Institución Universitaria

Abiotic stress detection using spectral information for crop monitoring

Manuel Mauricio Goez Mora

Instituto Tecnológico Metropolitano
Faculty of Engineering
Medellín, Colombia
2023

Abiotic stress detection using spectral information for crop monitoring

Manuel Mauricio Goez Mora

Research work submitted as requirement to obtain the degree of:
Doctor en Ingeniería

Advisor:
Prof. Maria Constanza Torres Madroñero, MSc, PhD

Research Area:
Smart Machine and Pattern Recognition

Research Group:
Grupo de investigación en Automática, Electrónica y Ciencias Computacionales (AEyCC).

Instituto Tecnológico Metropolitano
Faculty of Engineering
Medellín, Colombia
2023

Abstract

Remote sensing is one of the technologies with the potential for precision agriculture applications. Remote sensing systems include passive sensors, such as multispectral and hyperspectral sensors, which measure the energy reflected or emitted by a surface along the electromagnetic spectrum. Remote sensing allows monitoring large areas in less time than regular soil analysis processes. Several studies have demonstrated the potential of spectral data to crop stress conditions. However, most of these studies are limited to spectral signatures taken in situ. Some works estimate crop conditions from multispectral and hyperspectral images, but most use vegetation indeces, which do not take full advantage of the spatial and spectral data captured by spectral cameras. Despite the continuing development of precision agriculture based on remote sensing, there is still ample scope for further studies to meet the agricultural sector's needs. This thesis focuses on the extracting information from spectral data to detect crop stress conditions. The study was developed in two scales. The first one seeks the spectral characterization of stressed crops from spectral signatures collected in situ. The second one studies the capacities and limitations of remotely captured spectral imagery for stress detection, considering spatial information. This work developed a framework for water and nutritional stress detection using crop signatures combining the capabilities of either band ratios, discriminative bands, or the full spectra with supervised classifiers to detect water and nutritional deficiencies from spectral signatures. In a second approach, this work studied the capabilities of spectral imaging for crop stress detection. The main objective of this stage was to integrate the spatial information provided by spectral imagery into the framework developed in the first stage. The proposed method was evaluated using images with various spatial and spectral resolutions. The results show that using the full spectral signature instead of vegetation indices significantly improves stress detection. Support vector machines or neural networks using complete spectral signatures obtained detection accuracies of up to 98 % for common bean, 88 % for maize, and 75 % for avocado crops. These percentages vary according to type, stress level, and genotype. The main challenge in using spectral signatures is data collection since it requires extensive fieldwork. As an alternative, we evaluated a methodology with multispectral images of only ten bands, which facilitates data acquisition, achieving 88 % and 70 % stress detection accuracy in common beans and maize.

Keywords: Remote Sensing, Spectral Response of Vegetation, Machine Learning, Dimensionality Reduction and Band Selection.

Acknowledgment

I express my deepest gratitude to Dr. Maria Torres-Madronero, my mentor and supervisor, for their invaluable guidance, tenacious support, and patience throughout my Ph.D. studies. Their extensive knowledge and experience have been a great source of inspiration for me in my academic research and daily life. Their steadfast commitment to the research and securing the necessary funding sources for the project made this work possible. Additionally, I extend my sincere appreciation to all my colleagues from the MIRP laboratory at the Instituto Tecnológico Metropolitano and my co-workers in program 70107 Intelligent Systems for Monitoring Permanent and Transitory Agricultural Crops - Convocatoria 852-2019 Ministerio de ciencia, tecnología e innovación.

Finally, I am sincerely grateful to my family for their unconditional support and motivation. I especially thank my wife, who has been a pillar in my moments of weakness and doubt and has constantly motivated and inspired me to reach new goals. I also thank my three-year-old son, Nicolas Goez, who has been my guiding light, filling me with hope and the desire to improve daily.

Contents

Abstract v

Acknowledgment vi

1 Introduction 1
1.1 Justification . . . 2
1.2 Objectives . . . 4
1.2.1 General Objective . . . 4
1.2.2 Specific Objectives . . . 4
1.3 Contributions . . . 4

2 State of Art 6
2.1 Precision Agriculture and Remote Sensing . . . 6
2.2 Abiotic Limiting Factors in Crops . . . 7
2.3 Vegetation Spectral Signature . . . 9
2.3.1 Vegetation indices and Band Ratios for Abiotic Stress Detection . . . 10
2.3.2 Parametric Models . . . 13
2.3.3 Nonparametric Models . . . 13
2.3.4 Feature Extraction/Selection Approach for Crop Stress Detection . . . 14
2.4 Spectral Image Classification . . . 16

3 Methodology 17
3.1 Crops Description . . . 19
3.1.1 Bean under phosphorus-deficiency stress (Cerete 2022) . . . 19
3.1.2 Maize under nitrogen-deficiency stress (Cerete 2022) . . . 22
3.1.3 Maize under nitrogen-deficiency stress (Rionegro 2022) . . . 23
3.1.4 Avocado under water-deficiency stress (Rionegro 2021) . . . 26
3.2 Spectral Signature Acquisition . . . 28
3.2.1 Preprocessing . . . 29
3.3 Multispectral Image Acquisition . . . 33
3.3.1 Image calibration . . . 33
3.3.2 Band Registration . . . 35
3.4 Hyperspectral Image Acquisition . . . 36

3.5 Datasets Description 37
3.5.1 Bean Cerete 37
3.5.2 Maize Cerete 38
3.5.3 Maiz Rionegro 39
3.5.4 Avocado Rionegro 40
3.6 Stress Detection using Machine Learning 42
3.6.1 Classification Methods 42
3.6.2 Characterization of Crops Under Stress Factors Using Spectrometry . 42
3.6.3 Feature Selection Techniques 44
3.6.4 Spatial Features 46

4 Spectral signature analysis **47**
4.1 Signature characterization 47
4.2 Band Selection 59
4.3 Feature Selection 69
4.4 Stress detection 73
4.5 Discussion 78

5 Spatial Feactures **80**
5.1 Training Process 80
5.2 Testing Process 85
5.3 Discussion 90

6 Conclusions and future work **91**

1 Introduction

The world population is projected to increase by more than a third or 2.3 billion people between 2009 and 2050. It represents an increasing demand for food production estimated at 70 % [1]. The traditional method to increase agricultural production is to expand cultivation areas [2, 3]. However, several adverse effects are presented with land-use change, such as loss of biodiversity, increased sensitivity of crops to disease, and increased use of pesticides, among others [1].

Spectrometry is increasingly used in precision agriculture for crop monitoring due to its benefits such as non-destructive test measurements and fast results [4]. Spectrometers, multispectral, and hyperspectral cameras *(all hereafter referred to as spectral sensors)* can be used for agricultural studies. Spectral sensors enable the measurement of radiation emitted and reflected by a surface across the electromagnetic spectrum. Spectrometers, a type of spectral sensor, can provide high-resolution measurements of the spectral response of a surface. These measurements are often taken directly on the surface being analyzed, allowing for precise and accurate data collection [5]. Instead, multispectral and hyperspectral cameras combined spectrometer capabilities with imagery, collecting spatial and spectral information [6]. These technologies offer great possibilities for vegetation studies since the plants have a characteristic spectral response that changes according to the phenological state, health, and environmental variables [7].

Water and nutritional stress detection from spectral data is usually based on vegetation indices. Although these have a low computational cost, they do not consider all the information the spectral signature provides along the electromagnetic spectrum. In the literature, parametric and nonparametric approaches are also found for crop stress estimation. Parametric models seek to estimate physical variables related to crops. However, these are computationally costly and difficult to validate. Finally, nonparametric techniques such as linear regression, principal component analysis, and machine learning techniques have been explored for crop stress detection [8].

Machine learning approaches are of interest for this work. This research evaluates the capabilities and limitation of spectral data by combining either band indices, discriminative bands, or the full spectra with machine learning approaches to detect abiotic stress in different crops (e.g., avocado, maize, and beans). Evaluated spectral data included spectrometry

(close-range spectral signatures) with a high spectral resolution, and spectral imagery with ten bands; the different kind of data allows to determine the effects of scale in the stress detection problem.

1.1. Justification

The agricultural sector must face several challenges, including optimizing land use, increasing production, reducing the effects of climate change, and decreasing the adverse effects of agricultural techniques such as monoculture, overuse of fertilizer, and inadequate irrigation systems [9]. Therefore, precision farming technologies, which integrate sensor and crop information systems, are increasingly used. Precision agriculture seeks to improve productivity and cost-effectiveness and reduce the environmental impact of excessive chemical use [10].

Crop monitoring allows yield and quality prediction, providing timely information to optimize crop performance. However, measuring moisture and nutrients has several difficulties due to the variability of the topography, droughts, floods, air currents, and crop cycles, among other factors [11]. It has been reported that crop condition detection results may vary from season to season due to abiotic conditions [12]. The works of [13, 14, 15] presented satisfactory results, above 85 % of success, for hydric stress estimation under controlled conditions.

Due to the need to optimize crops [16] and the constant development of technologies, there is a growing interest in developing and implementing precision agriculture technologies [17]. One of the technologies used in precision agriculture [18] is remote sensing [19]. Remote sensing includes systems, sensors, and techniques that remotely measure physical variables from a surface, enabling extensive area coverage and avoiding invasive measurements [5, 4]. Some remote sensing systems are passive sensors, such as multispectral and hyperspectral sensors, which measure the energy reflected or emitted by a surface along the electromagnetic spectrum. Multispectral sensors capture this information in tens of board bands. Instead, hyperspectral sensors measure radiance in hundreds of narrow bands [5, 4].

Remote sensing allows monitoring large areas in less time than regular soil analysis processes [11, 20]. Current remote sensing applications in precision agriculture are cover mapping, vegetation identification, and change detection [21]. It is also used to detect water sources, leaf color intensity changes, crop temperature changes, estimate nutrient abundance, and perform plantation mapping [22].

Several studies have demonstrated the potential of spectral data to detect cropscrop stress conditions [23, 24, 25]. Most of these studies are limited to spectral signatures taken in situ [26, 27], and some works estimate crop conditions from multispectral and hyperspectral images [28]. However, most use vegetation indices [29, 30, 31, 32], which fail to take full

advantage of the spatial and spectral data captured by spectral cameras [33, 34]. The advantages of integrating spatial information from spectral images into the classification and estimation process have been identified. It be related to texture information with which discriminating features can be added [35].

Despite the continuing development of precision agriculture based on remote sensing, there is still ample scope for further studies to meet the agricultural sector's needs. Decision-making technologies are required to predict and implement corrective actions to optimize crop yield [36]. Multispectral and hyperspectral imagery can provide helpful information for crop management. Besides, a contribution can be generated to develop new techniques to consider the spatial and spectral information provided by spectral imagery [33, 34], and establish the limits of these technologies to detect stress in crops.

1.2. Objectives

1.2.1. General Objective

To develop a method for detecting water and nutritional stress in crops based on in situ spectrometry and remote sensing imagery, in the visible to the near-infrared range, considering spatial-spectral information.

1.2.2. Specific Objectives

1. To characterize the spectral signature of crops under the effect of abiotic limiting factors, such as water and nutritional deficiency, using spectrometry in the visible to the near-infrared range and feature selection techniques.

2. To develop a method for detecting water and nutritional stress of crops based on spectral imagery considering spatial-spectral information using machine learning and computer vision methods.

3. To evaluate the method for detecting abiotic limiting factors from spectral data using real data from field crops.

1.3. Contributions

The following is a list of publications related to the development of this thesis, including publications obtained as part of the research internship:

1. Torres-Madronero, M. C., Goez, M., Guzman, M. A., Rondon, T., Carmona, P., Acevedo-Correa, C., ... & Casamitjana, M. (2022). Spectral library of maize leaves under nitrogen deficiency stress. Data, 8(1), 2. https://doi.org/10.3390/data8010002.

2. Lopez, A. C. C., Goez Mora, M. M., Torres-Madronero, M. C., & Velez-Reyes, M. (2023, June). Using hyperspectral unmixing for the analysis of very high spatial resolution hyperspectral imagery. In Algorithms, Technologies, and Applications for Multispectral and Hyperspectral Imaging XXIX (Vol. 12519, pp. 277-284). SPIE. https://doi.org/10.1117/12.2666074.

3. Carmona-Zuluaga, P., Torres-Madronero, M.C., Goez, M., Rondon, T., Guzman, M., Casamitjana, M. (2023). Abiotic Maize Stress Detection Using Hyperspectral Signatures and Band Selection. In: Narvaez, F.R., Urgiles, F., Bastos-Filho, T.F., Salgado-Guerrero, J.P. (eds) Smart Technologies, Systems and Applications. SmartTech-IC 2022. Communications in Computer and Information Science, vol 1705. Springer, Cham. https://doi.org/10.1007/978-3-031-32213-6-35.

The following manuscripts are currently under evaluation:

4. Low-cost clamp for the measurement of vegetation spectral signatures. C. Acevedo-Correa, M. Goez, M.C. Torres-Madronero, T. Rondon. Sometido a Hardware X (Noviembre 2023).

5. Characterization of crops under abiotic stress factors using spectral signatures on the visible to near-infrared. M. Goez, M.C. Torres-Madronero, T. Rondon, M.A. Guzman, M. Casamitjana, J.M Gonzalez. Sometido a International Journal of Remote Sensing (Abril 2024).

In addition, the following conferences were presented:

6. Use of normalized difference vegetation index to assess nitrogen status in maize plants and grain yield prediction. Manuel Guzman, Maria Casamitjana, Tatiana Rondon, Maria Torres, Carolina Zuluaga, Manuel Goez. VII CBRG Congresso Brasileiro de Recursos Geneticos. 8-11 Noviembre 2022.

2 State of Art

2.1. Precision Agriculture and Remote Sensing

Precision agriculture refers to using advanced systems and techniques to optimize agricultural activities. This research area expands the frontiers of agricultural science by integrating knowledge fields such as engineering, computer science, economics, and social sciences [37], which aims to address the sector's current challenges, such as economic pressure, market behavior, and policies promoting sustainable agriculture [38]. It achieves this by monitoring, estimating, and controlling agricultural conditions, using spatial data related to crops, soil properties, and environmental factors to improve agricultural yields [35].

Despite the advantages of precision agriculture, its use in crops is limited. The implementation of precision agriculture requires high labor costs and the installation and maintenance of specialized equipment [39]. In this way, remote sensing becomes attractive because it can acquire spatial data from large crop areas faster and more effectively [40]. Remote sensing allows for timely decision-making and has enormous potential for application and development for precision agriculture [41, 42].

Remote sensing has applications in cartography, climate monitoring, defense and security, planetary exploration, mining, and agriculture [5, 4]. Remote sensing systems include passive and active sensors; passive sensors, like hyperspectral and multispectral cameras, measure the energy reflected and emitted by a surface along the electromagnetic spectrum [5]. Multispectral systems have tens of spectral board bands. Instead, hyperspectral imagery collects hundreds of narrow bands. Usually, multispectral systems are less expensive than hyperspectral sensors. However, selecting multispectral systems requires knowing relevant spectral features [43]. On the other hand, hyperspectral image processing involves higher computational costs due to the high dimensionality of the data [44].

2.2. Abiotic Limiting Factors in Crops

Several factors limit the normal development of plants, affecting their growth, productivity, and even survival [45]. These factors may be evidenced early by altered spectral responses of the plant [46]. Limiting factors in vegetation are classified as (Figure **2-1**) biotic factors, associated with the adverse effects produced by living beings, such as fungi and pests; and abiotic factors, related to nutritional deficiencies, and the environmental effects caused by rain, drought, temperature, and wind [47].

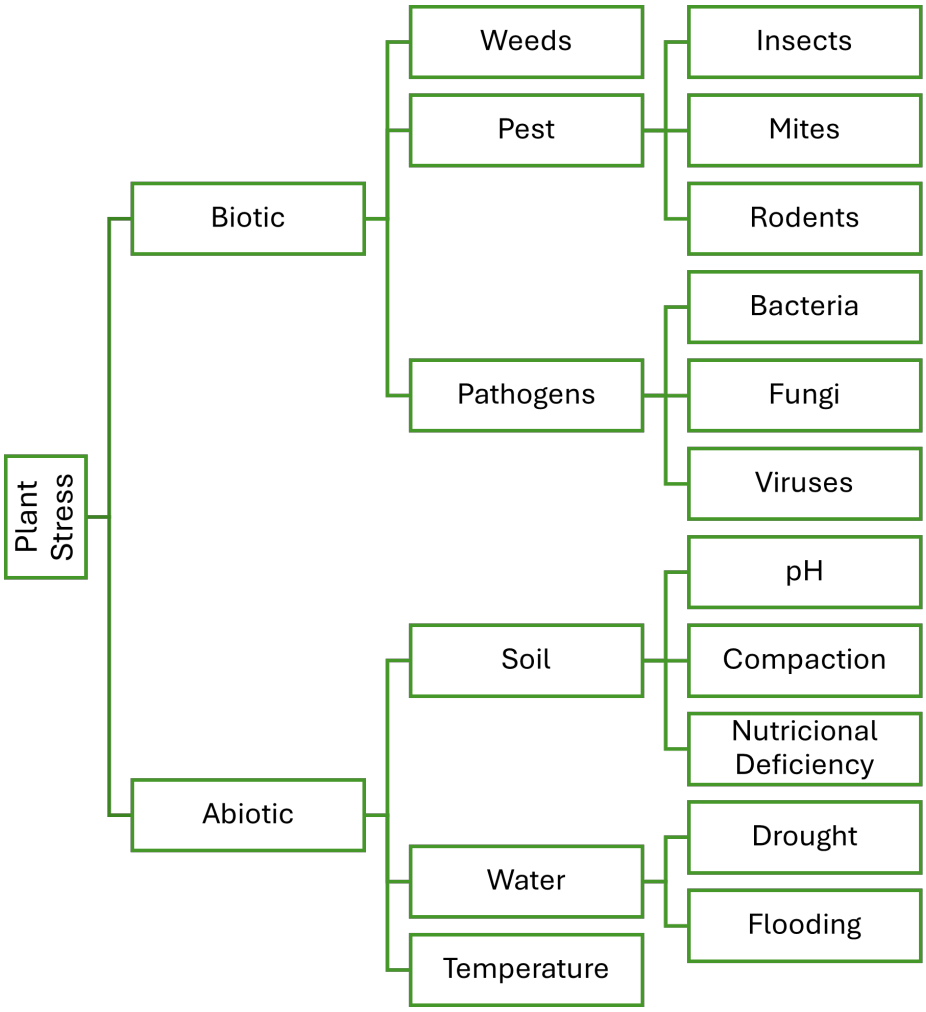


Figure 2-1: Most common abiotic and biotic factors of plant

The most common abiotic limiting factors in plants are water and nutritional deficiencies. Crop water stress occurs when atmospheric demand exceeds the soil water supply [48]. Water potential measurement is one indicator used to evaluate water stress [49]. However, hydric deficiencies are difficult to measure with conventional equipment, as the values can vary according to the time of day [49]. Water stress closes the stomata and impedes photosynthesis,

transpiration, and changes in leaf color and temperature [50]. On the other hand, nutritional stress is directly related to the photosynthetic process that causes changes in reflection, temperature, and vitality [51].

Soil moisture and nutrients are essential components for agricultural activities. However, controlling these variables has several difficulties due to many factors that influence the amount of water and nutrients, such as crop topology, droughts, floods, air currents, and crop cycles, among others [10]. Therefore, it is necessary to obtain up-to-date data to make decisions on crop management. The regular soil analysis procedure entails a high labor cost, which requires sampling and laboratory analysis. Therefore, few samples are taken throughout the crop, losing the level of detail of water and nutritional deficiencies [20]. Exploring new technologies that allow periodic evaluations is necessary, which will lead to new effective practical and rapid soil management methods..

Spectrometry, multispectral and hyperspectral imagery have been explored to detect water and nutritional stress in several crops. The following is a review of the most common approaches used for abiotic stress detection. Current spectral-based techniques can be classified into Vegetation indices (VI), which are easy to implement with a low associated computational cost, parametric methods for which a detailed study is required, and nonparametric methods for which prior data is necessary to perform the regression optimally.

2.3. Vegetation Spectral Signature

Information on plant components is obtained by analysis of reflectance spectral signatures acquired from the reflected solar spectrum. Considerable knowledge is available on the optical signatures of vegetation obtained through extensive field and laboratory spectroscopic analyses [52, 53]. Vegetation spectral signatures are a function of the optical properties of biogeochemical constituents, moisture status, and leaf elements' size, shape, and geometry [54]. The vegetation spectral signature represents the dynamic in the plants; these are a relationship among climate, plant structure, and vegetal function. Some intrinsic characteristics change the spectral response, like leaf anatomy, vascular tissues, photosynthetic processes, stomal, mesophyll, and canopy conductance. The spectral response of vegetation presents a complex but characteristic behavior (see Figure **2-2**). The visible region of the vegetation spectral signature has a low reflectance due to pigments, such as chlorophyll a and b, carotene, and xanthophyll. Due to chlorophyll, green vegetation presents absorption peaks around 420, 490, and 660 nm [55].The red border region between 690 and 720 nm is a critical feature of vegetation's spectral signature. Within this region, there is a significant increase in low red reflectance around 800 nm. This behavior is linked to leaves' internal structure and and their water content [55]. In the near-infrared region (700 to 1300 nm), plants have a high reflectance whose intensity depends on the internal leaf structure. The mid-infrared region (1300 to 2500 nm) is characterized by water absorption and the content of other biochemical compounds.

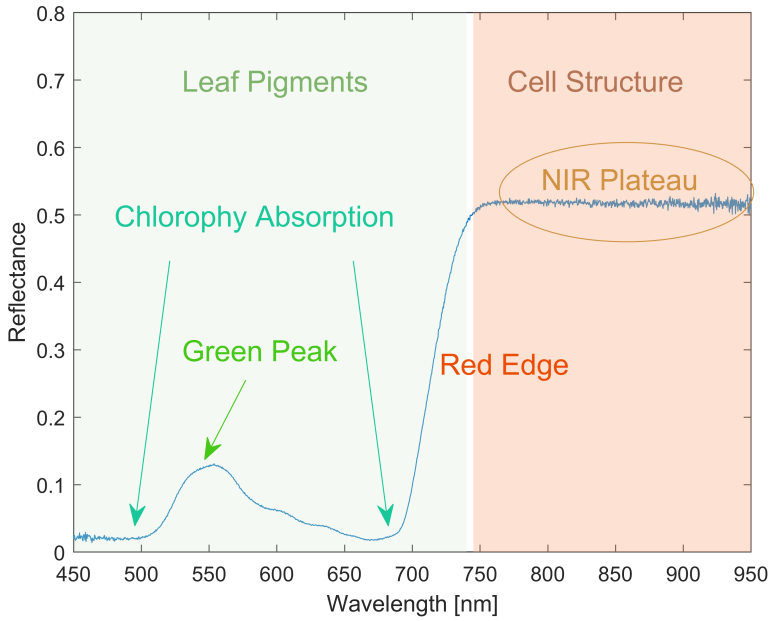


Figure 2-2: Vegetation spectral signature

A plant's spectral signature can change due to factors like growth stage, stress conditions, and irradiance. [21]. The spectral response and its relationship with different phenological and stress states and environmental and geographical conditions have been widely researched [56, 57, 58]. For example, the advantages and disadvantages of various optical proximal and remote sensing techniques for vegetation stress detection were evaluated in [58]. This study included proximal sensors, UAVs, and airborne and spaceborne sensors with different spatial resolutions. Behalf, [7] estimated irrigation levels with an RMSE closes to =0.015.

The analysis of vegetation spectral signatures is categorized into three approaches: methods based on vegetation indices [43], parametric models [7, 59, 60], and nonparametric models [23, 61, 62, 63] There methods Will be described bellow.

2.3.1. Vegetation indices and Band Ratios for Abiotic Stress Detection

Vegetation indices use limited spectral bands to compute band ratios correlated with phenological phenomena or stress factors. This approach uses between two and four bands and has a low computational cost [64]. An extended review is presented in [43]. Vegetation indices are used to predict leaf area, chlorophyll content, plant water content, biomass, and yield [43].

Using vegetation indices requires careful studies of the strengths and limitations for detecting and estimating plant conditions [35]. For instance, the normalized difference vegetation index (***NDVI***) is widely used in modern agriculture. This index is based on the reflectance in the near-infrared and part of the red [65]. Other indices commonly used in agriculture for the study of agronomic traits include: the crop water mass index (***CWMI***), which is highly related to the water content of the maize canopy [66]; the greenness (***GRE***) associated with plant chlorophyll content [67]; the plant senescence reflectance index (***PSRI***) and the photochemical reflectance index ***PRI***) correlated with irrigation and fertilizer levels [64]; chlorophyll index (***CI***) [68], modified soil adjusted vegetation index(***MSAVI***) and optimized soil adjusted vegetation index (***OSAVI***) [16] are used for the estimation of nitrogen status; among others. Table **2-1** summarizes vegetation indices and their use in abiotic stress detection.

Index	Equation	Type of crop	Data	Abiotic stress	Ref
Normalized Difference Vegetation Index [65]	$\boldsymbol{NDVI =} \frac{R_{NIR} - R_R}{R_{NIR} + R_R}$	Maize Wheat Maize	Spectrum Spectrum UAV	Water Water / Nitrogen Water	[22] [25] [29]
Ratio Vegetation Index [69]	$\boldsymbol{RVI =} \frac{R_{NIR}}{R_R}$	Maize	UAV	Nitrogen	[70]
Optimized Soil-adjusted Vegetation Index [71]	$\boldsymbol{OSAVI =} \frac{(R_{NIR} - R_R) * 1{,}16}{R_{NIR} + R_R + 0{,}16}$	Maize Wheat	Spectrum Spectrum	Water Water / Nitrogen	[22] [25]
Modified Chlorophyll Absorption in Reflectance Index [72]	$\boldsymbol{MCARI =} \frac{(R_{NIR} - R_R) * R_{NIR}}{R_R} - \frac{0{,}2(R_{NIR} - R_G) * R_{NIR}}{R_R}$	Wheat	Spectrum	Water / Nitrogen	[73]
Land Surface Wetness Index [74]	$\boldsymbol{LSWI =} \frac{R_{NIR} - R_{SWIR}}{R_{NIR} + R_{SWIR}}$	Wheat	Satellite	Water	[28]
Normalized Difference Water Index [75]	$\boldsymbol{NDWI =} \frac{R_{819} - R_{1243}}{R_{819} + R_{1243}}$	Potato	Spectrum	Water	[7]
Normalized Difference Infrared Index [76]	$\boldsymbol{NDII =} \frac{R_{819} - R_{1649}}{R_{819} + R_{1649}}$	Potato	Spectrum	Water	[7]
Moisture Stress Index [77]	$\boldsymbol{MSI =} \frac{R_{1600}}{R_{820}}$	Potato	Spectrum	Water	[7]
Red Edge Vegetation Stress Index [78]	$\boldsymbol{RVSI =} \frac{R_{714} - R_{752} - 2 * R_{733}}{2}$	Wheat	Spectrum	Nitrogen	[79]

Normalized Difference Red Edge Index [80]	$\boldsymbol{NDRE} = \frac{R_{NIR} - R_{RE}}{R_{NIR} + R_{RE}}$	Maize	UAV	Nitrogen	[70]
Canopy Chlorophyll Content Index [80]	$\boldsymbol{CCCI} = \frac{NDRE}{NDVI}$	Maize	UAV	Nitrogen	[70]
Green Normalized Difference Vegetation Index [81]	$\boldsymbol{GNDVI} = \frac{R_{NIR} - R_G}{R_{NIR} + R_G}$	Maize	Spectrum	Water	[22]
MERIS Terrestrial Chlorophyll Index [82]	$\boldsymbol{MTCI} = \frac{R_{NIR} - R_{RE}}{R_{RE} - R_R}$	Beans / Maize	Satellite	Water	[83]
Chlorophyll Estimation Red Edge Index [67]	$\boldsymbol{CI\text{-}rededge} = \frac{R_{NIR}}{R_R} - 1$	Maize	UAV	Nitrogen	[70]
Chlorophyll Estimation Green Index [84]	$\boldsymbol{CI\text{-}green} = \frac{R_{NIR}}{R_G} - 1$	Maize	UAV	Nitrogen	[70]
Carotenoid Reflectance Index [85]	$\boldsymbol{CRI} = \frac{1}{R_G} + \frac{1}{R_{NIR}}$	Potato	Satellite	Nitrogen	[86]

Table 2-1: Examples of vegetation indices and band ratios for abiotic stress detection. NIR: Near Infrared; RE: Red Edge; G: Green; R: Red; SWIR: Shortwave Infrared.

Vegetation indices' main advantage is their low computational cost; they can be easily calculated from the spectrum and multispectral and hyperspectral data. Despite their extensive use, vegetation indices use a limited subset of spectral bands [54]. They do not take advantage of all information from the spectral signatures or the spatial information provided by the spectral images.

2.3.2. Parametric Models

Parametric models are created by developing a response that describes the crop's characteristics based on its spectral response, which explores the light interaction with plants regarding their biophysical characteristics [59]. Biophysical models describe spectral variation as a function of land cover, leaf, and soil characteristics [59]. Estimating biophysical characteristics depends on an inversion process, which minimizes the difference between actual and simulated data.

Among the biophysical models are: the SAIL model [87], for bidirectional reflectance of the vegetation cover; PROSPECT [59], which models the optical properties of the leaves; and Hydrus-1D [88], a one-dimensional finite element model for simulating the movement of water, heat, and multiple solutes in soil. [60] compared vegetation indices and the PROSPECT model for estimating leaf chlorophyll content. In [89], a framework based on vegetation indices and biophysical model inversion is proposed to estimate leaf area index, total chlorophyll content, and water content from Terra and Aqua MODIS multispectral satellite data. Furthermore, [7] presented a study to assess water stress in potato crops using Hydrus 1-D.

Previous work has shown that the inversion model from spectral data is: (1) computationally expensive if the complex characteristics of the plants are included, (2) several of the parameters are interrelated, which makes it difficult to estimate them, and (3) the spatial resolution of the cameras limits the information available for the inversion model [21].

2.3.3. Nonparametric Models

Unlike parametric models, nonparametric models do not require assuming relationships between bands, knowing the data distribution, or including physical parameters. In this sense, new techniques to characterize crops based on methods such as regression, random forest, neural networks, k-nearest neighbors, and support vector machines are found in the literature [61, 63, 90, 91].

Previous studies have used regression models to estimate water stress. For instance, partial least-square regression is used to estimate water stress in potatoes [92], grapefruit [93, 94], leafy vegetables [92], wheat [95], and lettuce [96]. Other regression models, such as support vector [92] and ridge regression [97], are also found in crop monitoring applications. Most of these studies used spectral signatures [92, 95, 98], and a few spectral images [99]. The random forest is also explored in the literature to select a set of relevant spectral bands to estimate water [61] and nitrogen [100] deficiencies. Moreover, principal component analysis obtained a new representation space where water stress is detected for lettuce crops [101] and grass [23].

Most stress detection methods use spectral signatures measured in situ. Supervised classifier-based approaches are found in the literature for crop stress detection using spectral imagery, mainly for the detection of water deficiencies [90, 91, 102, 94, 103, 104, 105]. These approaches have two stages: first, a feature extraction/selection technique based on either spectral band selection or band ratios; second, a classification algorithm is applied to detect crop stress. Next, a brief review of methods for feature extraction/selection techniques and classification of spectral data is presented.

2.3.4. Feature Extraction/Selection Approach for Crop Stress Detection

Figure **2-3** (a) presents a graphical representation of the concept of spectral imagery. A spectral cube has two spatial dimensions (axes x and y) and a third dimension corresponding to spectral information. Figure **2-3** (c) represents the multispectral sensor bandwidth. Instead, hyperspectral imagery captures hundreds of narrow bands that obtained signature. as shown in Figure **2-3** (a). However, the high dimensionality of hyperspectral data involves a strong correlation among bands, The challenges posed by high-dimensional data, such as hyperspectral signatures, are called the *Curse of Dimensionality phenomenoon* in machine learning [104]. Usually, band ratios and spectral band selection approaches are used for spectral image processing to mitigate this phenomenon.

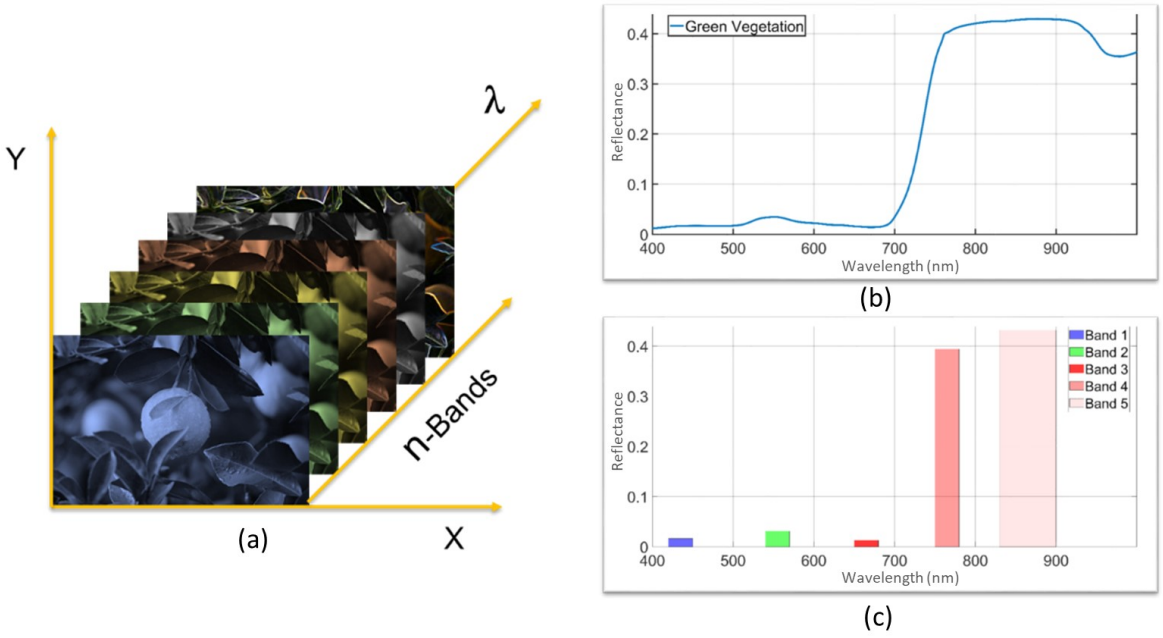


Figure 2-3: *S*pectral Signature Representation (a) Illustration of a Multispectral Cube. (b) Spectral Signature. (c) Bandwidth Representation.

As presented in previous sections, vegetation indices, and band ratios are agricultural applications' most common feature extraction approaches. However, subset band selection techniques can extract relevant and discriminative information from spectral data. Usually, these bands are selected such that changes between wavelengths are most prevalent [105]. Most subset band selection approaches require specifying the number of bands beforehand, but these techniques reduce the computational cost of hyperspectral data processing [106].

There are different techniques for band selection; these can be classified into supervised and unsupervised learning. Unsupervised techniques [106] seek to identify spectral bands with relevant statistical information and variability from hyperspectral data [107], such as maximum-variance principal component [108], fast density peak-based clustering [109], and information divergence [110]. Other techniques are based on iterative band removal, eliminating bands with little relevant information, such as Gaussian process regression [111], partitioned band image correlation, and capacitary discrimination [112].

In agricultural applications, some dimensional reduction techniques have been used for crop stress detection. For example, [90] used random forest and extreme gradient boosting for dimensional reduction and classification of hyperspectral imagery from grapefruit crops with water stress. However, more studies using band selection approaches for spectral characterization of crops are still needed because of spectral differences depending on the type of crop, season, and environmental conditions [113]; on the other hand, ideal bands should be selected according to the phenomenon under study (e.g., stress, diseases, phenology). For example, if the purpose of the study is to identify the presence of stress due to heavy metals, the ideal band is 1650 nm [114] , but for pigment detection, the outstanding bands are 450, 515, 550, 570, and 650nm for crops such as wheat, maize, rice, barley [115].

2.4. Spectral Image Classification

Machine learning in precision agriculture helps to detect crop stresses or to quantify their effects. The most common machine learning techniques used for vegetation spectral analysis are supervised approaches such as support vector machines (SVM) [6], random forest (RF) [61], neural networks (NN) [116, 117, 118], and deep learning [119]. SVM is widely used in remote sensing applications since obtains satisfactory results without representing a high computational cost [6]. The main challenge with applying supervised machine learning techniques in precision agriculture requires a significant amount of labeled data. Collecting labeled samples is generally difficult, expensive, and time-consuming [120].

Some supervised classifier examples applied to crop stress detection can be found in the literature [121, 122]. For example, an NN based on near-infrared reflectance is used in [121] and [122] for nitrogen stress detection in rice and wheat crops, respectively, and this classifier is used for water status estimation in vineyards in [123]. An SVM was used in [124] for water stress detection in wheat crops. Highlights in previous work the combination of machine learning with vegetation indices [123, 125].

Most of machine learning methods for detecting crop stress from spectral images only utilize spectral information. Only a few techniques incorporate spatial information into the analysis [104]. Several authors highlighted the importance of integrating spatial information in spectral image processing [126, 127, 128]. Spatial-spectral image processing takes advantage of both sources of information (spectral and spatial), improving the classification process [126, 127, 128] and mitigating isolated false points [16]. Studies have shown that the improvement in classification is provided by the pixels' spatial relationship [126, 127, 128]. Besides, it is essential to note that spatial information inclusion reduces the required labeled data [127]. Particularly within the scope of this study, to the best of our knowledge, only in [104] integrated spatial information through gray-level texture features for water content estimation on grapefruit.

High-performance machine learning techniques have recently been implemented in remote sensing and are growing steadily[129]. In particular, deep learning has shown promising results with high-performance indicators [130]. In particular, the detection of limiting factors in plants has shown increased incursion [131]. However, the need for a significant number of reliable, low-noise labeled data for proper training limits their practical application [131, 2, 16] so in this particular study they were not included as within the framework of the project the data acquisition protocols for both the spectrometer and multispectral cameras were developed, which introduces a level of uncertainty associated with the experimental trials, limiting the data available to ensure proper use of these techniques.

3 Methodology

Figure **3-1** presents the overall methodology for spectral signature analysis, which looks to detect stress on the plant using only spectral information (specific objective 1). Spectral analysis methodology considers the natural variation in the spectral response according to the species and genotypes [132]. An unsupervised feature selection process uses discriminative bands and vegetation index as input to determine the most relevant characteristics. Standard supervised classifiers were used for stress detection to generalize the approach to other studies. This first approach is a processing scheme that identifies stress conditions in spectral signatures.

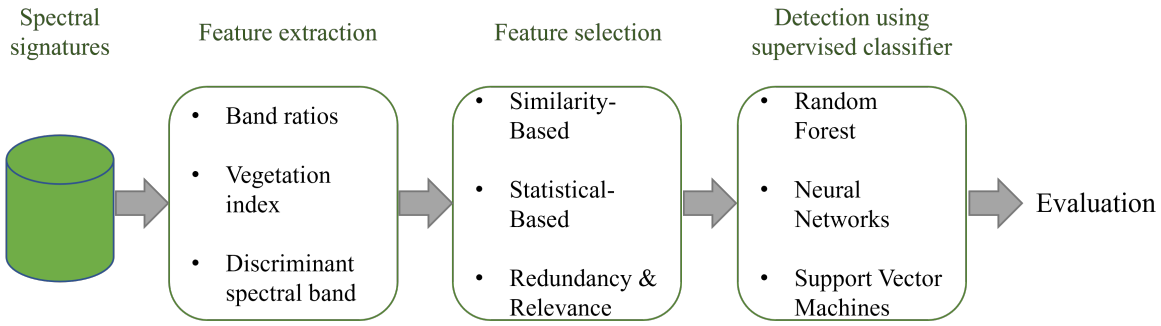


Figure 3-1: Spectral signatures processing methodology.

Once spectral signatures are analyzed, this work explores the capabilities of multispectral imagery. Multispectral cameras collect valuable spatial and spectral information (specific objective 2). Figure **3-2** presents a flowchart for the proposed approach for spectral image analysis for stress detection. Note that both methodologies are very similar (Figure **3-1** and **3-2**). However, the image analysis approach includes spatial information in the processing.

The following sections present the description of the experiments on maize, bean, and avocado crops, the acquisition of spectral data (signatures and images), their preprocessing (calibration, filtering, labeling), as well as the methods used for discriminant band selection, feature selection, and classification for stress detection. The vegetation indices employed in this study are those described in Table **2-1**. The results of the spectral signature analysis are presented in Chapter 4, and the images in Chapter 5. Then, both approaches are compared,

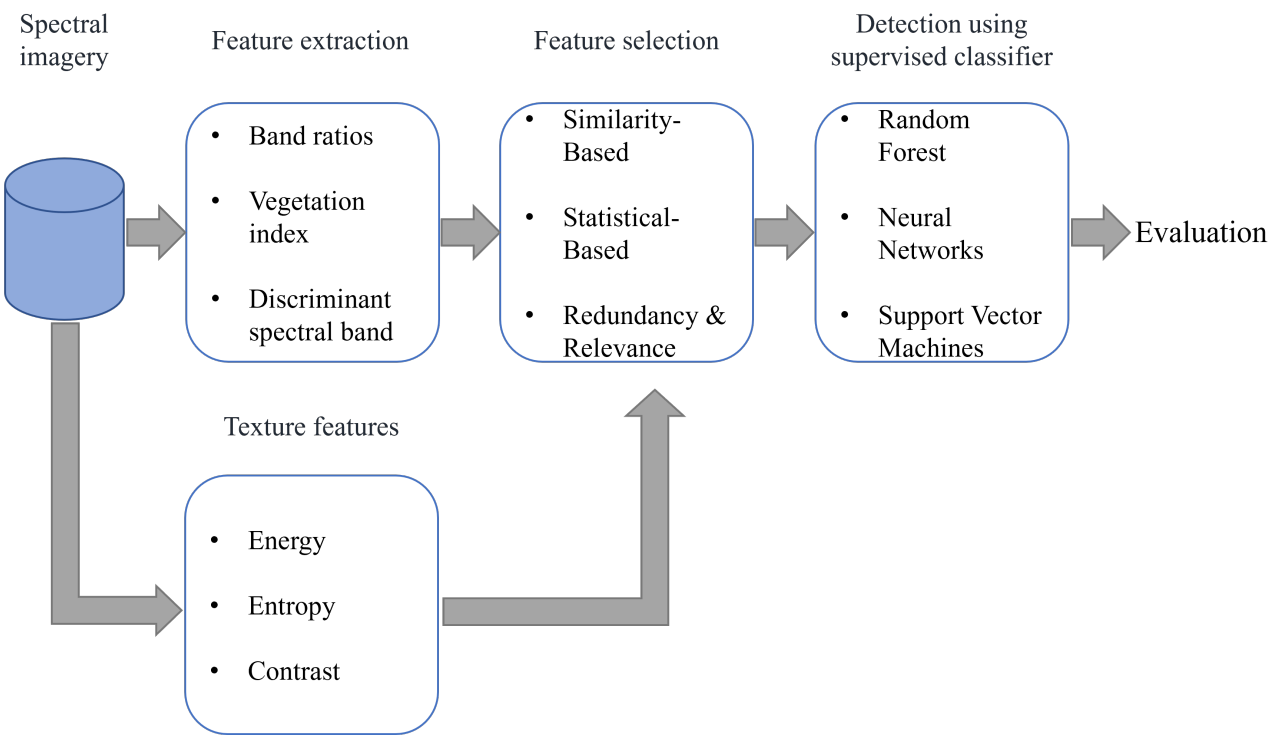


Figure 3-2: Spatial-spectral image processing methodology.

identifying potentials and limitations (specific objective 3).

Note: All performance tests were performed on a single computer with Intel(R) Core(TM) i9-9900K CPU @ 3.60GHz, 64.0 GB RAM HDD,1 TB; operating system: Windows 10 Pro 64-bit, MATLAB R2022b.

3.1. Crops Description

The datasets used in this project were acquired by the research program *Intelligent Systems for Monitoring Permanent and Transitional Crops* (grant RC 475-2020 Minciencias). They include spectral signatures and multispectral and hyperspectral imagery collected over common beans, maize, and avocado. The dataset was acquired over crops with some controlled conditions on trials established in Western Antioquia and North Cordoba (Colombia).For the experimental development of the crops, this project relied on the knowledge, advice, and experience of the co-investigators of the program associated with AGROSAVIA.

The first maize, common bean, and avocado trails were in La Selva Research Center of AGROSAVIA in Rionegro (Antioquia, Colombia) (06°08'06"N; 75°25'03"W, 2093 m.a.s.l.). This area has an average annual temperature of 17 °C, precipitation of 1917 mm, relative humidity of 78 %, daylight of 1726 h yr $^{-1}$, and evapotranspiration of 1202 mm. The study site is in the Low Montane Humid Forest (bh-MB) ecological life zone, in the Rionegro Association cartographic unit on a low alluvial terrace of the Rionegro River [133].

The other study site was in the Turipana Agricultural Research Center of AGROSAVIA, located in the Sinu Valley (Cerete, Cordoba, Colombia) (08°51'04"N; 75°49'05"W, 14 m.a.s.l.). This area has an average annual temperature of 29.7 °C, precipitation of 1280 mm, relative humidity of 80 %, daylight of 2190 h yr $^{-1}$, and evapotranspiration of 1329 mm. The study site is essentially flat without significant elevations [134].

3.1.1. Bean under phosphorus-deficiency stress (Cerete 2022)

The bean trial was conducted at the Turipana Agricultural Research Center. Two commercial genotypes of bush beans were studied (Table **3-1**). The plants were cultivated under two phosphorus pentoxide fertilization treatments: 100 % (T1) and 50 % (T2) of estimated crop requirements based on genotypes, expected yield, and initial soil conditions. The doses were determined by soil chemistry analysis before the experimental setup, obtaining 180 kg of N, 92 kg of P, and 148 kg of K. This information was used to establish the doses applied.

The studied bush beans include two commercial The soil was previously sterilized and freed of weed seeds and pathogens. Table **3-2** describes the soil's physical and chemical characteristics in the experimental trial's establishment plot under initial conditions before applying the treatments. Table **3-3** shows the fertilizer sources used for the mixtures and the base dose from which the restrictions were calculated. The quantity of fertilization applied was divided into two parts, 50 % at planting and 50 % 15 days after, due to the fertilization management recommendations of the crop. Crop fertilization was applied in bands along the planting row.

Id	Genotype	Description
G1	MUNGU	Commercial
G2	CAUPI	Commercial

Table 3-1: Description of Cerete bean genotypes.

	C. TEXTURE	pH	C.E dS.m-1	MO (%)	CO g/100g	P mg/kg
Substrate	Franco (F)	5.58	0.87	3.0	1.74	19.49

Table 3-2: Soil physical and chemical characteristics E.C.: electrical conductivity; MO: organic matter content; CO: organic carbon; P: total available phosphorous content.

The experimental design was arranged in a randomized complete block with a split-plot arrangement and three replications for 12 plots (Figure **3-3**(a)). Each plot included 300 plants of the same genotype, considering six furrows per plot of 10 m length spaced 0.2 m apart. The spacing between each replica was 2 m (Figure **3-3**). The distribution of genotypes within each replication was randomly arranged. The crop was established in the field on October 22, 2022.

Fertilizer	Kg/ha
UREA	151
Diammonium Phosphate	58
15-15-15	52
potassium sulfate	120

Table 3-3: Fractional dosage 50 % at planting and 50 % 15 days after planting.

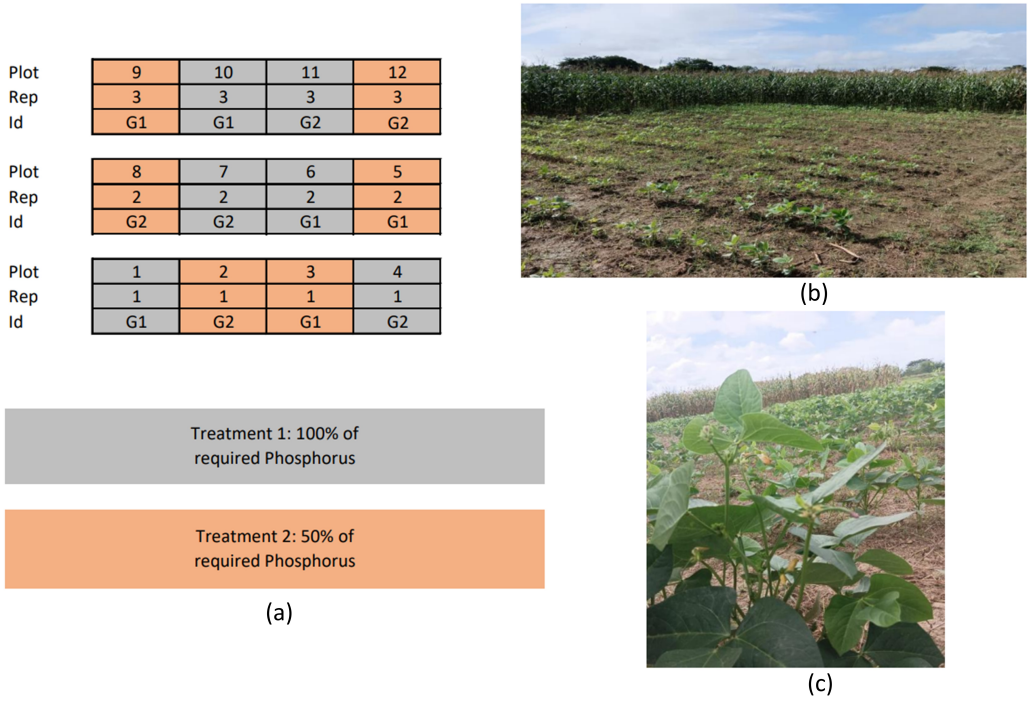


Figure 3-3: (a) Distribution (50 %, 75 %, 125 % y 150 %) of plots according to genotype, repetition, and treatment. (b) Bean Crop. (c) Photograph of a row of plants of the same genotype and treatment.

	Id	Genotype	Description
H	G1	V117	Experimental
	G2	P121	Experimental

Table 3-4: Description of Cerete maize genotypes.

	C. TEXTURE	pH	C.E dS.m-1	MO (%)	CO g/100g	NT (%)
Substrate	Franco (F)	5.58	0.87	3.0	1.74	0.15

Table 3-5: Soil physical and chemical characteristics E.C.: electrical conductivity; MO: organic matter content; CO: organic carbon; TN: total nitrogen content.

3.1.2. Maize under nitrogen-deficiency stress (Cerete 2022)

One of maize trials was conducted at the Turipana Agricultural Research Center. Like bean, the trial included two experimental genotypes (Table **3-4**). But the plants were cultivated under two nitrogen fertilization treatments: 100 % (T1) and 50 % (T2). The premliminary soil chemistry analysis established, 193.8 kg of N, 74.6 kg of P, and 178.8 kg of K as the basis to compute the dosages.

Table **3-5** describes the soil's physical and chemical characteristics before treatments, and Table **3-6** shows the fertilizer sources. The quantity of fertilization was divided into two applications : 70 % at planting and 30 % 30 days after.

The experimental design was also arranged in a randomized complete block with a split-plot arrangement and three replications for 12 plots (Figure **3-4**(a)). Each plot included 300 plants of the same genotype, considering six furrows per plot of 10 m length spaced 0.2 m apart. The spacing between each replica was 2 m (Figure **3-4**). The distribution of genotypes within each replication was randomly arranged. The crop was established in the field on September 28, 2022.

Fertilizer	Kg/ha
UREA	193.8
Diammonium Phosphate	74.16
Potassium Sulfate	176.6

Table 3-6: Fractional dosage is 70 % at planting and 30 % 30 days after planting.

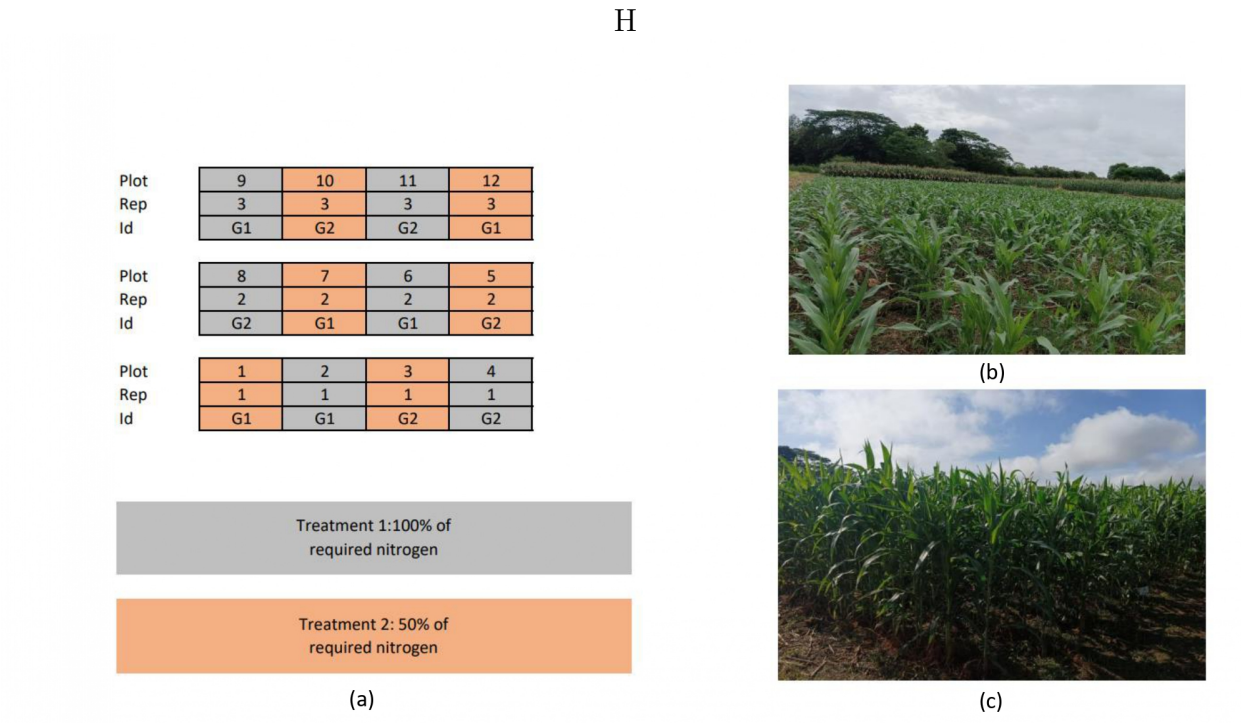


Figure 3-4: (a) Distribution of plots according to genotype, repetition, and treatment. (b) Maize Crop on 27/10/2022. (c) Maize Crop on 10/11/2022.

3.1.3. Maize under nitrogen-deficiency stress (Rionegro 2022)

The second trial was established at La Selva Research Center (Rionegro) in 2022. This trial incorporated ten genotypes, including experimental and commercial seeds from several companies, and different grain colors (see Table **3-7**). The plants were cultivated under different nitrogen fertilization treatments: 25 % (T1), 50 % (T2), 75 % (T3), and 100 % (T4) of the optimum level determined by soil analysis [135].

According to the experimental design, the soil used in the study was sterilized and treated to remove any unwanted weed seeds or pathogens. The physical and chemical properties of the soil before treatment application are presented in Table **3-8** . Table **3-9** provides information on the sources of fertilizer used in the mixtures and the base dose used to calculate the restrictions. The total amount of fertilizer was divided into two parts, with 70 % being applied at the time of planting and the remaining 30 % being applied 30 days.

The experimental design was arranged in a randomized complete block with a split-plot arrangement and three replications, where nitrogen doses were the main plots and genotypes

Id	Genotype	Description
G1	P3041	Commercial
G2	DK415UTPRO	Commercial
G3	DK7088	Commercial
G4	FNC8134	Commercial
G5	FNC8502	Commercial
G6	BIOMZn01	Commercial
G7	Synko	Commercial
G8	V114/P535	Experimental
G9	V114/P528	Experimental
G10	P535/V114	Experimental

Table 3-7: Description of Rionegro maize genotypes.

	C. TEXTURE	pH	C.E dS.m-1	MO (%)	Da g.cm-3	PT (%)
Substrate	Franco Limoso(FL)	5.58	0.34	18	0.55	67

Table 3-8: Soil physical and chemical characteristics E.C.: electrical conductivity; MO: organic matter content; Da: apparent density; PT: total porosity.

were the subplots, for a total of 120 plots (Figure **3-5**(a)). Each plot included eight plants of the same genotype, spaced 0.2 m apart, as shown in Figure **3-5**(b). The spacing between each plot was 0.8 m (Figure **3-5**(c)). The distribution of genotypes within each replication was randomly arranged. The crop was established in the field on April 5, 2022.

Fertilizer	Kg/ha
UREA	193.8
Diammonium Phosphate	74.16
Potassium Sulfate	176.6

Table 3-9: Fractional dosage 70 % at planting and 30 % 30 days after planting.

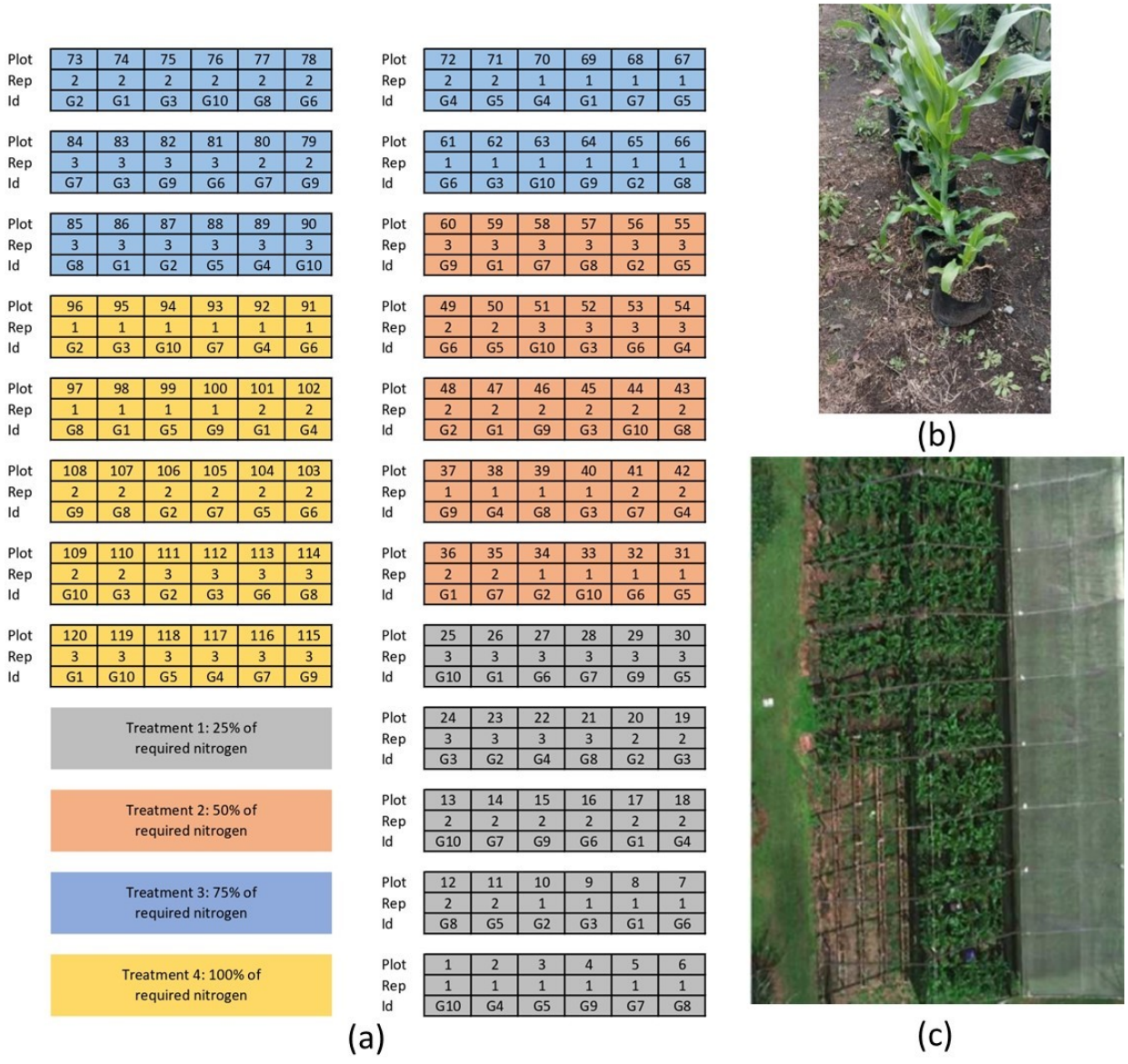


Figure 3-5: (a) Distribution of plots according to genotype, repetition, and treatment. (b) Photograph of a row of plants of the same genotype and treatment. (c) Aerial photograph of the experimental design area.

3.1.4. Avocado under water-deficiency stress (Rionegro 2021)

In 2021, the La Selva Research Center (Rionegro) conducted a trial of Avocado using two genotypes (Table **3-10**). The trial involved grafting Hass avocado onto these genotypes at 13 months, which were chosen for their tolerance to climatic conditions and biological irrigation in the region. The plants were grown under various water treatments, including 50 % (T1), 75 % (T2), 125 % (T3), and 150 % (T4) of the optimum level as determined by Dorado-Guerra [136] for the production zone of the Antioquia department. The optimum level for water was determined to be 222 ml per tree per day.

To ensure the irrigation levels applied for each experimental unit, the experiment was conducted under controlled conditions in a greenhouse and with the protection of an anti-trips net. The experimental design utilized in this study was a randomized complete block with a 23-factorial treatment arrangement consisting of two genotypes and four levels of water stress treatments (Figure **3-6** (a)). The experiment was replicated nine times, with each experimental unit consisting of one seedling.

To subject the avocado plants to the water stress treatment, 20-liter plastic plant pots were used after being disinfected with a 10 % hypochlorite solution. The plant pots were filled to 50 % with a commercial inert substrate to ensure their sanitary quality and physical and chemical properties (Table **3-11**). The seedlings were removed from the nursery bags and placed in the middle of the pots. The pots were finally filled with the same substrate, leaving approximately 5cm from the edge of the pot. After planting, each plant was fertilized to ensure the necessary nutrients for the crop, following the fertilization dose shown in Table **3-12**.

The plants were kept inside the greenhouse, separated by genotype, and the treatments were randomly marked. However, the treatments were applied one month after the plants were adapted to the greenhouse. During the adaptation period, the substrate moisture and field capacity were monitored to determine the appropriate water dose for all plants. Before starting the treatment application, the irrigation was stopped, and the soil tensiometers verified that the substrate had reached the field capacity.

Id	Genotype	Description
G1	ANSS88	Endemic
G2	ANGUI52	Endemic

Table 3-10: Description of Rionegro avocado genotypes.

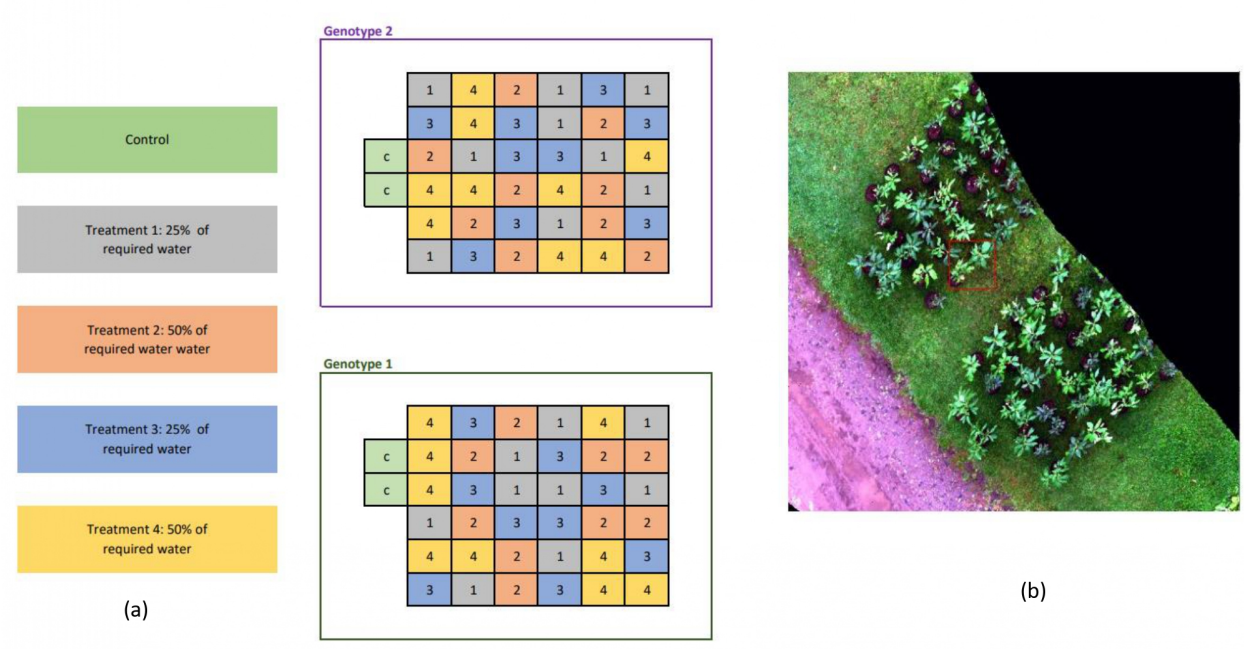


Figure 3-6: (a) Distribution of plots according to genotype and treatment. (b) Aerial photograph of the experimental design.

	C. TEXTURE	pH	C.E dS.m-1	MO (%)	Da g.cm-3	PT (%)
Substrate	Franco Limoso(FL)	5.58	0.34	18	0.55	67

Table 3-11: Soil physical and chemical characteristics E.C.: electrical conductivity; MO: organic matter content; Da: apparent density; PT: total porosity.

Fertilizer	Kg/ha
UREA	23.63
Diammonium Phosphate	20n-
15-15-15	26.67
potassium sulfate	1.3

Table 3-12: The fertilization dose was divided into two equal parts: half was applied when sowing, and the other 20 days after sowing. The remaining fertilizer was used three months later.

3.2. Spectral Signature Acquisition

The spectra were collected using a FLAME S VIR NIR spectrometer integrated with 2 m QP600-2-VIR-NIR optical fibers and an HL-2000-LL light source (all equipment, from Ocean Insight ®). The spectrometer's range is between 350 and 1000 nm, and it captures 2049 bands.

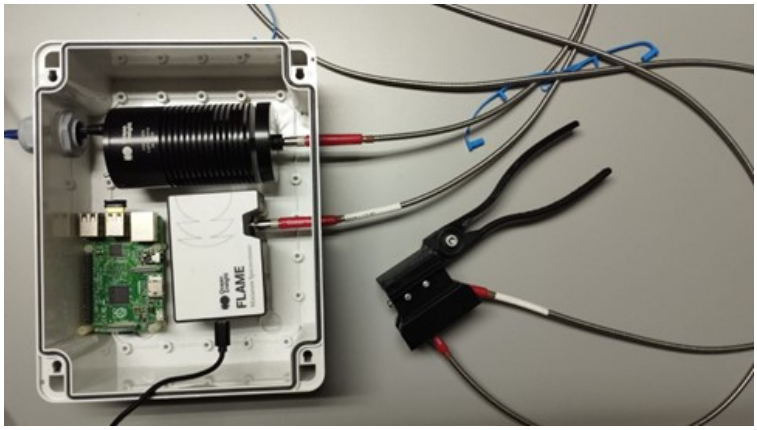


Figure 3-7: Spectrometer integration.

We used a customized clamp to integrate the spectrometer and light source (Figure **3-7**.). The tool was a 3D-printed clamp made of PLA material with a handle for easy manipulation. The clamp fixed the leaf and two optical fibers, with one connected to the light source and the other to the spectrometer (Figure **3-8**). The fiber probes were placed 5 mm from the surface, with an angle of 45° angles between the surface and the light, and the spectrometer fiber was in a perpendicular position to the surface. The captured spectrum resulted from an average of 10 measurements to improve the signal-to-noise ratio; the signatures were measured over the central area of the adaxial face of the upper side of the third leave of each plant. Black and white calibration spectra were collected before measurements on the leaves. OceanView 2.0 software (Ocean Insight, Orlando, FL, USA) was used to capture and calibrate the spectral signatures [135].

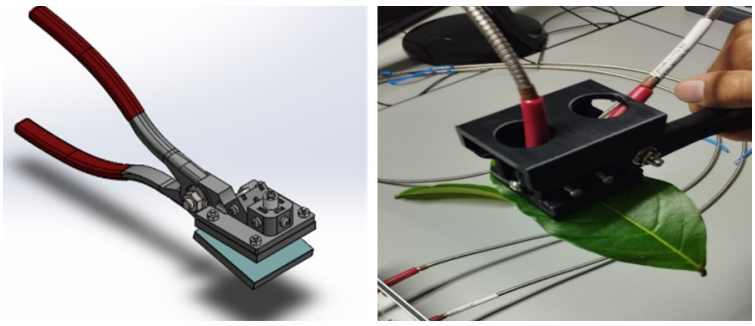


Figure 3-8: Customized clamp (a) design. (b) prototype.

3.2.1. Preprocessing

The acquired signatures presented atypical values along the spectral bands because of the inherent noise in the acquisition process (see Figure **3-9**). A graphical analysis helped to determine the spectral range with low noise levels, selecting only the spectral bands between 450 and 950 nm. Therefore, the high-noise bands were eliminated. In addition, a 10-point sliding window filter was applied to smooth and maintain the signature waveform improving the signal-to-noise ratio [137]. For instance, Figure **3-10** shows the effect of the window filter.

Once the signatures are filtered, it can be noted that some spectra have distortions and the effects of foreign particles; this results in unexpected shapes for vegetation signatures. A two - step automatic process was employed to remove the outliers. First, the standard deviation for each group per treatment and genotype are computed. Then, spectral signatures falling outside three standard deviations were eliminated, as shown in Figure **3-11**(a). This process is repeated, updating the standard deviation with the selected spectra. Figure **3-11**(b) shows the final results.

Figure **3-12** compares the original signatures and the resulting spectra after the preprocessing. Atypical acquisitions and overexposed signatures are successfully eliminated to obtain signatures without anomalies the main characteristics expected in vegetation signatures [132].

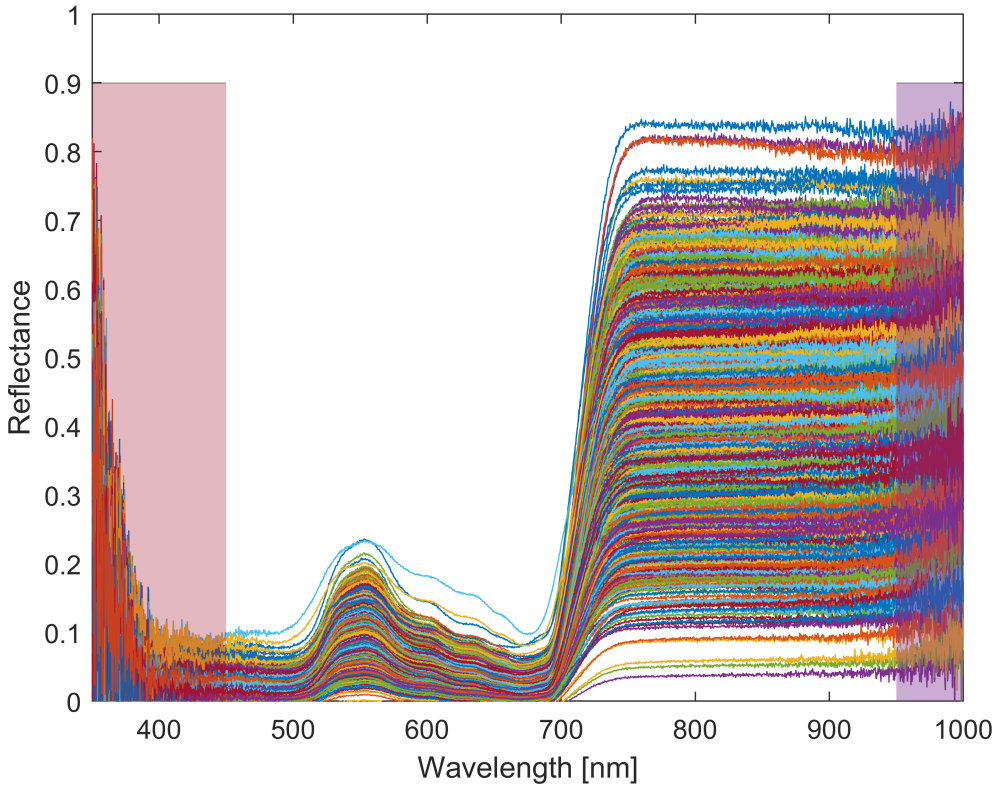


Figure 3-9: Highly noisy regions.

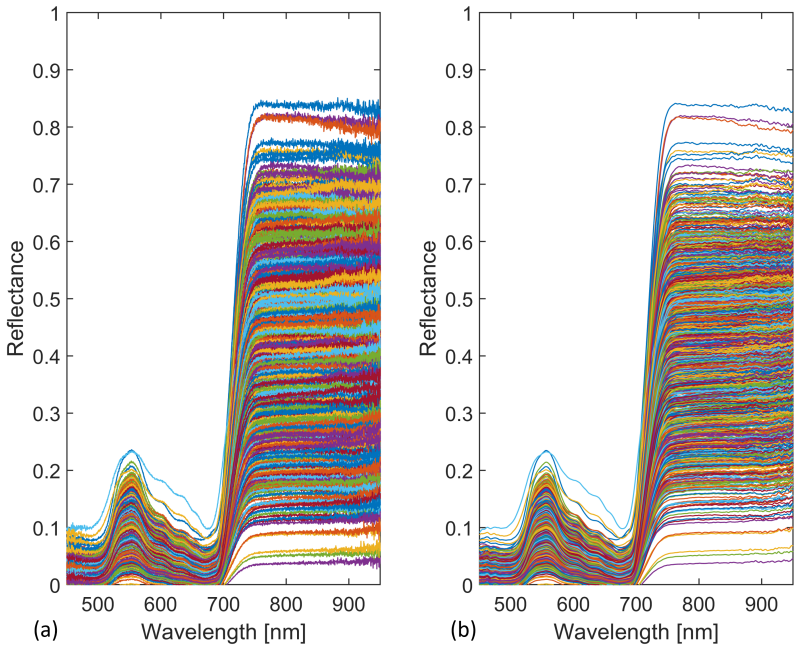


Figure 3-10: (a) Raw data (b) Filter data

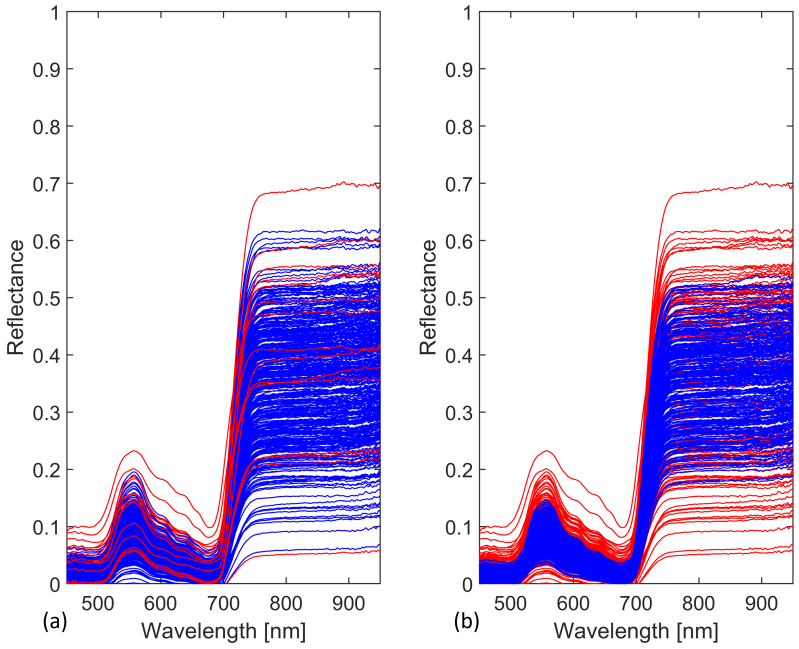


Figure 3-11: (a) Results after the first selection using the standard deviation (b) Results after the second step of selection using the standard deviation; the red lines indicate the signatures filtered by each stage.

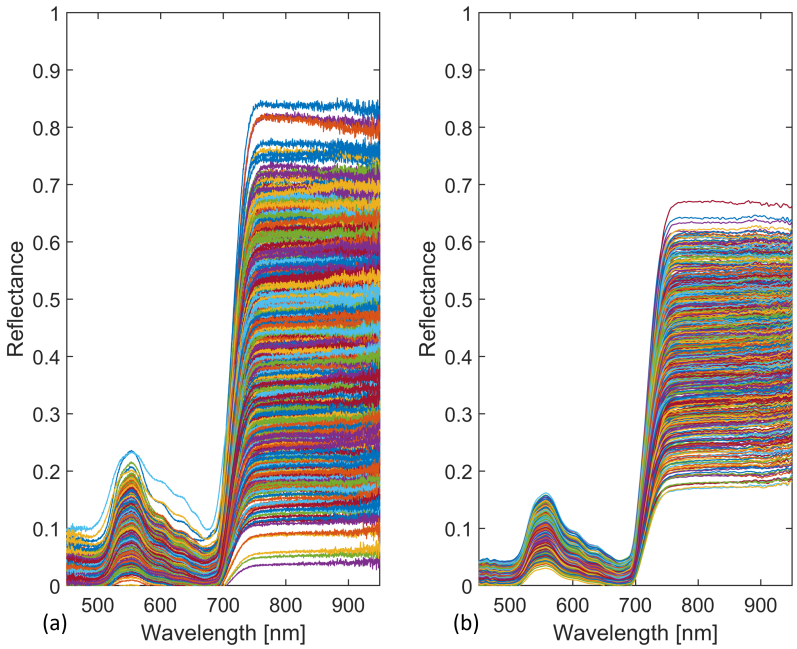


Figure 3-12: Comparison between (a) original signatures and (b) pre-processed signatures

.

3.3. Multispectral Image Acquisition

The multispectral images were collected using a MicaSense Dual Camera System (RedEdge-MX and RedEdge-MX Blue sensors). The acquired images have 1280 x 960 pixels (1.2 MP per band) and a radiometric resolution of 12 bits; these are stored as TIFF files. The MicaSense system captures ten spectral bands between 430 to 813 nm (Table **3-13**). The cameras were mounted on a tripod at a height of 2.10 meters with an image resolution of 1.4 mm/Pixel (Figure **3-14**) , integrating both cameras to capture the images simultaneously.

Bands	B1	B2	B3	B4	B5
RedEdge-MX red sensor [nm]	459-591	546-573	661-675	813-870	711-723
RedEdge-MX blue sensor [nm]	430-458	524-538	642-658	700-710	731-749

Table 3-13: RedEdge-Mx dual camera imaging system bands.

Figure 3-13: PLA-designed to integrate the tripod and cameras.

3.3.1. Image calibration

The MicaSense system includes a gray panel for image calibration. Before each image acquisition, a multispectral image is collected over the panel; this image is used to calibrate and transform raw values to reflectance. The image calibration includes building the spectral cube, calculating calibration factors using the panel, and changing it to reflectance. First, the ten images (one for each band) are sorted according to their wavelength (Table **3-14**); Then, the MicaSense radiometric calibration converts the raw values to radiance.

Figure 3-14: Setting up the tripod for image capture.

Chanel	1	2	3	4	5	6	7	8	9	10
Wavelength [nm]	440	475	531	560	650	668	705	717	740	842
Camera type	B	A	B	A	B	A	B	A	B	A
Band	1	1	2	2	3	3	4	5	5	4

Table 3-14: Multispectral cube bands organization (Camera type A: RedEdge-MX B: RedEdge-MX Blue).

The manufacturer recommends the employed procedure. Both steps (sorting bands and radiometric calibration) are applied to the image with and without a panel. Once the images are in radiance units, the panel is manually segmented, and the average radiance value is computed for each spectral band. The calibration factor is obtained by dividing MicaSense's average reflectance provided in the panel by MicaSense and the average value obtained from the panel in the captured image. Finally, the reflectance cube is obtained by multiplying the reflectance factor and the radiance images.

3.3.2. Band Registration

Due to the array configuration in the MicaSense sensors, the captured images have different fields of view for each spectral band. Then, it is necessary to perform a band registration. The manufacturer recommends capturing images from 30 meters away from the sensor, i.e., from drones. However, in the experimental design of this work, we explored the acquisition of images at close range, at a distance of 2 meters, using the tripod. This results in a more significant displacement effect in the field of view of the images. To remedy this effect, we first decided to acquire multispectral images of leaves arranged on a flat surface, decreasing the effects of a three-dimensional scene; and secondly, we performed the registration of all bands at the center band of the red sensor (band 5, 711-723 nm) using Matlab and its computer vision library. The registration process is based on the SURF technique for feature extraction in grayscale images, using Matlab functions detectSURFFeatures and extractFeatures. This procedure extracted features from both the image to be registered and the reference image, selected common points, calculated valid points using the matchFeatures function, estimated geometric transformation using the estimateGeometricTransform2D function, and performed image projection using the estimated transformation with the warp function. This automated process can be applied to the nine bands (excluding the fixed band) to register the complete cube see Figure **3-15**.

Figure 3-15: The preprocessed image of a leaf on a mesh for reference.

3.4. Hyperspectral Image Acquisition

The hyperspectral images were collected using a Dual HySpex Mjolnir VS-620 (Figure **3-16**). The system includes a V1240 sensor with a spectral range of 400-1000 nm, 200 bands with a bandwidth of 3.0 nm, and 12 bits of radiometric resolution. In addition, the hyperspectral system has the V620 sensor with a spectral range of 970-2500 nm, 300 bands with a bandwidth of 5.1nm, and a 16-bits resolution. The drone captured the image to 2140 m.a.s.l, approximately 40 m from the surface. The obtained pixel size was 0.02 m (Figure **3-17**). ®PARGE software was used for the orthorectification and DROACOR for atmospheric correction and reflectance retrieval.

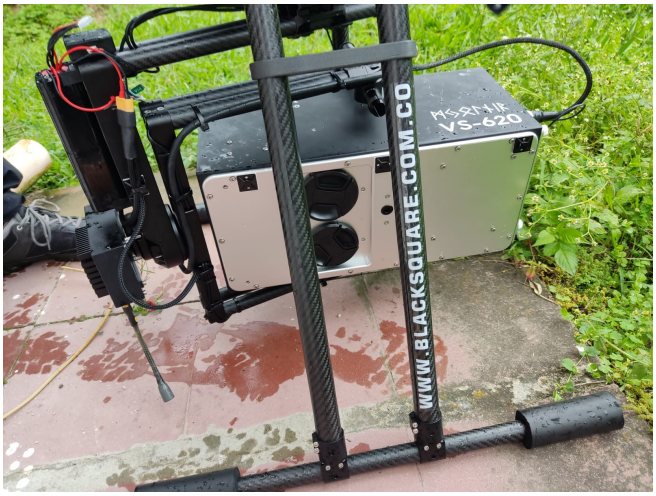


Figure 3-16: Drone equipped with HySpex Mjolnir hyperspectral camera.

Figure 3-17: Arrangement for drone image acquisition.

3.5. Datasets Description

3.5.1. Bean Cerete

For this work, spectral signatures and multispectral MicaSense images from the experiments of beans in Cerete are used. This experiment allows a simple framework for assessing the algorithms; since the experimental setup only includes two genotypes and two treatments. Twelve individuals were randomly selected from each plot with the same genotype, treatment, and repetition to acquire the spectra data. The signatures were measured over the central area of the adaxial face of each plant's upper third leaf, and ten spectral signatures were collected per leaf on 28 November 2022. The spectral library is obtained after applying the preprocessing procedure described in Section 3.2.1. Table **3-15** summarizes the spectra obtained by treatment and genotype after preprocessing.

	Genotype 1	Genotype 2
Treatment 1	310	301
Treatment 2	339	324
Total	649	625
Total Spectral DiammoniSignatureum	1274	

Table 3-15: Number of spectra after pre-processing.

To capture the multispectral images, twelve leaves per repetition and treatment were arranged in a 3 x 4 matrix on the floor to use the tripod. Once each image is captured, calibrated, and the bands recorded, the data is labeled manually, identifying the treatment and genotype. Finally, the images are cropped and organized into four photos, two corresponding to genotype one and two corresponding to genotype 2. Twelve training leaves and twelve evaluation leaves are selected for each genotype for each treatment. Table **3-16** summarizes the pixels obtained for each treatment and genotype.Figures **3-18** & **3-19** show an example of the segmented and rearranged images of the corresponding training and test sets.

	Genotype 1		Genotype 2	
	Train	Test	Train	Test
Treatment 1	23957	31222	25076	24914
Treatment 2	24082	29095	28786	27802
Total	108356		106578	
Total pixel	214934			

Table 3-16: Number of pixels per treatment and genotype.

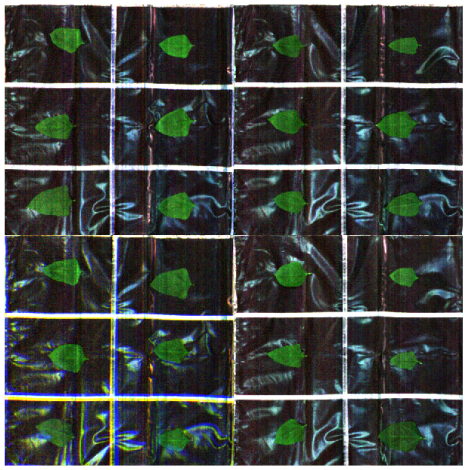


Figure 3-18: Training dataset genotype 1.

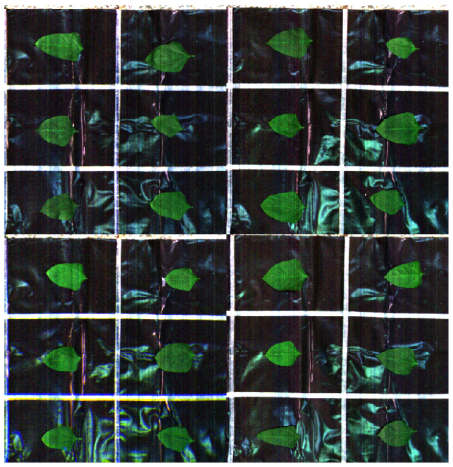


Figure 3-19: Testing dataset genotype 1.

3.5.2. Maize Cerete

For maize, this work used spectral data (signatures and multispectral image) from Cerete and Rionegro. For maize from the Cerete experiment, the spectral signatures were collected from twelve individuals randomly selected from each plot with the same genotype, treatment, and repetition on 11 November 2022 (see Table **3-15**).

	Genotype 1	Genotype 2
Treatment 1	309	318
Treatment 2	318	325
Subtotal	627	643
Total Spectral DiammoniSignatureum	1274	

Table 3-17: Number of spectra acquired from maize plants after pre-processing.

The multispectral images were captured on October 24, 2022 at the end of the experiment in Cerete, since the third leaf of each plant had to be removed. Twelve leaves were arranged in a 4 x 3 array on the ground to use the tripod for data capture. Once the images were preprocessed, the data were cropped to organize a mosaic for each genotype. The data were divided to form a training dataset and an evaluation dataset. Each dataset includes 18 leaves per treatment. Table **3-18** shows the summary of collected data. Figure **3-20** displays the segmented and rearranged images of genotype one.

	Genotype 1		Genotype 2	
	Train	Test	Train	Test
Treatment 1	41732	41587	42178	44765
Treatment 2	42145	40808	43547	44214
Total	166272		174704	
Total pixel	340976			

Table 3-18: Number of pixels per treatment and genotype.

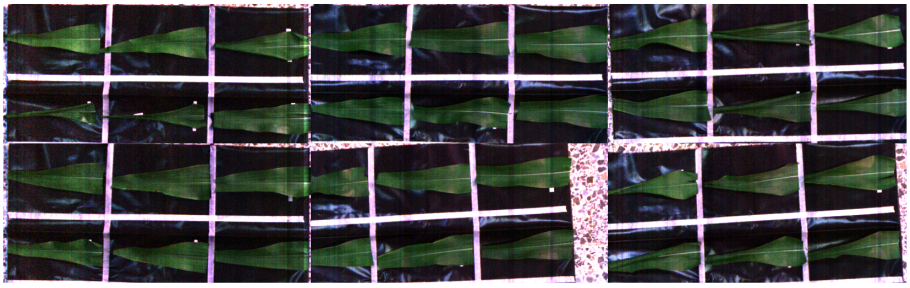


Figure 3-20: Dataset genotype 1.

3.5.3. Maiz Rionegro

The maize experiment in Rionegro allowed the evaluation of the proposed methodology with different stress levels and a more significant number of genotypes. experimentThis experiment used both spectral signatures and multispectral images. The spectral signatures were collected at 20 spectral signatures per leaf on 23 June 2022 (see Table **3-19**) when the maize was in stages V10 and V12 before flowering [138, 139, 135].

	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10
Treatment 1	52	52	51	51	48	54	51	43	52	52
Treatment 2	48	50	48	53	49	53	51	51	48	50
Treatment 3	51	51	51	51	53	48	50	54	54	53
Treatment 4	50	52	50	53	50	49	54	51	50	51
Subtotal	201	205	200	208	200	204	206	199	204	206
Total Spectral DiammoniSignatureum	2033									

Table 3-19: Number of spectra acquired from Maize plants after pre-processing

floorLeave samples were arranged in a grid pattern on the floor for image collection. This work shows that images from genotype 1 are only present to avoid an unnecessary extension on the document. Information for the four treatments and a summary of the data is provided in Table **3-20**.

	Genotype 1	
	Train	Test
Treatment 1	47013	50674
Treatment 2	54353	46334
Treatment 3	38024	29268
Treatment 4	46055	37622
Total pixel	349343	

Table 3-20: Number of pixels for treatment and genotype.

Figure 3-21: Training dataset for genotype 1.

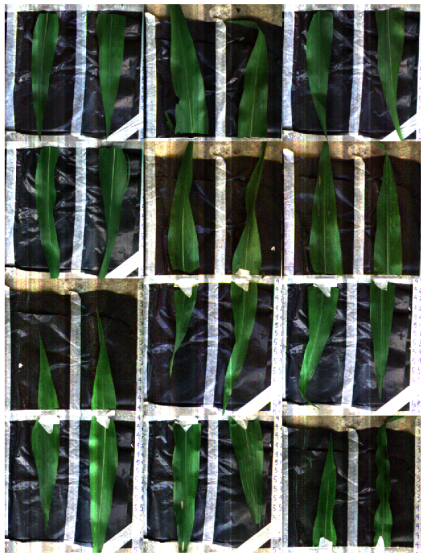


Figure 3-22: Testing dataset for genotype 1.

3.5.4. Avocado Rionegro

In the case of the avocado crop, only the spectral analysis is presented, based on the signatures acquired with the hyperspectral system described in section 3.4. The analysis of multispectral images is omitted because this was one of the first experiments performed. The procedure for data acquisition with spectrometers and tripod cameras was not perfected at the time of the avocado trials. The avocado plants were relocated outside the greenhouse to capture the hyperspectral image (Figure **3-6** (b)). The flight was performed on November 18, 2021. Once the images were calibrated and orthorectified, the data were subsampled to the same wavelengths as the spectral data from the maize and bean experiments. The selection of pixels corresponding to avocado leaves and their corresponding labeling between genotypes and treatments was performed manually. The data summary is presented in Table

3-21 after segmentation.

	Genotype 1	Genotype 2
Treatment1	1138	933
Treatment2	860	1244
Treatment3	1166	1003
Treatment4	1208	1225
Subtotal	4372	4405
Total	8777	

Table 3-21: Number of pixels acquired from avocado plants

3.6. Stress Detection using Machine Learning

3.6.1. Classification Methods

Although, as explored in the theoretical framework, there are several stress detection techniques, this work used classical supervised classifiers and evaluation metrics to assess the proposed methodologies for stress detection using spectral signatures and multispectral imagery. Classification techniques included in this work are: (1) Support Vector Machines ***(SVM)***, a robust method with acceptable overall performance on data with high dimensions [140]; (2) Random Forest ***(RF)*** that is widely used in nonlinear problems with multiple variables, characteristics present in the case study, with low sensitivity to variance in the training data and overfitting [141]; and, (3) the Fully connected Neural Networks ***(NN)*** specifically the multi-layer perceptron, which obtains the desired classification results with high dimensionality datasets [142]. The implementations of MATLAB 2022a were employed for the three classifiers. In the case of SVM, a 5th-degree polynomial was used as kernel, with an offset equal to 0.1, and the optimization routine of iterative single data algorithm (all included in the function template SVM). The RF parameters were set to 40 for the minimum leaf size, PCA as the categorial predictor, a maximum tree depth of 50, and a maximum number of splits of 9. Finally, the function fitness for the NN was used, with 25 fully connected layers and the optimized bayesopt.

3.6.2. Characterization of Crops Under Stress Factors Using Spectrometry

This study developed an algorithm to identify specific points of vegetation signatures by using local maxima and minima of the spectral signature and its first derivative. These points include the red edge change, chlorophyll absorption, green peak, and NIR Plateau. The results of this characterization are depicted in Figure **2-2**. The average of samples collected from healthy plants (on control crops) was computed for each type to determine the relevant points in the signature. The first derivative was then calculated from this average signature. For the points of maximum chlorophyll uptake, the points in the blue and red regions with the derivative closest to zero were chosen. The start and end of the red edge were also selected. The first band where the derivative is close to zero in the 750 to 950 nm range for the Plateau NIR was selected. These points may vary depending on the crop, genotype, nutritional health, and other factors, as stated in the literature [55]. This characterization aims to establish an initial set of relevant bands based on the expected behavior of a spectral signature.

In a second approach, automatic band selection methods were used to characterize the spectral signatures, band selection allow reducing the amount of data to process, mitigating the

curse of dimensionality effects. Also, selected bands will enable the design of specific spectral systems datafor a field of action, reducing instrument costs and making precision agriculture feasible without decreasing the data's classification potential. Band selection techniques can currently be classified into six groups considering the methods used: ranking, sparsity, search, clustering, learning, and hybrids [143].

In this work, the method used is a Similarity-Based Unsupervised Band [144] of the clustering group. It has shown a desired behavior, even without a priori information. The selected method identifies the most distinctive and informative bands based on the measurement of band similarity, showing a better performance in terms of information preservation and class separability than other widely used algorithms [144]. The algorithm modifies the normalized cut criterion technique and the maximum-variance principal component analysis, limiting highly correlated features and minimizing the sensitivity to noise bands. This process has a low computational consumption [145]. The procedure is performed by iterating the algorithm, until the total number of bands is reached. The incidence results of the selected bands are presented in a histogram, showing the frequency that each band is selected as representative (an example of a histogram is presented in Figure **3-23**). The histogram identifies relevant bands and spectral regions for each dataset.

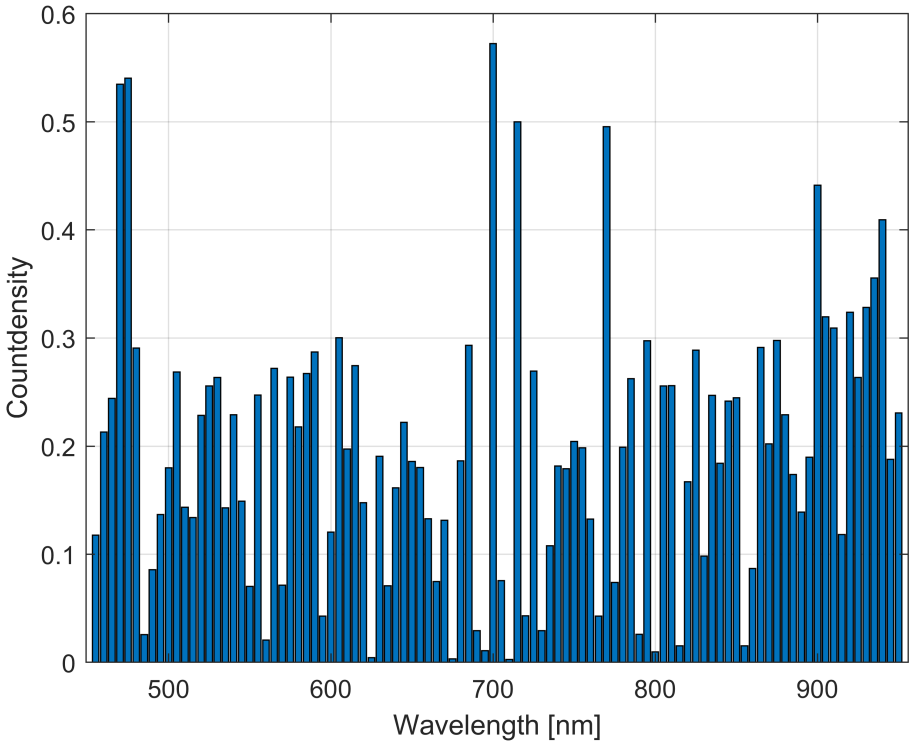


Figure 3-23: Region of relevant bands.

Once the histogram has been analyzed, the number of bands is chosen for each dataset using the overall accuracy obtained by ***(SVM)***, ***(RF)***, and ***(NN)***. These results will be plotted on a coordinate scale depicting the classification percentage, as shown in Figure **3-24**.

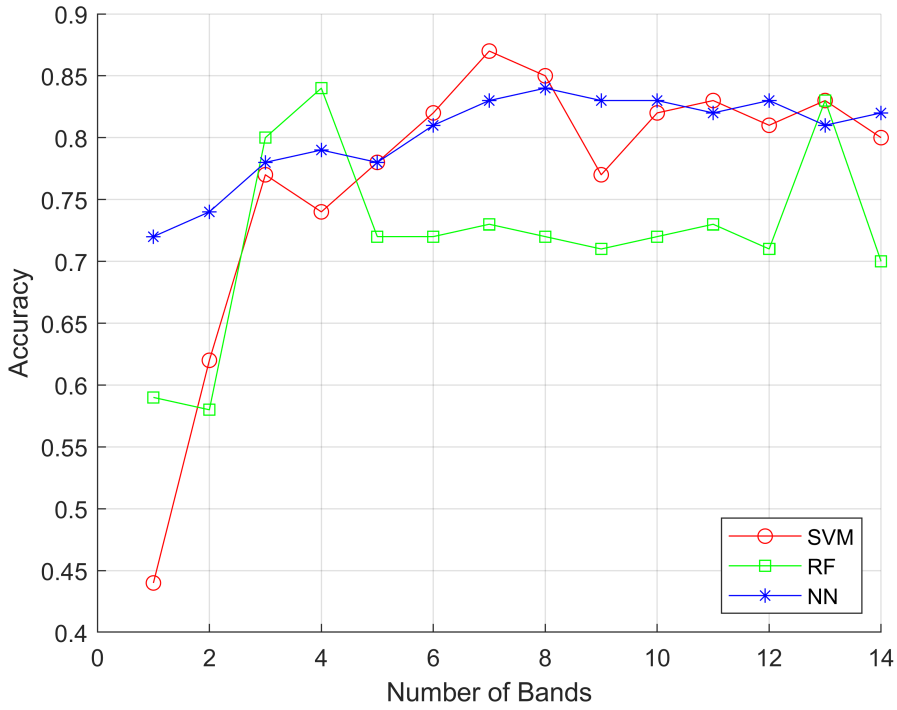


Figure 3-24: Classification overall accuracy per classifier and number of bands.

3.6.3. Feature Selection Techniques

A new representation space is built using the relevant bands and the vegetation indexes (see Table **2-1**) for spectral signature - for multispectral images, spatial-derived features are included. In this new representation space, irrelevant features may persist in the data, and these need to be removed. In addition, many mining algorithms must work better with fewy features or attributes. Consequently, this work explores feature selection techniques before using the classification algorithm. The main objectives of feature selection are to improve model performance and to provide faster and more cost-effective models while avoiding overfitting or overtraining [146]. Three techniques were chosen and evaluated to enhance the process of selecting features based on their effectiveness and suitability for use in data classification.

The first technique used for feature selection was the rank features for classification using minimum redundancy maximum relevance ***(MRMR)*** [147]. This algorithm determines an

optimal feature set that is mutually and maximally dissimilar and can effectively represent the classification variable. It minimizes edundancy and maximizes a feature set's relevance using the variables' mutual information. This information between two variables measures how much uncertainty can be reduced by knowing the other variable [148].

A second feature selection technique explored was the univariate feature ranking for regression using F-Tests ***(FSRF)***. This technique uses statistical data to determine the importance of each predictor; a small p-value of the individual test indicates that the corresponding predictor is relevant [149]. Previously, this approach was applied in identification and classification processes with a low computational cost and provided pertinent information [150].

Finally, this work used the univariate feature ranking for classification using chi-square tests ***(FSCC)***. The chi-square test helps solve feature selection problems by testing the relationship between different features. It is typically used when the feature and target variables are both categorical. The chi-square test measures the degree of association between two categorical variables and can help determine how strongly they are related.

A bar graph visually represents the importance score. The results are normalized and standardized to present a relevant diagram according to the methods previously shown in Figure **3-25**. After analyzing the available data, a list of relevant characteristics is obtained and optimized for classification.

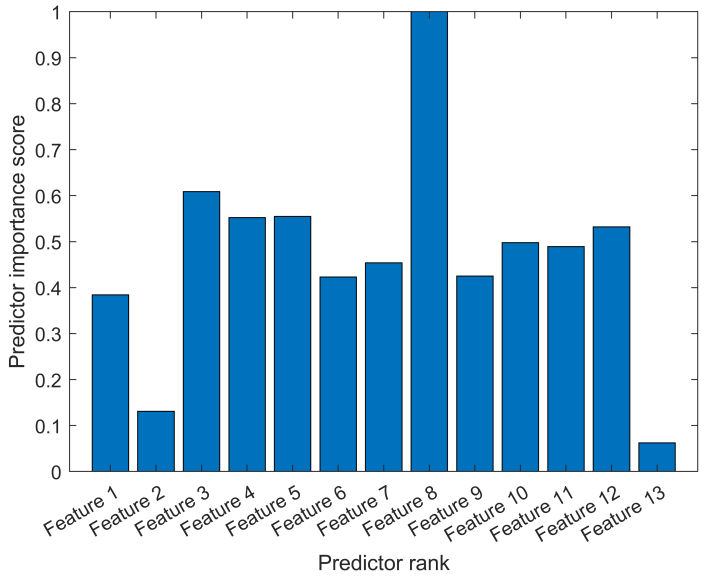


Figure 3-25: Weighted Feature Score.

3.6.4. Spatial Features

Texture features are employed to use spatial information in images effectively. These features have traditionally been used in spatial-spectral classification processes. We will incorporate this field's typical spatial texture features, such as energy, entropy, contrast, and correlation [151].

The energy was computed using equation 3-1 where P represents the grayscale level, and u and v represent the coordinates. The Energy or Second Angular Momentum measures image homogeneity. Its value is high when an image is uniform or has similar pixels. It is likewise sensitive to edges and noise [151].

$$\textbf{Energy} = \sum_{u}\sum_{v}(P_{uv})^2 \tag{3-1}$$

On other hand, the entropy is related to how much information is required to compress an image [151]. It quantifies the loss of information or message in a transmitted signal and the amount of information contained within an image to compute the entropy was used the following Equation 3-2.

$$\textbf{Entropy} = -\sum_{u}\sum_{v}(P_{uv})\log\left(P_{uv}\right) \tag{3-2}$$

The contrast feature measures the change of intensity between a pixel and its neighborhood, as defined by the window size in the image (Equation 3-3). [152]

$$\textbf{Contrast} = \sum_{u}\sum_{v}(P_{uv})(u-v)^2 \tag{3-3}$$

Finally, the correlation computed by Equation 3-4 refers to the degree of linear dependence between the gray levels of adjacent pixels [152]. Digital image correlation, on the other hand, is an optical technique that quantifies image variations by utilizing image tracking and registration methods.

$$\textbf{Correlation} = \frac{\sum_u \sum_v (uv)(P_{uv}) - (\mu_u - \mu_v)}{\sigma_u \sigma_v} \tag{3-4}$$

We standardized the image dataset by normalizing the results using a 3 x 3-pixel sliding window. This was done to assist in the classification of data. It was excluded if a neighboring pixel had no information or did not belong to a vegetation signature.

4 Spectral signature analysis

This chapter presents stress detection using only spectral signatures, according to the methodology of Figure 3-1. First, the spectral characterization of each crop is presented, followed by the selection of discriminant bands, the feature extraction process combining bands as vegetation indices, and finally, stress detection using classification methods

4.1. Signature characterization

The signature characterization was performed by establishing spectral bands and ranges where a particular behavior is expected for vegetation. For instance, the maximum absorption band for the red and blue regions, the green area peak, the Red Edge's beginning and end, and the beginning of the near-infrared (NIR) plateau. These bands and ranges can change according to the vegetation type, genotype, state of health, and other factors [153]. The first derivative of the average of captured signatures from healthy plants was used to establish the characteristic bands and ranges. Figure **4-1** presents the first derivative from the average signature obtained from each crop. The last local minimum occurs in blue range wavelengths (the blue point in the figure) is one of the maximum absorption points of chlorophyll. Similarly, the last local minimum of the visible range is defined as the maximum absorption point of chlorophyll in the red range. The first local maximum after the Red Edge is the NIR plateau, and the green peak is the local maximum obtained at wavelengths in the visible range. Due to the lack of a restrictive mathematical formalism for identifying the Red Edge's start and end points, an algorithm facilitated the detection of the maximum slope in the first derivative of the spectral signature's visible range from 450 nm to 680 nm. Figure **4-1** considers the wavelength at which this magnitude is exceeded as the starting point of the Red Edge and the wavelength where the derivative returns to less than that magnitude at the endpoint.

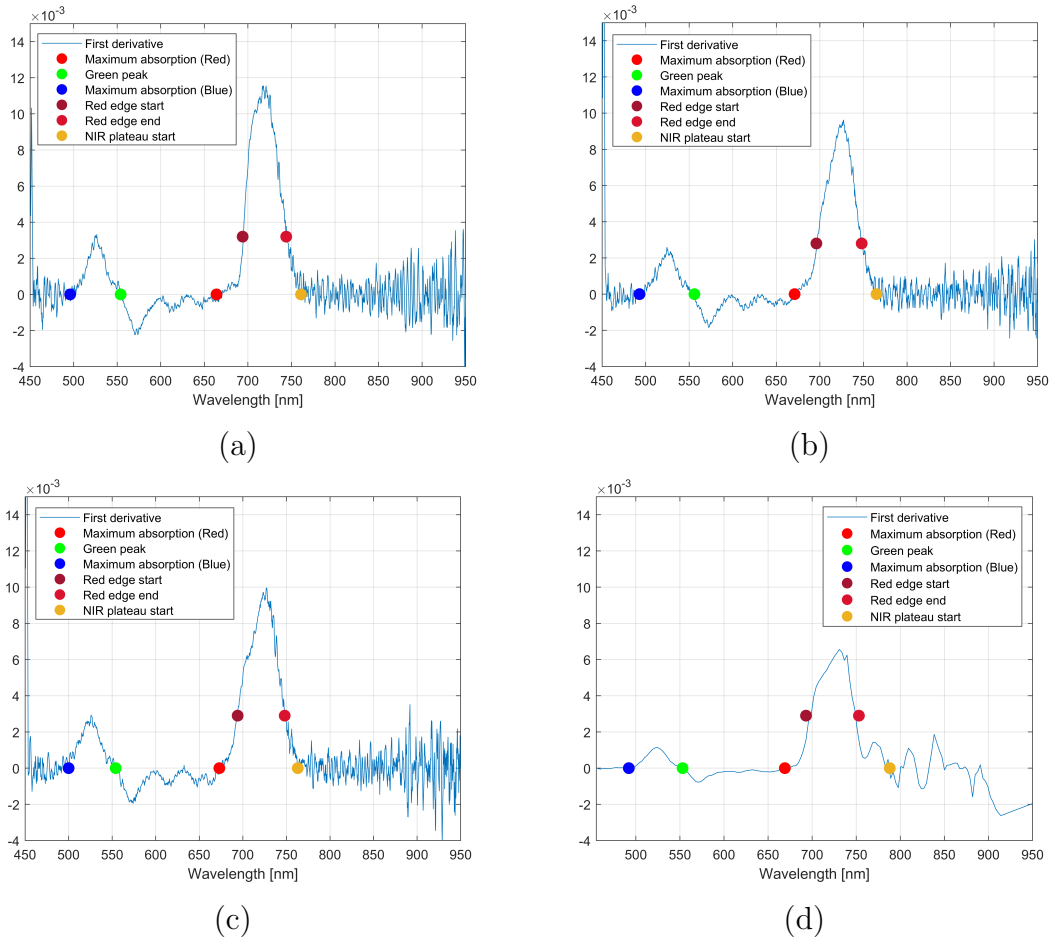


Figure 4-1: First derivative of average signature for healthy (control) plants of (a) Bean Cerete, (b) Maize Cerete, (c) Maize Rionegro, and (d) Avocado Rionegro.

Figure **4-2** highlights the identified local minima and maxima over the mean spectral signatures for each crop, and Table **4-1** summarizes these characteristic points. Near the points of maximum chlorophyll absorption, a considerable standard deviation, about 50 % of the obtained reflectance value, is deteced. Table **4-1** listed the characteristic values of the wavelengths, which were similar among the crops. However, the maize crops evidenced an increment of the reflectance compared to the spectral response of beans and avocados. These findings suggest that these crops can be accurately distinguished. In addition, the similarity observed between Figures **4-2** (b) and (c) is noteworthy. Despite differences in time, location, and genotype, the mean spectral signature of the optimal treatments displays similar amplitude, waveform, and range for the maize crops. This indicates that the data collection technique employed can produce reproducible information. Despite that, the data collected in Rionegro showed a lower standard deviation. In the avocado study, there is a noticeable change in the spectral signature, with a decay in the plateau of the near-infrared around 900 nm and irregular curves at the end of the Red Edge. The amplitude values are also significantly lower in the

signature, which may indicate changes in the calibration (Table **4-1**). The spectral signature was created using multispectral cameras, but there were some differences due to the changes in light caused by slight cloudiness during the picture-taking process. As a result, a smoothed spectral signature can be observed, which is evident in its derivative Figure **4-1**.

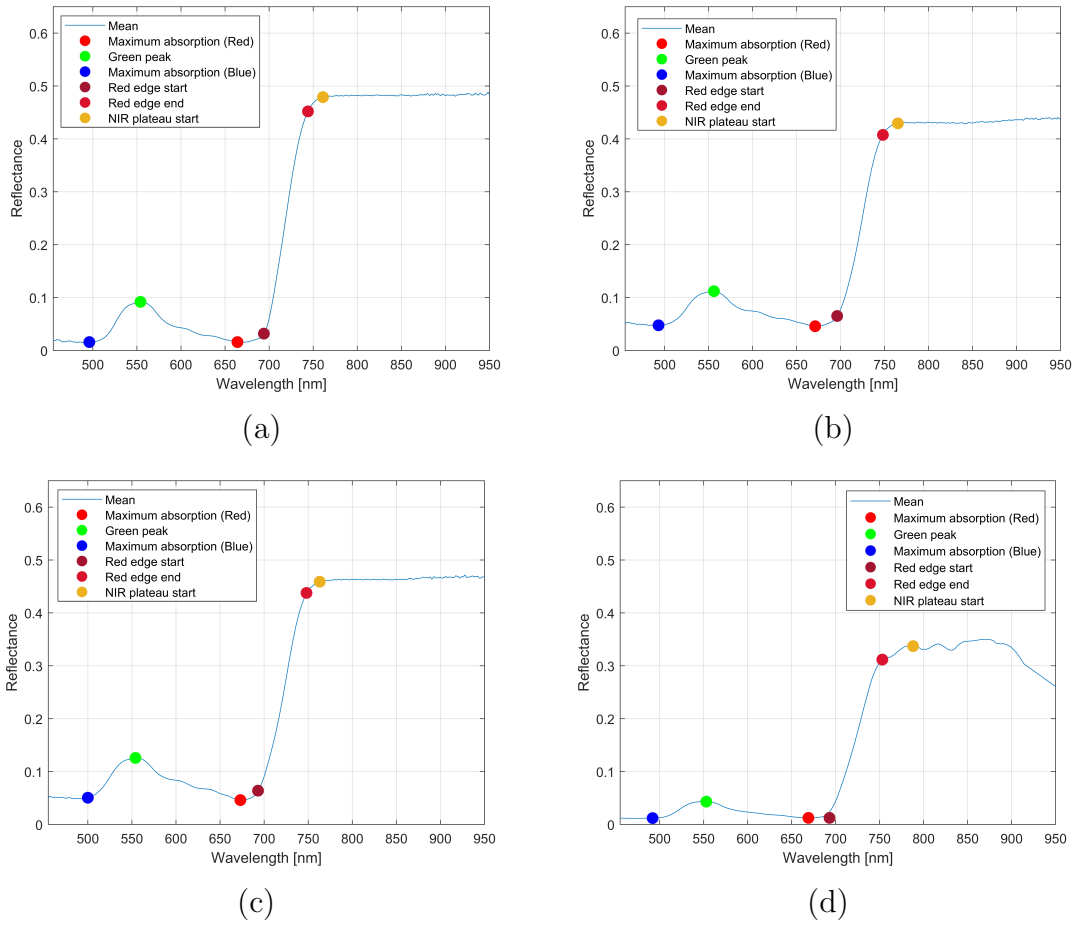


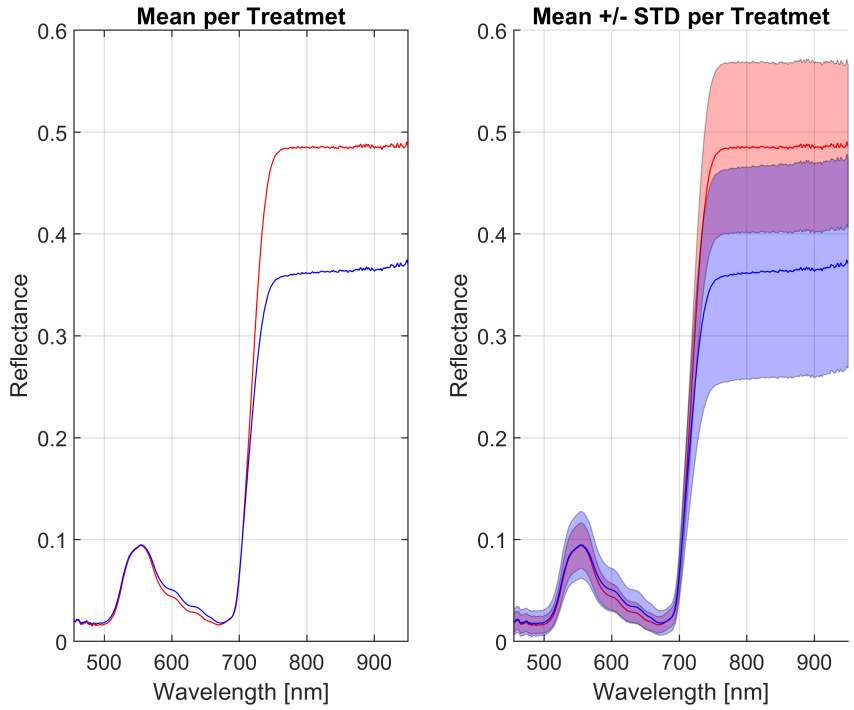
Figure 4-2: Representative points over average signature for healthy (control) plants of (a) Bean Cerete, (b) Maize Cerete, (c) Maize Rionegro, and (d) Avocado Rionegro.

Feature	Wavelengths [nm]	Reflectance	STD
Bean Cerete			
MAC Blue	496	$16.11e-3$	$7.56e-3$
Green Peak	554	$92.54e-3$	$21.94e-3$
MAC Red	664	$16.88e-3$	$8.58e-3$
Red Edge Change	[694 744]	-	-
NIR Plateau	[761 950]	-	-
Maize Cerete			
MAC Blue	493	$47.74e-3$	$14.22e-3$
Green Peak	556	$112.29e-3$	$23.26e-3$
MAC Red	671	$45.98e-3$	$13.35e-3$
Red Edge Change	[696 748]	-	-
NIR Plateau	[765 950]	-	-
Maize Rionegro			
MAC Blue	500	$50.68e-3$	$12.07e-3$
Green Peak	554	$125.93e-3$	$22.75e-3$
MAC Red	673	$46.19e-3$	$13.35e-3$
Red Edge Change	[693 748]	-	-
NIR Plateau	[763 950]	-	-
Avocado Rionegro			
MAC Blue	492	$12.16e-3$	$3.53e-3$
Green Peak	553	$43.36e-3$	$14.11e-3$
MAC Red	669	$12.60e-3$	$3.31e-3$
Red Edge Change	[693 753]	-	-
NIR Plateau	[788 900]	-	-

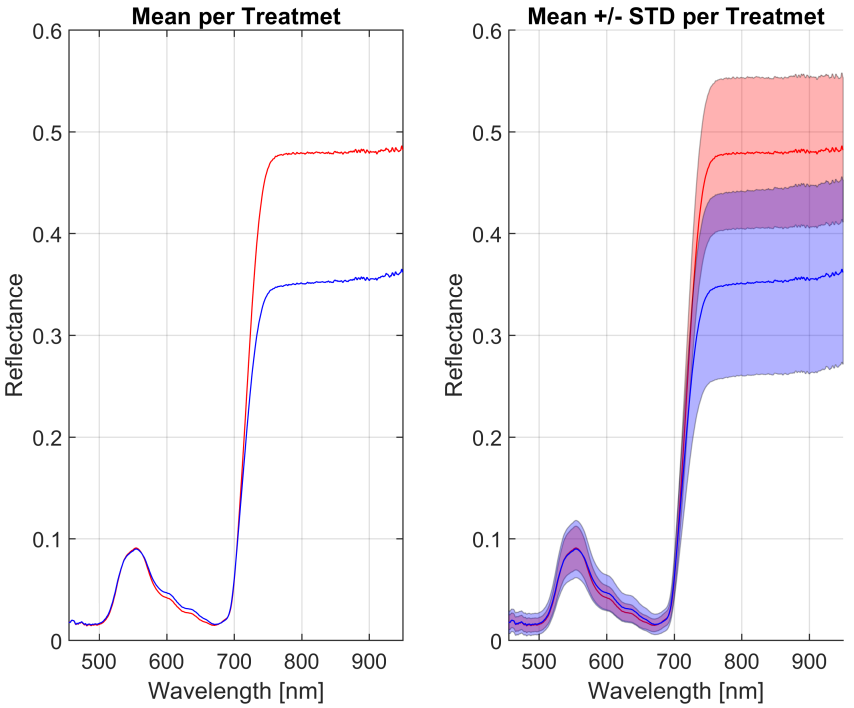
Table 4-1: Representative bands and spectral regions for Bean, Maize, and Avocado crops. (MAC: Maximum absorption of light by chlorophyll) Representative bands and spectral regions for Bean, Maize, and Avocado crops. (MAC: Maximum absorption of light by chlorophyll).

Finally, to understand the spectral behavior of each crop, mean and standard deviation signatures are calculated for each dataset. Figure **4-3** shows the mean and standard deviation (STD) for common bean crops for each genotype and treatment. A visual difference can be appreciated between both treatments, mainly at the end of Red Edge and the wavelengths corresponding to the red color, where the effect of stress generated by phosphorus deficiency in this variety is appreciated in the non-visible range but not a significant impact on the pigmentation of the leaf. Although the difference in the mean values presented between the treatments of genotype 2 (b) was lower compared to genotype 1 (a), an evident segmentation prevailed between the treatments with the same characteristics in both genotypes at the end of Red Edge. The wavelengths corresponding to the red color allowed us to conclude that the effect of phosphorus deficiency is equivalent to the genotypes evaluated in the experiment.

Figure **4-4** illustrates the mean and standard deviation behavior for maize Cerete under nutritional stress. The spectral signature across the entire wavelength spectrum analyzed for both genotypes showed a slight attenuation. However, despite this effect, the treatments displayed a low level of dispersion, except in the near-infrared range, where a slightly wider spread between treatments is observed. Unfortunately, this statistical dispersion poses significant challenges in differentiating the treatments from the currently available data.

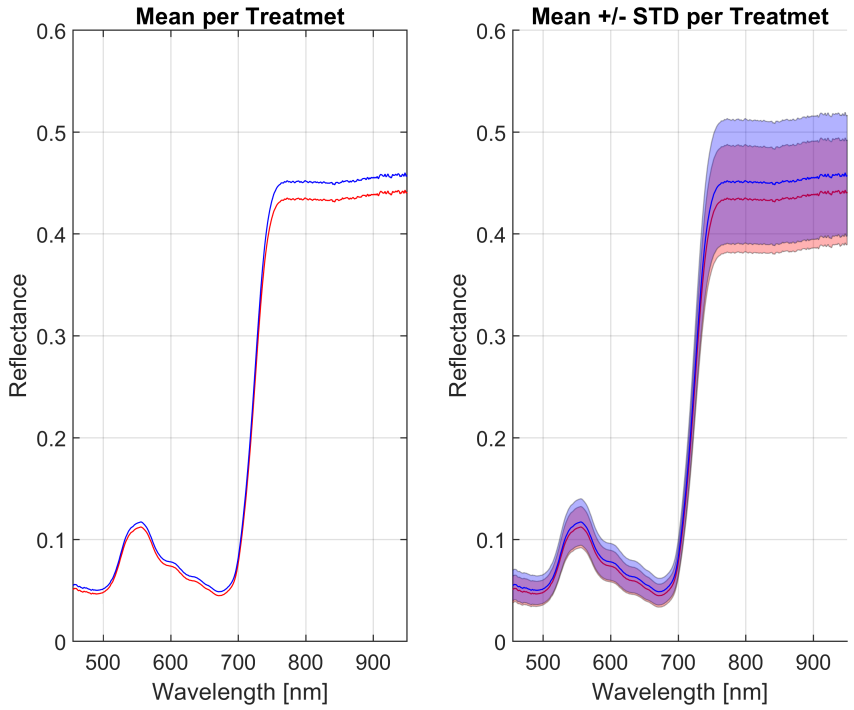


(a)

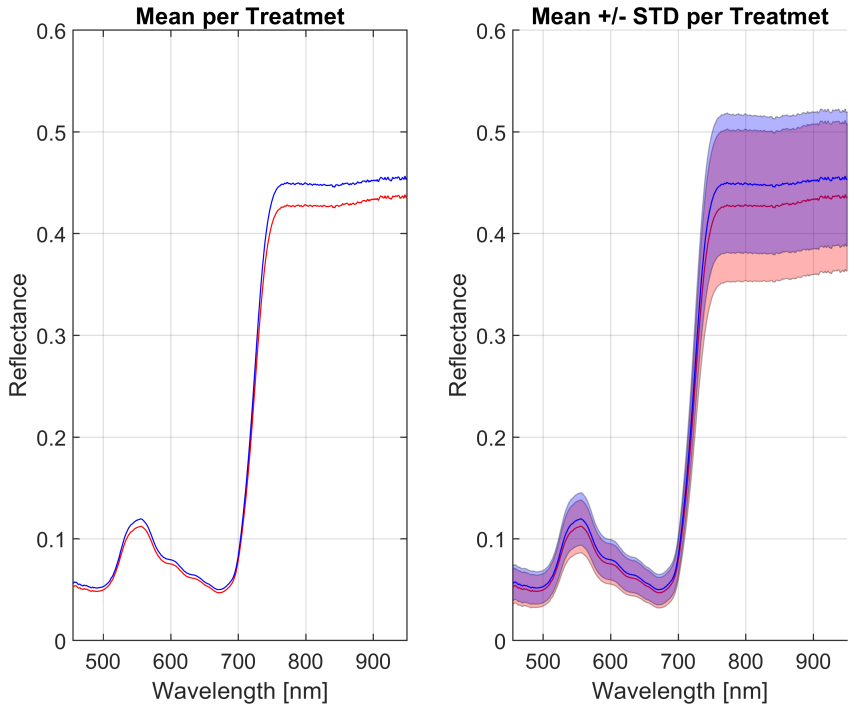


(b)

Figure 4-3: Comparison of the average spectral signatures for treatments (T1:Red, T2:Blue) on Bean Cerete (a) Genotype 1 and (b) Genotype 2.



(a)



(b)

Figure 4-4: Comparison of the average spectral signatures for treatments (T1:Red, T2:Blue) on Maize Cerete (a) Genotype 1 and (b) Genotype 2.

For Rionegro maize, of the ten genotypes studied, the discussion will focus on three genotypes that showed different characteristics in the classification of stress levels. The analysis was performed using all treatment levels, as well as a comparison using the extreme treatments, assimilating a binary comparison on the presence of the stress factor. Figure **4-5** illustrates the average signature of genotype one treatment, where treatment two significantly reduced amplitude across the entire spectrum. Meanwhile, treatments 1, 2, and 3 overlap in the visible range but differ in the near-infrared range. The signatures for genotype 3 show a higher amplitude in treatment 1 in the visible range Figure **4-6**. Despite overlapping standard deviation data, there is a different order of signatures from high to low amplitude compared to the near-infrared region. For genotype 10, the signatures'mean and standard deviation values are visually displayed with low dispersion and overlapping statistical values. The treatments with the lowest amplitude in the visible range have the highest amplitude in the near-infrared range, as shown in Figure **4-7**.

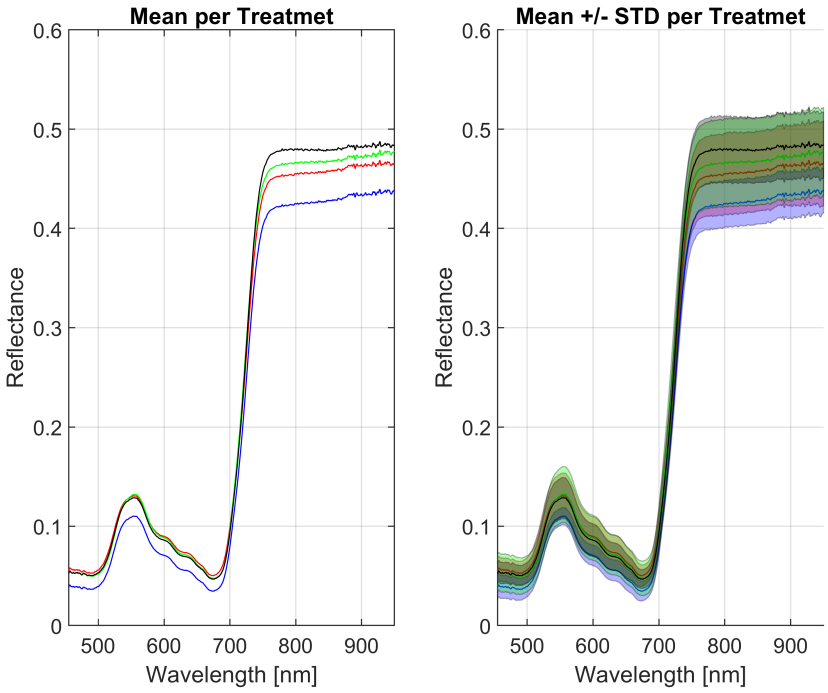


Figure 4-5: Comparison of average spectral signatures for treatments (T1:Red, T2:Blue, T3:Green, T4:Black.) on Maize Rionegro Genotype 1.

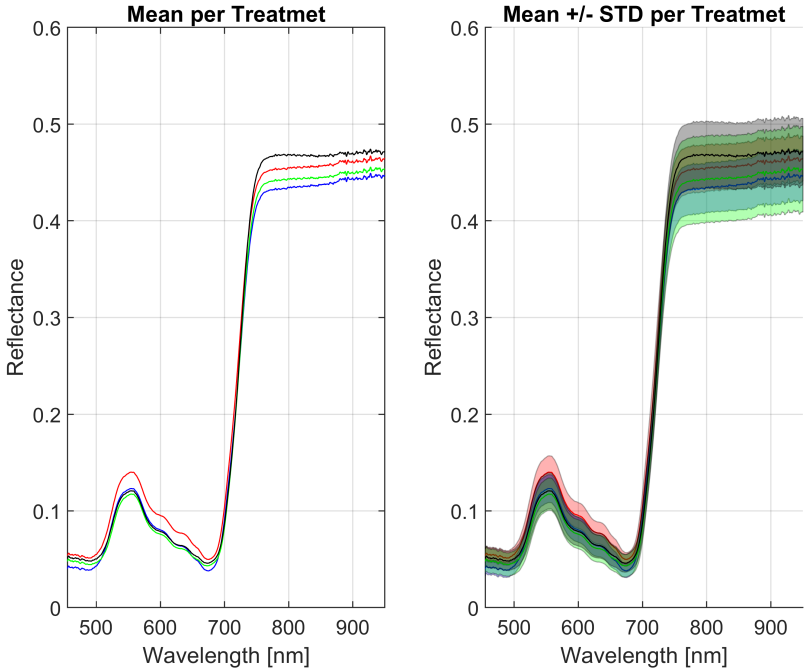


Figure 4-6: Comparison of average spectral signatures for each treatments (T1:Red , T2:Blue, T3:Green, T4:Black.) on Maize Rionegro Genotype 3.

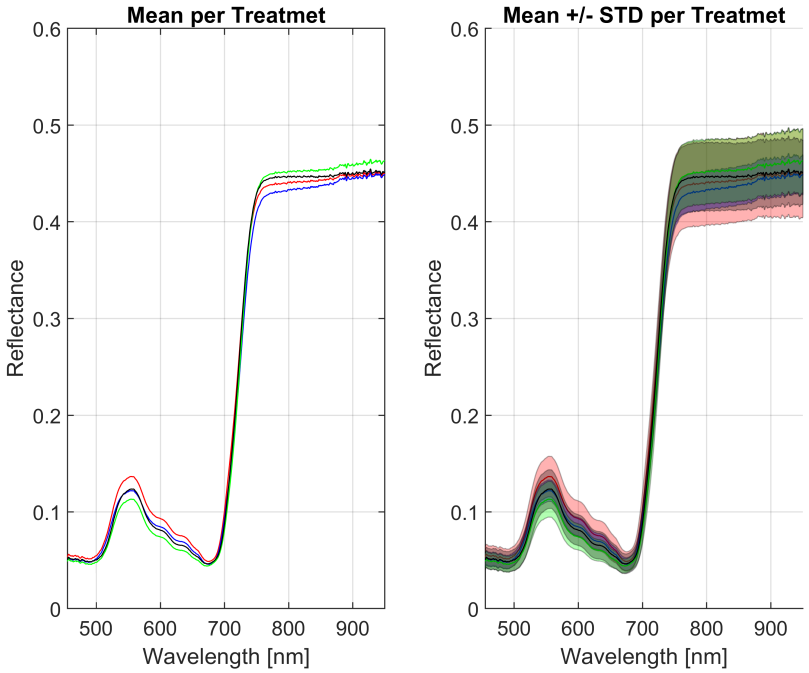
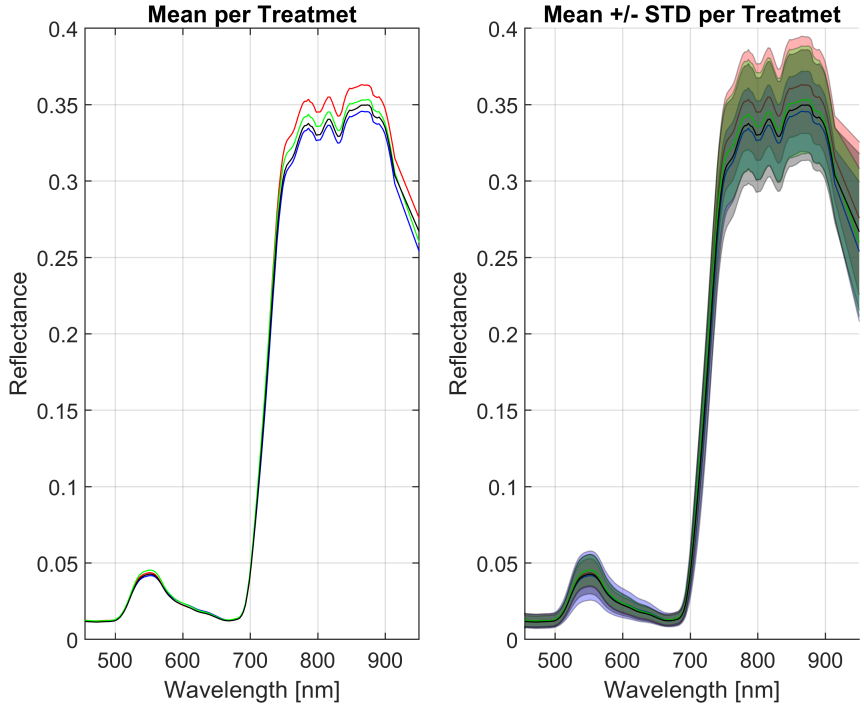
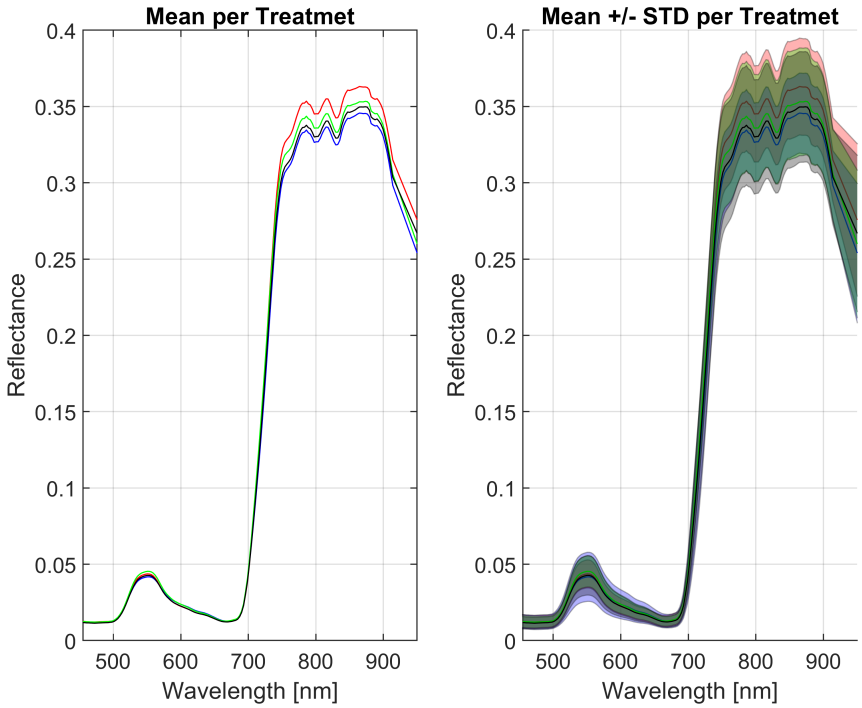


Figure 4-7: Comparison of average spectral signatures for each treatments (T1:Red , T2:Blue, T3:Green, T4:Black.) on Maize Rionegro Genotype 10.

Figure **4-8** display the average signatures for the four treatments in genotypes 1 and 2, the data shows low dispersion, which might be due to the constant change of leaves in the seedlings, making it challenging to identify the effect of water stress on the leaves. Since this test is more demanding due to the low variance of the signatures before the effects of stress, the study decided to include the response obtained by evaluating only the treatment of higher water deficit and overdosing of water, such as treatment 1 and treatment 4. Figure **4-9** demonstrates the outcome of this decision and shows if the proposed procedure produces similar results. There was a notable difference in amplitude values in both genotypes in the spectrum, particularly in genotype 2. There was a clear separation between the signatures in the visible spectrum.

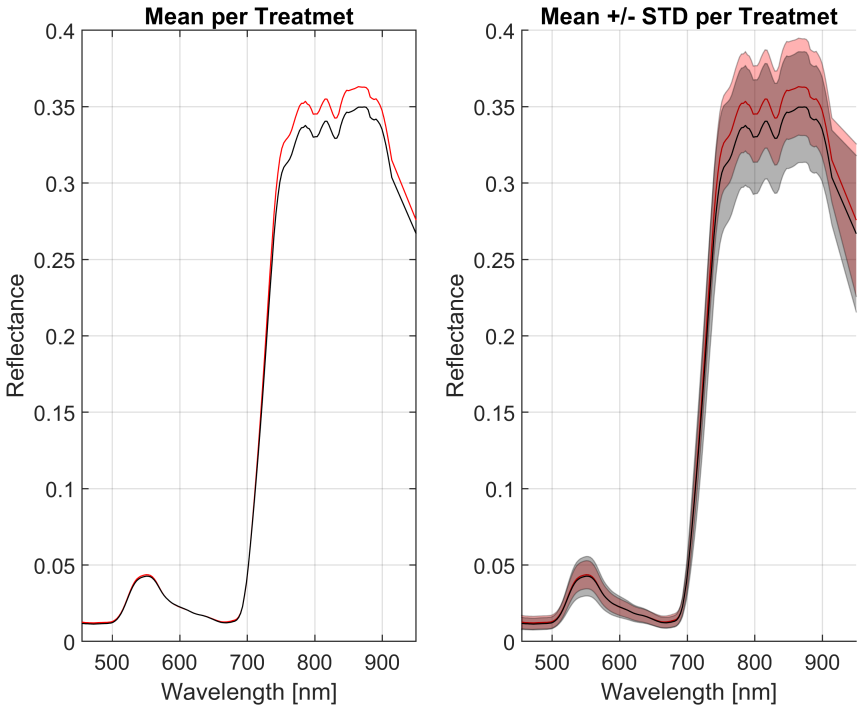


(a)

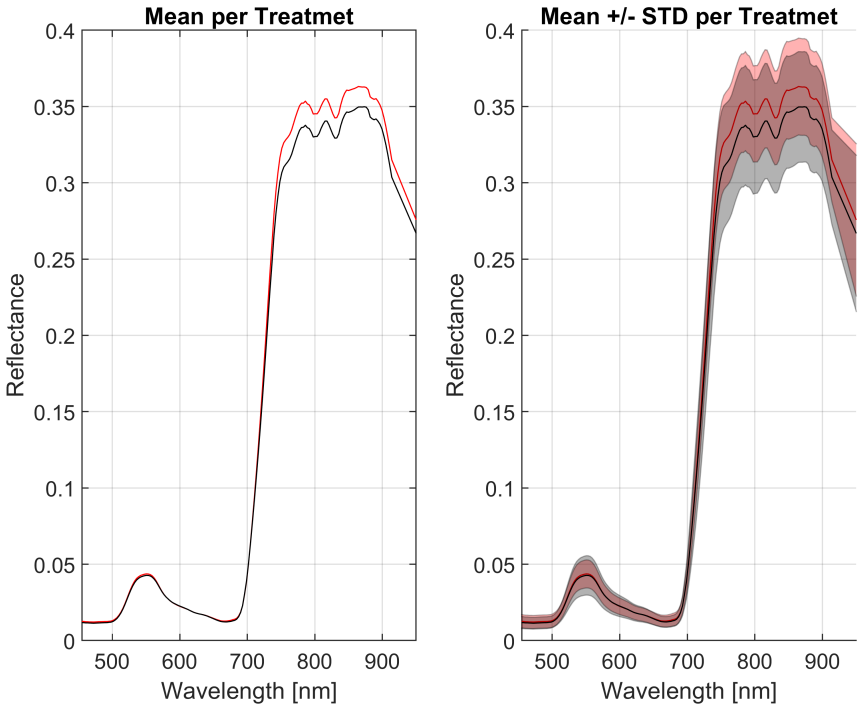


(b)

Figure 4-8: Compares the average spectral signatures for treatments (T1:Red, T2:Blue, T3:Green, T4:Black) on Avocado Rionegro (a) Genotype 1, (b) Genotype 2.



(a)



(b)

Figure 4-9: Compares the average spectral signatures for treatments (T1:Red, T4:Black) on Avocado Rionegro (a) Genotype 1, (b) Genotype 2.

4.2. Band Selection

Figure **4-10** illustrates the wavelength ranges of zones of interest. These ranges were determined based on the band selection criteria outlined in Table **4-2**. The selection process involves identifying wavelength regions with the highest concentration of relevant bands. These regions were categorized as appropriate if they exceeded 40 % of the maximum normalized value, had a width no greater than 30 nm, and were tabulated based on their impact. For the Bean Cerete (a), the wavelength ranges are short, and the regions of most significant impact are found in the Red Edge, NIR plateau start, and the 895-950 nm range, where there is a substantial increase in the noise level. For Maize Cerete (b), bands with significant relevance are generally observed in 720 and 770 nm wavelengths, with a low dispersion of relevant bands in these regions. The chosen bands present a low range of wavelengths, with most of the pertinent information concentrated in the 930-950 nm range. It's important to note that the regions corresponding to the green color don't add any discriminating information. Regarding the Maize Rionegro (c), the bands of interest are highly variable and distributed throughout the spectrum studied. It's worth mentioning that this crop had the greatest participation in the zone close to the green color, and there is a continuity of potentially relevant information in the range of the Red Edge. For Avocado crop (d), similarities are observed in both distributions for evaluating all treatments and in the case of extreme treatments. It's noteworthy that for this crop, a different behavior was presented compared to others, as there is a significant concentration of bands of interest after 700 nm.

Datasets	R1[nm]	R2[nm]	R3[nm]	R4[nm]	R5[nm]	R6[nm]	R7[nm]
Bean Cerete	770-775	700-720	470-485	895-950	550-555	—	—
Maize Cerete	930-950	710-735	460-470	685-690	785-790	—	—
Maize Rionegro	780-800	750-770	450-465	705-725	940-950	620-645	555-560
Avocado All Treatments	880-915	805-840	740-770	715-725	520-540	—	—
Avocado T1 & T4	880-915	820-865	760-775	720-725	540-545	—	—

Table 4-2: Spectral regions of interest for Bean, Maize, and Avocado crops.

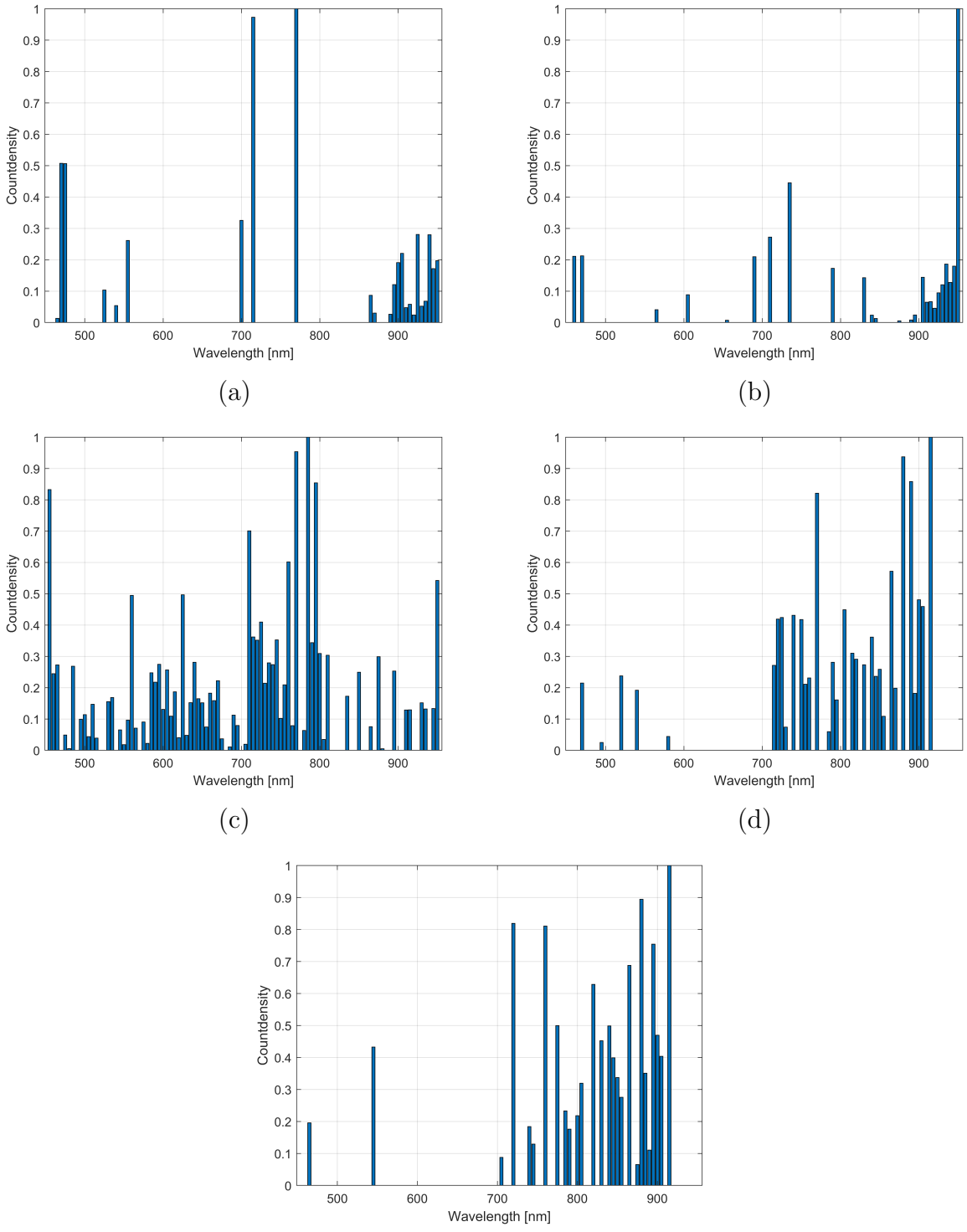
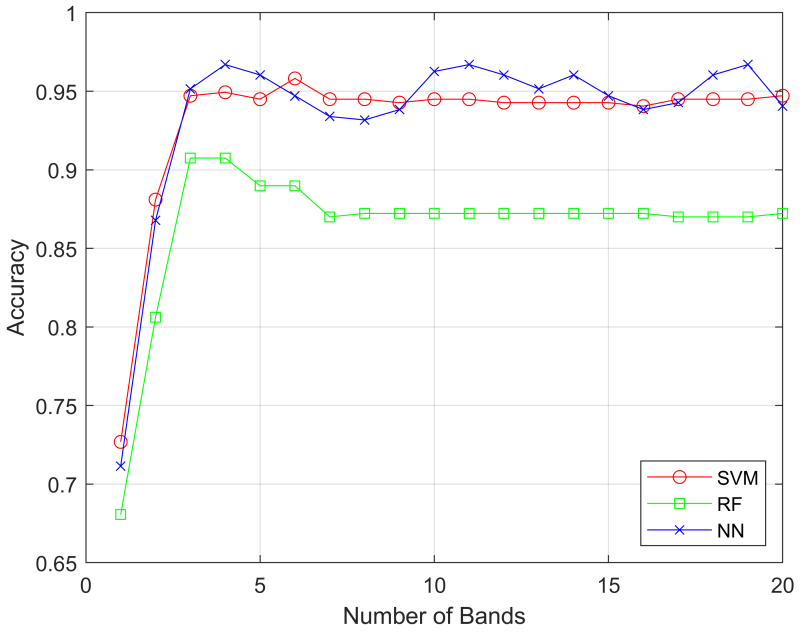


Figure 4-10: Spectral regions of interest according to band selection algorithms of (a) Bean Cerete, (b) Maize Cerete, (c) Maize Rionegro, (d) Avocado Rionegro and (d) Avocado Rionegro extreme treatments.

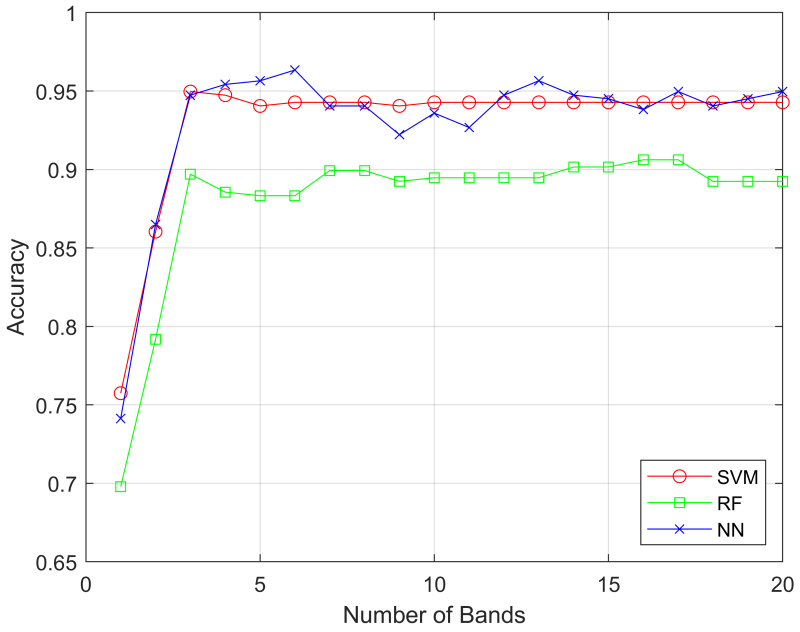
The supervised classification of the bean and corn crops used 30 % of the available data for training to evaluate the impact of increasing the number of bands. Due to the complexity of the avocado trial to stress levels, training data increased to 40 %. In the Bean Cerete Figure **4-11** , the classification accuracy was high, indicating that the data was easily classifiable. Genotype 1 (a) exhibited a local maximum with four bands, while genotype 2 required seven bands to reach a stabilized accuracy value. The results show that the classification process is more accurate for genotype 1 (a), while genotype 2 requires more bands. Nonetheless, both genotypes achieved an appropriate accuracy rate.

Figure **4-12** shows the results for Maize Cerete. For genotype 1, it was a clear inflection point in accuracy when 13 bands were used; as well as, it genotype shows slightly more favorable behavior with a lower number of bands. RF classifier has presented an unfavorable behavior for both genotypes. Unlike the previous crop, we can see a low accuracy, which could be due to the soft effect of stress on this kind of crop. The results suggest that stress has a minimal impact on the spectral signatures in the visible spectrum. However, other effects that are not discernible in the evaluated spectrum may occur at the biophysical level. These effects may have an impact on plant growth or yield. For Maize Rionegro, Figure **4-13** illustrates a classification process that displays a significant and constant increase in accuracy around four bands to genotype 3 (b), suggesting a minimum number of bands to consider, with a gradual increase in accuracy as more bands are added, eventually achieving results close to 80 % accuracy. The difference between the ***SVM*** and ***NN*** classifiers was less pronounced than in genotype 1 (a). For genotype 10 (c), higher accuracy was achieved compared to previous cases despite the treatments not having the lowest dispersion visually. Interestingly, the ***NN*** classifier performed better for this dataset, although all genotypes had a persistent oscillation.

Figures **4-14** and **4-15** indicate that the classification accuracy for Avocado dataset with all treatments is low, even using a 40 % of the labeled data for training; these is due to the low variability of the spectral signatures obtained. However, the information from the extreme treatments, which are highly differentiable even with a restricted number of bands, results in Figure **4-15** showing over 70 % accuracy. Interestingly, compared to the previous case studies, a significant difference emerged when comparing the performance of the random forest classifier in the analysis with all and extreme treatments.

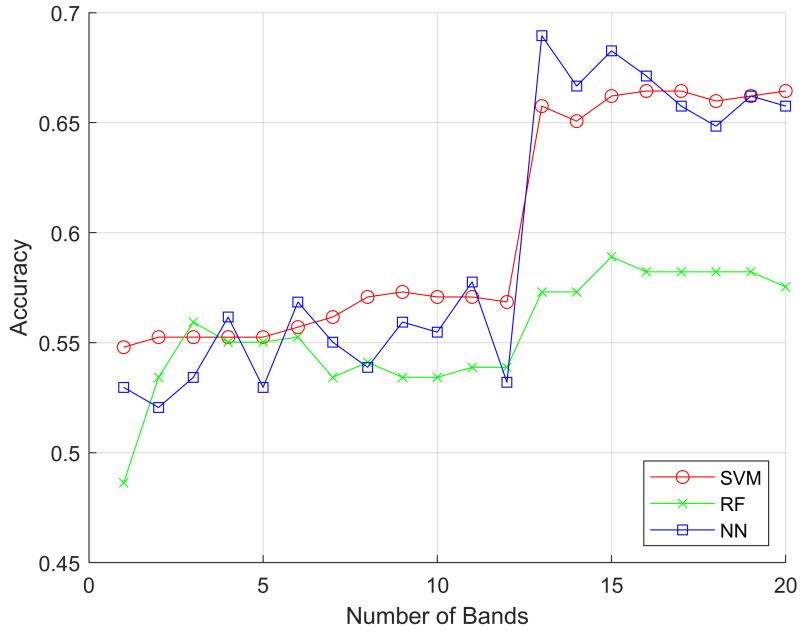


(a)

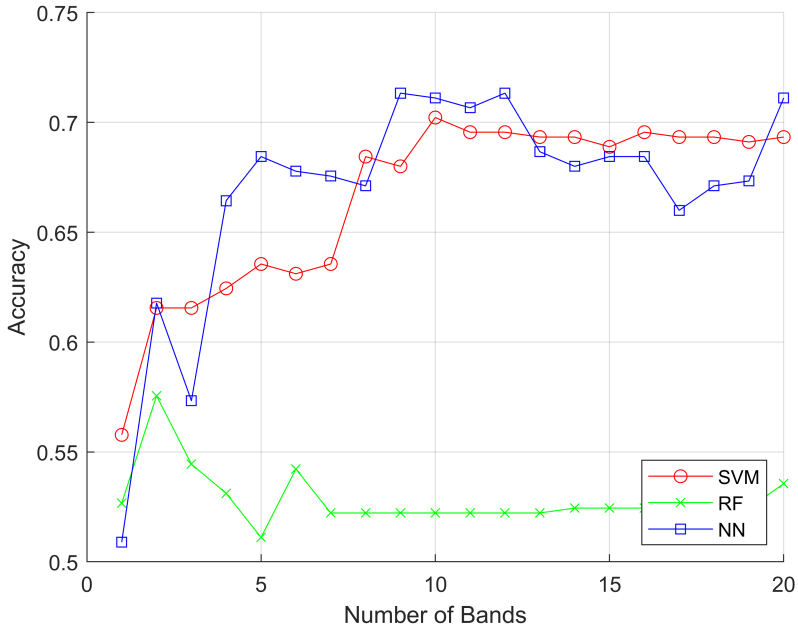


(b)

Figure 4-11: Overall classification accuracy according to number of bands for Bean Cerete (a) Genotype 1, (b) Genotype 2.

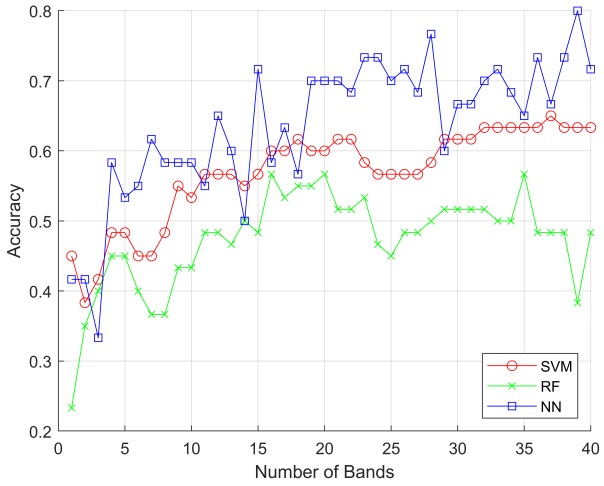


(a)

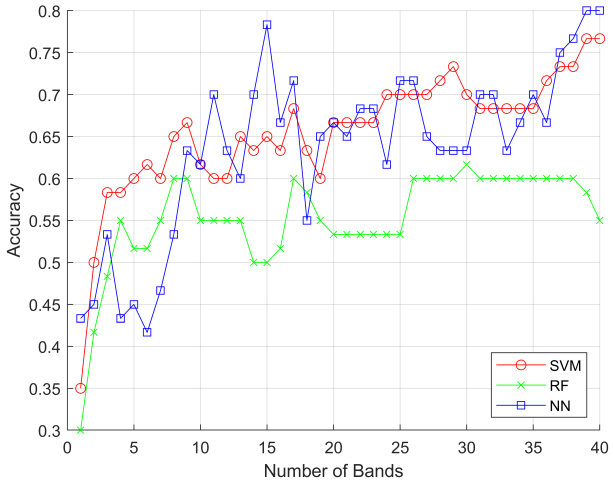


(b)

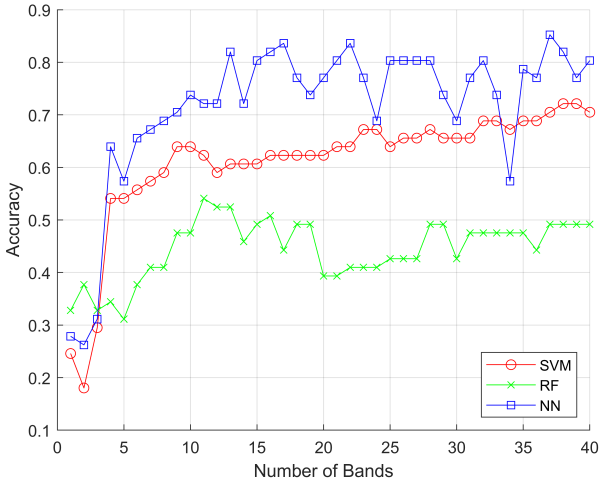
Figure 4-12: Overall classification accuracy according to number of bands for Maize Cerete (a) Genotype 1, (b) Genotype 2.



(a)

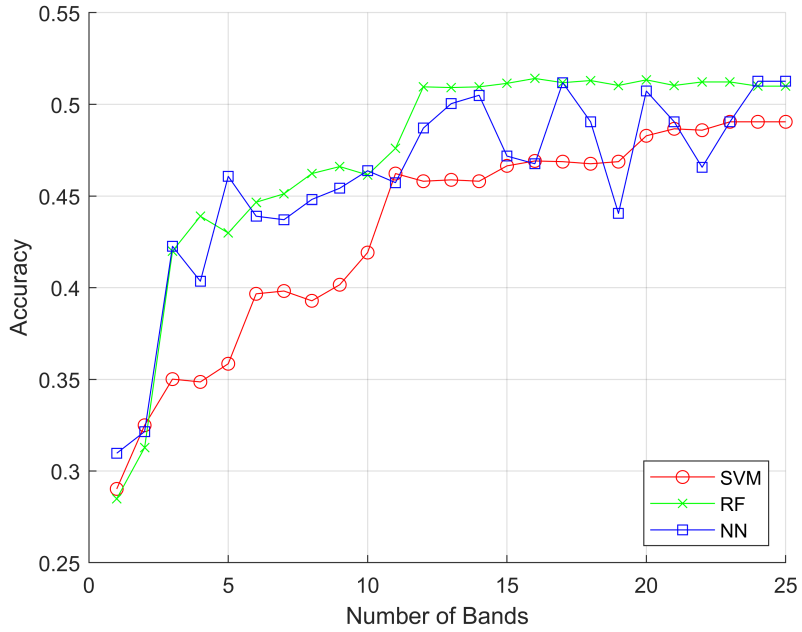


(b)

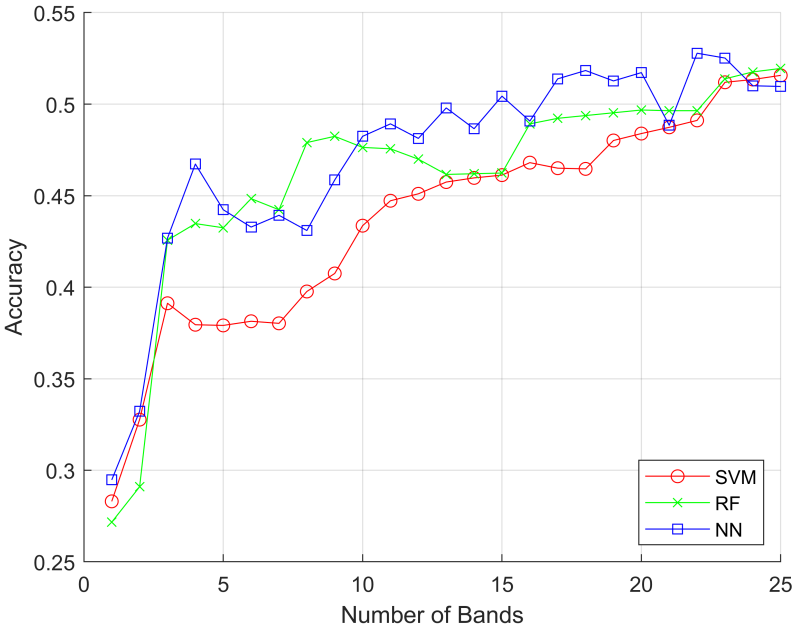


(c)

Figure 4-13: Overall classification accuracy according to number of bands for Maize Rionegro (a) Genotype 1, (b) Genotype 3 , (b) Genotype 10.

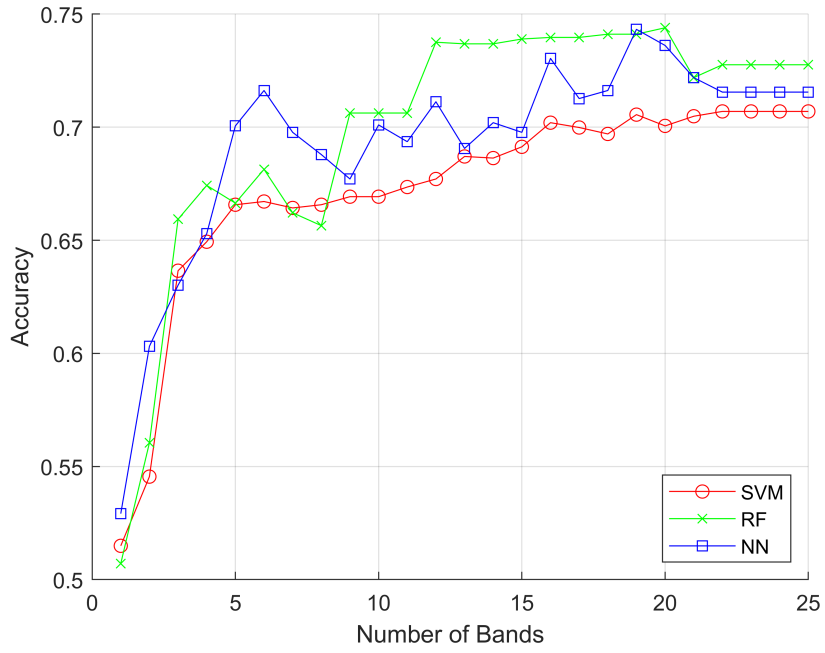


(a)

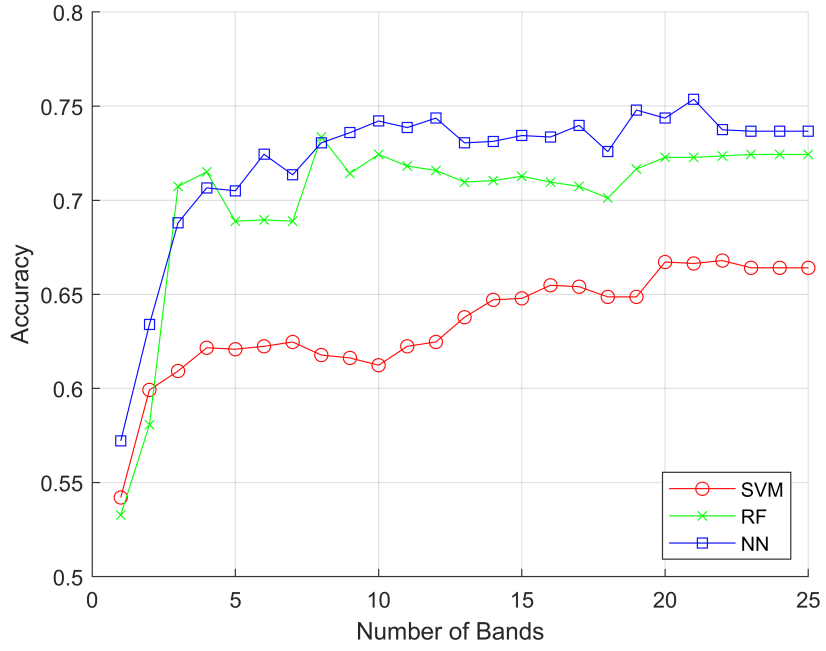


(b)

Figure 4-14: Overall classification accuracy according to number of bands for Avocado All Treatments (a) Genotype 1, (b) Genotype 2.



(a)



(b)

Figure 4-15: Overall classification accuracy according to number of bands for Avocado T1 & T4 (a) Genotype 1, (b) Genotype 2.

After selecting the number of bands, Table **4-3** was made to identify the corresponding wavelengths. Both genotypes showed satisfactory results for the Cerete bean. The relevant band of 551 nm was added for genotype 2, where the green peak is found. In the case of Cerete maiz, band 946 nm was the most relevant in both genotypes. It can be seen that the increase in the classification percentage over 10 % to ***NN*** and ***SVM*** of genotype 1 was due to the inclusion of band 563, a band very close to Green Peak (Table **4-3**). Thus, the band 946 is the most pertinent in both genotypes. Additionally, it is observed that the inclusion of band 563 nm, increased the classification accuracy.

Table **4-3** displays the bands chosen for the ten genotypes of Maize Rionegro. There is a higher concentration of bands in the 700 nm to 900 nm range, which includes the Red Edge and NIR plateau. This is quite different from the Maize Cerete, where the presence of bands near 950 nm had a greater impact. However, the selected bands are spread out across the entire spectrum that has been studied, with a relatively low number of bands in the green peak.

Table **4-3** presents the selected bands for avocado in both cases, using all treatments and only the extreme ones. Results showed a consistent concentration of influential bands between 710 and 913 nm. For genotype 1, 24 bands were considered relevant for all treatments, while the extreme cases could be represented with only 19 bands. For genotype 2, the number of essential bands differed, with 23 bands selected for all treatments and only 8 for the extreme cases. However, in the representation obtained in Table **4-3**, we can see a relationship among the selected bands for all and only extreme treatments.

Band [nm]	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	B13
Bean Cerete													
Genotype 1	766	714	472	943	-	-	-	-	-	-	-	-	-
Genotype 2	769	714	467	945	898	936	551	-	-	-	-	-	-
Maize Cerete													
Genotype 1	946	706	457	732	827	925	948	685	937	905	921	913	563
Genotype 2	946	730	467	785	902	685	942	604	934	931	-	-	-
Maize Rionegro													
Genotype 1	886	720	483	793	808	946	767	863	738	555	459	450	474
Genotype 2	927	474	946	783	798	767	450	847	751	731	706	460	808
Genotype 3	450	711	792	779	798	767	912	751	729	705	458	482	859
Genotype 4	751	713	456	783	802	888	858	767	792	723	556	450	927
Genotype 5	927	718	450	780	798	768	863	909	738	557	462	501	912
Genotype 6	783	710	450	785	798	768	941	847	751	731	705	459	863
Genotype 7	927	450	721	782	802	768	847	877	751	731	705	497	555
Genotype 8	784	573	751	779	805	767	847	879	450	731	705	461	456
Genotype 9	927	715	450	784	808	767	859	905	751	731	553	497	705
Genotype 10	722	927	765	784	806	946	450	847	738	704	555	458	767
Avocado All Treatments													
Genotype 1	876	724	913	899	768	887	469	835	745	710	538	803	716
Genotype 2	864	716	768	913	902	887	518	879	736	812	788	899	538
Avocado T1 & T4													
Genotype 1	915	756	716	913	876	893	844	902	463	803	864	541	826
Genotype 2	815	716	915	899	879	759	544	913	-	-	-	-	-

Table 4-3: Selected bands from spectral signatures for Bean, Maize; highlighted in cyan are the bands contained between 700 nm to 900 nm for Maize Rionegro, and Avocado crops; highlighted in green are the bands considered relevant to the avocado studies shared by all treatments with respect to the extreme treatments

4.3. Feature Selection

After selecting the appropriate wavelength bands, a new feature cube is created by integrating various vegetation indices, such as ***NDVI***, ***MTCI***, ***CI Green***, ***CI Rededge***, ***RVI***, ***NDRE***, ***CCCI***, and ***CRI***. These indices were chosen for their consistent relevance in the studied cases. Figure **4-16** displays the graphs of bean and maize Cerete crops, which show the significance of the predictors in a normalized manner. The graph demonstrates that for bean crop genotypes (a, b), the ***CI Green*** indicator is predominant, and vegetation index features are usually highly involved in the most relevant characteristics. On the other hand, for Maize Cerete (c, d), the relevant characteristics show behavior that is very distinct when comparing genotypes. This indicates that the bands can provide more relevant discriminant information in some specific cases than the vegetation indices.

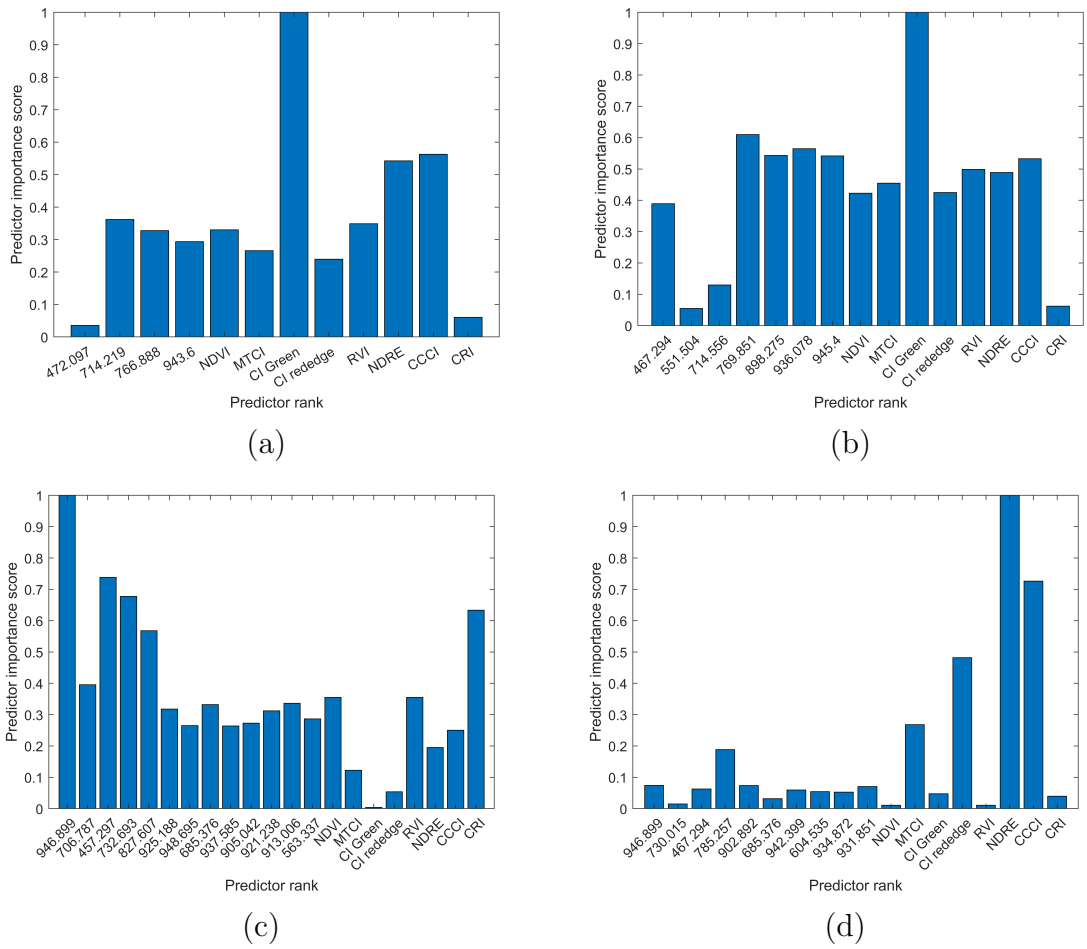


Figure 4-16: Rank for features extracted from spectral signatures of (a) Bean Cerete genotype 1, (b) Bean Cerete genotype 2, (c) Maize Rionegro genotype 1 , (d) Maize Rionegro genotype 2.

Table **4-4** presents the results of the feature selection conducted for different crops, each having eight main characteristics. The analysis of bean crops suggested that vegetation indices played a crucial role in determining the most essential features, with ***CI Green*** being the most significant impact factor. This observation indicates that the plant's green color could be an essential feature that needs to be considered while analyzing the crop. On the other hand, for genotype 2, individual bands made a more significant contribution, and predominant characteristics were highlighted in green. However, for Cerete maize, the results were more dispersed, making it difficult to define any relevant characteristics common to all studied genotypes.

Moving on to maize Rionegro, vegetation indices were highly present, but some cases indicated no significant genetic differences, such as in genotype 4 and genotype 6. This observation suggests that certain vegetation indices cannot be generalized as having predominant importance for all plant species because genetic variability results in specific characteristics. However, some characteristics were reported to have a high incidence in cyan color.

Finally, in the avocado study, the incidence of some bands and the prevalence of vegetation indices, particularly CI Red Edge and MTCI, were observed in all treatments. This indicates that each crop will have unique properties and physicochemical responses in the visible spectrum; these characteristics are highlighted in magenta.

A correlation matrix/plot has been created using a color scale to represent the correlation values. Each cell of the matrix will have a circle with a color corresponding to the lateral scale. The correlation was established between pairs of vegetation indices. The circle's size represents the correlation's magnitude, with larger circles indicating a more substantial relationship (positive or negative). Figure **5-7** illustrates some cases studied where it is indicated that their behavior is variable depending on the bands that make up the vegetation index, so it will also be susceptible to the spectral response of each genetic material to the levels of stress evaluated, also seen in cases such as ***MTCI*** - ***CI redegde*** and***NDRE*** -***CCI***, the latter being a relationship of dependence.

Features	F1	F2	F3	F4	F5	F6	F7	F8
Bean Cerete								
G1	***CI Green***	***CCCI***	***NDRE***	714nm	***RVI***	***NDVI***	766nm	943nm
G2	***CI Green***	769nm	936nm	945nm	898nm	***CCCI***	***RVI***	***NDRE***
Maize Cerete								
G1	946nm	457nm	732nm	***CRI***	827nm	706nm	***RVI***	***NDVI***
G2	***NDRE***	***CCCI***	***CI Rededge***	***MTCI***	785 nm	931nm	946nm	467nm
Maize Rionegro								
G1	***CCCI***	***NDRE***	751 nm	738nm	863nm	739nm	450 nm	793nm
G2	***NDRE***	***CCCI***	751nm	731nm	767nm	783nm	798nm	***MTCI***
G3	***CCCI***	***NDRE***	***CI Green***	705nm	711nm	***CRI***	912nm	450nm
G4	***CI Green***	450nm	459nm	***NDRE***	556nm	713nm	***MTCI***	792nm
G5	***CCCI***	***NDRE***	798nm	780nm	768nm	***CI Green***	459nm	***MTCI***
G6	***CI Green***	***NDRE***	***CCCI***	***NDVI***	705nm	710nm	459nm	452nm
G7	***NDRE***	***CCCI***	***MTCI***	751nm	768nm	***CRI***	721nm	877nm
G8	***NDRE***	***CCCI***	450nm	779nm	847nm	767nm	751nm	805nm
G9	***NDRE***	***CCCI***	905nm	784nm	767nm	808nm	***CI Green***	450nm
G10	***CCCI***	***NDRE***	704nm	555nm	***CRI***	***CI Green***	765nm	806nm
Avocado All Treatments								
G1	***CI Rededge***	***MTCI***	724nm	736nm	***NDRE***	710nm	***CCCI***	538nm
G2	***CI Rededge***	***MTCI***	***CI Green***	576nm	***CRI***	716nm	518nm	727nm
Avocado T1 & T4								
G1	***CI Rededge***	***MTCI***	716nm	541nm	***CRI***	***CCCI***	739nm	***NDRE***
G2	***CI Green***	***CI Rededge***	***MTCI***	544nm	***CRI***	716nm	915nm	759nm

Table 4-4: Selected features from spectral signatures for Bean, Maize, and Avocado crops.

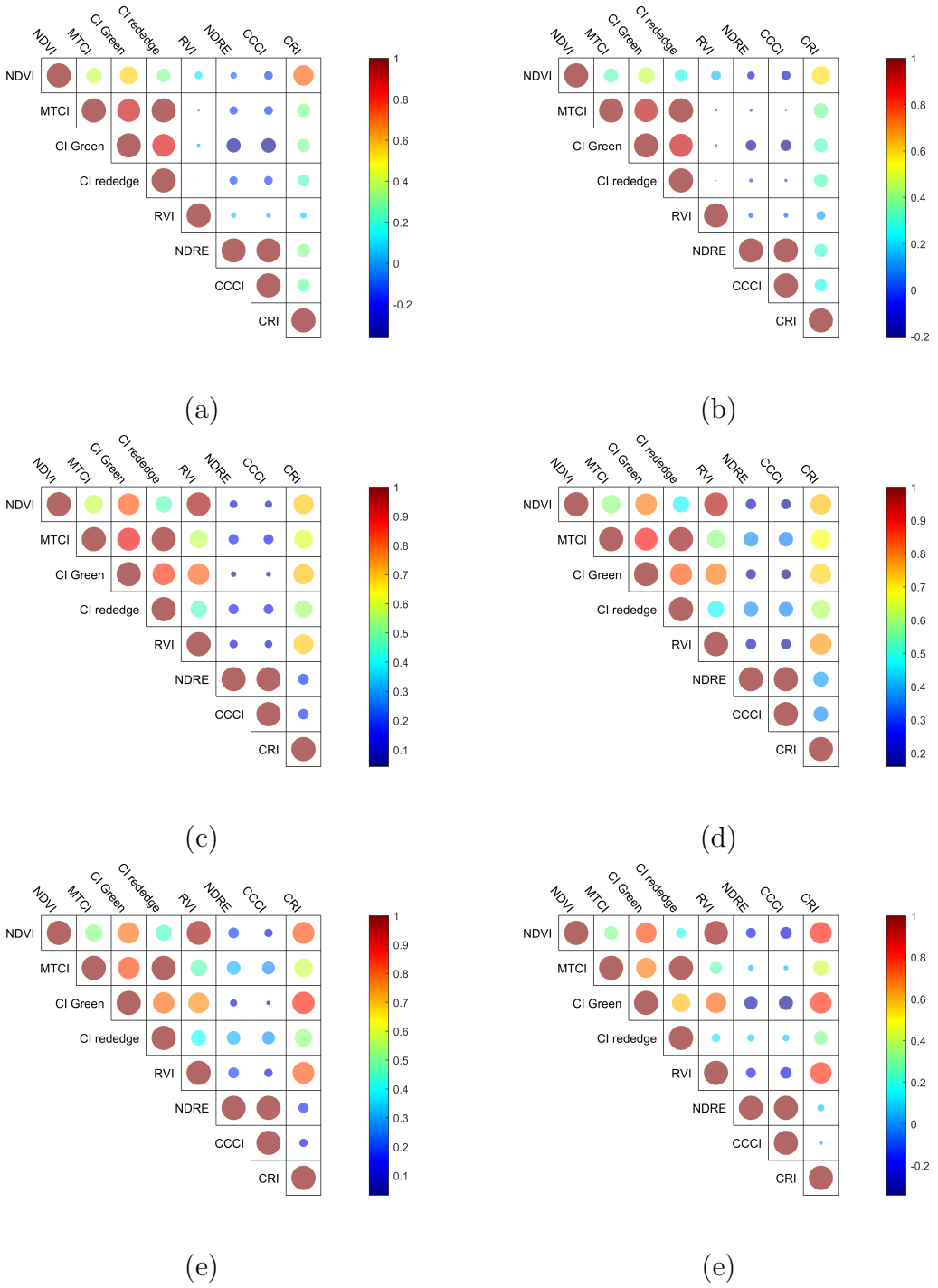


Figure 4-17: Correlation of vegetation indeces calculated from spectral signatures of (a) Bean Cerete genotype 1,(b) Bean Cerete genotype 2 (c) Maize Cerete genotype 1,(d) Maize Cerete genotype 2, (e) Maize Rionegro genotype 4, (f) Maize Rionegro genotype 6.

4.4. Stress detection

This section displays the stress detection results obtained through machine learning techniques. Bean Cerete showed a mean value accuracy of 95 % using the full spectra among the classifiers. However, the vegetation indices ***NDVI***, ***CCI*** and ***CI Green***, which displayed an average accuracy of 69 %, with a maximum of 85 %, presented a low response. Both genotypes showed an increase in computational cost when using the complete spectra, as seen in Tables **4-5** and **4-6**. The discriminative bands, regions of interest, or selected features produced results that were very close to those obtained with the full spectra Where the results were over 88 %. It is noteworthy that the nutritional stress case study in beans showed a high level of accuracy, and the features obtained through the procedures produced comparable evaluation results, with no significant differences in the selected classification techniques' performance.

	Accuracy			Computation Time [μSeg]		
	SVM	RF	NN	SVM	RF	NN
Selected Bands	0.96	0.88	0.95	1208	355	642
Selected Features	0.95	0.92	0.95	1270	358	590
Region Bands	0.96	0.89	0.97	1197	363	578
Region Bands & CI-Green	0.96	0.91	0.95	1169	348	587
NDVI	0.72	0.68	0.74	1165	492	543
CCI	0.53	0.54	0.52	1139	428	591
CI-Green	0.85	0.79	0.84	1109	372	571
All bands	0.98	0.90	0.97	7535	1929	4226

Table 4-5: Overall accuracy classification and computational cost for stress detection usign SVM, RF, and NN with spectral signatures for Bean Cerete Genotype 1.

	Accuracy			Computation Time [μSeg]		
	SVM	RF	NN	SVM	RF	NN
Selected Bands	0.94	0.88	0.95	1748	744	758
Selected Features	0.92	0.89	0.94	1409	426	650
Region Bands	0.95	0.90	0.94	1312	420	668
Region Bands & CI-Green	0.95	0.90	0.93	1271	373	701
NDVI	0.72	0.66	0.70	1664	645	846
CCI	0.55	0.55	0.53	1190	401	651
CI-Green	0.78	0.75	0.79	1200	416	638
All bands	0.96	0.89	0.95	7078	1890	4208

Table 4-6: Overall accuracy classification and computational cost for stress detection usign SVM, RF, and NN with spectral signatures for Bean Cerete 2.

The Cerete Maize study revealed that detecting stress in genotype 1 was challenging as the classification efficiency was low, with levels below 76 % Table **4-7**. Vegetation indices demonstrated similar behavior to the full spectrum when using the ***RF*** model. However, there was a slight difference between the proposed and conventional methods, especially when only selected bands were used, which may be due to the exclusion of the 563 nm Table **??**. In genotype 2 Table **4-8**, despite not being selected by the relevant feature algorithms, the ***NDVI*** vegetation index achieved the highest classification percentage in ***RF*** . In general, the crop showed low levels in the classification.

	Accuracy			Computation Time [μSeg]		
	SVM	RF	NN	SVM	RF	NN
Select Bands	0.70	0.52	0.70	1311	454	678
Features	0.67	0.60	0.67	1313	424	659
Region Bands	0.59	0.55	0.66	1362	476	662
Region Bands & NDRE	0.68	0.60	0.66	1349	444	649
NDVI	0.49	0.48	0.50	1227	467	656
NDRE	0.65	0.57	0.65	1214	421	617
CCCI	0.66	0.57	0.66	1210	430	626
All bands	0.74	0.55	0.76	9268	2354	4721

Table 4-7: Overall accuracy classification and computational cost for stress detection usign SVM, RF, and NN with spectral signatures for Maize Cerete Genotype 1.

	Accuracy			Computation Time [μSeg]		
	SVM	RF	NN	SVM	RF	NN
Select Bands	0.64	0.56	0.67	1730	553	793
Features	0.66	0.58	0.68	1310	432	688
Region Bands	0.56	0.54	0.54	1376	483	702
Region Bands & NDRE	0.56	0.55	0.55	1329	430	681
NDVI	0.53	0.60	0.61	1733	700	914
NDRE	0.51	0.55	0.53	1240	427	668
CCCI	0.51	0.52	0.59	1231	542	669
All bands	0.71	0.54	0.69	7292	2121	4507

Table 4-8: Overall accuracy classification and computational cost for stress detection usign SVM, RF, and NN with spectral signatures for Maize Cerete Genotype 2.

The results for Maize Rionegro genotype 1 in which 37 bands were selected Table **4-9**. Due to the complexity of the classification process, the use of reduced dimension methods showed slightly better results than the vegetation indices, with similar computational times. Notably, the selected feature set presented an accuracy rate with the ***NN*** model of 68 %, below the result obtained with all spectral information. The ***RF*** and ***NN*** methods were negatively affected when the whole spectrum was used, resulting in lower accuracy and more than three times the processing time required by the selected methods, as shown in Table 4-9. The reduced datasets that provided the best performance were Regional Bands & ***NDRE*** and Selected Bands.

	Accuracy			Computation Time [μSeg]		
	SVM	RF	NN	SVM	RF	NN
Select Bands	0.65	0.48	0.66	2815	369	634
Features	0.59	0.53	0.54	2791	386	564
Region Bands	0.47	0.39	0.59	2672	380	557
Region Bands & NDRE	0.65	0.51	0.59	2581	354	587
NDVI	0.35	0.37	0.37	2500	347	674
NDRE	0.51	0.49	0.52	2664	351	598
CCCI	0.50	0.50	0.51	2425	349	660
All bands	0.71	0.45	0.57	17137	1267	2363

Table 4-9: Overall accuracy classification and computational cost for stress detection usign SVM, RF, and NN with spectral signatures for Maize Rionegro Genotype 1.

According to the results presented in Table **4-10**, using reduced dimension methods on genotype 3 decreased the accuracy compared to the ***SVM*** classifier. However, it was more efficient than using the full spectrum for ***RF*** and ***NN*** methods. For genotype 3, the regional

banding and NDRE method performed well. In general, there is an apparent difficulty in classifying the data in the particular conditions of the study. Genotype 10 presented the best classification rates with a total of 38 selected bands Table **4-11**, particularly for this genotype, the worst performance of the vegetation indices in comparison, particularly the ***NDVI*** index, which in the literature is expected to be a discriminant feature presented the worst results for all three cases.

	Accuracy			Computation Time [μSeg]		
	SVM	RF	NN	SVM	RF	NN
Select Bands	0.70	0.59	0.72	3628	555	656
Features	0.71	0.57	0.75	2964	540	473
Region Bands	0.69	0.50	0.71	3008	401	495
Region Bands & NDRE	0.75	0.60	0.70	3036	440	481
NDVI	0.35	0.32	0.34	3132	531	579
NDRE	0.48	0.47	0.47	3223	936	683
CCCI	0.47	0.45	0.47	2957	421	456
All bands	0.81	0.48	0.77	7704	972	1362

Table 4-10: Overall accuracy classification and computational cost for stress detection usign SVM, RF, and NN with spectral signatures for Maize Rionegro Genotype 3.

	Accuracy			Computation Time [μSeg]		
	SVM	RF	NN	SVM	RF	NN
Select Bands	0.72	0.50	0.81	3811	591	679
Features	0.69	0.62	0.70	2990	646	511
Region Bands	0.71	0.49	0.75	2950	400	464
Region Bands & NDRE	0.81	0.62	0.76	2956	407	470
NDVI	0.37	0.30	0.35	3368	561	596
NDRE	0.49	0.48	0.50	3478	921	645
CCCI	0.52	0.43	0.52	2921	401	483
All bands	0.88	0.59	0.79	7653	1890	1296

Table 4-11: Overall accuracy classification and computational cost for stress detection usign SVM, RF, and NN with spectral signatures for Maize Rionegro Genotype 10.

The Tables **4-12** and **4-13** illustrate the results of the avocado classification analysis. There is difficulty in ranking the data even with complete spectrum information when attempting to rank stress levels, highlighting that, as with the maize crop, the ***NDVI*** indices could have been a better performer. However, when the effect of water deficit versus excess on the crop is assessed discretely between the extreme cases, there is a significant increase in the ranking

performance. Consistently for both genotypes, the results show a significant improvement of about 20 % in the accuracy obtained for the reduced feature spaces relative to using all bands in the cases considered, especially with the selected bands and the top features according to the feature extraction methods. On this dataset, all classification models presented comparable results.

	Accuracy All			Accuracy T1&T4		
	SVM	RF	NN	SVM	RF	NN
Select Bands	0.48	0.51	0.49	0.71	0.71	0.73
Features	0.46	0.50	0.49	0.72	0.70	0.73
Region Bands	0.40	0.46	0.46	0.67	0.66	0.72
Region Bands & CI-Green	0.41	0.48	0.47	0.69	0.67	0.70
NDVI	0.27	0.29	0.32	0.51	0.51	0.53
CCI	0.31	0.33	0.36	0.54	0.57	0.65
CI-Green	0.33	0.75	0.33	0.55	0.59	0.65
All bands	0.53	0.51	0.50	0.73	0.72	0.75

Table 4-12: Overall accuracy classification for stress detection using SVM, RF, and NN with spectral signatures of Avocado Rionegro Genotype 1.

	Accuracy All			Accuracy T1&T4		
	SVM	RF	NN	SVM	RF	NN
Select Bands	0.52	0.50	0.51	0.63	0.71	0.74
Features	0.44	0.49	0.40	0.68	0.73	0.71
Region Bands	0.40	0.44	0.45	0.60	0.67	0.72
Region Bands & CI-Green	0.41	0.45	0.46	0.60	0.67	0.72
NDVI	0.30	0.27	0.32	0.57	0.52	0.57
CCI	0.37	0.34	0.40	0.57	0.62	0.67
CI-Green	0.38	0.33	0.41	0.57	0.61	0.66
All bands	0.54	0.50	0.57	0.73	0.72	0.74

Table 4-13: Overall accuracy classification for stress detection using SVM, RF, and NN with spectral signatures of Avocado Rionegro Genotype 2.

4.5. Discussion

This section describes the results obtained for stress detection using the proposed methodology (Figure **3-1**). The results compared the overall accuracy and computational time for stress detection using (a) selected bands, (b) selected features, (d) regions bands, (e) regions bands combined with the most relevant vegetation index, (f) most relevant index for each crop (according to the results in section 4.3), and (g) the results using the full spectra. Table **4-5** presents the results obtained for Bean Cerete. The best accuracy was obtained using the full spectra with the ***SVM*** classifier as the stress detector, with an accuracy of 98 % for Genotype 1 and 96 % for Genotype 2. However, the computational cost to use the full spectrum is the highest. The closest results obtained with the full spectra can be obtained using one of the feature selections (selected bands, selected features, or regions of interest), with accuracies up to 97 % for Genotype 1 and 95 % for Genotype 2 using the ***NN***, which significantly reduces the computational cost. It can be noted in Table Table **4-5** that the lowest performance was obtained when a vegetation index was used with either genotype 1 or 2.

For Maize Cerete, a lower performance was obtained than the bean trail. Again, the best overall accuracy was obtained using the full spectrum, but for Genotype 1, this result was obtained using ***NN*** with 76 %, and for Genotype 2, by ***SVM*** with 71 %. One more time, using the full spectrum increases the computational cost. For Genotype 1, the ***SVM*** or ***NN*** using selected bands obtained the closest results to the best, with a 70 % overall accuracy. For Genotype 2, the same classifiers achieve 66 % and 68 % of overall accuracy, respectively. The results for Maize Rionegro were very similar to those of the trial in Cerete. The best detector was the ***SVM*** using the full spectrum, obtaining accuracies up to 71 % for Genotype 1, 81 % for Genotype 3, and 88 % for Genotype 10. This evidences the dependence of results on the genotypes because there is a significant difference between the overall accuracies obtained. For the experiment of Maize Riongero, the results obtained with any subset of bands are very low. Similar to the result with beans, using vegetation indexes provided bad stress detection results.

Finally, Tables **4-12** and **4-13** present the results for Avocado Rionegro. These results include the use of all treatments ("All" columns in Tables **4-12** and **4-13**) and the use of only the extreme treatments. For all treatments, it can be noted that lower performance is obtained with either detector: ***SVM***, ***RF***, or ***NN***. For Genotype 1, the best results were lower than 53 % using the full spectrum and ***SVM***. For Genotype 2, the best results were 58 % using the full spectrum but ***NN***. However, for the extreme treatments, the overall accuracies were up to 75 % and 74 % by using ***NN*** and the full spectrum. Close results can be obtained using selected relevant bands and ***NN***. The lower results obtained using all levels of treatment can be related to the need for more labeled data to train the models. For all crops, the best detection results were obtained using the full spectra. In most cases, ***SVM*** outperformed

the other classifier. The lowest performances were obtained using the ***RF*** classifier or by using only the vegetation index. The result provided good support for considering the need to use the full information of spectral signatures or to use lower bands (reducing the overall accuracies a little) to improve computational cost.

5 Spatial Feactures

Data was collected using two Micasense cameras to measure the impact of stress on crop leaves. To classify the images of individual leaves, they were divided into a grid. The images obtained were used to create mosaics for both the training and test data. The training data were presented first, with 40 % of the data being used to train the models. The results were then validated using the test set. To identify the spatial characteristics of different crops and treatments, textural descriptors such as energy, contrast, entropy, and correlation were used based on spectral signature analysis. During the training process, the most relevant textural features were evaluated.

5.1. Training Process

Overall accuracy for the crops included in this analysis (i.e., Bean Cerete, Maize Cerete, and Maize Rionegro) are presented in Figure **5-1**. For Cerete bean, Figure **5-1** (a,b) shows that accuracies higher than 95 % can be obtained using 8 features for Genotype 1 and 4 for Genotype 2 with any of the classifiers. On the other hand, for Maize Cerete, Figures **5-1** (c, d) show lower performance than bean trails. Using 6 features can achieve 95 % of accuracy for both genotypes. Finally, in Figure **5-1** (e), the results were included with Rionegro Maize Genotype 1 under the four treatments. For this crop, more features were required to obtain accuracies higher than 90 %. The RF classifier was the detector with the best results, particularly for maize trails. They have employed feature extractors to obtain various data sets, including band information, vegetation indices, and spatial textures. We then integrated all the spatial-spectral information and selected the main features using feature extractor relevance projectors.

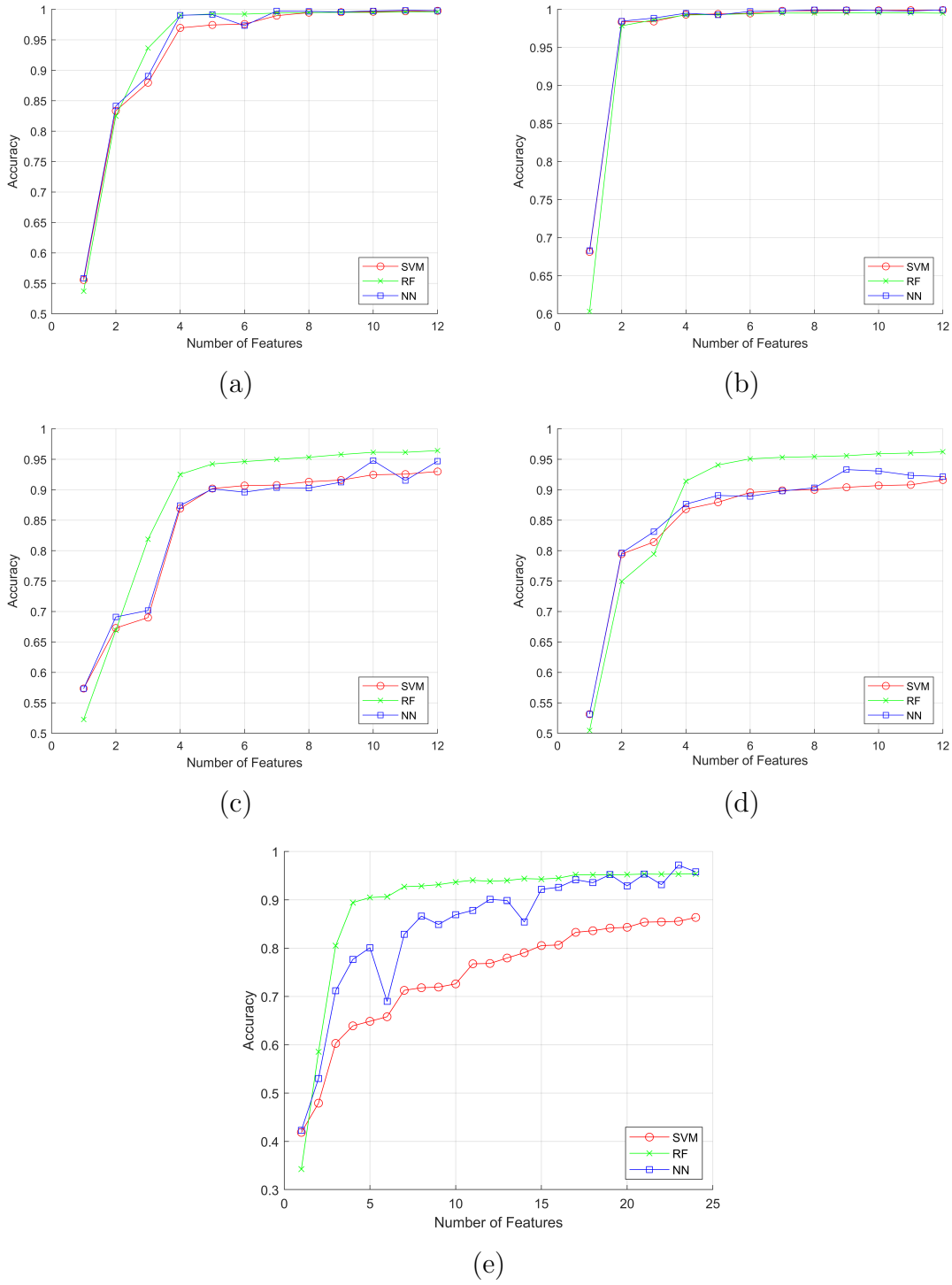


Figure 5-1: Overall classification accuracy according to number of features (a) Bean Cerete genotype 1,(b) Bean Cerete genotype 2 (c) Maize Cerete genotype 1,(d) Maize Cerete genotype 2, (e) Maize Rionegro genotype 1.

The classification results using bands from Micasense cameras, vegetation indices, selected textures, all textures, and the main features chosen from the feature extractors are presented in the Table **5-1**. Generally, good classification indices are observed, although it is noted that the ratings are lower compared to other methods and vegetation indices. It is worth highlighting that the data sets that include spatial features, which can be obtained from the image, stand out. The ***SVM*** classifier applied to the primary features is graphically presented in the Figures **5-2** and **5-3** . The left figure shows the training data, differentiated by the colors yellow and cyan for treatments 1 and 2, respectively. In contrast, the correct figure indicates the model's estimation, where atypical pixels can be observed in the classification within the leaf.

In particular, for Rionegro maize Figure **5-4** , the four treatments are presented in color yellow, golden , cyan, and blue, corresponding to treatments 1, 2, 3, and 4, respectively. Difficulties are observed in the classification process, particularly to treatments 2 and 3.

Crop		Bean C.G1	Bean C. G2	Maize C.G1	Maize C.G2	Maize R.G1
Band	SVM	0.81	0.97	0.80	0.84	0.75
	RF	0.97	0.98	0.83	0.82	0.88
	NN	0.87	0.94	0.68	0.83	0.60
VI	SVM	0.75	0.79	0.70	0.70	0.78
	RF	0.98	0.98	0.84	0.75	0.92
	NN	0.86	0.87	0.73	0.71	0.61
Textures	SVM	0.99	0.97	0.96	0.95	0.93
	RF	0.99	0.99	0.95	0.98	0.99
	NN	0.99	0.99	0.95	0.94	0.95
All Features	SVM	0.99	0.99	0.95	0.96	0.93
	RF	0.99	0.99	0.95	0.97	0.99
	NN	0.99	0.99	0.95	0.95	0.94
Principal Features	SVM	0.89	0.99	0.94	0.95	0.80
	RF	0.99	0.99	0.95	0.98	0.98
	NN	0.98	0.99	0.92	0.93	0.94

Table 5-1: Results Train Data for Bean and Maize crop.

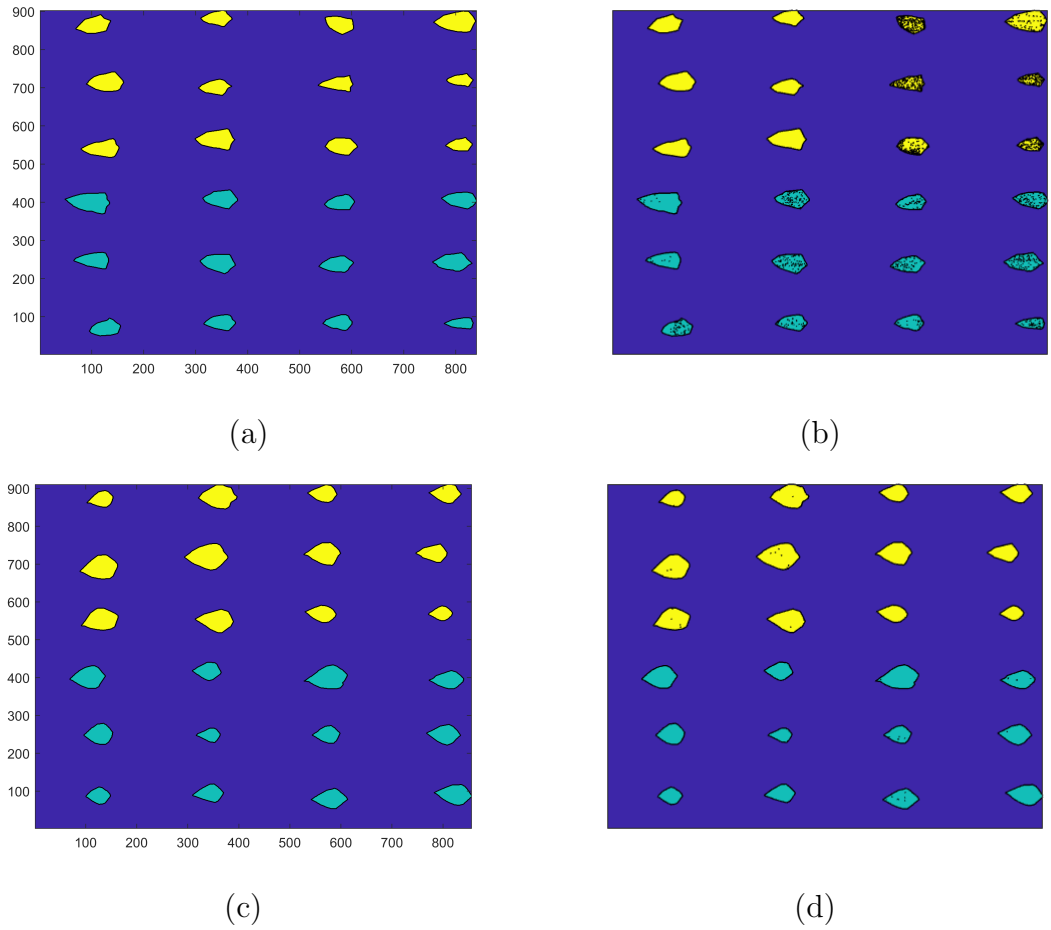


Figure 5-2: Model train accuracy Bean Certe (a) Test Data Label genotype 1,(b) Model Estimation genotype 1 (c) Test Data Label genotype 2 (d) Model Estimation genotype 2.

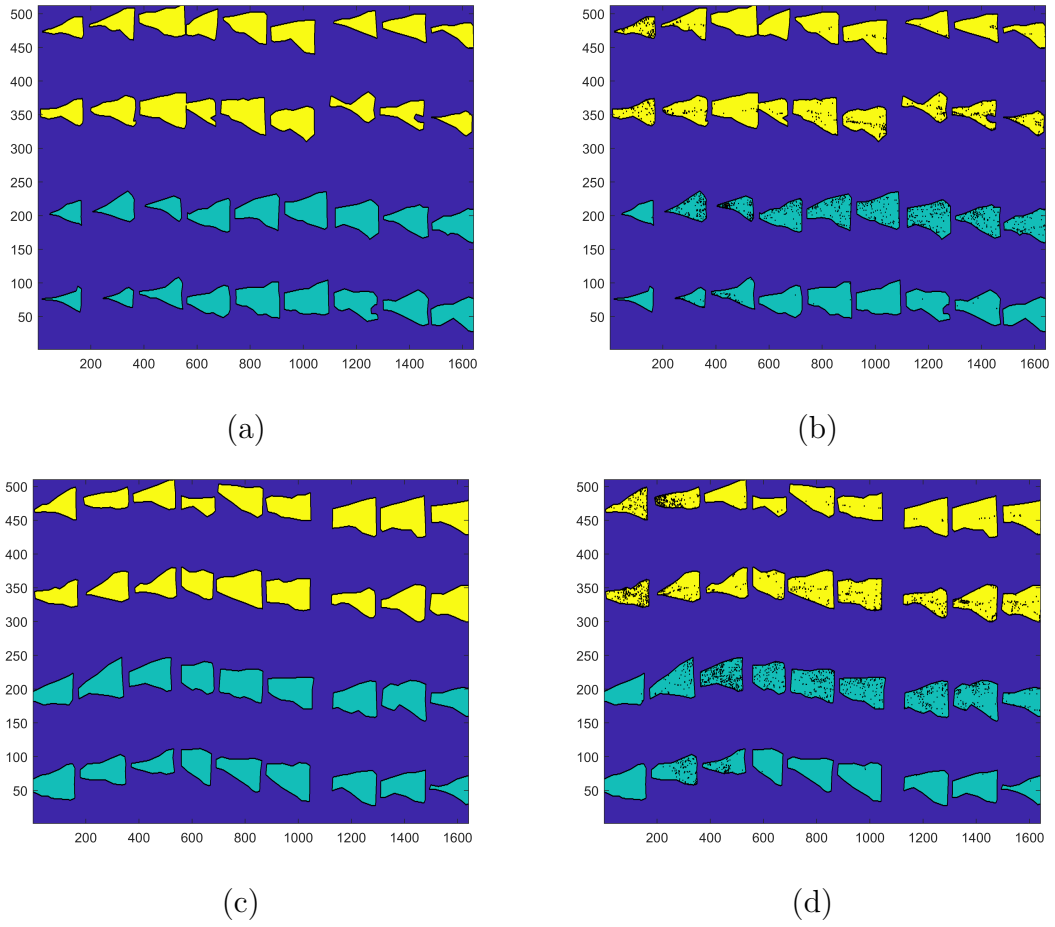
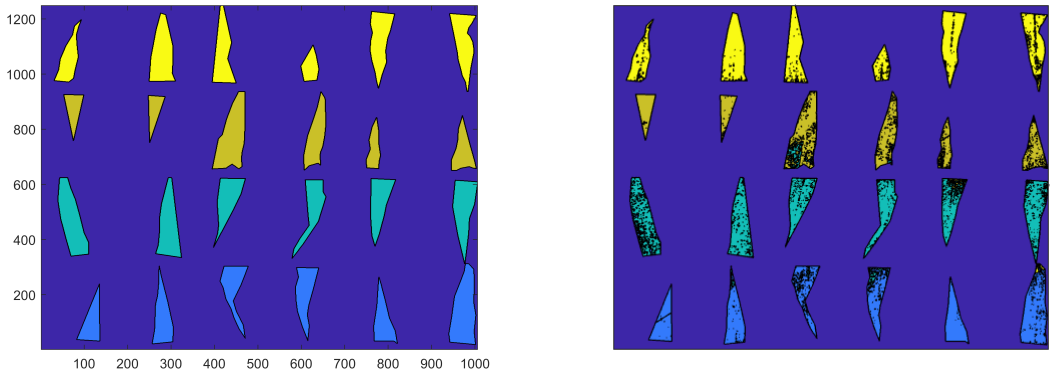


Figure 5-3: Model train accuracy Maize Certe (a) Test Data Label genotype 1,(b) Model Estimation genotype 1 (c) Test Data Label genotype 2 (d) Model Estimation genotype 2.



(a) (b)

Figure 5-4: Model train accuracy Maize Rionegro (a) Test Data Label genotype 1,(b) Model Estimation genotype 1.

5.2. Testing Process

Table **5-2** presents the classification results obtained using bands provided by Micasense cameras, applying vegetation indices and selected textures. The training model was created using the training set and then applied to the test data. Additionally, Table **5-3** shows the time taken for each data set and their respective classifiers. It is worth noting that the results for Cerete beans remained similar to those obtained in the previous chapter. Similarly, low results were observed for Cerete Maize, which was consistent with previous findings. In the case of Rionegro Maize, an improvement was observed when texture analysis data was included, indicating the importance of status features. This is particularly significant considering that the band information was obtained from only eight channels available on the sensors used for data capture.

Based on the evaluation of the datasets, it has been found that the ***RF*** classifier is highly efficient in terms of computational cost and produces generally good results. On the other hand, although practical, the ***SVM*** classifier can be computationally expensive for the same datasets. This information can help select the appropriate classifier for future data analysis and classification tasks.

The results of ***NN*** model simulation are displayed in Figure **5-5**. It seems that for genotype 1, the results are satisfactory, with only a few pixels showing anomalies. However, regarding genotype 2, significant challenges exist in identifying treatment 2. The graph in Figure **5-6** shows the results obtained at various stress levels. Although the extreme treatment produced a higher identification rate, identifying was more challenging than treating 2. In treatment 2, only isolated pixels were correctly identified, and most of them were estimated to be objects

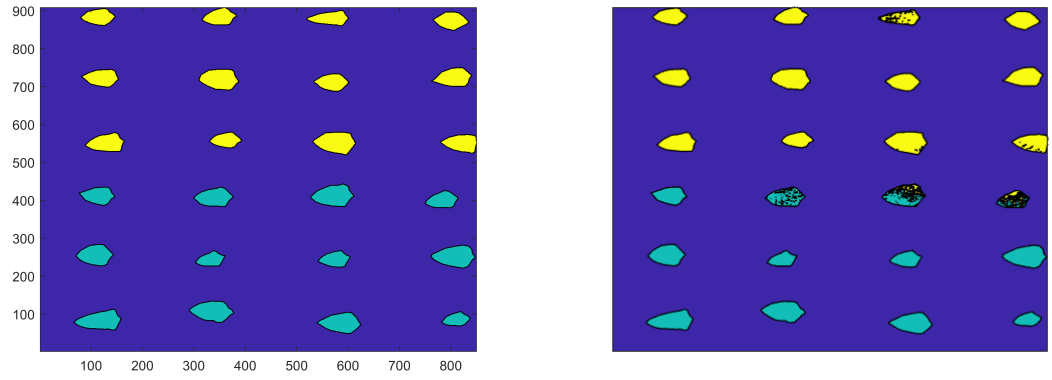
from treatment 3.

Crop		Bean C.G1	Bean C. G2	Maize C.G1	Maize C.G2	Maize R.G1
Band	SVM	0.61	0.76	0.58	0.51	0.45
	RF	0.57	0.77	0.55	0.61	0.38
	NN	0.63	0.89	0.59	0.59	0.46
VI	SVM	0.70	0.69	0.62	0.64	0.27
	RF	0.62	0.66	0.59	0.56	0.38
	NN	0.76	0.54	0.58	0.65	0.29
Textures	SVM	0.77	0.75	0.64	0.63	0.74
	RF	0.88	0.78	0.65	0.62	0.56
	NN	0.83	0.70	0.59	0.66	0.70
All Features	SVM	0.79	0.62	0.63	0.68	0.72
	RF	0.85	0.77	0.67	0.70	0.57
	NN	0.82	0.70	0.65	0.66	0.68
Principal Features	SVM	0.78	0.74	0.64	0.66	0.74
	RF	0.79	0.78	0.63	0.69	0.56
	NN	0.80	0.72	0.65	0.68	0.62

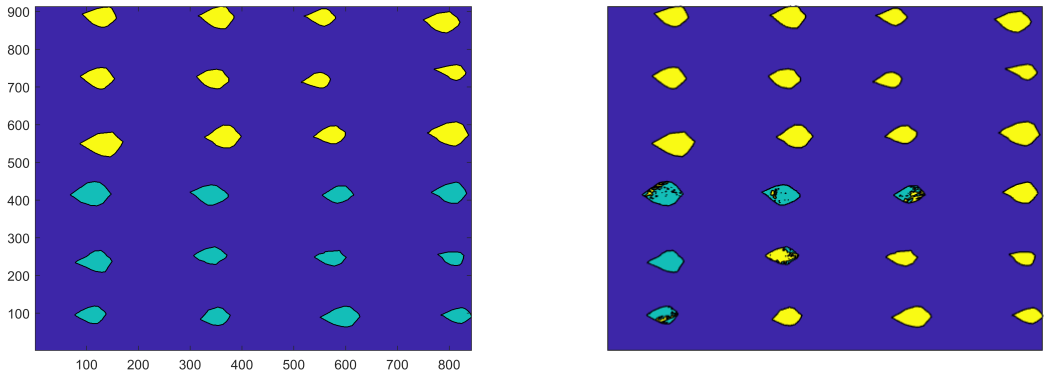
Table 5-2: Results Test Data for Bean and Maize crop, highlighted in green the best results for Bean Cerete, highlighted in cyan the best results for Maize Cerete, highlighted in magenta the best results for Maize Rionegro.

Crop		Bean C.G1	Bean C. G2	Maize C.G1	Maize C.G2	Maize R.G1
Band	SVM	306.52	52.32	143.98	280.85	173.45
	RF	0.93	0.22	0.42	1.18	238.35
	NN	70.61	47.62	68.33	255.62	143.37
VI	SVM	122.55	26.60	71.12	203.35	806.89
	RF	0.69	0.18	0.48	43.43	172.32
	NN	72.56	41.38	110.64	38.42	152.45
Textures	SVM	220.03	3.01	89.93	479.28	960.99
	RF	1.04	0.33	4.51	52.54	296.691
	NN	82.09	13.59	14.54	216.39	155.39
All Features	SVM	843	130.73	1124.04	968.47	1428.80
	RF	1.78	0.51	5.31	345.76	510.11
	NN	97.21	57.46	419.36	123.29	181.89
Principal Features	SVM	147.95	3.22	92.36	462.87	659.50
	RF	0.43	0.14	0.27	636.06	170.60
	NN	70.21	48.43	43.83	382.60	170.94

Table 5-3: Results Time Test Data for Bean and Maize crop.

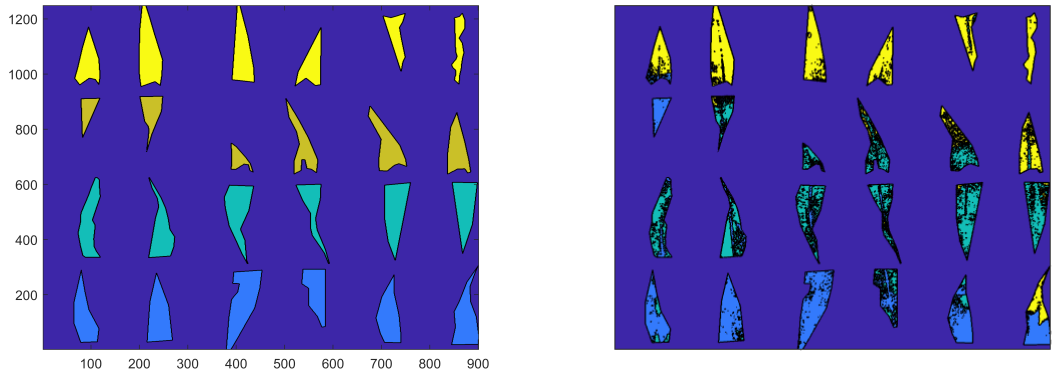


(a) (b)



(c) (d)

Figure 5-5: NN Model test using Principal Feactures data Bean Certe (a) Test Data Label genotype 1,(b) Model Estimation genotype 1 (c) Test Data Label genotype 2 (d) Model Estimation genotype 2.



(a) (b)

Figure 5-6: NN Model test using Principal Feactures data Maize Rionegro (a) Test Data Label genotype 1,(b) Model Estimation genotype 1.

A confusion matrix Figure **5-7** is used to abstract the numerical information in correspondence with the above figures. It is a tool used to evaluate the performance of classification models. It is a table that visualizes the number of correct and incorrect predictions made by the model compared to the actual observations. The diagonal cells correspond to correctly classified observations, while the off-diagonal cells correspond to incorrectly classified observations. The positive predictive value and false discovery rate are shown at the right and bottom of the plot, respectively, and the actual positive rate and false negative rate are shown at the bottom right of the plot. Identifying the pixels corresponding to treatment 3 of Maize Rionegro is notoriously difficult, with only a lower rate of 19 % of samples properly identified.

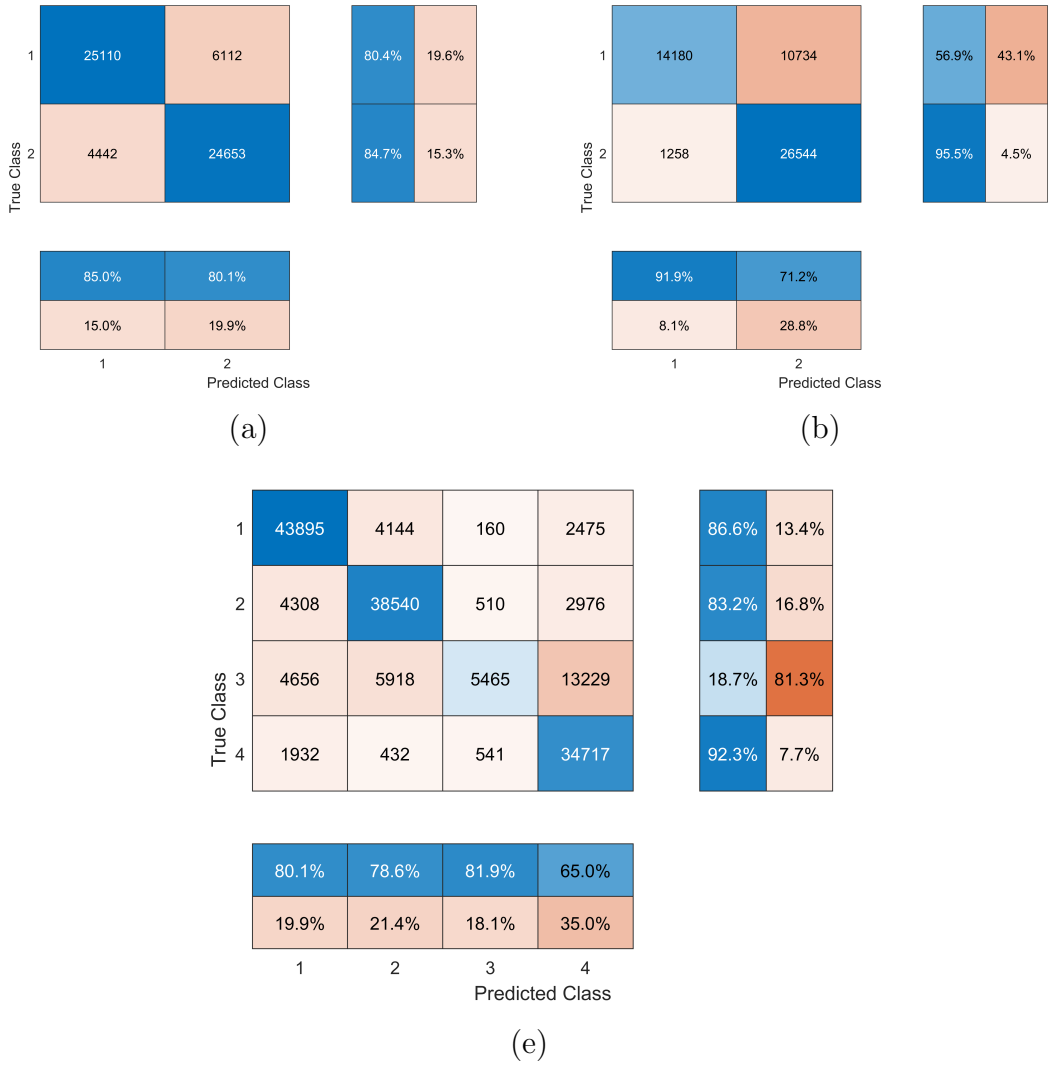


Figure 5-7: Confusion Matriz using NN Model test with Principal Feactures set (a) Bean Cerete genotype 1,(b) Bean Cerete genotype 2 (c) Maize Rionegro genotype 1. classes 1, 2, 3, 4 correspond to treatments 1, 2, 3 and 4.

5.3. Discussion

The approach allows the incorporation of the spatial relationship between pixels, completing spectral information. The spatial information was based on texture metrics, which obtained a new representation space. The findings presented in this section demonstrate the benefits of incorporating spatial information. However, the multispectral camera's capabilities limited the spectral data to ten channels. There were instances where this limitation adversely affected the identification process, such as Cerete bean. Nonetheless, the genotypes consistently scored classifications over 78 % (Table **5-2**). In the case study of Rionegro maize, genotype 1 represents a higher classification rate than the results obtained from the spectrometer.

The results evidence that more than relying solely on conventional methods, such as those based on vegetation indices, are needed to detect stress levels. This underscores the limitations of utilizing a restricted amount of information from samples and the challenges of applying generalized indices to all crops without considering where the most pertinent information is. It is easier to identify stress in vegetation by considering this aspect.

It has been observed that incorporating spatial information of images through texture characteristics significantly improves classification accuracy for both training and testing datasets. This is especially true for multi-spectral images where the spectral information is limited. Without incorporating spatial information, the classification measures are substantially lower when compared to the methods that consider it.

In this thesis project, classifying sets containing spatial and spectral information posed difficulties, such as identifying the selection of spatially relevant information, the limitations and characteristics of the sensors, the variability between leaves in shape and size, and external agents such as tears, malformations, or biotic agents. However, the proposed methodology showed a favorable impact when validated quantitatively. This opens up opportunities to further develop the methodology into a general framework that can be applied to any dataset.

6 Conclusions and future work

The process standardization under stress conditions in vegetation crops has posed a formidable challenge owing to each case's distinctive impact on the crop's spectral signatures and the stress factor. Despite certain statistical similarities in some cases, the spectral response of the plant can vary significantly in each instance. Regrettably, no process suggested thus far, including customized vegetation indexes or those seeking specific effects, can be generalized due to the variable outcomes based on the unique characteristics of the object of study. As discussed in Chapter 4, the statistical data presented in the case studies indicate that genetic variability is a crucial factor in stress arrest. This is because the response in the spectral signature varies according to the type and level of stress the crop is exposed to and the particular response of each genotype to these effects. In addition, environmental factors such as dew or impurities in the samples and changes in illumination can also affect the distribution of the signature characteristics. Therefore, developing agile methodologies, like the one presented in this work, is crucial for breaking down knowledge barriers and equipping us with the necessary technological tools to tackle this issue.

The proposed method enables the characterization of the spectral signature of crops under the influence of abiotic limiting factors. It identifies the zones and bands of interest where the most relevant information is preserved. The proposed method aims to characterize the spectral signature of crops that are under the influence of abiotic limiting factors. It identifies the areas and bands of interest that preserve the most relevant information. Despite increasing noise intensity due to limitations of the applied sensor, the literature highlights that NIR bands provide important information for the detection processes. This information is discussed in chapter 4. However, some limitations of the method need to be considered. Firstly, there is high variability within individuals, which may affect the accuracy of the results. Secondly, the effect of biotic stress on the spectrum is low. For example, in the case of maize, the regions of interest were similar in both experiments, although in one experiment, there was a classification.

The results indicate that the developed method outperforms the evaluated vegetation indices. This is accomplished through a thorough band selection process based on consistently superior spectral signature characterization. The full spectrum helps significantly enhance the indices' stress estimation performance. When it comes to images, having fewer bands necessitates the inclusion of spatial information. Each crop exhibits distinct behavior, and

pre-processing the images is more challenging due to illumination variations that must be considered. In some cases, even with well-characterized signatures, sufficient heterogeneity must be required to enable an adequate classification. The integration of spatial information is effective in enhancing the classification process, especially in cases where spectral information alone cannot detect stress in crops under stress levels. The integration of spatial-spectral information provides a clear advantage in detecting stress. However, spatial information is susceptible and can be affected by other crop affectations, and the texture can vary depending on the climatic conditions in which the samples are obtained. The proposed methodology provides a robust and comprehensive approach for classifying crops under abiotic stress and can significantly benefit the agricultural industry.

Even though the technification and industrialization of agriculture are areas of interest for the scientific community, there still needs to be a significant gap when providing access to these technologies, mainly due to the high cost and the difficulty of acquiring expert staff. This is where the methodological work done in this project can positively impact by helping to select suitable sensors for particular crops. It is essential to consider that agriculture processes today focus on genotypes with optimal conditions for productivity and geographical conditions of the environment, so if optimization of technologies and sensors can be established for a particular crop, the proposed scheme can serve as a valuable tool for selecting the appropriate bands and regions of interest. This, in turn, can aid in developing new, cost-effective devices and reduce implementation costs by targeting the spectrum to the specific regions of interest in each crop.

Multispectral sensors can be designed for specific crops by examining the spectral signature, selecting regions and bands of interest, and exploring the potential of vegetation indices. These sensors can be tailored with defined bandwidths and directed towards the particular study case, streamlining information acquisition and integrating spatial information where the most pertinent data is concentrated to classify the data under particular conditions. However, there are risks associated with designing sensors for specific genotypes. This is because of the low heterogeneity that the crop may have in biotic stress and the geographical difficulties that may arise. It is not desirable to lose enough generality to identify risk factors that may compromise crop productivity and the excellent use of the soil. Therefore, it is necessary to evaluate whether obtaining a set of sensors with specific bands is feasible before establishing sensor ranges. This requires submitting the genotype evaluated to the different stresses that may arise.

The following contributions were made in the implementation of this project: (I) development of a data collection protocol and a low-cost clamp for measuring vegetation spectral signatures, these were presented in the paper "Low-cost clamp for the measurement of vegetation spectral signatures"; (II) a spectral library database presented in "Spectral Library

of Maize Leaves under Nitrogen Deficiency Stress", which details the preprocessing steps to remove noise and outliers. The library includes three datasets captured at different growth stages of 10 tropical maize genotypes. The spectral signatures collected were in the visible to near-infrared range (450-950 nm), which is a significant contribution as one of the shortcomings identified is the lack of comprehensive libraries to test classification systems. The preliminary results of the characterization section identification signatures of the same under-stress effects are submitted in paper Characterization of crops under abiotic stress factors using spectral signatures in the visible to near infrared. The research conducted in this study has opened up a substantial opportunity for further exploration and discourse on the obtained results. This could entail formalizing and modifying the proposed methodology to cater to a broader range of circumstances and developing a tool to examine the characterization of main features. Such measures would enable the identification of bands of interest and create streamlined methods for their identification.

Reliable data capture is necessary. Even though semi-controlled conditions are considered, environmental changes or biotic risk factors may affect the process execution. This difficulty generated the need to statistically validate the labeled data beforehand, which limited its availability and affected the possibility of exploring other types of machine learning approaches, such as deep learning, which have proven to have the potential for this type of task [154, 152]. The study started with the assumption that the values of nitrogen, phosphorus, and water were controlled in the nutritional treatment. Soil measurements were performed, but the high cost of actual quantification of N and P in the plant did not temporarily compare the exact values that the plant was absorbing for each individual.

A future stage of deepening in a specific crop can be planned to guarantee adequate rates of training data with a broader scenario and meticulous control of external variables. Within the framework of the implementation of the project, future work can be carried out in the following areas: construction of other spectral libraries with signatures and images that allow the study to be extended to other ML techniques, development of customized devices based on the relevant spectral regions selected, extension of the methodology to different types of crops, and inclusion of novel selection techniques that can take advantage of the proposed scheme, Overall, the results of this study suggest ample room for further research and analysis, which could significantly advance how data is analyzed and utilized in various applications. The proposed measures can increase accuracy, efficiency, and repeatability, making them essential to the field of study. They also promote the application of precision agriculture on a cost-benefit basis.

Bibliography

[1] F. O. de las Naciones Unidas para la Agricultura y la Alimentación, "La agricultura mundial en la perspectiva del año 2050," 2009.

[2] J. G. M. Esgario, R. A. Krohling, and J. A. Ventura, "Deep learning for classification and severity estimation of coffee leaf biotic stress," *Computers and Electronics in Agriculture*, vol. 169, p. 105162, 2020.

[3] A. Singh, B. Ganapathysubramanian, S. Sarkar, and A. Singh, "Deep learning for plant stress phenotyping: Trends and future perspectives," *Trends in Plant Science*, vol. 23, 1 2018.

[4] R. A. Chowengerdt, *Remote sensing: models and methods for image processing*. elsevier, 2006.

[5] J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis,AnalysisBerlin/Heidelberg, Germany: springer*. 2022.

[6] L. Z. Huo and P. Tang, "Spectral and spatial classification of hyperspectral data using svms and gabor textures," *International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 46, pp. 1708–1711, 2011.

[7] D. Dutta, P. K. Das, S. Paul, T. Khemka, M. K. Nanda, and V. K. Dadhwal, "Spectral response of potato crop to accumulative moisture stress estimated from hydrus-1d simulated daily soil moisture during tuber bulking stage," *Journal of the Indian Society of Remote Sensing*, vol. 44, pp. 363–371, 2016.

[8] M. M. Islam, S. Matsushita, R. Noguchi, and T. Ahamed, "Development of remote sensing-based yield prediction models at the maturity stage of boro rice using parametric and nonparametric approaches," *Remote Sensing Applications: Society and Environment*, vol. 22, p. 100494, 4 2021.

[9] V. Nath and S. E. Levinson, "Future directions of precision agriculture," *SpringerBriefs in Computer Science*, pp. 55–56, 2014.

[10] S. Liaghat and S. Balasundram, "A review: The role of remote sensing in precision agriculture," *Journal, American Sciences, Biological Publications, Science*, vol. 5, pp. 50–55, 2010.

[11] J. A. Santanello, C. D. Peters-Lidard, M. E. Garcia, D. M. Mocko, M. A. Tischler, M. S. Moran, and D. P. Thoma, "Using remotely-sensed estimates of soil moisture to infer soil texture and hydraulic properties across a semi-arid watershed," *Remote Sensing of Environment*, vol. 110, pp. 79–97, 2007.

[12] J. A. Domínguez, J. Kumhálová, and P. Novák, "Assessment of the relationship between spectral indices from satellite remote sensing and winter oilseed rape yield," *Agronomy Research*, vol. 15, pp. 55–68, 2017.

[13] S. Elsayed and W. Darwish, "Hyperspectral remote sensing to assess the water status, biomass, and yield of maize cultivars under salinity and water stress," *Bragantia*, vol. 76, pp. 62–72, 2017.

[14] C. Stanton, M. J. Starek, N. Elliott, M. Brewer, M. M. Maeda, and T. Chu, "Unmanned aircraft system-derived crop height and normalized difference vegetation index metrics for sorghum yield and aphid stress assessment," *Journal of Applied Remote Sensing*, vol. 11, p. 026035, 2017.

[15] C. Zhang, J. Liu, J. Shang, and H. Cai, "Capability of crop water content for revealing variability of winter wheat grain yield and soil moisture under limited irrigation," *Science of the Total Environment*, vol. 631-632, pp. 677–687, 2018.

[16] H. Li, Z. Issaka, Y. Jiang, P. Tang, and C. Chen, "Overview of emerging technologies in sprinkler irrigation to optimize crop production," *International Journal of Agricultural and Biological Engineering*, vol. 12, pp. 1–9, 2019.

[17] J. M. Bioucas-dias, A. Plaza, G. Camps-valls, P. Scheunders, N. M. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *EEE Geoscience and remote sensing magazine*, vol. 1, pp. 6–36, 2013.

[18] A. Sharma, A. Jain, P. Gupta, and V. Chowdary, "Machine learning applications for precision agriculture: A comprehensive review," *IEEE Access*, vol. 9, pp. 4843–4873, 2021.

[19] J. D. Rudd, G. T. Roberson, and J. J. Classen, "Application of satellite, unmanned aircraft system, and ground-based sensor data for precision agriculture: A review," *2017 ASABE Annual International Meeting*, pp. 1–8, 2017.

[20] M. Ozdogan, Y. Yang, G. Allez, and C. Cervantes, "Remote sensing of irrigated agriculture: Opportunities and challenges," *Remote Sensing*, vol. 2, pp. 2274–2304, 2010.

[21] S. L. Ustin, A. A. Gitelson, S. Jacquemoud, M. Schaepman, G. P. Asner, J. A. Gamon, and P. Zarco-Tejada, "Retrieval of foliar information about plant pigment systems from high resolution spectroscopy," *Remote Sensing of Environment*, vol. 113, pp. S67–S77, 2009.

[22] R. Ferguson, D. Rundquist, D. Shannon, D. Clay, and N. Kitchen, "Remote sensing for site-specific crop management," vol. 69, pp. 647–664, 2018.

[23] A. J. Foster, V. G. Kakani, J. Ge, M. Gregory, and J. Mosali, "Discriminant analysis of nitrogen treatments in switchgrass and high biomass sorghum using leaf and canopy-scale reflectance spectroscopy," *International Journal of Remote Sensing*, vol. 37, pp. 2252–2279, 2016.

[24] K. C. DeJonge, B. S. Mefford, and J. L. Chávez, "Assessing corn water stress using spectral reflectance," *International Journal of Remote Sensing*, vol. 37, pp. 2294–2312, 2016.

[25] G. G. Peteinatos, A. Korsaeth, T. W. Berge, and R. Gerhards, "Using optical sensors to identify water deprivation, nitrogen shortage, weed presence and fungal infection in wheat," *Agriculture (Switzerland)*, vol. 6, 2016.

[26] M. Fang, W. Ju, W. Zhan, T. Cheng, F. Qiu, and J. Wang, "A new spectral similarity water index for the estimation of leaf water content from hyperspectral data of leaves," *Remote Sensing of Environment*, vol. 196, pp. 13–27, 2017.

[27] F. Rasheed, S. Delagrange, and F. Lorenzetti, "Detection of plant water stress using leaf spectral responses in three poplar hybrids prior to the onset of physiological effects," *International Journal of Remote Sensing*, vol. 41, pp. 5127–5146, 2020.

[28] N. Dangwal, N. R. Patel, M. Kumari, and S. K. Saha, "Monitoring of water stress in wheat using multispectral indices derived from landsat-tm," *Geocarto International*, vol. 31, pp. 682–693, 2016.

[29] Y. Ge, G. Bai, V. Stoerger, and J. C. Schnable, "Temporal dynamics of maize plant growth, water use, and leaf water content using automated high throughput rgb and hyperspectral imaging," *Computers and Electronics in Agriculture*, vol. 127, pp. 625–632, 2016.

[30] S. Chou, J. M. Chen, H. Yu, B. Chen, X. Zhang, H. Croft, S. Khalid, M. Li, and Q. Shi, "Canopy-level photochemical reflectance index from hyperspectral remote sensing and leaf-level non-photochemical quenching as early indicators of water stress in maize," *Remote Sensing*, vol. 9, pp. 1–17, 2017.

[31] S. Kokhan and A. Vostokov, "Using vegetative indices to quantify agricultural crop characteristics," *Journal of Ecological Engineering*, vol. 21, pp. 120–127, 2020.

[32] S. Chen, Y. Chen, J. Chen, Z. Zhang, Q. Fu, J. Bian, T. Cui, and Y. Ma, "Retrieval of cotton plant water content by uav-based vegetation supply water index (vswi)," *International Journal of Remote Sensing*, vol. 41, pp. 4389–4407, 2020.

[33] S. M. Borzov and O. I. Potaturkin, "Spectral-spatial methods for hyperspectral image classification. review," *Optoelectronics, Instrumentation and Data Processing*, vol. 54, pp. 582–599, 2018.

[34] K. V. Kale, M. M. Solankar, D. B. Nalawade, R. K. Dhumal, and H. R. Gite, "A research review on hyperspectral data processing and analysis algorithms," *Proceedings of the National Academy of Sciences India Section A - Physical Sciences*, vol. 87, pp. 541–555, 2017.

[35] A. Chlingaryan, S. Sukkarieh, and B. Whelan, "Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review," *Computers and Electronics in Agriculture*, vol. 151, pp. 61–69, 2018.

[36] A. Damm, E. Paul-Limoges, E. Haghighi, C. Simmer, F. Morsdorf, F. D. Schneider, C. van der Tol, M. Migliavacca, and U. Rascher, "Remote sensing of plant-water relations: An overview and future perspectives," *Journal of Plant Physiology*, vol. 227, pp. 3–19, 2018.

[37] D. Zhao, L. Huang, J. Li, and J. Qi, "A comparative analysis of broadband and narrowband derived vegetation indices in predicting lai and ccd of a cotton canopy," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 62, pp. 25–33, 2007.

[38] J. V. Stafford, "Implementing precision agriculture in the 21st century," *Journal of Agricultural and Engineering Research*, vol. 76, pp. 267–275, 2000.

[39] J. Lowenberg-Deboer and B. Erickson, "Setting the record straight on precision agriculture adoption," *Agronomy Journal*, vol. 111, pp. 1552–1569, 2019.

[40] N. Wang, N. Zhang, and M. Wang, "Wireless sensors in agriculture and food industry - recent development and future perspective," *Computers and Electronics in Agriculture*, vol. 50, pp. 1–14, 2006.

[41] S. K. Seelan, S. Laguette, G. M. Casady, and G. A. Seielstad, "Remote sensing applications for precision agriculture: A learning community approach," *Remote Sensing of Environment*, vol. 88, pp. 157–169, 2003.

[42] M. Borengasser, W. S. Hungate, and R. Watkins, *Hyperspectral remote sensing: principles and applications, CRC press*. 2007.

[43] J. Xue and B. Su, "Significant remote sensing vegetation indices: A review of developments and applications," *Journal of Sensors*, vol. 2017, 2017.

[44] G. ElMasry, M. Kamruzzaman, D. W. Sun, and P. Allen, "Principles and applications of hyperspectral imaging in quality evaluation of agro-food products: A review," *Critical Reviews in Food Science and Nutrition*, vol. 52, pp. 999–1023, 2012.

[45] J. C. M. S. Moura, C. A. V. Bonine, J. de Oliveira Fernandes Viana, M. C. Dornelas, and P. Mazzafera, "Abiotic and biotic stresses and changes in the lignin content and composition in plants," *Journal of Integrative Plant Biology*, vol. 52, pp. 360–376, 2010.

[46] R. D. Jackson, "Remote sensing of biotic and abiotic plant stress," *Annual Review of Phytopathology*, vol. 24, pp. 265–287, 1986.

[47] S. S. Hussain, M. Ali, M. Ahmad, and K. H. Siddique, "Polyamines: Natural and engineered abiotic and biotic stress tolerance in plants," *Biotechnology Advances*, vol. 29, pp. 300–311, 2011.

[48] J. Glinski, *Soil physical conditions and plant roots,CRC press*. 2018.

[49] A. Blum, *Plant breeding for stress environments,CRC press*. 2018.

[50] H. eric Nilsson, "Remote sensing and image analysis in plant," *Annu. Rev. Phytopathol.*, vol. 15, pp. 489–527, 1995.

[51] H. M. Kalaji, W. BÄ... ba, K. Gediga, V. Goltsev, I. A. Samborska, M. D. Cetner, S. Dimitrova, U. Piszcz, K. Bielecki, K. Karmowska, K. Dankov, and A. KompaÅ‚a-BÄ... ba, "Chlorophyll fluorescence as a tool for nutrient status identification in rapeseed plants," *Photosynthesis Research*, vol. 136, pp. 329–343, 2018.

[52] S. L. Ustin, D. A. Roberts, J. A. Gamon, G. P. Asner, and R. O. Green, "Using imaging spectroscopy to study ecosystem processes and properties," *BioScience*, vol. 54, pp. 523–534, 6 2004.

[53] E. Benâ€Dor, R. G. Taylor, J. Hill, J. A. M. DemattÃª, M. L. Whiting, S. Chabrillat, and S. Sommer, *Imaging Spectrometry for Soil Applications*, vol. 97, pp. 321–392. Academic Press, 2008.

[54] D. Eamus, A. Huete, and Q. Yu, *Vegetation Dynamics*. Cambridge University Press, 12 2015.

[55] S. D. A. Skidmore L. Kumar, K. Schmidt, "Imaging spectroscopy and vegetation science," *Imaging Spectrometry*, pp. 111–155, 2001.

[56] Y. Chen, J. P. Guerschman, Z. Cheng, and L. Guo, "Remote sensing for vegetation monitoring in carbon capture storage regions: A review," *Applied Energy*, vol. 240, pp. 312–326, 2019.

[57] V. Lawley, M. Lewis, K. Clarke, and B. Ostendorf, "Site-based and remote sensing methods for monitoring indicators of vegetation condition: An australian review," *Ecological Indicators*, vol. 60, pp. 1273–1283, 2016.

[58] A. Gholizadeh and V. Kopačková, "Detecting vegetation stress as a soil contamination proxy a review of optical proximal and remote sensing techniques," *International Journal of Environmental Science and Technology*, 2019.

[59] S. Jacquemoud, W. Verhoef, F. Baret, C. Bacour, P. J. Zarco-Tejada, G. P. Asner, C. Francois, and S. L. Ustin, "Prospect + sail models: A review of use for vegetation characterization," *Remote Sensing of Environment*, vol. 113, pp. S56–S66, 2009.

[60] C. Wu, Z. Niu, Q. Tang, and W. Huang, "Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation," *Agricultural and Forest Meteorology*, vol. 148, pp. 1230–1241, 2008.

[61] A. Chemura, O. Mutanga, and T. Dube, "Remote sensing leaf water stress in coffee (coffea arabica) using secondary effects of water absorption and random forests," *Physics and Chemistry of the Earth*, vol. 100, pp. 317–324, 2017.

[62] M. Corti, P. M. Gallina, D. Cavalli, and G. Cabassi, "Hyperspectral imaging of spinach canopy under combined water and nitrogen stress to estimate biomass, water, and nitrogen content," *Biosystems Engineering*, vol. 158, pp. 38–50, 2017.

[63] H. Liu, H. Zhu, and P. Wang, "Quantitative modelling for leaf nitrogen content of winter wheat using uav-based hyperspectral data," *International Journal of Remote Sensing*, vol. 38, pp. 2117–2134, 2017.

[64] K. A. Al-Gaadi, V. Patil, E. K. Tola, R. Madugundu, and S. Marey, "In-season assessment of wheat crop health using vegetation indices based on ground measured hyper spectral data," *American Journal of Agricultural and Biological Science*, vol. 9, pp. 138–146, 2014.

[65] C. J. Tucker, "Red and photographic infrared linear combinations for monitoring vegetation, remote sensing of environment," vol. 8, pp. 127–150, 2019.

[66] L. Winterhalter, B. Mistele, S. Jampatong, and U. Schmidhalter, "High-throughput sensing of aerial biomass and above-ground nitrogen uptake in the vegetative stage of well-watered and drought stressed tropical maize hybrids," *Crop Science*, vol. 51, pp. 479–489, 2011.

[67] A. A. Gitelson, Y. Gritz, and M. N. Merzlyak, "Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves," *Journal of Plant Physiology*, vol. 160, pp. 271–282, 1 2003.

[68] W. C. Bausch and R. Khosla, "Quickbird satellite versus ground-based multi-spectral data for estimating nitrogen status of irrigated maize," *Precision Agriculture*, vol. 11, pp. 274–290, 2010.

[69] C. F. Jordan, "Derivation of leaf-area index from quality of light on the forest floor," *Ecology*, vol. 50, pp. 663–666, 7 1969.

[70] M. J. López-Calderón, J. Estrada-ávalos, V. M. Rodríguez-Moreno, J. E. Mauricio-Ruvalcaba, A. R. Martínez-Sifuentes, G. Delgado-Ramírez, and E. Miguel-Valle, "Estimation of total nitrogen content in forage maize (zea mays l.) using spectral indices: Analysis by random forest," *Agriculture*, vol. 10, p. 451, 2020.

[71] G. Rondeaux, M. Steven, and F. Baret, "Optimization of soil-adjusted vegetation indices," *Remote Sensing of Environment*, vol. 55, pp. 95–107, 1996.

[72] X. H. Yang, J. F. Huang, F. M. Wang, X. Z. Wang, Q. X. Yi, and Y. Wang, "Modified chlorophyll absorption continuum index for chlorophyll estimation," *Journal of Zhejiang University: Science*, vol. 7, pp. 2002–2006, 2006.

[73] W. Zhu, Z. Sun, Y. Huang, J. Lai, J. Li, J. Zhang, B. Yang, B. Li, S. Li, K. Zhu, Y. Li, and X. Liao, "Improving field-scale wheat lai retrieval based on uav remote-sensing observations and optimized vi-luts," *Remote Sensing*, vol. 11, 2019.

[74] X. Xiao, Q. Zhang, D. Hollinger, J. Aber, and M. Berrien, "Modeling gross primary production of an evergreen needleleaf forest using modis and climate data," *Ecological Applications*, vol. 15, pp. 954–969, 2005.

[75] S. K. McFeeters, "The use of the normalized difference water index (ndwi) in the delineation of open water features," *International Journal of Remote Sensing*, vol. 17, pp. 1425–1432, 1996.

[76] K. Uto and Y. Kosugi, "Hyperspectral manipulation for the water stress evaluation of plants," *Contemporary Materials*, vol. 1, pp. 18–25, 2012.

[77] E. R. Hunt and B. N. Rock, "Detection of changes in leaf water content using near- and middle-infrared reflectances," *Remote Sensing of Environment*, vol. 30, pp. 43–54, 1989.

[78] R. Merton, "Monitoring community hysteresis red-edge vegetation stress index calculating red-edge inflection from second-derivative spectra red-edge vegetation stress index," *Stress: The International Journal on the Biology of Stress*, 1994.

[79] Y. Zhao, A. B. Potgieter, M. Zhang, B. Wu, and G. L. Hammer, "Predicting wheat yield at the field scale by combining high-resolution sentinel-2 satellite imagery and crop modelling," *Remote Sensing*, vol. 12, 2020.

[80] E. Barnes, T. Clarke, S. Richards, P. Colaizzi, J. Haberland, M. Kostrzewski, P. Waller, C. Choi, E. Riley, T. Thompson, R. J. Lascano, H. Li, and M. Moran, "Coincident

detection of crop water stress, nitrogen status and canopy density using ground-based multispectral data," 2000.

[81] A. A. Gitelson, Y. J. Kaufman, and M. N. Merzlyak, "Use of a green channel in remote sensing of global vegetation from eos- modis," *Remote Sensing of Environment*, vol. 58, pp. 289–298, 1996.

[82] J. Dash and P. J. Curran, "The meris terrestrial chlorophyll index," *International Journal of Remote Sensing*, vol. 25, pp. 5403–5413, 12 2004.

[83] J. Dash, P. J. Curran, M. J. Tallis, G. M. Llewellyn, G. Taylor, and P. Snoeijf, "Validating the meris terrestrial chlorophyll index (mtci) with ground chlorophyll content data at meris spatial resolution," *http://dx.doi.org/10.1080/01431160903376340*, vol. 31, pp. 5513–5532, 2010.

[84] A. A. Gitelson, A. V. na, V. Ciganda, D. C. Rundquist, and T. J. Arkebauer, "Remote estimation of canopy chlorophyll content in crops," *Geophysical Research Letters*, vol. 32, pp. 1–4, 4 2005.

[85] A. Gitelson, Y. Zur, O. C. Photochemistry, â€¦, and undefined 2002, "Assessing carotenoid content in plant leaves with reflectance spectroscopy," *Wiley Online Library*, vol. 75, pp. 272–281, 5 2002.

[86] D. Goffart, F. B. Abdallah, Y. Curnel, V. Planchon, P. Defourny, and J. P. Goffart, "In-season potato crop nitrogen status assessment from satellite and meteorological data," *Potato Research*, vol. 65, pp. 729–755, 9 2022.

[87] B. Bayat, C. van der Tol, and W. Verhoef, "Remote sensing of grass response to drought stress using spectroscopic techniques and canopy reflectance model inversion," *Remote Sensing*, vol. 8, pp. 1–24, 2016.

[88] J. Simunek, M. van Genuchten, and M. Sejna, "The hydrus software package for simulating the two- and three-dimensional movement in variably-saturated porous media, technical manual," *Technical manual*, p. 2:258, 2012.

[89] R. Houborg, H. Soegaard, and E. Boegh, "Combining vegetation index and model inversion methods for the extraction of key vegetation biophysical parameters using terra and aqua modis reflectance data," *Remote Sensing of Environment*, vol. 106, pp. 39–58, 2007.

[90] T. Poblete, S. Ortega-Farías, M. A. Moreno, and M. Bardeen, "Artificial neural network to predict vine water status spatial variability using multispectral information obtained from an unmanned aerial vehicle (uav)," *Sensors (Switzerland)*, vol. 17, 2017.

[91] K. Loggenberg, A. Strever, B. Greyling, and N. Poona, "Modelling water stress in a shiraz vineyard using hyperspectral imaging and machine learning," *Remote Sensing*, vol. 10, pp. 1–14, 2018.

[92] N. Panigrahi and B. S. Das, "Canopy spectral reflectance as a predictor of soil water potential in rice," *Water Resources Research*, vol. 54, pp. 2544–2560, 2018.

[93] B. J. Gutiérrez-Rodríguez, J. O. Argüello-Tovar, and O. L. García-Navarrete, "Use of vis-nir-swir spectroscopy for the prediction of water status in soybean plants in the colombian piedmont plains," *DYNA (Colombia)*, vol. 86, pp. 125–130, 2019.

[94] M. Zovko, U. Žibrat, M. Knapič, M. B. Kovačić, and D. Romić, "Hyperspectral remote sensing of grapevine drought stress," *Precision Agriculture*, vol. 20, pp. 335–347, 2019.

[95] B. Bruning, H. Liu, C. Brien, B. Berger, M. Lewis, and T. Garnett, "The development of hyperspectral distribution maps to predict the content and distribution of nitrogen and water in wheat (triticum aestivum)," *Frontiers in Plant Science*, vol. 10, pp. 1–16, 2019.

[96] L. Cotrozzi and J. J. Couture, "Hyperspectral assessment of plant responses to multiâ€stress environments: Prospects for managing protected agrosystems," *Plants, People, Planet*, pp. 244–258, 2019.

[97] J. Blancon, D. Dutartre, M. H. Tixier, M. Weiss, A. Comar, S. Praud, and F. Baret, "A high-throughput model-assisted method for phenotyping maize green leaf area index dynamics using unmanned aerial vehicle imagery," *Frontiers in Plant Science*, vol. 10, pp. 1–16, 2019.

[98] J. R. Rodríguez-Pérez, C. O. nez, A. B. González-Fernández, E. Sanz-Ablanedo, J. B. Valenciano, and V. Marcelo, "Leaf water content estimation by functional linear regression of field spectroscopy data," *Biosystems Engineering*, vol. 165, pp. 36–46, 2018.

[99] K. C. Tung, C. Y. Tsai, H. C. Hsu, Y. H. Chang, C. H. Chang, and S. Chen, "Evaluation of water potentials of leafy vegetables using hyperspectral imaging," *IFAC-PapersOnLine*, vol. 51, pp. 5–9, 2018.

[100] J. Gao, T. Liang, J. Yin, J. Ge, Q. Feng, C. Wu, M. Hou, J. Liu, and H. Xie, "Estimation of alpine grassland forage nitrogen coupled with hyperspectral characteristics during different growth periods on the tibetan plateau," *Remote Sensing*, vol. 11, 2019.

[101] M. ángel Lara, B. Diezma, L. Lleó, J. M. Roger, Y. Garrido, M. I. Gil, and M. Ruiz-Altisent, "Hyperspectral imaging to evaluate the effect of irrigationwater salinity in lettuce," *Applied Sciences (Switzerland)*, vol. 6, 2016.

[102] M. S. M. Asaari, S. Mertens, S. Dhondt, D. Inzé, N. Wuyts, and P. Scheunders, "Analysis of hyperspectral images for detection of drought stress and recovery in maize plants in a high-throughput phenotyping platform," *Computers and Electronics in Agriculture*, vol. 162, pp. 749–758, 2019.

[103] S. Hamzeh, A. A. Naseri, S. K. AlaviPanah, H. Bartholomeus, and M. Herold, "Assessing the accuracy of hyperspectral and multispectral satellite imagery for categorical and quantitative mapping of salinity stress in sugarcane fields," *International Journal of Applied Earth Observation and Geoinformation*, vol. 52, pp. 412–421, 2016.

[104] Y. Shao, H. Zhou, L. Jiang, Y. Bao, and Y. He, "Using reflectance and gray-level texture for water content prediction in grape vines," *Transactions of the ASABE*, vol. 60, pp. 207–213, 2017.

[105] R. Khanna, L. Schmid, A. Walter, J. Nieto, R. Siegwart, and F. Liebisch, "A spatio temporal spectral framework for plant stress phenotyping," *Plant Methods*, vol. 15, pp. 1–18, 2019.

[106] S. Jia, G. Tang, J. Zhu, and Q. Li, "A novel ranking-based clustering approach for hyperspectral band selection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, pp. 88–102, 2016.

[107] S. Jia, Z. Ji, Y. Qian, and L. Shen, "Unsupervised band selection for hyperspectral imagery classification without manual band removal," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, pp. 531–543, 2012.

[108] A. Martínez-Usó, F. Pla, J. M. Sotoca, and P. García-Sevilla, "Clustering-based hyperspectral band selection using information measures," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, pp. 4158–4171, 2007.

[109] A. Rodriguez and A. Laio, "Clustering by fast search and find of density peaks," *Science*, vol. 344, pp. 1492–1496, 2014.

[110] L. Wang, H. C. Li, B. Xue, and C. I. Chang, "Constrained band subset selection for hyperspectral imagery," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, pp. 2032–2036, 2017.

[111] J. Verrelst, J. P. Rivera, A. Gitelson, J. Delegido, J. Moreno, and G. Camps-Valls, "Spectral band selection for vegetation properties retrieval using gaussian processes regression," *International Journal of Applied Earth Observation and Geoinformation*, vol. 52, pp. 554–567, 2016.

[112] A. Datta, S. Ghosh, and A. Ghosh, "Band elimination of hyperspectral imagery using partitioned band image correlation and capacitory discrimination," *International Journal of Remote Sensing*, vol. 35, pp. 554–577, 2014.

[113] M. M. Awad, B. Alawar, and R. Jbeily, "A new crop spectral signatures database interactive tool (cssit)," *Data*, vol. 4, pp. 1–14, 2019.

[114] T. Slonecker, "Analysis of the effects of heavy metals on vegetation hyperspectral reflectance properties," *Hyperspectral Remote Sensing of Vegetation*, pp. 561–578, 11 2012.

[115] P. S. Thenkabail, I. Mariotto, M. K. Gumma, E. M. Middleton, D. R. Landis, and K. F. Huemmrich, "Selection of hyperspectral narrowbands (hnbs) and composition of hyperspectral twoband vegetation indices (hvis) for biophysical characterization and discrimination of crop types using field reflectance and hyperion/eo-1 data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, pp. 427–439, 2013.

[116] J. Behmann, A. K. Mahlein, T. Rumpf, C. Römer, and L. Plümer, "A review of advanced machine learning methods for the detection of biotic stress in precision crop protection," *Precision Agriculture*, vol. 16, pp. 239–260, 2015.

[117] X. Cai, F. Nie, H. Huang, and F. Kamangar, "Heterogeneous image feature integration via multi-modal spectral clustering," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1977–1984, 2011.

[118] A. Mucherino, P. Papajorgji, and P. M. Pardalos, "A survey of data mining techniques applied to agriculture," *Operational Research*, vol. 9, pp. 121–140, 2009.

[119] H. Data, Y. Chen, Z. Lin, Y. Chen, Z. Lin, X. Zhao, and S. Member, "Deep learning-based classification of deep learning-based classi fi cation of hyperspectral data," vol. 7, pp. 1–14, 2015.

[120] S. D. Afandi, Y. Herdiyeni, L. B. Prasetyo, W. Hasbi, K. Arai, and H. Okumura, "Nitrogen content estimation of rice crop based on near infrared (nir) reflectance using artificial neural network (ann)," *Procedia Environmental Sciences*, vol. 33, pp. 63–69, 2016.

[121] X. E. Pantazi, D. Moshou, R. Oberti, J. West, A. M. Mouazen, and D. Bochtis, "Detection of biotic and abiotic stresses in crops by using hierarchical self organizing classifiers," *Precision Agriculture*, vol. 18, pp. 383–393, 2017.

[122] M. Romero, Y. Luo, B. Su, and S. Fuentes, "Vineyard water status estimation using multispectral imagery from an uav platform and machine learning algorithms for irrigation scheduling management," *Computers and Electronics in Agriculture*, vol. 147, pp. 109–117, 2018.

[123] D. Moshou, X. E. Pantazi, D. Kateris, and I. Gravalos, "Water stress detection based on optical multisensor fusion with a least squares support vector machine classifier," *Biosystems Engineering*, vol. 117, pp. 15–22, 2014.

[124] J. Liu, Z. Wu, Z. Wei, L. Xiao, and L. Sun, "Spatial-spectral kernel sparse representation for hyperspectral image classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, pp. 2462–2471, 2013.

[125] V. Mazzia, L. Comba, A. Khaliq, M. Chiaberge, and P. Gay, "Uav and machine learning based refinement of a satellite-driven vegetation index for precision agriculture," *Sensors 2020, Vol. 20, Page 2530*, vol. 20, p. 2530, 4 2020.

[126] L. Wang, S. Hao, Y. Wang, Y. Lin, and Q. Wang, "Spatial-spectral information-based semisupervised classification algorithm for hyperspectral imagery," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, pp. 3577–3585, 2014.

[127] M. Cui, S. Prasad, W. Li, and L. M. Bruce, "Locality preserving genetic algorithms for spatial-spectral hyperspectral image classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, pp. 1688–1697, 2013.

[128] A. Plaza, J. A. Benediktsson, J. W. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, A. Gualtieri, M. Marconcini, J. C. Tilton, and G. Trianni, "Recent advances in techniques for hyperspectral image processing," *Remote Sensing of Environment*, vol. 113, pp. S110–S122, 2009.

[129] S. Jia, S. Jiang, Z. Lin, N. Li, M. Xu, and S. Yu, "A survey: Deep learning for hyperspectral image classification with few labeled samples," *Neurocomputing*, vol. 448, pp. 179–204, 8 2021.

[130] M. E. Paoletti, J. M. Haut, J. Plaza, and A. Plaza, "Deep learning classifiers for hyperspectral imaging: A review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 158, pp. 279–317, 12 2019.

[131] H. Singh, S. Tyagi, and P. Kumar, *Emerging Technology in Modelling and Graphics*, vol. 937. Springer Singapore, 2020.

[132] M. A. Cochrane, "Using vegetation reflectance variability for species level classification of hyperspectral data," *http://dx.doi.org/10.1080/01431160050021303*, vol. 21, pp. 2075–2087, 1 2010.

[133] IGAC, *Estudio General de Suelos y Zonificación de Tierras Departamento de Antioquia*. Instituto Geográfico Agustín Codazzi (IGAC); Imprenta Nacional de Colombia, 2007.

[134] C. Sánchez, D. Jaraba, J. Medina, J. Martínez, and A. Martínez, "Requerimientos hidricos del ají dulce (capsicum annum l.) bajo riego por goteo en el valle del sinú medio," *dialnet.unirioja.es*, pp. 11–20, 2003.

[135] M. Torres-Madronero, M. Goez, M. Guzman, M. A. Rondon, T. Carmona, P. Acevedo-Correa, and C. Casamitjana, "Spectral library of maize leaves under nitrogen deficiency stress," *Data*, 2020.

[136] D. Dorado, G. Luis, C. Grajales, G. Alexander, and R. Roa, "Requerimientos hídricos del cultivo de aguacate (persea americana) variedad hass en zonas productoras de colombia," *Corporación colombiana de investigación agropecuaria-AGROSAVIA*, 2017.

[137] L. W. Kuswidiyanto, H. H. Noh, and X. Han, "Plant disease diagnosis using deep learning based on aerial hyperspectral images: A review," *Remote Sensing 2022, Vol. 14, Page 6031*, vol. 14, p. 6031, 11 2022.

[138] S. W. Ritchie, j. j. Hanway, and G. Benson, "How a corn plant develops,," 1986.

[139] M. R. Alam, S. Nakasathien, M. S. H. Molla, M. A. Islam, M. Maniruzzaman, M. A. Ali, E. Sarobol, V. Vichukit, M. M. Hassan, E. S. Dessoky, E. M. A. El-Ghany, M. Brestic, M. Skalicky, S. V. Jagadish, and A. Hossain, "Kernel water relations and kernel filling traits in maize (zea mays l.) are influenced by water-deficit condition in a tropical environment," *Frontiers in Plant Science*, vol. 12, 10 2021.

[140] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*. Springer New York, 2009.

[141] W. L. S. sinica and 2002, "Regression tress with unbiased variable selection and interaction detection," *JSTOR*, 2002.

[142] K. He, X. Zhang, S. Ren, J. S. P. of the IEEE, and undefined 2015, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," *openaccess.thecvf.com*, 2015.

[143] W. Sun and Q. Du, "Ieee geoscience and remote sensing magazine hyperspectral band selection a review weiwei sun and qian du," *IEEE Geoscience and Remote Sensing Magazine*, vol. 7, 2019.

[144] Q. Du and H. Yang, "Similarity-based unsupervised band selection for hyperspectral image analysis," *IEEE Geoscience and Remote Sensing Letters*, vol. 5, pp. 564–568, 10 2008.

[145] Q. Wang, F. Zhang, and X. Li, "Optimal clustering framework for hyperspectral band selection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, pp. 5910–5922, 10 2018.

[146] S. Beniwal and J. Arora, "Classification and feature selection techniques in data mining," *International Journal of Engineering Research and Technology*, vol. 1, 08 2012.

[147] G. Darbellay, I. V. I. T. on Information, and undefined 1999, "Estimation of the information by an adaptive partitioning of the observation space," *IEEE Transactions on Information Theory*, 1999.

[148] C. Ding and H. Peng, "Minimum redundancy feature selection from microarray gene expression data," *Journal of Bioinformatics and Computational Biology*, vol. 3, pp. 185–205, 4 2005.

[149] S. Waugh, "Extending and benchmarking cascade-correlation: extensions to the cascade-correlation architecture and benchmarking of feed-forward supervised artificial neural," 1995.

[150] A. Frank, "Uci machine learning repository. irvine, ca: University of california, school of information and computer science," 2010.

[151] P. Mohanaiah, P. Sathyanarayana, L. Gurukumar, and A. Professor, "Image texture feature extraction using glcm approach," *International Journal of Scientific and Research Publications*, vol. 3, 2013.

[152] Y. Wang, S. Tan, X. Jia, L. Qi, S. Liu, H. Lu, C. Wang, W. Liu, X. Zhao, L. He, J. Chen, C. Yang, X. Wang, J. Chen, Y. Qin, J. Yu, and X. Ma, "Estimating relative chlorophyll content in rice leaves using unmanned aerial vehicle multi-spectral images and spectral textural analysis," *Agronomy 2023, Vol. 13, Page 1541*, vol. 13, p. 1541, 6 2023.

[153] P. K. Sethy, C. Pandey, Y. K. Sahu, and S. K. Behera, "Hyperspectral imagery applications for precision agriculture - a systemic survey," *Multimedia Tools and Applications 2021 81:2*, vol. 81, pp. 3005–3038, 11 2021.

[154] J. An, W. Li, M. Li, S. Cui, and H. Yue, "Identification and classification of maize drought stress using deep convolutional neural network," *Symmetry 2019, Vol. 11, Page 256*, vol. 11, p. 256, 2 2019.