CLASSIFICATION OF WHITE ROOT DISEASE SEVERITY LEVELS USING RUBBER TREE SPECTRAL DATA

By

SITI SARIPA RABIAH BINTI MAT LAZIM

© COPYRIGHT UPM

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Doctor Philosophy

March 2025

IKP 2025 1

All material contained within the thesis, including without limitation text, logos, icons, photographs, and all other artwork, is copyright material of Universiti Putra Malaysia unless otherwise stated. Use may be made of any material contained within the thesis for non-commercial purposes from the copyright holder. Commercial use of material may only be made with the express, prior, written permission of Universiti Putra Malaysia.

Copyright © Universiti Putra Malaysia

© COPYRIGHT UPM

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

CLASSIFICATION OF WHITE ROOT DISEASE SEVERITY LEVELS USING RUBBER TREE SPECTRAL DATA

By

SITI SARIPA RABIAH BINTI MAT LAZIM

March 2025

Chairman : Associate Professor Nazmi bin Mat Nawi, PhD
Institute : Plantation Studies

In Southeast Asia, white root disease (WRD) is one of the most damaging diseases for rubber plantations, which can cause up to 50% reduction in natural rubber productivity. The disease can be controlled if it is detected earlier, where affected trees are treated immediately. Early detection means detecting the disease before any visible symptoms of infection appear (pre-symptomatic phase). However, the existing detection technique is labor-intensive, costly, and time-consuming because it depends on skilled personnel and laboratory analysis. Therefore, there is a need for an accurate and reliable in-situ measurement system to detect the disease at an early stage. Thus, the objective of this study was to explore the capability of a low-cost visible shortwave near-infrared (VSNIR) in combination with machine learning as an early detection method for WRD from rubber trees. A total of 134 rubber trees at different severity levels namely, healthy (S0), light (S1), moderate (S2), severe (S3), and very severe infection (S4) rubber trees were used in the study. Samples of leaves, latex, and soil

© COPYRIGHT UPM

around rubber trees were collected from both healthy and affected trees. For the leaf samples, a visible shortwave near-infrared (VSNIR) spectrometer was used to record the spectral data. At the same time, SPAD value was measured using a SPAD meter, moisture content (MC) was measured using a moisture content analyzer, and leaf color was measured using a colorimeter. For the latex samples, the total soluble solid content (TSC) was determined using standard methods ISO 124:2001 in the laboratory. Finally, for the soil samples, the pH and moisture content (MCs) properties around the rubber trees were determined using a Portable Soil Analyzer with multiprobes. The calibration and prediction models were developed to correlate the spectral data with the chemical properties of the samples using the partial least square (PLS) regression method. Feature selection algorithms such as ANOVA, Chi-Square and RELIEF were used to select the most significant wavelength for classification models. Then, four classifiers, namely, artificial neural networks (ANN), support vector machine (SVM), random forest (RF), and k-nearest neighbor (kNN) were applied to classify spectral data into different severity levels of WRD. The result shows that the leaf sample from S4 had the lowest SPAD value of 47 and the lowest MC of 41.95 % w.b. The result presented SPAD value and MC, which indicates an inverse relationship with the severity levels of WRD. The L* value increased from S0 (32.97 ± 3.37) to S4 (38.43 ± 4.97), while the a* value was decreased from S0 (-1.58 ± 1.92) to S4 (-2.16 ± 3.04) and the b* value was increased from S0 (12.64 ± 3.44) to S4 (17.79 ± 6.39). The result shows that the leaves are lighter, greener, and have a more yellowish appearance with increasing disease severity. For the soil samples, the average pH value remains relatively consistent across all severity levels, ranging from 6.20 (S0) and 6.40 (S4), while MCs decreased as the disease severity increased. The MC value dropped significantly from S0 (33.50 % w.b.) to S4 (21.70 % w.b.). For the latex samples, the

© COPYRIGHT UPM

TSC value was decreased from S0 (1.90%) to S3 (1.58%), this indicates that as the severity of WRD increases, the latex becomes more diluted. The crop parameters obtained from the samples at different severity levels significantly difference at ($p<0.05$) except for soil samples. In term of PLS regression models, the results showed that the prediction of chlorophyll content from the leaf sample had the highest coefficient determination (R^2), with a calibration model value of 0.99 and a prediction model value of 0.99. For the MC prediction of rubber leaf samples, the calibration model gave an R^2 value of 0.61, whereas the prediction model gave an R^2 value of 0.65. For the color prediction of the rubber leaf samples, R^2 values for calibration models for L*, a*, and b* were 0.88, 0.85, and 0.89, respectively. For the prediction models, R^2 values for L*, a*, and b* were 0.95, 0.95, and 0.94, respectively. The result showed that the SPAD value of the leaf samples had higher R^2 than other crop parameters. Thus, SPAD value could be the parameter that VSNIR can predict across all severity levels. The RELIEF feature selection method was the most effective for identifying the significant wavelengths that distinguish between different disease severity levels. Additionally, the kNN classifier was the most accurate for classifying WRD severity, enabling reliable early disease detection. In conclusion, utilising VSNIR is a technology that shows promising in detecting WRD in rubber trees at an early stage by analysing changes the SPAD value, leaf color and MC levels. Therefore, VSNIR technology offers possible applications for controlling the disease's spread within the farm.

Keywords: Disease severity, Latex, Machine learning, Rubber trees, Spectroscopy

SDG: GOAL 4: Zero Hunger, GOAL 9: Industry, Innovation and Infrastructure

© COPYRIGHT UPM

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

KLASIFIKASI TAHAP KETERUKAN PENYAKIT AKAR PUTIH MENGGUNAKAN DATA SPEKTRA POKOK GETAH

Oleh

SITI SARIPA RABIAH BINTI MAT LAZIM

Mac 2025

Pengerusi : Profesor Madya Nazmi bin Mat Nawi,PhD
Institut : Kajian Perladangan

Penyakit akar putih (WRD) merupakan salah satu penyakit yang paling berbahaya di Asia Tenggara yang menyebabkan penggurangkan produktiviti getah asli sehingga 50% akibat jangkitan kulat *Rigidiporus micropores.* Penyakit ini boleh dikawal jika ianya dapat dikesan dengan lebih awal dan pokok-pokok yang dijangkiti dirawat dengan segera. Walau bagaimanapun, teknik pengesanan yang ada sekarang memerlukan tenaga kerja yang intensif, mahal dan memakan masa kerana ianya bergantung kepada kakitangan yang terlatih dan analisis di makmal. Oleh itu terdapat satu keperluan untuk system pengukuran in-situ yang tepat bagi mengesan penyakit ini pada peringkat awal. Pengesana awal bermaksud mengesan penyakit WRD sebelum pokok getah mengalami sebarang gejela jangkitan kelihatan (fasa pra-simptomik). Oleh itu, objektif ini adalah untuk mengkaji keupayaan spektroskopi gelombang pendek inframerah dekat yang boleh dilihat (VSNIR) dengan kombinasi dengan pembelajaran mesin sebagai kaedah pengesanan penyakit WRD pada peringkat awal di pokok getah. Sebanyak 134 bilangan pokok getah telah digunakan dalam kajian ini, mewakili lima

© COPYRIGHT UPM

tahap jangkitan berbeza: sihat (S0), ringan (S1), sederhana (S2), teruk (S3) dan sangat teruk dijangkiti (S4). Sampel daun, lateks dan tanah (dikumpul disekeliling pokok getah) telah di ambil dari pokok yang sihat dan pokok dijangkiti oleh WRD. Bagi sampel daun, spektrometer gelombang pendek inframerah dekat yang boleh dilihat (VSNIR) digunakan untuk merekod data spektrum sampel daun. Pada masa yang sama, nilai SPAD telah diukur menggunakan meter SPAD, kandungan kelembapan (MC) diukur menggunakan penganalisis kadungan kelembapan, dan akhirnya warna daun diukur menggunakan kolorimeter. Bagi sampel lateks, kandungan pepejal terlarut (TSC) ditentukan menggunalan kaedah standard ISO 124:2001 di makmal. Bagi sampel tanah, fifat kimia pH dan kandungan kelembapan (MC_S) di sekitar pokok getah ditentukan menggunakan meter penganalisis tanah mudah alih. Model penentukuran dan ramalan telah dirumuskan untuk memghubungkan data spektrum dengan sifat kimia sampel menggunakan kaedah separa kuasa dua terkecil (PLS). Algoritma pemilihan ciri seperti ANOVA, Chi-Square, dan RELIEF digunakan untuk memlilih panjang gelombang yang paling signifikan. Kemudian, empat pengkelas digunakan, iaitu rangkaian neural buatan (ANN), mesin sokongan vector (SVM), hutan rawak (RF) dan k-jiran terdekat (kNN) untuk mengelaskan tahap jangkitan WRD yang berbeza. Hasil kajian menunjukkan bahawa sampel daun dari S4 mempunyai nilai SPAD yang terendah iaitu 47 dan kandungan MC yang terendah iaitu 41.95% w.b.. Ini menunjukkan bahawa nilai SPAD dan MC mempunyai berkaitan secara songsang dengan keterukan WRD. Nilai L* meningkat dari S0 (32.97 ± 3.37) ke S4 (38.43 ± 4.97), manakala nilai a* menurun dari S0 (-1.58 ± 1.92) ke S4 (2.16 ± 3.04) dan nilai b* meningkat dari S0 (12.64 ± 3.44) ke S4 (17.79 ± 6.39). Hasil kajian ini menunjukkan bahawa daun menjadi lebih cerah, lebih hijau dan kekuningan dengan peningkatan tahap keterukan penyakit. Bagi sampel tanah, nilai purata pH kekal

© COPYRIGHT UPM

konsisten di semua tahap keterukan, iaitu antara 6.20 (S0) dan 6.40 (S4), manakala MCs menurun apabila tahap keterukan penyakit meningkat. Nilai MCs menurun dengan ketara dari S0 (33.50 % w.b.) ke S4 (21.70 % w.b.). Bagi sampel lateks, nilai TSC menurun dari S0 (1.90%) ke S3 (1.58%), ini menunjukkan bahawa apabila tahap keterukan WRD meningkat, lateks menjadi lebih cair. Ciri-ciri kimia yang diperoleh daripada sampel pada tahap keterukan yang berbeza menunjukkan perbezaan yang signifikan pada ($p<0.05$) kecuali untuk sampel tanah.Ramalan nilai SPAD mempunyai pekali korelasi (R^2) tertinggi, dengan nilai model kalibrasi sebanyak 0.99 dan nilai model ramalan sebanyak 0.99. Manakala bagi MC pada sampel daun getah, model kalibrasi memberikan nilai r^2 sebanyak 0.61 manakala model ramalan memberikan nilai R^2 sebanyak 0.65. Bagi ramalan warna sampel daun getah, R^2 untuk model kalibrasi bagi L*, a* dan b* masing adalah 0.88, 0.85 dan 0.89. Hasil kajian menunjukkan bahawa kandungan klorofil dalam sampel daun getah mempunyai pekali korelasi yang lebih tinggi berbanding sifat kimia yang lain. Oleh itu, nilai SPAD adalah faktor yang dapat diramal oleh VSNIR merentasi semua tahap jangkitan WRD. Bagi algoritma pemilihan ciri, kaedah RELIEF adalah keadah terbaik untuk mengenal pasti gelombang yang ketara untuk membezakan tahap keterukan penyakit pada pokok getah berbanding ANOVA dan Chi-Square. kNN dikenali sebagai pengelas terbaik. Secara ringkasnya, penggunaan VSNIR adalah teknik yang menunjukkan potensi dalam mengesan WRD pada pokok getah pada peringkat awal dengan mengalisis perubahan nilai SPAD, warna daun dan tahap MC pada daun getah. Pendekatan ini menawarkan aplikasi yang berpotensi untuk mematau penyebaran penyakit di ladang.

Kata Kunci: tahap keterukan, getah, mesin pembelajaran, pokok getah, spektroskopi

SDG: GOAL 4: Sifar kebuluran, GOAL 9: Industri, Innovasi and Infrastruktur

© COPYRIGHT UPM

ACKNOWLEDGEMENTS

First and foremost, I’m grateful to Allah, the Almighty, for giving me this opportunity and for granting me the capability to proceed successfully.

I would like to express my great appreciation to my supervisor, Assoc. Prof. Dr. Nazmi Bin Mat Nawi. Thank you for your never-ending encouragement and valuable, constructive suggestions while I was completing this research. He provided sound advice, good teaching, and friendly company and shared a lot of her expertise, research insight, and ideas.

With a great deal of luck, I had an excellent Supervisory Committee. I owe an immense debt to Assoc. Prof. Dr. Zulkefly bin Sulaiman and Ts. Dr. Anas bin Mohd Mustafah, thank you for your encouragement, insightful comments, and critical review. This thesis could not have been done without their advice. Special thanks are extended to Dr. Mohd Nasaruddin Mohd Aris, the Director of Research Station RRIM, Malaysia Rubber Board, for his invaluable support in granting permission to use the rubber tree samples for this study. His willingness to provide access to the essential material has significantly contributed to the success of this research.

I am deeply indebted to my beloved parents, Mr. Mat Lazim Bin Mohamed and Mdm. Mek Semah Binti Muhammad, who gave unconditional support in finishing my study. Special gratitude to my beloved labmates Ms. Najidah binti Abdullah, Mdm Siti Nooradzah binti Adam, Ms. Nik Norasyikin Binti Raja Ismail, and Ms. Nur Nadhirah Binti Bashid, their continuous guidance and encouragement throughout the completion of my PhD journey.

© COPYRIGHT UPM

I extend my appreciation to the Jabatan Perkhidmatan Awam (JPA) under the Excellent Student Program scholarship for sponsoring my study and the Fundamental Research Grant Scheme (FRGS) 5540386 Ministry of Higher Education, Malaysia, for the grant that supported this project.

© COPYRIGHT UPM

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

Nazmi bin Mat Nawi, PhD
Associate Professor
Faculty of Engineering
Universiti Putra Malaysia
(Chairman)

Anas bin Mohd Mustafah, PhD
Senior lecturer
Faculty of Engineering
Universiti Putra Malaysia
(Member)

Zulkefly bin Sulaiman, PhD
Associate Professor
Faculty of Agriculture
Universiti Putra Malaysia
(Member)

ZALILAH MOHD SHARIFF, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 12 June 2025

© COPYRIGHT UPM

TABLE OF CONTENTS

© COPYRIGHT UPM

© COPYRIGHT UPM

© COPYRIGHT UPM

LIST OF TABLES

© COPYRIGHT UPM

LIST OF FIGURES

© COPYRIGHT UPM

© COPYRIGHT UPM

LIST OF APPENDICES

© COPYRIGHT UPM

LIST OF ABBREVIATIONS

a*	Redness
ANN	Artificial Neural Network
ANOVA	Analysis of Variance
b*	Yellowness
BLAST	Basic Local Alignment Search Tool
BOC	Baseline of Correction
CIELAB	Commission International on Illumination
CTAB	Cetyltrimethylammonium bromide
DNA	Deoxyribonucleic Acid
EI	Electrical Impedance
ELISA	Enzyme-linked Immunosorbent-assay
FS	Feature Selection
GDP	Gross Domestic Product
IGRS	International Rubber Study Group
IR	Infrared
ITS	International Transcribed Spacer
KNN	k-Nearest Network
L*	Lightness
LSD	Fisher's Least Significant Difference
MC	Moisture Content
MIR	Mid Infrared
MS	Mass Spectroscopy
MSC	Multiple Scatter Correction
NMR	Nuclear Magnetic Resonance
OA	Overall Accuracy

© COPYRIGHT UPM

PCR	Polymerase Chain Reaction
PDA	Potato Dextrose Agar
PLS	Partial Least Square
R^2	Coefficients of Determination
RF	Random Forest
RM	Ringgit Malaysia
RMSEC	Root Mean Square of Calibration
RMSEP	Root Mean Square Error of Prediction
ROC	Receiver Operating Curve
RRIM	Rubber Research Institute Malaysia
SD	Standard Deviation
SG	Savitzky-Golay
SNV	Standard Normal Variate
SVM	Support Vector Machine
SWNIR	Short-wave Near Infrared
TSC	Total Solid Content
UVNIR	Ultraviolet-Near Infrared
Vis	Visible
WRD	White Root Disease
%	Percentange
CO_2	Carbon Dioxide
H_2O	Water

© COPYRIGHT UPM

© COPYRIGHT UPM

CHAPTER 1

INTRODUCTION

1.1 Background of the Study

Malaysia is ranked as the sixth largest producer of natural rubber, with the production value of 3.48 million tons per year of natural rubber (MRB, 2024). The rubber plantation significantly contributes to Malaysia's gross domestic product (GDP) growth, contributing 1.8 billion MYR per annum. The export of natural rubber decreased from MYR 7,219 million in 2022 to MYR 6,032 million in 2023 (MRB, 2024). About 1.135 million ha of land in Malaysia has been allocated for rubber plantation (MRB, 2024). However, the rubber tree is susceptible to white root disease (WRD), which is considered as the primary issue of natural rubber production (Andrew et al., 2021). According to a recent survey, rubber plantations experienced up to 40% loss of their rubber trees due to WRD infection (Saidi et al., 2023).

A necrotrophic fungus, *rigidoporus micropores* cause WRD. It is the most damaging agent for rubber trees because it can kill them slowly (Khairuzzaman et al., 2017). The pathogen lives and remains in the soil for a long time by producing white mycelia. Healthy rubber planted near infected trees can be infected directly by root contact between them (Chaiharn et al., 2019). The infection mechanism of this fungus involves the release of extracellular enzymes that can decay the root. Consequently, it directly affects plants' vascular circulation (Kaewchai et al., 2010). As a result, the nutrients and water utilization are blocked, which can adversely affect photosynthesis (Nakaew et al., 2017).

© COPYRIGHT UPM

After being infected by pathogens, crops typically do not exhibit obvious visual symptoms in the early stages. However, a series of physiological and biochemical changes begin to occur, leading to different spectral responses. Leaf discoloration is a main above-ground symptom of WRD and is often considered the primary visual indicator for identifying the infection. Unfortunately, when the leaf discoloration happens, the infection have progressed to an advanced stage. Dead branches and dieback symptoms are signs of severely infected rubber trees, which eventually destroy the canopy (Fernando et al., 2012). Dead trees typically have orange-red basidiocarps developed at their collars.

Conventionally, visual inspection is one of the most common methods employed by experts to detect WRD infection in rubber plantations (Omorusi et al., 2014). Although the visual inspection technique is the most common method, it might become costly as it requires a continuous inspection by an expert (Ray et al., 2017). The visual observation methods can lead to inaccurate predictions due to potential biases, even conducted by experts, and are impractical for monitoring large farms. In addition, molecular methods such as PCR are widely used to amplify specific DNA sequences of *Rigidoporus micropores* to detectable levels (Andrew et al., 2021). The presence of the pathogen is confirmed by visualizing the amplified DNA through gel electrophoresis. Additionally, these methods require consumable reagents specifically designed to detect pathogens. While molecular techniques are powerful tools for confirming the presence of WRD, they are unsuitable for preliminary screening of large plant samples due to the lengthy process involved.

WRD has been recognized as the primary cause of death for rubber trees. However, current detection and control methods are still not practical. The only way to stop this

© COPYRIGHT UPM

disease from spreading to other areas within the plantation is to remove any remaining infected trees (Mahyudin et al., 2023). The methods of trenching, plowing, harrowing, clearing and burning of the soil are used to achieve this isolation before new seedlings are planted. Therefore, early identification and detection of WRD infection are essential to minimize output losses and lower plantation management expenses. These techniques represent a major advancement in detecting WRD in rubber plantations. After being infected by pathogens, crops typically do not exhibit obvious visual symptoms in the early stages. However, a series of physiological and biochemical changes begin to occur, leading to different spectral responses. Currently, the spectroscopic method is rapid and non-destructive technology that has been made for the agricultural sector to address the need for the detection and monitoring plant disease (Ali et al., 2019). Several studies have shown that visible shortwave near-infrared (VSNIR) spectroscopy helps identify plant diseases (Prasetyo et al., 2024; Zhao et al., 2024; Atanassova et al., 2019), physiological dysfunction (Shakeel et al., 2021; Mogollón et al., 2020; Dhaliwal et al., 2024), and membrane damage in plant tissue (Du et al., 2019; Haynes et al., 2024; Babellahi, 2020). Nevertheless, future research is needed to determine how effective this approach to detect WRD disease infection in rubber trees plantation.

1.2 Problems Statement

WRD is a destructive disease of rubber trees. WRD causes major losses in natural rubber production (Cerda et al., 2017). Therefore, minimizing agricultural losses requires early disease detection (Devaraj et al., 2019). From the literature, leaves samples (Hadzli et al., 2014; Abdullah, 2007; Hashim, 2010;), latex samples

© COPYRIGHT UPM

(Longsaward et al., 2023), and soil around rubber trees (Saidi et al., 2023) are the most utilized samples used by the scouting crew to detect the presence of WRD. This approach can detect the disease distribution easily because it depends mainly on observing the appearance of specific disease symptoms (Parker et al., 1995). For example, WRD infection tree cannot produce sufficient chlorophyll due to nutrient deficiencies. As the results, the infection trees show the yellowing in the leaves, also known as chlorosis. The discolouration of leaves determines chlorophyll's concentration and the status of plant health (Killiny et al., 2017).

Conventionally, various methods are currently used in the rubber plantation to detect the presence of WRD. These methods can be divided into three groups: visual inspection (Saidi et al., 2023), serological (Dalimunthe et al., 2017), and molecular (Andrew et al., 2021). These approaches have been conducted to effectively detect the presence of WRD but, there is no effective detection method. When the disease symptoms begin to appear, more than half of the internal tissue is already rotten (Akpaja & Ogbebor, 2020). The infection also causes rotten internal tissue, which leads to stem fracture and tree die. Thus, an early detection system which can detect the presence of the disease before the symptoms appear is needed by the industry. An early detection of WRD refers to identifying the disease before it reaches stage 1.

Currently, the spectroscopic method for the detection of plant disease complies with non-invasive, rapid, sensitive, and precise diseases, which have been considered for developing and designing early-stage infections (Kumar et al., 2021). Several studies have described the ability of the VSNIR spectroscopy technique to detect fungal disease (Liu & Li, 2022; Atanassova et al., 2023; Prasetyo et al., 2024), water stress

© COPYRIGHT UPM

(Das et al., 2021; Haynes et al., 2024; Zeng et al., 2022) and disease severity assessment (Luo et al., 2021; Khan et al., 2021; Mandal et al., 2022). However, studies on applying this technique on WRD detection in rubber trees are yet to be explored. VSNIR spectroscopy is scientifically appropriate for detecting WRD due to its ability to capture spectral information from both visible light and shortwave infrared regions. This range can detect changes in the plant's physiological state, such as variations in leaf water content, chlorophyll levels, and overall stress responses, which can indicate early signs of infection before visible symptoms appear.

When dealing with spectral data obtained from spectroscopic measurements, spectral data often have many variables (wavelengths), leading to high dimensionality. This can complicate data analysis and increase computational complexity. High dimensionality can also lead to overfitting in machine learning models, where the model becomes too specific to the training data and performs poorly on new data. Addressing these challenges often requires dimensionality reduction to optimize data processing time, reduce dimensionality, and enhance data generalization by reducing overfitting and prediction. Data reduction offers numerous benefits, such as reducing data storage requirements (Ayesha et al., 2020), decreasing computation time (Zebari et al., 2020), and eliminating redundant (Ray et al., 2021), irrelevant (Wang et al, 2021), and noisy data (Ougiaroglou et al., 2024). It can enhance data quality and improve algorithm performance, particularly for algorithms that struggle with high-dimensional data. Reducing dimensions simplifies classification processes and boosts efficiency and accuracy (Jia et al., 2022).

© COPYRIGHT UPM

Feature selection (FS) techniques are frequently applied for dimensional reduction on spectral data (Zang et al., 2023). FS refers to identifying and preserving the dataset's most relevant attributes or variables. These approaches have been found to produce high accuracy and can provide better prediction of experimental data (Vinutha et al., 2023). Thus, to ensure efficient data processing and reliable prediction, spectral data handling using computational intelligence techniques is needed in this study.

The use of machine learning (ML) techniques in crop production systems has grown significantly in recent years, particularly for the identification of plant diseases (Das et al., 2022). ML is the term for computational algorithms that can learn from data and carry out classification tasks. These algorithms are useful for finding patterns and trends in huge amounts of data, including spectral data (Zhang & Wang, 2023). This study used k-nearest neighbor (kNN), support vector machines (SVMs), random forest (RF) and artificial neural networks (ANNs).

It is essential to highlight that it can be challenging to recognize the presence of WRD infection on rubber trees in its early stages in the field by visual symptoms. The basidiocarp appears at a late stage, making treatment ineffective by then. Therefore, an expert assessment is necessary. Laboratory analysis must be conducted on the rubber trees to confirm the infection. This highlights the need for more practical systems for WRD disease identification. Thus, it is now crucial to create a non-destructive method that provides a powerful, non-destructive tool for early detection of WRD in rubber trees, which can significantly reduce the economic and environmental impact of the disease.

© COPYRIGHT UPM

1.3 Objectives of the Study

The goal of this study was to investigate the capability of a low-cost and visible short-wave near-infrared spectrometer (200 nm – 1200 nm) in combination with FS methods and machine learnings of ANN, SVM, RF and kNN to be employed as an early detection method for identifying the presence of WRD in rubber trees based on leaves, latex, and soil around rubber trees. The specific objectives were:

i. To characterize leaves and latex properties of healthy and infected rubber trees.

ii. To characterize the properties of soil around healthy and infected rubber trees.

iii. To determine the spectral reflectance of the leaf samples for healthy and infected rubber trees at different severity levels.

iv. To compare the performance of different machine learning methods using different feature selection techniques for classifying disease severity levels in rubber trees.

v. To determine the best classification model for early detection of WRD using machine learning techniques.

1.4 Scope and Limitations of the Study

This investigation was conducted under particular circumstances and limitations. The following is a list of this research's limitations and scope:

i. The rubber tree samples were collected from two different age groups: 5-year-old trees at Sungai Buloh and 10-year-old trees at Kota Tinggi.

© COPYRIGHT UPM

ii. The reflectance spectra of rubber leaf samples using spectroscopy were measured from 200 nm to 1200.

iii. Only the reflectance spectrum of leaves were measured, as the leaf samples showed the highest correlation with disease severity compared to soil and latex samples.

1.5 Significant of the Study

This study investigated the applicability of reflectance spectral data for classifying WRD disease into different severity levels. This approach can detect WRD infection in rubber trees before visible symptoms appear. Early detection helps prevent the spread of WRD, minimize crop loss, enable timely intervention with precision treatments, and reduce the need for excessive pesticide use. Feature selection eliminates irrelevant or redundant information, ensuring that machine learning models focus on the most significant data. This leads to a more accurate classification of disease severity and better generalization across diverse datasets. Algorithms such as Chi-square, RELIEF, and ANOVA have proven effective in maintaining high classification accuracy in plant disease identification. Machine learning algorithms can efficiently classify disease severity across large datasets, which is especially useful in precision agriculture and large-scale farming. These algorithms can learn patterns from spectral or other sensor data to differentiate between healthy plants and those affected by diseases at varying severity levels.

© COPYRIGHT UPM

CHAPTER 2

LITERATURE REVIEW

Given the critical need for early detection of WRD in rubber plantations as discussed in Chapter 1, it is essential to evaluate existing detection methods to determine their effectiveness and limitations. Chapter 2, therefore, presents a comprehensive review of existing methods for detecting WRD, highlighting their effectiveness and limitations. Additionally, it aims to identify gaps in current knowledge and explore opportunities for improving detection techniques.

2.1 Overview of the Rubber Industry

A rubber tree (*Hevea brasiliensis)* is native to South America, produces latex, a milky sap that is the source of natural rubber (NR). NR is a polymeric hydrocarbon with exceptional tensile characteristics that is mainly made up of poly (cis-1,4-isoprrene) chains. The Amazon River basin is the original home of a rubber tree where latex can be harvested after six years of growth (Nair & Nair, 2021). NR can coagulate in air when the rubber particle stabilizes (Mekonnen et al., 2019). Ammonia is anticoagulant that can be added to NR latex to prevent coagulation (Nugraha et al., 2024).

In 2023, global natural rubber production reached nearly 28.8 million mt (MRB, 2024). This represents a significant rise from 2000, when worldwide production was approximately 6.8 million metric tons. Currently accounting for over 90% of global rubber production, the top producers are Thailand, Indonesia, Vietnam, China, India, Cote D'Ivoire, and Malaysia (Ho et al., 2023). The coronavirus (COVID-19) outbreak

© COPYRIGHT UPM

in 2020 had a major negative effect on the sector, resulting in a five percent annual decline in global production to 27.4 million metric tons. However, production recovered in 2021 and 2022, reaching a value of 29.6 million metric tons globally in 2022.

NR is a vital raw material in various industries and contributes significantly in the manufacturing of modern goods. Because of its special qualities, which include high strength, remarkable fatigue resistance, high resilience, low damping, and little sensitivity to strain effects in dynamic applications, is highly valued in engineering applications (Sethulekshmi et al., 2022). These characteristics make NR particularly suitable for the automotive tyre industry, which accounts for nearly 70% of global NR consumption. NR is mainly utilized in the manufacture of tyres for trucks, earthmovers, and aeroplanes because of its high degree of flexing, which produces a lot of heat that is difficult for big tyres to escape (Singha et al., 2019).

2.2 Rubber Plantation in Malaysia

The Euphorbiaceae family's rubber tree, *Hevea brasiliensis*, is one of the world's most significant industrial plantations in terms of economic importance. Around 14 million hectares of land are planted with rubber worldwide at the moment (IRGS, 2021). Only *Hevea brasiliensis*, out of the ten species in the genus Hevea, is grown for rubber on a commercial basis. Approximately 13 million tons of rubber were produced worldwide in 2020 (IRGS, 2021). Thailand, Indonesia, Vietnam, and Malaysia account for 70% of the world's natural rubber supply, with rubber-growing nations in

© COPYRIGHT UPM

Asia (88%), Africa (10%), and Latin America (2%), contributing the most to global natural rubber production (IRGS, 2021).

NR has distinct physicochemical properties that are unmatched by synthetic polymers, which is why its yearly worldwide demand is rising. The military, industrial, transportation, aviation, medical, and consumer sectors all have a significant demand for natural rubber. Rubber wood, aside from latex, is a great raw material for making wood products including plywood, particle board, sawn wood, and furnishings (Hussin et al., 2022).

Nevertheless, since the 1980s, 1.47 million hectares fewer have been used for rubber production. Consequently, due to a lack of raw materials, several processing facilities were forced to close (Killmann, 2001). A main obstacle in rubber cultivation leading to decline in rubber yield is the presence of pathogens. These pathogens adversely affect crop quality and quantity, contributing to the deterioration of rubber yields. Fungi-related diseases are prevalent in tropical areas. According to a recent survey, the occurrence of WRD in the field varied between 5 % and 40 % (Saidi et al., 2023). WRD poses a significant threat in West Africa and Thailand, causing approximately 50% of yield losses in mature plantations (Nakaew et al., 2017).

2.3 Typical Disease of Rubber Tree

Diseases are one of the limiting factors for rubber tree production because they can reduce the quality of latex and cause significant losses in the rubber plantation industry (Mahyudin et al., 2023). Rubber trees are prone to various fungal diseases at all parts

© COPYRIGHT UPM

of the plant, including leaf, stem, panel, and root. However, two critical organs, leaf and root, are the most affected, and severe infection may significantly reduce yield.

2.3.1 Typical Leaf Disease of Rubber Tree

Leaf disease is challenging in rubber plantations in rubber-producing countries. Leaf fall disease is believed to be caused by many pathogens that can attack individual rubber individually trees or the whole plantation area and cause considerable yield losses. In rubber, leaf diseases are mainly caused by Ascomycete fungi. These pathogens cause diseases that manifest symptoms, which include die-back, blight, powdery mildew, leaf spots and curls, and anthracnose. More than ten diseases are identified in rubber nurseries and plantations worldwide, caused by different genera and species of pathogens. However, the most reported in the literature review are SALB*, Fusicoccum* leaf Blight, and *pestaloptiosis* leaf fall.

2.3.1.1 South America Leaf Blight

South American leaf Blight (SALB) is caused by *Pseudocercospora ulei* (Henn.) SALB severely decreased the yield of South American rubber plantations by more than 90%. This fungus is constricted to Central and South America and is not found in rubber plantations in Asia or Africa (Mahyudin et al., 2023). However, climatic conditions in Asian and African rubber-growing areas are analogous to those in America, making planted rubber extremely vulnerable to SALB if it is introduced in those areas.

© COPYRIGHT UPM

SALB infection causes repeated defoliation, shoot die-back, and mortality of older trees. It affects petioles, stems, shoots, flowers, and fruits. However, young leaves are the most susceptible. Heavily infected young tree leaves will drop off, while on surviving matured leaves, dark-colored raised bodies called perithecia are observed (Figure 2.1).

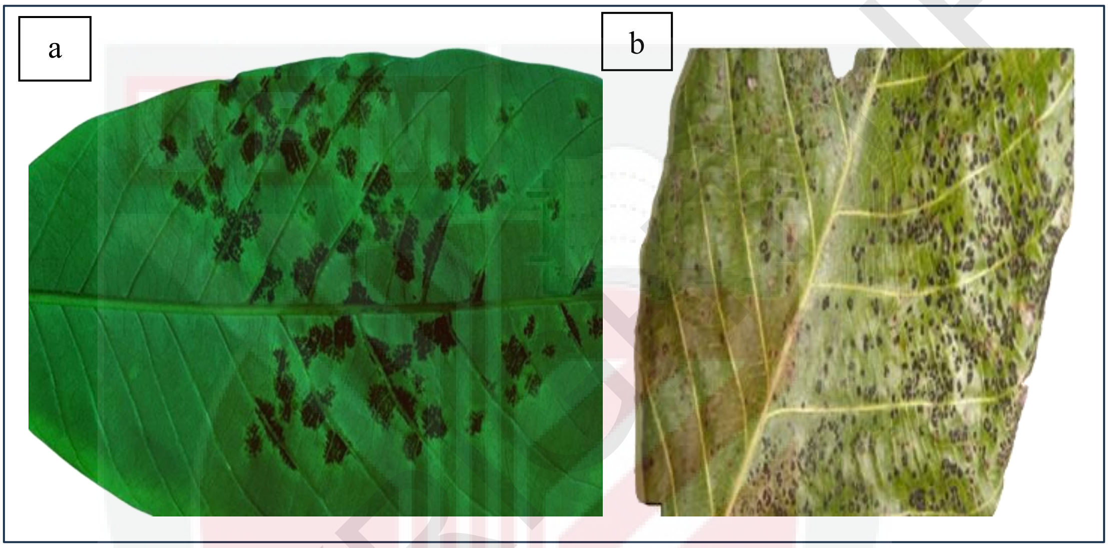


Figure 2.1: (a) Infection of Pseudocerospora ulei on Matured Leaves. (b) The Perithecia on the Surface of Leaves

2.3.1.2 *Fusicoccum* Leaf Blight

Fusicoccum leaf blight is caused by a fungus called *Neofusicoccum ribis* (Solpot et al., 2023). This pathogen is known for its ability to thrive in moist environments, making rubber plantations in tropical and subtropical regions particularly susceptible. The fungus spreads through spores, which are dispersed by wind, rain splash, and possibly insects. High humidity. Frequent rainfall and dense canopy cover create ideal conditions for the pathogen to infect and proliferate on rubber trees.

The pathogen causes large lesions with concentric brownish zones and rusty brown pin-head size spots on foliar parts of the rubber tree (Figure 2.2). The disease lesions measuring 0.5 – 3.0 cm in diameter resemble anthracnose with wave-like concentric brownish zones. It begins as rusty brown pin head-sizes spots, later developing into sinuous brownish lesions. The fungus also produces characteristic black pycnidia embedded within the lesion. Subsequently, infected young leaves drop odd after four months of infection.

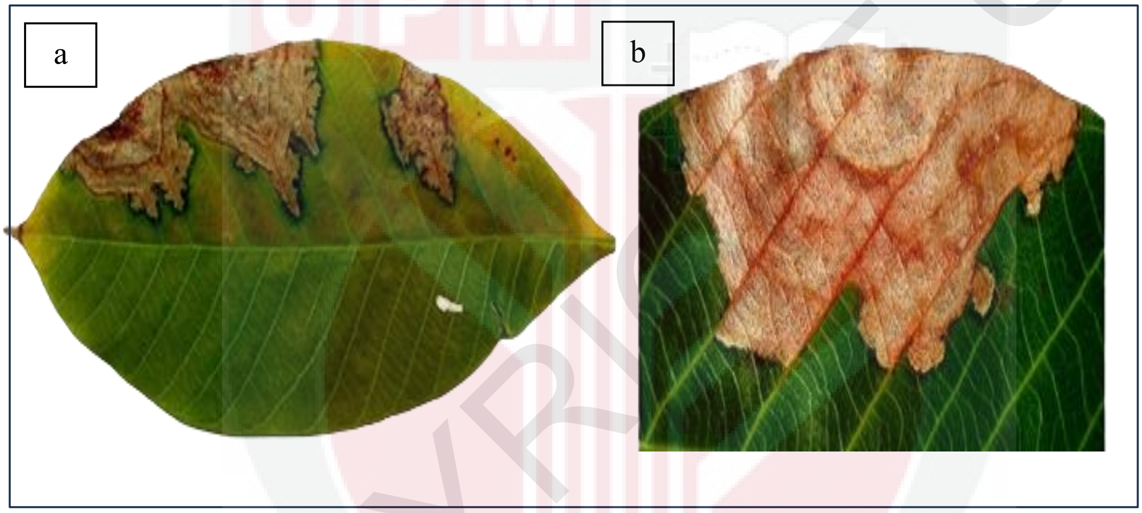


Figure 2.2: (a and b): Lessions of Fusicoccum Leaf Blight

2.3.1.3 Pestaloptiopsis Leaf Fall

Rubber leaf fall disease caused by *Pestaloptiopsis* sp is a new disease in rubber plantations (Damiri et al., 2022). This disease was first reported as an outbreak in rubber plantations in Malaysia in 1987 and 2003. Apart from Malaysia, this new disease has been identified as attacking rubber plantations in Indonesia, Sri Lanka, India and Thailand. *Pestaloptiopsis* sp is an airborne pathogen that spreads very

© COPYRIGHT UPM

quickly and can result in the fall of rubber leaves so that the plant winters throughout the year and decreases rubber production (Damiri et al., 2022).

The characteristic of the disease are dark to light brown spots observed on matured leaves only (Figure 2.3). The spots continue to widen to 1 – 2 cm, and then the tissue around the spots undergoes necrosis. Severe attacks will cause leaf fall until the plant crown is bare (Hadi et al., 2024). This disease causes continuous defoliation up to 75 – 90%, the canopy becomes thinner, and production drops to 25 – 45%.

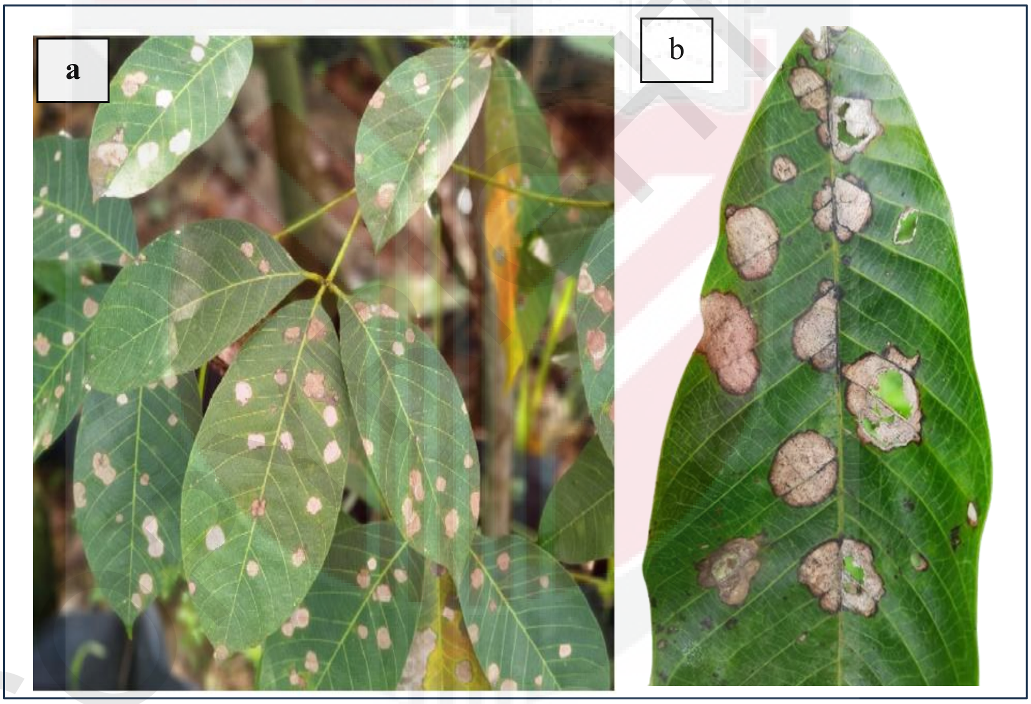


Figure 2.3: (a and b) Infected Rubber Leaves by *Pestaloptiopsis* Leaf Fall Disease

2.3.2 Typical Root Disease of Rubber Tree

Root diseases are the most destructive conditions that affect rubber trees' overall health, productivity, and lifespan. These diseases are often caused by pathogenic fungi that infect and degrade the tree's root system, leading to poor nutrient and water absorption, stunted growth, and, in severe cases, the tree's death. WRD is the most prevalent and damaging root disease of rubber trees, and it is responsible for significant losses in rubber plantations compared to root diseases.

2.3.2.1 White Root Disease

WRD is caused by *Rigidiporus micropores* fungus. It kills the host plant, later shifting to a saprotrophic phase that degrades the wood. WRD infection is established through root contact with a disease source, such as roots of infected neighbouring rubber trees, a dead stump, or infected wood debris buried in the ground (Chaiharn et al., 2019). WRD infection affects the xylem, which causes a significant problem in water and nutrient absorption of the tree.

The above-ground symptoms include foliar discoloration and tree defoliation. A change in canopy color indicated that the pathogen attack had already progressed to a late stage. Severely infected rubber trees exhibit dieback symptoms with dying branches until the whole canopy is destroyed due to the pathogen killing the host tree (Fernando et al., 2012). Orange-red basidiocarps are usually visible at the collar of dead trees. More than half of the internal tissue is already rotten when the disease symptoms appear, as shown in Figure 2.4.

© COPYRIGHT UPM

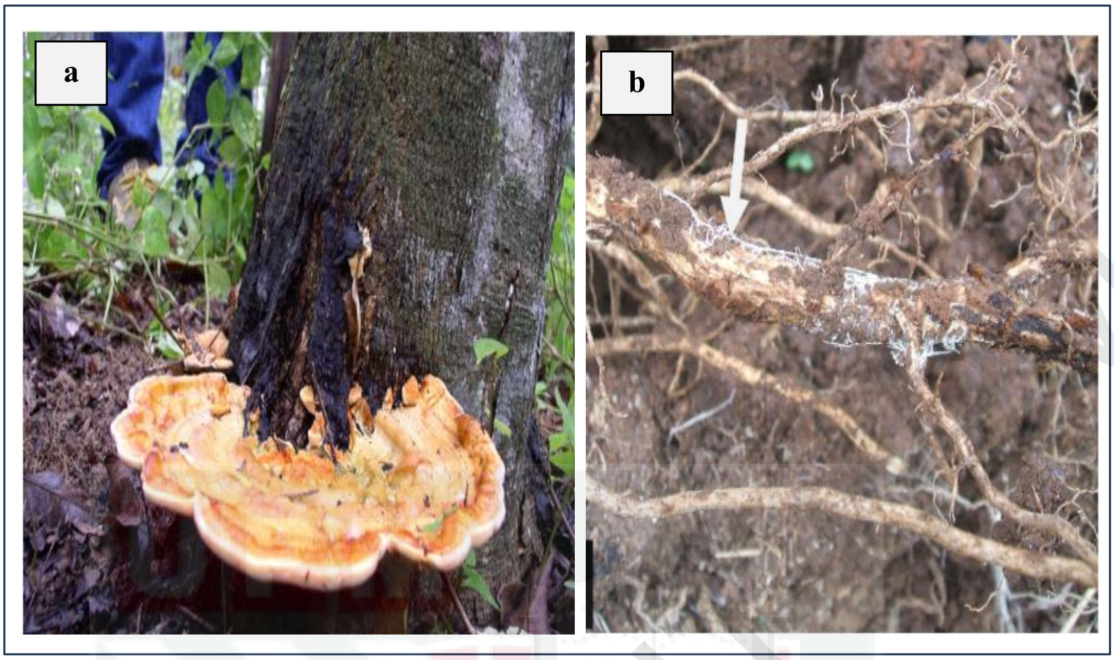


Figure 2.4: (a) Rhizomorph at the Root and (B) Fruiting Body at the Color of the Dead Stem of *Rigidoporus Micropores*

WRD is particularly serious compared to other diseases in rubber plantations because it targets the root system, which is crucial for water and nutrient absorption. Once the roots become infected, the tree's health deteriorates rapidly, often resulting in death. The disease spreads speedily through root contact between infected and healthy trees. Additionally, WRD is challenging to detect early, as symptoms typically appear only when the disease has progressed significantly. When signs like yellowing leaves or stunted growth are noticeable, the roots are already severely infected. This contrasts with other diseases, where early symptoms are more evident, allowing for quicker intervention.

2.3.2.2 Brown Root Disease

Brown root disease is caused by *Phellinus noxius* fungus. The name brown root disease refers to a brown to black mycelial crust formed by the fungus on the surface of infected roots and stem bases (Figure 2.5). *Phellinus noxius* fungus is a facultative parasite that obtains its nutrients from dead and dying plant tissue. The enzyme secretion from *Phellinus noxius* breaks down the lamella and cell wall of the plant.

Figure 2.5: A Brown-Black Mycelial Crust Formed by Fungus on the Surface of the Infected Root and Stem Base

The symptoms of the brown root disease on the rubber tree are typical as caused by many root disease pathogen, which are slowing the plant growth, yellowing and wilting of leaves, branch dieback and plant death. The above-ground symptoms are caused by the root and butt rot that constraint the uptake and transport of water and nutrient from soil (Mazlan et al., 2019). The visible rot at the fallen trees is another

© COPYRIGHT UPM

general sign that the tree is affected by the disease; the deadwood starts to discolor reddish-brown and finally becomes crumbly, dry, and white.

2.4 Characteristics of White Root Disease

2.4.1 Symptoms of White Root Disease

The literature review reported that the symptoms of WRD on a rubber tree could be recognized through its physical appearance. Foliar discolouration and tree defoliation are among the above-ground signs that are thought to be the leading indicators of root rot disease infection, including WRD. The discoloration of leaves is caused by the disappearance of chlorophyll (Mittelberger et al., 2017). The leaf color is often used to indicate the healthy status of a plant. These leaves then curve downward and buckle, resembling typical wilting symptoms. At later stages, leaves become orange–reddish (the colour may show slight variations depending on the clone) and defoliate within a short period. Rubber trees that are severely infected show dieback symptoms, such as losing branches that eventually destroy the entire canopy because the infection kills the host tree. The collar of deceased trees typically displays orange-red basidiocarps. The airborne basidiospores that the basidiocarps create spread and infect tree wounds or stumps (Figure 2.6 and Figure 2.7).

© COPYRIGHT UPM

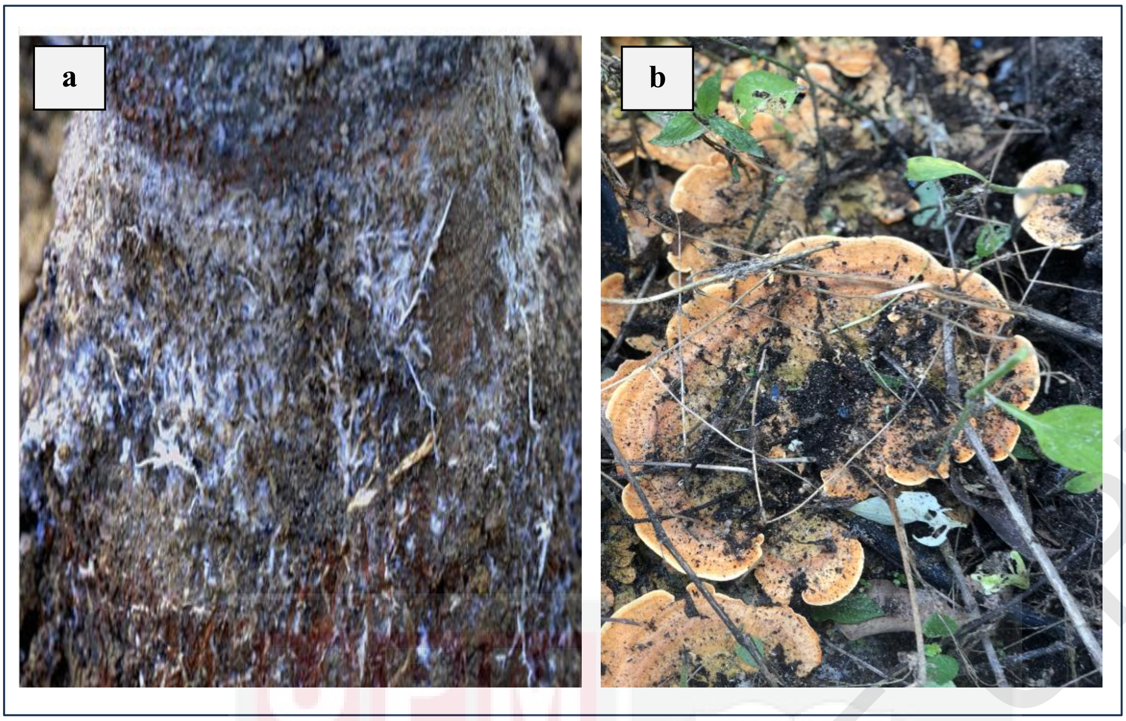


Figure 2.6: (a) Strands Rhizomorphs on the Root Collar: (b) Basidocarps of *Rigidoporus micropores*

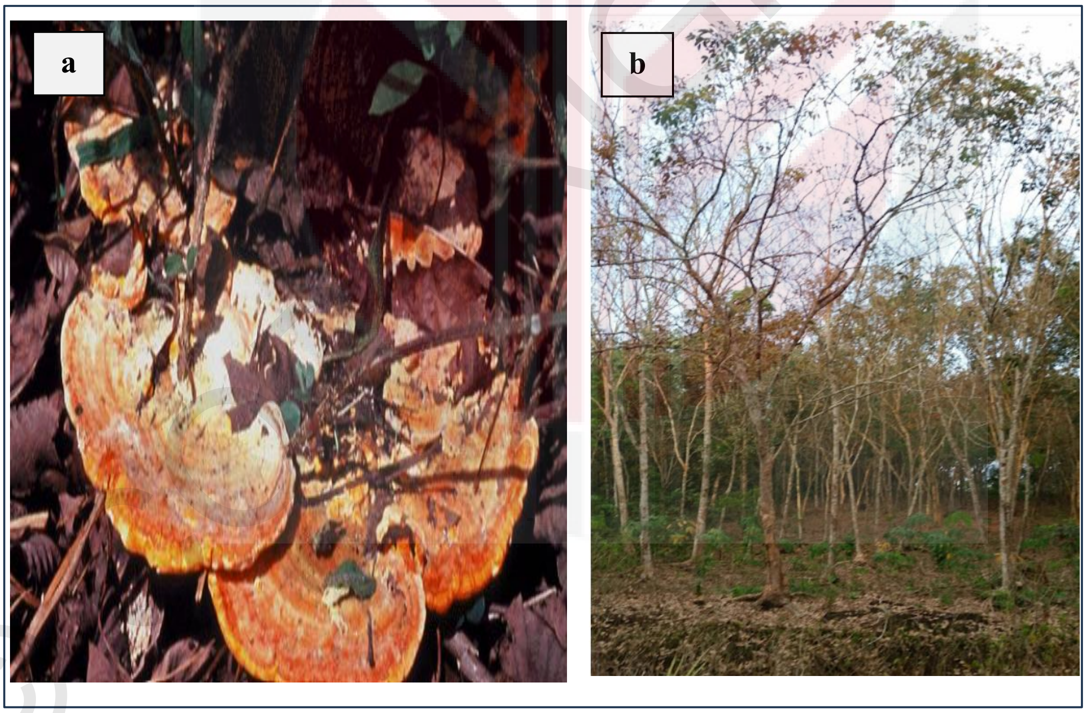


Figure 2.7: Field Symptoms of WRD: (a) Orange-red Basidiocarp at the Base of the Tree, (b) Rubber Tree Infected by WRD Showing Die-back Symptoms

Symptoms of WRD could also be determined through laboratory analysis based on latex samples (Khairuzzaman et al., 2017, & Yusuf et al., 2018). Typically, latex is composed of 50–80% water, 25–45% dry rubber, and 2–5% non-rubber particles (Zhao et al., 2010; Julrat et al., 2012). Total soluble solids (TSS) in latex is the term used to describe the amount of dissolved solids in the liquid; this is frequently stated as a percentage or in degrees Brix. When assessing the viscosity and stability of latex, among other qualities, this metric is crucial. The total weight of the final residue is used to determine the total solid content of that latex in the TS determination, which involves properly weighing a thin layer of latex and drying it in an oven until there is no further weight loss (Tillekeratne et al., 19880). The primary disadvantage of this approach is its time-consuming nature, and sample destruction is another matter to consider. The annealing process has been attempted to be accelerated with a microwave oven; however, the direct approach is only appropriate for usage in a strictly regulated environment.

Other than that, chlorophyll concentration and leaf color both significantly influence plant growth and appearance (Ma et al., 2024), according to studies by Iglesias et al. Dai et al. (2016), genes play a role in the variance of color and chlorophyll. Typically, the values of three variables from the CIELAB color system laboratory—L* (lightness), a* (redness), and b* (yellowness)—determine the color of the leaf. (Arabhosseini et al., 2011). The three most significant pigments in leaves are anthocyanin, carotenoid, and chlorophyll content (Croft & Chen, 2017). Green hue is typically attributed to chlorophyll, a pigment that is necessary for the conversion of light into chemical energy (Croft et al., 2017). Carotenoid is primarily responsible for the yellow color that results from the breakdown of chlorophyll, while enhanced

© COPYRIGHT UPM

anthocyanin production is the primary cause of the red hue. Chlorophyll is commonly employed as a reaction to environmental stress and nitrogen fertilizer administration. It primarily determines the photosynthetic rate and primary productivity in plants. Accordingly, chlorophyll concentration may be a valuable diagnostic marker for research on plant growth (Hotta et al., 1997).

2.4.2 Mechanism of Disease Infection Cycles

Necrotrophic *Rigidoporus micropores* destroy the host plant before transitioning to a saprotrophic phase that breaks down the wood. WRD infection is spread by root contact with a disease source, such as the roots of nearby rubber trees that are affected, a dead stump, or underground infected woody debris. Rhizomorphs, or flattened white mycelial strands that are 1 to 2 mm thick that grow in the soil and along the surface of the roots, sometimes extending several meters, are how the disease spreads locally. *Rigidoporus micropores* grow autotrophically as they progress. Saidi et al. (2023) state that for the rhizomorphs to continue to be alive and retain their virulence, they must remain in contact with the host. But according to Nadris et al. (1987), rhizomorphs can nevertheless infect the roots of healthy trees despite growing quickly in the soil without any wood substrate. After a rubber tree becomes infected and decays, it serves as a secondary source of inoculum, and the disease centers spread by infecting other nearby trees one after the other.

The infective hyphae of *Rigidoporus microporus* are the mechanism of infection, as described in multiple instances by Saidi et al. (2023) (Figure 2.8). The hyphae of *Rigidoporus microporus* produce significant cell damage and decay when they reach

© COPYRIGHT UPM

the tree taproot deep in the soil. The white rot of the wood is caused by the degradation of lignin through the activity of different hydrolases in the cell walls. In general, disease infection occurs in three stages: degradation, colonization, and penetration. The pathogen colonizes the tissues after penetrating the root system. The pathogen's mycelium then breaks down the cell structures of the host. The *Rigidoporus micropores* pathogen responsible for root rot necessitates frequent penetration and colonization of the host cell wall. The pathogen colonizes the tissues after penetrating the root system. The pathogen's mycelium then breaks down the cell structures of the host. The *Rigidoporus micropores* pathogen responsible for root rot necessitates frequent penetration and colonization of the host cell wall. To carry out its disease-infecting activities, *Rigidoporus micropores* either mechanically penetrates through colonized natural apertures or wounds or enzymatically breaks down the tissue, resulting in the differentiation of specialized structures (Nicole et al., 1986). After then, it releases extracellular enzymes that can break down the host plant's wood and collar (Siri-Udom et al., 2017). It has the potential to seriously harm a large area.

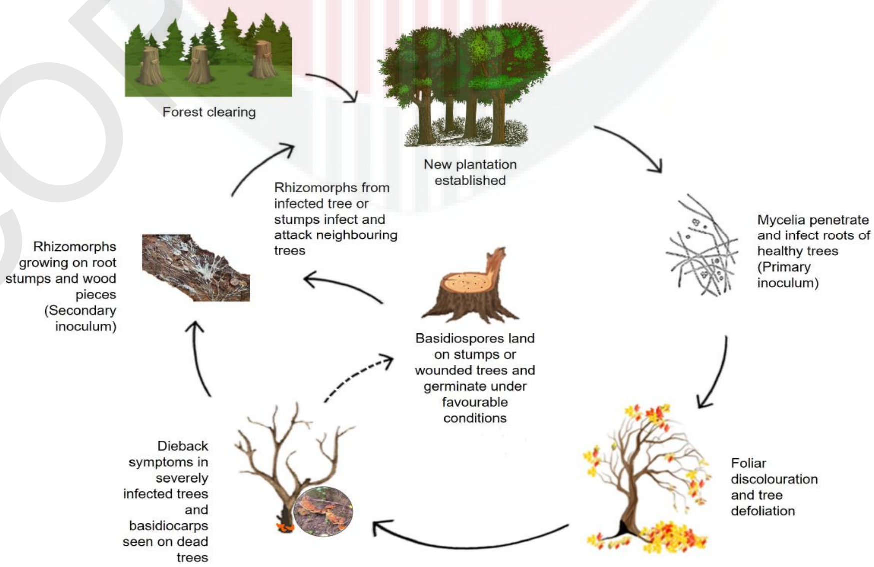


Figure 2.8: Mechanism of WRD Disease Infection Cycle

2.4.3 Severity Level of White Root Disease

Severity is a measure of a disease's intensity. It is the proportion of the appropriate host tissue or organ that is exhibiting disease signs (Del et al., 2024). Assessing disease severity is crucial in mitigating yield losses in plant pathology. A qualified professional visually inspects plant tissue to determine the severity of plant diseases (Wang & Wang, 2017). The severity of an illness can generally be categorized into five levels: healthy, mild, moderate, severe, and extremely severe infections (Emmanuel et al., 2021). This technique involves scanning the rubber's lateral roots for the pathogen's rhizomorph. Table 2.1 presents the scoring technique used to evaluate the prevalence and severity of WRD in rubber trees.

Table 2.1: Scoring and description of white root disease

Score	Category/rating of infection	Description
S0	No infection	No rhizomorph on stem base or lateral roots
S1	Light infection	Rhizomorph present, tissue penetrated
S2	Moderate infection	Portion of one lateral root penetrated
S3	Severe infection	More than one lateral root penetrated
S4	Very severe infection	Rhizomorph present on surface

2.4.4 Prevention and Control of White Root Disease

Traditionally, growers are advised to carry out routine checks to look for white rhizomorphs on the ground and foliar symptoms in older plants (Saidi et al., 2023). Placing a stick around the tree will reveal the presence of rhizomorph growth. If there is a positive infection, mycelium will grow on the stick, which can be observed after

© COPYRIGHT UPM

a few weeks (Omorusi et al., 2014). Methods such as mulching around a tree base followed by root exposure were recommended for rubber trees less than two years old to induce rhizomorph growth on the surface-contaminated rhizosphere (Fernando et al., 2011; Guyot & Flori, 2002).

Guyot and Flori (2002) propose that WRD can be managed both prior to and after planting. For recently infected trees and nearby healthy ones, recommended treatments include excavation, collar inspections, removal and cleaning of contaminated materials, and the application of chemicals such as quintozene or penconazole to protect the rubber tree collar (Lim et al., 1990). To reduce fungal inoculum in the soil, affected roots and stumps can be burned. Omorusi et al. (2014) suggest digging an isolation trench around the infected tree, cutting any roots that cross into it, and collecting and burning the debris. The trench should then be filled with loose dirt, and excavation should continue annually to remove and burn any new roots. This process can be repeated until the fungus on the affected trees stops spreading and its nutrient supply is depleted (Saidi et al., 2023). However, large-scale burning, as with other plantation practices, is difficult to control and may negatively affect soil nutrient stocks, soil carbon balance, and contribute to greenhouse gas emissions. Although this method can be somewhat effective, it is physically demanding when applied on a large plantation scale.

The effectiveness of fungicides in managing WRD in rubber trees depends on various factors, including the nature and application methods of the fungicides. Typically, chemical fungicides are used as soil treatments to reduce the inoculum of the soil-borne pathogen. A range of fungicides has been applied against WRD in rubber trees,

© COPYRIGHT UPM

such as sulfur amendments, hexaconazole, propiconazole (Wastie, 1957), triadimefon, triadimenol, PCNB (Jayasuriya & Thennakoon, 2007), phenol, and PCP (Jayaratne et al., 1997; Jayasinghe et al., 1995). The same treatments are often applied to neighboring trees as well. However, the prolonged use of chemicals can be costly, harmful to the environment and human health, and may lead to the development of resistance in the target fungus. These limitations have driven researchers to seek alternative management strategies.

Therefore, early plant pathogen detection is essential for plant health monitoring as part of management efforts to limit plant disease transmission and stop new ones from forming (Buja et al., 2021). When there are apparent signs, it is too late to receive curative treatment for WRD, therefore early detection is significantly more effective and financially feasible than the corrective method. Previously, spectroscopy was broadly employed for the early identification of diseases based on plant discoloration (Li et al., 2024). Therefore, WRD detection using spectroscopy could help to reduce the cost of fungicide application and avoid risks of health hazards and environmental pollution.

2.5 Existing Techniques for Detecting White Root Disease in Rubber Plantations

Numerous methods, covering from traditional to modern methodologies, have been employed to identify WRD in rubber trees. Table 2.2 illustrates the common technique for detecting WRD in rubber trees, which include visual assessment, laboratory testing, and contemporary methodologies. By visually inspecting the trees, infected rubber

© COPYRIGHT UPM

trees are identified based on the presence of pathological signs. For example, fungus fruit bodies appear on the above ground. This method makes it simple to identify the spread of the disease because it depends mainly on the observation of the appearance of specific disease symptoms (Parker et al., 1995). In the early stages of infection, however, this approach is costly, labor-intensive, time-consuming, and inefficient L(Akpaja and Ogbebor, 2020).

Table 2.2: Techniques for detection of plant disease

Disease detection in plants	References
Conventional method	Akpaja & Ogbebor, (2020)
Serological method	Dalimunthe and Wahyuni, (2017) Nakaew et al., (2015) Syafaah, (2019)
Molecular method	Andrew et al., (2021) Go et al., (2015) Kaewchai et al., (2010)
Imaging method	Khairulzzaman et al., (2017) Mohd et al., (2014) Azmi et al., (2016)

Laboratory-based techniques such as molecular and serological testing are commonly used to identify plant diseases. Polymerase chain reaction (PCR), enzyme-linked immunosorbent assay (ELISA), and fluorescence in situ hybridization (FISH) are the most common methods used for the detection and identification of plant diseases. These methods are based on directly detecting the pathogen or associated molecules, such as DNA or toxins (Panwar et al., 2023; Srivastava et al., 2020). These methods have high sensitivity, which enables early-stage disease detection, and high specificity, which ensures precise identification of fungal, bacteria, and viral pathogens (Sharma et al., 2024; Verma et al., 2022). However, these methods are labor-intensive and destructive techniques that require extensive sample preparation and are generally

© COPYRIGHT UPM

performed in a laboratory by highly skilled personnel (Loo et al., 2022; Chauhan et al., 2024; Jayamohan et al., 2021). For example, PCR method involves sample homogenization, extraction and purification of the DNA material, as well as gel electrophoresis and gel imaging, which add to the cost and duration of analysis of individual samples. Furthermore, PCR requires primer preparation and consequently determination of the DNA sequence of the pathogen of interest. Despite great progress in optimization of PCR-based disease diagnostics over the last few decades, this technique remains expensive for routine detection and identification of plant diseases. Also, PCR is highly sensitive to contamination from environmental DNA, can be inhibited by small quantities of organic solvents.

Imaging techniques have gained significant attention in detecting plant disease, including WRD. For example, Sulaiman & Saad (2020) classified the healthy and infected WRD . The accuracy from the selected most optimized models were greater than 70%. The selected most optimized models were then used to classify between healthy trees and white root infected trees based on single input categories. However, the main challenges in applying these imaging methods at scale remain the high costs, the need for specialized equipment, and the complexity of data interpretation (Laghari et al., 2022).

Spectroscopy techniques were also used for detection and identifying the presence of WRD in rubber plantations, which are listed in Table 2.3. As highlighted in the literature, accurate detection of WRD is critical for effective management. However, it remains a significant challenge for rubber tree growers. The most up-to-date research on early detection of WRD is still limited. But, early detection is crucial to

© COPYRIGHT UPM

minimize potential losses. Non-invasive approaches have gained considerable attention recently due to the limitations of existing methods (Sankaran et al., 2010; Lelong et al., 2010). Among these, spectroscopic techniques stand out because the high-resolution spectra they produce facilitate easier statistical analysis, making them a vital tool in plant disease diagnosis.

Table 2.3: Studies on WRD detection on rubber trees that demonstrate the application of several data analysis methods

Method	Statistical methods	Research gap	Reference
VIS-NIR spectrometer	t-test and normality test to discriminate between healthy and infected WRD	A moderate classification accuracy of 78% was obtained.	Ismail et al., (2013)
NIR spectroscopy	Receiver operating curve (ROC) for discriminating WRD levels	The overlapping spectral responses of mildly infected samples led to a moderate classification accuracy of 84%.	Mohd et al., (2014)
VIS-NIR spectroscopy	Statistical analysis (box plot)	Focus on visual symptoms only. Could provide a more comprehensive and accurate diagnosis of WRD.	Hashim et al., (2015)
Impedance Spectroscopy	Normality test to classify healthy and unhealthy WRD	Focuses on discriminating between healthy and infected WRD. However, the aspect of early detection is crucial for effective disease management.	Mohd Azmi et al., (2016)
NIR spectroscopy	skewness method to detect healthy and unhealthy samples	Only infected and healthy disease can discriminate successfully	Khairulzzaman et al., (2017).
Impedance Spectroscopy	t-test and normality test to discriminate healthy and infected WRD	Focuses on discriminating between healthy and infected WRD. However, the aspect of early detection is crucial for effective disease management.	Sulaiman, (2019)

© COPYRIGHT UPM

2.6 Spectroscopic Techniques in Detecting Plant Diseases

Spectroscopy is related to radiation energy and involves the interaction of the electromagnetic spectrum with matter at a certain wavelength (Alander et al., 2013). A spectrum series that serves as a representation of wavelength or frequency is typically used to indicate the spectral wavelength, as illustrated in Figure 2.9. A spectroscopic method is a broad approach that has a variety of disciplines and applications, particularly in the non-destructive field. In addition, several spectroscopic techniques, such as fluorescence and impedance spectroscopy, are discussed under the nature of interaction discipline for the identification of plant diseases.

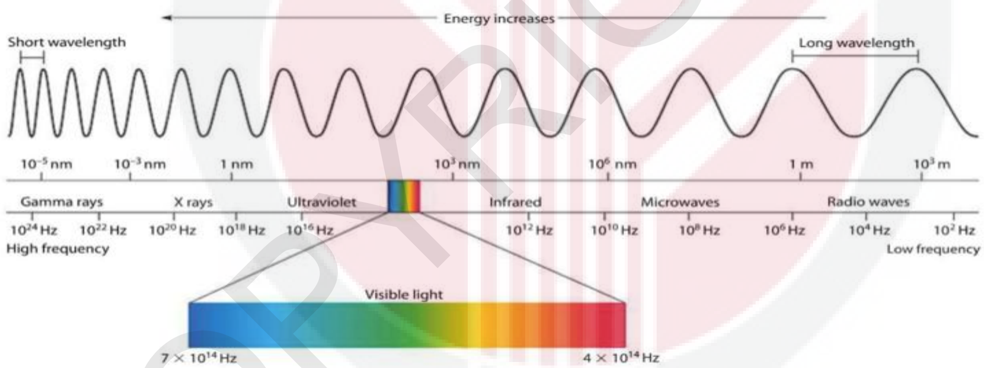


Figure 2.9: The Spectral Wavelength is Denoted by a Series Spectrum

Molecular and atomic spectroscopy techniques are widely distinguished according to how they are used. Additionally, they can be categorized according on how they interact. Visible (VIS), infrared (IR), nuclear magnetic resonance (NMR), mass spectroscopy (MS), and electrical impedance (EI) are a few types of molecular spectroscopy. Conversely, fluorescence spectroscopy (FS) is a component of atomic spectroscopy. Most of these spectroscopic methods have been extensively used to

© COPYRIGHT UPM

identify plant diseases. For example, the disease known as olive leaf spot (OLS) in olive trees was found using VIS/IR spectroscopy. spectroscopy, on the other hand, is used to discover diseases that impact the molecular structures and characteristics of plants, according to several studies. Spectroscopic analysis is a technology that has shown promise in the identification of plant diseases. When considered collectively, numerous studies have supported the use of spectroscopic techniques in assessing and tracking variations in the quality of food and agricultural products, chief among them being the early identification of plant diseases. Due to its advantages over other cutting-edge approaches, spectroscopy has lately been used to identify plant diseases. These advantages include affordability, simplicity, and speed. As a result, the use of spectroscopic techniques in plant disease diagnosis becomes essential.

2.6.1 Applications of Spectroscopic in Plant Disease Detection

A variety of methods are used to identify the symptoms or pathogens associated with various diseases. However, these methods are time-consuming and require highly trained personnel and specific equipment. As a result, the demand for rapid methods to detect the presence of plant disease at an early stage has been increasing. The spectroscopic approach is a unique disease monitoring method that has been used to detect diseases and stress due to various factors in plants and trees. Table 2.4 summarize a few studies on plant disease detection using the spectroscopic approach.

For example, Rivera-Romero et al. (2020) explored the use of visible and near-infrared spectroscopy for the early detection of powdery mildew (PM) infection in cucurbit plants. Their results showed that structure-independent pigment and red-edge

© COPYRIGHT UPM

reflection spots, with R^2 value of 0.211 and 0.1893, respectively, were linked to the best early detection predictions made throughout the sampling days. This work shows that reflectance spectra can be used to detect PM illness early in cucurbit plants that have comparable leaf features.

Table 2.4: Studies on the application of spectroscopy on plant disease detection

Spectroscopy	Wavelength region (nm)	Plant	Disease	Results	References
VSNIR	200 - 1200	Cucurbit	powdery mildew	R^2 = 0.211	Rivera et al., 2020
VSNIR	530 - 900	Citrus	Huanglong bing	QDA=92 %	Garcia et al., 2013
IR	450 - 1000	Tomatoe s and guava	plant-parasitic nematode infestation	Classifica tion accuracie s= 97%	San et al., 2020
Mid-infrared	500 – 4000	Oilseed rape	Healthy and infected	classifica tion accuracie s = 80%.	Zhang et al., 2017
VIS-NIR	350 - 1050	Red-berried	GLD	Classifica tion accuracy = 99%.	Sinha et al., 2019
NIRS	1000- 2500	Cassava	Cassava frog skin disease	Accuracy = 80%	Freiteas et al., 2020
VIS-NIR	550 - 1100	Tomato	Gray mold	Accuracy = 98%	Najjar & Abu, 2020
VIS-NIR	350- 2500	Peanut	Leaf spot	R^2 = 0.68	Guo et al., 2023
FT IR	2500 - 5000	Oil palm	Ganoderma infection	Accuracy = 90%	Liaghat et al. (2014)
IR	350 - 2500	Potato	Potato late blight	Accuracy = 83 %	Gold et al., 2020

© COPYRIGHT UPM

Sankaran et al. (2011) found Huanglongbing (HLB) in a citrus grove by using VIS/NIR spectroscopy. Their findings showed that QDA (Quadratic Discriminant Analysis) had an overall 92% accuracy rate in differentiating between leaves with HLB infection that were healthy and those that were symptomatic. In further research, VIS-NIR and mid-infrared spectroscopy methods were used by Sankaran and Ehsani (2013) to compare the classification of citrus leaves affected by canker and HLB illnesses. They discovered that mid-infrared spectra were better suited for diagnosing HLB, while VIS-NIR was useful for detecting canker. It is crucial to comprehend the viability of both spectroscopic methods because VIS-NIR provides a non-destructive way to diagnose diseases.

Similar to this, San-Blas et al. (2020) investigated the use of machine learning algorithms in conjunction with various infrared spectra analyses to identify plant-parasitic nematode infestation in tomatoes and guava. According to the authors, the model performed better than 97% of the time overall. The authors further claimed that by eliminating soil and root sampling, these methods could demonstrate significant benefits and shorten the processing time from field to laboratory. In a different investigation, Zhang et al. (2017) used mid-infrared spectroscopy to determine the spectral changes between oilseed rape leaves that were healthy and infected. Support vector machines (SVM) and extreme learning machines were also used by the authors, and the results demonstrated their efficacy for classification accuracies over 80%.

A further study by Sinha et al. (2019) assessed the suitability of VIS-NIR spectroscopy as a quick, reliable, and non-destructive technique for identifying GLD in a red-berried wine grape cultivar at different phenological stages. By employing multilinear

© COPYRIGHT UPM

regression and partial least square regression techniques, they discovered substantial differences between healthy and diseased leaves in both the visible (351, 377, 501, 526, 626, and 676 nm) and near-infrared (701, 726, 826, 901, 951, 976, 1001, 1027, 1052, and 1076 nm) areas. Using a quadratic discriminant analysis (QDA) classifier, the study's overall classification accuracy ranged from 75% to 99%.

In 2020, Freiteas et al. explored the use of NIRS in the diagnosis of cassava frog skin disease. Support vector machine with linear kernel, bayesian generalized linear model, parallel random forest, extreme learning machine, high dimensional discriminant analysis and PLS model were the six classification models used in this study. These models showed an overall accuracy of more than 80%, demonstrating remarkable efficiency in differentiating between unhealthy and healthy accessions. According to these findings, non-invasive radiography presents a feasible substitute for conventional diagnostic techniques in the identification of cassava frogskin disease. Its advantages include early and precise detection, rapidity, and affordability. Another study by Najjar et al. (2020) employed VIS-NIR spectroscopy to detect grey mould on tomato fruits in the early stages. They found that VISNIR could completely distinguish between infected and healthy samples before symptoms appeared, with overall accuracy of 98%. This suggests that VIS-NIR is a promising technique for the early detection of plant tissue diseases.

UV-NIR spectroscopy was utilized by Cheng et al. (2019) to identify leaf spots on peanuts, which have a major effect on peanut production. To create a hyperspectral vegetation index tailored to leaf spot detection, they first selected sensitive bands. According to their findings, when the disease index (DI) rose, canopy spectral

© COPYRIGHT UPM

reflectance dramatically dropped in the near-infrared regions (NIR) ($r < -0.90$). The spectral indicator that was most useful in identifying leaf spots in peanuts was LSI: NDSI (R938, R761), with regression model R^2 values as high as 0.68. This study shows that detecting leaf spot illnesses in peanuts may be done with accuracy, dependability, and efficiency using hyperspectral data processing.

Liaghat et al. (2014) used Fourier transform infrared (FT-IR) spectroscopy in the spectral region of 2550 to 25060 nm to construct a statistical model to classify the four categories (healthy, mild, moderate, and severe) of Ganoderma infection in oil palm trees. After spectral preprocessing, statistical multivariate classification models were created using linear discriminant analysis to test the viability of FT-IR spectroscopy. The classification model distinguished between healthy and diseased oil palm leaves with a high degree of accuracy—roughly 90%.

Gold et al. (2020) investigated the spectral response *Phytophthora infestans*, the causative agent of potato late blight. In this study, the spectral response of four potato cultivars, including two cultivars sharing a common genotypic background with the exception of a single resistance gene, following inoculation with *P. infestans* clonal lineage US-23 were collected using spectroradiometer. Three-class discrimination between control, pre-symptomatic, and post-symptomatic samples was performed using random forest classfier. The result showed a higher accuracy of 85%.

In light of these points of view, spectroscopy has shown to be among the best methods for identifying olive leaf spot infections at greater severity levels, offering insightful advice and proof regarding the yield of rubber trees.

© COPYRIGHT UPM

2.7 Spectral Data Analysis

SWNIR instruments are constructed up of huge spectrum data sets that require dependable, quick, robust, and effective data processing techniques in order to extract pertinent information. Usually, the relevant information embedded within a large amount of spectral data overlaps with overtones band of NIR spectra. The number of physical, chemical, and structural variables influence factors on overlapped band NIR spectral which is difficult to select useful analytical information (Zahir et al., 2022).

2.7.1 Pre-processing of Spectral Data

The spectral data of solid samples are influenced by the physical properties (texture, color, and form) of the samples. This led to the challenging process of the classifying diseases. Moreover, the spectral data typically include background information such as path, length fluctuation, random noise, light scattering, and sample information. To obtain robust, accurate, and reliable calibration models, spectral data must be pre-processed or pre-treated before a model can be created. To eliminate any unnecessary data, including interactions, noise, uncertainty, variability, and unidentified features, it is imperative to apply spectrum pre-processing techniques. Recently, several pre-processing methods have been developed. Nicolai and colleagues (2007) divided the pre-processing methods into the subsequent group:

2.7.1.1 Smoothing

Noise from NIR spectra can be removed randomly using smoothing techniques. Zhao et al. (2021) claim that optimizing the signal-to-noise ratio also requires smoothing.

© COPYRIGHT UPM

Moving average filters and the Savitzky-Golay algorithm were the most often utilized smoothing techniques (Gajbhiye et al., 2020). Smoothing eliminates irrelevant data while improving the spectra's appearance (Xie et al., 2022).

2.7.1.2 Standardization

Standardization procedures can be used to divide the spectrum by the wavelength standard deviation at each wavelength. Typically, the variance of each wavelength is standardized to 1, giving each variable in the model an equivalent amount of influence (Osco et al., 2022).

2.7.1.3 Normalization

Since the degree of scattering depends on the radiation's wavelength, particle size, and refractive index, a normalization technique is used to account for additive (baseline shift) and multiplicative (tilt) effects in the spectral data. These effects are caused by physical phenomena like non-uniform scattering throughout the spectrum. The two most widely used normalizing methods are standard normal variate (SNV) and multiple scatter correction (MSC) (Chang et al., 2015). By linearizing each spectrum to an "ideal" spectrum of the sample—which, in reality, corresponds to the average spectrum—MSC could mitigate the impacts of scattering. Every spectrum in SNV is standardized to a zero mean and a unit variance.

Diffuse reflectance spectroscopy provides valuable structural and chemical information, but its accuracy can be greatly hampered by light scattering (mostly due to particle size) and non-linearity in the instrument response (Wan et al., 2020). The

© COPYRIGHT UPM

idea behind scattering theory is that the reflectance signal should be affected by scattering in a multiplicative way. As a result, there will be a significant, fluctuating background from differential scattering at every wavelength in the recorded spectra. To lessen this light effect on diffuse reflectance NIR spectra, the MSC algorithm was created. The scattering at the sample's jth wavelength was calculated by regression each spectrum onto the mean spectrum, where it may be described by (Dhanoa et al., 2023);

$$x_j = a + b\bar{x}_j + \epsilon_j \qquad Eq. 2.1$$

where for every j wavelength in the samples, a and b remain constant. The scaled deviations about the regression yield the scatter-corrected spectra, as may be seen below;

$$x_j, MSC = \frac{(x_j, raw - a)}{b} \qquad Eq. 2.2$$

SNV eliminates the particles size, multiplicative interferences of scatter, and the changes of light distance similar to MSC (Barnes et al., 1989). Both multiplicative and additive scatter effects can be corrected with SNV. To eliminate gradient variations on an individual spectrum basis, each object is transformed individually using the following equation (Barnes et al., 1989):

$$x_{i,snv} = \frac{x_i - \bar{x}}{\sqrt{\frac{\sum(x_i - \bar{x})^2}{(n-1)}}} \qquad Eq. 2.3$$

© COPYRIGHT UPM

Where $x_{i,snv}$ is the transformed spectrum, x_i is the original spectral, $\bar{x}$ is the average value of the variable, and n is the number of the variables in the spectrum.

2.7.2 Multivariate Statistical Technique for Spectroscopic Analysis

The primary goal of multivariate analysis techniques is to create models that can precisely identify, predict, or classify the characteristics and attributes of unidentified samples (Blanco & Vllarroya, 2020). To create prediction models for specific attributes of unknown samples (quantitative analysis) or to establish classification procedures for unknown samples (qualitative analysis), multivariate analysis allows samples with comparable characteristics to be classified. The steps listed in Table 2.5 are involved in the procedure.

Table 2.5: The process involved developing models for prediction

Steps	Purpose
Selecting the samples for calibration	To choose a group of samples that accurately reflect the entire population.
Finding the target parameter through the use of the reference technique	to accurately and precisely calculate the measured property's value. The calibration model's quality is determined by the value's quality.
Capture the spectral data	To acquire physical-chemical data in a repeatable way
Applying appropriate pre-treat	To apply multivariate techniques to determine the spectrum-property relationship.
Verifying the model	to make sure that in samples that haven't gone through the calibration process, the model predicts the attribute of interest appropriately.
Predicting random samples	Accurately predicting unfamiliar samples.

(Source: Lazim (2019))

© COPYRIGHT UPM

The common methods for quantitative analysis are established based on linear regression, which models a collection of independent and most likely related information (spectral data) to a set of dependent variables of crop quality indicators, including color, moisture content, sugar content, and chlorophyll content. In VSNIR spectroscopy regression, the two most used multivariate regression techniques are principal component regression (PCR) and partial least-square (PLS) (Kamal & Sulaiman, 2021).

PCR utilizes the principal components (PCs) generated by PCA to perform regression on the property of the sample to be predicted. In contrast, PLS identifies the directions of maximum variability by incorporating both the spectral data and the target property information, with the resulting new axes referred to as 'latent variables (LVs)' or PLS factors. The simplest quantitative multivariate analysis method is multiple linear regression (MLR) (Sau which usually uses fewer than five spectral wavelengths. MLR assumes concentration to be a function of absorbance, which entails the knowledge of the concentrations of not only the target analysis but also all other components contributing to the overall signal.

Multivariate analytical techniques are mainly applied to establish a correlation between the (*n*x1) vector of observed response values y ('Y-variables'; quality attributes of interest, such as chlorophyll content and moisture content) and the *n*XN spectral matric X ('X-variables'). With n the number of spectra and N the number of wavelengths (Lakshmi et al., 2021).

© COPYRIGHT UPM

2.7.2.1 Multiple Linear Regression (MLR)

MLR method, a value of y is estimated by a linear combination of the spectral data at each single wavelength. The *N* regression coefficients are approximated by minimizing the error between predicted and observed response values in the least square sense. Several variables for the MLR equation can be identified by a stepwise multiple linear regression (SMLR) technique (Zare et al., 2019). SMLR first selects the variable with the biggest correlation with the *Y*-data, then the second variable that results in the best improvement of the accuracy of the prediction, and then the third variable and so on until no further significant improvement can be obtained (Shams et al., 2021). MLR models normally do not perform well because of the regularly high co-linearity of the spectra, and easily lead to over-fitting and a loss of robustness of the calibration models (Schwartz et al., 2011).

2.7.2.2 Principle Component Regression (PCR)

PCR has a two-step procedure firstly, decomposes the X by a PCA analysis and then fits a MLR model. This model uses a small number of PCs instead of the original variables as predictors (Malavi et al., 2024). The advantage of the PCR compared to MLR is that the *X*-variables (PCs) are unrelated, the noise can be filtered, and require only a small number of PCs. The disadvantages of the PCR is that the PCs are well-arranged based on decreasing explained variance of the spectral matrix, and that the first principal components which are used for the regression model are not essentially the most informative concerning the response variable (Abdolmaleki et al., 2015).

© COPYRIGHT UPM

2.7.2.3 Partial Least Square (PLS)

PLS was introduced to overcome the disadvantage of PCR method. In PLS regression, an orthogonal basis of LVs is constructed one by one in such way that are oriented along directions of maximal covariance between spectral matrix X and the response vector y (Xu et al., 2020). This procedure ensures that the LVs are well-ordered according to their relevance for predicting the *Y*-variable. In the regression model, the interpretation of the relationship between X-data and Y-data is then simplified as this relationship is built based on the smallest possible number of LVs. PLS models are slightly better compared to the PCR because they exclude latent variables that are less significant to describe the variance of the quality parameter, thus requiring fewer LVs compared to PCR for producing a similar model performance (Jong, 1993). The PLS method is commonly applied for spectral data of non-destructive biological material when the various *X*-variables express common information, i.e., when there is a large amount of correlation, or even co-linearity (Fu & Ying, 2014).

2.7.3 Data Reduction Techniques using Feature Selection

Dimensionality reduction is crucial especially when experiencing a high-dimensional feature space, therefore it is commonly used in machine learning where a new space of reduced dimensions is produced by mapping the original feature space onto it. Data reduction involves transforming high-dimensional data into low-dimensional representations while preserving the data's original meaning as much as possible (Vinutha et al., 2023). This study demonstrates the application of data reduction to identify significant wavelengths, echoing similar research that emphasized the

© COPYRIGHT UPM

potential of feature selection methods to differentiate between severity levels of WRD in rubber trees. Data reduction offers numerous benefits, such as reducing data storage requirements (Ayesha et al., 2020) decreasing computation time (Zebari et al., 2020), and eliminating redundant (Ray et al., 2021), irrelevant (Wang et al, 2021), and noisy data (Ougiaroglou et al., 2024). It can enhance data quality and improve algorithm performance, particularly for algorithms that struggle with high-dimensional data.

Feature selection (FS) is one of the data reduction techniques that aims to identify the most relevant features in a dataset to improve model performance and reduce complexity (Rostami et al. 2021). Reducing dimensions simplifies classification processes and boosts efficiency and accuracy (Jia et al., 2022). In FS, the features are evaluated and ranked using the information-theoretic measures, and then those with the highest ranks are selected (Das et al., 2020). Filter and wrapper methods are two popular approaches for FS, and they differ in how evaluate the importance of features (Bashir et al., 2022). The filter methods use statistical techniques to evaluate the relevance of each feature independently of any machine learning algorithm (Saarela & Jauhiainen, 2021). Therefore, the methods in this approach are typically fast and suitable for high-dimensional datasets (Bommert et al., 2020). Wrapper methods evaluate feature subsets by training and testing a specific machine-learning model.

Among the various methods for FS, analysis of variance (ANOVA), relevant feature interactions (RELIEF), and Chi-square are used due to their effectiveness and simplicity.

© COPYRIGHT UPM

2.7.3.1 Feature Selection using ANOVA

ANOVA is a statistical approach used to analyze the differences among the groups and their associated procedures. ANOVA assesses the statistical significance of the differences between the means of multiple groups. It uses the F-test to determine whether the means of different groups are significantly different from each other. Shakil et al. (2023) explore the effectiveness of feature selection methods in automating the recognition of disease in dragon fruits. Specifically, it evaluates the use of ANOVA-based feature selection techniques in improving disease detection accuracy by eliminating irrelevant and redundant features in a dataset. The findings demonstrate impressive performance, achieving a 96.20% accuracy in identifying dragon fruit diseases, and better than previous research, which reported accuracy below 95%.

2.7.3.2 Feature selection using RELIEF

RELIEF is a feature selection algorithm that evaluates the quality of features based on their ability to distinguish between instances that are near each other. It assigns weights to features, giving higher weights to those that contribute more to the distinction between classes. This foundational paper introduced the RELIEF algorithm and demonstrated its effectiveness in handling noisy and multi-class datasets. The authors showed that RELIEF could identify relevant features even in the presence of redundant features.

© COPYRIGHT UPM

2.7.3.3 Feature Selection using Chi-square

The Chi-Square test is a statistical test used to determine if there is a significant association between two categorical variables. In feature selection, it evaluates the independence of features concerning the target variable. FS is a critical step in machine learning, and methods like ANOVA, Chi-Square, and RELIEF play significant roles in improving model performance and interpretability.

2.8 Classification of Disease Severity Levels

This section describes the review of various techniques for the classification of diseases using machine learning. Plant diseases can significantly impact agricultural productivity, leading to economic losses and food security levels in crops such as rubber trees are crucial for effective management. This section focuses on using machine learning to classify the severity levels of disease in plants.

2.8.1 Classification using Artificial Neural Network

An artificial neural network (ANN) is an information processing paradigm that is inspired by the structure and function of the human brain (Schmidgall et al., 2024). ANN consists of layers of interconnected "neurons". Neuron is the building component of the ANN designed to simulate the function of the biological neuron (Shobika et al., 2024). The arriving signals, called inputs, multiplied by the connection weights (adjusted) are first summed (combined) and then passed through a transfer function to produce the output for that neuron. The activation function is the weighed sum of the neuron's inputs and the most used transfer function is the sigmoid function

© COPYRIGHT UPM

that process and transmit information as shown in Figure 2.10 (Singh et al., 2024) Feedback is another type of connection where the output of one layer routes back to the input of a previous layer, or to same layer. Feedback architecture has connections from output to input neurons (Figure 2.11) (Miyashita, 2024). Each neuron has one additional weight as an input that will allow an additional degree of freedom when trying to minimize the training error.

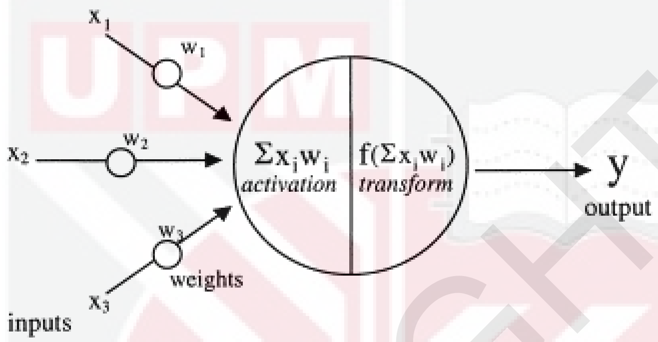


Figure 2.10: Model of an Artificial Neuron

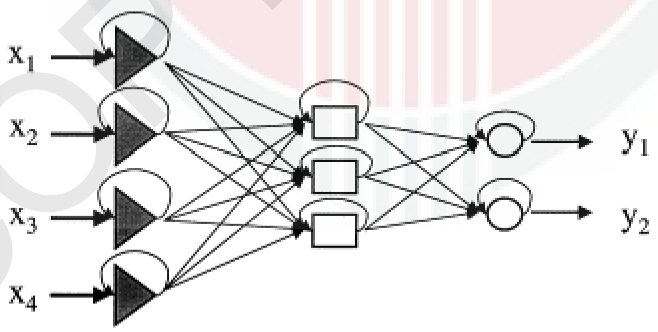


Figure 2.11: Feedback Network

The crucial procedure for developing an accurate process model using ANN is the training and validation processes. For this case, training feature sets are used to train

© COPYRIGHT UPM

the ANN model while testing feature sets are used to validate the accuracy of the ANN model. The proper network design must be well-developed before data processing by ANN model which includes network type and the training process. The optimum properties based on the selected phase are obtained and conducted simultaneously with the network training phase. The connection of weights is updated during the training phase until a certain iteration number has been defined. There are four important features in ANN which are several hidden layers, the number of the input layer, the number of the output layer and the mean square error. Several inputs are the number of features that are used for extracting image information. The function of the input layer is to receive input information from the outside which is the images to be processed and passed to the middle layer. The function of the middle layer is for information transformation which is known as the internal information processing layer depending on the information change capacity. Mean square error is used as a performance measurement of ANN. After the retrieving process is finished, the output layer delivers the output of processed information to the user which is the classification of the diseases.

Abdullah et al. (2020) classified the diseases in rubber tree leaves using multilayer perception neural networks. After the preprocessing using a median filter, the color features were extracted using RGB distribution indices to calculate the region of interest (i.e. lesion spots). The optimized models were then evaluated and validated by analyzing various performance indexes. According to the authors, the proposed method is the best model for identifying bird eye spots and tritium diseases. However, the effectiveness of this classification can be increased by providing a better processor and a high-resolution camera.

© COPYRIGHT UPM

2.8.2 Classification using Support Vector Machine

One classifier that maps input data into a new high-dimensional feature space and offers a classification between two classes is the Support Vector Machine (SVM) (Ghaddar & Naoum, 2018). With the widest margin for separating positive and negative classes, it has an ideal separation hyperplane (Figure 2.12). The largest separation is indicated by the best hyperplane obtained from the classification. The hyperplane that has maximized the distance between it and the closest data point on either side is known as the maximum-margin hyperplane. There are two categories for SVM: linear and nonlinear. A hyperplane is used in linear SVM to divide the training sample based on the decision function. Every dot or piece of data in a nonlinear SVM is swapped out with a nonlinear kernel function. To categorize the different types of diseases, SVM is used to train the extracted features that are kept in the feature database. Using feature extractions from the segmented pictures of the input photos, any input image fed into the SVM system can be classified into one of the illness classifications.

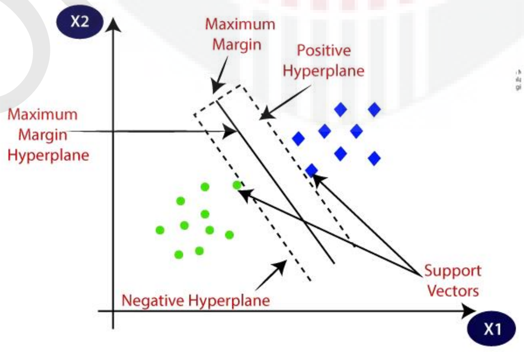


Figure 2.12: Ideal Separation Hyperplane of SVM

© COPYRIGHT UPM

2.8.3 Classification using k-Nearest Neighbor

kNN is supervised ML technique commonly applied to pattern recognition. It is predicated on the nearest neighbor rule, which is applied in ML applications to classify data. In k-NN classifier, k is just a number that tells the algorithm how many nearby points (neighbors) to look when it makes a decision (Cunningham & Delany, 2021) (Figure 2.13). This method involves training the test pattern using the classifier and classifying the test pattern according to how similar it is to each training pattern. The object is allocated to the most widely used class labels among its k-NN. The use of k-NN in plant disease detection has gained significant traction due to its simplicity and effectiveness. KNN is a non-parametric, lazy learning algorithm that classifies data points based on the majority class of their nearest neighbors.

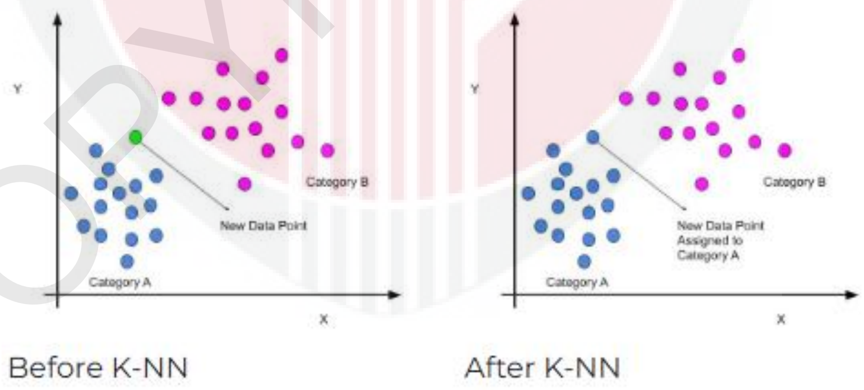


Figure 2.13: Architecture of kNN

kNN has been applied successfully in various studies for plant disease classification using spectroscopy techniques. A study demonstrated using kNN with spectroscopy to classify diseases in plants based on spectral data, achieving high classification accuracy by analyzing signatures (Li et al., 2018). While kNN is effective in handling spectral data, it is sensitive to the choice of k and can be computationally expensive with large datasets. The choice of k affects the classification accuracy, and many require optimization for different datasets and plant species.

2.9 Summary

Rubber cultivation has economic importance worldwide, and Malaysia also plays a significant role in world rubber market, even though is contribution is declining The WRD has been known as major problem in Southeast Asia that resulted in a loss in terms of natural rubber productivity up to 50%. It is a silent killer, where the above ground symptoms are produced one the roots are fully damaged. The main constraint to control WRD disease are regularly do checking on trees by expert workers and chemical applied on disease trees and their effectiveness depends on reliable and early detection of pathogen. In this context, the utilization of a feasible approach in achieving sustainable rubber production through the VSNIR spectroscopy and improving plant growth may help to develop an effective control strategy for managing WRD disease.

© COPYRIGHT UPM

CHAPTER 3

METHODOLOGY

The previous chapter discussed WRD as a significant challenge in the rubber industry, where it can lead to losses in natural rubber productivity of up to 50%. The review highlighted the current common methods for detecting WRD and the limitations of traditional techniques. To address these challenges, Chapter 3 presents methodologies designed to characterize the crop parameters of leaves, latex, and surrounding soil of rubber trees at various stages of WRD severity. The methods outlined in this chapter focus on measuring the crop parameters of the sample, followed by the use of spectral techniques to capture data that correlates with these properties. By combining traditional chemical analysis with advanced spectral data and chemometrics, the study aims to optimize regression models that can predict WRD severity levels. The chapter further explores data pre-processing, statistical analysis, and regression modelling, which are crucial steps to ensure the accurate interpretation of spectral measurements. Figure 3.1 illustrates the methodology processes.

© COPYRIGHT UPM

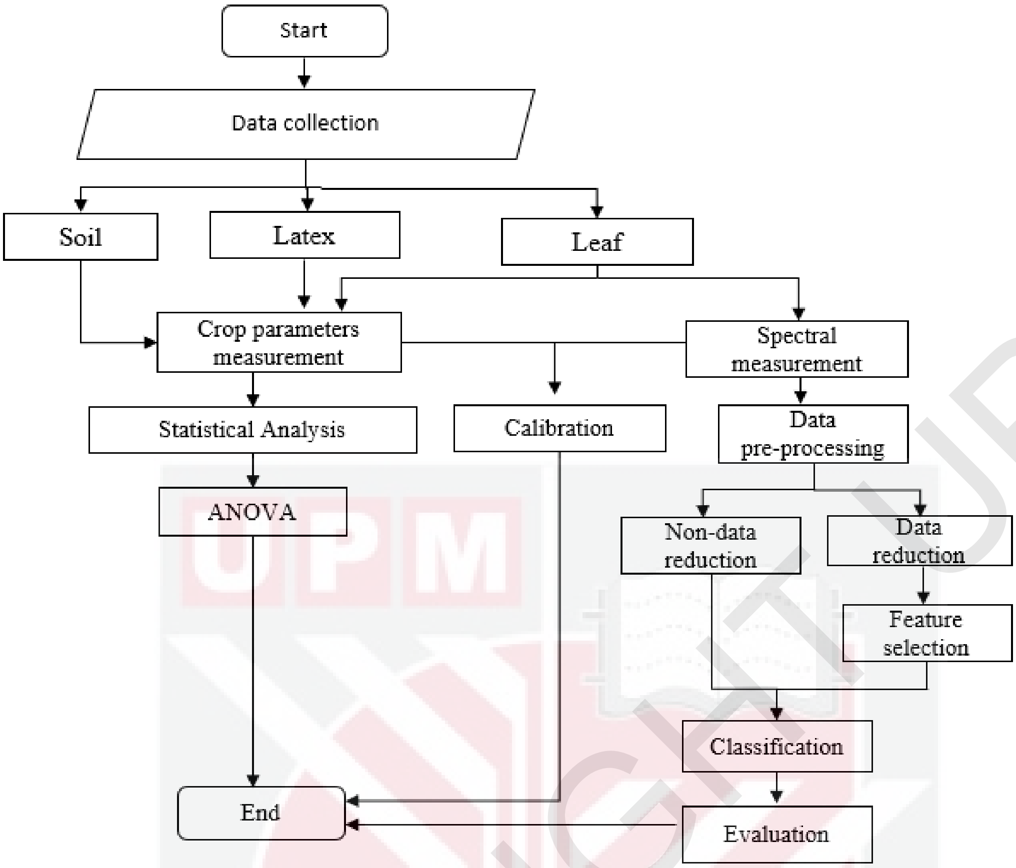


Figure 3.1: Flowchart of the Methodology

3.1 Study Area

The fieldwork of this study took place at the rubber estate of the Rubber Research Institute Malaysia (RRIM) in Kota Tinggi, Johor, Sungai Buloh, Malaysia. A randomized complete block design (RCBD) was employed to minimize the effect of variability. In parallel with data collection, meteorological information, including temperature and humidity, was obtained from World Weather Online. A total of 163 rubber tree samples representing five different severity levels, namely, healthy (37), light infection (22), moderate infection (23), severe infection (28), and very severe (24).

The trial plot at Sungai Buloh was chosen for a preliminary study. This preliminary study was a pilot project to gather essential data and insights before starting a more extensive study in Kota Tinggi, Johor. RRIM Sungai Buloh was located at 1° 47.778' N and 103° 51.343' in Selangor. In total, 50 rubber trees were used for this preliminary study, with 10 rubber trees for each severity category. The severity is illustrated below (Table 3.1). This preliminary study aims to test the hypothesis systematically, determining whether the proposed explanation or prediction holds under the specific conditions at Sungai Buloh. Through various experiments, observations, and analyses conducted at this site, the validation of the hypothesis can assess while identifying any potential flaws, limitations, or areas requiring further investigation. Overall, the selection of Sungai Buloh for this preliminary study is a crucial step in the scientific process, offering valuable insight and laying the foundation for more comprehensive research efforts to explore the hypothesis in greater detail.

Table 3.1: Typical severity level WRD infection and their description

Score	Severity Level	Description
S0	No infection	No rhizomorph on stem base or lateral roots
S1	Light infection	Rhizomorph present, tissue penetrated
S2	Moderate infection	Portion of one lateral root penetrated
S3	Severe infection	More than one lateral root penetrated
S4	Very severe infection	Rhizomorph present on surface

The trial plot at RRIM Kota Tinggi is located at 1° 47.778' N and 103° 51.343' in Johor. The total area of this field is 9.4 hectares, with soil classified as the Jerengau series, a Class 1 soil series suitable for rubber plantations. This area was divided into

© COPYRIGHT UPM

four zones: three affected and one healthy area, as illustrated in Figure 3.2. This zone area was confirmed by experienced RRIM staff. Each zone consisted of five circular areas, and a circular area consisted of five trees using a single tree as a benchmark, as shown in Figure 3.3. A total of 84 rubber trees were used in this study, with 21 trees in each zone. The tree with the most severe infection served as the benchmark for the study.

Figure 3.2: Study Area in Kota Tinggi, Johor

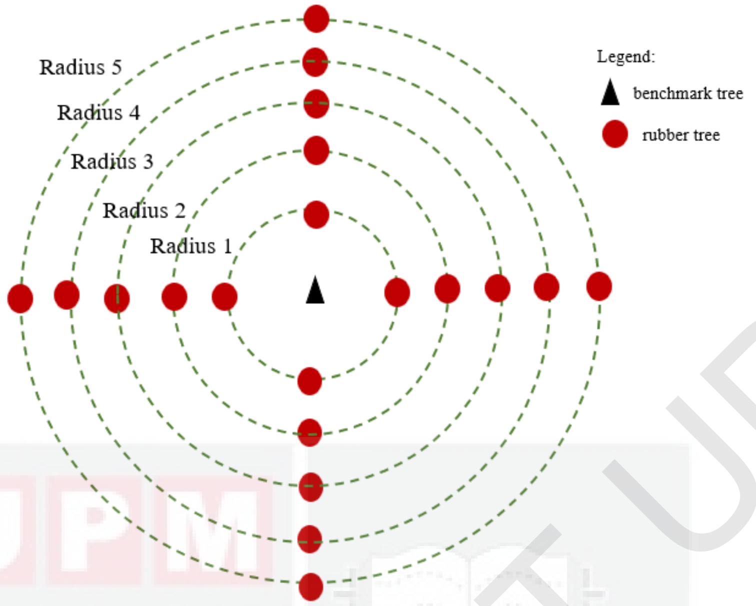


Figure 3.3: Circular Area for Each Zone Sampling

3.2 White Root Disease Assessment

The study collected three types of samples—leaves, latex, and soil—from healthy and affected trees by WRD at different severity levels (light, moderate, severe, and very severe). Experienced staff from RRIM assessed the WRD infection at five different severity levels. Traditionally, the severity of WRD in rubber trees is identified via the scoring method (Maiden et al., 2022) and segmentation method (Sulaiman et al., 2017).

A trained expert visually inspects the leaves of each tree to score WRD plant disease severity using a five-scoring scale: S0 for healthy, S1 for light infection, S2 for moderate infection, S3 for severe infection, and S4 for very severe infection (Maiden et al., 2022) (Table 3.1). The second method determined the severity levels by observing tree morphology and the presence of the *Rigidoporus micropores* as shown in Figure 3.4. This approach involved evaluating the fungus infection based on the number of root segmentations (a, b, c, d) affected. For example, if three out of four segmentations are infected by WRD fungus, the tree is classified as S3, indicating it cannot produce latex and is likely to die.

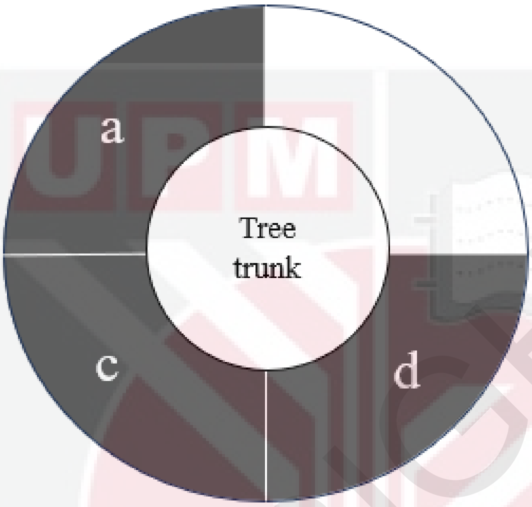


Figure 3.4: Segmentation of the Root Zone (top view) for WRD Infection

3.3 Characterization of Disease Infection on Leaf Samples

This study aimed to evaluate the application of VSNIR spectroscopy for early detection of WRD in rubber trees. This comparative study seeks to determine the most effective rubber tree sample type for early WRD detection using VSNIR spectroscopy. The rubber tree samples were provided by RRIM, and all physiochemical properties were measured following standard procedures. Three types of samples were analyzed in this study: leaf samples, latex samples, and soil samples. The preparation methods for each sample type are detailed below.

© COPYRIGHT UPM

3.3.1 Leaf Samples Collection

This study utilized a total of 600 rubber leaves, which were collected from 134 rubber trees. The leaf samples were plucked from the outer periphery of the canopy, based on severity level by well-trained farm workers according to standard practice in the field as shown in Figure 3.5. For consistency and minimize variability, leaves were collected from bottom, middle and upper canopy of each tree. Nine leaves per tree were collected from different direction (north, south, east and west) to account for any potential differences due to sunlight exposure or shading. The leaf samples were collected in the morning to preserve their color before prolonged exposure to sunlight. After that, the leaf samples were sealed in plastic bags and brought to the laboratory. Then, the crop parameters of leaves such as moisture content, color, and SPAD value were measured. All the measurements were recorded in three replications.

3.3.2 Spectroscopic measurement

Leaf reflectance measurement was conducted ultraviolet-near infrared (UVNIR) spectrometer, (Ocean Optic HR4000, Ocean Optics Inc experimental, Dunedin, Florida). This visible and shortwave near-infrared spectrometer (UVNIRS) operates within a wavelength range of 200 nm to 1200 nm, featuring an optical resolution of 0.05 nm. The scanning distance was kept constant by positioning the probes on a holder at a 90° angle. The measurement was performed by placing the probe in the different leaf surfaces. The selected regions of interest to be measured, the main vein, vein, and petiole as shown in Figure 3.6. All the spectral measurements were carried out in the black box.

© COPYRIGHT UPM

Figure 3.5: Leaves Samples Representing Five Different Severity Diseases

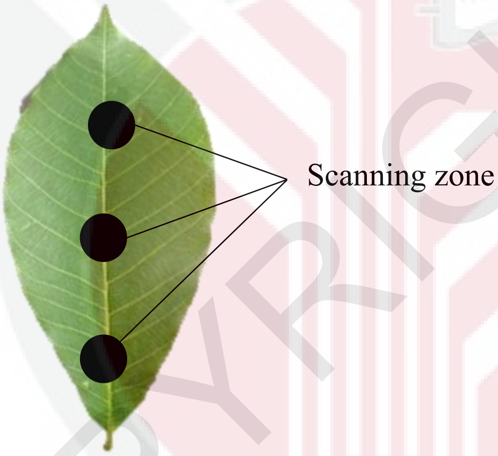


Figure 3.6: The Scanning Points on the Rubber Leaves

3.3.2.1 Ligh-proof Measurement Box

In this study, a lightproof measuring box (30 cm x 30 cm x 10 cm) equipped with black cloth was used to prevent stray light from reaching the samples (Figure 3.7). This box was designed to enclose the light sources, sensors, and samples from ambient light. To further minimize the impact of background surfaces on spectral data, all interior surfaces of the box were covered with black cloth (Lazim et al., 2022).

© COPYRIGHT UPM

Figure 3.7: The Lightproof Measuring Box

3.3.2.2 Illumination System

The illumination source in this study was provided by a tungsten halogen light source (HL-2000 12VDC (15W), Ocean Optics Inc., USA). The lamp was powered by an AC/DC adapter connected to a universal power supply. It provided a constant light source, covering both the visible and NIR regions (Wu et al., 2008). The light source was connected to the optical resolution of 0.05 nm outside of the box and then was set at a constant scanning distance of 1 cm from the sample for the spectral measurement as shown in Figure 3.8. The light source was turned on for 30 minutes before taking spectral measurements to warm up and stabilize the spectral data, as recommended by the manufacturer. In this study, a white reference (WS-1 Diffuse Reflectance Standard) and a black reference were used before collecting spectral measurements. The function is to act as reference optical radiation (light) that illuminates objects during reflectance measurement.

© COPYRIGHT UPM

Figure 3.8: Experimental Setup for Spectral Measurements

3.3.2.3 Spectral Acquisition

SpectraSuite is a modular, Java-based spectroscopy software platform compatible with 32 and 64-bit Windows, Macintosh, and Linux operating systems. It can control any Ocean Optics USB spectrometer and device (Ocean Optics, Inc., 2009).

3.3.3 SPAD Value Measurement

The chlorophyll content is crucial for plant growth and development. Physiologically, it facilitates photosynthesis, the process through which solar energy is converted into chemical energy (Golhani et al., 2019). Chlorophyll content was quantified using a SPAD (Soil-Plant Analyses Development) meter – 502 (Minolta Co. Ltd., Osaka, Japan) (Figure 3.9) an optical device designed to measure the relative chlorophyll index of plant leaves. The amount of light absorbed by chlorophyll pigment in the leaf indicates its chlorophyll concentration. The SPAD value was recorded at the same leaf position where spectral measurement had been scanned.

Figure 3.9: Measuring Leaf Chlorophyll using SPAD Meter

3.3.4 Color Measurement

The leaf surface of the sample was measured using a handheld colorimeter (NR20XE, Shenzhen 3nh Technology Co. Ltd., China) (Figure 3.10), which has a measuring aperture diameter of 20 mm. Before taking the color measurement, the colorimeter was calibrated using a standard white tile. The color measurement of the leaf was measured at the same position where spectral measurement had been scanned.

© COPYRIGHT UPM

Figure 3.10: Leaf Color was Measured using a Handheld Colorimeter

3.3.5 Moisture Content Measurement

The ML-50 moisture analyzer (A & D Company Ltd., Tokyo, Japan) (Figure 3.11) operates using electric current as its power source. It measures 32 cm in length, 21.5 cm in width, and 17.3 cm in height, and weighs approximately 6 kg. This device is essentially a precision balance featuring an 8.5 cm wide circular pan, capable of weighing samples up to 51 g with an accuracy of 0.005 g. It dries samples using a 400 W halogen lamp (reddish-brown color), with drying temperatures adjustable between 50 and 200 °C in 1 °C increments. The moisture content can be measured with an accuracy of 0.5% for samples weighing between 1 and 5 g and 0.1% for samples over 5 g. The instantaneous drying rate is displayed with a precision of 0.01% per minute during the drying process.

Figure 3.11: Leaf Moisture Content Measured using Moisture Analyzer

3.4 Latex Samples Collection

WRD leads to a decline in latex concentration and a shift in latex chemical composition, which alters leaf internal structure and light-absorbing properties, producing different spectral data (Yadav et al., 2023). Latex samples were collected directly from the tree trunk at the usual tap (Figure 3.12). A minimum of 60 ml of latex was obtained from each tree to measure the total solid content (TSC) (Figure 3.13). A preservative was added to the samples to prevent coagulation and maintain freshness.

© COPYRIGHT UPM

Figure 3.12: The Trunk of the Rubber Tree is Typically Tapped to Extract Latex

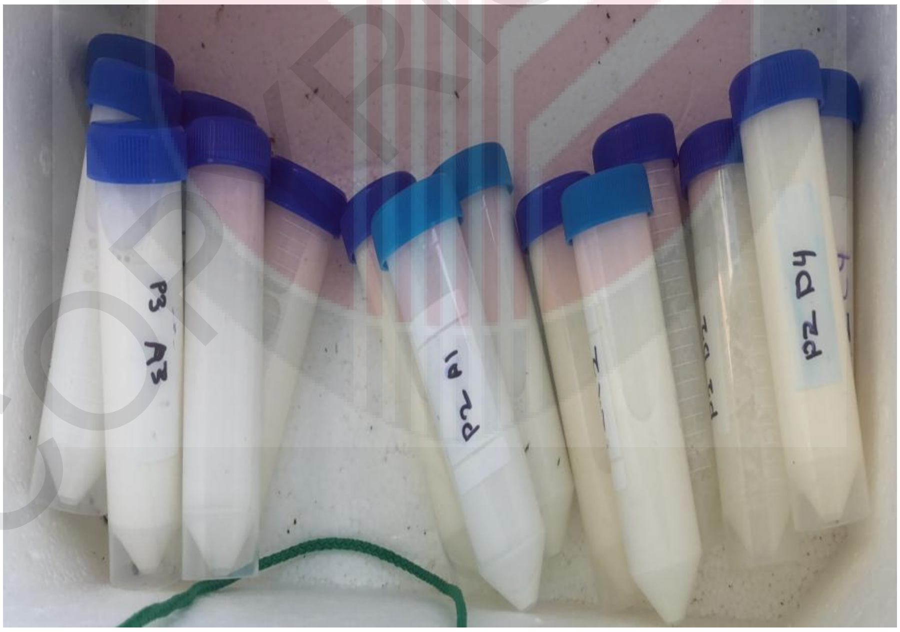


Figure 3.13: Latex samples were collected from rubber trees

© COPYRIGHT UPM

3.4.1 Total Solid Content Measurement

Each rubber latex sample was measured using an 80-cm diameter petri dish and weighed on an electric balance with a resolution of 0.0001 g. The TSC measurement of rubber latex was measured by the factory following the standard methods ISO 124:2001. The average of three replicates was calculated for each sample. The process for the measurement of TSC is presented on Figure 3.14.

The TSC of rubber latex was measured based on a wet basis method (%, w.b.) using the Eq. (3.1):

$$\text{Total solids content} = \frac{M_1}{M_0} \times 100 \qquad 3.1$$

Where, M_0= weight of wet sample and M_1= weight of the dried sample

© COPYRIGHT UPM

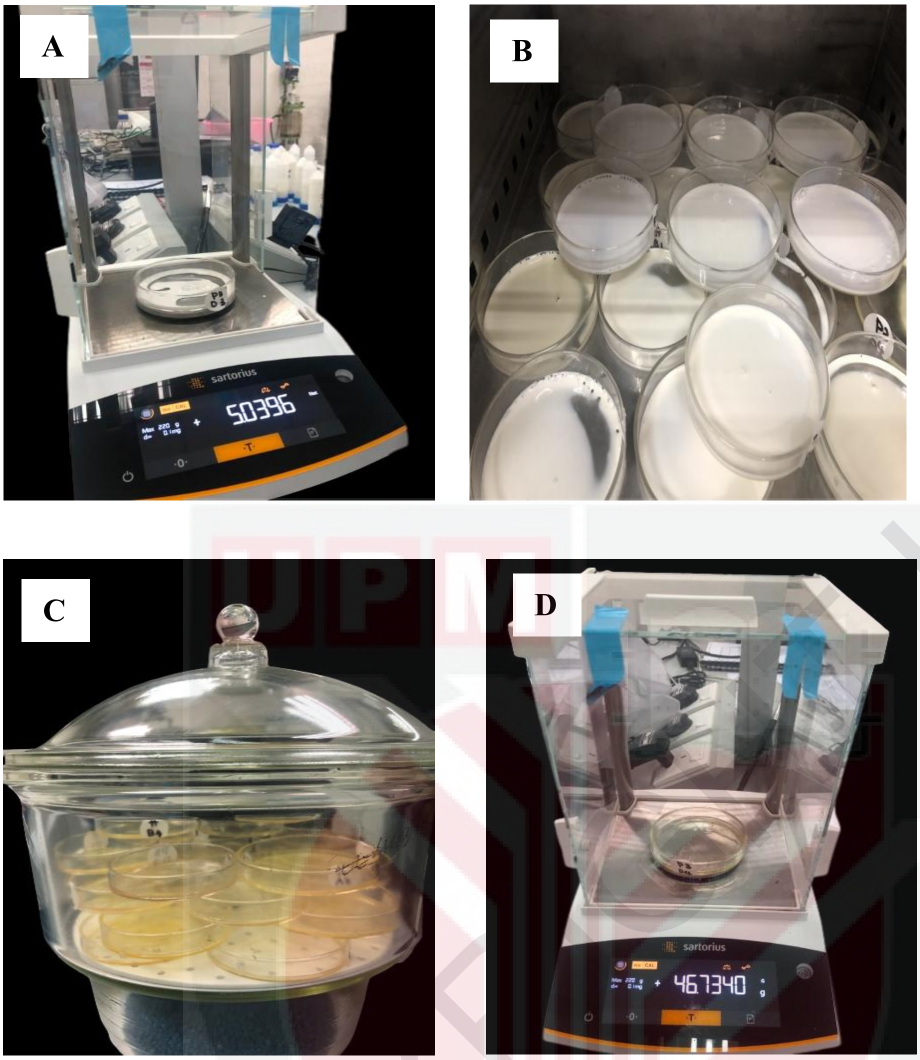


Figure 3.14: Procedure for the Measurement of TSC: (a) Mass Measurement Before Oven-Dried, (B) Latex Samples Were Oven-Dried at 70 °C for 16h, (C) The Samples Were Cooled in a Desiccator for 30 Minutes, (D) Mass Measurement After Oven-dried

3.5 Soil Collected Surrounding from Rubber Tree

The plot at RRIM Kota Tinggi has soil classified as the Jerengau series, which is a Class 1 soil. The Jerengau series typically consists of 66% sand, with clay and silt each making up 17%. The pH and MC measurements for each soil around rubber trees were collected using a Portable Soil Analyzer with multiprobes (SEM2260, China) with a probe diameter of 3 mm. Handheld soil meters are portable devices that can measure

© COPYRIGHT UPM

various soil parameters on-site, providing quick and convenient information for agricultural assessments. This device uses a probe to measure the electrical conductivity or dielectric of the soil which correlated with soil moisture content. The pH meter uses a probe to measure the acidity or alkalinity of the soil. The pH value indicated the concentration of hydrogen ions in the soil.

Before conducting pH and MC measurements, any debris, stones, or plant material within a 10 cm diameter around the root was cleared. Then, probes were directly inserted into the soil at 10 cm depth, and the readings for pH and MC were displayed on the screen (Figure 3.15). The average values of three replications were calculated for all rubber tree samples

Figure 3.15: The pH and MC Measurements of Soil Around Rubber Trees Were Collected using a Portable Soil Analyzer With Multiprobe

© COPYRIGHT UPM

3.6 Molecular Identification of White Root Disease Pathogen

Molecular identification is a scientific approach to determine and characterize the identity of fungi based on their unique molecular signature. This process involves three key steps: isolation, PCR amplification, and DNA sequencing. The overall procedure is illustrated in Figure 3.16.

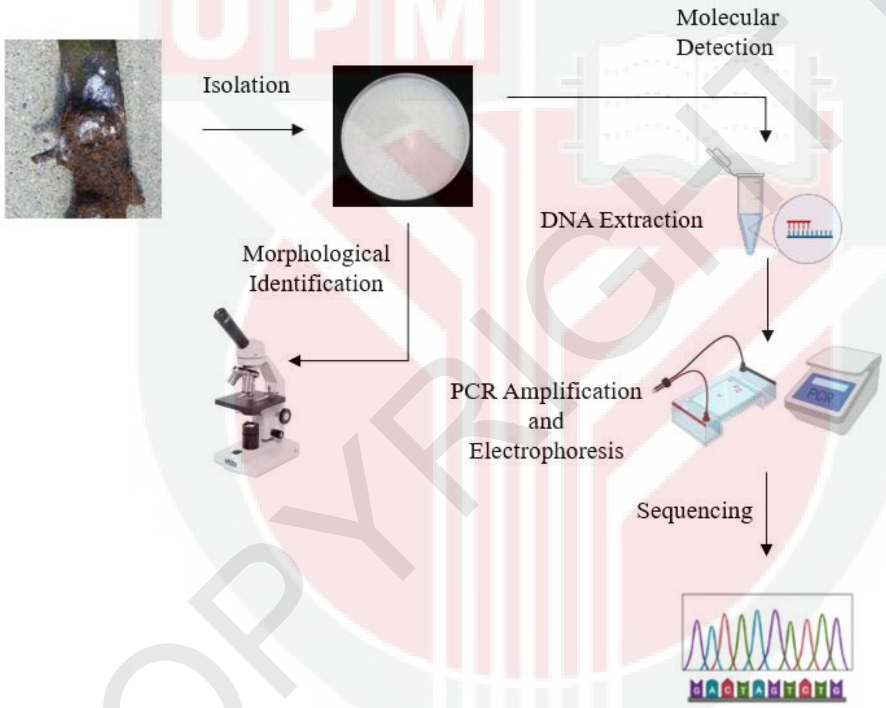


Figure 3.16: Overview of the Molecular Identification of White Root Disease

3.6.1 Isolation of *Rigidoporus microporus*

The root-rot pathogens of WRD were isolated and identified using the method described by Luo et al. (2019). Mycelia of the *Rigidoporus microporus* were obtained from the root tissue of rubber trees. The samples were brought back to the Processing

© COPYRIGHT UPM

and Product Development Laboratory Universiti Putra Malaysia (UPM). Soil particles, dirt, and insects were eliminated from the root tissue and fruiting bodies by running tap water. The cleaned samples were cut into smaller pieces around 5 cm in length. The small pieces of the cleaned sample were then surface sterilized before being placed on potato agar (PDA) (Difco™, Becton Dickinson and Company, Maryland, NY, USA), amended with streptomycin sulphate (300 mg/L) for bacterial suppression agar plates with four pieces of sample in one plate. The plates were then incubated at 27 ± 2° for 24 to 48 hours to allow the mycelium from the sample to grow. The process for isolation of *Rigidoporus microporus* is presented in Figure 3.17. For *Rigidoporus microporus* isolation process was replicated three times.

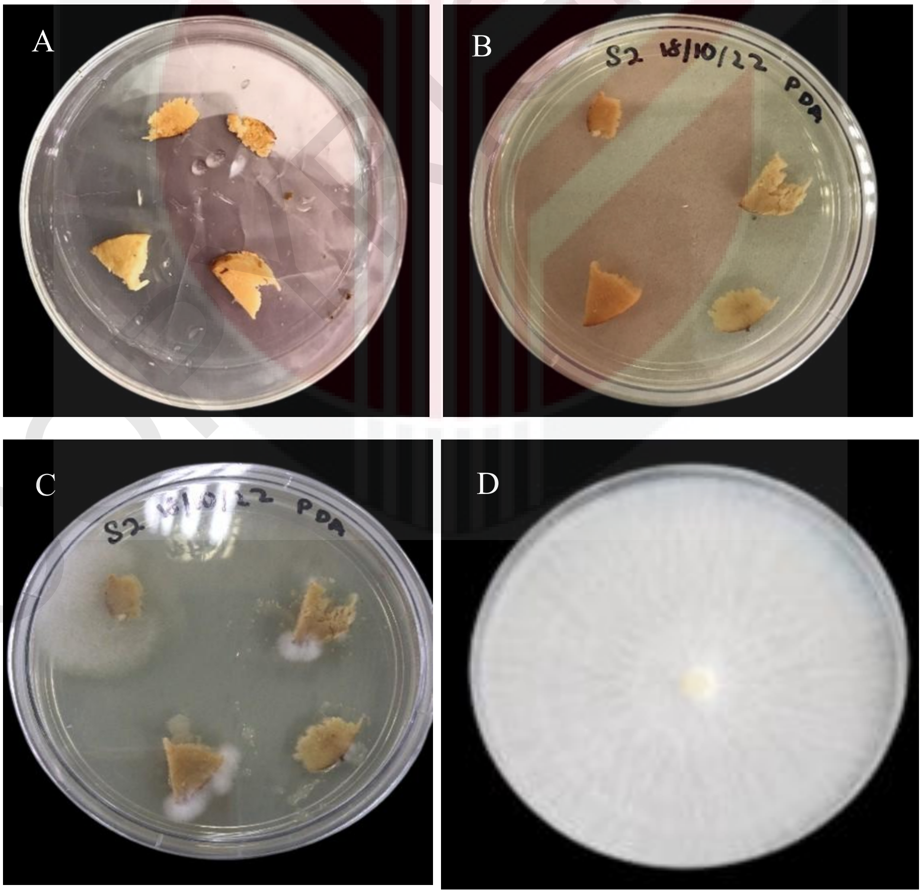


Figure 3.17:The procedure for the *Rigidoporus microporus* Isolation; (A) Fresh Root, (B) Root Samples on PDA Agar Plate, (C) The Plates were Incubated at 27 ° C (D) Mycelium from Sample Grow

© COPYRIGHT UPM

3.6.2 DNA Sequencing

3.6.2.1 DNA Extraction

Hyphal tip isolation was grown on PDA and incubated 25°C for 10 days. The mycelium was scraped from the surface of PDA by an aseptic technique. DNA was extracted with Cetyltrimethylammonium bromide (CTAB buffer as described by Gupta et al., (2020) by grinding mycelium with liquid nitrogen to fine powder. The mycelial powder was suspended in 600 µl (CTAB), and incubated at 65°C for 30 min. During the incubation period, the tube was vortexed every 10 min. After that, the tube was cooled for a few minutes, and then 600 µl chloroform: isoamyl alcohol (CIA; 24:1, v/v) was added, gently mixed, and centrifuged at 7,000 rpm for 5 min at 40°C. The aqueous layer was transferred to new tubes and extracted again with the CIA. Finally, DNA was precipitated by adding 300 µl isopropanol, mixed well incubated at room temperature for 30 min, and centrifuged at 12,000 rpm for 10 min. The supernatant was decanted. The DNA pellet was suspended in 40 µl ddH2O.

3.6.2.2 PCR Amplification

The International transcribed spacer (ITS) regions were amplified using the universal primers ITS 1 (5′TCC GTA GGT GAA CCT GCG G 3′) and ITS 4 (5′TCC TCC GCT TAT TGA TAT GC 3′) (White *et al*., 1990). The amplifications were performed in reaction volumes of 50 µl containing 5 µl of 10x PCR buffer, 3 µl of 25 mM MgCl2, 1 µl of 10 mM dNTP, 0.3 µl of 1.5 U *Taq* DNA polymerase, 1 µl of each primer, 1 µl genomic DNA and 37.7 µl of ddH2O. PCR was carried out in a MyGeneTM Series Peltier Thermal Cycler (Model MG96G) using the following program: 2 min initial

© COPYRIGHT UPM

denaturation at 94 °C followed by 35 cycles of 50 sec denaturation at 94 °C, 1 min annealing at 55 °C, 1 min extension at 72 °C, and final extension for 5 min at 72 °C. Amplification products were separated by 1% agarose gel in 1xTAE buffer strained with ethidium bromide which was included in the agarose gel and visualized under UV fluorescence. PCR products were purified using the AxyPrepTM DNA Gel Extraction Kit (Axygen Scientific, Inc. USA) according to the manufacturer's instructions. DNA was sequenced by Shanghai Invitrogen Biolotech Co., Ltd. (Shanghai, P.R China).

3.7 Data Analysis

3.7.1 Descriptive Statistics

In this study, descriptive statistics were used to describe the basic features of a series collection of data. It summarizes the large data samples and quantitative analysis of data including ranges means, standard deviation and variances. The mean or average of the data is the most used in this study to describe the central tendency of the data collection. The standard deviation was used to describe detailed estimate of dispersion due to the presence of outlier which can greatly affect the range of the data. standard deviation denotes the relation of the data sets to the meaning of the samples.

The ranges of the data collection were measured by calculating the minimum and maximum values. The minimum value of the data sets is the smallest observation whereas the maximum values of the largest observation within the data set of the whole samples, both values were used to represent the ranges of the data sets.

© COPYRIGHT UPM

3.7.2 Two-way Analysis of Variance (ANOVA)

The significant difference between WRD severity level and quality parameters on the optimal harvesting data was obtained by using a two-way analysis of variance (ANOVA). The Tukey test was used to analyze significant differences in the physicochemical properties (chlorophyll content, moisture content, color) across different levels of WRD severity, with a significance level set at $p \leqslant 0.05$. The data were presented as means and standard errors, analyzed using SAS software (version 9.4, SAS Institute, Cary, NC, USA).

3.7.3 Noise Removal

Due to imperfections, noticeable scattering noises that can affect measurement accuracy might be present at the beginning and end of the spectral data (Cui & Reiss, 2010). The first 300 nm and the last 200 nm were removed from the original reflectance spectral data to avoid a low signal-to-noise ratio. Only the wavelength range between 500 and 1000 nm was used for further analysis.

3.7.4 Pre-processing of the Spectral Data

The spectra of solid samples are influenced by their physical properties with scattering phenomena (which is wavelength-dependent and non-linear) being the most common factor for causing errors in absorbance values. Thus, pre-processing methods should be used to minimize the influence of irrelevant information on spectra to be able to develop more simple and robust models (Tsagkaris et al., 2023). The pre-processing of spectral data is a key part of spectral analysis used to improve the quality and

© COPYRIGHT UPM

accuracy of the regression models (Shi et al., 2023). The commonly used pre-processing methods have been reviewed in Section 2.6.

In this study, several pre-processing methods was investigated including smoothing by moving average, multiplicative scatter correction (MSC), baseline of correction (BOC), first and second derivatives, standard normal variate (SNV) transformation and mean normalization. BOC was found to be the best pre-processing technique for this study. The pre-processing processes were implemented using Unscrambler, V 10.4 Software (Camo process As, Oslo Norway).

3.7.5 Partial Least Square (PLS)

PLS regression and PCR are the two multivariate regression techniques that are most typically applied to spectrum data (Nantongo et al., 2024). Due to the absence of latent variables, which are less crucial for explaining the variation of the quality metrics, PLS models perform marginally better than PCR (Polson, 2021). Employing new axes known as LVs, PLS was able to determine the directions of highest variability by taking into account both spectral and measured property information (Sorochan et al., 2022). In order to lower the dimensionality and compress the original spectral data, the LVs were viewed as new eigenvectors of the original spectra (Agustika et al., 2022).

PLS is a popular multivariate factor analysis technique that is frequently applied to the analysis of spectral data for predictive analysis (Buvé et al., 2022). It needs a calibration step when a model is built from multiple important factors, chosen, for instance, using the popular cross-validation leave-one-out technique. Using spectral

© COPYRIGHT UPM

data (matrix X), PLS analysis was used to create a regression model that would forecast the physio-chemical quality parameters of rubber trees (matrix Y). In this work, the spectra were interpreted using this method, and rubber parameter calibration and prediction models were created. Generally, an acceptable PLS model has a maximum of ten LVs (Wang et al., 2022). As a result, eleven LVs in total were used in this investigation. This study used Unscrambler, version 10.32 (Camo Process AS, Oslo, Norway) for PLS analysis.

The aim of the PLS regression model in this study was to develop a calibration method (for spectral data analysis) to predict the properties of a sample (e.g., moisture content, chlorophyll content, etc.) y_i in the objects i=1,2,….., I from a set of spectral absorption x_{ik} at wavelength channels k=1,2…., K using a linear predictor equation as demonstrated in Eq 3.2 (Nawi 2014):

$$\hat{y} = b0 + \sum_{K=1}^{K} xijbk \qquad (3.2)$$

The predictor equation's parameters are statistically estimated using data from a set of I calibration samples, y = (y_i, i = 1, 2….I)^T and X = (x_{ik}, I = 1, 2…., K) Once acquired, this predictor equation can be applied to a fresh sample in order to transform X data into y estimations.

PLS is a technique for bilinear regression that takes a small number of linear combinations of the KX variables, or "factors," such as t_a, a = 1, 2, 3 …, A, and utilizes them as regressors for the variable y. Furthermore, these regression factors also model

© COPYRIGHT UPM

the X variables themselves. Consequently, it is possible to identify outliers with anomalous spectra xi in the calibration set or in subsequent data sets.

PLS differs from PCR in that it actively uses the y variable to define the regression component ta, which describes as much of the covariance between X and y as feasible after the preceding a-1 factors have been estimated and eliminated (Nawi, 2014). When compared to PCR, PLS's robust feature enables it to build an effective model with a small number of LVs, accelerating analysis and utilizing less computer power.

3.7.6 Calibration and Validation Methods

The mathematical procedure of calibration is necessary to connect spectral data to the targeted crop components. An important part of analytical measurement is played by multivariate calibration models. Making ensuring that the samples chosen for examination for calibration development reflect the range of spectral variation present in the entire population is the fundamental step in any calibration process (Foley et al. 1998).

Full cross-validation (leave-one-out) was employed in the PLS model's development to assess the calibration model's quality and avoid overfitting (Wang et al., 2005). In leave-one-out cross-validation, a calibration model was created for the subset of the dataset that remained after one sample was eliminated. The prediction residual was then computed using the eliminated samples. Once every sample had been excluded, the procedure was repeated with further subgroups, and the variance of each prediction residual was finally determined.

© COPYRIGHT UPM

The samples in the validation set were used to further validate the performances of the established calibration equations. The PLS calibration model's predictive power was assessed using the external validation method. The process of external validation established an equation's predictive capacity using a sample set that wasn't utilized throughout the calibration development. The obtained calibration equations' performances were further verified through the use of the. Two sets of samples were separated prior to calibration. A prediction equation (calibration set) was created using one portion of the samples (75%), and a validation set (25%) of the samples was utilized to confirm the prediction. One sample out of every four from the complete sample set was chosen for validation, being careful to include samples that encompassed the full range of high-quality data in each set. The model used for both calibration and validation was Unscrambler version 10.32 (Camo Process AS, Oslo, Norway).

The crop parameters of the samples, wavelength ranges, and pro-processing techniques on reflectance data were all tuned for the PLS model's performance. Using various pre-processing techniques, the improvement in the performance of calibration/prediction models in predicting rubber leaves crop parameters was evaluated.

3.7.7 Statistical Assessment for Model Accuracy

Rubber tree crop parameters measurements were associated with numerous highly correlated spectra, which led to the development of PLS regression; however, the prediction power of this method needs to be evaluated. According to Lopez et al.

© COPYRIGHT UPM

(2023), the models shouldn't be overly optimistic in the sense that they match the calibration data set too well while remaining insufficiently robust to take future sample variances into account. Standard error of calibration (SEC), root mean square error of calibration (RMSEC), coefficient of determination for calibration (R^2), standard error of prediction (SEP), root mean square error of prediction (RMSEP), coefficient of determination for prediction (R^2), and bias are common statistical parameters used to evaluate the models Althoff & Rodrigues, 2021). The following is a mathematical definition of SEP:

$$\mathrm{SEP}=\sqrt{\frac{1}{N-1}\sum\left(x_i-y_i-b\right)^2} \qquad (3.3)$$

Where b is bias and x_i -y_i is the difference between the reference method (y_i) on sample I and the expected response (x_i) based on spectral data. The term "bias" describes the discrepancy between the actual and expected values' means. The definition of bias is:

$$\mathrm{bias}=\frac{1}{N}\sum\left(x_i-y_i\right) \qquad (3.4)$$

Where N is the test sets of total number of samples.

The difference between the reference method values and the projected response in the spectral data is used to determine the root mean square error (RMSE), which is computed using Equation 3.5 (Agyeman et al., 2022). The RMSECV and RMSEP can be computed by extending this formula. Similar calculations are made for RMSEP and RMSECV; the only distinction is that RMSEP is determined when internal or external

© COPYRIGHT UPM

validation is performed, whereas RMSECV is acquired when cross-validation is employed (Cunha & Luna, 2020).

$$\text{RMSE}=\sqrt{\frac{1}{N}\sum\left(x_i-y_i\right)^2} \qquad (3.5)$$

SEP and bias are combined in RMSE, a single connection provided by:

$$\text{RMSE}^2=\text{SEP}^2+\text{bias}^2 \qquad (3.6)$$

With an increasing number of samples, the RMSEP trends towards the SEP when the bias is negligible. In general, the RMSEP provides a more accurate assessment of the calibration's prediction abilities than the SEP. The validation set is used to compute the SEP and RMSEP.

An additional valuable metric is the R^2 value. In essence, R^2 shows how much of the response variable's variance in the calibration or validation set can be explained. How well the measured reaction matches the expected response is indicated by the R^2 value. (Nawi, 2014) defines the R^2 as follows:

$$R^2=\frac{\sum(x_i-\tilde{x})\left(y_i-\tilde{y}\right)}{\sqrt{\sum(x_i-\tilde{x})^2\sum(y_i-\tilde{y})^2}} \qquad (3.7)$$

© COPYRIGHT UPM

in which ỹ represents the mean predicted value of samples in the prediction/validation population and x̄ is the mean of reference for the calibration population. For both prediction and calibration models, an appropriate model should have a high coefficient of determination and low SEC, SEP, and RMSEP. The best calibration equations in this investigation were selected using the lowest SEP values and the highest R^2 values.

3.8 Data Reduction Techniques

Reducing the dimensionality of data involves translating it from a high-dimensional space into a new one with a considerably lesser dimensionality. This procedure bears close relation to the notion of (lossy) information compression. Dimensionality reduction is the process of feature selection. Selecting only representative spectral bands for subsequent classification was made possible by the critical process of feature selection. Feature selection has several advantages, such as shortening processing times, making data easier to interpret, increasing algorithmic classification accuracy, and extracting more information features (Houssein et al., 2023).

The two main categories of feature selection techniques are feature subset selection techniques and feature raking techniques. The raking approach involves ranking features according to certain criteria that orderly weight their relevance. The features that are most crucial to the outcome are listed highest in the ranking features list, while those that are less significant are ranked lower. Different techniques employed various strategies to determine the important wavelength. In this research, three different feature selection techniques were applied. The study included feature selection techniques such as ANOVA, Chi-Square, and RELIEF to determine the relevant

© COPYRIGHT UPM

spectral bands that may be used to differentiate the severity level of WRD in rubber trees.

This feature selection based on scoring or ranking involves assigning scores to features in a dataset and selecting the top-ranked feature according to these scores. The process of feature selection in this study involves implementing a 10-fold cross-validation approach and selecting 50 features or wavelengths that give higher values of score. The open source software Orange was used to perform feature selection. Orange mining software was originally developed by scientists at University of Ljubljana in 1997 using the Python programming language. Orange software is a data visualization and analysis tool that provides a user-friendly interface for machine learning tasks, including feature selection and steps involve dragging and connecting widgets on canvas as shown in Figure 3.18.

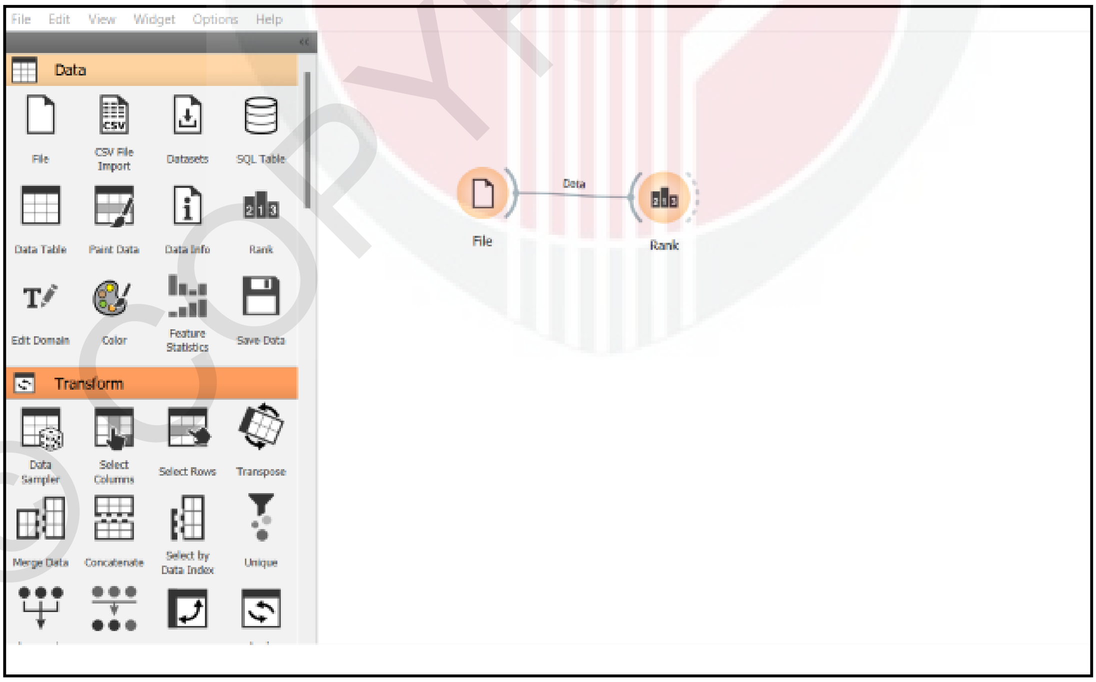


Figure 3.18: The widget of Orange software

© COPYRIGHT UPM

3.8.1 Feature Selection using ANOVA Technique

ANOVA (Analysis of Variance) is commonly used in statistics to analyze the differences among group means in a sample. In the context of feature selection, ANOVA can be employed to assess the significance of the relationship between each feature and the target variable. The formula for ANOVA involves calculating the F-score can be defined as:

$$F(\mu) = \frac{S_{MSB}^2(\mu)}{S_{MSW}^2(\mu)} \qquad (3.8)$$

Where $S_{MSB}^2(\mu)$ is Mean Square Between (MSB) and $S_{MSW}^2(\mu)$is the mean square within (MSW), and they are defined as;

$$S_{MSB}^2 = \frac{1}{df_{MSB}} \sum_{i=1}^{K} g_i \left[\frac{\sum_{i=1}^{g_i} f_\mu(i,j)}{g_i} - \frac{\sum_{i=1}^{K} \sum_{j=1}^{g_i} f_\mu(i,j)}{\sum_{i=1}^{K} g_i} \right]^2 \qquad (3.9)$$

$$S_{MSW}^2 = \frac{1}{df_{MSW}} \sum_{i=1}^{K} \sum_{j=1}^{g_i} \left[f_\mu(i,j) - \frac{\sum_{i=1}^{K} \sum_{j=1}^{g_i} f_\mu(i,j)}{\sum_{i=1}^{K} g_i} \right]^2 \qquad (3.10)$$

df_{MSB} stands for the degrees of freedom for MSB in these formulas, while df_{MSW} stands for the degree for MSW. K and G stand for all samples and the number of groups, respectively. Gi stands for the quantity of samples in each of the i-th groups. The frequency of the *j*-th sample in the *i*-th groups is represented by $f_\mu(i,j)$.

© COPYRIGHT UPM

Features with a higher F-statistic value or lower p-values are considered more significant in the context of ANOVA-based feature selection. ANOVA helps in identifying which features are significantly impacting the target variables making it a powerful tool for feature selection in classification modelling.

3.8.2 Feature Selection using Chi-square Technique

Chi-square is a frequently used statistical method in feature selection. Chi-square analyzes data that is categorical. By contrasting the observed frequency distribution with the expected distribution based on the independence assumption, it evaluates the degree of independence between the feature and the target variable. By applying Eq. 3.11 to compare the observed frequencies in each category with the anticipated frequencies if the feature and target were independent, the Chi-square statistic is computed.

$$X^2 = \sum_{i=1}^{2}\sum_{j=1}^{k}\frac{(A_{ij-Eij})^2}{E_{ij}} \quad (3.11)$$

where E_{ij} is the predicted frequency, k is the number of classes, A_{ij} is the number of patterns in the ith interval, and j is the class.

Features with a lower p-values (typically $p < 0.05$) are considered more significant in the context of Chi-square based feature selection. Chi-square helps in identifying which features are significantly impacting the target variables.

© COPYRIGHT UPM

3.8.3 Feature Selection using RELIEF Technique

Kira and Rendell (1992) provided a description of the Relief algorithm. Relief is the distance-based method's procedure, in which a feature's weight is determined by how separable a feature is from a target instance's class. The closest same-class data instances are referred to as "hits" and the closest different-class data instances are referred to as "missed" from the target instance (Akhter et al., 2021). Relief uses the distance difference between the miss and hit values to determine how much one class of data varies from another, taking into account one nearest neighbor (NN) from the target instance. According to Rosario and Thangadurai (2015), the feature weight value in relief ranges from -1 to 0. A feature with a negative weight produces noise and has no bearing on the output prediction. Greater positive weight indicates that the characteristic provides more pertinent information about the output. The Relief formula is as follows:

$$\omega_t=\omega_t-\sum_{x\in H}\frac{\mathrm{diff}(t, S_i,x)}{r*d}+\sum_{c\neq \mathrm{class}(S_i)}\left[\frac{p(c)}{1-p\big(\mathrm{class}(S_i)\big)}\sum_{x\in M(c)}\mathrm{diff}(t,S_i,x)\right]/(r*d) \qquad (3.12)$$

where S_i is the chosen data sample and ωt is the weight of feature t. The data points S, x d , and r reflect near hits and near misses, the number of nearest samples, the number of iterations, the frequency of occurrence of class C, and the chance that sample Si belongs to a class, respectively. One way to compute diff(*) is as follows.

where ωt means the weight of feature t, S_i is the selected data sample S, x represents the data:

© COPYRIGHT UPM

$$diff\ (t, S_i, S_j) = \left|\frac{S_{it} - S_{jt}}{max_t - min_t}\right|, \text{if the features are continuous values}$$

(3.13)

$$diff(t, S_i, S_j) = \begin{Bmatrix} 0\ S_a = S_{jt} \\ 1\ S_a \neq S_{jt} \end{Bmatrix}, if\ the\ features\ are\ discete\ value \quad (3.14)$$

3.9 Classification Models

The classifiers that were developed to allocate the data set based on the severity level of WRD were then developed using the ideal spectral bands that were discovered through feature selection. Utilizing the training accuracy (%) and validation (%) obtained from the classification processes, the optimal spectral bands' accuracy in differentiating between samples of infected leaves at varying disease severity levels was measured.

To identify the best and most appropriate classifier model to differentiate between diseased leaf samples at varying disease severity, a number of machine learning classifications are tested. The classification is carried out with Orange software. The Orange software implements various classifier algorithms, including random forest (RF), support vector machine (SVM), K-nearest neighbor (kNN) artificial neural network (ANN), and others, to conduct classification on spectral data. When selecting four ML classifiers, data characteristics, computational resources, interpretability, and performance metrics are considered. Each model suits different types of data, from high-dimensional (ANN, SVM) to more straightforward (kNN, RF) datasets (Silva & Eugenio, 2020). kNN and RF are less demanding in terms of training resources while ANN and SVM require more computational power, especially for large datasets

© COPYRIGHT UPM

(Cebekhulu et al., 2022). In terms of interpretability, kNN and RF provide more interpretability while ANN and SVM can be harder to explain, especially with more complex versions (Elshawi et al., 2019). Each algorithm has its strength based on the nature of the data. RF is a solid all-rounder, while ANN and SVM shine in specialized tasks, and kNN is a simple but effective for smaller datasets.

The entire set of pre-processed spectral data, with wavelengths ranging from 450 to 1000 nm, has been used for categorization. Thirty percent of the remaining data set is gathered as testing data, with the remaining seventy percent designated as training data. To randomize the models, every classifier is trained and tested ten times on every dataset. As a result, the model's consistency and durability may be seen. Since the confusion matrix is more frequently employed to gauge a machine learning classification model's effectiveness, the performance of the models.

3.9.1 Classification using Support Vector Machine Classifier

This study applied the SVM classifier as one of the used classifiers due to its robustness against high-dimensional data and its capability of excellent performance when dealing with high-dimensional data. The algorithm was developed by Vapnik, (1999) based on statistical learning theory that works by mapping vectors into high dimensional space and using the SVM to find linear separating hyperplanes with maximal margin. In a classification system, linear SVM can be used to optimize the marginal distance calculated perpendicular to the hyperplanes along the line to identify hyperplanes that distinguish between the different levels in the training data set (Huang et al., 2014).

© COPYRIGHT UPM

This equation is obtained from assumed N training data pairs, {*xi, yi*}, I = 1,……,N.

Here, n is the dimensional space, x1; x2,…. xn are the input variables, and x(i) is represented as either +1 or -1. The formula for x(i) is {x1(i), [x2(i),….. xn(i)]}T. In the presence of scalar b and vector w, which are represented in the equation as given below by Mao (2004), the training patterns are separable linearly:

$$w.x + b = 0 \qquad (3.15)$$

where w and b represent the weight vector and scalar vector, respectively. The weight vector (w) obtained from the equation as below;

$w *= argmin\frac{1}{2}\|w\|^2$ (3.17)

The linearly separable classes were exposed to

$$y(i)\lfloor wT\ x(i) + b\rfloor \geq 1, i = 1,2, \ldots \ldots N. \qquad (3.16)$$

The efficiency of the significant frequencies produced by the features selection models was evaluated in this study using the radial basis function (RBF) kernel classification of SVM technique. The Grid Search approach is used to get the data that needs to be optimized in the RBF kernel parameter for maximum performance. Gamma = 10 and c = 16 are the ideal RBF kernel parameters for the classification job in this study (Samsudin et al., 2015).

© COPYRIGHT UPM

By applying the "kernel trick," one can significantly improve the classification skills of conventional Support Vector Machines (SVMs) by converting the initial feature space into a higher-dimensional feature space. SVMs are useful in a variety of applications because they address overfitting issues that arise in high-dimensional spaces and are based on global optimization.

3.9.2 Classification using Artificial Neural Network Classifier

The structure and operation of the human brain, which is modeled as a network of processing units called neurons, served as the model for artificial neural networks (ANNs) (Schmidgall et al., 2024). They are made up of many interconnected processing units (neurons), and the weights assigned to each link represent its strength. The multi-layered perceptron (MLP) structure in backpropagation (BP) is frequently employed in this type of research (Zorkafli et al., 2019).

A layer of input neurons, a layer of output neurons, plus one or more hidden layers typically make up an ANN structure. Information moves from the ((l-1)th layer to the lth layer in an MLP since there is no connection between the neurons in a particular layer. The input nodes allow external datasets to enter the network and perform non-linear changes; the output nodes produce the output values. The data that the input nodes receive is processed by hidden nodes that have the proper non-linear transfer functions. The mathematical notation for the model is;

$$y_t = \alpha_0 + \sum_{j=1}^{n} \alpha_j f\left(\sum_{i=1}^{m} \beta_{ij} y_{t-i} + \beta_{0j}\right) + \varepsilon_t \qquad [i = 1, \dots . m \text{ and } j = 1, \dots . . n]$$

(3.17)

© COPYRIGHT UPM

The weights from the input to hidden nodes are indicated by $ß_{ij}$, while the weights from the hidden to output nodes are indicated by a_j. The numbers m and n represent the number of input and hidden nodes, respectively. A sigmoid transfer function is represented by f, and the weights of the arcs that lead from the bias terms are represented by α_0 and $ß_{oj}$, which are always equal to 1.

This study used a variety of input factors, as shown in Figure 3.19, such as color, moisture content, and chlorophyll content as the input layer. We employed a technique called "N-fold cross-validation," in which data are split into two sets at random: a training set comprising 70% of the total data, and a cross-validation set including the remaining 30% of the data. By using the "Leave N out" option, networks were trained repeatedly while being excluded from various data segments for N training runs. When evaluating a model's robustness on tiny datasets, this training process is quite helpful. In this case, we separated the 26 data into six equal-sized subsets (N = 6). Next, we trained the net N times, excluding one subgroup from each training run but utilizing only the excluded subset to calculate the error criterion each time. The 6-fold cross-validation is chosen because it strikes a balance between the computational efficiency and the reliability of the validation process (Boukhlifa & Chibani, 2024). It is particularly useful to ensure the model generalizes well while managing computational costs effectively (Mira et al., 2024).

© COPYRIGHT UPM

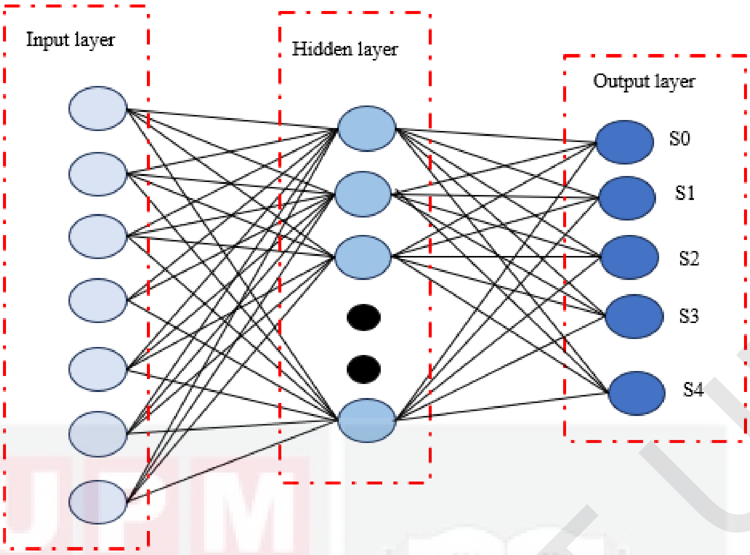


Figure 3.19: The Selected Multi-layer Neural Network Topology

Every node in the subsequent layer has a connected weight to every other node, which is computed to provide the sum of the nodes (x), which represents the node's activation input value function. Until the activation function of the output node is obtained, the value of x is first computed, and then the activation function of the node. The activation function of a hidden layer comprising many nodes is utilized to process the data that the input nodes receive. Several activation functions do exist; sigmoid, step, linear, and sign function (Sharma & Athaiya, 2017). In this study, sigmoid function is used as the activation function due to its popularity applied in previous studies.

Back-propagation technique was used in this study to run ANN because of its benefits, which include low memory requirements and quick processing (Pahlavan et al., 2012). To train data, this method underwent four stages: initialization, activation, weight training, and iteration. In this instance, gradient descent, conjugate gradient, and Levenberge-Marquardt methods can all be used to minimize errors. The gradient

© COPYRIGHT UPM

descent technique, which is stable when a small learning rate is utilized, such as when utilizing a variable learning rate and adding a momentum factor, is what the back-propagation algorithm uses in this application. The learning rate and epoch momentum parameters in this process were set at 0.3, 0.2, and 500, respectively.

In this study, a back-propagation algorithm was applied to run ANN due to its advantages; little memory required and fast processing (Pahlavan et al., 2012). This algorithm went through four steps; initialization, activation, weight training, and iteration; to train data. In this case, the error minimization can be obtained by several procedures: gradient descent, conjugate gradient, and Levenberge-Marquardt. In this application, the back-propagation algorithm used the gradient descent technique because it is stable when a small learning rate us used like using a variable learning rate and adding a momentum term. In this process, the parameters of the learning rate, and momentum of epoch were set as 0.3, 0.2, and 500, respectively. To prevent oscillations in the error and accelerate convergence, the previous study proposed these parameter settings.

3.9.3 Classification using k-Nearest Neighbor Classifier

The supervised machine learning technique known as k-nearest neighbor, or kNN, is widely used. This is a straightforward classifier technique that uses k-values to classify an unknown sample by counting how many of its closest neighbors are known samples. Using cross-validation, kNN can determine the right value for k and carry out a distance weighting. When k is small, the algorithm may overfit and become more susceptible to noisy data points. However, a high number of k may result in

© COPYRIGHT UPM

underfitting, when the algorithm fails to detect specific patterns within the data. Therefore, choosing the right number for k is crucial to getting the best results.

The kNN method frequently employs the Euclidean distance as a distance metric (Karadağ et al., 2020). The measurement of the straight-line distance in a Euclidean space between two points is called the Euclidean distance. It is computed in mathematics using the Pythagorean theorem. Between these two places, the Euclidean distance is computed as follows:

$$distance = \sqrt{\sum_{i=1}^{k} (x_i - y_i)^2} \qquad (3.18)$$

3.9.4 Classification using Random Forest Classifier

The Random Forest (RF) Classifier is an ensemble learning technique that generates a large number of decision trees during training and outputs the class that represents the mean of all classes (classification). This technique is referred to as "bagging" or "bootstrapped sampling." It entails taking replacement samples of the data at random to create several subsets that are then used to train different trees. The trees vote to choose the final classification task by a majority.

3.9.5 Performance Evaluation of Classifiers

Four statistical indices are computed using the confusion matrix to understand performance metrics including accuracy, precision, recall, specificity, and F1-score. An overview of a classification algorithm's prediction outcomes is called a confusion

© COPYRIGHT UPM

matrix. Count values are used to describe the number of accurate and inaccurate forecasts, broken down by class. Equations (2) through (7) were used to construct the different performance parameters, which may be used to assess the validity of a classification model by looking at its elements in the confusion matrix.

Models that can accurately identify diseases with high-risk levels will demonstrate their accuracy. Accuracy with a formula;

$$Overall\ Accuracy = \frac{TP + TN}{TP + EP + TN + FN} \quad (3.19)$$

where false negative (FN), false positive (FP), true positive (TP), and true negative (TN) are found. The F1-score calculates the test's accuracy and the evaluation formula;

$$F1 - score = \frac{2 * TP}{2 * TP + FP + FN} \quad (3.20)$$

$$sensivity = \frac{TN}{TP + FN} \quad (3.21)$$

$$Recall = \frac{TN}{TP + FP} \quad (3.22)$$

3.10 Statistical Analysis for Mean Comparison

The Statistical Analysis System software (SAS version 9.2, Institute. Inc., Cary, N. C.) was used to conduct statistical analysis. The mean classification accuracies of several

© COPYRIGHT UPM

classifiers with and without data reduction techniques in classifying WRD severity levels were compared using an ANOVA at the 5% level of significance.

3.11 Summary

In this study, five different levels of WRD severity, namely S0, S1, S2, S3, and S4, were employed. The classification of rubber trees into severity levels was performed by expert workers from RRIM. A total of 163 rubber trees were used in this study. For each tree, three samples were collected, encompassing latex, leaves, and soil around the tree. For the leaves samples, the spectral data and crop parameters (SPAD value, color, and MC) were measured. For the soil samples, the MC and pH values of the soil around the rubber trees were measured. In the case of latex, the study focused on determining the total soluble content. A Partial Least Squares (PLS) model was developed to ascertain the spectral relation among leaves. Feature selection was implemented to select the most relevant wavelength from the dataset to enhance the accuracy of classification, aiding in the discrimination of plants at various severity levels.

© COPYRIGHT UPM

CHAPTER 4

RESULTS AND DISCUSSION

In the previous chapter, the method in this study was outlined for five different severity levels of WRD, namely S0, S1, S2, S3, and S4, for a sample of 134 rubber trees. This chapter discusses the statistical properties of rubber leaf samples, latex samples, and soil at different WRD severity levels. Additionally, this chapter presents the results of classifying rubber leaves according to the five WRD severity levels.

4.1 Molecular Identification of White Root Disease Pathogen

Molecular identification focuses on identifying the *Rigidoporus micropores* 's genetic material (DNA) using PCR techniques. The process involves using short DNA sequences known as primers that match the DNA of the fungus. The result from the basic local alignment search tool (BLAST) search showed that all internal transcribed spacer (ITS) sequences of the white root disease pathogen isolate from RRIM were 90% identical to *Rigidoporus micropores*, as all performed the highest level of similarity of 99% with isolate SEG from Malaysia with reference pathogenic strain *Rigidoporus micropores* (Acc. No*:* MF574815.1). It can be used as a reference pathogenic for molecular confirmation as it shares the higher percentages that identify a similarity with host and country origin.

© COPYRIGHT UPM

4.2 Non discrimination of Age in Plant Disease

In this study, rubber tree samples were collected from two different age groups: 5-year-old trees at Sungai Buloh and 10-year-old trees at Kota Tinggi. The t-test is used to determine if there are statistically significant differences between disease severity levels and the age of rubber trees. Statistical analysis using t-test showed that the incidence of WRD does not significantly differ across the ages of rubber plants, as indicated by the lack of significant variation in infection rates between age groups p-value (0.001) more than 0.05. This means that both young and mature plants are equally susceptible to the disease.

Rigidoporus microporus is primarily a soil-borne pathogen, with its spores and mycelium living in the soil, making it a persistent threat to all rubber plants growing in infected soil (Farid et al., 2023). Since the fungus spreads through direct root contact between infected and healthy plants, the risk of infection is tied to environmental factors like root proximity rather than the plant's age. This explains why young and mature plants are equally susceptible to the disease, as the pathogen's ability to transfer and infect does not depend on the plant’s age but rather on root contact in contaminated soil (Go et al., 2023). Consequently, statistical analysis using t-test shows no significant difference in WRD incidence across different age groups.

4.3 Statistical Characteristics of the Crop Parameters

Statistical characteristics of the crop parameters of rubber plants involved describing the mean, maximum, minimum, and SD and relationships of crop parameters with

© COPYRIGHT UPM

WRD severity level. This study involves rubber leaves, latex, and soil samples from different severity levels.

4.3.1 Statistical Characteristics of Leaves Sample

A total of 600 rubber leaves representing 5 different severity levels were used in this study. Each rubber leaf was measured for SPAD value, moisture content, and color. These crop parameters were described by means, standard deviation, and ranges of quality parameters of the samples.

4.3.1.1 Statistical Characteristics of the SPAD Value

The statistical characteristics of the SPAD value for rubber leaf samples are shown in Table 4.1. Results showed that there is has significant difference in SPAD value among different severity levels with p-value (<.001) less than 0.05. The result presented for the SPAD value indicates an inverse relationship between the leaf content of SPAD value and the severity level of WRD. Healthy leaves or S0 contain a higher amount of SPAD value and vice versa. The highest mean value of SPAD value was found in the healthy (S0) leaves with an average value of 60, followed by S1, S2, and S3 with values of 56, 54, and 51, respectively. While the lowest value of SPAD value is at S4 with an average value of 47.

© COPYRIGHT UPM

Table 4.1: Characteristics for SPAD value of rubber leaves samples at different severity levels

Severity level	Max	Min	Mean	SD
S0	70.00	35.40	[a]60.00	7.60
S1	62.70	48.90	[a]56.00	5.81
S2	61.60	20.80	[b]51.00	8.90
S3	66.00	15.90	[c]50.60	11.05
S4	63.30	22.70	[d]47.00	9.79

All figures were obtained from means ± SD of three replications. *

This result showed that SPAD value was decreased from healthy to very severe-infected WRD severity levels. This correlates to the mechanism of WRD disease on rubber trees. In a severe infection, there is no development of new leaves and no increase in height or girth that will consequently lead to growth inhibition. These are all attributed to the inability of the plant to perform photosynthesis due to low SPAD value.

Similar results were found in the study by Goh et al. (2016), who discovered when the disease worsens or throughout the pathogenesis of Ganoderma in oil palm seedlings, SPAD values decrease. Chang et al. (2015) also observed that a decline in the amount of SPAD value was brought on by an increase in disease in different stages of plant growth. Significantly lower SPAD value may be linked to fungal infection-induced damage to the oil palm seedlings' vascular and root systems, which may have decreased their ability to absorb nutrients, particularly nitrogen (N) and magnesium (Mg) (Goh et al., 2016). According to Melo et al. (2014), there is a positive correlation between the amount of SPAD value in plants and nitrogen and magnesium, which are linked to the manufacture of SPAD value.

© COPYRIGHT UPM

4.3.1.2 Statistical Characteristics of the Moisture Content

The statistical characteristics of the MC for rubber leaf samples are shown in Table 4.2. For the MC, the findings from this study indicate an inverse relationship between the moisture content of leaves and the severity level of WRD. Healthy leaves or S0 contain more MC and vice versa. Results showed that there is has significant difference in moisture content among different severity levels with a p-value (<.0001) less than 0.05. The highest moisture content value was found at the healthy (S0) leaves with an average value of 56, followed by S1, S2 and S3 with values of 54, 54, and 53, respectively. While the lowest value of MC was at S4 with an average value of 41.

Table 4.2: Statistical characteristics for MC of rubber leaves samples at different severity levels

Severity level	Max (% w.b.)	Min (% w.b.)	Mean (% w.b.)	SD
S0	65.91	17.05	^{a}56.29	8.22
S1	67.00	37.42	^{b}54.47	8.29
S2	77.08	33.18	^{b}54.03	8.75
S3	59.25	45.59	^{b}53.21	3.34
S4	53.04	33.18	^{b}41.95	7.91

All figures were obtained from means ± SD of three replications

This result showed that the mean values of MC decreased from S0 to S4. A decrease in the mean value of MC indicates early plant stress (Azadnia et al., 2023). This is because the rubber trees lose water when the disease progresses. When WRD infects a tree, it releases enzymes that damage the tissue, preventing water and nutrients from reaching the tree's upper branches, including the leaves. This damages the cells' organelles, proteins, and pigments (Khaled et al., 2020). According to Buxton & Casler (1993), WRD causes lignin to break down into water and CO_2 in the stem

© COPYRIGHT UPM

tissues' cell walls. Fungal activities are the consequence of lignin breakdown. This has a direct impact on plants' vascular circulation, which restricts their ability to absorb nutrients and water. As a result, the water content of the tree's upper section will decrease, especially near leaves.

This finding aligns with Wu et al. (2023), who highlighted that healthy strawberry leaves have more MC compared to infected leaves. This study also has similar findings to Khaled et al. (2020), who showed that the water content decreased as the disease progressed in the palm oil. Insufficient water availability can lead to reduced plant growth and impaired photosynthesis.

4.3.1.3 Statistical Characteristics of the Color

The color coordinates of rubber samples at different severity levels are presented in Table 4.3. There were changes in the values of all coordinates. S4 had the higher mean value of L* (38.43 ± 4.97), followed by S3 (36.98 ± 5.24), S2 (36.89 ± 4.85), S1 (33.70 ± 1.88) and S0 (32.97 ± 3.37). For the mean value of a*, S0 showed the higher value (-1.58 ± 1.92) followed by S1 (-1.59 ± 2.38), S2 (-1.61 ± 3.18), S3 (-1.78 ± 3.15) and S4 (-2.16 ± 3.04). For the mean value of b*, S0 showed the higher value (12.64 ± 3.44) followed by S1 (13.95 ± 2.05), S2 (16.14 ± 5.99), S3 (16.39 ± 6.67) and S4 (17.79 ± 6.39).

© COPYRIGHT UPM

Table 4.3: Color parameter for rubber leaves samples at different severity levels

Severity levels	Color parameter		
	L*	a*	b*
S0	$^{a}32.97 \pm 3.37$	$^{a}-1.58 \pm 1.92$	$^{a}12.64 \pm 3.44$
S1	$^{a}33.70 \pm 1.88$	$^{a}-1.59 \pm 2.38$	$^{a}13.95 \pm 2.05$
S2	$^{b}36.89 \pm 4.85$	$^{a}-1.61 \pm 3.18$	$^{b}16.14 \pm 5.99$
S3	$^{b}36.98 \pm 5.24$	$^{a}-1.78 \pm 3.15$	$^{b}16.39 \pm 6.67$
S4	$^{b}38.43 \pm 4.97$	$^{a}-2.16 \pm 3.04$	$^{b}17.79 \pm 6.39$

All figures were obtained from means ± SD of three replications

Results showed that there is has significant difference in L* value among different severity levels with a p-value (<.0001) less than 0.05The L* value represents the lightness or brightness of a color in a CIELAB color space. It ranges from 0 (black) to 100 (white). This result showed that the L* values of WRD severity levels increased from S0 to S4. As the disease progresses, the rubber leaves undergo chlorosis, where the leaf loses chlorophyll content, leading to increased lightness. Chlorophyll is responsible for the green color and helps the leaf absorb light. The leaves become yellow or pale when chlorophyll is lost (Ghouri et al., 2022).

Results showed that there is not significant difference in a* value among different severity levels with a p-value (0.36) more than 0.05. The a* value typically refers to the red-green axis in the CIELAb color space. In this space, negative a* values indicate a shift toward green. This result showed the a* values of WRD severity levels decreased from S0 to S4 in the negative region. As the severity of WRD progresses, the a* value becomes progressively more negative, indicating a reduction in the intensity of the green color. As the infection worsened, the leaves exhibited less greenness, which could be attributed to chlorophyll degradation due to the pathogen's

© COPYRIGHT UPM

impact on the plant tissue (Zhang et al., 2021). SD value is higher than the mean value, typically indicating that the data is highly dispersed because of the presence of outliers or extreme values. The presence of dark spot on leaves can lead to outliers in a* value measurements because these spots often indicating of some stress on leaves. These spots can have significant effect on how light interact with leaf surface.

Results showed that there is has significant difference in b* value among different severity levels with a p-value (<.0001) less than 0.05The b* value typically refers to the blue-yellow axis in the CIELAb color space. In this space, positive b* values indicate a shift toward yellow. This result showed the b* values of WRD severity levels increased from S0 to S4 in the positive region. The rising b* value correlates with visible symptoms of disease severity, such as chlorosis or yellowing of the leaves. Initially, healthy leaves exhibit low b* values. However, as the disease severity escalates, the gradual increase in the b* value reflects the degradation of green chlorophyll and the relative prominence of yellow pigments.

These trends suggest a clear relationship between disease progression and colorimetric changes in rubber leaves, providing a visual indicator that could potentially be used for early disease detection and assessment. The increasing L*, a*, and b* values directly result from the plant's physiological decline, leading to significant implications for plant health monitoring and management practices.

4.3.2 Statistical Soil Properties of Samples

This study used 134 soil rubber trees representing 5 different severity levels. pH and MC for each soil sample around rubber tree soil were measured. The chemical

© COPYRIGHT UPM

properties of the samples were described by means, standard deviation, and maximum and minimum quality parameters.

4.3.2.1 Statistical Characteristics of the pH

Table 4.4 shows the statistical characteristics of the pH of soil at different severity levels. Results showed that there is no significant difference in pH value among different severity levels with a p-value (0.070) more than 0.05. The healthy trees (S0) showed the lowest mean value of pH compared to unhealthy trees (S1, S2, S3, and S4).

Table 4.4: Statistical characteristics pH soil samples at different severity levels

Severity level	Max	Min	Mean	SD
S0	6.50	5.20	^{a}6.20	0.29
S1	6.54	6.23	^{a}6.40	0.11
S2	6.70	6.11	^{a}6.30	0.19
S3	7.00	6.10	^{a}6.40	0.31
S4	6.50	6.30	^{a}6.40	0.08

All figures were obtained from means ± SD of three replications

The mean pH value for S0 shows the lowest value (6.20) compared to the S1 (6.40), S2 (6.30) S3 (6.40) and S4 (6.40). It indicates that the plants appeared to be most healthy when the soil was distinctly acidic. Soil pH strongly influences microbial communities (Matthies et al., 1997). This finding aligned with research conducted by Go et al. (2021), which focused on examining the radial growth of *Rigidoporus microporus* at various pH values. The study revealed that the highest mycelial growth for *Rigidoporus microporus* isolates occurred at pH 6.5, while the lowest mycelial

© COPYRIGHT UPM

growth was observed at pH 3.5. Consistent with the findings of Rodesuchit et al. (2012) and Wahyuni et al. (2018), neutral pH favoured the growth of *Rigidoporus microporous*, whereas fungal growth was impeded at low pH value. Dede et al. (2011) also noted higher occurrences of WRD in areas with elevated pH values within rubber plantations.

In practical applications, estate farmers have traditionally amended the soil with sulfur to hinder the growth of *Rigidoporus microporus*, aiming to reduce the reliance on pesticides or predate the availability of fungicides (Mahyudin et al, 2023). This sulfur addition, with minimal adverse effects on plants, elevates soil acidity and is believed to contribute to slowing down the growth of *Rigidoporus microporus*. This practical is commonly known as acidification of soil. Changes in soil pH can affect the availability of essential nutrients for fungi (Song et al., 2023). The acidic soil creates high levels of hydrogen ions (H+) that can have direct toxic effects on fungi (Bolan et al., 2023). The increased concentration of H+ ions can disturb cellular processes, interfere with membrane integrity, and impede fungal metabolic activities, ultimately inhibit growth (Fincheira et al., 2023).

4.3.2.2 Statistical Characteristics of the Moisture Content

Table 4.5 shows the statistical characteristics of soil MCs at different severity levels. Results showed that there is has significant difference in MCs among different severity levels with a p-value (<.0001) less than 0.05. A significant difference in MCs was observed between healthy and unhealthy rubber trees. However, the differences in MC across varying severity levels of WRD were not statistically significant.

© COPYRIGHT UPM

The mean value of MC for S0 shows the highest value (33.50% w.b.) compared to the S1 (23.30 w.b), S2 (24.14 w.b) S3 (21.36 w.b) and S4 (21.70 w.b). The MCs in the soil is a critical factor that influences the water status of plants. Inadequate MCs in plants disrupts the balance between water uptake and loss, thereby affecting normal physiological functions (Bi et al., 2021; Suresh et al., 2018). Insufficient MCs of can inhibit the absorption of essential nutrients by plant roots, potentially leading to nutrient deficiencies and weakening the plant's ability to resist pathogens. (Sinha et al., 2021). Conversely, enough MC in the soil surrounding healthy trees creates an environment conducive to plant growth. Proper soil moisture plays a crucial role in facilitating nutrient uptake and promoting overall plant well-being.

Table 4.5: Statistical characteristics of moisture soil samples at different severity levels

Severity level	Max (% w.b.)	Min (% w.b.)	Mean (% w.b.)	SD
S0	47.9	24.30	^{a}33.50	6.70
S1	27.90	20.40	^{b}23.30	2.82
S2	52.90	15.60	^{b}24.14	6.80
S3	26.10	15.30	^{b}21.36	3.51
S4	26.2	18.90	^{b}21.70	3.02

All figures were obtained from means ± SD of three replications

In a research conducted by Yang et al. (2023), it was noted that elevated soil MC in the rhizosphere of a pine forest was associated with diminished levels of bacterial communities and reduced abundance of fungi. Deng et al. (2023) reported that a decrease in MC led to a decline in chlorophyll content, consequently causing alterations in the spectral characteristics of the leaf. The chlorophyll content also

© COPYRIGHT UPM

decreased resulting in changes to spectral. These changes in spectral characteristics are expected to align with a corresponding shift in Section 4.5.

In summary, soil properties can influence the composition of microbial communities and inhibit disease occurrence. However, using soil properties as reliable predictors of disease occurrence requires a nuanced understanding of the specific interactions involved and consideration of additional factors influencing plant health.

4.3.3 Statistical Characteristics of Latex Samples

4.3.3.1 Statistical Characteristics of Total Solid Content

Table 4.6 shows the statistical data of TSC obtained from the referenced method used in this study. It consists of the maximum value, mean value, minimum value, and standard deviation. At the S4 levels, the trees no longer produce latex, which explains the absence of TSC data. Results showed that there is has significant difference in TSC among different severity levels with a p-value (0.0001) less than 0.05. TSC indicates the percentage of solid particles contained in the latex.

Table 4.6: Total soluble solid characteristics at different severity levels

Severity level	Max (% w/w)	Min (% w/w)	Mean (% w/w)	SD
S0	2.24	1.37	^{a}1.90	0.29
S1	2.04	1.76	^{a}1.85	0.20
S2	2.24	0.83	^{b}1.71	0.43
S3	1.82	1.71	^{b}1.58	0.16

All figures were obtained from means ± SD of three replications.

© COPYRIGHT UPM

WRD can affect the overall health and physiological functions of the tree and potentially reduce latex production. S0 shows the higher value of TSC while S3 shows the lowest value of TSC. The difference in value in TSC may be indicative of the presence of WRD. Disease infection in rubber trees can result in disturbances to regular plant processes. One common response to this condition is the breakdown of chlorophyll, the green pigment responsible for photosynthesis (Mokhatar et al., 2011). As chlorophyll breaks down, leaves may turn yellow indicating a decline in photosynthesis activity (Guo et al., 2023). A decrease in photosynthesis limits the plant's ability to produce carbohydrates, which are essential for the synthesis of latex and other plant components. Carbohydrates produced through photosynthesis serve as precursors for the formation of latex (Gianfagna, 1995). A reduction in photosynthesis results in lower carbohydrate availability for latex biosynthesis. The overall impact is a reduction in the quantity and quality of latex produced by stresses trees.

4.4 Identification of Outliers

The leading causes of outliers in experiments are sample acquisition errors, operator error, and spectrometer instrument performance. Identifying them helps ensure that the data analysis is based on accurate and reliable data. Selecting the best representative sample set for calibration or identifying samples that differ significantly from most of the remaining data is always a difficulty for VSNIR users. It is necessary to examine data variances to identify outliers (Heidari et al., 2019). For this study, two spots were found outside the 95% confidence ellipse by PCA of the spectral data (Figure 4.1). The loading plot and score plots for the first and second principal components (PC1 and PC2, respectively) are displayed in Figure 4.1. According to the

© COPYRIGHT UPM

labels on the score plot, PC1 accounts for 84% and PC2 for 13% of the overall variance in the data. Two data points in total were identified as outliers.

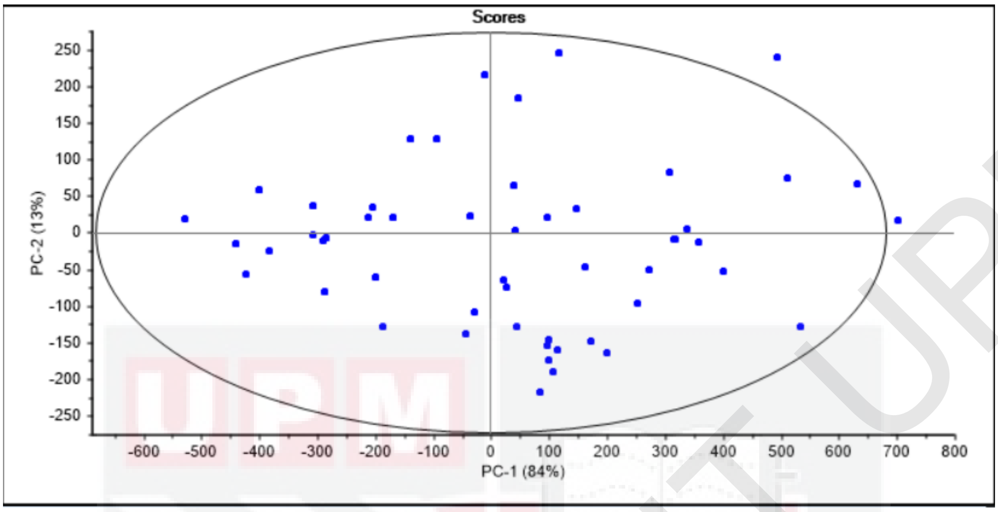


Figure 4.1: PCA Detection Outlier of the Spectral Data. The Blue Line Represents a 95% Confidence an Ellipse

4.5 Spectral Curve for Reflectance Measurement of Rubber Leaves

The reflectance data for all leaf samples were collected using VSNIR spectroscopy under constant experimental setups. Figure 4.2 shows the reflectance spectra data after removing the noise. The overall pattern of the mean reflectance spectra was similar for all severity levels. The figure demonstrates that the severity level of the disease can be identified.

In this study, the Mann-Whitney U test was applied at a significance level of 0.05 to analyze the reflectance spectra across the five severity levels of WRD. The Mann-Whitney U test compares two groups or datasets to determine if there is a significant difference in their probability distributions. In this case, no significant difference was found between the groups. The advantage of this method is that it can be used for small

© COPYRIGHT UPM

set sample size and does not require normality assumption (Izzuddin et al., 2017). Null hypothesis of the test is that there is no difference in the mean of the populations from which the two sample groups were taken (the population are identical) and alternative hypothesis is that there is difference in the mean of two populations (the populations are not identical).

Using this method, the difference between reflectance spectra was investigated by pairing the classes a pair at a time. In this case, there are ten possible pairs that were analyzed using the Mann-Whitney U Test which were S0 versus S1, S0 versus S1, S0 versus S2, S0 versus S3, S2 versus S4, S1 versus S2, S1 versus S3, S1 versus S4, S2 versus S3 and S3 versus S4. The results are p-values for each pair where the p-values show the values of significant of difference for each wavelength. The significant wavelengths that have p value lower than 0.05.

The spectra for the five severity disease levels closely overlap with each one in the wavelength region NIR region (700 nm to 900 nm). While spectral reflectance in the visible region (450 nm to 700 nm) varies and easily differentiates the severity levels of samples. These results indicate that the visible regions are crucial wavelengths to discriminate the WRD at severity levels.

When the plants are stressed, such as by disease, their absorption of incident light changes and produces a spectral signature that is different from the signature of healthy plants (Martinelli et al., 2015). This response is probably due to decreased chlorophyll content, changes in other pigments, foliar internal structure, and changes in absorption consequently influencing the reflectance of stressed plants.

© COPYRIGHT UPM

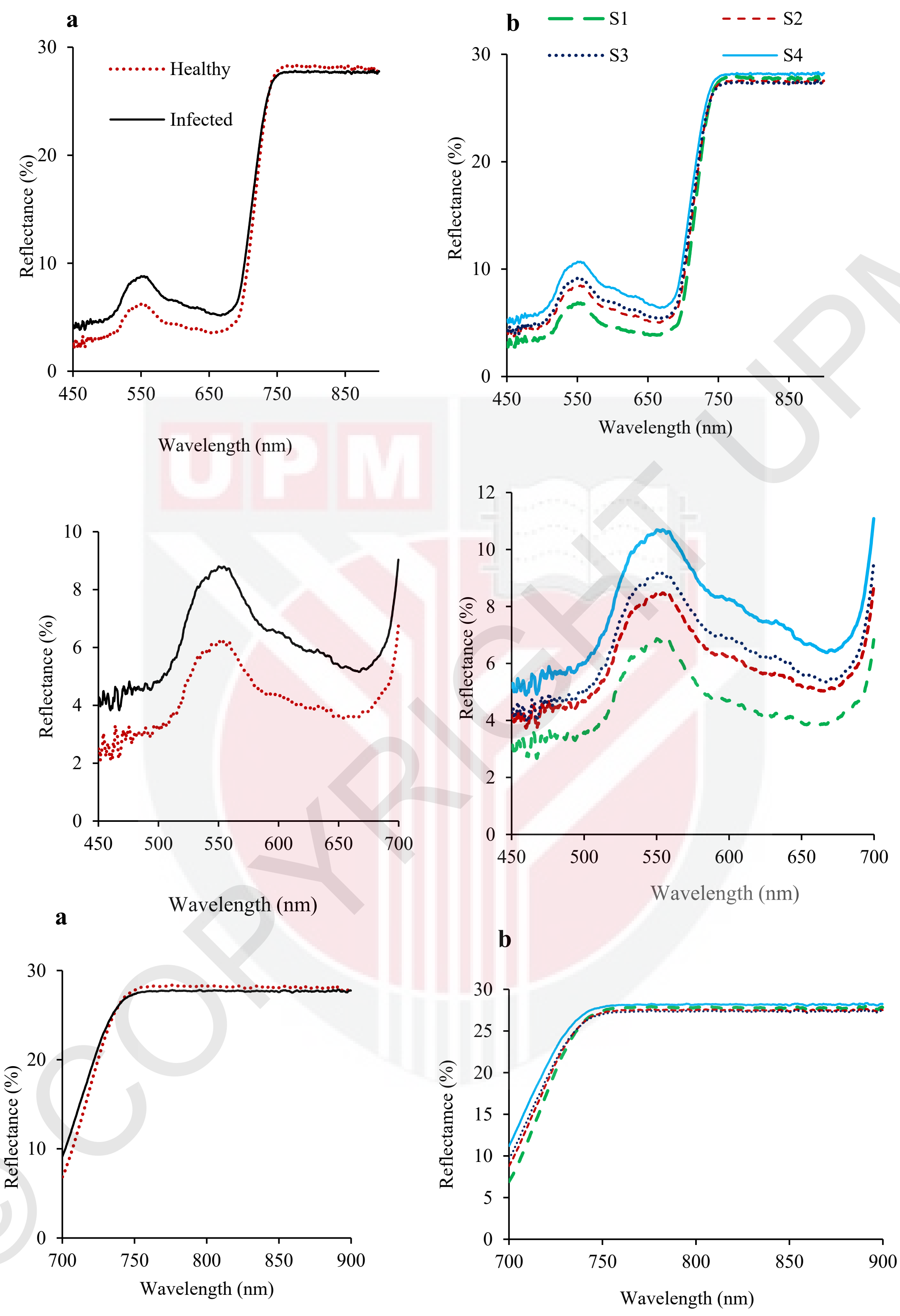


Figure 4.2: The Spectral Reflectance of Rubber Leaves (a) Healthy and Infected WRD (b) Infected WRD at Different Severity Levels Across Different Wavelength Regions

© COPYRIGHT UPM

The color in the leaves is created by a mixture of pigments, including green chlorophyll. Chlorophylls are the most common pigments found in the leaves which influence the coloration of the whole visible reflectance spectra of the leaves. Chlorophyll is found in chloroplast, which absorbs light to drive the photosynthesis process into which the chlorophyll converts CO_2 and H_2O into simple sugar with the presence of sunlight.

4.6 Effect of Different Pre-processing on Data Accuracy

The purpose of pre-processing spectral data is to remove physical phenomena in the spectra to improve subsequent multivariate regression. The PLS models were developed based on raw spectra ranging from 400 nm to 950 nm and spectra with SG smoothing, SNV, and MSC pre-treatment based on BOC, respectively. From Table 4.7, it can be concluded that model performance was better using the spectra pre-processing method. SG is the best pre-processing since previous studies revealed that spectral preprocessing has emerged as an effective technique to enhance spectral responsiveness. The SG algorithm is also widely used to smooth the spectrum and eliminate noise (Savitzky & Golay, 1964; Feng et al., 2023).

Table 4.7: Results of PLS model performance for prediction of chlorophyll

Pre-processing	Calibration		Prediction	
	R^2	RMSEC	R^2	RMSEP
Raw data	0.59	6.21	0.62	5.84
BOC	0.56	5.87	0.63	5.53
SG	0.99	0.57	0.99	0.76
SNV	0.57	0.56	0.77	0.51
BOC + SNV	0.57	0.56	0.76	0.52

© COPYRIGHT UPM

The correlation analysis between the chlorophyll value and the spectral data were analyzed. The R^2 value was increased after pre-processing data were applied. The correlation of chlorophyll value with spectra after BOC method was changed a little bit from the correlation using the raw spectrum. The correlations were lowest, with R^2 varying from 0.06 to 0.32 for calibration model, 0.12 to 0.32 for prediction model respectively. The effect of SG and SNV pre-processing method on the PLS model was almost the same with small difference of R^2 value. In the Vis region, the performance of SG pre-processing is slightly better than SNV, whereas in VSNIR region the opposite trend can be observed.

In terms of the overall pre-treatment, SG showed the best pre-processing effect with the highest R^2 compared to the other test spectra pre-processing methods for all FS. Therefore, the subsequent data analysis will be performed using spectra pre-processed with SG to maintain the uniformity of the data.

4.7 Prediction of Crop Parameters of Rubber Leaves Samples

The PLS regression model analyzes spectral data that relates to disease severity, directly interpreting changes in spectral reflectance to assess the level of severity (Hou et al., 2022). These properties among the first change when a tree is affected by WRD. By using PLS regression, these early symptoms of disease can be predicted accurately (Xuan et al., 2022).

PLS regression models were built and applied to calibration and prediction data sets. The models' performances were evaluated using the coefficient of determination (R^2),

© COPYRIGHT UPM

root mean square error of calibration (RMSEC), and root mean square error of prediction (RMSEP).

4.7.1 Prediction of SPAD Value

The spectral reflectance was pre-treated with SG pre-processing method and outlier removal were performed before construct PLS model. Figure 4.3 shows the prediction of SPAD value of rubber leaf, the calibration model gave R^2 value of 0.99 and RMSEP of 0.711, whereas the prediction model yielded R^2 value of 0.99 and RMSEP of 0.64. The prediction and calibration model accuracies obtained in this study were excellent fit between predicted value and the actual measured. RMSEP and RMSEC present the average difference between predicted chlorophyll content and actual measured.

The result indicates the SPAD value has a strong relationship to determine the healthy status of plants. Chlorophyll is the primary pigment in plant photosynthesis, reflecting the plant's photosynthesis ability, growth stage and nitrogen status (Liu et al., 2019). This finding aligns with a study reported by Jiang et al. (2022). This study focused on developing hyperspectral vegetation indices to estimate leaf chlorophyll content in mangrove forests experiencing pest and disease stress. The model demonstrated a strong correlation with chlorophyll content.

© COPYRIGHT UPM

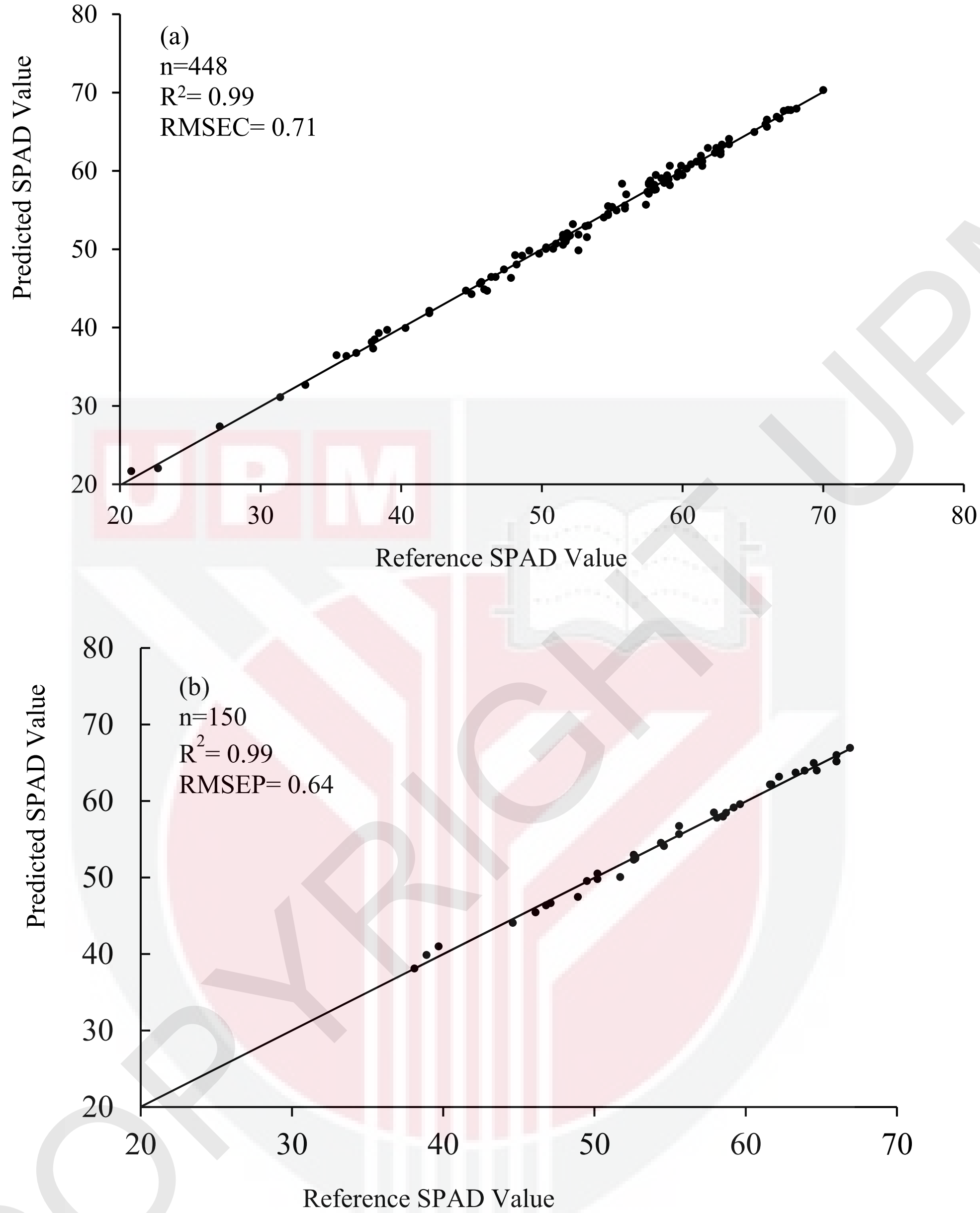


Figure 4.3: Scatter Plots of Reference Versus Predicted SPAD Value for Reflectance Spectral Data; (a) Calibration (b) Prediction Model

4.7.2 Prediction of Moisture Content

The MC of the leaf is an indicator of essential physiological processes such as transpiration and respiration of plants. Plant diseases such as fungal infection can lead

© COPYRIGHT UPM

to water loss from the affected leaves. The pathogen blocks water-conducting vessels. Figure 4.4 shows the MC prediction model of rubber leaf samples; the calibration gave an R^2 value of 0.61 and an RMSEC value of 3.34, whereas the prediction model showed an R^2 value of 0.65 and an RMSEP value of 2.00. The prediction and calibration model accuracies obtained in this study were moderate fit between the predicted value and the actual measured. RMSEP and RMSEC present the average difference between predicted chlorophyll content and actual measured. A lower value indicates better predictive accuracy.

This result is lower than the finding reported by Tang et al. (2023). This study explores the application of EI for predicting MC in tomato leaves. Another study was conducted by Nie et al. (2017) in the detection of MC in rapeseed leaves using terahertz spectroscopy. These results demonstrated excellent performance of R^2 above 0.80. Due to the presence of absorption peaks in IR spectrum at 970, 1200, 1440 and 1950 nm wavelength associated with water, the prediction of MC in VSNIR probability will lower (Raj et al., 2021).

© COPYRIGHT UPM

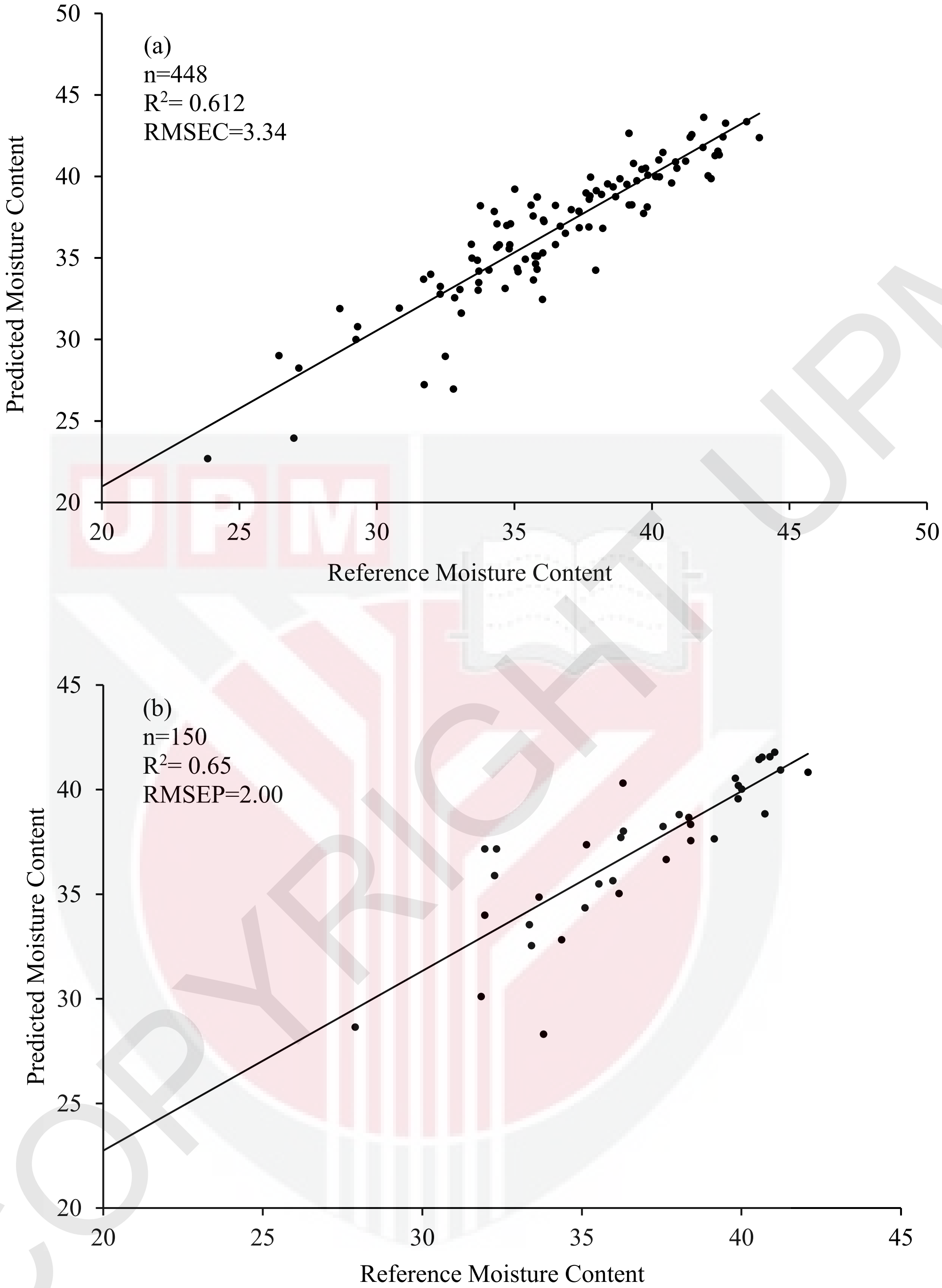


Figure 4.4: Scatter Plots of Reference Versus Predicted Moisture Content for Reflectance Spectral Data; (a) Calibration (b) Prediction Model

© COPYRIGHT UPM

4.7.3 The Prediction of Color

A color reader measures the color of an object or material based on human perception and gives a precise color value in a standardized color system (Munsell, 2022). While VISNIR provided information about the reflectance of each wavelength offering insight into the chemical composition, molecular structure and physical properties of the material (Pisello et al., 2022). In addition the output of color reader is often a single value or a few values that describe the perceived color, hue, brightness and saturation (Emery et al., 2021). While VISNIR gives information about the chemical and physical properties of the sample (Wang et al., 2021).

.The prediction of color in relation to plant disease is a common approach used in plant pathology to assess the impact of diseases on plant tissue. Color changes in leaves are an indicator of physiological response to disease. Figure 4.5 to Figure 4.7 shows the L*, a* and b* prediction model of rubber leaves samples. The calibration gave R^2 value of 0.88 and RMSEC value of 1.09, whereas the prediction model showed R^2 value of 0.95 and RMSEP value of 1.44 for the L*. The calibration gave R^2 value of 0.85 and RMSEC value of 1.16, whereas the prediction model showed R^2 value of 0.95 and RMSEP value of 0.78 for the a*. The calibration gave R^2 value of 0.89 and RMSEC value of 0.78, whereas the prediction model showed R^2 value of 0.94 and RMSEP value of 0.62 for the b*.

Leaf color and chlorophyll content play a critical role in plant growth and contribute greatly to the appearance of plants. The result yielded from this study demonstrated the excellent prediction of leaf color based on VSNIR with R^2 above 0.80. From this

© COPYRIGHT UPM

finding, it can be concluded that color is one of the parameters that can be used to detect the presence of WRD in leaves. This finding is parallel with the study reported by Tsai et al. (2021). This study explored leaf color as a morpho-physiological index for early screening of plant stress in rabbit-eye blueberries.

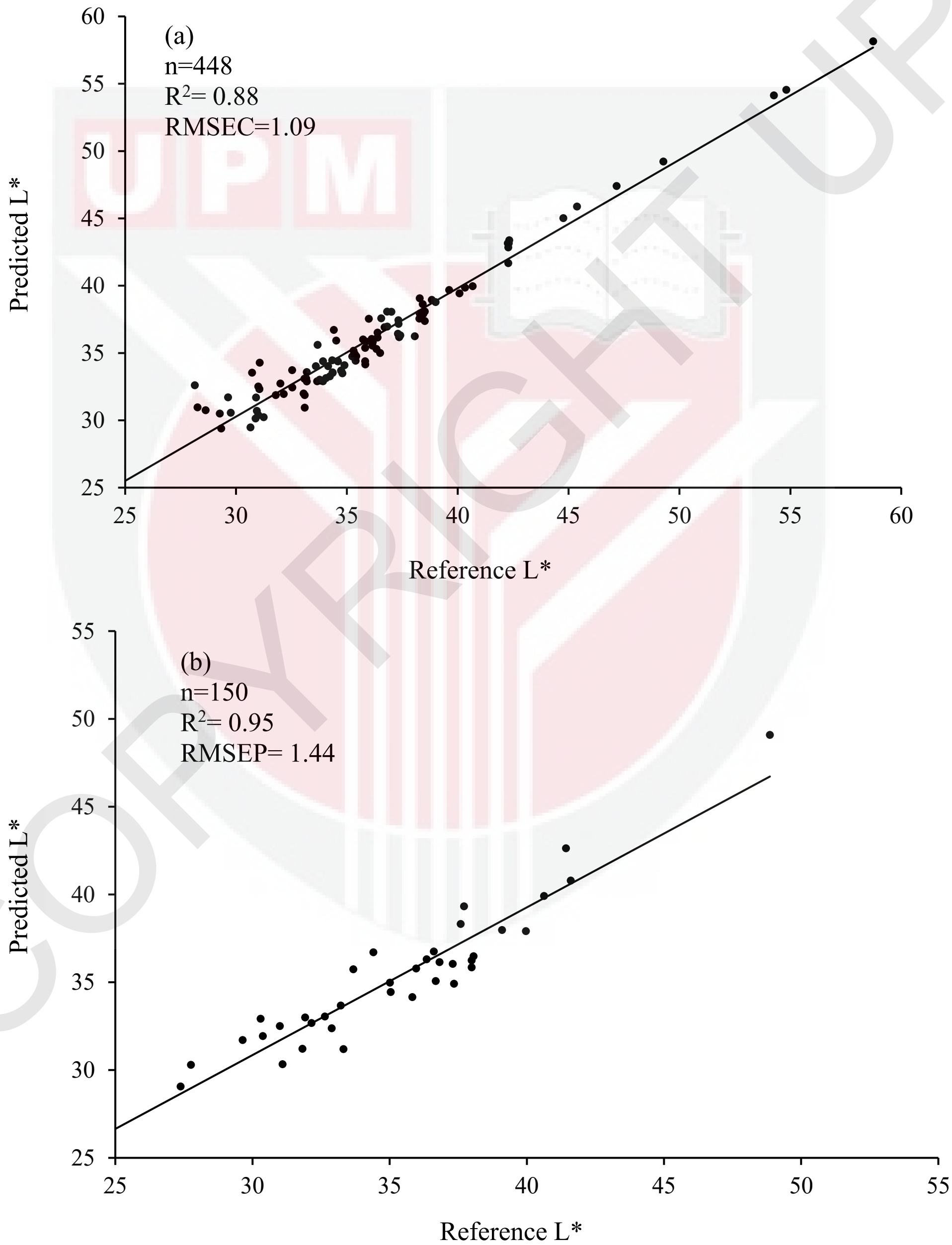


Figure 4.5: Scatter Plots of Reference Versus Predicted L* for Reflectance Spectral Data; (a) Calibration (b) Prediction Model

© COPYRIGHT UPM

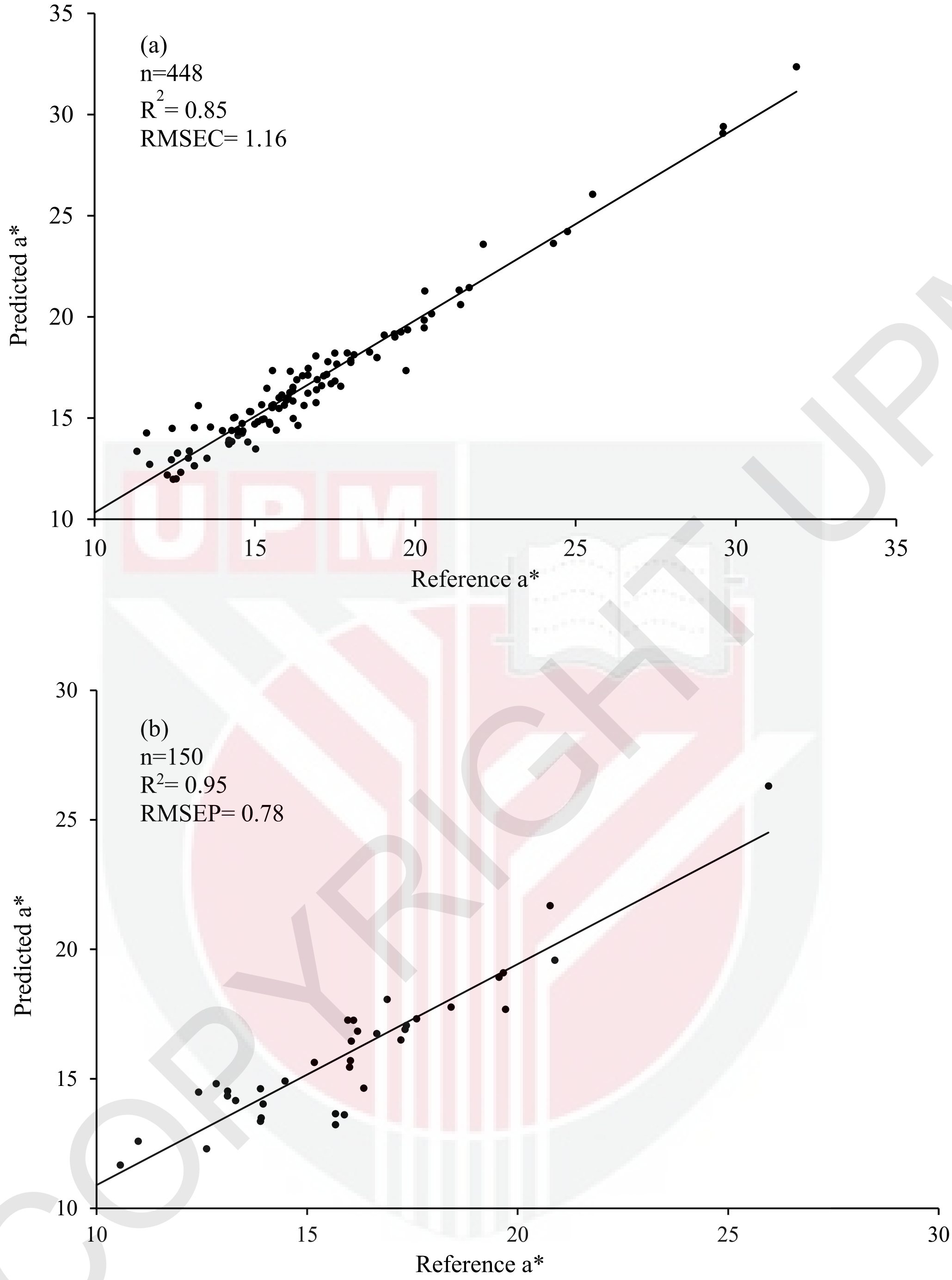


Figure 4.6: Scatter Plots of Reference Versus Predicted a* for Reflectance Spectral Data; (a) Calibration (b) Prediction model

© COPYRIGHT UPM

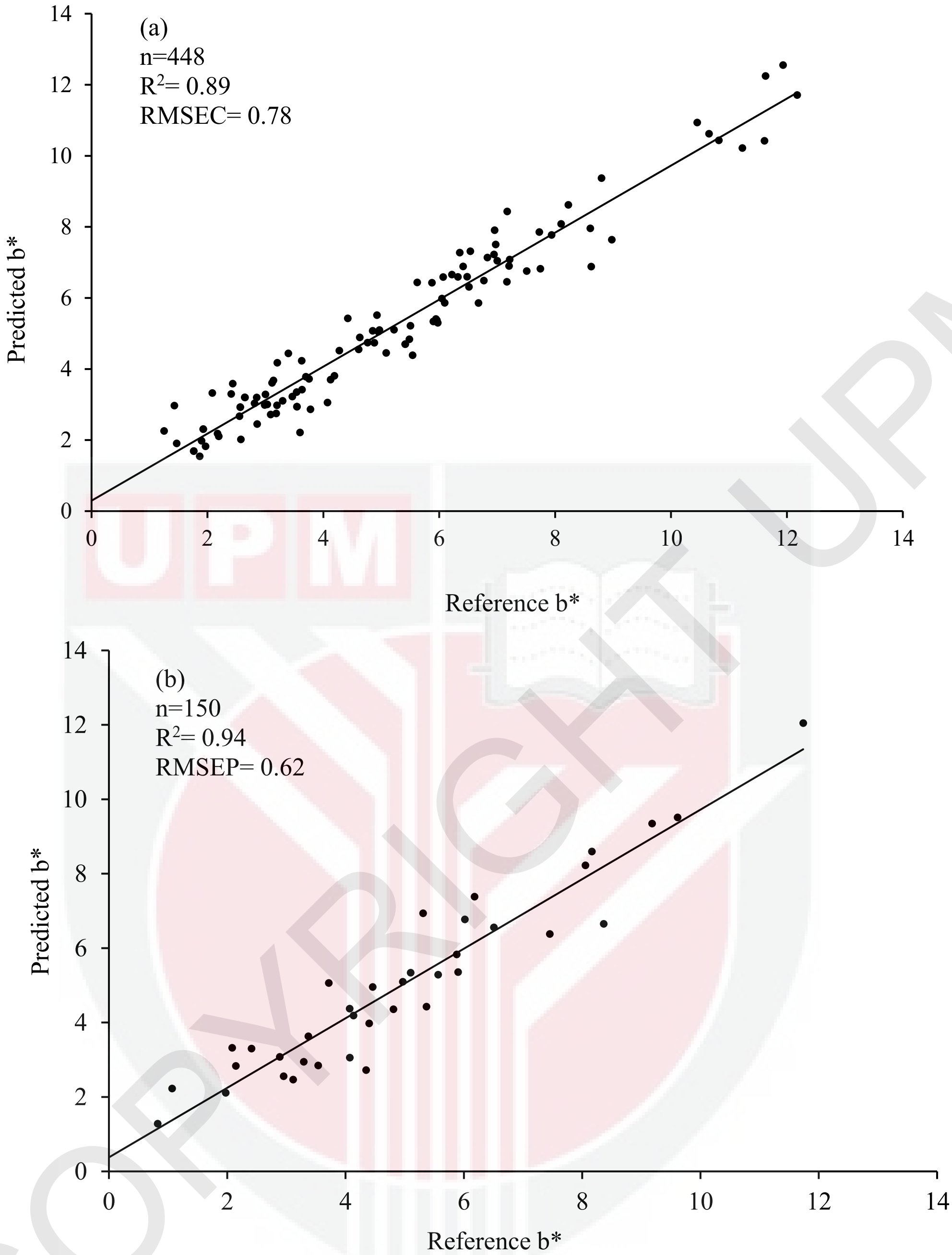


Figure 4.7: Scatter Plots Of Reference Versus Predicted b* For Reflectance Spectral Data; (a) Calibration (b) Prediction model

© COPYRIGHT UPM

In summary, this study demonstrated the ability of VSNIR spectroscopy to predict the chemical composition of rubber leaves. It was found that the SPAD value was the best crop parameter to be predicted by VSNIR from all WRD severity levels since rubber leaves achieved higher regression coefficient compared to other crop parameters. SPAD value mainly determines the photosynthetic rate and primary productivity in plants and is widely used as a response to the environment stress and nitrogen fertilizer application. Therefore, chlorophyll content could be used as an important diagnostic indicator for plant disease studies (Hotta et al., 1997).

4.8 Results of Data Reduction using Feature Selection

Most spectral reflectance is composed of noise, overlapping bands, and background, which makes obtaining the optimum wavelength for discriminating different levels of WRD in rubber trees a great challenge. Thus, feature selection techniques were used to find significant wavelengths to reduce dimensionality and optimize data processing time. Three feature selection algorithms (ANOVA, Chi-Square, and RELIEF) were applied to select the top significant wavelength, which could serve as input data for classification.

ANOVA is a statistical method that assesses the variance with different groups to identify features with significant differences. In the context of feature selection, ANOVA helps identify features that are relevant for classification. The Chi-square is used to assess the independence of categorical variables. in feature selection, it can identify features with a significant relationship with the target variables. RELIEF is a

© COPYRIGHT UPM

feature selection algorithm that evaluates the importance of features by considering the difference in feature values between instances of the same and different classes.

Table 4.8 summarizes the significant wavelength from the spectral reflectance data for the discrimination of different levels of WRD in rubber leaves. The spectral data was divided into three regions (visible, SWNIR, VSNIR). Different spectral regions interact differently with materials and substances (Stamford et al., 2024). The reflectance in the visible region is sensitive to the color of the vegetation (Zahir et al., 2024) while the SWNIR region is sensitive to water and very useful for detecting vegetation health and VSNIR region captures a broader spectrum which includes information about the color and the health of vegetation, water bodies and soil properties (Chowdhury et al. 2024).

The results showed that the significant wavelengths were different using different feature selection models. Given that the variation of various plant physiological parameters could induce strong responses in specific spectral ranges, the unbalance of the spectral features selected from the visible and VSNIR range could be attributed to the sensitivity of different physiological parameters to disease infection. The models of wrapper embedded and filter decided the best wavelengths based on the ranker searching method. Thus, these models generate the results in ranking order based on the predictive model score. Consequently, the selected wavelengths were affected by different working operations of different algorithm models applied.

© COPYRIGHT UPM

Table 4.8: The significant wavelength obtained from three feature selection models

Spectrum range	Wavelength (nm)	Significant wavelength (nm)		
		ANOVA	Chi-Square	RELIEF
Visible	450 - 699	663, 664,665,667, 668, 669	686, 655, 657, 659, 660, 685	532, 535,536 538,539,540
SWNIR	700 - 900	710, 712, 713, 714, 715, 716	700, 731, 733, 734, 735	700, 701, 702
VSNIR	450 – 900	664, 665, 667, 668, 669	686, 660, 654, 655, 667, 666,	450, 451,655, 653, 654, 650

Several wavelengths were repeatedly selected as top features by different feature selection methods in the VIS region (668 nm), SWNIR (700 nm), and VSNIR region (668, 655, 667 nm). Reflectance at 500 nm is dependent on the combined absorption of chlorophyll a, chlorophyll b, and carotenoids, while reflectance around 680 nm is solely controlled by chlorophyll (Datt, 1998). The concentration of chlorophyll and carotenoid contents in the leaves may be altered during the disease progression of WRD, as indicated by discoloration and wilting of the foliage. The changes of leaf pigment contents may be caused by the damage to the stem tissues and accumulation of toxic acids in plants during infection with *Rigidoporus micropores*.

From the result given from three feature selection methods, it can be concluded that the frequency of selected wavelengths in SWNIR (700 nm to 900 nm) is relatively lower than those selected in the visible region (450 nm to 699 nm). Therefore, the wavelength from the visible region is mostly highly significant in discriminating the WRD at different severity levels. The spectral selected wavelengths match well with Figure 4.2. This finding lines with a study conducted by Zhang et al. (2019) which confirms that the visible regions were most frequently used in the detection of plant diseases and pests. The sensitivity of spectral features in the visible region was linked

© COPYRIGHT UPM

to leaf internal structure and rapid degradations and biosynthesis of pigments, including chlorophyll, carotenoids, and anthocyanin (Choudhury et al., 2023). It is like a previous study regarding fungal infection in wheat leaf in the blue, green and red regions were also used for important feature discrimination of durum wheat leaf samples (Atanassova et al., 2023). Similarly, Zheng et al. (2018) reported that visible wavelengths are the optimal spectral for detecting yellow rust in wheat at different growths based on canopy spectral reflectance.

In the SWNIR range, no significant variation in reflectance was observed across the five severity levels, a trend that can be seen in various diseases. However, some studies have indicated that specific pathogens can alter plant cellular components, such as enzymes and proteins, leading to changes in the reflectance spectra within the SWNIR region (Wei et al., 2021). In this study, diseased leaves were collected from a field naturally infested with pathogens. Managing the pathogen within each leaf was difficult, resulting in localized disease spread. Enzymatic changes varied depending on the disease. Additionally, this study categorized leaves based on disease severity rather than specific disease types. Unlike foliar pathogens, which directly affect plant leaves, soilborne pathogens usually damage the roots or vascular systems first, only later manifesting as foliar symptoms. This distinction may explain why the wavelengths identified in this study for detecting stem rot differ from those reported in previous research on foliar diseases in various crops and trees (Gold et al., 2020; Conrad et al., 2020; Heim et al., 2018).

It can be concluded that the significant wavelengths identified by the RELIEF model encompass the entire range of wavelengths. In contrast, the wavelengths selected by

© COPYRIGHT UPM

the ANOVA and Chi-Square models were concentrated in the first half of the full spectrum. As a result, classification accuracies based on the wavelengths chosen by the RELIEF model are expected to outperform those based on the ANOVA and Chi-Square models. In summary, notable differences were observed in the wavelength ranges selected by the different feature selection models (ANOVA, Chi-Square, and RELIEF).

4.9 Different Feature Selection with Pre-processing

The correlation analysis between the chlorophyll value and the spectral data after BOC, SG, SNV, and SNV + BOC with different FS were analyzed (Table 4.9). The R^2 value was increased after pre-processing data were applied. The correlation of chlorophyll value with spectra after BOC method was changed a little bit from the correlation using the raw spectrum. The correlations were lowest, with R^2 varying from 0.06 to 0.32 for the calibration model, 0.12 to 0.32 for the prediction model respectively. The effect of SG and SNV pre-processing methods on the PLS model was almost the same with small difference of R^2 value. In the Vis region, the performance of SG pre-processing is slightly better than SNV, whereas in SWNIR region the opposite trend can be observed.

In terms of the overall pre-treatment, SG showed the best pre-processing effect with the highest R^2 compared to the other test spectra pre-processing methods for all FS. Therefore, the subsequent data analysis will be performed using spectra pre-processed with SG to maintain the uniformity of the data.

© COPYRIGHT UPM

Table 4.9: Result of PLS model performance for prediction of chlorophyll

Feature selection	Wavelength bands	Pre-processing	Calibration		Prediction	
			R^2	RMSEC	R^2	RMSEP
ANOVA	450-699	Raw	0.06	14.10	0.12	9.07
		BOC	0.21	13.54	0.22	9.69
		Savitzky_Golay	0.99	0.55	0.99	0.57
		SNV	0.99	0.01	0.99	0.01
		BOC + SNV	0.47	0.18	0.66	0.18
	700 - 900	Raw	0.22	12.85	0.29	8.4
		BOC	0.38	12.62	0.43	9.12
		Savitzky_Golay	0.99	0.83	0.99	0.89
		SNV	0.99	0.07	0.99	0.07
		BOC + SNV	0.99	0.07	0.99	0.07
	450- 900	Raw	0.20	13.11	0.29	8.42
		BOC	0.32	12.56	0.24	9.59
		Savitzky_Golay	0.99	0.71	0.99	0.64
		SNV	0.99	0.02	0.99	0.02
		BOC + SNV	0.99	0.02	0.99	0.02
Chi-Square	450-699	Raw	0.07	14.15	0.02	9.87
		BOC	0.28	12.99	0.17	10.06
		Savitzky_Golay	0.99	0.94	0.99	0.87
		SNV	0.79	0.01	0.80	0.02
		BOC + SNV	0.79	0.01	0.80	0.02
	700 - 900	Raw	0.07	14.15	0.02	9.87
		BOC	0.28	12.99	0.18	10.06
		Savitzky_Golay	0.99	0.94	0.99	0.87
		SNV	0.79	0.01	0.80	0.02
		BOC + SNV	0.79	0.01	0.80	0.02
	450- 900	Raw	0.06	14.15	0.03	9.86
		BOC	0.28	12.97	0.18	10.04
		Savitzky_Golay	0.99	0.95	0.99	0.87
		SNV	0.88	0.01	0.82	0.02
		BOC + SNV	0.79	0.01	0.80	0.02
RELIEF	450-699	Raw	0.05	14.27	0.14	9.24
		BOC	0.12	14.34	0.20	10.41
		Savitzky_Golay	0.99	0.96	0.99	0.79
		SNV	0.99	0.03	0.99	0.03
		BOC + SNV	0.99	0.03	0.99	0.03
	700-900	Raw	0.15	13.49	0.10	12.01
		BOC	0.26	13.84	0.38	9.45
		Savitzky_Golay	0.99	0.19	0.99	0.22
		SNV	0.99	0.01	0.99	0.01
		BOC + SNV	0.26	13.83	0.38	9.45
	450-900	Raw	0.20	13.09	0.27	8.55
		BOC	0.25	13.10	0.41	8.45
		Savitzky_Golay	0.99	0.54	0.99	0.69
		SNV	0.96	0.01	0.81	0.03
		BOC + SNV	0.96	0.01	0.81	0.03

© COPYRIGHT UPM

4.10 White Root Disease Classification using Different Classifiers with Feature Selection

Variable selection is a crucial step in spectral data analysis and model development, focusing on identifying and extracting key variables from large volumes of spectral data. In this section, the effect of applying different FS techniques (ANOVA, Chi-Square, and RELIEF) on different classifier models (SVM, ANN, kNN and RF) was evaluated (Figure 4.8). The performance evaluation based on accuracy, precision, recall and F1-score provides a comprehensive view of how well classification models are performing, especially after applying FS.

The results show that accuracy, precision, recall, and F1-score were higher when using data at the selected wavelength than when using data at the whole wavelength. This can be explained by the fact that FS techniques remove unneeded, irrelevant features from data that do not contribute to the accuracy of a predictive model (Omuya et al., 2021). Khan et al. (2020) stated that the FS technique improved classification accuracy.

© COPYRIGHT UPM

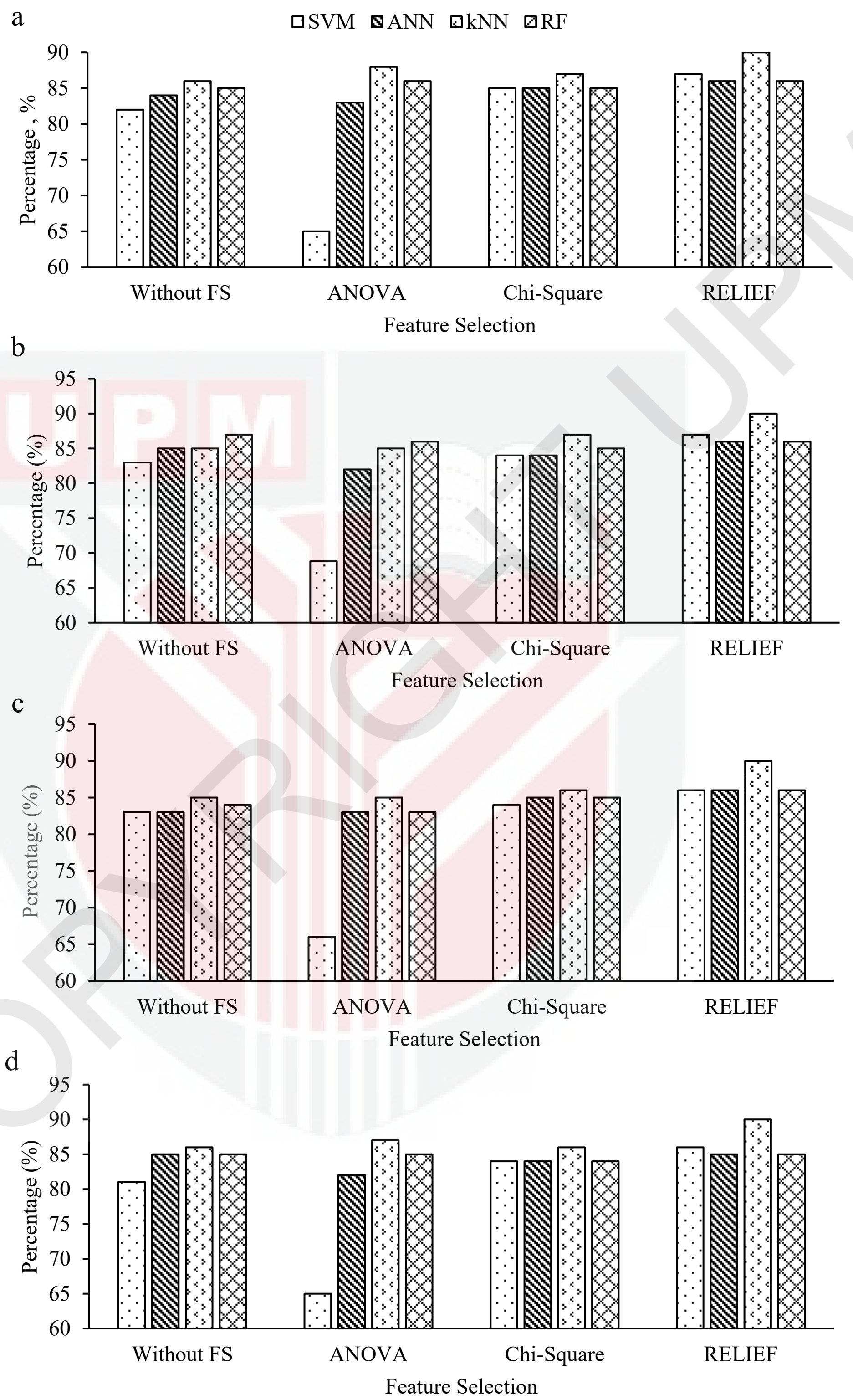


Figure 4.8: Performance Evaluation of Different Classification Techniques (a) Accuracy (b) Precision (c) F1-score (d) Recall

© COPYRIGHT UPM

The classification models using the full range of wavelength yielded the lowest accuracy with the SVM classifier, while other classifiers developed with selected spectral features performed significantly better. These findings are consistent with the reports by Fan et al. (2021) and Zhou et al. (2022). Notably, the performance of the same classifier varied depending on the specific spectral wavelength features chosen for analysis. Among the classifiers tested, kNN emerged as the top performer, delivering strong results with and without FS techniques. The classification accuracies for different feature selection methods were as follows: ANOVA (88%), Chi-Square (87%), and RELIEF (90%). These findings highlight that FS are highly effective for detecting WRD at varying severity levels.

The precision classification is a performance metric that measures the proportion of true positive instances predicted as positive. A high value means the model is good at identifying positive cases. Precision values range from 68% to 68% for ANOVA, indicating a moderate ability to identify positive cases across different feature subsets. Recall values range from 84% to 87% for Chi-Square, demonstrating an improvement in the classifier's ability to identify true positives when Chi-Square was applied. Recall values range from 86% to 90% for RELIEF, showing the highest performance, with RELIEF providing the best feature subset for improving precision.

Recall measures the ability of a classifier model to identify all relevant instances within a dataset. It focuses on the proportion of true positive cases that are correctly predicted by the model out of all actual positive cases. A high value means the model is good at identifying positive cases—the recall values for a particular classifier under different FS methods. Recall values range from 65% to 85% for ANOVA, indicating

© COPYRIGHT UPM

a moderate ability to identify positive cases across different feature subsets. Recall values range from 84% to 85% for Chi-Square, demonstrating an improvement in the classifier's ability to identify true positives when Chi-Square was applied. Recall values range from 85% to 90% for RELIEF, showing the highest performance, with RELIEF providing the best feature subset for improving recall.

Based on the results of this study, it can be concluded that RELIEF is the best feature selection method, as it yields the highest classification accuracy (90%) among the FS techniques. Additionally, kNN emerges as the best classifier overall, demonstrating superior performance both with and without feature selection. This combination of RELIEF and kNN provides an effective approach for accurately detecting WRD with varying severity levels.

4.11 Means Comparison Performance of Classification Algorithms

The mean classification accuracies of different classifiers (SVM, ANN, kNN and RF) were tested using ANOVA. The ANOVA depicted that there was a significant effect of classifiers on the model, indicating their significant roles in classification, and the mean comparison revealed that the performance of the classifiers was significantly different from each other (p-value = 0.024 which is less than 0.05 level of significance) (Table 4.10 and Table 4.11). These results align with findings from previous studies, reinforcing the importance of applying feature selection in plant disease classification. The outcomes highlight the effectiveness of the proposed FS technique, which could significantly aid in the classification of plant diseases. Furthermore, this study

© COPYRIGHT UPM

underscores the value of dimension reduction in achieving optimal grouping of spectral data within each disease severity level.

Table 4.10: ANOVA for classification accuracy at different feature selections

Source	Degree of Freedom	Sum of square	Mean Square	F-Value	Pr> F
Model	3	23.25	7.75	4.54	0.024
Error	12	20.50	1.71		
Corrected error	15	43.75			

Table 4.11: Means of classification accuracy for different classifiers

Classifier	Mean (%)
kNN	^{a}87.75
RF	^{b}85.5
ANN	^{c}84.50
SVM	^{d}79.75

4.12 Summary

The results of this study demonstrate that VSNIR spectroscopy can be effectively used for early detection of WRD. The spectral data revealed different patterns corresponding to different severity levels of WRD. Among the various parameters measured, chlorophyll emerged as the most reliable indicator for detecting the presence of WRD in rubber trees, showing a stronger correlation with the spectral data than color and MC. The RELIEF feature selection method proved to be the most effective for identifying the key wavelengths that distinguish between different disease severity levels. Additionally, the kNN classifier was found to be the most accurate for classifying WRD severity, enabling reliable early detection of the disease.

© COPYRIGHT UPM

CHAPTER 5

CONCLUSION AND RECOMMENDATION FOR FUTURE RESEARCH

Early detection and identification of WRD in rubber plantations are essential to minimize production losses. Current methods used to detect the presence of WRD from rubber trees are costly and labor-intensive. They can only be used under laboratory conditions, requiring complicated procedures, sample preparation, and a controlled environment. Thus, this study aims to develop a non-destructive technique to detect and identify WRD using VSNIR at an early stage.

To address these challenges, this study explores the use of VSNIR spectroscopy as a non-destructive approach for the early detection of WRD in rubber trees. VSNIR spectroscopy is a promising tool for collecting spectral data from plant leaves to detect disease-related changes in plant physiology. By leveraging the interaction of light with plant tissues, VSNIR spectroscopy can identify spectral signatures associated with the onset of WRD, which typically affects the roots and can indirectly affect the above-ground foliage.

In addition to spectral data collection, this study integrates computational intelligence techniques to optimize processing time and improve the efficiency of disease detection. Specifically, the research focuses on reducing data dimensionality and enhancing model generalization through data reduction methods, which help minimise overfitting. This is achieved using feature selection techniques such as ANOVA (Analysis of Variance), Chi-square, and RELIEF, which are employed to identify the

© COPYRIGHT UPM

most relevant spectral features for classifying different levels of WRD severity in rubber trees.

By comparing the performance of different machine learning classifiers (ANN, SVM, kNN, RF) —both with and without the application of data reduction models—the study aims to identify the most effective method for classifying the severity of WRD in rubber plantations. This approach not only aims to improve the accuracy of WRD detection at early stages but also reduces the computational burden, making the technique more practical for real-time field applications.

The outcomes of this study offer valuable insights into the application of VSNIR spectroscopy and computational intelligence in plant disease diagnostics. Specifically, it enhances the understanding of how spectral data from rubber leaves can be utilized to detect WRD and classify its severity in a non-destructive manner. This contribution extends existing knowledge on plant disease analysis through spectroscopy, with the potential to revolutionize the management of WRD in rubber plantations, leading to better early detection, faster decision-making, and, ultimately, reduced losses in rubber production. The major conclusions from this study are listed in the section below (Section 5.1).

5.1 Conclusion

All objectives of this study were achieved. The literature on WRD in rubber plantations has primarily focused on identifying infected versus non-infected trees, with limited research on methods for early detection. Early detection would involve identifying the presence of the pathogen or initial infection before visible symptoms

© COPYRIGHT UPM

appear. Meanwhile, the VSNIR spectroscopy technique to detect *Rigidoporus micropores* infection in rubber plantations has not yet been explored.

A total of 134 rubber trees were used in the study, representing five severity levels: healthy (S0), light (S1), moderate (S2), severe (S3), and very severe infection (S4). Samples of leaves, latex, and soil around rubber trees were collected from both healthy and affected trees. The result shows that the leaf sample from S4 had the lowest SPAD value of 47 and the lowest MC of 41.95 % w.b.. The result for the SPAD value indicates an inverse relationship between the chlorophyll content and MC with the severity level of WRD. The L* value increased from S0 (32.97 ± 3.37) to S4 (38.43 ± 4.97), while the a* value was decreased from S0 (-1.58 ± 1.92) to S4 (-2.16 ± 3.04) and the b* value was increased from S0 (12.64 ± 3.44) to S4 (17.79 ± 6.39). The result shows that the leaves are lighter and have a more yellowish appearance with increasing disease severity. For the soil samples, the average pH value remains relatively consistent across all severity levels, ranging from 6.20 (S0) and 6.40 (S4), while MCs decreased as the disease severity increased. The MCs value dropped significantly from S0 (33.50 % w.b.) to S4 (21.70% w.b.). For the latex samples, the TSC value was decreased from S0 (1.90%) to S3 (1.58%), this indicates that as the severity of WRD increases, the latex becomes more diluted. The chemical properties obtained from the samples at different severity levels significantly difference at ($p<0.05$) except for soil samples.

PLS regression models were built and applied to calibration and prediction data sets. From the result, chlorophyll content from the leaf sample had the highest R^2 with a calibration model value of 0.99 and a prediction model value of 0.99. For the MC

© COPYRIGHT UPM

prediction of rubber leaf samples, the calibration model gave an R^2 value of 0.61, whereas the prediction model gave an R^2 value of 0.65. For the color prediction of the rubber leaf samples, R^2 values for calibration models for L*, a*, and b* were 0.88, 0.85, and 0.89, respectively. For the prediction models, R^2 values for L*, a*, and b* were 0.95, 0.95, and 0.94, respectively. The result showed that the SPAD value of the leaf samples had a stronger relationship than other chemical properties. Based on this study, it can be concluded that the rubber leaf samples were the best in determining the healthy status of plants since the PLS regression model gave the highest value compared to other parameters.

The data reduction using FS (ANOVA, Chi-Square and RELIEF) were applied to identify the most significant wavelength in classifying different levels of WRD severity levels. The findings of this study indicated that there were no significant difference between FS techniques in classifying WRD disease with p-value > 0.024 at 0.05 significant level. The result revealed that RELIEF is the best feature selection method, as it yields the highest classification accuracy (90%) among the FS techniques. Additionally, kNN emerges as the best classifier overall, demonstrating superior performance both with and without feature selection. This combination of RELIEF and kNN provides an effective approach for accurately detecting WRD with varying severity levels.

5.2 Recommendation for Future Research

Based on the outcomes of this experiment, some recommendations for future research are proposed. In the future, The results found in this study indicated that the spectral

© COPYRIGHT UPM

data using VSNIR measurements is useful and demonstrated great potential to be used as alternative method to detect WRD at an early stage in rubber trees. This study found that computational intelligence; and data reduction algorithms; are quite important to be used for spectral data analysis to reduce the dimensionality and increase the classification accuracy. The findings from this study could be useful as inputs in the development of a sensing system for classifying WRD severity levels.

Based on the experience obtained from this work, the following further research is recommended to:

1. Extending the study to cover the stem of the rubber tree for VSNIR measurement. The stem is often the first part of the tree to show signs of stress or disease. Incorporating the stem into VSNIR analysis could allow for early detection of issues such as pest infestations, fungal infections, or drought stress, which may not be immediately visible in the leaves. Early detection can aid in timely interventions, potentially reducing losses and improving plantation management.
2. To deepen the understanding of plant physiological responses to WRD, future research should include an analysis of nitrogen content. As a vital nutrient, nitrogen is closely linked to plant health and is essential for the synthesis of chlorophyll.
3. Further research should focus on the development of spectral indices specifically designed to detect WRD-induced physiological changes. This could include exploring SWIR-based indices for detecting alterations in latex chemistry and red-edge indices for identifying chlorophyll degradation or leaf stress associated with WRD.

© COPYRIGHT UPM

REFERENCES

Abd Razak, S. B. (2017). Morphological features of Rigidoporus microporus isolated from infected Malaysian rubber clones. *Malaysian Journal of Microscopy*, *13*(1).

Azmi, A. F., Hashim, H., Sulaiman, M. S., Wahid, N. B., Khairuzzaman, N. A., & Mohd Sampian, A. F. (2016). Determination of White Root Infected Rubber Tree through Dielectric Characteristic in Latex. In *Advanced Materials Research*,1134, 231-235.

Azadnia, R., Rajabipour, A., Jamshidi, B., & Omid, M. (2023). New approach for rapid estimation of leaf nitrogen, phosphorus, and potassium contents in apple-trees using Vis/NIR spectroscopy based on wavelength selection coupled with machine learning. *Computers and Electronics in Agriculture*, *207*, 107746.

Ayesha, S., Hanif, M. K., & Talib, R. (2020). Overview and comparative study of dimensionality reduction techniques for high dimensional data. *Information Fusion*, *59*, 44-58.

Atanassova, S., Yorgov, D., Veleva, P., & Nedyalkova, S. (2023, December). Visible and near-infrared spectroscopy for detection of fungal diseases on Durum wheat. In *AIP Conference Proceedings,* 2889, (1).

Atanassova, S., Nikolov, P., Valchev, N., Masheva, S., & Yorgov, D. (2019, February). Early detection of powdery mildew (Podosphaera xanthii) on cucumber leaves based on visible and near-infrared spectroscopy. In *AIP conference proceedings*, 2075(1). AIP Publishing.

Ariff, E. A. R. E., Suratman, M. N., & Abdullah, S. (2011, September). Stomatal conductance, chlorophyll content, diameter and height in different growth stages of rubber tree (Hevea brasiliensis) saplings. In *2011 IEEE Symposium on Business, Engineering and Industrial Applications (ISBEIA)*, 84-88. IEEE.

Arabhosseini, A., Padhye, S., Huisman, W., van Boxtel, A., & Müller, J. (2011). Effect of drying on the color of tarragon (Artemisia dracunculus L.) leaves. *Food and Bioprocess Technology*, *4*(7), 1281-1287.

Andrew, B., Ahmad, K., Ismail, S. I., Ahmad, M. F., Ahmed, O. H., & Yun, W. M. (2021). Disease prevalence and molecular characterisation of Rigidoporus microporus associated with white root rot disease of rubber tree (Hevea brasiliensis) in Malaysia. *Journal of Rubber Research*, *24*, 175-186.

Althoff, D., & Rodrigues, L. N. (2021). Goodness-of-fit criteria for hydrological models: Model calibration and performance assessment. *Journal of Hydrology*, *600*, 126674.

© COPYRIGHT UPM

Ali, M. M., Bachik, N. A., Muhadi, N. A., Yusof, T. N. T., & Gomes, C. (2019). Non-destructive techniques of detecting plant diseases: A review. *Physiological and Molecular Plant Pathology*, *108*, 101426.

Albersheim, P., & Valent, B. S. (1974). Host-pathogen interactions: VII. Plant pathogens secrete proteins which inhibit enzymes of the host capable of attacking the pathogen. *Plant physiology*, *53*(5), 684-687.

Alander, J. T., Bochko, V., Martinkauppi, B., Saranwong, S., & Mantere, T. (2013). A Review of Optical Nondestructive Visual and Near-Infrared Methods for Food Quality and Safety. *International Journal of Spectroscopy*.

Akpaja, E. O., & Ogbebor, N. O. (2020). Field evaluation of Hevea brasiliensis clones for the incidence of white root disease in Nigeria. *Journal of Plantation Crops*, *49*(1), 1-6.

Akhter, S., Ahmed, S., Kabir, A., & Shoyaib, M. (2021). A relief based feature subset selection method. *Dhaka Univ. J. Appl. Sci. Eng.*, *6*(2), 7-13.

Ahmad, N., Asif, H. M. S., Saleem, G., Younus, M. U., Anwar, S., & Anjum, M. R. (2021). Leaf image-based plant disease identification using color and texture features. *Wireless Personal Communications*, *121*(2), 1139-1168.

Agyeman, P. C., Khosravi, V., Kebonye, N. M., John, K., Borůvka, L., & Vašát, R. (2022). Using spectral indices and terrain attribute datasets and their combination in the prediction of cadmium content in agricultural soil. *Computers and Electronics in Agriculture*, *198*, 107077.

Agustika, D. K., Mercuriani, I., Purnomo, C. W., Hartono, S., Triyana, K., Iliescu, D. D., & Leeson, M. S. (2022). Fourier transform infrared spectrum pre-processing technique selection for detecting PYLCV-infected chilli plants. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *278*, 121339.

Abdullah, N. E., Rahim, A. A., Hashim, H., & Kamal, M. M. (2007). Classification of rubber tree leaf diseases using multilayer perceptron neural network. In *2007, 5th Student Conference on Research and Development*, IEEE, 1-6.

Abdolmaleki, A., B Ghasemi, J., Shiri, F., & Pirhadi, S. (2015). Application of multivariate linear and nonlinear calibration and classification methods in drug design. *Combinatorial chemistry & high throughput screening*, *18*(8), 795-808.

Ahmed, H. G. M. D., Zeng, Y., Yang, X., Anwaar, H. A., Mansha, M. Z., Hanif, C. M. S., Ikram, K., Ullah, A., & Alghanem, S. M. S. (2020). Conferring drought-tolerant wheat genotypes through morpho-physiological and chlorophyll indices at seedling stage. *Saudi Journal of Biological Sciences*, *27*(8), 2116–2123.

Ajeng, A. A., Abdullah, R., Malek, M. A., Chew, K. W., Ho, Y. C., Ling, T. C., Lau, B. F., & Show, P. L. (2020). The effects of biofertilizers on growth, soil

© COPYRIGHT UPM

fertility, and nutrients uptake of oil palm (Elaeis guineensis) under greenhouse conditions. *Processes*, *8*(12), 1–16.

Anjum, A., Jaggi, S., Varghese, E., Lall, S., Bhowmik, A., & Rai, A. (2016). Identification of Differentially Expressed Genes in RNA-seq Data of Arabidopsis thaliana: A Compound Distribution Approach. *Journal of Computational Biology*, *23*(4), 239–247.

Anli, M., Baslam, M., Tahiri, A., Raklami, A., Symanczik, S., Boutasknit, A., Ait-El-Mokhtar, M., Ben-Laouane, R., Toubali, S., Ait Rahou, Y., Ait Chitt, M., Oufdou, K., Mitsui, T., Hafidi, M., & Meddich, A. (2020). Biofertilizers as Strategies to Improve Photosynthetic Apparatus, Growth, and Drought Stress Tolerance in the Date Palm. *Frontiers in Plant Science*, *11*.

Azzeme, A. M., Abdullah, S. N. A., Aziz, M. A., & Wahab, P. E. M. (2016). Oil palm leaves and roots differ in physiological response, antioxidant enzyme activities and expression of stress-responsive genes upon exposure to drought stress. *Acta Physiologiae Plantarum*, *38*(2), 1–12.

Ba, W., Jin, X., Lu, J., Rao, Y., Zhang, T., Zhang, X., Zhou, J. & Li, S. (2023). Research on predicting early Fusarium head blight with asymptomatic wheat grains by micro-near infrared spectrometer. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *287*, 122047.

Babellahi, F. (2020). The potential use of non destructive optical-based techniques for early detection of chilling injury and freshness in horticultural commodities.

Balasubramaniam, S., Joe, C. V., Manthiramoorthy, C., & Kumar, K. S. (2024). ReliefF based feature selection and Gradient Squirrel search Algorithm enabled Deep Maxout Network for detection of heart disease. *Biomedical Signal Processing and Control*, *87*, 105446.

Barnes, R. J., Dhanoa, M. S., & Lister, S. J. (1989). Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Applied spectroscopy*, *43*(5), 772-777.

Basal, O., Szabó, A., & Veres, S. (2020). Physiology of soybean as affected by PEG-induced drought stress. *Current Plant Biology*, *22*.

Bashir, S., Khattak, I. U., Khan, A., Khan, F. H., Gani, A., & Shiraz, M. (2022). A novel feature selection method for classification of medical data using filters, wrappers, and embedded approaches. *Complexity*, *2022*(1), 8190814.

Basyuni, M., Amri, N., Putri, L. A. P., Syahputra, I., & Arifiyanto, D. (2017). Characteristics of Fresh Fruit Bunch Yield and the Physicochemical Qualities of Palm Oil during Storage in North Sumatra, Indonesia. *Indonesian Journal of Chemistry*, *17*(2), 182.

Bayona-Rodriguez, C. J., & Romero, H. M. (2019). Physiological and agronomic behavior of commercial cultivars of oil palm (Elaeis guineensis) and OxG hybrids (Elaeis oleifera x Elaeis guineensis) at rainy and dry seasons.

© COPYRIGHT UPM

Australian Journal of Crop Science, *13*(3), 424–432.

Bhaskarla, V., Zinta, G., Ford, R., Jain, M., Varshney, R. K., & Mantri, N. (2020). Comparative root transcriptomics provide insights into drought adaptation strategies in chickpea (Cicer arietinum L.). *International Journal of Molecular Sciences*, *21*(5).

Bi, J., Hou, D., Zhang, X., Tan, J., Bi, Q., Zhang, K., Liu, Y., Wang, F., Zhang, A., Chen, L., liu, G., liu, Z., Yu, X., & Luo, L. (2021). A novel water-saving and drought-resistance rice variety promotes phosphorus absorption through root secreting organic acid compounds to stabilize yield under water-saving condition. *Journal of Cleaner Production*, *315*, 127992.

Bogati, K., & Walczak, M. (2022). The Impact of Drought Stress on Soil Microbial Community, Enzyme Activities and Plants. In *Agronomy* 12(1). MDPI.

Bolan, N., Sarmah, A. K., Bordoloi, S., Bolan, S., Padhye, L. P., Van Zwieten, L. Sooriyakumar, P., Khan, B., Ahmad, M., Solaiman, Z., Rinklebe, J., Wang, H., Singh, B., & Siddique, K. H. (2023). Soil acidification and the liming potential of biochar. *Environmental Pollution*, *317*, 120632.

Bommert, A., Sun, X., Bischl, B., Rahnenführer, J., & Lang, M. (2020). Benchmark for filter methods for feature selection in high-dimensional classification data. *Computational Statistics & Data Analysis*, *143*, 106839.

Boukhlifa, G., & Chibani, Y. (2024, April). A Lightweight CNN Design Based on Convolutional Autoencoder for Tomato Disease Identification. In *2024 8th International Conference on Image and Signal Processing and their Applications (ISPA)* (pp. 1-8). IEEE.

Brum, M., Oliveira, R. S., López, J. G., Licata, J., Pypker, T., Chia, G. S., Tinôco, R. S., & Asbjornsen, H. (2021a). Effects of irrigation on oil palm transpiration during ENSO-induced drought in the Brazilian Eastern Amazon. *Agricultural Water Management*, *245*.

Brum, M., Oliveira, R. S., López, J. G., Licata, J., Pypker, T., Chia, G. S., Tinôco, R. S., & Asbjornsen, H. (2021b). Effects of irrigation on oil palm transpiration during ENSO-induced drought in the Brazilian Eastern Amazon. *Agricultural Water Management*, *245*.

Brum, M., Oliveira, R. S., López, J. G., Licata, J., Pypker, T., Chia, G. S., Tinôco, R. S., & Asbjornsen, H. (2021c). Effects of irrigation on oil palm transpiration during ENSO-induced drought in the Brazilian Eastern Amazon. *Agricultural Water Management*, *245*.

Buja, I., Sabella, E., Monteduro, A. G., Chiriacò, M. S., De Bellis, L., Luvisi, A., & Maruccio, G. (2021). Advances in plant disease detection and monitoring: From traditional assays to in-field diagnostics. *Sensors*, *21*(6), 2129.

Buvé, C., Saeys, W., Rasmussen, M. A., Neckebroeck, B., Hendrickx, M., Grauwet, T., & Van Loey, A. (2022). Application of multivariate data analysis for food quality investigations: An example-based review. *Food Research International*, *151*, 110878.

Buxton, D. R., & Casler, M. D. (1993). Environmental and genetic effects on cell wall composition and digestibility. *Forage cell wall structure and digestibility*, 685-714.

Cai, F., Zhang, Y., Mi, N., Ming, H., Zhang, S., Zhang, H., & Zhao, X. (2020). Maize (Zea mays L.) physiological responses to drought and rewatering, and the associations with water stress degree. *Agricultural Water Management*, *241*.

Cao, H. X., Sun, C. X., Shao, H. B., & Lei, X. T. (2011). Effects of low temperature and drought on the physiological and growth changes in oil palm seedlings. *African Journal of Biotechnology*, *10*(14), 2630–2637.

Cao, J. (2016). Analysis of the Prefoldin gene family in 14 plant species. *Frontiers in Plant Science*, *7*(MAR2016).

Carles, C. C., & Fletcher, J. C. (2009). The SAND domain protein Ultrapetala1 acts as a trithorax group factor to regulate cell fate in plants. *Genes and Development*, *23*(23), 2723–2728.

Cebekhulu, E., Onumanyi, A. J., & Isaac, S. J. (2022). Performance analysis of machine learning algorithms for energy demand–supply prediction in smart grids. *Sustainability*, *14*(5), 2546.

Cerda, R., Avelino, J., Gary, C., Tixier, P., Lechevallier, E., & Allinne, C. (2017). Primary and secondary yield losses caused by pests and diseases: Assessment and modeling in coffee. *PloS one*, *12*(1).

Chaiharn, C., Sujada, N., Pathom-aree, W., & Lumyong, S. (2019). Biological control of Rigidoporus microporus the casue of white root disease in rubber using PGPRs in vivo. *Chiang Mai J. Sci*, *46*, 850-866.

Chaiharn, C., Sujada, N., Pathom-aree, W., & Lumyong, S. (2019). Biological control of Rigidoporus microporus the casue of white root disease in rubber using PGPRs in vivo. *Chiang Mai J. Sci*, *46*, 850-866.

Chan, P. L., Rose, R. J., Abdul Murad, A. M., Zainal, Z., Leslie Low, E. T., Ooi, L. C. L., Ooi, S. E., Yahya, S., & Singh, R. (2014). Evaluation of reference genes for quantitative real-time PCR in oil palm elite planting materials propagated by tissue culture. *PLoS ONE*, *9*(6).

Chandhini, S., & Rejish Kumar, V. J. (2019). Transcriptomics in aquaculture: current status and applications. In *Reviews in Aquaculture* 11(4), 1379–1397.

Chauhan, N., Singh, C., Chauhan, D., Chaudhary, J., & Rawat, P. (2024). Modern techniques and developments in the detection of foodborne pathogens. In *Biosensors for Foodborne Pathogens Detection*, 17-55. Academic Press.

Choi, J., Baldwin, T. M., Wong, M., Bolden, J. E., Fairfax, K. A., Lucas, E. C., Cole,

© COPYRIGHT UPM

R., Biben, C., Morgan, C., Ramsay, K. A., Ng, A. P., Kauppi, M., Corcoran, L. M., Shi, W., Wilson, N., Wilson, M. J., Alexander, W. S., Hilton, D. J., & de Graaf, C. A. (2019). Haemopedia RNA-seq: A database of gene expression during haematopoiesis in mice and humans. *Nucleic Acids Research*, *47*(D1), D780–D785.

Choudhury, B. U., Narzari, R., Zafar, M., Singh, N., Mishra, V. K., & Prabhakar, M. (2023). Spectral library of crops and discrimination of major vegetables grown in the eastern Himalayan ecosystem: A proximal hyperspectral remote sensing approach. *Ecological Informatics*, *77*, 102263.

Choudhury, D., Ponneganti, S., Radhakrishnanand, P., Murty, U. S., & Banerjee, S. (2023). Selective laser sintering additive manufacturing of solid oral dosage form: Effect of laser power and hatch spacing on the physico-technical behaviour of sintered printlets. *Applied Materials Today*, *35*, 101943.

Chowdhury, M., Anand, R., Dhar, T., Kurmi, R., Sahni, R. K., & Kushwah, A. (2024). Digital insights into plant health: exploring vegetation indices through computer vision. In *Applications of Computer Vision and Drone Technology in Agriculture 4.0* (pp. 7-30). Singapore: Springer Nature Singapore.

Chuanlei, Z., Shanwen, Z., Jucheng, Y., Yancui, S., & Jia, C. (2017). Apple leaf disease identification using genetic algorithm and correlation based feature selection method. *International Journal of Agricultural and Biological Engineering*, *10*(2), 74-83.

Conrad, A. O., Li, W., Lee, D. Y., Wang, G. L., Rodriguez-Saona, L., & Bonello, P. (2020). Machine learning-based presymptomatic detection of rice sheath blight using spectral profiles. *Plant Phenomics*.

Cottage, A., Edwards, Y. J. K., & Elgar, G. (2001). SAND, a new protein family: From nucleic acid to protein structure and function prediction. *Comparative and Functional Genomics*, *2*(4), 226–235.

Croft, H., & Chen, J. M. (2017). Leaf pigment content. *Reference Module in Earth Systems and Environmental Sciences. Oxford: Elsevier Inc*, 1-22.

Croft, H., Chen, J. M., Luo, X., Bartlett, P., Chen, B., & Staebler, R. M. (2017). Leaf chlorophyll content as a proxy for leaf photosynthetic capacity. *Global change biology*, *23*(9), 3513-3524.

Cuff, J. A., Clamp, M. E., Siddiqui, A. S., Finlay, M., & Barton, G. J. (1996). *JPred: a consensus secondary structure prediction server*. Livingston and Barton.

Culman, M., de Farias, C. M., Bayona, C., & Cabrera Cruz, J. D. (2019). Using agrometeorological data to assist irrigation management in oil palm crops: A decision support method and results from crop model simulation. *Agricultural Water Management*, *213*, 1047–1062.

© COPYRIGHT UPM

Cunha, C. L., Torres, A. R., & Luna, A. S. (2020). Multivariate regression models obtained from near-infrared spectroscopy data for prediction of the physical properties of biodiesel and its blends. *Fuel*, *261*, 116344.

Cunningham, P., & Delany, S. J. (2021). K-nearest neighbour classifiers-a tutorial. *ACM computing surveys (CSUR)*, *54*(6), 1-25.

Dai, W., Girdthai, T., Huang, Z., Ketudat-Cairns, M., Tang, R., & Wang, S. (2016). Genetic analysis for anthocyanin and chlorophyll contents in rapeseed. *Ciência Rural*, *46*(5), 790-795.

Dalimunthe, C. I., Tistama, R., & Wahyuni, S. (2017, December). Detection of white root disease (Rigidoporus microporus) in various soil types in the rubber plantations based on the serological reaction. In *IOP Conference Series: Earth and Environmental Science* 97(1), 012043, IOP Publishing.

Damiri, N., Pratama, Y., Febbiyanti, T. R., Rahim, S. E., Astuti, D. T., & Purwanti, Y. (2022). Pestalotiopsis sp. infection causes leaf fall disease of new arrivals in several clones of rubber plants. *Biodiversitas Journal of Biological Diversity*, *23*(8).

Das, A. K., Kumar, S., Jain, S., Goswami, S., Chakrabarti, A., & Chakraborty, B. (2020). An information-theoretic graph-based approach for feature selection. *Sādhanā*, *45*, 1-9.

Das, B., Sahoo, R. N., Pargal, S., Krishna, G., Verma, R., Viswanathan, C., & Gupta, V. K. (2021). Evaluation of different water absorption bands, indices and multivariate models for water-deficit stress monitoring in rice using visible-near infrared spectroscopy. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *247*, 119104.

Das, S., Pattanayak, S., & Behera, P. R. (2022). Application of machine learning: a recent advancement in plant diseases detection. *Journal of Plant Protection Research*, 122-135.

Datt, B. (1998). Remote sensing of chlorophyll a, chlorophyll b, chlorophyll a+ b, and total carotenoid content in eucalyptus leaves. *Remote Sensing of Environment*, *66*(2), 111-121.

Davidson, N. M., & Oshlack, A. (2014). Corset: Enabling differential gene expression analysis for de novo assembled transcriptomes. *Genome Biology*, *15*(7).

de Andrade, S. A. L., Domingues Jr, A. P., & Mazzafera, P. (2015). Photosynthesis is induced in rice plants that associate with arbuscular mycorrhizal fungi and are grown under arsenate and arsenite stress. *Chemosphere*, *134*, 141-149.

Dede, A. P. O., Akpaja, E. O., & Galillee, J. E. (2011). Effect of pH on the growth of the white root rot pathogen, Rigidoporus lignosus (Klotzsch) Imazeki, on selected para rubber sustaining soils in Nigeria. *Africa Science*, *12*, 175-179.

Del, E. M., Mahlein, A. K., & Bock, C. H. (2024). Plant disease quantification.

© COPYRIGHT UPM

In *Agrios' Plant Pathology*, 211-225. Academic Press.

Deng, J., Pan, S., Zhou, M., Gao, W., Yan, Y., Niu, Z., & Han, W. (2023). Optimum sampling window size and vegetation index selection for low-altitude multispectral estimation of root soil moisture content for Xuxiang Kiwifruit. *Agricultural Water Management*, *282*, 108297.

Devaraj, A., Rathan, K., Jaahnavi, S., & Indira, K. (2019, April). Identification of plant disease using image processing technique. In *2019 International Conference on Communication and Signal Processing (ICCSP)* ,749-0753. IEEE.

Dey, A. K., Sharma, M., & Meshram, M. R. (2016). Image processing based leaf rot disease, detection of betel vine (Piper BetleL.). *Procedia Computer Science*, *85*, 748-754.

Dhaliwal, S. S., Sharma, V., Shivay, Y. S., Gupta, R. K., Verma, V., Kaur, M., ... & Hossain, A. (2024). Assessment and detection of biotic and abiotic stresses in field crops through remote and proximal sensing techniques—evidence from earlier findings. *Arabian Journal of Geosciences*, *17*(6), 188.

Dhanoa, M. S., López, S., Sanderson, R., Lister, S. J., Barnes, R. J., Ellis, J. L., & France, J. (2023). Methodology adjusting for least squares regression slope in the application of multiplicative scatter correction to near-infrared spectra of forage feed samples. *Journal of Chemometrics*, *37*(11).

Diago, M. P., Correa, C., Millán, B., Barreiro, P., Valero, C., & Tardaguila, J. (2012). Grapevine yield and leaf area estimation using supervised classification methodology on RGB images taken under field conditions. *Sensors*, *12*(12), 16988-17006.

Dissanayake, K., & Md Johar, M. G. (2021). Comparative study on heart disease prediction using feature selection techniques on classification algorithms. *Applied Computational Intelligence and Soft Computing*, *2021*(1), 5581806.

Du, X. L., Li, X. Y., Liu, Y., Zhou, W. H., & Li, J. L. (2019). Genetic algorithm optimized non-destructive prediction on property of mechanically injured peaches during postharvest storage by portable visible/shortwave near-infrared spectroscopy. *Scientia Horticulturae*, *249*, 240-249.

Dubey, S. R., & Jalal, A. S. (2016). Apple disease classification using color, texture and shape features from images. *Signal, Image and Video Processing*, *10*(5), 819-826.

Dussert, S., Guerin, C., Andersson, M., Joet, T., Tranbarger, T. J., Pizot, M., Sarah, G., Omore, A., Durand-Gasselin, T., & Morcillo, F. (2013). Comparative Transcriptome Analysis of Three Oil Palm Fruit and Seed Tissues That Differ in Oil Content and Fatty Acid Composition. *Plant Physiology*, *162*(3), 1337–1358.

Elshawi, R., Al-Mallah, M. H., & Sakr, S. (2019). On the interpretability of machine

© COPYRIGHT UPM

learning-based model for predicting hypertension. *BMC medical informatics and decision making*, *19*, 1-32.

Emery, K. J., Parthasarathy, M. K., Joyce, D. S., & Webster, M. A. (2021). Color perception and compensation in color deficiencies assessed with hue scaling. *Vision research*, *183*, 1-15.

Enciso, M., Carrascosa, J. P., Sarasa, J., Martínez-Ortiz, P. A., Munné, S., Horcajadas, J. A., & Aizpurua, J. (2018). Development of a new comprehensive and reliable endometrial receptivity map (ER Map/ER Grade) based on RT-qPCR gene expression analysis. *Human Reproduction*, *33*(2), 220–228.

Falqueto, A. R., da Silva Júnior, R. A., Gomes, M. T. G., Martins, J. P. R., Silva, D. M., & Partelli, F. L. (2017). Effects of drought stress on chlorophyll a fluorescence in two rubber tree clones. *Scientia Horticulturae*, *224*, 238-243.

Fan, H., Xiao, Y., Yang, Y., Xia, W., Mason, A. S., Xia, Z., Qiao, F., Zhao, S., & Tang, H. (2013). RNA-Seq Analysis of Cocos nucifera: Transcriptome Sequencing and De Novo Assembly for Subsequent Functional Genomics Approaches. *PLoS ONE*, *8*(3).

Fan, F., Changwei, Z., Xiaojun, Z., Di, W., Zhi, T., & Yishen, X. (2021, October). Feature Wavelength Selection in Near-Infrared Spectroscopy Based on Genetic Algorithm. In *2021 International Conference on Sensing, Measurement & Data Analytics in the era of Artificial Intelligence (ICSMD)*,1-5. IEEE.

Faraldi, M., Mangiavini, L., Conte, C., Banfi, G., Napoli, N., & Lombardi, G. (2022). A novel methodological approach to simultaneously extract high-quality total RNA and proteins from cortical and trabecular bone. *Open Biology*, *12*(5).

Farid, A. M., Terhem, R., Aswad, R. M., Agustini, L., Ho, W. M., Indrayadi, H., & Hidayati, N. (2023). Diseases of Acacia and control measures in the tropics. In *Forest microbiology*, 375-400. Academic Press.

Feng, Z., Ding, X., Zhang, H., He, L., Duan, J., Ma, X., Zhao, C., Yang, G., & Feng, W. (2023). Spectroscopic detection of wheat yellow mosaic virus infection based on invariant shape spectral processing and machine learning. *Ecological Indicators*, *154*, 110750.

Fernando, T. H. P. S., Jayasinghe, C. K., & Siriwardene, D. (2013). Development of an early detection method for white root disease of rubber: a preliminary investigation.

Fernando, T. H. P. S., Jayasinghe, C. K., Wijesundera, R. L. C., & Siriwardene, D. (2012). Development of an early detection method for white root disease caused by Rigidoporus microporus. In *Proceedings of International Forestry and Environment Symposium*,17.

Fincheira, P., Jofré, I., Espinoza, J., Levío-Raimán, M., Tortella, G., Oliveira, H. C., Diez, M. C., Quiroz, A., & Rubilar, O. (2023). The efficient activity of plant essential oils for inhibiting Botrytis cinerea and Penicillium expansum: Mechanistic insights into antifungal activity. *Microbiological Research*, *277*, 127486.

Foley, W. J., McIlwee, A., Lawler, I., Aragones, L., Woolnough, A. P., & Berding, N. (1998). Ecological applications of near infrared reflectance spectroscopy–a tool for rapid, cost-effective prediction of the composition of plant and animal tissues and aspects of animal performance. *Oecologia*, *116*(3), 293-305.

Freitas, E. L., Brito, A. C., de Oliveira, S. A. S., & de Oliveira, E. J. (2020). Early diagnosis of cassava frog skin disease in powdered tissue samples using near-infrared spectroscopy. *European Journal of Plant Pathology*, *156*, 547-558.

Gafur, M. Abd., & Putra, E. T. S. (2019). Effect of Drought Stress in Physiological Oil Palm Seedling (Elaeis guineensis Jacq.) using Calcium Application. *Asian Journal of Biological Sciences*, *12*(3), 550–556.

Gajbhiye, P., Mingchinda, N., Chen, W., Mukhopadhyay, S. C., Wilaiprasitporn, T., & Tripathy, R. K. (2020). Wavelet domain optimized Savitzky–Golay filter for the removal of motion artifacts from EEG recordings. *IEEE Transactions on Instrumentation and Measurement*, *70*, 1-11.

Gambino, G., Perrone, I., & Gribaudo, I. (2008). A rapid and effective method for RNA extraction from different tissues of grapevine and other woody plants. *Phytochemical Analysis*, *19*(6), 520–525.

Garcia, F., Sankaran, S., Maja, J. M., Lee, W. S., Rasmussen, J., & Ehsani, R. (2013). Comparison of two aerial imaging platforms for identification of Huanglongbing-infected citrus trees. *Computers and Electronics in Agriculture*, *91*, 106-115.

Gazis, R., & Chaverri, P. (2010). Diversity of fungal endophytes in leaves and stems of wild rubber trees (Hevea brasiliensis) in Peru. *fungal ecology*, *3*(3), 240-254.

Geiger, J. P., Rio, B., Nicole, M., & Nandris, D. (1986). Biodegradation of Hevea brasiliensis wood by Rigidoporus lignosus and Phellinus noxius 1. *European Journal of forest pathology*, *16*(3), 147-159.

Ghaddar, B., & Naoum, S. J. (2018). High dimensional data classification and feature selection using support vector machines. *European Journal of Operational Research*, *265*(3), 993-1004.

Ghouri, F., Ali, Z., Naeem, M., Ul-Allah, S., Babar, M., Baloch, F. S., ... & Shahid, M. Q. (2022). Effects of silicon and selenium in alleviation of drought stress in rice. *Silicon*, *14*(10), 5453-5461.

Gianfagna, T. (1995). Natural and synthetic growth regulators and their use in horticultural and agronomic crops. In *Plant hormones: physiology, biochemistry and molecular biology*,751-773. Dordrecht: Springer Netherlands.

© COPYRIGHT UPM

Go, W. Z., Chin, K. L., H'ng, P. S., Wong, M. Y., Lee, C. L., & Khoo, P. S. (2023). Exploring the biocontrol efficacy of Trichoderma spp. against Rigidoporus microporus, the causal agent of white root rot disease in rubber trees (Hevea brasiliensis). *Plants*, *12*(5), 1066.

Go, W. Z., Chin, K. L., H'ng, P. S., Wong, M. Y., Luqman, C. A., Surendran, A., Tan, G. H., Lee, C. L., Khoo, P. S., & Kong, W. J. (2021). Virulence of Rigidoporus microporus isolates causing white root rot disease on rubber trees (*Hevea brasiliensis*) in Malaysia. *Plants*, *10*(10), 2123.

Go, W. Z., H'ng, P. S., Wong, M. Y., Tan, G. H., Chuah, A. L., Salmiah, U. Maminska, R. T., Soni, O., Wong, W. Z., Chin, K. L., & Chai, E. W. (2015). Occurrence and characterisation of mycoflora in soil of different health conditions associated with white root rot disease in Malaysia rubber plantation. *Journal of Rubber Research*, *18*, 3.

Godlewska, K., Biesiada, A., Michalak, I., & Pacyga, P. (2019). The effect of plant-derived biostimulants on white head cabbage seedlings grown under controlled conditions. *Sustainability (Switzerland)*, *11*(19).

Godlewska, K., Michalak, I., Tuhy, L., & Chojnacka, K. (2016). Plant Growth Biostimulants Based on Different Methods of Seaweed Extraction with Water. *BioMed Research International*, *2016*.

Gold, K. M., Townsend, P. A., Chlus, A., Herrmann, I., Couture, J. J., Larson, E. R., & Gevens, A. J. (2020). Hyperspectral measurements enable pre-symptomatic detection and differentiation of contrasting physiological effects of late blight and early blight in potato. *Remote Sensing*, *12*(2), 286.

Golhani, K., Balasundram, S. K., Vadamalai, G., & Pradhan, B. (2019). Estimating chlorophyll content at leaf scale in viroid-inoculated oil palm seedlings (Elaeis guineensis Jacq.) using reflectance spectra (400 nm–1050 nm). *International Journal of Remote Sensing*, *40*(19), 7647-7662.

Gomez-Garcia, D., Rodriguez-Morales, F., Welch, S., & Leuschen, C. (2022). High-Throughput Phenotyping of Wheat Canopy Height Using Ultrawideband Radar: First Results. *IEEE Geoscience and Remote Sensing Letters*, *19*.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*(7), 644–652.

Grace, A. (2020). *A beginner's guide to RT-PCR, qPCR and RT-qPCR*. http://portlandpress.com/biochemist/articlepdf/42/3/48/917384/bio20200034.pdf

Grossman, L., & Moldave, K. (1967). *Nucleic acids*. Academic Press.

© COPYRIGHT UPM

Gu, Y., Deng, Z., Paredez, A. R., Debolt, S., Wang, Z.-Y., & Somerville, C. (2008). *Prefoldin 6 is required for normal microtubule dynamics and organization in Arabidopsis*.

Guidi, L., & Tattini, M. (2021). Antioxidant defenses in plants: A dated topic of current interest. In *Antioxidants* 10(6). MDPI.

Guo, B., Liu, M., Yang, H., Dai, L., & Wang, L. (2023). Brassinosteroids Regulate the Water Deficit and Latex Yield of Rubber Trees. *International Journal of Molecular Sciences*, *24*(16), 12857.

Gupta, D. R., Surovy, M. Z., Mahmud, N. U., Chakraborty, M., Paul, S. K., Hossain, M. S., ... & Islam, M. T. (2020). Suitable methods for isolation, culture, storage and identification of wheat blast fungus Magnaporthe oryzae Triticum pathotype. *Phytopathology Research*, *2*, 1-13.

Guyot, J., & Flori, A. (2002). Comparative study for detecting Rigidoporus lignosus on rubber trees. *Crop Protection*, *21*(6), 461-466.

Hadidi, A., Czosnek, H., & Barba, M. (2004). DNA microarrays and their potential applications for the detection of plant viruses, viroids, and phytoplasmas. *Journal of Plant Pathology*, 97-104.

Hadzli, H., Noor, E. A., Suhaimi, S., & Marwani, I. (2014). Classification of Hevea brasiliensis leaf diseases through imaging technique and neural network. *Journal of Applied Science and Agriculture*, *9*(11 Special), 358-365.

Harith-Fadzilah, N., Haris-Hussain, M., Ghani, I. A., Zakaria, A., Amit, S., Zainal, Z., Azmi, W. A., Jalinas, J., & Hassan, M. (2020). Physical and physiological monitoring on red palm weevil-infested oil palms. *Insects*, *11*(7), 1–12.

Harun, M. Y., Che Yunus, M. A., Ismail, M. H. S., & Morad, N. A. (2016). A comparative investigation on the effect of thermal treatments on the mechanical properties of oil palm fruitlet components. *Journal of the Taiwan Institute of Chemical Engineers*, *60*, 582–587.

Hasan, S., Jahan, S., & Islam, M. I. (2022). Disease detection of apple leaf with combination of color segmentation and modified DWT. *Journal of King Saud University-Computer and Information Sciences*, *34*(9), 7212-7224.

Hashim, H., Haron, M. A., Osman, F. N., & Al Junid, S. A. M. (2010, May). Classification of rubber tree leaf disease using spectrometer. In *2010 Fourth Asia International Conference on Mathematical/Analytical Modelling and Computer Simulation*(pp. 302-306). IEEE.

Hashim, H., Sampian, A. F. M., Khairuzzaman, N. A., Abdullah, N. E., Sulaiman, M. S., Aris, M. N. M., & Abd Hamid, N. R. (2015, November). Discriminant analysis with visible lighting properties for White root disease infected rubber tree. In *IOP Conference Series: Materials Science and Engineering* (Vol. 99, No. 1, p. 012023). IOP Publishing.

© COPYRIGHT UPM

Hawkins, T., & Kihara, D. (2007). Function Prediction Of Uncharacterized Proteins. In *Journal of Bioinformatics and Computational Biology* ,5(1).

Haynes, R. S., Lucieer, A., Brodribb, T. J., Tonet, V., & Cimoli, E. (2024). Predicting key water stress indicators of Eucalyptus viminalis and Callitris rhomboidea using high-resolution visible to short-wave infrared spectroscopy. *Plant, Cell & Environment*.

Heim, R. H. J., Wright, I. J., Chang, H. C., Carnegie, A. J., Pegg, G. S., Lancaster, E. K., Falster, D. S., & Oldeland, J. (2018). Detecting myrtle rust (Austropuccinia psidii) on lemon myrtle trees using spectral signatures and machine learning. *Plant pathology*, *67*(5), 1114-1121.

Henson, I. E., Roslan Noor, M., Haniff Harun, M., Yahya, Z., & Nor Aishah Mustakim, S. (2005). Stress Development and Its Detection In Young Oil Palms In North Kedah, Malaysia. In *Journal of Oil Palm Research*. 17.

Hitimana, E., & Gwun, O. (2014). Automatic estimation of live Coffee leaf infection based on Image processing techniques.

Ho, K. C., Chan, M. K., Chen, Y. M., & Subhramaniyun, P. (2023). Treatment of rubber industry wastewater review: Recent advances and future prospects. *Journal of Water Process Engineering*, *52*, 103559.

Hong, M., Tao, S., Zhang, L., Diao, L. T., Huang, X., Huang, S., Xie, S. J., Xiao, Z. D., & Zhang, H. (2020). RNA sequencing: new technologies and applications in cancer research. In *Journal of Hematology and Oncology*, 13(1). BioMed Central Ltd.

Horsfall, J. G. (Ed.). (2012). *Plant disease: an advanced treatise: How Plants Suffer from Disease*. Elsevier.1-18.

Hotta, Y., Tanaka, T., Takaoka, H., Takeuchi, Y., & Konnai, M. (1997). New physiological effects of 5-aminolevulinic acid in plants: the increase of photosynthesis, chlorophyll content, and plant growth. *Bioscience, biotechnology, and biochemistry*, *61*(12), 2025-2028.

Hou, C., Zhuang, J., Tang, Y., He, Y., Miao, A., Huang, H., & Luo, S. (2021). Recognition of early blight and late blight diseases on potato leaves based on graph cut segmentation. *Journal of Agriculture and Food Research*, *5*, 100154.

Hou, B., Hu, Y., Zhang, P., & Hou, L. (2022). Potato late blight severity and epidemic period prediction based on Vis/NIR spectroscopy. *Agriculture*, *12*(7), 897.

Houssein, E. H., Oliva, D., Celik, E., Emam, M. M., & Ghoniem, R. M. (2023). Boosted sooty tern optimization algorithm for global optimization and feature selection. *Expert Systems with Applications*, *213*, 119015.

Hu, T., Chitnis, N., Monos, D., & Dinh, A. (2021). Next-generation sequencing technologies: An overview. *Human Immunology*, *82*(11), 801–811.

© COPYRIGHT UPM

Huang, M. L., Hung, Y. H., Lee, W. M., Li, R. K., & Jiang, B. R. (2014). SVM-RFE based feature selection and Taguchi parameters optimization for multiclass SVM classifier. *The Scientific World Journal*, *2014*(1), 795624.

Husin, N. A., Khairunniza-Bejo, S., Abdullah, A. F., Kassim, M. S. M., Ahmad, D., & Azmi, A. N. N. (2020). Application of Ground-Based LiDAR for Analysing Oil Palm Canopy Properties on the Occurrence of Basal Stem Rot (BSR) Disease. *Scientific Reports*, *10*(1).

Hussin, M. H., Abd Latif, N. H., Hamidon, T. S., Idris, N. N., Hashim, R., Appaturi, J. N., ... & Sedliačik, J. (2022). Latest advancements in high-performance bio-based wood adhesives: A critical review. *Journal of Materials Research and Technology*, *21*, 3909-3946.

Iglesias, C., Mayer, J., Chavez, L., & Calle, F. (1997). Genetic potential and stability of carotene content in cassava roots. *Euphytica*, *94*(3), 367-373.

International Rubber Study Group (IRSG), 2021. Rubber statistical bulletin. IRSG, Singapore 76 (1), 1–3.

Islam, M., Dinh, A., Wahid, K., & Bhowmik, P. (2017, April). Detection of potato diseases using image segmentation and multiclass support vector machine. In *2017 IEEE 30th canadian conference on electrical and computer engineering (CCECE)*, 1-4. IEEE.

Ismail, F. A., Hashim, H., Madzhi, N. K., Abdullah, N. E., Sam, R., Latib, S., & Sulaiman, M. S. (2013, December). Fuzzification of rubber tree white root disease based on leaf's discolouration. In *2013 International Conference on Electrical, Electronics and System Engineering (ICEESE)*, 54-59. IEEE.

Izzuddin, M. A., Seman Idris, A., Nisfariza, M. N., Nordiana, A. A., Shafri, H. Z. M., & Ezzati, B. (2017). The development of spectral indices for early detection of Ganoderma disease in oil palm seedlings. *International Journal of Remote Sensing*, *38*(23), 6505-6527.

Jaleel, C. A., Manivannan, P., Wahid, A., Farooq, M., Somasundaram, R., & Panneerselvam, R. (2009). Drought stress in plants: a review on morphological characteristics and pigments composition. *International Journal Of Agriculture & Biology*, *11*, 100–105.

Jayamohan, H., Lambert, C. J., Sant, H. J., Jafek, A., Patel, D., Feng, H., Beeman, M., Mahmood, T., Nze, U., & Gale, B. K. (2021). SARS-CoV-2 pandemic: a review of molecular diagnostic tools including sample collection and commercial response with associated advantages and limitations. *Analytical and Bioanalytical Chemistry*, *413*, 49-71.

Jayaratne, R. (1997). Economical and less hazardous fungicide for the management of white root disease.

© COPYRIGHT UPM

Jayasinghe, C. K., Jayasuriya, K. E., & Fernando, T. H. P. S. (1995). Penta chloro phenol-effective and economical fungicide for the management of white root disease caused by *Rigidoporus lignosus* in Sri Lanka.

Jayasuriya, K. E., & Thennakoon, B. I. (2007). Biological control of Rigidoporus microporus, the cause of white root disease in rubber. *Cey. J. Sci.(Bio. Sci.)*, *36*(1), 9-16.

Jia, H. sen, Guan, C. F., Zhang, J. S., He, C. X., Yin, C. J., & Meng, P. (2022). Drought effects on tree growth, water use efficiency, vulnerability and canopy health of Quercus variabilis-Robinia pseudoacacia mixed plantation. *Frontiers in Plant Science*, *13*.

Jia, W., Sun, M., Lian, J., & Hou, S. (2022). Feature dimensionality reduction: a review. *Complex & Intelligent Systems*, *8*(3), 2663-2693.

Jiang, Q., Wu, G., Tian, C., Li, N., Yang, H., Bai, Y., & Zhang, B. (2021). Hyperspectral imaging for early identification of strawberry leaves diseases with machine learning and spectral fingerprint features. *Infrared Physics & Technology*, *118*, 103898.

Jiang, X., Zhen, J., Miao, J., Zhao, D., Shen, Z., Jiang, J., ... & Wang, J. (2022). Newly-developed three-band hyperspectral vegetation index for estimating leaf relative chlorophyll content of mangrove under different severities of pest and disease. *Ecological Indicators*, *140*, 108978.

Jin, J., Sun, Y., Qu, J., Syah, R., Lim, C. H., Alfiko, Y., Rahman, N. E. B., Suwanto, A., Yue, G., Wong, L., Chua, N. H., & Ye, J. (2017). Transcriptome and functional analysis reveals hybrid vigor for oil biosynthesis in oil palm. *Scientific Reports*, *7*(1).

Judkins, W. P., & Wander, I. W. (1950). Correlation between leaf color, leaf nitrogen content, and growth of apple, peach, and grape plants. *Plant physiology*, *25*(1), 78.

Julrat, S., Chongcheawchamnan, M., Khaorapapong, T., Patarapiboolchai, O., Kririksh, M., & Robertson, I. D. (2012). Single-frequency-based dry rubber content determination technique for in-field measurement application. *IEEE Sensors Journal*, *12*(10), 3019-3030.

Kaewchai, S. L. F. C., Lin, F. C., Wang, H. K., Soytong, K., & Soytong, K. (2010). Characterization of Rigidoporus microporus isolated from rubber trees based on morphology and ITS sequencing. *Journal of Agricultural Technology*, *6*(2), 289-298.

Kaewchai, S., & Soytong, K. (2010). Application of biofungicides against *Rigidoporus microporus* causing white root disease of rubber trees. *Journal of Agricultural Technology*, *6*(2), 349-363.

Kamal, M., Munawar, A. A., & Sulaiman, M. I. (2021, February). Comparison of principal component and partial least square regression method in NIRS data

© COPYRIGHT UPM

analysis for cocoa bean quality assessment. In *IOP Conference Series: Earth and Environmental Science*, 667(1), 012058. IOP Publishing.

Kaniapan, S., Hassan, S., Ya, H., Nesan, K. P., & Azeem, M. (2021). *The Utilisation of Palm Oil and Oil Palm Residues and the Related Challenges as a Sustainable Alternative in Biofuel, Bioenergy, and Transportation Sector: A Review*.

Karadağ, K., Tenekeci, M. E., Taşaltın, R., & Bilgili, A. (2020). Detection of pepper fusarium disease using machine learning algorithms based on spectral reflectance. *Sustainable Computing: Informatics and Systems*, *28*, 100299.

Keskin, A., & Birinci, A. (2010). Analysis of the factors affecting the instrument and machinery assets in enterprises that deal with agricultural production: The case of Erzurum Province. *African Journal of Agricultural Research*, *5*(8), 600–605.

Keskin, A., Tumer, E. I., & Birinci, A. (2010). Analysis of the factors affecting the instrument and machinery assets in enterprises that deal with agricultural production: The case of erzurum province. *African Journal of Agricultural Research*, *5*(8), 600–605.

Ketta, H. A., & Hewedy, O. A. E. R. (2021). Biological control of Phaseolus vulgaris and Pisum sativum root rot disease using Trichoderma species. *Egyptian Journal of Biological Pest Control*, *31*, 1-9.

Khairuzzaman, N. A., Hashim, H., Mohd Sampian, A. F., & Yussuf, U. R. (2017). White root disease infected rubber tree detection through statistical inference of Dry Rubber sheet Near Infrared (NIR) spectrum measurement method. *Journal of Mechanical Engineering*, (2), 167-177.

Khaled, Alfadhl Yahya, Samsuzana Abd Aziz, Siti Khairunniza Bejo, Nazmi Mat Nawi, Diyana Jamaludin, and Nur Ul Atikah Ibrahim. "A comparative study on dimensionality reduction of dielectric spectral data for the classification of basal stem rot (BSR) disease in oil palm." *Computers and electronics in agriculture* 170 (2020): 105288.

Khan, M. A., Akram, T., Sharif, M., Javed, K., Raza, M., & Saba, T. (2020). An automated system for cucumber leaf diseased spot detection and classification using improved saliency method and deep features selection. *Multimedia Tools and Applications*, *79*, 18627-18656.

Khan, I. H., Liu, H., Li, W., Cao, A., Wang, X., Liu, H., ... & Yao, X. (2021). Early detection of powdery mildew disease and accurate quantification of its severity using hyperspectral images in wheat. *Remote Sensing*, *13*(18), 3612.

Khayatnezhad, M., & Gholamin, R. (2020). *The Effect of Drought Stress on the Superoxide Dismutase and Chlorophyll Content in Durum Wheat Genotypes*. www.als-journal.com

© COPYRIGHT UPM

Killiny, N., & Nehela, Y. (2017). One target, two mechanisms: the impact of 'Candidatus Liberibacter asiaticus' and its vector, Diaphorina citri, on citrus leaf pigments. *Molecular Plant-Microbe Interactions*, *30*(7), 543-556.

Kira, K., & Rendell, L. A. (1992). A practical approach to feature selection. In *Machine learning proceedings 1992*, 249-256. Morgan Kaufmann.

Körner, C. (2015). Paradigm shift in plant growth control. In *Current Opinion in Plant Biology* 25,107–114. Elsevier Ltd.

Kralik, P., & Ricchi, M. (2017). A basic guide to real time PCR in microbial diagnostics: Definitions, parameters, and everything. In *Frontiers in Microbiology* 8(2).

Kranz, J. (1988). Measuring plant disease. In *Experimental techniques in plant disease epidemiology* ,35-50. Springer, Berlin, Heidelberg.

Kumar, R., Pathak, S., Prakash, N., Priya, U., & Ghatak, A. (2021). Application of Spectroscopic Techniques in Early Detection of Fungal Plant Pathogens. In *Diagnostics of Plant Diseases*. London, UK: IntechOpen.

Laghari, A. A., & Yin, S. (2022). How to collect and interpret medical pictures captured in highly challenging environments that range from nanoscale to hyperspectral imaging. *Current Medical Imaging*, *54*(36582065), 1.

Lakshmi, K., Mahaboob, B., Venkateswararao, P., Krishna, Y. H., Sreeharireddy, P., & Balajiprakash, G. (2021). On mathematical and statistical aspects of linear models. In *Intelligent and Reliable Engineering Systems* ,98-106. CRC Press.

Lazim, S. S. R. M., Nawi, M. N., Bejo, S. K., Shariff, A. R. M., & Abdullah, N. (2022). Prediction and classification of soluble solid contents to determine the maturity level of watermelon using visible and shortwave near infrared spectroscopy. *International Food Research Journal*, *29*(6).

Lazim, Siti Saripa Rabiah (2019) Non-destructive assessment of the internal quality of watermelon using ultraviolet near-infrared spectroscopy. *Masters thesis, Universiti Putra Malaysia*.

Lee, F. C., Yeap, W. C., Appleton, D. R., Ho, C. L., & Kulaveerasingam, H. (2022). Identification of drought responsive Elaeis guineensis WRKY transcription factors with sensitivity to other abiotic stresses and hormone treatments. *BMC Genomics*, *23*(1).

Lee, F. C., Yeap, W. C., Appleton, D. R., Ho, C. L., & Kulaveerasingam, H. (2022b). Identification of drought responsive Elaeis guineensis WRKY transcription factors with sensitivity to other abiotic stresses and hormone treatments. *BMC Genomics*, *23*(1).

Lei, X., Xiao, Y., Xia, W., Mason, A. S., Yang, Y., Ma, Z., & Peng, M. (2014). RNA-seq analysis of oil palm under cold stress reveals a different C-repeat binding factor (CBF) mediated gene expression pattern in Elaeis guineensis compared

© COPYRIGHT UPM

to other species. *PLoS ONE*, *9*(12).

Leroux, M. R., Fä Ndrich, M., Klunker, D., Siegers, K., Lupas, A. N., Brown, J. R., Schiebel, E., Dobson, C. M., Hartl, F. U., Fändrich, M., & Klunker, D. (1999). MtGimC, a novel archaeal chaperone related to the eukaryotic chaperonin cofactor GimC/prefoldin. In *The EMBO Journal*, 18(23).

Li, M., Zhang, L., Jiang, L. L., Zhao, Z. B., Long, Y. H., Chen, D. M., ... & Liu, Y. J. (2024). Label-free Raman microspectroscopic imaging with chemometrics for cellular investigation of apple ring rot and nondestructive early recognition using near-infrared reflection spectroscopy with machine learning. *Talanta*, *267*, 125212.

Li, R., Reddy, V. A., Jin, J., Rajan, C., Wang, Q., Yue, G., Lim, C. H., Chua, N. H., Ye, J., & Sarojam, R. (2017). Comparative transcriptome analysis of oil palm flowers reveals an EAR-motif-containing R2R3-MYB that modulates phenylpropene biosynthesis. *BMC Plant Biology*, *17*(1).

Li, Y., Sun, Y., Jiang, J., & Liu, J. (2019). Spectroscopic determination of leaf chlorophyll content and color for genetic selection on Sassafras tzumu. *Plant methods*, *15*(1), 1-11.

Li, Y., Zhang, L., Li, R., Zhang, M., Li, Y., Wang, H., Wang, S., & Bao, Z. (2019). Systematic identification and validation of the reference genes from 60 RNA-Seq libraries in the scallop Mizuhopecten yessoensis. *BMC Genomics*, *20*(1).

Liaghat, S., Mansor, S., Ehsani, R., Shafri, H. Z. M., Meon, S., & Sankaran, S. (2014). Mid-infrared spectroscopy for early detection of basal stem rot disease in oil palm. *Computers and electronics in agriculture*, *101*, 48-54.

Lim, T. K., Hamm, R. T., & Mohamad, R. B. (1990). Persistency and volatile behaviour of selected chemicals in treated soil against three basidiomycetous root disease pathogens. *International Journal of Pest Management*, *36*(1), 23-26.

Liu, J., Han, J., Chen, X., Shi, L., & Zhang, L. (2019). Nondestructive detection of rape leaf chlorophyll level based on Vis-NIR spectroscopy. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *222*, 117202.

Liu, Y., Li, D., Li, H., Jiang, X., Zhu, Y., Cao, W., & Ni, J. (2022). Design of a Phenotypic Sensor About Protein and Moisture in Wheat Grain. *Frontiers in Plant Science*, *13*, 881560.

Liu, Z., & Li, Y. (2022). Fungi classification in various growth stages using shortwave infrared (SWIR) spectroscopy and machine learning. *Journal of Fungi*, *8*(9), 978.

Livak, K. J., & Schmittgen, T. D. (2001a). Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. *Methods*, *25*(4), 402–408.

© COPYRIGHT UPM

Livak, K. J., & Schmittgen, T. D. (2001b). Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. *Methods*, *25*(4), 402–408.

Longsaward, R., Pengnoo, A., Kongsawadworakul, P., & Viboonjun, U. (2023). A novel rubber tree PR-10 protein involved in host-defense response against the white root rot fungus Rigidoporus microporus. *BMC Plant Biology*, *23*(1), 157.

Loo, K. Y., Law, J. W. F., Tan, L. T. H., Pusparajah, P., Letchumanan, V., & Lee, L. H. (2022). Diagnostic techniques for rapid detection of Vibrio species. *Aquaculture*, *561*, 738628.

Lopes Filho, W. R. L., Rodrigues, F. H. S., Ferreira, I. V. L., Correa, L. O., Cunha, R. L., & Pinheiro, H. A. (2021). Physiological responses of young oil palm (Elaeis guineensis Jacq.) plants to repetitive water deficit events. *Industrial Crops and Products*, *172*.

Lopes, U. P., Alonzo, G., Onofre, R. B., Mello, P. P., Gadoury, D. M., Vallad, G. E., & Peres, N. A. (2023). Effective management of powdery mildew in cantaloupe plants using nighttime applications of UV light. *Plant Disease*, *107*(8), 2483-2489.

Lopez, E., Etxebarria-Elezgarai, J., Amigo, J. M., & Seifert, A. (2023). The importance of choosing a proper validation strategy in predictive models. A tutorial with real examples. *Analytica Chimica Acta*, 341532.

Luo, L., Chang, Q., Wang, Q., & Huang, Y. (2021). Identification and severity monitoring of maize dwarf mosaic virus infection based on hyperspectral measurements. *Remote Sensing*, *13*(22), 4560.

Luo, L., Guo, C., Wang, L., Zhang, J., Deng, L., Luo, K., ... & Yang, M. (2019). Negative plant-soil feedback driven by re-assemblage of the rhizosphere microbiome with the growth of Panax notoginseng. *Frontiers in Microbiology*, *10*, 1597.

Ma, X., Liu, J., Li, H., Wang, W., Liu, L., Wang, P., ... & Qu, F. (2024). Greenhouse covering cultivation promotes chlorophyll accumulation of tea plant (Camellia sinensis) by activating relevant gene expression and enzyme activity. *BMC Plant Biology*, *24*(1), 455.

Mahajan, S., & Pandit, A. K. (2023). Hybrid method to supervise feature selection using signal processing and complex algebra techniques. *Multimedia Tools and Applications*, *82*(6), 8213-8234.

Mahmood, T., Khalid, S., Abdullah, M., Ahmed, Z., Shah, M. K. N., Ghafoor, A., & Du, X. (2020). Insights into drought stress signaling in plants and the molecular genetic basis of cotton drought tolerance. In *Cells* ,*9*(1). MDPI.

Mahyudin, M. M., Noran, A. S., Ismail, M. Z. H., & Ahmad, K. (2023). Diseases of rubber trees: Malaysia as a case study. In *Forest Microbiology*, 401-414. Academic Press.

© COPYRIGHT UPM

Mahyudin, M. M., Noran, A. S., Ismail, M. Z. H., & Ahmad, K. (2023). Diseases of rubber trees: Malaysia as a case study. In *Forest Microbiology* 401-414. Academic Press.

Maiden, N. A., Syd Ali, N., Ahmad, K., Atan, S., & Wong, M. Y. (2022). Growth and physiological responses of Hevea brasiliensis to Rigidoporus microporus infection. *Journal of Rubber Research*, *25*(3), 213-221.

Malavi, D., Nikkhah, A., Alighaleh, P., Einafshar, S., Raes, K., & Van Haute, S. (2024). Detection of saffron adulteration with Crocus sativus style using NIR-hyperspectral imaging and chemometrics. *Food Control*, *157*, 110189.

Malaysia Rubber Board. (2024). About us. https://www.lgm.gov.my/webv2/pdfViewer/nrStatistic

Malaysian Rubber Board (MRB), 2021. Natural Rubber: Planted Area of Natural Rubber on Estates and Smallholdings in Malaysia ('000 Hectares). Pocket Book July 2021, p. 4.

Mandal, N., Das, D. K., Sahoo, R. N., Adak, S., Kumar, A., Viswanathan, C., ... & Das, B. (2022). Assessing rice blast disease severity through hyperspectral remote sensing. *Journal of Agrometeorology*, *24*(3), 241-248.

Masazhar, A. N. I., & Kamal, M. M. (2017, November). Digital image processing technique for palm oil leaf disease detection using multiclass SVM classifier. In *2017 IEEE 4th International Conference on Smart Instrumentation, Measurement and Application (ICSIMA)*, 1-6). IEEE.

Masood, R., Khan, S. A., & Khan, M. N. A. (2016). Plants disease segmentation using image processing. *International Journal of Modern Education and Computer Science*, *8*(1), 24.

Matthies, C., Erhard, H. P., & Drake, H. L. (1997). Effects of pH on the comparative culturability of fungi and bacteria from acidic and less acidic forest soils. *Journal of Basic Microbiology*, *37*(5), 335-343.

McNew, G. l., (1960). The nature, origin and evolution of parasitism. In "plant Pathology: An Advanced Treatise" (J. G. Hornsfall and A. E. Dimond, ends.) 2,9-69. Academic Press, New York.

Mekonnen, T. H., Ah-Leung, T., Hojabr, S., & Berry, R. (2019). Investigation of the co-coagulation of natural rubber latex and cellulose nanocrystals aqueous dispersion. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, *583*, 123949.

Minocha, R., Majumdar, R., & Minocha, S. C. (2014). Polyamines and abiotic stress in plants: a complex relationship1. *Frontiers in Plant Science*, *5*(2553), 1–17.

Mira, J. L., Barba, J., Romero, F. P., Escolar, M. S., Caba, J., & López, J. C. (2024). Benchmarking of computer vision methods for energy-efficient high-accuracy olive fly detection on edge devices. *Multimedia Tools and*

© COPYRIGHT UPM

Applications, *83*(34), 81785-81809.

Mittelberger, C., Yalcinkaya, H., Pichler, C., Gasser, J., Scherzer, G., Erhart, T., Sandra, S., Holzer. B., Janik, K., Robatscher, P. & Müller, T. (2017). Pathogen-induced leaf chlorosis: products of chlorophyll breakdown found in degreened leaves of Phytoplasma-infected apple (malus× domestica borkh.) and apricot (Prunus armeniaca L.) trees relate to the pheophorbide a oxygenase/phyllobilin pathway. *Journal of Agricultural and Food Chemistry*, *65*(13), 2651-2660.

Miyashita, Y. (2024). Cortical layer-dependent signaling in cognition: Three computational modes of the canonical circuit. *Annual Review of Neuroscience*, *47*(1), 211-234.

Mo, Y., Wan, R., & Zhang, Q. (2012). *Application of Reverse Transcription-PCR and Real-Time PCR in Nanotoxicity Research*, 99–112.

Mogollón, R., Contreras, C., da Silva Neta, M. L., Marques, E. J. N., Zoffoli, J. P., & de Freitas, S. T. (2020). Non-destructive prediction and detection of internal physiological disorders in'Keitt'mango using a hand-held Vis-NIR spectrometer. *Postharvest Biology and Technology*, *167*, 111251.

Mohd, S. S., Hadzli, H., Noor, E. A., Siti, S. R., Nurbaiti, W., Azrie, F., & Faridatul, A. I. (2014). Early detection of white-root disease for rubber tree based on leaf discoloration with neural network technique. *Journal of Applied Science and Agriculture*, *9*(11), 341-351.

Mokhatar, S. J., Daud, W. N., & Zamri, N. M. (2011). Evaluation of hevea brasiliensis (latex timber clone: RRIM 2001 and RRIM 3001) in relation to different water stress. *American Journal of Agricultural and Biological Sciences*, *6*(1), 122-127.

Molero, G., & Reynolds, M. P. (2020). Spike photosynthesis measured at high throughput indicates genetic variation independent of flag leaf photosynthesis. *Field Crops Research*, *255*.

Monir, M. U., Aziz, A. A., & Yousuf, A. (2022). Integrated technique to produce sustainable bioethanol from lignocellulosic biomass. *Materials Letters: X*, *13*.

Montazeaud, G., Langrume, C., Moinard, S., Goby, C., Ducanchez, A., Tisseyre, B., & Brunel, G. (2021). Development of a low cost open-source ultrasonic device for plant height measurements. *Smart Agricultural Technology*, *1*, 100022.

Monzon, J. P., Jabloun, M., Cock, J., Caliman, J. P., Couëdel, A., Donough, C. R., Vui, P. H. V., Lim, Y. L., Mathews, J., Oberthür, T., Prabowo, N. E., Edreira, J. I. R., Sidhu, M., Slingerland, M. A., Sugianto, H., & Grassini, P. (2022). Influence of weather and endogenous cycles on spatiotemporal yield variation in oil palm. *Agricultural and Forest Meteorology*, *314*.

Mosele, F., Remon, J., Mateo, J., Westphalen, C. B., Barlesi, F., Lolkema, M. P., Normanno, N., Scarpa, A., Robson, M., Meric-Bernstam, F., Wagle, N.,

© COPYRIGHT UPM

Stenzinger, A., Bonastre, J., Bayle, A., Michiels, S., Bièche, I., Rouleau, E., Jezdic, S., Douillard, J. Y., … André, F. (2020). Recommendations for the use of next-generation sequencing (NGS) for patients with metastatic cancers: a report from the ESMO Precision Medicine Working Group. In *Annals of Oncology*, 31(11),1491–1505. Elsevier Ltd..

Morais, C. L., Lima, K. M., Singh, M., & Martin, F. L. (2020). Tutorial: multivariate classification for vibrational spectroscopy in biological samples. *Nature Protocols*, *15*(7), 2143-2162.

Motato-Vásquez, V., Pires, R. M., Vitali, V. M. V., & Gugliotta, A. D. M. (2016). Cultural and ligninolytic activity studies of some polypores (Basidiomycota) from brazilian Atlantic Forest, São Paulo State, Brazil. *Hoehnea*, *43*, 289-300.

Munsell, A. H. (2022). A Color Notation: a measured color system, based on the three qualities Hue, Value and Chroma. DigiCat.

Mutiara Amanah, D., Laksmita, N.-H., & Santi, P. (2019). *Physiological responses of bio-silica-treated oil palm seedlings to drought stress*. *87*(1), 20–30..

Mwebaze, E., & Owomugisha, G. (2016, December). Machine learning for plant disease incidence and severity measurements from leaf images. In *2016 15th IEEE international conference on machine learning and applications (ICMLA)* (pp. 158-163). IEEE.

Nair, K. P., & Nair, K. P. (2021). Rubber (Hevea brasiliensis). *Tree Crops: Harvesting Cash from the World's Important Cash Crops*, 287-332.

Najihah, T. S., Ibrahim, M. H., Razak, A. A., Nulit, R., & Megat, P. E. W. (2019). Effects of water stress on the growth, physiology and biochemical properties of oil palm seedlings. *AIMS Agriculture and Food*, *4*(4), 854–868.

Najjar, K., & Abu, N. (2021). Visible/near-infrared (VIS/NIR) spectroscopy technique to detect gray mold disease in the early stages of tomato fruit: VIS/NIR spectroscopy for detecting gray mold in tomato. *Journal of microbiology, biotechnology and food sciences*, *11*(2), e3108-e3108.

Nakaew, N., Rangjaroen, C., & Sungthong, R. (2015). Utilization of rhizospheric Streptomyces for biological control of Rigidoporus sp. causing white root disease in rubber tree. *European Journal of Plant Pathology*, *142*, 93-105.

Nantongo, J. S., Serunkuma, E., Burgos, G., Nakitto, M., Davrieux, F., & Ssali, R. (2024). Machine learning methods in near infrared spectroscopy for predicting sensory traits in sweetpotatoes. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *318*, 124406.

Nawi , (2014). Development of New Measurement Methods to determine Sugarcane Quality from Stalk Samples. PhD thesis, University of Southern Queensland, Australia.

Nemeskéri, E., & Helyes, L. (2019). Physiological responses of selected vegetable crop species to water stress. In *Agronomy*, 9(8). MDPI AG.

Neto, J. C. R., Vieira, L. R., de Aquino Ribeiro, J. A., de Sousa, C. A. F., Júnior, M. T. S., & Abdelnur, P. V. (2021). Metabolic effect of drought stress on the leaves of young oil palm (Elaeis guineensis) plants using UHPLC–MS and multivariate analysis. *Scientific Reports*, *11*(1).

Nie, P., Qu, F., Lin, L., Dong, T., He, Y., Shao, Y., & Zhang, Y. (2017). Detection of water content in rapeseed leaves using terahertz spectroscopy. *Sensors*, *17*(12), 2830.

Nobles, M. K. (1965). Identification of cultures of wood-inhabiting Hymenomycetes. *Canadian journal of botany*, *43*(9), 1097-1139.

Noireung, P. (2020). First Record of Phytopythium vexans causing root rot on Mandarin (Citrus reticulate L. cv. Sainampueng) in Thailand. *Plant Pathology & Quarantine*, *10*(1), 85–90.

Norhidayu, A., Nur-Syazwani, Radzil, R., & Balu, A. (2017). The Production of Crude Palm Oil in Malaysia. In *International Journal of Economics and Management Journal homepage* ,11(S3).

Nuanlaong, S., Wuthisuthimethavee, S., Azzeme, A. M., & Suraninpong, P. (2022). Optimized Method for the Identification of Candidate Genes and Molecular Maker Development Related to Drought Tolerance in Oil Palm (Elaeis guineensis Jacq.). *Plants*, *11*(17).

Nugraha, M. A. R., Lin, Y. R., Dewi, N. R., Huang, H. T., Nan, F. H., & Hu, Y. F. (2024). Effects of Taiwanese indigenous cinnamon (Cinnamomum osmophloeum) leaf hot-water extract on nonspecific immune responses, resistance against Vibrio parahaemolyticus, nonviable cells, and haemocyte subpopulations in white shrimp (Penaeus vannamei). *Fish & Shellfish Immunology*, 109680.

Nurwahyuni, E., & Putra, E. T. S. (2021). The role of calcium in drought stress response induced through antioxidant activity in oil palm (Elaeis guineensis Jacq.) seedlings. *E-Journal Menara Perkebunan*, *89*(1).

Omorusi, V. I. (2012). Effects of White Root Rot Disease on Hevea brasiliensis (Muell. Arg.). Challenges and Control Approach. *Plant Science*, *14*.

Omorusi, V. I., Eguavoen, O. I., Ogbebor, N. O., Bosah, B. O., Orumwense, K., & Ijie, K. (2014). Control of white root rot disease in rubber plantations in Nigeria. *International Journal of Microbiology and Immunology Research*, *3*(4), 046-051.

Omuya, E. O., Okeyo, G. O., & Kimwele, M. W. (2021). Feature selection for classification using principal component analysis and information gain. *Expert Systems with Applications*, *174*, 114765.

© COPYRIGHT UPM

Onofre, R. B., Gadoury, D. M., Stensvand, A., Bierman, A., Rea, M., & Peres, N. A. (2021). Use of ultraviolet light to suppress powdery mildew in strawberry fruit production fields. *Plant Disease*, *105*(9), 2402-2409.

Osco, L. P., Furuya, D. E. G., Furuya, M. T. G., Correa, D. V., Goncalvez, W. N., Junior, J. M., Borges, M., Moraes, M. C. B., Michereff, M. F. F., Aquino, M. F. S., Lauman, R. A., Lisenberg, V., Ramos, A. P. M., & de Castro Jorge, L. A. (2022). An impact analysis of pre-processing techniques in spectroscopy data to classify insect-damaged in soybean plants with machine and deep learning methods. *Infrared Physics & Technology*, *123*, 104203.

Ougiaroglou, S., Papadimitriou, N., & Evangelidis, G. (2024). Reduction Through Homogeneous Clustering: Variations for Categorical Data and Fast Data Reduction. *SN Computer Science*, *5*(6), 671.

Osmolovskaya, N., Shumilina, J., Kim, A., Didio, A., Grishina, T., Bilova, T., Keltsieva, O. A., Zhukov, V., Tikhonovich, I., Tarakhovskaya, E., Frolov, A., & Wessjohann, L. A. (2018). Methodology of drought stress research: Experimental setup and physiological characterization. In *International Journal of Molecular Sciences*, 19(12). MDPI AG.

Ozguven, M. M., & Adem, K. (2019). Automatic detection and classification of leaf spot disease in sugar beet using deep learning algorithms. *Physica A: Statistical Mechanics and its Applications*, *535*, 122537.

Pabinger, S., Rödiger, S., Kriegner, A., Vierlinger, K., & Weinhäusel, A. (2014). A survey of tools for the analysis of quantitative PCR (qPCR) data. In *Biomolecular Detection and Quantification* 1(1), 23–33. Elsevier GmbH.

Padmavathi, K. (2012). Investigation and monitoring for leaves disease detection and evaluation using image processing. *International Research Journal of Engineering Science, Technology and Innovation*, *1*(3), 66-70.

Pahlavan, R., Omid, M., & Akram, A. (2012). Energy input–output analysis and application of artificial neural networks for predicting greenhouse basil production. *Energy*, *37*(1), 171-176.

Pangaribuan, I. F., Rahmawati, N., Akoeb, E. N., Yenni, Y., Sujadi, & Pangaribuan, Y. (2024). Reactions of Drought-Affected Oil Palm Variety Seedlings during the Nursery Phase: Bud and Root Responses. *IOP Conference Series: Earth and Environmental Science*, *1308*(1).

Panwar, S., Duggirala, K. S., Yadav, P., Debnath, N., Yadav, A. K., & Kumar, A. (2023). Advanced diagnostic methods for identification of bacterial foodborne pathogens: Contemporary and upcoming challenges. *Critical Reviews in Biotechnology*, *43*(7), 982-1000.

Paper, C., Rakib, M., Rashid, M., Upper, G., Rot, B. S., Palm, O., Rakib, M., & Rashid, M. (2017). *Sustainable practices in oil palm industry in*. *October*.

Patankar, H. v., Al-Harrasi, I., Al-Yahyai, R., & Yaish, M. W. (2019). Functional

© COPYRIGHT UPM

characterization of date palm aquaporin gene pdpip1;2 confers drought and salinity tolerance to yeast and arabidopsis. *Genes*, *10*(5).

Paterson, R. R. M., & Lima, N. (2017). Climate change affecting oil palm agronomy, and oil palm cultivation increasing climate change, require amelioration. *Ecology and Evolution*, *September*, 1–10.

Pawar, P., Turkar, V., & Patil, P. (2016, August). Cucumber disease detection using artificial neural network. In *2016 International Conference on Inventive Computation Technologies (ICICT)*, 3,1-5. IEEE.

Pearson, L. A., D'Agostino, P. M., & Neilan, B. A. (2021). Recent Developments In Quantitative PCR For Monitoring Harmful Marine Microalgae. *Harmful Algae*, *108*, 102096.

Pinit, S., Chadchawan, S., & Chaiwanon, J. (2020). A simple high-throughput protocol for the extraction and quantification of inorganic phosphate in rice leaves. *Applications in Plant Sciences*, *8*(10).

Pisello, A., De Angelis, S., Ferrari, M., Porreca, M., Vetere, F. P., Behrens, H., Behrens, Maria, C. & Perugini, D. (2022). Visible and near-InfraRed (VNIR) reflectance of silicate glasses: Characterization of a featureless spectrum and implications for planetary geology. *Icarus*, *374*, 114801.

Polson, N., Sokolov, V., & Xu, J. (2021). Deep learning partial least squares. *arXiv preprint arXiv:2106.14085*.

Pradiko, I. (2022). Transpiration Of Oil Palm (Elaeis Guineensis Jacq.) Based On Sap Flow Measurement: The Relation To Soil And Climate Variables. *Journal of Oil Palm Research*.

Pradiko, I., Rahutomo, S., Farrasati, R., Ginting, E. N., Hidayat, F., & Syarovy, M. (2022). Transpiration Of Oil Palm (Elaeis Guineensis Jacq.) Based On Sap Flow Measurement: The Relation To Soil And Climate Variables. *Journal of Oil Palm Research*.

Prasetyo, E. W., Amanah, H. Z., Farras, I., Pahlawan, M. F. R., & Masithoh, R. E. (2024). Detection of Fusarium spp. infection in potato (Solanum tuberosum L.) during postharvest storage through visible–near-infrared and shortwave–near-infrared reflectance spectroscopy. *Open Agriculture*, *9*(1), 20220295.

Purnomo, H., Okarda, B., Dermawan, A., Ilham, Q. P., Pacheco, P., Nurfatriani, F., & Suhendang, E. (2020). Reconciling oil palm economic development and environmental conservation in Indonesia: A value chain dynamic approach. *Forest Policy and Economics*, *111*.

Qaim, M., Sibhatu, K. T., Siregar, H., & Grass, I. (2020). Environmental, economic, and social consequences of the oil palm boom. In *Annual Review of Resource Economics*, 12,321–344. Annual Reviews Inc..

Ranjan, A., Sinha, R., Bala, M., Pareek, A., Singla-Pareek, S. L., & Singh, A. K.

© COPYRIGHT UPM

(2021). Silicon-mediated abiotic and biotic stress mitigation in plants: Underlying mechanisms and potential for stress resilient agriculture. *Plant Physiology and Biochemistry*, *163*, 15–25.

Raj, R., Walker, J. P., Vinod, V., Pingale, R., Naik, B., & Jagarlapudi, A. (2021). Leaf water content estimation using top-of-canopy airborne hyperspectral data. *International Journal of Applied Earth Observation and Geoinformation*, *102*, 102393.

Ramcharan, A., Baranowski, K., McCloskey, P., Ahmed, B., Legg, J., & Hughes, D. P. (2017). Deep learning for image-based cassava disease detection. *Frontiers in plant science*, *8*, 1852.

Ray, P., Reddy, S. S., & Banerjee, T. (2021). Various dimension reduction techniques for high dimensional data analysis: a review. *Artificial Intelligence Review*, *54*(5), 3473-3515.

Ray, P., Reddy, S. S., & Banerjee, T. (2021). Various dimension reduction techniques for high dimensional data analysis: a review. *Artificial Intelligence Review*, *54*(5), 3473-3515.

Rivera, C. A., Palacios-Hernández, E. R., Trejo-Durán, M., Rodríguez-Liñán, M. D. C., Olivera-Reyna, R., & Morales-Saldaña, J. A. (2020). Visible and near-infrared spectroscopy for detection of powdery mildew in Cucurbita pepo L. leaves. *Journal of Applied Remote Sensing*, *14*(4), 044515-044515.

Roeber, V. M., Bajaj, I., Rohde, M., Schmülling, T., & Cortleven, A. (2021). Light acts as a stressor and influences abiotic and biotic stress responses in plants. In *Plant Cell and Environment* 44(3), 645–664. Blackwell Publishing Ltd.

Rodesuchit, A., Suchatgul, S., Klaewklong, B., & Damnoi, S. (2012). Efficacy of fertilizers to control white root disease of rubber caused by Rigidoporus microporus at the early planting stages. *Rubber Thai Journal*, *1*, 62-72.

Rosario, S. F., & Thangadurai, K. (2015). RELIEF: feature selection approach. *International journal of innovative research and development*.

Rostami, M., Berahmand, K., Nasiri, E., & Forouzandeh, S. (2021). Review of swarm intelligence-based feature selection methods. *Engineering Applications of Artificial Intelligence*, *100*, 104210.

Saarela, M., & Jauhiainen, S. (2021). Comparison of feature importance measures as explanations for classification models. *SN Applied Sciences*, *3*(2), 272.

Saidi, N. B., Al-Obaidi, J. R., & Fisol, A. F. B. C. (2023). Rigidoporus microporus and the white root rot disease of rubber. *Forest Pathology*, *53*(1), e12794.

Salgado, F. F., da Silva, T. L. C., Vieira, L. R., Silva, V. N. B., Leão, A. P., Costa, M. M. do C., Togawa, R. C., de Sousa, C. A. F., Grynberg, P., & Souza, M. T. (2022). The early response of oil palm (Elaeis guineensis Jacq.) plants to water deprivation: Expression analysis of miRNAs and their putative target genes,

© COPYRIGHT UPM

and similarities with the response to salinity stress. *Frontiers in Plant Science*, *13*.

Sampian, A. F. M., Hashim, H., Kamal, M. M., Abdullah, N. E., Yussuf, U. R., Khairuzzaman, N. A., & Azmi, A. F. M. (2016, October). Vision system for detection of white root disease infection based on capacitance properties. In *2016 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS)*, 186-189. IEEE.

Samsudin, S. H., Shafri, H. Z., Hamedianfar, A., & Mansor, S. (2015). Spectral Feature Selection And Classification Of Roofing Materials Using Field Spectroscopy Data. *Journal of Applied Remote Sensing*, *9*(1), 095079-095079.

San, E., Paba, G., Cubillán, N., Portillo, E., Casassa-Padrón, A. M., González-González, C., & Guerra, M. (2020). The use of infrared spectroscopy and machine learning tools for detection of Meloidogyne infestations. *Plant Pathology*, *69*(8), 1589-1600.

Sander, F., & Goulson, A. R. (1976). A Rapid Method for Determining Sequences in DNA by Primed Synthesis with DNA Polymerase. In *J. Mol. Bid*, 94.

Sankaran, S., Mishra, A., Ehsani, R., & Davis, C. (2010). A review of advanced techniques for detecting plant diseases. *Computers and Electronics in Agriculture*, *72*(1), 1-13.

Sannakki, S. S., & Rajpurohit, V. S. (2015). Classification of pomegranate diseases based on back propagation neural network. *International Research Journal of Engineering and Technology (IRJET)*, *2*(2).

Sauerbrei, W., Perperoglou, A., Schmid, M., Abrahamowicz, M., Becher, H., Binder, H., ... & TG2 of the STRATOS initiative Michal Abrahamowicz Heiko Becher Harald Binder Daniela Dunkler Frank Harrell Georg Heinze Aris Perperoglou Geraldine Rauch Patrick Royston Willi Sauerbrei. (2020). State of the art in selection of variables and functional forms in multivariable analysis—outstanding issues. *Diagnostic and prognostic research*, *4*, 1-18.

Savitzky, A., & Golay, M. J. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry*, *36*(8), 1627-1639.

Schmidgall, S., Ziaei, R., Achterberg, J., Kirsch, L., Hajiseyedrazi, S., & Eshraghian, J. (2024). Brain-inspired learning in artificial neural networks: a review. *APL Machine Learning*, *2*(2).

Schwartz, G., Eshel, G., & Ben Dor, E. (2011). Reflectance spectroscopy as a tool for monitoring contaminated soils. *Soil Contam*, *6790*.

Sethulekshmi, A. S., Saritha, A., & Joseph, K. (2022). A comprehensive review on the recent advancements in natural rubber nanocomposites. *International Journal of Biological Macromolecules*, *194*, 819-842.

© COPYRIGHT UPM

Shabbir, I., Abd Samad, M. Y., Othman, R., Wong, M. Y., Sulaiman, Z., Jaafar, N. M., & Bukhari, S. A. H. (2021). Evaluation of bioformulation of Enterobacter sp. UPMSSB7 and mycorrhizae with silicon for white root rot disease suppression and growth promotion of rubber seedlings inoculated with Rigidoporus microporus. *Biological control*, *152*, 104467.

Shabbir, I., bin Abd Samad, M. Y., Othman, R., Wong, M. Y., bin Sulaiman, Z., & Bukhari, S. A. H. (2020). White root rot disease suppression in rubber plant with microbial co-inoculants and silicon addition. *Rhizosphere*, 100221.

Shakeel, Q., Bajwa, R. T., Rashid, I., Aslam, H. M. U., Iftikhar, Y., Mubeen, M., Li, G., & Wu, M. (2022). Concept and application of infrared thermography for plant disease measurement. *Trends in Plant Disease Assessment*, 109-125.

Shakil, R., Islam, S., Shohan, Y. A., Mia, A., Rajbongshi, A., Rahman, M. H., & Akter, B. (2023). Addressing agricultural challenges: An identification of best feature selection technique for dragon fruit disease recognition. *Array*, *20*, 100326.

Shams, S. R., Jahani, A., Kalantary, S., Moeinaddini, M., & Khorasani, N. (2021). The evaluation on artificial neural networks (ANN) and multiple linear regressions (MLR) models for predicting SO2 concentration. *Urban Climate*, *37*, 100837.

Sharma, S., Sharma, S., & Athaiya, A. (2017). Activation functions in neural networks. *Towards Data Sci*, *6*(12), 310-316.

Sharma, G., Dwibedi, V., Seth, C. S., Singh, S., Ramamurthy, P. C., Bhadrecha, P., & Singh, J. (2024). Direct and indirect technical guide for the early detection and management of fungal plant diseases. *Current Research in Microbial Sciences*, 100276.

Shearman, J. R., Jantasuriyarat, C., Sangsrakru, D., Yoocha, T., Vannavichit, A., Tragoonrung, S., & Tangphatsornruang, S. (2013). Transcriptome analysis of normal and mantled developing oil palm flower and fruit. *Genomics*, *101*(5), 306–312.

Shi, X., Song, J., Wang, H., Lv, X., Zhu, Y., Zhang, W., Bu, W., & Zeng, L. (2023). Improving soil organic matter estimation accuracy by combining optimal spectral preprocessing and feature selection methods based on pXRF and vis-NIR data fusion. *Geoderma*, *430*, 116301.

Shivaraj, S. M., Sharma, Y., Chaudhary, J., Rajora, N., Sharma, S., Thakral, V., Ram, H., Sonah, H., Singla-Pareek, S. L., Sharma, T. R., & Deshmukh, R. (2021). Dynamic role of aquaporin transport system under drought stress in plants. In *Environmental and Experimental Botany*, 184. Elsevier B.V.

Shobika, G. V., Siddharth, S. P., Vishwa, K. D., Hemachandran, K., & Rege, M. (2024). ANN Model for Analytics. In *Artificial Intelligence and Knowledge Processing* (pp. 195-207). CRC Press.

Shu, B., Zhang, J., Cui, G., Sun, R., Sethuraman, V., Yi, X., & Zhong, G. (2018). Evaluation of reference genes for real-time quantitative PCR analysis in larvae

© COPYRIGHT UPM

of spodoptera litura exposed to azadirachtin stress conditions. *Frontiers in Physiology*, *9*(APR).

Silva, I., & Eugenio Naranjo, J. (2020). A systematic methodology to evaluate prediction models for driving style classification. *Sensors*, *20*(6), 1692.

Simova-Stoilova, L., Vaseva, I., Grigorova, B., Demirevska, K., & Feller, U. (2010). Proteolytic activity and cysteine protease expression in wheat leaves under severe soil drought and recovery. *Plant Physiology and Biochemistry*, *48*(2–3), 200–206.

Singh, M. K., & Chand, S. (2024). Hybrid sigmoid activation function and transfer learning assisted breast cancer classification on histopathological images. *Multimedia Tools and Applications*, *83*(20), 59043-59060.

Singh, V., & Misra, A. K. (2017). Detection of plant leaf diseases using image segmentation and soft computing techniques. *Information processing in Agriculture*, *4*(1), 41-49.

Singha, N. R., Mahapatra, M., Karmakar, M., & Chattopadhyay, P. K. (2019). Processing, characterization and application of natural rubber based environmentally friendly polymer composites. *Sustainable polymer composites and nanocomposites*, 855-897.

Sinha, R., Irulappan, V., Patil, B. S., Reddy, P. C. O., Ramegowda, V., Mohan-Raju, B., Rangappa, K., Singh, H. K., & Senthil-Kumar, M. (2021). Low soil moisture predisposes field-grown chickpea plants to dry root rot disease: evidence from simulation modeling and correlation analysis. *Scientific reports*, *11*(1), 6568.

Sinha, R., Khot, L. R., Rathnayake, A. P., Gao, Z., & Naidu, R. A. (2019). Visible-near infrared spectroradiometry-based detection of grapevine leafroll-associated virus 3 in a red-fruited wine grape cultivar. *Computers and electronics in agriculture*, *162*, 165-173.

Sinnarong, N., Kuson, S., Nunthasen, W., Puphoung, S., & Souvannasouk, V. (2022). The potential risks of climate change and weather index insurance scheme for Thailand's economic crop production. *Environmental Challenges*, *8*.

Siri-Udom, S., Suwannarach, N., & Lumyong, S. (2017). Applications of volatile compounds acquired from Muscodor heveae against white root rot disease in rubber trees (Hevea brasiliensis Müll. Arg.) and relevant allelopathy effects. *Fungal biology*, *121*(6-7), 573-581.

Slatko, B. E., Gardner, A. F., & Ausubel, F. M. (2018). Overview of Next-Generation Sequencing Technologies. *Current Protocols in Molecular Biology*, *122*(1).

Somwong, S., & Chongcheawchamnan, M. (2018). A portable system for rapid measurement of dry rubber content with contaminant detection feature. *IEEE Sensors Journal*, *18*(20), 8329-8337.

© COPYRIGHT UPM

Song, B., Li, Y., Yang, L., Shi, H., Li, L., Bai, W., & Zhao, Y. (2023). Soil acidification under long-term N addition decreases the diversity of soil bacteria and fungi and changes their community composition in a semiarid grassland. *Microbial ecology*, *85*(1), 221-231.

Sorochan Armstrong, M. D., de la Mata, A. P., & Harynuk, J. J. (2022). Review of variable selection methods for discriminant-type problems in chemometrics. *Frontiers in Analytical Science*, *2*, 867938.

Soytong, K., & Kaewchai, S. (2014). Biological control of white root of rubber trees using Chaetomium cupreum. *J Agric Tech*, *10*(1), 93-103.

Srivastava, K. R., Awasthi, S., Mishra, P. K., & Srivastava, P. K. (2020). Biosensors/molecular tools for detection of waterborne pathogens. *Waterborne Pathogens*, 237-277.

Stalpers, J. A. (1978). Identification of wood-inhabiting fungi in pure culture. *Studies in mycology*, (16).

Stamford, J., Aciksoz, S. B., & Lawson, T. (2024). Remote sensing techniques: hyperspectral imaging and data analysis. In *Photosynthesis: Methods and Protocols* (pp. 373-390). New York, NY: Springer US.

Stark, R., Grzelak, M., & Hadfield, J. (2019). RNA sequencing: the teenage years. In *Nature Reviews Genetics*, 20(11), 631–656. Nature Publishing Group.

Subramaniam, V., Loh, S. K., & Aziz, A. A. (2021). GHG analysis of the production of crude palm oil considering the conversion of agricultural wastes to by-products. *Sustainable Production and Consumption*, *28*, 1552–1564.

Sujatha, R., Chatterjee, J. M., Jhanjhi, N. Z., & Brohi, S. N. (2021). Performance of deep learning vs machine learning in plant leaf disease detection. *Microprocessors and Microsystems*, *80*, 103615.

Sulaiman, M. S. (2019). *White root disease auto-detection system for rubber trees based on dynamic electro-biochemical latex properties* (Doctoral dissertation, Universiti Teknologi MARA).

Sulaiman, M. S., Madzhi, N. K., Sajari, R., & Abdullah, N. E. (2017). Statistical discrimination of latex between healthy and white root infected Hevea brasiliensis tree based on protein concentration. *Journal of Fundamental and Applied Sciences*, *9*(5S), 818-829.

Sulaiman, M. S., & Saad, Z. (2020). Classification of healthy and white root disease infected rubber trees based on relative permittivity and capacitance input properties using LM and SCG artificial neural network. *Indonesian Journal of Electrical Engineering and Computer Science*, *19*(1), 222-228.

Suman, T., & Dhruvakumar, T. (2015). Classification of paddy leaf diseases using shape and color features. *IJEEE*, *7*(01), 239-250.

© COPYRIGHT UPM

Syafaah, A. (2019). *Screening of rubber rootstocks for the white root disease tolerance with defense-related genes and enzymes* (Doctoral dissertation, Prince of Songkla University).

Sytykiewicz, H., Łukasik, I., Goławska, S., & Chrzanowski, G. (2019). Aphid-triggered changes in oxidative damage markers of nucleic acids, proteins, and lipids in maize (Zea mays L.) seedlings. *International Journal of Molecular Sciences*, *20*(15).

Tahmasebi, A., Ashrafi-Dehkordi, E., Shahriari, A. G., Mazloomi, S. M., & Ebrahimie, E. (2019). Integrative meta-analysis of transcriptomic responses to abiotic stress in cotton. *Progress in Biophysics and Molecular Biology*, *146*, 112–122.

Tamburini, E., Ferrari, G., Marchetti, M. G., Pedrini, P., & Ferro, S. (2015). Development of FT-NIR models for the simultaneous estimation of chlorophyll and nitrogen content in fresh apple (Malus domestica) leaves. *Sensors*, *15*(2), 2662-2679.

Tang, L., Gao, S., Wang, W., Xiong, X., Han, W., & Li, X. (2023). Moisture Content Detection of Tomato Leaves Based on Electrical Impedance Spectroscopy. *Communications in Soil Science and Plant Analysis*, 1-15.

Taylor, S. C., Nadeau, K., Abbasi, M., Lachance, C., Nguyen, M., & Fenrich, J. (2019). The Ultimate qPCR Experiment: Producing Publication Quality, Reproducible Data the First Time. In *Trends in Biotechnology*, 37(7),761–774. Elsevier Ltd.

Thorat, A., Kumari, S., & Valakunde, N. D. (2017, December). An IoT based smart solution for leaf disease detection. In *2017 International Conference on Big Data, IoT and Data Science (BID)*, 193-198. IEEE.

Tezara, W., Domínguez, T. S. T., Loyaga, D. W., Ortiz, R. N., Chila, V. H. R., & Ortega, M. J. B. (2021). Photosynthetic activity of oil palm (Elaeis guineensis) and interspecific hybrid genotypes (Elaeis oleifera × Elaeis guineensis), and response of hybrids to water deficit. *Scientia Horticulturae*, *287*..

Tian, L., Wang, Z., Xue, B., Li, D., Zheng, H., Yao, X., Cao, W., & Cheng, T. (2023). A disease-specific spectral index tracks Magnaporthe oryzae infection in paddy rice from ground to space. *Remote sensing of environment*, *285*, 113384.

Tian, L., Xue, B., Wang, Z., Li, D., Yao, X., Cao, Q., Zhu, Y., Cao, W., & Cheng, T. (2021). Spectroscopic detection of rice leaf blast infection from asymptomatic to mild stages with integrated machine learning and feature selection. *Remote sensing of environment*, *257*, 112350.

Tillekeratne, L. M. K., Karunanayake, L., Kumara, P. S., & Weeraman, S. (1988). A rapid and accurate method for determining the dry rubber content and total solid content of NR latex. *Polymer Testing*, *8*(5), 353-358.

© COPYRIGHT UPM

Tsagkaris, A. S., Bechynska, K., Ntakoulas, D. D., Pasias, I. N., Weller, P., Proestos, C., & Hajslova, J. (2023). Investigating the impact of spectral data pre-processing to assess honey botanical origin through Fourier transform infrared spectroscopy (FTIR). *Journal of Food Composition and Analysis*, *119*, 105276.

Tsai, M. H., & Li, K. T. (2021). Leaf color as a morpho-physiological index for screening heat tolerance and improved water use efficiency in rabbiteye blueberry (Vaccinium virgatum Aiton). *Scientia Horticulturae*, *278*, 109864.

Vainberg, I. E., Lewis, S. A., Rommelaere, H., Ampe, C., Vandekerckhove, J., Klein, H. L., & Cowan, N. J. (1998). Prefoldin, a Chaperone that Delivers Unfolded Proteins to Cytosolic Chaperonin *et al.* In *‡ and Hartl*, 93. Lorimer.

Vapnik, V. N. (1999). An overview of statistical learning theory. *IEEE transactions on neural networks*, *10*(5), 988-999.

Verma, N. V., Shukla, M., Kulkarni, R., Srivastava, K., Claudic, B., Savara, J., ... & Pandya, A. (2022). Emerging extraction and diagnostic tools for detection of plant pathogens: recent trends, challenges, and future scope. *ACS Agricultural Science & Technology*, *2*(5), 858-881.

Vincke, D., Eylenbosch, D., Jacquemin, G., Chandelier, A., Pierna, J. A. F., Stevens, F., ... & Vermeulen, P. (2023). Near infrared hyperspectral imaging method to assess Fusarium Head Blight infection on winter wheat ears. *Microchemical Journal*, *191*, 108812.

Vinutha, M. R., Chandrika, J., Krishnan, B., & Kokatnoor, S. A. (2023). EPCA—enhanced principal component analysis for medical data dimensionality reduction. *SN Computer Science*, *4*(3), 243.

Vogado, N. O., Liddell, M. J., Laurance, S. G. W., Campbell, M. J., Cheesman, A. W., Engert, J. E., Palma, A. C., Ishida, F. Y., & Cernusak, L. A. (2020). The effects of an experimental drought on the ecophysiology and fruiting phenology of a tropical rainforest palm. *Journal of Plant Ecology*, *13*(6), 744–753.

Wahyuni, M., Simanjuntak, J. H., & Sitompul, I. O. (2018). Efektivitas fungisida berbahan aktif heksakonazol terhadap penyakit jamur akar putih bibit tanaman karet (Hevea brasiliensis). *Agrotekma: Jurnal Agroteknologi dan Ilmu Pertanian*, *3*(1), 1-10.

Wan, X., Li, G., Zhang, M., Yan, W., He, G., Awelisah, Y. M., & Lin, L. (2020). A review on the strategies for reducing the non-linearity caused by scattering on spectrochemical quantitative analysis of complex solutions. *Applied Spectroscopy Reviews*, *55*(5), 351-377.

Wang, B., Kumar, V., Olson, A., & Ware, D. (2019). Reviving the transcriptome studies: An insight into the emergence of single-molecule transcriptome sequencing. In *Frontiers in Genetics*, l0(ssue APR). Frontiers Media S.A.

Wang, C., Reis, M. G., Waterhouse, G. I., Hemar, Y., & Reis, M. M. (2021). Prediction of dairy powder functionality attributes using diffuse reflectance in

© COPYRIGHT UPM

the visible and near infrared (Vis-NIR) region. *International Dairy Journal*, *117*, 104981.

Wang, G., Sun, Y., & Wang, J. (2017). Automatic image-based plant disease severity estimation using deep learning. *Computational intelligence and neuroscience*, *2017*.

Wang, H., Chu, X., Chen, P., Li, J., Liu, D., & Xu, Y. (2022). Partial least squares regression residual extreme learning machine (PLSRR-ELM) calibration algorithm applied in fast determination of gasoline octane number with near-infrared spectroscopy. *Fuel*, *309*, 122224.

Wang, L., Lee, M., Ye, B., & Yue, G. H. (2020). Genes, pathways and networks responding to drought stress in oil palm roots. *Scientific Reports*, *10*(1).

Wang, Y., Huang, H., Rudin, C., & Shaposhnik, Y. (2021). Understanding how dimension reduction tools work: an empirical approach to deciphering t-SNE, UMAP, TriMAP, and PaCMAP for data visualization. *Journal of Machine Learning Research*, *22*(201), 1-73.

Wang, Y., Htwe, Y. M., Ihase, L. O., Shi, P., Cao, H., & Lei, X. (2018). Pollen germination genes differentially expressed in different pollens from Dura, Pisifera and Tenera oil palm (Elaeis guineensis jacq.). *Scientia Horticulturae*, *235*, 32–38.

Wastie, R. L. (1975). Diseases of rubber and their control. *PANS Pest Articles & News Summaries*, *21*(3), 268-288.

Watanabe, T. (2010). *Pictorial atlas of soil and seed fungi: morphologies of cultured fungi and key to species*. CRC press.

Watt, M. S., Bartlett, M., Soewarto, J., de Silva, D., Estarija, H. J. C., Massam, P., Cajes, D., Yorston, E. G., Dobbie, K., Fraser, S., Dungey, H. S., & Buddenbaum, H. (2023). Previsual and early detection of myrtle rust on rose apple using indices derived from thermal imagery and visible-to-short-infrared spectroscopy. *Phytopathology®*, *113*(8), 1405-1416.

Wei, X., Johnson, M. A., Langston Jr, D. B., Mehl, H. L., & Li, S. (2021). Identifying optimal wavelengths as disease signatures using hyperspectral sensor and machine learning. *Remote Sensing*, *13*(14), 2833.

Wu, F., Chen, J. J., Ji, X. H., Vlasák, J., & Dai, Y. C. (2017). Phylogeny and diversity of the morphologically similar polypore genera Rigidoporus, Physisporinus, Oxyporus, and Leucophellinus. *Mycologia*, *109*(5), 749-765.

Wu, G., Fang, Y., Jiang, Q., Cui, M., Li, N., Ou, Y., Diao, Z., & Zhang, B. (2023). Early identification of strawberry leaves disease utilizing hyperspectral imaging combing with spectral features, multiple vegetation indices and textural features. *Computers and Electronics in Agriculture*, *204*, 107553.

© COPYRIGHT UPM

Wu, Q., & Xu, H. (2020). Design and development of an on-line fluorescence spectroscopy system for detection of aflatoxin in pistachio nuts. *Postharvest Biology and Technology*, *159*, 111016.

Xie, S., Ding, F., Chen, S., Wang, X., Li, Y., & Ma, K. (2022). Prediction of soil organic matter content based on characteristic band selection method. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *273*, 120949.

Xiong, D., Chen, J., Yu, T., Gao, W., Ling, X., Li, Y., Peng, S., & Huang, J. (2015). SPAD-based leaf nitrogen estimation is impacted by environmental factors and crop leaf characteristics. *Scientific reports*, *5*(1), 1-12.

Xu, G., Li, F., Luo, Y., Xie, J., & Guo, X. (2020). Soil total nitrogen estimation of alpine grassland using visible/near-infrared spectra: a comparison of multivariate techniques with different spectral transformations. *Journal of Applied Remote Sensing*, *14*(1), 015501-015501.

Xu, H. R., Ying, Y. B., Fu, X. P., & Zhu, S. P. (2007). Near-infrared spectroscopy in detecting leaf minor damage on tomato leaf. *Biosystems Engineering*, *96*(4), 447-454.

Xu, P., Wu, G., Guo, Y., Yang, H., & Zhang, R. (2017). Automatic wheat leaf rust detection and grading diagnosis via embedded image processing system. *Procedia Computer Science*, *107*, 836-841.

Xu, S., & Yu, Z. (2020). Environmental control on transpiration: A case study of a desert ecosystem in northwest china. *Water (Switzerland)*, *12*(4). https://doi.org/10.3390/W12041211.

Xuan, G., Li, Q., Shao, Y., & Shi, Y. (2022). Early diagnosis and pathogenesis monitoring of wheat powdery mildew caused by blumeria graminis using hyperspectral imaging. *Computers and Electronics in Agriculture*, *197*, 106921.

Yadav, N., Garg, V. K., Chhillar, A. K., & Rana, J. S. (2023). Recent advances in nanotechnology for the improvement of conventional agricultural systems: A review. *Plant Nano Biology*, *4*, 100032.

Yang, K., Wang, H., Luo, L., Zhu, S., Huang, H., Wei, Z., Zhu, Y., Guo, L., & He, X. (2023). Effects of different soil moisture on the growth, quality, and root rot disease of organic Panax notoginseng cultivated under pine forests. *Journal of Environmental Management*, *329*, 117069.

Yang, X., Li, J., Ji, C., Wei, Z., Zhao, T., & Pang, Q. (2022). Overexpression of an aquaporin gene EsPIP1;4 enhances abiotic stress tolerance and promotes flowering in Arabidopsis thaliana. *Plant Physiology and Biochemistry*, *193*, 25–35.

Yang, X., Lu, M., Wang, Y., Wang, Y., Liu, Z., & Chen, S. (2021). Response mechanism of plants to drought stress. In *Horticulturae* (Vol. 7, Issue 3).

© COPYRIGHT UPM

MDPI AG.

Yono, D., Purwanti, E., Sahara, A., Nugroho, Y. A., Tanjung, Z. A., Aditama, R., Dewi, C., Sihotang, A. E., Utomo, C., & Liwang, T. (2019). Physiology and genotyping of adaptive and sensitive oil palm progenies under unwatered stress condition. *IOP Conference Series: Earth and Environmental Science*, *293*(1).

Youssefian, L., Saeidian, A. H., Palizban, F., Bagherieh, A., Abdollahimajd, F., Sotoudeh, S., Mozafari, N., Farahani, R. A., Mahmoudi, H., Babashah, S., Zabihi, M., Zeinali, S., Fortina, P., Salas-Alanis, J. C., South, A. P., Vahidnezhad, H., & Uitto, J. (2021). Whole-Transcriptome Analysis by RNA Sequencing for Genetic Diagnosis of Mendelian Skin Disorders in the Context of Consanguinity. *Clinical Chemistry*, *67*(6), 876–888.

Yu, D., Zhang, J., Tan, G., Yu, N., Wang, Q., Duan, Q., Qi, X., Cheng, M., Yan, C., Wei, Z., Yu, Z., Huang, W., & Li, C. (2019). An easily-performed high-throughput method for plant genomic DNA extraction. *Analytical Biochemistry*, *569*, 28–30.

Zahir, S. A. D. M., Jamlos, M. F., Omar, A. F., Jamlos, M. A., Mamat, R., Muncan, J., & Tsenkova, R. (2024). Review–Plant nutritional status analysis employing the visible and near-infrared spectroscopy spectral sensor. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *304*, 123273.

Zahir, S. A. D. M., Omar, A. F., Jamlos, M. F., Azmi, M. A. M., & Muncan, J. (2022). A review of visible and near-infrared (Vis-NIR) spectroscopy application in plant stress detection. *Sensors and Actuators A: Physical*, *338*, 113468.

Zandalinas, S. I., Mittler, R., Balfagón, D., Arbona, V., & Gómez-Cadenas, A. (2018). Plant adaptations to the combination of drought and high temperatures. In *Physiologia Plantarum*,162, (1), 2–12.

Zang, Z., Xu, Y., Lu, L., Geng, Y., Yang, S., & Li, S. Z. (2023). Udrn: unified dimensional reduction neural network for feature selection and feature projection. *Neural Networks*, *161*, 626-637.

Zare, S., Shamsi, S. R. F., & Abtahi, S. A. (2019). Weakly-coupled geo-statistical mapping of soil salinity to Stepwise Multiple Linear Regression of MODIS spectral image products. *Journal of African Earth Sciences*, *152*, 101-114.

Zebari, R., Abdulazeez, A., Zeebaree, D., Zebari, D., & Saeed, J. (2020). A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction. *Journal of Applied Science and Technology Trends*, *1*(1), 56-70.

Zeng, Y., Chen, M., Hao, D., Damm, A., Badgley, G., Rascher, U., Johnson, J. E., Dechant, B., Siegmann, B., Ryu, Y., Qui, H., Krieger, V., Panigada, C., Celesti, M., Miglietta, F., Yang, Xi., & Berry, J. A. (2022). Combining near-infrared radiance of vegetation and fluorescence spectroscopy to detect effects of abiotic changes and stresses. *Remote Sensing of Environment*, *270*, 112856.

© COPYRIGHT UPM

Zhang, A., Jin, L., Yarra, R., Cao, H., Chen, P., & John Martin, J. J. (2022). Transcriptome analysis reveals key developmental and metabolic regulatory aspects of oil palm (Elaeis guineensis Jacq.) during zygotic embryo development. *BMC Plant Biology*, *22*(1).

Zhang, C., Feng, X., Wang, J., Liu, F., He, Y., & Zhou, W. (2017). Mid-infrared spectroscopy combined with chemometrics to detect Sclerotinia stem rot on oilseed rape (Brassica napus L.) leaves. *Plant Methods*, *13*, 1-9.

Zhang, F., & Zhou, G. (2019). Estimation of vegetation water content using hyperspectral vegetation indices: A comparison of crop water indicators in response to water stress treatments for summer maize. *BMC ecology*, *19*, 1-12.

Zhang, J., Huang, Y., Pu, R., Gonzalez-Moreno, P., Yuan, L., Wu, K., & Huang, W. (2019). Monitoring plant diseases and pests through remote sensing technology: A review. *Computers and Electronics in Agriculture*, *165*, 104943.

Zhang, J., Sui, C., Liu, H., Chen, J., Han, Z., Yan, Q., Liu, S., & Liu, H. (2021). Effect of chlorophyll biosynthesis-related genes on the leaf color in Hosta (Hosta plantaginea Aschers) and tobacco (Nicotiana tabacum L.). *BMC Plant Biology*, *21*(1), 1-14.

Zhang, K., Zhang, Y., Chen, G., & Tian, J. (2009). Genetic analysis of grain yield and leaf chlorophyll content in common wheat. *Cereal research communications*, *37*(4), 499-511.

Zhang, Y., & Wang, Y. (2023). Recent trends of machine learning applied to multi-source data of medicinal plants. *Journal of Pharmaceutical Analysis*.

Zhao, T., Hu, Y., & Cheng, L. (2021). Deep-DRM: a computational method for identifying disease-related metabolites based on graph deep learning approaches. *Briefings in Bioinformatics*, *22*(4), bbaa212.

Zhao, X. Y., Liu, G. Y., Sui, Y. T., Xu, M., & Tong, L. (2021). Denoising method for Raman spectra with low signal-to-noise ratio based on feature extraction. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *250*, 119374.

Zhao, Z. M., Jin, X. D., Zhang, L., & Yu, X. L. (2010). A novel measurement system for dry rubber content in concentrated natural latex based on annular photoelectric sensor. *International Journal of Physical Sciences*, *5*(3), 251-260.

Zhao, X., Liu, Y., Huang, Z., Li, G., Zhang, Z., He, X., Du, H., Wang, M., & Li, Z. (2024). Early diagnosis of Cladosporium fulvum in greenhouse tomato plants based on visible/near-infrared (VIS/NIR) and near-infrared (NIR) data fusion. *Scientific Reports*, *14*(1), 20176.

Zheng, Q., Huang, W., Cui, X., Dong, Y., Shi, Y., Ma, H., & Liu, L. (2018). Identification of wheat yellow rust using optimal three-band spectral indices in different growth stages. *Sensors*, *19*(1), 35.

© COPYRIGHT UPM

Zhou, L., Yarra, R., Jin, L., & Cao, H. (2020). Genome-wide identification and expression analysis of MYB gene family in oil palm (Elaeis guineensis Jacq.) under abiotic stress conditions. *Environmental and Experimental Botany*, *180*.

Zhou, Quan, Wenqian Huang, and Xi Tian. (2020). Feature wavelength selection based on the combination of image and Spectrum for aflatoxin B1 concentration classification in single maize kernels. *Agriculture* 12, no. 3 (2022): 385.

Zorkafli, M. F., Osman, M. K., Isa, I. S., Ahmad, F., & Sulaiman, S. N. (2019). Classification of cervical cancer using hybrid multi-layered perceptron network trained by genetic algorithm. *Procedia Computer Science*, *163*, 494-501.

© COPYRIGHT UPM

APPENDICES

APPENDIX A

DNA Sequencing Process

DNA Extraction

Mycelium Preparation

Hyphal tip isolation is done to grow the fungal culture on PDA, and the mycelium is scrapped off aseptically after 10 days of incubation at 25°C.

Cell Lysis and CTAB Buffer

The mycelium is ground using liquid nitrogen into a fine powder. The powder is suspended in 600 µl of CTAB buffer, which helps lyse the cell and release the DNA. The mixture is incubated at 65°C for 30 minutes, with vortexing every 10 minutes to enhance cell disruption.

Extraction with Chlorofom Alcohol (CIA)

After incubation, 600 µl of CIA is added to remove proteins and other impurities. The mixture is centrifuged at 7000 rpm for 5 minutes at 40°C, separating it into two layers. The aqueous layer containing the DNA is collected, and the process is repeated.

DNA Precipitation with Isopropanol

The DNA is precipitated by adding 300 µl of isopropanol and incubating at room temperature for 30 minutes. The mixture is centrifuged at 12,000 rpm for 10 minutes to pellet the DNA. The supernatant is decanted, and the DNA pellet is washed and resuspended in 40 µl of ddH_2O.

© COPYRIGHT UPM

APPENDIX B

DNA Sequencing Results

1. Raw sequenced nucleotide data

```
CTACCTGATTTGAGGTCAGTATCAAAAGTGTAAAGGCCACTTTGAAAAAGAGGCCAGAGATTAGAA
GCCGGCAAAAACACAGGCACCAACAGTTAGAACGGCGATGAAGTGAGACTATCACACCGTGAACCG
TCGGGCAAACTACGTGACCGAGCTAATGCATTCAAGGAGAGTCTAACCTCTTTCATTATAGAAAGAG
ACACGACAAGCTCCCAAATCCAATCCACAACAAAAAAATTGAAAATGGAAATTACCCCAACCTCCAA
CCGGGCTGGTCCTCCGAATACCCAGGGACCCCAGGGGGCTTCCAAAAATCCAAGAATCCCTGAATTC
TGCCATTCCCATTACTTATCCCCTTTCCCTGCCTTCTTCCTCCATGCCAAAACCCAAAAAACCCTTGGTG
AAAATTGGATTTAACCTTATAATTAACCCACCAATACCAAAAAATTTTAAGTGGGTTGGAATGATTAA
AAAAAGANNAAGGCCAAAAAAAGGTCCCCCCTCTTGAAATGCCCCAATGGGGGTTGAAATGGATTG
GCAAGGCTGCCCCTGTTCTCCCCTGGTAAAAAAAAAACCCACCACCACCCCCAACCCCATTCCTTAAT
GATCCTTCCCCCAGGTCCCCTACCGAAACCTTGGTACCATTTTTTTACTTTCCAA
```

2. Clean nucleotide sequence used for data analysis

```
CTACCTGATTTGAGGTCAGTATCAAAAGTGTAAAGGCCACTTTGAAAAAGAGGCCAGAGATTAGAA
GCCGGCAAAAACACAGGCACCAACAGTTAGAACGGCGATGAAGTGAGACTATCACACCGTGAACCG
TCGGGCAAACTACGTGACCGAGCTAATGCATTCAAGGAGAGTCTAACCTCTTTCATTATAGAAAGAG
ACACGACAAGCTCCCAAATCCAATCCACAACAAAAAAATTGAAAATGGAAATTACCCCAACCTCCAA
CCGGGCTGGTCCTCCGAATACCCAGGGACCCCAGGGGGCTTCC
```

4. Confirmation of similarity by looking at the data of reference pathogen MF574815.1

```
ACCESSION   MF574815
VERSION     MF574815.1
KEYWORDS    .
SOURCE      Rigidoporus microporus (Fomes lignosus)
  ORGANISM  Rigidoporus microporus
            Eukaryota; Fungi; Dikarya; Basidiomycota; Agaricomycotina;
            Agaricomycetes; Polyporales; Meripilaceae; Rigidoporus.
REFERENCE   1  (bases 1 to 486)
  AUTHORS   Goh,Y.K., Marzuki,N.F., Liew,Y.A., Cheng,C.R., Tan,S.S., Tey,S.H.,
            Goh,Y.K. and Goh,K.J.
  TITLE     Biological control of Rigidoporus microsporus with Cladobotryum
            semicirculare
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 486)
  AUTHORS   Goh,Y.K., Marzuki,N.F., Liew,Y.A., Cheng,C.R., Tan,S.S., Tey,S.H.,
            Goh,Y.K. and Goh,K.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JUL-2017) Pest and Diseases, Advanced Agriecological
            Research Sdn Bhd, No. 11 Jalan Teknologi 3/6, Taman Sains Selangor,
            Kota Damansara, Petaling Jaya, Selangor 47810, Malaysia
COMMENT     ##Assembly-Data-START##
            Sequencing Technology :: Sanger dideoxy sequencing
            ##Assembly-Data-END##
FEATURES             Location/Qualifiers
     source          1..486
                     /organism="Rigidoporus microporus"
                     /mol_type="genomic DNA"
                     /isolate="AA0001"
                     /host="Hevea brasiliensis (rubber)"
                     /db_xref="taxon:219653"
                     /country="Malaysia"
                     /collected_by="LGM"
```

3. Blast result

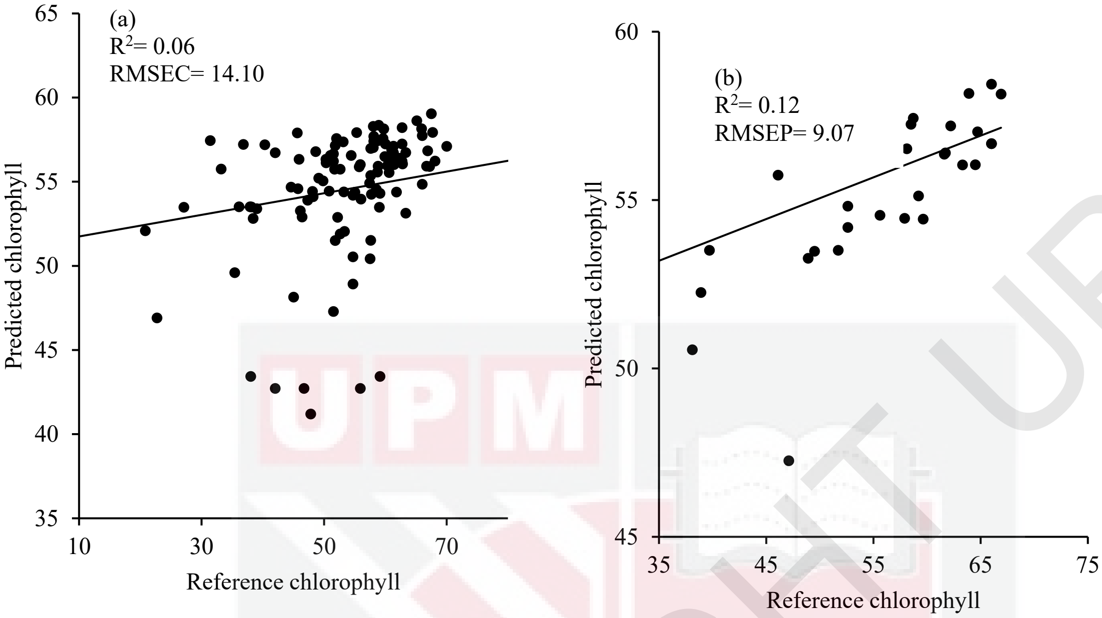
Sequences producing significant alignments Download Select columns Show 100

select all *100 sequences selected* GenBank Graphics Distance tree of results MSA Viewer

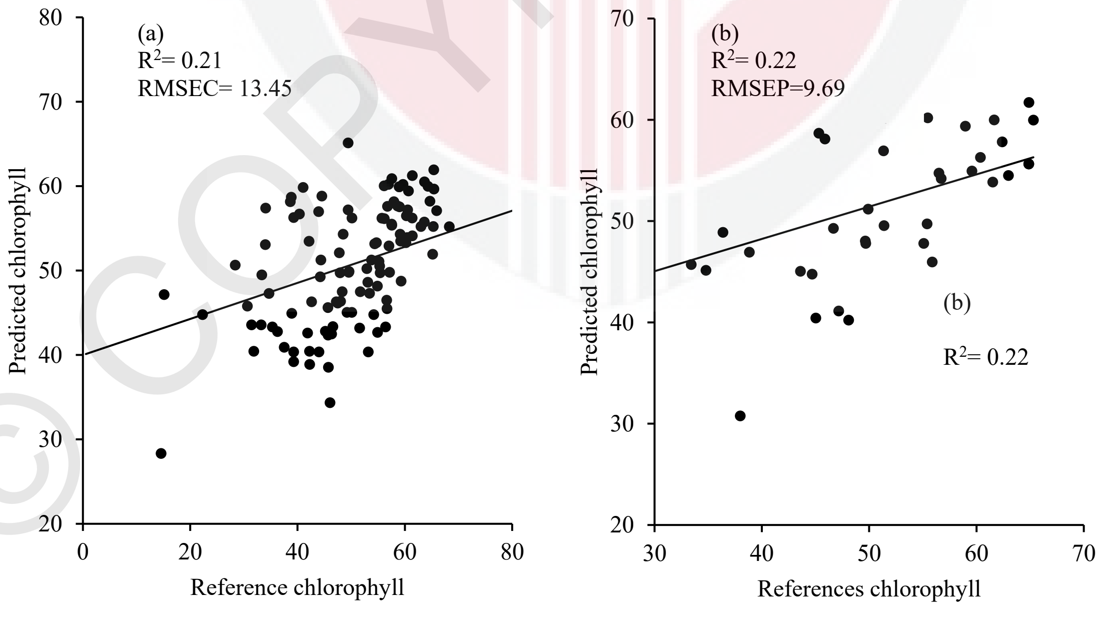
Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
Rigidoporus microporus isolate AA0001 internal transcribed spacer 1, partial sequence; 5.8S ribosomal RNA gene ...	Rigidoporus micr...	449	449	97%	2e-121	94.52%	486	MF574815.1
Abundisporus roseoalbus voucher Dai 12269 18S ribosomal RNA gene, partial sequence; internal transcribed spac...	Abundisporus ros...	449	449	97%	2e-121	94.50%	660	KC415908.1
Rigidoporus microporus isolate RL21 internal transcribed spacer 1, partial sequence; 5.8S ribosomal RNA gene an...	Rigidoporus micr...	446	446	97%	2e-120	94.20%	679	MN103602.1
Rigidoporus microporus culture CBS:173.33 strain CBS 173.33 small subunit ribosomal RNA gene, partial sequenc...	Rigidoporus micr...	446	446	97%	2e-120	94.20%	666	MH855397.1
Rigidoporus microporus isolate SEG small subunit ribosomal RNA gene, partial sequence; internal transcribed spa...	Rigidoporus micr...	446	446	97%	2e-120	94.20%	669	MG199653.1
Rigidoporus microporus voucher SL1059 internal transcribed spacer 1, partial sequence; 5.8S ribosomal RNA gene...	Rigidoporus micr...	446	446	97%	2e-120	94.20%	557	OR636143.1
Rigidoporus microporus strain RL 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ri...	Rigidoporus micr...	446	446	97%	2e-120	94.20%	669	KM246744.1
Rigidoporus microporus isolate VM small subunit ribosomal RNA gene, partial sequence; internal transcribed space...	Rigidoporus micr...	446	446	97%	2e-120	94.20%	674	MZ453082.1
Rigidoporus microporus isolate R7 small subunit ribosomal RNA gene, partial sequence; internal transcribed space...	Rigidoporus micr...	446	446	97%	2e-120	94.20%	674	MZ453079.1
Rigidoporus microporus isolate R4 small subunit ribosomal RNA gene, partial sequence; internal transcribed space...	Rigidoporus micr...	446	446	97%	2e-120	94.20%	669	MZ453078.1
Rigidoporus microporus isolate RM3 internal transcribed spacer 1, partial sequence; 5.8S ribosomal RNA gene and...	Rigidoporus micr...	446	446	97%	2e-120	94.20%	568	MZ453077.1
Rigidoporus microporus isolate R10 small subunit ribosomal RNA gene, partial sequence; internal transcribed spac...	Rigidoporus micr...	446	446	97%	2e-120	94.20%	673	MZ453074.1
Rigidoporus microporus isolate FRIM642 internal transcribed spacer 1, partial sequence; 5.8S ribosomal RNA gene...	Rigidoporus micr...	446	446	97%	2e-120	94.20%	650	HQ400708.1

APPENDIX C

PLS Results

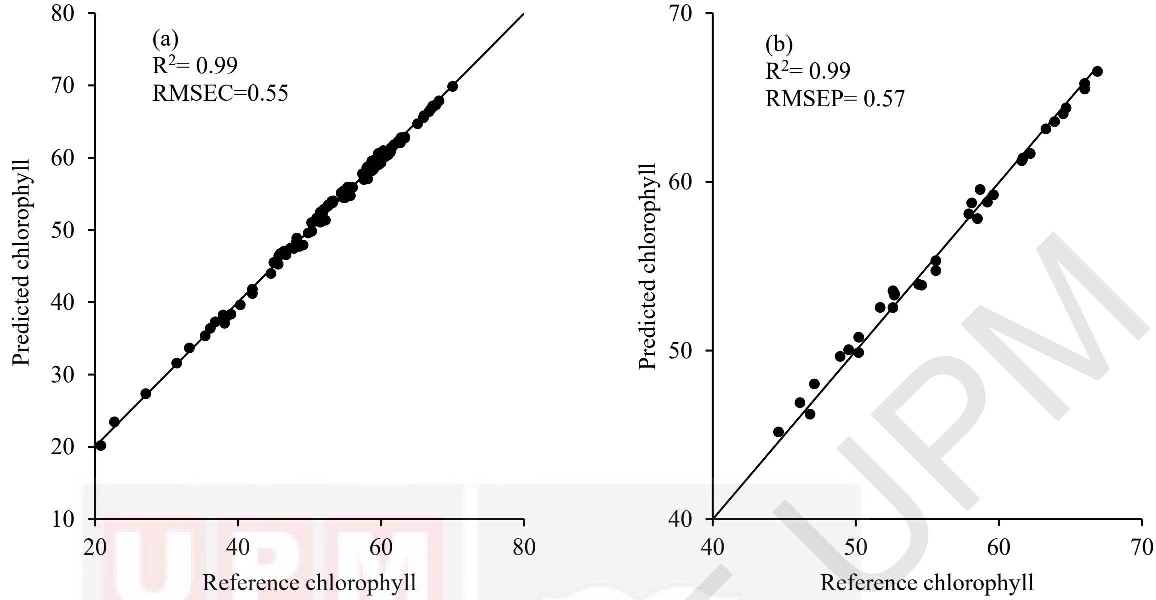


Scatter plots of chlorophyll content for ANOVA FS using raw data reflectance of 450 –699 nm spectral data; (a) calibration (b) prediction model

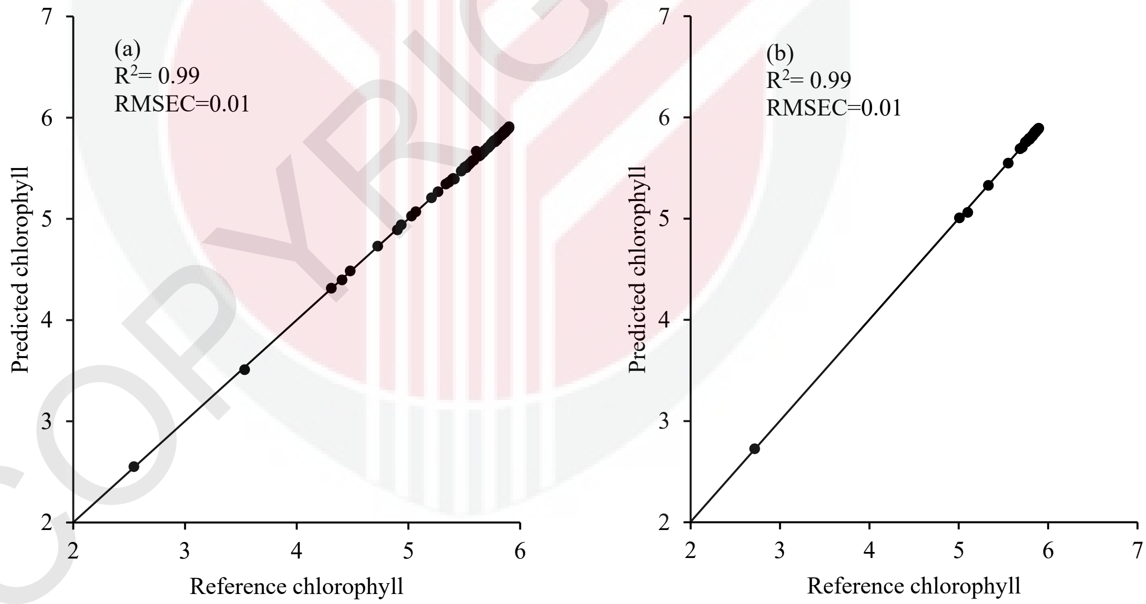


Scatter plots of chlorophyll content for ANOVA FS using BOC pre-processing data reflectance of 450 – 699 nm spectral data; (a) calibration (b) prediction model.

© COPYRIGHT UPM

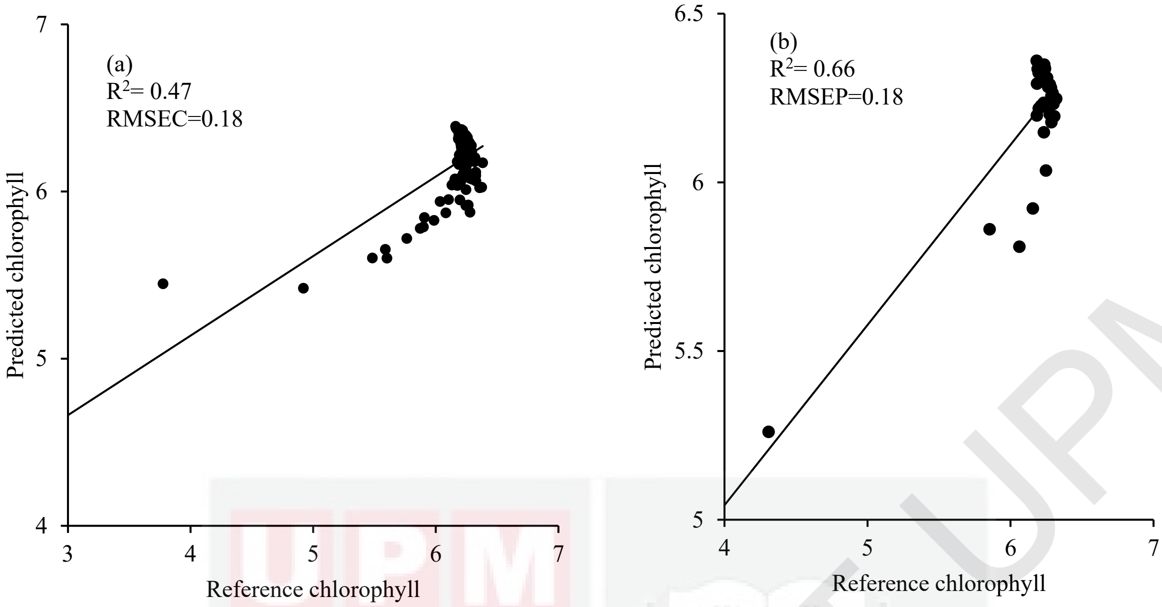


Scatter plots of chlorophyll content for ANOVA FS using SG pre-processing data reflectance of 450 – 699 nm spectral data; (a) calibration (b) prediction model

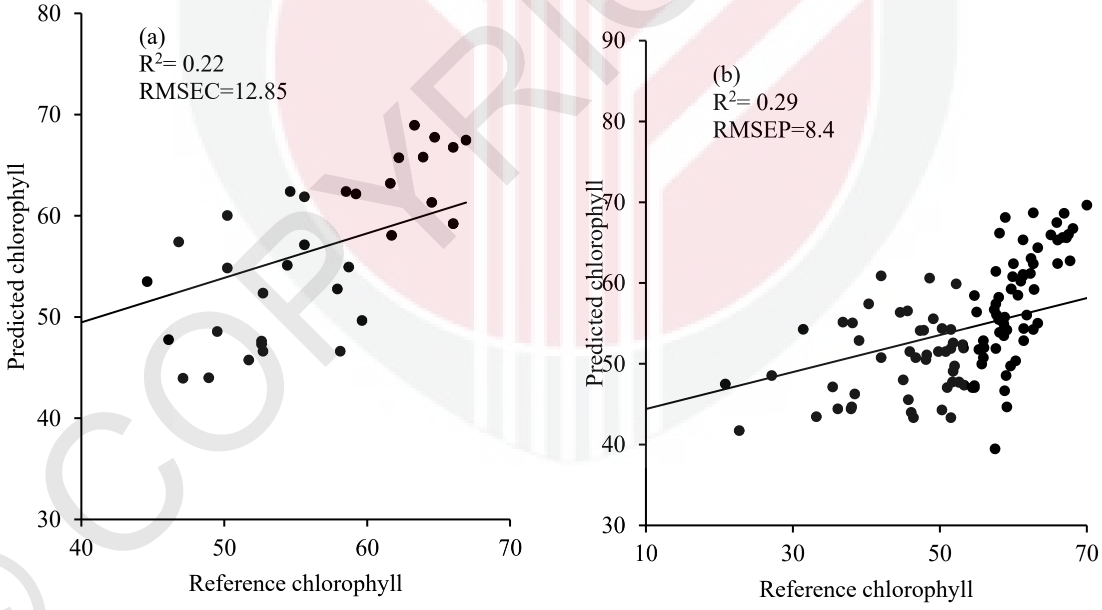


Scatter plots of chlorophyll content for ANOVA FS using SNV pre-processing data reflectance of 450 – 699 nm spectral data; (a) calibration (b) prediction model

© COPYRIGHT UPM

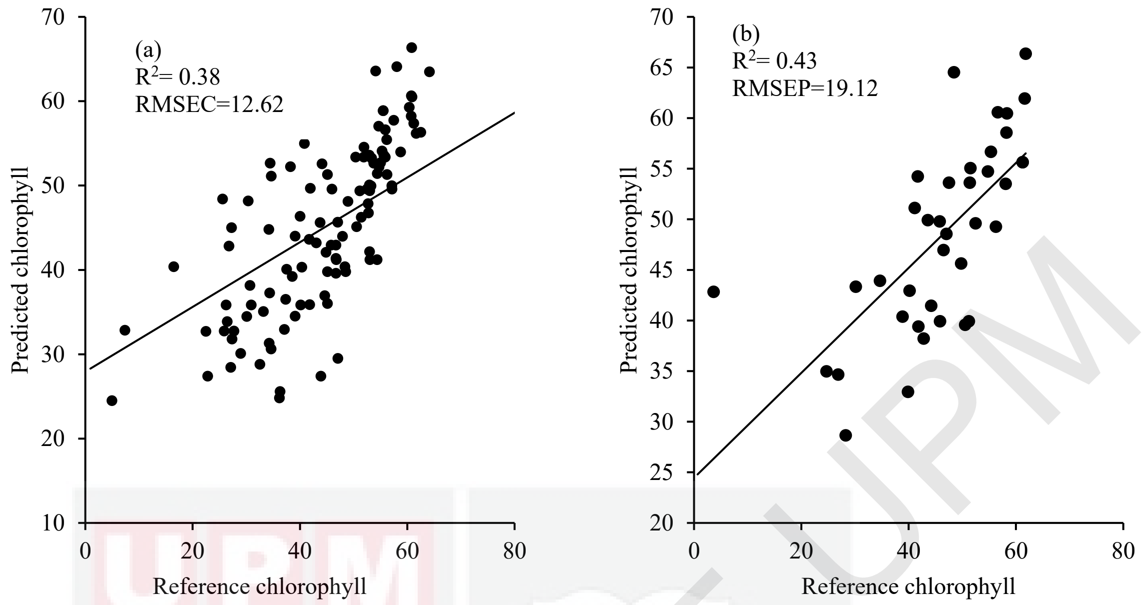


Scatter plots of chlorophyll content for ANOVA FS using BOC + SNV pre-processing data reflectance of 450 – 699 nm spectral data; (a) calibration (b) prediction model

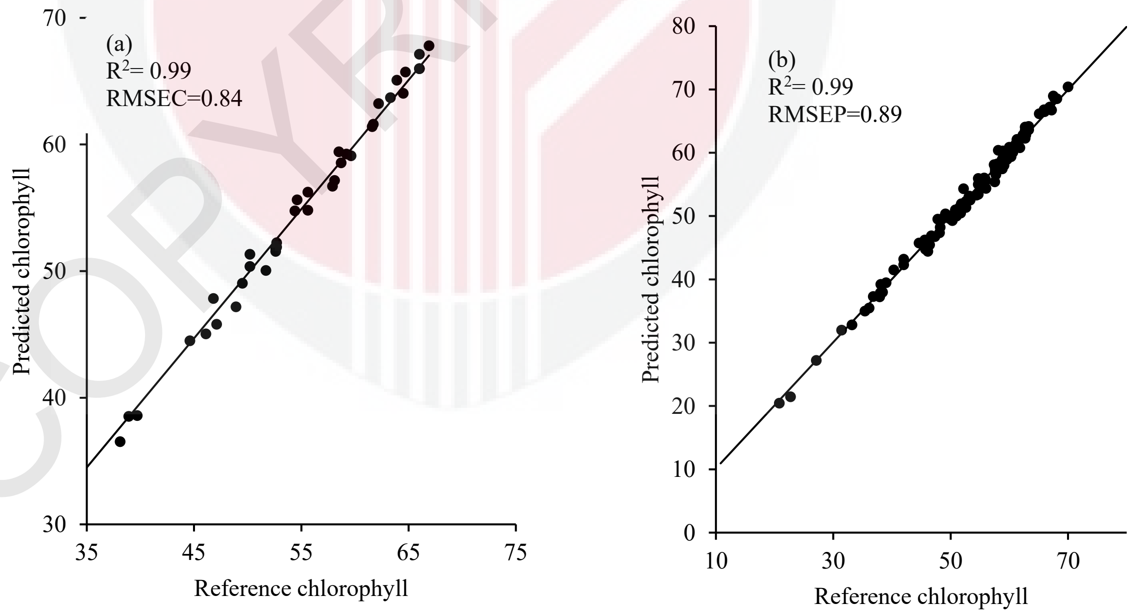


Scatter plots of chlorophyll content for ANOVA FS using raw data reflectance of 700 - 900 nm spectral data; (a) calibration (b) prediction model

© COPYRIGHT UPM

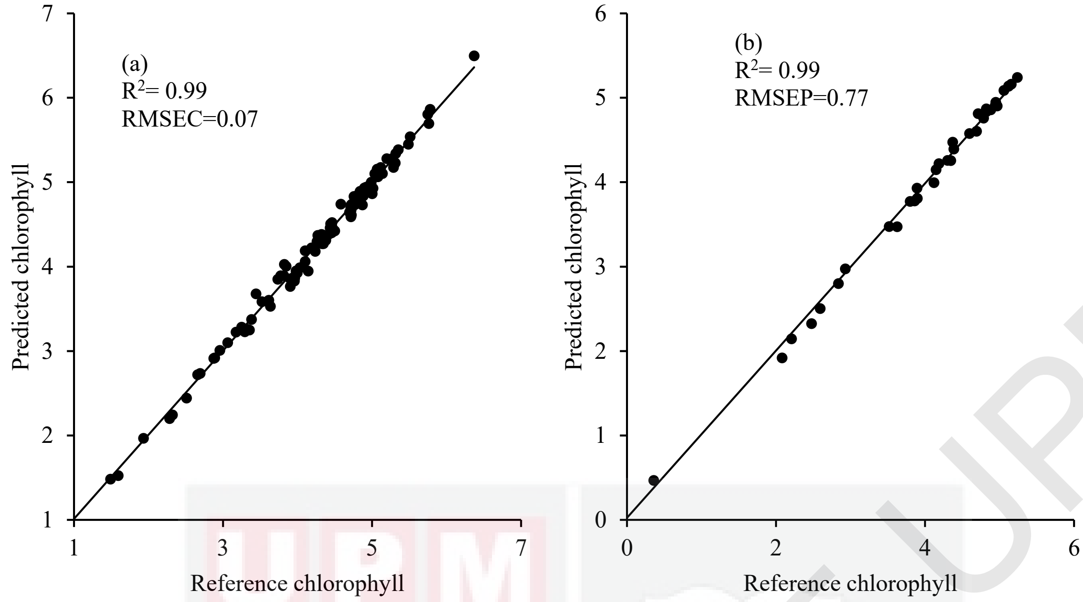


Scatter plots of chlorophyll content for ANOVA FS using BOC pre-processing data reflectance of 700 –900 nm spectral data; (a) calibration (b) prediction model.

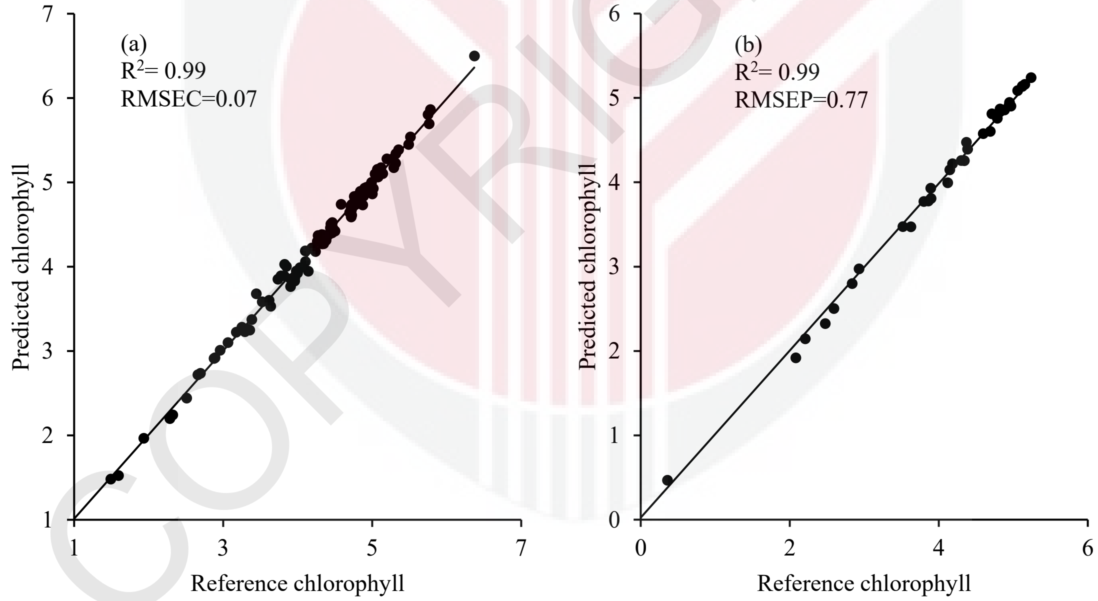


Scatter plots of chlorophyll content for ANOVA FS using SG pre-processing data reflectance of 700 –900 nm spectral data; (a) calibration (b) prediction model

© COPYRIGHT UPM

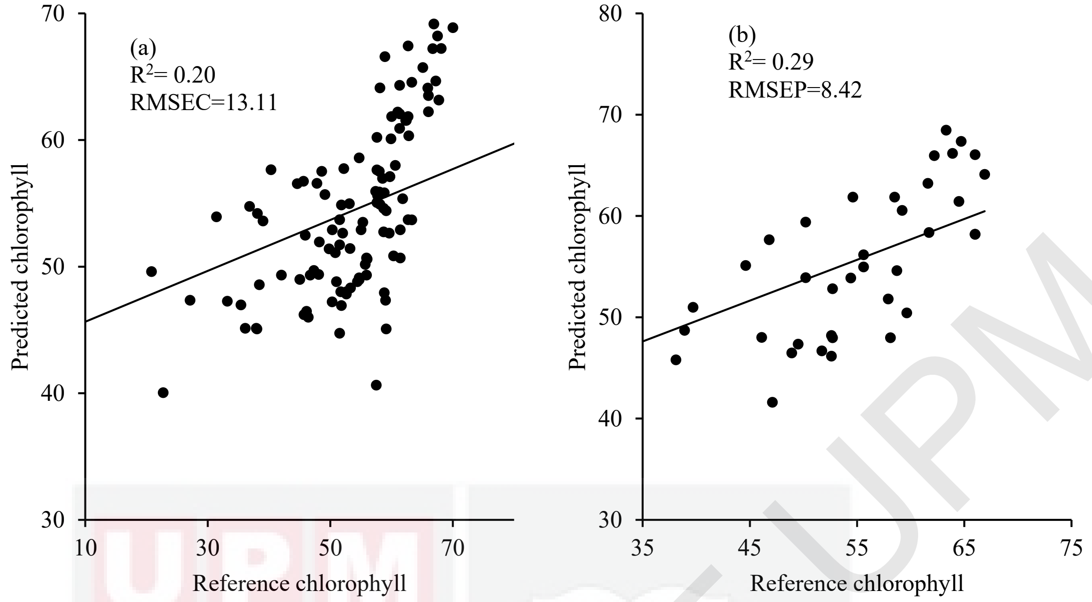


Scatter plots of chlorophyll content for ANOVA FS using SNV pre-processing data reflectance of 700 –900 nm spectral data; (a) calibration (b) prediction model

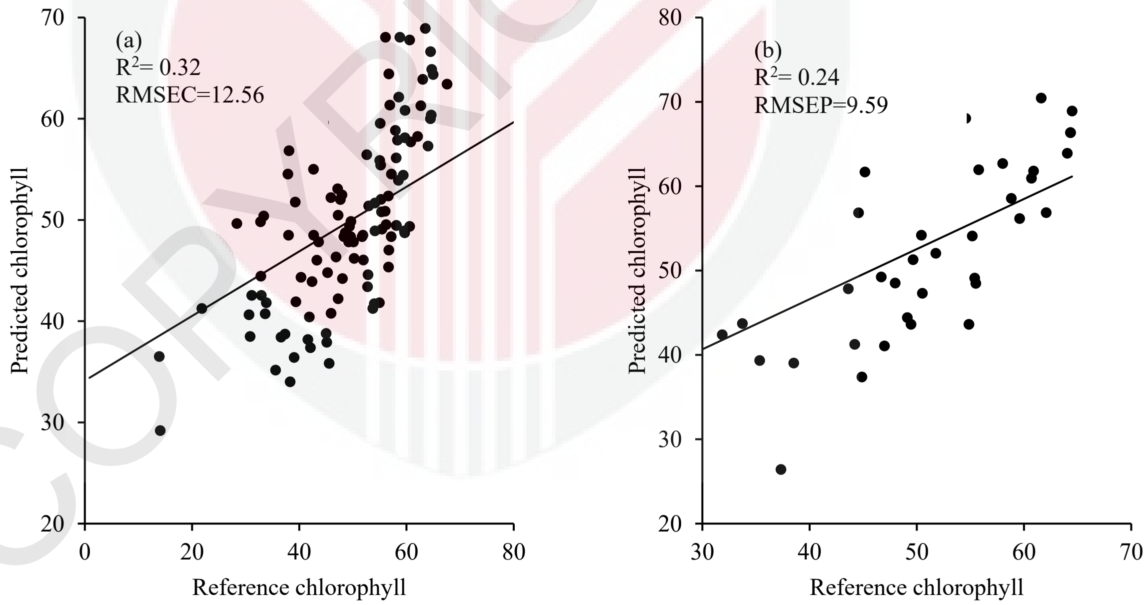


Scatter plots of chlorophyll content for ANOVA FS using BOC + SNV pre-processing data reflectance of 700 –900 nm spectral data; (a) calibration (b) prediction model

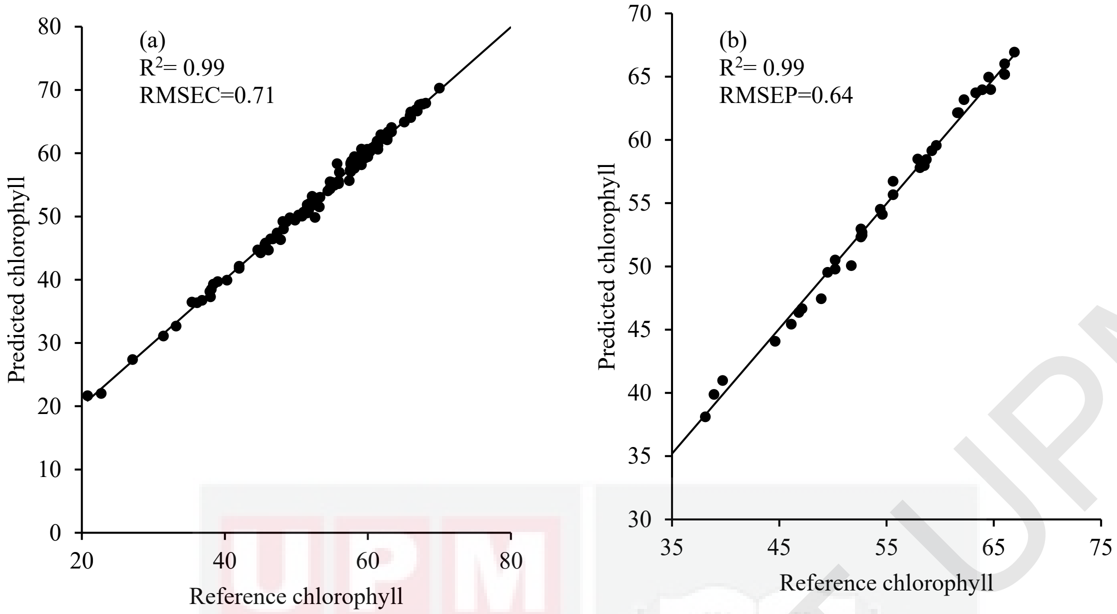
© COPYRIGHT UPM



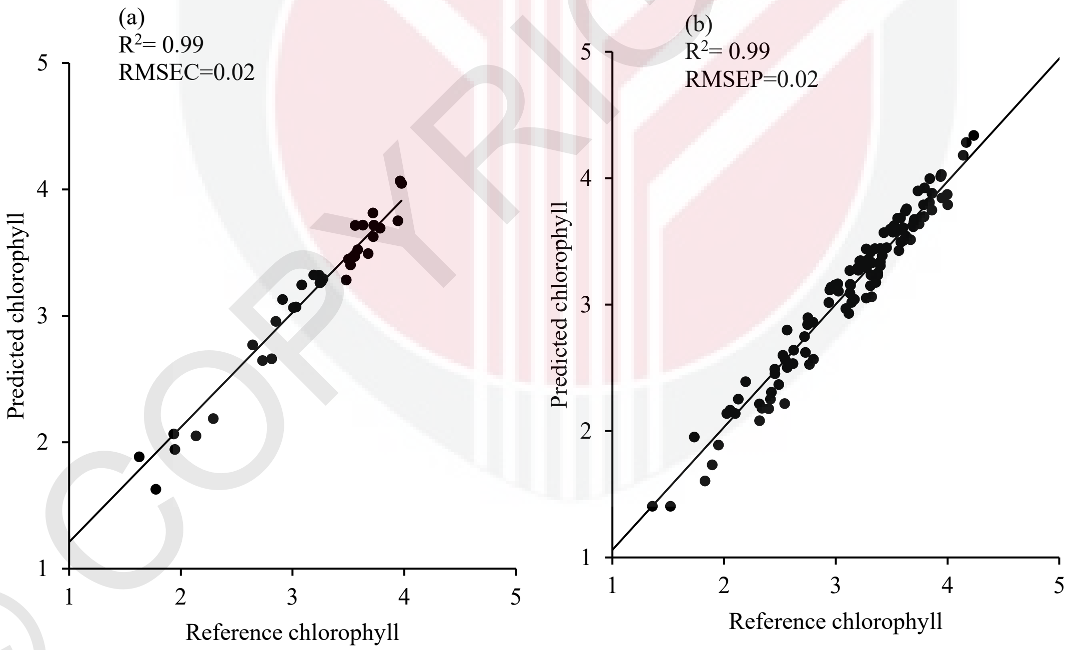
Scatter plots of chlorophyll content for ANOVA FS using raw data reflectance of 450 –900 nm spectral data; (a) calibration (b) prediction model.



Scatter plots of chlorophyll content for ANOVA FS using BOC pre-processing data reflectance of 450 –900 nm spectral data; (a) calibration (b) prediction model

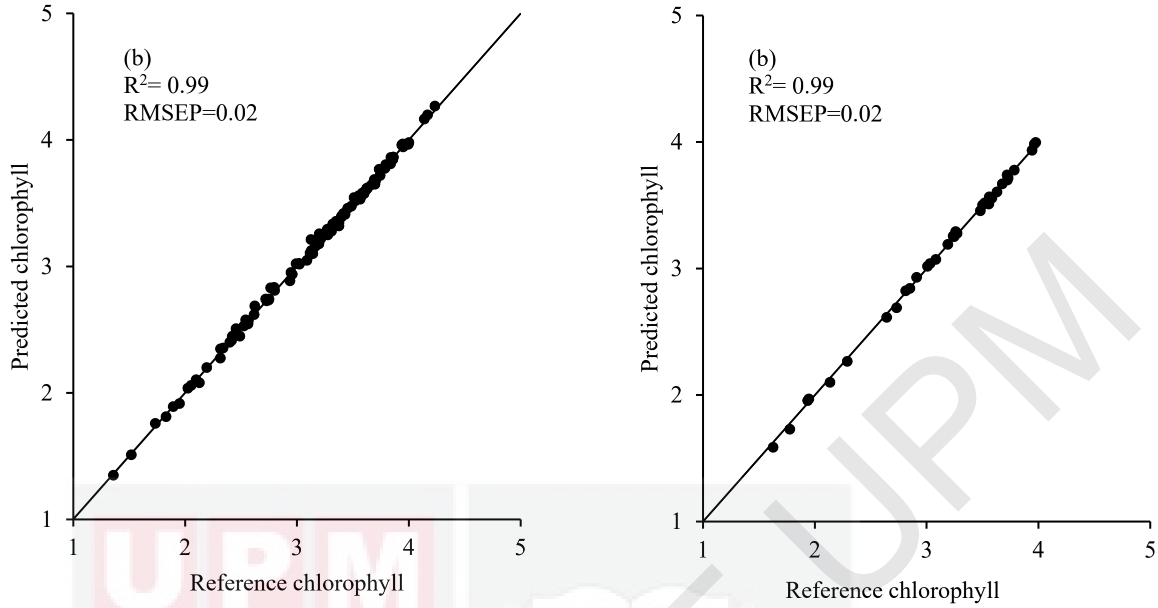


Scatter plots of chlorophyll content for ANOVA FS using SG pre-processing data reflectance of 450 –900 nm spectral data; (a) calibration (b) prediction model

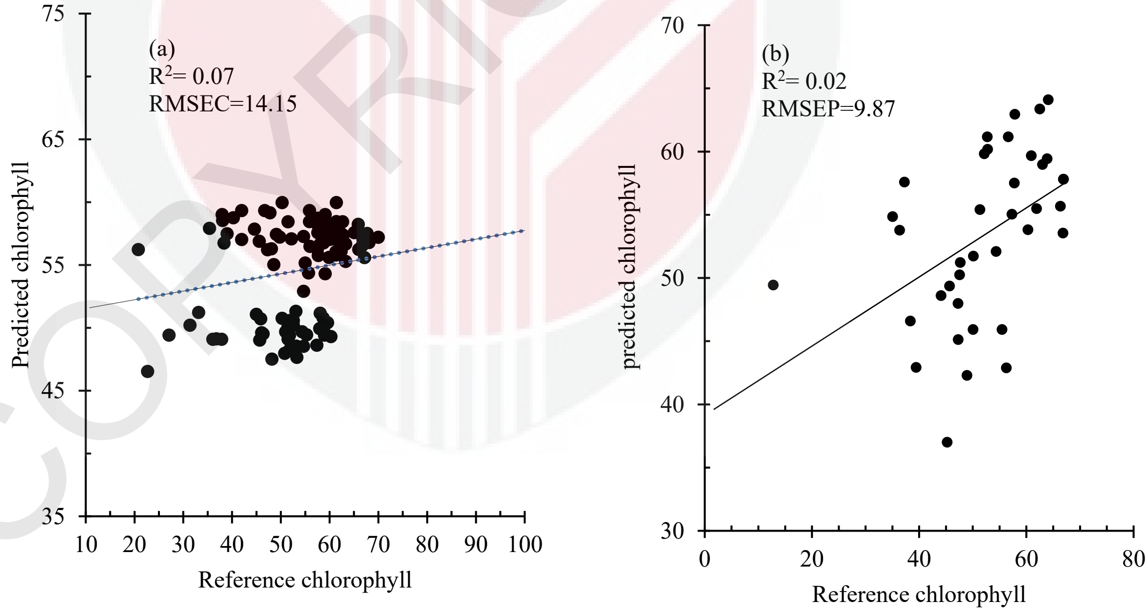


Scatter plots of chlorophyll content for ANOVA FS using SNV pre-processing data reflectance of 450 –900 nm spectral data; (a) calibration (b) prediction model

© COPYRIGHT UPM

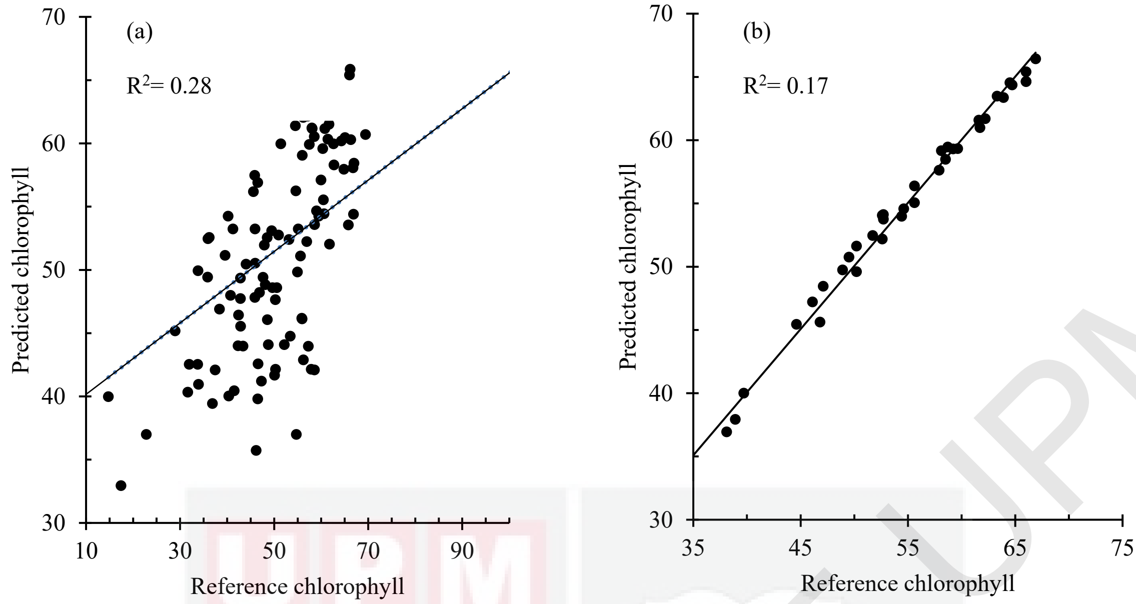


Scatter plots of chlorophyll content for ANOVA FS using BOC + SNV pre-processing data reflectance of 450 –900 nm spectral data; (a) calibration (b) prediction model

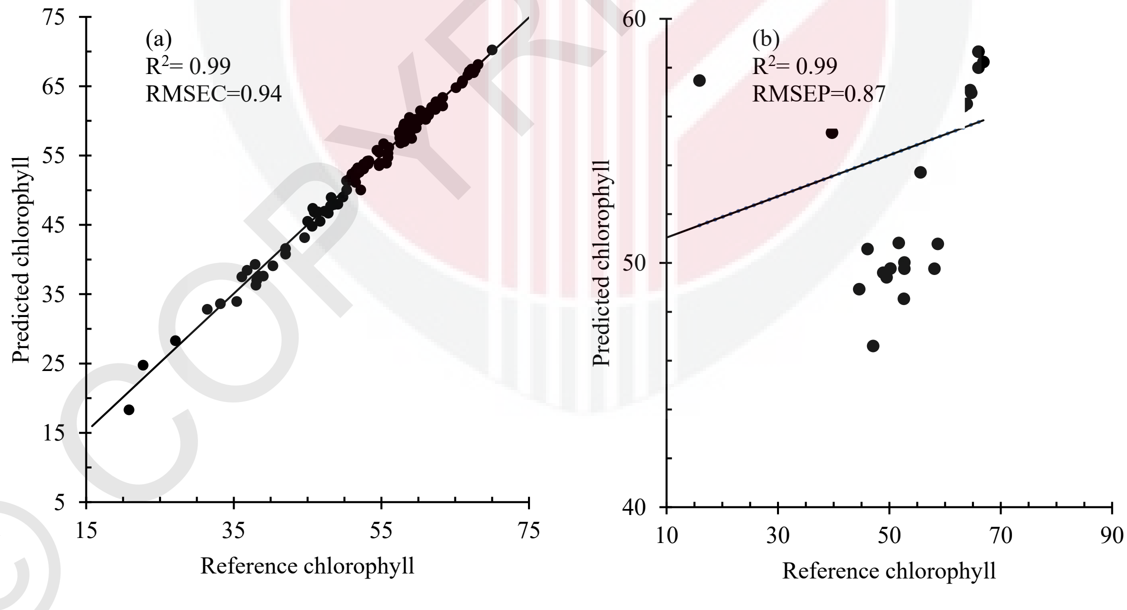


Scatter plots of chlorophyll content for Chi Square FS based on raw data reflectance spectral for 450 – 699 nm data; (a) calibration (b) prediction model

© COPYRIGHT UPM

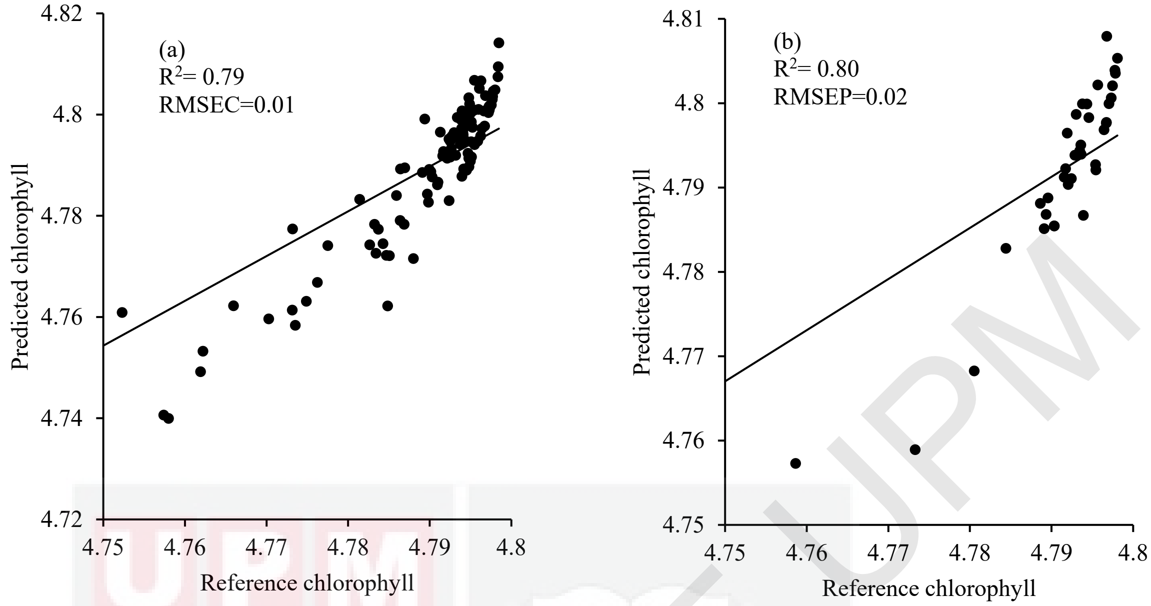


Scatter plots of chlorophyll content for Chi Square FS using BOC pre-processing data reflectance spectral data for 450 – 699 nm; (a) calibration (b) prediction model

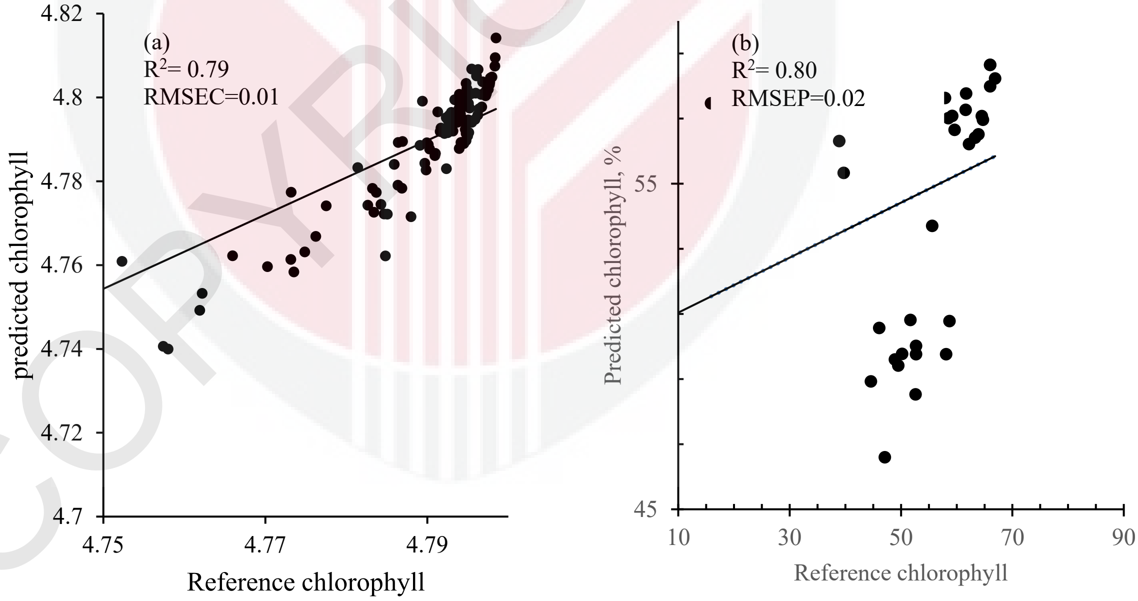


Scatter plots of chlorophyll content for Chi Square FS using SG pre-processing data reflectance of 450 – 699 nm spectral data; (a) calibration (b) prediction model

© COPYRIGHT UPM

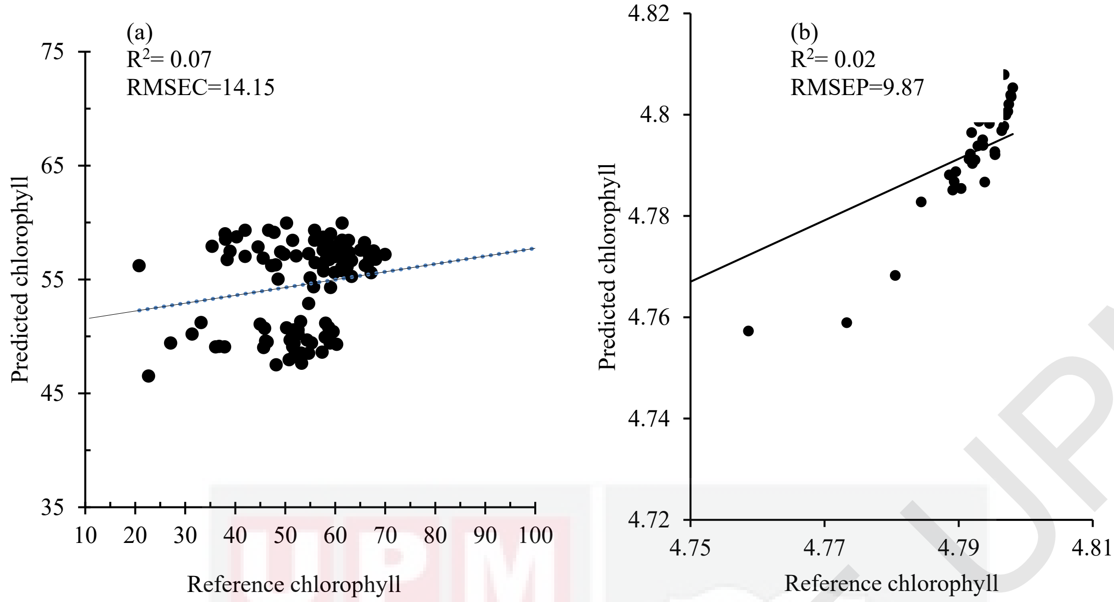


Scatter plots of chlorophyll content for Chi Square FS using SNV pre-processing data reflectance of 450 – 699 nm spectral data; (a) calibration (b) prediction model

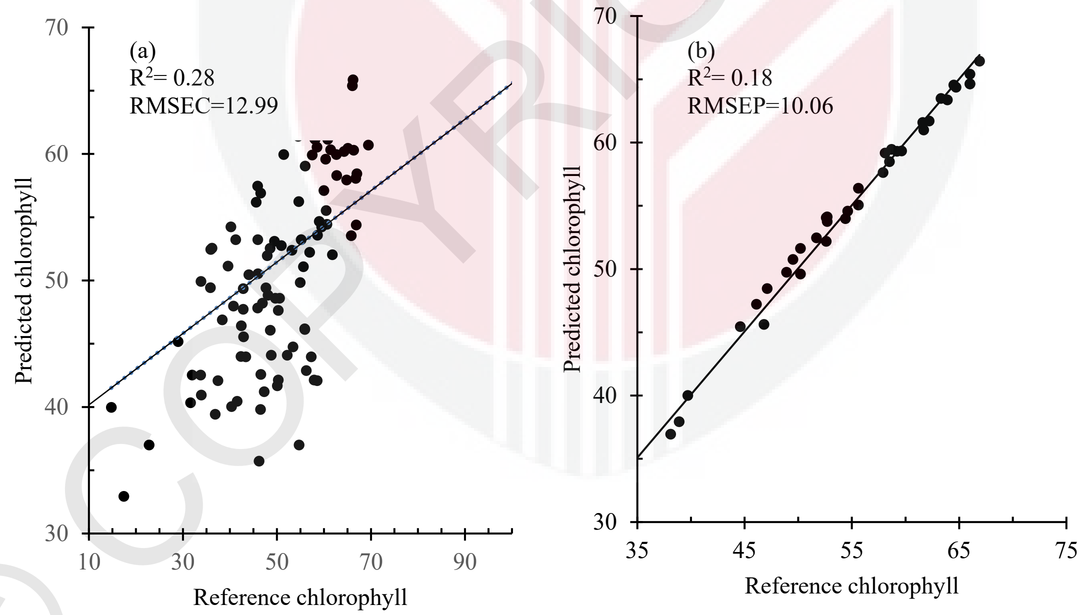


Scatter plots of chlorophyll content for Chi Square FS using SNV + BOC pre-processing data reflectance of 450 – 699 nm spectral data; (a) calibration (b) prediction model

© COPYRIGHT UPM

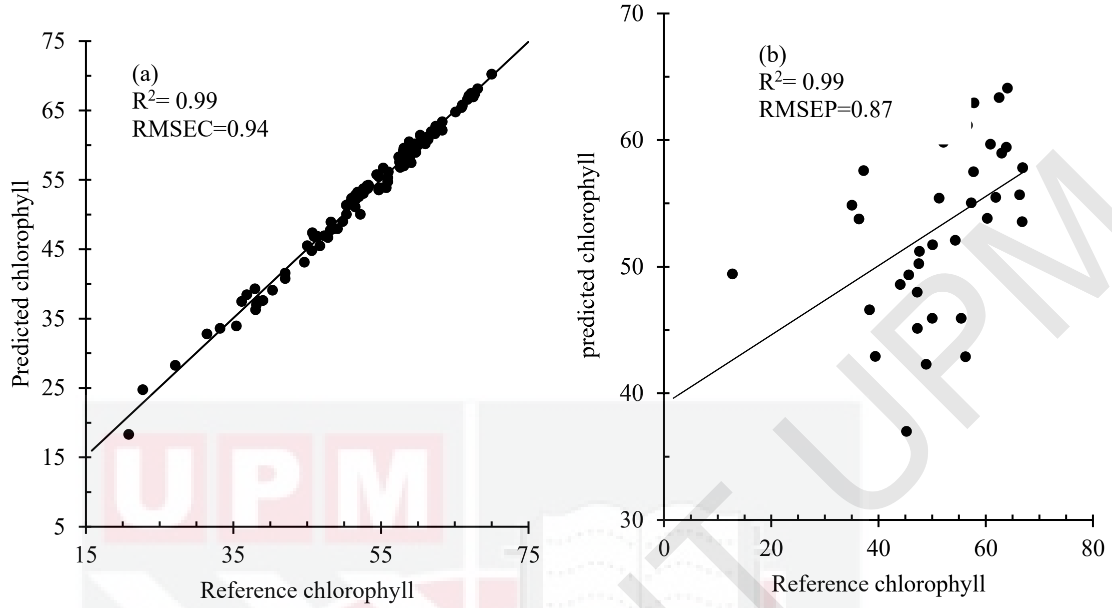


Scatter plots of chlorophyll content for Chi Square FS based on raw data reflectance spectral for 700 - 900 nm data; (a) calibration (b) prediction model

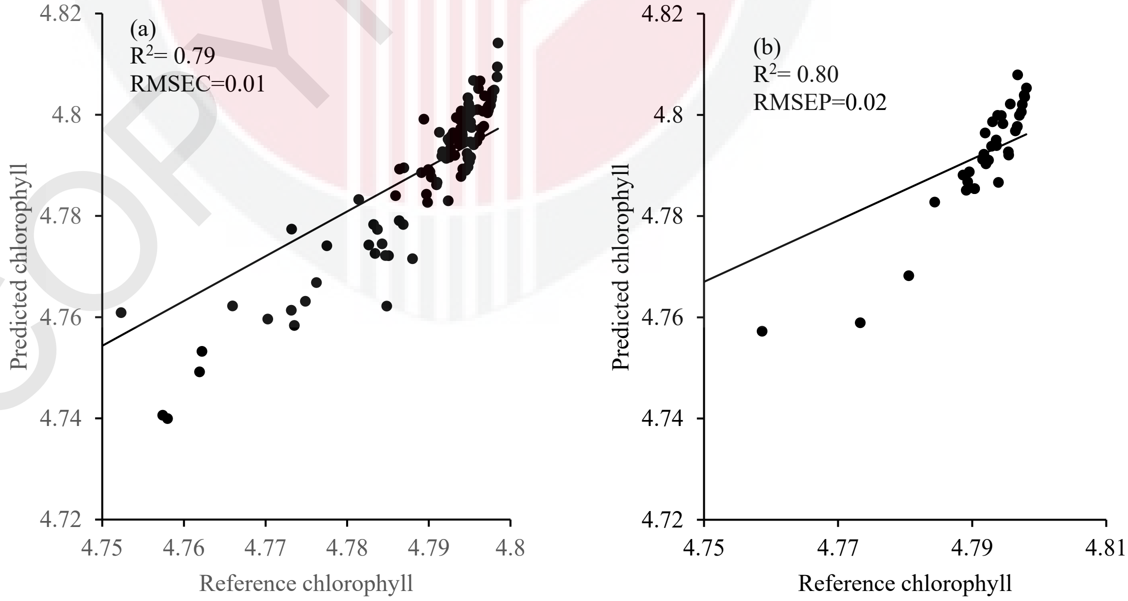


Scatter plots of chlorophyll content for Chi Square FS using BOC pre-processing data reflectance spectral data for 700 - 900 nm; (a) calibration (b) prediction model

© COPYRIGHT UPM

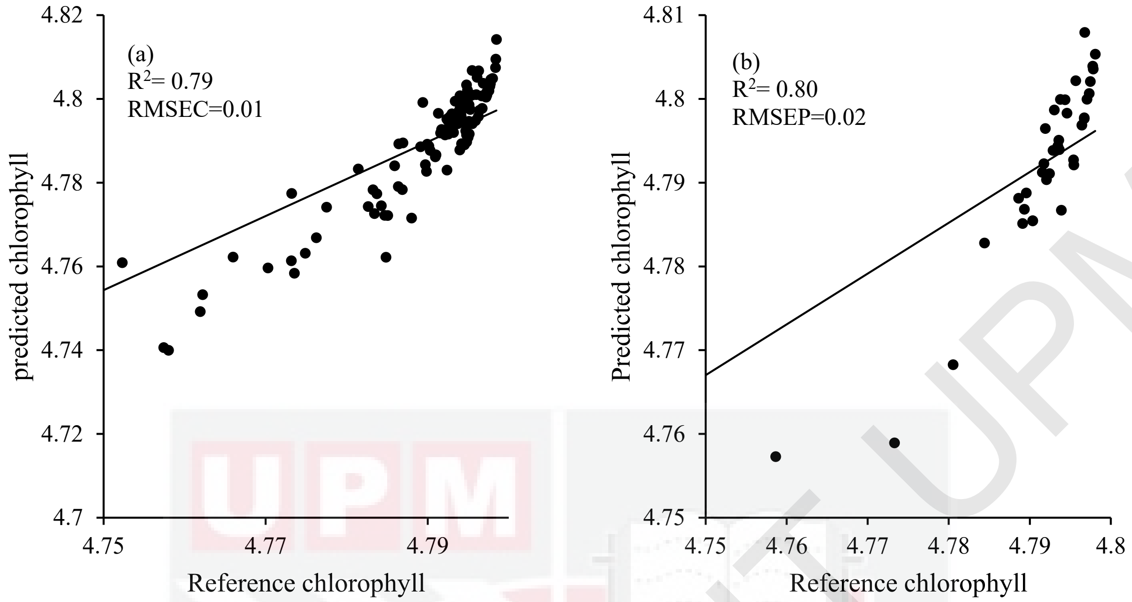


Scatter plots of chlorophyll content for Chi Square FS using SG pre-processing data reflectance of 700 - 900 nm spectral data; (a) calibration (b) prediction model

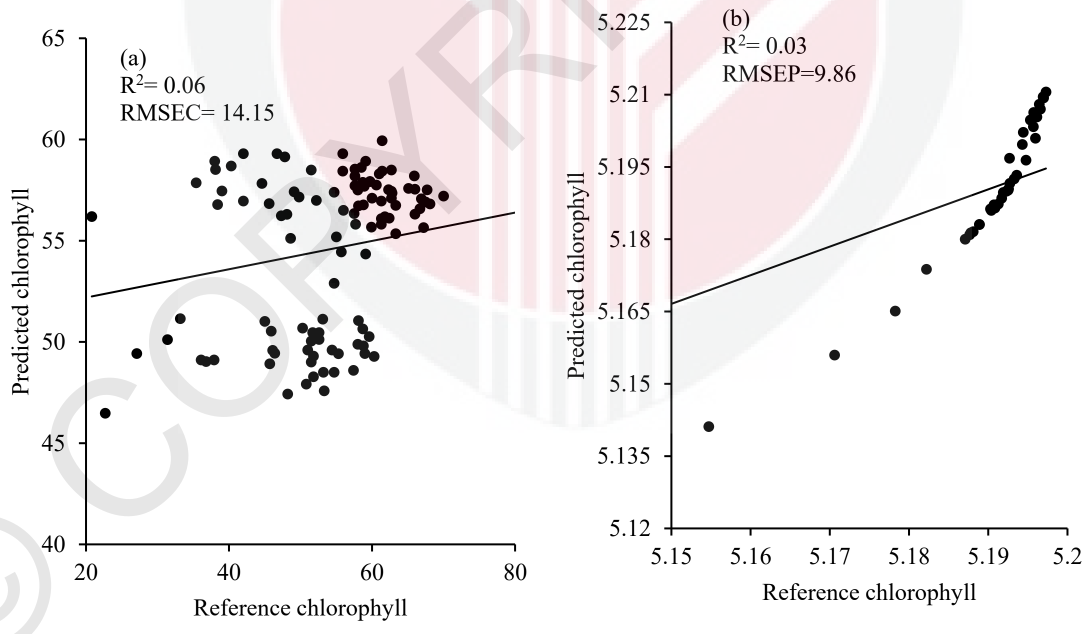


Scatter plots of chlorophyll content for Chi Square FS using SNV pre-processing data reflectance of 700 - 900 nm spectral data; (a) calibration (b) prediction model.

© COPYRIGHT UPM

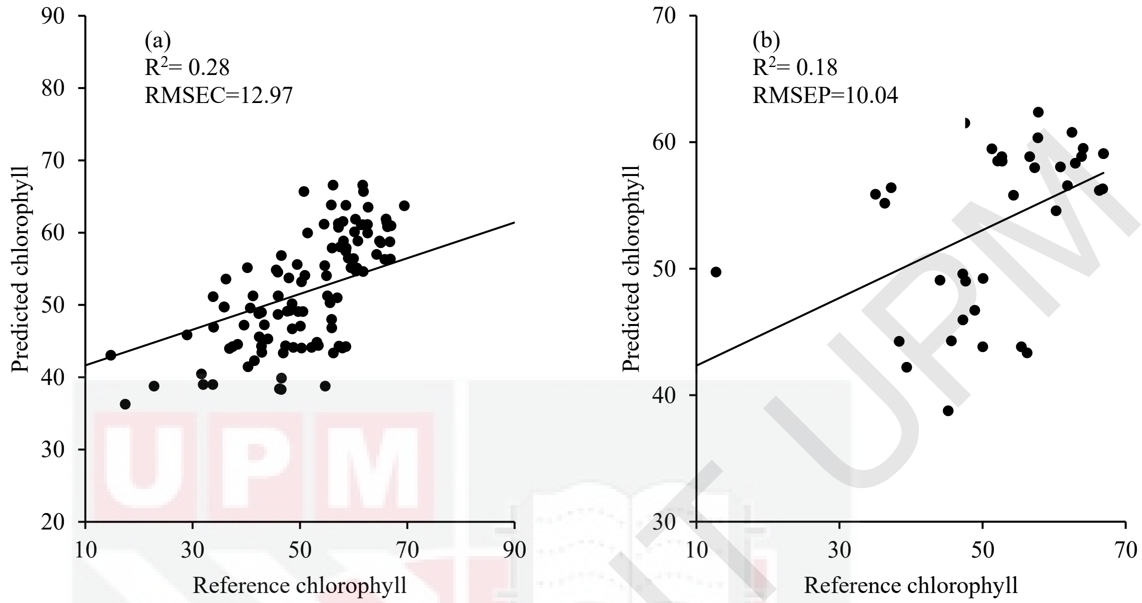


Scatter plots of chlorophyll content for Chi Square FS using SNV + BOC pre-processing data reflectance of 700 - 900 nm spectral data; (a) calibration (b) prediction model

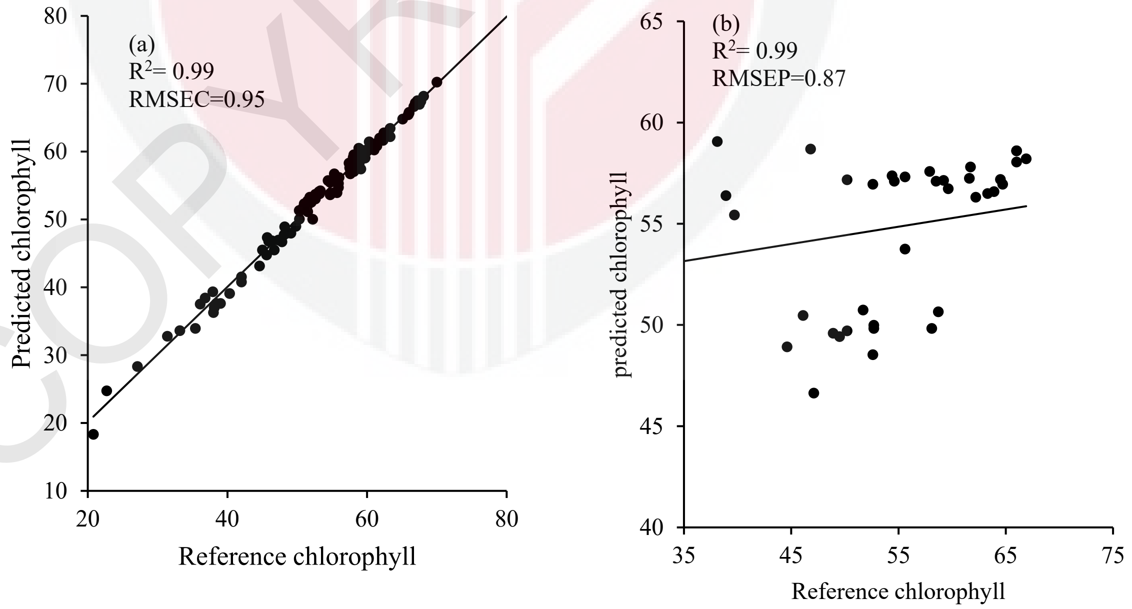


Scatter plots of chlorophyll content for Chi Square FS using raw data reflectance of 450 – 900 nm spectral data; (a) calibration (b) prediction model

© COPYRIGHT UPM

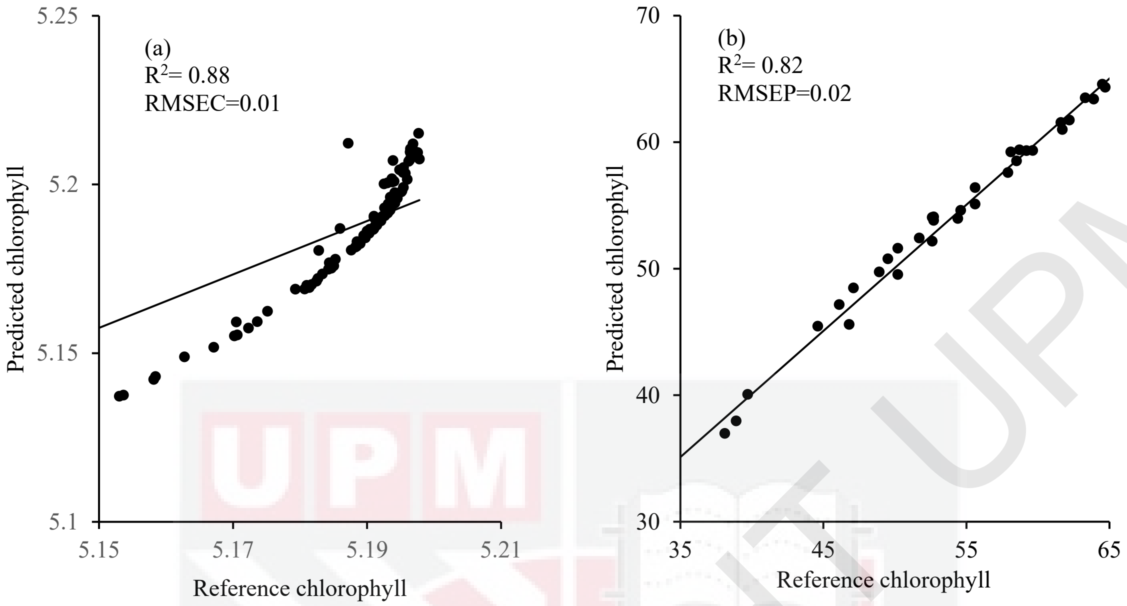


Scatter plots of chlorophyll content for Chi Square FS using BOC pre-processing data reflectance of 450 – 699 nm spectral data; (a) calibration (b) prediction model

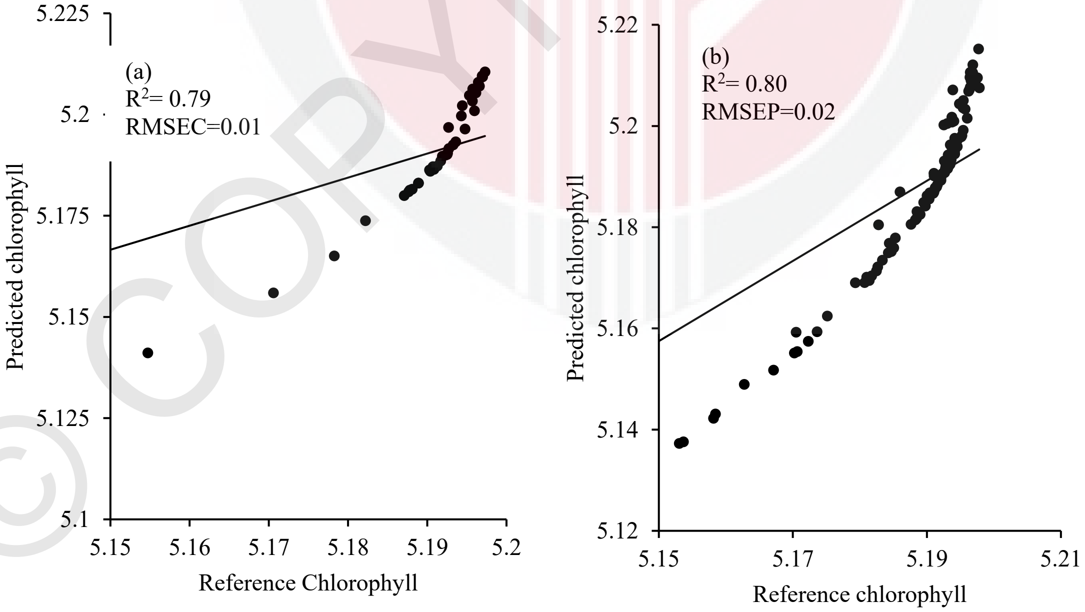


Scatter plots of chlorophyll content for Chi Square FS using SG pre-processing data reflectance of 450 – 699 nm spectral data; (a) calibration (b) prediction model

© COPYRIGHT UPM

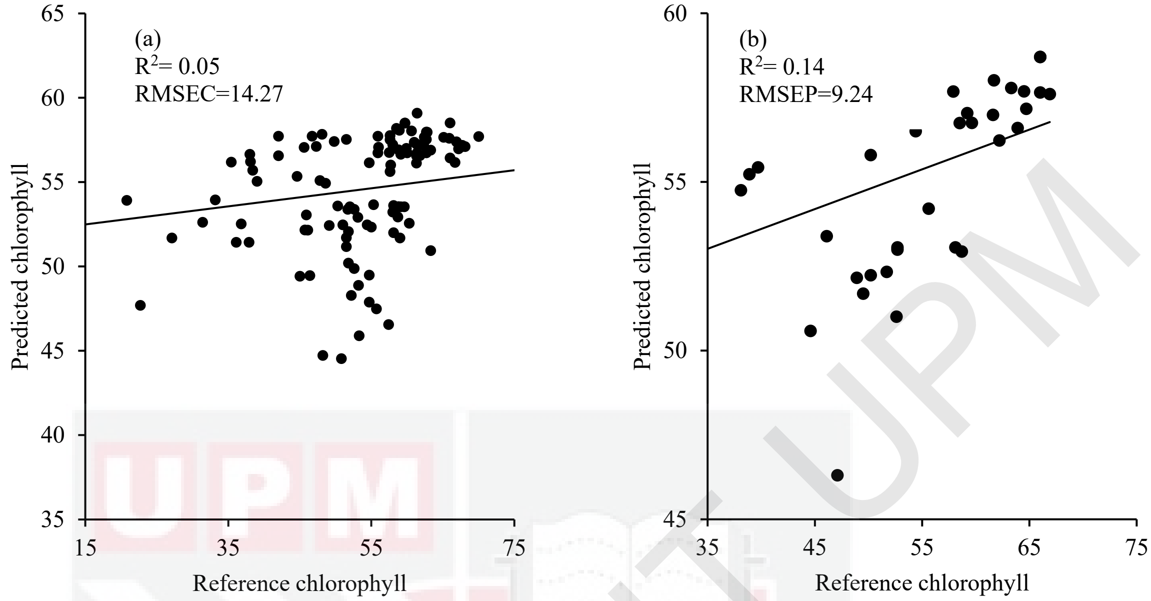


Scatter plots of chlorophyll content for Chi Square FS using SNV pre-processing data reflectance of 450 – 699 nm spectral data; (a) calibration (b) prediction model

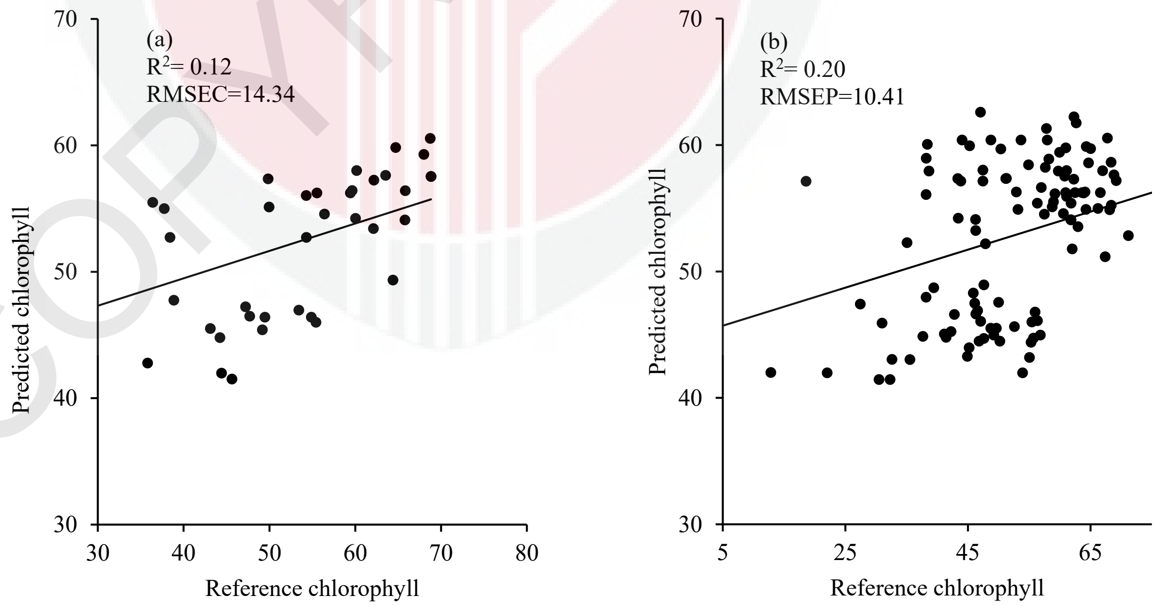


Scatter plots of chlorophyll content for Chi Square FS using SNV+ BOC pre-processing data reflectance of 450 – 699 nm spectral data; (a) calibration (b) prediction model

© COPYRIGHT UPM

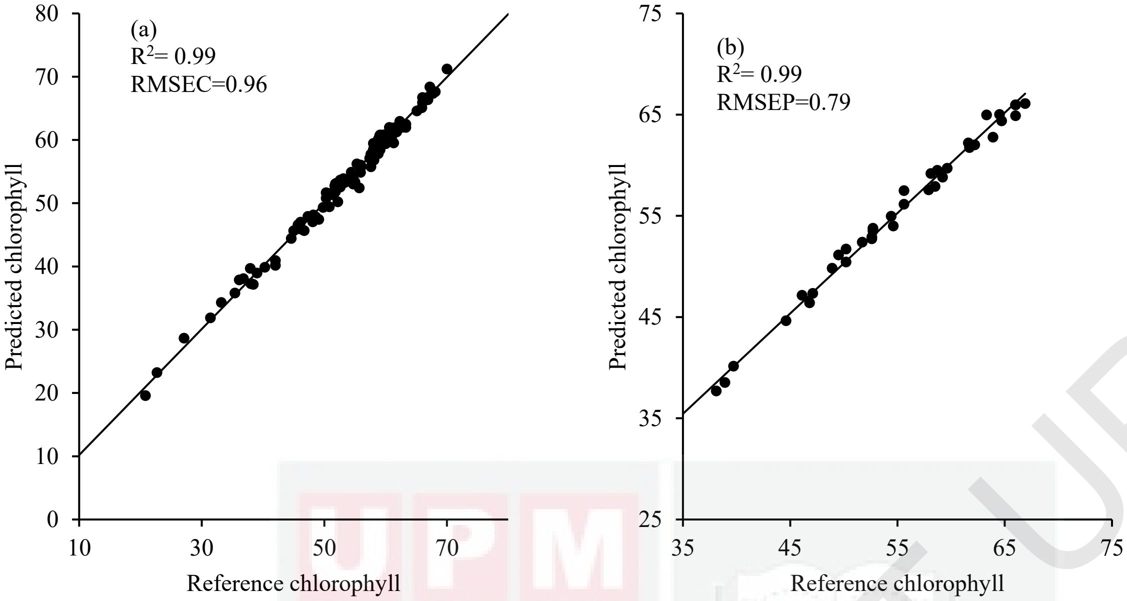


Scatter plots of chlorophyll content for RELIEF FS using raw data reflectance of 450 -699 nm spectral data; (a) calibration (b) prediction model

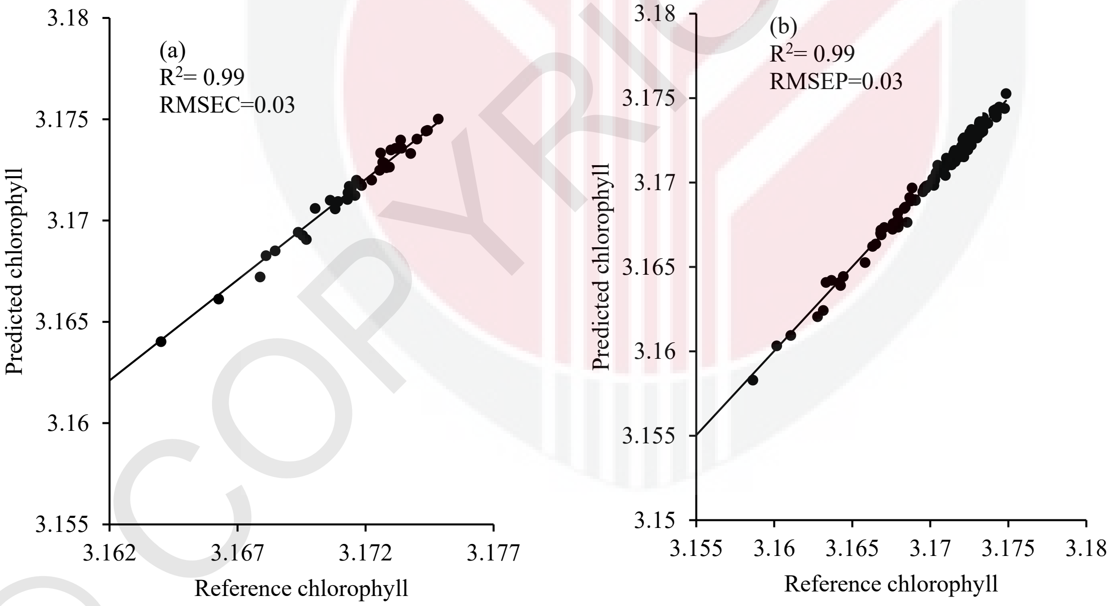


Scatter plots of chlorophyll content for RELIEF FS using BOC pre-processing data reflectance of 450 – 699 nm spectral data; (a) calibration (b) prediction model

© COPYRIGHT UPM

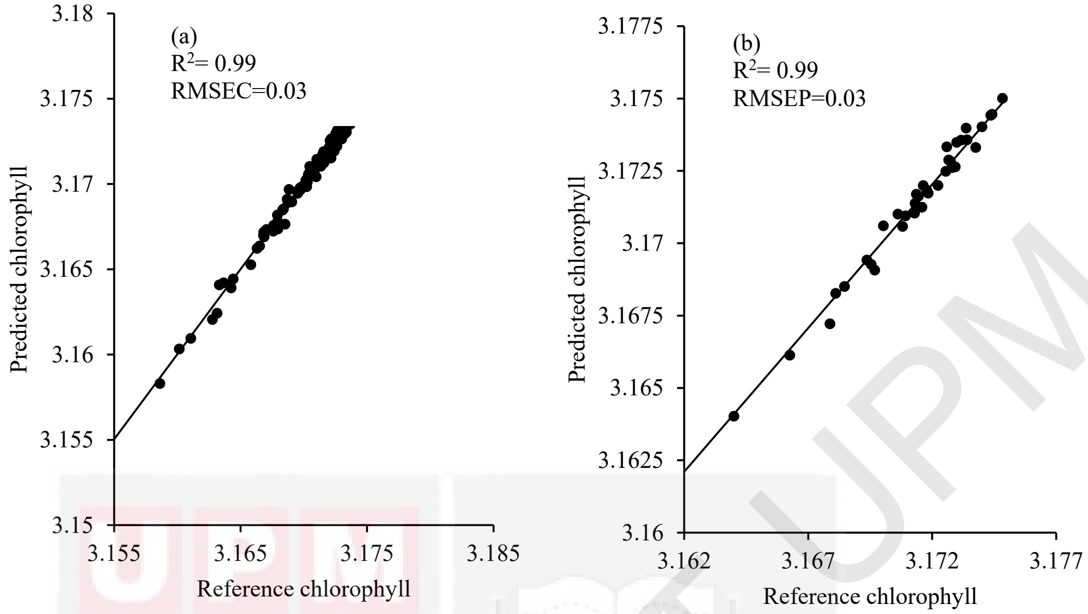


Scatter plots of chlorophyll content for RELIEF FS using SG pre-processing data reflectance of 450 – 699 nm spectral data; (a) calibration (b) prediction model

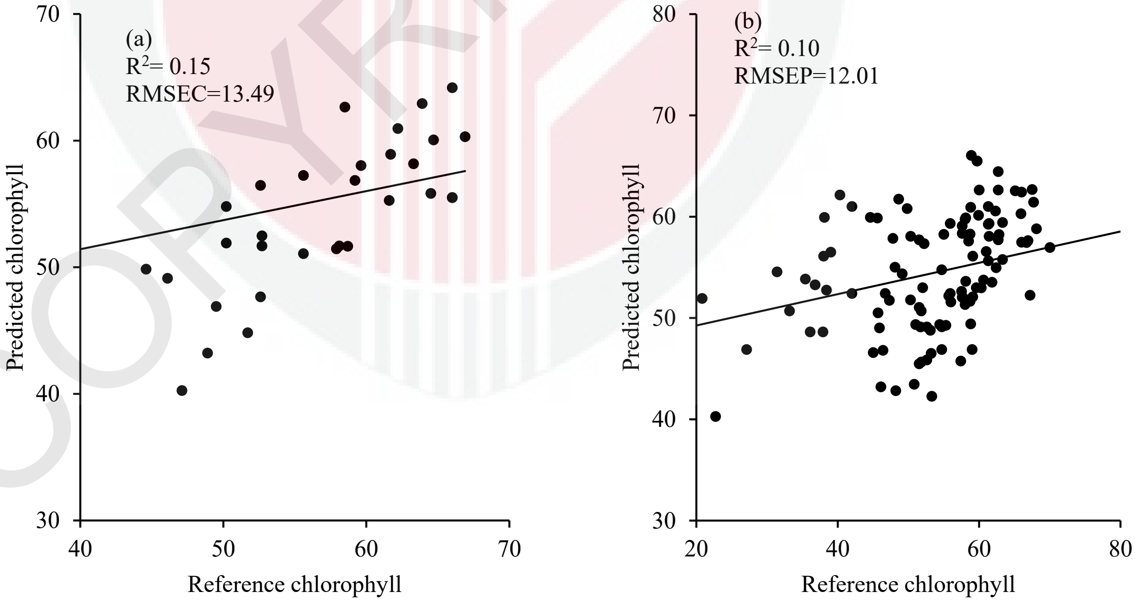


Scatter plots of chlorophyll content for RELIEF FS using SNV pre-processing data reflectance of 450 – 699 nm spectral data; (a) calibration (b) prediction model

© COPYRIGHT UPM

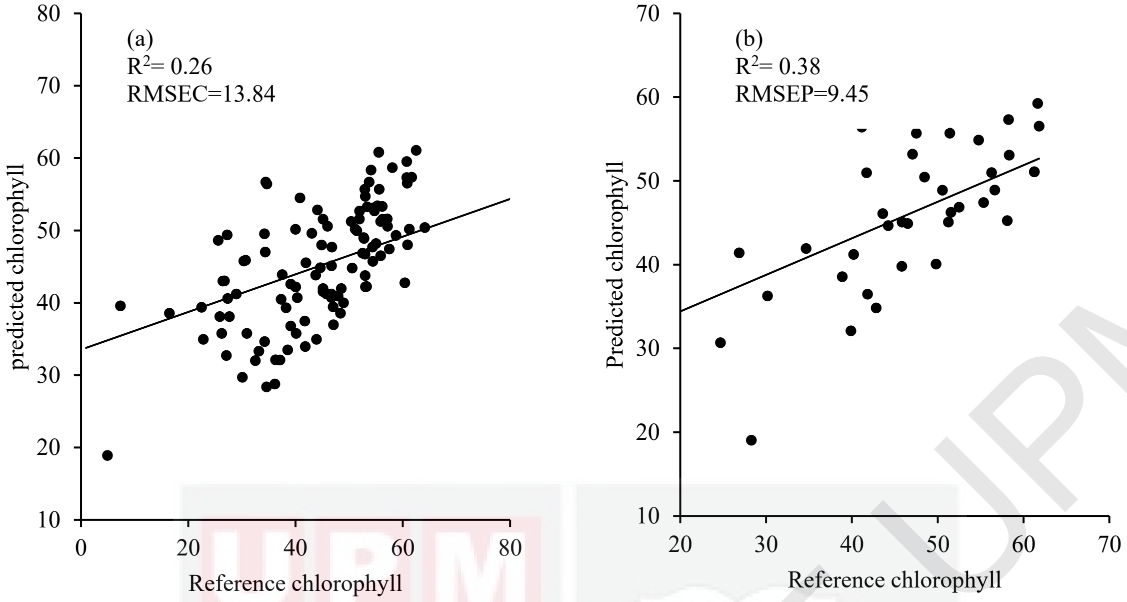


Scatter plots of chlorophyll content for RELIEF FS using BOC + SNV pre-processing data reflectance of 700 – 900 nm spectral data; (a) calibration (b) prediction model

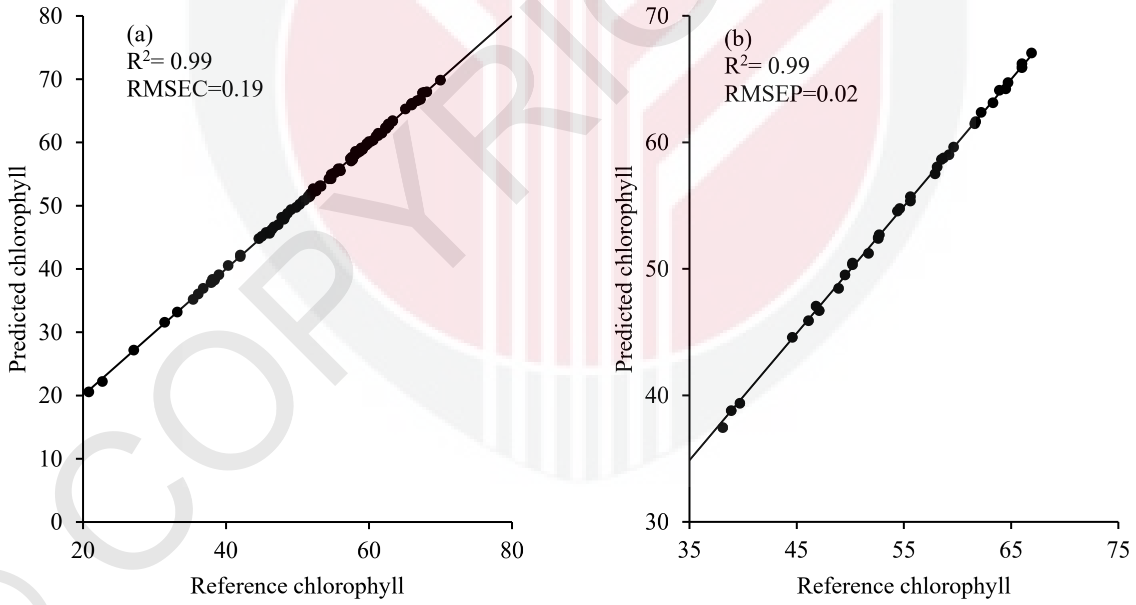


Scatter plots of chlorophyll content for RELIEF FS using raw data reflectance of 700 – 900 nm spectral data; (a) calibration (b) prediction model

© COPYRIGHT UPM

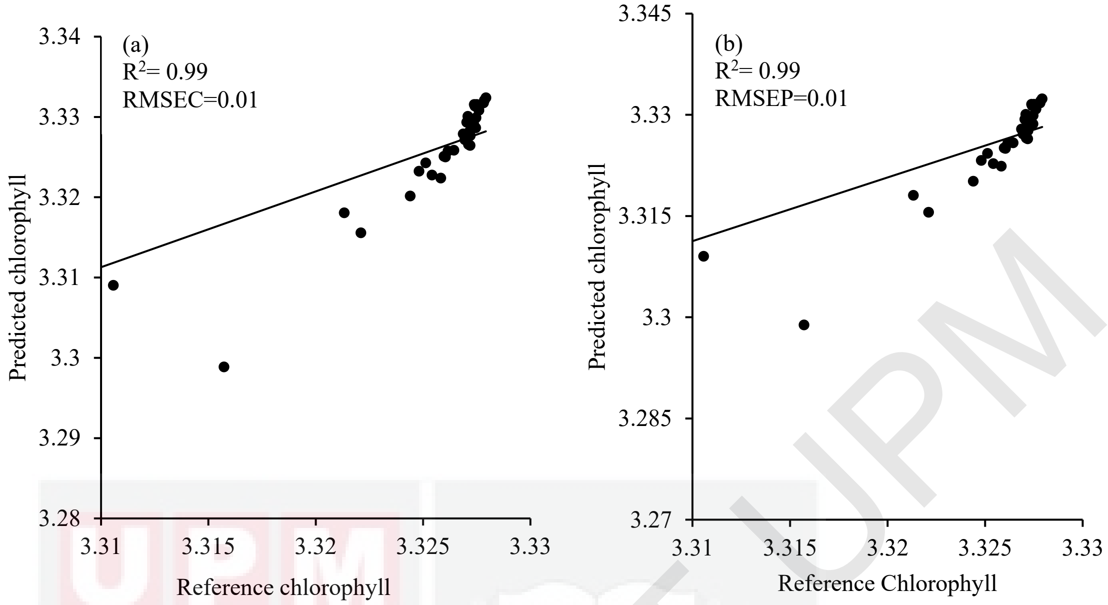


Scatter plots of chlorophyll content for RELIEF FS using BOC pre-processing data reflectance of 700 – 900 nm spectral data; (a) calibration (b) prediction model

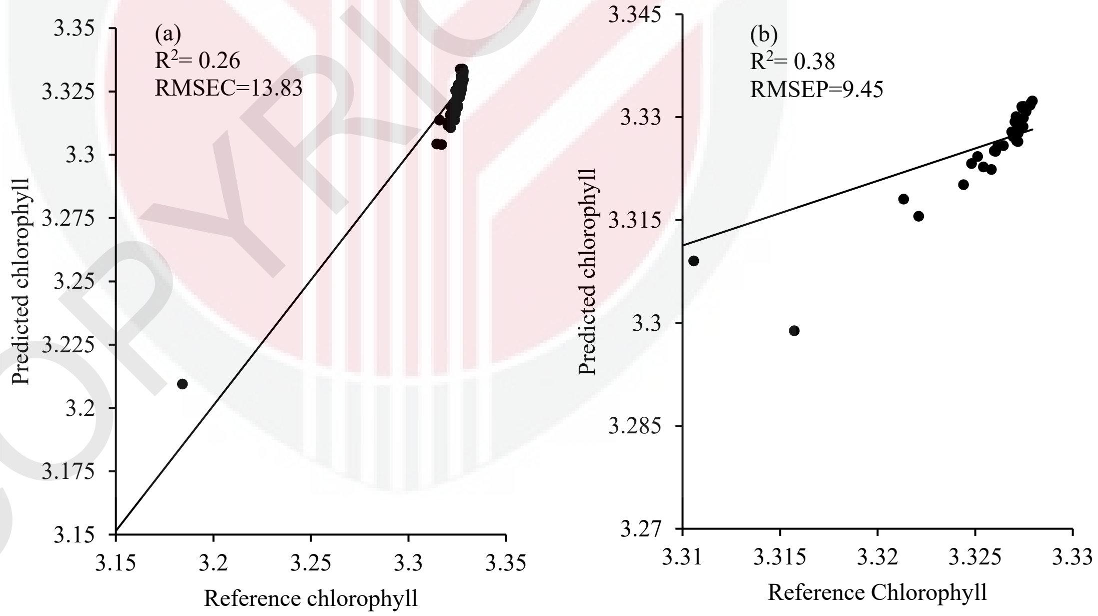


Scatter plots of chlorophyll content for RELIEF FS using SG pre-processing data reflectance of 700 – 900 nm spectral data; (a) calibration (b) prediction model

© COPYRIGHT UPM

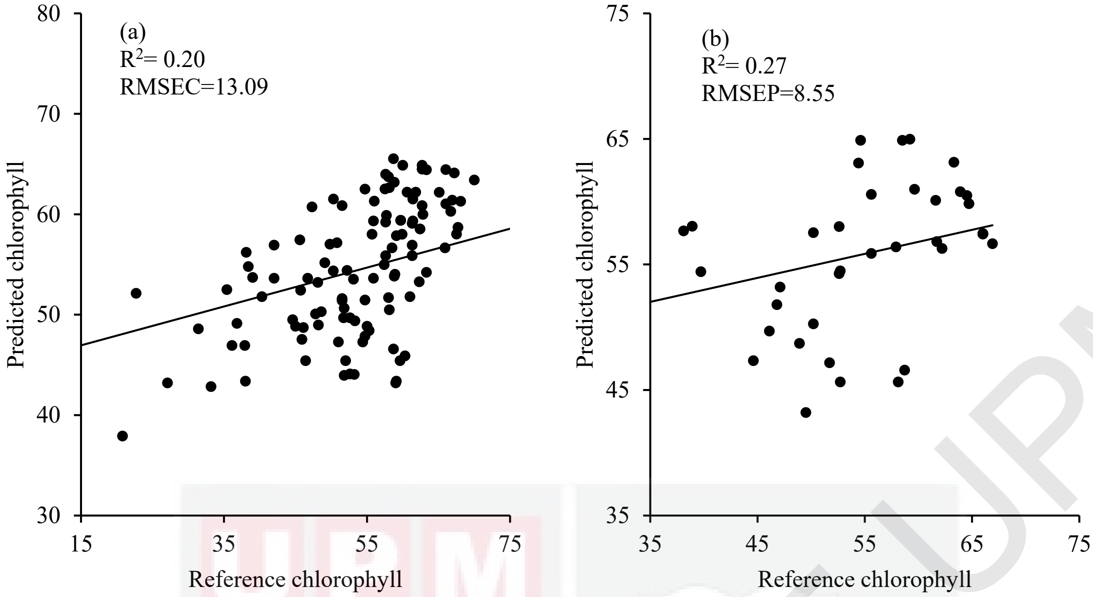


Scatter plots of chlorophyll content for RELIEF FS using SNV pre-processing data reflectance of 700 – 900 nm spectral data; (a) calibration (b) prediction model

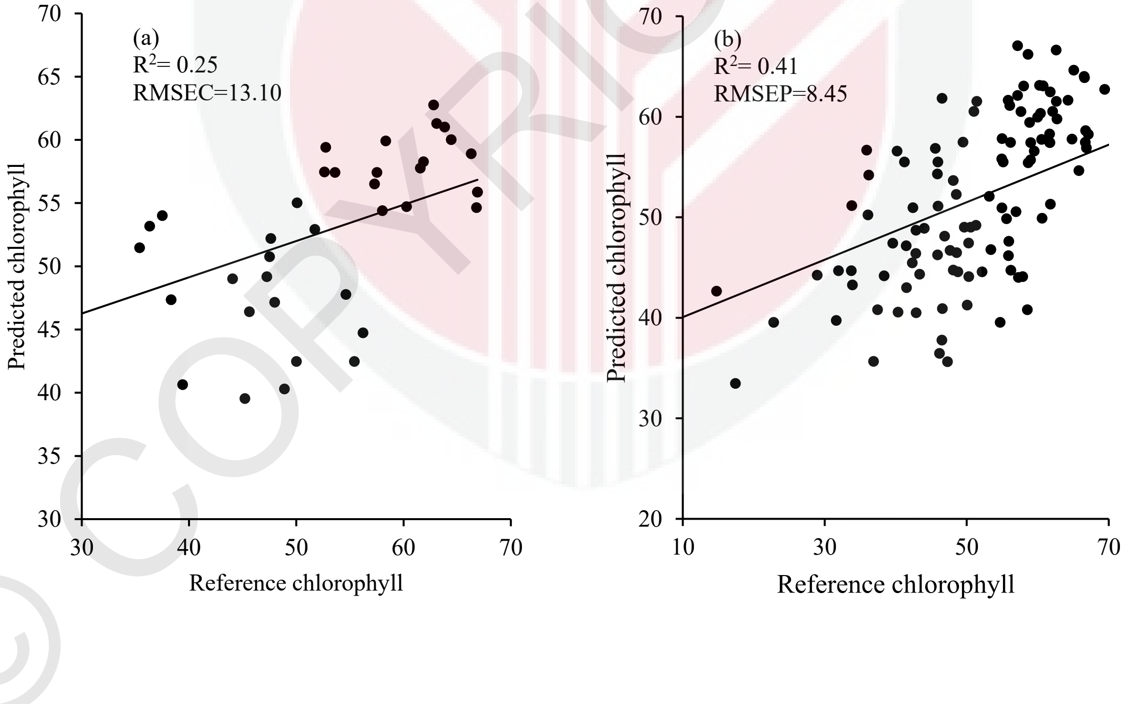


Scatter plots of chlorophyll content for RELIEF FS using BOC + SNV pre-processing data reflectance of 700 – 900 nm spectral data; (a) calibration (b) prediction model

© COPYRIGHT UPM

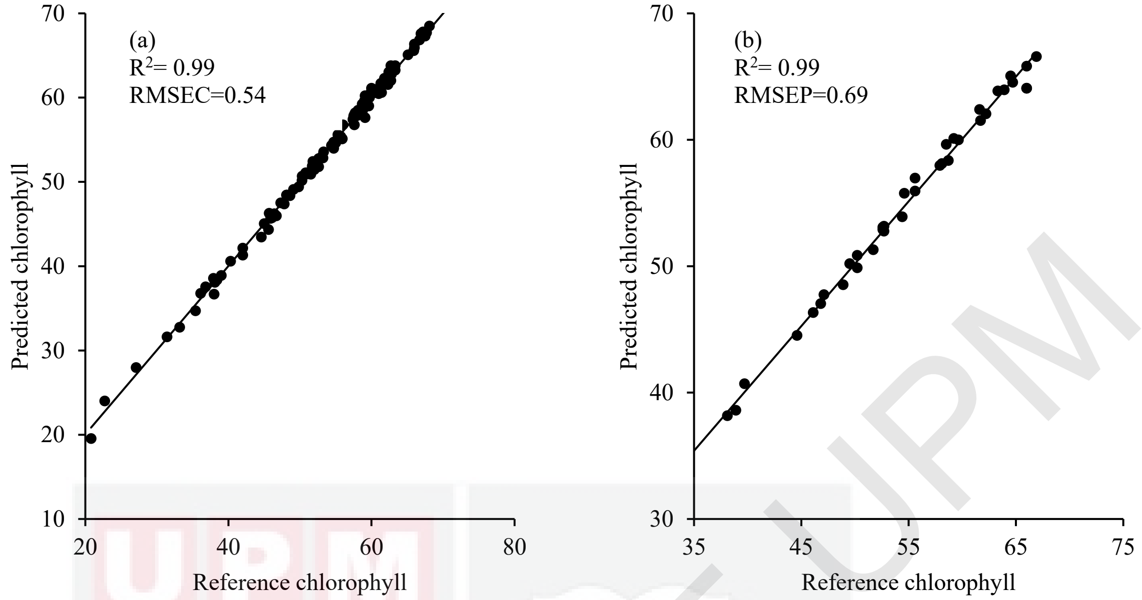


Scatter plots of chlorophyll content for RELIEF FS using raw data reflectance of 450 – 900 nm spectral data; (a) calibration (b) prediction model

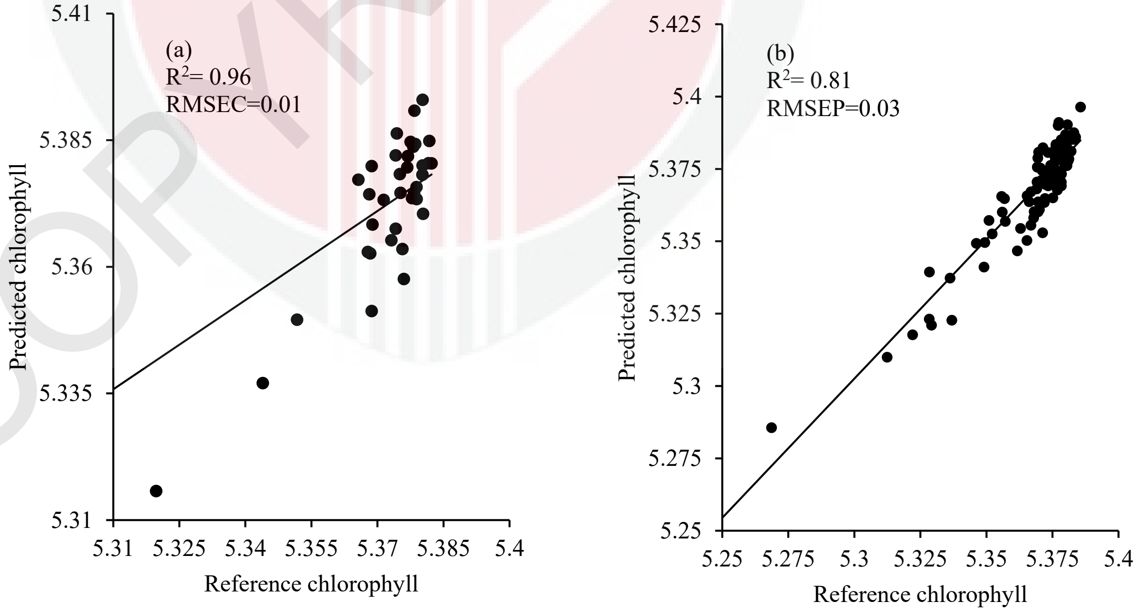


Scatter plots of chlorophyll content for RELIEF FS using BOC pre-processing data reflectance of 450 – 900 nm spectral data; (a) calibration (b) prediction model

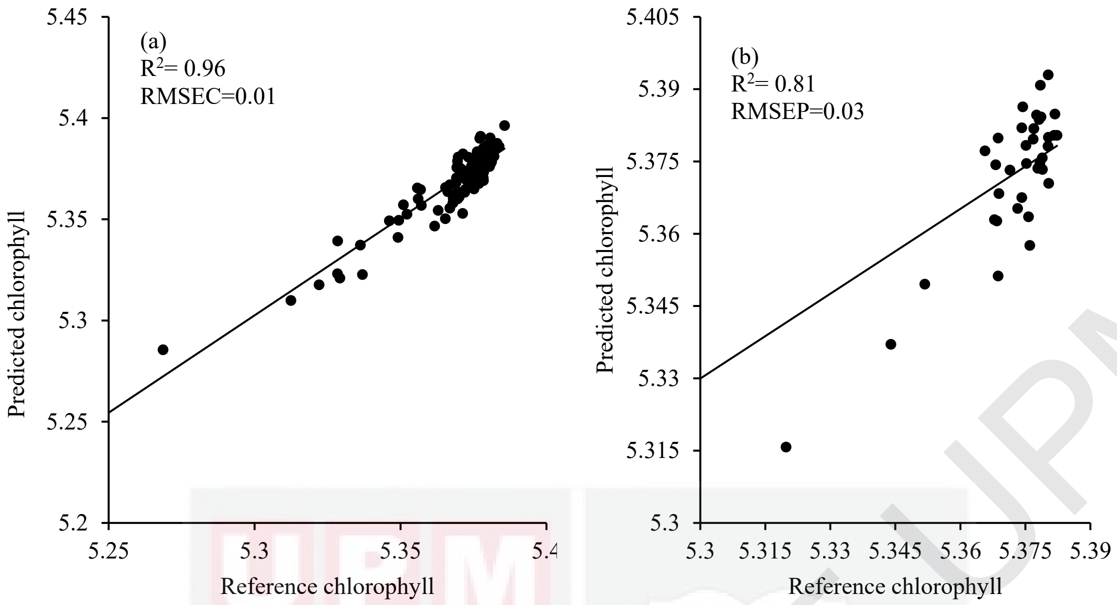
© COPYRIGHT UPM



Scatter plots o Scatter plots of chlorophyll content for RELIEF FS using SG pre-processing data reflectance of 450 – 900 nm spectral data; (a) calibration (b) prediction model



© COPYRIGHT UPM



Scatter plots of chlorophyll content for RELIEF FS using SNV + BOC pre-processing data reflectance of 450 – 900 nm spectral data; (a) calibration (b) prediction model

APPENDIX D

ANOVA TABLE

ANOVA for SPAD value of leaves at different WRD severity levels

Source	Degree of Freedom	Sum of square	Mean Square	F-Value	Pr> F
Model	4	55191.84	13797.96	700.06	<.0001
Error	595	11727.18	19.71		
Corrected error	599	66919.03			

ANOVA for MC of leaves at different WRD severity levels

Source	Degree of Freedom	Sum of square	Mean Square	F-Value	Pr> F
Model	4	2870.28	717.57	10.91	<.0001
Error	595	36162.72	65.75		
Corrected error	599	39033.00			

ANOVA for L* of leaves at different WRD severity levels

Source	Degree of Freedom	Sum of square	Mean Square	F-Value	Pr> F
Model	4	2875.96	718.99	37.55	<.0001
Error	595	11391.61	19.14		
Corrected error	599	14267.58			

ANOVA for a* of leaves at different WRD severity levels

Source	Degree of Freedom	Sum of square	Mean Square	F-Value	Pr> F
Model	4	32.27	8.06	1.09	0.36
Error	595	4344.95	7.40		
Corrected error	599	4377.23			

© COPYRIGHT UPM

ANOVA for b* of leaves at different WRD severity levels

Source	Degree of Freedom	Sum of square	Mean Square	F-Value	Pr> F
Model	4	2441.11	610.27	21.33	<0.0001
Error	595	17019.93	28.60		
Corrected error	599	19461.05			

ANOVA for pH value of soil at different WRD severity levels

Source	Degree of Freedom	Sum of square	Mean Square	F-Value	Pr> F
Model	3	0.36	0.09	2.26	0.07
Error	79	3.19	0.04		
Corrected error	83	3.55			

ANOVA for MCs of soil at different WRD severity levels

Source	Degree of Freedom	Sum of square	Mean Square	F-Value	Pr> F
Model	4	1607.30	401.83	15.14	<.0001
Error	79	2096.44	26.53		
Corrected error	83	3703.74			

ANOVA TSC of latex at different WRD severity levels

Source	Degree of Freedom	Sum of square	Mean Square	F-Value	Pr> F
Model	3	2.71	0.90	7.35	0.0001
Error	120	14.76	0.12		
Corrected error	123	17.47			

© COPYRIGHT UPM

BIODATA OF STUDENT

The student, Siti Saripa Rabiah Binti Mat Lazim, was born on September 6, 1992, in Kota Bharu, Kelantan. She received her primary education at Sekolah Kebangsaan Raja Abdullah and her secondary education at Sekolah Menengah Kebangsaan Melor, both in Kota Bharu, Kelantan. She did her foundation at the Centre of Foundation Studies for Agricultural Science, Universiti Putra Malaysia (UPM). She then furthered her tertiary education at Universiti Putra Malaysia and graduated with a Bachelor of Engineering (Agricultural and Biosystem) in 2015. She completed her Master of Science (Agricultural Mechanization & Automation) in 2019 at UPM. Subsequently, after she received Senate Certification, she registered as a Ph.D candidate in Doctorate of Philosophy at UPM.

© COPYRIGHT UPM

LIST OF PUBLICATIONS

Lazim, S. S. R. M., Nawi, N. M., Sulaiman, Z., & Mustafah, A. M. (2023, December). Performance Analysis of Machine Learning Algorithms for Classification of Infection Severity Levels on Rubber Leaves. In *2023 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)* (pp. 1-6). IEEE.

Lazim, S. S. R. M., N. M. Nawi, Z. Sulaiman, A. M. Mustafah, B. O. Soyoye, N. Abdullah. Prediction and Classification of Chlorophyll Contents to Determine the Severity Levels of White Root Disease Infection on Rubber Leaves. Agricultural Mechanization in Asia, Africa and Latin America, ISSN: 00845841, Volume 54, Issue 12, December, 2023.

CONFERENCE

Lazim, S. S. R. M. (2021). Non-destructive Methods for Early Detection of Plant Disease in Rubber Plantation- A review. 1st International Conference on Plantation Technology.

Lazim, S. S. R. M. (2023). Performance Analysis of Machine Learning Algorithm of Infection Severity Levels on Rubber Leaves. 2nd International Conference on Plantation Technology.

© COPYRIGHT UPM