



BEN-GURION UNIVERSITY OF THE NEGEV  
FACULTY OF ENGINEERING SCIENCES  
DEPARTMENT OF INDUSTRIAL ENGINEERING AND MANAGEMENT

# **EARLY DETECTION OF FUSARIUM INFECTION IN CORN USING SPECTRAL ANALYSIS**

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE M.Sc DEGREE

By: Tidhar Sandovsky

Advisors: Prof. Y. Edan, Prof. V. Alchanatis

February, 2021



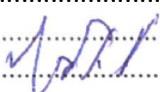
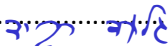
BEN-GURION UNIVERSITY OF THE NEGEV  
FACULTY OF ENGINEERING SCIENCES  
DEPARTMENT OF INDUSTRIAL ENGINEERING AND MANAGEMENT

# EARLY DETECTION OF FUSARIUM INFECTION IN CORN USING SPECTRAL ANALYSIS

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE M.Sc DEGREE

By: Tidhar Sandovsky

Advisors: Prof. Yael Edan, Prof. Victor Alchanatis

AUTHOR: .....  ..... DATE: 13/05/21  
SUPERVISOR: .....  ..... DATE: 13/05/21  
SUPERVISOR: .....  ..... DATE: 13/05/21  
CHAIRMAN OF GRADUATE STUDIES COMMITTEE: .....  ..... DATE: 18/5/2021

אוניברסיטת בן-גוריון בנגב

הפקולטה למדעי ההנדסה

המחלקה להנדסת תעשייה וניהול

## זיהוי מוקדם של פוזריום בתירס

### בעזרת ניתוח ספקטרלי

מגישה: תדהר סנדובסקי

מנחים: פר' יעל אידן, פר' ויקטור אלחנתי

מגישה: תדהר  
13/05/21 תאריך: .....  
מנחה: .....  
13/05/21 תאריך: .....  
מנחה: .....  
13/05/21 תאריך: .....  
יו"ר ועדת הוראה: .....  
18/5/2021 תאריך: .....  
גילי ירד

## ABSTRACT

This thesis presents a method based on leaf spectral reflectance to classify between infected and healthy corn plants, before infection can be visually identified on the leaves. Focus was on Fusarium, a fungus with a global distribution, capable of infecting a wide range of crop plants, including cereals such as maize, wheat or barley. It may rapidly result in very high crop losses and quality reductions. Moreover, mycotoxins, potentially generated by these fungi, are poisonous and harmful for both human and animal nutrition.

Data of 249 plants was collected from a controlled experiment conducted at Evogene's research facilities in Central Israel. Corn seeds were germinated in a greenhouse for a period of 3 weeks. Two types of spectral measurements were performed 21 days after sowing: (a) point spectral measurements with a field spectro-radiometer (FieldSpec, ASD) in the range of 300-2500 nm and (b) hyper-spectral images with a pushbroom type system (SPECIM) in the range of 400-1000 nm. After the last sampling day, the plants' roots were exposed and the disease infection level was manually examined by an agronomist and classified into 6 levels, according to the state of the roots. These levels were then aggregated into infected and healthy plants resulting with an imbalanced data set which included 154 infected and 96 healthy plants.

Several classification algorithms were developed and evaluated using the ASD data and the hyper-spectral images. To simplify the algorithm development, stepwise regression was used to determine the significant wavelengths for Fusarium detection using the point spectral measurements. The most significant wavelengths for disease detection were wavelengths 400, 440, 630, 700nm that are mainly influenced by leaf pigment content and 750, 820, 900 nm that depend on the leaf biology structure. Using these wavelengths, a random forest classifier was used to classify between infected and healthy plants using features from the hyperspectral images. The classifier was evaluated using the leave one out cross validation method.

Fusarium infection was successfully classified with 82% accuracy (67.5% True Positive and 8.7% False Positive) and 0.79 AUC. These results prove the feasibility of detecting Fusarium at a stage which currently is not possible without destroying the plant.

**Keywords:** hyperspectral imaging, spectral imaging, hyperspectral analysis, disease detection, Fusarium, random forest, corn

## Publications

### Journal

**Sandovsky, T.,** Y. Edan, V. Alchanatis. Early detection of Fusarium infection in corn using spectral analysis. (Appendix 1, Submitted January 2021).

### Conference papers and presentations

**Sandovsky, T.,** Y. Edan, S. Gad, Z. Etzioni, T. Nacson, V. Alchanatis. 2019. Early detection of Fusarium infection in corn using spectral analysis. The European Conference on Precision Agriculture, Montpellier, France. (Appendix 2)

## תקציר

תזה זו מציגה שיטה המבוססת על החזר ספקטרלי של עלי תירס לסיווג בין צמחי תירס נגועים ובריאים, לפני שניתן לזהות חזותית על העלים. התמקדנו בפטריית הפוזריום (*Fusarium*), פטרייה עם תפוצה עולמית, המסוגלת להדביק מגוון רחב של צמחי יבול, כולל דגנים כמו תירס, חיטה או שעורה. הפטרייה עלולה לגרום להפסדי יבול גבוהים מאוד ולהפחתת איכות היבול. יתר על כן, רעלנים, **mycotoxins**, נוצרים באופן פוטנציאלי על ידי פטריות אלה. רעלים אלה מזיקים הן לתזונה אנושית והן לבעלי חיים.

נתונים של 249 צמחים נאספו מניסוי מבוקר שנערך במתקני המחקר של אבוג'ן במרכז ישראל. זרעי תירס נבטו בחממה לתקופה של 3 שבועות. שני סוגים של מדידות ספקטרליות בוצעו 21 יום לאחר הזריעה: (א) מדידות ספקטרליות נקודתיות באמצעות ספקטרום-רדיומטר שדה (**FieldSpec, ASD**) בטווח של 300-2500 ננומטר ו- (ב) תמונות היפר-ספקטרליות עם מערכת מסוג **pushbroom (SPECIM)** בטווח של 400-1000 ננומטר. לאחר יום הדגימה האחרון נחשפו שורשי הצמחים ורמת זיהום המחלה נבדקה ידנית על ידי אגרונום וסווגה ל-6 רמות, על פי מצב השורשים. את 6 רמות המחלה הפכנו לשתי רמות לצמחים נגועים ובריאים, ולמרות זאת, סט הנתונים לא היה מאוזן וכלל 154 צמחים נגועים ו-96 צמחים בריאים.

כמה אלגוריתמי סיווג פותחו והוערכו באמצעות נתוני המדידות ספקטרליות הנקודתיות והתמונות ההיפר-ספקטרליות. ע"מ לבנות מסווג פשוט יותר, נעשה שימוש ברגרסיה בשלבים לקביעת אורכי הגל המשמעותיים לזיהוי פוזריום באמצעות מדידות ספקטרליות נקודתיות. אורכי הגל המשמעותיים ביותר לגילוי המחלה היו אורכי גל 400, 440, 630, 700 ננומטר המושפעים בעיקר מתכולת פיגמנט העלה (בעיקר כלורופיל) ו- 750, 820, 900 ננומטר התלויים במבנה הביולוגיה של העלה ותכולת מים. באמצעות אורכי גל אלה נבנה מסווג יער אקראי לסיווג בין צמחים נגועים ובריאים באמצעות תכונות הנבנו מתמונות היפר-ספקטרליות. המסווג הוערך על ידי שימוש בשיטת אימות הצלב.

זיהום הפוזריום סווג בהצלחה בדיוק של 82% ( $TP\ 67.5\% - FP\ 8.7\%$ ) ו- $AUC = 0.7$ . תוצאות אלו מוכיחות את היתכנות סיווג הפוזריום בשלב שאינו אפשרי כרגע מבלי להרוס את הצמח.

**מילות מפתח:** תמונה היפר-ספקטרלית, ניתוח ספקטרלי, זיהוי מחלות, פוזריום, יער אקראי, תירס

## **Acknowledgments**

This research was partially supported by the Israeli Innovation Authority, Phenomics project, fund number 580653293 and by the Rabbi W. Gunther Plaut Chair in Manufacturing Engineering at Ben-Gurion University of the Negev.

I gratefully thank:

- All researchers for their professional guidance: Prof. Stanly Rotman (BGU), Mr. Guy Fargon (BGU) and Ms. Shahar Gad (BGU).
- The Volcani Center, ARO and Ben-Gurion University for their financial support during the years 2018-2020.
- Thanks to Tslil Nacson for the technical support of the various sensors along the long sampling days.
- Evogene company for conducting the experiments for my research in their facilities as part of the Phenomics project. I would like to thank to Adi Etzioni who led the experiments and the technical team from Evogene (Noa Guedj, Aviv Himmel, Tomer Mahler, Sahar Repael) for their help in running the experiments.

I wish to express my sincere gratitude to my academic advisors, Prof. Yael Edan and Prof. Victor Alchanatis, who provided me guidance, support and insights during the research.

**Thank you all**

# TABLE OF CONTENTS

<b>1.</b>	<b>Introduction .....</b>	<b>1</b>
1.1	Problem descriptopn .....	1
1.2	Objectives.....	2
1.3	Research contribution & innovation.....	2
<b>2.</b>	<b>Literature review.....</b>	<b>3</b>
2.1	Precision agriculture .....	3
2.2	Fusarium infection .....	3
2.3	Diseases detection.....	5
2.4	Image Processing.....	7
2.5	Spectral sensors.....	7
2.6	Classification.....	11
<b>3.</b>	<b>Methods.....</b>	<b>13</b>
3.1	Overview .....	13
3.2	Experiments.....	13
3.3	Algorithms .....	17
3.4	Performance measures .....	18
<b>4.</b>	<b>Spectro-Radiometer (ASD) reflectance analysis .....</b>	<b>20</b>
4.1	Data.....	20
4.2	Methods .....	20
4.3	Results .....	25
4.4	Sensitivity analysis .....	34
<b>5.</b>	<b>Hyperspectral images analysis .....</b>	<b>37</b>
5.1	Data.....	37
5.2	Methods .....	37
5.3	Results .....	43
5.4	Sensitivity analysis .....	49
<b>6.</b>	<b>Conclusions, limitations, future work &amp; Contributions.....</b>	<b>51</b>
6.1	Conclusions.....	51
6.2	Research limitations.....	51
6.3	Future work .....	51
6.4	Contributions.....	52
<b>7.</b>	<b>Bibliography .....</b>	<b>53</b>
<b>8.</b>	<b>Appendices.....</b>	<b>61</b>
8.1	Early detection of fusarium infection in corn using spectral analysis.....	61
8.2	ECPA Conference paper .....	75
8.3	Convolutional neural networks (CNN) .....	83
8.4	Feature selection VIS/NIR range.....	90
8.5	Feature selection VIS/NIR/SWIR range .....	93
8.6	Multispectral features tuning random forest.....	101
8.7	Link to Raw data .....	110
8.8	Link to code .....	110

## LIST OF FIGURES

Figure 1: Sensors used for disease detection (Mahlein, 2016) .....	4
Figure 2: HSV color space.....	7
Figure 3: Typical spectral reflectance curve for vegetation image (source: gsp.humboldt.edu) .....	8
Figure 4: Spectral reflectance (a) of corn (b) of wheat.....	9
Figure 5: Example of hyperspectral image acquired from a green leaf .....	10
Figure 6: RGB image of plant.....	14
Figure 7: Plant labelling protocol for different disease levels according to plant root status. ....	14
Figure 8: Experiment setting (a) ASD measurement (b) Hyperspectral images measurement .....	16
Figure 9: Flowchart of the development process .....	17
Figure 10: Confusion matrix .....	18
Figure 11: ROC curve.....	19
Figure 12: Spectralon plate.....	20
Figure 13: Spectral response of white reference .....	20
Figure 14: (a) Spectral reflectance of corn plant .....	21
Figure 15: Logistic regression curve showing probability to be infected versus the spectral reflectance .....	23
Figure 16: The accuracy of the RF classifier for the different parameters' combinations.....	24
Figure 17: Mean and the 95% confidence interval of leaves' spectral reflectance of the plants .....	25
Figure 18: Density plot distributions of the predicted values of the infected and the healthy plants of logistic regression classifier using 7 wavelengths in the VIS/NIR range .....	29
Figure 19: Density plot distributions of the predicted values of the 'healthy' and the 'infected' plants of random forest classifier using all wavelengths in the VIS/NIR range .....	30
Figure 20: Density plot distributions of the predicted values of the infected and the healthy plants of logistic regression classifier using 7 wavelengths in the VIS/NIR/SWIR range .....	31
Figure 21: Density plot distributions of the predicted values of the infected and the healthy plants of random forest classifier using all wavelengths in the VIS/NIR/SWIR range .....	33
Figure 22: Sensitivity analysis of classification with increasing number of wavelengths in the in the VIS/NIR range (a) using SR data (b) using the first derivative of all individual spectra .....	36
Figure 23: Sensitivity analysis of classification classifiers with increasing number of wavelengths in the VIS/NIR/SWIR range (a) using SR data (b) using the first derivative of all individual spectra .....	36
Figure 24: SNR calculation flowchart.....	37
Figure 25: signal-to-noise ratio for each band .....	38
Figure 26: Hyperspectral images pre-processing flowchart .....	38
Figure 27: Diameter spectralon plate.....	40
Figure 28: Histogram of the white reference pixels, band 200 .....	40
Figure 29: Hyperspectral images pre-processing steps.....	41
Figure 30: (a) All plant pixels (b) pixels at the edges of the plant (c) pixels not in edges .....	42



Figure 31: Mean and the 95% confidence interval of leaves' spectral reflectance of the plants from ASD.....	44
Figure 32: Mean and the 95% confidence interval of leaves' spectral reflectance of the plants from hyperspectral images.....	45
Figure 33: Mean and the 95% confidence interval of the standard deviation of the leaves' spectral reflectance of the plants .....	46
Figure 34: Mean and the 95% confidence interval of the height/width ratio of the plants .....	46
Figure 35: The accuracy of the RF algorithm for each parameters combination .....	47
Figure 36: Density plot distributions of the predicted values of the infected and the healthy infected plants .....	47
Figure 37: The plant's average spectral signature from the hyperspectral image (blue), and from the hyperspectral image after recalculating all bands by 50nm (orange) and 150nm (green) bandwidths .....	49
Figure 38: The importance of the features .....	49
Figure 39: Convolutional neural networks layers .....	83
Figure 40: Activation functions in convolutional layer .....	84
Figure 41: An illustration of the architecture of the CNN .....	85
Figure 42: The training and validation .....	86
Figure 43: Density plot distributions of the predicted values of the infected and the healthy plants of CNN model (a) train set (b) validation set (c) test set .....	87
Figure 44: Density plot distributions of the predicted values of the infected and the healthy plants of CNN model (a) samples collected 14 days after sowing (b) samples collected 21 days after sowing .....	89

## LIST OF TABLES

Table 1: Examples of hyperspectral imaging applications in agriculture .....	12
Table 2: Plant material from the two experiments according to the different treatments .....	13
Table 3: Distribution of disease levels by treatment, experiment 1 .....	15
Table 4: Distribution of disease levels by treatment, experiment 2 .....	15
Table 5: Significant wavelengths in the VIS/NIR range .....	26
Table 6: The aggregated significant wavelengths for disease detection in the VIS/NIR range .....	27
Table 7: Significant wavelengths according to the various methods for disease detection in the VIS/NIR/SWIR range .....	27
Table 8: Aggregation of the significant wavelengths in the VIS/NIR/SWIR range for disease detection .....	28
Table 9: Confusion matrix of logistic regression classifier using 7 wavelengths in the VIS/NIR range	29
Table 10: Logistic regression classifier using 7 wavelengths results in the VIS/NIR range .....	29
Table 11: Random forest classification classifier results using all wavelengths in the VIS/NIR range	30
Table 12: Confusion matrix of logistic regression classifier using 7 wavelengths in the VIS/NIR/SWIR range.....	32
Table 13: Logistic regression classifier using 7 wavelengths results in the VIS/NIR/SWIR range.....	32
Table 14: Random forest classifier results using all wavelengths in the VIS/NIR/SWIR range .....	33
Table 15: Sensitivity analysis by changing the threshold value .....	35
Table 16: The number of images from which the bandwidth image was calculated .....	43
Table 17: Confusion matrix of RF classifier using the features .....	48
Table 18: Comparison between the different algorithms .....	48
Table 19: Classification results of the RF classifier for different multispectral bands .....	50
Table 20: Classification results of CNN model for train, validation and test set.....	87
Table 21: Confusion matrix of CNN model, test set .....	87
Table 22: Classification results of CNN model for different multispectral bands .....	88
Table 23: Confusion matrix of CNN model, V1 growth stage plants .....	88
Table 24: Classification results of CNN model for plant on different growth stage.....	89

# 1. INTRODUCTION

## 1.1 PROBLEM DESCRIPTOPN

Over the next five decades the global demand for food is projected to increase by 1.5–2 times, with the world's population crossing the six billion mark and expected to increase by another three billion (Seelan et al., 2003). Current production rates will not satisfy the demands of the world's population by 2050 (An et al., 2016). The production of high-quality food must increase with reduced inputs, along with addressing the challenges of global environmental change (Tester & Langridge 2010).

Plant diseases cause major production and economic losses in agriculture, it is estimated that the crop losses due to plant pathogens in United States result in about 33 billion dollars every year (Savary et al., 2012). To achieve food security is one of the major challenges for plant science and crop improvement in the 21st century.

*Fusarium* is a large genus of filamentous fungi, part of a group often referred to as hyphomycetes, widely distributed in soil and associated with plants which has remarked worldwide as a disease of economic importance (Windels, 2000). It is a phytopathogenic fungus with a global distribution, capable of infecting a wide range of crop plants, including cereals such as maize, wheat or barley (Jurado et al., 2006). The pathogen is capable of causing a variety of diseases: head blight or 'scab' on wheat, barley, rice, oats, Gibberella stalk and ear rot disease on maize (*Zea*) and soybean.

USDA ranks *Fusarium* Head Blight (FHB) caused by *Fusarium graminearum* in wheat as the worst plant disease to hit the US since the rust epidemics in the 1950s (Windels, 2000). Combined direct and secondary economic losses of FHB between 1993 and 2001 were \$7.67 billion (Goswami & Kistler, 2004) due to very high crop losses and quality reductions. Moreover, mycotoxins, potentially generated by these fungi, are poisonous and harmful both in human and animal nutrition (Browne et al. 2005). These compounds may be present even after removal of mycelium (Falasconi et al., 2005). Since most of them are resistant to physical and chemical treatments, they usually stay in the food during processing and storage (Falasconi et al., 2005).

The early detection of the fungal species producing mycotoxins or of the mycotoxins themselves has become very important to prevent the human and animal risk deriving from the entry of mycotoxins into the food chain (Falasconi et al., 2005). Traditional methods used for plant assessment are still time-consuming, labor intensive and destructive in nature (Busemeyer et al., 2013). Introduction of modern technologies to improve crop yield, provide information to enable better in-field management decisions, reduce chemical and fertilizer costs through more efficient application, permit more accurate farm records, increase profit margin and reduce pollution (Li et al., 2014).

The appearance of Fusarium-infected ears and plants largely changes during the development of this disease, mostly due to degradation of chlorophyll contents and pronounced water. These changes lead to pronounced variations in spectral properties of infected grains and total ears losses (Bauriegel et al., 2011). Both fungi and bacteria usually cause damages at molecular, cellular and/or tissue levels, which, in turn, can be detected as changes in the spectral signatures (West et al., 2003).

Utilization of spectral analysis for detection of fungal and bacterial diseases has been investigated (e.g., Lu et al., 2018; Zarco-Tejada et al., 2018). Moreover, some applications of hyper-spectral fungal detection are already working (Bauriegel et al., 2011). Research on Fusarium infections in wheat found occurrence of head blight can easily be recognized by spectral analysis during BBCH-stage 71–85 using hyper-spectral imaging (Bauriegel et al., 2011). To the best of our knowledge, there are no known works of early detection of Fusarium in corn.

## **1.2 OBJECTIVES**

The research objective is to develop an algorithm to classify between infected and healthy corn leaves using a hyperspectral sensor. Specific objectives are to:

1. Find significant wavelengths for Fusarium detection.
2. Develop an algorithm to classify between Fusarium infected and healthy plants.

## **1.3 RESEARCH CONTRIBUTION & INNOVATION**

Computer vision has significantly developed with innovative applications in agriculture (Abdulridha et al., 2019; Hariharan et al., 2019). Agriculture provides a great challenge for developing such technology due to the high object variability caused by the biological nature along with the unknown, dynamic and unstructured environment. This research is innovating in two aspects:

1. Application of hyperspectral imaging in the field of agriculture for organic diseases detection. This will enable detection at early stages which has been studied very little in the past.
2. Development of intelligent algorithms combined with advanced statistical algorithms to detect diseases at an early stage.

In our pre-research, we have found no such combination. We truly believe that this research provides an important development in disease detection contributing to the advance of precision agriculture.

## 2. LITERATURE REVIEW

### 2.1 PRECISION AGRICULTURE

Over the next five decades the global demand for food is projected to increase by 1.5–2 times, with the world's population crossing the six billion mark and expected to increase by another three billion (Seelan et al., 2003). Current production rates will not satisfy the demands of the world's population by 2050 (An et al., 2016). The production of high-quality food must increase with reduced inputs, but this accomplishment will be particularly challenging in the face of global environmental change and with the reduction of land availability (Tester & Langridge 2010).

Precision agriculture promotes variable management practices within a field, according to site conditions (Seelan et al., 2003). It enables better management by implementing the right decision in the right place at the right time (Mulla, 2013). Precise agriculture is defined as “a management strategy that uses information technologies to bring data from many sources for decision-making related to crop production” (Seelan et al., 2003).

The progress in sensor and information technologies together with the expansion of geographic information systems opens new opportunities for precision agriculture and plant phenotyping (Mahlein, 2016). A survey of ranging and imaging techniques for precision agriculture sensing for detection and morphological features and sensing for physiological assessment includes ultrasonic, time-of-flight, and light detection and ranging (LiDAR), structured light, color, RGBD, stereo vision cameras, thermal cameras, multispectral and hyperspectral cameras and spectroscopy sensors (Figure 1).

### 2.2 FUSARIUM INFECTION

Fusarium is a phytopathogenic fungus with a global distribution, capable of infecting a wide range of crop plants, including important main food crops such as maize, wheat or barley. Infection can be seed borne and systemic in the crop from seedling to harvest, or starting during the pollination where the silks are infected by the airborne conidia (Morales-Rodriguez et al., 2007). These pathogens survive in the soil, in infected plant debris, and inside apparently healthy seed and can affect the embryo and pericarp without visible symptoms. The disease was recorded most frequently under hot, wet climatic conditions where significant yield losses and mycotoxin accumulation in grain were reported (McLeod, 1993). Fusarium infection has been studied for decades, most of the work on *Fusarium graminearum* has focused on *Fusarium head blight*, mainly of small-grain crops, but some of the results obtained are also relevant to maize (Munkvold, 2003).

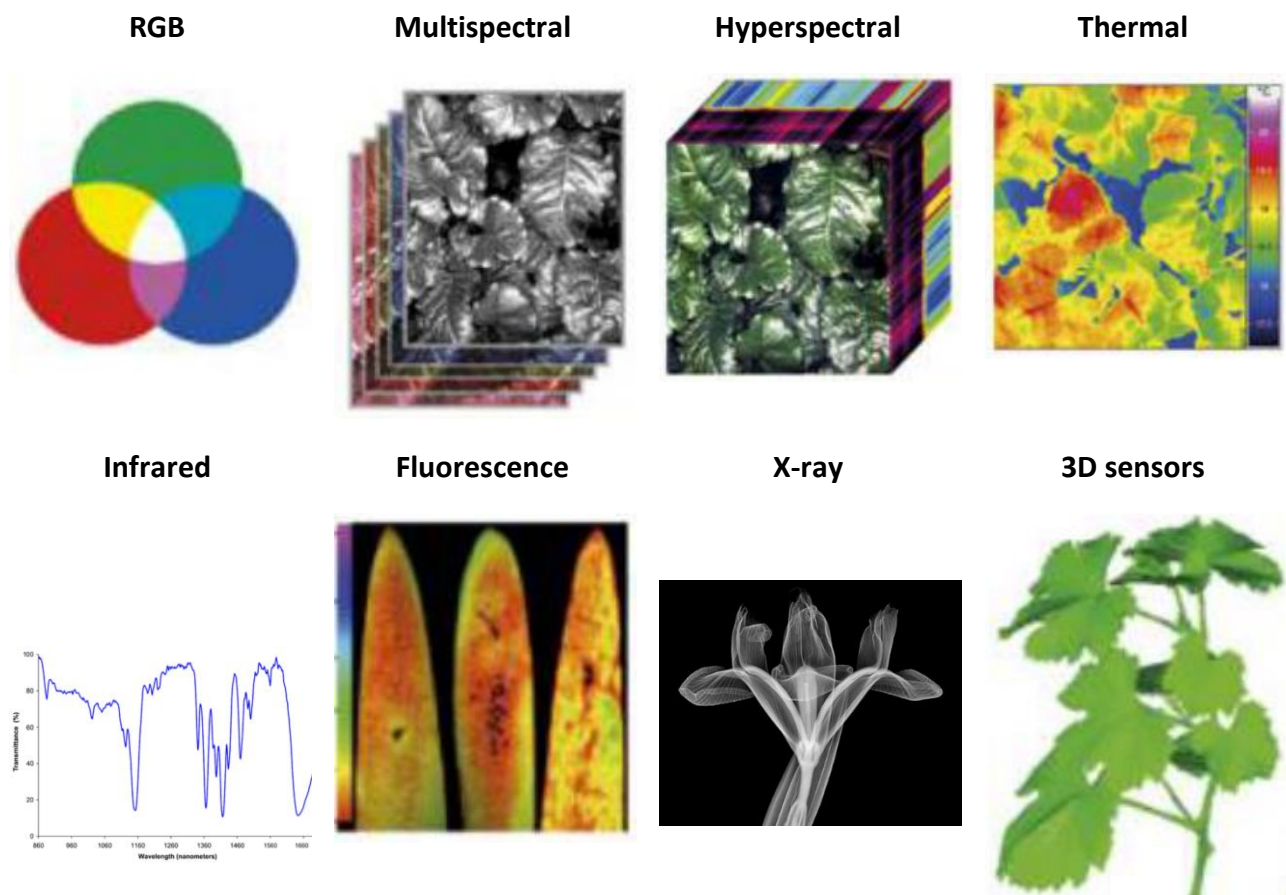


Figure 1: Sensors used for disease detection (Mahlein, 2016)

### 2.2.1 RISKS FROM FUSARIUM INFECTION

*Fusarium graminearum* (also known as *Gibberella zeae*) is the causal agent of *Gibberella ear rot*, *seedling blight* during germination, *root and stalk rot* of maize and *Fusarium head blight* of small grain crops (Goswami & Kistler 2004). Entry of *Fusarium* spp. into maize ears can occur by growth of mycelium down silks to the kernels and cob (rachis) from spores germinating on the silks or through wounds through the husk caused by insects or birds (Reid et al., 2002). When there is an epidemic in the field at a young stage, it can eliminate the entire field. Moreover, there is a critical health risk in the late stage of the rash, since *Fusarium* produces Mycotoxins. *Fusarium graminearum* is known to produce two important mycotoxins, *deoxynivalenol* (DON) and *zearalenone*, which can contaminate the diseased grain. Humans and animals, including dairy cows and beef cattle, are sensitive to those toxins. The fungal toxin zearalenone has estrogenic properties and produces many reproductive disorders in animals. To protect consumers from mycotoxicosis, by the end of 2003, approximately 100 countries (covering approximately 85% of the world's inhabitants) had specific regulations or detailed guidelines for mycotoxins in food (Van Egmond et al., 2007).

### 2.2.2 DISEASE DETECTION BY SPECTRAL ANALYSIS

The appearance of Fusarium-infected ears and plants, respectively, largely changes during the development of this disease, mostly due to degradation of chlorophyll contents and pronounced water losses. Hence, these changes lead to pronounced variations in spectral properties of infected grains and total ears (Bauriegel et al., 2011). Both fungi and bacteria usually cause damages at molecular, cellular and/or tissue levels, which, in turn, can be detected as changes in the spectral signatures (West et al., 2003). Thus, spectral analysis and, especially, spectral imaging has been found suitable for the detection of Fusarium infection (Bauriegel et al., 2011).

## 2.3 DISEASES DETECTION

It is estimated that the crop losses due to diseases in United States result is about 33 billion dollars every year (Pimentel, Zuniga, & Morrison, 2005). Early information on crop health and disease detection can enable to control diseases through proper management strategies such as pesticide applications, fungicide applications, and disease-specific chemical applications; and can improve productivity. Disease detection techniques can be broadly classified into direct and indirect methods.

### 2.3.1 DIRECT METHODS

After the onset of plant disease symptoms, the presence of disease in plants can be verified using several disease detection techniques. The most significant advance, especially in virus detection is the enzyme linked immunosorbent assay (ELISA) (Clark & Adams, 1977). The method enables the highly sensitive detection of a number of morphologically different viruses in purified preparations and in unclarified extracts of herbaceous hosts and of infected crop plants (López et al., 2003). Virus concentrations are estimated by photometric measurement of the color intensity of the hydrolyzed substrate (Clark & Adams, 1977). Various applications of ELISA are detection of potato viruses S, X, and Y (Bantari & Goodwin, 1985), detection of apple chlorotic leafspot virus (Flegg & Clark, 1979), detecting the red core disease of strawberries (Mohan, 1988) etc.

The polymerase chain reaction (PCR), developed by Carrie Molis (who won the Nobel Prize for his invention) in the early 1980s, is a method widely used in molecular biology to make many copies of a specific DNA segment. PCR is a method widely used in molecular biology for rapid replication of DNA segment. PCR has many applications for the in vitro detection and diagnosis of disease pathogens in agriculture e.g. to determine genetic variability of bean golden mosaic virus (Gilbertson et al., 1991), detection of Fusarium in cotton tissue (Moricca et al., 1998).

These molecular techniques are time-consuming and labor-intensive, and require an elaborate procedure, especially during sample preparation, collection and extraction, to obtain reliable and accurate results on plant disease detection. In addition, these techniques require consumable reagents that must be tailored to detect each specific pathogen (Sankaran et al., 2010).

### 2.3.2 INDIRECT METHODS

In indirect methods, plant diseases are identified by detecting the impact of pathogen on the physiological plant response and not through direct identification of the pathogen. When plants become diseased, they can display a range of symptoms such as colored spots, or streaks that can occur on the leaves, stems, and seeds of the plant. These visual symptoms continuously change their color, shape and size as the disease progresses.

Common methods for detection of plant diseases include visual plant disease estimation by human raters, e.g., estimate severity of Fusarium head blight of wheat (Stack & McMullen, 1998), assessment of Andean bean diseases using visual keys (Stonehouse, 1994). The disease is detected based on characteristic plant disease symptoms (e.g., lesions, blight, galls, tumors, cankers, wilts, rots, or damping-off) or visible signs of a pathogen (e.g., uredinospores of Pucciniales, mycelium or conidia of Erysiphales) (Mahlein, 2016). Reliability and accuracy are benchmarks for the performance of visual assessment ratings. Visual estimation has become more accurate and reliable due to the availability of detailed guidelines and standards used for assessment training (Bock et al., 2010). Nevertheless, visual estimation is always subject to an individual's experience and can be affected by temporal variation. This variation causes significant interrater variability and changes in interrater repeatability (Nutter et al., 1993). These time-consuming methods demand experienced individuals with well-developed skills in diagnosis and disease detection and are thus subject to human bias.

Despite availability of these techniques, there is a demand for a fast, sensitive, and selective method for detection of plant diseases. It is desirable that the plant disease detection tool should be rapid, specific to a particular disease, and sensitive for detection at the early onset of the symptoms (López et al., 2003). New technologies offer opportunity to assess disease with greater objectivity (reliability, precision, and accuracy). Various spectroscopic and imaging techniques have been studied for the detection of symptomatic and asymptomatic plant diseases (West et al., 2003). The methods include multispectral or hyperspectral imaging (detailed in 2.5.2 section), infrared spectroscopy (Purcell et al., 2009; Spinelli et al., 2006), fluorescence imaging (Chaerle et al., 2007; Moshou et al., 2005), fluorescence spectroscopy (Lins, Belasque et al., 2009; Marcassa et al., 2006), RGB (Neumann et al., 2014) and X-ray imaging (Karunakaran et al., 2004; Narvankar et al., 2009), thermal imaging or thermography (Lindenthal et al., 2007) and 3D sensors (Paulus et al., 2014).

These sensors assess the optical properties of plants within different regions of the electromagnetic spectrum and are able to utilize information beyond the visible range. They enable to detect early changes in plant physiology due to biotic stresses, since disease can cause modifications in tissue color, leaf shape, transpiration rate, canopy morphology, and plant density as well as variation in the interaction of solar radiation with plants (West et al., 2003).



## 2.4 IMAGE PROCESSING

Color information is very important for feature extraction. There are different color spaces, such as RGB (red-green-blue) and HSV (hue saturation value). The RGB color format is common in digital images since it possesses compatibility with computer displays. However, a major drawback of the RGB space is that it is not perceptually uniform. Therefore, uniform quantization of RGB space gives perceptually redundant bins and perceptually uniform holes in the color space. Ordinary distance functions defined in RGB space will be unsatisfactory since perceptual distance is a function of position in RGB space (Xin, Zhou, & Zheng, 2006). The HSV color space is commonly used in agriculture image analysis.

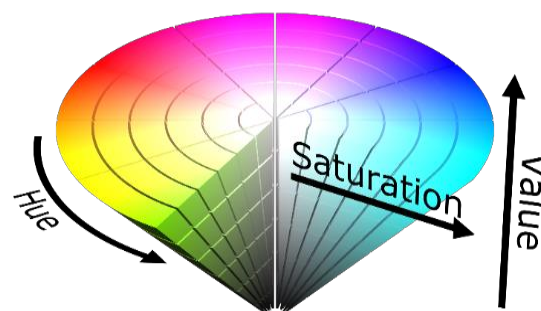


Figure 2: HSV color space

One aspect of image processing is to split the image into multiple distinct regions based on their similarity concerning some spatial or spectral properties, to reduce the image complexity for further analysis (Wu & Sun, 2013). Image segmentation algorithms have been developed based on threshold, clustering, morphological operations, edge or contour detection and watershed transformation. All these techniques can adequately perform segmentation on a single wavelength image or a monochromatic image with a scalar intensity value (spatial properties). Typically, the easiest way to perform image segmentation in the spatial domain is by using threshold methods (Mishra et al., 2017). These methods use different color spaces for thresholding (Zemmour et al. 2019).

## 2.5 SPECTRAL SENSORS

Reflectance is defined as the ratio between the returned radiation by that surface and the radiation received to the surface. The ratio is influenced by the chemical characteristics of the object, the micro topographical surface of the object and the incident's angle of the light source.

In the earth's atmosphere, water vapor absorbs many wavelengths of IR energy, while others are not absorbed. Those sections of the electromagnetic spectrum that it does not absorb, are like windows in the atmosphere, allowing electromagnetic energy to flow freely in and out of the system.

It has been suggested that spectral derivatives have important advantages over spectral reflectance, such as their ability to reduce variability due to changes in illumination reflectance (Blackburn, 2007). A Savitzky–Golay filter is a digital filter that can be applied to a set of digital data points for the purpose of smoothing the data, that is, to increase the precision of the data without distorting the signal tendency (Savitzky & Marcel, 1964).

### 2.5.1 VEGETATION SPECTRAL BEHAVIOR

Spectral information is related to the chemical composition of the plant, i.e., its physiological status (Behmann et al., 2015). The wavelengths are defined as follows:

- Blue - wavelengths under 500nm
- Green - wavelengths between 500nm and 600nm
- Red - wavelengths between 600nm and 700nm
- NIR - wavelengths between 700nm and 1000nm
- SWIR - wavelengths above 1000nm.

The visible range (VIS 400 to 700 nm) is mainly influenced by leaf pigment content, the near-infrared reflectance (NIR 700 to 1,000 nm) depends on the leaf structure, internal scattering processes, and on the absorption by leaf water, and the short-wave infrared (1,000 to 2,500 nm) is influenced by the composition of leaf chemicals and water (Jacquemoud & Ustin, 2001).

Various combinations between these wavelengths are defined as Index Vegetation (Xue & Su 2017). The index is constructed as a numerical value that defines the power of a general phenomenon, which is too complex to break down into specific known characteristics, found to be well correlated with various vegetation parameters including green leaf area, biomass, percent green cover, productivity, and photosynthetic activity (Huete, 1988). The vegetation indices should maximize their sensitivity to the physical parameters of the plant and normalize external influences (such as the angle of the sun), normalize internal effects (such as the return of radiation not from the measured object) (Bannari et al., 2009). A specific biophysical index that can be measured in order to conduct quality control of the index. The most common of these indices utilize red and near infrared canopy reflectance's or radiances in the form of ratios (normalized difference and ratio vegetation indices) or in linear combination.

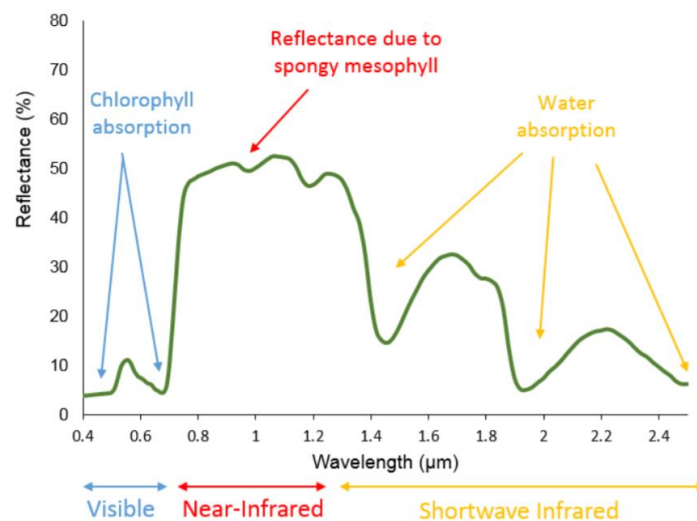
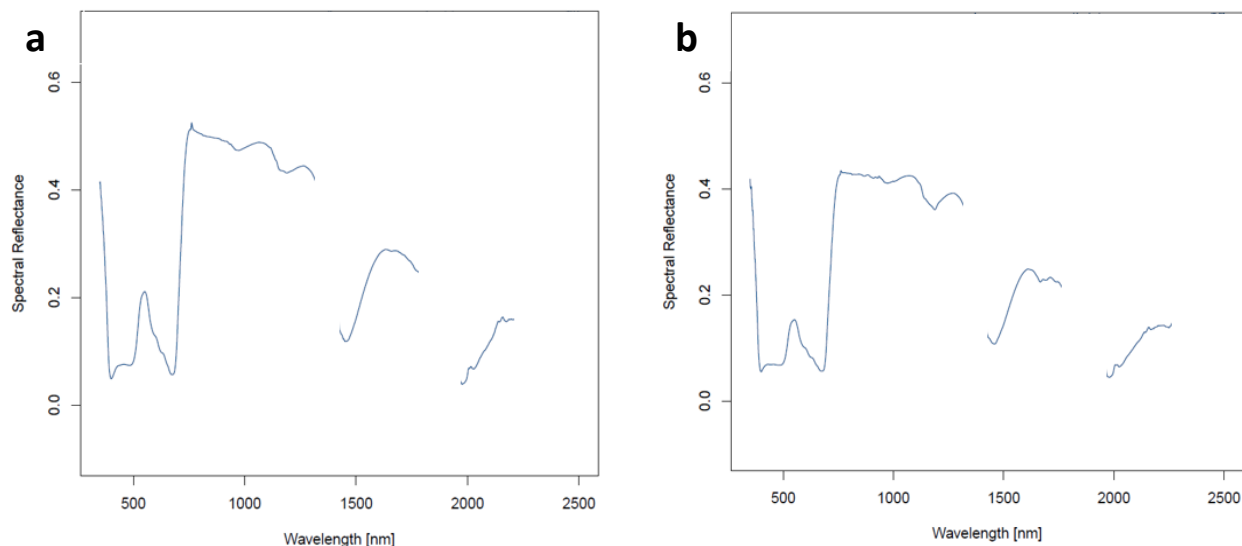


Figure 3: Typical spectral reflectance curve for vegetation image (source: [gsp.humboldt.edu](http://gsp.humboldt.edu))

Each material reflects differently the electromagnetic radiation (Figure 4). For vegetation, there is a known difference between the red spectrum (around 650 nm) and the NIR spectrum (around 750 nm) reflected radiation due to chlorophyll characteristics.



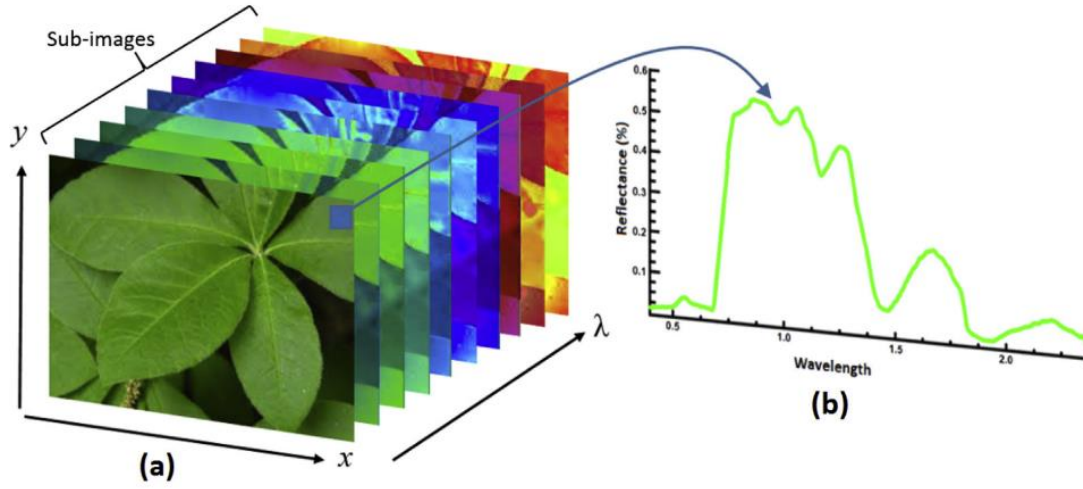
**Figure 4: Spectral reflectance (a) of corn (b) of wheat**

### 2.5.2 HYPERSPECTRAL IMAGING

Hyperspectral sensing is the ability to acquire information from an area by separating the spectral dimension of the electromagnetic radiation returning from the area into narrow and continuous bands (Shaw & Manolakis 2002). An HSI (Hyper Spectral Imaging) system integrates a spectrograph that records reflectance in a wide range of the spectrum, including the ultraviolet (UV), visible (VIS) and near-infrared (NIR) into a digital sensor. Data is generated in the form of a 3D spatial map of spectral variation: the first two dimensions provide the spatial information and a third-dimension accounts for the spectral information (Mishra et al., 2017) (Figure 5). For each wavelength there is a separate matrix indicating the spectral reflectance at each point in the image.

Being an integration of imaging and conventional spectroscopy, hyperspectral images obtain complementary information from both domains. The conventional spectroscopy exploits the fact that all materials reflect, absorb, and emit electromagnetic energy, at specific wavelengths, in distinctive patterns related to their molecular composition (Manolakis & Shaw 2002). The imaging provides spatial information and enables the use of fast and efficient image algorithms for classification. In combination, hyperspectral images have the potential to extract integrated spatial and spectral information related to the plant's functional dynamics regarding both structure and physiology (Mishra et al., 2017). The application of HSI can be found in diverse domains of research such as medical (Lu & Fei 2014), remote sensing (Goetz, 2009), food quality and safety control (Feng

& Sun 2012), archaeology and art conservation (Liang, 2012), crime scene detection (Schuler et al. 2012), etc.



**Figure 5: Example of hyperspectral image acquired from a green leaf**  
**(a) Stack of narrow band sub- images forming a 3-D hypercube**  
**(b) Reflectance spectrum of a particular pixel (Mishra et al., 2017)**

#### 2.5.2.1 Noise extraction

Hyperspectral images can be very noisy, and the amount of noise may differ from band to band (Karami, Heylen, & Scheunders, 2015). Therefore, noise reduction is an important preprocessing step to analyze the information in the hyperspectral image (Mishra et al., 2017). The signal-to-noise ratio (SNR) is a measure that compares the level of a desired signal to the level of background noise. The definition models SNR based on the mean signal of the time series and the standard deviation of the noise in the time series,  $SNR = \frac{\bar{s}}{\sigma_N}$ . As such, the global signal level, comprised of the baseline and activation, is related to the noise.

#### 2.5.2.2 Hyperspectral imaging in agriculture

In agriculture, vegetation monitoring has been studied using hyperspectral images for many years (Blackburn, 2007), and has motivated the use of HSI for exploring plants at close range. Individual plant pixels captured by HSI provide vast spectral information (multiple wavelengths). Various emerging applications of hyperspectral imaging (Table 1) HSI are abiotic stress detection (Gowen et al. 2009; Rapaport et al. 2015), biotic stress detection (Bauriegel et al., 2011; Bravo et al., 2003), species identification (Kumar et al. 2010), foliar biochemistry estimation (Yu et al., 2014) and ripeness classification (Lleó et al., 2011).

## 2.6 CLASSIFICATION

### 2.6.1 GENERAL

Classification algorithms are predictive calculations used to assign data to preset categories by analyzing sets of training data. Classification algorithms in machine learning use input training data to predict the likelihood that subsequent data will fall into one of the predetermined categories (Kotsiantis et al., 2006). There are various types of classification algorithms: logistic regression, naïve bayes classifier, K-nearest neighbors, decision tree (include random forest), support vector machines (SVM) and convolutional neural network (CNN) etc.

### 2.6.2 RANDOM FOREST

Random forest (RF) algorithm is an ensemble method for classification and regression that operates by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees (Akar & Güngör, 2012). This strategy makes the random forest algorithm good in accuracy, it is also very fast and robust against overfitting, and it is possible to form many trees (Akar & Güngör, 2012). The significant advantages of random forest algorithms using multispectral images is the accuracy (Ghose, et al. 2010). The performance of RF algorithm using multispectral images revealed their advent (e.g., detect the effects of oil pollution on vegetation (Ozigis et al., 2020); detect tobacco mosaic virus, Zhu et al., 2016) and hence it was employed in this research.

To initialize the RF algorithm, three parameters must be defined:

- The maximum depth of the trees.
- The number of features to consider when looking for the best split.
- The number of trees in the forest.

Applications	GT	Plants	Domain	Platform	Spectral Range (nm)	Features	Data Analysis	Reference
Abiotic stress Detection	Water status	Grapevine	Greenhouse, vineyard	Ground-based system	400-1,700		PLS	(Rapaport et al., 2015)
Abiotic stress Detection	Freeze	Mushroom	Laboratory	Ground-based system	400-1,000		PCA	(Gowen et al. 2009)
Biotic stress Detection	Infection estimated	Wheat	Laboratory	Ground-based system	400-1,000	HBI	PCA	(Bauriegel et al., 2011)
Biotic stress Detection	Infection estimated	Wheat	Ambient conditions	Self-made buggy	460-900	543, 630, 750, 861 nm	F-test	(Bravo et al., 2003)
Foliar biochemistry estimation	Chemical analysis	Pepper	Laboratory	Mobile platform	400-1,000	992, 756, 749, 918, 909, 921, 758, 912 nm	PLS	(Yu et al., 2014)
Foliar biochemistry estimation	Nitrogen Content	Rice	Field	Ground-based system	400-1000		PLS	(Onoyama, Ryu, Suguri, & Iida, 2013)
Phenotyping	Host-pathogen interactions	Barley	Greenhouse	Ground-based system	400-1000		SiVM	(Kuska et al., 2015)
Phenotyping	Biochemical traits	Maize	Greenhouse		550-1750		PLS	(Ge, Bai, Stoerger, & Schnable, 2016)
Ripeness classification	Ripeness	Persimmon	Laboratory	Ground-based system	400-1000	518, 711, 980 nm	SPA	(Wei, Liu, Qiu, Shao, & He, 2014)
Ripeness classification	Ripeness	Peach	Laboratory		400-1000	Ind2	ANOVA	(Lleó et al., 2011)
Species identification	Species	Eucalyptus	Laboratory	Ground-based system	400-2400	550, 630, 800 nm	F-test	(Kumar et al. 2010)
Species identification	Species	Tree - honey locust, maple, oak	Urban environment	Airborne	400-980		LiDAR	(Sugumaran & Voss, 2007)

**Table 1: Examples of hyperspectral imaging applications in agriculture**

## 3. METHODS

### 3.1 OVERVIEW

This section presents:

- The data collection process which included two experiments.
- The algorithms used to derive significant wavelengths for Fusarium detection
- The method developed to extract information from hyperspectral images
- Classification algorithms developed.

Spectral measurements were collected using two different systems:

- Spectro-Radiometer (ASD), that provides the spectral reflectance of the plant at a single point.
- A hyperspectral camera that provides a multi-dimensional image of the spectral reflectance of the whole plant.

### 3.2 EXPERIMENTS

Data was collected from two experiments conducted at Evogene's research facilities in central Israel (31° 52' 55.80" N 34° 50' 30.77" E), in June 2019 and in November 2019.

#### 3.2.1 PLANT MATERIAL

Commercial corn seeds (7210) were inoculated with Fusarium and germinated in a greenhouse for a period of 3 weeks. In each experiment, three treatments were applied at sowing in order to create variation in the disease severity level of the plants. Three levels of infection were created by drenching with spores of Fusarium Graminearum. In addition, a control group was sown, these seeds were not infected and used as a control group. The seeds were germinated in 380 ml pots. The number of seeds that sowed in each experiment and in each treatment described in Table 2. Also the number of seeds germinated and sampled.

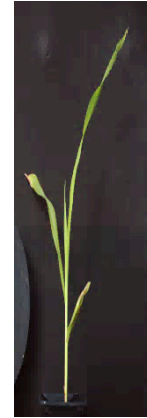
Exp. 1 - June 2019			Exp. 2 - November 2019		
Treatment	Sown	Germination	Treatment	Sown	Germination
Fus 10 <sup>5</sup>	24	21	Fus 10 <sup>6</sup>	50	34
Fus 10 <sup>3</sup>	24	23	Fus 10 <sup>5</sup>	50	43
Fus 10 <sup>2</sup>	24	20	Fus 10 <sup>4</sup>	50	36
Untreated	24	24	Untreated	50	49
Total	96	87	Total	200	162

**Table 2: Plant material from the two experiments according to the different treatments**

In each experiment the greenhouse temperature was maintained at  $24\pm 2^{\circ}\text{C}$  during the day and at  $20\pm 2^{\circ}\text{C}$  during the night. Drip irrigation with no fertilization (inhibits fusarium infection) was applied, with the irrigation frequency defined according to plot weight at 50% water content.

### 3.2.2 EXPERIMENTAL AND MEASUREMENT PROCEDURE

In each experiment, the plants were sampled 21 days after seeding, defined as stage V2 according to the BBCH scale (Lancashire et al., 1991). The sampling was conducted between 09:00 to 12:30 local time (UTC-2). At each sampling day, all plants were sampled with both the spectro-radiometer and with the hyperspectral camera. One day after the sampling day, 22 days after seeding, the roots of the plants were exposed and the disease severity level was manually classified into 5 levels by an agronomist expert where 0 is healthy, 5 is an infected plant (Figure 7).



**Figure 6: RBG image of plant 21 days after seeding**



**Figure 7: Plant labelling protocol for different disease levels according to plant root status.**

Due to the small number of samples and the imbalance between the samples at each level (described in Table 3, the number of infestation levels were aggregated to infected and healthy: levels 0 and 1 were aggregated to healthy, and levels 2 to 5 were aggregated to infected.



### 3.2.2.1 Experiment 1

This experiment performed in June 2019 included 96 plants of which 88 germinated.

Treatment	Number of sown plants	Number of germinated plants	Number of plants that labeled as DS = 0	Number of plants that labeled as DS = 1	Number of plants that labeled as DS = 2	Number of plants that labeled as DS = 3	Number of plants that labeled as DS = 4	Number of plants that labeled as DS = 5
Fus 10 <sup>5</sup> / 6	24	21	0	3	8	9	1	0
Fus 10 <sup>3</sup> / 5	24	23	0	1	11	9	2	0
Fus 10 <sup>2</sup> / 4	24	20	1	4	5	7	2	1
Untreated	24	24	10	12	2	0	0	0
Total	96	88	11	20	26	25	5	1

**Table 3: Distribution of disease levels by treatment, experiment 1**

The disease severity distribution (Table 3) reveals imbalance between the different disease levels of the plants.

Each plant was sampled once with the ASD and twice with the hyperspectral camera. After all the plants were sampled once, they were all sampled again. This enabled to enlarge the dataset for training the classification algorithms.

### 3.2.2.2 Experiment 2

The experiment performed in November 2019 included 200 sown plants sown, of which 162 germinated.

Treatment	Number of sown plant	Number of germinated plants	Number of plants that labeled as DS = 0	Number of plants that labeled as DS = 1	Number of plants that labeled as DS = 2	Number of plants that labeled as DS = 3	Number of plants that labeled as DS = 4	Number of plants that labeled as DS = 5
Fus 10 <sup>5</sup> / 6	50	34	1	4	8	12	9	0
Fus 10 <sup>3</sup> / 5	50	43	3	1	18	11	8	2
Fus 10 <sup>2</sup> / 4	50	36	3	4	9	11	9	0
Untreated	50	49	46	3	0	0	0	0
Total	200	162	53	12	35	34	26	2

**Table 4: Distribution of disease levels by treatment, experiment 2**

In this experiment there was also no balance between the different disease levels of the plants. There are more plants with high disease severity level (2- 5) than plants with low disease level (0, 1). Despite the aggregation of the level of disease, there is no balance between the 2 classes- infected and healthy.

### 3.2.3 SPECTRAL MEASUREMENT SYSTEM

#### 3.2.3.1 Spectro-Radiometer (ASD) reflectance

Spectral measurements were performed by point spectral measurements with a spectro-radiometer (ASD FieldSpec 4 hi-res, ASD Inc. Malvern Panalytical, Boulder, Colorado, USA) with a 25° field of view in the range of 350-2500 nm with 3 nm spectral resolution in the visible and near- infrared (VNIR) range and 8 nm in the short- wave infrared (SWIR) range. All measurements were conducted on the last leaf that was fully developed, at fixed locations at approximately 10 mm distance from the leaf surface, resulting in an effective sampling area of about 60 mm<sup>2</sup>

A 100 mm diameter spectralon plate was used as a white reference, to obtain reflectance curves. Each recorded spectrum was the average of four sequential measurements at the same spot, where each measurement was the average of 30 full spectrum scans. Total acquisition time for each leaf was about 10 s.

#### 3.2.3.2 Hyperspectral images

Spatial spectral measurements were performed by the V10E hyperspectral camera (SPECIM, Oulu, Finland). With a CMOS detector in the VIS and VNIR ranges of 400-1,000 nm. Full spectral range can be acquired with 150 fps at 1,312 spatial location sand up to 100 Hz with higher spatial resolution of 1,775 pixels. The spectral resolution is 3 nm, and the maximum imaging size is 1,312 × 1,775 (spatial pixels) pixels, each pixel size is 8 × 8 μm. All measurements were conducted at locations at approximately 1,500 mm distance from the plant with black background (black Foam Sheet). A 100 mm diameter spectralon plate was used as a white reference, to obtain reflectance curves in each image. Total acquisition time for each plant was about 40s.

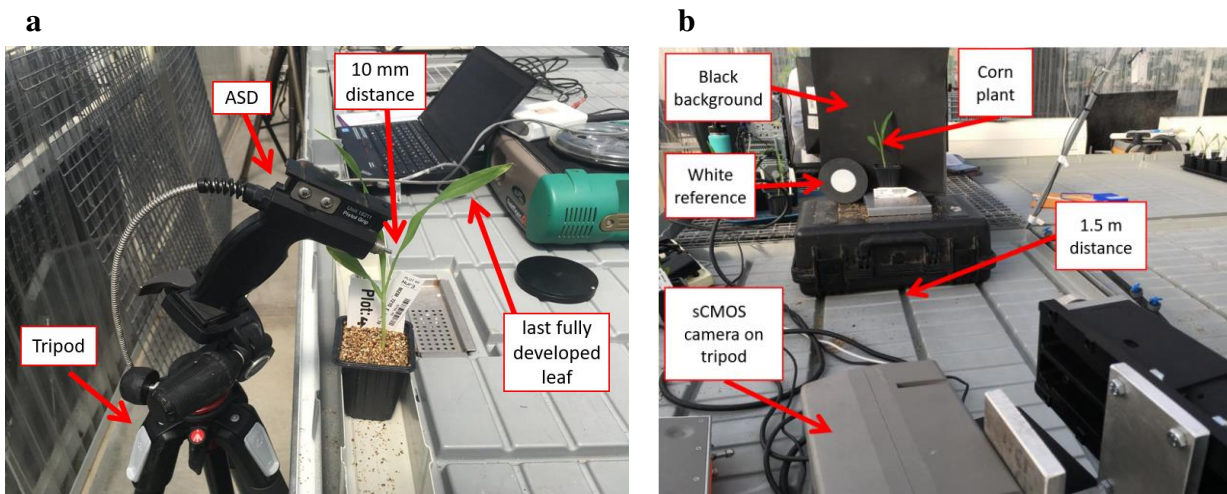


Figure 8: Experiment setting (a) ASD measurement (b) Hyperspectral images measurement

### 3.3 ALGORITHMS

The image processing and analysis were performed in Python, RStudio, and MATLAB environments.

The different algorithms (Figure 9), for both the data from ASD and the images from the hyper-spectral camera, were developed using data combined from the two experiments. The following analyses were performed with details in ASD (chapter 4) and hyperspectral camera (chapter 5).

#### 3.3.1 SPECTRO-RADIOMETER REFLECTION ANALYSIS (ASD)

The ASD data was used to: (1) derive the most significant wavelengths for Fusarium detection; (2) develop algorithms to classify between infected and healthy plants (based on these wavelengths).

(1) Significant wavelengths were determined with several methods:

1. T-test
2. Analysis of Variance (Anova)
3. Stepwise regression
4. Random forest
5. Partial Least Squares Regression (PLSR)

(2) Two algorithms to classify between infected and healthy plants:

1. Logistic regression
2. Random forest

#### 3.3.2 HYPERSPECTRAL IMAGES ANALYSIS

The hyper-spectral image was used to develop algorithms to classify between infected and healthy plants. Random forest algorithms were composed using different features from the multispectral images.

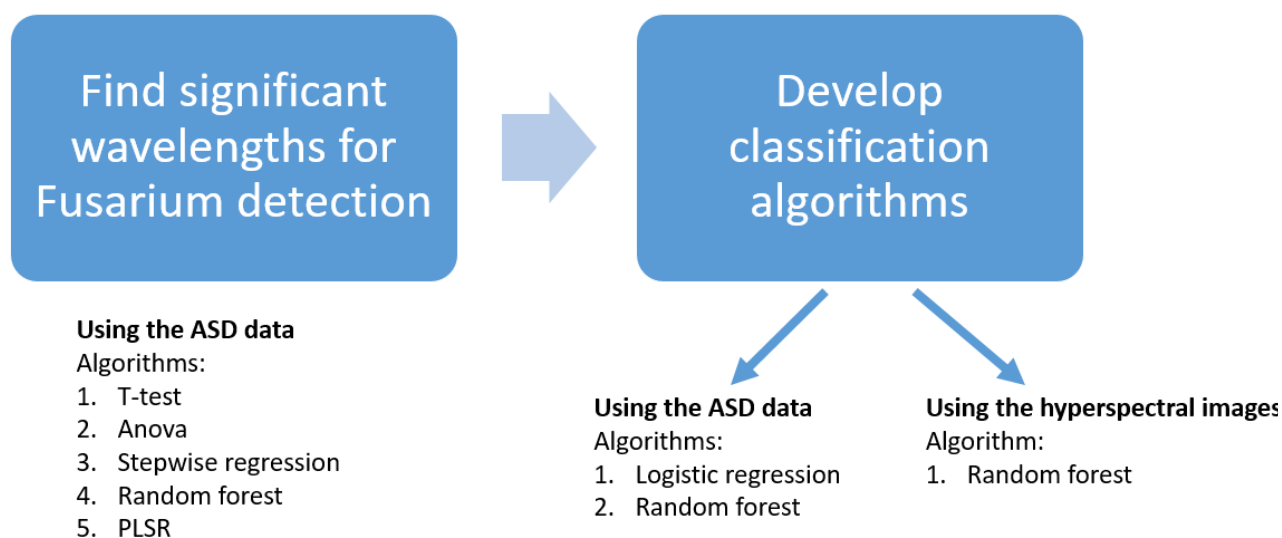


Figure 9: Flowchart of the development process

### 3.4 PERFORMANCE MEASURES

The performance of the various classifiers was assessed using leave-one-out [LOO] cross validation method. LOO cross-validation is a special case of cross-validation where the number of folds equals the number of instances in the data set. is applied once for each instance, using all other instances as a training set and using the selected instance as a single-item test set similar to jack-knife estimation (Efron, 1982). In this method all observations are used both to build the classifier and to evaluate it. The classifier is examined with many observations, however there is no final classifier for examination. In each iteration a different classifier is built, fitted to the data in that iteration.

The classifiers performance was evaluated by two metrics: accuracy and AUC. The metrics were calculated using the following values:

- True Positive (TP)- the proportion of actual positives that were correctly identified. The healthy plants which were predicted as healthy.
- False Positive (FP)- The rate of infected plants which were predicted as healthy.
- True Negative (TN)- the proportion of actual negatives that were correctly identified. The infected plants which were predicted as healthy.
- False Negative (FN)- The rate of healthy plants which were predicted as infected.

		Ground truth	
		0	1
Prediction	0	TP	FP
	1	FN	TN

**Figure 10: Confusion matrix**

**Accuracy:** Accuracy is defined as the fraction of predictions that the model classified correctly

$$(Eq.1). Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

**Equation 1**

In addition, a 95% confidence interval is calculated for the accuracy to determine whether there is a statistically significant difference between the different models (Eq.2).

$$confidence\ interval : [ acc \pm 1.96 * \sqrt{\frac{acc * (1 - acc)}{n}} ]$$

**Equation 2**

**AUC:** The Area under the Curve (AUC) of Receiver Operating Characteristic (ROC) curve (Tsenkova et al, 2009). The ROC curve is the plot of True Positive (TP) on y-axis versus False Positive (FP) on the x-axis. Each point on the curve describes the TP and FP for a specific probability threshold of classification. The AUC is the area under this curve, a value that quantifies how the classifier is capable of distinguishing between classes. The higher the AUC, the better the classifier is at predicting the actual value (in our research, distinguishing between infected plants and healthy plants).

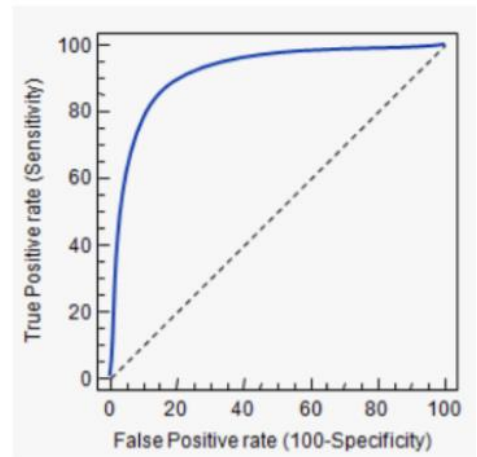


Figure 11: ROC curve

## 4. SPECTRO-RADIOMETER (ASD) REFLECTANCE ANALYSIS

### 4.1 DATA

The point spectral reflectance of a total of 162 infected and 97 healthy plants were used for the classifier development from the two experiments. Each sample includes 2150 wavelengths from the spectral reflectance of the plant in the range 350- 2,500 nm.

### 4.2 METHODS

The difference between the classifiers using the entire spectrum of ASD (VIS/NIR/SWIR range, 350 - 2,500 nm) was tested, and compared to classifiers that used only the existing wavelengths in the hyperspectral camera (VIS/NIR range, 400 - 1,000 nm), using the spectral reflectance and using the smoothed first derivative of the spectral reflectance.

#### 4.2.1 PRE- PROCESSING

##### 4.2.1.1 Noise reduction

The water vapor atmospheric absorption windows were removed from the data by analyzing the image obtained from the spectralon plate (Figure 12). Accordingly the following wavelengths of the water vapor atmospheric absorption windows were removed from the spectral data: 1350-1410nm, 1800-1950nm and 2250-2500nm (Figure 13).

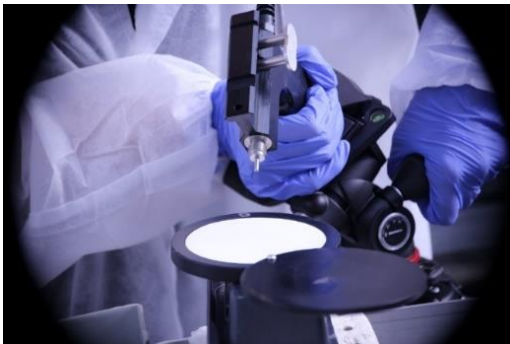


Figure 12: Spectralon plate

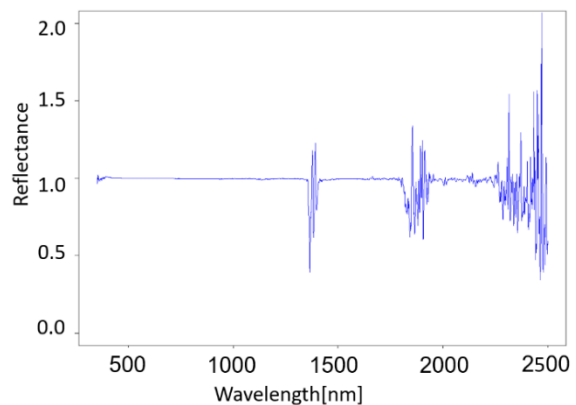
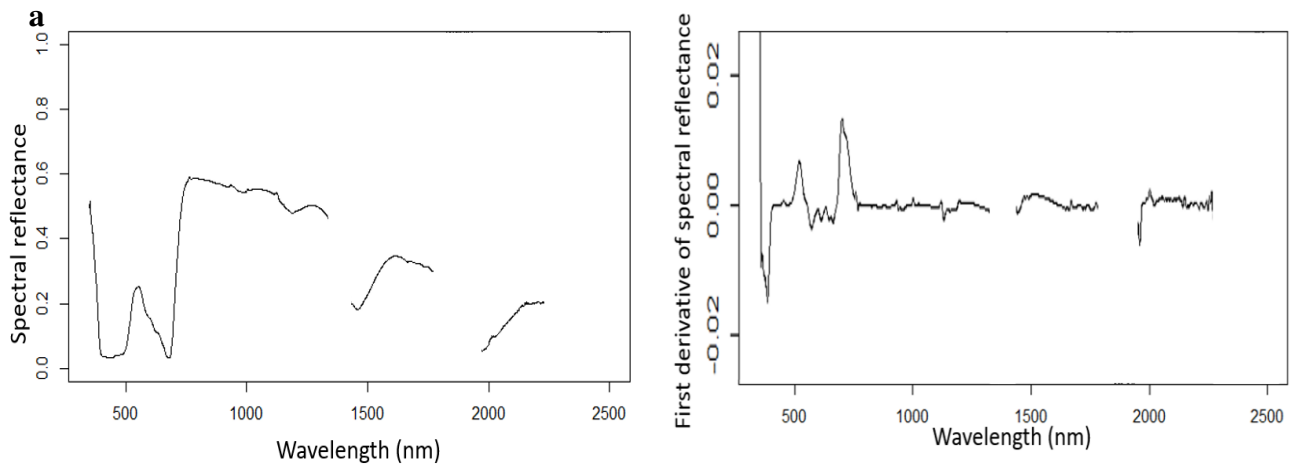


Figure 13: Spectral response of white reference

##### 4.2.1.2 Savgol

Spectral derivatives were used to reduce variability due to changes in illumination reflectance (Blackburn, 2007). To minimize noise from the environment, the first derivative of the spectral reflectance was calculated, and simultaneous smoothing was applied with a 2 order polynom for

smoothing and a filter length of 11 (window in which the smoothing and derivation was performed) wavelengths (using the Savgol function implemented with R software<sup>1</sup>).



**Figure 14: (a) Spectral reflectance of corn plant  
(b) the smoothed first derivative of the spectral reflectance**

#### 4.2.2 WAVELENGTH SELECTION

The significant wavelengths for disease detection were selected using T-test, Anova, random forest, Partial Least Squares (PLS) and stepwise regression.

The importance of each of the wavelengths (the significant wavelengths obtained from the different methods) was analyzed using stepwise regression on the spectral reflectance (SR data) and the first derivative of the spectral reflectance of all individual spectra (R1D data). Stepwise regression considers the relationship between the different wavelengths which together are significant for Fusarium detection. Each wavelength was rated for its level of importance.

##### 4.2.2.1 T-test

T-test was used to determine at which wavelengths is the most significant difference expressed (Chaves et al., 2009), between Fusarium infected and healthy plants using Eq.3.

##### Equation 3

$$t = \sqrt{\frac{\bar{X}_1 - \bar{X}_2}{s}},$$

When the variance of the two data sets,  $s_1$  and  $s_2$ , cannot be assumed to be equal (the size of the two samples may or may not be equal), the variance of the set,  $s$ , is estimated by equation 4.

<sup>1</sup> <https://www.rdocumentation.org/packages/pracma/versions/1.9.9/topics/savgol>

**Equation 4**

$$s = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}},$$

where  $s_i^2$  is the unbiased estimator of the variance of each of the two sets with  $n_i$  the number of samples in set  $i$ .

**4.2.2.2 Analysis of variance (Anova)**

Anova was used to determine at which wavelengths the most significant difference between the six different disease levels occurs. The F statistic was applied to test whether the population means were different was calculated as Eq.5.

**Equation 5: Analysis of variance**

$$F = \frac{MS_{Between}}{MS_{Within}}$$

Where MSE (Mean Square Between) denote the sample variance between groups and MSW (Mean Square Within) is the sample variance within groups.

**4.2.2.3 Random forest**

The wavelengths' importance within the random forest algorithm, was calculated by the "Gini impurity" metric (Zhi, et al. 2018). Gini impurity is a measure of how often a randomly chosen element from a set would be incorrectly labeled if it was randomly labeled according to the distribution of labels in the subset. If we have  $C$  total classes and  $p(i)$  is the probability of picking a point from the data set with class  $i$ , then the Gini Impurity is calculated as:

**Equation 6: GINI impurity**

$$G = \sum_{i=1}^C p(i) \times (1 - p(i))$$

The quality of each split was determined by weighting the impurity of each branch. The weight was defined as the number of elements the branch was divided into by the total number of the elements. The perfect split will reach impurity of 0. The gain between the Gini impurity of data set and the Gini impurity after a specific split is defined the "Gini gain". The higher the Gini gain, the better the split.

**4.2.2.4 Partial least squares (PLS)**

PLS is a method that uses data compression to reduce the large number of collinear measured spectral variables, to a few orthogonal latent variables (LV's) (REF). It is employed in spectral remote sensing for studying vegetation and soil characteristics (Atzberger et al., 2010). The LVs represent the relevant structural information, which is present in the reflectance measurements to predict the dependent variable. PLS regression uses component projection successively to find latent structures



(REF). Visual inspection of score-plots and validation residual variance plots was used to find the optimal number of LVs, to prevent over-fitting. In most cases, this procedure can reduce the number of spectral variables to a few independent LVs.

In the PLS method the most influencing features were selected to maximize the spectrum variance. To understand which bands contribute more, the variable importance in projection (VIP) score was calculated (Wold, 1993). VIP scores summarize the influence of individual X-variables on the PLS algorithm. VIP scores are calculated as the weighted sum of squares of the PLS weights, which take into account the amount of explained y-variance in each extracted latent variable (dimension/component) (Chong & T, 2005).

#### 4.2.2.5 Stepwise regression

Stepwise methods for variable selection are widely used in applications. (Yamashita, et al. 2007) used stepwise regression and applied the Akaike Information Criterion (AIC). AIC estimates the relative amount of information lost by a given model: the less information a model loses, the higher the quality of that model. In estimating the amount of information lost by a model, AIC deals with the trade-off between the goodness of fit of the model and the simplicity of the model.

#### Equation 7

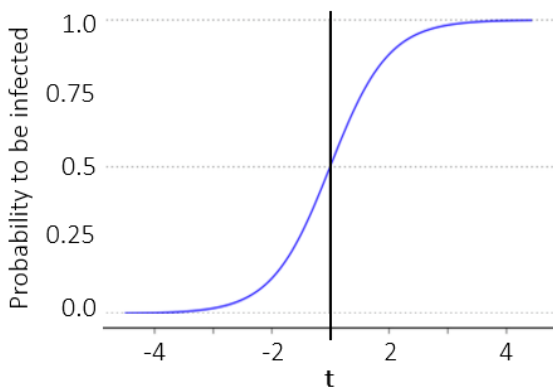
$$AIC = 2k - 2\ln(\hat{L})$$

where,  $k$  is the number of estimated parameters in the model and  $\hat{L}$  is the maximum value of the likelihood function for the model.

### 4.2.1 CLASSIFICATION ALGORITHMS

#### 4.2.1.1 Logistic regression

Logistic regression is a statistical algorithm that uses a logistic function to model a binary dependent variable (Figure 15). A binary logistic algorithm has a dependent variable with two possible values, which is represented by an indicator variable, where the two values are labeled "0" and "1".



$$t = \beta_0 + \beta_1\omega_1 + \dots + \beta_n\omega_n$$

$$P = \frac{1}{1 + e^{-t}}$$

#### Equation 8

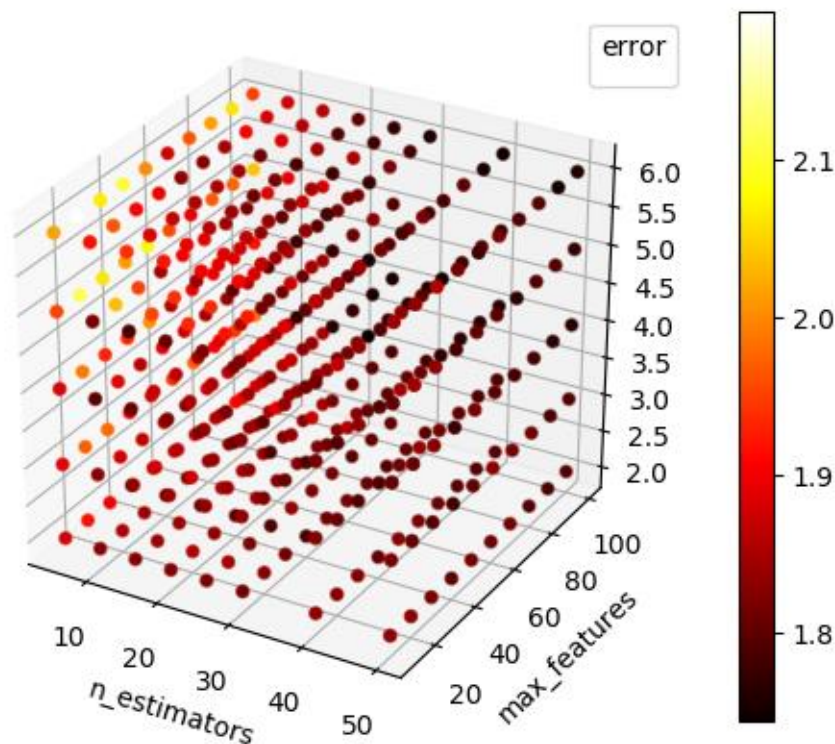
Figure 15: Logistic regression curve showing probability to be infected versus the spectral reflectance

#### 4.2.1.2 Random forest

All combinations of the following values of the three different parameters that define the RF algorithm were examined:

- max\_depth – The values that tested were [3, 4, 5, 6, 7].
- max\_features – The values that tested were [10, 20, 30, 40, 50, 60, 70, 80, 90, 100].
- n\_estimators – The values that tested were [5, 10, 15, 20, 25, 30, 40, 50].

Result reveal that the best results (Figure 16) derived according to accuracy and AUC were achieved for the following values: max\_depth=6, max\_features=50 and n\_estimators=40.



**Figure 16: The accuracy of the RF classifier for the different parameters' combinations**

#### 4.2.2 SENSITIVITY ANALYSIS

##### 4.2.2.1 Probability classification threshold

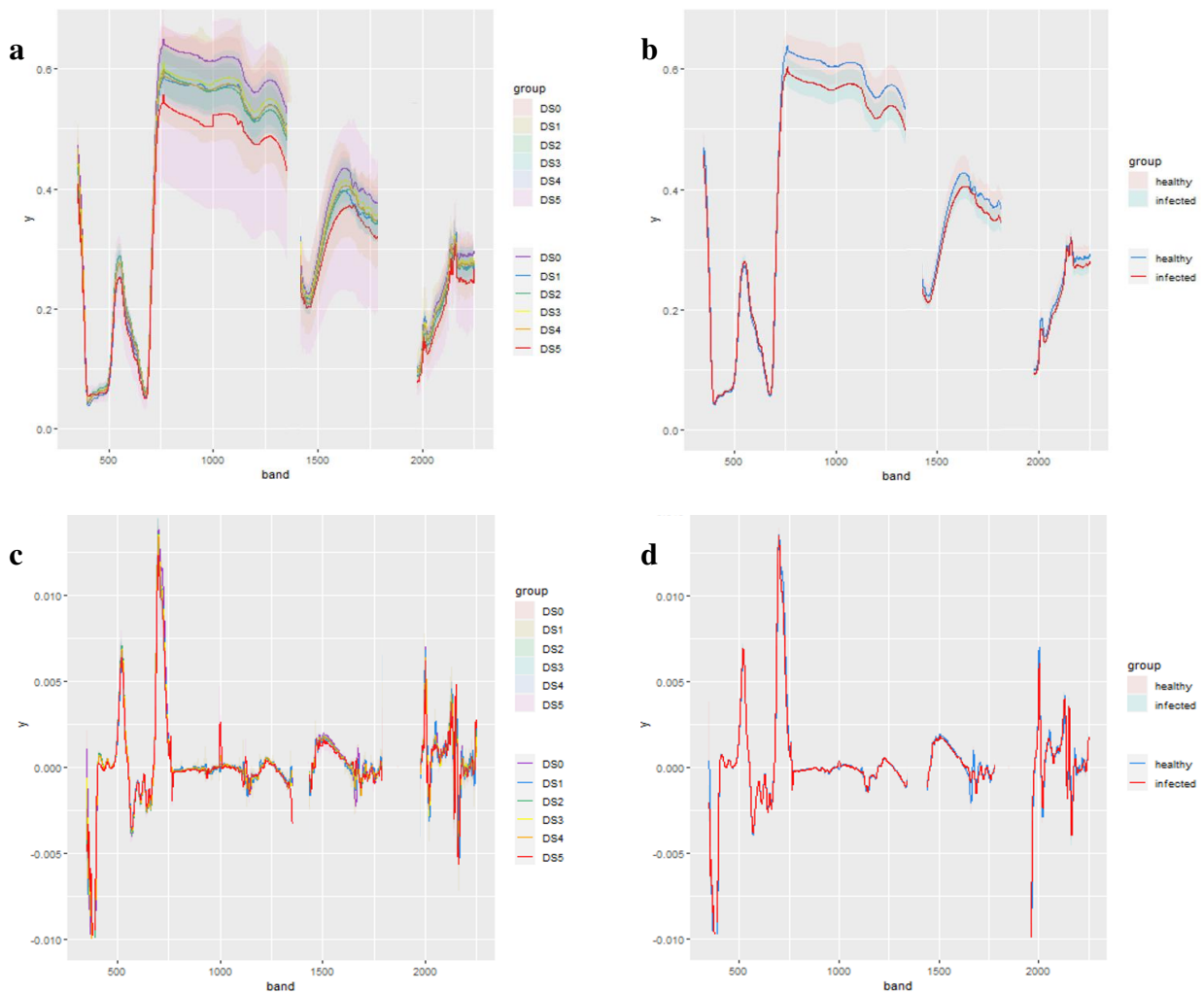
The classifiers return the probabilities of belonging to class 1, denoted as 'infected plant'. The 'healthy' plants were derived as those with a predicted probability below 0.5; the 'infected' plants were derived as those with a predicted probability over 0.5. Sensitivity analysis was conducted to examine the exact threshold that minimizes the accuracy of each classification algorithm.

#### 4.2.2.2 Number of wavelengths

Sensitivity analyzes were performed for classifiers that used a different number of the significant wavelengths. Between a classifier that uses one wavelength and a classifier that uses 12 wavelengths, and a different number of wavelengths between them.

### 4.3 RESULTS

Figure 17 revealed the 95% confidence interval of each leaves' spectral reflectance of the plants\ of the first derivative of leaf spectral reflectance of each disease level. There is no significant different using the reflectance or using the first derivative of one wavelength between the different disease levels for both 2 and 6 classes.



**Figure 17: Mean and the 95% confidence interval of leaves' spectral reflectance of the plants (a) according to 6 different disease levels (b) according to 2 different disease levels Mean and the 95% confidence interval of first derivative of leaf spectral reflectance of the plants (c) according to 6 different disease levels (d) according to 2 different disease levels**

#### 4.3.1 WAVELENGTH SELECTION

##### 4.3.1.1 High resolution spectral data in the VIS/NIR range

A classifier that used the wavelengths in the range of the hyperspectral camera [400- 1000nm] was examined with wavelengths at 10 nm intervals. Results reveal that some wavelengths are significant in the different tests, with T-test (2 classes) and Anova (6 classes) showing similar results (Table 5). Details are presented in Appendix 4.

The significant wavelengths derived from the SR data analysis are:

- 400nm as determined in the T-test, Anova and Random forest.
- 780nm as determined in the T-test method and Anova.

The significant wavelengths derived from the R1D data analysis are:

- 410nm as determined in the T-test method and PLS.
- 950nm as determined in the T-test method, Anova, Random forest and PLS.

	Significant wavelengths [nm]	
Statistic test	SR data analysis	R1D data analysis
T-test	940, 820, 900, 850, 780, 880, 400	950, 750, 410, 810, 720, 900, 780
Anova	940, 820, 850, 880, 780, 750, 400	950, 750, 410, 720, 810, 440, 680
Random forest	700, 600, 520, 460, 760, 640, 400	740, 950, 860, 430, 400, 550, 670
PLS	400, 700, 470, 500, 440, 630, 600	770, 950, 740, 410, 460, 550, 680

**Table 5: Significant wavelengths in the VIS/NIR range according to the various methods for disease detection**

Ranked results from the combination of the wavelengths from the different methods (Table 6) reveal that 400nm was found as the most significant wavelength for disease detection using the SR data, (it was found significant according to three methods: T-test method, Anova and Random forest). This wavelength is mainly influenced by leaf pigment content. 770nm was found as the most significant wavelength for disease detection using the R1D data (it was found significant according to PLS). This wavelength depends on the leaf biology structure.

SR data:

Wavelength [nm]	Importance
400	6.021
700	3.038
750	2.979
440	2.978
900	2.968
820	2.862
630	2.726
500	2.714
460	2.650
940	2.466
730	2.368
760	2.28z
640	2.107

R1D data:

Wavelength [nm]	Importance
770	6.014
410	5.536
460	4.041
860	3.683
700	3.497
720	3.393
480	3.198
900	3.024
550	2.903
450	2.288
680	1.904
780	1.887
430	1.655

**Table 6: The aggregated significant wavelengths for disease detection in the VIS/NIR range**

#### 4.3.1.2 High resolution spectral data in the VIS/NIR/SWIR range

A classifier that used the wavelengths in the range of the hyperspectral camera [350- 2500] was examined with wavelengths at 10 nm intervals. Results reveal that some wavelengths are significant in the different tests, with T-test and Anova showing similar results (Table 7). The significant wavelength derived from the SR data analysis was 400nm as determined in the random forest and PLS algorithm. Details are presented in Appendix 5.

The significant wavelengths derived from the R1D data analysis are:

- 750nm as determined in the T-test, Anova and PLS.
- 1000nm as determined in the T-test method, Anova and PLS.
- 1150nm as determined in the T-test Anova and Random forest.

	Significant wavelengths [nm]	
Statistic test	SR data analysis	R1D data analysis
T-test	940, 380, 820, 1120, 900, 850, 780	1670, 1280, 1150, 750, 1000, 2050, 950
Anova	940, 820, 850, 880, 780, 380, 750	1670, 1000, 950, 2050, 750, 1150, 410
Random forest	360, 2050, 1500, 550, 400, 2000, 460	1150, 950, 740, 400, 2050, 1670, 430
PLS	400, 1960, 1420, 670, 700, 470, 1670	430, 750, 1420, 1150, 1000, 950, 1670

**Table 7: Significant wavelengths according to the various methods for disease detection in the VIS/NIR/SWIR range**

Ranked results from the combination of the wavelengths from the different methods (Table 8) reveal that 750nm was found as the most significant wavelength for disease detection using the SR data, which was found significant only according to Anova. But it has a high correlation to 740nm which was found as significant in Step regression. 740nm was found as the most significant wavelength for disease detection using the R1D data, which was found significant according to three methods (T-test, Anova and PLS. These wavelengths are mainly influenced by leaf pigment content.

SR data:

Wavelength [nm]	Importance
750	6.746
740	6.582
400	5.647
870	4.360
850	4.195
550	3.487
1120	3.287
2050	3.222
1960	2.893
420	2.586
1670	2.386
940	2.220
1420	2.124

R1D data:

Wavelength [nm]	Importance
740	6.772
710	5.372
1000	4.221
1670	4.164
1260	3.756
410	2.938
840	1.832
950	1.633
670	1.571
430	1.406
1150	-
2050	-
750	-

**Table 8: Aggregation of the significant wavelengths in the VIS/NIR/SWIR range for disease detection**

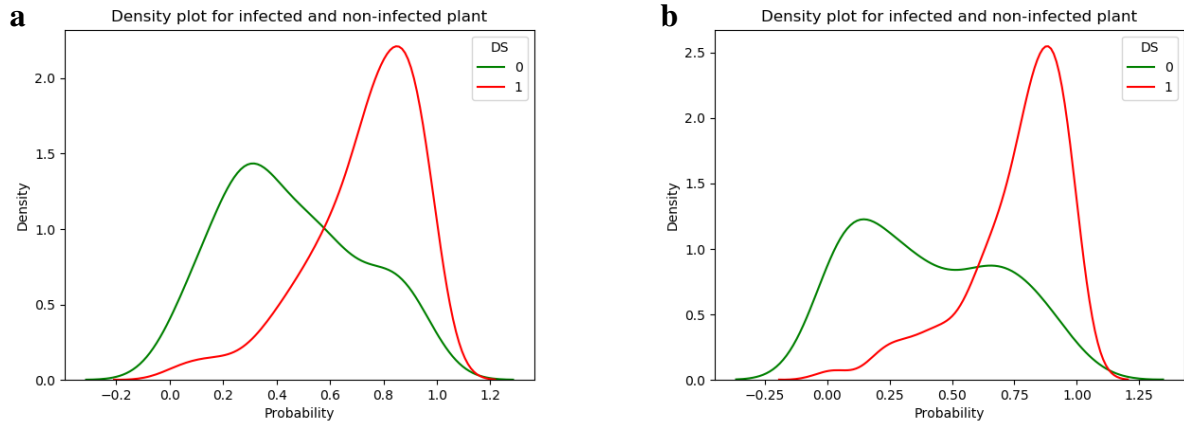
## 4.3.2 CLASSIFICATION ALGORITHMS

### 4.3.2.1 High resolution spectral data in the VIS/NIR range

A classifier that used 7 wavelengths was compared to a classifier that used the entire range of wavelengths.

#### 4.3.2.1.1 Using 7 wavelengths

The SR data and the first derivative of all individual spectral were used to build a logistic regression classifier that discriminates between infected and healthy plants using the 7 most significant wavelengths.



**Figure 18: Density plot distributions of the predicted values of the infected and the healthy plants of logistic regression classifier using 7 wavelengths in the VIS/NIR range (a) using SR data (b) using the first derivative of all individual spectra**

There is difference between the density of the predicted probability of the infected and the density of the healthy plants as expected (Figure 18). The density of the predicted probability of the 'infected' centered around predicted probability 1, while the 'healthy' center lower.

	SR data		R1D data	
	Healthy	Infected	Healthy	Infected
Predicted as 'healthy'	<b>TP</b> – 57	<b>FP</b> – 21	<b>TP</b> – 60	<b>FP</b> – 21
Predicted as 'infected'	<b>FN</b> – 40	<b>TN</b> – 141	<b>FN</b> – 37	<b>TN</b> – 141
	<i>sensiyivity</i> $= \frac{TP}{TP + FN}$ $= \frac{57}{97} = 0.587$	<i>specificity</i> $= \frac{TN}{TN + FP}$ $= \frac{141}{162} = 0.87$	<i>sensiyivity</i> $= \frac{TP}{TP + FN}$ $= \frac{60}{97} = 0.618$	<i>specificity</i> $= \frac{TN}{TN + FP}$ $= \frac{141}{162} = 0.87$

**Table 9: Confusion matrix of logistic regression classifier using 7 wavelengths in the VIS/NIR range**

The classification of the infected plants is consistently better than healthy plant classification (Table 9) in both data sets, SR and R1D. This is probably due to the imbalanced data set (there are 162 infected plants, compared to 97 infected plants).

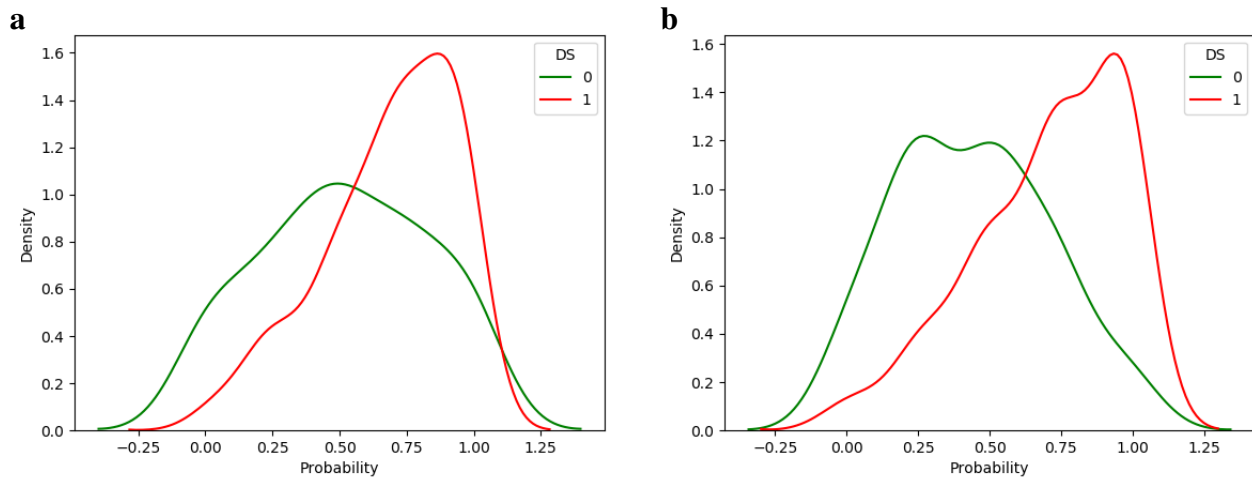
Performance measure	SR data		R1D data	
<b>AUC</b>	0.729		0.744	
<b>Accuracy</b>	$\frac{198}{259} = 0.764$		$\frac{201}{259} = 0.776$	
	<b>95% confidence interval</b>		<b>95% confidence interval</b>	
	0.713	0.816	0.725	0.826

**Table 10: Logistic regression classifier using 7 wavelengths results in the VIS/NIR range**

The classifier that used the R1D data is better than the classifier that used the SR data by 101.5% in accuracy. But, there is no significant difference (Table 10).

#### 4.3.2.1.2 Model that uses all wavelengths

The SR data and the first derivative of all individual spectral (61 features total) were used to build a random forest classifier.



**Figure 19: Density plot distributions of the predicted values of the 'healthy' and the 'infected' plants of random forest classifier using all wavelengths in the VIS/NIR range (a) using SR data (b) using the first derivative of all individual spectra**

There is difference between the density of the predicted probability of the infected and the density of the healthy plant as expected (Figure 19). The density of the predicted probability of the 'infected' centered around predicted probability 1, while the healthy center lower. In addition, the prediction of the infected plants is consistently better than healthy plant prediction. The bias probably exists due to the imbalanced data set (there are 162 infected plants, compared to 97 healthy plants).

Visual analysis reveals (Figure 19) that if the threshold would be set to be higher than 0.5 (the cut point between the graphs of the infected and healthy plants), the accuracy could be better. Sensitivity analyses for different thresholds are detailed in section 4.4.1, revealing that a threshold of 0.543 in SR data improves the results by 1.7% in accuracy, and a threshold of 0.57 in R1D data improves the results by 1.1% in accuracy.

Performance measure	SR data		R1D data	
<b>AUC</b>	0.65		0.757	
<b>Accuracy</b>	$172/259 = 0.664$		$181/259 = 0.698$	
	<b>95% confidence interval</b>		<b>95% confidence interval</b>	
	0.606	0.721	0.643	0.754

**Table 11: Random forest classification classifier results using all wavelengths in the VIS/NIR range**



The classifier that using the R1D data is better by accuracy than classifier that using the SR data by 105.23%. But, there is no significant difference (Table 11).

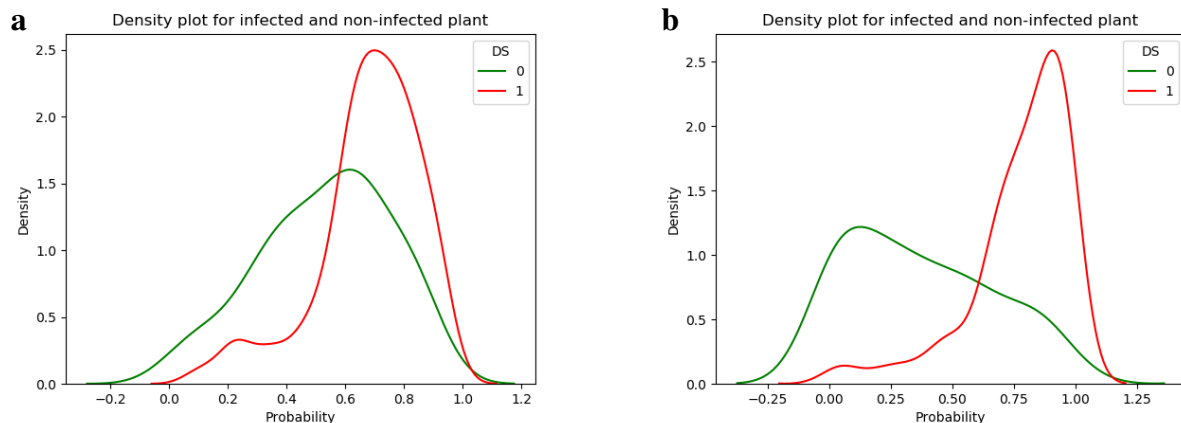
In addition, results reveal that there is no advantage to using the wavelengths in the entire range compared to using only 7 wavelengths (Table 10 \_+ Table 11). This may be due to the low number of samples which leads to overfitting when using a large number of features. For the SR data, the classifier that used 7 wavelengths is better than the classifier that used all wavelengths by 15.1% by accuracy. But, there is no significant difference. For the R1D data, the classifier that used the 7 wavelengths is better than the classifier that used all wavelengths by 111.05% as define by accuracy. But, there is no significant difference.

#### 4.3.2.2 High resolution spectral data in the VIS/NIR/SWIR range

Classifier that used 7 wavelengths was compared to a classifier that used the entire range of wavelengths.

##### 4.3.2.2.1 Using 7 wavelengths

The SR data and the first derivative of all individual spectral were used to build a logistic regression classifier that using the 7 most significant wavelengths: 400, 550nm that are mainly influenced by leaf pigment content, 740, 750, 850, 870nm that depend on the leaf biology structure and 1120nm that is influenced by the composition of leaf chemicals and water.



**Figure 20: Density plot distributions of the predicted values of the infected and the healthy plants of logistic regression classifier using 7 wavelengths in the VIS/NIR/SWIR range (a) using SR data (b) using the first derivative of all individual spectra**

Here also, we received the same trend as in the VIS/NIR range. There is difference between the density of the predicted probability of the infected and the density of the healthy plant as expected (Figure 20). The density of the predicted probability of the infected centered around predicted probability 1, while the healthy center lower. In addition, the prediction of the infected plants is consistently better than healthy plant prediction. The bias may be existing because the data set is not balanced (there are 162 infected plants, compared to 97 healthy plants).

Visual analysis reveals (Figure 20) that if the threshold would be set to be higher than 0.5 (the cut point between the graphs of the infected and healthy plants), the accuracy could be better. Sensitivity analyses for different thresholds are detailed in section 4.4.1, revealing that a threshold of 0.635 in R1D data improves the results by 3.8% in accuracy.

	SR data		R1D data	
	Healthy	Infected	Healthy	Infected
Predicted as 'healthy'	<b>TP</b> – 43	<b>FP</b> – 21	<b>TP</b> – 63	<b>FP</b> – 18
Predicted as 'infected'	<b>FN</b> – 54	<b>TN</b> – 141	<b>FN</b> – 34	<b>TN</b> – 144
	<i>sensitivity</i> $= \frac{TP}{TP + FN}$ $= \frac{43}{97} = 0.443$	<i>specificity</i> $= \frac{TN}{TN + FP}$ $= \frac{141}{162} = 0.87$	<i>sensitivity</i> $= \frac{TP}{TP + FN}$ $= \frac{63}{97} = 0.649$	<i>specificity</i> $= \frac{TN}{TN + FP}$ $= \frac{144}{162} = 0.89$

**Table 12: Confusion matrix of logistic regression classifier using 7 wavelengths in the VIS/NIR/SWIR range**

The classification of the infected plants is consistently better than healthy plant classification (Table 9) in both data sets, SR and R1D. This is probably due to the imbalanced data set (there are 162 infected plants, compared to 97 healthy plants).

Performance measure	SR data		R1D data	
<b>AUC</b>	0.657		0.769	
<b>Accuracy</b>	$184/259 = 0.71$		$207/259 = 0.8$	
	<b>95% confidence interval</b>		<b>95% confidence interval</b>	
	0.655	0.765	0.75	0.848

**Table 13: Logistic regression classifier using 7 wavelengths results in the VIS/NIR/SWIR range**

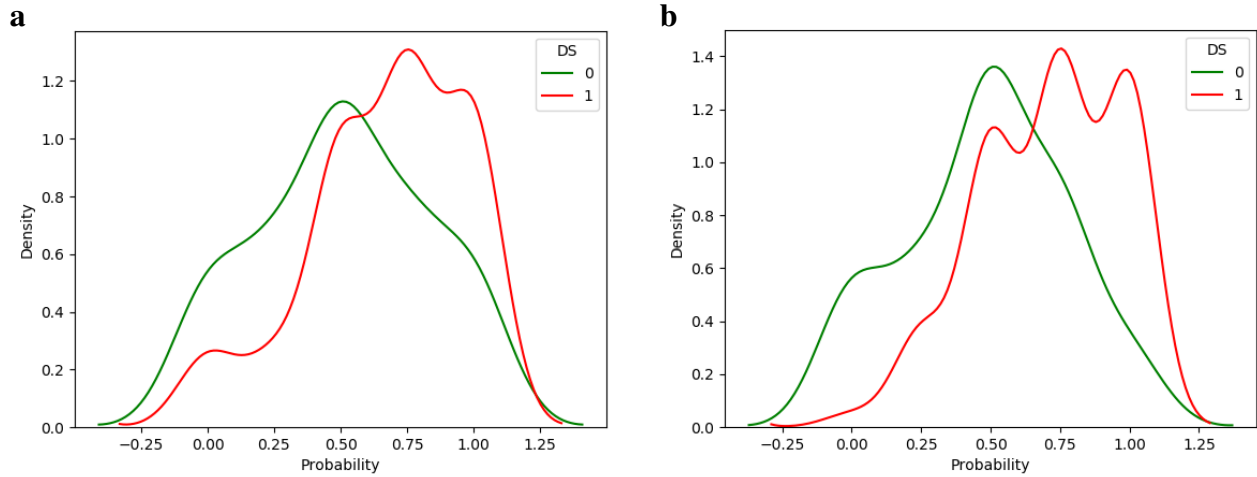
The classifier that used the R1D data is better than the classifier that used the SR data by 112.5% in the accuracy. But, there is no significant difference (Table 10).

In the SR data the classifier that used the 7 wavelengths in the VIS/NIR range is better than the classifier that used the 7 wavelength from the VIS/NIR/SWIR range by 107.6% in the accuracy. But, there is no significant difference.

In the R1D data the classifier that used the 7 wavelength in the VIS/NIR/SWIR range is better than the classifier that used the 7 wavelength in the VIS/NIR range by 102.98% in accuracy. But, there is no significant difference.

#### 4.3.2.2.2 Using all wavelengths

The SR data and the first derivative of all individual spectral (169 features total) were used to build a random forest classifier that discriminates between infected and healthy plants.



**Figure 21: Density plot distributions of the predicted values of the infected and the healthy plants of random forest classifier using all wavelengths in the VIS/NIR/SWIR range (a) using SR data (b) using the first derivative of all individual spectra**

There is difference between the density of the predicted probability of the infected (centered around 1) and healthy plants centered lower as expected (Figure 21). In addition, classification of the infected plants is consistently better than healthy plant prediction. The bias may be due to the unbalanced data set (there are 162 infected plants, compared to 97 healthy plants).

Performance measure	SR data		R1D data	
<b>AUC</b>	0.647		0.656	
<b>Accuracy</b>	$161/259 = 0.62$		$170/259 = 0.656$	
	<b>95% confidence interval</b>		<b>95% confidence interval</b>	
	0.56	0.68	0.59	0.71

**Table 14: Random forest classifier results using all wavelengths in the VIS/NIR/SWIR range**

The classifier that used the R1D data is better by accuracy than classifier that using the SR data by  $\frac{170}{161} * 100 = 105.6\%$ . But, there is no significant difference (Table 14).

Results reveal that there is no advantage to using the wavelengths in the entire range compared to using only 7 wavelengths (Table 13 + Table 14). This may be due to the low number of samples causing

overfitting when using a large number of features. In the SR data the classifier that used the 7 wavelength is better by accuracy than classifier that using the entire range by  $\frac{184}{161} * 100 = 114.28\%$ . But, there is no significant difference. In the R1D data the classifier that using the 7 wavelength is better by accuracy than classifier that using the entire range by  $\frac{207}{170} * 100 = 121.76\%$ . But, there is no significant difference.

In addition, it can be seen that there is no significant difference between the classifier that use the all wavelengths in the High resolution spectral data in the VIS/NIR/SWIR range to the classifier that use the High resolution spectral data in the VIS/NIR range (Table 11 + Table 14).

In the SR data the classifier that using the entire range from the VIS/NIR range is better by accuracy than classifier that using the entire range from the VIS/NIR/SWIR range  $\frac{172}{161} * 100 = 106.8\%$ . But, there is no significant difference. In the R1D data the classifier that using the entire range from the VIS/NIR range is better by accuracy than classifier that using the entire range from the VIS/NIR/SWIR range  $\frac{181}{170} * 100 = 106.47\%$ . But, there is no significant difference.

**There is not enough data to determine that there is a significant difference between the classifiers that used also the SWIR range compared to the classifiers that used only the VIS and NIR ranges.**

## 4.4 SENSITIVITY ANALYSIS

### 4.4.1 PROBABILITY CLASSIFICATION THRESHOLD

Analysis reveals (Table 17) that for each classifier different threshold should be defined for classification between infected and healthy plants in order to maximize the accuracy. Etc. the random forest using all wavelengths in the VIS/NIR/SWIR range classifier revealing that a threshold of 0.499 in SR data improves the results by 6.8% in accuracy.

### 4.4.2 NUMBER OF WAVELENGTHS

#### 4.4.2.1 High resolution spectral data in the VIS/NIR range

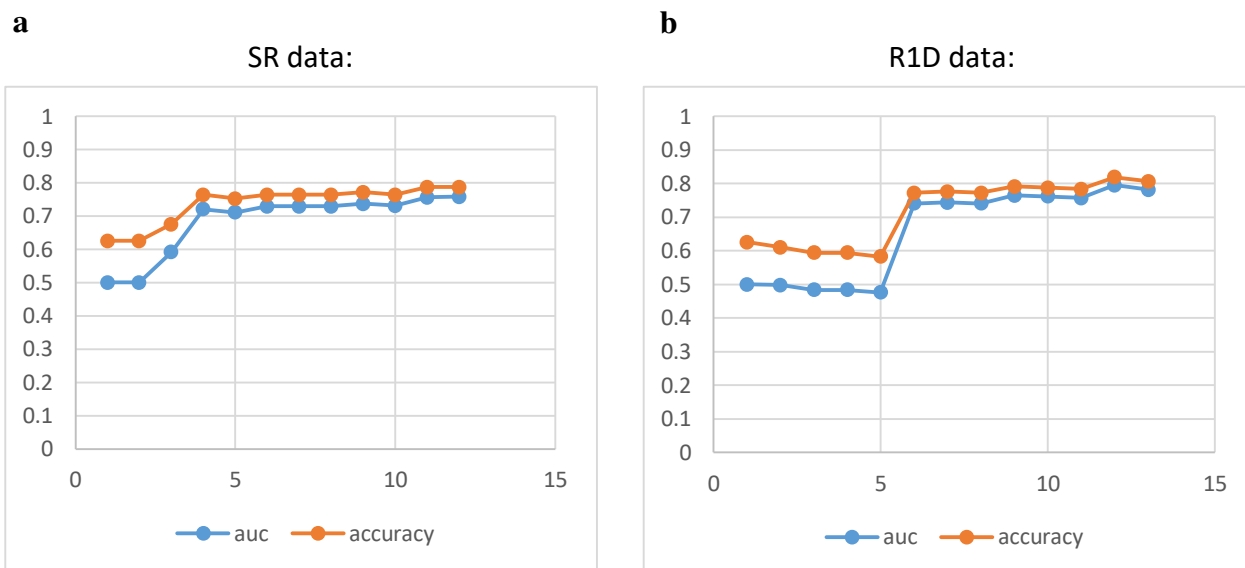
Results reveal same trend for the SR data and the first derivative data. using 4 wavelengths significantly improves the results of classifiers than classifiers that using less than 4 wavelengths (Figure 22).

#### 4.4.2.2 High resolution spectral data in the VIS/NIR/SWIR range

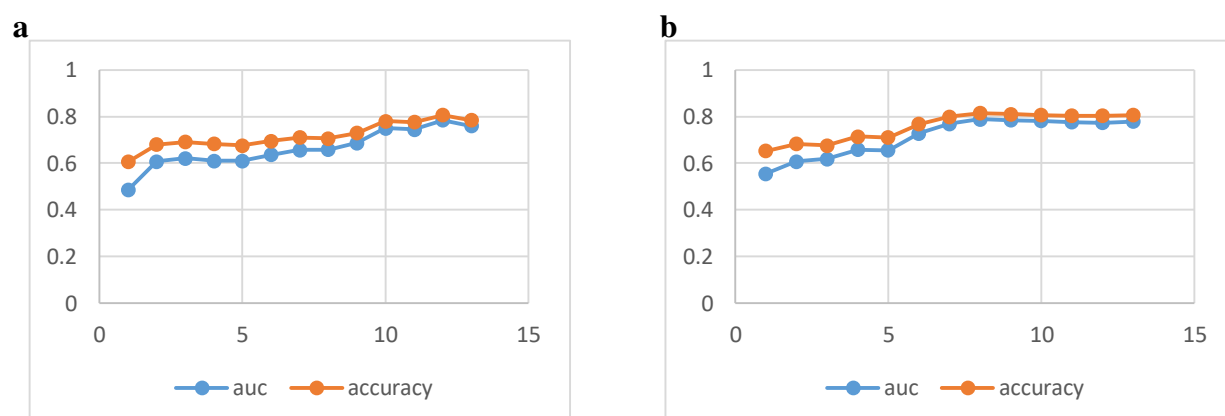
Results reveal same trend for the SR data and the first derivative data. The more wavelengths the model is based on, the better it is classified (in both, accuracy and AUC) (Figure 23).

Algorithm	Data analyses	Accuracy with 0.5 threshold	Classification threshold that maximizes accuracy	Accuracy	Improvement rate
Logistic regression using 7 wavelengths in the VIS/NIR range	SR	$\frac{198}{259} = 0.764$	0.51	$\frac{199}{259} = 0.768$	$\frac{199}{198} * 100 = 100.5\%$
	R1D	$\frac{201}{259} = 0.776$	0.52	$\frac{203}{259} = 0.783$	$\frac{203}{201} * 100 = 100.9\%$
Random forest using all wavelengths in the VIS/NIR range	SR	$\frac{172}{259} = 0.664$	0.543	$\frac{175}{259} = 0.676$	$\frac{175}{172} * 100 = 101.74\%$
	R1D	$\frac{181}{259} = 0.689$	0.57	$\frac{183}{259} = 0.706$	$\frac{183}{181} * 100 = 101.1\%$
Logistic regression using 7 wavelengths in the VIS/NIR/SWIR range	SR	$\frac{184}{259} = 0.71$	0.5	-	-
	R1D	$\frac{207}{259} = 0.8$	0.635	$\frac{215}{259} = 0.83$	$\frac{215}{207} * 100 = 103.86\%$
Random forest using all wavelengths in the VIS/NIR/SWIR range	SR	$\frac{161}{259} = 0.62$	0.499	$\frac{172}{259} = 0.664$	$\frac{172}{161} * 100 = 106.8\%$
	R1D	$\frac{170}{259} = 0.656$	0.499	$\frac{176}{259} = 0.62$	$\frac{176}{170} * 100 = 103.5\%$

**Table 15: Sensitivity analysis by changing the threshold value**



**Figure 22: Sensitivity analysis of classification with increasing number of wavelengths in the in the VIS/NIR range (a) using SR data (b) using the first derivative of all individual spectra**



**Figure 23: Sensitivity analysis of classification classifiers with increasing number of wavelengths in the VIS/NIR/SWIR range (a) using SR data (b) using the first derivative of all individual spectra**

## 5. HYPERSPECTRAL IMAGES ANALYSIS

### 5.1 DATA

Hyperspectral images from the two experiments were used for the classifier development (208 infected and 126 healthy plants). Each image included the spectral reflectance of all wavelengths in the range 400- 1,000 nm with a total of 840 bands. The length of each image was 1600 pixels; the width of the images was changed manually to include the entire plant (different boundaries were set for each image according to the acquired plant image).

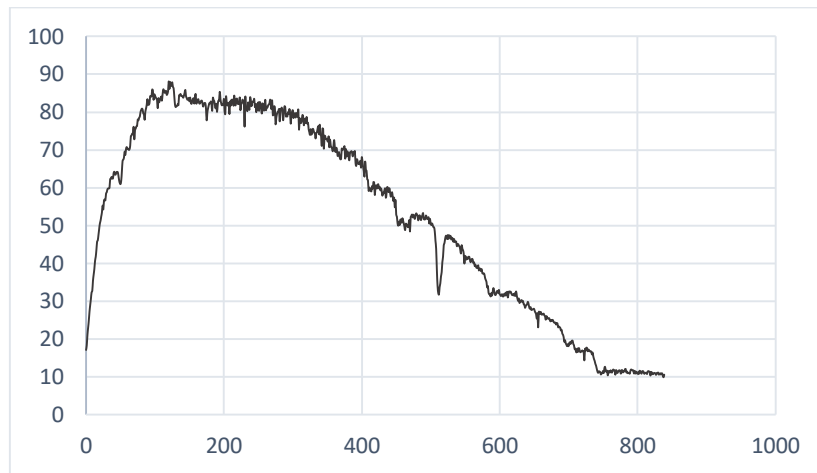
### 5.2 METHODS

#### 5.2.1 NOISE EXTRACTION

The SNR was calculated as follows (Figure 24) for each image for each wavelength based on the mean and standard deviation calculated from  $10 \times 10$  (100) pixels from the white reference in the images. The average SNR was calculated from the aggregation of all images for each band.

```
Input: images with hyper cube with the range of 400 to 1000 nm
Output: images with hyper cube with all wavelengths with SNR > 12
for i from 400 to 1000:
    for image in images:
        WR = 10*10 pixels from the center of the white reference
        mean[i, image] = mean(WR pixels in wavelength i)
        std[i, image] = std(WR pixels in wavelength i)
        SNR[i, image] = mean / std
    SNR[i] = mean(SNR[i, :])
    if (SNR[i]) < 12:
        removed from the spectral data in all images
return images hyper cube with all wavelengths with SNR > 12
```

Figure 24: SNR calculation

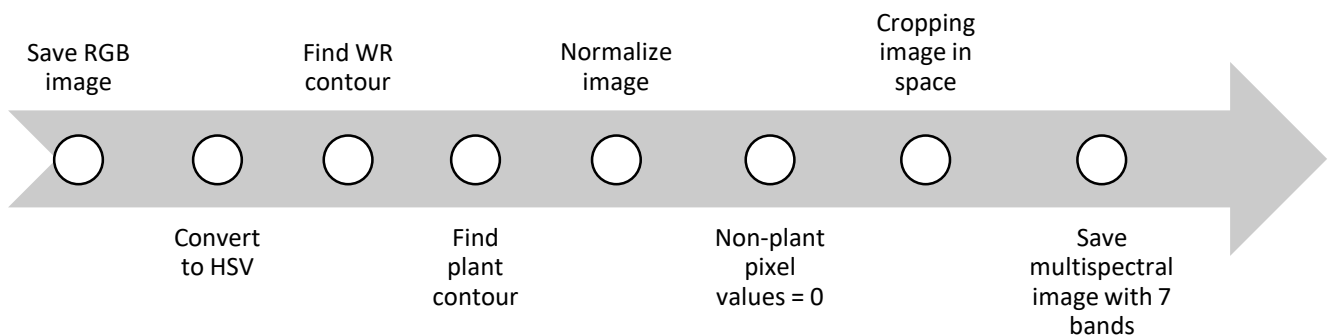


**Figure 25: signal-to-noise ratio for each band**

The threshold was set to  $SNR = 12$ , this value was determined by the noise in a black reference image, where the reflectance should be 0. The bands (wavelengths) with a lower SNR 740- 840 (930- 1,005nm) were removed from the spectral data.

### 5.2.2 PRE- PROCESSING

Pre-processing included several steps for additional noise reduction and reduction of the image size (Figure 26).



**Figure 26: Hyperspectral images pre-processing flowchart**

#### 5.2.2.1 Segmentation

Segmentation was performed using seven steps as detailed below:

- Step 1 - Save the RGB images from the hyperspectral images.
- Step 2 - Converted to HSV color space (hue, saturation, value).
- Step 3 - Multiple masks, a low threshold and a high threshold were applied for hue, saturation and value in order to find the pixels with these values using “inRange” functions from “cv2”



library. Any pixel within these thresholds was set to 1 and the remaining pixels was set to zero.

The HSV threshold of the plants includes hues of green and brown:

- Lower threshold – [10, 50, 15]
- Upper threshold – [70, 255, 255]

The HSV threshold of the white reference board includes:

- Lower threshold – [0, 0, 100]
- Upper threshold – [140, 255, 255]

Step 4 - From the mask the contours of all the areas in the image that are within the defined threshold were found with “findContours” functions from “cv2” library.

Step 5 - Each contour has the size of the area it delimits. In each image the bigger contour was the plant/ white reference board.

Step 6 - All the pixels outside the counter were converted to 0 in all bands.

#### 5.2.2.2 Normalize image

The hyperspectral images not only acquire information about the plants but also reflects the influence of nuisance signals coming from illumination effects, the detector sensitivity and the transmission properties of the optics (Geladi, Burger, & Lesstanderas, 2004). Spectral calibration was conducted to compensate for these effects by using black and white references (Mishra et al., 2017).

The black image was acquired when the camera shutter is completely turned off. The white reference was obtained on each image by a white surface board, a 100 mm diameter spectralon plate, which has a uniform, stable and high reflectance surface. These two references were used to correct the raw images by using the following equation:  $I_R = \frac{I_{raw} - I_{dark}}{I_{white} - I_{dark}}$ .

where,  $I_R$  is the calibrated reflectance image,  $I_{raw}$  is the raw intensity image measured,  $I_{dark}$  is the average intensity of 800 pixels for each column and band of the black reference, and  $I_{white}$  is the most common value of the white reference intensity for each band.

Input: hyperspectral images + dark image

Output: normalized hyperspectral images

for i from 400 to 1000:

for w from 1 to 840 (images width):

$$dark_{i,w} = \text{mean}(\text{dark image}_w)$$

for image in images:

WR = white reference pixels

for i from 400 to 1000:

$$white_i = \text{mode}(WR_i)$$

for w from 1 to 840 (images width):

for h from 1 to 1400 (image length):

$$normalized\ image_{w,h} = \frac{image_{i,w,h} - dark_{i,w}}{white_i - dark_{i,w}}$$

return normalized images

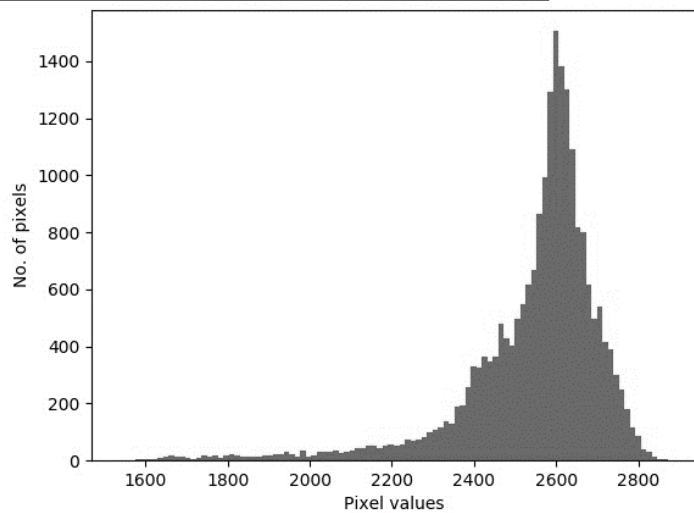

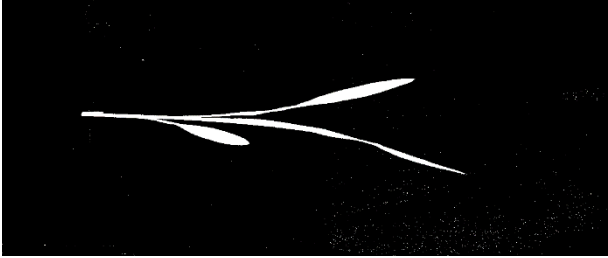




Figure 27: Diameter spectralon plate

Figure 28: Histogram of the white reference pixels, band 200

Step 7 - Each pixel in the image normalized with this equation. The hyperspectral images were cropped to the minimum rectangle containing the contour (that contain all plant pixels) and the non-plant pixels were set to zero.

	
<p>Step 1+2- Save RGB image from the hyperspectral cube and convert the image to HSV color space</p>	<p>Step 3- Applied multiple masks</p>
	
<p>Step 4+5- Find the contour of the plant in image</p>	<p>Step 6+7- Normalize image and remove background</p>

**Figure 29: Hyperspectral images pre-processing steps**

### 5.2.3 DATA ANALYSIS

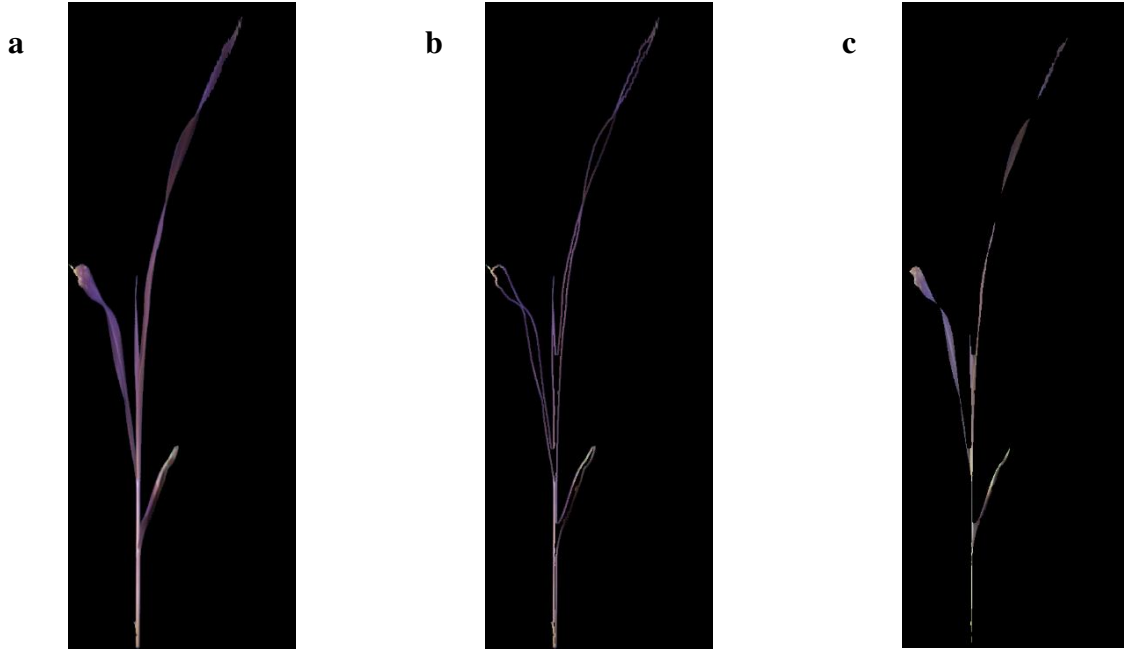
For each hyper-spectral image of a plant, the different pixels of the plant in the image were examined. Visual analysis was performed in order to examine whether the data from the ASD is similar to the data from the hyper-spectral camera. The average spectral signature of the plant was examined to examine if it is similar to the distribution obtained from a single point in the plant from ASD. For each wavelength, the average value of all the pixels of the plant in the image was calculated and compared to the ASD value.

The smooth first derivative of the average spectral signature also was calculated from the average spectral signature. The first derivative of the spectral reflectance was calculated, and simultaneous smoothing was applied with a filter order of 2 and a filter length of 11 wavelengths (using the “savgol\_filter” function from “scipy” library).

### 5.2.4 FEATURES

Features 1- 5 were calculated for each channel of the multispectral image. Features 1-3 were examined based on common features in the literature. The first feature does not use the spatial information contained in the multispectral image. Features 2- 7 used the spatial information

contained in the multispectral image. Features 4-5 are based on the Fusarium infection process. The infection process starts from the roots and continues to the edges of the leaves (Figure 30). Features 6- 7 are related to the shape of the plant (equal in all channels), therefore were calculated once for each plant.



**Figure 30: (a) All plant pixels (b) pixels at the edges of the plant (c) pixels not in edges**

1.  $Average(\lambda)$  – average reflectance of all plant pixels in the image at wavelength  $\lambda$ .
2.  $SD(\lambda)$  – standard deviation of all plant pixels in the image at wavelength  $\lambda$ .
3.  $Coefficient\ of\ Variation(\lambda)$  ( $CV$ ) –  $\frac{\sigma}{\mu}$  the average reflectance of all plant pixels in the image at wavelength  $\lambda$  divided by the standard deviation of all plant pixels in the image at wavelength  $\lambda$ .
4.  $Discriminability(\lambda)$  ( $Disc$ ) –  $\frac{\mu_{edges} - \mu_{center}}{\mu_{all}}$  the average reflectance of pixels in the edge of the plant ( $\mu_{edges}$ ) subtraction of the pixels in the center of the plan ( $\mu_{center}$ ) divided by the average reflectance of all plant pixels ( $\mu_{all}$ ) in the image at wavelength  $\lambda$ .
5.  $ER(\lambda)$  –  $\frac{\mu_{edges}}{\mu_{center}}$  the average reflectance of pixels in the edge of the plant ( $\mu_{edges}$ ) divided by the pixels in the center of the plan ( $\mu_{center}$ ) in the image at wavelength  $\lambda$ .
6. LAI – the number of the plant's pixels (B. Zhang, et al. 2012)
7. Height/Width Ratio - the ration between the height to the width of the plant

Features 6- 7 are related to the shape of the plant and therefore were calculated once for each plant (not calculated for each channel).

8. LAI – defined as the area of single sided leaves per area of soil (B. Zhang, et al. 2012).
9. Height/Width Ratio – the ration between the height to the width of the plant

Total 37 features ( $5_{indices} \times 7_{bands} + 2 = 37$ ) for each multispectral image.

### 5.2.5 CLASSIFICATION ALGORITHMS

The various features from the multispectral image were introduced into a RF algorithm to classify between infected and health plants. Combinations of three different parameters that define the RF algorithm were examined:

- max\_depth – The values that tested were [2, 3, 4, 5, 6, 7].
- max\_features – The values that tested were [5, 10, 15, 20, 25, 30, 35].
- n\_estimators – The values that tested were [5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 65, 70].

### 5.2.6 SENSITIVITY ANALYSIS

The classifier was evaluated using data with different bandwidths, which simulate a lower spectral resolution than that of the camera. Images with bandwidths of 10, 30, 50, and 150 nm were examined. The narrower the bandwidth, the higher the cost of the sensors. If the classifier is found to be insensitive to bandwidth, we can lower the cost of sensors significantly.

For each bandwidth, each wavelength, from the 7 wavelengths, Gaussian image calculated using the images from all wavelengths within the bandwidth. When the bandwidth included wavelengths that are not within the camera range [400- 1,000nm], the image was calculated with a cut Gaussian and divided by the weight of the images.

Bandwidth	Number of images
10	15
30	41
50	67
150	423

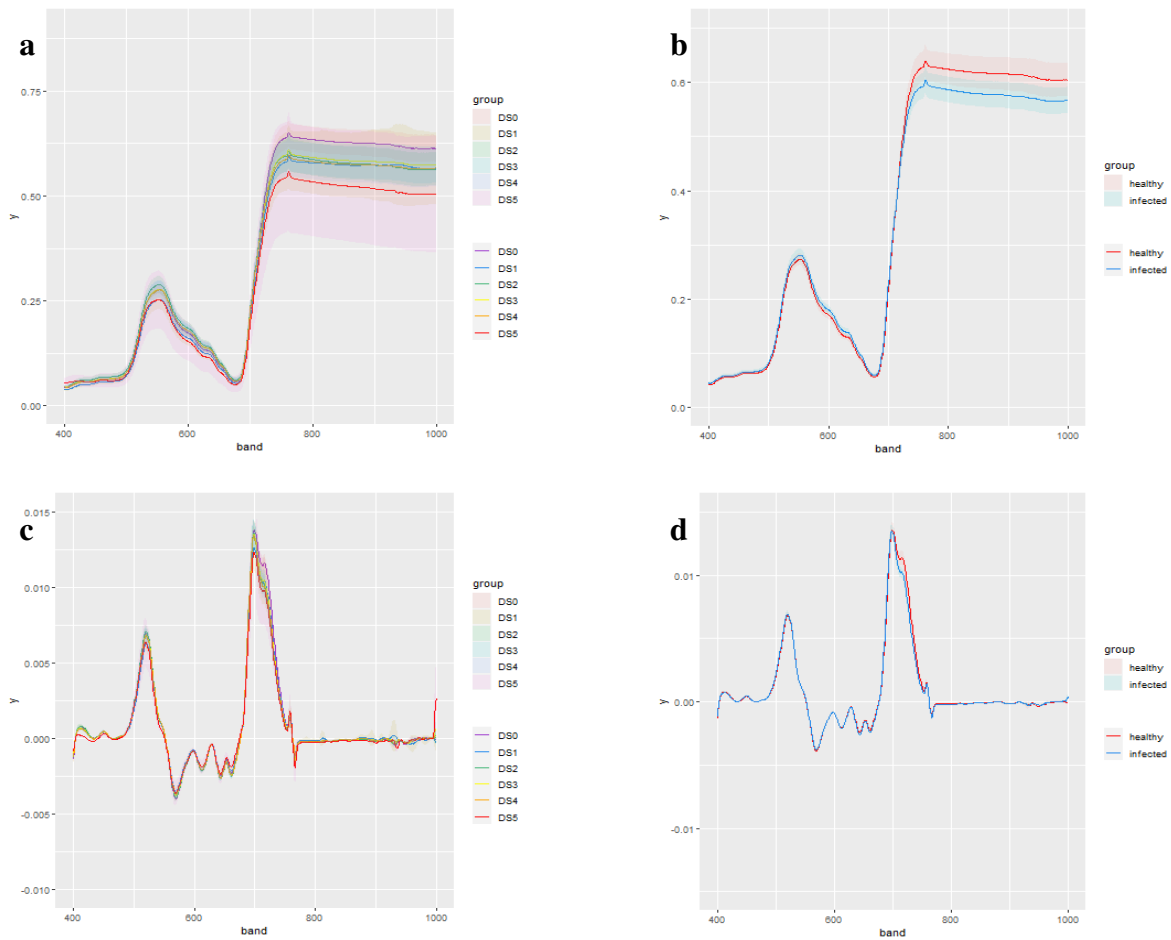
Table 16: The number of images from which the bandwidth image was calculated

## 5.3 RESULTS

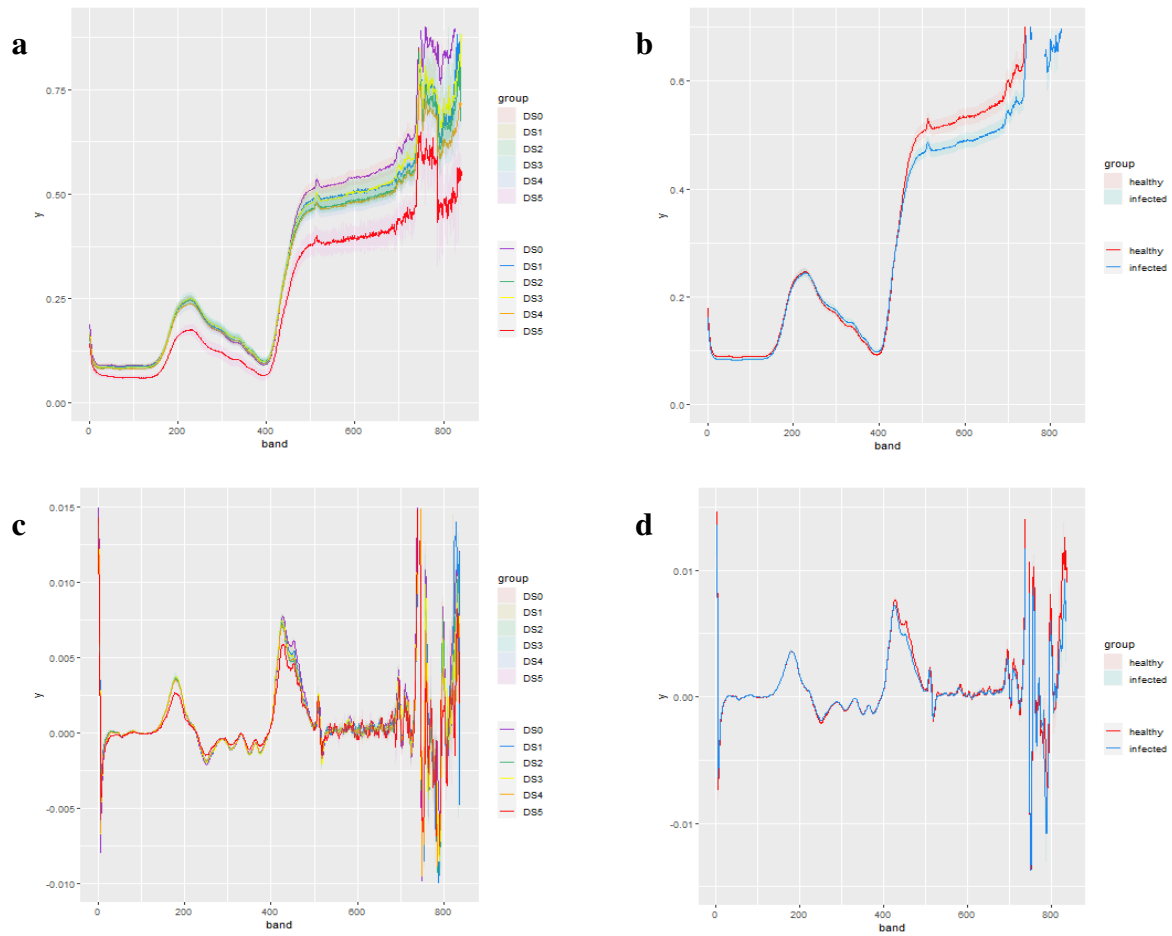
### 5.3.1 DATA ANALYSES

Visual analysis reveals there is a correlation between the distribution obtained from a single point in plants from the ASD, and the average distribution of the pixels in the plant obtained from the hyperspectral images (Figure 31+ Figure 32).

The significant wavelengths for disease detection from the ASD analysis available in the hyperspectral camera [VIS/NIR range], were used to build algorithm to classify between infected and healthy.



**Figure 31: Mean and the 95% confidence interval of leaves' spectral reflectance of the plants from ASD (a) according to 6 different disease levels (b) according to 2 different disease levels**  
**Mean and the 95% confidence interval of first derivative of leaf spectral reflectance of the plants from ASD (c) according to 6 different disease levels (d) according to 2 different disease level**



**Figure 32: Mean and the 95% confidence interval of leaves' spectral reflectance of the plants from hyperspectral images**

(a) according to 6 different disease levels (b) according to 2 different disease levels  
Mean and the 95% confidence interval of first derivative of leaf spectral reflectance of the plants from hyperspectral images  
(c) according to 6 different disease levels (d) according to 2 different disease level

Input: hyperspectral images

Output: average spectral signature of the images

for image in images:

plant = all plant pixels in the image

for i from 400 to 1000:

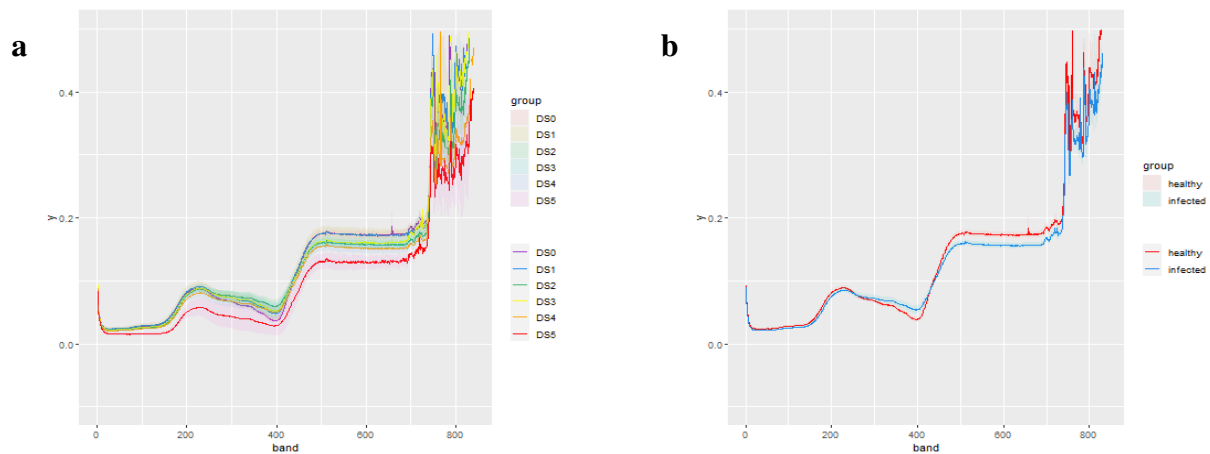
$$x_i = \text{mean}(\text{plant})$$

plant average spectral signature =  $x$

average spectral signature of the images = plant average spectral signature

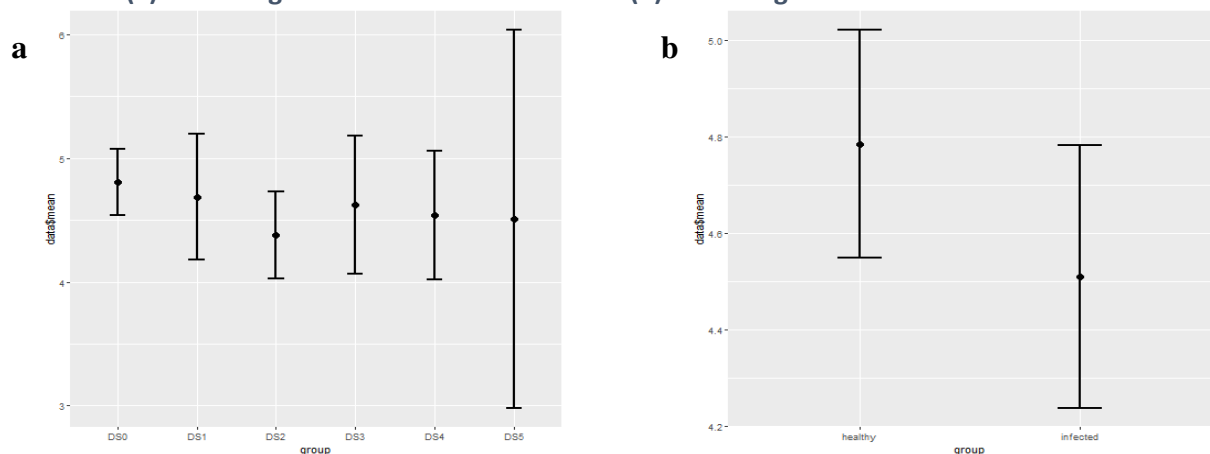
The average standard deviation and the geometric feature height/width ratio of the plants in the images were examined using Anova analysis (Figure 33).

There is an adjustment, the higher the reflectance the higher standard deviation is (Figure 34). In addition, according to Anova method there is no significant different between the height/ width ratio of the different disease severity levels.



**Figure 33: Mean and the 95% confidence interval of the standard deviation of the leaves' spectral reflectance of the plants**

(a) according to 6 different disease levels (b) according to 2 different disease levels



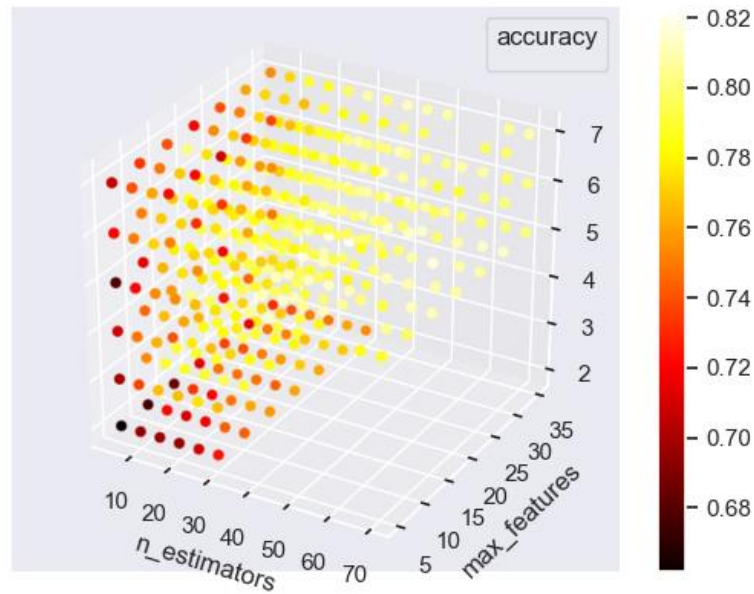
**Figure 34: Mean and the 95% confidence interval of the height/width ratio of the plants**  
(a) according to 6 different disease levels (b) according to 2 different disease levels

## 5.3.2 CLASSIFICATION ALGORITHMS

### 5.3.2.1 Tuning parameters

Results (Appendix 6) reveal that best accuracy and AUC results are obtained for using the combination of 6 for the following combination of parameters: maximum depth of the trees, 15 features for the best split and 50 different trees by both (Figure 35).

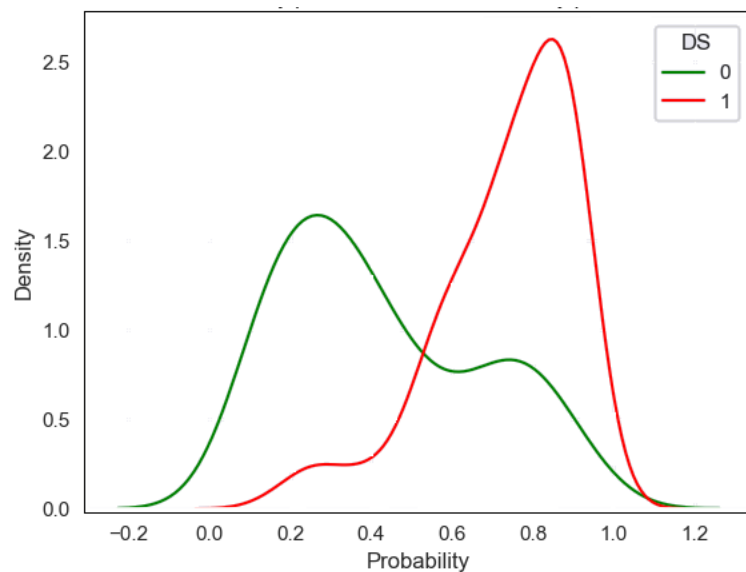




**Figure 35: The accuracy of the RF algorithm for each parameters combination**

#### 5.3.2.2 Results

The features were used to build a random forest classifier between infected and healthy plants using the 7 most significant wavelengths using the above parameters.



**Figure 36: Density plot distributions of the predicted values of the infected and the healthy infected plants**

There is difference between the density of the predicted probability of the infected and the density of the healthy plant as expected (Figure 36). The density of the predicted probability of the infected centered around predicted probability 1, while the healthy center lower.

	Healthy	Infected
Predicted as 'healthy'	<b>TP</b> – 85	<b>FP</b> – 18
Predicted as 'infected'	<b>FN</b> – 41	<b>TN</b> – 190
	$Sensitivity = \frac{TP}{TP + FN} = \frac{85}{126} = 0.675$	$specificity = \frac{TN}{TN + FP} = \frac{190}{208} = 0.913$

**Table 17: Confusion matrix of RF classifier using the features**

The classification of the infected plants is consistently better for healthy plant classification (Table 17). This is probably due to the imbalanced data set (there are 208 infected plants, compared to 126 healthy plants).

Algorithm	Data analyses	Accuracy	AUC
Logistic regression using 7 wavelengths in the VIS/NIR range	ASD - SR	$\frac{198}{259} = 0.764$	0.729
		0.713    0.816	
	ASD - R1D	$\frac{201}{259} = 0.776$	0.744
		0.725    0.826	
Random forest using all wavelengths in the VIS/NIR range	ASD - SR	$\frac{172}{259} = 0.664$	0.65
		0.606    0.721	
	ASD - R1D	$\frac{181}{259} = 0.689$	0.757
		0.643    0.754	
Logistic regression using 7 wavelengths in the VIS/NIR/SWIR range	ASD - SR	$\frac{184}{259} = 0.71$	0.657
		0.655    0.765	
	ASD - R1D	$\frac{207}{259} = 0.8$	0.769
		0.75    0.848	
Random forest using all wavelengths in the VIS/NIR/SWIR range	ASD - SR	$\frac{161}{259} = 0.62$	0.647
		0.56    0.68	
	ASD - R1D	$\frac{170}{259} = 0.656$	0.656
		0.59    0.71	
<b>Random Forest</b>	Features from the multispectral images	$\frac{275}{334} = 0.823$	<b>0.794</b>
		<b>0.782    0.86</b>	

**Table 18: Comparison between the different algorithms**

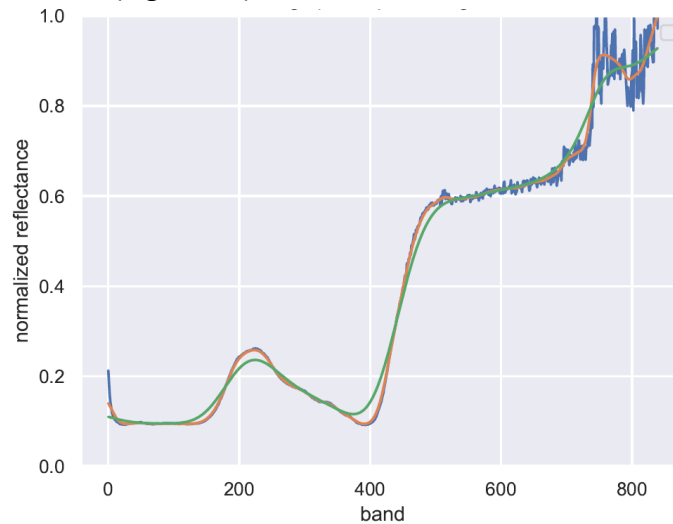
### 5.3.2.3 Significant wavelengths

The importance is calculated according to the average decrease in impurity on trees. Results reveal that the most significant feature for disease detection is LAI, the area of the plant, then the height/width ratio and then the other features calculated from the various wavelengths from the multispectral images (Figure 38). Logistic regression classifier using these two features achieved an accuracy of 63.5% and AUC of 0.64.

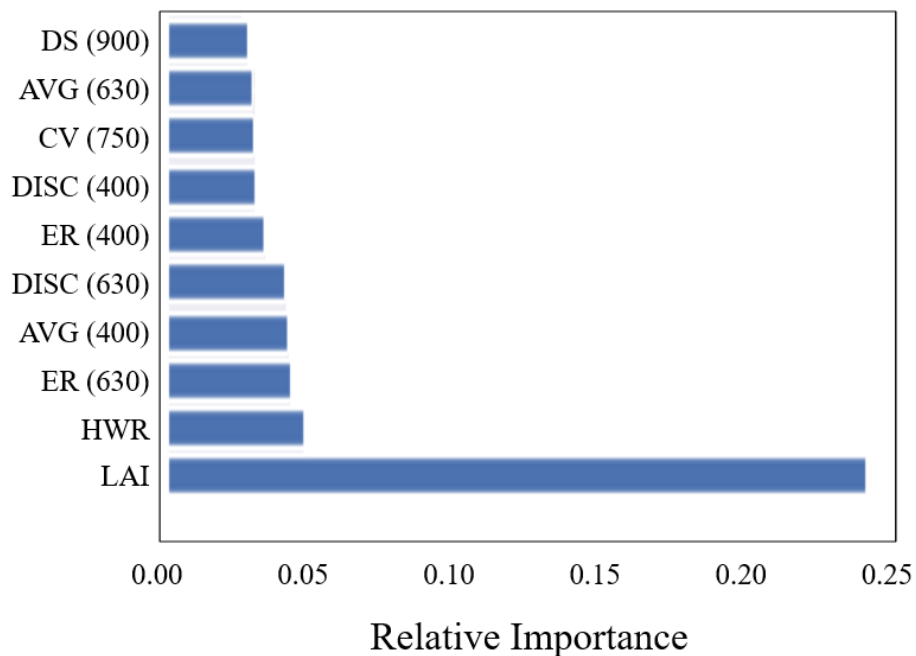
## 5.4 SENSITIVITY ANALYSIS

The wider the bandwidth the average spectral signature of the plant is smoother (Figure 37).

Results reveal that the classifier is not sensitive to the bandwidth of the images (Table 19). In addition, it can be seen that the smoothing does not significantly change the average plant spectral signature for 50nm bandwidth (Figure 37).



**Figure 37:** The plant's average spectral signature from the hyperspectral image (blue), and from the hyperspectral image after recalculating all bands by 50nm (orange) and 150nm (green) bandwidths



**Figure 38:** The importance of the features

Bandwidth	Number of images	Accuracy	AUC
Hyperspectral	1	$\frac{275}{334} = 0.823$	0.794
10nm	15	$\frac{275}{334} = 0.823$	0.794
30nm	41	$\frac{275}{334} = 0.823$	0.794
50nm	67	$\frac{275}{334} = 0.823$	0.794
150nm	211	$\frac{275}{334} = 0.823$	0.794

**Table 19: Classification results of the RF classifier for different multispectral bands**

## **6. CONCLUSIONS, LIMITATIONS, FUTURE WORK & CONTRIBUTIONS**

### **6.1 CONCLUSIONS**

The stepwise regression algorithm was used to determine the significant wavelengths for Fusarium detection using point spectral measurements from the ASD. The wavelengths that allow differentiation between spectra of infected and healthy corn plants: 400, 440, 630, 700, 750, 820 and 900nm. A random forest classifier that used features from the hyperspectral images that were calculated from the seven wavelengths yielded classification accuracy of 82%, (67.5% True Positive and 8.7% False Positive) and 0.79 AUC.

The classifier yielded better results in predicting both infected and healthy plants from the classifier used the ASD data. These results prove it is feasible to identify Fusarium at a stage which currently is not possible without destroying the plant.

### **6.2 RESEARCH LIMITATIONS**

There is not enough data to determine whether the spectral reflectance or the first derivative of the spectral reflectance is better for classifying between infected and healthy plants. Additionally, there is not enough data to determine that there is a significant difference between the classifier that used also the SWIR range compared to the classifier that used only the VIS and NIR ranges. There is a correlation between the distribution obtained from a single point in plants from the ASD, and the average distribution of the pixels in the plant obtained from the hyperspectral images.

The classification of the infected plants is consistently better for healthy plant classification. This is probably due to the imbalanced data set. An improved classifier can be constructed with a balanced data set (there are statistics methods for balancing a data set, for example removing some of the diseased plants from the data set). However, due to the small dataset this was not conducted as part of this research.

### **6.3 FUTURE WORK**

There is a statistically significant difference between the different disease levels in analysis of the distribution of the hyperspectral images. We would like to consider the choice of wavelengths from the analysis of the images themselves.

More samples should be collected over the course of the growing stages to provide a better understanding of the spectral discrimination over time. To develop a more generic classifier it is necessary to collect images in different lighting conditions, from different areas on the plants and of different varieties. Increasing the data set will also enable to build a stronger classifier that in this case was based on leave-one-out due to the small datasets.

## 6.4 CONTRIBUTIONS

This research proved the feasibility of detecting Fusarium in infected plants before the visual symptoms appear. Additionally, the importance of hyperspectral imaging has been demonstrated. However, to lower computational and sensory costs the most important wavelengths were derived and prove sufficient for disease classification. The algorithms and methods developed can be potentially used in the future to create infestation maps for site specific management of the disease. However, the algorithms presented are limited to the corn variety, the growing stage and the imaging and lighting conditions. It needed to imaging under different conditions and make the algorithm more generic, as well as to examine the possibility of imaging the corn plants from above in order to test the feasibility of hovering imaging.

## 7. BIBLIOGRAPHY

- Abdulridha, J., Ehsani, R., Abd-elrahman, A., & Ampatzidis, Y. (2019). A remote sensing technique for detecting laurel wilt disease in avocado in presence of other biotic and abiotic stresses. *Computers and Electronics in Agriculture*, 156(December 2018), 549–557. <https://doi.org/10.1016/j.compag.2018.12.018>
- Akar, Ö., & Güngör, O. (2012). Classification of multispectral images using Random Forest algorithm. *Journal of Geodesy and Geoinformation*, 90(462), 105–112. <https://doi.org/10.9733/jgg.241212.1>
- An, N., Palmer, C. M., Baker, R. L., Markelz, R. J. C., Ta, J., Covington, M. F., ... Weinig, C. (2016). Plant high-throughput phenotyping using photogrammetry and imaging techniques to measure leaf length and rosette area. *Computers and Electronics in Agriculture*, 127, 376–394. <https://doi.org/10.1016/j.compag.2016.04.002>
- Atzberger, C., Guérif, M., Baret, F., & Werner, W. (2010). Comparative analysis of three chemometric techniques for the spectroradiometric assessment of canopy chlorophyll content in winter wheat. *Computers and Electronics in Agriculture*, 73(2), 165–173. <https://doi.org/10.1016/j.compag.2010.05.006>
- Bannari, A., Morin, D., Bonn, F., & Huete, A. R. (2009). A review of vegetation indices. *Taylor & Francis*, 13(April), 95–120. <https://doi.org/10.1080/02757259509532298>
- Banttari, E., & Goodwin, P. (1985). Detection of potato viruses S, X, and Y by enzyme-linked immunosorbent assay on nitrocellulose membranes (dot-ELISA). *Plant Disease*. Retrieved from <http://www.csa.com/partners/viewrecord.php?requester=gs&collection=ENV&recid=1083666>
- Bauriegel, E., Giebel, A., Geyer, M., Schmidt, U., & Herppich, W. B. (2011). Early detection of Fusarium infection in wheat using hyper-spectral imaging. *Computers and Electronics in Agriculture*, 75(2), 304–312. <https://doi.org/10.1016/j.compag.2010.12.006>
- Behmann, J., Mahlein, A. K., Rumpf, T., Römer, C., & Plümer, L. (2015). A review of advanced machine learning methods for the detection of biotic stress in precision crop protection. *Precision Agriculture*, 16(3), 239–260. <https://doi.org/10.1007/s11119-014-9372-7>
- Blackburn, G. A. (2007). Hyperspectral remote sensing of plant pigments. *Journal of Experimental Botany*, 58(4), 855–867. <https://doi.org/10.1093/jxb/erl123>
- Bock, C. H., Poole, G. H., Parker, P. E., & Gottwald, T. R. (2010). Plant disease severity estimated visually, by digital photography and image analysis, and by hyperspectral imaging. *Critical Reviews in Plant Sciences*, 29(2), 59–107. <https://doi.org/10.1080/07352681003617285>
- Bravo, C., Moshou, D., West, J., McCartney, A., & Ramon, H. (2003). Early disease detection in wheat fields using spectral reflectance. *Biosystems Engineering*. [https://doi.org/10.1016/S1537-5110\(02\)00269-6](https://doi.org/10.1016/S1537-5110(02)00269-6)
- Browne, R. A., Murphy, J. P., Cook, B., Devaney, D., Walsh, E. J., Griffey, C. A., & Al, E. (2005). Evaluation of components of Fusarium head blight resistance in soft red winter wheat germ plasm using a detached leaf assay. *Plant Disease*, 89(4), 404–411. <https://doi.org/10.1094/PD-89-0404>

- Busemeyer, L., Mentrup, D., Möller, K., Wunder, E., Alheit, K., Hahn, V., ... Ruckelshausen, A. (2013). Breedvision - A multi-sensor platform for non-destructive field-based phenotyping in plant breeding. *Sensors (Switzerland)*, 13(3), 2830–2847. <https://doi.org/10.3390/s130302830>
- Chaerle, L., Lenk, S., Hagenbeek, D., Buschmann, C., & Van Der Straeten, D. (2007). Multicolor fluorescence imaging for early detection of the hypersensitive reaction to tobacco mosaic virus. *Journal of Plant Physiology*, 164(3), 253–262. <https://doi.org/10.1016/j.jplph.2006.01.011>
- Chaves, R., Ramírez, J., Górriz, J. M., López, M., Álvarez, I., & Segovia, F. (2009). SVM-based computer-aided diagnosis of the Alzheimer ' s disease using t -test NMSE feature selection with feature correlation weighting. *Elsevier*, 461, 293–297. <https://doi.org/10.1016/j.neulet.2009.06.052>
- Chong, I., & T, C. J. (2005). Performance of some variable selection methods when multicollinearity is present. *Chemometrics and Intelligent Laboratory Systems*, 78, 103–112. <https://doi.org/10.1016/j.chemolab.2004.12.011>
- Clark, M. F., & Adams, A. N. (1977). Characteristics of the microplate method of enzyme linked immunosorbent assay for the detection of plant viruses. *Journal of General Virology*, 34(3), 475–483. <https://doi.org/10.1099/0022-1317-34-3-475>
- Falasconi, M., Gobbi, E., Pardo, M., Della Torre, M., Bresciani, A., & Sberveglieri, G. (2005). Detection of toxigenic strains of *Fusarium verticillioides* in corn by electronic olfactory system. *Sensors and Actuators, B: Chemical*, 108(1-2 SPEC. ISS.), 250–257. <https://doi.org/10.1016/j.snb.2004.09.046>
- Feng, Y. Z., & Sun, D. W. (2012). Application of Hyperspectral Imaging in Food Safety Inspection and Control: A Review. *Critical Reviews in Food Science and Nutrition*, 52(11), 1039–1058. <https://doi.org/10.1080/10408398.2011.651542>
- Flegg, C. L., & Clark, M. F. (1979). The detection of apple chlorotic leafspot virus by a modified procedure of enzyme-linked immunosorbent assay (ELISA). *Annals of Applied Biology*. <https://doi.org/10.1111/j.1744-7348.1979.tb07413.x>
- Ge, Y., Bai, G., Stoerger, V., & Schnable, J. C. (2016). 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*, 127, 625–632. <https://doi.org/10.1016/j.compag.2016.07.028>
- Geladi, P., Burger, J., & Lestander, T. (2004). Hyperspectral imaging : calibration problems and solutions. *Chemometrics and Intelligent Laboratory Systems*, 72, 209–217. <https://doi.org/10.1016/j.chemolab.2004.01.023>
- Gilbertson, R. L., Rojas, M. R., Russell, D. R., & Maxwell, D. P. (1991). Use of the Asymmetric Polymerase Chain-Reaction and Dna Sequencing To Determine Genetic-Variability of Bean Golden Mosaic Geminivirus in the Dominican-Republic. *Journal of General Virology*, 72(1991), 2843–2848. <https://doi.org/10.1099/0022-1317-72-11-2843>
- Goetz, A. F. H. (2009). Three decades of hyperspectral remote sensing of the Earth: A personal view. *Remote Sensing of Environment*, 113(SUPPL. 1), S5–S16. <https://doi.org/10.1016/j.rse.2007.12.014>
- Goswami, R. S., & Kistler, H. C. (2004). Heading for disaster: *Fusarium graminearum* on cereal crops. *Molecular Plant Pathology*, 5(6), 515–525. <https://doi.org/10.1111/J.1364-3703.2004.00252.X>



- Gowen, A. A., Taghizadeh, M., & O'Donnell, C. P. (2009). Identification of mushrooms subjected to freeze damage using hyperspectral imaging. *Journal of Food Engineering*, 93(1), 7–12. <https://doi.org/10.1016/j.jfoodeng.2008.12.021>
- Hariharan, J., Fuller, J., Ampatzidis, Y., & Abdulridha, J. (2019). Finite Difference Analysis and Bivariate Correlation of Hyperspectral Data for Detecting Laurel Wilt Disease and Nutritional Deficiency in Avocado. *Remote Sensing*.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *IEEE Conf. on Computer Vision and Patt. Recognition*, 770–778.
- Huete, A. R. (1988). A Soil-Adjusted Vegetation Index (Savi). *Remote Sensing of Environment*.
- Jacquemoud, S., & Ustin, S. L. (2001). Leaf optical properties: a state of the art. *Proceedings of the 8th International Symposium Physical Measurements & Signatures in Remote Sensing, 8-12 January 2001, CNES, Aussois, France*, 223–232.
- Jurado, M., Vázquez, C., Marín, S., Sanchis, V., & Teresa González-Jaén, M. (2006). PCR-based strategy to detect contamination with mycotoxigenic *Fusarium* species in maize. *Systematic and Applied Microbiology*. <https://doi.org/10.1016/j.syapm.2006.01.014>
- Karami, A., Heylen, R., & Scheunders, P. (2015). Band-Specific Shearlet-Based Hyperspectral Image Noise Reduction. *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, 53(9), 5054–5066.
- Karunakaran, C., Jayas, D. S., & White, N. D. G. (2004). Identification of wheat kernels damaged by the red flour beetle using X-ray images. *Biosystems Engineering*, 87(3), 267–274. <https://doi.org/10.1016/j.biosystemseng.2003.12.002>
- Kotsiantis, S. B., Zaharakis, I. D., & Pintelas, P. E. (2006). Machine learning a review of classification and combining techniques.pdf.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In *Papers published at the Neural Information Processing Systems Conference* (pp. 1–9).
- Kumar, L., Skidmore, A. K., & Mutanga, O. (2010). Leaf level experiments to discriminate between eucalyptus species using high spectral resolution reflectance data: Use of derivatives, ratios and vegetation indices. *Geocarto International*, 25(4), 327–344. <https://doi.org/10.1080/10106040903505996>
- Kuska, M., Wahabzada, M., Leucker, M., Dehne, H. W., Kersting, K., Oerke, E. C., ... Mahlein, A. K. (2015). Hyperspectral phenotyping on the microscopic scale: Towards automated characterization of plant-pathogen interactions. *Plant Methods*, 11(1). <https://doi.org/10.1186/s13007-015-0073-7>
- Lancashire, D., Bleiholder, H., Van Den Boom, T., Langeluddeke, P., Stauss, R., Weber, E. Hack, H., & Witzemberger, A. (1991). A uniform decimal code for growth stages of crops and weeds. *Annals of Applied Biology*, 119(3), 561–601. <https://doi.org/10.1111/j.1744-7348.1991.tb04895.x>
- Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521, 436–444. <https://doi.org/10.1038/nature14539>

- LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (n.d.-a). backpropagation applied to handwritten zip code recognition. *Neural Computation*, 4, 541–551.
- LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (n.d.-b). Handwritten Digit Recognition with a Back-Propagation Network. *Advances in Neural Information Processing Systems*, 2, 396–404.
- LECUN, Y., BOTTOU, L. ´EON, Bengio, Y., & Haffner, P. (1998). Gradient-Based Learning Applied to Document Recognition. *Proceedings of the IEEE*, 86(11), 2278–2324.
- Li, L., Zhang, Q., & Huang, D. (2014). A review of imaging techniques for plant phenotyping. *Sensors (Switzerland)*, 14(11), 20078–20111. <https://doi.org/10.3390/s141120078>
- Liang, H. (2012). Advances in multispectral and hyperspectral imaging for archaeology and art conservation. *Applied Physics A: Materials Science and Processing*, 106(2), 309–323. <https://doi.org/10.1007/s00339-011-6689-1>
- Lindenthal, M., Steiner, U., Dehne, H.-W., & Oerke, E.-C. (2007). Effect of Downy Mildew Development on Transpiration of Cucumber Leaves Visualized by Digital Infrared Thermography. *Phytopathology*, 95(3), 233–240. <https://doi.org/10.1094/phyto-95-0233>
- Lins, E. C., Belasque, J., & Marcassa, L. G. (2009). Detection of citrus canker in citrus plants using laser induced fluorescence spectroscopy. *Precision Agriculture*, 10(4), 319–330. <https://doi.org/10.1007/s11119-009-9124-2>
- Lleó, L., Roger, J. M., Herrero-langreo, A., Diezma-iglesias, B., & Barreiro, P. (2011). Comparison of multispectral indexes extracted from hyperspectral images for the assessment of fruit ripening, 104, 612–620. <https://doi.org/10.1016/j.jfoodeng.2011.01.028>
- López, M. M., Bertolini, E., Olmos, A., Caruso, P., Gorris, M. T., Llop, P., ... Cambra, M. (2003). Innovative tools for detection of plant pathogenic viruses and bacteria. *International Microbiology*, 6(4), 233–243. <https://doi.org/10.1007/s10123-003-0143-y>
- Lu, G., & Fei, B. (2014). Medical hyperspectral imaging: a review. *Journal of Biomedical Optics*, 19(1), 010901. <https://doi.org/10.1117/1.JBO.19.1.010901>
- Lu, J., Ehsani, R., Shi, Y., Ana Isabel, C. De, & Wang, S. (2018). Detection of multi-tomato leaf diseases (late blight, target and bacterial spots) in different stages by using a spectral-based sensor. *Scientific Reports*, 8(1), 1–11. <https://doi.org/10.1038/s41598-018-21191-6>
- Mahlein, A.-K. (2016). Plant Disease Detection by Imaging Sensors – Parallels and Specific Demands for Precision Agriculture and Plant Phenotyping. *APS*, 100(February), 241–251.
- Manolakis, D. G., & Shaw, G. (2002). Detection algorithms for hyperspectral imaging applications. *IEEE Signal Processing Magazine*, 19(1), 29–43. <https://doi.org/10.1109/79.974724>
- Marcassa, L. G., Gasparoto, M. C. G., Belasque, J., Lins, E. C., Dias Nunes, F., & Bagnato, V. S. (2006). Fluorescence spectroscopy applied to orange trees. *Laser Physics*, 16(5), 884–888. <https://doi.org/10.1134/s1054660x06050215>
- Mcleod, L. (1993). Fusarium ear blight ( scab ) in small grain cereals — a review. *Review Literature And Arts Of The Americas*, 207–238. <https://doi.org/10.1111/j.1365-3059.1995.tb02773.x>
- Mishra, P., Asaari, M. S. M., Herrero-Langreo, A., Lohumi, S., Diezma, B., & Scheunders, P. (2017). Close

- range hyperspectral imaging of plants: A review. *Biosystems Engineering*, 164, 49–67. <https://doi.org/10.1016/j.biosystemseng.2017.09.009>
- Mohan, S. B. (1988). Evaluation of antisera raised against *Phytophthora fragariae* for detecting the red core disease of strawberries by enzyme-linked immunosorbent assay (ELISA). *Plant Pathology*. <https://doi.org/10.1111/j.1365-3059.1988.tb02066.x>
- Morales-Rodriguez, I., De Yañez-Morales, M. J., Silva-Rojas, H. V., García-De-Los-Santos, G., & Guzmán-De-Peña, D. A. (2007). Biodiversity of *Fusarium* species in Mexico associated with ear rot in maize, and their identification using a phylogenetic approach. *Mycopathologia*, 163(1), 31–39. <https://doi.org/10.1007/s11046-006-0082-1>
- Moricca, S., Ragazzi, A., Kasuga, T., & Mitchelson, K. R. (1998). Detection of *Fusarium oxysporum* f.sp. *vasinfectum* in cotton tissue by polymerase chain reaction. *Plant Pathology*, 47(4), 486–494. <https://doi.org/10.1046/j.1365-3059.1998.00262.x>
- Moshou, D., Bravo, C., Oberti, R., West, J., Bodria, L., McCartney, A., & Ramon, H. (2005). Plant disease detection based on data fusion of hyper-spectral and multi-spectral fluorescence imaging using Kohonen maps. *Real-Time Imaging*, 11(2), 75–83. <https://doi.org/10.1016/j.rti.2005.03.003>
- Mulla, D. J. (2013). Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems Engineering*, 114(4), 358–371. <https://doi.org/10.1016/j.biosystemseng.2012.08.009>
- Munkvold, G. P. (2003). Epidemiology of *Fusarium* diseases and their mycotoxins in maize ears Gary. *European Journal of Plant Pathology*, 109(2), 705–713. <https://doi.org/10.1023/A:1026078324268>
- Nair, V., & Hinton, G. E. (2010). Rectified Linear Units Improve Restricted Boltzmann Machines. *Rawat, W., & Wang, Z. (2017). Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review. Communicated by Vincent Vanhoucke, 2449, 2352–2449. <https://doi.org/10.1162/NECO.2010.1162.3>, 807–814.*
- Narvankar, D. S., Singh, C. B., Jayas, D. S., & White, N. D. G. (2009). Assessment of soft X-ray imaging for detection of fungal infection in wheat. *Biosystems Engineering*, 103(1), 49–56. <https://doi.org/10.1016/j.biosystemseng.2009.01.016>
- Neumann, M., Hallau, L., Klatt, B., Kersting, K., & Bauckhage, C. (2014). Erosion band features for cell phone image based plant disease classification. *Proceedings - International Conference on Pattern Recognition*, 3315–3320. <https://doi.org/10.1109/ICPR.2014.571>
- Nutter, F. W., Gleason, M. L., Jenco, J. H., & Christians, N. C. (1993). Assessing the accuracy, intra-rater repeatability, and inter-rater reliability of disease assessment systems. *Ecology and Epidemiology*, 83, 312–323.
- Onoyama, H., Ryu, C., Suguri, M., & Iida, M. (2013). Potential of hyperspectral imaging for constructing a year-invariant model to estimate the nitrogen content of rice plants at the panicle initiation stage. *IFAC Proceedings Volumes (IFAC-PapersOnline)* (Vol. 4). IFAC. <https://doi.org/10.3182/20130828-2-SF-3019.00054>
- Ozigis, M. S., Kaduk, J. D., & Jarvis, C. H. (2020). Detection of oil pollution impacts on vegetation using multifrequency SAR, multispectral images with fuzzy forest and random forest. *Environmental Pollution*, 256, 1–17. <https://doi.org/10.1016/j.envpol.2019.113360>

- Paulus, S., Behmann, J., Mahlein, A. K., Plümer, L., & Kuhlmann, H. (2014). Low-cost 3D systems: Suitable tools for plant phenotyping. *Sensors (Switzerland)*, 14(2), 3001–3018. <https://doi.org/10.3390/s140203001>
- Pimentel, D., Zuniga, R., & Morrison, D. (2005). Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological Economics*, 52(3 SPEC. ISS.), 273–288. <https://doi.org/10.1016/j.ecolecon.2004.10.002>
- Purcell, D. E., O'shea, M. G., Johnson, R. A., & Kokot, S. (2009). Near-infrared spectroscopy for the prediction of disease ratings for fiji leaf gall in sugarcane clones. *Applied Spectroscopy*, 63(4), 450–457. <https://doi.org/10.1366/000370209787944370>
- Ranzato, M. A., Huang, F., Boureau, Y., & Lecun, Y. (2007). Unsupervised Learning of Invariant Feature Hierarchies with Applications to Object Recognition. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, 1–8.
- Rapaport, T., Hochberg, U., Shoshany, M., Karnieli, A., & Rachmilevitch, S. (2015). Combining leaf physiology, hyperspectral imaging and partial lefor grapevine water status assessmentast squares-regression (PLS-R). *ISPRS Journal of Photogrammetry and Remote Sensing*, 109, 88–97. <https://doi.org/10.1016/j.isprsjprs.2015.09.003>
- Reid, L. M., Woldemariam, T., Zhu, X., Stewart, D. W., & Schaafsma, A. W. (2002). Effect of inoculation time and point of entry on disease severity in *Fusarium graminearum*, *Fusarium verticillioides*, or *Fusarium subglutinans* inoculated maize ears. *Canadian Journal of Plant Pathology*, 24(2), 162–167. <https://doi.org/10.1080/07060660309506991>
- Sankaran, S., Mishra, A., Ehsani, R., & Davis, C. (2010). A review of advanced techniques for detecting plant diseases. *Computers and Electronics in Agriculture*, 72(1), 1–13. <https://doi.org/10.1016/j.compag.2010.02.007>
- Savary, S., Ficke, A., Aubertot, J. N., & Hollier, C. (2012). Crop losses due to diseases and their implications for global food production losses and food security. *Food Security*, 4(4), 519–537. <https://doi.org/10.1007/s12571-012-0200-5>
- Savitzky, A., & Marcel, J. E. G. (1964). Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, 36(8), 1627–1639. <https://doi.org/10.1021/ac60214a047>
- Schuler, R. L., Kish, P. E., & Plese, C. A. (2012). Preliminary Observations on the Ability of Hyperspectral Imaging to Provide Detection and Visualization of Bloodstain Patterns on Black Fabrics. *Journal of Forensic Sciences*, 57(6), 1562–1569. <https://doi.org/10.1111/j.1556-4029.2012.02171.x>
- Seelan, S. K., Laguet, S., Casady, G. M., & Seielstad, G. A. (2003). Remote sensing applications for precision agriculture: A learning community approach. *Remote Sensing of Environment*, 88(1–2), 157–169. <https://doi.org/10.1016/j.rse.2003.04.007>
- Shaw, G., & Manolakis, D. (2002). Signal processing for hyperspectral image exploitation. *IEEE Signal Processing Magazine*, 19(1), 12–16. <https://doi.org/10.1109/79.974715>
- Simonyan, K., & Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. *ArXiv:1409.1556*, 1–14.
- Spinelli, F., Noferini, M., & Costa, G. (2006). Near infrared spectroscopy (NIRs): Perspective of fire blight detection in asymptomatic plant material. *Acta Horticulturae*, 704, 87–90.

<https://doi.org/10.17660/ActaHortic.2006.704.9>

- Stack, R., & McMullen, M. (1998). A visual scale to estimate severity of Fusarium head blight of wheat. *NDSU Extension Service*.
- Stonehouse, J. (1994). Assessment of Andean bean diseases using visual keys. *Plant Pathology*, 43(3), 519–527. <https://doi.org/10.1111/j.1365-3059.1994.tb01586.x>
- Sugumaran, R., & Voss, M. (2007). Object-Oriented Classification of LIDAR-Fused Hyperspectral Imagery for Tree.pdf.
- Szegedy, C., Liu, W., Jia, Y., Reed, S., Sermanet, P., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1–12.
- Tester, M., & Langridge, P. (2010). Breeding Technologies to Increase Crop Production in a Changing World. *Science*, 818(5967), 818–822. <https://doi.org/10.1126/science.1183700>
- Tsenkova, R., Meilina, H., Kuroki, S., & Burns, D. H. (2009). Near infrared spectroscopy using short wavelengths and leave-one-cow-out cross-validation for quantification of somatic cells in milk. *Journal of Near Infrared Spectroscopy*, 17(6), 345–351. <https://doi.org/10.1255/jnirs.868>
- Van Egmond, H. P., Schothorst, R. C., & Jonker, M. A. (2007). Regulations relating to mycotoxins in food : PPPerspectives in a global and European context. *Analytical and Bioanalytical Chemistry*, 389(1), 147–157. <https://doi.org/10.1007/s00216-007-1317-9>
- Wei, X., Liu, F., Qiu, Z., Shao, Y., & He, Y. (2014). Ripeness Classification of Astringent Persimmon Using Hyperspectral Imaging Technique. *Food and Bioprocess Technology*, 7(5), 1371–1380. <https://doi.org/10.1007/s11947-013-1164-y>
- West, J. S., Bravo, C., Oberti, R., Lemaire, D., Moshou, D., & McCartney, H. A. (2003). The potential of optical canopy measurement for targeted control of field crop diseases. *Annual Review of Phytopathology*, 41(1), 593–614. <https://doi.org/10.1146/annurev.phyto.41.121702.103726>
- Windels, C. E. (2000). Economic and Social Impacts of Fusarium Head Blight: Changing Farms and Rural Communities in the Northern Great Plains. *Phytopathology*, 90(1), 17–21. <https://doi.org/10.1094/PHYTO.2000.90.1.17>
- Wu, D., & Sun, D. (2013). Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment : A review — Part I : Fundamentals. *Innovative Food Science and Emerging Technologies*, 19, 1–14. <https://doi.org/10.1016/j.ifset.2013.04.014>
- Xin, D., Zhou, X., & Zheng, H. (2006). Contour Line Extraction from Paper-based Topographic Maps. *Journal of Information and Computing Science*, 1(5), 275–283.
- Xue, J., & Su, B. (2017). Significant remote sensing vegetation indices: a review of developments and applications. *Journal of Sensors*, Vol.2017, 17p. <https://doi.org/10.1155/2017/1353691>
- Yamashita, T., Yamashita, K., & Kamimura, R. (2007). A Stepwise AIC Method for Variable Selection in linear regression. *Taylor & Francis*, 0926, 2395–2403. <https://doi.org/10.1080/03610920701215639>
- Yu, K. Q., Zhao, Y. R., Li, X. L., Shao, Y. N., Liu, F., & He, Y. (2014). Hyperspectral imaging for mapping of total nitrogen spatial distribution in pepper plant. *PLoS ONE*, 9(12), 1–19.

<https://doi.org/10.1371/journal.pone.0116205>

- Zarco-Tejada, P. J., Camino, C., Beck, P. S. A., Calderon, R., Hornero, A., Hernández-Clemente, R., ... Navas-Cortes, J. A. (2018). Previsual symptoms of *Xylella fastidiosa* infection revealed in spectral plant-trait alterations. *Nature Plants*, 4(7), 432–439. <https://doi.org/10.1038/s41477-018-0189-7>
- Zhang, B., Wu, D., Zhang, L., & Jiao, Q. (2012). Application of hyperspectral remote sensing for environment monitoring in mining areas. *Environmental Earth Sciences*, 65(3), 649–658. <https://doi.org/10.1007/s12665-011-1112-y>
- Zhi, T., Luo, H., & Liu, Y. (2018). A Gini Impurity-Based Interest Flooding Attack Defence Mechanism in NDN. *IEEE*, 22(3), 2018–2021.

## 8. APPENDICES

### 8.1 EARLY DETECTION OF FUSARIUM INFECTION IN CORN USING SPECTRAL ANALYSIS

#### Early detection of Fusarium infection in corn using spectral analysis

##### Abstract

*Fusarium* has been pronounced worldwide as a disease of economic importance. There is an increased need to develop tools for early detection of infections and their spatial location. This work presents a non-destructive method for early detection of *Fusarium* infection by spectral analysis in the range of 400– 1000nm. A commercial corn seeds infected by drenching in a suspension containing spores of *Fusarium Graminearum*. The corn plants growth in an experimental greenhouse conditions and sampled with two spectral systems: (1) a single point and (2) A hyperspectral camera. A stepwise regression method was used to determine the significant wavelengths for Fusarium detection. Dimensionality reduction enable to build a simpler model and reduce sensor cost dramatically. Using these wavelengths, a random forest model was used to classify between infected and healthy plant using features derived from the hyperspectral images. Fusarium infection was successfully classified during the VS growth stage with 82% truth rate, 67.5% in sensitivity and 91.3% in specify. This is a significant result enabling identification at a stage which currently is not possible without destroying the plant and can be potentially used in the future to create infestation maps for site specific management of the disease.

**Keywords:** hyperspectral, spectral analysis, disease detection, fusarium, random forest

##### Introduction

The global demand for food is projected to increase by 1.5- 2 times with the world's population crossing the six billion mark and expected to increase by another three billion over the next five decades (Seelan, Laguette, Casady, & Seielstad, 2003). Plant diseases cause major production and economic losses in agriculture. It is estimated that the crop losses due to plant pathogens in United States result in about 33 billion dollars every year (Savary, Ficke, Aubertot, & Hollier, 2012).

*Fusarium* is a phytopathogenic fungus with a global distribution, capable of infecting a wide range of crop plants, including cereals such as maize, wheat or barley (Jurado, Vázquez, Marín, Sanchis, & Teresa González-Jaén, 2006). It may rapidly result in very high crop losses and quality reductions (Windels, 2000). Moreover, mycotoxins, potentially generated by these fungi, are poisonous and harmful for both human and animal nutrition (Browne et al. 2005). These compounds may be present even after removal of mycelium and since most of them are resistant to physical and chemical treatments, they usually stay in the food during processing and storage (Falasconi et al., 2005).

Early detection of the fungal species producing mycotoxins or of the mycotoxins themselves has become very important to prevent the human and animal risk deriving from entry of mycotoxins into the food chain (Falasconi et al., 2005). Traditional methods used for plant disease detection are still time-consuming, labor intensive and destructive in nature (Busemeyer et al., 2013). Various modern sensing technologies and imaging techniques have been studied for disease detection. The methods include multispectral or hyperspectral imaging (detailed in Hyper- spectral imaging chapter), infrared spectroscopy (Purcell et al., 2009; Spinelli et al., 2006), fluorescence imaging (Chaerle et al., 2007; Moshou et al., 2005), fluorescence spectroscopy (Lins, Belasque et al., 2009; Marcassa et al., 2006), RGB imaging (Neumann et al., 2014), X-ray imaging (Karunakaran et al., 2004; Narvankar et al., 2009), thermal imaging or thermography (Lindenthal et al., 2007), 3D sensors (Paulus et al., 2014).



The appearance of *Fusarium* in infected ears and plants, largely changes during the development of this disease, mostly due to degradation of chlorophyll contents and pronounced water losses. Hence, these changes lead to pronounced variations in spectral properties of infected grains and total ears. Both fungi and bacteria usually cause damages at molecular, cellular and/or tissue levels, which, in turn, can be detected as changes in the spectral signatures (West et al., 2003). Utilization of spectral analysis for detection of fungal and bacterial diseases has been investigated (Ray et al., 2017). According to (Bauriegel et al., 2011) in the spectral range of 400-1000 nm the spectral patterns of healthy and diseased ear of wheat during BBCH-stage 71–85 were significantly different.

Different classification algorithms have been employed for disease detection using multispectral imaging including Random Forests (Zhu et al., 2016), Support Vector Machine (SVM) (Rumpf et al., 2010) and Maximum Likelihood Classification (MLC) algorithms (Ozdarici-ok et al., 2015). Results reveal RF yields higher classification accuracies than other methods (Akar & Güngör, 2012). Random forest is an ensemble method for classification and regression that operates by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. The significant advantages of random forest algorithms using multispectral images is the accuracy (Lowe & Arun, 2015). The performance of RF algorithm using multispectral images revealed their advent (e.g., detect the effects of oil pollution on vegetation (Ozigis et al., 2020); detect tobacco mosaic virus (Zhu et al., 2016)).

The objective of this study was to develop a method to detect infected and healthy corn plants at early growth stages, based on their leaves spectral reflectance.

## Materials and Methods

### Experimental and Plant material

Two experiments were conducted (Table 1). Commercial corn seeds (7210) were infected and germinated in a greenhouse for a period of 3 weeks, at Evogene's research facilities (31° 52' 55.80" N 34° 50' 30.77" E). Three levels of infection were induced by drenching the seeds in a suspension containing [ $10^2 - 10^6$ ] spores of *Fusarium Graminearum* in order to create a wide range in the disease status of the plants. In addition, a control group of seeds, that were not infected, was also created. The seeds were germinated in 380 ml pots. Each experiment included seeds with 3 different treatments and a control group. In all treatment groups the same number of seeds were sown. The more spores the seeds contained, the worse the disease severity and fewer seeds germinated.

The temperature in the greenhouse was maintained at  $24 \pm 2^\circ\text{C}$  during the day and at  $20 \pm 2^\circ\text{C}$  during the night. Drip irrigation with no fertilization (fertilizer inhibits *Fusarium* infection) was applied, and irrigation frequency was defined according to plot weight at 50% water content.

Exp. 1 - June 2019			Exp. 2 - November 2019		
Treatment	Sown	Germination	Treatment	Sown	Germination
Fus $10^5$	24	21	Fus $10^6$	50	34
Fus $10^3$	24	23	Fus $10^5$	50	43
Fus $10^2$	24	20	Fus $10^4$	50	36
Untreated	24	24	Untreated	50	49
Total	96	87	Total	200	162

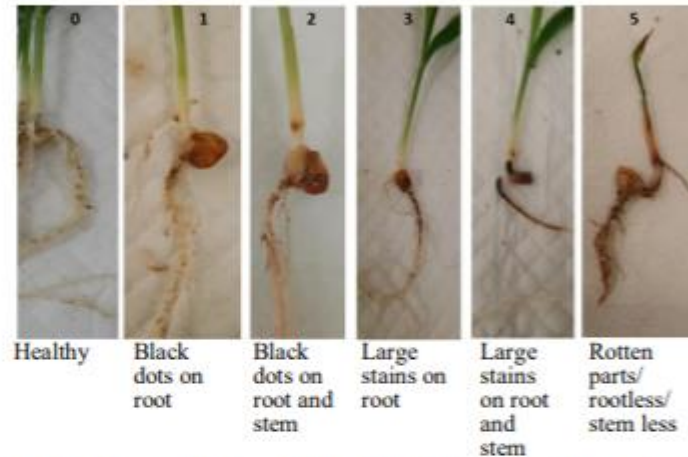


Table 1: Plant material from the two experiments according to the different treatments

### Measuring procedure

The spectral reflectance of the leaves of the plants was measured 21 days after seeding, when they reached stage V2, according to developmental stages of growth defined by the BBCH scale (Lancashire et al., 1991). Spectral measurements were conducted between 09:00 AM to 12:30PM local time (UTC-2).

Figure 1: Labelling protocol of plants to different levels of disease according to plant status after 21 days



22 days after seeding, in the day after the spectral sampling, the plants were taken out of the pots and the disease infection level was visually assessed by an agronomist expert, by visually inspecting the roots of the plant, according to a protocol similar to (Manandhar et al., 2016). The disease severity was classified from level 0 to 5 where level 0 represented a healthy and level 5 represented a severely infected plant, according to the state of the roots of each plant (Figure 1).

### Spectral measurement system

All plants were sampled with two systems (Figure 3): (1) Spectro-Radiometer (ASD), that provided the spectral reflectance of one of the plant's leaves at a single point and (2) A hyperspectral camera that provides a hyper cube that contains the spectral reflectance of each point of the plant.

#### *Spectro-Radiometer (ASD)*

Spectral measurements were performed by point spectral measurements with a field spectro-radiometer (FieldSpec 4 hi-res, ASD Inc. Malvern Panalytical, Boulder, Colorado, USA) in the range of 350- 2500 nm, with 3 nm spectral resolution in VNIR range and 8 nm in SWIR range. All measurements were conducted at approximately 1cm distance from the leaf surface, using an optical fiber with a 25° field of view, resulting to an effective sampling area of about 0.6 cm<sup>2</sup>. Data were acquired from the last leaf that was fully developed. A 10 cm diameter spectralon plate was used as a white reference, to obtain reflectance curves. Each recorded spectrum was the average of four sequential measurements at the same spot, where each measurement was the average of 30 full spectrum scans. Total acquisition time for each leaf was about 10s.

#### *Hyperspectral camera*

Spatial spectral measurements were performed by the V10E hyperspectral camera (SPECIM, Oulu, Finland). With a CMOS detector in the VIS and VNIR ranges of 400-1,000 nm. Full spectral range can be acquired with 150 fps at 1,312 spatial location sand up to 100 Hz with higher spatial resolution of 1,775 pixels. The spectral resolution is 3 nm, and the maximum imaging size is  $1,312 \times 1,775$  (spatial pixels) pixels, each pixel size is  $8 \times 8 \mu\text{m}$ . All measurements were conducted at locations at approximately 1,500 mm distance from the plant with black background (black Foam Sheet). A 100 mm diameter spectralon plate was used as a white reference, to obtain reflectance curves in each image. Total acquisition time for each plant was about 40s. In addition, a one black reference hyperspectral image was acquired when the camera's lens was completely covered.

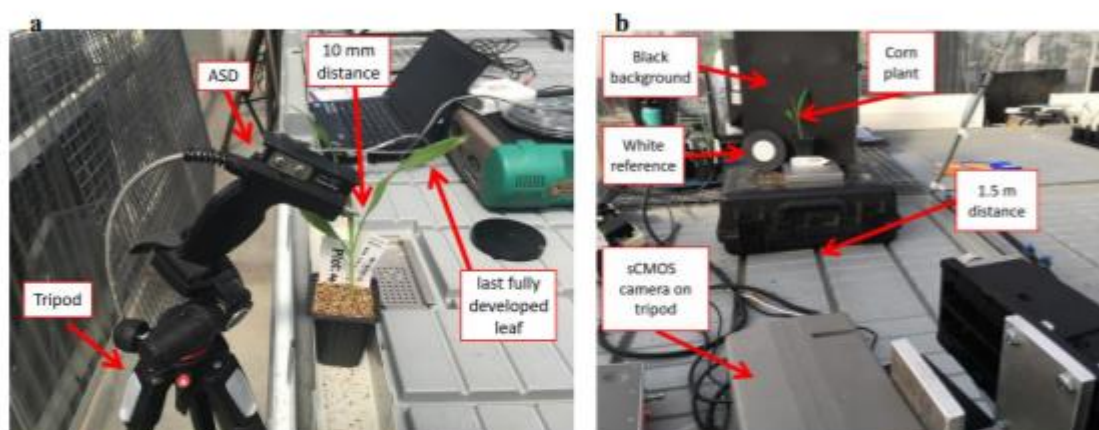


Figure 3: Experimental setting (a) ASD measurement (b) Hyperspectral camera measurement

### Data analysis

The image processing, statistical analysis and classification models were developed in the Python environment using spectral, sklearn, numpy and pandas Packages.

#### *Overview*

The point spectral measurements from the ASD were used to determine the significant wavelengths for Fusarium detection using the stepwise regression. Using the derived significant wavelengths, two classification models were developed: (1) logistic regression model that used the point spectral measurements from the ASD, (2) random forest model that used features from the hyperspectral images.

#### *Spectro-Radiometer (ASD)*

The seven most significant wavelengths (corresponding to most multispectral cameras which have 7 bands) were derived using a stepwise regression model applied with the Akaike Information Criterion (AIC, Yamashita et al. 2007). AIC estimates the relative amount of information lost by a given model: the less information a model loses, the higher the quality of that model.

Logistic regression was used to develop an algorithm to classify between healthy and infected plants using the 7 wavelengths reflectance from the ASD (model 1).



### Hyperspectral images

Pre-processing included several steps for noise reduction (Karami, Heylen, & Scheunders, 2015; Mishra et al., 2017) ) and reduction of the image size by cropping the image to the minimum size that contain all plant's pixels (Figure 4). Segmentation was performed in order to find the White Reference (WR) and plant contour. The camera measures the radiance which is captured as an integer number from 0 to 4,000. The recorded number is affected by the spectral reflectance of the plant, but also reflects the influence of nuisance signals coming from illumination effects, the detector sensitivity and the transmission properties of the optics (Geladi, Burger, & Lesstanderas, 2004). Spectral calibration was conducted to compensate for these effects in order to calculate the spectral reflectance by using black and white references (Mishra et al., 2017). The black reference hyperspectral image and the white reference in each image were used to correct the raw images by using Eq. 1.

$$I_R = \frac{I_{raw} - I_{dark}}{I_{white} - I_{dark}} \quad \text{Eq. 1}$$

Where,  $I_R$  is the calibrated reflectance image,  $I_{raw}$  is the raw intensity image measured,  $I_{dark}$  is the intensity recorded by the sensor when no light enters the camera and  $I_{white}$  is the mode value of the white reference intensity for each wavelength.

Reflectance images of the 7 most significant wavelengths were calculated and saved as multispectral cubes with 7 bands.

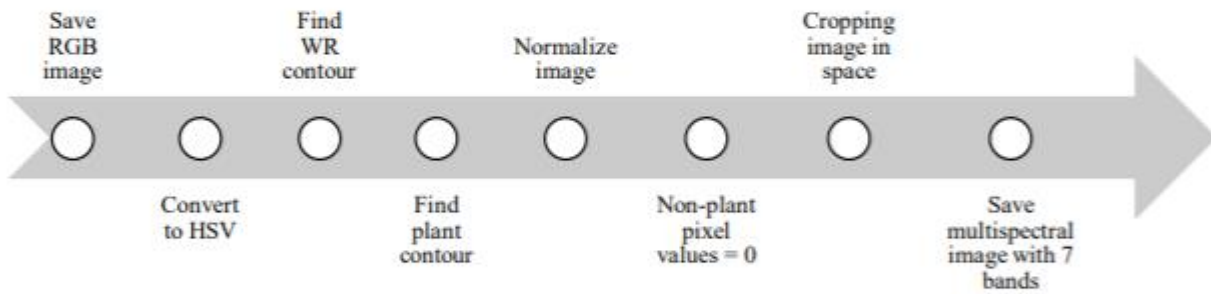


Figure 4: Hyperspectral images pre-processing flowchart

After images pre-processing, several features were calculated from the multispectral images as detailed below.

Five different features were calculated for each channel of the multispectral image. features 1-3 were examined based on common features in the literature. The first feature does not use the spatial information contained in the multispectral image. features 2- 7 used the spatial information contained in the multispectral image. features 4-5 are based on the Fusarium infection process which starts from the roots and continues to the edges of the leaves. features 6- 7 are related to the shape of the plant (equal in all wavelengths), therefore were calculated once for each plant from the first wavelength (400nm).

- (1)  $Average(\lambda)$  – average reflectance of all plant pixels in the image at wavelength  $\lambda$ .
- (2)  $SD(\lambda)$  – standard deviation of all plant pixels in the image at wavelength  $\lambda$ .

- (3) *Coefficient of Variation*( $\lambda$ ) (*CV*) –  $\frac{\sigma}{\mu}$  the average reflectance of all plant pixels in the image at wavelength  $\lambda$  divided by the standard deviation of all plant pixels in the image at wavelength  $\lambda$ .
- (4) *Discriminability*( $\lambda$ ) (*Disc*) –  $\frac{\mu_{edges} - \mu_{center}}{\mu_{all}}$  the average reflectance of pixels in the edge of the plant ( $\mu_{edges}$ ) subtraction of the pixels in the center of the plan ( $\mu_{center}$ ) divided by the average reflectance of all plant pixels ( $\mu_{all}$ ) in the image at wavelength  $\lambda$ .
- (5) *ER*( $\lambda$ ) –  $\frac{\mu_{edges}}{\mu_{center}}$  the average reflectance of pixels in the edge of the plant ( $\mu_{edges}$ ) divided by the pixels in the center of the plan ( $\mu_{center}$ ) in the image at wavelength  $\lambda$ .
- (6) *LAI* – the number of the plant's pixels (B. Zhang, et al. 2012)
- (7) *Height/Width Ratio* - the ration between the height to the width of the plant

This resulted in a total of 37 features ( $5_{features} \times 7_{bands} + 2 = 37$ ) that were calculated from each multi-spectral image. A random forest model was developed to classify between healthy and infected plants using these 37 features (model 2). Each decision tree was built over a random extraction of the observations from the dataset and a random extraction of the features (wavelengths). Since not every tree sees all the features or all the observations, this guarantees that the trees are de-correlated and therefore less prone to over-fitting (Ghojogh & Crowley, 2019). Combinations of three different parameters that define the RF model were examined by accuracy and AUC: the maximum depth of the trees, the number of features to looking for the best split and the number of trees in the forest.

#### *Performance measures*

The performance of the two models was evaluated with the leave-one-out cross validation method (Tsenkova, Meilina, Kuroki, & Burns, 2009), with AUC on the ROC curve and accuracy. The ROC curve is created by plotting the true positive rate (TPR) against the false positive rate (FPR) for various threshold settings. The area under the curve (AUC) is the definite integral of ROC and is a known metric to evaluate the model quality (Tsenkova et al, 2009). The 95% confidence interval of the accuracy was calculated to examine if there is a statistically significant difference between the models.

#### *Sensitivity analysis*

The model was evaluated using data with different bandwidths, which simulate a lower spectral resolution than that of the camera. Images with bandwidths 10, 30, 50, and 150 nm were examined. For each bandwidth, each wavelength, from the 7 wavelengths, a Gaussian image was calculated using the images from all wavelengths within the bandwidth. When the bandwidth included wavelengths that are not within the camera range [400- 1,000nm], the image was calculated with a cut Gaussian and divided by the weight of the images.

## **Results and Discussion**

Due to the small number of samples, the number of disease levels were aggregated from five to two: levels 0 and 1 were aggregated to healthy, and levels 2 to 5 were aggregated to infected (Figure 2).



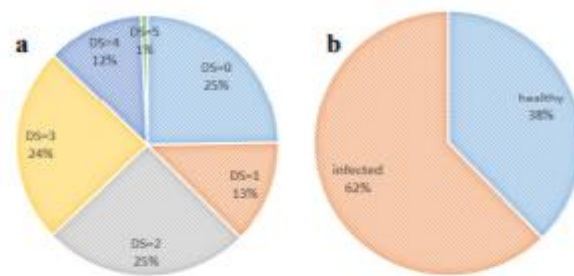


Figure 2: Distribution of the disease level of the samples from ASD

(a) by 6 disease levels (b) by 2 disease levels

Visual analysis reveals there is a correlation between the distribution obtained from a one point in plants from the ASD, and the average distribution of the pixels in the plant obtained from the hyperspectral images (Figure 6).

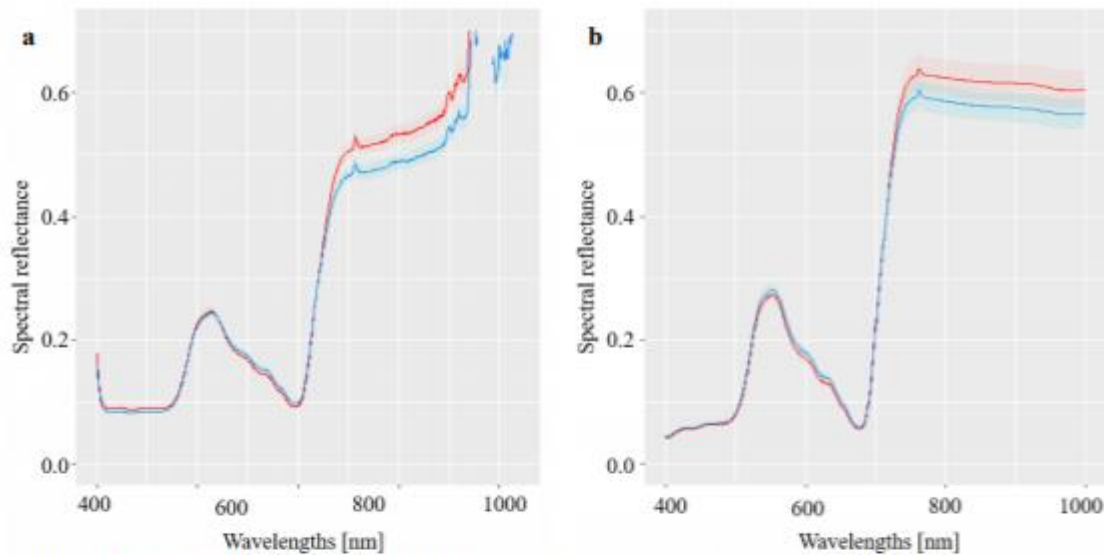


Figure 6: Mean and the 95% confidence interval of leaves' spectral reflectance of the healthy plants in blue and the infected ones in red (a) hyper-spectral images (b) ASD data

Wavelength [nm]	Importance
400	6.021
700	3.038
750	2.979
440	2.978
900	2.968
820	2.862
630	2.726

Table 2: The aggregated 7 most significant wavelengths for disease detection in the VIS/NIR range

#### *Spectro-Radiometer (ASD)*

Stepwise regression ranked results (Table 2) reveal that 400nm was found as the most significant wavelength for disease detection with additional wavelengths as noted. 400, 440, 630, 700nm are mainly influenced by leaf pigment content and 750, 820, 900nm depends on the leaf biology structure.

The reflectance of these individual spectra were integrated into a logistic regression model to classified between healthy and infected plants (model 1). The model was able to predict the state of the plant for 76% of the plants. The prediction of the infected plants is consistently better than healthy plants prediction (correctly classified 87% of the infected plants vs. to 56% of healthy plants).

#### *Hyperspectral images*

Figure 5 shows the different steps in images pre-processing flowchart, from saving the RGB image to save the normalized multispectral cube and cropped.

A random forest model was built based on the 37 features (Model 2) from the multispectral cube. The random forest includes 50 different trees, each tree in maximum depth of 6, looking for 15 of features to the best split. These parameters define the RF and were chosen to maximize accuracy and AUC (Figure 7).

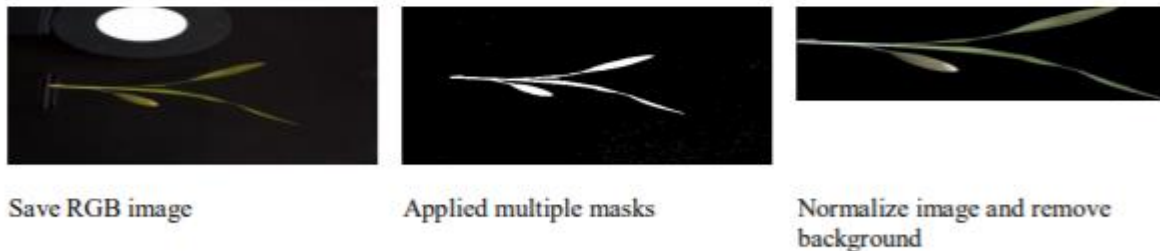


Figure 5: Hyperspectral images pre-processing

The importance is calculated according to the average decrease in impurity on trees. Results reveal that the most significant feature for disease detection is LAI, the area of the plant, then the height/width ratio and then the other features calculated from the various wavelengths from the multispectral images (Figure 8). The 2 geometric ones are the most important. Logistic regression model using these two features achieved an accuracy of 63.5% and AUC of 0.64.

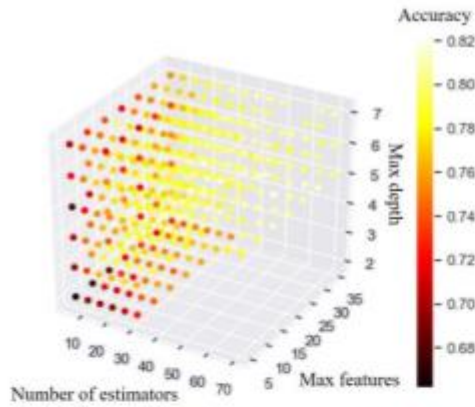


Figure 7: The accuracy of the RF model for each parameters combination

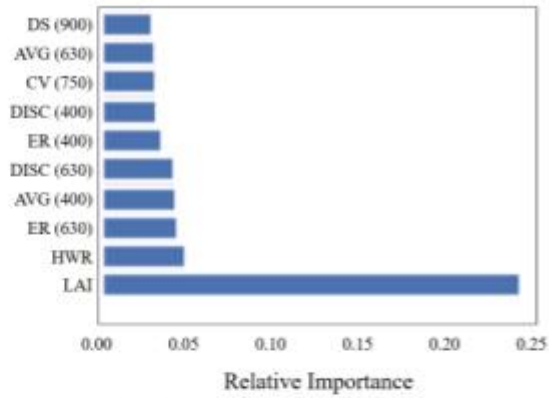


Figure 8: Features importance

Table 3 summarizes the classification results of the models with AUC and accuracy depicted for models' comparison. The model which was based on the hyperspectral features (model 2) was able to predict the state of the plant for 82% of the plants compared to 76% in model 1. This result is very significant since identification was carried out at an early stage where it was not possible to classify between healthy and infected plants without destroying the plant (Chemical testing or taken the plant out and visually assessed the roots).

Model	Performance measure		
	AUC	Accuracy	
<b>Model 1</b>			
Logistic Regression - ASD	0.729	0.76	Confidence interval
			0.713    0.816
<b>Model 2</b>			
Random Forest - features from MSI	0.794	0.82	Confidence interval
			0.782    0.864

Table 3: Summary of classification results of the models

The density plots of the predicted probability for the healthy plants and for the infected plants (Figure 9) reveal the same trend for the two models. There is difference between the density of the predicted probability of the infected and the density of the healthy plant as expected. Prediction of the infected plants is consistently better than healthy plants prediction (correctly classified 87% of the infected plants vs. to 56% of healthy plants in model 1 and 91.3% vs. 67.5% in model 2). Model 2 is better in predicting both healthy and infected plants (67.5% vs. 58.7% in sensitivity and 91.3% vs. 87% in specificity).



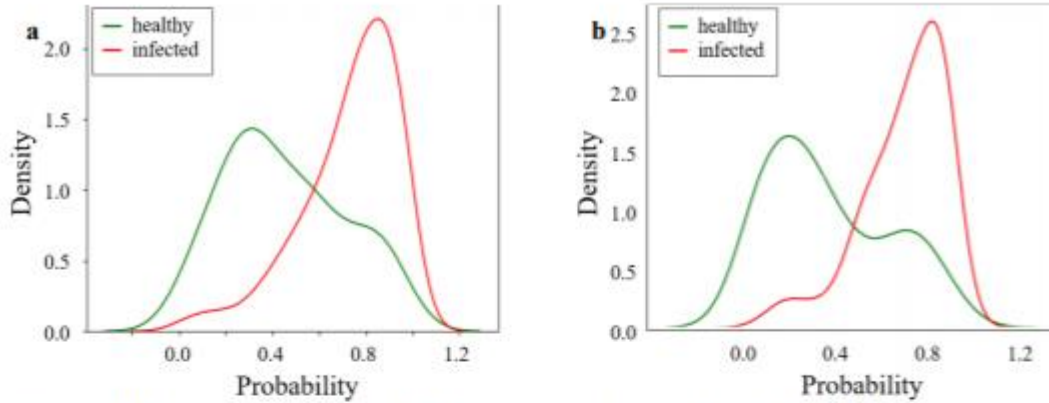


Figure 9: Density plot distributions of the predicted values of the healthy and the infected plants of (a) model 1 (b) model 2

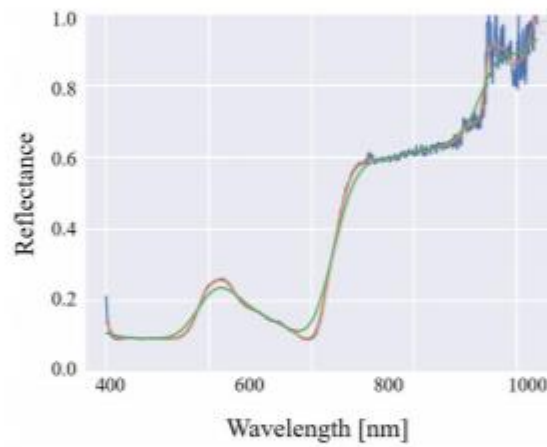


Figure 10: The average spectral signature from the hyperspectral image (blue), and after recalculating the image for 50nm bandwidth (orange), and for 150nm bandwidth (green)

#### *Sensitivity analysis*

Result reveal that the wider the bandwidth the average spectral signature of the plant is smoother. The smoothing does not significantly change the average plant spectral signature for 50nm bandwidth in the wavelengths with the SNR ratio bigger than 12 (Figure 10).

The 37 features were calculated using the smoothed images with the different bandwidths. Model 2 (based on the features that calculated from the multispectral cube without smoothing) was evaluated using these features (features that calculated from the 50nm bandwidth multispectral cube and the 150nm bandwidth multispectral cube). The classification results of the model did not change for any of the bandwidths.



The model is not sensitive to the bandwidth of the images. The model has a large number of weights to learn. The 2 geometric features are the most important in the model, these features are not affected by bandwidth.

## Conclusions

Non-destructive and rapid methods to assess diseases, yield and other traits for large numbers of samples is essential for precision agriculture and high throughput phenotyping.

To the best of our knowledge, early detection of Fusarium infection without distractive phenotyping is still a big challenge that might have a massive impact on the future agriculture economy. Therefore, classification accuracy of 82% applying spectral analysis is a step towards early detection of the disease.

The evaluated wavelength range 400- 1000 nm reliably detected corn plants infected with Fusarium. Dimensionality reduction enable to build a simpler model and reduce sensor cost dramatically. The stepwise regression model was used to determine the significant wavelengths for Fusarium detection using point spectral measurements from the ASD. The wavelengths that allow differentiation between spectra of infected and healthy corn plants: 400, 440, 630, 700, 750, 820 and 900nm. The random forest model that used features from the hyperspectral images calculated from the seven wavelengths yielded better results in predicting both healthy and infected plants.

The narrower the bandwidth, the higher the cost of the sensors. The model is found to be insensitive to bandwidth, so we can lower the cost of so the sensors significantly.

For future work, more samples should be collected over the course of the growing stages to provide a better understanding of the spectral discrimination over time. In addition, it is necessary to collect images in different lighting conditions, from different area and of different varieties in order to build a model that will be more generic. In this research because of the small dataset we used the leave one out method to evaluate the models. Increasing the data set will allow to split the data to train and test set and build a stronger model.

## Acknowledgments

This research was partially supported by the Israeli Innovation Authority, fund number 580653293 and by Rabbi W. Gunther Plaut Chair in Manufacturing Engineering at Ben-Gurion University of the Negev. We would like to thank the technical team from Evogene (Noa Guedj, Aviv Himmel, Tomer Mahler) for their help in running the experiments.

## References

- Akar, Ö., & Güngör, O. (2012). Classification of multispectral images using Random Forest algorithm. *Journal of Geodesy and Geoinformation*, 90(462), 105–112. <https://doi.org/10.9733/jgg.241212.1>
- Bannari, A., Morin, D., Bonn, F., & Huete, A. R. (2009). A review of vegetation indices. *Taylor & Francis*, 13(April), 95–120. <https://doi.org/10.1080/02757259509532298>
- Bauriegel, E., Giebel, A., Geyer, M., Schmidt, U., & Herppich, W. B. (2011). Early detection of Fusarium infection in wheat using hyper-spectral imaging. *Computers and Electronics in*

*Agriculture*, 75(2), 304–312. <https://doi.org/10.1016/j.compag.2010.12.006>

- Browne, R. A., Murphy, J. P., Cook, B., Devaney, D., Walsh, E. J., Griffey, C. A., & Al, E. (2005). Evaluation of components of Fusarium head blight resistance in soft red winter wheat germ plasm using a detached leaf assay. *Plant Disease*, 89(4), 404–411. <https://doi.org/10.1094/PD-89-0404>
- Busemeyer, L., Mentrup, D., Möller, K., Wunder, E., Alheit, K., Hahn, V., ... Ruckelshausen, A. (2013). Breedvision - A multi-sensor platform for non-destructive field-based phenotyping in plant breeding. *Sensors (Switzerland)*, 13(3), 2830–2847. <https://doi.org/10.3390/s130302830>
- Chaerle, L., Lenk, S., Hagenbeek, D., Buschmann, C., & Van Der Straeten, D. (2007). Multicolor fluorescence imaging for early detection of the hypersensitive reaction to tobacco mosaic virus. *Journal of Plant Physiology*, 164(3), 253–262. <https://doi.org/10.1016/j.jplph.2006.01.011>
- Falasconi, M., Gobbi, E., Pardo, M., Della Torre, M., Bresciani, A., & Sberveglieri, G. (2005). Detection of toxigenic strains of *Fusarium verticillioides* in corn by electronic olfactory system. *Sensors and Actuators, B: Chemical*, 108(1-2 SPEC. ISS.), 250–257. <https://doi.org/10.1016/j.snb.2004.09.046>
- Geladi, P., Burger, J., & Lestander, T. (2004). Hyperspectral imaging : calibration problems and solutions. *Chemometrics and Intelligent Laboratory Systems*, 72, 209–217. <https://doi.org/10.1016/j.chemolab.2004.01.023>
- Ghojogh, B., & Crowley, M. (2019). The Theory Behind Overfitting, Cross Validation, Regularization, Bagging, and Boosting: Tutorial. Cornell University.
- Jurado, M., Vázquez, C., Marín, S., Sanchis, V., & Teresa González-Jaén, M. (2006). PCR-based strategy to detect contamination with mycotoxigenic *Fusarium* species in maize. *Systematic and Applied Microbiology*. <https://doi.org/10.1016/j.syapm.2006.01.014>
- Karami, A., Heylen, R., & Scheunders, P. (2015). Band-Specific Shearlet-Based Hyperspectral Image Noise Reduction. *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, 53(9), 5054–5066.
- Karunakaran, C., Jayas, D. S., & White, N. D. G. (2004). Identification of wheat kernels damaged by the red flour beetle using X-ray images. *Biosystems Engineering*, 87(3), 267–274. <https://doi.org/10.1016/j.biosystemseng.2003.12.002>
- Lancashire, D., Bleiholder, H., Van Den Boom, T., Langeluddeke, P., Stauss, R., Weber, E. Hack, H., & Witzinger, A. (1991). A uniform decimal code for growth stages of crops and weeds. *Annals of Applied Biology*, 119(3), 561–601. <https://doi.org/10.1111/j.1744-7348.1991.tb04895.x>
- Lindenthal, M., Steiner, U., Dehne, H.-W., & Oerke, E.-C. (2007). Effect of Downy Mildew Development on Transpiration of Cucumber Leaves Visualized by Digital Infrared



- Thermography. *Phytopathology*, 95(3), 233–240. <https://doi.org/10.1094/phyto-95-0233>
- Lins, E. C., Belasque, J., & Marcassa, L. G. (2009). Detection of citrus canker in citrus plants using laser induced fluorescence spectroscopy. *Precision Agriculture*, 10(4), 319–330. <https://doi.org/10.1007/s11119-009-9124-2>
- Lowe, B., & Arun, K. (2015). Multispectral Image Analysis Using Random Forest. *International Journal on Soft Computing*. <https://doi.org/10.5121/ijsc.2015.6101>
- Manandhar, H. K., Timila, R. D., Sharma, S., Joshi, S., Manandhar, S., Gurung, S. B., ... Sthapit, B. R. (2016). *A field guide for identification and scoring methods of diseases in the mountain crops of Nepal*.
- Marcassa, L. G., Gasparoto, M. C. G., Belasque, J., Lins, E. C., Dias Nunes, F., & Bagnato, V. S. (2006). Fluorescence spectroscopy applied to orange trees. *Laser Physics*, 16(5), 884–888. <https://doi.org/10.1134/s1054660x06050215>
- Mishra, P., Asaari, M. S. M., Herrero-Langreo, A., Lohumi, S., Diezma, B., & Scheunders, P. (2017). Close range hyperspectral imaging of plants: A review. *Biosystems Engineering*, 164, 49–67. <https://doi.org/10.1016/j.biosystemseng.2017.09.009>
- Moshou, D., Bravo, C., Oberti, R., West, J., Bodria, L., McCartney, A., & Ramon, H. (2005). Plant disease detection based on data fusion of hyper-spectral and multi-spectral fluorescence imaging using Kohonen maps. *Real-Time Imaging*, 11(2), 75–83. <https://doi.org/10.1016/j.rti.2005.03.003>
- Narvankar, D. S., Singh, C. B., Jayas, D. S., & White, N. D. G. (2009). Assessment of soft X-ray imaging for detection of fungal infection in wheat. *Biosystems Engineering*, 103(1), 49–56. <https://doi.org/10.1016/j.biosystemseng.2009.01.016>
- Neumann, M., Hallau, L., Klatt, B., Kersting, K., & Bauckhage, C. (2014). Erosion band features for cell phone image based plant disease classification. *Proceedings - International Conference on Pattern Recognition*, d, 3315–3320. <https://doi.org/10.1109/ICPR.2014.571>
- Ozdarici-ok, A., Ok, A. O., & Schindler, K. (2015). Mapping of Agricultural Crops from Single High-Resolution Multispectral Images—Data-Driven Smoothing vs. Parcel-Based Smoothing. *Remote Sensing*, 5611–5638. <https://doi.org/10.3390/rs70505611>
- Ozigis, M. S., Kaduk, J. D., & Jarvis, C. H. (2020). Detection of oil pollution impacts on vegetation using multifrequency SAR , multispectral images with fuzzy forest and random forest. *Environmental Pollution*, 256, 1–17. <https://doi.org/10.1016/j.envpol.2019.113360>
- Paulus, S., Behmann, J., Mahlein, A. K., Plümer, L., & Kuhlmann, H. (2014). Low-cost 3D systems: Suitable tools for plant phenotyping. *Sensors (Switzerland)*, 14(2), 3001–3018. <https://doi.org/10.3390/s140203001>
- Purcell, D. E., O'shea, M. G., Johnson, R. A., & Kokot, S. (2009). Near-infrared spectroscopy for the prediction of disease ratings for fiji leaf gall in sugarcane clones. *Applied*

- Spectroscopy*, 63(4), 450–457. <https://doi.org/10.1366/000370209787944370>
- Ray, M., Ray, A., Dash, S., Mishra, A., Achary, K. G., Nayak, S., & Singh, S. (2017). Fungal disease detection in plants: Traditional assays, novel diagnostic techniques and biosensors. *Biosensors and Bioelectronics*, 87(August 2016), 708–723. <https://doi.org/10.1016/j.bios.2016.09.032>
- Rumpf, T., Mahlein, A., Steiner, U., Oerke, E., Dehne, H., & Plümer, L. (2010). Early detection and classification of plant diseases with Support Vector Machines based on hyperspectral reflectance. *Computers and Electronics in Agriculture*, 74(1), 91–99. <https://doi.org/10.1016/j.compag.2010.06.009>
- Savary, S., Ficke, A., Aubertot, J. N., & Hollier, C. (2012). Crop losses due to diseases and their implications for global food production losses and food security. *Food Security*, 4(4), 519–537. <https://doi.org/10.1007/s12571-012-0200-5>
- Seelan, S. K., Laguet, S., Casady, G. M., & Seielstad, G. A. (2003). Remote sensing applications for precision agriculture: A learning community approach. *Remote Sensing of Environment*, 88(1–2), 157–169. <https://doi.org/10.1016/j.rse.2003.04.007>
- Spinelli, F., Noferini, M., & Costa, G. (2006). Near infrared spectroscopy (NIRS): Perspective of fire blight detection in asymptomatic plant material. *Acta Horticulturae*, 704, 87–90. <https://doi.org/10.17660/ActaHortic.2006.704.9>
- Tsenkova, R., Meilina, H., Kuroki, S., & Burns, D. H. (2009). Near infrared spectroscopy using short wavelengths and leave-one-cow-out cross-validation for quantification of somatic cells in milk. *Journal of Near Infrared Spectroscopy*, 17(6), 345–351. <https://doi.org/10.1255/jnirs.868>
- West, J. S., Bravo, C., Oberti, R., Lemaire, D., Moshou, D., & McCartney, H. A. (2003). The potential of optical canopy measurement for targeted control of field crop diseases. *Annual Review of Phytopathology*, 41(1), 593–614. <https://doi.org/10.1146/annurev.phyto.41.121702.103726>
- Windels, C. E. (2000). Economic and Social Impacts of Fusarium Head Blight: Changing Farms and Rural Communities in the Northern Great Plains. *Phytopathology*, 90(1), 17–21. <https://doi.org/10.1094/PHTO.2000.90.1.17>
- Xue, J., & Su, B. (2017). Significant remote sensing vegetation indices: a review of developments and applications. *Journal of Sensors*, Vol.2017, 17p. <https://doi.org/10.1155/2017/1353691>
- Yamashita, T., Yamashita, K., & Kamimura, R. (2007). A Stepwise AIC Method for Variable Selection in linear regression. *Taylor & Francis*, 0926, 2395–2403. <https://doi.org/10.1080/03610920701215639>
- Zhang, B., Wu, D., Zhang, L., & Jiao, Q. (2012). Application of hyperspectral remote sensing for environment monitoring in mining areas. *Environmental Earth Sciences*, 65(3), 649–658.



## 8.2 ECPA CONFERENCE PAPER

### Early detection of Fusarium infection in corn using spectral analysis

T. Sandovsky<sup>1</sup>, Y. Edan<sup>1</sup>, S. Gad<sup>1</sup>, A. Etzioni<sup>2</sup>, T. Nacson<sup>3</sup>, V. Alchanatis<sup>3</sup>

<sup>1</sup> *Department of Industrial Engineering & Management, Ben-Gurion University of the Negev, Beer Sheva 8410501, Israel*

<sup>2</sup> *Evogene, 13, Gad Feinstein, Rehovot 76120, Israel*

<sup>3</sup> *Institute of Agricultural Engineering, Agricultural Research Organization (ARO), Volcani Center, Rishon-LeZion, Israel*  
[victor@volcani.agri.gov.il](mailto:victor@volcani.agri.gov.il)

#### Abstract

This work presents a non-destructive methodology for early detection of Fusarium infection, by spectral analysis in the 350– 2500nm range. Corn plants in greenhouse conditions were analyzed using spectral analysis. The Lasso model was used to differentiate infected from non-infected plants based on the first derivative of leaf spectral reflectance. Fusarium infection was successfully recognized in plants at V2 growth stage with 74% success rate. This result enables infection detection at a stage which currently is not possible without destroying the plant, which can be further applied to map the disease in field scale.

**Keywords:** multispectral, spectral analysis, disease detection, fusarium

#### Introduction

Plant diseases cause major production and economic losses in agriculture. It is estimated that the crop losses due to plant pathogens in the USA amount to about 33 billion dollars every year (Savary et al, 2012). In recent years, Fusarium has been pronounced worldwide as a disease of economic importance (Windels, 2000). Fusarium is a phytopathogenic fungus with a global distribution, capable of infecting a wide range of crop plants, including cereals such as maize, wheat or barley (Jurado et al, 2006). It may rapidly result in very high crop losses and quality reductions. Moreover, mycotoxins, potentially generated by these fungi, are poisonous and harmful for both human and animal nutrition (Browne et al. 2005). These compounds may be present even after removal of mycelium and since most of them are resistant to physical and chemical treatments, they usually stay in the food during processing and storage (Falasconi et al., 2005).

Early detection of the fungal species producing mycotoxins or of the mycotoxins themselves has become very important to prevent the human and animal risk deriving from entry of mycotoxins into the food chain (Falasconi et al., 2005). Traditional methods used for plant assessment are still time-consuming, labor intensive and destructive in nature (Busemeyer et al., 2013). Introduction of modern sensing technologies can improve crop yield, provide information to enable better in-field management decisions, reduce chemical and fertilizer costs through more efficient application, permit more accurate farm records, increase profit margin and reduce pollution (Li et al. 2014).

The appearance of Fusarium in infected ears and plants, largely changes during the development of this disease, mostly due to degradation of chlorophyll contents and pronounced water losses. Hence, these changes lead to pronounced variations in spectral properties of infected grains and total ears. Both fungi and bacteria usually cause damage

at molecular, cellular and/or tissue levels, which, in turn, can be detected as changes in the spectral signatures (West et al., 2003).

Utilization of spectral analysis for detection of fungal and bacterial diseases is currently extensively under investigation with several working applications (Ray et al., 2017). According to Bauriegel et al. (2011), in the spectral range of 400-1000 nm, there was a significant difference between the spectral patterns of healthy and diseased ears of wheat during BBCH-stage 71–85 (Lancashire et al. 1991). There are no known works of early detection of *Fusarium* in corn.

The objective of this study was to develop a method to detect infected and non- infected corn plants at early growth stages, before it can be visually identified on the leaves, based on leaf spectral reflectance.

## **Material and Methods**

### Plant material

Corn seeds were infected and germinated in a greenhouse for a period of 3 weeks, at Evogene's research facilities in central Israel (31° 52' 55.80" N 34° 50' 30.77" E). Twelve bio pesticide microbial treatments were applied in two commercial corn varieties (P2088, 7210), on infected and non-infected plants in order to create variation in disease status of the various plants. Infection was created by drench with  $10^4$  spores of *Fusarium Graminearum*. Four seeds were germinated in 380 ml pots. Five pots were repeated for each treatment. 140 pots were sampled, each pot contained 4 plants, but during the experiment only one plant (the same plant each time) was sampled.

The greenhouse temperature was maintained at  $24\pm 2^\circ\text{C}$  during the day and at  $20\pm 2^\circ\text{C}$  during the night. Drip irrigation with no fertilization (inhibits fusarium infection) was applied, and irrigation frequency was defined according to plot weight at 50% water content.

### Experimental and measuring procedure

The plants were sampled four times during a two-week period (from 11 days after seeding, stage V1, until 22 days after seeding, stage V2) between 09:00 to 12:30 local time (UTC-2). Developmental stages of growth were graded according to the BBCH scale (Lancashire et al. 1991). At each sampling day, all plants were sampled with a field spectro-radiometer. Data were acquired from the last leaf that was fully developed. After the last sampling day, the roots of the plants were exposed and the disease infection level was manually examined by an agronomist expert. The disease was classified from level 0 to 5 where 0 is healthy, 5 is an infected plant, according to the state of the roots of the different plants (Figure 1). Due to the small number of samples, the number of infestation levels were aggregated from 6 to 2: levels 0 and 1 were aggregated to non-infected, and levels 2 to 5 were aggregated to infected.

### Spectral measurement system

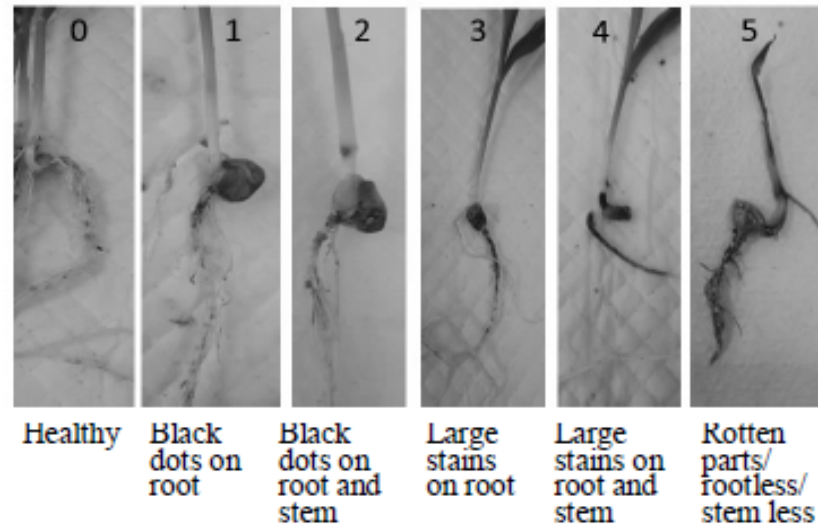
Spectral measurements were performed by point spectral measurements with a field spectro-radiometer (ASD FieldSpec 4 hi-res, ASD Inc. Malvern Panalytical, Boulder, Colorado, USA) with a 25° field of view in the range of 350-2500 nm with 3 nm spectral resolution in the visible and near- infrared (VNIR) range and 8 nm in the short- wave infrared (SWIR) range. All measurements were conducted at fixed locations at approximately 10 mm distance from the leaf surface, resulting to an effective sampling



area of about 60 mm<sup>2</sup>. A 100 mm diameter spectralon plate was used as a white reference, to obtain reflectance curves.

Each recorded spectrum was the average of four sequential measurements at the same spot, where each measurement was the average of 30 full spectrum scans. Total acquisition time for each leaf was about 10 s.

Figure 1:  
Labelling  
protocol of  
plants to  
different levels  
of disease  
according to  
plant status.



#### Data analysis

##### *Normalized Difference Indices*

The normalized difference vegetation index (NDVI) (Rouse et al. 1974) is the most widely used for retrieval of vegetation canopy biophysical properties (Jiang et al. 2006). This index normalizes the differences between two selected bands. It has been related to crop variables such as biomass, leaf area, plant cover, leaf gap fraction, nitrogen and chlorophyll in cereals (Hansen & Schjoerring 2003). NDVI values can range from (-1) to (1) while the normalization is effective in standardizing the spectral response to observed targets.

The generic form of normalized difference indices (NDI) was used in this work, where all combinations of wavelengths  $\{i, j\}$  in  $\{350 \dots 2500\text{nm}\}$  were substituted in Eq. 1.

$$NDI = \frac{R_i - R_j}{R_i + R_j} \quad (1)$$

##### *Partial Least Squares Regression (PLSR)*

PLS is a method that uses data compression to reduce the large number of collinear measured spectral variables, to a few orthogonal latent variables (LV's). It is employed in remote sensing for studying vegetation and soil characteristics (Atzberger et al. 2010). The LVs represent the relevant structural information, which is present in the reflectance measurements to predict the dependent variable. In principle, PLS regression uses component projection successively to find latent structures. Visual inspection of score-plots and validation residual variance plots was used to find the optimal number of LVs, to prevent over-fitting. In most cases, this procedure can reduce the number of spectral variables to a few independent LVs. Validation of the models was carried out by the area under the curve (AUC) of receiver operating characteristic curve (ROC) – explained below.

##### *Performance measures*

A receiver operating characteristic curve (ROC) is a graphical plot that illustrates the diagnostic ability of a binary classifier system as its discrimination threshold is varied.

The ROC curve is created by plotting the true positive rate (TPR) against the false positive rate (FPR) for various threshold settings. The area under the curve (AUC) is the definite integral of ROC and is a known metric to evaluate the model quality. The performance of all the models was evaluated with the leave-one-out cross validation method (Tsenkova et al, 2009).

## Results and Discussion

Data analyses were performed on the data from the last sampling day, 22 days after seeding, stage V2, closest to the day when the manual classification of disease level was performed. Different models were built for the P2088 variety.

It has been suggested that spectral derivatives have important advantages over spectral reflectance, such as their ability to reduce variability due to changes in illumination reflectance (Blackburn, 2007). In this work, the first derivative of the spectral reflectance was calculated, and simultaneous smoothing was applied (using the Savgol function implemented with R software<sup>1</sup>) (Savitzky & Marcel, 1964), in order to minimize noise from the environment.

The first derivative of all individual spectra was subjected to a partial least squares regression (PLS) to build a model that discriminates between non- infected and infected plants (model 1).

In addition, the confidence intervals of the mean for the non- infected plants and the infected ones were calculated, to find wavelength ranges where there is a difference between the mean of the non- infected and the infected plants.

Figure 2 shows the spectral reflectance and the first derivative of the spectral reflectance of infected and non-infected plants in the entire spectral range. The graph shows that there are several wavelength ranges where there is a statistically significant difference between the mean of the non- infected and the infected plants.

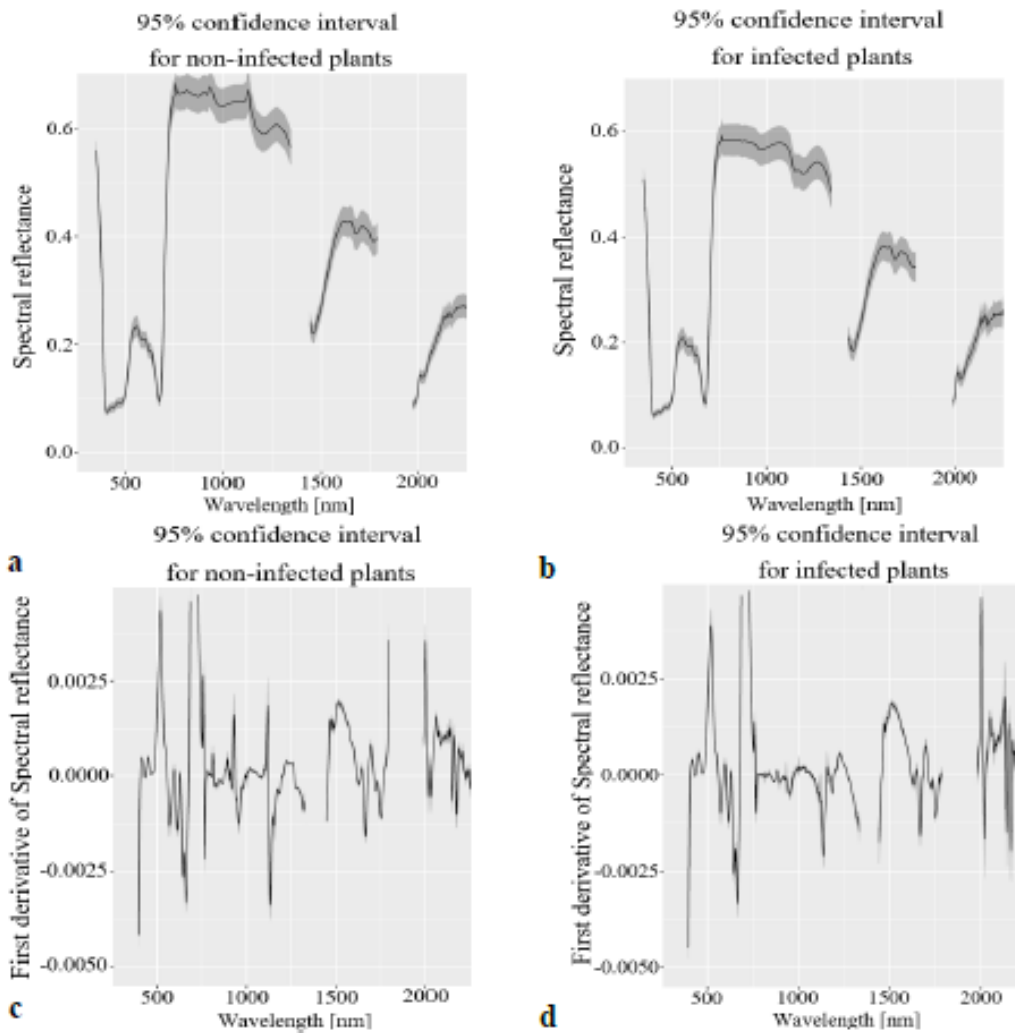
Although there was a significant difference in reflectance between infected and non-infected plants, this measure was not used despite its low computation complexity. Instead the derivative of reflectance was used since it is known to minimize noise and (Blackburn, 2007) and revealed more significant results.

Using the first derivative, the wavelengths with the most significant differences between non- infected and infected plants were derived using the statistical T-test. The most significant wavelengths were 715, 755, 920, 950, 975, 1160, 1170 and 1280 nm. These wavelengths marginally include spectral regions sensitive to leaf [Chl] (715nm) while all remaining wavelengths are near-infrared (NIR), sensitive to leaf parenchyma conditions. Using the first derivative at these wavelengths, a classification model was built to separate non- infected from infected plants. The least absolute shrinkage and selection operator (Lasso) classification method was used (Tibshirani, 1996). The Lasso classifier model simultaneously conducts prediction and variable selection, selecting the most relevant wavelengths to discriminate between non- infected and infected plants (model 2). Classification results are depicted in table 1, model 2.

---

<sup>1</sup> <https://www.rdocumentation.org/packages/pracma/versions/1.9.9/topics/savgol>





**Figure 2:** Mean and the 95% confidence interval of leaves' spectral reflectance of (a) non-infected plants and (b) infected plants. Mean and the 95% confidence interval of first derivative of leaf spectral reflectance of (c) non-infected and (d) of infected plants.

An additional model was built, based on NDIs. Figure 3 shows the correlation between the normalized difference index (NDI) to the plant visually assessed level of the disease, for every two-band combination in a defined spectral range.

The areas with the higher correlation were for the following NDIs:

$$\frac{R_{890}-R_{900}}{R_{890}+R_{900}}, \frac{R_{950}-R_{1040}}{R_{950}+R_{1040}}, \frac{R_{1020}-R_{1120}}{R_{1020}+R_{1120}}, \frac{R_{950}-R_{1245}}{R_{950}+R_{1245}}, \frac{R_{1155}-R_{1220}}{R_{1155}+R_{1220}}.$$

Logistic regression models were examined for these indices. A logistic model is a widely used statistical model that uses a logistic function to model a binary dependent variable. The model which maximized the value of the AUC was based on the NDI with wavelengths 1155, 1220nm. Results are depicted in table 1, model 3.

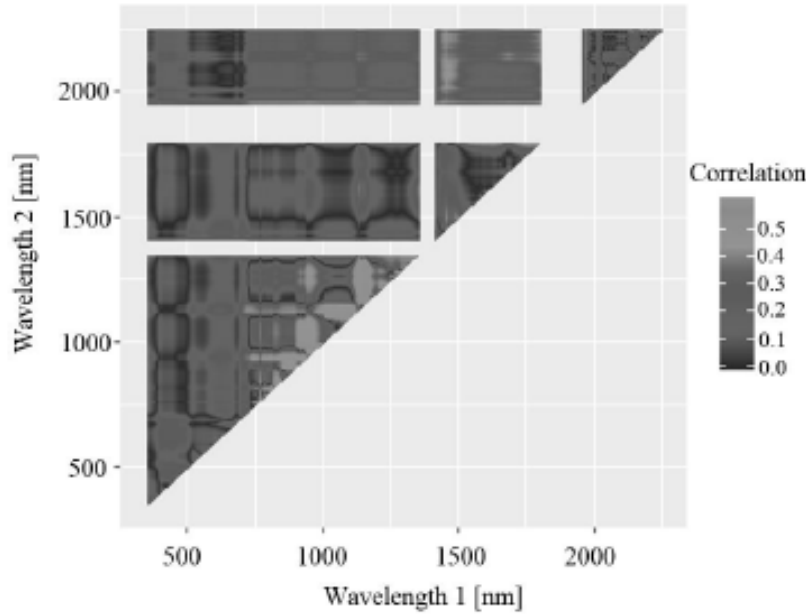


Figure 3: Correlation map of all possible two-band combinations by NDIs with the visually assessed level of the disease.

Table 1 summarizes the classification results of the three models with AUC and accuracy depicted for model comparison. There is no significant difference between models 1, 2 and 3. The best model according to the truth rate values is model 2 followed by models 1 and 3. Model 3 yielded best results for AUC values. Furthermore, models 2 and 3 are simpler than model 1, since in model 1 all the wavelengths in the range 350- 2500nm are included, whereas in models 2 and 3, only significant wavelengths are included. Model 2 was able to predict the state of the plant for 74% of the plants. This result is very significant since identification was carried out at an early stage where it was not possible to classify between non- infected and infected plants without destroying the plant.

Table 1: Summary of classification results of the three models.

Model	Performance measure	
	AUC	Truth rate
Model 1 (PLS)	0.744	$\frac{50}{70} = 0.71$
Model 2 (derivative of single wavelength)	0.740	$\frac{52}{70} = 0.74$
Model 3 (NDI)	0.799	$\frac{49}{70} = 0.70$

Figure 4 shows the density plots of the predicted probability for the non- infected plants and for the infected plants, for model 2. The three models show the same trend, while the prediction of the non- infected plants is consistently better than infected plant prediction. Another trend that appears in all the models are the 2 peaks in the distributions, on both sides of the threshold. This behavior should be examined in future work.

For future work, more samples should be collected over the course of the growing stages to provide a better understanding of the spectral discrimination over time. Additionally, further research should focus on developing measures to identify at earlier stages. The point sampling technique that was used in this work cannot describe the spatial variability of the leaves spectral reflectance. Hyper-spectral images (HSI) can be useful to capture the complementary information from both domains - spatial domain and spectral domain.

Undergoing research is aimed at utilizing a hyperspectral imaging system in order to classify vegetation can add an important data source in order to achieve these goals successfully.

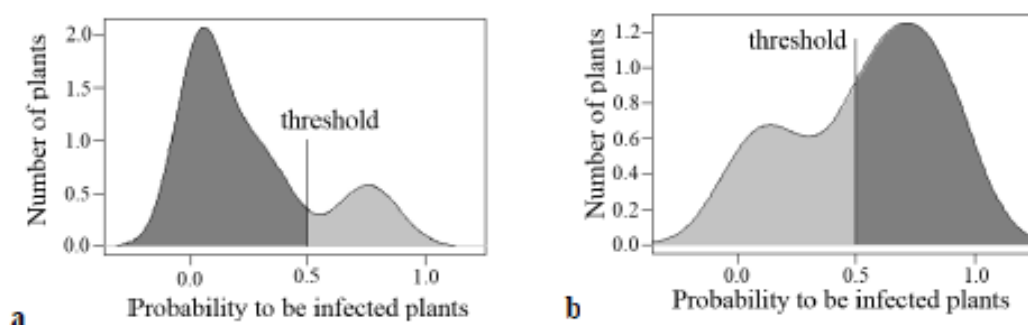


Figure 4: (a) Density plot distributions of the predicted values of the non- infected plants for model 2. The plants with predicted probability below 0.5 were correctly classified. (b) Density plot distributions of the predicted values of the infected plants for model 2.

## Conclusion

To the best of the authors' knowledge, early non-destructive detection of Fusarium infection is still a big challenge that might have a large impact on the future agriculture economy. Therefore, achieving a classification accuracy of 74% of infected plants by applying spectral analysis is a step towards early detection of the disease.

The evaluated wavelength range 350- 2500 nm reliably detected corn plants infected with Fusarium. The Lasso classifier model was applied to identify distinct wavelengths that allow differentiation between spectra of infected and non-infected corn leaves: 715, 755, 920, 950, 975, 1160, 1170 and 1280 nm.

## Acknowledgments

This research was partially supported by the Israeli Innovation Authority, fund number 580653293 and by Rabbi W. Gunther Plaut Chair in Manufacturing Engineering at Ben-Gurion University of the Negev. We would like to thank the technical team from Evogene (Noa Guedj, Aviv Himmel, Tomer Mahler) for their help in running the experiments.

## References

- Atzberger, C., Guérif, M., Baret, F. & Werner, W. (2010). Comparative analysis of three chemometric techniques for the spectroradiometric assessment of canopy chlorophyll content in winter wheat. *Computers and Electronics in Agriculture*, 73(2), 165-173.
- Bauriegel, E., Giebel, A., Geyer, M., Schmidt, U., Herppich, W.B. (2011). Early Detection of Fusarium Infection in Wheat Using Hyper-Spectral Imaging. *Computers and Electronics in Agriculture* 75(2): 304–12.
- Blackburn, G. A. (2007). Hyperspectral remote sensing of plant pigments. *Journal of Experimental Botany*, 58(4), 855–867.



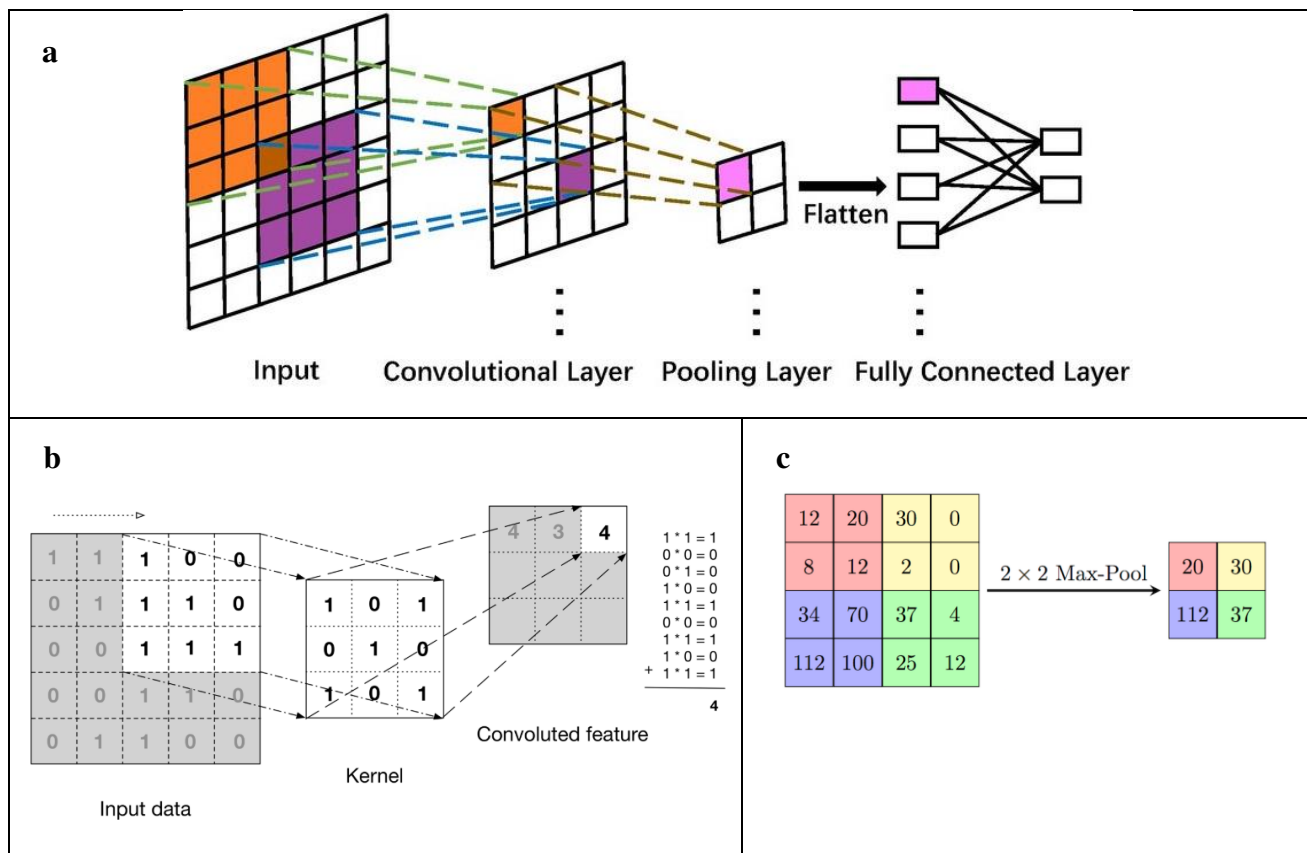
- Browne, R. A., Murphy, J. P., Cooke, B., Devaney, D., Walsh, E.J., Griffey, C.A., et al. (2005). Evaluation of components of *Fusarium* head blight resistance in soft red winter wheat germ plasm using a detached leaf assay. *Plant Disease*, 89(4), 404–411.
- Busemeyer, L., Mentrup, D., Möller, K., Wunder, E., Alheit, K., Hahn, V., et al. (2013). Breedvision - A multi-sensor platform for non-destructive field-based phenotyping in plant breeding. *Sensors (Switzerland)*, 13(3), 2830–2847.
- Falasconi, M., Gobbi, E., Pardo, M., Della Torre, M., Bresciani, A., & Sberveglieri, G. (2005). Detection of toxigenic strains of *Fusarium verticillioides* in corn by electronic olfactory system. *Sensors and Actuators, B: Chemical*, 108, 250–257.
- Hansen, P. M., & Schjoerring, J. K. (2003). Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sensing of Environment*.
- Jurado, M., Vázquez, C., Marín, S., Sanchis, V., & Teresa González-Jaén, M. (2006). PCR-based strategy to detect contamination with mycotoxigenic *Fusarium* species in maize. *Systematic and Applied Microbiology*.
- Lancashire, P., Bleiholder, H., Van Den Boom, T., Langeluddeke, P., Stauss, R., Weber, E., et al. (1991). A Uniform Decimal Code for Growth Stages of Crops and Weeds. *Annals of Applied Biology* 119(3): 561–601.
- Li, Lei, Qin Zhang, & Danfeng Huang. (2014). A Review of Imaging Techniques for Plant Phenotyping. *Sensors (Switzerland)* 14(11): 20078–111.
- Ray, M., Ray, A., Dash, S., Mishra, A., Achary, K. G., Nayak, S., & Singh, S. (2017). Fungal disease detection in plants: Traditional assays, novel diagnostic techniques and biosensors. *Biosensors and Bioelectronics*, 87(August 2016), 708–723.
- Rouse, R. W. H., Haas, J. A. W., Schell, J. A., & Deering, D. W. (1974). Monitoring Vegetation Systems in the Great Plains With ERTS. *Proceedings Third ERTS-1 Symposium*, 351.
- Savary, S., Ficke, A., Aubertot, J. N., & Hollier, C. (2012). Crop losses due to diseases and their implications for global food production losses and food security. *Food Security*, 4(4), 519–537.
- Savitzky, A., & Marcel, J. E. G. (1964). Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, 36(8), 1627–1639.
- Tibshirani, R. (1996). Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267–288.
- Tsenkova, R., Meilina, H., Kuroki, S., & Burns, D. H. (2009). Near infrared spectroscopy using short wavelengths and leave-one-cow-out cross-validation for quantification of somatic cells in milk. *Journal of Near Infrared Spectroscopy*, 17(6), 345–351.
- West, J. S., Bravo, C., Oberti, R., Lemaire, D., Moshou, D., & McCartney, H. A. (2003). The potential of optical canopy measurement for targeted control of field crop diseases. *Annual Review of Phytopathology*, 41(1), 593–614.
- Windels, C. E. (2000). Economic and Social Impacts of *Fusarium* Head Blight: Changing Farms and Rural Communities in the Northern Great Plains. *Phytopathology*, 90(1), 17–21.
- Zhangyan, J., Alfredo, R., Jin, C., Yunhao, C., Jing, L., Guangjian, Y., et al. (2006). Analysis of NDVI and Scaled Difference Vegetation Index Retrievals of Vegetation Fraction. *Remote Sensing of Environment* 101(3): 366–78.

## 8.3 CONVOLUTIONAL NEURAL NETWORKS (CNN)

### 8.3.1 LITERATURE REVIEW:

In the last years, convolutional neural networks (CNN) have become one of the most promising methods for both general image classification tasks (He et al. 2016). CNN are a special type of neural networks that present a series of convolutional layers especially designed to cope with inputs in the form of multidimensional arrays (image patches) (Lecun et al., 2015).

A CNN consists of an input and an output layer, as well as multiple hidden layers. The hidden layers of a CNN typically consist of a series of convolutional layers, pooling layers and fully connected layers.



**Figure 39: Convolutional neural networks layers**  
**(a) Neuronal network structure (b) Convolutional layer (c) Max-Pool layer**

#### 8.3.1.1 Convolutional Layers

The convolutional layers serve as feature extractors, they learn the feature representations of their input images. The neurons in the convolutional layers are arranged into feature maps. Each neuron in a feature map has a receptive field, which is connected to a neighborhood of neurons in the previous layer via a set of trainable weights (Lecun et al., 2015). Inputs are convolved with the learned weights in order to compute a new feature map, and the convolved results are sent through a nonlinear activation function. All neurons within a feature map have weights that are constrained to be equal; however, different feature maps within the same convolutional layer have different

weights so that several features can be extracted at each location (Lecun et al., 1998). Nonlinear activation functions allow for the extraction of nonlinear features.

The standard way to model a neuron's output  $f(x)$  as a function of its input  $x$  is with:

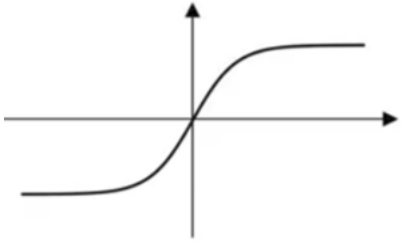
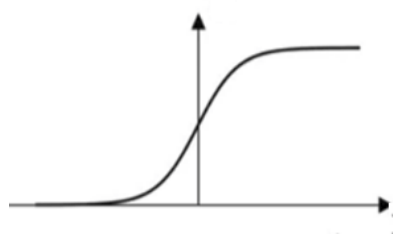
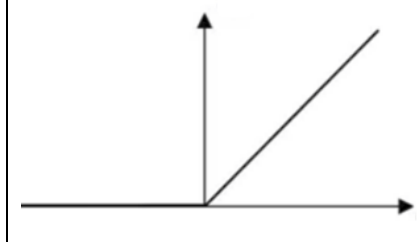
<p>Tanh:</p> $f(x) = \tanh(x)$	<p>Sigmoid:</p> $f(x) = (1 + e^{-x})^{-1}$	<p>ReLU:</p> $f(x) = \max(0, x)$
		

Figure 40: Activation functions in convolutional layer

Traditionally, the sigmoid and hyperbolic tangent functions were used; recently, **rectified linear units (ReLU) have become the most used** (Nair & Hinton, 2010). Deep convolutional neural networks with ReLUs train several times faster than their equivalents with tanh units (Krizhevsky, Sutskever, & Hinton, 2012).

#### 8.3.1.2 Pooling Layers

Pooling layers in CNNs summarize the outputs of neighboring groups of neurons in the same kernel map. The purpose of the pooling layers is to reduce the spatial resolution of the feature maps and thus achieve spatial invariance to input distortions and translations (LeCun et al., 1989a; 1989b). Initially, it was common practice to use average pooling aggregation layers to propagate the average of all the input values, of a small neighborhood of an image to the next layer (LeCun et al., 1989a; 1989b). However, in more recent models (Szegedy et al., 2015), **max pooling** aggregation layers propagate the maximum value within a receptive field to the next layer (Ranzato et al., 2007).

#### 8.3.1.3 Fully Connected Layers

Several convolutional and pooling layers are usually stacked on top of each other to extract more abstract feature representations in moving through the network. The fully connected layers that follow these layers interpret these feature representations and perform the function of high-level reasoning (Simonyan & Zisserman, 2015).

### 8.3.2 METHODS

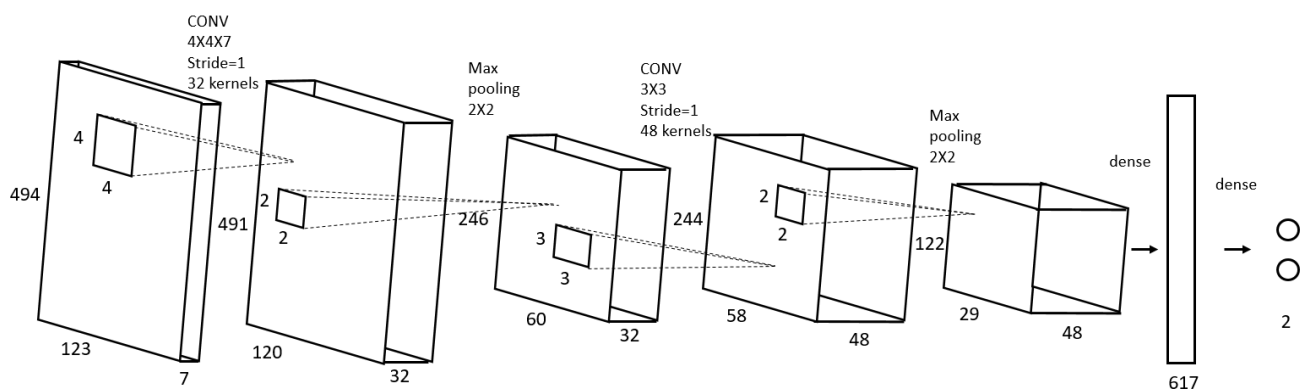
Data analyses were performed on the data from the two experiment from the second sampling day, 21 days after seeding, stage V2, closest to the day when the manual classification of disease level was performed.

Convolutional neural network to disease detection.

The input for the convolutional neural network were the 7 bands (with high SNR value) selected from the data analysis in the ASD. All images were padded with zeros so that they were all at the maximum width and maximum height from the images. All images had the same spatial and spectral size (width- 368, height- 1480, bands- 7). The 7 bands selected were 7, 64, 332, 429, 497, 592, 700 corresponding to the selected wavelengths: 400, 440, 630, 700, 750, 820, 900nm.

A number of different networks with different structures were examined to best fit the small data set. The model is selected according to the accuracy of the validation set.

### 8.3.2.1 The Architecture



**Figure 41: An illustration of the architecture of the CNN**

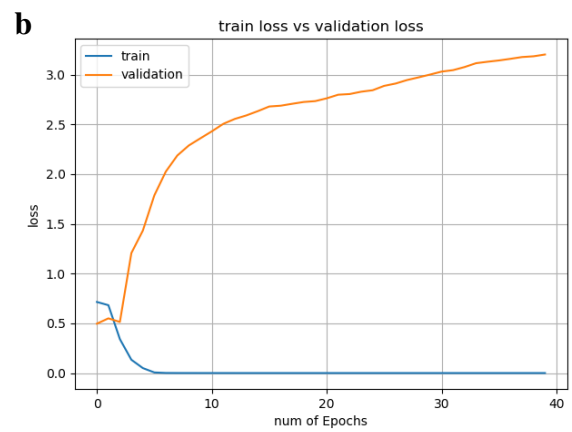
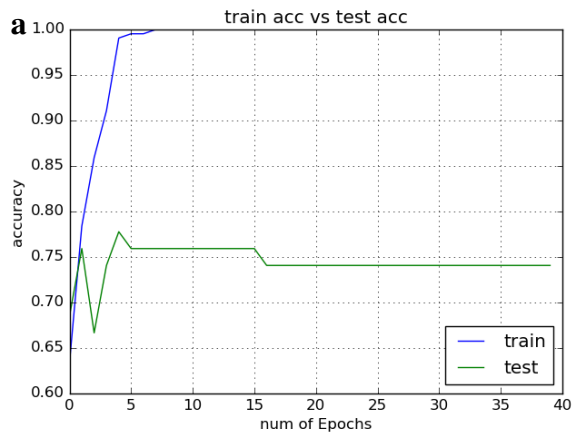
The input to the CNN is an multi spectral image of size  $123 \times 292 \times 7$  with “resize” functions from “cv2” library in order to resize the images in space. All images in all the data sets- training, validation and test were converted to this size.

The first convolutional layer filters the  $123 \times 292 \times 7$  input image with 32 kernels of size  $4 \times 4 \times 7$  with a stride of 1 pixels (this is the distance between the receptive field centers of neighboring neurons in a kernel map). The second convolutional layer takes as input the (response-normalized and max pooled) output of the first convolutional layer and filters it with 48 kernels of size  $3 \times 3 \times 32$ . The dense layer 617 neurons. In the last layer there are 2 neurons – ‘healthy’ or infected.

The chosen parameters for the training model were:

Epocs=40, steps\_per\_epoch=213 size of train set, batch\_size=1 image size.

### 8.3.3 RESULTS



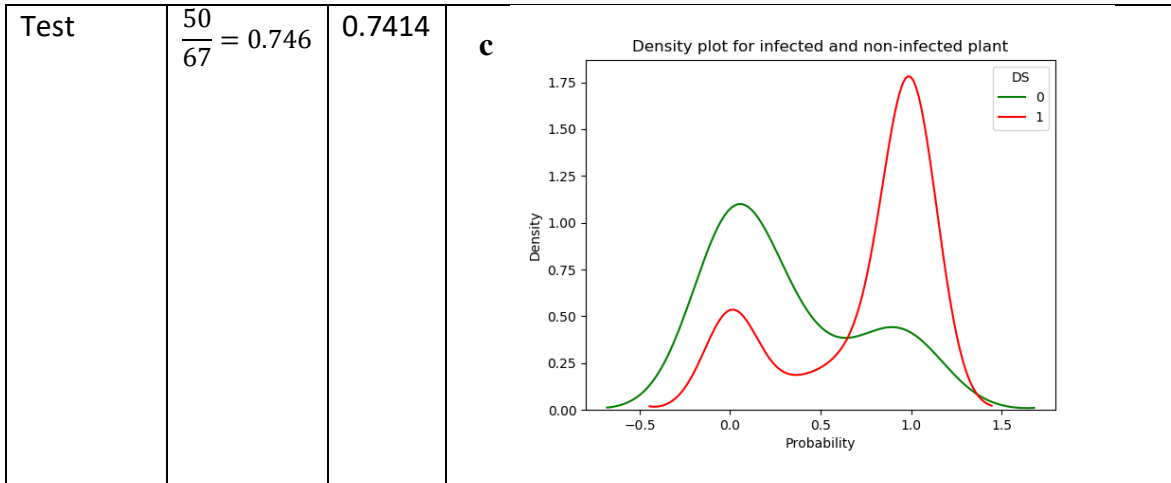
**Figure 42: The training and validation**

**(a) accuracy curve during the training of the CNN (b) loss curve during the training of the CNN**

Results reveal that there is overfitting. The model is significantly better in classifying the observations on the train set than the observations in the validation set. The model trains a large number of parameters, and there is not enough data in order to train them.

	Accuracy	AUC	Density plot
Train	$\frac{213}{213} = 1$	1	<b>a</b>
Validation	$\frac{41}{54} = 0.759$	0.75	<b>b</b>





**Figure 43: Density plot distributions of the predicted values of the infected and the healthy plants of CNN model (a) train set (b) validation set (c) test set**

**Table 20: Classification results of CNN model for train, validation and test set**

Results reveal that there is a difference between the density of the ‘infected’ and ‘healthy’ plant. In addition, it can be seen that in all sets the probability predicted in the CNN model is centered around probability 0 and probability 1.

In addition, the result shows the same trend, implying overfitting.

	Healthy	Infected
Predicted as ‘healthy’	<b>TP</b> – 16	<b>FP</b> – 11
Predicted as ‘infected’	<b>FN</b> – 6	<b>TN</b> – 34
	$Sensitivity = \frac{TP}{TP + FN} = \frac{16}{22} = 0.727$	$specificity = \frac{TN}{TN + FP} = \frac{34}{45} = 0.755$

**Table 21: Confusion matrix of CNN model, test set**

Result reveal that the prediction of the infected plants is better than ‘healthy’ plant prediction. The bias may be existing because the imbalanced dataset (there are 208 images of infected plants, compared to 126 healthy plants). The model trained on 133 images of infected plants and only 80 images of ‘healthy’ ones (images in train set).

#### 8.3.4 SENSITIVITY ANALYSIS

In this section the sensitivity of the CNN model was tested.

##### 8.3.4.1 other bandwidth

Result reveal that the model is not sensitive to the bandwidth of the images. In addition, it can be seen that the smoothing does not significantly change the average plant spectral signature. In addition, the model has a large number of weights to learn. **It might be that the reason that the different bandwidth given the similar results is due to lack in amount of images.**

bandwidth	Number of images	Accuracy	AUC
Hyperspectral	1	$\frac{50}{67} = 0.746$	0.7414
10nm	15	$\frac{50}{67} = 0.746$	0.7414
30nm	41	$\frac{50}{67} = 0.746$	0.7414
50nm	67	$\frac{50}{67} = 0.746$	0.7414
150nm	211	$\frac{51}{67} = 0.761$	0.7641

**Table 22: Classification results of CNN model for different multispectral bands**

### 8.3.4.2 other grothe stage

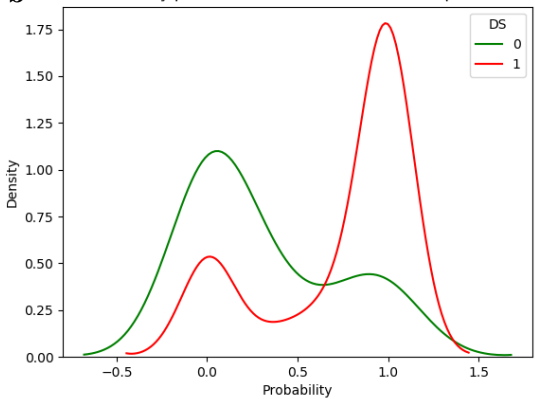
In the previous sections, the models used only the data that were collected 21 days after sowing (stage V2), the day before the disease level was classified for train, validation and test sets.

	Healthy	Infected
Predicted as healthy	<b>TP</b> – 7	<b>FP</b> – 5
Predicted as infected	<b>FN</b> – 55	<b>TN</b> – 93
	$sensiyivity = \frac{TP}{TP + FN} = \frac{7}{62} = 0.113$	$specificity = \frac{TN}{TN + FP} = \frac{93}{98} = 0.949$

**Table 23: Confusion matrix of CNN model, V1 growth stage plants**

The model classified most plants from growth stage V1, both infected and healthy , as infected ( $\frac{55+93}{62+98} = \frac{148}{160} = 0.925$ ). Compared to 12 plants classified as healthy (0.075).

	Accuracy	AUC	Density plot
Test set (stage V1)	$\frac{100}{160} = 0.625$	0.5309	<p><b>a</b></p>

Test set (stage V2)	$\frac{50}{67} = 0.746$	0.7414	<b>b</b> 
------------------------	-------------------------	--------	---

**Figure 44: Density plot distributions of the predicted values of the infected and the healthy plants of CNN model (a) samples collected 14 days after sowing (b) samples collected 21 days after sowing**

**Table 24: Classification results of CNN model for plant on different growth stage**

Results reveal that predictions of diseases in growth stage V2 is consistently better by

$$\frac{0.746}{0.625} * 100 = 119.4\%$$

than the prediction of the plant on test set in growth stage V1.

## 8.4 FEATURE SELECTION VIS/NIR RANGE

	SR Data				R1D Data			
wl	T- test	Anova	RF	PLS	T- test	Anova	RF	PLS
X350	1.214639	0.455834	0.01568	0.129494	1.404316	3.302868	0.008929	-0.00588
X355	1.48658	1.165456	0.02914	0.114706	2.028443	7.94577	0.015574	-0.1084
X360	1.793315	2.2155	0.030521	-0.06724	1.020193	2.251775	0.009104	-0.05966
X365	1.907073	2.488658	0.040685	-0.1052	1.24316	2.030829	0.009724	-0.00144
X370	1.670149	1.481783	0.015801	0.00397	2.041145	3.841308	0.009931	0.042449
X375	1.565832	1.217072	0.003382	-0.0024	1.599369	1.059946	0.006845	0.017229
X380	1.866435	2.382579	0.033535	-0.22549	0.762353	0.000778	0.004007	-0.00452
X385	1.762411	2.570225	0.055348	-0.23311	2.265301	3.584447	0.001911	0.078734
X390	1.495711	1.967526	0.064646	-0.2574	1.988102	3.353248	0.007157	0.063978
X395	0.420348	0.029594	0.030219	0.263558	2.261954	5.630637	0.013156	0.097872
X400	0.571791	0.68568	0.037825	0.639938	1.704956	2.778357	0.009123	0.057739
X405	0.456314	0.438018	0.010988	0.544963	0.95521	1.918978	0.007215	-0.06427
X410	0.100033	0.054359	0.008479	0.332447	2.528121	9.619236	0.018028	-0.13105
X415	0.111312	0.002174	0.005938	0.198373	1.93562	6.648879	0.009312	-0.08708
X420	0.354568	0.118708	0.001051	0.06201	2.30379	9.055153	0.016239	-0.09348
X425	0.550512	0.337938	0.001039	-0.03824	2.171927	8.347939	0.009177	-0.08176
X430	0.694405	0.567141	0.005959	-0.10694	1.970153	6.919245	0.020665	-0.06126
X435	0.747298	0.678421	0	-0.12899	1.141012	2.535156	0.008366	-0.03275
X440	0.783327	0.743522	0.005991	-0.14521	0.961923	1.415234	0.010852	-0.07456
X445	0.790823	0.769278	0.008718	-0.15924	0.198707	0.0005	0.002016	-0.03243
X450	0.781049	0.76198	0.008832	-0.16748	0.505965	0.088961	0.004659	-0.03162
X455	0.71331	0.664789	0.017994	-0.1566	0.641568	0.185353	0.002162	-0.04346
X460	0.709972	0.667653	0	-0.16771	0.355968	0.259271	0.002243	-0.10254
X465	0.726749	0.699529	0.001952	-0.18598	0.006672	0.020163	0.004913	-0.06998
X470	0.705496	0.669014	0.004257	-0.18419	0.366258	0.098683	0.002871	-0.02835
X475	0.725813	0.700505	0.014584	-0.19009	0.367354	0.090494	0.006816	-0.0109
X480	0.720258	0.687295	0.006148	-0.17773	0.612368	0.447576	0.008092	0.021972
X485	0.683128	0.623907	0.005933	-0.15958	1.086066	0.997535	0.002394	0.003916
X490	0.640944	0.565282	0.009	-0.15654	1.181918	1.102684	0.003295	0.004541
X495	0.542686	0.434779	0.007457	-0.14826	1.069117	0.921946	0.002173	0.006705
X500	0.405849	0.276525	0.009924	-0.13239	0.897587	0.689067	0.003223	0.015187
X505	0.257673	0.140978	0.006675	-0.11236	0.810793	0.56244	0.004109	0.023233
X510	0.086699	0.041112	0.006162	-0.07421	0.479018	0.178938	0.002731	0.020016
X515	0.01748	0.015763	0	-0.05126	0.129475	0.007253	0.000992	0.018225
X520	0.019762	0.01332	0.019185	-0.02819	0.258096	0.08726	0.000788	0.008714
X525	0.105331	0.035619	0.006184	-0.03113	0.544667	0.33324	0.000945	0.000344
X530	0.17753	0.062978	0.013181	-0.02936	0.656348	0.460024	0.004273	-0.00103
X535	0.22314	0.084103	0.013956	-0.02867	0.654013	0.422837	0.001265	0.005266
X540	0.241659	0.092066	0.002502	-0.02562	0.434959	0.163876	0.00567	0.032435
X545	0.236596	0.089369	0.008008	-0.01757	0.335461	0.047914	0.011189	0.041304
X550	0.251876	0.091406	0.004907	-0.01115	0.363342	0.019785	0.011854	0.039329
X555	0.23855	0.07842	0.006504	0.00346	1.689244	5.524424	0.005929	0.09351

X560	0.200847	0.05445	0.002771	0.017653	1.191323	1.99133	0.009817	0.034746
X565	0.137358	0.028265	0	0.028079	1.119394	1.507193	0.002639	0.029396
X570	0.037559	0.004402	0.002601	0.047422	1.146186	1.654938	0.006496	0.032525
X575	0.063168	0.001468	0.007226	0.067918	0.877863	0.957505	0.000549	0.013525
X580	0.118404	0.00901	0.002807	0.07197	0.522127	0.362247	0.002648	-0.00844
X585	0.151592	0.016925	0.002324	0.071233	0.390526	0.346448	0.002017	0.004285
X590	0.167022	0.024981	0.003413	0.078845	0.156236	0.129737	0.001846	-0.00123
X595	0.175441	0.0271	0.000227	0.075483	0.280248	0.02358	0.006981	-0.0336
X600	0.188573	0.030262	0.021453	0.078669	0.363383	0.105897	0.004169	-0.03102
X605	0.19287	0.032481	0.007825	0.074351	0.18472	0.105239	0.003698	-0.0234
X610	0.229334	0.049894	0.007193	0.082206	0.287184	0.20311	0.003026	-0.01058
X615	0.248823	0.063755	0.007257	0.08286	0.083406	0.048383	0.00035	-0.01889
X620	0.258896	0.072219	8.58E-05	0.082335	0.068399	0.005867	0.007449	-0.02031
X625	0.266222	0.080474	0.000251	0.087304	0.123089	0.093251	0.003137	0.012605
X630	0.265154	0.084465	0.017104	0.096158	0.571751	0.243151	0.00514	-0.04029
X635	0.260295	0.078564	0.00226	0.091109	0.444127	0.562726	0.00572	-0.06123
X640	0.239199	0.055306	0.012127	0.067555	0.746625	1.106357	0.002884	-0.07273
X645	0.151342	0.014167	0.013869	0.019233	0.927257	1.194434	0.003473	-0.06732
X650	0.072143	0.001103	0.011891	-0.0104	0.882673	0.661809	0.004452	-0.05515
X655	0.022912	5.64E-06	0.013475	-0.01597	0.745655	0.225457	0.001552	-0.03035
X660	0.072292	0.004456	0.009263	-0.02192	0.991586	0.675646	0.004361	-0.03392
X665	0.256664	0.05026	0.007304	-0.03221	1.320498	1.404372	0.004761	-0.03683
X670	0.459952	0.1628	0.001384	-0.03665	1.478955	1.767297	0.003867	-0.02576
X675	0.58374	0.25774	0.028176	-0.02672	1.344566	1.062033	0.011824	0.020569
X680	0.595161	0.249505	0	0.001593	1.190913	2.180402	0.006393	0.087841
X685	0.366669	0.068796	0.009323	0.040457	1.212017	2.018284	0.007538	0.07355
X690	0.090927	0.049778	0.008742	0.158496	0.715877	0.883989	0.006076	0.065481
X695	0.181189	0.120068	0.004169	0.232292	0.073531	0.01687	0.008422	0.042882
X700	0.021258	0.026153	0.008317	0.258031	0.793851	0.407607	0.004404	0.01257
X705	0.31028	0.015116	0.025654	0.23556	1.375684	1.605603	0.005433	-0.01906
X710	0.589787	0.158686	0.00831	0.196875	1.853748	3.040643	0.004217	-0.04399
X715	0.864465	0.440918	0.007641	0.150584	2.393685	4.826275	0.006999	-0.06086
X720	1.157541	0.882769	0.007888	0.095905	2.516216	5.764489	0.004374	-0.07775
X725	1.353529	1.296701	0.004776	0.053013	2.731951	6.967827	0.005088	-0.09314
X730	1.509244	1.69052	0.000798	0.017419	2.69819	7.50237	0.02259	-0.10903
X735	1.60318	1.988935	0.002606	-0.01051	2.78659	8.263208	0.013508	-0.12302
X740	1.674056	2.213881	0.000906	-0.02968	2.983264	9.352843	0.04271	-0.13156
X745	1.724777	2.380727	0	-0.04433	3.288814	10.97526	0.034064	-0.13957
X750	1.759719	2.491228	0.012376	-0.05248	3.586002	12.65693	0.012651	-0.14289
X755	1.782434	2.563141	0.013886	-0.05682	3.32578	7.767457	0.006435	-0.08123
X760	1.811086	2.602452	0.000906	-0.05101	1.25725	0.340285	0.007048	0.07859
X765	1.797902	2.562716	0.002938	-0.03585	0.715057	0.00291	0.006765	-0.07
X770	1.822277	2.690126	0.007984	-0.06164	0.449097	1.163005	0.008879	-0.14817
X775	1.829872	2.716984	0.007355	-0.06338	0.911224	0.683495	0.003623	-0.03462
X780	1.83653	2.737395	0.006534	-0.06367	1.425242	1.105888	0.010723	-0.00906
X785	1.844798	2.756993	0	-0.06323	1.881525	1.276396	0.00444	0.032652

X790	1.856021	2.776283	0.003439	-0.06173	1.23247	0.294516	0.00266	0.039249
X795	1.856786	2.77813	0	-0.05961	1.315896	0.399196	0.002534	0.014163
X800	1.864008	2.793702	0.00875	-0.05904	1.553576	0.759327	0.003568	-0.00114
X805	1.864883	2.796505	0.010753	-0.05739	1.197889	0.192213	0.002876	0.003178
X810	1.875889	2.815425	0.002724	-0.0571	2.485418	2.469365	0.002401	-0.00027
X815	1.914057	2.886697	0.001048	-0.06154	2.439736	2.653021	0.004538	0.004723
X820	1.906819	2.873503	0.002422	-0.05836	2.190381	2.150476	0.000569	0.006998
X825	1.90109	2.862409	0.006365	-0.05592	1.474982	0.955908	0.006743	-0.00607
X830	1.900043	2.862364	0	-0.05392	1.514837	0.825933	0.007412	-0.01948
X835	1.886505	2.833011	0	-0.05013	2.172912	2.119547	0.007502	-0.01662
X840	1.878435	2.820839	0	-0.04806	1.577185	0.958506	0.009653	-0.0295
X845	1.872913	2.809685	0.000157	-0.04594	0.251887	0.015163	0.011297	-0.0307
X850	1.869715	2.796566	0	-0.04188	0.159025	0.08455	0.006767	0.040827
X855	1.868158	2.792186	0.005938	-0.03922	0.581983	0.13171	0.008532	-0.00238
X860	1.862983	2.772891	0.005516	-0.03421	0.095999	0.216074	0.0248	0.04245
X865	1.856437	2.74641	0.003191	-0.02786	0.385362	0.021788	0.010914	0.026426
X870	1.855119	2.737535	7.15E-05	-0.02548	0.823433	4.95E-06	0.005853	-0.0173
X875	1.853684	2.725281	0.000182	-0.0233	0.455631	0.066514	0.004701	-0.01715
X880	1.855272	2.724361	0.002412	-0.02262	1.215158	0.442386	0.008659	-0.04367
X885	1.860959	2.74094	0	-0.0244	1.604614	0.532927	0.007145	0.000136
X890	1.860693	2.720501	0.000699	-0.01921	1.823153	0.559657	0.009887	0.036189
X895	1.879051	2.732262	0	-0.01576	2.097599	1.295397	0.002977	0.024756
X900	1.895184	2.759089	0.004613	-0.01518	1.875407	1.282008	0.007002	-0.00024
X905	1.890111	2.740647	0	-0.01154	2.026629	1.292676	0.002413	-0.0025
X910	1.908782	2.780369	0.000165	-0.01354	2.238057	2.579169	0.004624	-0.01664
X915	1.907661	2.768762	0.00421	-0.01208	1.976764	2.98993	0.00618	-0.01135
X920	1.89752	2.746751	0.003564	-0.01048	0.20043	0.001008	0.006805	-0.08726
X925	1.91811	2.801974	0	-0.01686	2.23707	1.906357	0.005927	-0.01971
X930	1.975806	2.91029	0	-0.02919	1.749533	1.052802	0.003402	-0.00616
X935	2.040689	3.034862	9.69E-05	-0.04244	1.237674	0.254313	0.011467	0.03418
X940	2.005154	2.911263	0.000979	-0.02632	2.443831	7.86466	0.020505	0.126623
X945	2.006352	2.917166	0.004411	-0.02804	2.587923	8.807992	0.032914	0.127934
X950	1.985187	2.84716	0	-0.01476	4.042657	14.21182	0.030359	0.141896
X955	1.961854	2.795748	0.005636	-0.00688	2.259052	2.909926	0.004522	0.042501
X960	1.917728	2.683304	0.00301	0.010545	3.050458	6.758889	0.012071	0.073937
X965	1.857263	2.523398	0.001763	0.034112	2.070064	1.790593	0.005824	0.010642
X970	1.848072	2.532303	0.004952	0.03007	2.918671	6.049407	0.002069	0.057886
X975	1.823243	2.45712	0	0.042498	1.920268	2.207732	0.014575	0.005438
X980	1.816126	2.448419	0.000812	0.044384	1.363941	0.798183	0.004722	-0.03213
X985	1.800709	2.417918	0.003121	0.046973	2.070259	2.178553	0.005339	-0.00471
X990	1.780311	2.363233	0.001588	0.054862	2.55104	4.202555	0.010243	0.032634
X995	1.776913	2.364552	0	0.055569	1.918713	4.151435	0.0118	0.071131
X1000	1.781067	2.36873	0.005724	0.055635	3.538966	13.05229	0.020094	0.191501

## 8.5 FEATURE SELECTION VIS/NIR/SWIR RANGE

	SR Data				R1D Data			
wl	T- test	Anova	RF	PLS	T- test	Anova	RF	PLS
X350	1.214639	0.455834	0.001124	-0.01528	1.404316	3.302868	0.000114	-0.00023
X355	1.48658	1.165456	0.000894	0.042607	2.028443	7.94577	0.00258	-0.01952
X360	1.793315	2.2155	0.006162	0.035075	1.020193	2.251775	0.003331	-0.02753
X365	1.907073	2.488658	0.008755	0.016114	1.24316	2.030829	0.000764	-0.02461
X370	1.670149	1.481783	0.00424	0.006357	2.041145	3.841308	0	-0.00348
X375	1.565832	1.217072	0.002507	0.025299	1.599369	1.059946	0.006055	-0.01892
X380	1.866435	2.382579	6.51E-05	0.007186	0.762353	0.000778	0.000774	-0.01867
X385	1.762411	2.570225	0.002794	0.027519	2.265301	3.584447	0.002721	0.001349
X390	1.495711	1.967526	0.004083	0.044819	1.988102	3.353248	0.000171	0.008986
X395	0.420348	0.029594	0.004295	0.111251	2.261954	5.630637	0	0.023202
X400	0.571791	0.68568	0.000748	0.153114	1.704956	2.778357	0.001859	0.006904
X405	0.456314	0.438018	0.004627	0.135984	0.95521	1.918978	0.010468	-0.03958
X410	0.100033	0.054359	0.001475	0.093596	2.528121	9.619236	0.000537	-0.05982
X415	0.111312	0.002174	0	0.055021	1.93562	6.648879	0	-0.05151
X420	0.354568	0.118708	0	0.022573	2.30379	9.055153	7.32E-05	-0.05242
X425	0.550512	0.337938	0.002973	-0.00159	2.171927	8.347939	0.008899	-0.04996
X430	0.694405	0.567141	0.003668	-0.01782	1.970153	6.919245	0.004937	-0.04008
X435	0.747298	0.678421	0.002742	-0.02631	1.141012	2.535156	0.002665	-0.04285
X440	0.783327	0.743522	0.002262	-0.03352	0.961923	1.415234	0.002006	-0.06566
X445	0.790823	0.769278	0.001865	-0.03784	0.198707	0.0005	0.000161	-0.0307
X450	0.781049	0.76198	0.001209	-0.03923	0.505965	0.088961	0.001766	-0.02548
X455	0.71331	0.664789	0.000343	-0.03998	0.641568	0.185353	0.000266	-0.0313
X460	0.709972	0.667653	0.003439	-0.04155	0.355968	0.259271	0.000314	-0.0517
X465	0.726749	0.699529	0.003058	-0.04421	0.006672	0.020163	0.003403	-0.04686
X470	0.705496	0.669014	0.00089	-0.04664	0.366258	0.098683	0.003045	-0.03673
X475	0.725813	0.700505	0.000424	-0.04769	0.367354	0.090494	0	-0.01621
X480	0.720258	0.687295	0	-0.04639	0.612368	0.447576	0.000285	-0.00333
X485	0.683128	0.623907	0.00037	-0.0439	1.086066	0.997535	0.002122	-0.0021
X490	0.640944	0.565282	0.000612	-0.04106	1.181918	1.102684	0	0.004452
X495	0.542686	0.434779	0.000276	-0.03421	1.069117	0.921946	9.03E-05	0.01081
X500	0.405849	0.276525	0.000178	-0.02252	0.897587	0.689067	0	0.015667
X505	0.257673	0.140978	0.000604	-0.00683	0.810793	0.56244	0.000367	0.018002
X510	0.086699	0.041112	0.00318	0.00956	0.479018	0.178938	4.81E-05	0.017106
X515	0.01748	0.015763	0.000281	0.02251	0.129475	0.007253	0	0.015198
X520	0.019762	0.01332	8.89E-05	0.02944	0.258096	0.08726	0.000444	0.009694
X525	0.105331	0.035619	9.74E-05	0.028738	0.544667	0.33324	0	0.004545
X530	0.17753	0.062978	0.001161	0.026107	0.656348	0.460024	0	0.001933
X535	0.22314	0.084103	0.002905	0.023696	0.654013	0.422837	0	0.003024
X540	0.241659	0.092066	0	0.022246	0.434959	0.163876	0.000362	0.01752
X545	0.236596	0.089369	0.005118	0.023245	0.335461	0.047914	0.000172	0.019815
X550	0.251876	0.091406	0	0.021518	0.363342	0.019785	6.08E-05	-0.00595
X555	0.23855	0.07842	0	0.022324	1.689244	5.524424	0	0.017335

X560	0.200847	0.05445	0	0.023182	1.191323	1.99133	0	0.003891
X565	0.137358	0.028265	0.000197	0.023788	1.119394	1.507193	0.00014	0.004484
X570	0.037559	0.004402	0.000738	0.02577	1.146186	1.654938	0.00155	0.002953
X575	0.063168	0.001468	0	0.027532	0.877863	0.957505	0	-0.0029
X580	0.118404	0.00901	0.000505	0.026615	0.522127	0.362247	0.000178	-0.01175
X585	0.151592	0.016925	0.001939	0.024904	0.390526	0.346448	0.000398	-0.00271
X590	0.167022	0.024981	0.000246	0.025353	0.156236	0.129737	0.000216	-0.00065
X595	0.175441	0.0271	0.000305	0.024664	0.280248	0.02358	0.00118	-0.01409
X600	0.188573	0.030262	0.000103	0.024787	0.363383	0.105897	0	-0.02077
X605	0.19287	0.032481	0.002559	0.023085	0.18472	0.105239	0.001153	-0.0203
X610	0.229334	0.049894	0.001295	0.020815	0.287184	0.20311	0	-0.01878
X615	0.248823	0.063755	0.000792	0.017378	0.083406	0.048383	0	-0.01985
X620	0.258896	0.072219	0.001896	0.014642	0.068399	0.005867	0	-0.0155
X625	0.266222	0.080474	0.001089	0.013491	0.123089	0.093251	0	0.000665
X630	0.265154	0.084465	8.06E-05	0.013971	0.571751	0.243151	0	-0.00996
X635	0.260295	0.078564	0.000321	0.012579	0.444127	0.562726	0.004023	-0.02499
X640	0.239199	0.055306	0.000204	0.006533	0.746625	1.106357	0.003946	-0.02941
X645	0.151342	0.014167	0.00016	-0.00415	0.927257	1.194434	0.003416	-0.0305
X650	0.072143	0.001103	0.000221	-0.01583	0.882673	0.661809	0.003652	-0.03195
X655	0.022912	5.64E-06	0.000617	-0.02299	0.745655	0.225457	8.97E-05	-0.02676
X660	0.072292	0.004456	0.000537	-0.03376	0.991586	0.675646	0.000708	-0.02374
X665	0.256664	0.05026	0.003094	-0.04778	1.320498	1.404372	0	-0.0179
X670	0.459952	0.1628	0.002443	-0.05655	1.478955	1.767297	0.002448	-0.00673
X675	0.58374	0.25774	0.000493	-0.05727	1.344566	1.062033	0	0.024851
X680	0.595161	0.249505	0.001109	-0.05076	1.190913	2.180402	0.00284	0.047849
X685	0.366669	0.068796	4.88E-05	-0.03379	1.212017	2.018284	0.002089	0.036188
X690	0.090927	0.049778	0.000744	0.006654	0.715877	0.883989	6.84E-05	0.035528
X695	0.181189	0.120068	0.002899	0.03975	0.073531	0.01687	0	0.02655
X700	0.021258	0.026153	0.003265	0.049145	0.793851	0.407607	0.000135	0.012844
X705	0.31028	0.015116	0.002554	0.045085	1.375684	1.605603	0.000468	0.000112
X710	0.589787	0.158686	0	0.036525	1.853748	3.040643	0.001492	-0.00992
X715	0.864465	0.440918	0	0.025314	2.393685	4.826275	0.000386	-0.01644
X720	1.157541	0.882769	7.56E-05	0.01166	2.516216	5.764489	0	-0.02353
X725	1.353529	1.296701	0.002806	0.001531	2.731951	6.967827	0.010204	-0.02941
X730	1.509244	1.69052	0.002549	-0.00652	2.69819	7.50237	0	-0.03565
X735	1.60318	1.988935	0	-0.01147	2.78659	8.263208	0.015549	-0.04056
X740	1.674056	2.213881	0.002475	-0.01542	2.983264	9.352843	0.008214	-0.04426
X745	1.724777	2.380727	0	-0.01832	3.288814	10.97526	0.003948	-0.04843
X750	1.759719	2.491228	0.00291	-0.02044	3.586002	12.65693	0.002079	-0.05047
X755	1.782434	2.563141	0	-0.02183	3.32578	7.767457	0	-0.03176
X760	1.811086	2.602452	0.000163	-0.02395	1.25725	0.340285	0.000737	0.016737
X765	1.797902	2.562716	0.000251	-0.02262	0.715057	0.00291	0.005809	-0.01244
X770	1.822277	2.690126	0	-0.02369	0.449097	1.163005	0.004525	-0.04014
X775	1.829872	2.716984	0	-0.024	0.911224	0.683495	0.000153	-0.00551
X780	1.83653	2.737395	0.000792	-0.02414	1.425242	1.105888	0.000316	0.00942
X785	1.844798	2.756993	8.67E-05	-0.02432	1.881525	1.276396	0.001643	0.019477



X790	1.856021	2.776283	0	-0.02475	1.23247	0.294516	0	0.025883
X795	1.856786	2.77813	0.000745	-0.02452	1.315896	0.399196	0.004556	0.020866
X800	1.864008	2.793702	0.000558	-0.02472	1.553576	0.759327	0.001329	0.017293
X805	1.864883	2.796505	0.000173	-0.02452	1.197889	0.192213	9.46E-05	0.019584
X810	1.875889	2.815425	0.000227	-0.02503	2.485418	2.469365	0.001973	0.001928
X815	1.914057	2.886697	9.74E-05	-0.02755	2.439736	2.653021	0.002977	0.00691
X820	1.906819	2.873503	0	-0.02677	2.190381	2.150476	0	0.011894
X825	1.90109	2.862409	0	-0.02633	1.474982	0.955908	0	0.013681
X830	1.900043	2.862364	0	-0.02578	1.514837	0.825933	0	0.004922
X835	1.886505	2.833011	0	-0.0249	2.172912	2.119547	0.000677	-0.00129
X840	1.878435	2.820839	0	-0.0241	1.577185	0.958506	0.000324	-0.00216
X845	1.872913	2.809685	0	-0.02352	0.251887	0.015163	0.001353	0.003733
X850	1.869715	2.796566	0	-0.02309	0.159025	0.08455	0.001006	0.025389
X855	1.868158	2.792186	0	-0.02255	0.581983	0.13171	8.09E-05	0.019289
X860	1.862983	2.772891	0	-0.02174	0.095999	0.216074	0	0.023418
X865	1.856437	2.74641	0.000253	-0.02092	0.385362	0.021788	0.000163	0.017088
X870	1.855119	2.737535	0.001504	-0.02074	0.823433	4.95E-06	0.000946	-0.01482
X875	1.853684	2.725281	0	-0.02129	0.455631	0.066514	0.00011	-0.01784
X880	1.855272	2.724361	0.00077	-0.02189	1.215158	0.442386	0.004275	-0.01344
X885	1.860959	2.74094	0.001093	-0.02235	1.604614	0.532927	0	0.011012
X890	1.860693	2.720501	0	-0.02257	1.823153	0.559657	0.004	0.010975
X895	1.879051	2.732262	0	-0.02409	2.097599	1.295397	0.001164	0.008974
X900	1.895184	2.759089	0	-0.02483	1.875407	1.282008	0.00026	0.028385
X905	1.890111	2.740647	8.40E-05	-0.0242	2.026629	1.292676	0.000411	0.017408
X910	1.908782	2.780369	0	-0.02543	2.238057	2.579169	0.000728	-0.00289
X915	1.907661	2.768762	0.000694	-0.02597	1.976764	2.98993	0.00014	-0.01385
X920	1.89752	2.746751	0.000233	-0.02566	0.20043	0.001008	0.004159	-0.02109
X925	1.91811	2.801974	0	-0.02672	2.23707	1.906357	0	-0.01155
X930	1.975806	2.91029	0	-0.03157	1.749533	1.052802	0.000867	-0.01295
X935	2.040689	3.034862	6.00E-05	-0.03877	1.237674	0.254313	0.000249	0.00587
X940	2.005154	2.911263	6.51E-05	-0.03623	2.443831	7.86466	0.006304	0.056108
X945	2.006352	2.917166	0.000175	-0.0355	2.587923	8.807992	0.015978	0.043439
X950	1.985187	2.84716	0	-0.03385	4.042657	14.21182	0.003585	0.045762
X955	1.961854	2.795748	0	-0.03177	2.259052	2.909926	0.000281	0.023245
X960	1.917728	2.683304	0	-0.02791	3.050458	6.758889	0.005748	0.022583
X965	1.857263	2.523398	7.36E-05	-0.02455	2.070064	1.790593	0.001504	0.015211
X970	1.848072	2.532303	0.000545	-0.02187	2.918671	6.049407	0.001053	0.027918
X975	1.823243	2.45712	0	-0.02094	1.920268	2.207732	0.00216	-0.00412
X980	1.816126	2.448419	8.70E-05	-0.01991	1.363941	0.798183	0.000521	-0.00078
X985	1.800709	2.417918	0	-0.01815	2.070259	2.178553	0	0.010047
X990	1.780311	2.363233	9.02E-05	-0.01708	2.55104	4.202555	0.000298	0.006208
X995	1.776913	2.364552	7.59E-05	-0.01665	1.918713	4.151435	0	0.012619
X1000	1.781067	2.36873	0	-0.01642	3.538966	13.05229	0.003565	0.051343
X1005	1.633176	1.869948	0.000133	-0.00194	3.934992	16.49739	0.010242	0.062123
X1010	1.633451	1.874454	0.000846	-0.00271	2.960292	13.45959	0.001451	-0.06237
X1015	1.648362	1.924358	0.000115	-0.00235	2.278189	4.563567	0.006371	-0.03746

X1020	1.64872	1.914029	0.000196	-0.00346	1.961364	1.976421	0.000981	-0.0271
X1025	1.654826	1.934823	0	-0.00435	1.804916	3.409487	0.000404	-0.00917
X1030	1.652242	1.919742	0.000334	-0.00378	0.684256	3.701277	0.003326	-0.00276
X1035	1.645183	1.893964	0.000155	-0.00379	2.238328	3.759576	0.004637	-0.03006
X1040	1.665675	1.948946	0	-0.0049	2.989644	8.441571	0.001191	-0.01536
X1045	1.65473	1.920403	0.001496	-0.00373	3.390393	14.82344	0.01744	0.043868
X1050	1.648881	1.895253	0	-0.00352	0.007814	0.469033	0.001408	-0.00185
X1055	1.647564	1.887801	0.000282	-0.00364	2.465364	5.646282	0.000289	-0.01032
X1060	1.662084	1.926498	0	-0.00429	3.273536	9.642568	0.001649	-0.03775
X1065	1.664416	1.928346	0	-0.0045	0.810805	3.874698	0.001594	0.026631
X1070	1.658796	1.901624	0	-0.00442	0.609037	1.838977	0.000323	-0.0134
X1075	1.651779	1.876462	0	-0.00437	2.704774	5.680701	0.003148	-0.0563
X1080	1.669683	1.930945	0	-0.00624	1.968952	1.496864	0	-0.00258
X1085	1.658924	1.881177	0.000118	-0.00451	0.164471	1.94438	0.01982	0.063788
X1090	1.665884	1.895348	7.65E-05	-0.00465	4.013006	12.90058	0.005133	-0.02472
X1095	1.692673	1.96001	0	-0.00635	2.265628	1.960161	0.000808	0.00794
X1100	1.697925	1.93988	0.000849	-0.00693	1.713723	0.617909	0	0.020125
X1105	1.714841	1.968566	0	-0.0074	2.165018	1.386821	0.001427	0.009844
X1110	1.733291	1.993719	0	-0.0086	3.346849	7.673845	0.002223	-0.01045
X1115	1.798393	2.152824	0	-0.01349	2.84707	4.91767	8.60E-05	-0.00105
X1120	1.879605	2.291891	0	-0.01937	1.458194	0.814702	0	-0.00981
X1125	1.898448	2.360207	0.001189	-0.02377	0.005187	0.006913	0.000954	-0.02634
X1130	1.859186	2.217433	0	-0.01756	1.152296	0.394459	0.003422	-0.00248
X1135	1.79672	2.10584	8.62E-05	-0.01503	4.093894	13.73156	0.002545	0.033796
X1140	1.763082	1.998476	0	-0.00788	1.30267	2.278333	0.000147	0.044409
X1145	1.783028	2.051116	9.05E-05	-0.01216	0.975643	0.382875	0.001817	-0.00257
X1150	1.760972	2.001166	0	-0.01149	3.594381	10.56021	0.008759	0.025203
X1155	1.684948	1.811196	0	-0.0071	3.169465	5.639309	0	0.018367
X1160	1.651162	1.772455	0.000178	-0.00342	2.600663	3.336525	8.41E-05	0.018454
X1165	1.608733	1.697721	0	0.000175	2.462276	2.475635	0.000549	0.023352
X1170	1.588945	1.668841	0	0.003103	2.39401	3.655031	0.001575	0.052783
X1175	1.576385	1.645108	0.001447	0.005153	0.32175	1.288558	0.00151	0.029241
X1180	1.583888	1.681348	0	0.004848	1.106881	0.176138	0	0.01494
X1185	1.576508	1.663739	0.000323	0.006217	1.03591	0.078856	0.000306	-0.01388
X1190	1.581013	1.688937	0.000106	0.005284	1.120385	0.369915	0.003461	-0.02407
X1195	1.582739	1.691898	0	0.004082	0.041006	0.033658	0	-0.01442
X1200	1.590898	1.707154	0.001616	0.001093	0.422659	0.204399	8.20E-05	-0.05008
X1205	1.594425	1.726461	0	-0.00111	0.424672	0.065779	0	-0.03951
X1210	1.598535	1.732678	0	-0.00152	1.049188	0.639847	0.000212	-0.01765
X1215	1.597695	1.733359	0	-0.0017	0.472757	0.265887	0.001093	-0.0203
X1220	1.592443	1.716062	0	-0.00179	1.01212	2.172508	0.003288	0.008151
X1225	1.578001	1.67015	0.000187	-0.00077	0.791522	1.617972	0.001032	0.006154
X1230	1.575233	1.658196	0.003562	-0.00034	0.339059	0.001011	0.0017	-0.00716
X1235	1.566124	1.633505	0.000713	0.000369	0.616246	0.060118	0.000472	0.00438
X1240	1.569443	1.638056	0	0.000376	1.108577	0.197451	0.004737	-0.02107
X1245	1.569978	1.637707	0.000389	-0.00026	1.518825	1.632475	0	-0.03154

X1250	1.568046	1.628281	0.000173	-0.00015	1.413485	0.560184	0.009763	0.015089
X1255	1.576005	1.643445	0	-0.0005	1.957074	1.630931	0.001672	0.004154
X1260	1.578847	1.63716	0	-0.00102	0.191558	0.654102	0.001845	0.025472
X1265	1.568001	1.592378	0	-0.0004	0.207925	1.94662	0	-0.01094
X1270	1.569698	1.590618	0	-0.00086	0.491091	0.817515	0	-0.06162
X1275	1.575868	1.616211	7.48E-05	-0.00106	1.790113	7.67421	0.000861	-0.04557
X1280	1.585659	1.65049	0.000364	-0.00167	1.395503	4.299919	0.000439	-0.03608
X1285	1.593936	1.672208	0	-0.00186	4.126288	10.56426	0	-0.04303
X1290	1.598335	1.672853	0.001656	-0.00221	1.743698	0.647661	0.003091	0.015877
X1295	1.606687	1.690891	0	-0.0029	2.086987	4.199487	0.002675	-0.01502
X1300	1.619145	1.716929	0	-0.00397	0.074136	0.568701	0.000829	0.020448
X1305	1.617352	1.697369	8.13E-05	-0.0042	0.221585	0.340523	0	0.012909
X1310	1.624099	1.702035	9.56E-05	-0.00432	0.88706	0.067981	0.000172	0.011169
X1315	1.628421	1.699369	0.000821	-0.00492	0.321437	0.51184	0.000319	0.001493
X1320	1.627537	1.691383	5.18E-05	-0.00629	0.276873	0.021276	0.001271	0.012594
X1325	1.644093	1.723174	0	-0.00722	0.935069	0.109402	0.000578	0.00598
X1330	1.663478	1.742716	0	-0.00984	0.938754	0.161371	0.001352	-0.0082
X1335	1.675215	1.763149	0	-0.01066	0.978228	0.984952	0	0.006229
X1340	1.665849	1.725916	0.000101	-0.01106	0.575587	0.013106	0.001405	0.036784
X1345	1.712793	1.850986	0	-0.01391	1.741342	3.116734	0	0.008543
X1415	1.496145	1.319229	0.001914	-0.01584	1.610213	4.584212	0.005504	-0.01472
X1420	1.563224	1.638187	0.000776	-0.06705	1.846105	3.396312	0.000338	0.005531
X1425	1.491551	1.180055	0.000769	-0.03301	1.087851	1.887091	0.001373	0.062376
X1430	1.32031	0.824404	0.000809	-0.02481	0.862686	0.525954	0	0.007021
X1435	1.452116	1.160427	0.002237	-0.02729	1.789426	1.350327	0.002605	0.040316
X1440	1.178059	0.672287	0.002245	-0.01587	2.398801	4.681259	0.0009	0.015153
X1445	1.24741	0.738993	0	-0.01635	0.34993	0.023347	0.000149	0.024026
X1450	1.228792	0.708683	0.00056	-0.00889	0.828206	1.267147	0.001208	0.045898
X1455	1.155829	0.582456	5.04E-05	-0.0025	3.261748	6.762893	0.003384	0.053152
X1460	1.150124	0.592568	0	0.003337	1.543535	2.312345	0.001307	0.012353
X1465	1.196163	0.658599	0	0.002683	1.779817	2.006377	9.58E-05	-0.00766
X1470	1.207199	0.662709	0	-0.00418	0.976455	0.159747	0	-0.00638
X1475	1.199402	0.6485	0	0.002928	0.02183	0.000723	0.002872	0.016698
X1480	1.146666	0.581857	0.001111	0.004936	0.787235	0.482976	0.001562	-0.00058
X1485	1.193347	0.661824	0.000107	0.004527	0.760612	0.198346	0.004888	0.018942
X1490	1.118782	0.544158	0	0.011456	0.358773	0.180537	0.001447	0.020004
X1495	1.123258	0.570271	0	0.011923	0.822306	0.655997	0.004893	0.002114
X1500	1.11289	0.557713	0	0.014399	1.207605	1.092021	0.002655	0.00781
X1505	1.122337	0.585741	0.005533	0.01551	1.392621	1.399124	0	0.002846
X1510	1.133876	0.600135	0.002198	0.015873	1.365509	1.105406	0	-0.00158
X1515	1.139431	0.61461	0.000817	0.015609	1.35894	1.464258	0	-0.00229
X1520	1.144122	0.630774	0.000121	0.015767	1.480779	1.767582	0.001848	0.002397
X1525	1.159977	0.66146	0	0.016285	1.542797	1.650052	0.000927	0.002221
X1530	1.16402	0.669335	0	0.016502	1.294627	1.012923	0.000746	0.003759
X1535	1.167409	0.67614	0	0.017222	1.448721	1.495048	0	-0.00115
X1540	1.176636	0.696882	0.005429	0.016817	1.757671	2.453972	0.001591	-0.01094

X1545	1.19443	0.732379	0	0.016043	1.916236	2.790356	0.000742	-0.00206
X1550	1.199596	0.73885	0.001726	0.016527	1.429059	1.18071	0.000237	0.007186
X1555	1.208971	0.757731	5.46E-05	0.01692	2.005161	3.143496	0.000394	-0.00525
X1560	1.225338	0.786852	0	0.016262	2.093703	3.329327	0	-0.00726
X1565	1.238887	0.812305	0.00026	0.015785	2.021144	2.920933	0.000249	0.000392
X1570	1.247308	0.828865	0	0.015292	1.738168	2.557654	0.000121	-0.01948
X1575	1.257378	0.852336	0	0.014616	2.10927	3.453674	0	-0.00954
X1580	1.264619	0.861944	0	0.015288	1.098491	0.541446	0.000822	0.009768
X1585	1.263799	0.864136	0.001417	0.015792	2.390349	6.803728	0.005165	-0.03129
X1590	1.292466	0.92853	0	0.01437	2.385963	5.675395	0.0001	-0.02411
X1595	1.288362	0.919493	7.23E-05	0.014733	1.518835	1.94372	0.000107	0.006979
X1600	1.298595	0.94398	0.000501	0.01535	0.856608	0.331954	7.28E-05	0.004224
X1605	1.290747	0.923954	7.77E-05	0.016313	1.660738	2.542888	0.000894	-0.00511
X1610	1.309405	0.972443	0	0.017372	1.151017	1.984797	0.000278	-0.00219
X1615	1.305463	0.969426	0.001894	0.019191	0.272267	0.037415	0.000478	0.002265
X1620	1.28534	0.93094	0	0.02053	0.330974	0.127789	0.000416	-0.00676
X1625	1.285001	0.930306	6.71E-05	0.021299	0.007945	0.00647	7.61E-05	-0.00228
X1630	1.266571	0.894483	0.000464	0.023952	0.019313	1.52E-05	0	9.84E-05
X1635	1.272934	0.910204	0	0.024549	1.145473	0.669051	7.35E-05	-0.01106
X1640	1.229043	0.840042	0	0.027777	2.679842	5.125338	0.000975	0.025353
X1645	1.170197	0.741595	8.87E-05	0.031694	2.752669	6.481872	0.001215	0.012916
X1650	1.118155	0.669126	7.56E-05	0.035718	1.772281	2.56882	0	-0.01176
X1655	1.065253	0.577881	0.003171	0.036124	3.197445	9.46694	0.001122	-0.00355
X1660	0.950115	0.450475	0.000117	0.037508	3.506012	11.36468	0.000801	-0.02105
X1665	0.685673	0.153097	0.000422	0.043983	4.419173	27.66274	0.015093	0.028275
X1670	0.685601	0.148963	0.000154	0.048425	4.844629	27.67574	0.007151	-0.04962
X1675	0.919156	0.392502	0	0.03768	4.36003	24.64212	0.002268	-0.02615
X1680	0.979254	0.486224	8.42E-05	0.034849	0.349942	0.464368	0.001436	0.016605
X1685	0.971213	0.471222	0.000307	0.035924	0.629506	1.142356	0.004459	-0.02049
X1690	1.045259	0.601676	0	0.028524	1.740061	3.801076	0.002466	-0.00585
X1695	1.098627	0.687324	7.99E-05	0.027764	1.517458	2.198815	0	0.012454
X1700	1.153827	0.777419	0	0.026384	3.407051	14.69137	0.002568	-0.01819
X1705	1.233793	0.951971	0	0.019233	3.310144	10.97008	0.00671	-0.01884
X1710	1.313669	1.090566	0.000393	0.015931	1.900889	1.86555	0.0007	0.017141
X1715	1.285611	1.023263	0	0.017425	0.765447	1.046584	0.000924	0.012336
X1720	1.338496	1.157163	0.000916	0.0133	0.116054	0.083606	0.001524	-0.00864
X1725	1.327933	1.124888	0	0.014798	0.554256	0.081493	0	-0.04936
X1730	1.314397	1.109256	0.001193	0.01547	2.339988	3.867666	0.007872	0.017247
X1735	1.298959	1.087955	0	0.01364	0.717386	2.676303	0	0.02889
X1740	1.288327	1.015746	0	0.014402	1.48516	0.312319	0.001957	0.005051
X1745	1.337029	1.111855	0	0.012066	3.054132	7.956321	0.004882	-0.03839
X1750	1.377677	1.188174	0	0.008108	0.424013	1.520978	0	0.023043
X1755	1.321036	1.045957	0.000159	0.012472	3.079931	11.68213	0.002032	0.029706
X1760	1.311689	1.024617	8.06E-05	0.008657	0.92482	1.235381	0.000141	0.010844
X1765	1.328411	1.052763	0.000254	0.010429	1.192166	0.87654	9.85E-05	0.021298
X1770	1.346543	1.088549	0	0.010972	0.211309	0.157028	0.001174	-0.00696

X1775	1.342274	1.057839	6.47E-05	0.005535	0.749039	0.481406	0.000453	-0.00578
X1780	1.362758	1.08803	9.07E-05	0.003562	0.21074	0.744945	0.004211	0.020844
X1785	1.335724	1.010969	0	0.00199	0.743489	0.001497	0.000733	-0.04772
X1790	1.40014	1.140732	0	0.00038	2.812899	6.965308	0.002035	-0.02409
X1795	1.471852	1.222139	7.77E-05	-0.00725	0.746328	0.51115	0	-0.00769
X1955	2.76118	5.056534	0.005369	-0.06954	0.288428	0.486749	0.007191	-0.03159
X1960	2.389354	3.28082	0.000688	-0.08882	3.192071	8.526071	0.014523	0.025024
X1965	1.931303	1.708664	0	-0.01754	2.975005	5.341935	0.002566	0.037168
X1970	1.676576	1.223967	0.000881	-0.01891	2.314297	3.13171	0.001904	-0.0052
X1975	1.492391	0.915008	0.000403	0.000167	2.89872	5.494158	0.00065	0.014783
X1980	1.366696	0.760413	0.001358	-0.00848	0.232254	0.084838	0	-0.05166
X1985	1.339921	0.71306	0.004001	-0.00616	0.451458	1.303792	0.000646	0.021428
X1990	1.280008	0.607505	0.003866	0.000296	0.337003	0.00276	0.004318	-0.02057
X1995	1.260992	0.5396	0.000411	0.000823	1.869539	1.99818	0.001123	0.005432
X2000	1.447384	0.884587	0.004487	-0.00225	2.539441	3.989223	0.004832	-0.00095
X2005	1.735543	1.332621	0.002974	-0.01744	1.578755	0.68329	0.000781	-0.02804
X2010	1.709575	1.344817	0.000117	-0.02513	0.617863	0.675758	0.000244	-0.0151
X2015	1.59317	1.035591	0	-0.00728	2.328139	4.427902	0.001043	0.055275
X2020	1.543999	1.037355	0.000787	-0.01073	2.263908	2.264037	0.000195	0.025305
X2025	1.317	0.695068	0.003734	0.012073	2.256819	2.827239	0.001742	0.004402
X2030	1.244916	0.629251	0.000153	0.008513	2.930444	6.948491	0.012967	0.036282
X2035	1.116311	0.443528	5.89E-05	0.027469	0.004405	0.045381	0.003525	0.00863
X2040	1.163511	0.519597	0	0.018267	1.190637	0.586503	0.002021	0.001895
X2045	1.125192	0.450251	0.001878	0.026302	0.927522	1.914745	0.000186	0.022481
X2050	1.020153	0.29147	0.012218	0.026492	1.857896	0.857525	0.000208	-0.00427
X2055	1.226979	0.525352	0.005579	0.024095	3.485751	13.05889	0.000914	0.01577
X2060	1.344762	0.791798	0.001299	0.022499	0.931122	1.766504	0.00462	-0.02503
X2065	1.194205	0.571098	8.55E-05	0.021808	1.985928	3.594337	0	0.018938
X2070	1.217094	0.622192	0	0.025114	1.277859	1.724369	7.61E-05	-0.04549
X2075	1.21198	0.609738	0.00171	0.019495	1.256329	0.981966	6.03E-05	0.04449
X2080	1.139079	0.526518	0	0.035021	1.047772	0.692685	0.001998	0.044659
X2085	1.118821	0.515209	0.00145	0.02956	0.339038	0.129993	0	0.000851
X2090	1.079672	0.453288	0.001615	0.033995	1.889006	2.349124	0.001762	-0.00743
X2095	1.19652	0.631073	0.001239	0.026424	1.214468	0.710758	0.001681	-0.02491
X2100	1.098701	0.461907	0.001548	0.036028	2.108011	5.479173	0.007423	0.063929
X2105	1.064519	0.439468	0.000768	0.037722	2.103057	6.062727	0.004877	-0.03577
X2110	1.131062	0.539592	0.002747	0.036082	0.732749	0.078908	0.000336	0.003537
X2115	1.125333	0.49852	0	0.037522	0.413578	0.01567	0	-0.01581
X2120	1.048563	0.433958	0.000172	0.039744	1.725143	2.263987	0.000667	-0.00843
X2125	0.869628	0.223868	0.000988	0.036487	0.55703	1.061052	0.000242	-0.01346
X2130	0.909955	0.218799	0.000751	0.042785	1.180583	0.964693	0.001404	0.006524
X2135	0.949431	0.309963	0.0027	0.026317	0.653079	0.502764	0.000821	0.006566
X2140	0.916548	0.277057	0.000136	0.031804	1.560137	2.426378	0.000627	-0.02625
X2145	1.050236	0.357282	0	0.04099	1.035371	0.813794	0.002108	-0.00623
X2150	0.886768	0.289133	0.000923	0.028792	1.329979	1.191874	0.001007	-0.01946
X2155	0.638273	0.070997	0.002706	0.029988	0.228712	0.001216	0	-0.00799

X2160	0.911018	0.226272	0.000171	0.029803	1.393935	0.887808	0	0.040248
X2165	0.970269	0.30468	0.002247	0.039185	0.867623	2.645982	0.001637	-0.0389
X2170	1.310374	0.893579	0.006423	0.005551	0.119208	0.00052	0.004386	-0.00939
X2175	0.986573	0.369415	0.001535	0.025288	1.393143	0.303821	0	0.016607
X2180	1.172792	0.726923	0.001397	0.015902	3.042038	12.28852	0.000835	-0.02973
X2185	1.237095	0.825314	0.000228	0.005708	1.651819	5.086292	0.002015	0.054057
X2190	1.015363	0.380159	0.000197	0.021592	0.066579	0.004407	0.001982	0.016897
X2195	1.21233	0.757106	0.001162	0.024617	0.834367	1.020553	0.003158	0.028101
X2200	1.183909	0.639426	0.000227	0.029807	0.859266	0.318054	0.001757	-0.02656
X2205	1.211226	0.729675	0	0.019486	0.975421	0.823218	0.000975	0.029873
X2210	1.146882	0.576934	0.001324	0.029406	0.036038	0.210823	0.001077	0.048613
X2215	1.15918	0.588348	0	0.035365	0.647934	1.043403	0.00053	-0.01696
X2220	1.268506	0.8222	0	0.019342	1.607858	3.163358	0.019905	-0.08454
X2225	1.174298	0.657165	0.004439	0.015438	1.474376	0.986746	0.000943	-0.00754
X2230	1.175082	0.709316	0	0.020442	0.951972	0.974686	0.004853	-0.01429
X2235	1.18347	0.701511	0.000963	0.016688	0.262745	0.053531	0	-0.00874
X2240	1.121304	0.588565	0	0.014619	0.527122	0.421196	0.002537	0.032627
X2245	1.135828	0.633393	0	0.024709	0.851524	1.119424	0.000512	-0.0007

## 8.6 MULTISPECTRAL FEATURES TUNING RANDOM FOREST

n_estimators	max_features	max_depth	error	AUC	Accuracy
50	15	6	59	0.794032	0.823353
65	30	5	62	0.78995	0.814371
70	15	6	61	0.789225	0.817365
50	10	7	60	0.7885	0.820359
70	30	5	62	0.788385	0.814371
45	15	6	61	0.78766	0.817365
25	35	5	62	0.786821	0.814371
45	10	6	61	0.786096	0.817365
40	35	7	63	0.785981	0.811377
50	30	5	62	0.785256	0.814371
30	35	7	64	0.785142	0.808383
30	15	4	62	0.783692	0.814371
40	10	5	62	0.783692	0.814371
45	15	5	62	0.783692	0.814371
65	20	5	64	0.783578	0.808383
70	20	5	64	0.783578	0.808383
50	35	7	64	0.783578	0.808383
70	35	7	64	0.783578	0.808383
65	10	6	63	0.782853	0.811377
50	25	7	63	0.782853	0.811377
30	20	4	63	0.782853	0.811377
35	15	5	63	0.782853	0.811377
40	30	6	63	0.782853	0.811377
65	10	7	62	0.782128	0.814371
70	10	7	62	0.782128	0.814371
45	30	6	64	0.782013	0.808383
70	20	6	64	0.782013	0.808383
50	15	5	63	0.781288	0.811377
50	10	6	63	0.781288	0.811377
35	10	5	63	0.781288	0.811377
40	15	5	63	0.781288	0.811377
65	35	7	65	0.781174	0.805389
40	25	7	64	0.780449	0.808383
45	35	7	64	0.780449	0.808383
65	25	5	64	0.780449	0.808383
70	25	5	64	0.780449	0.808383
70	25	7	64	0.780449	0.808383
70	15	5	63	0.779724	0.811377
40	35	5	63	0.779724	0.811377
20	35	5	65	0.779609	0.805389
65	25	7	64	0.778884	0.808383
35	30	5	64	0.778884	0.808383
30	35	5	64	0.778884	0.808383

40	15	6	64	0.778884	0.808383
35	35	7	66	0.77877	0.802395
35	35	5	63	0.778159	0.811377
45	5	6	63	0.778159	0.811377
45	10	7	63	0.778159	0.811377
50	10	5	63	0.778159	0.811377
65	15	5	63	0.778159	0.811377
30	15	5	65	0.778045	0.805389
40	25	6	65	0.778045	0.805389
40	35	6	65	0.778045	0.805389
50	30	6	65	0.778045	0.805389
70	10	6	64	0.77732	0.808383
65	15	6	64	0.77732	0.808383
25	30	4	64	0.77732	0.808383
50	20	5	66	0.777205	0.802395
65	25	6	66	0.777205	0.802395
70	30	6	66	0.777205	0.802395
45	25	5	66	0.777205	0.802395
35	30	6	65	0.77648	0.805389
40	15	7	65	0.77648	0.805389
70	30	7	67	0.776366	0.799401
45	35	5	64	0.775755	0.808383
40	10	7	64	0.775755	0.808383
40	25	5	66	0.775641	0.802395
25	30	5	66	0.775641	0.802395
25	30	6	66	0.775641	0.802395
30	15	7	66	0.775641	0.802395
45	25	7	66	0.775641	0.802395
65	30	6	66	0.775641	0.802395
70	35	6	66	0.775641	0.802395
50	35	6	65	0.774916	0.805389
70	25	6	67	0.774802	0.799401
45	20	6	66	0.774077	0.802395
25	35	6	66	0.774077	0.802395
30	35	6	66	0.774077	0.802395
35	35	6	66	0.774077	0.802395
35	15	7	66	0.774077	0.802395
45	15	7	66	0.774077	0.802395
50	20	6	66	0.774077	0.802395
65	15	7	66	0.774077	0.802395
65	30	7	68	0.773962	0.796407
25	15	4	65	0.773352	0.805389
45	10	5	65	0.773352	0.805389
50	25	5	67	0.773237	0.799401
65	20	6	67	0.773237	0.799401
45	30	5	67	0.773237	0.799401



30	30	6	67	0.773237	0.799401
25	15	7	67	0.773237	0.799401
45	30	7	67	0.773237	0.799401
65	35	5	66	0.772512	0.802395
30	30	5	66	0.772512	0.802395
25	10	6	66	0.772512	0.802395
35	15	6	66	0.772512	0.802395
35	25	6	66	0.772512	0.802395
45	35	6	66	0.772512	0.802395
30	10	7	66	0.772512	0.802395
50	30	7	68	0.772398	0.796407
30	30	3	65	0.771787	0.805389
35	10	7	65	0.771787	0.805389
65	10	5	65	0.771787	0.805389
50	35	5	65	0.771787	0.805389
40	30	5	67	0.771673	0.799401
40	20	6	67	0.771673	0.799401
30	25	6	67	0.771673	0.799401
45	25	6	67	0.771673	0.799401
50	15	7	67	0.771673	0.799401
35	10	6	66	0.770948	0.802395
30	35	4	67	0.770108	0.799401
30	10	6	67	0.770108	0.799401
35	20	6	67	0.770108	0.799401
70	35	5	66	0.769383	0.802395
50	5	7	66	0.769383	0.802395
10	15	7	66	0.769383	0.802395
35	25	5	68	0.769269	0.796407
25	35	7	68	0.769269	0.796407
65	35	6	68	0.769269	0.796407
70	15	7	68	0.769269	0.796407
30	15	6	67	0.768544	0.799401
35	20	5	69	0.768429	0.793413
40	20	5	69	0.768429	0.793413
10	15	4	66	0.767819	0.802395
65	5	6	66	0.767819	0.802395
25	20	6	68	0.767705	0.796407
30	20	6	68	0.767705	0.796407
15	35	7	70	0.76759	0.790419
30	30	4	67	0.76698	0.799401
45	20	5	69	0.766865	0.793413
25	15	6	69	0.766865	0.793413
30	25	4	68	0.76614	0.796407
15	35	4	68	0.76614	0.796407
25	30	7	70	0.766026	0.790419
35	30	7	70	0.766026	0.790419

20	15	4	67	0.765415	0.799401
15	35	5	69	0.765301	0.793413
25	20	4	68	0.764576	0.796407
25	35	4	68	0.764576	0.796407
40	10	6	68	0.764576	0.796407
25	25	6	70	0.764461	0.790419
25	30	3	67	0.763851	0.799401
25	25	4	69	0.763736	0.793413
20	35	4	69	0.763736	0.793413
25	15	5	69	0.763736	0.793413
20	30	6	69	0.763736	0.793413
40	30	7	71	0.763622	0.787425
70	10	5	68	0.763011	0.796407
50	5	6	68	0.763011	0.796407
50	25	6	70	0.762897	0.790419
70	20	7	70	0.762897	0.790419
20	15	5	70	0.762897	0.790419
30	25	5	70	0.762897	0.790419
20	35	6	70	0.762897	0.790419
25	10	7	70	0.762897	0.790419
35	25	7	70	0.762897	0.790419
15	20	5	71	0.762057	0.787425
20	35	7	71	0.762057	0.787425
30	10	4	68	0.761447	0.796407
15	15	4	68	0.761447	0.796407
45	20	7	70	0.761332	0.790419
65	20	7	70	0.761332	0.790419
30	30	7	72	0.761218	0.784431
30	35	3	69	0.760607	0.793413
25	20	5	71	0.760493	0.787425
30	25	7	71	0.760493	0.787425
20	20	5	72	0.759654	0.784431
25	25	7	72	0.759654	0.784431
70	5	7	69	0.759043	0.793413
25	25	3	69	0.759043	0.793413
10	35	4	69	0.759043	0.793413
30	10	5	69	0.759043	0.793413
20	10	6	69	0.759043	0.793413
20	20	4	71	0.758929	0.787425
20	25	5	71	0.758929	0.787425
25	25	5	71	0.758929	0.787425
10	15	6	71	0.758929	0.787425
15	30	6	72	0.758089	0.784431
70	5	6	69	0.757479	0.793413
15	20	6	73	0.75725	0.781437
25	10	3	68	0.756754	0.796407

20	30	4	70	0.756639	0.790419
20	15	7	72	0.756525	0.784431
35	20	7	72	0.756525	0.784431
40	20	7	72	0.756525	0.784431
30	20	5	73	0.755685	0.781437
15	15	6	73	0.755685	0.781437
15	30	4	70	0.755075	0.790419
45	5	7	70	0.755075	0.790419
50	20	7	72	0.75496	0.784431
30	5	6	72	0.75496	0.784431
10	20	3	71	0.754235	0.787425
15	30	5	72	0.753396	0.784431
25	20	7	72	0.753396	0.784431
45	5	5	69	0.752785	0.793413
65	5	7	71	0.752671	0.787425
20	20	3	71	0.752671	0.787425
15	35	3	71	0.752671	0.787425
20	35	3	71	0.752671	0.787425
40	5	6	71	0.752671	0.787425
10	15	5	73	0.752556	0.781437
20	30	3	70	0.751946	0.790419
30	5	4	70	0.751946	0.790419
70	5	5	71	0.751107	0.787425
15	30	3	71	0.751107	0.787425
15	20	4	73	0.750992	0.781437
30	20	7	73	0.750992	0.781437
20	25	7	75	0.750878	0.775449
20	5	4	70	0.750382	0.790419
50	5	5	70	0.750382	0.790419
20	30	5	74	0.750153	0.778443
15	30	7	76	0.750038	0.772455
10	20	4	73	0.749428	0.781437
20	10	7	73	0.749428	0.781437
10	20	5	75	0.749313	0.775449
20	15	6	75	0.749313	0.775449
20	20	6	75	0.749313	0.775449
30	10	3	70	0.748817	0.790419
20	10	4	72	0.748703	0.784431
25	10	5	72	0.748703	0.784431
10	35	5	72	0.748703	0.784431
40	5	7	72	0.748703	0.784431
20	25	4	74	0.748588	0.778443
10	30	7	76	0.748474	0.772455
10	35	7	76	0.748474	0.772455
30	25	3	73	0.747863	0.781437
25	35	3	73	0.747863	0.781437

15	35	6	75	0.747749	0.775449
65	5	5	72	0.747138	0.784431
25	20	3	72	0.747138	0.784431
10	30	4	72	0.747138	0.784431
15	15	5	76	0.746909	0.772455
15	25	7	76	0.746909	0.772455
25	5	4	71	0.746413	0.787425
30	20	3	73	0.746299	0.781437
15	25	5	76	0.745345	0.772455
10	10	3	71	0.744849	0.787425
10	30	6	75	0.74462	0.775449
15	15	7	75	0.74462	0.775449
20	25	6	77	0.744505	0.769461
30	15	3	72	0.744009	0.784431
15	20	3	74	0.743895	0.778443
10	20	6	78	0.743666	0.766467
20	30	7	78	0.743666	0.766467
15	10	3	71	0.743284	0.787425
20	10	3	71	0.743284	0.787425
25	10	4	73	0.74317	0.781437
35	5	5	73	0.74317	0.781437
40	5	5	74	0.742331	0.778443
20	5	6	74	0.742331	0.778443
35	5	7	74	0.742331	0.778443
25	5	6	76	0.742216	0.772455
5	30	7	80	0.740423	0.760479
25	15	3	73	0.740041	0.781437
10	30	3	75	0.738362	0.775449
10	35	3	75	0.738362	0.775449
15	20	7	77	0.738248	0.769461
20	20	7	77	0.738248	0.769461
10	25	5	79	0.738133	0.763473
10	15	3	74	0.737637	0.778443
10	25	3	76	0.737523	0.772455
15	10	6	76	0.737523	0.772455
10	35	6	78	0.737408	0.766467
5	35	7	82	0.737179	0.754491
15	10	5	75	0.736798	0.775449
20	10	5	75	0.736798	0.775449
10	25	4	77	0.736683	0.769461
15	25	4	77	0.736683	0.769461
10	10	6	77	0.736683	0.769461
15	10	7	77	0.736683	0.769461
15	25	3	76	0.735958	0.772455
5	35	3	76	0.735958	0.772455
35	5	6	76	0.735958	0.772455

30	5	5	75	0.735234	0.775449
20	5	7	77	0.735119	0.769461
15	5	6	76	0.734394	0.772455
25	5	7	76	0.734394	0.772455
30	5	7	76	0.734394	0.772455
10	30	5	78	0.73428	0.766467
5	15	4	81	0.733326	0.757485
20	15	3	76	0.73283	0.772455
30	30	2	77	0.73199	0.769461
20	25	3	77	0.73199	0.769461
25	30	2	76	0.731265	0.772455
5	30	5	80	0.731036	0.760479
25	25	2	77	0.730426	0.769461
25	5	5	78	0.729586	0.766467
10	25	6	81	0.728632	0.757485
10	10	5	79	0.727183	0.763473
15	25	6	81	0.727068	0.757485
5	15	5	83	0.726954	0.751497
5	35	5	82	0.726229	0.754491
25	5	3	75	0.725847	0.775449
5	20	3	79	0.725618	0.763473
15	15	3	78	0.724893	0.766467
10	10	4	78	0.724893	0.766467
5	30	4	80	0.724779	0.760479
15	5	7	80	0.724779	0.760479
5	20	4	82	0.724664	0.754491
10	25	7	83	0.723825	0.751497
30	20	2	80	0.723214	0.760479
10	20	7	82	0.7231	0.754491
15	5	4	79	0.722489	0.763473
20	30	2	78	0.721764	0.766467
30	25	2	80	0.72165	0.760479
25	35	2	80	0.72165	0.760479
15	10	4	80	0.72165	0.760479
20	25	2	81	0.72081	0.757485
5	10	6	83	0.720696	0.751497
20	5	3	78	0.7202	0.766467
25	15	2	80	0.720085	0.760479
5	25	7	87	0.717338	0.739521
15	5	3	78	0.717071	0.766467
30	5	3	78	0.717071	0.766467
10	5	6	83	0.716003	0.751497
30	35	2	82	0.715278	0.754491
5	35	4	84	0.715163	0.748503
10	10	7	84	0.715163	0.748503
30	15	2	81	0.714553	0.757485

15	5	5	82	0.713713	0.754491
20	5	5	82	0.713713	0.754491
5	35	6	88	0.71337	0.736527
5	15	7	87	0.712645	0.739521
15	25	2	83	0.71131	0.751497
5	30	3	85	0.711195	0.745509
5	10	3	82	0.710585	0.754491
5	10	4	85	0.709631	0.745509
20	20	2	84	0.708906	0.748503
20	35	2	84	0.708906	0.748503
15	35	2	83	0.708181	0.751497
10	5	7	87	0.707952	0.739521
5	30	6	89	0.707837	0.733533
5	10	7	89	0.707837	0.733533
15	20	2	84	0.707341	0.748503
5	25	5	89	0.706273	0.733533
5	20	5	90	0.705433	0.730539
15	30	2	83	0.705052	0.751497
20	15	2	84	0.704212	0.748503
10	20	2	84	0.704212	0.748503
10	5	4	84	0.704212	0.748503
25	20	2	86	0.702534	0.742515
5	20	6	93	0.701351	0.721557
10	30	2	84	0.701084	0.748503
5	5	6	93	0.698222	0.721557
10	25	2	87	0.697001	0.739521
10	35	2	87	0.695437	0.739521
5	15	3	87	0.695437	0.739521
5	25	4	93	0.695093	0.721557
5	25	2	88	0.694597	0.736527
5	20	7	93	0.693529	0.721557
5	15	6	93	0.691964	0.721557
5	35	2	90	0.688225	0.730539
5	25	3	92	0.686546	0.724551
5	10	5	95	0.685592	0.715569
5	25	6	97	0.685478	0.709581
10	15	2	89	0.684371	0.733533
10	5	3	88	0.682082	0.736527
30	10	2	86	0.679067	0.742515
25	10	2	86	0.677503	0.742515
5	5	7	98	0.675252	0.706587
10	5	5	93	0.673191	0.721557
5	20	2	95	0.671513	0.715569
5	5	4	97	0.669834	0.709581
15	15	2	93	0.668498	0.721557
5	30	2	95	0.668384	0.715569



30	5	2	92	0.652129	0.724551
5	5	5	107	0.648924	0.679641
10	10	2	93	0.648161	0.721557
20	10	2	94	0.644193	0.718563
15	10	2	96	0.639385	0.712575
5	5	3	101	0.636752	0.697605
25	5	2	97	0.635417	0.709581
5	15	2	105	0.625572	0.685629
15	5	2	101	0.61485	0.697605
10	5	2	102	0.614011	0.694611
20	5	2	102	0.607753	0.694611
5	10	2	108	0.599588	0.676647
5	5	2	113	0.586004	0.661677

## 8.7 LINK TO RAW DATA

<https://drive.google.com/drive/folders/1SqIH22jXSoXbkM2O66yyRf7GxMtMRmMx?usp=sharing>

## 8.8 LINK TO CODE

<https://drive.google.com/drive/folders/1QUBAvBVfV3GPvQ8EnjJpRYCXz-d3s2rg?usp=sharing>