

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA E
INFORMÁTICA INDUSTRIAL**

WALTER OSWALDO CUTIPA FLORES

**MODELOS DE APRENDIZADO DE MÁQUINA PARA DETECTAR
ESTRESSE ABIÓTICO E CLOROFILA TOTAL EM FOLHAS DE
TRIGO USANDO ESPECTROSCOPIA ÓPTICA**

DISSERTAÇÃO

CURITIBA

2024

WALTER OSWALDO CUTIPA FLORES

**MODELOS DE APRENDIZADO DE MÁQUINA PARA
DETECTAR ESTRESSE ABIÓTICO E CLOROFILA TOTAL EM
FOLHAS DE TRIGO USANDO ESPECTROSCOPIA ÓPTICA**

**Machine Learning Models to Detect Abiotic Stress and Total
Chlorophyll in Wheat Leaves Using Optical Spectroscopy**

Dissertação apresentada como requisito para obtenção do título de Mestre em Engenharia Elétrica e Informática Industrial, do Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial, da Universidade Tecnológica Federal do Paraná (UTFPR).

Orientador: Prof. Dr. José Luís Fabris
Coorientador: Prof. Dr. André Eugênio Lazaretti

CURITIBA

2024



[4.0 Internacional](https://creativecommons.org/licenses/by/4.0/)

Esta licença permite compartilhamento, remixe, adaptação e criação a partir do trabalho, mesmo para fins comerciais, desde que sejam atribuídos créditos ao(s) autor(es).

Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.



Ministério da Educação
Universidade Tecnológica Federal do Paraná
Campus Curitiba



WALTER OSWALDO CUTIPA FLORES

**MODELOS DE APRENDIZADO DE MÁQUINA PARA DETECTAR ESTRESSE ABIÓTICO E CLOROFILA
TOTAL EM FOLHAS DE TRIGO USANDO ESPECTROSCOPIA ÓPTICA**

Trabalho de pesquisa de mestrado apresentado como requisito para obtenção do título de Mestre Em Ciências da Universidade Tecnológica Federal do Paraná (UTFPR). Área de concentração: Fotônica Em Engenharia.

Data de aprovação: 08 de Outubro de 2024

Dr. Jose Luis Fabris, Doutorado - Universidade Tecnológica Federal do Paraná

Dr. Luiz Felipe Ribeiro Barrozo Toledo, Doutorado - Instituto de Tecnologia para Desenvolvimento (Lactec)

Dr. Uilian Jose Dreyer, Doutorado - Universidade Tecnológica Federal do Paraná

Documento gerado pelo Sistema Acadêmico da UTFPR a partir dos dados da Ata de Defesa em 08/10/2024.

This work is dedicated to my mother, my source
of motivation.

ACKNOWLEDGEMENTS

I sincerely want to thank my esteemed advisors, Dr. José Luís Fabris, Dr. André Eugênio Lazzaretti, and Dra. Márcia Müller, whose patience, insightful guidance, and constant support have been fundamental in the development of this research. Their profound knowledge and experience have been key to the progress of this investigation, and receiving their guidance has been a privilege.

I also wish to extend my gratitude to my colleagues at the laser laboratory, whose invaluable support and assistance have greatly contributed.

I am deeply thankful to my beloved family for their support, especially to my mother, a woman of few words but a true example to follow.

I am deeply grateful to the Federal University of Technology – Paraná for providing the resources and environment necessary for the successful completion of this project.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001

"Curiosity is more important than knowledge."

(Albert Einstein)

RESUMO

FLORES, Walter Oswaldo. **Modelos de Aprendizado de Máquina para Detectar Estresse Abiótico e Clorofila Total em Folhas de Trigo Usando Espectroscopia Óptica**. 2024. 118 f. Dissertação (Mestrado em Engenharia Elétrica e Informática Industrial) – Universidade Tecnológica Federal do Paraná. Curitiba, 2024.

A detecção do estresse abiótico e a estimativa da clorofila é fundamental para determinar a saúde das plantas. Existem diferentes métodos para estimar e detectar esses valores nas plantas com espectroscopia óptica. Os métodos podem ser destrutivos (espectroscopia de absorção) ou não destrutivos (espectroscopia de reflexão). Assim, o objetivo principal desta pesquisa é desenvolver modelos de inteligência artificial para detectar estresse abiótico e estimar a clorofila total em trigo por meio de espectroscopia de reflexão difusa. Foi realizado um experimento para gerar quinze condições de estresse diferentes, incluindo iluminação e quantidade de água, a partir de sementes de trigo que germinaram e cresceram até o sétimo dia. Com as amostras de folhas de trigo, foram obtidos espectros de reflexão e absorção. Na análise para a classificação por estresse hídrico (3ml, 5ml e 10ml), foram comparados vários modelos MiniRocket com técnicas de pré-processamento, como o filtro Savitzky-Golay, a remoção de contínuo (CR), a remoção de tendência (DT), a normalização dos espectros de reflexão por área, a variante padrão normal (SNV) e a correção multiplicativa de dispersão (MSC). Os resultados mostram que os modelos que empregam SNV e CR alcançaram precisão perfeita para dados de validação. Para a classificação por estresse luminoso (LED verde, vermelho, azul, branco e um escuro), os modelos avaliados com diferentes técnicas de pré-processamento atingiram precisões que variam entre 0.841 e 0.986, onde o melhor modelo obtido para a classificação do estresse luminoso alcançou uma precisão de 0.986 para dados de validação. A previsão do conteúdo de clorofila em folhas de trigo a partir de espectroscopia de reflexão difusa multiangular foi abordada utilizando modelos de aprendizado de máquina que incluem rede neural profunda (DNN) e MiniRocket, revelando que tanto DNN quanto MiniRocket apresentaram alto desempenho na previsão do conteúdo de clorofila. MiniRocket mostrou tempos de processamento mais curtos (tão baixos quanto 2.667 s e até 6.840 s), mas com um coeficiente de determinação de 0.928 para o melhor modelo. DNN teve tempos de processamento mais longos (variando de 76.50 s a 152.35 s), mas alcançou coeficientes de determinação de 0.993 no conjunto de dados de validação. Finalmente, conclui-se que a pesquisa demonstrou a eficácia de combinar a espectroscopia de reflexão multiangular com técnicas de aprendizado de máquina para detectar estresse e estimar o conteúdo de clorofila em folhas de trigo. Essa abordagem fornece um método mais preciso e robusto, abrindo novas perspectivas e possíveis direções para futuras pesquisas em análise de plantas e aplicações espectrais.

Palavras-chave: Espectroscopia de Refletância Difusa. Clorofila. Aprendizado de máquina. Estresse Abiótico. Minirocket.

ABSTRACT

FLORES, Walter Oswaldo. **Machine Learning Models to Detect Abiotic Stress and Total Chlorophyll in Wheat Leaves Using Optical Spectroscopy**. 2024. 118 p. Dissertation (Master's Degree in Electrical Engineering and Industrial Informatic) – Universidade Tecnológica Federal do Paraná. Curitiba, 2024.

The detection of abiotic stress and the estimation of chlorophyll are fundamental for determining plant health. There are different methods to estimate and detect these values in plants using optical spectroscopy. The methods can be destructive (absorption spectroscopy) or non-destructive (reflection spectroscopy). Thus, the main objective of this research is to develop artificial intelligence models to detect abiotic stress and estimate total chlorophyll in wheat through diffuse reflection spectroscopy. An experiment was conducted to generate fifteen different stress conditions, including lighting and water amount, from wheat seeds that germinated and grew until the seventh day. With the wheat leaf samples, reflection and absorption spectra were obtained. In the analysis for water stress classification (3ml, 5ml, and 10ml), several MiniRocket models were compared with preprocessing techniques such as the Savitzky-Golay filter, Continuum Removal (CR), De-trending (DT), Normalization of Reflectance Spectra by Area, Standard Normal Variate (SNV), and Multiplicative Scatter Correction (MSC). The results show that models using SNV and CR achieved perfect precision for validation data. For light stress classification (green, red, blue and white LEDs and one dark), models evaluated with different preprocessing techniques achieved precisions ranging from 0.841 to 0.986, where the best model obtained for light stress classification achieved a precision of 0.986 for validation data. The prediction of chlorophyll content in wheat leaves from multi-angular diffuse reflection spectroscopy was addressed using machine learning models including deep neural network (DNN) and MiniRocket, revealing that both DNN and MiniRocket achieved high performance in predicting chlorophyll content. MiniRocket showed shorter processing times (as low as 2.667 s and up to 6.840 s) but with a coefficient of determination of 0.928 for the best model. DNN had longer processing times (ranging from 76.50 s to 152.35 s) but achieved coefficients of determination of 0.993 for the validation dataset. Finally, it is concluded that the research demonstrated the effectiveness of combining multi-angular reflection spectroscopy with machine learning techniques to detect stress and estimate chlorophyll content in wheat leaves. This approach provides a more accurate and robust method, opening new perspectives and possible directions for future research in plant analysis and spectral applications.

Keywords: Diffuse Reflectance Spectroscopy. Chlorophyll. Machine learning. Abiotic Stress. Minirocket.

LIST OF FIGURES

Figure 1 – Map of this document, summarizing the structure and content of the Chapters	20
Figure 2 – Chemical structures of chlorophyll a and chlorophyll b in acetone	23
Figure 3 – Characteristic spectra of chlorophyll a and b. Chlorophyll a exhibits a prominent peak in both the blue light and red light bands	24
Figure 4 – Characteristic graph of chlorophyll a and b absorption at the peaks of 663 nm for chlorophyll a and 645 nm for chlorophyll b	25
Figure 5 – The characteristic graph of the reflection spectrum in wheat leaves is shown, with a peak intensity around 550 nm (green band) and a plateau around 750 nm	26
Figure 6 – Water distillation equipment used in the experiment (Evoxx water distiller operating at 127V or 220V, 750W).	41
Figure 7 – HR4000 Spectrometer with internal components and numbered operational steps. The figure illustrates the complete process from light input to spectral conversion	42
Figure 8 – Digital spectrophotometer Kasuaki IL-593-S, with a wavelength range of 190 to 1100 nm, (KASUAKI, 2024).	43
Figure 9 – SpectraSuite Software	44
Figure 10 – Physical implementation of the reflectance spectroscopy setup in the laboratory Laser at UTFPR. The configuration includes a halogen light source, a silver mirror, an iris diaphragm, and a lens. Light is directed onto a sample in a Petri dish, and the reflected light is collected by a fiber collimator, connected to an HR4000 spectrometer for analysis	45
Figure 11 – Diagram of the reflectance spectroscopy setup in the laboratory Laser at UTFPR. The configuration includes a halogen light source, a silver mirror, an iris diaphragm, and a lens. Light is directed onto a sample in a Petri dish, and the reflected light is collected by a fiber collimator, connected to an HR4000 spectrometer for analysis	46
Figure 12 – Physical implementation of the absorbance spectroscopy configuration. A halogen light source (Philips 12345SL, 3100 K) and two optical fibers connected to a cuvette holder (CUV-ALL-UV, Ocean Optics) with 4.5 mL polystyrene cuvettes (K42-045, Olen) and the UV-Vis spectrometer (HR4000, Ocean Optics)	47
Figure 13 – Diagram of the absorbance spectroscopy configuration. A halogen light source (Philips 12345SL, 3100 K) and two optical fibers connected to a cuvette holder (CUV-ALL-UV, Ocean Optics) with 4.5 mL polystyrene cuvettes (K42-045, Olen) and the UV-Vis spectrometer (HR4000, Ocean Optics)	48
Figure 14 – Absorption spectra acquisition process. To obtain the absorption spectra, the software performs the following steps: First, the intensity of light transmitted through a cuvette with 80% acetone (R_λ) is measured, as shown in Figure 14(a). For D_λ , the signal measured in the absence of incident light, the light source is blocked (Figure 14(b))	49
Figure 15 – Graph of the spectrum for the intensity of light transmitted through a cuvette with 80% acetone (R_λ), starting from Figure 14(a). The dark reference spectrum, shown in Figure 14(b), represents the spectrum obtained under these conditions for D_λ , which indicates the absence of incident light	49

Figure 16 – Stress containers with different LED lighting types: Green LED, Red LED, Blue LED, No LED, and White LED	50
Figure 17 – LED strips installed on the lids of the containers, arranged in a spiral pattern.	51
Figure 18 – Spectra of LED lights in containers	51
Figure 19 – Diagram of the digital timer connected to a power distributor. Visual configuration for observing the time intervals in which the LED strips are turned on and off during the 7-day experiment.	52
Figure 20 – Stress Distribution in Containers for Water Stress Experiments (3 ml, 5 ml, 10 ml)	53
Figure 21 – Distribution of DHT11 Sensors for Monitoring Temperature and Humidity of Containers for Plant Stress	54
Figure 22 – Implementation of a monitoring system in stress containers for plants, composed of DHT11 sensors, Arduino Uno, Raspberry Pi 3B+	55
Figure 23 – Obtaining samples from different positions and angles	56
Figure 24 – Interpolated Wavelength Data	59
Figure 25 – Overview of the classification process	62
Figure 26 – Overview of the regression process	63
Figure 27 – Deep Learning Model Architecture with ReLU, Dropout 0.2, and dense layers (256, 64, 32 neurons) for Chlorophyll prediction	67
Figure 28 – Graph of data collected by five temperature sensors in containers under various lighting conditions (green, red, blue, white LEDs, and no lighting). Each sensor recorded 10,502 measurements over a period of 7 days	72
Figure 29 – Graph of data collected by five Humidity sensors in containers under various lighting conditions (green, red, blue, white LEDs, and no lighting). Each sensor recorded 10,502 measurements over a period of 7 days	74
Figure 30 – Temperature on day 4, over 24 hours.	76
Figure 31 – Humidity on day 4, over 24 hours	76
Figure 32 – Containers with wheat leaves after 7 days of growth under 5 ml distilled water stress	77
Figure 33 – Wheat leaf samples after 7 days of growth under 9 different stress conditions. (a) Green LED, 3 ml (33(a)), (b) Green LED, 5 ml (33(b)), (c) Green LED, 10 ml (33(c)), (d) Red LED, 3 ml (33(d)), (e) Red LED, 5 ml (33(e)), (f) Red LED, 10 ml (33(f)), (g) Blue LED, 3 ml (33(g)), (h) Blue LED, 5 ml (33(h)), (i) Blue LED, 10 ml (33(i))	78
Figure 34 – Wheat leaf samples after 7 days of growth under various stress conditions. (a) Without LED, 3 ml (34(a)), (b) Without LED, 5 ml (34(b)), (c) Without LED, 10 ml (34(c)), (d) White LED, 3 ml (34(d)), (e) White LED, 5 ml (34(e)), (f) White LED, 10 ml (34(f))	79
Figure 35 – Multi-angular reflection spectra of fresh wheat leaves stressed under three LED colors: (a) Green LED, 3 ml (35(a)), (b) Green LED, 5 ml (35(b)), (c) Green LED, 10 ml (35(c)), (d) Red LED, 3 ml (35(d)), (e) Red LED, 5 ml (35(e)), (f) Red LED, 10 ml (35(f)), (g) Blue LED, 3 ml (35(g)), (h) Blue LED, 5 ml (35(h)), and (i) Blue LED, 10 ml (35(i))	80
Figure 36 – Multi-angular reflectance spectra of fresh wheat leaves, stressed without LED and with white LED in different volumes of distilled water. (a) No LED, 3 ml (36(a)), (b) No LED, 5 ml (36(b)), (c) No LED, 10 ml (36(c)), (d) White LED, 3 ml (36(d)), (e) White LED, 5 ml (36(e)), (f) White LED, 10 ml (36(f))	81

Figure 37 – Standard deviation of reflectance spectra for wheat leaf samples after 7 days of growth under 15 different stress conditions (SC). Each subplot Figure 37(a) to 37(o) represents the standard deviation of reflectance measurements across wavelengths from 450 nm to 780 nm for one of the 15 stress conditions, including water and light stress	82
Figure 38 – Extracted samples showing different stress conditions across containers: (a) Green LED light for 3 ml, 5 ml, 10 ml of distilled water, (b) Red LED light for 3 ml, 5 ml, 10 ml of distilled water, (c) Blue LED light for 3 ml, 5 ml, 10 ml of distilled water, (d) No light for 3 ml, 5 ml, 10 ml of distilled water, and (e) White LED light for 3 ml, 5 ml, 10 ml of distilled water	84
Figure 39 – Absorption spectra of 15 liquid chlorophyll samples. The spectra show intensity as a function of wavelength (in nm) for each sample, representing a different stress condition (S.C. 01 to S.C. 15)	86
Figure 40 – Visual comparison of the distributions of Chlorophyll a, Chlorophyll b, and Total Chlorophyll measurements obtained with two different equipment setups. Each boxplot illustrates the variation in the measurements and allows for the evaluation of the impact of different techniques on the results, according to the data presented in Table 11	88
Figure 41 – Comparison of absorption spectra obtained using two different equipment setups: a spectrophotometer IL-593-S providing two spectral indices (645 nm and 663 nm) and an HR4000 spectrometer that provides a curve describing intensity at different wavelengths	89
Figure 42 – Confusion matrices showing the performance of different models for water stress classification as presented in Table 13 (for the test dataset). The images are confusion matrices corresponding to the following models: (a) Model 26, (b) Model 9, (c) Model 10, (d) Model 1, (e) Model 20, and (f) Model 4. Each subimage visualizes the model’s performance in classifying different levels of water stress, as detailed in the associated tables (Table 14, Table 15, Table 16, Table 17)	92
Figure 43 – Confusion matrices showing the performance of different models for light stress classification for the test dataset. (a) Model 01: with an accuracy of 0.986, (b) Model 08: with an accuracy of 0.971, (c) Model 29: with an accuracy of 0.971, (d) Model 14: with an accuracy of 0.971, (e) Model 12: with an accuracy of 0.971, (f) Model 18: with an accuracy of 0.971	96
Figure 44 – Comparação do conteúdo total de clorofila previsto versus medido para diferentes modelos de regressão MiniRocket utilizando dados de teste	100
Figure 45 – Performance of DNN Models for Estimating Chlorophyll Content in Leaves with Test Data	103

LIST OF TABLES

Table 1 – List of research analysis in plant stress based on Vis-NIR spectroscopy	32
Table 2 – Chlorophyll Estimation Using Artificial Intelligence	38
Table 3 – Chlorophyll Estimation Using Artificial Intelligence (continue)	39
Table 4 – Stress conditions and details for each container. The conditions include the type of light (LED) applied and the amount of water stress (in ml) for each container.	53
Table 5 – Statistical results of data collected by five temperature sensors in containers under various lighting conditions (green, red, blue, white LEDs, and no lighting). Each sensor recorded 10,502 measurements over a period of 7 days	73
Table 6 – Correlation between temperature sensors	73
Table 7 – Statistical results of data collected by five Humidity sensors in containers under various lighting conditions (green, red, blue, white LEDs, and no lighting). Each sensor recorded 10,502 measurements over a period of 7 days	75
Table 8 – Correlation Matrix between humidity sensors	75
Table 9 – Chlorophyll measurements obtained with the Spectrophotometer IL-593-S. The table shows the mass (in grams) and the optical densities at A_{645} and A_{663} , as well as the concentrations of chlorophyll a, b, and total (in $\text{mg} \cdot \text{g}^{-1}$) for each sample	85
Table 10 – Comparison of chlorophyll measurements obtained with the Digital Spectrophotometer IL-593-S and the HR4000 spectrometer. OD values at A_{645} and A_{663} , and concentrations of chlorophyll a, b, and total for both devices are shown. The percentage of total chlorophyll is shown in Pct. for each device.	87
Table 11 – Descriptive statistics of chlorophyll measurements obtained with the Digital Spectrophotometer IL-593-S and the HR4000 spectrometer. The table presents the mean, median, standard deviation, variance, maximum, and minimum for OD values (A_{645} and A_{663}) and chlorophyll concentrations a, b, and total	87
Table 12 – Pearson correlation between chlorophyll measurements from two different equipment setups	90
Table 13 – Performance metrics of the MiniRocket models for classification tasks. The table presents the accuracy and processing time for various models using different preprocessing techniques on two datasets (Data1 and Data2). The accuracy values indicate the performance of each model for the test dataset	91
Table 14 – Metrics for test data: Model 26, 9 and 10	93
Table 15 – Metrics for test data: Model 1	93
Table 16 – Metrics for test data: Model 20	93
Table 17 – Metrics for test data: Model 4	93
Table 18 – Comparison of Models for Light Stress Classification Based on Different Data Preprocessing Techniques, Accuracy (for the test dataset), and Execution Time	95
Table 19 – Metrics for test data 1	97
Table 20 – Metrics for test data 8	97
Table 21 – Metrics for test data 29	97

Table 22 – Performance metrics of various MiniRocket regression models on different datasets. The table shows R^2 , root mean squared error (RMSE), and mean absolute error (MAE) for both the test and train sets, along with the time taken for each model. Models are evaluated using different combinations of data preprocessing techniques and feature sets	99
Table 23 – Performance metrics of various DNN regression models on different datasets. The table shows R^2 , root mean squared error (RMSE), and mean absolute error (MAE) for both the test and train sets, along with the time taken for each model. Models are evaluated using different combinations of data preprocessing techniques and feature sets	102

LIST OF ACRONYMS

ABBREVIATIONS

abs.	Absorbance
art.	Article
cap.	Charter
fig.	Figure
sec.	Section
tab.	Table

INITIALISM

AC	Alternate Current
ANN	Artificial Neural Networks
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil
CNPq	Conselho Nacional de Desenvolvimento Científico e Tecnológico
CR	Continuum Removal
CPGEI	Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial
DC	Direct Current
DNN	Deep Neural Network
EWT	Equivalent Water Thickness
FM	Fluorometric Methods
FW	Fresh Weight
GPR	Gaussian Process Regression
HPLC	High-Performance Liquid Chromatography
HRS	Hyperspectral Reflectance Spectroscopy
HI	Hyperspectral Imaging
IA	Artificial Intelligence
ISO	International Organization for Standardization
LED	Light Emitting Diode
LDA	Linear Discriminant Analysis
MARS	Multivariate Adaptive Regression Splines
MAE	Mean Absolute Error
MSC	Multiplicative Scatter Correction
MDR	Modified Spectral Index
NIRS	Near-Infrared Reflectance Spectroscopy
NDVI	Normalized Difference Vegetation Index
OD	Optical Density
OLSR	Ordinary Least Squares Regression

OSAVI	Optimized Soil Adjusted Vegetation Index
PCA	Principal Component Analysis
PLSR	Partial Least Squares Regression
PPV	Proportion of positive values
PRI	Photochemical Reflectance Index
RF	Random Forest
RGB	Red, Green, Blue
RMSE	Root Mean Squared Error
RDVI	Renormalized Difference Vegetation Index
SMLR	Stepwise Multiple Linear Regression
SNV	Standard Normal Variate
SVR	Support Vector Regression
SVM	Support Vector Machine
TP	True Positives
UV	Ultraviolet
VAC	Voltage Alternating Current
USB	Universal Serial Bus
XGBoost	Extreme Gradient Boosting

ACRONYMS

Chl a	Chlorophyll a
Chl b	Chlorophyll b
Chl total	Total Chlorophyll
CIred edge	Red Edge Chlorophyll Index
DHT11	Digital Humidity and Temperature Sensor Model 11
HR4000	High-Resolution Fiber Optic Spectrometer Model 4000
MINIROCKET	Minimally Random Convolutional Kernel Transform
Nor. Area	Normalization by Area
ReLU	Rectified Linear Unit
RMSprop	Root Mean Square Propagation
SPAD	Soil and Plant Analysis Development
VIS	Visible Light
VNIR	Visible-Near Infrared Spectroscopy

LIST OF SYMBOLS

LATIN LETTERS

°C	Degree Celsius	[C]
cm	Centimeter	[cm]
cm ²	Square centimeter	[cm ²]
H	Hour	[H]
Hz	Hertz	[s ⁻¹]
liter	Liter	[L]
mA	Milliampere	[mA]
mg·g ⁻¹	Milligrams per gram	[mg·g ⁻¹]
ml	Milliliter	[ml]
nm	Nanometer	[nm]
OD	Optical Density	
rpm	Revolutions per minute	[rpm]
s	Seconds	[s]
VAC	Volts in Alternating Current	[V]
VDC	Volts in Direct Current	[V]
W	Watt	[W]

GREEK LETTERS

λ	Wavelength	[λ]
μ m	Micrometer	[μ m]
σ	Standard deviation	[σ]
f	Focal length	[f]

SUBSCRIPTS

A_λ	Absorbance measured at a wavelength of λ nm
R_λ	Reflectance measured at a wavelength of λ nm

CONTENTS

1	INTRODUCTION	18
1.1	GENERAL OBJECTIVE	18
1.2	SPECIFIC OBJECTIVES	19
1.3	CONTRIBUTIONS	19
1.4	STRUCTURE OF THIS DOCUMENT	20
2	LITERATURE REVIEW	22
2.1	WHEAT PHYSIOLOGY	22
2.2	METHODS FOR QUANTIFYING CHLOROPHYLL IN LEAVES	22
2.2.1	Absorption Spectrophotometry	24
2.2.2	Reflectance Spectroscopy	25
2.2.3	Fluorometric Methods	26
2.2.4	Portable Chlorophyll Meters (SPAD)	26
2.2.5	Image-based Methods	27
2.2.6	Chemical Methods	27
2.3	DETECTION OF STRESS IN PLANTS BY REFLECTANCE SPECTROSCOPY	27
2.4	CHLOROPHYLL ESTIMATION BY REFLECTANCE SPECTROSCOPY	32
3	METHODS	40
3.1	EQUIPMENT	40
3.1.1	Water Destiller	40
3.1.2	HR4000 Spectrometer	41
3.1.3	Digital Spectrophotometer IL-593-S	43
3.1.4	SpectraSuite Software	44
3.1.5	Reflectance Spectroscopy Configuration	45
3.1.6	Absorbance Spectroscopy Configuration	47
3.2	ENVIRONMENTAL CONSTRUCTION FOR WHEAT STRESS	50
3.2.1	Light Stress	50
3.2.1.1	Lighting Time Control	52
3.2.2	Water Stress	52
3.2.3	Stress Conditions	53
3.2.4	Temperature and Humidity Monitoring	54
3.3	DATABASE OF REFLECTION AND ABSORPTION SPECTRA	56
3.3.1	Acquisition of Reflection Spectra	56
3.3.2	Methodology for Chlorophyll Extraction and Quantification	56
3.4	PRE-PROCESSING	57
3.4.1	Processing Spectral Data by Wavelength Intervals	58
3.4.2	Savitzky-Golay Filter	59
3.4.3	Continuum Removal (CR)	60
3.4.4	De-trending (DT)	60
3.4.5	Normalization of Reflectance Spectra by Area	60
3.4.6	Standard Normal Variate (SNV)	61
3.4.7	Multiplicative Scatter Correction (MSC)	61

3.5	MACHINE LEARNING MODEL	62
3.5.1	Classification Models	63
3.5.1.1	MiniRocket Classifier	63
3.5.2	Regression Models	65
3.5.2.1	MiniRocket Regressor	65
3.5.2.2	DNN Regressor	66
3.6	EVALUATION METRICS	67
3.6.1	Evaluation Metrics for Model Classifiers	68
3.6.2	Regression Model Evaluation Metrics	69
3.6.3	Comparison of Values from Different Equipment	70
4	RESULTS AND DISCUSSIONS	71
4.1	WATER AND LIGHT STRESS ENVIRONMENT FOR WHEAT LEAVES	71
4.1.1	Result of Temperature Conditions in Stress Environment	71
4.1.2	Result of Humidity Conditions in Stress Environment	73
4.1.3	Interaction Between Temperature and Humidity Under Stress Conditions . .	75
4.2	SPECTRAL DATABASE FOR REFLECTION AND ABSORPTION .	77
4.2.1	Analysis of the Reflectance Spectra	77
4.2.2	Analysis of the Absorption Spectra	83
4.2.2.1	Results for Spectrophotometer IL-593-S	84
4.2.2.2	Results for HR4000 Spectrometer	85
4.2.2.3	Comparison of Total Chlorophyll Results	86
4.3	MACHINE LEARNING MODELS TO DETECT WATER AND LIGHT STRESS	90
4.3.1	Water Stress Classification using AI	90
4.3.2	Light Stress Classification using AI	94
4.4	MACHINE LEARNING MODELS TO ESTIMATE THE AMOUNT OF TOTAL CHLOROPHYLL	98
4.4.1	MiniRocket-Based Model for Total Chlorophyll Estimation in Wheat Leaves	98
4.4.2	Deep Neural Network-Based Model for Total Chlorophyll Estimation in Wheat Leaves	101
5	FINAL CONCLUSIONS AND FUTURE WORKS	105
5.1	COMPLIANCE WITH SPECIFIC OBJECTIVES	106
5.2	FUTURE WORKS	107
	REFERENCES	108

1 INTRODUCTION

The current climate changes are making agriculture face increasingly larger challenges (PEDERSEN *et al.*, 2022). Additionally, the growing world population increases the demand for food (VALIN *et al.*, 2014), making plant health analysis increasingly required and essential for boosting agricultural production. Furthermore, techniques such as neural networks and machine learning have emerged as powerful and reliable tools for analyzing datasets that could not be efficiently managed with traditional methods. The ability of artificial intelligence (AI) to process large amounts of data and extract relevant patterns allows for early detection of abiotic or biotic stress, enabling better crop yields (JHA *et al.*, 2019).

Among the most studied problems in agriculture is the early detection of plant stress. Chlorophyll in leaves is a fundamental indicator for determining plant health. Traditional methods for quantifying chlorophyll content in leaves use optical equipment and techniques such as absorbance spectroscopy and Reflectance spectroscopy. The main differences between these techniques lie in that absorbance spectroscopy is a destructive method that requires obtaining a solution for subsequent analysis. On the other hand, reflectance spectroscopy is a non-destructive technique that allows analysis only with the incidence of light on the leaf. The procedure for quantifying chlorophyll through reflectance spectroscopy in leaves consists of a reading through spectrometer adapters or portable equipment that obtain the count directly on the leaf or through a specific angle of light that will be reflected towards the interrogator.

The ability to detect chlorophyll count as well as detect a certain stress in leaves using a multi-angular setup (obtaining reflection spectra for different angles of light incidence) can significantly improve the efficiency in chlorophyll quantification and stress detection in a more robust manner. This dissertation focuses on the analysis of stress detection and chlorophyll estimation in wheat leaves using multi-angular reflectance spectroscopy. The following subsections describe the objectives, fundamental contributions, and structural scheme.

1.1 GENERAL OBJECTIVE

Develop machine learning models to detect abiotic stress and total chlorophyll in wheat leaves using optical spectroscopy.

1.2 SPECIFIC OBJECTIVES

To achieve the general objective mentioned in Section 1.1, several specific objectives have been established. These objectives serve as tangible reference points to guide the research towards the development of an efficient, safe, and simple procedure for detecting stress and estimating total chlorophyll in wheat leaves through optical spectroscopy. Each specific objective addresses a fundamental aspect of the research process, including the construction of the stress environment, obtaining reflection and absorption spectra, as well as analysis with artificial intelligence and validation.

1. Create a water and light stress environment for wheat leaves.
2. Produce a spectral database for reflection, for samples obtained under different angles of light incidence, and absorption.
3. Discover machine learning models to detect water and light stress, with reflection spectra for samples obtained under different angles of light incidence.
4. Discover machine learning models to estimate the amount of total chlorophyll, with reflection spectra for samples obtained under different angles of light incidence.

1.3 CONTRIBUTIONS

This section highlights three main contributions of the research:

- **Light stress identification through artificial intelligence techniques with reflection spectra for samples obtained under different angles of light incidence:** An experiment was implemented to stress wheat leaves with 4 containers with different types of lighting and one container without lighting. The stress was applied from germination to its growth in 7 days. By acquiring diffuse reflection spectra, a model was trained to classify the type of light stress to which the wheat plants were subjected.
- **Water stress identification through artificial intelligence techniques with reflection spectra for samples obtained under different angles of light incidence:** An experiment was implemented to stress wheat leaves, where 3 containers were set up with different amounts of water at 3 ml, 5 ml, and 10 ml, from germination to growth in 7 days. By

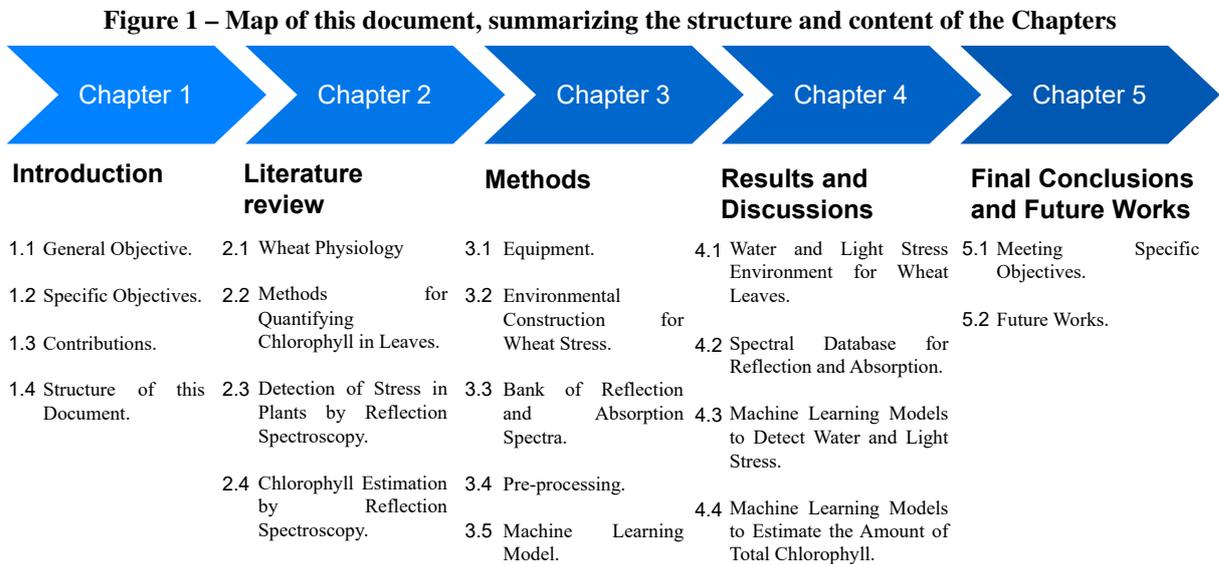
acquiring diffuse reflection spectra, a model was trained to classify the type of stress to which the plants were subjected, enabling the detection and classification of water stress in plants.

- **Chlorophyll amount estimation with reflection spectra for samples obtained under different angles of light incidence:** The experiment was implemented allowing spectrum acquisition at different angles of light incidence, these diffuse reflection spectra are related to a specific amount of chlorophyll with which a model can be trained to predict the chlorophyll amount through sample acquisition at different angles.

These three contributions are fundamental to advancing knowledge in the field of crops and stress detection in plants through optical spectroscopy.

1.4 STRUCTURE OF THIS DOCUMENT

Figure 1, presents a visual map that illustrates the structure of this research work, providing an overview of the organization and different components that make up this research work.



Source: Own authorship (2024).

This document is organized as follows:

- Chapter 1, sets the context by introducing the research problem and the general and specific objectives, as well as the structure of the dissertation.

- Chapter 2, presents a description of wheat physiology, considerations on chlorophyll, and a literature review of research work related to the work carried out with reference to stress detection and chlorophyll estimation through reflectance spectroscopy. Section 2.2, highlights the different methods to estimate chlorophyll in leaves. Section 2.3, describes the scientific work that addresses the detection of different types of stress in leaves, highlighting the spectroscopy technique used, the spectral region employed in the approach, the type of abiotic stress detected, the plant species, and the type of analysis performed. Section 2.4, describes the scientific work that addresses chlorophyll estimation in leaves through reflectance spectroscopy.
- In Chapter 3, the entire methodology applied in our research is shown. Section 3.2, describes the construction of a stress environment for wheat plants, which considers 15 different stress conditions, as well as the required electronic implementation for time control (on and off), and monitoring of temperature and humidity. Section 3.3, shows the data collection method for constructing our reflection and absorption spectrum database. Section 3.4, describes the preprocessing methods applied and the filter concepts used. Section 3.5, presents the models used for stress classification by water and light and total chlorophyll content estimation in leaves. Section 3.6, describes the metrics that will validate the classification and regression models.
- Chapter 4, offers the presentation of the results obtained in the research through the methodology explained in Chapter 3. In Section 4.1, the results related to the environmental conditions of the containers (temperature and humidity) are shown. In Section 4.2, the analysis of the data obtained for reflection and absorption spectra is presented, as well as the results for the classification of abiotic stress (water and light) and estimation of chlorophyll for multi-angular reflection spectra.
- In Chapter 5, an analysis of the research findings is presented, providing valuable conclusions derived from the research.

2 LITERATURE REVIEW

The objective of this chapter is to describe important aspects of wheat, the chemical structure of chlorophyll and its characterization, as well as to review the scientific literature in search of methods and techniques to quantify chlorophyll in leaves, detect stress in plant leaves, and estimate total chlorophyll using reflectance spectroscopy. The review emphasizes the spectroscopy technique used, the type of abiotic stress, particularly water and light stress, the types of plants, and the statistical analysis or artificial intelligence technique applied.

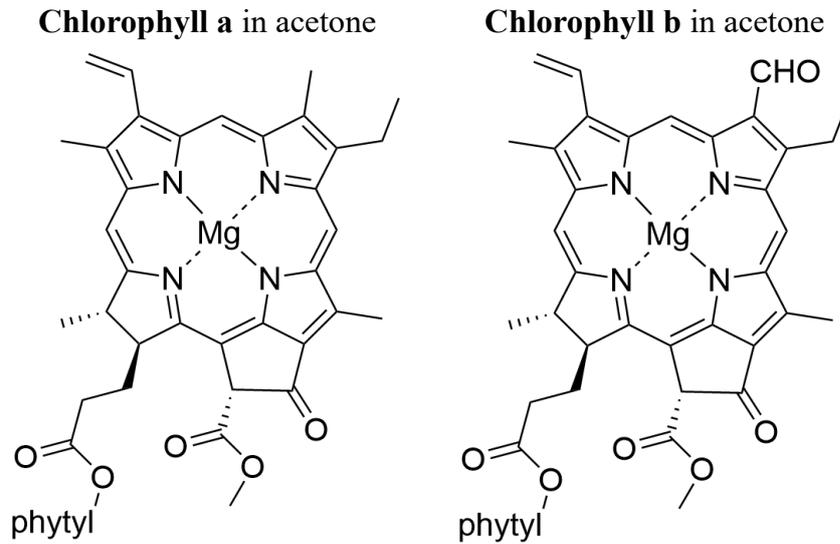
2.1 WHEAT PHYSIOLOGY

Wheat (*Triticum aestivum L.*) is one of the most important food crops worldwide, with an optimal growth temperature between 15 and 20°C. Heat stress negatively affects plant growth and development, including photosynthesis in the leaves (ULLAH *et al.*, 2022). Wheat productivity is affected by various biotic and abiotic factors. Heat stress and drought are two significant abiotic factors that limit wheat production. Drought, characterized by prolonged periods of water scarcity, affects the ability of plants to grow and reproduce under low water availability. Heat stress negatively impacts photosynthesis, accelerates leaf senescence, reduces leaf area, and decreases yield (TYAGI; PANDEY, 2022).

2.2 METHODS FOR QUANTIFYING CHLOROPHYLL IN LEAVES

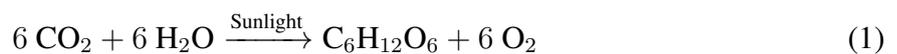
Chlorophyll is a crucial biochemical component in photosynthesis, allowing plants and other photosynthetic organisms to convert solar energy into oxygen. The determination of chlorophyll content in plants can be used for the implementation of corrective measures for plant development. The atomic structure of chlorophyll a and b in acetone is shown in Figure 2. Chlorophyll a is the most abundant form and gives plants their green color, while other forms such as chlorophyll b, c, and d are present in smaller amounts (YSI Environmental, 2024).

Figure 2 – Chemical structures of chlorophyll a and chlorophyll b in acetone



Source: (KOBAYASHI *et al.*, 2013).

In Equation 1, the process by which plants, algae, and some bacteria convert sunlight into chemical energy is shown. In this reaction, carbon dioxide (CO₂) and water (H₂O) are the reactants. As a result of this reaction, glucose (C₆H₁₂O₆), a sugar that plants use as an energy source and material for their growth, and oxygen (O₂) are produced and released into the environment. Thus, photosynthesis allows plants to transform carbon dioxide and water, using the energy emitted by sunlight, into glucose and oxygen (LÓPEZ; SILVA, 2024; YSI Environmental, 2024).

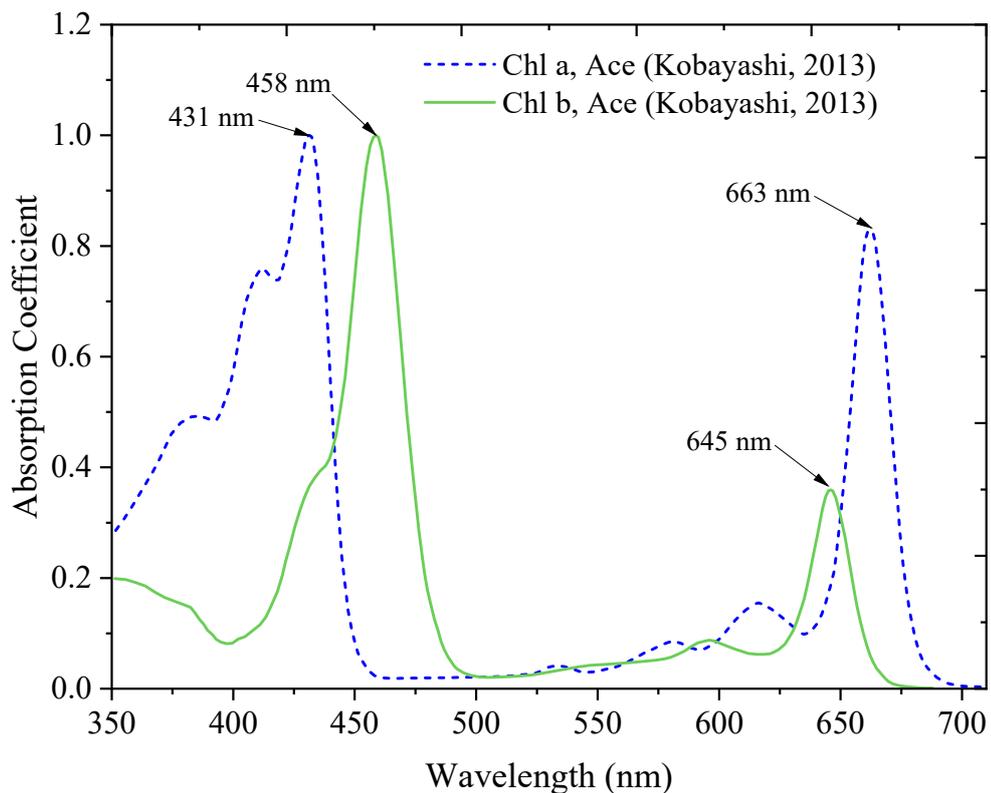


Total chlorophyll content is the sum of chlorophyll a and b, where chlorophyll a participates in light harvesting and the conversion of absorbed photon energy into chemical energy. chlorophyll a is indispensable for oxygenic photosynthesis, being the only member of the chlorophyll family present in all organisms that perform oxygenic photosynthesis. Depending on its protein environment, it can function as a light harvester or as a redox participant in the capture of electronic excitations and electron transport (BJÖRN *et al.*, 2009). Chlorophyll b, although less abundant than chlorophyll a, is crucial for light capture in plants as it is converted to chlorophyll a for subsequent degradation. chlorophyll b is synthesized from chlorophyll a and catabolized after being reconverted to chlorophyll a (SHIMODA *et al.*, 2012), (WILLSTÄTTER, 1915), (EGGINK *et al.*, 2001).

In Figure 3, the characteristic absorption spectra of chlorophyll a and b are shown.

Chlorophyll a and b exhibit absorption peaks at different wavelengths within the visible spectrum. Chlorophyll a shows a prominent peak in the 431 and 663 nm, while chlorophyll b has peaks in the 458 and 645 nm. These peaks indicate the wavelengths of light that chlorophylls absorb most efficiently, which is essential for maximizing the efficiency of the photosynthetic process (TANIGUCHI; LINDSEY, 2021; KOBAYASHI *et al.*, 2013).

Figure 3 – Characteristic spectra of chlorophyll a and b. Chlorophyll a exhibits a prominent peak in both the blue light and red light bands



Source: Based on (KOBAYASHI *et al.*, 2013).

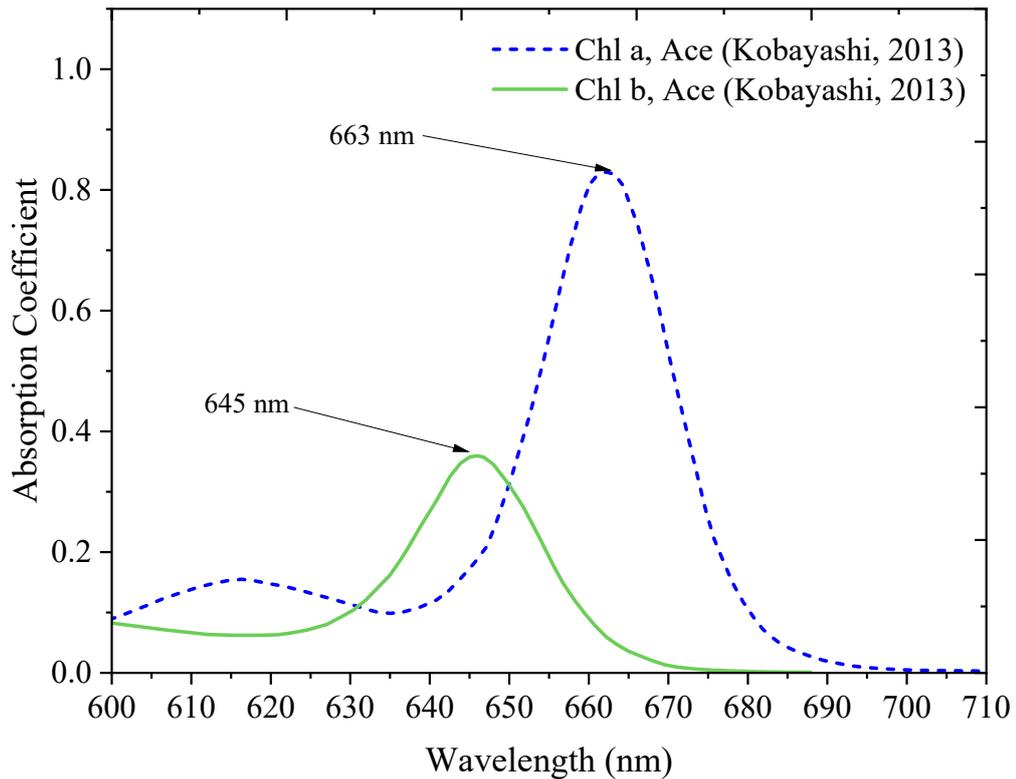
Various methods and techniques are employed to measure chlorophyll concentration (SMITH; BENITEZ, 1955), each with its advantages and respective applications:

2.2.1 Absorption Spectrophotometry

Absorption spectrophotometry uses different wavelengths to quantify chlorophyll a, chlorophyll b, and carotenoids in plant extracts, where the solvent used affects the absorption coefficients (SEELY; JENSEN, 1965; SHINANO *et al.*, 1996; SUMANTA *et al.*, 2014). Mackinney (1941) describe a procedure for determining chlorophyll concentrations in plant extracts such as oats and mallows, showing good agreement between calculated and experimental values

for chlorophyll absorption (MACKINNEY, 1941). Also, by measuring absorbance at 663 nm and 645 nm (as shown in the Figure 4), chlorophyll content can be quantified using specific absorption coefficients and equations (ARONOFF, 1950), (MACKINNEY, 1941), a widely used but destructive method involving leaf maceration.

Figure 4 – Characteristic graph of chlorophyll a and b absorption at the peaks of 663 nm for chlorophyll a and 645 nm for chlorophyll b

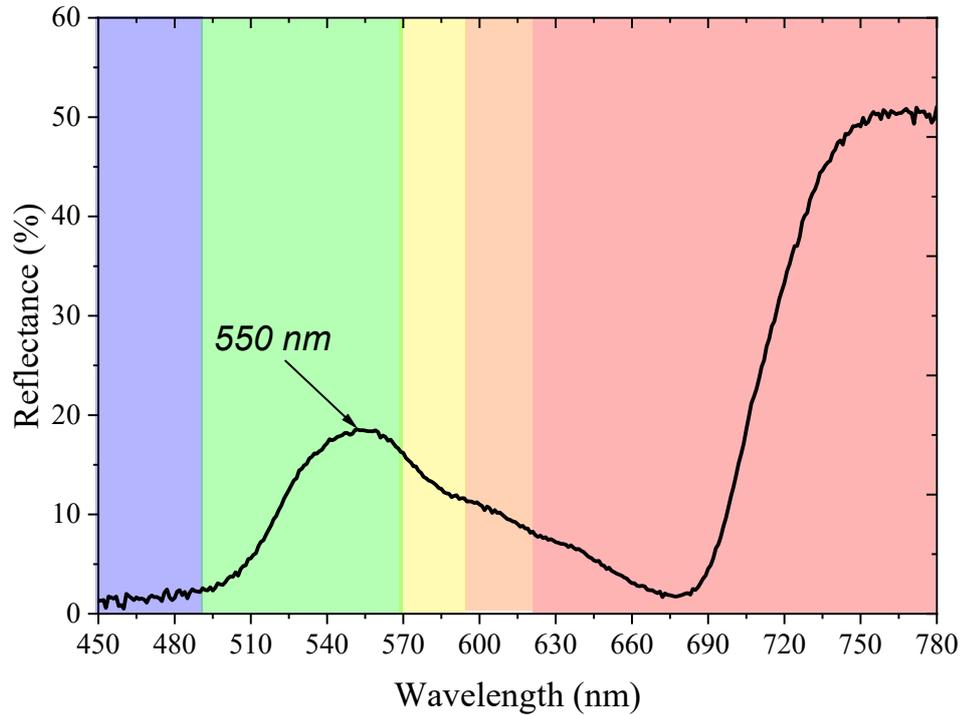


Source: Based on (KOBAYASHI *et al.*, 2013).

2.2.2 Reflectance Spectroscopy

The estimation of chlorophyll through reflectance spectroscopy as a non-destructive method is a highly relevant and currently studied topic. Various methods have been used to estimate chlorophyll through spectral bands and spectral ranges (KANDPAL; KUMAR, 2023), (MAIN *et al.*, 2011). The characteristic graph of the reflection spectrum in wheat leaves is shown in Figure 5.

Figure 5 – The characteristic graph of the reflection spectrum in wheat leaves is shown, with a peak intensity around 550 nm (green band) and a plateau around 750 nm



Source: Own authorship (2024).

2.2.3 Fluorometric Methods

By measuring chlorophyll-induced fluorescence, data can be obtained on the plant's ability to capture and convert sunlight into chemical energy. This method is widely used to measure photosynthetic efficiency and chlorophyll content (MURCHIE; LAWSON, 2013), (SWOCZYNA *et al.*, 2022).

2.2.4 Portable Chlorophyll Meters (SPAD)

The SPAD meter is a device used to measure chlorophyll concentration in plant leaves. The SPAD technology works by measuring the amount of light passing through a leaf at two specific wavelengths (usually around 650 nm and 940 nm), allowing for rapid and non-invasive real-time measurements (KANDPAL; KUMAR, 2023).

2.2.5 Image-based Methods

Chlorophyll estimation through hyperspectral images captures images across a wide range of wavelengths to analyze chlorophyll distribution and quantity (LI *et al.*, 2022), (GAO *et al.*, 2021), (DENG *et al.*, 2024). Studies demonstrate chlorophyll detection through short-range hyperspectral imaging, enabling precise chlorophyll mapping in canopies under varying light intensities (ZHU *et al.*, 2024).

RGB Image Analysis: The use of RGB cameras to estimate chlorophyll content through image analysis based on the three primary colors: red (R), green (G), and blue (B). This method is a fast and non-invasive way to determine chlorophyll content in leaves using RGB images (YADAV *et al.*, 2010).

2.2.6 Chemical Methods

High-Performance Liquid Chromatography (HPLC) is an analytical technique used to separate, identify, and quantify components in a liquid mixture (LOUGH; WAINER, 1995). It is widely used in chemistry, biochemistry, and pharmacology for analyzing complex mixtures, such as separating compounds in drugs, foods, chemicals, and biological samples (LYNCH; WEINER, 1979), (HANSEN; REUBSAET, 2015). HPLC is an essential tool in research and quality control laboratories due to its precision and ability to provide detailed information on sample composition (SHIOI *et al.*, 1983), (HANSEN; REUBSAET, 2015), (MEYER, 2010).

2.3 DETECTION OF STRESS IN PLANTS BY REFLECTANCE SPECTROSCOPY

The detection of abiotic stress in plants is crucial for efficient agricultural management, especially under climate change conditions and limited water resources. Several studies demonstrate the application of optical spectroscopy and machine learning techniques to monitor and predict the water status of different plant species using leaf spectral reflectance. The research consolidated in Table 1 synthesizes related studies on stress detection in different plant species, such as vines, sunflowers, rice, and wheat, employing methodologies to estimate critical physiological parameters under various water stress conditions. Techniques like reflectance spectroscopy in the visible and near-infrared (Vis/NIR) range and hyperspectral data analysis combined with artificial intelligence algorithms allow non-destructive estimation of parameters such as water

content and other types of abiotic stress. The following paragraphs detail related research on the application of reflectance spectroscopy, some statistical data processing methods, and the use of artificial intelligence, highlighting the findings in each investigation.

Tunca *et al.* (2023) conducted a study in Samsun, Turkey, during the 2020 and 2021 sorghum growing seasons. They explored different levels of water stress on sorghum spectral reflectance, leaf area index (LAI), and the accuracy of machine learning models (XGBoost, RF, SVM) in estimating crop water content (CWC). Proximal hyperspectral data was collected using an ASD Field Spec Pro spectrometer, covering the 325-1075 nm range and involving non-destructive sampling of sorghum leaves. The results highlighted significant variations in spectral reflectance due to irrigation treatments. The study underscores the utility of vegetation indices such as CL_Rededge and EVI in improving CWC estimates. Spectral reflectance values were obtained from a height of 4.5 m above ground level, and spectral measurements were captured with a solar incidence angle of approximately 45°. However, this study's limitation lies in potential variations in results for different solar incidence angles, and it uses the hyperspectral range.

Estrada *et al.* (2023) evaluated four wheat genotypes in pots under semi-controlled conditions in Chile and Spain. Reflectance was measured with a FieldSpec 3 spectrometer (Analytical Spectral Devices ASD Inc., Boulder, CO, USA) in the 350 to 2,500 nm range, along with a contact probe (ASD Inc., Boulder, CO, USA) equipped with a 5 W halogen light. The authors applied three treatments: control, water stress, and combined water stress with heat shock. Changes in the genotypes' spectral signature, in response to environmental fluctuations, were associated with variations in stomatal conductance under water stress and combined water and heat stress. It has the limitation of using a contact probe, which is subject to specific distance and angle requirements.

Zhang *et al.* (2021) explored the classification of greenhouse tomato plants under varying water stress conditions. They used two tomato varieties grown in greenhouses over two seasons, subjected to three different irrigation treatments. Spectral reflections from the tomato canopy were collected using visible and near-infrared spectroscopy. The successive projections algorithm (SPA) was employed to identify six optimal wavelength bands, and the analysis used a multilayer perceptron classifier (MLPC) and the one-vs.-rest classifier (ORC). The results showed that MLPC achieved better classification accuracy than ORC, especially with the full spectrum instead of selected optimal bands. The equipment used was a portable ASD

FieldSpec® HandHeld™ 2 spectrometer (ASD Inc., USA). The collected wavelength range was 325–1075 nm, with a 25° field angle, and the optical input of the spectrometer was maintained at 3–4 inches. It has the limitation of a specific angle and distance for sample acquisition.

Das *et al.* (2021) conducted a study to evaluate leaf water content (LWC) in various rice genotypes under water-deficit stress using visible-near infrared spectroscopy (VNIR). Using leaf samples from ten rice genotypes, they found that spectral bands centered around 1400 nm showed the best correlation with LWC at different water stress levels. Additionally, artificial intelligence techniques such as Partial Least Squares Regression (PLSR), Support Vector Machine Regression (SVR), and Random Forests (RF) were applied. They used a contact probe with an internal light source coupled to a spectroradiometer (Analytical Spectral Devices, Boulder, CO) in the 350–2500 nm spectral range. It has the limitation of potential variations in angles.

Marín-Ortiz *et al.* (2020) investigated the effects of *Fusarium oxysporum* infection and water stress on tomato plants (*Solanum lycopersicum*) using reflectance spectroscopy with the Ocean Optics HR2000 equipment operating in the visible (Vis) and near-infrared (NIR) ranges, 380-1000 nm. Samples were obtained from leaves affected by the disease and subjected to water stress, analyzing physiological changes using techniques such as Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). This approach allowed them to correlate gas exchange parameters and chlorophyll fluorescence with spectral response, highlighting the applicability of spectroscopy for early disease detection non-destructively. The study may be affected by variations in incident angle for obtaining reflection spectra.

Ihuoma and Madramootoo (2019) evaluated the sensitivity of various spectral indices for monitoring water stress in tomato plants under different irrigation regimes in greenhouse conditions. Using hyperspectral data, they acquired spectral reflectance from tomato leaves and calculated indices such as the Normalized Difference Vegetation Index (NDVI) and the Photochemical Reflectance Index (PRI) centered at 550 nm, among others. The results highlighted that PRI550, Water Index (WI), and OSAVI were the most sensitive for distinguishing levels of water stress. They used a miniature fiber optic spectrometer (Blue-wave, Stellar Net Inc., FL, USA) in the 200 to 1150 nm spectral range, calibrating the probe at 45° relative to a white reference standard, at a distance of approximately 64 mm. Spectral canopy measurements were taken under clear sky conditions between 10:00 and 15:00 h to ensure maximum solar intensity when the sun was shining directly on the plants. The spectrometer was placed 30 cm above the plant canopy so that only the canopy was visible. The experiment has the limitation of a specific

distance and time for sample acquisition.

El-Hendawy *et al.* (2019) present the potential of existing and novel spectral indices for estimating leaf water status and grain yield of spring wheat under different deficit irrigation rates. Using hyperspectral reflectance spectroscopy, they evaluated parameters such as leaf water potential (LWP), relative water content (RWC), and equivalent water thickness (EWT), as well as grain yield (GY). Spectral wavelengths sensitive to these parameters were found primarily in the NIR and SWIR regions, with peaks around 351, 518, and 687 nm in VIS, 762, 974, 1100, and 1240 nm in NIR.

Neto *et al.* (2017) employed and validated models to estimate the water and chlorophyll content in sunflower leaves under progressive water stress. Using reflectance spectroscopy in the visible/near-infrared (Vis/NIR) region and chemometric techniques, specifically partial least squares regressions (PLSR), they achieved high determination coefficients and low mean errors for water ($R^2 = 0.8386$ and -0.40 mg g^{-1} on a dry basis) and chlorophyll ($R^2 = 0.8097$ and 0.09 mg g^{-1}), respectively. These results indicate that spectrometry has potential as an alternative method for quantifying the water and chlorophyll status in sunflower leaves. They used a miniature spectrometer (JAZ-EL350, Ocean Optics, Dunedin, Florida, USA) coupled with a tungsten-halogen light source in the Vis/NIR wavelength range (500-1039 nm), with a reflection probe (R400-7-VIS-NIR, Ocean Optics, Dunedin, Florida, USA) to collect the reflected light from the sunflower leaves, connected to the spectrometer and light source, respectively. The other end of the probe was inserted into an anodized aluminum holder, positioned vertically at 90 degrees in relation to the leaves for specular reflectance measurements. The study has a limitation related to the equipment configuration considering an angle in relation to the sample.

Chemura *et al.* (2017) evaluated the ability of selected spectral bands in the VIS/NIR range to predict water content in coffee plants using the random forests algorithm. They conducted an experiment where coffee plants were exposed to different levels of water stress, and the reflectance and water content of the plants were measured. Variable selection methods were used to identify 11, 16, and 22 relevant spectral bands through cross-correlation, reflectance difference, and reflectance sensitivity, respectively. The selected bands were integrated into the random forests algorithm to successfully predict water content, demonstrating that reflectance sensitivity obtained the best performance ($r = 0.87$, $\text{RMSE} = 4.91\%$, and $\text{pBias} = 0.9\%$). These results indicate that water content in coffee plants can be reliably predicted using VIS/NIR spectroscopy and machine learning algorithms like random forests. Reflectance was measured

using an Apogee VIS-NIR spectrometer (Apogee Instruments, Inc., Logan, UT, USA) with a spectral range of 400 to 900 nm. Each reading consisted of an average of three spectral scans, taken 15 cm above the coffee leaf of interest at a 30-degree angle. The experiment has a limitation due to variations caused by different angles.

Maimaitiyiming *et al.* (2017) investigated the early detection of physiological responses in grapevines under different levels of water stress using reflectance spectroscopy. Conducted in a vineyard in Mount Vernon, MO, during the 2014 and 2015 growing seasons, they implemented three irrigation treatments: no irrigation, full irrigation, and intermediate. They analyzed various spectral indices, including new normalized difference spectral indices (NDSI), to correlate with stomatal conductance (G_s), a crucial indicator of water stress. The study highlighted that NDSI(R603,R558) was effective in estimating G_s ($R^2 = 0.720$), outperforming traditional indices. Additionally, they identified NDSI(R685,R415) as optimal for non-photochemical quenching (NPQ) ($R^2 = 0.681$). They used partial least squares regression (PLSR) to develop predictive models for G_s , although NDSI performed better. The Variable Importance in Projection (VIP) analysis showed crucial wavelengths for estimating G_s . The study was conducted with an Apogee VIS-NIR spectrometer (Apogee Instruments, Inc., Logan, UT, USA) with a spectral range of 400 to 900 nm, taken 15 cm above the coffee leaf of interest at a 30-degree angle. The experiment has a limitation as it was conducted with a specific angle, which could vary at different angles.

González-Fernández *et al.* (2015) evaluated spectroscopic methods to estimate water content in commercial grapevine leaves of three grape varieties (Mencía, Merlot, and Tempranillo) in vineyards of El Bierzo, Spain. Applying the continuum removal (CR) technique to transform the spectral data, they used partial least squares regression (PLSR) and ordinary least squares regression (OLSR) models to fit and validate the data. The results showed that using CR improved the accuracy of water content estimates, with the most accurate models being those that used PLSR on the CR-transformed spectrum. They used a portable ASD FieldSpec 4 spectroradiometer with a plant probe attachment (Analytical Spectral Devices, Inc., Boulder, Colorado, USA), ranging from 350 to 2500 nm. The experiment has a limitation as it requires a probe to measure directly on the leaf, which could vary at different angles.

Table 1 – List of research analysis in plant stress based on Vis-NIR spectroscopy

Research by	Spectroscopy techniques	Spectral Region (nm)	Abiotic Stress Type	Plant Species	Infected Area	Analysis
Tunca <i>et al.</i> (2023)	Hyperspectral reflectance	325–1075	Water	Sorghum	leaves	Random Forest (RF), Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost), Redes Neuronales Artificiales (ANN), Regresión Lineal y Logística
Estrada <i>et al.</i> (2023)	Reflectance	350-2500	Water (WS), water and heat shock (WS+T)	Wheat	Leaves	Statistical analyses, ANOVA and regression
Zhang <i>et al.</i> (2021)	Reflectance VNIR	325–1075	Water	tomato	Leaves	Multilayer perceptron classifier (MLPC) and ORC classifier
Das <i>et al.</i> (2021)	Reflectance VNIR	350-2500	Water	Rice	Leaves	Partial Least Squares Regression (PLSR), Support Vector Machine Regression (SVR), Gaussian Process Regression (GPR), Multivariate Adaptive Regression Splines (MARS), Random Forest (RF), Stepwise Multiple Linear Regression (SMLR)
Marín-Ortiz <i>et al.</i> (2020)	Reflectance	380-1000	Water	Tomato	leaves	Análisis de Componentes Principales (PCA) y Análisis Discriminante Lineal (LDA).
Ihuoma and Madramootoo (2019)	Reflectance	200–1150	Water	Tomato	Leaves	Spectral vegetation indices
El-Hendawy <i>et al.</i> (2019)	Hyperspectral reflectance	350–2500	Water	Wheat	Leaves	Partial least square regression (PLSR)
Neto <i>et al.</i> (2017)	Reflectancia VNIR	500-1039	Water	Sunflower	Leaves	Chemometrics, Specifically partial least squares regressions (PLSR)
Chemura <i>et al.</i> (2017)	Reflectance	400-900	Water	Coffee	leaves	Random Forests(RF)
Maimaitiyiming <i>et al.</i> (2017)	reflectance	350-2500	Water	Grapevine	Leaves	Partial Least Squares Regression (PLSR) and Variable Importance in Projection (VIP)
González-Fernández <i>et al.</i> (2015)	Reflectance	350–2500	Water	Grape	Leaves	Partial least squares regression (PLSR), ordinary least squares regression (OLSR)

Source: Own authorship (2024).

2.4 CHLOROPHYLL ESTIMATION BY REFLECTANCE SPECTROSCOPY

Estimating chlorophyll through Reflectance spectroscopy is an important and non-destructive method for estimating chlorophyll in plants. Many methods for estimating chlorophyll

using spectral bands have been notably utilized, and the use of machine learning methods is also showing considerable progress. Studies have conducted a comprehensive review of techniques for estimating chlorophyll content in leaves, comparing destructive and non-destructive methods, concluding and recommending the use of artificial intelligence and machine learning techniques such as Random Forest and Support Vector Machine for accurate chlorophyll estimation (KANDPAL; KUMAR, 2023).

In Table 2, several studies are presented that address chlorophyll quantification using non-destructive methods based on reflectance spectroscopy applied to the leaves of various plant species. These methods allow for the measurement of chlorophyll levels using the optical properties of the leaves to estimate chlorophyll concentration, and the works included cover a variety of plant species.

Singh *et al.* (2022) utilized secondary hyperspectral data from maize leaf reflectance, collected in the spectral range of 350 to 2500 nm in field and greenhouse experiments under different nitrogen conditions. They measured the water, chlorophyll, nitrogen (N), phosphorus (P), and potassium (K) content in maize leaves using standard methods. Six machine learning regression algorithms (Random Forest, Support Vector Regression, k-Nearest Neighbours, Multilayer Perceptron, Gradient Boosting Regression, and Partial Least Squares Regression) were employed to develop models that predict these parameters from the leaf reflectance data. Explainable artificial intelligence methods were used to identify the optimal wavelengths for each parameter. They found that wavelengths in the short-wave infrared (SWIR) region were optimal for estimating water content, while the red-edge band was optimal for chlorophyll. For their experiment, they used a tabletop spectroradiometer (FieldSpec4, Malvern Panalytical Ltd.) with a contact probe. It has the limitation of using a contact probe, which is subject to specific distance and angle requirements.

Liu *et al.* (2022) conducted a detailed study using the plant species *Toona sinensis*, known for its economic value and drought resistance. They employed near-infrared reflectance spectroscopy (NIRS) to non-destructively analyze the plant's leaves, using a spectral range of 1100 to 2500 nm. This technique allowed the identification of key physiological features such as chlorophyll content using specific spectral bands like 1420 nm, 1694 nm, and 2160 nm. The analysis was carried out using partial least squares regression (PLSR), which proved effective in accurately predicting these indicators. The equipment used included a portable fiber optic contact probe from a field spectrometer (LF-2500, Spectral Evolution, USA), It has the limitation of

using a contact probe, which is subject to specific distance and angle requirements.

Li *et al.* (2021) developed a new modified spectral index (MDR) for estimating equivalent water thickness (EWT) in leaves of various plant species, covering a wide range of growth stages. The study used multi-angular spectral reflectance measurements in the 400 to 2500 nm range of leaves from 18 plant species, including trees, shrubs, and lianas. This non-destructive method employed advanced spectroscopy techniques to capture data from different angles and conditions, both in the lab and the field, and validated the MDR index as an effective and accurate tool for estimating water content in leaves, without the need for re-parameterization for each species. For the experiment, they used a goniometer system, an ASD FieldSpec spectroradiometer (Analytical Spectral Devices FieldSpec 4, Boulder, CO, USA), taking leaf measurements in the principal plane from -60° to 60° (with a 10° interval) under three different zenith angles of incidence (30° , 40° , and 50°), finding a new spectral index tolerant to different angles. The study did not delve into techniques such as artificial intelligence.

Urbanovich *et al.* (2021) employed the random forest algorithm to predict chlorophyll concentrations (chlorophylls a and b) in plant leaves using reflection spectra ranging from 400 to 2500 nm. They used 276 leaf samples from 39 plant species, with a significant focus on the sycamore maple (*Acer pseudoplatanus* L.). They evaluated several models, with the best performance achieved by those using the intensity and derivatives of the reflection spectra in the visible wavelength range. Comparison with traditional methods highlighted the superior predictive accuracy of the random forest. This study underscores the usefulness of machine learning methods for the non-invasive assessment of chlorophyll content in plant leaves. The research refers to the ECOS database, and the equipment used was an ASD FieldSpec spectroradiometer, with limitations for other conditions of obtaining multi-angular reflection spectra.

Mahajan *et al.* (2021) used visible near-infrared (VNIR) spectroscopy to assess the foliar nutrient status of mango (*Mangifera indica* L.). Leaf samples were collected from mango orchards in the North Goa and South Goa districts of India during the post-harvest season. Spectral measurements were taken in the 350 to 1050 nm wavelength range, with preprocessing techniques such as smoothing and correction applied to improve data quality. Subsequently, a detailed chemical analysis was conducted to determine the nutrient content in the samples. They found that a combined approach of PLSR and machine learning models, such as Cubist, SVR, and Elastic Net, significantly improved predictive accuracy for nutrients like nitrogen, phosphorus, and potassium, among others. They used a visible near-infrared spectroradiometer

(GER1500, Spectra Vista Corp., Poughkeepsie, NY, USA) for non-contact observations, and mango leaf measurements were conducted inside a black box to reduce the impact of scattered light. The study is limited by measurements at different angles of light incidence for different angles.

Osco *et al.* (2019) conducted a study on the hyperspectral response of lettuce plants (*Lactuca sativa* L.) under water stress. The researchers used a non-destructive technique to measure the spectral response of the leaves in the 325 to 1075 nm range. Samples were taken from the plant leaves, and the data were analyzed using artificial neural networks (ANN). Spectral bands in the visible region, specifically in the blue (380-460 nm) and red (640-680 nm) ranges, were crucial for determining chlorophyll content. The study demonstrated that it is possible to detect early stages of water stress in lettuce plants with high accuracy using this methodology, which has potential for application in other crops and species in agricultural fields. The spectral response of the lettuce was measured using a FieldSpec HandHeld ASD spectroradiometer, with the equipment placed near the leaves at a 45° angle relative to the plant height, and a halogen lamp positioned at a 45° angle on the other side. The equipment was calibrated with a Lambertian surface plate (Spectralon® plate). The study is limited by measurements at different angles of light incidence.

Imanishi *et al.* (2010) employed a non-destructive method using a spectroradiometer with a leaf-clip attachment. This device allowed analysis of the reflectance and absorptance spectra in the 325 to 1075 nm wavelength range of leaf samples. The primary objective was to develop optimal reflectance and absorptance indices for estimating chlorophyll content, comparing them with previously published indices. Among the evaluated indices, the Datt reflectance index and the Ciganda red-edge chlorophyll index stood out for their effectiveness in accurately estimating chlorophyll content. They used the FieldSpec HandHeld spectroradiometer with a 10 mm spot size plant probe and leaf-clip attachments (Analytical Spectral Devices Inc., USA). The study is limited by measurements at different angles of light incidence.

Ciganda *et al.* (2009) investigated maize content using a non-destructive extraction method. The spectral range used (720-730 nm in the red edge and 770-800 nm in the near infrared), the reflectance-based spectroscopy technique, specifically on collar or ear leaves, the analysis technique (Red Edge Chlorophyll Index, CIred edge), and the spectral bands used to determine chlorophyll content (red edge and near infrared). Their objective was to establish a non-destructive and rapid methodology for estimating total chlorophyll content in a maize

canopy, based on the chlorophyll content of a single leaf. They used an Ocean Optics USB2000 spectrometer with an Ocean Optics tungsten halogen light source LS-1 and a leaf clip with a 60° angle relative to the bifurcated fiber optic. The study is limited by measurements at different angles of light incidence.

Gitelson *et al.* (2003) collected leaves of Norway maple (*Acer platanoides* L.), horse chestnut (*Aesculus hippocastanum* L.), beech (*Fagus sylvatica* L.), and wild vine (*Parthenocissus tricuspidata* L.) to study chlorophyll content. The spectral reflectance of the leaves was measured in the range of 400 to 800 nm using reflectance spectroscopy. The analyses were based on regression of spectral indices to determine chlorophyll content, identifying spectral bands between 520-550 nm, 695-705 nm, 525-555 nm, 695-725 nm, 520-585 nm, and 695-740 nm as the most relevant for this determination. They used the Hitachi 150-20 spectrophotometer and the Shimadzu 2101 PC spectrophotometer to obtain the reflectance spectra. The study is limited by measurements at different angles of light incidence.

Richardson *et al.* (2002) evaluated non-destructive methods for estimating foliar chlorophyll content in paper birch (*Betula papyrifera*). They used optical methods based on light absorbance and reflectance, evaluating reflectance indices that correlate with foliar chlorophyll. They found that these methods provide reliable estimates of relative chlorophyll content, highlighting specific reflectance indices that outperformed handheld chlorophyll absorbance meters. This study emphasizes the usefulness of reflectance spectroscopy as a non-destructive and rapid tool for assessing chlorophyll status in individual leaves. For the study, they used a UniSpec Spectral Analysis System (PP Systems, Haverhill, Massachusetts, USA) with a 2.3 mm diameter foreoptic (0.042 cm²) and a 6.8 W internal halogen lamp. The leaves were held in a black PVC leaf clip at a 60° angle to the foreoptic. The study is limited by measurements at different angles of light incidence.

Maccioni *et al.* (2001) investigated the directional reflectance (R) properties of nadir-illuminated leaves from four different plant species using a non-destructive method. Spectra were acquired in the 380 to 780 nm wavelength range. The technique used was directional reflectance spectroscopy, focusing on logarithmic correlations between chlorophyll content and reflectance in key spectral bands, such as 550 nm (green band) and near 700 nm (red edge). They found that normalized internal reflectance in these bands showed a linear relationship with the logarithm of chlorophyll content, emerging as a more precise index for chlorophyll determination compared to traditional vegetation index methods. The study may vary under different conditions of light

incidence.

Datt (1999) studied the visible and near-infrared reflectance properties of leaves from various Eucalyptus species. They applied a multiple scatter correction technique to reduce scattering effects in the reflectance spectra, thereby improving chlorophyll content estimations. They found that the 710 nm reflectance band is sensitive for determining chlorophyll content. Additionally, they proposed a new index $\frac{R_{850}-R_{710}}{R_{850}-R_{680}}$ as an effective indicator for remote estimation of chlorophyll content in plants. Reflectance measurements on leaves were made using a Geophysical Environmental Research Intelligent Infrared Spectroradiometer (GERIRIS–Mark IV). The study is limited by measurements at different angles of light incidence.

Gitelson and Merzlyak (1998) studied remote detection of chlorophyll concentration in leaves of higher plants. They collected leaves of maple, chestnut, cotoneaster, tobacco, fig, oleander, hibiscus, vine, and rose. They used a non-destructive method of reflectance and transmittance spectrophotometry in a spectral range of 350 to 1100 nm, with samples taken from the leaves. They developed algorithms to estimate chlorophyll using reflectance ratios and the "green" normalized difference vegetation index (NDVI green). They identified spectral bands of 530 to 630 nm and near 700 nm as highly sensitive for determining chlorophyll content. For reflectance spectra measurements, they used a Hitachi 150-20 spectrophotometer equipped with an integrating sphere attachment, which is the limitation concerning their proposal.

Gitelson and Merzlyak (1994) conducted a detailed study to quantitatively estimate chlorophyll-a content in chestnut (*Aesculus hippocastanum* L.) and maple (*Acer platanoides* L.) leaves during the autumn season. Using the technique of reflectance spectroscopy in the 400-750 nm range, they employed a non-destructive method to extract specific samples from leaves collected at the Botanical Garden of Moscow State University. During the analysis, they identified several crucial spectral bands for accurately determining chlorophyll content, particularly at 550 nm, 705 nm, 675 nm, 400 nm, and above 730 nm. These bands provided sensitive indicators of changes in chlorophyll content. For reflectance spectra measurements, a Hitachi 150-20 spectrophotometer equipped with an integrating sphere attachment was used, which is the limitation concerning our proposal.

Chappelle *et al.* (1992) developed an algorithm called Ratio Analysis of Reflectance Spectra (RARS) to remotely estimate the concentrations of chlorophyll a, chlorophyll b, and carotenoids in soybean leaves. They used reflectance spectra in the range of 300-750 nm, focusing particularly on the spectral bands at 675 nm, 650 nm, and 500 nm. The leaf samples were analyzed

using non-destructive techniques. RARS manipulates the reflectance spectra by dividing them by a reflection spectrum, thereby amplifying the pigment-specific absorption bands. This approach allowed the correlation of pigment concentrations with the selected spectral bands, thereby achieving an estimation of the photosynthetic pigments. To obtain the reflectance spectra, they used a LICOR Model 1800 integrating sphere radiometer with a resolution of 6 nm, The use of an integrating sphere is a limitation because the setup requires additional equipment.

Table 2 – Chlorophyll Estimation Using Artificial Intelligence

Research by	Spectroscopy techniques	Spectral Region (nm)	metodo de estracion	Plant Species	Infected Area	Analysis
Singh <i>et al.</i> (2022)	Reflectance Hyperspectral	350-2500	Non-destructive	Maize	Leaves	Machine Learning Regression Algorithms (Random Forest, Support Vector Regression, k-Nearest Neighbours, Multilayer Perceptron, Gradient Boosting Regression, Partial Least Squares Regression) and Explainable Artificial Intelligence (XAI)
Liu <i>et al.</i> (2022)	NIRS	1100 - 2500	Non-destructive	Toona sinensis	Leaves	Partial least squares regression (PLSR)
Li <i>et al.</i> (2021)	Reflection	400–2500	Non-destructive	18 plant species	Leaves	Modified spectral index (MDR)
Urbanovich <i>et al.</i> (2021)	Reflection	400–2500	Non-destructive	Sycamore maple	Leaves	For Chl-a, b: Random forest
Mahajan <i>et al.</i> (2021)	(VNIR)	350–1050	Non-destructive.	Mango	leaves	For nutrient status: Regresión de Componentes Principales (PCR), Regresión de Mínimos Cuadrados Parciales (PLSR), Regresión de Vectores de Soporte (SVR), Redes Elásticas (Elastic Net)
Oscó <i>et al.</i> (2019)	Reflectance Hyperspectral	325–1075	Non-destructive	Lettuce	Leaves	Artificial Neural Networks (ANN)
Imanishi <i>et al.</i> (2010)	reflectance / Absc	325-1075	Non-destructive / destructive	Flowering cherries	Leaves	Development of optimal reflectance and absorptance indices
Ciganda <i>et al.</i> (2009)	reflectance	720–730 and 770–800	Non-destructive	Maize	Leaves	Red Edge Chlorophyll Index $CI_{red\ edge} = \left(\frac{RNIR}{R_{red\ edge}} \right) - 1$
Gitelson <i>et al.</i> (2003)	Reflectance / Absorbance	400-800	Non-destructive / destructive	Norway maple, horse chestnut, beech, wild vine	Leaves	Regression analysis using spectral indices

Source: Own authorship (2024).

Table 3 – Chlorophyll Estimation Using Artificial Intelligence (continue)

Research by	Spectroscopy techniques	Spectral Region (nm)	metodo de estracion	Plant Species	Infected Area	Analysis
(RICHARDSON <i>et al.</i> , 2002)	Reflectance / Absorbance	306-1138	Non-destructive / destructive	Paper birch	Leaves	SPAD, CCI, Chl NDI, RII, D730
Maccioni <i>et al.</i> (2001)	reflectance	380-780	Non-destructive	Croton, Spotted Eleagnus, Japanese Pittosporum, and Benjamin Fig	Leaves	$\frac{R_{780} \cdot R_{710}}{R_{780} \cdot R_{680}}$
Datt (1999)	Vis-NIRS	400-2500	Non-destructive	Eucalyptus	Leaves	$\text{Reflectance : } Chl.a = 0.0885 \left[\frac{R_{850}/R_{710}}{R_{850}/R_{680}} \right]^{2.9564} + b = 0.126 \left[\frac{R_{850}/R_{710}}{R_{850}/R_{680}} \right]^{2.9065}$ $\text{1st derivative } Chl.a = 0.0813 \frac{D1(754)}{D1(704)} + 0.0102, Chl.a + b = 0.1144 \frac{D1(754)}{D1(704)} + 0.0157$ $\text{2nd Derivative : } Chl.a = 0.021 \frac{D2(712)}{D2(688)} + 0.0242, Chl.a + b = 0.0295 \frac{D2(712)}{D2(688)} + 0.0354$
Gitelson and Merzlyak (1998)	Reflectance and transmittance	400-750	Non-destructive	Chestnut, Tobacco, Maple, Cotoneaster	Leaves	$NDVI = [R_{nir} - R_{green}] / [R_{nir} + R_{green}]$
Gitelson and Merzlyak (1994)	Reflectance	400-750	Non-destructive	Chestnut and Maple	Leaves	$\text{For } Chl-a \left(\frac{R_{750} - R_{705}}{R_{750} + R_{705}} \right), \frac{R_{750}}{R_{555}}, \int_{705}^{750} \left[\frac{R(\lambda)}{R_{555}} - 1 \right] d\lambda$ $\text{and } \frac{R_{750}}{R_{705}}, \int_{705}^{750} \left[\frac{R(\lambda)}{R_{705}} - 1 \right] d\lambda.$
Chappelle <i>et al.</i> (1992)	Reflectance	300-750	Non-destructive	Soybean	Leaves	Ratio Analysis of Reflectance Spectra (RARS)

Source: Own authorship (2024).

Table 1–3 provides a comprehensive overview of some studies, referencing research where certain limitations were identified: (i) lack of tolerance for a multi-angular setup; (ii) accuracy depending on the use of a probe or spectrometer adapter to obtain reflectance spectra; (iii) measurements dependent on distance for obtaining reflectance spectra; (iv) measurements reliant on the calibration and measurement angle of the reflectance spectra; (v) lack of exploration of artificial intelligence tools. Then, the development of a technique not subject to the observed limitations is intended and will be addressed in this work.

3 METHODS

The Chapter 3 discusses the methodologies used to create different stress conditions in wheat leaves. Five different cardboard containers were constructed, each equipped with LED lighting: red, green, blue, and white light, as well as one container without light. Each box contains three containers where the seeds are placed after soaking (16 hours) in germination trays. During seven days, 3 ml, 5 ml and 10 ml of distilled water are added daily. The LED lighting is controlled by a timer configured to be on for 18 hours and off for 6 hours (MCKINNEY; SANDO, 1930; KONDRATEVA *et al.*, 2021). Additionally, the boxes are equipped with a DH11 sensor to monitor temperature and humidity in each box. The wheat progresses from an initial germination stage until it develops leaves. Fresh leaves are then used to obtain reflection spectra from each sample using an HR4000 spectrometer, resulting in a total of 23 spectra for each sample through reflectance spectroscopy, generating a total of 345 diffuse reflectance spectra. The data from the 15 samples were used to calculate the chlorophyll content through absorbance spectroscopy with the IL-593-Sr spectrophotometer (following the extraction and quantification methodology to obtain chlorophyll values). The reflection spectra are used to identify water and light stress, and a regression model based on time series and deep learning was used to estimate the chlorophyll content.

3.1 EQUIPMENT

This Section 3.1 presents a description of the equipment used, as well as the equipment configurations made for the experiments. The equipment used belongs to the Laser Laboratory¹ at the Federal University of Technology - Paraná (UTFPR), except for the Digital Spectrophotometer IL-593-S, which belongs to the Plant Ecophysiology Laboratory - Department of Crop Science and Plant Protection - Agricultural Sciences Sector - UFPR.

3.1.1 Water Destiller

The water distillation equipment Figure 6 was used for the experiment, as it is of great importance because the quality of the water used can significantly influence the experimental

¹ <https://lablaser.jlfabris.com/>

results. According to the equipment's characteristics, it can remove a wide range of mineral salts and impurities from the water. The equipment has a distillation capacity of one liter per hour, operates at 127 VAC or 220 VAC, has a power of 750 W , a frequency of 50/60 Hz , a stainless steel tank capacity of 4.0 liters , and a collection reservoir capacity of 4.0 liters (EVOXX LTDA, 2024). The presence of salts and impurities can alter the results and mask the effects of stress. Distilled water ensures that the plants are only exposed to the controlled conditions of the experiment without external interferences that could compromise the integrity of the collected data.

Figure 6 – Water distillation equipment used in the experiment (Evoxx water distiller operating at 127V or 220V, 750W).



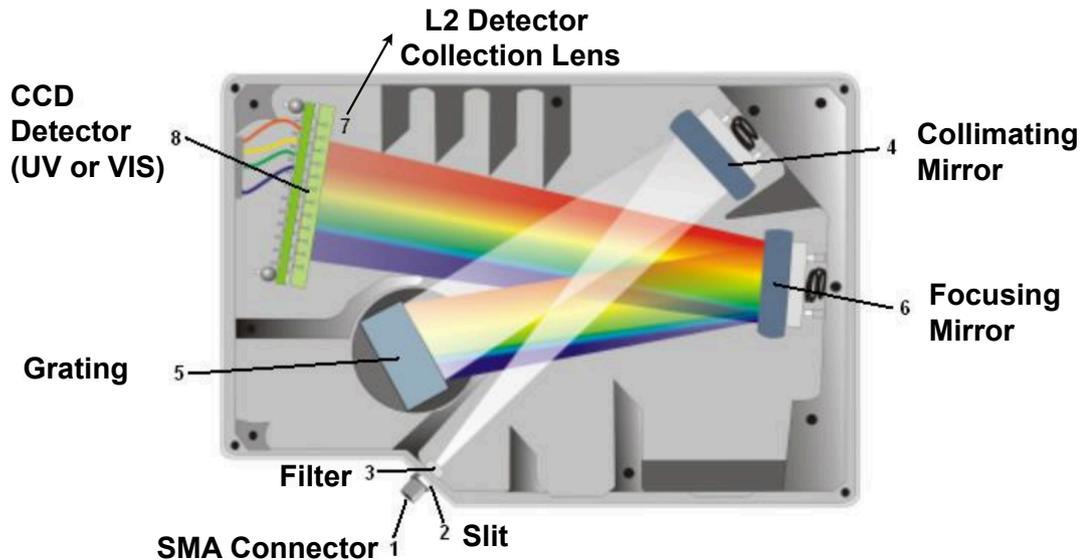
Source: Own authorship (2024).

3.1.2 HR4000 Spectrometer

The HR4000 Spectrometer is an electro-optical device certified with ISO 9001:2008 standards, ensuring quality in both its design and manufacturing processes. It offers compatibility across various platforms and operating systems, making it versatile and easily accessible in different environments. The HR4000 spectrometer has an energy consumption of 450 mA and 5 VDC. The HR4000 optical fiber spectrometer has an optical resolution of up to 0.025 nm (FWHM), sensitivity of 100 photons per count at 800 nm and range from 200 to 1100 nm. It

connects to a computer through a USB port or a serial port. When connected via USB, the device is powered by the host computer, eliminating the need for an external power source. It can be controlled by the OceanView software, a Java-based spectroscopy platform that operates on Windows, Macintosh, and Linux operating systems.

Figure 7 – HR4000 Spectrometer with internal components and numbered operational steps. The figure illustrates the complete process from light input to spectral conversion



Source: Based on (HR40000, 2012).

The HR4000 spectrometer functions by guiding light through a series of components starting with the SMA connector, a slit, a filter, a collimator mirror, a grating, a focusing mirror, an optional detector collection lens, and a CCD detector. These components work together to ensure that the light is collimated, diffracted, focused, and finally converted into a digital signal that can be analyzed to determine the spectral characteristics of the incoming light. The functionality, as detailed in Figure 7, is as follows: The SMA Connector (1): used to secure the optical fiber to the spectrometer's input. The light coming from the input fiber enters the spectrometer's optical bench through this connector; Slit (2): a rectangular aperture that regulates the amount of light entering the optical bench, it is mounted directly behind the SMA connector. The aperture size controls the amount of light and spectral resolution; Filter (3): Light passes through the filter before entering the optical bench. Bandpass and longpass filters are available to restrict radiation to certain wavelength regions; Collimating Mirror (4): focuses the light towards the spectrometer's diffraction grating. The light enters the spectrometer, passes through the SMA connector, the slit, and the filter, and then reflects off the collimating mirror towards the grating; Grating (5): a diffraction grating with different groove densities, allowing to specify

the wavelength coverage and resolution in the spectrometer; Focusing Mirror (6): receives the light reflected from the grating and focuses it onto the CCD detector or the L2 detector collection lens; L2 Detector Collection Lens (7): Focuses light from a tall slit onto the shorter elements of the CCD detector. It is an optional component that improves efficiency by reducing the effects of scattered light; CCD Detector (8): collects the light received from the focusing mirror or the L2 detector collection lens and converts the optical signal into a digital signal. Each pixel in the CCD detector responds to the wavelength of light that hits it, creating a digital response that the spectrometer transmits to the application (HR40000, 2012).

3.1.3 Digital Spectrophotometer IL-593-S

A digital spectrophotometer Kasuaki IL-593-S with Wavelength accuracy ± 1 nm , repeatability ≤ 0.2 nm and range of 190 to 1100 nm, is an automatic device operating in the UV-VIS range. It offers a backlit LCD screen to display the measured values, USB outputs for analysis by multiple wavelengths, and DNA/protein analysis, compatible with cuvettes from 5 mm to 100 mm, a spectrophotometer for chemical and biological analysis (KASUAKI, 2024). The device was used to quantify chlorophyll. The measurements were performed at the Plant Ecophysiology Laboratory - Department of Crop Science and Plant Protection - Agricultural Sciences Sector - UFPR.

Figure 8 – Digital spectrophotometer Kasuaki IL-593-S, with a wavelength range of 190 to 1100 nm, (KASUAKI, 2024).



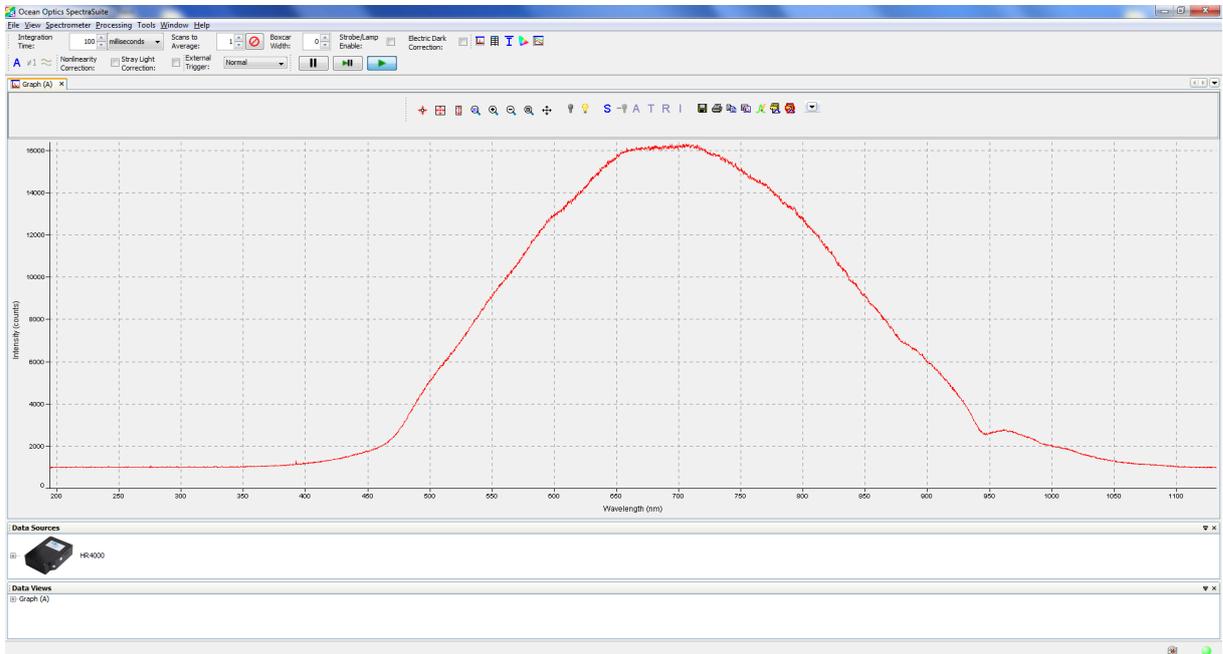
Source: Based on (KFEQUIPAMENTOS, 2024).

3.1.4 SpectraSuite Software

Java-based spectroscopy software, capable of controlling any USB spectrometer from Ocean Optics equipment, facilitates the management of multiple USB spectrometers, each with different acquisition parameters, with the ability to display spectra in real-time (Ocean Optics, Inc., 2007). Important parameters:

- **Integration Time:** Determines the light detection period. The ability to increase this time is useful for weak sources.
- **Scans to Average:** Sets the number of spectral acquisitions before averaging and sending the results to SpectraSuite. Its use reduces the noise without impairing the spectral resolution.
- **Boxcar Width:** Averages data from adjacent points on the CCD, smoothing the graph. Its use reduces the noise, however impairing the spectral resolution.

Figure 9 – SpectraSuite Software

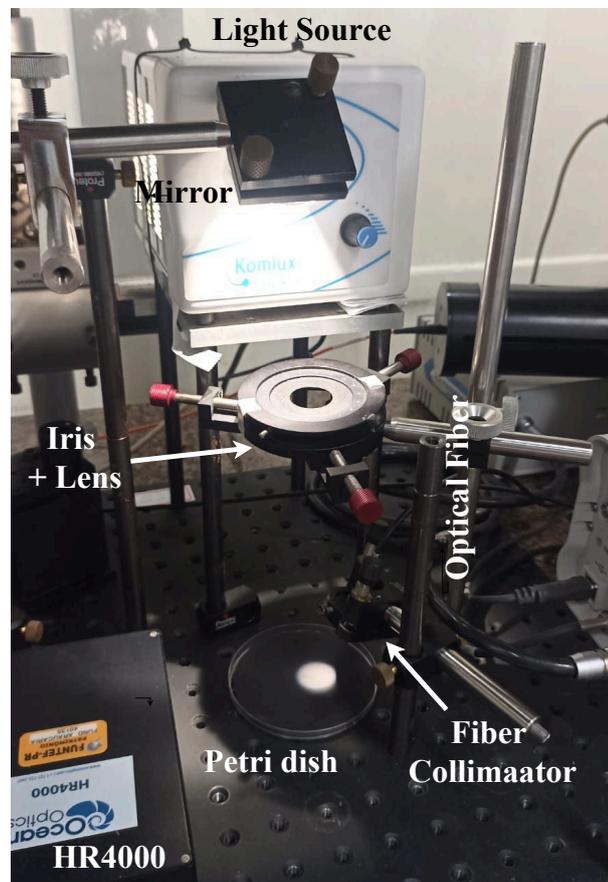


Source: Own authorship (2024).

3.1.5 Reflectance Spectroscopy Configuration

We used the reflectance spectroscopy configuration to non-destructively analyze the chemical properties of our wheat leaves. The amount of reflected light will allow us to predict the type of stress the leaves were subjected to, as well as the total chlorophyll content, through diffuse reflection spectra in samples taken from multiple angles.

Figure 10 – Physical implementation of the reflectance spectroscopy setup in the laboratory Laser at UTFPR. The configuration includes a halogen light source, a silver mirror, an iris diaphragm, and a lens. Light is directed onto a sample in a Petri dish, and the reflected light is collected by a fiber collimator, connected to an HR4000 spectrometer for analysis



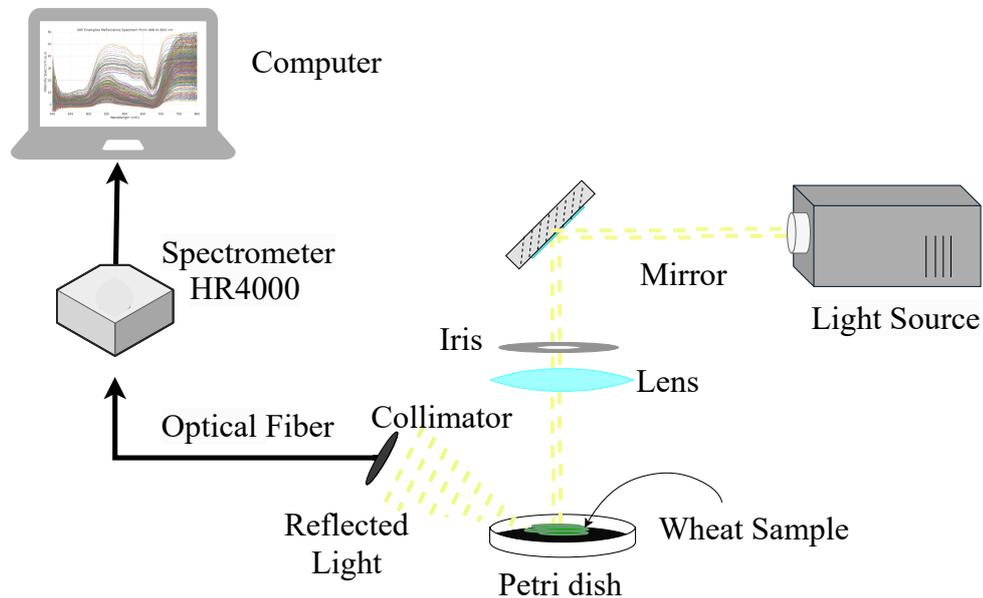
Source: Own authorship (2024).

The equipment configuration, as shown in Figure 10, includes a halogen white light source (Komlux HL7, power consumption 127 VAC and 75 W). The light is incident on a silver mirror (ThorLabs) to direct it towards the iris diaphragm (ThorLabs, 2.5 cm diameter) and lens (ThorLabs, 10.0 cm focal length), which together direct the light towards the sample contained in a Petri dish with an area of 55 cm² (SCD10100, Sorfa) against a black background. The reflected light is collected by a fiber collimator (F220SMA-532, ThorLabs) connectors, 350 - 700 nm spectral range and a focal length of $f = 10.90$ mm. This fiber collimator is connected

to an optical fiber (F600-VIS-NIR, StellarNet), which, in turn, is connected to a spectrometer (HR4000, described in Subsection 3.1.2) acting as an interrogator to collect the light reflected from the sample.

In Figure 11, the diagram of the Reflectance Spectroscopy Configuration implemented in the experiment is shown. For reflection spectra under different stress conditions, the spectra were obtained using the SpectraSuite software described in Subsection 3.1.4. The spectra data was saved in text files ("txt" format) for subsequent analysis using Python in the Colab environment².

Figure 11 – Diagram of the reflectance spectroscopy setup in the laboratory Laser at UTFPR. The configuration includes a halogen light source, a silver mirror, an iris diaphragm, and a lens. Light is directed onto a sample in a Petri dish, and the reflected light is collected by a fiber collimator, connected to an HR4000 spectrometer for analysis



Source: Own authorship (2024).

To calculate the percentage of relative reflectance of the samples at a specific wavelength, the software uses Equation 2, where (D_λ) is the dark background intensity (or dark intensity) at the wavelength. For the experiment, the Petri dish with a black background was used. (R_λ), the Reflection Spectrum, is taken with light incident on the Petri dish with a white background. (S_λ) is the intensity of the sample at each wavelength. To calculate the percentage of relative diffuse reflectance, first, the dark background is corrected by subtracting D_λ from S_λ and R_λ . The intensity ratio is obtained by dividing the corrected sample intensity by the corrected reference intensity. To get the percentage value, multiply by 100 (Ocean Optics, Inc., 2005), resulting in

² <https://colab.research.google.com/>

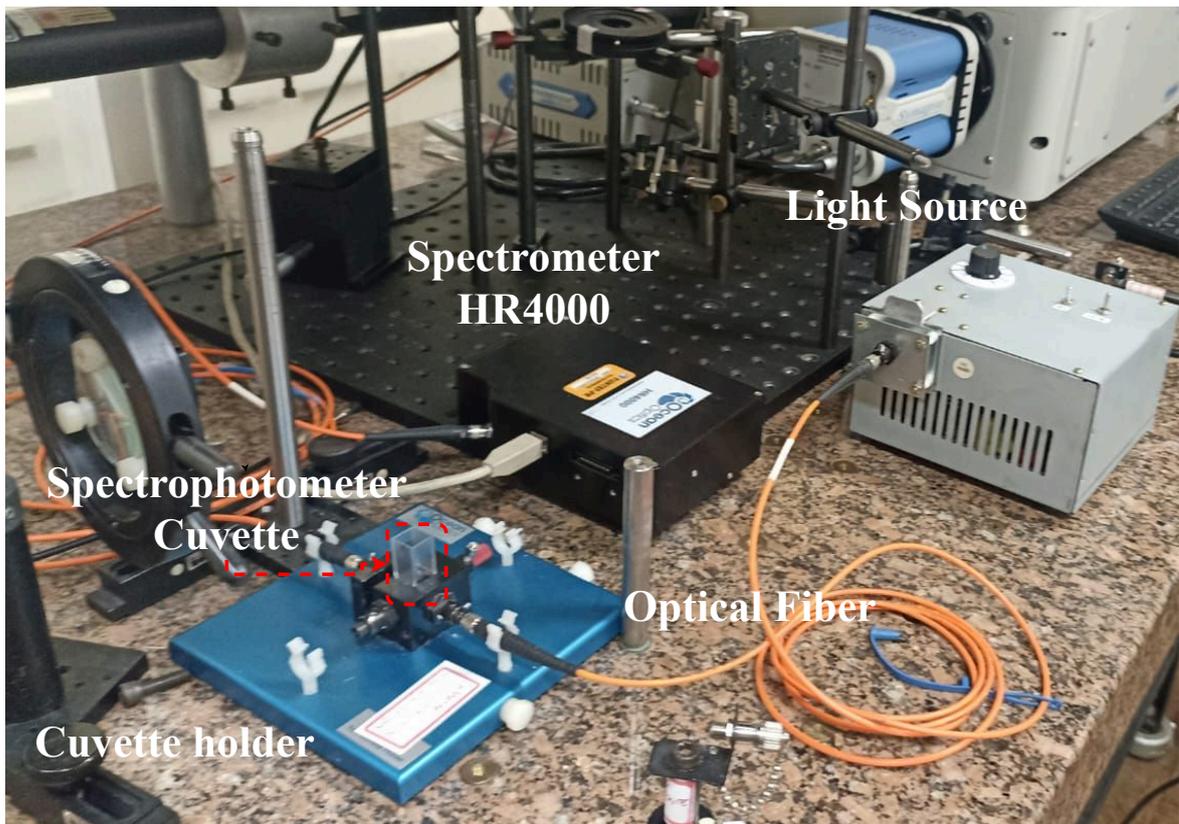
the function that gives us the corrected spectrum in percentage:

$$\%R_{\lambda} = \frac{S_{\lambda} - D_{\lambda}}{R_{\lambda} - D_{\lambda}} \times 100\%. \quad (2)$$

3.1.6 Absorbance Spectroscopy Configuration

This section details the implementation of absorbance spectroscopy tests, where the HR4000 spectrometer described in Subsection 3.1.2 was used. Figure 12 shows the actual laboratory setup for absorbance spectroscopy, including a diagram of the equipment used for the experiment (Figure 13) and the corresponding Equation 3.

Figure 12 – Physical implementation of the absorbance spectroscopy configuration. A halogen light source (Philips 12345SL, 3100 K) and two optical fibers connected to a cuvette holder (CUV-ALL-UV, Ocean Optics) with 4.5 mL polystyrene cuvettes (K42-045, Olen) and the UV-Vis spectrometer (HR4000, Ocean Optics)

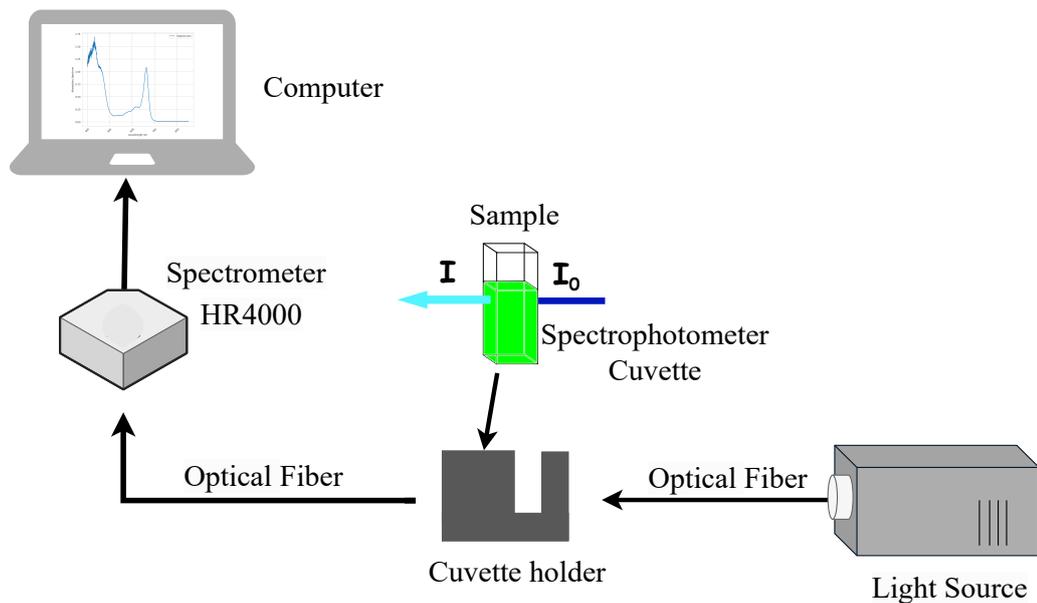


Source: Own authorship (2024).

Based on the principle that atoms, molecules, or ions absorb specific wavelengths of light, the following equipment configuration was used for this work. For UV-Vis spectroscopy, a standard halogen light source (Philips 12345SL, 3100 K) was employed. Two optical fibers with a core diameter of 200 μm (P200-2-UV-Vis, Thorlabs) were used for light transmission.

A cuvette holder (CUV-ALL-UV, Ocean Optics) was used to connect the two fibers, and the cuvettes used during the analyses were polystyrene with a volume of 4.5 mL (K42-045, Olen). These were connected to a UV-Vis fiber optic spectrometer (HR4000, Ocean Optics), which allows for obtaining UV-Vis optical spectra in the range between 200 nm and 1100 nm. The configuration used for UV-Vis spectroscopy analyses can be seen in Figure 13.

Figure 13 – Diagram of the absorbance spectroscopy configuration. A halogen light source (Philips 12345SL, 3100 K) and two optical fibers connected to a cuvette holder (CUV-ALL-UV, Ocean Optics) with 4.5 mL polystyrene cuvettes (K42-045, Olen) and the UV-Vis spectrometer (HR4000, Ocean Optics)

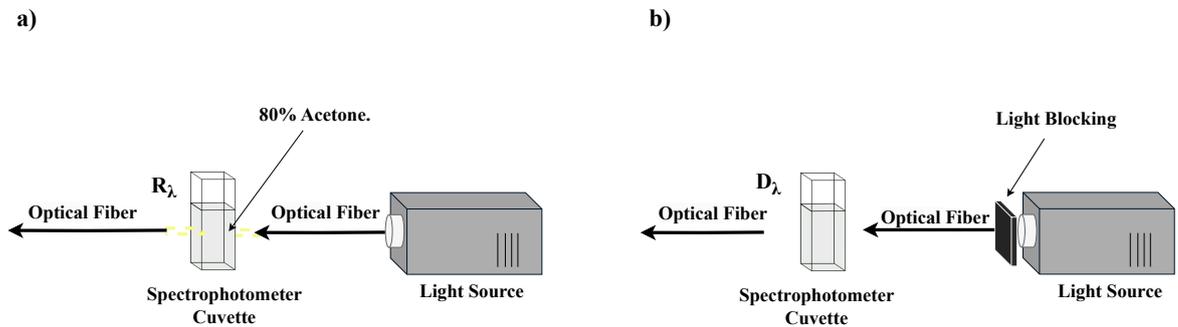


Source: Own authorship (2024).

To obtain absorption spectra, the software performs the following steps: first, we measure R_λ , which is the intensity of light transmitted through the cuvette with acetone 80%, as shown in Figure 14.a and 15. For D_λ , which represents the signal measured in the absence of incident light, the light source is blocked in the experiment (Figure 14b), resulting in the spectrum shown in Figure 15. It is important to subtract this background signal to correct for background noise and obtain an accurate measurement of the transmitted signal. The software provides the spectrum after the respective calculation via its graphical interface (Ocean Optics, Inc., 2005).

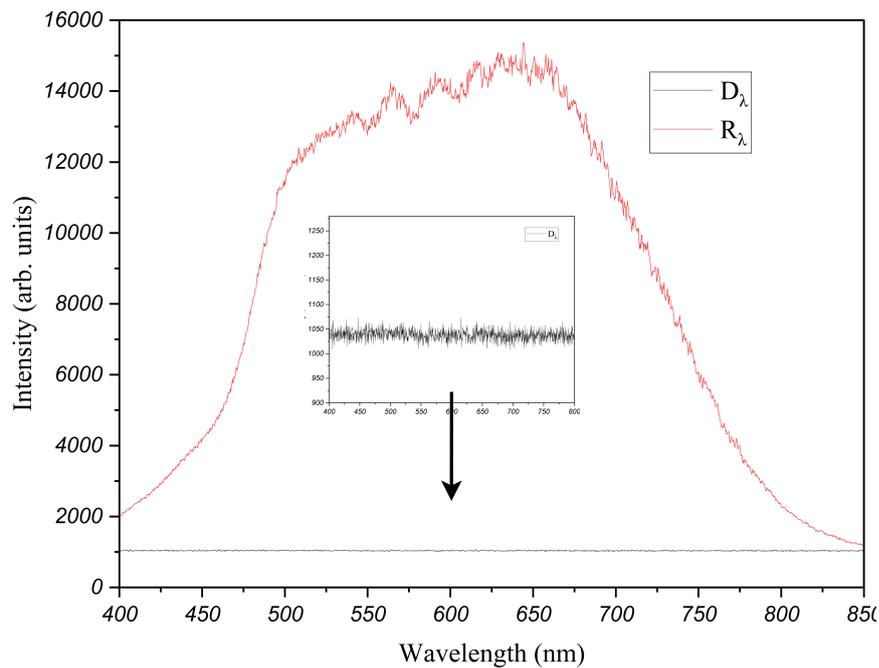
The software described in section 3.1.4 acquires the reflection spectrum and the dark spectrum (Figure 15) to provide the absorption spectrum.

Figure 14 – Absorption spectra acquisition process. To obtain the absorption spectra, the software performs the following steps: First, the intensity of light transmitted through a cuvette with 80% acetone (R_λ) is measured, as shown in Figure 14(a). For D_λ , the signal measured in the absence of incident light, the light source is blocked (Figure 14(b))



Source: Own authorship (2024).

Figure 15 – Graph of the spectrum for the intensity of light transmitted through a cuvette with 80% acetone (R_λ), starting from Figure 14(a). The dark reference spectrum, shown in Figure 14(b), represents the spectrum obtained under these conditions for D_λ , which indicates the absence of incident light



Source: Own authorship (2024).

Equation 3 is used to calculate optical density (OD), which is a measure of how much a sample absorbs light at a specific wavelength. The equation is based on the Lambert-Beer law (SWINEHART, 1962). The general formula for absorbance (A), previously known as optical density, provides a way to calculate the OD of a sample at a specific wavelength, taking into account both the background signal and the reference signal to obtain a more accurate and

reliable measure of OD (Ocean Optics, Inc., 2005). The equation is:

$$A_{\lambda} = -\log_{10} \left(\frac{S_{\lambda} - D_{\lambda}}{R_{\lambda} - D_{\lambda}} \right). \quad (3)$$

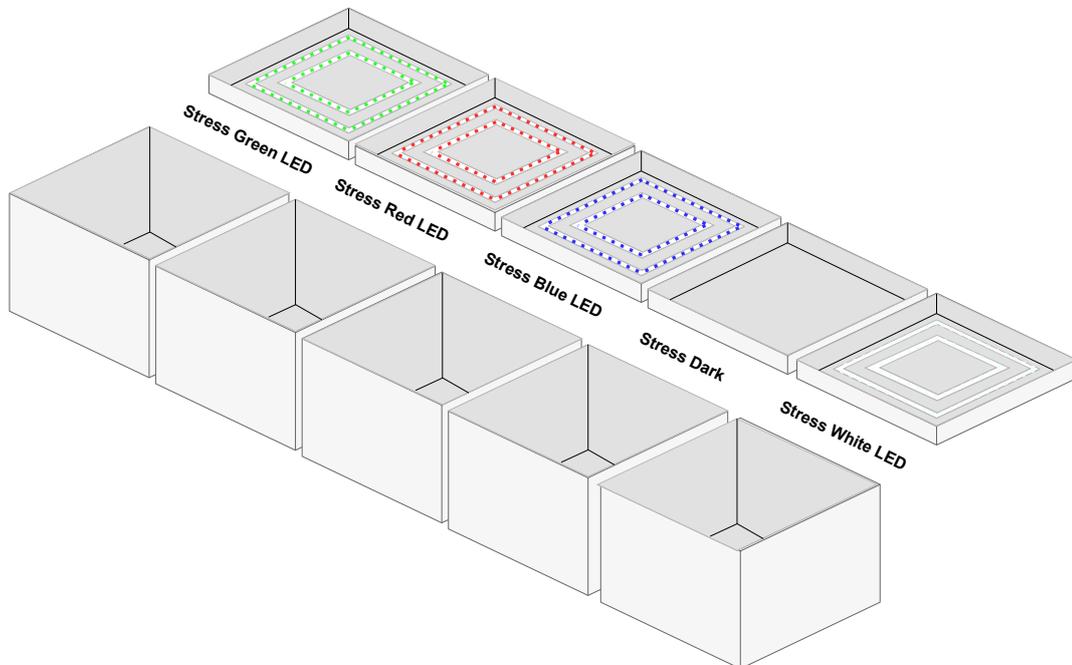
3.2 ENVIRONMENTAL CONSTRUCTION FOR WHEAT STRESS

In this section 3.2, environments were developed to create different stress conditions. It includes five cardboard containers, four with different types of LED lighting (green, red, blue, and white) and one container without lighting. The on/off control is managed by a timer controller. It includes five DHT11 sensors to monitor temperature and humidity throughout the experiment in each container.

3.2.1 Light Stress

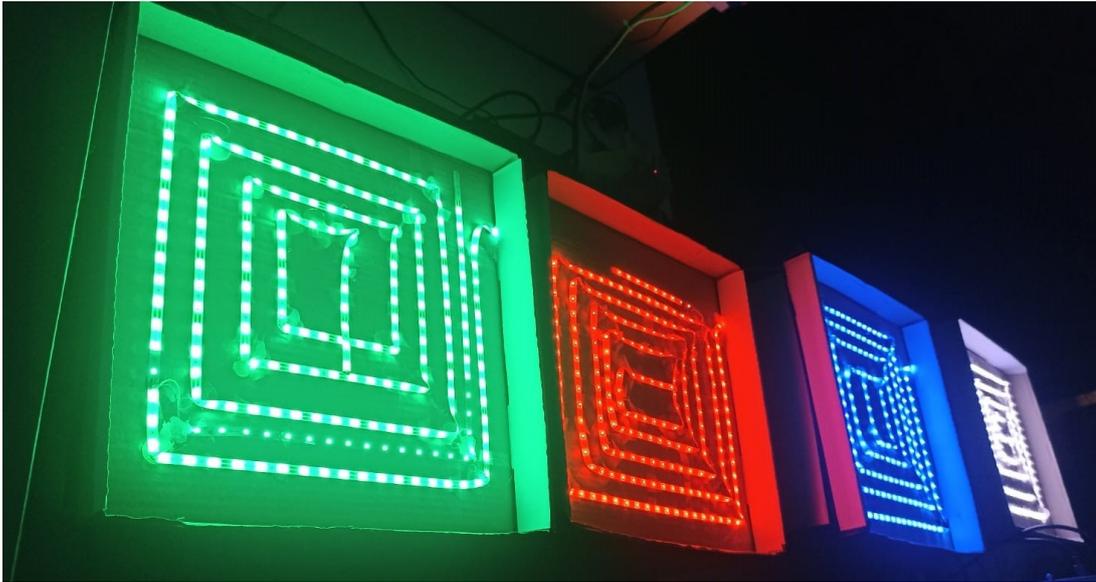
The five cardboard containers have the following dimensions: 36 x 35 x 23 cm, with a lid of 36.5 x 36.5 x 4 cm. LED strips of 12V DC, approximately 2 meters in length, were adhered in a spiral pattern on the top of the lid. The setup is shown in Figures 16 and 17.

Figure 16 – Stress containers with different LED lighting types: Green LED, Red LED, Blue LED, No LED, and White LED



Source: Own authorship (2024).

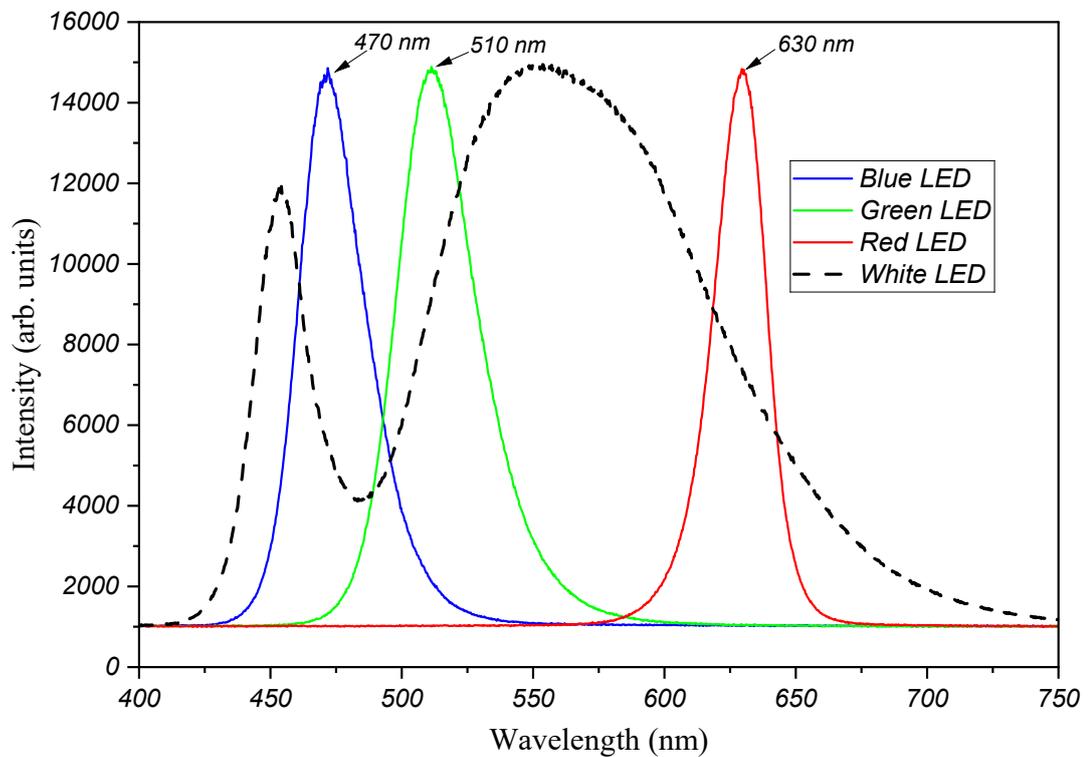
Figure 17 – LED strips installed on the lids of the containers, arranged in a spiral pattern.



Source: Own authorship (2024).

The spectral characteristics of the LED sources were measured with the HR4000 spectrometer. For the red LED, the peak wavelength is at 630 nm. For green LED emission, the peak wavelength is at 510 nm, and for blue LED emission, it is at 470 nm, as shown in Figure 18.

Figure 18 – Spectra of LED lights in containers

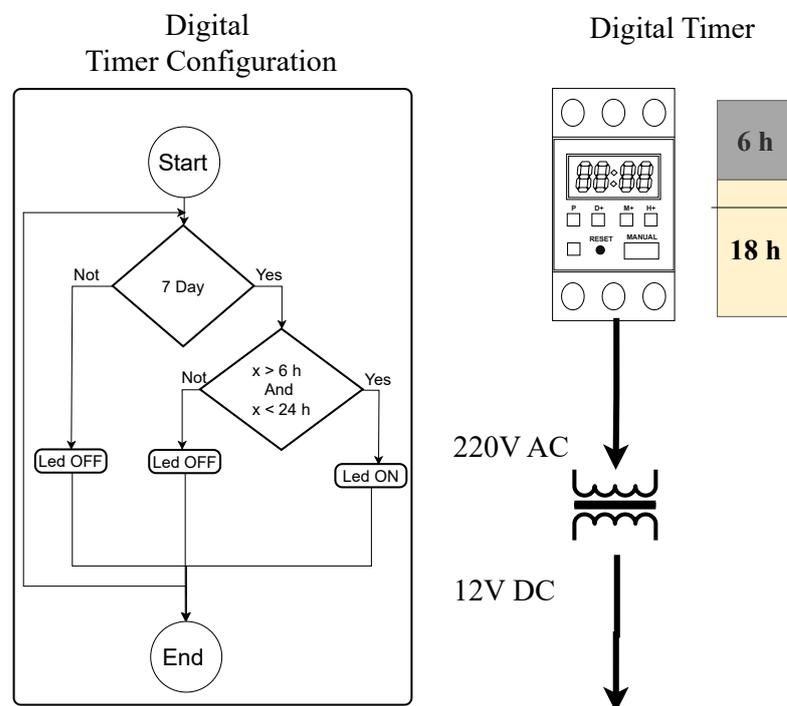


Source: Own authorship (2024).

3.2.1.1 Lighting Time Control

A digital timer is connected to a power distributor responsible for electrical current distribution. The power distributor is connected to various power supplies that convert 220VAC to 12VDC. Figure 19 also shows the time configuration of the digital timer, providing a detailed flow diagram that allows for understanding and visualizing how the time intervals for turning the LED strips on and off are programmed throughout the 7-day experiment.

Figure 19 – Diagram of the digital timer connected to a power distributor. Visual configuration for observing the time intervals in which the LED strips are turned on and off during the 7-day experiment.

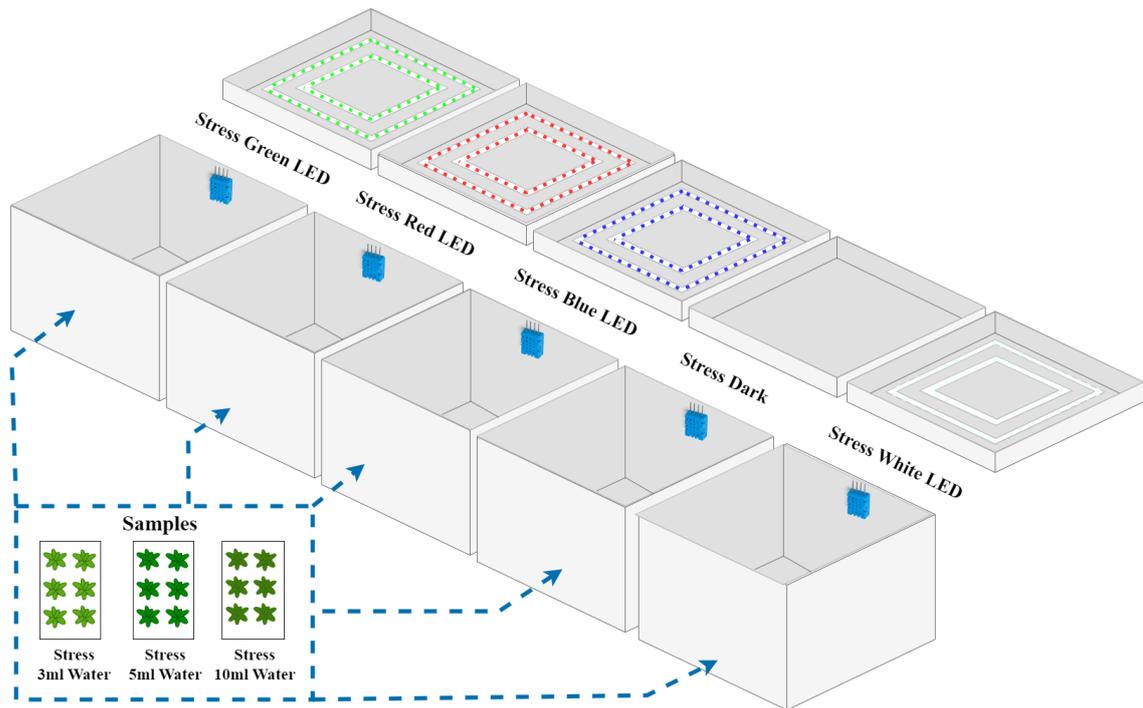


Source: Own authorship (2024).

3.2.2 Water Stress

In the experiment, wheat seeds were first soaked in distilled water for 18 hours (CHATTHA, 2017), washed and cleaned. Twenty seeds were placed in each of the five cardboard containers, each with three plastic containers and germination paper at the base. Daily, the seeds were stressed with 3 ml, 5 ml, or 10 ml of distilled water for 7 days. The distilled water was obtained from an Evox distiller described in section 3.1.1. Figure 20 shows the containers and distribution of the experiment, with each cardboard container having the respective light stress, totaling 15 different stress conditions.

Figure 20 – Stress Distribution in Containers for Water Stress Experiments (3 ml, 5 ml, 10 ml)



Source: Own authorship (2024).

3.2.3 Stress Conditions

Table 4 presents the details of fifteen stress conditions across different containers in the experiment. Each stress condition is characterized by two main variables: the color of LED light used (green, red, blue, or white) or light (dark) and the amount of water applied (in milliliters). The container number where each stress condition was carried out is also listed.

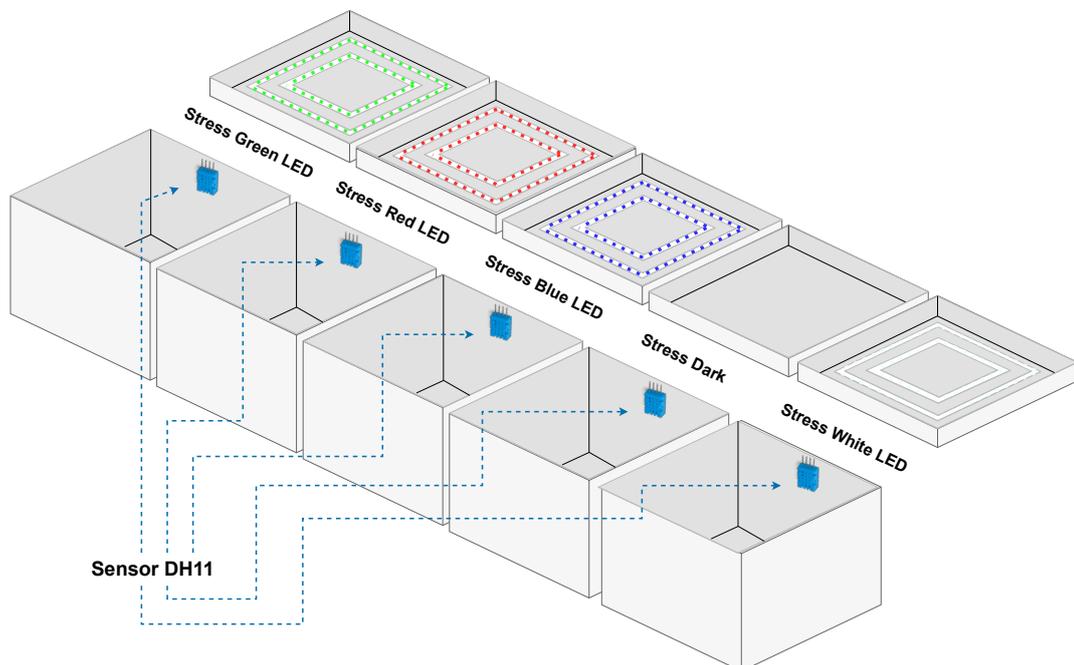
Table 4 – Stress conditions and details for each container. The conditions include the type of light (LED) applied and the amount of water stress (in ml) for each container.

	Stress Light LED	Water Stress (ml)	Container
Stress condition 01	Green	3.00	01
Stress condition 02	Green	5.00	01
Stress condition 03	Green	10.00	01
Stress condition 04	Red	3.00	02
Stress condition 05	Red	5.00	02
Stress condition 06	Red	10.00	02
Stress condition 07	Blue	3.00	03
Stress condition 08	Blue	5.00	03
Stress condition 09	Blue	10.00	03
Stress condition 10	Dark	3.00	04
Stress condition 11	Dark	5.00	04
Stress condition 12	Dark	10.00	04
Stress condition 13	White	3.00	05
Stress condition 14	White	5.00	05
Stress condition 15	White	10.00	05

3.2.4 Temperature and Humidity Monitoring

The components used for the implementation of the monitoring system include five DHT11 sensors distributed as shown in Figure 21. These sensors were chosen for their ability to measure temperature and humidity in a single electronic component. An Arduino board was used to process the analog signals from the DHT11 sensors. It was programmed to read the data from the sensors at one-minute intervals and send it to the Raspberry Pi for storage.

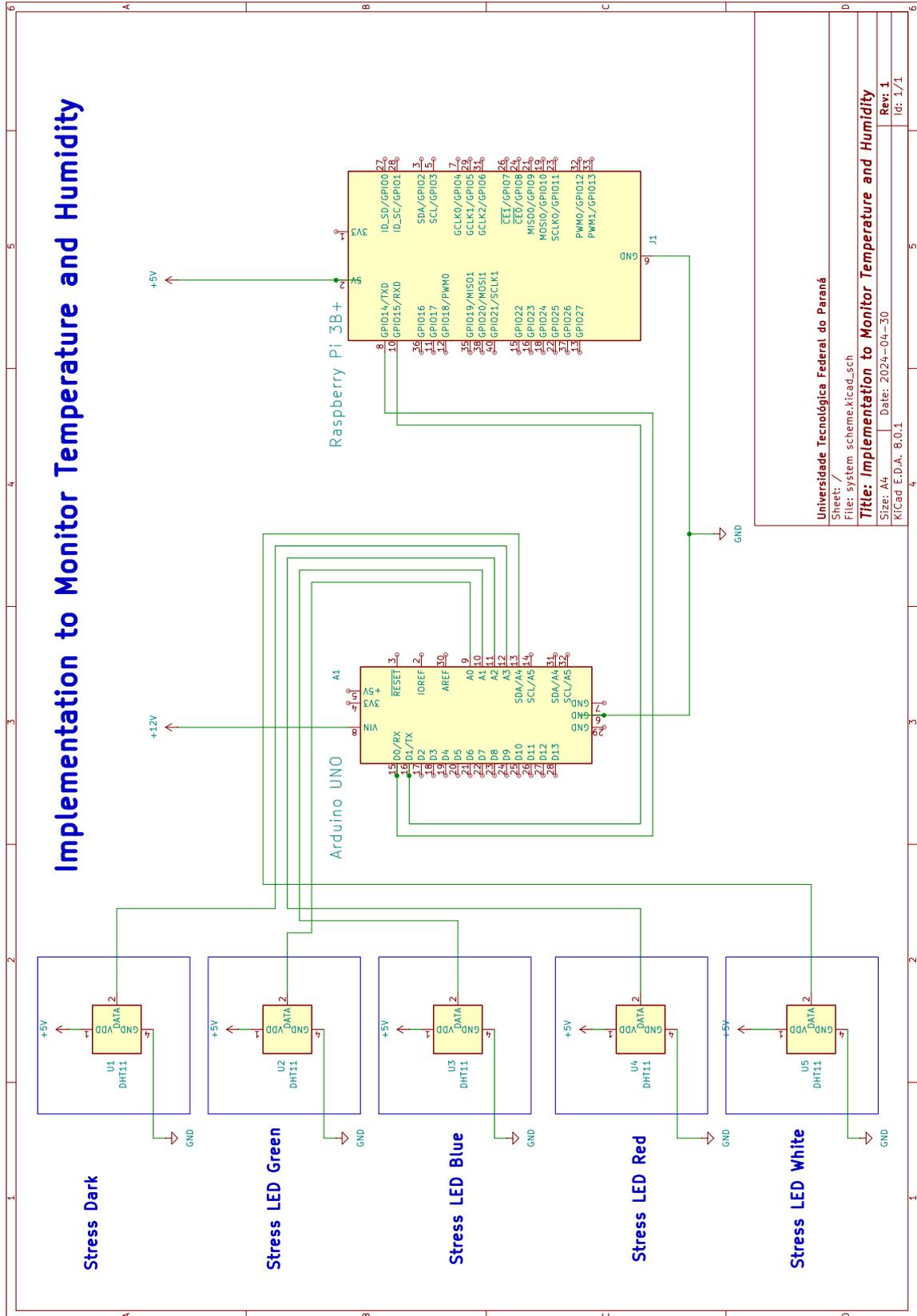
Figure 21 – Distribution of DHT11 Sensors for Monitoring Temperature and Humidity of Containers for Plant Stress



Source: Own authorship (2024).

Figure 22, shows the implementation of the system and its connections. The DHT11 sensors were connected to the Arduino board following the appropriate pin scheme and configured to continuously measure the temperature and humidity in the experimental environment. The Arduino board was programmed to read data from the DHT11 sensors every minute. This code also handled sending the data to the Raspberry Pi via a serial connection. Integration with the Raspberry Pi was achieved using Python scripts to receive and store the data from the Arduino board. The temperature and humidity monitoring system ran throughout the study period, recording data every minute, allowing for a detailed dataset on the environmental conditions in the experimental setup.

Figure 22 – Implementation of a monitoring system in stress containers for plants, composed of DHT11 sensors, Arduino Uno, Raspberry Pi 3B+

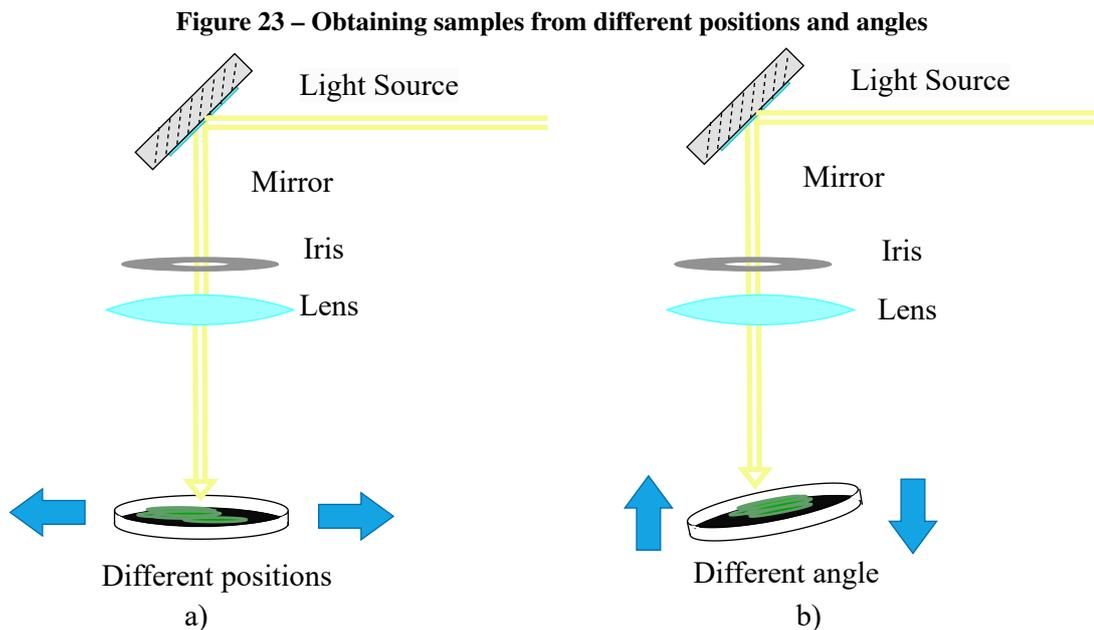


3.3 DATABASE OF REFLECTION AND ABSORPTION SPECTRA

For the construction of the database of reflection and absorption spectra according to the equipment configurations described in Subsection 3.3.1 and 3.3.2, we first obtained the diffuse reflection spectra by modifying the angle at which the light strikes the sample. For the absorption spectra, we constructed it considering two types of processes to obtain the total chlorophyll from the samples, which are detailed in the following subsections.

3.3.1 Acquisition of Reflection Spectra

The spectra were obtained in a multi-angular setup as shown in Figure 23, varying the angle of the petri dish for a fixed light beam, generating diffuse reflected light. For each sample, 23 diffuse reflection spectra were taken at different light incidence angles, for 15 different stress conditions, making a total of 345 spectra.



Source: Own authorship (2024).

3.3.2 Methodology for Chlorophyll Extraction and Quantification

Through absorbance spectroscopy, the chlorophyll content can be obtained. In our experiment, using fresh leaves as samples, the following procedure was carried out to obtain the chlorophyll content ($\text{mg}\cdot\text{g}^{-1}$ FW). 5 ml of acetone 80% were placed in a mortar and the

leaves were macerated using a pestle. The solution was then centrifuged at 12,000 rpm for 10 minutes, and the supernatant was transferred to a cuvette (ARNON, 1949). For quantification, two different devices were used. First, the equipment (digital spectrophotometer IL-593-S) described in section 3.1.3 was used. Readings were taken in the spectrophotometer at wavelengths of 645 nm and 663 nm. Afterward, the liquid samples were preserved on ice to be measured later by the equipment (HR4000 spectrometer) described in 3.1.2, a spectrometer that returns intensities over a range of wavelengths.

Using the values obtained from the spectrophotometer and spectrometer, chlorophyll content can be quantified following the protocol described in (WITHAM *et al.*, 1971). Based on these optical density (OD) readings, the contents of chlorophyll a, b, and total will be calculated in mg/g, independently for each device:

$$\text{Chlorophyll a} = \frac{12.7 \cdot \text{Abs663} - 2.69 \cdot \text{Abs645}}{1000 \cdot \text{FW}} \cdot V \text{ (mg} \cdot \text{g}^{-1}\text{)}, \quad (4)$$

$$\text{Chlorophyll b} = \frac{22.9 \cdot \text{Abs645} - 4.68 \cdot \text{Abs663}}{1000 \cdot \text{FW}} \cdot V \text{ (mg} \cdot \text{g}^{-1}\text{)}, \quad (5)$$

$$\text{Total chlorophyll} = \frac{8.02 \cdot \text{Abs663} + 20.2 \cdot \text{Abs645}}{1000 \cdot \text{FW}} \cdot V \text{ (mg} \cdot \text{g}^{-1}\text{)}, \quad (6)$$

in which:

- Abs645 and Abs663 are the optical densities at 645 and 663 nm, respectively;
- V is the volume in *ml* acetone 80% ;
- FW is the fresh weight of material used in the extract.

3.4 PRE-PROCESSING

In Section 3.4, we address various techniques and methods for processing and analyzing spectroscopic data. It begins with the selection of the optical data range from the HR4000 spectrometer, limiting the data from 400 nm to 780 nm for pigment analysis. It also describes the procedure used to obtain integer wavelengths (645 nm and 663 nm) in the spectrometer described in Subsection 3.1.2, which will be used for the calculations required in Subsection 3.3.2. Techniques such as the Savitzky-Golay filter for smoothing spectral signals and Continuum

Removal are also addressed, which highlight specific spectral features by normalizing each spectrum by its maximum reflectance value, eliminating unrelated intensity variations. To correct systematic variations due to optical interferences, techniques such as Multiplicative Scatter Correction (MSC), which adjusts spectra by removing scattering effects, and Standard Normal Variate (SNV), which standardizes spectral data to correct dispersion variations, are applied. These techniques are used as preprocessing in the spectral data.

3.4.1 Processing Spectral Data by Wavelength Intervals

The first step in data processing will be the selection of the data range. According to (HECHT, 2017), Maximum Wavelength of the Band red color is defined at 780 nm. Knowing that the analysis is for pigment, and also considering (BAURIEGEL *et al.*, 2011; WENDAR *et al.*, 2014), they used a minimum range of 400 nm for their analyses, so we limit our data to the range of 400 nm to 780 nm.

According to the methodology for quantifying chlorophyll described in Subsection 3.3.2, two specific integer wavelengths are required for the respective calculation (645 nm and 663 nm). In our experiment, data was obtained in wavelength steps of 0.27 nm within the spectral range of 400 to 780 nm; therefore, the procedure to set the wavelengths in steps of 1 nm is necessary. The wavelength values obtained through processing in Python apply functions to extract the integer part and return the average intensities at each integer wavelength.

- Grouping columns by their integer values: For each column j , we determine its group based on the integer part of the value in the first row:

$$\text{Group}(k) = \{j \mid \lfloor \text{wavelength}_j \rfloor = k\}.$$

- Calculating the mean for each group of columns: For each row i and each group k :

$$\bar{R}_{ik} = \frac{1}{|\text{Group}(k)|} \sum_{j \in \text{Group}(k)} R_{ij},$$

where $|\text{Group}(k)|$ is the number of columns in group k .

Figure 24 – Interpolated Wavelength Data

Original Wavelength Data

wavelength nm	400.03	400.30	400.57	400.83	401.10	401.37	849.98	850.23	850.48	850.73	850.98
------------------	--------	--------	--------	--------	--------	--------	-------	--------	--------	--------	--------	--------

Data Wavelength

wavelength nm	400	401	402	403	404	405	846	847	848	849	850
Spectrum 1	$\bar{R}_{1,400}$	$\bar{R}_{1,401}$	$\bar{R}_{1,402}$	$\bar{R}_{1,403}$	$\bar{R}_{1,404}$	$\bar{R}_{1,405}$	$\bar{R}_{1,846}$	$\bar{R}_{1,847}$	$\bar{R}_{1,848}$	$\bar{R}_{1,849}$	$\bar{R}_{1,850}$

Source: Own authorship (2024).

3.4.2 Savitzky-Golay Filter

The Savitzky-Golay filtering method is a signal processing technique used in spectroscopy to improve the accuracy of data without distorting the essential features of the original spectra. This technique smooths the signal while maintaining an appropriate signal-to-noise ratio. It is especially effective in preserving the high-frequency components of the signal and peaks in spectra (SAVITZKY; GOLAY, 1964; ZAHIR *et al.*, 2022).

For a data point y_i , the Savitzky-Golay filter fits a polynomial of order p to all data points within a window of length $2m + 1$ centered on y_i . The general equation for the smoothed value \hat{y}_i is:

$$\hat{y}_i = \sum_{j=-m}^m c_j y_{i+j},$$

where:

- \hat{y}_i is the smoothed value at point i ;
- y_{i+j} are the data points within the window;
- c_j are the filter coefficients, which depend on the order of the polynomial p and the window size $2m + 1$.

3.4.3 Continuum Removal (CR)

The *Continuum Removal* technique is used in spectral analysis to highlight specific spectral features of a spectrum. This technique normalizes each spectrum by dividing it by its maximum value, allowing for clearer observation of the features of interest by eliminating intensity variations that are not related to the spectral features (CLARK *et al.*, 1990).

Given an original spectrum $R(\lambda)$, the *Continuum Removal* process is carried out by dividing the signal value at each point by the maximum value $\max(R(\lambda))$ of the entire spectrum. This process is fundamental in spectral analysis to eliminate the continuum or background component and is expressed as:

$$R_{\text{Norm}}(\lambda) = \frac{R_{\lambda}}{\max(R(\lambda))},$$

where:

- $R(\lambda)$ is the reflectance at wavelength λ ;
- $\max(R(\lambda))$ is the maximum reflectance value across the entire spectrum.

3.4.4 De-trending (DT)

De-trending is a process used to remove unwanted trends or variations from spectra, particularly useful when dealing with nonlinear baselines or systematic variations. To adjust or smooth the original spectrum $R(\lambda)$, the goal is to find a function $\hat{R}(\lambda)$ that represents the adjusted spectrum (ZAHIR *et al.*, 2022; KIYONO; TSUJIMOTO, 2016), and is expressed as:

$$\hat{R}(\lambda) = R(\lambda) - \text{DT}(R(\lambda)),$$

where:

- $\text{DT}(R(\lambda))$ is the function resulting from the de-trending process;
- The result \hat{R} is a matrix of the same size as R , where each row \hat{R}_i has been detrended (the linear trend has been removed).

3.4.5 Normalization of Reflectance Spectra by Area

Normalization by area in the reflection spectrum is performed by dividing each reflectance value by the area under the curve corresponding to that row within the analyzed

range (JIAO *et al.*, 2020). In a specific wavelength range (W_{\min} to W_{\max}), the equation for Area-Normalized Reflectance is applied:

$$\text{Area-Normalized Reflectance}(i, \lambda) = \frac{R(i, \lambda)}{\int_{W_{\min}}^{W_{\max}} R(i, \lambda) d\lambda}, \quad (7)$$

where:

- $R(i, \lambda)$ is the reflectance value for sample i at wavelength λ ;
- $\int_{W_{\min}}^{W_{\max}} R(i, \lambda) d\lambda$ is the area under the reflectance curve for sample i over the wavelength range from W_{\min} to W_{\max} .

3.4.6 Standard Normal Variate (SNV)

This is a preprocessing technique used to correct scattering variations and improve comparability between different samples. This technique is particularly useful in spectroscopy, allowing for the correction of samples. SNV standardizes the spectral data by adjusting for differences in baseline shifts and scaling effects due to sample presentation (ZAHIR *et al.*, 2022; BARNES *et al.*, 1989). The SNV procedure is applied to each spectrum in a collection of spectra individually, minimizing differences due to scattering and other unwanted variations (YANG *et al.*, 2020; MISHRA *et al.*, 2020). The mathematical representation is:

$$\text{SNV}(X_i) = \frac{X_i - \bar{X}}{\sigma_X}, \quad (8)$$

where:

- X_i is the signal intensity in the original spectrum at point i ,
- \bar{X} is the mean of the signal intensities across the entire spectrum,
- σ_X is the standard deviation of the signal intensities in the spectrum.

3.4.7 Multiplicative Scatter Correction (MSC)

This method can be used as a preprocessing step to eliminate optical interferences (MALEKI *et al.*, 2007), used to correct the scattering effect in measured spectra resulting from physical variations in the samples (CHEN; THENNADIL, 2012). First, a straight line (a first-degree polynomial) is fitted between the mean spectrum and the current spectrum.

This correction normalizes each point in the current spectrum, eliminating systematic variations due to light scattering. The equation for MSC correction is:

$$\text{corrected_spectrum}[i] = \frac{\text{spectrum}[i] - \text{fit}[1]}{\text{fit}[0]},$$

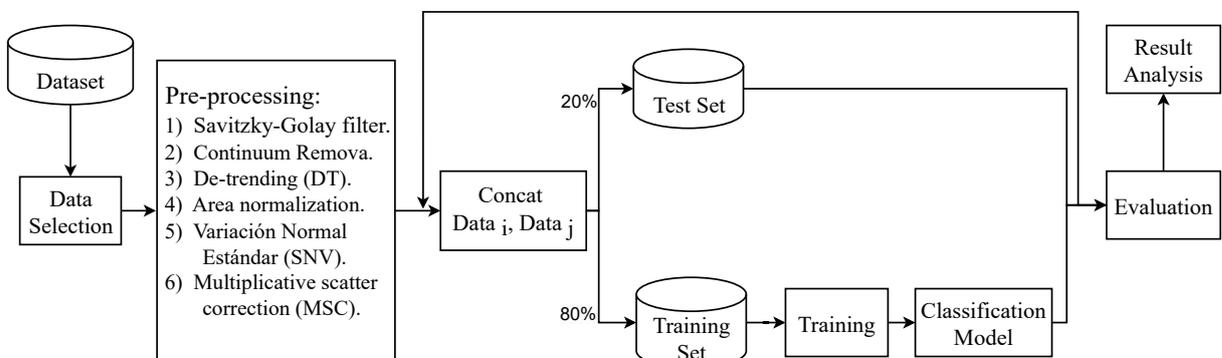
where:

- $\text{spectrum}[i]$ is the value of the original spectrum at position i ;
- $\text{fit}[1]$ is the intercept of the linear fit;
- $\text{fit}[0]$ is the slope of the linear fit.

3.5 MACHINE LEARNING MODEL

Figure 25, illustrates the comprehensive process of training a supervised classification model in the context of machine learning. The flow begins with the selection of relevant data from the initial dataset, followed by a series of preprocessing techniques to enhance data quality. Subsequently, the preprocessed data are concatenated to form a combined dataset (all possible options and their individual responses as performed by the researchers Urbanovich *et al.* (2021)). This new dataset is then split into 80% for training and 20% for testing. The model is trained on the training set and then evaluated on the testing set to determine its performance. Finally, the evaluation results are analyzed to understand the effectiveness of the model and make necessary adjustments. This process ensures the creation of a model with various types of preprocessing, selecting the one that achieves the best result.

Figure 25 – Overview of the classification process

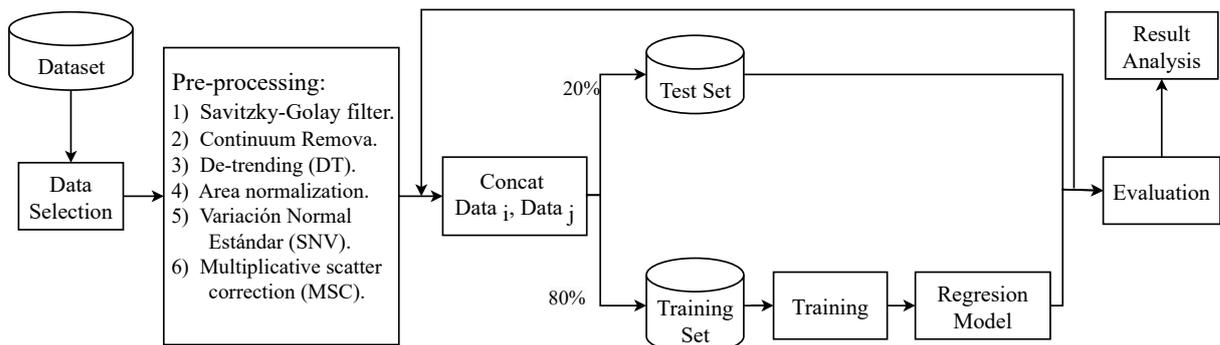


Source: Own authorship (2024).

Figure 26, shows a schematic of the process for selecting the best preprocessing for a regression model. First, the dataset is selected and subjected to several preprocessing stages,

including the Savitzky-Golay filter, continuum removal (Continuum Removal), detrending (Detrending, DT), area normalization, standard normal variate (SNV), and multiplicative scatter correction (MSC). Then, the preprocessed data are divided into two subsets: 80% for training and 20% for testing. The training data are used to train the regression model, while the testing data are used to evaluate the model's performance. Finally, the evaluation results are analyzed to determine the effectiveness of the regression model.

Figure 26 – Overview of the regression process



Source: Own authorship (2024).

3.5.1 Classification Models

We chose MiniRocket, a variant of the Random Convolutional Kernel Transform (ROCKET), as our classification model. MiniRocket offers several advantages over other methods. Used in time series classification applications, MiniRocket simplifies the process by eliminating the need to normalize input time series, making it easier to apply across various fields without compromising performance (DEMPSTER *et al.*, 2021).

3.5.1.1 MiniRocket Classifier

Tsai is an open-source deep learning package developed on top of PyTorch and fast.ai, specializing in advanced techniques for time series tasks such as classification and regression (OGUIZA, 2023). MiniRocket significantly improves the efficiency of the process by adjusting specific convolutions that maintain the effectiveness of the original method but with much faster computation speeds. This approach uses feature transformations based on random convolutions over the time series, followed by linear classifier training (DEMPSTER *et al.*, 2021; BONDUGULA *et al.*, 2023).

MiniRocket uses the features of a small and fixed set of two-value kernels, as well as the proportion of positive values (PPV) index, to significantly improve the transformation through four fundamental optimizations.

- i. Calculate the PPV for W and $-W$ simultaneously: Where $C = X * W - b$, the PPV is given by:

$$PPV(C) = \frac{1}{n} \sum [c > 0].$$

The PPV is between 0 and 1, and its complement, the proportion of negative values (PNV), is defined as $1 - PPV(X * W - b) = PNV(X * W - b)$. Therefore, calculating the PPV also yields the PNV, and vice versa, as both are equivalent. The convolution operation is associative, meaning $X * -W = -(X * W)$. Thus, by calculating the PPV for a kernel W , the PNV for $-W$ is also obtained. In practice, it is only necessary to perform the convolution for W , as the inverted kernel $-W$ is implicit. MiniRocket uses only kernels with weights $\alpha = -1$ and $\beta = 2$, avoiding the use of inverted kernels with weights $\alpha = 1$ and $\beta = -2$ (DEMPSTER *et al.*, 2021; DEMPSTER, 2022).

- ii. Leverage the convolution output to generate multiple features, by using the same kernel and dilation to extract various features. With a kernel W and a dilation d , $C = X * W_d$ is initially calculated. Subsequently, the output of this convolution, C , is used to generate different features for various bias values. In this way, multiple features can be obtained with the computational cost of a single convolution operation (DEMPSTER *et al.*, 2021; DEMPSTER, 2022).
- iii. Avoid multiplications in the convolution operation. Restricting the kernel weights to two values, α and β , allows replacing multiplications in the convolution with additions. This is achieved by precomputing $A = \alpha X$ and $B = \beta X$, thus eliminating the need for multiplications during the convolution (DEMPSTER *et al.*, 2021; DEMPSTER, 2022).
- iv. Process all kernels (almost) at once for each dilation. By restricting the kernel weights to two values, it is possible to calculate almost all kernels simultaneously for each dilation, using the convolution output, C_α , with $\alpha = -1$ for all weights, and adjusting C_α (DEMPSTER *et al.*, 2021; DEMPSTER, 2022).

3.5.2 Regression Models

For regression models in the analysis of spectral data, two main approaches have been considered: MiniRocket, which is based on time series, and Deep Learning, due to the complexity of the data analyzed. MiniRocket is a powerful and efficient tool for time series regression, standing out for its ability to handle large volumes of data in shorter times. On the other hand, in our experiment, Deep Learning models have also been considered, offering the advantage of identifying complex and non-linear patterns in spectral data, which is crucial for the analysis in our proposal.

3.5.2.1 MiniRocket Regressor

Minimally Random Convolutional Kernel Transform (MiniRocket) applications demonstrate that it surpasses traditional approaches in both accuracy and efficiency, achieving very acceptable scores in short times of seconds. The algorithm shows robustness even in data-limited scenarios (ALAGOZ, 2024). MiniRocket is adapted as a regressor by transforming the input signal into a high-dimensional feature vector (9,996 dimensions), similar to its use in classification. The process involves dilated convolutions with predefined kernels, applying selected dilations to capture different temporal scales. The results are compared with threshold values, generating binary vectors. These are averaged to produce a feature descriptor. This descriptor is then used as input for a regression model, allowing the prediction of continuous values in time series (ALAGOZ, 2024; DEMPSTER *et al.*, 2021; OGUIZA, 2023; SCHLEGEL *et al.*, 2022).

MiniRocket outperforms Rocket in terms of speed, being up to 75 times faster on large datasets and almost deterministic while maintaining nearly identical accuracy levels. This method offers unparalleled efficiency compared to other methods with similar accuracy, including Rocket, making it significantly faster. Additionally, MiniRocket achieves superior accuracy compared to methods with comparable computational costs (ALAGOZ, 2024; DEMPSTER *et al.*, 2021; OGUIZA, 2023; SCHLEGEL *et al.*, 2022). The parameters used were: `greater_is_better=False`, indicating that lower MSE values are better, and `mean_squared_error`, where the model will focus on minimizing the mean squared error during training and will be evaluated on how well it minimizes this error.

3.5.2.2 DNN Regressor

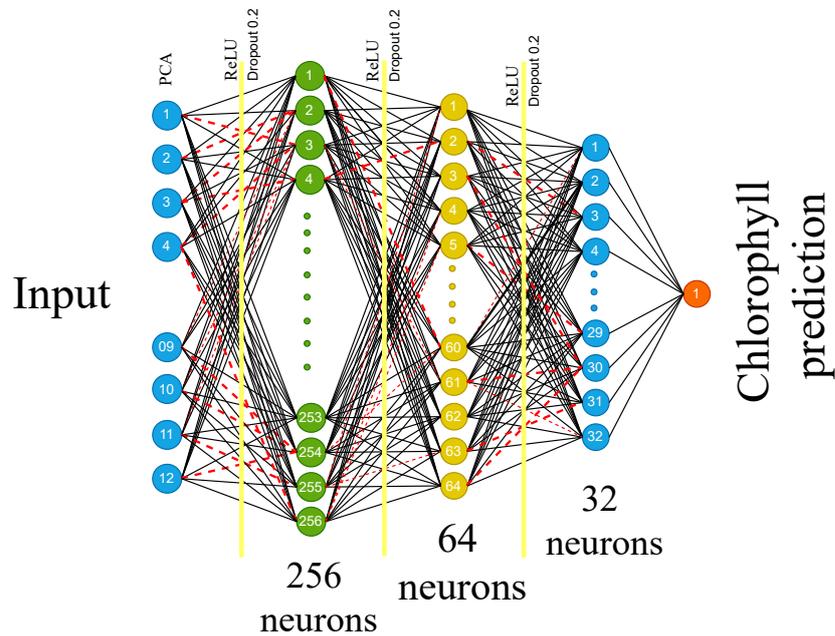
DNNs can approximate complex functions through deep non-linear network structures, enabling them to learn essential features from datasets, also due to their ability to identify patterns in complex datasets (CAO *et al.*, 2019; GOODFELLOW *et al.*, 2016). In the DNN architecture, we encounter terms such as the cost function (or loss function), which is intended for statistical estimation. Minimizing this cost function facilitates obtaining a maximum likelihood estimate (GOODFELLOW *et al.*, 2016). Optimization in the training of deep models is crucial as it allows for adjusting model parameters to improve performance. Additionally, dropout is used to prevent overfitting. Due to their high capacity to learn and represent complex tasks (GOODFELLOW *et al.*, 2016), deep models also benefit from the Rectified Linear Unit (ReLU) activation function, which helps the model learn useful representations of the input data (HE *et al.*, 2018; GOODFELLOW *et al.*, 2016). The approximation capability of DNNs with ReLU is analyzed through the following Equation 9:

$$\text{DNN}_1^m = \left\{ f : f = \sum_{i=1}^m \alpha_i \text{ReLU}(w_i x + b_i) + \beta \right\}, \quad (9)$$

where $\alpha_i, b_i, \beta \in \mathbb{R}$ and $w_i \in \mathbb{R}^{1 \times d}$. This equation describes the set of functions produced by a deep neural network (DNN) with a single hidden layer, where $\text{ReLU}(w_i x + b_i)$ acts as the activation function (HE *et al.*, 2018).

In Figure 27, the DNN model architecture is illustrated, consisting of 12 principal components (PCA) as input and several dense layers with ReLU activation functions and dropout layers to prevent overfitting. The architecture includes an initial dense layer with 256 units, followed by dense layers with 64 and 32 units, and a single-unit output layer. The model is compiled with the RMSprop optimizer and a learning rate of 0.001, minimizing the mean squared error (MSE) and evaluating with mean absolute error (MAE) and mean squared error (MSE) metrics. It is trained for 700 epochs with a batch size of 32 and validated using a separate dataset.

Figure 27 – Deep Learning Model Architecture with ReLU, Dropout 0.2, and dense layers (256, 64, 32 neurons) for Chlorophyll prediction



Source: Own authorship (2024).

3.6 EVALUATION METRICS

In predictive models, evaluating their performance in both classification and regression tasks is of utmost importance. For classification models, we used metrics such as accuracy, precision, recall, and F1-score, as well as Macro Average, Weighted Average, and the confusion matrix, to measure the effectiveness of the models in making correct categorizations. In contrast, for regression models, we used Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2), which quantify the variation between predicted values and actual values. Both categories of metrics ensure the correct evaluation of model performance, allowing for a precise assessment of its ability in both classification and regression tasks.

In our study, we also considered the measurement of chlorophyll using two different devices. To evaluate the relationship between the measurements obtained with each device, we employed Pearson's correlation, a statistical measure that quantifies the strength and direction of the linear relationship between the results.

3.6.1 Evaluation Metrics for Model Classifiers

In the study, metrics for multiclass classifiers were considered (HOSSIN; SULAIMAN, 2015). Our samples contain three levels of water stress as well as five types of illumination to which the wheat leaves were subjected. Additionally, the experiment ensured that the classes were balanced, meaning that the same number of reflection spectra were obtained for each sample. According to this, the following metrics were used:

- **Accuracy:** A metric used to measure the performance of a classifier (PETHE *et al.*, 2024; ZHANG *et al.*, 2020; RIZWAN *et al.*, 2019).

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{True Positives} + \text{False Positives} + \text{False Negatives} + \text{True Negatives}}.$$

- **Precision:** Measures the proportion of correctly predicted positive patterns relative to the total predicted positive patterns, defined as the ratio between true positives and the total predicted positives (PETHE *et al.*, 2024; RIZWAN *et al.*, 2019). It is calculated as:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}.$$

- **Recall:** A performance measure defined as the proportion of true positives (TP) relative to the total of true positives (TP) and false negatives (FN) (PATHARKAR *et al.*, 2024; ZHENG; JIN, 2020).

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}.$$

- **F1-Score:** The weighted harmonic mean of Precision and Recall (HOSSIN; SULAIMAN, 2015; PETHE *et al.*, 2024).

$$\text{F1-Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}.$$

- **Confusion Matrix:** A comparison of estimated class hits and errors with the actual class for each instance in the dataset, allowing us to analyze how the model performs in each specific class (KRSTINIĆ *et al.*, 2023; LUQUE *et al.*, 2021).

The Macro Average and the Weighted Average were used to calculate precision, recall, and F1-Score for each class separately, applying the equations in (10) to (15) (FLORES *et al.*, 2024; ZHOU *et al.*, 2021):

$$\text{Macro Precision} = \frac{1}{N} \sum_{i=1}^N \text{Precision}_i, \quad (10)$$

$$\text{Macro Recall} = \frac{1}{N} \sum_{i=1}^N \text{Recall}_i, \quad (11)$$

$$\text{Macro F1 Score} = \frac{1}{N} \sum_{i=1}^N \text{F1 Score}_i, \quad (12)$$

$$\text{Weighted Precision} = \frac{\sum_{i=1}^N \text{True Positives}_i}{\sum_{i=1}^N (\text{True Positives}_i + \text{False Positives}_i)}, \quad (13)$$

$$\text{Weighted Recall} = \frac{\sum_{i=1}^N \text{True Positives}_i}{\sum_{i=1}^N (\text{True Positives}_i + \text{False Negatives}_i)}, \quad (14)$$

$$\text{Weighted F1 Score} = \frac{\sum_{i=1}^N 2 \times \text{Precision}_i \times \text{Recall}_i}{\sum_{i=1}^N (\text{Precision}_i + \text{Recall}_i)}. \quad (15)$$

3.6.2 Regression Model Evaluation Metrics

To evaluate the performance of the model, the performance metrics are calculated: Root Mean Squared Error (RMSE), R-squared (R^2) score, and Mean Absolute Error (MAE) (STEURER *et al.*, 2021; CHICCO *et al.*, 2021; GONZÁLEZ-SOPENA *et al.*, 2021; RAINIO *et al.*, 2024).

- **Root Mean Squared Error (RMSE):** Square root of the MSE:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}. \quad (16)$$

- **Mean Absolute Error (MAE):** Average of the absolute errors:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|. \quad (17)$$

- **R^2 (Coefficient of Determination):** Proportion of the variance in the dependent variable that is explained by the model:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}. \quad (18)$$

where y_i represents the observed values, \hat{y}_i represents the values predicted by the model, \bar{y} is the mean of the observed values, and n is the total number of observations.

3.6.3 Comparison of Values from Different Equipment

Pearson correlation allows for an analysis to determine how related the values obtained by two different equipment from (Digital Spectrophotometer IL-593-S and HR4000 Spectrometer) the same sample are (ARMSTRONG, 2019):

$$r_{xy} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}. \quad (19)$$

4 RESULTS AND DISCUSSIONS

This Chapter analyzes the water and light stress environment for wheat leaves. Section 4.1 addresses the results obtained under different stress conditions in the containers for temperature and humidity, providing an analysis of the environmental factors in each cardboard container. Section 4.2 explores the spectral databases, focusing on analyzing reflectance and absorption spectra. The results obtained from two instruments, the IL-593-S spectrophotometer and the HR4000 spectrometer, are detailed, and a comparison between the two is made to determine the total chlorophyll content. The Chapter concludes with the search for machine learning models in Section 4.3 and 4.4. These models are used to classify water and light stress and estimate the total chlorophyll content using multi-angular diffuse reflection spectra, demonstrating the capability of artificial intelligence to predict and quantify the impact on wheat plants under this characteristic.

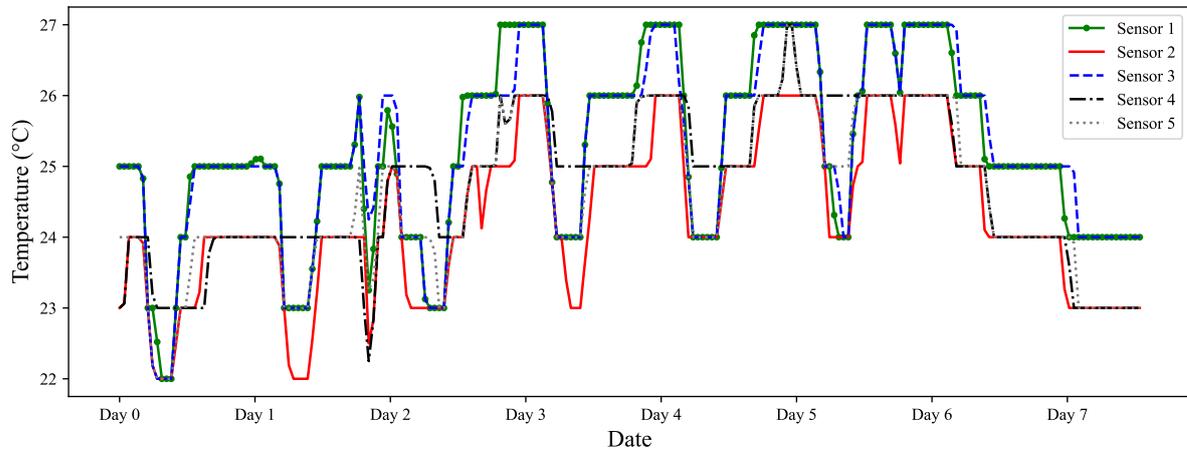
4.1 WATER AND LIGHT STRESS ENVIRONMENT FOR WHEAT LEAVES

To ensure that the containers with different stress conditions received uniform treatment, both temperature and humidity were monitored using the DHT11 sensor installed in the five containers, obtaining readings throughout the experiment.

4.1.1 Result of Temperature Conditions in Stress Environment

In Figure 28, the data collected by five temperature sensors, each located in a container subjected to different lighting conditions (with green, red, blue, and white LEDs, and one without lighting), is shown. Each sensor recorded a total of 10,502 measurements obtained over a period of 7 days. Additionally, Table 5 shows the statistical summary of the five temperature sensors.

Figure 28 – Graph of data collected by five temperature sensors in containers under various lighting conditions (green, red, blue, white LEDs, and no lighting). Each sensor recorded 10,502 measurements over a period of 7 days



Source: Own authorship (2024).

Sensor Temperature 1 (Container with Stress Light Green) recorded a mean of 25.3 °C with a standard deviation of 1.3 °C. The temperature values ranged from a minimum of 22.0 °C to a maximum of 27.0 °C, with a median of 25.0 °C. The temperature measurements under this green light condition are considered similar compared to other containers.

Sensor Temperature 2 (Container with Stress Light Red) showed a mean of 24.3 °C with a standard deviation of 1.1 °C. The values ranged from a minimum of 22.0 °C to a maximum of 26.0 °C, with a median of 24.0 °C. The lower mean and standard deviation compared to the previous sensor suggest that red light conditions present lower maximum temperatures compared to other environments.

For Sensor Temperature 3 (Container with Stress Light Blue), the mean was 25.3 °C and the standard deviation was 1.3 °C. The temperature values fluctuated between a minimum of 22.0 °C and a maximum of 27.0 °C, with a median of 25.0 °C. The similarity in the mean and median with Sensor Temperature 1 suggests a comparable response in the containers under blue and green light.

Sensor Temperature 4 (Container with Stress Light Dark) reported a mean of 24.7 °C with a standard deviation of 1.1 °C. The values ranged from a minimum of 22.0 °C to a maximum of 27.0 °C, with a median of 25.0 °C. Despite the slightly lower mean, the minimum and maximum ranges are similar compared to other containers.

Finally, Sensor Temperature 5 (Container with Stress Light White) showed a mean of 24.6 °C with a standard deviation of 1.1 °C. The values ranged from a minimum of 22.0 °C to a

maximum of 27.0 °C, with a median of 24.0 °C. This suggests a relatively stable response, similar to red light in terms of variability and in the range of minimum and maximum temperatures compared to other containers.

Table 5 – Statistical results of data collected by five temperature sensors in containers under various lighting conditions (green, red, blue, white LEDs, and no lighting). Each sensor recorded 10,502 measurements over a period of 7 days

Sensor Temperature	Container with Stress	Mean (°C)	Standard Deviation	Minimum (°C)	Median (°C)	Maximum (°C)
1	Light Green	25.3	1.3	22.0	25.0	27.0
2	Light Red	24.3	1.1	22.0	24.0	26.0
3	Light Blue	25.3	1.3	22.0	25.0	27.0
4	Dark	24.7	1.1	22.0	25.0	27.0
5	Light White	24.6	1.1	22.0	24.0	27.0

Source: Own authorship (2024).

In Table 6, the correlation between temperature sensors is shown, providing insights into the relationships between the measurements. A variety of significant correlations between temperature sensors are observed, with a correlation of 0.91 between Sensor 1 and Sensor 5, indicating a consistent relationship in temperature measurements. On the other hand, Table 6 shows the correlation of 0.97 between Sensor 1 and Sensor 3 reflects a similar temperature scenario for the experiment.

Table 6 – Correlation between temperature sensors

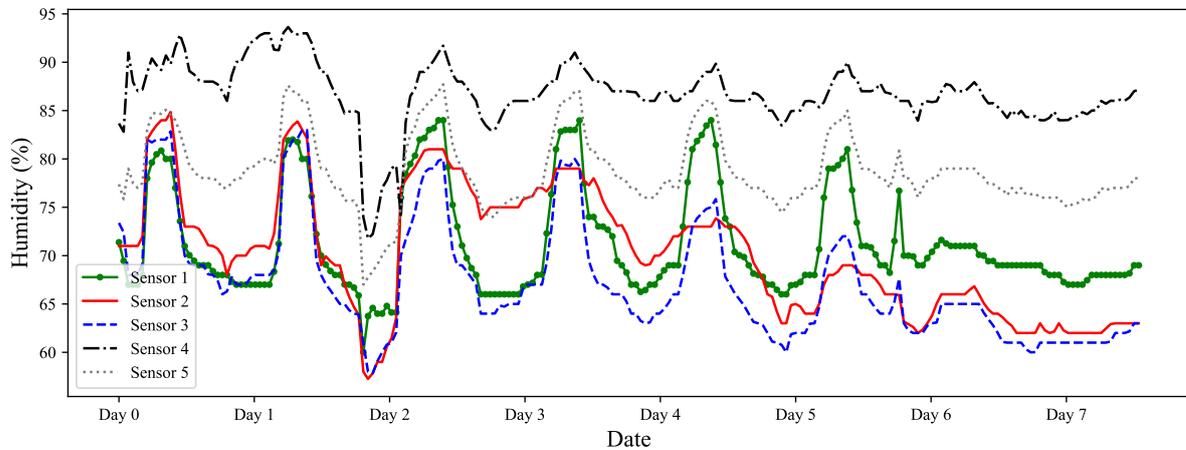
	Sensor 1	Sensor 2	Sensor 3	Sensor 4	Sensor 5
Sensor 1	1.00	0.93	0.97	0.69	0.91
Sensor 2	0.93	1.00	0.93	0.79	0.92
Sensor 3	0.97	0.93	1.00	0.67	0.92
Sensor 4	0.69	0.79	0.67	1.00	0.88
Sensor 5	0.91	0.92	0.92	0.88	1.00

Source: Own authorship (2024).

4.1.2 Result of Humidity Conditions in Stress Environment

In Figure 29, the data collected by five humidity sensors (DHT11), each in a container with light stress, is presented. Each sensor recorded a total of 10,502 measurements over 7 days, with the descriptive statistical results detailed in Table 7.

Figure 29 – Graph of data collected by five Humidity sensors in containers under various lighting conditions (green, red, blue, white LEDs, and no lighting). Each sensor recorded 10,502 measurements over a period of 7 days



Source: Own authorship (2024).

Sensor Humidity 1 (Stress Light Green) recorded a mean of 71.2 with a standard deviation of 5.5. The humidity values ranged from a minimum of 60% to a maximum of 84%, with a median of 69%. This indicates moderate variability in the humidity measurements under green light conditions.

Sensor Humidity 2 (Stress Light Red) showed a mean of 70.3 with a standard deviation of 6.5. The values ranged from a minimum of 57% to a maximum of 85%, with a median of 70%. The higher standard deviation suggests greater variability in the humidity response under red light compared to the previous sensor.

For Sensor Humidity 3 (Stress Light Blue), the mean was 67.3 and the standard deviation was 6.1. The humidity values fluctuated between a minimum of 57% and a maximum of 83%, with a median of 66%. The similarity in the mean and median with Sensor Humidity 1 suggests a comparable response under blue and green light.

Sensor Humidity 4 (Stress Light Dark) reported a mean of 86.8 with a standard deviation of 3.4. The values ranged from a minimum of 70% to a maximum of 94%, with a median of 87%. This high presence of humidity is due to the absence of lighting in the container.

Finally, Sensor Humidity 5 (Stress Light White) showed a mean of 78.8 with a standard deviation of 3.9. The values ranged from a minimum of 67% to a maximum of 88%, with a median of 78%. This suggests a stable response to white light conditions, similar to red light in terms of variability.

Table 7 – Statistical results of data collected by five Humidity sensors in containers under various lighting conditions (green, red, blue, white LEDs, and no lighting). Each sensor recorded 10,502 measurements over a period of 7 days

Sensor Humidity	Container with Stress	Mean	Standard Deviation %	Minimum %	Median %	Maximum %
1	Light Green	71.2	5.5	60	69	84
2	Light Red	70.3	6.5	57	70	85
3	Light Blue	67.3	6.1	57	66	83
4	Dark	86.8	3.4	70	87	94
5	Light White	78.8	3.9	67	78	88

Source: Own authorship (2024).

In the correlation analysis between the different humidity sensors used, significant relationships are observed, indicating the presence of consistent patterns in the measurements. Table 8 presents the correlation coefficients between the sensors, where the correlation between Sensor 1 and Sensor 5 is 0.93, indicating a robust relationship between these humidity measurements. These results provide a reference for the environmental conditions in our experiment.

Table 8 – Correlation Matrix between humidity sensors

	Sensor 1	Sensor 2	Sensor 3	Sensor 4	Sensor 5
Sensor 1	1.00	0.64	0.87	0.51	0.93
Sensor 2	0.64	1.00	0.84	0.58	0.70
Sensor 3	0.87	0.84	1.00	0.63	0.88
Sensor 4	0.51	0.58	0.63	1.00	0.79
Sensor 5	0.93	0.70	0.88	0.79	1.00

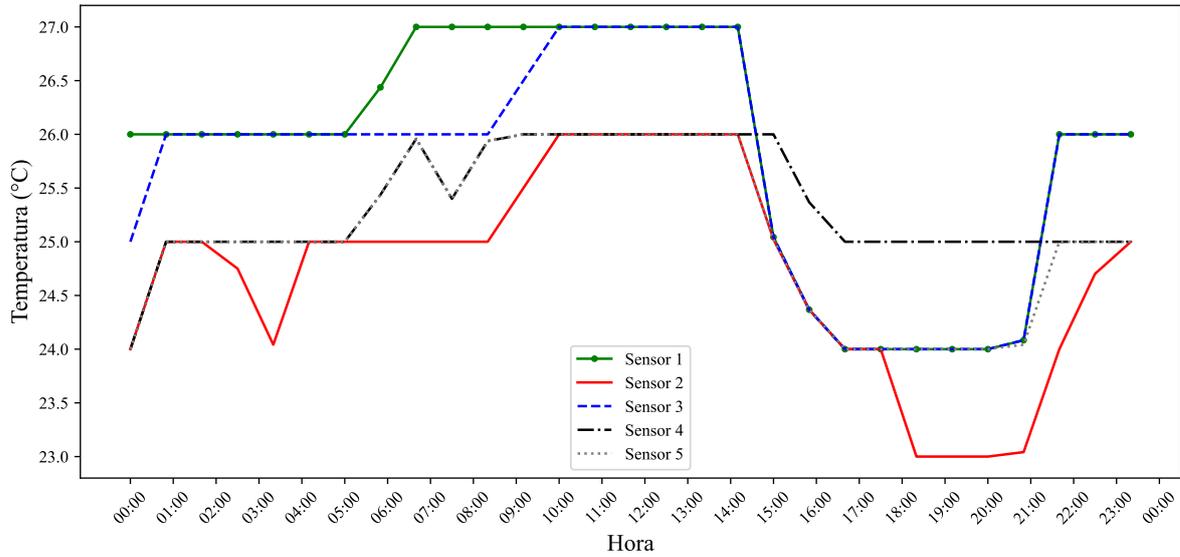
Source: Own authorship (2024).

4.1.3 Interaction Between Temperature and Humidity Under Stress Conditions

Figures 30 and 31 show the environmental conditions in our tests with DHT11 sensors in the containers, which are designed to measure both humidity and temperature simultaneously. An inverse relationship between temperature and relative humidity is observed. As the temperature increases, the air's capacity to retain moisture varies because warm air can hold more water vapor before reaching saturation. At higher temperatures, relative humidity tends to decrease. Conversely, at lower temperatures, the air has a reduced capacity to hold moisture, resulting in higher relative humidity (INGRAHAM *et al.*, 1974; ZHA *et al.*, 2017).

During the experiment, continuous readings of relative humidity and temperature were recorded at intervals of approximately 1 minute, allowing us to analyze how these variables responded to changes in internal environmental conditions. Figures 30 and 31 show the temperature and humidity in the containers over 24 hours on day 4 of the experiment.

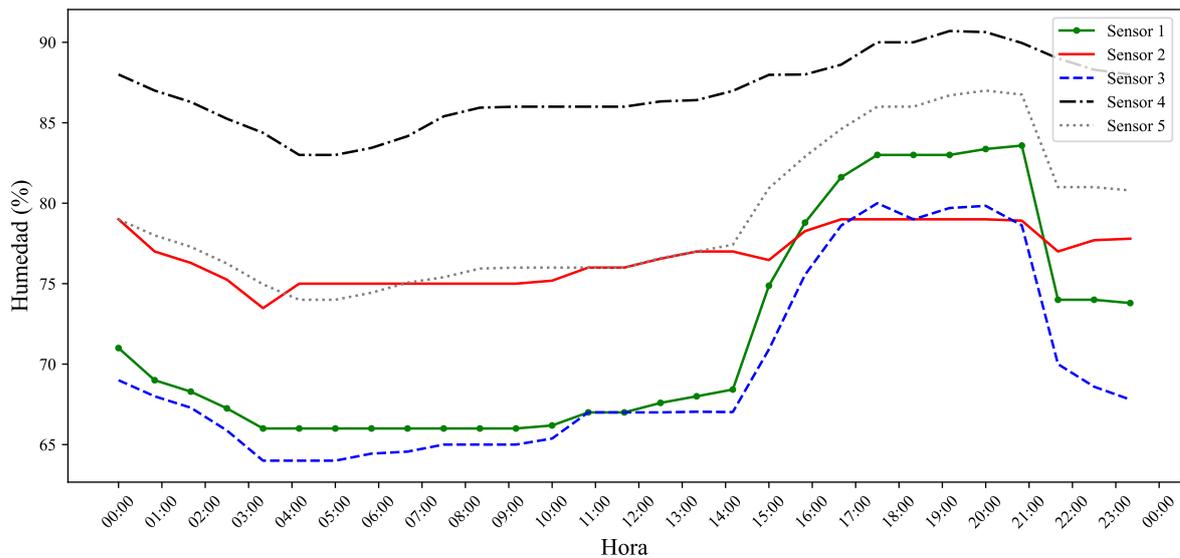
Figure 30 – Temperature on day 4, over 24 hours.



Source: Own authorship (2024).

These data reflect the consistency in humidity measurements under different stress conditions. The similarity in the minimum and maximum ranges indicates common environmental conditions, with the experiment observing that the temperature and humidity conditions were as similar as possible in each container.

Figure 31 – Humidity on day 4, over 24 hours



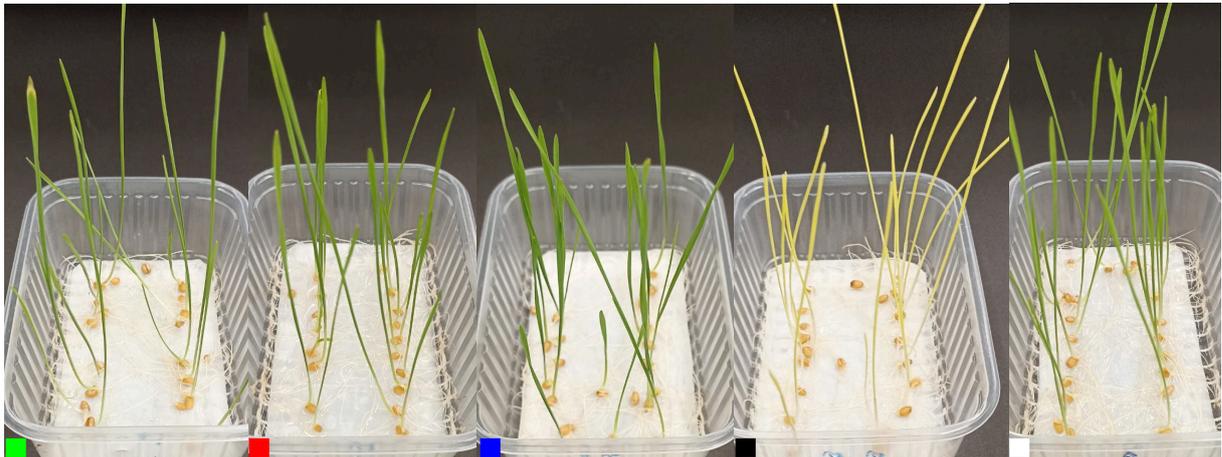
Source: Own authorship (2024).

4.2 SPECTRAL DATABASE FOR REFLECTION AND ABSORPTION

In this Section, the collection of reflection and absorption spectra from samples under different stress conditions in wheat leaves is explored. The process begins with wheat seeds soaked for 18 hours, followed by placement in plastic containers with a double-layer germination paper. After germination and growth, the leaves were exposed for 7 days to different combinations of LED light colors (green, red, blue, and white) and distilled water in volumes of (3 ml, 5 ml, and 10 ml). Figure 32 shows the samples under 5 ml water stress for different types of light stress. After cutting the fresh leaves from the containers, they were placed in a Petri dish (area of 55 cm²) with a black background (in Figure 33 and 34). In Figure 35 and 36 the corresponding diffuse reflectance spectroscopy measurement. Each set of images represents a specific combination of LED color and distilled water volume, allowing a preliminary visual assessment of the impact of these stress conditions on the health and appearance of the leaves.

Finally, an analysis of the absorption spectra of the leaves was carried out. To obtain a comparison of the results, we used a spectrophotometer (described in Subsection 3.1.3) and a spectrometer (described in Subsection 3.1.2) to measure the total chlorophyll content, and the results were compared.

Figure 32 – Containers with wheat leaves after 7 days of growth under 5 ml distilled water stress



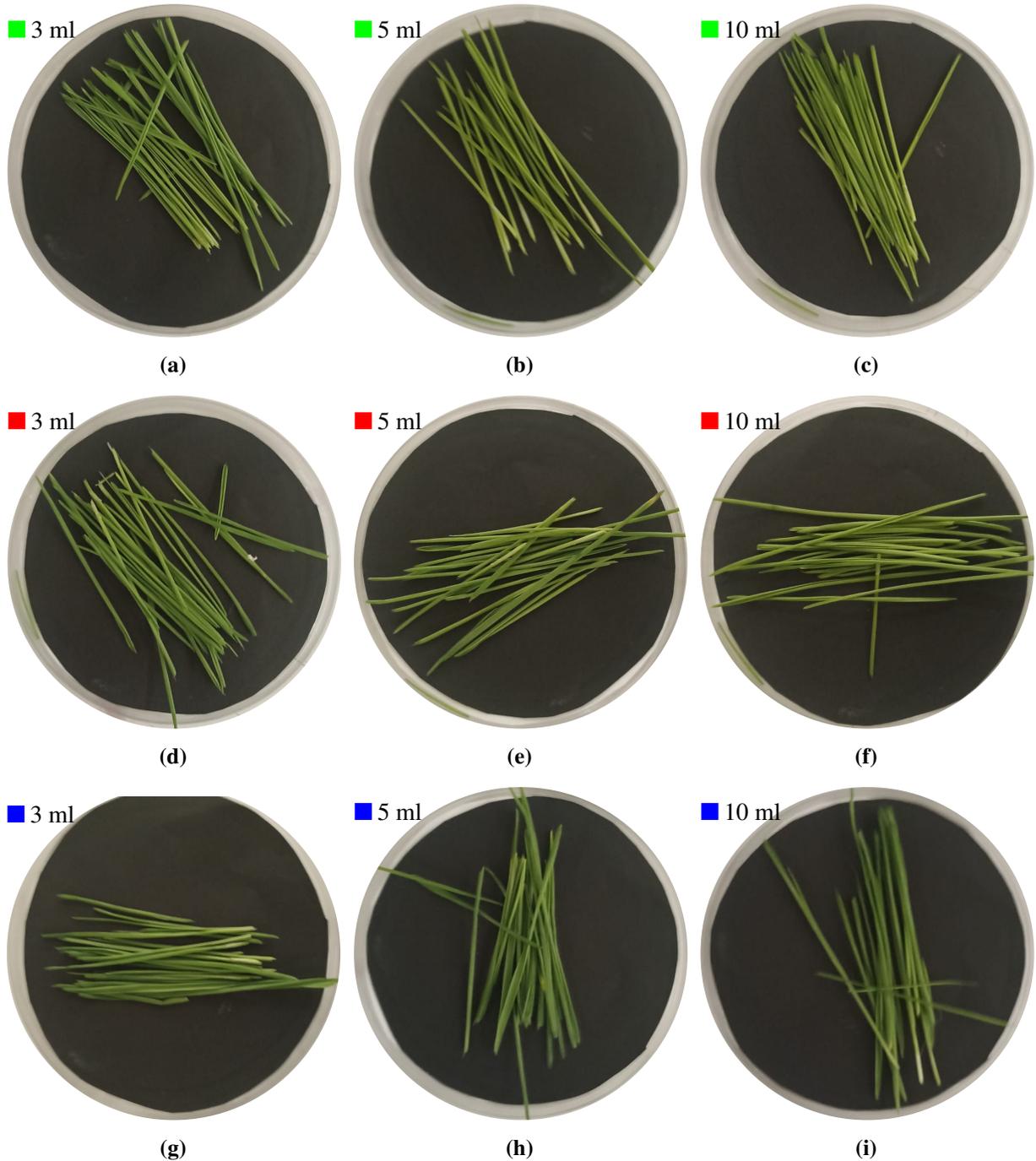
Source: Own authorship (2024).

4.2.1 Analysis of the Reflectance Spectra

To perform the reflectance spectra measurements, samples of wheat leaves were obtained, focusing only on the green region without roots. As shown in Figure 33, wheat leaf

samples are presented after 7 days of growth under 9 different stress conditions. Each subplot shows a specific combination of LED color and distilled water volume.

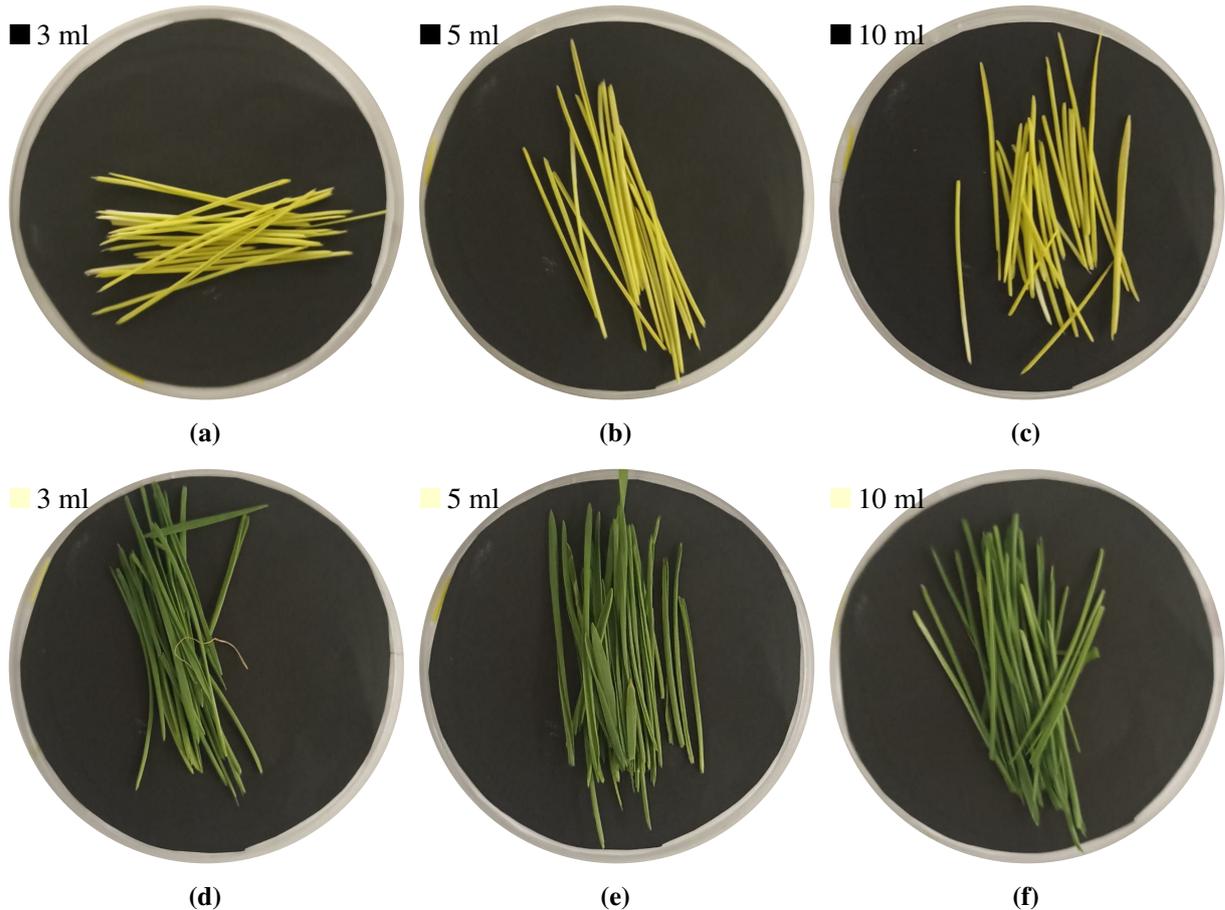
Figure 33 – Wheat leaf samples after 7 days of growth under 9 different stress conditions. (a) Green LED, 3 ml (33(a)), (b) Green LED, 5 ml (33(b)), (c) Green LED, 10 ml (33(c)), (d) Red LED, 3 ml (33(d)), (e) Red LED, 5 ml (33(e)), (f) Red LED, 10 ml (33(f)), (g) Blue LED, 3 ml (33(g)), (h) Blue LED, 5 ml (33(h)), (i) Blue LED, 10 ml (33(i))



In Figure 34, wheat leaf samples are shown after 7 days of growth under various stress conditions. The first row displays the conditions without illumination LED, with different volumes of distilled water: (a) 3 ml (Figure 34(a)), (b) 5 ml (Figure 34(b)), and (c) 10 ml

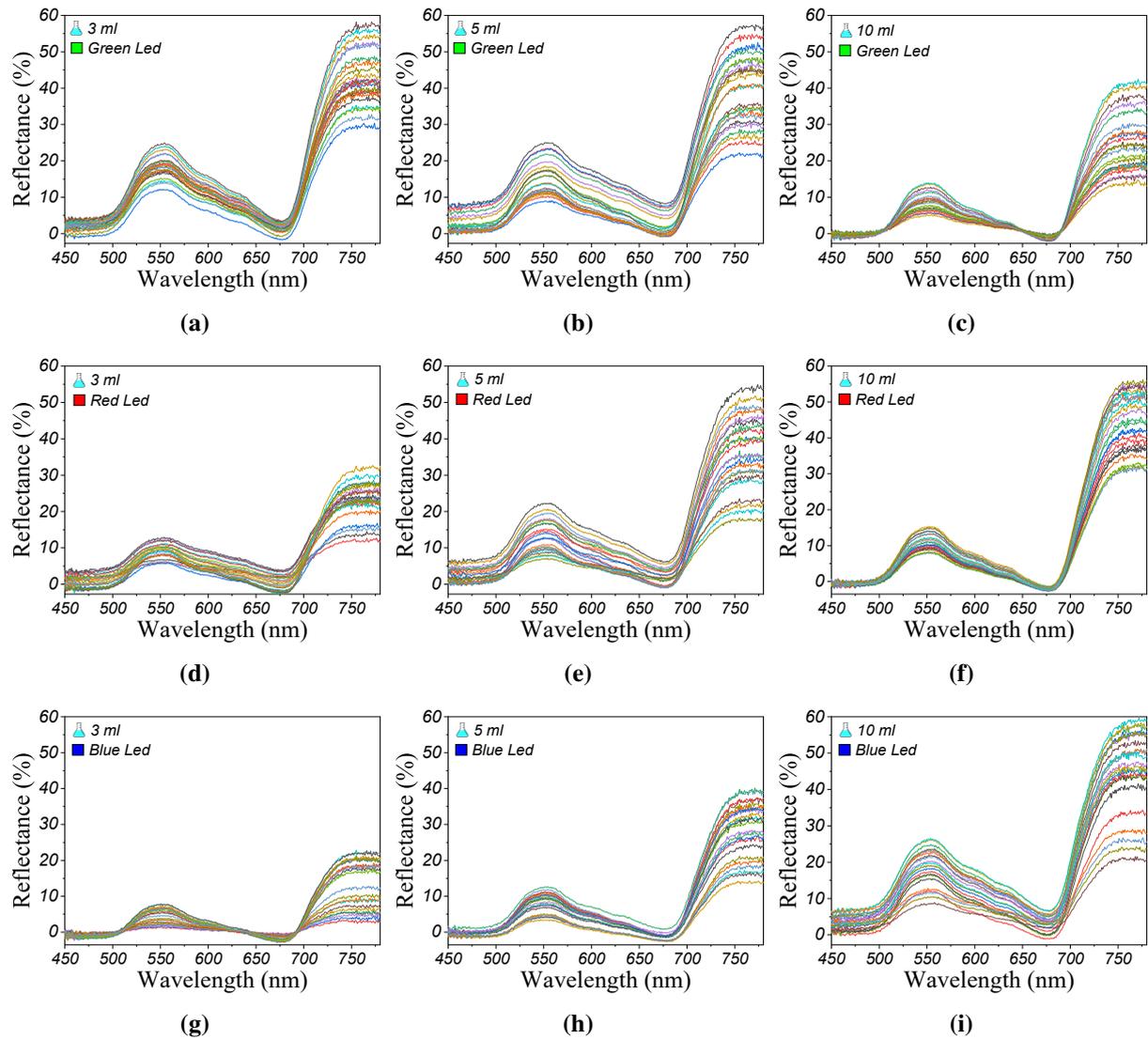
(Figure 34(c)). The second row presents the conditions with white LED, also with different volumes: (d) 3 ml (Figure 34(d)), (e) 5 ml (Figure 34(e)), and (f) 10 ml (Figure 34(f)). This organization facilitates a visual comparison of the effect of the presence or absence of LED lighting.

Figure 34 – Wheat leaf samples after 7 days of growth under various stress conditions. (a) Without LED, 3 ml (34(a)), (b) Without LED, 5 ml (34(b)), (c) Without LED, 10 ml (34(c)), (d) White LED, 3 ml (34(d)), (e) White LED, 5 ml (34(e)), (f) White LED, 10 ml (34(f))



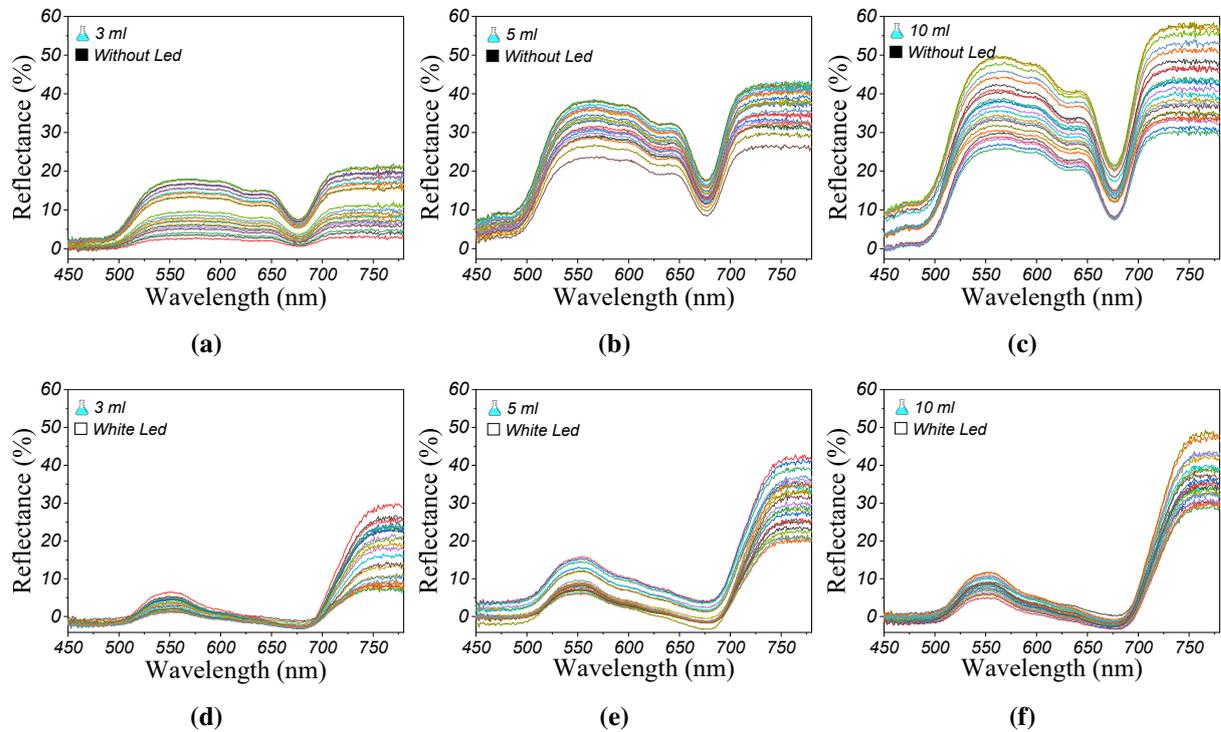
To evaluate how different stress conditions affect wheat leaves, a spectral analysis was conducted using multi-angular reflectance spectroscopy. Figure 35 presents the reflection spectra of fresh wheat leaves subjected to three LED colors (green, red, and blue) with three different volumes (3 ml, 5 ml, and 10 ml). A total of 23 diffuse reflection spectra were obtained at random angles of light incidence on the Petri dish containing the wheat leaf samples. Due to the multi-angular nature of the measurements, variations in spectral intensities were observed. These amplitude differences pose a challenge for the models, which must extract features from the complex patterns present in our samples.

Figure 35 – Multi-angular reflection spectra of fresh wheat leaves stressed under three LED colors: (a) Green LED, 3 ml (35(a)), (b) Green LED, 5 ml (35(b)), (c) Green LED, 10 ml (35(c)), (d) Red LED, 3 ml (35(d)), (e) Red LED, 5 ml (35(e)), (f) Red LED, 10 ml (35(f)), (g) Blue LED, 3 ml (35(g)), (h) Blue LED, 5 ml (35(h)), and (i) Blue LED, 10 ml (35(i))



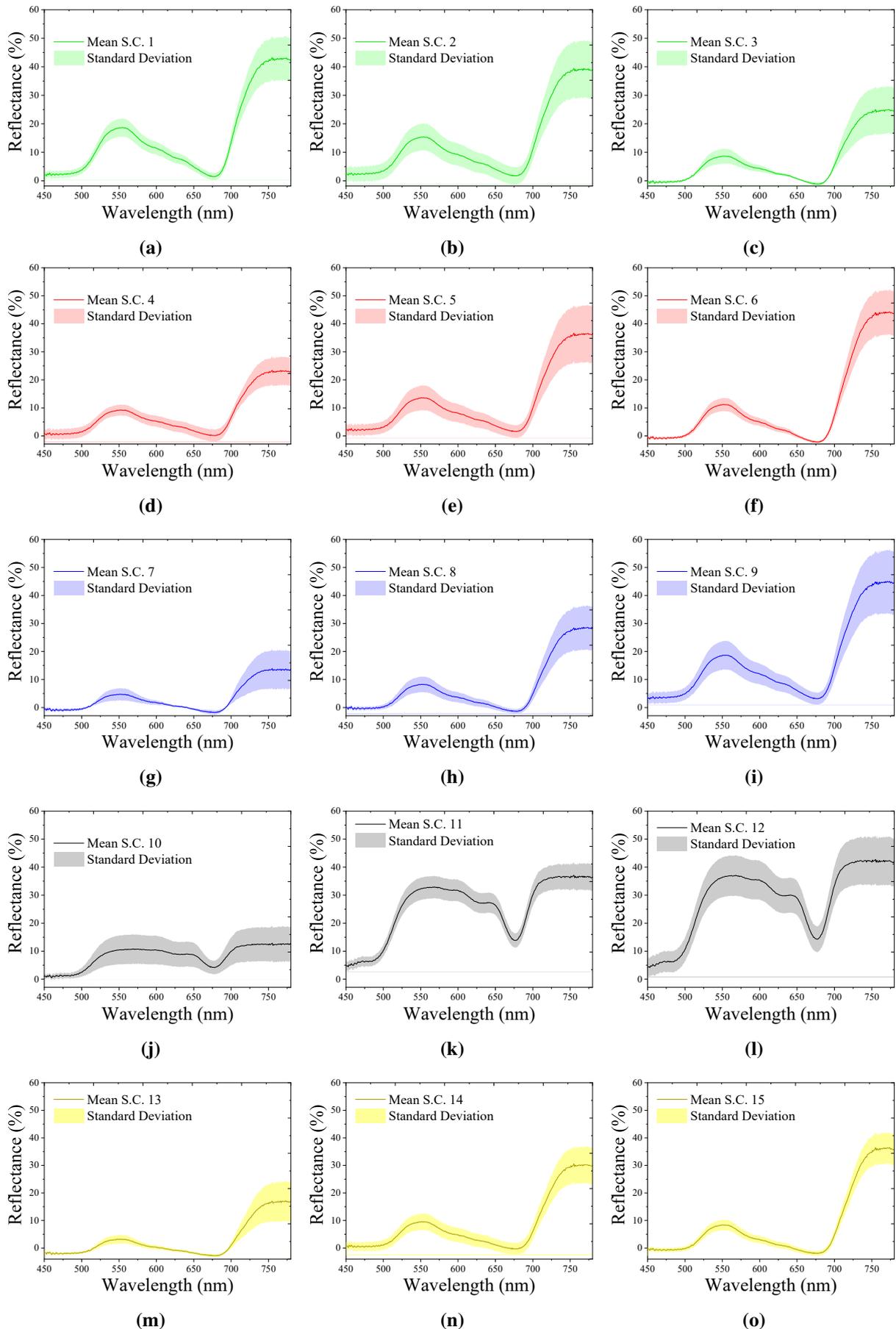
This Figure 36 presents the multi-angular reflection spectra of fresh wheat leaves subjected to different stress conditions, with and without exposure to white LED illumination, using different volumes of distilled water. A total of 23 diffuse reflection spectra were obtained at different angles of light incidence on the Petri dish containing the fresh wheat leaf samples.

Figure 36 – Multi-angular reflectance spectra of fresh wheat leaves, stressed without LED and with white LED in different volumes of distilled water. (a) No LED, 3 ml (36(a)), (b) No LED, 5 ml (36(b)), (c) No LED, 10 ml (36(c)), (d) White LED, 3 ml (36(d)), (e) White LED, 5 ml (36(e)), (f) White LED, 10 ml (36(f))



The Figure 37 shows the standard deviation as a function of wavelength for each of the 15 stress conditions applied to wheat leaf samples after 7 days of growth, and for samples taken in a multi-angular setup. Each graph presents the standard deviation along with the average of the diffuse reflectance data for each stress condition, covering a wavelength range from 450 nm to 780 nm. The Figure 37(a) to 37(o) represent these stress conditions individually, allowing for visualization of how the dispersion of the data obtained varies for the different stress conditions. The comparison between these standard deviation curves provides a general overview of how the intensities vary for each sample taken in a multi-angular setup.

Figure 37 – Standard deviation of reflectance spectra for wheat leaf samples after 7 days of growth under 15 different stress conditions (SC). Each subplot Figure 37(a) to 37(o) represents the standard deviation of reflectance measurements across wavelengths from 450 nm to 780 nm for one of the 15 stress conditions, including water and light stress

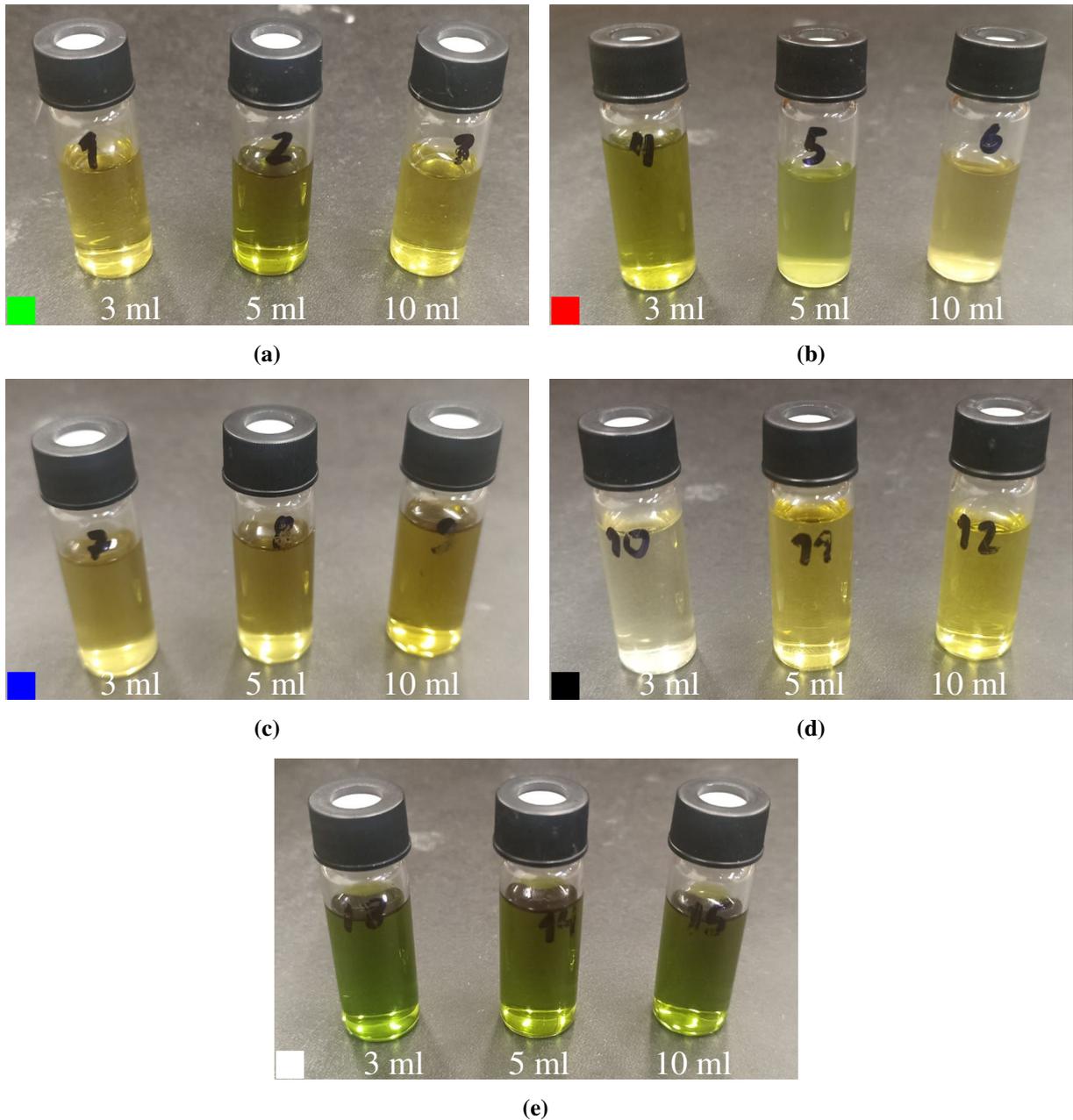


In Figure 37, the minimum standard deviation is about 0.2%, corresponding to stress condition 3, and the maximum standard deviation is about 12%, corresponding to stress condition 9.

4.2.2 Analysis of the Absorption Spectra

Following the reflection measurements described in Subsection 4.2.1 and with the extraction methodology (described in Subsection 3.3.2), 15 samples were extracted, as shown in Figure 38. Here, Figure 45(a) shows the samples for 3 ml, 5 ml, 10 ml stressed in the first container (Green LED Light) of the experiment, Figure 45(b) shows the samples for 3 ml, 5 ml, 10 ml stressed in the second container (Red LED Light), Figure 45(c) shows the samples for 3 ml, 5 ml, 10 ml stressed in the third container (Blue LED Light), Figure 45(d) shows the samples for 3 ml, 5 ml, 10 ml stressed in the fourth container (No Light), and Figure 45(e) shows the samples for 3 ml, 5 ml, 10 ml stressed in the fifth container (White LED Light). It can be highlighted that, compared to leaves with a non-distinguishable green color except for the container without light, a notable difference in green color tones can be observed in the liquid extraction. The lighter green color is notable for the samples stressed without any illumination, while the greener color is observed for the samples stressed with white light. However, this does not necessarily indicate the total amount of chlorophyll just by the green color tone in the sample, as color perception is subjective.

Figure 38 – Extracted samples showing different stress conditions across containers: (a) Green LED light for 3 ml, 5 ml, 10 ml of distilled water, (b) Red LED light for 3 ml, 5 ml, 10 ml of distilled water, (c) Blue LED light for 3 ml, 5 ml, 10 ml of distilled water, (d) No light for 3 ml, 5 ml, 10 ml of distilled water, and (e) White LED light for 3 ml, 5 ml, 10 ml of distilled water



Source: Own authorship (2024).

4.2.2.1 Results for Spectrophotometer IL-593-S

Table 9 shows the results obtained for various samples in terms of chlorophyll. The table includes the mass of each sample (in grams), the optical densities measured at 645 nm (A_{645}) and 663 nm (A_{663}) using the Spectrophotometer IL-593-S (described in Subsection 3.1.3),

and the concentrations of chlorophyll a, b, and total (using Equations (4)–(6)), expressed in milligrams per gram ($\text{mg} \cdot \text{g}^{-1}$).

Table 9 – Chlorophyll measurements obtained with the Spectrophotometer IL-593-S. The table shows the mass (in grams) and the optical densities at A_{645} and A_{663} , as well as the concentrations of chlorophyll a, b, and total (in $\text{mg} \cdot \text{g}^{-1}$) for each sample

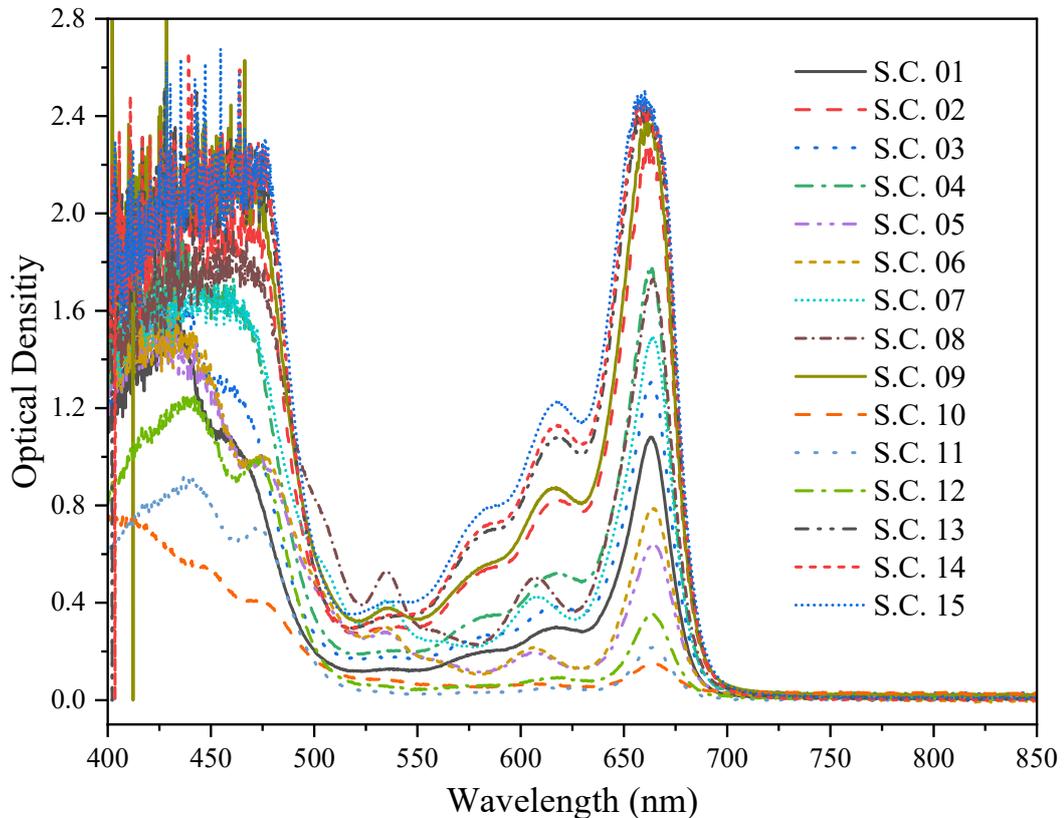
S.C.	Mass (g)	A_{645}	A_{663}	Chlorophyll a ($\text{mg} \cdot \text{g}^{-1}$)	Chlorophyll b ($\text{mg} \cdot \text{g}^{-1}$)	Total Chlorophyll ($\text{mg} \cdot \text{g}^{-1}$)
01	0.080	0.404	1.143	0.8393	0.2439	1.083
02	0.220	1.097	2.252	0.5829	0.3314	0.914
03	0.086	0.483	1.351	0.9220	0.2755	1.197
04	0.110	0.699	1.851	0.9831	0.3338	1.317
05	0.130	0.132	0.504	0.2325	0.0255	0.258
06	0.140	0.262	0.828	0.3504	0.0759	0.426
07	0.113	0.711	1.864	0.9628	0.3344	1.297
08	0.136	0.839	2.035	0.8672	0.3562	1.223
09	0.204	1.195	2.291	0.6343	0.4079	1.042
10	0.116	0.021	0.090	0.0468	0.0026	0.049
11	0.149	0.047	0.219	0.0891	0.0017	0.091
12	0.172	0.082	0.352	0.1235	0.0067	0.130
13	0.228	1.452	2.323	0.5613	0.4907	1.052
14	0.214	1.502	2.327	0.5961	0.5492	1.145
15	0.218	1.597	2.334	0.5813	0.5883	1.169

Source: Own authorship (2024).

4.2.2.2 Results for HR4000 Spectrometer

For the HR4000 spectrometer described in Subsection 3.1.2, it provides all intensities in the range of 200-1100 nm. We obtained 15 absorption spectra using the configuration for absorbance spectroscopy (Subsection 3.1.6). In Figure 39, the spectra obtained with the mentioned spectrometer are shown in the range from 400 to 850 nm, with a reference point at 750 nm (all spectra were vertically shifted with 0.00 at that reference point). In the region from 600 nm to 700 nm, which is our region of interest, we observe a notable progressive variation in intensity, yielding different intensity levels. This ensures the variability in chlorophyll quantification, allowing for adequate training in the proposed models.

Figure 39 – Absorption spectra of 15 liquid chlorophyll samples. The spectra show intensity as a function of wavelength (in nm) for each sample, representing a different stress condition (S.C. 01 to S.C. 15)



Source: Own authorship (2024).

4.2.2.3 Comparison of Total Chlorophyll Results

Table 10 presents the quantification of chlorophyll using a Digital Spectrophotometer IL-593-S and the HR4000 spectrometer. The table is organized into two main sections, each corresponding to one of the devices. For the 15 different stress conditions, OD values at the peaks of 645 nm (A_{645}) and 663 nm (A_{663}) are reported, as well as the concentrations of chlorophyll a, chlorophyll b, and total chlorophyll derived after applying the corresponding calculation described in 3.3.2. These data allow us to compare how correlated both devices are in measuring chlorophyll concentrations.

Table 11 summarizes the descriptive statistics for chlorophyll data obtained using two different instruments. For the Digital Spectrophotometer IL-593-S, the chlorophyll measurements (Chl a, Chl a b, and Chl total) show variable means and medians, with standard deviations and variances indicating moderate to high levels of dispersion in the readings. Extreme values cover a significant range, suggesting some variability in the measurements. In contrast, the HR4000

Table 10 – Comparison of chlorophyll measurements obtained with the Digital Spectrophotometer IL-593-S and the HR4000 spectrometer. OD values at A_{645} and A_{663} , and concentrations of chlorophyll a, b, and total for both devices are shown. The percentage of total chlorophyll is shown in Pct. for each device.

S.C.	Digital Spectrophotometer IL-593-S						HR4000 spectrometer					
	A_{645}	A_{663}	Chl a	Chl b	Chl T	Pct.	A_{645}	A_{663}	Chl a	Chl b	Chl T	Pct.
01	0.404	1.143	0.839	0.244	1.083	48.6%	0.481	1.072	0.770	0.375	1.145	51.4%
02	1.097	2.252	0.583	0.331	0.914	47.9%	1.269	2.241	0.569	0.422	0.991	52.1%
03	0.483	1.351	0.922	0.275	1.197	48.5%	0.575	1.278	0.854	0.418	1.271	51.5%
04	0.699	1.851	0.983	0.334	1.317	48.6%	0.823	1.753	0.911	0.484	1.395	51.4%
05	0.132	0.504	0.233	0.026	0.258	41.0%	0.229	0.626	0.282	0.089	0.371	59.0%
06	0.262	0.828	0.350	0.076	0.426	49.6%	0.287	0.789	0.330	0.103	0.433	50.4%
07	0.711	1.864	0.963	0.334	1.297	53.0%	0.701	1.473	0.744	0.405	1.149	47.0%
08	0.839	2.035	0.867	0.356	1.223	52.4%	0.821	1.707	0.716	0.398	1.113	47.6%
09	1.195	2.291	0.634	0.408	1.042	47.8%	1.363	2.347	0.641	0.496	1.136	52.2%
10	0.021	0.090	0.047	0.003	0.049	37.4%	0.044	0.125	0.063	0.018	0.082	62.6%
11	0.047	0.219	0.089	0.002	0.091	47.9%	0.061	0.214	0.086	0.013	0.099	52.1%
12	0.082	0.352	0.124	0.007	0.130	46.8%	0.114	0.347	0.119	0.029	0.148	53.2%
13	1.452	2.323	0.561	0.491	1.052	47.5%	1.678	2.388	0.566	0.598	1.163	52.5%
14	1.502	2.327	0.596	0.549	1.145	47.2%	1.749	2.420	0.608	0.671	1.279	52.8%
15	1.597	2.334	0.581	0.588	1.169	47.3%	1.845	2.440	0.597	0.707	1.304	52.7%

Source: Own authorship (2024).

shows slightly lower means for Chl a and Chl total, with standard deviations and variances reflecting moderate to low variability in the chlorophyll readings. The comparison between both instruments highlights differences in measurements and dispersion, revealing a similarity between the results from these devices.

Table 11 – Descriptive statistics of chlorophyll measurements obtained with the Digital Spectrophotometer IL-593-S and the HR4000 spectrometer. The table presents the mean, median, standard deviation, variance, maximum, and minimum for OD values (A_{645} and A_{663}) and chlorophyll concentrations a, b, and total

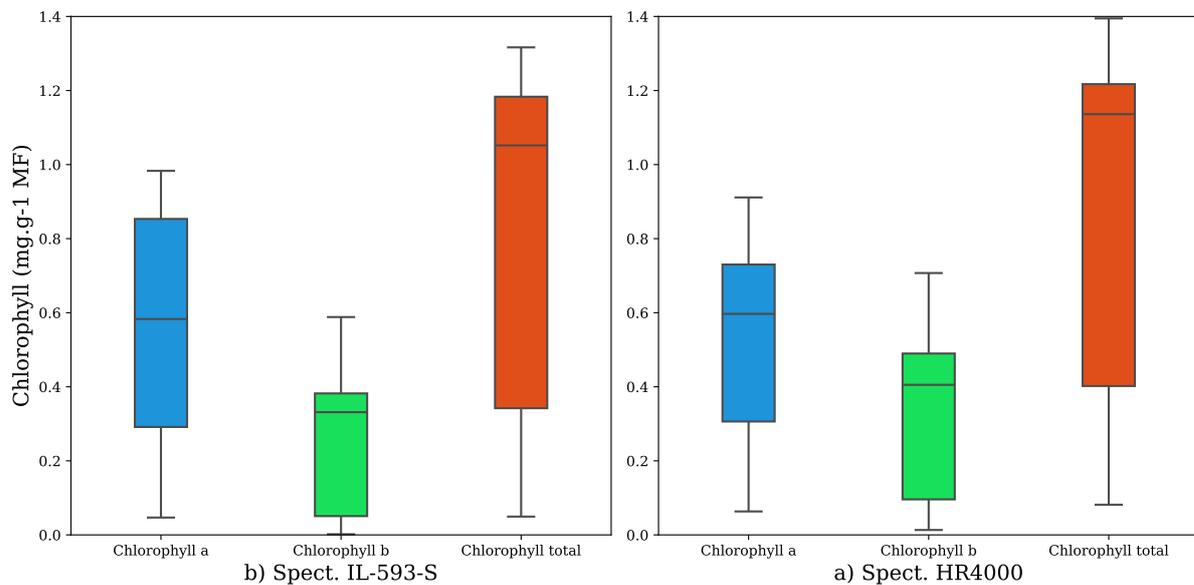
Variable	Mean	Median	Standard Deviation	Variance	Maximum	Minimum
Digital Spectrophotometer IL-593-S						
A_{645}	0.702	0.699	0.558	0.311	1.597	0.021
A_{663}	1.451	1.851	0.857	0.735	2.334	0.090
Chl a	0.558	0.583	0.325	0.106	0.983	0.047
Chl b	0.268	0.331	0.203	0.041	0.588	0.002
Chl total	0.826	1.052	0.483	0.233	1.317	0.049
HR4000 Spectrometer						
A_{645}	0.803	0.701	0.635	0.403	1.845	0.044
A_{663}	1.415	1.473	0.852	0.726	2.440	0.125
Chl a	0.524	0.597	0.280	0.078	0.911	0.063
Chl b	0.348	0.405	0.239	0.057	0.707	0.013
Chl total	0.872	1.136	0.489	0.239	1.395	0.082

Source: Own authorship (2024).

In Figure 40, with the information from Table 11, a visual comparison is provided between the distributions of chlorophyll measurements obtained with two different equipment setups. Each boxplot displays the distribution of values for Chlorophyll a, Chlorophyll b, and

Total Chlorophyll, allowing for a comparison of how different techniques affect the measurements and assessing the variations between the equipment. Using the spectrophotometer described in Subsection 3.1.3, the same sample was measured five times to calculate the standard deviation of the measurements, which was $0.001 \text{ mg} \cdot \text{g}^{-1}$ for total chlorophyll. It is important to note that, for the sake of the comparison, equivalent chlorophyll levels were obtained. .

Figure 40 – Visual comparison of the distributions of Chlorophyll a, Chlorophyll b, and Total Chlorophyll measurements obtained with two different equipment setups. Each boxplot illustrates the variation in the measurements and allows for the evaluation of the impact of different techniques on the results, according to the data presented in Table 11



Source: Own authorship (2024).

Figure 41 shows a comparison of the absorption spectra obtained using two different equipment setups. In particular, it compares the results from a spectrophotometer for spectral indices at 645 nm and 663 nm with the absorption spectra obtained from the HR4000 spectrometer. Each Figure 41(a) to 41(o) represents a specific spectrum measured under different stress conditions. Figure 41 is also organized in a grid of subfigures. This layout facilitates direct visual comparison between the spectra and spectral indices, where it can be observed that the values are highly correlated with some variations.

Figure 41 – Comparison of absorption spectra obtained using two different equipment setups: a spectrophotometer IL-593-S providing two spectral indices (645 nm and 663 nm) and an HR4000 spectrometer that provides a curve describing intensity at different wavelengths

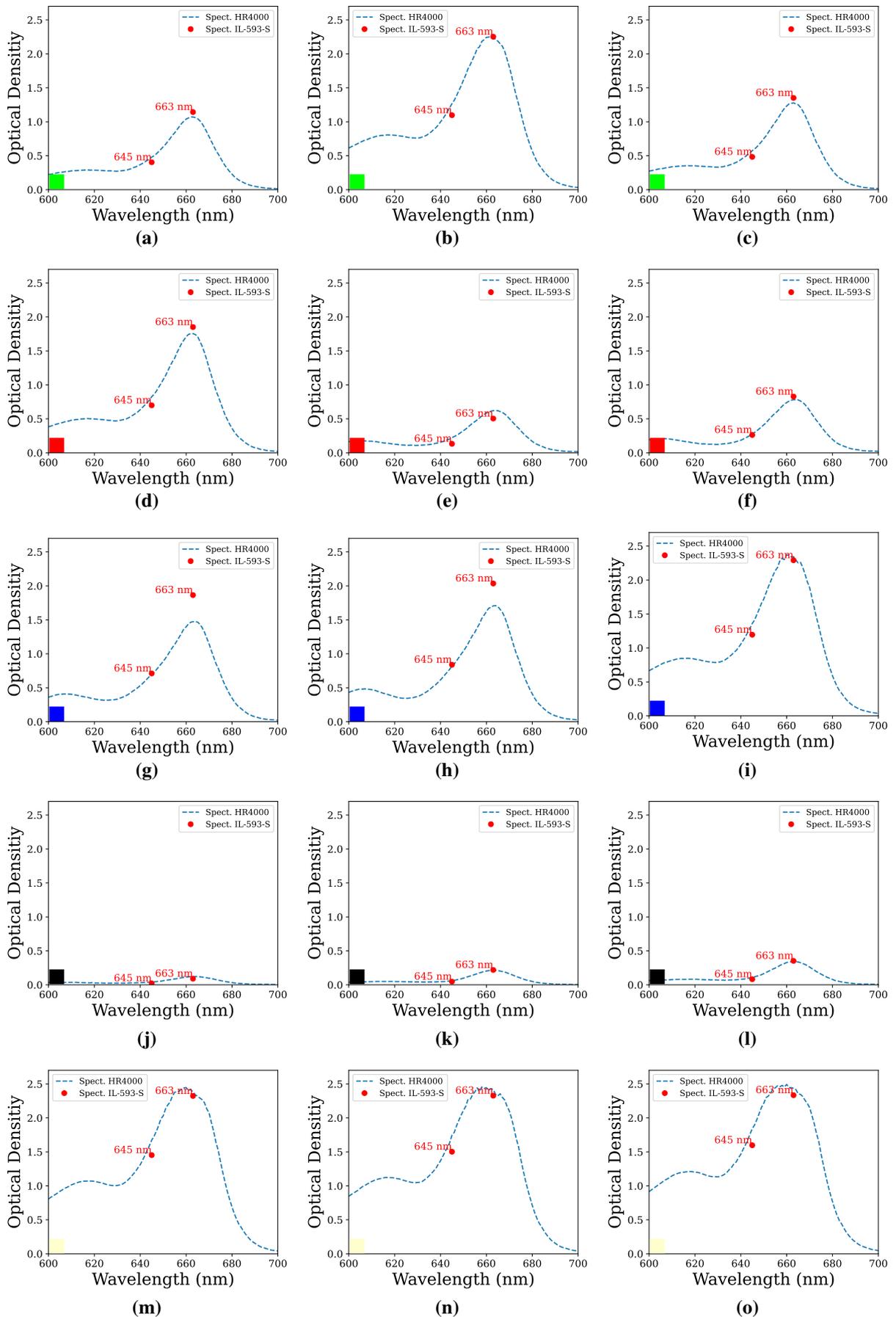


Table 12 presents the Pearson correlation coefficients for Chlorophyll a, b, and total measurements taken by two different equipment setups. The high correlation values suggest that the chlorophyll measurements obtained by the two equipment setups are highly consistent and closely related.

Table 12 – Pearson correlation between chlorophyll measurements from two different equipment setups

Measurement	Pearson Correlation
Chl a	0.9831
Chl b	0.9895
Chl total	0.9859

Source: Own authorship (2024).

The results presented in Table 11 and Figure 41 reaffirm the high consistency between the chlorophyll measurements obtained using the two equipment setups. The Pearson correlation coefficients, all above 0.98, suggest strong agreement in the readings of Chlorophyll a, b, and total, regardless of the equipment used. These results support the reliability of the collected data and the validity of subsequent analyses, as the variability in measurements, as highlighted in Table 11, also reflects different levels of total chlorophyll, a crucial factor for training and validating the artificial intelligence-based model. These varying levels adequately capture the necessary variability for a robust and precise analysis of chlorophyll concentrations in test data.

4.3 MACHINE LEARNING MODELS TO DETECT WATER AND LIGHT STRESS

In Section 4.3, we will explore the application of machine learning in the classification of water and light stress in wheat leaves. The first subsection will focus on water stress classification using AI, presenting the results obtained for different models where the preprocessing technique mainly varies. The second subsection will address light stress classification, highlighting how AI models for different preprocessing techniques can improve classification results. Both subsections will illustrate the effectiveness of these models for detecting the applied stress.

4.3.1 Water Stress Classification using AI

Table 13 and Figure 42 presents a comparison of the performance of various MiniRocket models evaluated on different datasets with preprocessing techniques (described in Section 3.4). Model 26, using SNV and Continuum Removed preprocessing, achieved perfect accuracy with a runtime of 5.96 seconds. Similarly, Model 9, with Continuum Removed and MSC preprocessing,

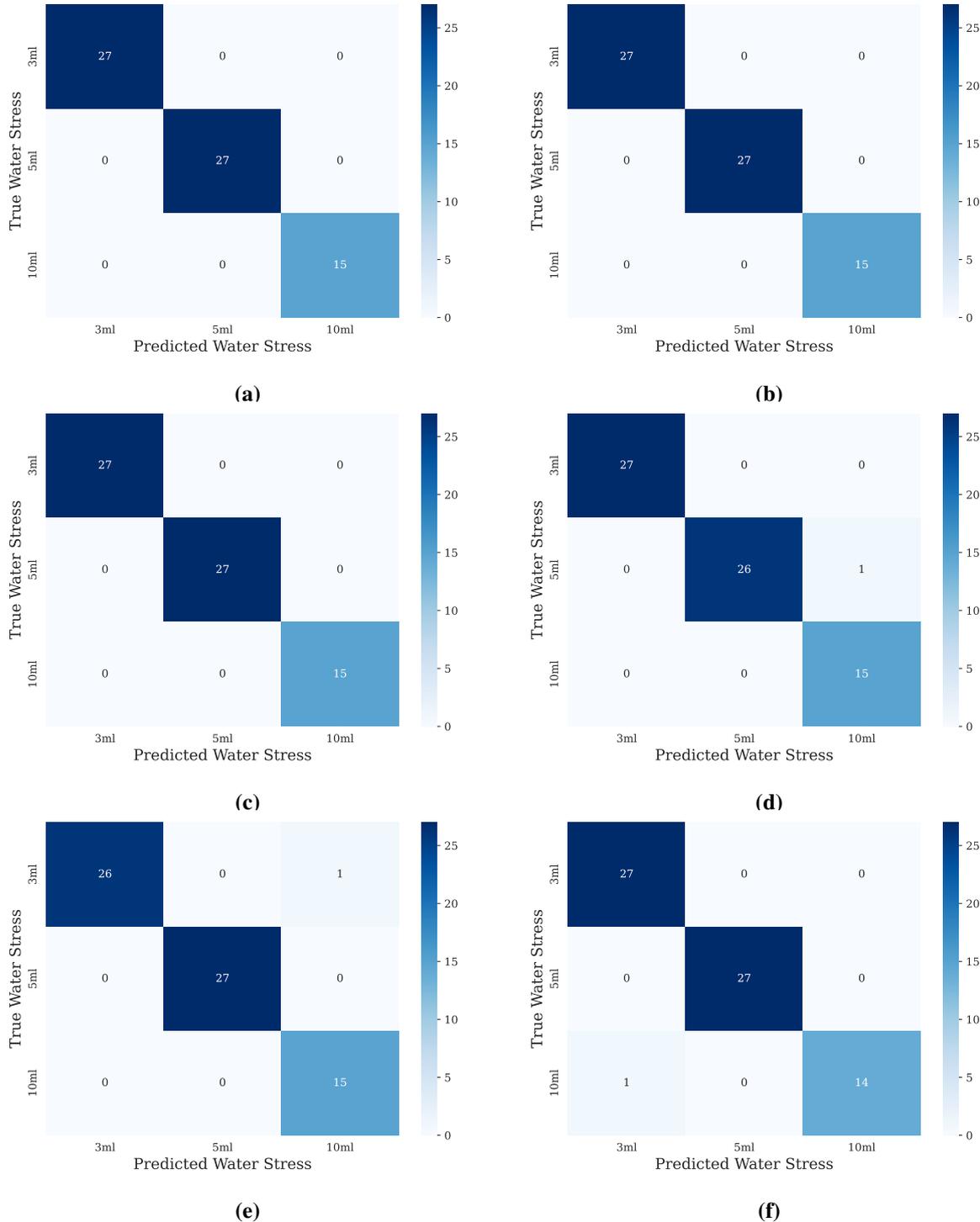
also achieved perfect accuracy but with a faster runtime of 3.10 seconds, Model 10 achieved a perfect accuracy. Other models, such as Models 01, 20, 4, 32, 11 and 23 demonstrated high accuracies of 0.986.

Table 13 – Performance metrics of the MiniRocket models for classification tasks. The table presents the accuracy and processing time for various models using different preprocessing techniques on two datasets (Data1 and Data2). The accuracy values indicate the performance of each model for the test dataset

Model	Data1	Data2	Accuracy (Test Dataset)	Training Time (s)
Model 26	SNV	Continuum Removed	1.000	5.963
Model 09	Continuum Removed	MSC	1.000	3.104
Model 10	Continuum Removed	SNV	1.000	3.996
Model 01	-	Continuum Removed	0.986	100.139
Model 20	MSC	Continuum Removed	0.986	3.014
Model 04	-	SNV	0.986	2.263
Model 32	Nor. Area	Continuum Removed	0.986	3.213
Model 11	Continuum Removed	Nor. Area	0.986	3.073
Model 23	MSC	Nor. Area	0.986	3.888
Model 35	Nor. Area	SNV	0.971	3.101
Model 34	Nor. Area	MSC	0.971	3.075
Model 29	SNV	Nor. Area	0.971	5.221
Model 22	MSC	SNV	0.971	3.080
Model 12	Continuum Removed	Smoothed	0.971	5.772
Model 42	Smoothed	Nor. Area	0.957	4.270
Model 07	Continuum Removed	-	0.957	3.550
Model 41	Smoothed	SNV	0.957	3.198
Model 17	Detrended	Nor. Area	0.957	3.038
Model 38	Smoothed	Continuum Removed	0.957	3.060
Model 37	Smoothed	-	0.957	2.207
Model 36	Nor. Area	Smoothed	0.957	5.370
Model 24	MSC	Smoothed	0.957	6.157
Model 33	Nor. Area	Detrended	0.957	4.847
Model 06	-	Smoothed	0.957	2.676
Model 40	Smoothed	MSC	0.942	3.243
Model 03	-	MSC	0.942	4.770
Model 30	SNV	Smoothed	0.942	4.119
Model 14	Detrended	Continuum Removed	0.942	3.049
Model 25	SNV	-	0.942	3.241
Model 28	SNV	MSC	0.942	4.689
Model 31	Nor. Area	-	0.928	2.276
Model 08	Continuum Removed	Detrended	0.928	3.093
Model 05	-	Nor. Area	0.928	2.300
Model 39	Smoothed	Detrended	0.928	4.553
Model 18	Detrended	Smoothed	0.928	5.976
Model 19	MSC	-	0.913	2.265
Model 16	Detrended	SNV	0.913	3.107
Model 27	SNV	Detrended	0.884	4.384
Model 15	Detrended	MSC	0.870	4.450
Model 21	MSC	Detrended	0.870	4.757
Model 13	Detrended	-	0.826	2.262
Model 02	-	Detrended	0.812	2.219

Source: Own authorship (2024).

Figure 42 – Confusion matrices showing the performance of different models for water stress classification as presented in Table 13 (for the test dataset). The images are confusion matrices corresponding to the following models: (a) Model 26, (b) Model 9, (c) Model 10, (d) Model 1, (e) Model 20, and (f) Model 4. Each subimage visualizes the model’s performance in classifying different levels of water stress, as detailed in the associated tables (Table 14, Table 15, Table 16, Table 17)



Source: Own authorship (2024).

The performance metrics on the test data for the best models presented in Table 13, as seen in the tables (Table 14, Table 15, Table 16, and Table 17). Models 26, 9, and 10 exhibit

exceptional performance, with precision, recall, and F1 scores all reaching a perfect score of 1.00 in the 3 ml, 5 ml, and 10 ml classes. This results in an overall accuracy of 1.00, with consistently high macro and weighted averages also at 1.00, based on a total support of 69 instances. Model 11 shows equally high performance, achieving an accuracy of 0.99, with slightly lower precision scores for the 10 ml class, but maintaining solid macro and weighted averages (see Section 3.6). Models 20 and 4 also demonstrate robust performance, consistently achieving high precision and recall across all classes, with an overall accuracy of 0.99.

Table 14 – Metrics for test data: Model 26, 9 and 10

Class	Precision	Recall	F1-score	Support
3 ml	1.00	1.00	1.00	27
5 ml	1.00	1.00	1.00	27
10 ml	1.00	1.00	1.00	15
Accuracy		1.00		
Macro Avg	1.00	1.00	1.00	69
Weighted Avg	1.00	1.00	1.00	69

Table 15 – Metrics for test data: Model 1

Class	Precision	Recall	F1-score	Support
3 ml	1.00	0.96	0.98	27
5 ml	1.00	1.00	1.00	27
10 ml	0.94	1.00	0.97	15
Accuracy		0.99		69
Macro Avg	0.98	0.99	0.98	69
Weighted Avg	0.99	0.99	0.99	69

Table 16 – Metrics for test data: Model 20

Class	Precision	Recall	F1-score	Support
3 ml	1.00	0.96	0.98	27
5 ml	1.00	1.00	1.00	27
10 ml	0.94	1.00	0.97	15
Accuracy		0.99		69
Macro Avg	0.98	0.99	0.98	69
Weighted Avg	0.99	0.99	0.99	69

Table 17 – Metrics for test data: Model 4

Class	Precision	Recall	F1-score	Support
3 ml	0.96	1.00	0.98	27
5 ml	1.00	1.00	1.00	27
10 ml	1.00	0.93	0.97	15
Accuracy		0.99		69
Macro Avg	0.99	0.98	0.98	69
Weighted Avg	0.99	0.99	0.99	69

According to the results, Model 26 using SNV and Continuum Removed demonstrated the best performance in the tests, achieving an accuracy of 1.000 on the test dataset, with a

training time of 5.963 seconds. This model stood out as the best among the best in terms of precision and training efficiency. The use of MiniRocket for classification tasks related to water stress (3ml, 5ml, 10ml) in leaves through reflection spectroscopy provided an ideal combination of high accuracy and reasonable training time.

4.3.2 Light Stress Classification using AI

The results of the models presented in Table 18 indicate a significant performance using different preprocessing techniques. The models achieved accuracies ranging from 0.841 to 0.986, highlighting the effectiveness of these techniques in spectral (multiangular) analysis tasks for classifying light stress in wheat leaves.

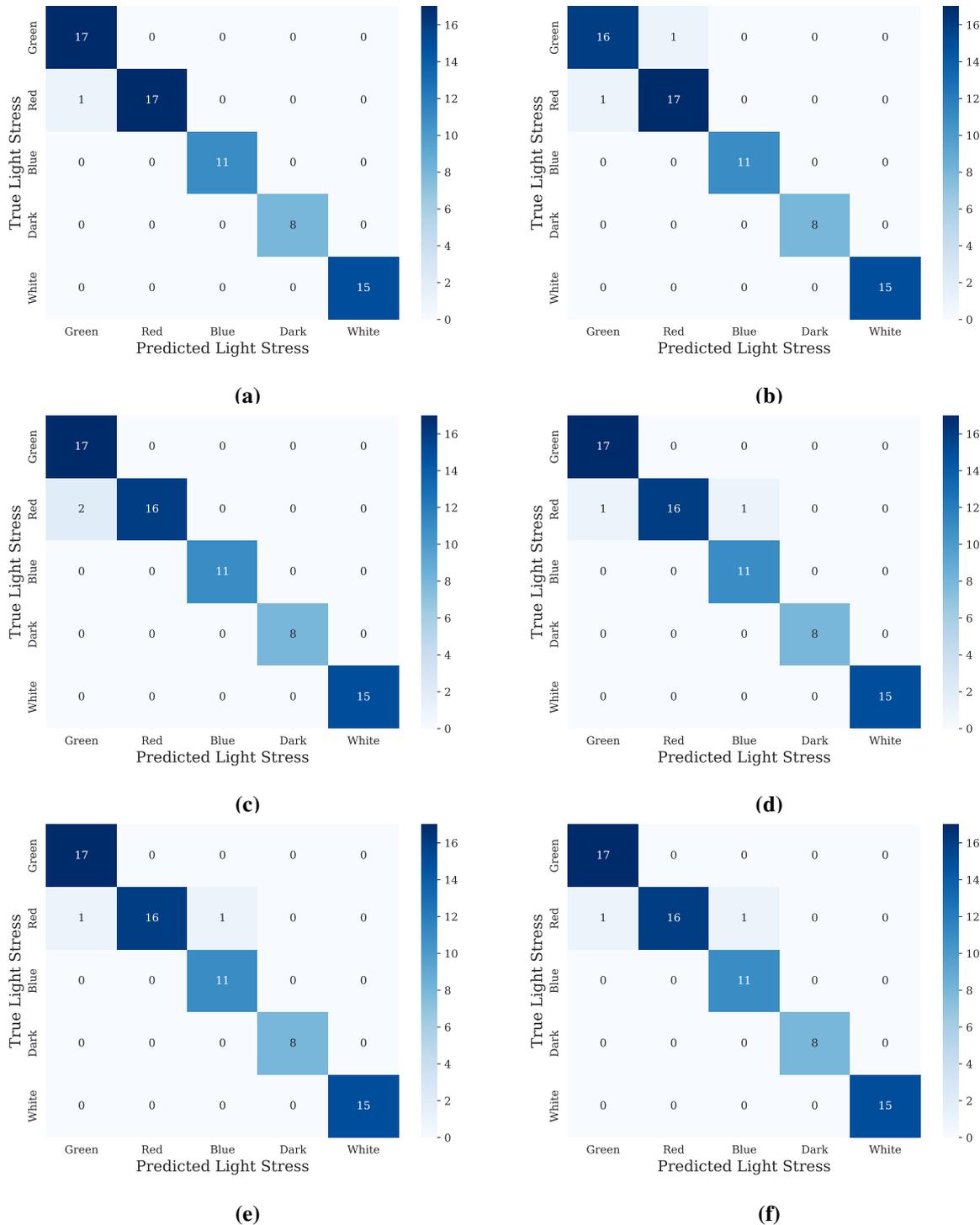
Table 18 – Comparison of Models for Light Stress Classification Based on Different Data Preprocessing Techniques, Accuracy (for the test dataset), and Execution Time

Model	Data1	Data2	Accuracy (Test Dataset)	Training Time (s)
Model 01	-	Continuum Removed	0.986	3.603
Model 08	Continuum Removed	Detrended	0.971	5.943
Model 29	SNV	Nor. Area	0.971	4.436
Model 14	Detrended	Continuum Removed	0.971	4.340
Model 12	Continuum Removed	Smoothed	0.971	3.085
Model 18	Detrended	Smoothed	0.971	5.408
Model 35	Nor. Area	SNV	0.971	3.208
Model 24	MSC	Smoothed	0.971	2.997
Model 36	Nor. Area	Smoothed	0.957	4.624
Model 37	Smoothed	-	0.957	3.595
Model 33	Nor. Area	Detrended	0.957	6.508
Model 32	Nor. Area	Continuum Removed	0.957	4.697
Model 38	Smoothed	Continuum Removed	0.957	5.416
Model 30	SNV	Smoothed	0.957	6.717
Model 39	Smoothed	Detrended	0.957	4.561
Model 40	Smoothed	MSC	0.957	5.318
Model 26	SNV	Continuum Removed	0.957	2.895
Model 42	Smoothed	Nor. Area	0.957	5.027
Model 07	Continuum Removed	-	0.957	4.323
Model 41	Smoothed	SNV	0.942	3.993
Model 17	Detrended	Nor. Area	0.942	4.652
Model 06	-	Smoothed	0.942	4.265
Model 28	SNV	MSC	0.942	7.153
Model 34	Nor. Area	MSC	0.942	2.997
Model 20	MSC	Continuum Removed	0.942	5.732
Model 31	Nor. Area	-	0.928	6.249
Model 23	MSC	Nor. Area	0.928	4.164
Model 09	Continuum Removed	MSC	0.928	3.006
Model 10	Continuum Removed	SNV	0.928	4.901
Model 11	Continuum Removed	Nor. Area	0.913	4.216
Model 03	-	MSC	0.913	2.951
Model 19	MSC	-	0.913	3.418
Model 15	Detrended	MSC	0.913	4.658
Model 27	SNV	Detrended	0.899	2.850
Model 16	Detrended	SNV	0.899	5.826
Model 25	SNV	-	0.899	4.897
Model 05	-	Nor. Area	0.899	4.168
Model 21	MSC	Detrended	0.899	4.665
Model 04	-	SNV	0.884	2.703
Model 22	MSC	SNV	0.870	3.246
Model 13	Detrended	-	0.841	5.323
Model 02	-	Detrended	0.841	4.480

Source: Own authorship (2024).

Figure 43 shows the confusion matrices for test data with the models that achieved the best performance in light stress classification.

Figure 43 – Confusion matrices showing the performance of different models for light stress classification for the test dataset. (a) Model 01: with an accuracy of 0.986, (b) Model 08: with an accuracy of 0.971, (c) Model 29: with an accuracy of 0.971, (d) Model 14: with an accuracy of 0.971, (e) Model 12: with an accuracy of 0.971, (f) Model 18: with an accuracy of 0.971



Source: Own authorship (2024).

Additional analysis with test metrics (Table 19, Table 20, and Table 21) reveals high and consistent performance across all classes, with most models achieving precision and recall scores above 90%, highlighting their reliability in classifying light stress levels in the test dataset.

Table 19 – Metrics for test data 1

Class	Precision	Recall	F1-score	Support
Green LED	0.94	1.00	0.97	17
Red LED	1.00	0.94	0.97	18
Blue LED	1.00	1.00	1.00	11
Dark	1.00	1.00	1.00	8
White LED	1.00	1.00	1.00	15
Accuracy		0.99		69
Macro Avg	0.99	0.99	0.99	69
Weighted Avg	0.99	0.99	0.99	69

Table 20 – Metrics for test data 8

Class	Precision	Recall	F1-score	Support
Green LED	0.94	0.94	0.94	17
Red LED	0.94	0.94	0.94	18
Blue LED	1.00	1.00	1.00	11
Dark	1.00	1.00	1.00	8
White LED	1.00	1.00	1.00	15
Accuracy		0.97		69
Macro Avg	0.98	0.98	0.98	69
Weighted Avg	0.97	0.97	0.97	69

Table 21 – Metrics for test data 29

Class	Precision	Recall	F1-score	Support
Green LED	0.94	1.00	0.97	17
Red LED	1.00	0.89	0.94	18
Blue LED	0.92	1.00	0.96	11
Dark	1.00	1.00	1.00	8
White LED	1.00	1.00	1.00	15
Accuracy		0.97		69
Macro Avg	0.97	0.98	0.97	69
Weighted Avg	0.97	0.97	0.97	69

The comparison of various models for light stress classification, using different data preprocessing techniques, indicates that Model 01, with the Continuum Removed technique, excels with an accuracy of 0.986 on the test dataset and a training time of 3.603 seconds. This model stands out as one of the best in terms of the combination of high accuracy and efficiency in training time.

These results for water stress classification, with an accuracy of 1.00, and light stress classification, with an accuracy of 0.986, highlight the model's effectiveness. However, the lower results reflect the importance of using preprocessing techniques and model selection in the analysis of diffuse reflectance spectra, demonstrating their impact on classification accuracy and robustness.

4.4 MACHINE LEARNING MODELS TO ESTIMATE THE AMOUNT OF TOTAL CHLOROPHYLL

In this section, the results are presented using two approaches for estimating the total chlorophyll content in wheat leaves: the MiniRocket model and a deep neural network (DNN) model. The first approach uses MiniRocket, an efficient method for analysis that highlights the fast training time of the models. The second approach relies on deep neural networks (DNN), a technique capable of learning complex nonlinear features from the data. This approach shows higher precision in estimating the data in the test set but also presents longer training times.

4.4.1 MiniRocket-Based Model for Total Chlorophyll Estimation in Wheat Leaves

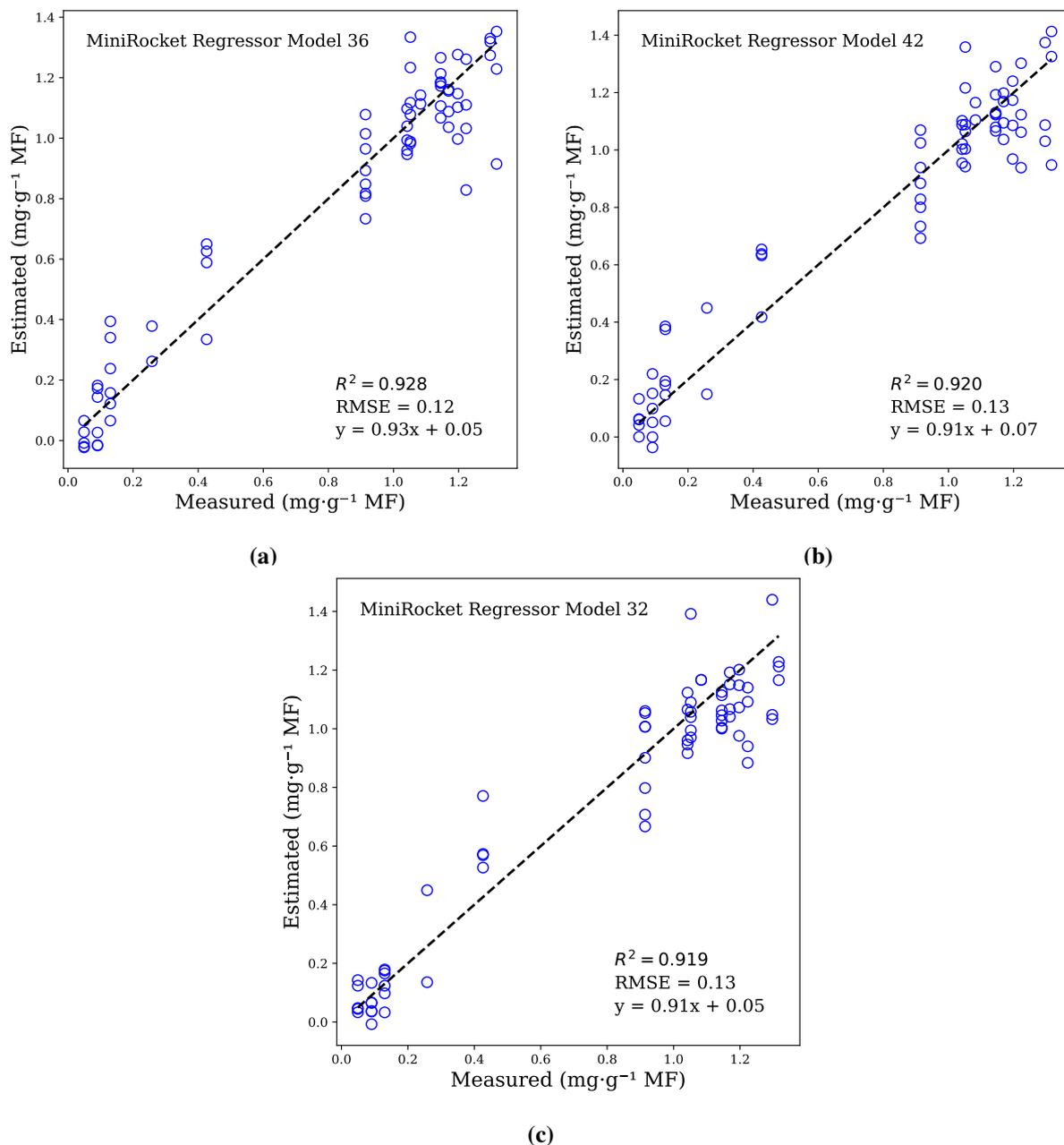
Table 22 presents the performance metrics for various MiniRocket regression models applied to different datasets. The table summarizes the coefficient of determination (R^2), the root mean square error (RMSE), and the mean absolute error (MAE) for both test and training sets, along with the training time for each model. The models were evaluated using different combinations of data preprocessing techniques described in Section 3.4, allowing an assessment of their predictive capability and computational efficiency.

Figure 44 shows the visual comparison between the predicted and measured total chlorophyll content for several MiniRocket regression models in the test set. Each subplot illustrates the relationship between predicted and measured values through a regression plot.

Table 22 – Performance metrics of various MiniRocket regression models on different datasets. The table shows R^2 , root mean squared error (RMSE), and mean absolute error (MAE) for both the test and train sets, along with the time taken for each model. Models are evaluated using different combinations of data preprocessing techniques and feature sets

Model	Data1	Data2	Result Test Set			Result Train Set			Training Time (s)
			R^2	RMSE	MAE	R^2	RMSE	MAE	
Model 36	Nor. Area	Smoothed	0.928	0.125	0.094	0.999	0.013	0.009	3.392
Model 42	Smoothed	Nor. Area	0.920	0.132	0.100	1.000	0.000	0.000	3.603
Model 32	Nor. Area	CR	0.919	0.132	0.103	0.992	0.041	0.028	3.364
Model 37	Smoothed	-	0.917	0.134	0.103	0.998	0.023	0.016	3.797
Model 12	CR	Smoothed	0.912	0.138	0.098	1.000	0.011	0.007	4.984
Model 18	Detrended	Smoothed	0.911	0.139	0.109	0.998	0.020	0.013	3.809
Model 38	Smoothed	CR	0.911	0.139	0.104	1.000	0.008	0.006	3.489
Model 24	msc	Smoothed	0.907	0.141	0.109	0.981	0.065	0.045	3.582
Model 05	-	Nor. Area	0.906	0.142	0.108	0.984	0.060	0.041	2.667
Model 40	Smoothed	MSC	0.905	0.143	0.110	0.975	0.074	0.051	3.621
Model 30	SNV	Smoothed	0.904	0.144	0.109	1.000	0.010	0.007	3.424
Model 11	CR	Nor. Area	0.903	0.145	0.113	0.994	0.036	0.024	3.534
Model 29	SNV	Nor. Area	0.899	0.147	0.114	0.995	0.035	0.024	5.648
Model 26	SNV	CR	0.896	0.150	0.117	0.994	0.036	0.024	3.402
Model 28	SNV	MSC	0.894	0.151	0.116	0.992	0.043	0.029	3.591
Model 10	CR	SNV	0.893	0.152	0.122	1.000	0.005	0.003	5.140
Model 19	MSC	-	0.891	0.153	0.120	0.993	0.038	0.025	2.750
Model 04	-	SNV	0.891	0.154	0.117	0.994	0.035	0.023	4.283
Model 22	MSC	SNV	0.890	0.154	0.119	0.993	0.040	0.027	3.359
Model 09	CR	MSC	0.889	0.155	0.123	0.995	0.034	0.023	4.346
Model 35	Nor. Area	SNV	0.888	0.155	0.121	0.994	0.038	0.026	5.806
Model 01	-	CR	0.888	0.156	0.120	0.996	0.031	0.021	3.576
Model 21	MSC	Detrended	0.887	0.156	0.118	0.986	0.055	0.037	5.094
Model 31	Nor. Area	-	0.887	0.156	0.127	1.000	0.009	0.006	4.533
Model 20	MSC	CR	0.885	0.158	0.123	0.995	0.032	0.022	3.387
Model 03	-	MSC	0.884	0.158	0.125	1.000	0.005	0.003	2.562
Model 06	-	Smoothed	0.883	0.159	0.116	0.997	0.024	0.016	5.066
Model 07	CR	-	0.883	0.159	0.124	0.996	0.031	0.021	2.844
Model 25	SNV	-	0.882	0.160	0.119	0.995	0.034	0.023	3.987
Model 34	Nor. Area	MSC	0.879	0.162	0.127	0.993	0.041	0.028	3.772
Model 33	Nor. Area	Detrended	0.877	0.163	0.131	1.000	0.000	0.000	6.058
Model 27	SNV	Detrended	0.875	0.165	0.129	0.990	0.047	0.032	6.078
Model 23	MSC	Nor. Area	0.873	0.165	0.126	0.994	0.038	0.026	4.851
Model 41	Smoothed	SNV	0.872	0.166	0.123	1.000	0.009	0.006	6.382
Model 17	Detrended	Nor. Area	0.872	0.167	0.133	0.986	0.055	0.039	3.408
Model 14	Detrended	CR	0.869	0.168	0.136	0.992	0.042	0.029	3.810
Model 15	Detrended	MSC	0.862	0.173	0.134	0.983	0.060	0.041	3.588
Model 08	CR	Detrended	0.862	0.173	0.137	1.000	0.000	0.000	3.364
Model 16	Detrended	SNV	0.861	0.173	0.136	0.990	0.046	0.031	4.030
Model 13	Detrended	-	0.847	0.182	0.139	0.980	0.066	0.045	2.579
Model 39	Smoothed	Detrended	0.838	0.187	0.121	0.998	0.019	0.014	5.874
Model 02	-	Detrended	0.836	0.188	0.136	0.979	0.068	0.046	6.840

Figure 44 – Comparação do conteúdo total de clorofila previsto versus medido para diferentes modelos de regressão MiniRocket utilizando dados de teste



Source: Own authorship (2024).

Table 22 and Figure 44 present the performance metrics of various MiniRocket regression models, with Model 36 standing out as the best. This model, which utilizes data normalized by area and smoothed (same dataset with different preprocessing concatenated), achieved a notable R^2 of 0.928 on the test set, with a root mean squared error (RMSE) of 0.125 and a mean absolute error (MAE) of 0.094. On the training set, Model 36 nearly achieved perfect accuracy with an R^2 of 0.999, and RMSE and MAE values of 0.013 and 0.009, respectively. Additionally, it demonstrated a relatively brief training time of 3.392 seconds, highlighting its

excellent balance of precision and efficiency.

4.4.2 Deep Neural Network-Based Model for Total Chlorophyll Estimation in Wheat Leaves

Table 23 presents the performance metrics for various DNN regression models applied to different datasets. It also shows the values of the coefficient of determination R^2 , root mean square error (RMSE), and mean absolute error (MAE) for both test and training sets. Additionally, the training time for each model is included. The models were evaluated using different combinations of data preprocessing techniques. The results indicate that models with preprocessing techniques and combinations using preprocessed datasets show better performance, with higher R^2 values and lower errors. It is noteworthy that the training time in all cases exceeds 1 minute.

Figure 45 illustrates the performance of the DNN models for estimating chlorophyll content in leaves using test data. In this figure, the best models are highlighted. The subfigures corresponding to the top models show the predictions made compared to the actual values, providing a clear view of the accuracy and effectiveness of each model in the estimation task. The selection of the models presented in the figure encompasses a variety of preprocessing configurations, allowing for an assessment of their impact on prediction performance.

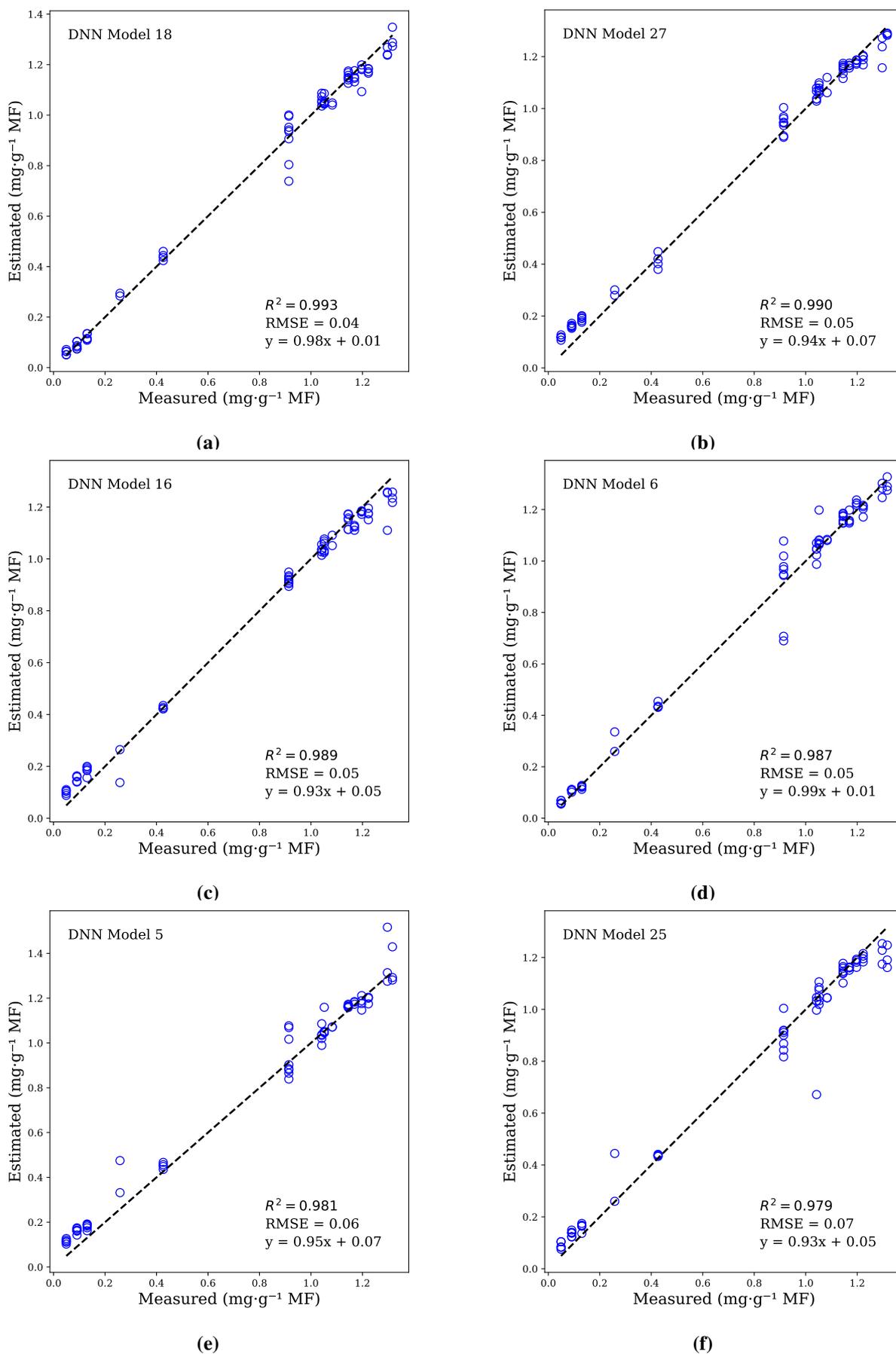
Table 23 – Performance metrics of various DNN regression models on different datasets. The table shows R^2 , root mean squared error (RMSE), and mean absolute error (MAE) for both the test and train sets, along with the time taken for each model. Models are evaluated using different combinations of data preprocessing techniques and feature sets

Model	Data1	Data2	Result Test Set			Result Train Set			Training Time (s)
			R^2	RMSE	MAE	R^2	RMSE	MAE	
Model 18	Detrended	Smoothed	0.993	0.040	0.027	0.996	0.030	0.024	81.55
Model 27	SNV	Detrended	0.990	0.045	0.037	0.993	0.040	0.031	80.74
Model 16	Detrended	SNV	0.989	0.048	0.037	0.991	0.044	0.036	84.55
Model 6	-	Smoothed	0.987	0.053	0.032	0.997	0.024	0.018	87.56
Model 5	-	Nor. Area	0.981	0.065	0.046	0.994	0.036	0.028	90.31
Model 25	SNV	-	0.979	0.067	0.041	0.993	0.040	0.031	80.76
Model 11	CR	Nor. Area	0.975	0.073	0.058	0.985	0.056	0.047	81.15
Model 36	Nor. Area	Smoothed	0.974	0.074	0.035	0.997	0.027	0.021	81.23
Model 31	Nor. Area	-	0.973	0.076	0.055	0.988	0.052	0.036	81.36
Model 22	MSC	SNV	0.973	0.077	0.048	0.994	0.036	0.028	78.57
Model 32	Nor. Area	CR	0.967	0.084	0.065	0.986	0.055	0.039	84.62
Model 15	Detrended	MSC	0.961	0.091	0.047	0.993	0.039	0.029	86.07
Model 4	-	SNV	0.956	0.098	0.063	0.986	0.056	0.041	88.33
Model 37	Smoothed	-	0.955	0.099	0.044	0.997	0.027	0.021	112.60
Model 30	SNV	Smoothed	0.954	0.099	0.057	0.986	0.056	0.038	85.16
Model 21	MSC	Detrended	0.954	0.100	0.043	0.994	0.036	0.029	78.70
Model 39	Smoothed	Detrended	0.953	0.101	0.046	0.996	0.029	0.021	118.38
Model 33	Nor. Area	Detrended	0.952	0.102	0.054	0.994	0.036	0.029	80.75
Model 3	-	MSC	0.952	0.102	0.074	0.976	0.072	0.058	108.37
Model 41	Smoothed	SNV	0.949	0.105	0.054	0.990	0.048	0.035	136.86
Model 42	Smoothed	Nor. Area	0.947	0.107	0.057	0.990	0.045	0.033	136.73
Model 19	MSC	-	0.944	0.110	0.069	0.988	0.052	0.041	76.50
Model 35	Nor. Area	SNV	0.941	0.113	0.061	0.990	0.047	0.034	79.42
Model 34	Nor. Area	MSC	0.939	0.115	0.051	0.993	0.038	0.031	78.90
Model 2	-	Detrended	0.934	0.119	0.055	0.990	0.046	0.037	148.25
Model 17	Detrended	Nor. Area	0.921	0.130	0.071	0.985	0.058	0.047	86.32
Model 13	Detrended	-	0.921	0.131	0.060	0.993	0.038	0.033	84.17
Model 40	Smoothed	MSC	0.911	0.139	0.067	0.984	0.059	0.039	152.35
Model 20	MSC	CR	0.909	0.140	0.084	0.991	0.044	0.036	76.64
Model 10	CR	SNV	0.901	0.146	0.086	0.990	0.047	0.039	78.98
Model 14	Detrended	CR	0.897	0.149	0.084	0.988	0.052	0.039	86.95
Model 23	MSC	Nor. Area	0.896	0.150	0.073	0.991	0.045	0.031	80.70
Model 29	SNV	Nor. Area	0.891	0.153	0.066	0.995	0.033	0.026	81.36
Model 28	SNV	MSC	0.883	0.159	0.127	0.902	0.146	0.125	81.54
Model 38	Smoothed	CR	0.882	0.159	0.099	0.983	0.060	0.044	120.37
Model 9	CR	MSC	0.882	0.160	0.092	0.965	0.087	0.073	83.55
Model 24	MSC	Smoothed	0.876	0.164	0.089	0.980	0.065	0.043	78.70
Model 12	CR	Smoothed	0.849	0.181	0.110	0.976	0.072	0.047	82.32
Model 8	CR	Detrended	0.847	0.182	0.095	0.985	0.057	0.040	82.63
Model 26	SNV	CR	0.840	0.186	0.114	0.989	0.050	0.040	80.94
Model 1	-	CR	0.720	0.246	0.154	0.991	0.044	0.036	140.05
Model 7	CR	-	0.709	0.250	0.152	0.986	0.056	0.048	83.39

Source: Own authorship (2024).

The comparison of performance metrics for various MiniRocket and DNN regression models on the datasets reveals significant results. In Table 22, Model 36, using normalized by area and smoothed data, showed an R^2 of 0.928 and RMSE and MAE of 0.125 and 0.094, respectively, on the test set, with a training time of 3.392 seconds. In comparison, Table 23 shows

Figure 45 – Performance of DNN Models for Estimating Chlorophyll Content in Leaves with Test Data



Source: Own authorship (2024).

that Model 18, using detrended and smoothed data, achieved an R^2 of 0.993, with RMSE and MAE of 0.040 and 0.027, respectively, on the test set, and a considerably higher training time of 81.55 seconds. These results highlight the superiority of Model 18 in terms of classification accuracy, albeit at the cost of a longer training time. The higher R^2 in the results reflects a better predictive capability of the model for estimating total chlorophyll values in wheat leaves using diffuse reflectance spectra.

5 FINAL CONCLUSIONS AND FUTURE WORKS

This dissertation concerns the detecting water and light stress and estimating total chlorophyll in wheat leaves using multi-angular reflectance spectroscopy supervised by artificial intelligence. Chapter 2 provided a literature review of related works, highlighting their contributions and final notes on the limitations of each method. The conclusion of the literature review chapter emphasized the lack of research on multi-angular reflectance spectroscopy measurements.

Chapter 3 presented methodologies ranging from setting up stress environments for containers and monitoring them, seed germination to growth over 7 days, methods for generating spectral reflectance and absorption databases, preprocessing techniques for spectral data, and artificial intelligence approaches for data analysis. These included MiniRocket, which is based on time series for water and light stress classification, and two approaches—MiniRocket and DNN—for estimating total chlorophyll in wheat leaves.

The analysis in Chapter 4 provided information for each stage of the dissertation, highlighting the robustness and reliability of the methodologies introduced in Chapter 3 through the data obtained from implementation. The results of environmental conditions for the containers ensured that temperature and humidity results were similar across containers, with the exception of the container without any lighting. Statistical analysis of the reflectance and absorption spectra provided a spectral database. In Section 4.3, MiniRocket models for water and light stress classification in wheat leaves are presented. MiniRocket models demonstrated exceptional capability for classifying water stress in wheat leaves, distinguished by its relatively short processing time. The impact of preprocessing and the use of two preprocessed datasets to improve classification significantly enhanced accuracy, achieving perfect results (100%) for the test dataset. For water stress classification, a perfect result was achieved, with model 26 reaching 100% accuracy on the test dataset. For light stress classification, solid performance was achieved, with model 01 reaching 98.6% accuracy on the test dataset, suggesting that the combination of preprocessing techniques may be fundamental to further improving precision. For estimating total chlorophyll in leaves, MiniRocket showed shorter processing times (as low as 2.667 s and up to 6.840 s) but with a coefficient of determination of 0.928 for the best model. DNN had longer processing times (ranging from 76.50 s to 152.35 s) but achieved nearly perfect coefficients of determination (0.993) for the test dataset of the best model. These models with

high accuracy enable the classification of water and light stress in wheat leaves using diffuse reflectance spectra. Additionally, the coefficient of determination close to one allows for a more precise estimation of total chlorophyll in wheat leaves using diffuse reflectance spectroscopy.

5.1 COMPLIANCE WITH SPECIFIC OBJECTIVES

In addition to the general conclusions drawn from this study, it is crucial to revisit the specific objectives outlined in Section 1.2 and assess how each has been achieved. Below is a summary of how each objective has been addressed and demonstrated in the research:

1. **Water and Light Stress Environments:** This objective was achieved through Section 4.1, where the construction and statistical monitoring results of these environments during the experimental period are presented. The results indicate that the environmental conditions for the containers are relatively similar, with differences noted for the container without any lighting.
2. **Spectral Database for Reflection and Absorption:** This objective is addressed in Section 4.2, which presents statistical results for diffuse reflection spectra for each sample, as well as a comparison of absorption spectra (obtained with an HR4000 spectrometer) with results from two spectral indices (spectrophotometers). Both methods of obtaining spectral indices for chlorophyll estimation and the degree of correlation of the results are discussed.
3. **Machine Learning Models for Detecting Water and Light Stress:** Demonstrated in Section 4.3, where water stress classification is discussed in Subsection 4.3.1 and light stress classification in Subsection 4.3.2. The comparison of different models for various types of abiotic stress in wheat leaves through multi-angular reflectance spectroscopy is presented.
4. **Machine Learning Models for Estimating Total Chlorophyll Amount:** Addressed in Section 4.4, where two approaches are presented: first, the MiniRocket time series-based model (Subsection 4.4.1) and second, the deep learning-based DNN model (Subsection 4.4.2), with a comparison of their results.

In conclusion, this research contributes to the advancement of methods for detecting stress and estimating total chlorophyll in wheat leaves using diffuse reflectance spectra measured

at different angles of incident light (multi-angular). A method tolerant to angular variations for estimating total chlorophyll and detecting stress in wheat leaves is proposed.

5.2 FUTURE WORKS

Looking ahead, future research efforts could focus on further improving the methodology proposed and applied in this study by expanding the dataset to include different plant species and generating robust models, with participation from the biological and/or agricultural fields.

In future work, the proposed models, such as MiniRocket, are expected to be key tools for monitoring systems, especially due to their speed and efficient processing capability. MiniRocket, with its exceptional speed, will enable the implementation of real-time monitoring systems.

The proposed models, such as MiniRocket and DNN, are expected to be applied to different plant species for disease prediction tasks using diffuse reflectance spectra obtained from leaves.

Additionally, the proposed models are anticipated to be used in applications for detecting agrochemicals through diffuse reflectance spectra obtained from leaves.

Likewise, the proposed models, such as MiniRocket, DNN, and others to be explored, may be applied for the detection of heavy metals using diffuse reflectance spectra obtained from leaves.

REFERENCES

- ALAGOZ, Celal. Robust and efficient atrial fibrillation detection from intracardiac electrograms using minirocket. **International Journal of Engineering Research and Development**, Kirikkale University, v. 16, n. 1, p. 432–447, 2024.
- ARMSTRONG, Richard A. Should pearson’s correlation coefficient be avoided? **Ophthalmic and Physiological Optics**, Wiley Online Library, v. 39, n. 5, p. 316–327, 2019.
- ARNON, Daniel I. Copper enzymes in isolated chloroplasts. polyphenoloxidase in beta vulgaris. **Plant physiology**, Oxford University Press, v. 24, n. 1, p. 1, 1949.
- ARONOFF, S. The absorption spectra of chlorophyll and related compounds. **Chemical reviews**, ACS Publications, v. 47, n. 2, p. 175–195, 1950.
- BARNES, RJ; DHANOA, Mewa Singh; LISTER, Susan J. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. **Applied spectroscopy**, SAGE Publications Sage UK: London, England, v. 43, n. 5, p. 772–777, 1989.
- BAURIEGEL, E.; GIEBEL, A.; GEYER, M.; SCHMIDT, U.; HERPPICH, W.B. Early detection of fusarium infection in wheat using hyper-spectral imaging. **Computers and Electronics in Agriculture**, v. 75, n. 2, p. 304–312, 2011. ISSN 0168-1699.
- BJÖRN, Lars Olof; PAPAGEORGIU, George C; BLANKENSHIP, Robert E; GOVINDJEE. A viewpoint: why chlorophyll a? **Photosynthesis research**, Springer, v. 99, p. 85–98, 2009.
- BONDUGULA, Rohit Kumar; UDGATA, Siba K; SIVANGI, Kaushik Bhargav. A novel deep learning architecture and minirocket feature extraction method for human activity recognition using ecg, ppg and inertial sensor dataset. **Applied Intelligence**, Springer, v. 53, n. 11, p. 14400–14425, 2023.
- CAO, Zhigang; MA, Ronghua; DUAN, Hongtao; XUE, Kun. Effects of broad bandwidth on the remote sensing of inland waters: Implications for high spatial resolution satellite data applications. **ISPRS Journal of Photogrammetry and Remote Sensing**, Elsevier, v. 153, p. 110–122, 2019.
- CHAPPELLE, Emmett W; KIM, Moon S; III, James E McMurtrey. Ratio analysis of reflectance spectra (rars): an algorithm for the remote estimation of the concentrations of chlorophyll a, chlorophyll b, and carotenoids in soybean leaves. **Remote sensing of environment**, Elsevier, v. 39, n. 3, p. 239–247, 1992.

CHATTHA, MU. Effect of seeding rate and seed soaking duration on productivity of relay-intercropped wheat in cotton. **Pakistan Journal of Science**, v. 69, n. 2, 2017.

CHEMURA, Abel; MUTANGA, Onesimo; DUBE, Timothy. Remote sensing leaf water stress in coffee (*coffea arabica*) using secondary effects of water absorption and random forests. **Physics and Chemistry of the Earth, Parts A/B/C**, Elsevier, v. 100, p. 317–324, 2017.

CHEN, Yi-Chieh; THENNADIL, Suresh N. Insights into information contained in multiplicative scatter correction parameters and the potential for estimating particle size from these parameters. **Analytica chimica acta**, Elsevier, v. 746, p. 37–46, 2012.

CHICCO, Davide; WARRENS, Matthijs J; JURMAN, Giuseppe. The coefficient of determination r-squared is more informative than smape, mae, mape, mse and rmse in regression analysis evaluation. **Peerj computer science**, PeerJ Inc., v. 7, p. e623, 2021.

CIGANDA, Verónica; GITELSON, Anatoly; SCHEPERS, James. Non-destructive determination of maize leaf and canopy chlorophyll content. **Journal of plant physiology**, Elsevier, v. 166, n. 2, p. 157–167, 2009.

CLARK, Roger N; GALLAGHER, Andrea J; SWAYZE, Gregg A. Material absorption band depth mapping of imaging spectrometer data using a complete band shape least-squares fit with library reference spectra. *In*: JPL PUBLICATION 90-54. **Proceedings of the second airborne visible/infrared imaging spectrometer (AVIRIS) workshop**. [S.l.], 1990. v. 90, p. 176–186.

DAS, Bappa; SAHOO, Rabi N; PARGAL, Sourabh; KRISHNA, Gopal; VERMA, Rakesh; VISWANATHAN, Chinnusamy; SEHGAL, Vinay K; GUPTA, Vinod K. 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**, Elsevier, v. 247, p. 119104, 2021.

DATT, B. Visible/near infrared reflectance and chlorophyll content in eucalyptus leaves. **International Journal of Remote Sensing**, Taylor & Francis, v. 20, n. 14, p. 2741–2759, 1999.

DEMPSTER, Angus; SCHMIDT, Daniel F; WEBB, Geoffrey I. Minirocket: A very fast (almost) deterministic transform for time series classification. *In*: **Proceedings of the 27th ACM SIGKDD conference on knowledge discovery & data mining**. [S.l.: s.n.], 2021. p. 248–257.

DEMPSTER, ANGUS HUGH. **Scalable Methods for Time Series Classification**. 2022. Phd Thesis (PhD Thesis) — Monash University, 2022.

DENG, Xianzhi; ZHANG, Zhixin; HU, Xiaolong; LI, Jinmin; LI, Shenji; SU, Chenye; DU, Shuai; SHI, Liangsheng. Estimation of photosynthetic parameters from hyperspectral images using optimal deep learning architecture. **Computers and Electronics in Agriculture**, v. 216, p.

108540, 2024. ISSN 0168-1699. Available at: <https://www.sciencedirect.com/science/article/pii/S0168169923009286>.

EGGINK, Laura L; PARK, Hyoungshin; HOOBER, J Kenneth. The role of chlorophyll b in photosynthesis: hypothesis. **BMC plant biology**, Springer, v. 1, p. 1–7, 2001.

EL-HENDAWY, Salah E; AL-SUHAIBANI, Nasser A; ELSAYED, Salah; HASSAN, Wael M; DEWIR, Yaser Hassan; REFAY, Yahya; ABDELLA, Kamel A. Potential of the existing and novel spectral reflectance indices for estimating the leaf water status and grain yield of spring wheat exposed to different irrigation rates. **Agricultural Water Management**, Elsevier, v. 217, p. 356–373, 2019.

ESTRADA, Félix; FLEXAS, Jaume; ARAUS, Jose Luis; MORA-POBLETE, Freddy; GONZALEZ-TALICE, Jaime; CASTILLO, Dalma; MATUS, Ivan A; MÉNDEZ-ESPINOZA, Ana Maria; GARRIGA, Miguel; ARAYA-RIQUELME, Carlos *et al.* Exploring plant responses to abiotic stress by contrasting spectral signature changes. **Frontiers in Plant Science**, Frontiers Media SA, v. 13, p. 1026323, 2023.

EVOXX LTDA. **Manual de Instruções: Destilador de Água EVOXX**. Paçandu, PR, Brasil, 2024. Razão Social: Indústria EVOXX Ltda.

FLORES, Walter Oswaldo Cutipa; CARVALHO, Vinícius; MARTINS, Victor Hugo; FABRIS, José Luís; MULLER, Marcia; LOPES, Heitor Silvério; LAZZARETTI, André Eugenio. Machine learning-based methods for force mapping with an optical fiber sensing system. **IEEE Sensors Letters**, v. 8, n. 7, p. 1–4, 2024.

GAO, Dehua; LI, Minzan; ZHANG, Junyi; SONG, Di; SUN, Hong; QIAO, Lang; ZHAO, Ruomei. Improvement of chlorophyll content estimation on maize leaf by vein removal in hyperspectral image. **Computers and Electronics in Agriculture**, v. 184, p. 106077, 2021. ISSN 0168-1699. Available at: <https://www.sciencedirect.com/science/article/pii/S0168169921000958>.

GITELSON, Anatoly; MERZLYAK, Mark N. Quantitative estimation of chlorophyll-a using reflectance spectra: Experiments with autumn chestnut and maple leaves. **Journal of Photochemistry and Photobiology B: Biology**, Elsevier, v. 22, n. 3, p. 247–252, 1994.

GITELSON, Anatoly A; GRITZ, Yuri; MERZLYAK, Mark N. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. **Journal of plant physiology**, Elsevier, v. 160, n. 3, p. 271–282, 2003.

GITELSON, Anatoly A; MERZLYAK, Mark N. Remote sensing of chlorophyll concentration in higher plant leaves. **Advances in Space Research**, Elsevier, v. 22, n. 5, p. 689–692, 1998.

GONZÁLEZ-FERNÁNDEZ, Ana Belén; RODRÍGUEZ-PÉREZ, José Ramón; MARABEL, Miguel; ÁLVAREZ-TABOADA, Flor. Spectroscopic estimation of leaf water content in commercial vineyards using continuum removal and partial least squares regression. **Scientia horticulturae**, Elsevier, v. 188, p. 15–22, 2015.

GONZÁLEZ-SOPEÑA, JM; PAKRASHI, V; GHOSH, B. An overview of performance evaluation metrics for short-term statistical wind power forecasting. **Renewable and Sustainable Energy Reviews**, Elsevier, v. 138, p. 110515, 2021.

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. **Deep Learning**. [S.l.]: MIT Press, 2016. <http://www.deeplearningbook.org>.

HANSEN, Steen Honoré; REUBSAET, Leon. High-performance liquid chromatography (hplc) and high-performance liquid chromatography–mass spectrometry (lc-ms). **Bioanalysis of Pharmaceuticals: Sample Preparation, Separation Techniques, and Mass Spectrometry**, Wiley Online Library, p. 123–172, 2015.

HE, Juncai; LI, Lin; XU, Jinchao; ZHENG, Chunyue. Relu deep neural networks and linear finite elements. **arXiv preprint arXiv:1807.03973**, 2018.

HECHT, Eugene. **Optics, Fifth Edition, Global Edition**. Edinburgh Gate, Harlow, Essex CM20 2JE, England: Pearson Education Limited, 2017. ISBN 978-1-292-09693-3.

HOSSIN, Mohammad; SULAIMAN, Md Nasir. A review on evaluation metrics for data classification evaluations. **International journal of data mining & knowledge management process**, Academy & Industry Research Collaboration Center (AIRCC), v. 5, n. 2, p. 1, 2015.

HR40000, Ocean Optics. **HR4000 Data Sheet Description**. [S.l.], 2012. 1–40 p.

IHUOMA, Samuel O; MADRAMOOTOO, Chandra A. Sensitivity of spectral vegetation indices for monitoring water stress in tomato plants. **Computers and Electronics in Agriculture**, Elsevier, v. 163, p. 104860, 2019.

IMANISHI, Junichi; NAKAYAMA, Aki; SUZUKI, Yoko; IMANISHI, Ayumi; UEDA, Nobuyuki; MORIMOTO, Yukihiro; YONEDA, Minoru. Nondestructive determination of leaf chlorophyll content in two flowering cherries using reflectance and absorbance spectra. **Landscape and Ecological Engineering**, Springer, v. 6, p. 219–234, 2010.

INGRAHAM, RH; GILLETTE, DD; WAGNER, WD. Relationship of temperature and humidity to conception rate of holstein cows in subtropical climate. **Journal of Dairy Science**, Elsevier, v. 57, n. 4, p. 476–481, 1974.

JHA, Kirtan; DOSHI, Aalap; PATEL, Poojan; SHAH, Manan. A comprehensive review on automation in agriculture using artificial intelligence. **Artificial Intelligence in Agriculture**, Elsevier, v. 2, p. 1–12, 2019.

JIAO, Yiping; LI, Zhichao; CHEN, Xisong; FEI, Shumin. Preprocessing methods for near-infrared spectrum calibration. **Journal of Chemometrics**, Wiley Online Library, v. 34, n. 11, p. e3306, 2020.

KANDPAL, Kishor Chandra; KUMAR, Amit. Migrating from invasive to noninvasive techniques for enhanced leaf chlorophyll content estimations efficiency. **Critical Reviews in Analytical Chemistry**, Taylor & Francis, p. 1–16, 2023.

KASUAKI. **IL-593-S Digital Spectrophotometer with UV-VIS Range**. 2024. Product Catalog.

KFEQUIPAMENTOS. **Digital Spectrophotometer IL-593-S**. 2024. <https://kfequipamentos.com.br/produto/espectrofotometro-digital-com-faixa-de-190-a-1100nm-uv-vis-automatico-e-largura-de-banda-de-2nm/>.
Accedido: 31-may-2024.

KIYONO, Ken; TSUJIMOTO, Yutaka. Nonlinear filtering properties of detrended fluctuation analysis. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 462, p. 807–815, 2016.

KOBAYASHI, Masami; AKUTSU, Shinya; FUJINUMA, Daiki; FURUKAWA, Hayato; KOMATSU, Hirohisa; HOTOTA, Yuichi; KATO, Yuki; KUROIWA, Yoshinori; WATANABE, Tadashi; OHNISHI-KAMEYAMA, Mayumi *et al.* Physicochemical properties of chlorophylls in oxygenic photosynthesis—succession of co-factors from anoxygenic to oxygenic photosynthesis. **Photosynthesis**, InTech London, UK, p. 47–90, 2013.

KONDRATEVA, N; FILATOV, D; BOLSHIN, R; KRASNOLUTSKAYA, M; SHISHOV, A; OVCHUKOVA, S; MIKHEEV, G. Determination of the effective operating hours of the intermittent lighting system for growing vegetables. *In*: IOP PUBLISHING. **IOP Conference Series: Earth and Environmental Science**. [S.l.], 2021. v. 935, n. 1, p. 012004.

KRSTINIĆ, Damir; ŠERIĆ, Ljiljana; SLAPNIČAR, Ivan. Comments on “mlcm: Multi-label confusion matrix”. **IEEE Access**, v. 11, p. 40692–40697, 2023.

LI, Xiao; SUN, Zhongqiu; LU, Shan; OMASA, Kenji. A multi-angular invariant spectral index for the estimation of leaf water content across a wide range of plant species in different growth stages. **Remote Sensing of Environment**, Elsevier, v. 253, p. 112230, 2021.

LI, Xunlan; WEI, Zhaoxin; PENG, Fangfang; LIU, Jianfei; HAN, Guohui. Estimating the distribution of chlorophyll content in cvcv infected lemon leaf using hyperspectral imaging.

Computers and Electronics in Agriculture, v. 198, p. 107036, 2022. ISSN 0168-1699. Available at: <https://www.sciencedirect.com/science/article/pii/S0168169922003532>.

LIU, Wenjian; LI, Yanjie; TOMASETTO, Federico; YAN, Weiqi; TAN, Zifeng; LIU, Jun; JIANG, Jingmin. Non-destructive measurements of toona sinensis chlorophyll and nitrogen content under drought stress using near infrared spectroscopy. **Frontiers in Plant Science**, Frontiers Media SA, v. 12, p. 809828, 2022.

LÓPEZ, Raquel Elisa da Silva; SILVA, Leonardo Lucchetti Caetano da (Ed.). **Saberes, Ciências e Plantas Mediciniais: uma abordagem multidisciplinar**. Rio de Janeiro, RJ: Farmanguinhos/Fiocruz, 2024. 486 p. Inclui Bibliografia. ISBN 978-65-980644-3-3.

LOUGH, W John; WAINER, Irving W. **High performance liquid chromatography: fundamental principles and practice**. [S.l.]: cRc press, 1995.

LUQUE, Amalia; MAZZOLENI, Mirko; CARRASCO, Alejandro; FERRAMOSCA, Antonio. Visualizing classification results: Confusion star and confusion gear. **IEEE Access**, IEEE, v. 10, p. 1659–1677, 2021.

LYNCH, Michael; WEINER, Ed. Hplc: high performance liquid chromatography. **Environmental Science & Technology**, ACS Publications, v. 13, n. 6, p. 666–671, 1979.

MACCIONI, Andrea; AGATI, Giovanni; MAZZINGHI, Piero. New vegetation indices for remote measurement of chlorophylls based on leaf directional reflectance spectra. **Journal of Photochemistry and Photobiology B: Biology**, Elsevier, v. 61, n. 1-2, p. 52–61, 2001.

MACKINNEY, Gr. Absorption of light by chlorophyll solutions. **Journal of biological chemistry**, Elsevier, v. 140, n. 2, p. 315–322, 1941.

MAHAJAN, Gopal Ramdas; DAS, Bappa; MURGAOKAR, Dayesh; HERRMANN, Ittai; BERGER, Katja; SAHOO, Rabi N; PATEL, Kiran; DESAI, Ashwini; MORAJKAR, Shaiesh; KULKARNI, Rahul M. Monitoring the foliar nutrients status of mango using spectroscopy-based spectral indices and pls-r-combined machine learning models. **Remote Sensing**, MDPI, v. 13, n. 4, p. 641, 2021.

MAIMAITIYIMING, Matthew; GHULAM, Abduwasit; BOZZOLO, Arianna; WILKINS, Joseph L; KWASNIEWSKI, Misha T. Early detection of plant physiological responses to different levels of water stress using reflectance spectroscopy. **Remote Sensing**, MDPI, v. 9, n. 7, p. 745, 2017.

MAIN, Russell; CHO, Moses Azong; MATHIEU, Renaud; O’KENNEDY, Martha M; RAMOELO, Abel; KOCH, Susan. An investigation into robust spectral indices for leaf

chlorophyll estimation. **ISPRS Journal of Photogrammetry and Remote Sensing**, Elsevier, v. 66, n. 6, p. 751–761, 2011.

MALEKI, MR; MOUAZEN, AM; RAMON, H; BAERDEMAEKER, J De. Multiplicative scatter correction during on-line measurement with near infrared spectroscopy. **Biosystems Engineering**, Elsevier, v. 96, n. 3, p. 427–433, 2007.

MARÍN-ORTIZ, Juan Carlos; GUTIERREZ-TORO, Nathalia; BOTERO-FERNÁNDEZ, Verónica; HOYOS-CARVAJAL, Lilliana María. Linking physiological parameters with visible/near-infrared leaf reflectance in the incubation period of vascular wilt disease. **Saudi Journal of Biological Sciences**, Elsevier, v. 27, n. 1, p. 88–99, 2020.

MCKINNEY, HH; SANDO, WJ. The behavior of winter wheat in artificial environments. **Science**, American Association for the Advancement of Science, v. 71, n. 1852, p. 668–670, 1930.

MEYER, Veronika R. **Practical high-performance liquid chromatography**. [S.l.]: John Wiley & Sons, 2010.

MISHRA, Puneet; BIANCOLILLO, Alessandra; ROGER, Jean Michel; MARINI, Federico; RUTLEDGE, Douglas N. New data preprocessing trends based on ensemble of multiple preprocessing techniques. **TrAC Trends in Analytical Chemistry**, Elsevier, v. 132, p. 116045, 2020.

MURCHIE, Erik H; LAWSON, Tracy. Chlorophyll fluorescence analysis: a guide to good practice and understanding some new applications. **Journal of experimental botany**, Oxford University Press UK, v. 64, n. 13, p. 3983–3998, 2013.

NETO, Antonio José Steidle; LOPES, Daniela C; PINTO, Francisco AC; ZOLNIER, Sérgio. Vis/nir spectroscopy and chemometrics for non-destructive estimation of water and chlorophyll status in sunflower leaves. **Biosystems engineering**, Elsevier, v. 155, p. 124–133, 2017.

Ocean Optics, Inc. **HR4000 and HR4000CG-UV-NIR Series High-Resolution Fiber Optic Spectrometers Installation and Operation Manual**. Version 02-1105. Dunedin, FL, USA, 2005. <https://www.oceanoptics.com>.

Ocean Optics, Inc. **SpectraSuite Spectrometer Operating Software Installation and Operation Manual**. Dunedin, FL, USA, 2007. Document Number 000-20000-300-02-0607, DRAFT 06/18/07. Available at: <mailto:Info@OceanOptics.com>.

OGUIZA, Ignacio. **tsai - A state-of-the-art deep learning library for time series and sequential data**. 2023. Github. Available at: <https://github.com/timeseriesAI/tsai>.

OSCO, Lucas Prado; RAMOS, Ana Paula Marques; MORIYA, Érika Akemi Saito; BAVARESCO, Lorrayne Guimarães; LIMA, Bruna Coelho de; ESTRABIS, Nayara; PEREIRA, Danilo Roberto; CRESTE, José Eduardo; JÚNIOR, José Marcato; GONÇALVES, Wesley Nunes *et al.* Modeling hyperspectral response of water-stress induced lettuce plants using artificial neural networks. **Remote Sensing**, Mdpi, v. 11, n. 23, p. 2797, 2019.

PATHARKAR, Abhidnya; HUANG, Jiajing; WU, Teresa; FORZANI, Erica; THOMAS, Leslie; LIND, Marylaura; GADES, Naomi. Eigen-entropy based time series signatures to support multivariate time series classification. **Scientific Reports**, Nature Publishing Group UK London, v. 14, n. 1, p. 16076, 2024.

PEDERSEN, Jiesper Tristan Strandsbjerg; VUUREN, Detlef van; GUPTA, Joyeeta; SANTOS, Filipe Duarte; EDMONDS, Jae; SWART, Rob. Ippc emission scenarios: How did critiques affect their quality and relevance 1990–2022? **Global Environmental Change**, Elsevier, v. 75, p. 102538, 2022.

PETHE, Yoginee Surendra; GOURISARIA, Mahendra Kumar; SINGH, Pradeep Kumar; DAS, Himansu. Fsboa: feature selection using bat optimization algorithm for software fault detection. **Discover Internet of Things**, Springer, v. 4, n. 1, p. 6, 2024.

RAINIO, Oona; TEUHO, Jarmo; KLÉN, Riku. Evaluation metrics and statistical tests for machine learning. **Scientific Reports**, Nature Publishing Group UK London, v. 14, n. 1, p. 6086, 2024.

RICHARDSON, Andrew D; DUGAN, Shane P; BERLYN, Graeme P. An evaluation of noninvasive methods to estimate foliar chlorophyll content. **New phytologist**, Wiley Online Library, v. 153, n. 1, p. 185–194, 2002.

RIZWAN, Muhammad; NADEEM, Aamer; SINDHU, Muddassar Azam. Analyses of classifier's performance measures used in software fault prediction studies. **IEEE Access**, IEEE, v. 7, p. 82764–82775, 2019.

SAVITZKY, Abraham; GOLAY, Marcel JE. Smoothing and differentiation of data by simplified least squares procedures. **Analytical chemistry**, ACS Publications, v. 36, n. 8, p. 1627–1639, 1964.

SCHLEGEL, Kenny; NEUBERT, Peer; PROTZEL, Peter. Hdc-minirocket: Explicit time encoding in time series classification with hyperdimensional computing. *In: IEEE. 2022 International Joint Conference on Neural Networks (IJCNN)*. [S.l.], 2022. p. 1–8.

SEELY, G.R.; JENSEN, R.G. Effect of solvent on the spectrum of chlorophyll. **Spectrochimica Acta**, v. 21, n. 10, p. 1835–1845, 1965. ISSN 0371-1951. Available at: <https://www.sciencedirect.com/science/article/pii/0371195165800959>.

SHIMODA, Yousuke; ITO, Hisashi; TANAKA, Ayumi. Conversion of chlorophyll b to chlorophyll a precedes magnesium dechelation for protection against necrosis in arabidopsis. **The Plant Journal**, Wiley Online Library, v. 72, n. 3, p. 501–511, 2012.

SHINANO, T; LEI, TT; KAWAMUKAI, T; INOUE, MT; KOIKE, T; TADANO, T. Dimethylsulfoxide method for the extraction of chlorophylls a and b from the leaves of wheat, field bean, dwarf bamboo, and oak. **Photosynthetica**, Praha, Academia., v. 32, n. 3, p. 409–415, 1996.

SHIOI, Yuzo; FUKAE, Rumiko; SASA, Tsutomu. Chlorophyll analysis by high-performance liquid chromatography. **Biochimica et Biophysica Acta (BBA)-Bioenergetics**, Elsevier, v. 722, n. 1, p. 72–79, 1983.

SINGH, Harpinder; ROY, Ajay; SETIA, Raj; PATERIYA, Brijendra. Estimation of chlorophyll, macronutrients and water content in maize from hyperspectral data using machine learning and explainable artificial intelligence techniques. **Remote Sensing Letters**, Taylor & Francis, v. 13, n. 10, p. 969–979, 2022.

SMITH, James H. C.; BENITEZ, Allen. Chlorophylls: Analysis in plant materials. *In: _____*. **Modern Methods of Plant Analysis / Moderne Methoden der Pflanzenanalyse: Volume 4**. Berlin, Heidelberg: Springer Berlin Heidelberg, 1955. p. 142–196. ISBN 978-3-642-64961-5. Available at: https://doi.org/10.1007/978-3-642-64961-5_6.

STEURER, Miriam; HILL, Robert J; PFEIFER, Norbert. Metrics for evaluating the performance of machine learning based automated valuation models. **Journal of Property Research**, Taylor & Francis, v. 38, n. 2, p. 99–129, 2021.

SUMANTA, Nayek; HAQUE, Choudhury Imranul; NISHIKA, Jaishee; SUPRAKASH, Roy. Spectrophotometric analysis of chlorophylls and carotenoids from commonly grown fern species by using various extracting solvents. **Res J Chem Sci**, v. 2231, p. 606X, 2014.

SWINEHART, Donald F. The beer-lambert law. **Journal of chemical education**, ACS Publications, v. 39, n. 7, p. 333, 1962.

SWOCZYNA, Tatiana; KALAJI, Hazem M; BUSSOTTI, Filippo; MOJSKI, Jacek; POLLASTRINI, Martina. Environmental stress-what can we learn from chlorophyll a fluorescence analysis in woody plants? a review. **Frontiers in Plant Science**, Frontiers Media SA, v. 13, p. 1048582, 2022.

TANIGUCHI, Masahiko; LINDSEY, Jonathan S. Absorption and fluorescence spectral database of chlorophylls and analogues. **Photochemistry and photobiology**, Wiley Online Library, v. 97, n. 1, p. 136–165, 2021.

TUNCA, Emre; KÖKSAL, Eyüp Selim; ÖZTÜRK, Elif; AKAY, Hasan; TANER, Sakine Çetin. Accurate estimation of sorghum crop water content under different water stress levels using machine learning and hyperspectral data. **Environmental Monitoring and Assessment**, Springer, v. 195, n. 7, p. 877, 2023.

TYAGI, Manya; PANDEY, Girish Chandra. Physiology of heat and drought tolerance in wheat: An overview. **Journal of Cereal Research**, v. 14, n. 1, p. 13–25, 2022.

ULLAH, Aman; NADEEM, Faisal; NAWAZ, Ahmad; SIDDIQUE, Kadambot HM; FAROOQ, Muhammad. Heat stress effects on the reproductive physiology and yield of wheat. **Journal of Agronomy and Crop Science**, Wiley Online Library, v. 208, n. 1, p. 1–17, 2022.

URBANOVICH, EA; AFONNIKOV, DA; NIKOLAEV, SV. Determination of the quantitative content of chlorophylls in leaves by reflection spectra using the random forest algorithm. **Vavilov Journal of Genetics and Breeding**, Institute of Cytology and Genetics, Siberian Branch of the Russian Academy ..., v. 25, n. 1, p. 64, 2021.

VALIN, Hugo; SANDS, Ronald D; MENSBRUGGHE, Dominique Van der; NELSON, Gerald C; AHAMMAD, Helal; BLANC, Elodie; BODIRSKY, Benjamin; FUJIMORI, Shinichiro; HASEGAWA, Tomoko; HAVLIK, Petr *et al.* The future of food demand: understanding differences in global economic models. **Agricultural economics**, Wiley Online Library, v. 45, n. 1, p. 51–67, 2014.

WENDAR, Huang WenDar Huang; KUANHUNG, Lin KuanHung Lin; MINGHUANG, Hsu MingHuang Hsu; MENGYUAN, Huang MengYuan Huang; ZHIWEI, Yang ZhiWei Yang; PIYU, Chao PiYu Chao; CHIMING, Yang ChiMing Yang. Eliminating interference by anthocyanin in chlorophyll estimation of sweet potato (*ipomoea batatas* L.) leaves. 2014.

WILLSTÄTTER, Richard. Chlorophyll. **Journal of the American Chemical Society**, ACS Publications, v. 37, n. 2, p. 323–345, 1915.

WITHAM, F. H.; BLAYDES, D. F.; DEVLIN, R. M. **Experiments in Plant Physiology**. New York: D. Van Nostrand Company, 1971. 55–58 p.

YADAV, Satya Prakash; IBARAKI, Yasuomi; GUPTA, S Dutta. Estimation of the chlorophyll content of micropropagated potato plants using rgb based image analysis. **Plant Cell, Tissue and Organ Culture (PCTOC)**, Springer, v. 100, p. 183–188, 2010.

YANG, Zhenfa; XIAO, Hang; ZHANG, Lei; FENG, Dejun; ZHANG, Faye; JIANG, Mingshun; SUI, Qingmei; JIA, Lei. Fast determination of oxides content in cement raw meal using nir spectroscopy combined with synergy interval partial least square and different preprocessing methods. **Measurement**, Elsevier, v. 149, p. 106990, 2020.

YSI Environmental. **The Basics of Chlorophyll Measurement**. Yellow Springs, OH: YSI Environmental, 2024. The document details the role of chlorophyll in photosynthesis, its importance in water quality measurement, and various methods of measuring chlorophyll concentration. It also includes information on YSI's optical sensors for chlorophyll detection and the impact of fouling on optical measurements.

ZAHIR, Siti Anis Dalila Muhammad; OMAR, Ahmad Fairuz; JAMLOS, Mohd Faizal; AZMI, Mohd Azraie Mohd; MUNCAN, Jelena. A review of visible and near-infrared (vis-nir) spectroscopy application in plant stress detection. **Sensors and Actuators A: Physical**, Elsevier, v. 338, p. 113468, 2022.

ZHA, Jianfeng; LI, Huaizhan; GUO, Guangli; WANG, Jintao. Influence of temperature and humidity on similar material and its control measures. **Environmental Earth Sciences**, Springer, v. 76, p. 1–11, 2017.

ZHANG, Ji; CHEN, Yu; ZHAI, Yongjie. Zero-shot classification based on word vector enhancement and distance metric learning. **IEEE Access**, IEEE, v. 8, p. 102292–102302, 2020.

ZHANG, WenYu; ZHANG, WeiXin; YANG, YuWang; HU, GuangYong; GE, DaoKuo; LIU, Hong; CAO, HongXin *et al.* A cloud computing-based approach using the visible near-infrared spectrum to classify greenhouse tomato plants under water stress. **Computers and Electronics in Agriculture**, Elsevier, v. 181, p. 105966, 2021.

ZHENG, Wanwan; JIN, Mingzhe. The effects of class imbalance and training data size on classifier learning: an empirical study. **SN Computer Science**, Springer, v. 1, p. 1–13, 2020.

ZHOU, Liang; ZHENG, Xiaoyuan; YANG, Di; WANG, Ying; BAI, Xuesong; YE, Xinhua. Application of multi-label classification models for the diagnosis of diabetic complications. **BMC medical informatics and decision making**, Springer, v. 21, n. 1, p. 182, 2021.

ZHU, Fengle; QIAO, Xin; ZHANG, Yuqian; JIANG, Jiandong. Analysis and mitigation of illumination influences on canopy close-range hyperspectral imaging for the in situ detection of chlorophyll distribution of basil crops. **Computers and Electronics in Agriculture**, Elsevier, v. 217, p. 108553, 2024.