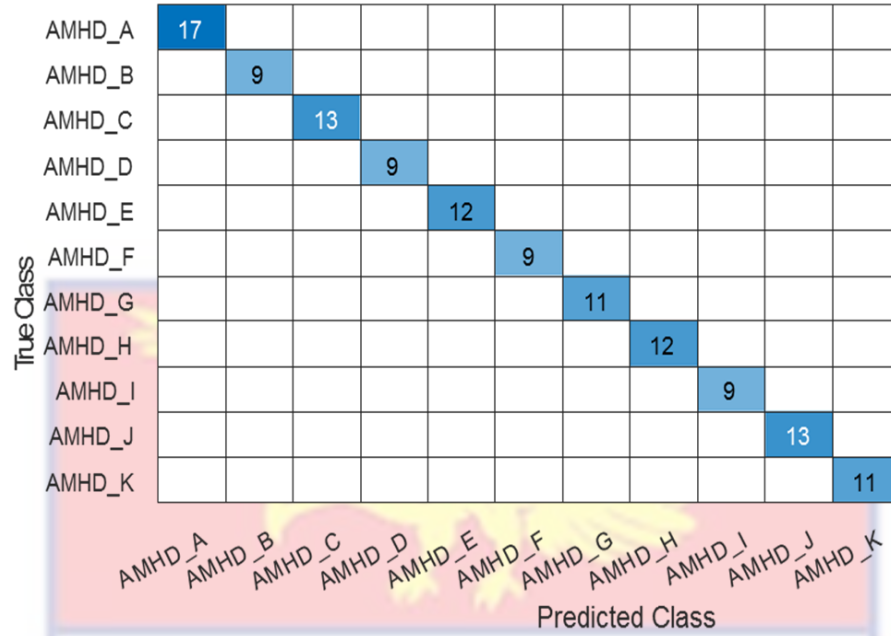


Figure 20: Confusion Matrix for the Prediction set of the AMHDs for the KNN model

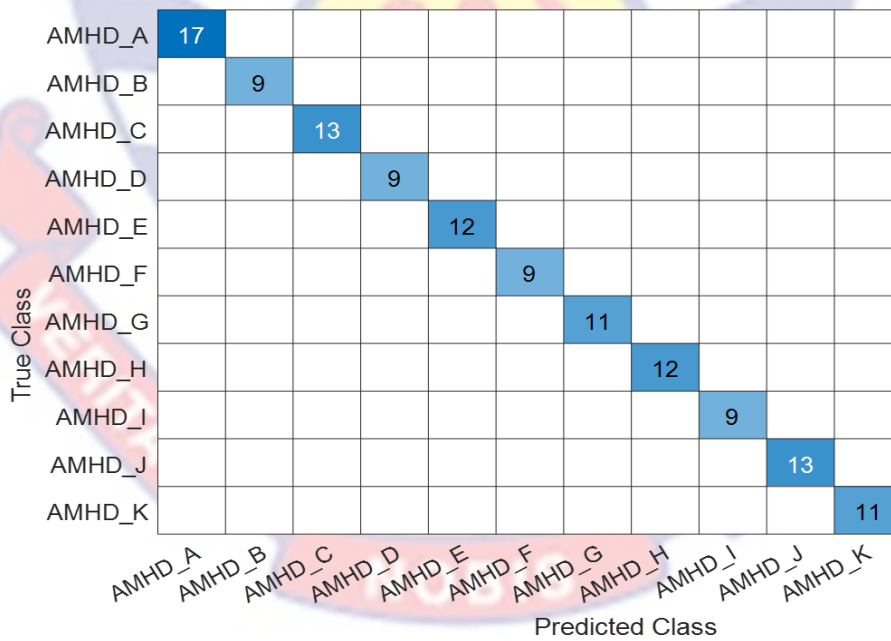
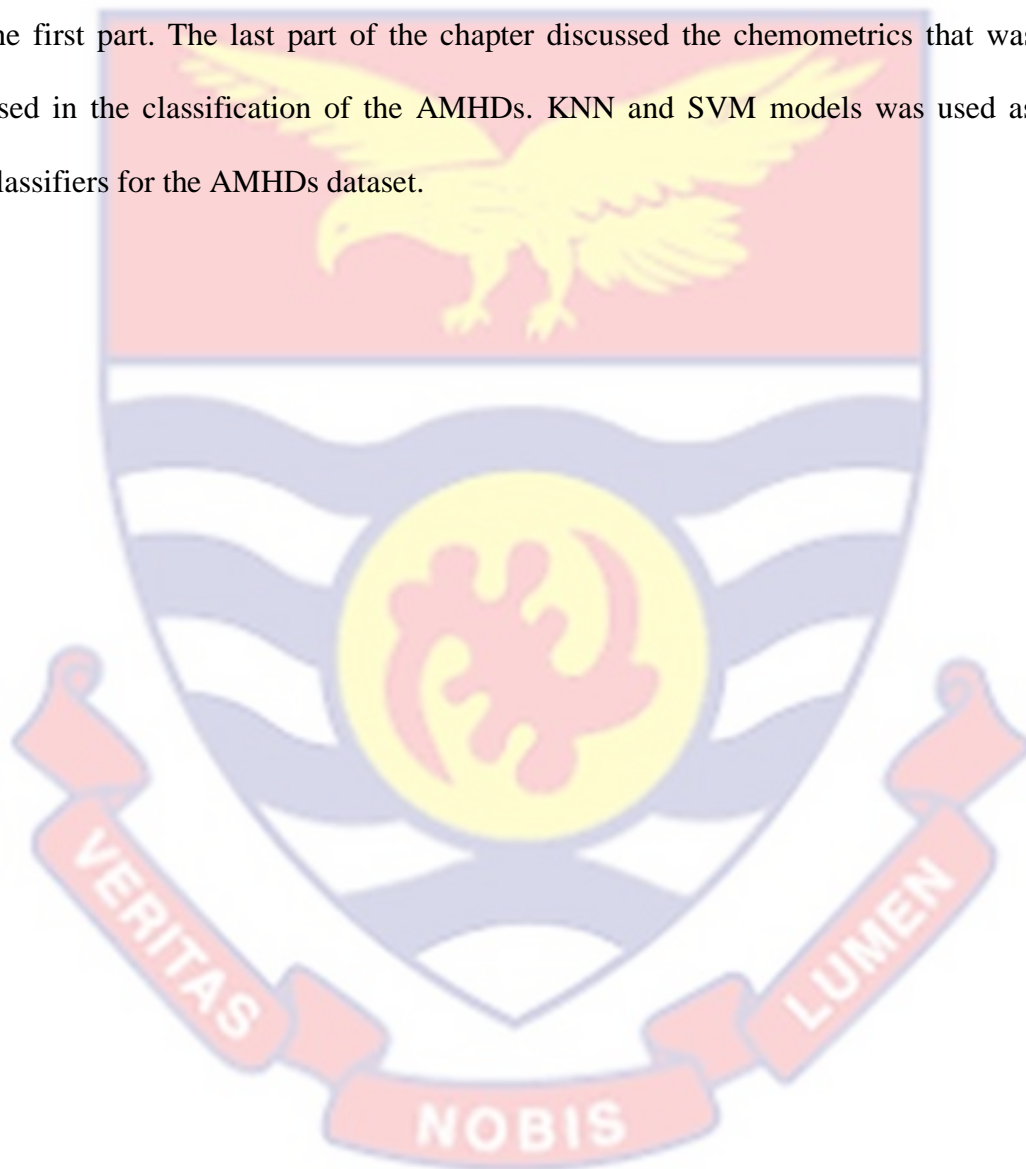


Figure 21: Confusion Matrix for the Prediction set of the AMHDs for the SVM model



### Chapter Summary

This chapter presented the analysis and discussions on results of LIAF spectra for the identification and classification of AMHDs. The spectral fingerprints and the photochemical constituents that may be present were analyzed and discussed in the first part. The last part of the chapter discussed the chemometrics that was used in the classification of the AMHDs. KNN and SVM models was used as classifiers for the AMHDs dataset.



## CHAPTER FIVE

### SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

#### Summary

In summary, eleven (11) commercial anti-malarial herbal drugs (AMHDs) were assessed using laser-induced autofluorescence (LIAF) technique combined with multivariate analysis. The evaluation of AMHDs was done using their fluorescence emission spectra at an excitation wavelength of 405 nm. Results showed varying LIAF spectra of AMHDs. The deconvolution of the standardized fluorescence spectra from each AMHDs explained the secondary metabolites that may be present.

Using Principal Component Analysis (PCA), two Principal Components (PCs) accounted for more than 97.93 % of cumulative variability. The two PCs were plotted to classify the AMHDs. A nonhierarchical clustering technique: K-means method was used to determine the number of clusters of AMHDs. Hierarchical Clustering Analysis (HCA) confirmed the clustering done by the K-means method and the classification of the AMHDs by PCA. Based on the two PCs, a model was generated and an accuracy of 100.00 % was obtained from K-Nearest Neighbor (KNN) and Support Vector Machine (SVM) classifier which could be used to identify unknown AMHDs.

#### Conclusions

Laser-induced autofluorescence (LIAF) technique was used in this study to ascertain the spectral signature of eleven (11) selected AMHDs from Ghana. AMHDs: AMHD\_A, AMHD\_B, AMHD\_C, AMHD\_G, and AMHD\_K exhibited

one major broad emission band within a spectral range of 460 to 800 nm which was characterized in the blue (455-492 nm) to far-red (680-800 nm) fluorescence. LIAF spectra of AMHD\_D, AMHD\_E, AMHD\_F, AMHD\_H, AMHD\_I and AMHD\_J exhibited two emission bands. A major emission band was within 460-650 nm and a minor emission band within 650-750 nm. Intensity-wise, all the fluorescence spectra differ depending on the AMHDs and this was because of the different biophysical properties each AMHD has.

A plot of standardized spectra was able to distinguish the AMHDs based on the structure of the spectral signature of AMHDs which made secondary metabolites equally important to be studied. Eight of the AMHD tend to have almost the same emission peak that runs within a range of 460-650 nm across their fluorescence spectra which depict high fluorescence within a wavelength range of 495-560 nm and can be characterized by the green (490-575 nm) fluorescence. Comparably, three of the AMHDs: AMHD\_D, AMHD\_G and AMHD\_K depicted high fluorescence within a wavelength range of 510–580 nm and was characterized by the fluorescence of green (492-575 nm) and yellow (575-597 nm). The second emission band was within a wavelength range of 650-750 nm which was characterized by the red to far-red (620-800 nm) fluorescence and appeared to peak more in AMHD\_D and AMHD\_F as compared to the other AMHDs.

When deconvoluted five hidden peaks were seen in the LIAF spectra of each AMHD. The constituents' quindoline, lignin and ferulate were attributed to hidden peaks in the blue fluorescence (Buschmann et al., 2001; Donaldson, 2018,

2020; Lang et al., 1991). Quindoline has been reported as an alkaloid derivative compound found in *Cryptolepis Sanguinolenta* and proven to possess antiplasmodial activity (Mariz et al., 2017; Odoh & Akwuaka, 2012; Shuvalov et al., 2019; Yao et al., 2012). The hidden peaks in the green-yellow fluorescence were attributed to compounds cryptolepine, anthocyanins, flavonoids and tannins. Cryptolepine, flavonoids and tannins are major phytochemical constituents in *Cryptolepis Sanguinolenta* which are proven to be most effective in the treatment of uncomplicated malaria issues (Forkuo et al., 2017; Kirby et al., 1995b; Mariz et al., 2017; Osei-Djarbeng et al., 2015; Roshchina, 2018; Roshchina et al., 2016, 2017; Schlechter et al., 2017). The hidden peaks at orange-red region were attributed to azulene, anthocyanins and derivatives of tannins, alkaloids and quinones (Donaldson, 2018, 2020; García-plazaola et al., 2015; Roshchina et al., 2017).

Three unsupervised pattern recognition methods were used to classify AMHDs based on their physio-chemical properties: PCA, K-mean Clustering, and HCA. In this analysis, the AMHD\_K was used as a standard against which all AMHDs samples were assessed due to the medicinal plant (*Cryptolepis Sanguinolenta*) used in its preparation. PCA Biplot was graphically presented to compare the presence of some secondary metabolites reported by Donaldson (2020), Lang et al. (1991), Mariz et al. (2017), and Roshchina et al. (2016) to be found in medicinal plants and individual AMHDs. Secondary metabolites (cryptolepine, tannin and quindoline) contributed very high in AMHD\_K and AMHD\_G as compared to other AMHDs. In the PCA and K-mean method, three

(3) clusters were observed, Cluster 1 consist of AMHD\_A, AMHD\_B, AMHD\_C, AMHD\_E, AMHD\_F, AMHD\_H AMHD\_I, and AMHD\_J; Cluster 2 consists of AMHD\_K and AMHD\_G and Cluster 3 also consist of AMHD\_D. PCA and K-mean method was therefore able to confirm the classification of AMHDs based on their physio-chemical constituent. HCA confirmed the clustering done by PCA and K-means method. The classification showed that AMHD\_G clustered most with AMHD\_K (used as standard) because the two samples were prepared from the same medicinal plant (*Cryptolepis Sanguinolenta*) resulting in common physio-chemical properties among them.

A supervised pattern recognition method: Supervised Vector Machine (SVM), K-Nearest Neighbor (KNN), Quadratic Discriminant Analysis (QDA), and Linear Discriminant Analysis (LDA) were used to generate the model to identify and classify unknown AMHDs. The accuracy results of the classifiers from PC data was 99.04 % for LDA while that of QDA was 99.52 %, KNN was 100 %, and SVM was 100 %. Accuracy results obtained from all the classifiers indicating all the supervised methods can be used to identify and classify unknown AMHDs. Classifier KNN and SVM, showed 100% accuracy, indicating classifiers' ability to predict perfectly unknown AMHDs.

### **Recommendations**

It is recommended that, in the scope of research in AMHDs in Ghana, different optical and chemometrics techniques could be employed. Food and Drug Authority (FDA) could adopt Laser-Induced Autofluorescence (LIAF) technique as a preliminary test technique before any quality assessment of AMHDs.

AMHDs manufacturing industries may adopt the LIAF technique to assess the consistency, substandard and counterfeit of their manufactured drugs. The LIAF technique could also be applied in the pharmaceutical industry in examining phytochemical constituents present when exposed to ultra-violet (UV) radiation.



## REFERENCES

- Adams, M. J. (2004). Chemometrics in analytical spectroscopy. In B. W. Neil (Ed.), *Analytica Chimica Acta* (2nd ed). The Royal Society of Chemistry. [https://doi.org/10.1016/s0003-2670\(97\)89584-0](https://doi.org/10.1016/s0003-2670(97)89584-0)
- Adeniyi, A., Asase, A., Ekpe, P. K., Asitoakor, B. K., Adu-Gyamfi, A., & Awekor, P. Y. (2018). Ethnobotanical study of medicinal plants from Ghana; confirmation of ethnobotanical uses, and review of biological and toxicological studies on medicinal plants used in Apra Hills Sacred Grove. *Journal of Herbal Medicine*, 14(October 2016), 76–87. <https://doi.org/10.1016/j.hermed.2018.02.001>
- Adusei-Mensah, F., Tikkanen-Kaukanen, C., Kauhanen, J., Henneh, I. T., Owusu Agyei, P. E., Akakpo, P. K., & Ekor, M. (2020). Erratum: Sub-chronic toxicity evaluation of top-three commercial herbal anti-malarial preparations in the Kumasi metropolis, Ghana (Bioscience Reports (2020) 40 (BSR-20192536) DOI: 10.1042/BSR20192536). *Bioscience Reports*, 40(7), 1–16. [https://doi.org/10.1042/BSR-20192536\\_COR](https://doi.org/10.1042/BSR-20192536_COR)
- Agatonovic-Kustrin, S., & Morton, D. W. (2018). The Current and Potential Therapeutic Uses of Parthenolide. In *Studies in Natural Products Chemistry* Vol. 58, pp. 61–91. Elsevier B.V. <https://doi.org/10.1016/B978-0-444-64056-7.00003-9>
- Ahmad, N., & Saleem, M. (2018). *Studying heating effects on desi ghee obtained from buffalo milk using fluorescence spectroscopy*. 1–17.

- Akram, S., Najam, R., Rizwani, G. H., & Abbas, S. A. (2015). Determination of heavy metal contents by atomic absorption spectroscopy (AAS) in some medicinal plants from Pakistani and Malaysian origin. In *Pak. J. Pharm. Sci* Vol. 28, Issue 5.
- Allen, C., Welford, W. T., & Oster, G. (1971). Physical Techniques In Biological Research. In G. Oster (Ed.), *Optical Techniques* (Second). Academic Press.
- Amuah, C. L. Y., Eghan, J. M., Anderson, B., Adueming, P. O. W., & Opoku-Ansah, J. (2017). Laser Induced Fluorescence in combination with Multivariate analysis classifies anti-malarial herbal plants. *Frontiers in Optics , OSA*.
- Amuah, L. Y., Eghan, J. M., Anderson, B., Adueming, P. O. W., & Opoku-Ansah, J. (2017). Laser induced fluorescence in combination with multivariate analysis classifies anti-malarial herbal plants. *Optics InfoBase Conference Papers, Part F66-F*. <https://doi.org/10.1364/FIO.2017.JTu2A.71>
- Andrade-Garda, J. M. (2009). *Basic Chemometric Techniques in Atomic Spectroscopy Series Editor* : (N. W. Barnett (ed.)). RSCPublishing.
- Ansah, J. O., Eghan, M. J., Anderson, B., & Boampong, J. N. (2014). Wavelength Markers for Malaria (*Plasmodium Falciparum*) Infected and Uninfected Red Blood Cells for Ring and Trophozoite Stages. *Applied Physics Research*, 6(2), 47. <https://doi.org/10.5539/apr.v6n2p47>
- Appiah, K., Opong, C., Mardani, H., Omari, R., Kpabitey, S., Amoatey, C., Onwona-Agyeman, S., Oikawa, Y., Katsura, K., & Fujii, Y. (2018).



Medicinal Plants Used in the Ejisu-Juaben Municipality, Southern Ghana:  
An Ethnobotanical Study. *Medicines*, 6(1), 1. <https://doi.org/10.3390/medicines6010001>

Asanga, E., State, C. R., Eseyin, O. A., Ebong, P., & Igile, G. O. (2017).

*Antiplasmodial Activity of Ethanol Extract and Fractions of Nauclea Latifolia Smith ( Rubiaceae ) Roots Antiplasmodial Activity of Ethanol Extract and Fractions of Nauclea Latifolia Smith ( Rubiaceae ) Roots. April.*

Bankole, A. E., Adekunle, A. A., & Sowemimo, A. A. (2015). *Phytochemical*

*screening and in vivo antimalarial activity of extracts from three medicinal plants used in malaria treatment in Nigeria.* <https://doi.org/10.1007/s00436-015-4747-x>

Baudelet, M. (n.d.). *Laser spectroscopy for sensing: Fundamentals, techniques and applications.*

Bello, O. M., Zaki, A. A., Khan, S. I., Fasinu, P. S., Ali, Z., Khan, I. A., Usman,

L. A., & Oguntoye, O. S. (2017). South African Journal of Botany  
Assessment of selected medicinal plants indigenous to West Africa for antiprotozoal activity. *South African Journal of Botany*, 113, 200–211. <https://doi.org/10.1016/j.sajb.2017.08.002>

Berg, R. A. Van Den, Hoefsloot, H. C. J., Westerhuis, J. A., Smilde, A. K., &

Werf, M. J. Van Der. (2006). *Centering , scaling , and transformations : improving the biological information content of metabolomics data.* 15, 1–15. <https://doi.org/10.1186/1471-2164-7-142>

- Biswas, P., Karn, A. K., Balasubramanian, P., & Kale, P. G. (2017). Biosensor for detection of dissolved chromium in potable water: A review. In *Biosensors and Bioelectronics*, Vol. 94, pp. 589–604. Elsevier Ltd. <https://doi.org/10.1016/j.bios.2017.03.043>
- Biswas, P., Karn, A. K., Pares, G., & Balasubramanian, P. (2017). Author's Accepted Manuscript. *Biosensors and Bioelectronics*. <https://doi.org/10.1016/j.bios.2017.03.043>
- Bolshov, M. A., Kuritsyn, Y. A., & Romanovskii, Y. V. (2015). Tunable diode laser spectroscopy as a technique for combustion diagnostics. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 106, 45–66. <https://doi.org/10.1016/j.sab.2015.01.010>
- Bordin-Aykroyd, S., Dias, R. B., & Lynch, E. (2019). *EC Dental Science Review Article Laser-Tissue Interaction*.
- Brauer, A. (2015). Laser-Induced Fluorescence (LIF) and Phosphorescence (LIP) Techniques. In *Supercritical Fluid Science and Technology*, Vol. 7, pp. 313–345. Elsevier B.V. <https://doi.org/10.1016/B978-0-444-63422-1.00005-5>
- Brinck, T., Bittererova, M., & Östmark, H. (2003). Electronic structure calculations as a tool in the quest for experimental verification of N<sub>4</sub>. In *Theoretical and Computational Chemistry* (Vol. 12, pp. 421–439). Elsevier. [https://doi.org/10.1016/s1380-7323\(03\)80017-1](https://doi.org/10.1016/s1380-7323(03)80017-1)
- Bugyei, K., Boye, G., & Addy, M. (2011). Clinical efficacy of a tea-bag formulation of *Cryptolepis sanguinolenta* root in the treatment of acute

uncomplicated falciparum malaria. *Ghana Medical Journal*, 44(1).  
<https://doi.org/10.4314/gmj.v44i1.68849>

Buschmann, C., Langsdorf, G., & Lichtenthaler, H. K. (2001). Imaging of the blue, green, and red fluorescence emission of plants: An overview. In *Photosynthetica* (Vol. 38, Issue 4, pp. 483–491). <https://doi.org/10.1023/A:1012440903014>

Butcher, D. J. (2005). *Interferences and Background Correction Types of Interferences for ETA-AAS*. 157–163.

Calatayud, J. M., & Icardo, M. C. (2004). Flow Injection Analysis - Clinical and Pharmaceutical Applications. In *Encyclopedia of Analytical Science: Second Edition* (pp. 76–89). Elsevier Inc. <https://doi.org/10.1016/B0-12-369397-7/00159-X>

Campbell, N. R., & Ingram, J. C. (2014). Characterization of 234U/238U Activity Ratios and Potential Inorganic Uranium Complexation Species in Unregulated Water Sources in the Southwest Region of the Navajo Reservation. In *Water Reclamation and Sustainability* (pp. 77–94). Elsevier Inc. <https://doi.org/10.1016/B978-0-12-411645-0.00004-3>

Canizo, B. V., Escudero, L. B., Pellerano, R. G., & Wuilloud, R. G. (2019). Quality monitoring and authenticity assessment of wines: Analytical and chemometric methods. In *Quality Control in the Beverage Industry: Volume 17: The Science of Beverages* (pp. 335–384). Elsevier. <https://doi.org/10.1016/B978-0-12-816681-9.00010-2>

- Chen, X., Zheng, B., & Liu, H. (2011). Optical and digital microscopic imaging techniques and applications in pathology. In *Analytical Cellular Pathology* (Vol. 34, Issues 1–2, pp. 5–18). NIH Public Access. <https://doi.org/10.3233/ACP-2011-0006>
- Chirayil, C. J., Abraham, J., Mishra, R. K., George, S. C., & Thomas, S. (2017). Instrumental Techniques for the Characterization of Nanoparticles. In *Thermal and Rheological Measurement Techniques for Nanomaterials Characterization* (Vol. 3, pp. 1–36). Elsevier. <https://doi.org/10.1016/B978-0-323-46139-9.00001-3>
- Cimanga, K., De Bruyne, T., Pieters, L., Vlietinck, A. J., & Turger, C. A. (1997). In vitro and in vivo antiplasmodial activity of cryptolepine and related alkaloids from *Cryptolepis sanguinolenta*. *Journal of Natural Products*, *60*(7), 688–691.
- Cox-Georgian, D., Ramadoss, N., Dona, C., & Basu, C. (2019). Therapeutic and medicinal uses of terpenes. In *Medicinal Plants: From Farm to Pharmacy* (pp. 333–359). Springer International Publishing. [https://doi.org/10.1007/978-3-030-31269-5\\_15](https://doi.org/10.1007/978-3-030-31269-5_15)
- Crosley, D. R. (1982). Laser-induced fluorescence in spectroscopy, dynamics, and diagnostics. *Journal of Chemical Education*, *59*(6), 446–455. <https://doi.org/10.1021/ed059p446>
- Da Silva, J. M., & Utkin, A. B. (2018). Application of laser-induced fluorescence in functional studies of photosynthetic biofilms. *Processes*, *6*(11). <https://doi.org/10.3390/pr6110227>

- Dey, A., Nanda, B., Nandy, S., Mukherjee, A., & Pandey, D. K. (2021). Implications of phytochemicals as disease-modifying agents against Huntington's disease (HD): Bioactivity, animal models and transgenics, synergism and structure–activity studies. In *Studies in Natural Products Chemistry* (Vol. 67, pp. 27–79). Elsevier B.V. <https://doi.org/10.1016/B978-0-12-819483-6.00002-3>
- Donaldson, L. (2018). *Imaging and Spectroscopy of Natural Fluorophores in Pine Needles*. 1–16. <https://doi.org/10.3390/plants7010010>
- Donaldson, L. (2020). *Autofluorescence in Plants*.
- Edwards, H. G. M. (2005). Modern Raman spectroscopy—a practical approach. In *Journal of Raman Spectroscopy* (Vol. 36, Issue 8). <https://doi.org/10.1002/jrs.1320>
- Eyring, M. B., & Martin, P. (2013). Spectroscopy in Forensic Science. In *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering*. Elsevier. <https://doi.org/10.1016/b978-0-12-409547-2.05455-x>
- Febir, L. G., Asante, K. P., Afari-Asiedu, S., Abokyi, L. N., Kwarteng, A., Ogotu, B., Gyapong, M., & Owusu-Agyei, S. (2016). Seeking treatment for uncomplicated malaria: Experiences from the Kintampo districts of Ghana. *Malaria Journal*, 15(1), 108. <https://doi.org/10.1186/s12936-016-1151-7>

- Forkuo, A. D., Ansah, C., Boadu, K. M., Boampong, J. N., Ameyaw, E. O., Gyan, B. A., & Ofori, M. F. (2016). Synergistic anti-malarial action of cryptolepine and artemisinin. *Malaria Journal*, *15*(1), 89.
- Forkuo, A. D., Ansah, C., Mensah, K. B., Annan, K., Gyan, B., Theron, A., & Wright, C. W. (2017). In vitro anti-malarial interaction and gametocytocidal activity of cryptolepine. *Malaria Journal*, *16*(1), 1–9.
- Forkuo, Arnold Donkor, Ansah, C., Mensah, K. B., Annan, K., Gyan, B., Theron, A., Mancama, D., & Wright, C. W. (2017). In vitro anti - malarial interaction and gametocytocidal activity of cryptolepine. *Malaria Journal*, 1–9. <https://doi.org/10.1186/s12936-017-2142-z>
- Frank Adusei-Mensah, Kauhanen, J., Haaranen, A., & Henneh, I. T. (2019). Post-Market Safety and Efficacy Surveillance of Herbal Medicinal Products from Users Perspective: A Qualitative Semi-Structured Interview Study in Kumasi, Ghana. *International Journal of Pharmaceutics & Pharmacology* *3*(1), 1–11. <https://doi.org/10.31531/2581-3080.1000136>
- Gaertn, L. (2017). *Pharmacognostic Studies And Antimicrobial Activity of Synedrella Nodiflora ( L.) Gaertn. July 2011.*
- García-plazaola, J. I., Fernández-marín, B., Duke, S. O., Hernández, A., López-arbeloa, F., & María, J. (2015). Plant Science Autofluorescence : Biological functions and technical applications. *Plant Science*, *236*, 136–145. <https://doi.org/10.1016/j.plantsci.2015.03.010>

- Giridhar, G., Manepalli, R. R. K. N., & Apparao, G. (2017). Confocal Raman Spectroscopy. In *Spectroscopic Methods for Nanomaterials Characterization* (Vol. 2). Elsevier Inc. <https://doi.org/10.1016/B978-0-323-46140-5.00007-8>
- Gomes, A. R., & Rocha-Santos, T. A. P. (2019). Bioassays | Enzyme Assays. In *Encyclopedia of Analytical Science* (pp. 271–277). Elsevier. <https://doi.org/10.1016/B978-0-12-409547-2.14331-8>
- Goodman, L. A., Kruskal, W., Rabe-Hesketh, S., & Everitt, B. (1979). Multivariate Analysis: Overview ordinal data. *Journal of the Royal Statistical Society, Series B*, 54, 109–151.
- Grimnes, S., & Martinsen, Ø. G. (2015). Data and Models. In *Bioimpedance and Bioelectricity Basics* (pp. 329–404). Elsevier. <https://doi.org/10.1016/b978-0-12-411470-8.00009-x>
- Gruber, M. (1956). Physical techniques in biological research optical techniques. In *Biochimica et Biophysica Acta* (Vol. 21, Issue 3). [https://doi.org/10.1016/0006-3002\(56\)90209-8](https://doi.org/10.1016/0006-3002(56)90209-8)
- Guo, S., Bocklitz, T., & Popp, J. (2016). Optimization of Raman-spectrum baseline correction in biological application. *Analyst*, 141(8), 2396–2404. <https://doi.org/10.1039/c6an00041j>
- Halina, A. (2005). Introduction to Laser Spectroscopy. *Introduction to Laser Spectroscopy*. <https://doi.org/10.1016/B978-0-444-51662-6.X5000-X>

- Hancock, T., & Smyth, C. (2009). Tree-Based Clustering and Extensions. *Comprehensive Chemometrics*, 2, 683–705. <https://doi.org/10.1016/B978-044452701-1.00069-7>
- Hanrahan, G., Casey, H., & Worsfold, P. J. (2004). Water Analysis - Freshwater. In *Encyclopedia of Analytical Science: Second Edition* (pp. 262–268). Elsevier Inc. <https://doi.org/10.1016/B0-12-369397-7/00655-5>
- Harborne, J. B. (1973). *Phytochemical methods*. CHAPMAN AND HALL. <https://doi.org/10.1007/978-94-009-5921-7>
- Haruna, C. R., Hou, M., Eghan, M. J., Kpiebaareh, M. Y., Tandoh, L., Eghan-Yartel, B., & Asante-Mensah, M. G. (2019). Cost-Based and Effective Human-Machine Based Data Deduplication Model in Entity Reconciliation. *2018 5th International Conference on Systems and Informatics, ICSAI 2018*, 1265–1270. <https://doi.org/10.1109/ICSAI.2018.8599375>
- Heinrich, M., Mah, J., & Amirkia, V. (2021). Alkaloids Used as Medicines: Structural Phytochemistry Meets Biodiversity-An Update and Forward Look. In *Molecules (Basel, Switzerland)* (Vol. 26, Issue 7). NLM (Medline). <https://doi.org/10.3390/molecules26071836>
- Hernández-Rodríguez, P., Baquero, L. P., & Larrota, H. R. (2018). Flavonoids: Potential Therapeutic Agents by Their Antioxidant Capacity. In *Bioactive Compounds: Health Benefits and Potential Applications*. Elsevier Inc. <https://doi.org/10.1016/B978-0-12-814774-0.00014-1>



- Hill, S. J., & Fisher, A. S. (2017). Atomic Absorption , Methods and Instrumentation. In *Encyclopedia of Spectroscopy and Spectrometry* (3rd ed., Issue 1998). Elsevier Ltd. <https://doi.org/10.1016/B978-0-12-803224-4.00099-6>
- Holak, W. (1974). Atomic Absorption in Food Analysis — Special Techniques for Traces of Heavy Metals. In *Methods in Radioimmunoassay, Toxicology, and Related Areas* (pp. 157–178). Springer US. [https://doi.org/10.1007/978-1-4684-3321-0\\_9](https://doi.org/10.1007/978-1-4684-3321-0_9)
- Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(6), 417–441. <https://doi.org/10.1037/h0071325>
- Hussain, C. M., & Keçili, R. (2020a). Spectroscopic techniques for environmental analysis. In *Modern Environmental Analysis Techniques for Pollutants* (pp. 133–161). Elsevier. <https://doi.org/10.1016/b978-0-12-816934-6.00006-0>
- Hussain, C. M., & Keçili, R. (2020b). Spectroscopic techniques for environmental analysis. In *Modern Environmental Analysis Techniques for Pollutants* (pp. 133–161). <https://doi.org/10.1016/B978-0-12-816934-6.00006-0>
- Huzortey, A. A., Anderson, B., & Owusu, A. (2017). A Composite Algorithm for Optimized Baseline Correction in Raman A Composite Algorithm for Optimized Baseline Correction in Raman Spectroscopy. *May 2018*, 10–12. <https://doi.org/10.1364/FIO.2017.FM3C.2>

- Islam, M. S., Matsuki, N., Nagasaka, R., Ohara, K., Takamitsu Hosoya, Ozaki, H., Ushio, H., & Hori, M. (2014). Rice Bran Antioxidants in Health and Wellness. In *Wheat and Rice in Disease Prevention and Health* (pp. 443–451). Elsevier Inc. <https://doi.org/10.1016/B978-0-12-401716-0.00034-9>
- Jaffrennou, C., Stephan, L., Giamarchi, P., Cabon, J. Y., Burel-Deschamps, L., & Bautin, F. (2007). Direct fluorescence monitoring of coal organic matter released in seawater. *Journal of Fluorescence*, *17*(5), 564–572. <https://doi.org/10.1007/s10895-007-0216-y>
- John, N., & George, S. (2017). Raman Spectroscopy. In *Spectroscopic Methods for Nanomaterials Characterization* (Vol. 2). Elsevier Inc. <https://doi.org/10.1016/B978-0-323-46140-5.00005-4>
- Johnson, R. A., & Wichern, D. W. (2007). *Applied Multivariate Statistics Analysis* (6th ed.). Upper Saddle River.
- Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *374*(2065). <https://doi.org/10.1098/rsta.2015.0202>
- Jothilakshmi, S., & Gudivada, V. N. (2016). Large Scale Data Enabled Evolution of Spoken Language Research and Applications. In *Handbook of Statistics* (Vol. 35, pp. 301–340). Elsevier B.V. <https://doi.org/10.1016/bs.host.2016.07.005>

- Karlinsey, J. M. (2012). Electrophoresis. In *Chemical Analysis of Food: Techniques and Applications* (pp. 375–405). Elsevier Inc. <https://doi.org/10.1016/B978-0-12-384862-8.00012-1>
- Keiser, G. (2016). *Light-Tissue Interactions* (pp. 147–196). [https://doi.org/10.1007/978-981-10-0945-7\\_6](https://doi.org/10.1007/978-981-10-0945-7_6)
- Kessler, R. W., & Kessler, W. (2020). Best Practice and Performance of Hardware in Process Analytical Technology (PAT). In *Comprehensive Chemometrics* (pp. 237–274). Elsevier. <https://doi.org/10.1016/b978-0-12-409547-2.14611-6>
- Khalaf, A. R., Alhusban, A. A., Al-Shalabi, E., Al-Sheikh, I., & Sabbah, D. A. (2019). Isolation and structure elucidation of bioactive polyphenols. In *Studies in Natural Products Chemistry* (Vol. 63, pp. 267–337). Elsevier B.V. <https://doi.org/10.1016/B978-0-12-817901-7.00010-1>
- Kirby, G. C., Paine, A., Warhurst, D. C., Noamese, B. K., & Phillipson, J. D. (1995a). In vitro and in vivo antimalarial activity of cryptolepine, a plant-derived indoloquinoline. *Phytotherapy Research*, 9(5), 359–363.
- Kirby, G. C., Paine, A., Warhurst, D. C., Noamese, B. K., & Phillipson, J. D. (1995b). In vitro and in vivo antimalarial activity of cryptolepine, a plant-derived indoloquinoline. *Phytotherapy Research*, 9(5), 359–363. <https://doi.org/10.1002/ptr.2650090510>
- Koenig, J. L. (1971). Raman Scattering of Synthetic Polymers-A Review. *Applied Spectroscopy Reviews*, 4(2), 233–305. <https://doi.org/10.1080/05704927108082605>

- Köhler, I., Jenett-Siems, K., Kraft, C., Siems, K., Abbiw, D., Bienzle, U., & Eich, E. (2002). Herbal remedies traditionally used against malaria in Ghana: Bioassay-guided fractionation of *Microglossa pyrifolia* (Asteraceae). *Zeitschrift Fur Naturforschung - Section C Journal of Biosciences*, 57(11–12), 1022–1027. <https://doi.org/10.1515/znc-2002-11-1212>
- Komlaga, G., Agyare, C., Dickson, R. A., Mensah, M. L. K., Annan, K., Loiseau, P. M., & Champy, P. (2015). Medicinal plants and finished marketed herbal products used in the treatment of malaria in the Ashanti region, Ghana. *Journal of Ethnopharmacology*, 172, 333–346.
- Komlaga, G., Cojean, S., Dickson, R. A., Beniddir, M. A., Suyyagh-Albouz, S., Mensah, M. L., & Loiseau, P. M. (2016). Antiplasmodial activity of selected medicinal plants used to treat malaria in Ghana. *Parasitology Research*, 115(8), 3185–3195.
- Kotu, V., & Deshpande, B. (2019). Model Evaluation. *Data Science*, 263–279. <https://doi.org/10.1016/b978-0-12-814761-0.00008-3>
- Krzyzowska, M., Tomaszewska, E., Ranoszek-Soliwoda, K., Bien, K., Orłowski, P., Celichowski, G., & Grobelny, J. (2017). Tannic acid modification of metal nanoparticles: Possibility for new antiviral applications. In *Nanostructures for Oral Medicine* (Vol. 1). Elsevier Inc. <https://doi.org/10.1016/B978-0-323-47720-8.00013-4>
- Kukula-Koch, W. A., & Widelski, J. (2017). Alkaloids. In *Pharmacognosy: Fundamentals, Applications and Strategy* (pp. 163–198). Elsevier Inc. <https://doi.org/10.1016/B978-0-12-802104-0.00009-3>

- Kumar, V., Coluccelli, N., & Polli, D. (2018). Coherent Optical Spectroscopy/ Microscopy and Applications. In *Molecular and Laser Spectroscopy: Advances and Applications* (pp. 87–115). Elsevier Inc. <https://doi.org/10.1016/B978-0-12-849883-5.00005-X>
- Lakowicz, J. R. (1999). *Principles of Fluorescence Spectroscopy*. Kluwer Academic\_Plenum.
- Lamirel, C. (2014). Optical Coherence Tomography. In *Encyclopedia of the Neurological Sciences* (pp. 660–668). Elsevier Inc. <https://doi.org/10.1016/B978-0-12-385157-4.00171-8>
- Lang, M., Stober, F., & Lichtenthaler, H. K. (1991). Fluorescence emission spectra of plant leaves and plant constituents. *Radiation and Environmental Biophysics*, 30(4), 333–347. <https://doi.org/10.1007/BF01210517>
- Larkin, P. J. (2018). Instrumentation and Sampling Methods. In *Infrared and Raman Spectroscopy* (pp. 29–61). Elsevier. <https://doi.org/10.1016/b978-0-12-804162-8.00003-3>
- Liu, B., & Brezinski, M. E. (2014). Optical Coherence Tomography. In *Comprehensive Biomedical Physics* (Vol. 4, pp. 209–226). Elsevier. <https://doi.org/10.1016/B978-0-444-53632-7.00416-0>
- Liu, Y.-J., Kyne, M., Wang, C., & Yu, X.-Y. (2020). Data mining in Raman imaging in a cellular biological system. *Computational and Structural Biotechnology Journal*, 18(2001–0370), 2920–2930. <https://doi.org/10.1016/j.csbj.2020.10.006>(<https://www.sciencedirect.com/science/a>

article/pii/S2001037020304281)

- Lodder, R. A. (2002). Handbook of Near-Infrared Analysis, 2nd ed., Revised and Expanded. Practical Spectroscopy Series Volume 27 Edited by D. A. Burns (NIR Resources) and E. W. Ciurczak (Purdue Pharma LP). Dekker: New York. 2001. xv + 814 pp. \$225.00. ISBN: 0-8247-0534-3. In *Journal of the American Chemical Society* (Vol. 124, Issue 19, pp. 5603–5604). <https://doi.org/10.1021/ja015320c>
- Lopes, J., Moreira, S., & Barbosa Neto, N. (2020). Selective Inner-Filter on the Fluorescence Response of Chlorophyll and Pheophytin Molecules Extracted from *Caesalpinia echinata* Leaves. *Journal of the Brazilian Chemical Society*, 31(1), 162–169. <https://doi.org/10.21577/0103-5053.20190150>
- MacArthur, R. L., Teye, E., & Darkwa, S. (2020). Predicting adulteration of Palm oil with Sudan IV dye using shortwave handheld spectroscopy and comparative analysis of models. *Vibrational Spectroscopy*, 110(May), 103129. <https://doi.org/10.1016/j.vibspec.2020.103129>
- Mariz, I. F. A., Pinto, S., Lavrado, J., Paulo, A., Martinho, J. M. G., & Maçõas, E. M. S. (2017). Cryptolepine and quindoline: Understanding their photophysics. *Physical Chemistry Chemical Physics*, 19(16), 10255–10263. <https://doi.org/10.1039/c7cp00455a>
- Markus, S., Hofkens, J., & Enderlein, J. (2013). Handbook of Fluorescence Spectroscopy and Imaging. In *Journal of Chemical Information and Modeling* (Vol. 53, Issue 9). WILEY-VCH Verlag & Co. KGaA,.

- MathWorks. (2014). *classperf Evaluate performance of classifier*. <https://www.mathworks.com/help/bioinfo/ref/classperf.html>
- MathWorks. (2019). *Compute confusion matrix for classification problem - MATLAB confusionmat - MathWorks J.* MathWorks Documentation. <https://www.mathworks.com/help/stats/confusionmat.html%0Ahttps://es.mathworks.com/help/stats/confusionmat.html>
- Mayinger, F. (2001). *Fundamentals of Holography and Interferometry*. [https://doi.org/10.1007/978-3-642-56443-7\\_3](https://doi.org/10.1007/978-3-642-56443-7_3)
- McCreery, R. L. (2005). Raman Spectroscopy for Chemical Analysis. In *Raman Spectroscopy for Chemical Analysis*. <https://doi.org/10.1002/0471721646>
- Mena, P., Calani, L., Bruni, R., & Del Rio, D. (2015). Bioactivation of High-Molecular-Weight Polyphenols by the Gut Microbiome. In *Diet-Microbe Interactions in the Gut: Effects on Human Health and Disease* (pp. 73–101). Elsevier Inc. <https://doi.org/10.1016/B978-0-12-407825-3.00006-X>
- Michalke, B., & Nischwitz, V. (2013). *Speciation and Element-Specific Detection*. 633–649. <https://doi.org/10.1016/B978-0-12-415806-1.00022-X>
- Morais, C. L. M., Santos, M. C. D., Lima, K. M. G., & Martin, F. L. (2019). Improving data splitting for classification applications in spectrochemical analyses employing a random-mutation Kennard-Stone algorithm approach. *Bioinformatics*, 35(24), 5257–5263. <https://doi.org/10.1093/bioinformatics/btz421>

- Moreno, J., & Peinado, R. (2012). Polyphenols. In *Enological Chemistry* (pp. 53–76). Elsevier. <https://doi.org/10.1016/B978-0-12-388438-1.00005-4>
- Nadeem, H., & Heindel, T. J. (2018). Review of noninvasive methods to characterize granular mixing. In *Powder Technology* (Vol. 332, pp. 331–350). Elsevier B.V. <https://doi.org/10.1016/j.powtec.2018.03.035>
- Nasdala, L., Beyssac, O., William Schopf, J., & Bleisteiner, B. (2012). Application of Raman-based images in the earth sciences. In *Springer Series in Optical Sciences* (Vol. 168). [https://doi.org/10.1007/978-3-642-28252-2\\_5](https://doi.org/10.1007/978-3-642-28252-2_5)
- Neath, R. C., & Johnson, M. S. (2010). Discrimination and classification. *International Encyclopedia of Education*, 135–141. <https://doi.org/10.1016/B978-0-08-044894-7.01312-9>
- Neilson, A. P., Goodrich, K. M., & Ferruzzi, M. G. (2017a). Bioavailability and Metabolism of Bioactive Compounds From Foods. In *Nutrition in the Prevention and Treatment of Disease* (Fourth Edi). Elsevier Inc. <https://doi.org/10.1016/B978-0-12-802928-2.00015-1>
- Neilson, A. P., Goodrich, K. M., & Ferruzzi, M. G. (2017b). Bioavailability and metabolism of bioactive compounds from foods. In *Nutrition in the Prevention and Treatment of Disease* (pp. 301–319). Elsevier. <https://doi.org/10.1016/B978-0-12-802928-2.00015-1>
- Ngwoke, K. (2018). Antioxidant , Anti-inflammatory , Analgesic Properties , and Phytochemical Characterization of Stem Bark Extract and Fractions of *Anthocleista nobilis*. *Pharmacognosy Research*, February. <https://doi.org>



/10.4103/pr.pr

- O'Connor, S. E. (2010). Alkaloids. In *Comprehensive Natural Products II* (pp. 977–1007). Elsevier. <https://doi.org/10.1016/B978-008045382-8.00013-7>
- Odoh, U. E., & Akwuaka, C. I. (2012). Pharmacognostic profile of root of *Cryptolepis sanguinolenta* ( Lindl .) Schlechter. *Pharmacognosy Journal*, 4(28), 40–44. <https://doi.org/10.5530/pj.2012.28.8>
- Okokon, J. E., Augustine, N. B., Mohanakrishnan, D., Okokon, J. E., Augustine, N. B., & Mohanakrishnan, D. (2017). Antimalarial , antiplasmodial and analgesic activities of root extract of *Alchornea laxiflora*. *Pharmaceutical Biology*, 0(0), 000. <https://doi.org/10.1080/13880209.2017.1285947>
- Oladeji, O. S., Oluyori, A. P., Bankole, D. T., & Afolabi, T. Y. (2020). Natural Products as Sources of Antimalarial Drugs: Ethnobotanical and Ethnopharmacological Studies. *Scientifica*, 2020. <https://doi.org/10.1155/2020/7076139>
- Oliveri, P., Casolino, M. C., & Forina, M. (2010). Chemometric Brains for Artificial Tongues. In *Advances in Food and Nutrition Research* (Vol. 61, Issue C). <https://doi.org/10.1016/B978-0-12-374468-5.00002-7>
- Ombinda-lemboumba, S. (2006). *Laser Induced Chlorophyll Fluorescence of Plant Material*. December.
- Opoku-Ansah, J., Eghan, M. J., Anderson, B., Boampong, J. N., & Buah-Bassuah, P. K. (2016). Laser-Induced Autofluorescence Technique for Plasmodium falciparum Parasite Density Estimation. *Applied Physics Research*, 8(2), 43. <https://doi.org/10.5539/apr.v8n2p43>

- Oppong, E., Agyare, C., Duah, Y., Mbeah, B., Asase, A., Sarkodie, J., Nettey, H., Adu, F., Boatema, P., Agyarkwa, B., Amoateng, P., Asiedu-gyekye, I., & Nyarko, A. (2019). Ethnomedicinal survey and mutagenic studies of plants used in Accra metropolis , Ghana. *Journal of Ethnopharmacology*, July, 112309. <https://doi.org/10.1016/j.jep.2019.112309>
- Oprea, C., & Ti, P. Ş. (2014). Performance Evaluation of the Data Mining Classification Methods. *Analele Universităţii Constantin Brâncuşi Din Târgu Jiu : Seria Economie, I*(Special number-Information society and sustainable development), 249–253.
- Osafo, N., Mensah, K. B., & Yeboah, O. K. (2017). Phytochemical and Pharmacological Review of *Cryptolepis sanguinolenta* (Lindl.) Schlechter. *Advances in Pharmacological Sciences*, 2017. <https://doi.org/10.1155/2017/3026370>
- Osei-Djarbeng, S., Agyekum-Attobra, E., Nkansah, R., Solaga, D., Osei-Asante, S., & Owusu-Dapaah, G. (2015). Medicinal Plants Constituting Antimalarial Herbal Preparations in the Ghanaian Market. *British Journal of Pharmaceutical Research*, 5(3), 153–162. <https://doi.org/10.9734/bjpr/2015/14896>
- Pamp, S., Arce-diego, & Luis, J. (2020). *Application of Classification Algorithms to Di ff use Reflectance Spectroscopy Measurements for Ex Vivo Characterization of Biological Tissues*. 22(736). <https://doi.org/10.3390/e22070736>

- Panche, A. N., Diwan, A. D., & Chandra, S. R. (2016). Flavonoids: An overview. In *Journal of Nutritional Science* (Vol. 5, pp. 1–15). Cambridge University Press. <https://doi.org/10.1017/jns.2016.41>
- Pandey, K. B., & Rizvi, S. I. (2009). Plant polyphenols as dietary antioxidants in human health and disease. *Oxidative Medicine and Cellular Longevity*, 2(5), 270–278. <https://doi.org/10.4161/oxim.2.5.9498>
- Panneton, B., Roger, J., Guillaume, S., & Louis Longchamps. (2008). *Effects of Preprocessing of Ultraviolet-Induced Fluorescence Spectra in Plant Fingerprinting Applications*. August. <https://doi.org/10.1366/000370208784909661>
- Parson, W. W. (2015). Modern optical spectroscopy: With exercises and examples from biophysics and biochemistry, second edition. In *Modern Optical Spectroscopy: With Exercises and Examples from Biophysics and Biochemistry, Second Edition*. <https://doi.org/10.1007/978-3-662-46777-0>
- Pasquini, C. (2003). Near infrared spectroscopy: Fundamentals, practical aspects and analytical applications. *Journal of the Brazilian Chemical Society*, 14(2), 198–219. <https://doi.org/10.1590/S0103-50532003000200006>
- Paulo, A., Paulo, A., Gomes, E. T., Steele, J., Warhurst, D. C., & Houghton, P. J. (2000). *Antiplasmodial Activity of Cryptolepis sanguinolenta Alkaloids from Leaves and Roots*. February. <https://doi.org/10.1055/s-2000-11106>

- Pietta, P., Minoggio, M., & Bramati, L. (2003). Plant polyphenols: Structure, occurrence and bioactivity. *Studies in Natural Products Chemistry*, 28, 257–312. [https://doi.org/10.1016/S1572-5995\(03\)80143-6](https://doi.org/10.1016/S1572-5995(03)80143-6)
- Press, W. H., & Teukolsky, S. A. (1990). Savitzky-Golay Smoothing Filters. *Computers in Physics*, 4(6), 669. <https://doi.org/10.1063/1.4822961>
- Qian, F., Wu, Y., & Hao, P. (2017). A fully automated algorithm of baseline correction based on wavelet feature points and segment interpolation. *Optics and Laser Technology*, 96, 202–207. <https://doi.org/10.1016/j.optlastec.2017.05.021>
- Qiu, S., Sun, H., Zhang, A. H., Xu, H. Y., Yan, G. L., Han, Y., & Wang, X. J. (2014). Natural alkaloids: Basic aspects, biological roles, and future perspectives. *Chinese Journal of Natural Medicines*, 12(6), 401–406. [https://doi.org/10.1016/S1875-5364\(14\)60063-7](https://doi.org/10.1016/S1875-5364(14)60063-7)
- Rencher, A. C. (2002). Methods of Multivariate Analysis. In *Methods of Multivariate Analysis*. <https://doi.org/10.1002/0471271357>
- Reyes, B. A. S., Dufourt, E. C., Ross, J., Warner, M. J., Tanquilut, N. C., & Leung, A. B. (2017). Selected Phyto and Marine Bioactive Compounds: Alternatives for the Treatment of Type 2 Diabetes. In *Studies in Natural Products Chemistry* (Vol. 55, pp. 111–143). Elsevier B.V. <https://doi.org/10.1016/B978-0-444-64068-0.00004-8>
- Rigler, R., & Widengren, J. (2018). Fluorescence-based monitoring of electronic state and ion exchange kinetics with FCS and related techniques: from T-jump measurements to fluorescence fluctuations. *European Biophysics*

*Journal*, 47(4), 479. <https://doi.org/10.1007/S00249-017-1271-1>

Rodríguez-García, C., Sánchez-Quesada, C., Toledo, E., Delgado-Rodríguez, M., & Gaforio, J. J. (2019). Naturally lignan-rich foods: A dietary tool for health promotion? In *Molecules* (Vol. 24, Issue 5). MDPI AG. <https://doi.org/10.3390/molecules24050917>

Roshchina, V. V. (2012). Vital Autofluorescence: Application to the Study of Plant Living Cells. *International Journal of Spectroscopy*, 2012, 1–14. <https://doi.org/10.1155/2012/124672>

Roshchina, V. V. (2018). *Fluorescing World of Plant Secreting Cells*.

Roshchina, V. V, Kuchin, A. V, & Yashin, V. A. (2016). *Autofluorescence of Plant Secretary Cells as Possible Tool for Pharmacy*. 2(2), 31–38. <https://doi.org/10.11648/j.ijpc.20160202.15>

Roshchina, V. V, Kuchin, A. V, & Yashin, V. A. (2017). *Application of Autofluorescence for Analysis of Medicinal Plants*. 2017.

Rost, F. (2016). Fluorescence microscopy, applications. In *Encyclopedia of Spectroscopy and Spectrometry* (pp. 627–631). Elsevier. <https://doi.org/10.1016/B978-0-12-803224-4.00147-3>

Roy, S., Sharma, P., Nath, K., Bhattacharyya, D. K., & Kalita, J. K. (2018). Pre-processing: A data preparation step. *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics*, 1–3(January), 463–471. <https://doi.org/10.1016/B978-0-12-809633-8.20457-3>

- Salkind, N. (2012). Matrix Algebra. *Encyclopedia of Research Design*.  
<https://doi.org/10.4135/9781412961288.n232>
- Schlechter, L., Osafo, N., Mensah, K. B., & Yeboah, O. K. (2017). *Phytochemical and Pharmacological Review of Cryptolepis sanguinolenta* *Phytochemical and Pharmacological Review of Cryptolepis sanguinolenta ( Lindl .) Schlechter. October.* <https://doi.org/10.1155/2017/3026370>
- Schrader, B. (1996). Infrared and Raman Spectroscopy, Method and Applications. In *Vibrational Spectroscopy* (Vol. 11, Issue 2). [https://doi.org/10.1016/0924-2031\(00\)00065-5](https://doi.org/10.1016/0924-2031(00)00065-5)
- Selden, A. C. (1973). Criticality in lasers and masers. *Physics Letters A*, 45(5), 389–390. [https://doi.org/10.1016/0375-9601\(73\)90251-X](https://doi.org/10.1016/0375-9601(73)90251-X)
- Seyrek, E., & Decher, G. (2012). Layer-by-Layer Assembly of Multifunctional Hybrid Materials and Nanoscale Devices. In *Polymer Science: A Comprehensive Reference, 10 Volume Set* (Vol. 7, pp. 159–185). Elsevier. <https://doi.org/10.1016/B978-0-444-53349-4.00182-5>
- Shuvalov, V. Y., Rupp, A. S., Fisyuk, A. S., & Kuratova, A. K. (2019). *Synthesis and Optical Properties of Alkaloid Quindoline , Its Structural Analogues and Substituted  $\delta$  -Carbolines*. 1696–1699. <https://doi.org/10.1002/slct.201803515>
- Simpson, D., & Amos, S. (2017). Other Plant Metabolites. In *Pharmacognosy: Fundamentals, Applications and Strategy* (pp. 267–280). Elsevier Inc. <https://doi.org/10.1016/B978-0-12-802104-0.00012-3>

- Sirohi, R. S. (2009). Optical methods of measurement: Wholefield techniques, second edition. In *Optical Methods of Measurement: Wholefield Techniques, Second Edition* (pp. 1–291).
- Souza, P. A. De, & Moreira, L. F. (2018). *Cacao — Theobroma cacao*. 3(2001), 69–76. <https://doi.org/10.1016/B978-0-12-803138-4.00010-1>
- Speight, J. G. (2017). Sources and Types of Inorganic Pollutants. In *Environmental Inorganic Chemistry for Engineers* (pp. 231–282). Elsevier. <https://doi.org/10.1016/B978-0-12-849891-0.00005-9>
- Srinivasan, M., Sudheer, A. R., & Menon, V. P. (2007). Ferulic acid: Therapeutic potential through its antioxidant property. In *Journal of Clinical Biochemistry and Nutrition* (Vol. 40, Issue 2, pp. 92–100). The Society for Free Radical Research Japan. <https://doi.org/10.3164/jcbrn.40.92>
- Strachan, C. J., Rades, T., Gordon, K. C., & Rantanen, J. (2007). Raman spectroscopy for quantitative analysis of pharmaceutical solids. *Journal of Pharmacy and Pharmacology*, 59(2), 179–192. <https://doi.org/10.1211/jpp.59.2.0005>
- Stuart, B. H. (2005). Infrared Spectroscopy: Fundamentals and Applications. In *Infrared Spectroscopy: Fundamentals and Applications*. <https://doi.org/10.1002/0470011149>
- Stuart, M. (1982). A geometric approach to principal components analysis. *American Statistician*, 36(4), 365–367. <https://doi.org/10.1080/00031305.1982.10483050>

- Subasi, A. (2020). Data preprocessing. In *Practical Machine Learning for Data Analysis Using Python* (pp. 27–89). Elsevier. <https://doi.org/10.1016/b978-0-12-821379-7.00002-3>
- Sun, C. W. (2015). Biophotonics for tissue oxygenation analysis. In *Biophotonics for Medical Applications* (pp. 301–320). Elsevier Inc. <https://doi.org/10.1016/B978-0-85709-662-3.00010-5>
- Tamokou, J. de D., Mbaveng, A. T., & Kuete, V. (2017). Antimicrobial Activities of African Medicinal Spices and Vegetables. In *Medicinal Spices and Vegetables from Africa: Therapeutic Potential Against Metabolic, Inflammatory, Infectious and Systemic Diseases* (pp. 207–237). Elsevier Inc. <https://doi.org/10.1016/B978-0-12-809286-6.00008-X>
- Tay, S. C. K., Dankwa, K., Gbedema, S. Y., & Archibald, A. (2011). *In vivo Antimalarial Activity Evaluation of Two Cryptolepis sanguinolenta Based Herbal Decoctions*. 2011(1), 2–5.
- Taylor, A. (1999). Biomedical Applications of Atomic Spectroscopy. In *Encyclopedia of Spectroscopy and Spectrometry* (pp. 174–182). Elsevier Ltd. <https://doi.org/10.1016/B978-0-12-374413-5.00110-X>
- Telle, H. H., & Donovan, A. G. U. R. J. (2007). Laser-induced Fluorescence Spectroscopy. In *Spectroscopy, Dynamics and Applications* (pp. 101–102).
- Teoh, E. S. (2016). Secondary Metabolites of Plants. In *Medicinal Orchids of Asia* (pp. 59–73). Springer International Publishing. [https://doi.org/10.1007/978-3-319-24274-3\\_5](https://doi.org/10.1007/978-3-319-24274-3_5)



- Thoppil, R. J., & Bishayee, A. (2011). Terpenoids as potential chemopreventive and therapeutic agents in liver cancer. In *World Journal of Hepatology* (Vol. 3, Issue 9, pp. 228–249). Baishideng Publishing Group Inc. <https://doi.org/10.4254/wjh.v3.i9.228>
- Titus, D., Jebaseelan, J. S. E., & Roopan, S. M. (2019). Nanoparticle characterization techniques. In *Green Synthesis, Characterization and Applications of Nanoparticles*. Elsevier Inc. <https://doi.org/10.1016/b978-0-08-102579-6.00012-5>
- Tranter, G. E. (1999). UV-Visible Absorption and Fluorescence Spectrometers. *Encyclopedia of Spectroscopy and Spectrometry*, 2, 2383–2389. <https://doi.org/10.1006/rwsp.2000.0088>
- Vaibhaw, Sarraf, J., & Pattnaik, P. K. (2020). Brain–computer interfaces and their applications. In *An Industrial IoT Approach for Pharmaceutical Industry Growth* (pp. 31–54). Elsevier. <https://doi.org/10.1016/b978-0-12-821326-1.00002-4>
- Valeur, B. (2001). Handbook of Analytical Techniques Single-Molecule Detection in Solution . Methods and Applications. In *Methods* (Vol. 8). <https://doi.org/10.1002/3527600248>
- Van Lith, R., & Ameer, G. A. (2016). Antioxidant Polymers as Biomaterial. In *Oxidative Stress and Biomaterials* (pp. 251–296). Elsevier. <https://doi.org/10.1016/B978-0-12-803269-5.00010-3>

- Vanlanduit, S., & Guillaume, P. (2009). Optical measurement techniques. In *Optics and Lasers in Engineering* (Vol. 47, Issues 3–4). <https://doi.org/10.1016/j.optlaseng.2008.06.011>
- Verpoorte, R. (2004). Alkaloids. In *Encyclopedia of Analytical Science: Second Edition* (pp. 56–61). Elsevier Inc. <https://doi.org/10.1016/B0-12-369397-7/00010-8>
- Virk, H. S. (2014). *Luminescence related phenomena and their applications : special topic volume with invited peer reviewed papers only*. [https://locate.coventry.ac.uk/openurl/COV/COV\\_VU1?u.ignore\\_date\\_coverage=true&rft.mms\\_id=996648129202011](https://locate.coventry.ac.uk/openurl/COV/COV_VU1?u.ignore_date_coverage=true&rft.mms_id=996648129202011)
- Wartewig, S. (2004). Front Matter. In *IR and Raman Spectroscopy*. <https://doi.org/10.1002/3527601635.fmatter>
- Watanabe, A., Furukawa, H., Miyamoto, S., & Minagawa, H. (2019). Non-destructive chemical analysis of water and chlorine content in cement paste using near-infrared spectroscopy. *Construction and Building Materials*, 196, 95–104. <https://doi.org/10.1016/j.conbuildmat.2018.11.114>
- Weiner, J. (2003). *LIGHT—MATTER INTERACTION*. John Wiley & Sons, Inc., Hoboken.
- Weselucha-birczy, A., Fr, A., Paciorek, K., Bajowska, A., Ko, A., & Bła, M. (2014). *Vibrational Spectroscopy Application of Raman spectroscopy to study of the polymer foams modified in the volume and on the surface by carbon nanotubes* &. 72, 50–56. <https://doi.org/10.1016/j.vibspec.2014>

02.009

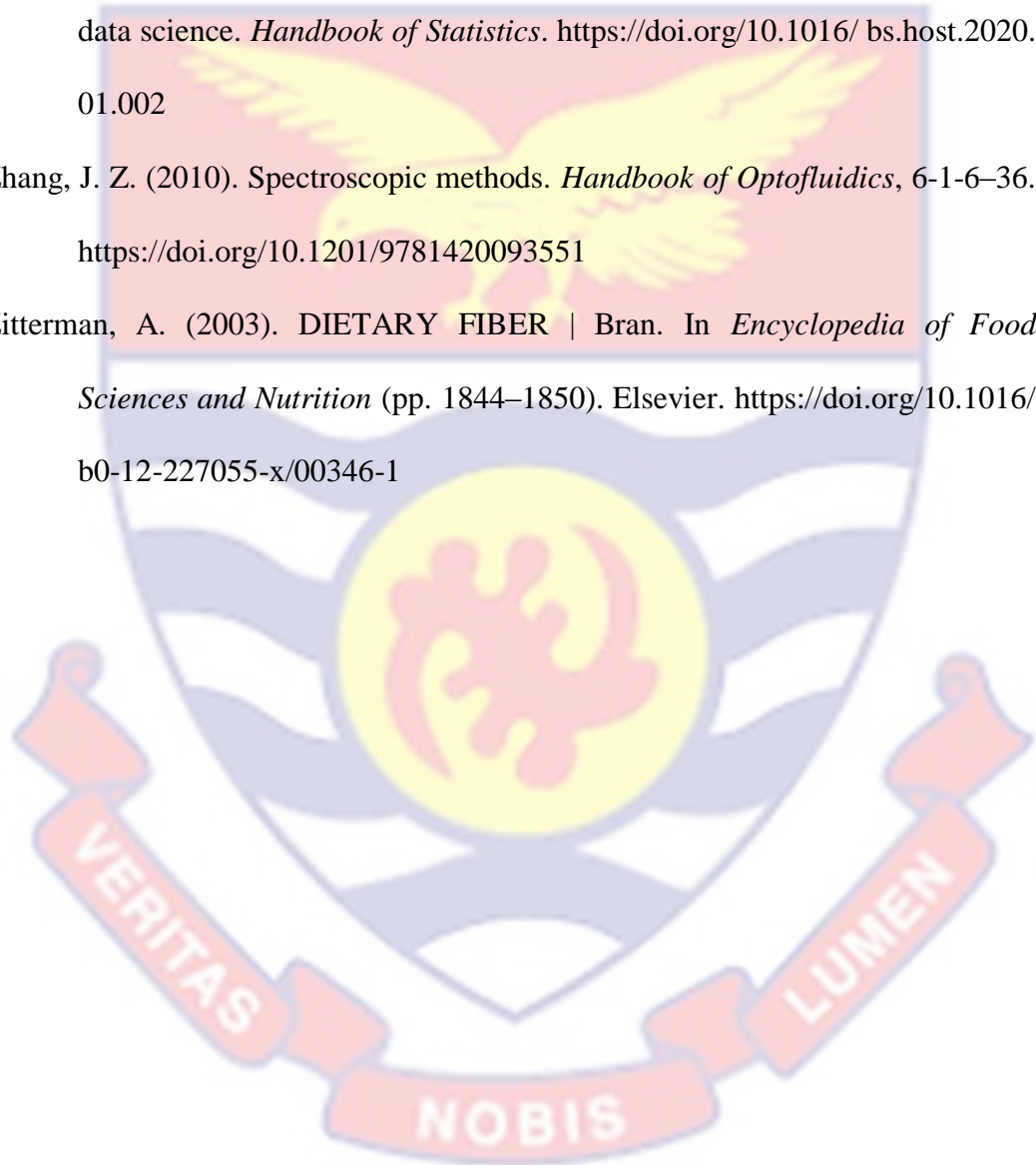
- WHO. (2017a). *WHO | Number of malaria deaths*. <http://www.who.int/gho/malaria/epidemic/deaths/en/>
- WHO. (2017b). *WHO | WHO traditional medicine strategy: 2014-2023*.
- WHO. (2018). *WHO | Malaria*. <http://www.who.int/mediacentre/factsheets/fs094/en/>
- Wilmot, D., Ameyaw, E. O., Amoako-Sakyi, D., Boampong, J. N., & Quashie, N. B. (2017). In vivo efficacy of top five surveyed Ghanaian herbal anti-malarial products. *Malaria Journal*, *16*(1), 103.
- Wolfbeis, O. S. (2004). Springer Series on Fluorescence: Methods and application. Fluorescence Spectroscopy in Biology. In *New York*.
- Wright, C. W., Addae-Kyereme, J., Breen, A. G., Brown, J. E., Cox, M. F., Croft, S. L., & Pollet, P. L. (2001). Synthesis and evaluation of cryptolepine analogues for their potential as new antimalarial agents. *Journal of Medicinal Chemistry*, *44*(19), 3187–3194.
- Wulandari, L., Retnaningtyas, Y., Nuri, & Lukman, H. (2016). Analysis of Flavonoid in Medicinal Plant Extract Using Infrared Spectroscopy and Chemometrics. *Journal of Analytical Methods in Chemistry*, 2016. <https://doi.org/10.1155/2016/4696803>
- Xu, M., Wang, J., & Gu, S. (2019). Rapid identification of tea quality by E-nose and computer vision combining with a synergetic data fusion strategy. *Journal of Food Engineering*, *241*, 10–17. <https://doi.org/10.1016/j.jfoodeng.2018.07.020>

Yao, V., Barku, A., & Dzotsi, E. Y. (2012). *Isolation and pharmacological activities of alkaloids from Cryptolepis sanguinolenta ( Lindl ) schlt. June 2016.*

Yeturu, K. (2020). Machine learning algorithms, applications, and practices in data science. *Handbook of Statistics*. <https://doi.org/10.1016/bs.host.2020.01.002>

Zhang, J. Z. (2010). Spectroscopic methods. *Handbook of Optofluidics*, 6-1-6-36. <https://doi.org/10.1201/9781420093551>

Zitterman, A. (2003). DIETARY FIBER | Bran. In *Encyclopedia of Food Sciences and Nutrition* (pp. 1844-1850). Elsevier. <https://doi.org/10.1016/b0-12-227055-x/00346-1>



APPENDICES

APPENDIX A

LIAF SPECTRA OF AMHDS SAMPLE

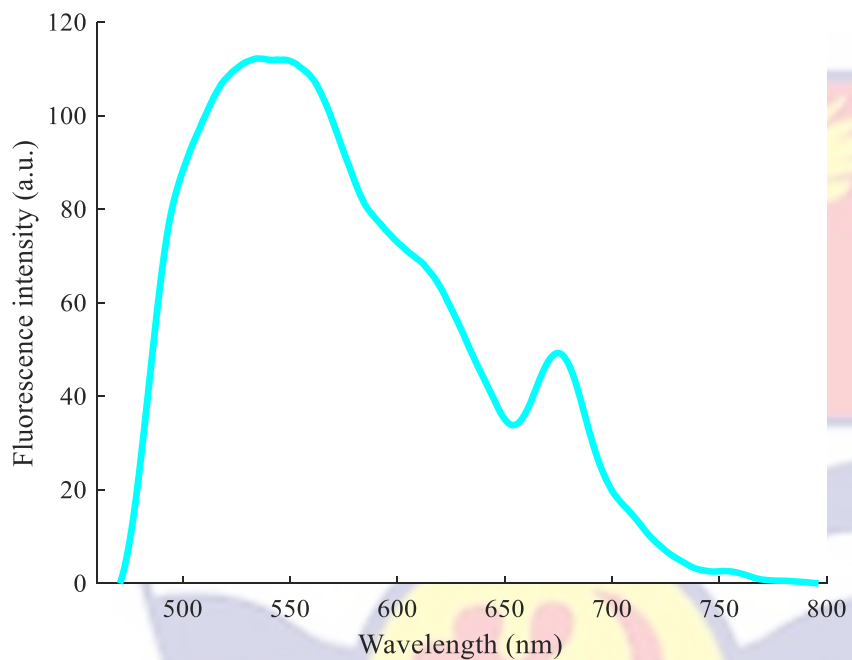


Figure 22: LIAF spectra of AMHD\_D

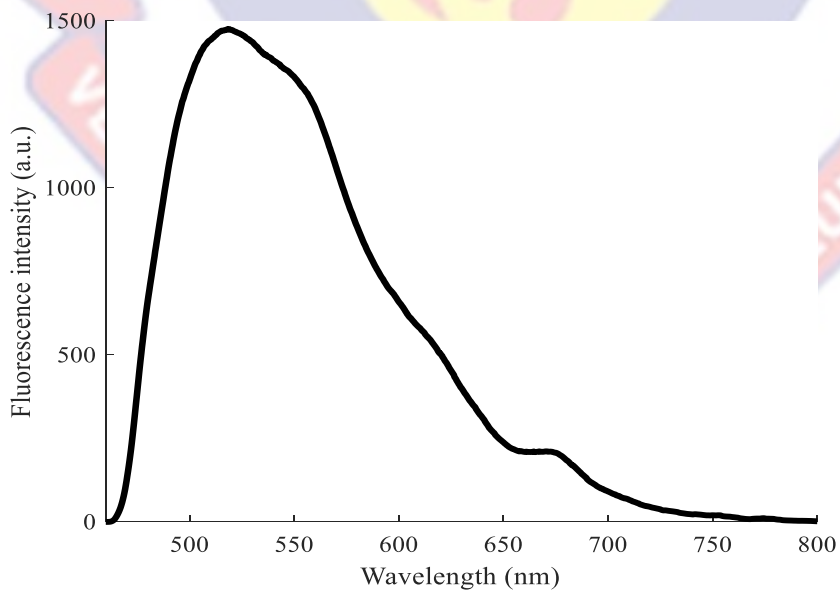


Figure 23: LIAF spectra of AMHD\_E

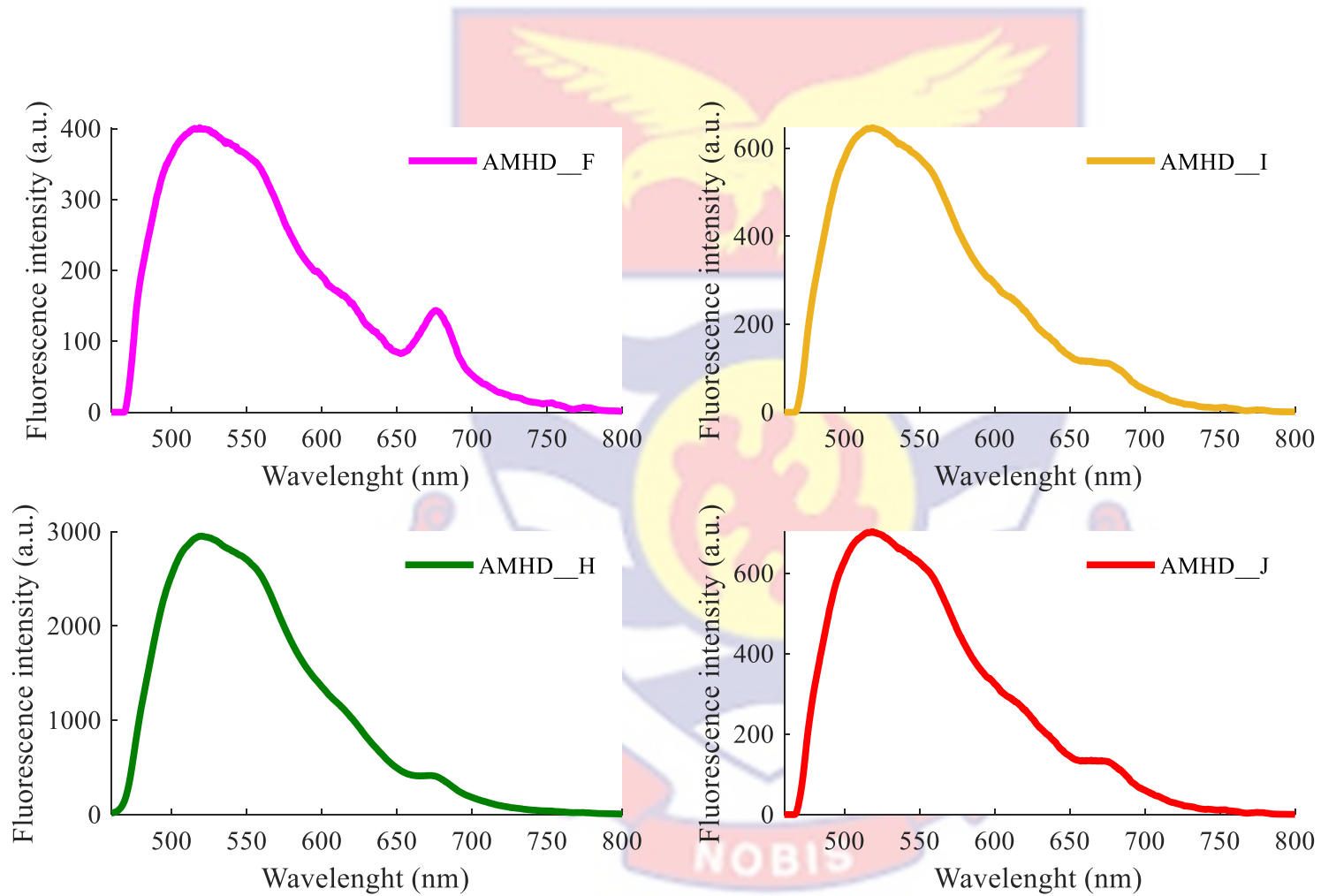


Figure 24: LIAF spectra of AMHD\_F, AMHD\_H, AMHD\_I and AMHD\_J

APPENDIX B

DECONVOLUTED PEAKS OF AMHDs SAMPLE

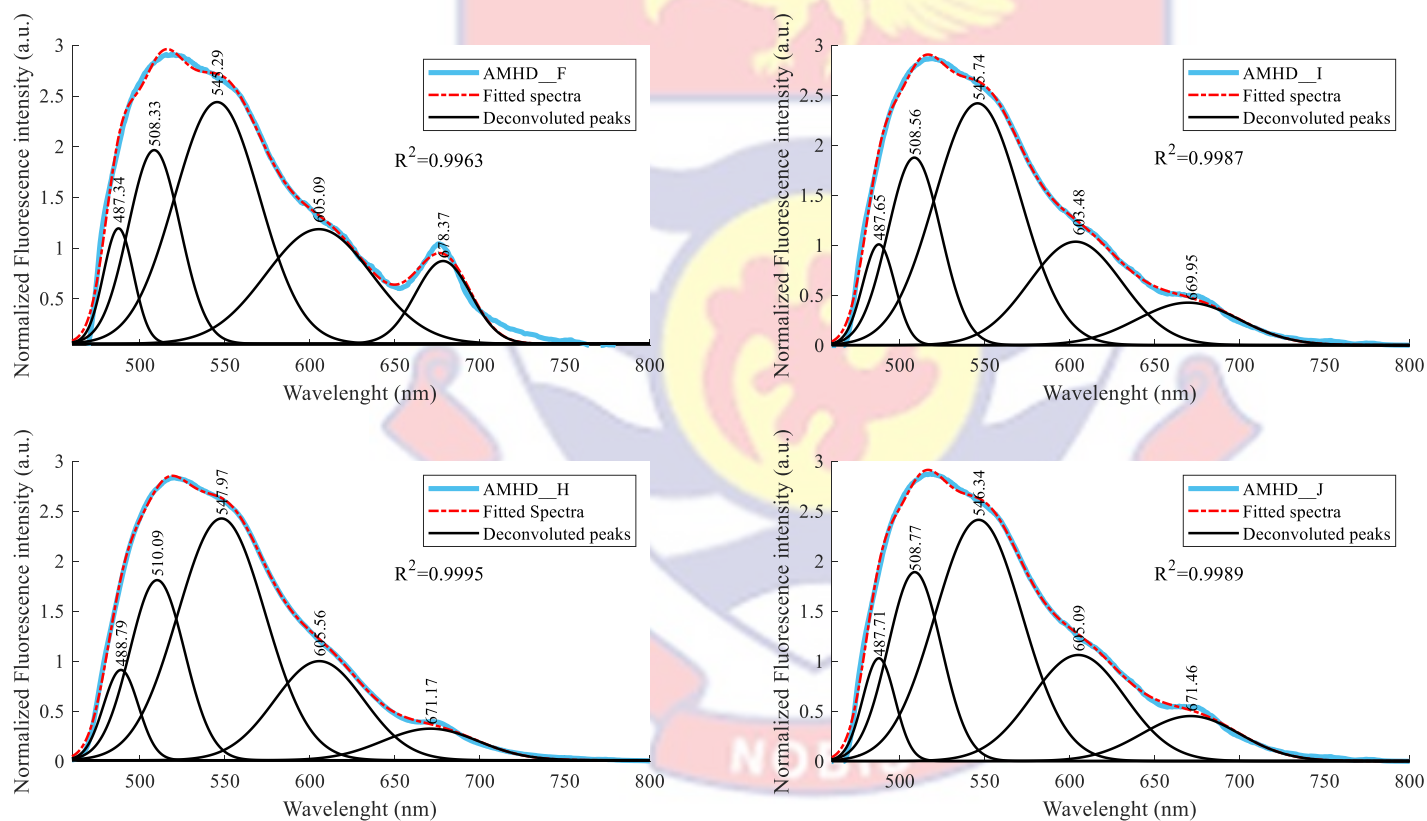


Figure 25: Deconvoluted LIAF spectra of AMHD\_F, AMHD\_H, AMHD\_I and AMHD\_J

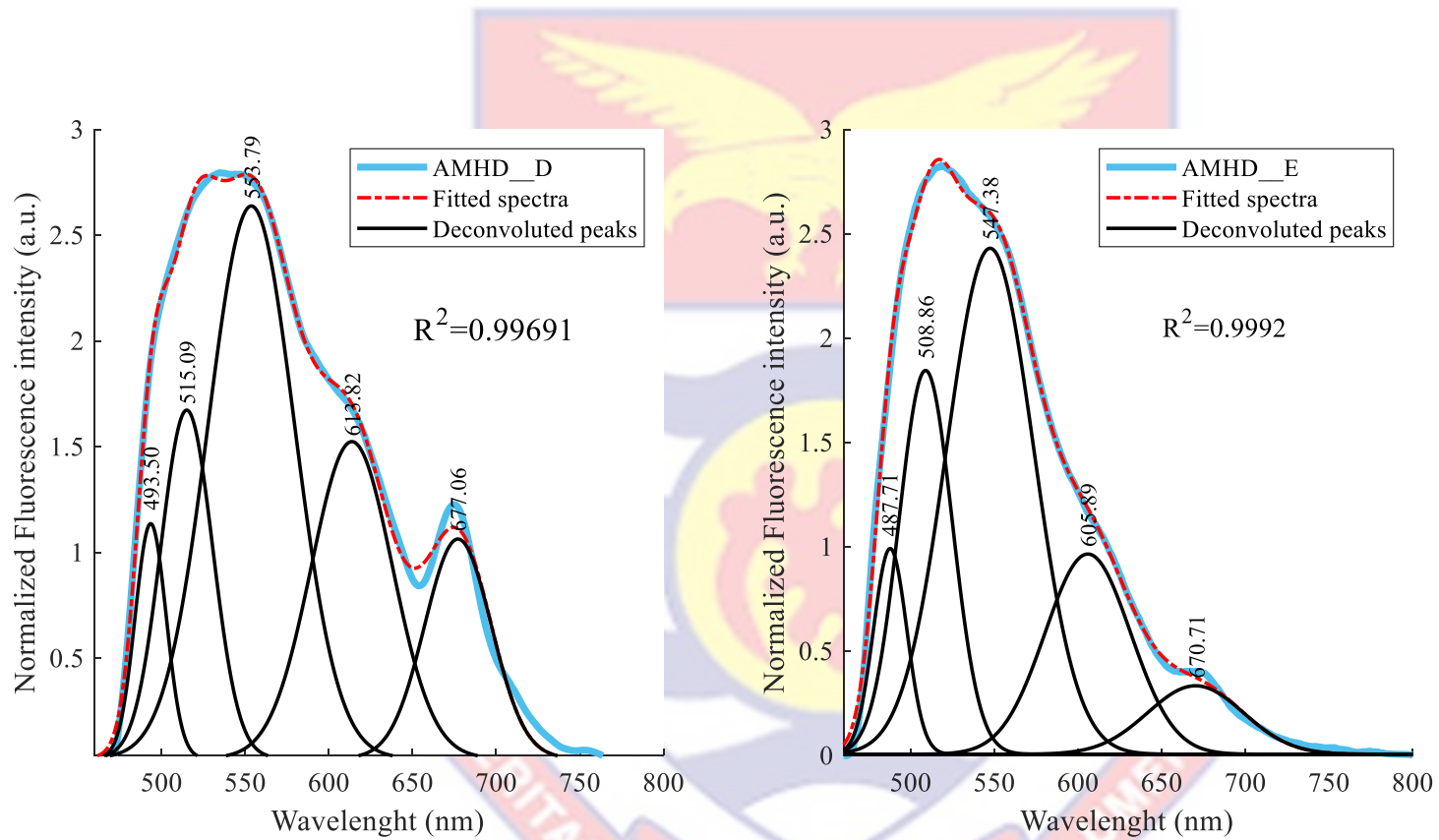


Figure 26: Deconvoluted LIAF spectra of AMHD\_D, and AMHD\_E