DETECCIÓN TEMPRANA Y DISCRIMINACIÓN DE ENFERMEDADES FÚNGICAS EN PLANTAS USANDO ESPECTROSCOPÍA *IN SITU*

JUAN CARLOS MARÍN ORTIZ

Universidad Nacional de Colombia

Facultad de Ciencias Agrarias

Doctorado en Ciencias Agrarias

Medellín, Colombia

2019

DETECCIÓN TEMPRANA Y DISCRIMINACIÓN DE ENFERMEDADES FÚNGICAS EN PLANTAS USANDO ESPECTROSCOPÍA *IN SITU*

JUAN CARLOS MARÍN ORTIZ

Presentado como requisito parcial para optar al título de:

Doctor en Ciencias Agrarias

Director:
PhD. VERONICA BOTERO FERNANDEZ

Co-Director:
PhD. LILLIANA MARÍA HOYOS CARVAJAL

Línea de Investigación:
Fitopatología

Universidad Nacional de Colombia

Facultad de Ciencias Agrarias

Medellín, Colombia

2019

Acknowledgements

I gratefully acknowledge the financial support of the Colombian Administrative Department of Science, Technology and Innovation (COLCIENCIAS) for financing my doctorate in Agricultural Sciences and doctoral internship (University of Brasilia, Brazil). In addition, I must thank the National University of Colombia for its support with the partial financing made with funds from the National Call for Support to Projects of Research and Artistic Creation of Colombian National University (2015-2016, 2017-2018).

I would like to express my special appreciation and thanks to my directors, professors Dra. Veronica Botero Fernández and Lilliana M. Hoyos Carvajal, you have been great mentors to me. I would like to thank you for your unconditional support in all stages of my doctoral training, in addition to encouraging my research and allowing me to grow as a research scientist. Your advice on both research as well as on my career have been priceless. I would also thank my evaluating committee, for performing this function at times when the time has been critical for them and for me. I also want to thank you for letting my defense be an enjoyable moment, and for your brilliant comments and suggestions, thanks so much to you. I would also like to thank undergraduate students who participated directly or indirectly in the experiments assembly, they made my doctoral project's weight a little "lighter".

A special thanks to my family. Words cannot express how grateful I am to my mother for all of the sacrifices that you have made on my behalf and his unconditional support. You have always been with me when it seems that all the lights go out. I would also like to thank my daughter, Susana, for all this time stolen and invested in what we call science. I hope to be able to return something of the lost in near future. At the end I would like express appreciation to my beloved wife Alejandra who spent sleepless nights with me and was always my support in the moments when there was no one to answer my questions. My frequent absences, the continuous stress, and a growing bad mood when the moments of delivery of results approached always you knew how handle them it in the best way. You showed me the strength of an integral woman and you were the "structure" that helped me to face my most difficult moments.

For these reasons, I cannot stop thinking that this achievement is not only mine, it is really of many (even those people not mentioned here, because of my poor memory), and without you it would not have been possible! Thank you very much to all!

Abstract

After infection, a plant develops symptoms that appear in different parts of plants; however, at moment in which these symptoms are visible, the plant can already be affected negatively. In addition, plants that remain asymptomatic are pathogens reservoirs, since they can remain infected for most of their development cycle, becoming a source of contamination for entire crop. After the symptoms onset, disease is verified using detection techniques, such as ELISA, Polymerase Chain Reaction, Immunofluorescence, Flow Cytometry, Fluorescence in situ and, Gaseous Metabolite Profiles, among others. However, despite the availability of these techniques, a diseases early detection system based on spectrometry techniques can help to reduce losses caused in crops and prevent a greater spread of disease, with more speed, sensitivity, selectivity and without requiring the samples destruction required for analysis. The aim of this study is to evaluate early detection of plants diseases caused by fungal infections using *in situ* reflectance spectroscopy. To achieve this, reflectance spectra were measured from leaves of *S. lycopersicum* infected with a fungus pathogenic strains at various times of pathogenesis before the symptoms of the disease were visible. Additionally, physiological analyzes were performed and were related to reflectance spectra of the infected and healthy plants in different infection periods; also, were developed disease prediction models based on Vis/NIR reflectance data before the visual expression of the symptoms using different multivariate statistical tools. In this study it was possible to characterize the spectral variation in leaves of *S. lycopersicum* L. infected with *F. oxysporum* during the incubation period. It was also possible to identify the relevant specific wavelengths in the range of 380-1000 nm that can be used as spectral signatures for the detection and discrimination of vascular wilt in *S. lycopersicum*. We watch that inoculated tomato plants increased their reflectance in the visible range (Vis) and decreased slowly in the near infrared range (NIRs) measured during incubation, showing marked differences with plants subjected to water stress in the VIS/NIR. Additionally, three ranges were found in the spectrum related to infection by *F. oxysporum* (510nm-520nm, 650nm-670nm and 700-750nm). Linear discriminant models on spectral reflectance data were able to differentiate between tomatos varieties inoculated with *F. oxysporum* from healthy ones with accuracies higher than 70% 9 days after inoculation (only with three explanatory variables). Additionally, it was possible to characterize and relate the spectral variance in leaves of *S. lycopersicum* infected with *F. oxysporum* with the physiological variation and pathogen concentration in tomato plants during the asymptomatic period of vascular wilt. Photosynthetic parameters derived from gaseous exchange analyzes in the tomato leaves correlated related with four bands in the visible range (Vis). Additionally, five specific bands also correlated highly correlated with the increase of *F. oxysporum* conidia concentration measured at root: 448-523nm, 624-696nm, 740-960nm, 973-976nm and 992-995nm. These wavelengths allowed classifying correctly 100% the plants inoculated with *F. oxysporum* of plants subjected to hydric stress and controls in the disease asymptomatic period. Finally, it was possible to develop logistic regression models to predict infection by *F. oxysporum* in plants, obtaining accuracies and areas under the curve greater than 0.9 for one of the tomato varieties evaluated. The results of this study will contribute to a better understanding of the optical properties of the plant during the development of fungal diseases. These methods will be applicable in development of precision crops, specifically in crop protection, differentiation, quantification, and disease early detection of plant; in addition to, the developed models which can be used as a basic input in the design of technological tools that allow the plant disease detection in real time.

Key words: early detection, diseases plants, fungal infection, spectroscopy, reflectance, *F. oxysporum*

Contents

List of figures

List of tables

List of annexed

List of abbreviations

Abbreviation	Meaning
A	CO_2 net assimilation rate
A/E	Transpiration efficiency
A/gs	Intrinsic water use efficiency
AIC	Akaike Information Criterion
AUC	Areas under the curve
BLRM	Binomial Logistic Regression Models
CFU	Colony Forming Unit
Ci	Intercellular CO_2 concentration
Ci/Ca	Internal CO_2 and atmospheric CO_2
CLS	Classical Least Squares
dpi	Days post infection
E	Transpiration rate
ELISA	Enzyme Linked Immunoassay
FAO	Food and Agriculture Organization of the United Nations
Fig.	Figure
FISH	Fluorescent In Situ Hybridization
Fo5	*Fusarium oxysporum* f. sp. *passiflorae*
FPR	False Positive Rate
Ft	Continuous fluorescence
GDP	Gross Domestic Product
GLM	Generalized Linear Models
gs	Stomatal conductance
IF	Immunofluorescence
ILS	Inverse Least Squares
LDA	Linear Discriminant Analysis
LRTest	Likelihood Ratio Test
mg	miligram
NIR	Near-Infrared Range
nm	Nanometers
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
PDA	Potato Dextrose Agar
PLS	Partial Least Squares
r2CU	Ragg and Uhler Pseudo R^2
ROC	Receiver Operating Characteristic
RSW	Relevant Specific Wavelength
SDI	Spectral Disease Indices
SNV	Standard Normal Variance
SVI	Spectral Vegetation Indices
TPR	True Positive Rate
vif	Variance-inflation factors

Vis	Visible
ϕ	Overdispersion coefficients
ΦPSII	Quantitative yield of PSII

1. GENERAL INTRODUCTION

1.1 Justification and research problem

The agricultural sector in Colombia has historically been an important sector in national economy, which had an average contribution about 7 % to the Gross Domestic Product (GDP), sectoral value added amounting to 366,737 million pesos in 2000-2008 period and 21% generated of country's jobs, according to National Administrative Department of Statistics (Romero, 2011). In addition, agricultural production of the country's main crops amounted to approximately 9,903,569 tons in 2012 (DANE, 2013). However, it has been estimated that crop diseases and pests cause approximately 30% of pre-harvest losses in developing countries or even more specifically in cereals (Koyshibayev and Muminjanov 2016); with insects, fungi, bacteria and viruses mainly responsible for this damage.

Although there are some methods which are currently used for the detection of diseases in plants, such as the enzyme linked immunoassay (ELISA), polymerase chain reaction (PCR), immunofluorescence (IF), flow cytometry, fluorescence In situ and gaseous metabolite profiles. They are generally used after the onset of symptoms, when the disease is at an advanced stage in the plant and dispersed in the crop. It possible use these methods before symptoms, but they require a lot of time, the destruction of plant samples, and usually they are expensive. This is the reasons a system of early detection of the disease can help to diminish losses caused in crops and prevent a greater spread of disease, quickly, sensitivity, selectivity and without requiring the destruction of samples for analysis (Chaerle and Van 2000).

The plant diseases early detected in plantations or in propagated materials, and selective eradication programs are methods used to prevent the advance of pathogens. Once the problem is established, control has traditionally been carried out with exclusively chemical methods. However, this is difficult because these methods generate resistance problems, added to low efficiency of some active principles (Moss, 2010). The high cost of control measures and environmental impact that they generate have provoked great interest in precision agricultural techniques in which the products application is carried out in the affected specific sites by the disease. These site-specific applications result in a reduction in use of pesticides and, therefore, can minimize costs and the ecological impact of agricultural crop production systems (Gebber and Adamchuk 2010).

In addition, it must be taken into account that in plants, disease management of their timely detection is fundamental and the non-destructive methods of productive units are desirable. Among the above,

there is remote sensing by spectroscopic techniques, used in initial stages of plant material propagation by direct reading of planting units, being useful in identification of primary foci of disease (Franke and Menz 2007; Franke et al. 2009; Marin et al. 2018). To implement these sensors within precision plant protection technologies, these have to be robust, low cost and preferably offering real-time detection (Zhang et al. 2002).

For these techniques based on spectroscopy to be applicable, it is necessary that specific pathosystem information about spectral signature of infectious processes be generated, as a primary input of knowledge that feeds the spectral libraries used to identify diseases. That is why we intend to carry out the study in tomato *(S. lycopersicum*), a plant with widely known agronomic characteristics and susceptible to range of pathogens, which allows contrasting different colonization processes of microorganisms. In addition, *Fusarium oxysporum* f. sp. Cubense is a widely accepted model organism for the plants pathogenicity study, and is morphologically diverse and widely distributed species of the genus Fusarium spp., which is currently considered a species complex of plant pathogenic fungi (Baayen et al. 2000). This species has many pathogenic strains that parasitize more than plants species 100 (Bosland 1998), characterized by producing colonies of rapid growth and very variable morphology, predominating two primordial types: one mycelial and one pionotal type (Garcés et al. 2001). In general, the majority of pathogenic strains produce symptoms of wilt, accompanied by partial yellowing of leaves and folding of disease plant buds. A characteristic that allows to diagnose and differentiate quickly this disease is the whitish, yellowish or brown coloration in vascular bundles; additionally, there may be dwarfing in shoots and decrease in the plant overall growth (Garcés et al. 1999).

1.2 Hypothesis

- The leaves reflectance of *S. lycopersicum* plants with systemic fungal infections differs in healthy leaves from early stages of disease, before the symptoms are visible.
- Fungal infections generate changes in *S. lycopersicum* plants physiology in early phases the disease, influencing their optical properties and allowing the use of in situ reflectance spectrometry for early detection.
- The plants disease presence can be accurately determined by means of multivariate calibration models of spectral data, based on changes generated by the phytopathogenic fungus in *S. lycopersicum* plants in initial stages in which disease is asymptomatic.

1.3 Objective

1.3.1 General

Detect in asymptomatic period the vascular wilt in plants from *in situ* reflectance spectroscopy in model *Solanum lycopersicum-Fusarium oxysporum*.

1.3.2 Specific

- Characterize the plants spectral changes in *S. lycopersicum* infected with *F. oxysporum* before the appearance of visible symptoms using VIS/NIR reflectance spectroscopy.
- Identify and relate the physiological and spectral responses before occurrence of visible symptoms caused by fungal infections in *Solanum lycopersicum*.
- Perform prediction models for vascular wilt detection in asymptomatic period of vascular wilt in *Solanum lycopersicum* using Vis/NIR reflectance spectral data.

1.4 List of academic products

1.4.1 Articles

Marín-Ortiz, J. C., Gutierrez-Toro N., Fernandez, V. B., Hoyos-Carvajal, L.M. (2019). Linking physiological parameters with visible/near-infrared leaf reflectance in incubation period of vascular wilt disease. Saudi Journal of Biological Sciences. Doi: https://doi.org/10.1016/j.sjbs.2019.05.007 (A1 Publindex Colciencias: Chapter 2, published, see annex 3)

Marín-Ortiz, J. C., Hoyos-Carvajal, L. M., Fernandez, V. B. (2018). Detection of asymptomatic *Solanum lycopersicum* L. plants infected with *Fusarium oxysporum* using reflectance VIS spectroscopy VIS. Revista Colombiana de Ciencias Hortícolas, 12(2): 436-446. (B Publindex Colciencias: preliminary results, published, see annex 1).

Marín-Ortiz, J. C., Hoyos-Carvajal, L. M., Fernandez, V. B. (2018). Detection of significant wavelength for identifying and classification of *Fusarium oxysporum* during the incubation period and water stress in *Solanum lycopersicum* plants using reflectance spectroscopy. Journal of Plant Protection Research (A2 Publindex Colciencias: Chapter 1- accepted, in press, see annex 2)

1.4.2 Academic events

Marín-Ortiz JC, Hoyos-Carvajal L, Botero V. 2017. Detection of *Fusarium oxysporum* from asymptomatic *Solanum lycopersicum* L plants by means of reflectance spectroscopy. Agro-Geoinformatics, Fairfax, Virginia. EEUU.

Marín-Ortiz JC, Hoyos-Carvajal L, Botero V. 2017. Respuestas iniciales de genotipos tolerantes y sensibles de tomate al ataque de Fusaruim obtenidas por espectroscopia de reflectancia. Congreso nacional de la Sociedad Colombiana de Control de Malezas y Fisiología Vegetal "COMALFI", Medellín, Colombia.

Marín-Ortiz JC, Hoyos-Carvajal L, Botero V. 2017. Respuestas espectrales ante el estrés biótico y abiótico en plántulas de tomate. Congreso nacional de la Sociedad Colombiana de Control de Malezas y Fisiología Vegetal "COMALFI", Medellín, Colombia.

Da Silva DJ; **Marín-Ortiz JC**; Pinto Diniz C; De Andrade I; Moraes Samerto ME; De Albuquerque Vieira HL; Grattapaglia D; Sonsin-Oliveira J; De Alencar-Figueiredo LF. 2017. Classificação de madeiras comerciais com a espectroscopia no infravermelho próximo portátil. Congresso Nacional de Botânica - 36ª Jornada Fluminense de Botânica,Centro de Convenções Sul América, Rio de Janeiro, Brasil.

Pereira Quintino D., **Marín-Ortiz J.C**., De Oliveira Soares W.R., Batista Pinho D., Corrêa Café Filho A., De Alencar-Figueiredo LF. 2017. Identificação de espécies de Colletotrichum utilizando a espectroscopia no infravermelho próximo portátil. 23º Congresso de Iniciação Científica da Unb e 14º do DF, Brasilia, Brasil.

Da Silva D.J., **Marín-Ortiz J.C**., Silva Carvalho M., Diego Knop H., De Alencar-Figueiredo L.F. 2017. Classificação de famílias de briófitas das três divisões Anthocerotophyta, Bryophyta e Marchantiophyta usando a espectroscopia no infravermelho próximo portátil. 23º Congresso de Iniciação Científica da Unb e 14º do DF, Brasilia, Brasil.

Pereira Quintino D., **Marín-Ortiz J.C**, Gerard Miller R.N., De Alencar-Figueiredo L.F. 2017. Discriminação das variedades de banana Cavendish (Musa acuminata, cv. Cavendish Grande Naine) e Calcutta 4 (Musa acuminata subsp. burmannicoides var. Calcutta) usando a espectroscopia do infravermelho próximo portátil. 23º Congresso de Iniciação Científica da Unb e 14º do DF, Brasilia, Brasil.

Da Silva D.J., **Marín-Ortiz J.C**., Diniz C., De Andrade Sá I., Moraes Samerto M.E., De Albuquerque Vieira H.L., Grattapaglia, D., Sonsin-Oliveira J., De Alencar-Figueiredo L.F. 2017. Discriminação de madeiras comerciais utilizando um espectrômetro portátil. 23º Congresso de Iniciação Científica da Unb e 14º do DF, Brasilia, Brasil.

1.4.3 Projects derived from this investigation

2017-2018**.** "Detección temprana de la marchitez vascular en plantas mediante espectroscopia de reflectancia en el modelo *Solanum lycopersicum-Fusarium oxysporum*". Financiado por la convocatoria nacional para el apoyo a proyectos de investigación y creación artística de la Universidad Nacional de Colombia 2017-2018. Función: Co investigador – Estudiante de doctorado.

2015-2016**.** Detección temprana e identificación de enfermedades bacterianas en plantas usando espectroscopia in situ. Financiación por la convocatoria del programa nacional de proyectos para el fortalecimiento de la investigación, la creación y la innovación en posgrados de la universidad Nacional de Colombia. Función: Co investigador – Estudiante de doctorado.

1.4.3 Doctoral internship

2016. Botanic department of Brasilia University, Brazil. Doctoral internship director: professor Lucio Flavio de Alencar Figueiredo (PhD). Aims: I) Design and installation of experiments, II) Spectral data collection with different spectroscopy, III) analysis of plants spectral data subjected to fungal, viral and nematoid diseases. Duration: 6 months.

1.5 Introduction references

Baayen R.P., O`Donnell K., Bonants P.J.M., Cigelnik E., Kroon L.P.N.M., Roebroeck J.A., waalwijk C. 2000. Gene genealogies and AFLP analysis in the *Fusarium oxysporum* complex identify monophyletic and nonmonophyletic *formae especiales* causing wilt and rot disease. Phytopathology, 90(8): 891-900. DOI: 10.1094/PHYTO.2000.90.8.891

Bosland P.W. 1988. *Fusarium oxysporum* a pathogen of many plant species. Advances in plant pathology, 6(1): 281- 289. DOI: https://doi.org/10.1016/B978-0-12-033706-4.50023-2

Chaerle L, Van Der Straeten D. 2000. Imaging techniques and the early detection of plant stress. Trends Plant Science, 5(11):495-501. DOI:https://doi.org/10.1016/S1360-1385(00)01781-7

DANE. 2011. DANE. Recovered from http://www. dane.gov.co

Franke J., Menz, G. 2007. Multi-temporal wheat disease detection by multi-spectral remote sensing. Precision Agriculture, 8(3): 161-172. DOI: 10.1007/s11119-007-9036-y

Franke J., Mewes T. Menz, G. 2009. Requirements on spectral resolution of remote sensing data for crop stress detection. Cape Town: IEEE International Geoscience and Remote Sensing Symposium, 2009.

Gebbers R., Adamchuk V.I. 2010. Precision agriculture and food security. Science, 327(5957): 828-831. DOI: 10.1126/science.1183899

Garcés E., Orozco A.M., Zapata A.C. 1999. Fitopatología en Flores. Acta Biológica Colombiana. 4(2): 5- 26.

Garcés E., Orozco A.M., Bautista G.R., Valencia H. 2001. *Fusarium oxysporum* EL HONGO QUE NOS FALTA CONOCER. Acta Biológica Colombiana, 6(1): 7-25.

Marín-Ortiz J.C., Hoyos-Carvajal L.M., Botero-Fernández V. 2018. Detection of asymptomatic *Solanum lycopersicum* L. plants infected with *Fusarium oxysporum* using reflectance VIS spectroscopy. Colombian Journal of Horticultural Sciences, 12(2): 436-446. DOI: 10.17584/rcch.2018v12i2.7293

Moss S.R. 2010. Non-chemical methods of weed control: Benefits and limitations. Pp. 14–19. In Seventeenth Australasian Weeds Conference.

Romero Y. 2011. Incidence of agricultural GDP in national GDP. Evolution and transformation. Journal of economic literature, 8(2):49-60.

Zhang Q., Li Q., Zhang G. 2012. Rapid Determination of Leaf Water Content Using VIS/NIR Spectroscopy Analysis with Wavelength Selection. Spectroscopy: An International Journal, 27 (2): 93–105. DOI: 10.1155/2012/276795.

CHAPTER 2. DETECTION OF SIGNIFICANT WAVELENGTHS FOR IDENTIFYING AND CLASSIFYING *Fusarium oxysporum* DURING THE INCUBATION PERIOD AND WATER STRESS IN *Solanum lycopersicum* PLANTS USING REFLECTANCE SPECTROSCOPY

ABSTRACT

Spectroscopy has become one of the most used non-invasive methods to detect plant diseases before symptoms are visible. In this study, it was possible to characterize the spectral variation in leaves of *S. lycopersicum* L. infected with *F. oxysporum* during the incubation period. It was also possible to identify the relevant specific wavelengths in the range of 380-1000 nm that can be used as spectral signatures for the detection and discrimination of vascular wilt in *S. lycopersicum*. It was observed that inoculated tomato plants increased their reflectance in the visible range (Vis) and decreased slowly in the near infrared range (NIRs) measured during incubation, showing marked differences with plants subjected to water stress in the Vis/NIR. Additionally, three ranges were found in the spectrum related to infection by *F. oxysporum* (510nm-520nm, 650nm-670nm and 700-750nm). Linear discriminant models on spectral reflectance data were able to differentiate between tomato varieties inoculated with *F. oxysporum* from healthy ones with accuracies higher than 70% 9 days after inoculation. The results showed the potential of reflectance spectroscopy to discriminate plants inoculated with *F. oxysporum* from healthy ones as well as those subjected to water stress in the incubation period of the disease.

Keywords: plant diseases, vascular wilt, hyperspectral reflectance, relevant specific wavelength, band selection

2.1 Introduction

Remote sensing and nearly sensing methods, like multi or hyperspectral sensors, have been widely applied in agriculture, livestock, industries, and even some sectors of pharmaceutical and human medicine (Huang et al. 2007; Hatfield *et al.* 2008; Berzaghi and Riovanto 2010; Teixeira et al. 2013; Ciurczak and Igne 2014). Additionally, remote sensing technology provides an alternative method that is unbiased and automatic for visual evaluation of plant diseases (Mahlein et al. 2012), even in the early stages of evolution when the symptoms are not visible (Khaled *et al*. 2018).

After symptom expression of a plant disease, the disease can be verified through detection techniques. Detection techniques of plant diseases that are currently available can be divided into four groups: 1) serological methods: Flow Cytometry, Enzyme Linked Immunosorbent Assay (ELISA) and immunofluorescence, 2) "Purely molecular" methods: Fluorescent *In Situ* Hybridization (FISH), Polymerase Chain Reaction (PCR) and DNA Arrays, 3) Disease detection based on biomarkers: Profiles of Metabolites in the Gas Phase and Profiles of Plant Metabolites, 4) Disease detection based on plant properties and stress, which includes: imaging techniques (hyperspectral and fluorescence images) and spectroscopic techniques (Vis/NIR), infrared (NIR), fluorescence and multispectral bands (Sankaran et al.

2010). In recent decades, research in this last group of technology has led to the development of methods based on spectroscopy for the detection of diseases and stress in plants. These methods are faster, non-destructive, sensitive and selective for detection of disease during its incubation phase.

These techniques are based on measuring the amount of radiation reflected on a surface in function of the wavelengths to produce unique reflectance spectra for each material. These spectra can be used as a "fingerprint" (spectral signature), since they sense healthy and diseased plants in different states of evolution, even in cases when the symptoms of the disease are not yet visible (Zhang et al. 2003; Huang et al. 2004; Larsolle and Muhammed 2007; Mahlein et al. 2010; Marin-Ortiz et al. 2018). Characteristics of spectral radiation, reflected, transmitted or absorbed by the leaves, can provide a deep understanding of the histological, physiological and biochemical responses to growth conditions and adaptations of plants to the environment. However, because of increased interest in remote sensing, leaf reflectance has been studied more than absorbance and transmittance as stress responses in plants (Gregory et al. 2001).

The efficiency of measuring spectral reflectance for detection of diseases is based on the identification of the most significant spectral wavelength, which is highly correlated with a specific disease (Song et al. 2011; Mahlein et al. 2013), since it is found in only some regions of the electromagnetic spectrum of interest. Jacquemoud and Ustin (2001) divided the spectral range from 400 nm to 2500 nm into three large bands. The first is the Vis range (~380-750 nm), in which photosynthetic pigments have a greater impact on spectral signature, characterized by low reflectance values and a maximum peak located ~ 550 nm. The second one includes the near infrared plateau (750-1100 nm), a high reflectance spot due to multiple dispersion within the leaf in relation to the fraction of air spaces within the tissue (internal structure) and/or its water content. Finally, the third one includes near infrared (1100-2500 nm), which is a low reflectance zone, due to high absorption of water mainly in tissues such as fresh leaves and dry matter.

Significant wavelengths have been identified as a base to develop many Spectral Vegetation Indices (SVE) (Robert et al. 2011), as well as a method to detect and analyze changes in physiological and biochemical parameters in plants (Merzlyak et al. 2003a, 2003b; Gitelson et al. 2007). However, it is not possible to perform a quantitative definition or identification of a particular disease based on common SVEs because the method lacks specificity for diseases (Mahlein et al. 2013). Therefore, it is necessary to determine the Relevant Specific Wavelengths (RSW) for the construction of Spectral Disease Indices

(SDI) that can be used to simplify the detection of diseases by spectral sensors, since each disease uniquely influences the spectral signature in a characteristic way.

In this research, we conducted experiments under semi-controlled conditions to identify important spectral wavelengths for the early detection of *F. oxysporum* in tolerant and susceptible plants of *Solanum lycopersicum* L. This organism model is widely accepted for the study of pathogenicity in plants (Baayen et al. 2000). First, we carried out a characterization of spectral variation in leaves of *S. lycopersicum* infected with *F. oxysporum* during the incubation period of the disease and subjected to mild water stress. Then we identified RSW on 380-1000 nm that could be used for detection of spectral signatures in *S. lycopersicum* plants infected with *F. oxysporum* before the expression of visible symptoms. Finally, we tested or discriminated infected *S. lycopersicum* plants with *F. oxysporum* from healthy plants as a test of the RSW identified in the previous step.

2.2 Materials and Methods

2.2.1 Biological material

The plants used in this study were maintained under semi-controlled greenhouse conditions, located at the National University of Colombia Medellín (Antioquia, Colombia). Average temperatures ranged from 18 and 24°C, relative humidity between 60 and 70%, and there was a photoperiod of 12 hours during the experiments. In this study was used two tomato cultivars, Ponderosa, which is susceptible to all races of *F. oxysporum* (Reis and Boiteu 2007), and the Santa Cruz, which is resistant to races 1 and 2 were used. The seeds were planted in germination trays of 86 wells with sterile peat as a substrate and kept in the greenhouse for the duration of the respective experiments. The plants were irrigated daily, fertilized once a week with a nutritious solution (Annexed 4) and a protective action fungicide was applied every 7 days, starting when the plants were 7 days old, according to the management plan. After 4 weeks of germination, the inoculation procedure was carried out and individual plants were placed in 900 cm3 plastic cups containing the same substrate as that used to plant the seeds. In this study, five treatments were evaluated: (I) tomato plants var Ponderosa inoculated at 4 weeks after germinating with a pathogenic strain of *F. oxysporum* (Fo5), (II) healthy plants var Ponderosa submitted to hydric stress sustaining 60% of field capacity, (III) healthy plants (var. Ponderosa) and substratum maintained with 100% field capacity as control treatment, (IV) plants of tomato var. Santa Cruz infected with Fo5, (V) healthy plants (var Santa Cruz) and substratum maintained at 100% field capacity. The plants subjected to these treatments were kept under almost the same conditions as the greenhouse during

the rest of the experiment. Fertilization with nutritious substances was increased to twice every week. *F. oxysporum* Fo5 was isolated from *Passiflora edulis* (passionfruit), which is hyghly pathogenic on tomato plants (Marin et al. 2018). This strain has an incubation period of 24 dpi on tomato plants.

2.2.2 Inoculation

Four-week-old tomato plants inoculated according to the modified procedure described by Ortiz and Hoyos (2016) with a suspension of spores of isolate Fo5. Ten milliliters of spore suspension of *F. oxysporum* were prepared in distilled water. The concentration of the spore suspension used was $1x10^6$ spores/ml. The tomato seedlings were removed gently from the nurseries and the roots were washed with tap water to remove the remains of peat. Then wounds (cuts) were made on the secondary roots of all the plants with a sterile scalpel and only the roots were immersed in 15 ml of the solution of distilled water and spores for 20 minutes. The same procedure performed on the inoculated plants. The control plants inoculated only with distilled water, and subjected to water stress. The seedlings submitted to the different treatments transplanted to the vessels with sterile peat of 900 cm3. To verify the efficiency of the inoculation (postulate three of Koch), a cross section from the neck of roots from five plants of each treatment was made (days 0, 12 and 24 dpi). The tissue was disinfected and diluted in distilled sterilized water 1:10 (w/v). The homogenized solution (100 uL) was placed on PDA + malachite green oxalate supplemented with 200 ppm chloramphenicol. During the first 4 days after sowing, observations made under a microscope and the *F. oxysporum* colonies that grew in the medium counted according to the following formula:

$$\#\ \text{Conidia / gram of plant}\ \frac{Colonies}{Plantingvolume+Dilutionfactor}$$

2.2.3 Spectroscopy

For the acquisition of Vis/NIR reflectance spectra, a HR2000 portable spectroscope (Ocean Optics, USA) with a tungsten halogen light source HL-2000-HP (wavelength range of 360-2400 nm), a diffuse reflectance standard model WS-1 (reflectivity> 98% in the range of 250-1500 nm) and a 600µm premium grade reflectance probe QR600-7-VIS-125F were used. The measurements were made with the optical fiber attached to the adaxial surface of the leaf, in which five spectra acquisition were taken for each leaf. Different parameters required for spectrometer calibration, such as integration time, average readings per measurement and interval time were determined at the beginning of the experiments.

2.2.4 Pathogenicity test

Destructive samplings were executed to confirm plant infection at 24 days post infection in the following way: a cut of 1 cm was made from the neck of the stem of each plant and each segment was cut into five equal parts, approximately 2mm long each. Then, the segments disinfected in 70% alcohol, 2% sodium hypochlorite and water. Finally, the five cuts of the stem of each plant were placed in Petri dishes with PDA medium + 300 ppm of Gentamicin. Seven days after sowing the five stem segments, each Petri dish and stem segment observed for the presence and growth of the pathogen.

2.2.5 Data analysis

The results obtained in this work are presented as a function of the treatment realized in two tomato varieties infected with a strain of *F. oxysporum*, and subjected to water stress, and their respective controls, through the incubation period of the disease (before the appearance of visible symptoms). Each treatment has 30 plants and the same to control. A comparison of the effects caused by the treatments was carried out on leaves in the same stage of growth and development. Spectra acquisitions were carried out every three days from the infection of plants in five places of the second leaf of each plant. Due to this design, we collected six groups of data with 150 spectra each (30 leaves per treatment and control plant), for each sampling day (900 spectra/day of sampling), and 1800 spectra total. For the analysis, we use all the individual spectra from each sheet (It not their averages)

Initially, a spectra selection was made to remove noise, either by being deformed and/or with reading error. The spectra that showed very different patterns confirmed with analysis of "outliers" identified in a Principal Component Analysis (PCA) without prior data treatment. After the elimination of the spectra with noise and previous treatments, several types of transformations applied to reduce the impact of the difference in illumination, cultivar of the plant or specific effects of the sensor. After several analyses, the standard normal variate transformation (SNV) was chosen as the best pre-treatment that allow a good grouping of the plants through the treatments, reducing spectral noise and eliminate background effects of Vis/NIR data. After performing the pre-treatment of raw data, an analysis of the variance was made from the absolute differences between the reflectance means of the plants of *S. lycopersicum* (two varieties) subjected to biotic stress (infected with *F. oxysporum*) and abiotic (water stress) with healthy plants, and standard deviations of reflectance. Subsequently, a binary classification of healthy leaves and diseased plants was made to test the detection and the later classification of disease by RSW. To reduce the information of the measured spectrum and obtain these RSW to

separate diseased and healthy plants, the RELIEF algorithm was used (Robnik-Šikonja and Kononenko 2003). Finally, the RSW identified in the previous step were used to perform Linear Discriminant Analysis (ADL) in order to characterize or separate tomato plants subjected to biotic stress (infected with Fo5) and abiotic (hydric stress) during the incubation period of disease. All statistical analyses were performed using R Software. The main libraries and functions used summarized in Table 1.

Table 1. Description of the main analyses used with the R software

Analysis	Library	Function	Description
Pre processing	mdatools, prospectr	prep.snv, gapDer	Standard normal variate (SNV) and second derivatives of different orders
Detection and elimination of extreme points	mvoutlier, rrcov	pcout, PcaHubert	Fast algorithm to identify multivariate outliers in high dimensionality data, using the Filzmoser algorithms (Maronna and Werner 2007), and ROBPCA (Hubert *et al*. 2005)
Data visualization	mdatools, ggplot2	mdaplot, ggplot, plot	Functions used for data visualization (scatter plots, bars, histograms)
Selection of LOE	dprep	relief	This function implements the RELIEF feature selection algorithm (Kira and Rendel 1992)
Discrimination of treatments	MASS	lda, predict	Functions used to calculate linear discriminant analyses and matrices of confusion

2.3 Results

2.3.1 Variation of spectral signatures

The Figure 1 show Changes in treatments analyzed as absolute differences between statistical average reflectance of infected plants and subjected to biotic stress, less the reflectance of no infected plants. Difference in reflectance increased with *F. oxysporum* colonization on two tomato cultivars evaluated (380-750 nm) 21-24 days post infection (dpi). On these days, the susceptible cultivar displayed changes at same time as visible alterations or symptoms occurred, while the tolerant plants revealed slight differences in spectra, and there were no visible symptoms of disease. In the infrared (750-1000 nm) the reflectance showed a sustained increase in the difference of the tolerant cultivar with their respective healthy controls until 21 dpi, decreasing markedly 24 dpi. In the susceptible cultivar, there was an increase 12 dpi and, then fell. Taking into account the standard deviations, it can be affirmed that only the mean varied, but there was no disparity in the dispersion of the data in each treatment. The highest differences on reflectance were on limit of red (750 nm), usually after the first week of infection. Besides, plants with water stress displayed different patterns, with a lack of reflectance from the first

week on visible and NIR, 750-1000 nm being the region with a high difference of reflectance, during which the assessment standard deviation was kept constant.

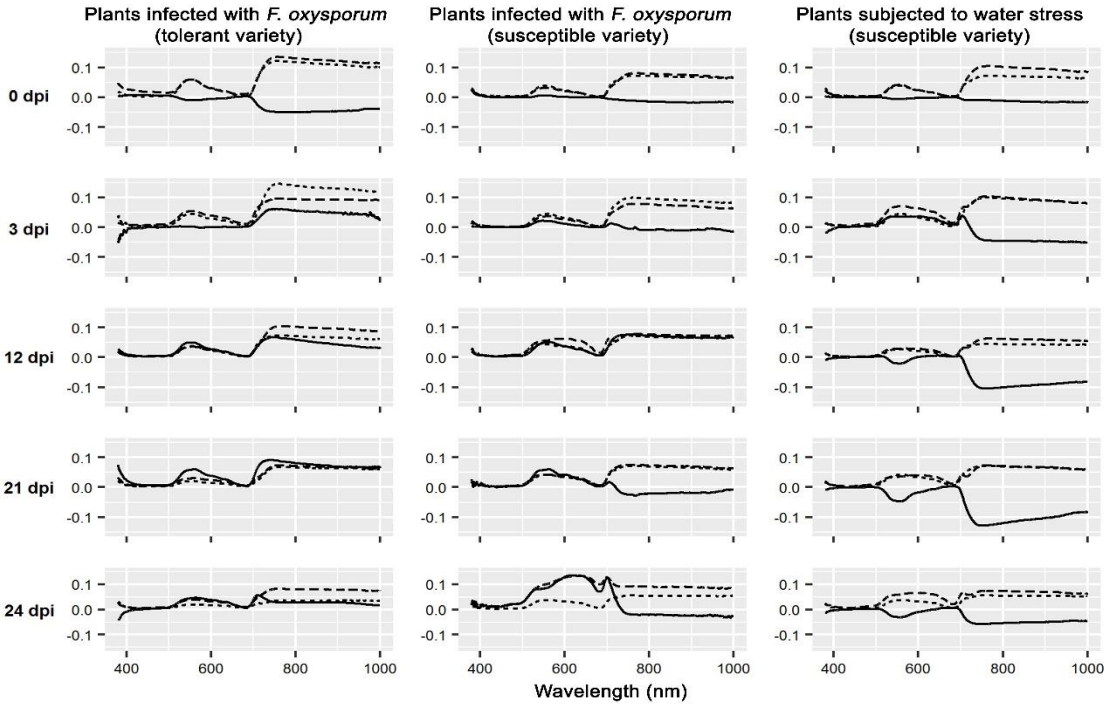

Figure 1. Differences between the average reflectance (____) and standard deviation of tomato leaves subjected to both types of stress compared to healthy controls: infected with *F. oxysporum* and subjected to water stress (--- standard deviation on control plants, _ _ _ standard deviation on treatment).

2.3.2 RSW for detection and classification of plants subjected to two types of stress

Since vascular wilt is a systemic disease in tomato plants, a binary classification of healthy leaves and diseased plants was made to test the detection and subsequent classification of the disease by SRW. To reduce the information of the measured spectrum and obtain these RSW the RELIEF algorithm was used to separate diseased and healthy plants (Fig. 2).

On day 0 (plants without infection and without water stress) weights for wavelengths were constant and close to "0" for the two types of stress evaluated in the measured range, without highlighting any relevant RSW (Fig. 2A and 2F). The specific wavelengths relevant to infection by *F. oxysporum* (biotic stress) were 510-520 nm, 650-660 nm and 750 nm, clarifying that the ranges of 510-520 nm, 650-660 nm were relevant from 3 dpi (Fig. 2B), while the wavelength of 750 nm began to be important

approximately 9-12 dpi (Fig. 2C and 2D). No RSW was observed in plants subjected to biotic stress in the infrared range measured (750-1000 nm).

The scores of these RSW increased with the period of incubation of the disease in the plant, but decreased markedly in the susceptible cultivar to 24 dpi (their weights tended to be constant), when the symptoms were observed on them, whereas the RSW in the tolerant plants continued with the same pattern observed during the previous days (Fig. 2E). The RSW for plants subjected to water stress were 750 nm and the range 900-950 nm, was observed after 18 dpi (Fig. 2I and 2J).

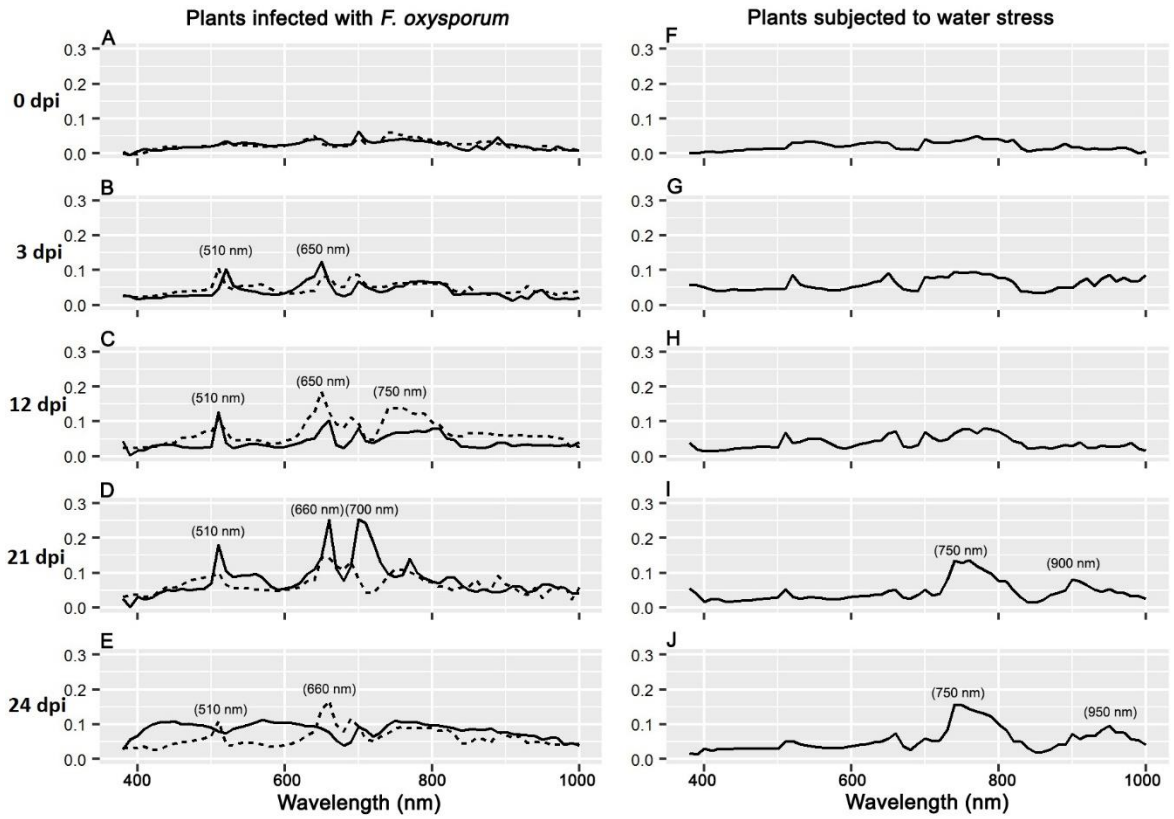

Figure 2. Relevant specific wavelengths for the two tomato varieties during the incubation period of *F. oxysporum* infection (A-E): 0 dpi (A), 3 dpi (B), 12 dpi (C), 21 dpi (D) and 24 dpi (E); Tolerant cultivar (dashed line) and susceptible cultivar (solid line). The LOER for the susceptible cultivar of tomato subjected to water stress illustrated in F-J.

The RSW with scores greater than 0.1 that selected after analysis with the RELIEF algorithm are summarized in Table 1.

Table 2. Relevant specific wavelengths (nm) selected for the detection and classification of plants of both tomato varieties (tolerant and susceptible) subjected to two types of stress, biotic (*F. oxysporum* - Fo5, during the incubation period) and water stress (ws).

	Biotic Stress (Fo5)		Abiotic Stress (ws)
Dpi	Tolerant	Susceptible	Susceptible
0	-	-	-
3	510	520, 650	-
6	510, 550, 710	520, 650	-
9	650, 760	520, 650	-
12	510, 650, 750, 880	510, 660	-
15	650, 740, 890	510, 650	-
18	-	-	750, 900
21	510, 650, 770	510, 660, 710	750, 950

2.3.3 Classification of tomato plants infected with *F. oxysporum*

Table 3 summarizes classification percentages of plants compared to their healthy controls during the incubation period of disease; this figure summarizes confusion tables in Linear Discriminant Analysis (LDA) using the RSW selected with the RELIEF algorithm. The tolerant cultivar of tomato infected with *F. oxysporum* exhibited a progressive increase in the classification percentage compared to its controls since infection, reaching the highest value (92.76%) 12 dpi and maintained little variation until 24 dpi. Otherwise, susceptible tomato plants showed intermediate values of correct classification, between 68% and 72% from 3 to 15 dpi, respectively, abruptly increasing on day 18 from coincident at the same time as early visual symptoms with early visual symptoms. Susceptible plants subjected to water stress showed a pattern similar to the one described above, maintaining constant classification percentages between 71.00 and 76.00%, but increasing rapidly on day 15 and reaching 93.00% to 21 dpi.

Table 3. Temporal variation on classification percentage for plants subjected to biotic stress (Fo5) and abiotic stress (ws)

	Days post infection							
Cultivar/stress	**0**	**3**	**6**	**9**	**12**	**15**	**18**	**21**
Tolerant / infected with *F. oxysporum*	60.5	70.0	80.3	87.5	92.7	92.3	86.0	90.2
Susceptible / infected with *F. oxysporum*	62.4	68.0	68.9	69.8	69.4	72.1	70.9	90.0
Susceptible / subjected to water stress	64.5	71.2	74.0	75.0	75.7	72.0	89.6	93.00

Histograms for each group in the first discriminant dimension were made to visualize post-infection during 24 dpi, when the separation of the groups of plants can be observed compared to their respective controls (Fig. 3). Tolerant tomato plants infected with *F. oxysporum* are clearly discriminated from control plants from 12 dpi (Fig. 3A). In addition, the plants of the susceptible cultivar achieved an acceptable classification after 12 dpi (Fig. 3B), although it was appreciably lower than the previous treatment. Regarding plants subjected to water stress, it should be noted that the water deficiency to which they were subjected was slight (field capacity at 60%), so a classification percentage >85% was observed for 18 and 21 dpi (Fig. 3C).

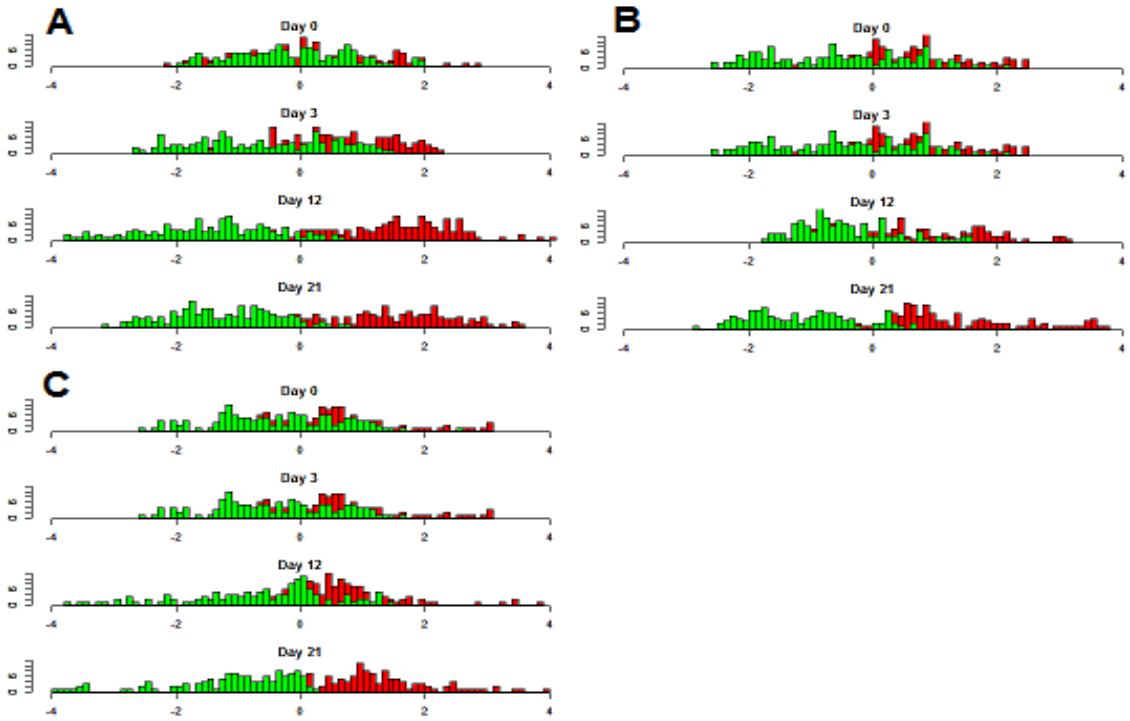

Fig. 3. Histograms for the first linear discriminant dimension of two tomato varieties infected with *F. oxysporum* -tolerant (A), susceptible (B) -, and subjected to water stress (C), on different post infection days. Red: Plants subject to biological or water stress; green: control plants (Wavelengths, selected with RELIEF algorithm, were 510nm, 650nm and 750 nm on data from plants infected with *F. oxysporum*; 750nm and 950nm for data on plants subjected to water stress).

2.4 Discussion

Spectral reflectance analyses are useful for detecting different types of biotic and abiotic plant stress due to changes in the light absorption incident in the Vis/NIR range of the electromagnetic spectrum (Sankaran et al. 2010; Khaled et al. 2014). Additionally, the relationships between the physiological, histological and biochemical changes generated in the plant by different pathogens affect the spectral

characteristics and can be detected before the expression of symptoms in the Vis/NIR regions. Currently there are few studies dedicated to the *S. lycopersicum-F. oxysporum* pathosystem specific using the reflectance spectroscopy technique focused on the early detection of the disease (Salman et al. 2012; Abu-Khalaf 2015; Marín et al. 2018). However, these did not investigate in depth early detection during the incubation period or the search for LOER related with disease.

The data variance set analyzed from the absolute differences between the reflectance means of *S. lycopersicum* plants subjected to stress by *F. oxysporum* (biotic) and water stress (abiotic) with healthy plants agrees with the theoretical basis proposed by Zhang *et al*. (2003). This is a basis for the use of spectroscopy in the discrimination of different forms of stress in plants, which suggests that the wave magnitude will typically vary in different lengths, increasing the reflectance in the Vis range and decreasing in the NIR (750-1100 nm) on plants infected with pathogens. This is possible due to the specific interaction of the light with different photosynthetic pigments of the plant. In the Vis light range, there are mainly transition processes, in which the atoms of the pigments absorb visible photons of light and can excite one of their electrons at higher energy levels. The region of the infrared (IR) generates a specific vibration of the molecules chemical bonds , and as the energy to excite an electron in a specific atom is constant, the energy to change the vibration of a particular chemical bond is also fixed (Weiner and Ping-Tong 2003). As each link is different, each one has a different way of vibrating, varying the percentage of absorbance and reflectance in each wavelength, so it is possible to carry out characterizations about of organic structure in different applications.

During the incubation period of the disease the difference of reflectance values in tomato varieties evaluated (tolerant and susceptible) fluctuated between 380 and 750 nm compared to the positive controls (Dif = $\lambda_{Fo5} - \lambda_{CON}$), indicating higher values in the reflectance of infected plants in this spectral range. This increase in reflectance in the Vis range (decrease in absorbance) suggests that the content of the different photosynthetic pigments is lower in leaves of infected plants rather than the healthy ones (Carter and Knapp 2001). This fact is related to physiological responses against stress produced by reported pathogens (Berger *et al*. 2007). On the other hand, the small differences between plants infected with Fo5 and their controls, in the range of 750-1000 nm, may indicate a minor disturbance in the hydric state of the leaves infected (Genc et al. 2013; Jin et al. 2077). These results contrast with the strong decrease in reflectance in plants subjected to water stress in the same range, from the start of data collection (day 3 of the beginning of the stimulus) to the death of the leaf (day 24), suggesting a strong relationship between water reduction in leaf tissues and wavelengths in this spectral range.

Additionally, plants with an increasing level of water deficiency showed a decrease in the reflectance, which has also been observed in plants with a decrease in the relative water content of the leaf and response to some organic compounds (Zhang et al. 2012; Genc et al. 2013).

This study provides evidence regarding the finding of specific wavelengths more relevant to vascular wilt in tomato plants caused by *F. oxysporum* that can help to improve the detection and discrimination of asymptomatic infected plants and water stress from hyperspectral data. These RSW are characterized by having high sensitivity and specificity in the pathosystem studied and can be used in the future to make definition indices that incorporate two or three bands of the spectrum in the Vis/NIR range. In general, the method most commonly used to detect RSW for the development of disease indices is by correlation with biophysical or biochemical traits (Hatfield *et al*. 2008). However, the use of the RELIEF feature selection algorithm offers many advantages, since it works with non-linear classes and efficiently separates the classes by performing an iterative process. In this process, each iteration of a "X" instance of the data set is chosen randomly and the weight of each characteristic becomes updated according to the distance from "X" to the nearest instance of the same class ("NearHit"), and in turn to the nearest instance of a different class ("NearMiss"). Finally, all the data are added in a class (Kira and Rendel 1992; Kononenko et al. 1997).

The detection of a specific relevant wavelength range around 510-520 nm for tomato vascular wilt caused by *F. oxysporum* coincides with the maximum absorbance peak of the carotenoids (Zur et al. 2000; Merzlyak et al. 2003). In plants, carotenoids fulfill different functions, mainly as light-gathering molecules and photoprotection (Demmig-Adams *et al*. 1996). Additionally, others studies show that carotenoids play a key role in the adaptation of plants to mild stress and other unfavorable factors (Strzalka *et al*. 2003). The majority of specific wavelengths relevant to the pathosystem *S. lycopersicum*-*F. oxysporum* was found in the range of 650-750nm, which is an absorption region for groups of molecules that have oxygen, mainly water and free -OH alcohol (possibly phenols). The phenols are the result of the secondary metabolism of plants, crucial for the functional aspects in the life of the same, with different functions as protector against pests, environmental stress and pathogens. This result can be related with an importance of the reflectance near 700 nm as a fundamental characteristic of green vegetation produced by a balance between biochemical and biophysical characteristics of plants (Gitelson and Merzlyak 1996, 1997). It has been observed that the displacement towards the blue of the red edge of the reflectance curve frequently accompanies the stress generated by pathogens in plants, whereby it could be used in the early detection of diseases, since an increase in reflectance around 700

nm can be a first indicator to detect cultures infected by pathogens. However, this relevant wavelength (700nm) is not specific to a disease, since in plants there can be an overlap with important nearby wavelengths, like 680nm, which is related to the chlorophyll content. With respect to water stress, there are decreases in reflectance in the Vis/NIR, which becomes more evident after the 15th day, considering that the plants were subject to 60% field capacity, which can be considered as a slight stress in the state of development of the plants. These RSW were present in the near-infrared region in which the main information is generated regarding the water absorption of the leaf (760-1100), with peaks in the ranges 750-760nm and 900nm-960nm. Water absorption characteristics, as a result of absorption by O-H bonds, can be found at approximately 760 nm, 970 nm, 1200 nm, 1450 nm and 1950 nm (Li et al. 2006).

It is important to note the lower magnitude on the reflectance spectra obtained in the plants of the tolerant cultivar. These results in the region suggest greater photosynthesis and subsequent synthesis of different types of polysaccharides compared to susceptible plants. In vitro studies from the 1980s and 1990s showed that tolerant phenotypes infected with *F. oxysporum* generated high contents of polysaccharides and callose, and induction of peroxidase, phytoalexins synthesis and inhibition of pathogens in dual crops (Storti et al. 1989). In contrast, moderately tolerant phenotypes had lower polysaccharide content and showed no hypersensitivity reaction when treated with the pathogen. The authors proposed that the presence of high levels of polysaccharides in incompatible interactions generally should be considered as evidence of direct inhibition of the fungus by these compounds and with their recognition by the plant, limit their defensive factors.

Our results support the hypothesis that the differences in spectral responses during the incubation period of the disease of evaluated cultivars (susceptible and tolerant) are due to physiological changes generated in the plant-pathogen recognition process and the generation of polysaccharides important to inhibiting the pathogen. These changes at different times of the incubation period may cause differences at the time that make it possible to discriminate each cultivar with percentages of classification greater than 80%: 9-12 dpi (tolerant) and 18-21dpi (susceptible), under these particular test conditions. The colonization of susceptible and tolerant plants is systemic and similar in terms of the amount of inoculum used for both. In tolerant plants, pathogen recognition occurs quickly and therefore important compounds are synthesized to suppress growth and spread of the pathogen. In contrast, susceptible plants have delayed responses. The plants respond to the invading pathogen with physical barriers, producing depositions in the cell walls, blockages of the xylem vessels, and by chemical defense, synthesizing antimicrobial substances (Fradin et al. 2006; Cregeen et al. 2015). The different

physical and chemical responses to the pathogen by the susceptible and tolerant varieties generate spectral changes that can be detected in real time with spectroscopic techniques and different types of multivariate analysis.

The ability to identify healthy tomato plants and those infected with *F. oxysporum*, or subjected to water stress with RSW seems to have been demonstrated in this work, although it is important to remember that other factors could also have an impact on the development of the disease, since the plants were maintained under semi-controlled conditions. There were varying percentages of success in the classification by increasing time after infection of the plant. Values between 85% -93% were reached in the varieties evaluated (although at different dpi). Previous studies on the detection and classification of plant diseases using Vis/NIR spectroscopy and different multivariate analysis techniques (including the Linear Discriminant Analysis, Partial Least Squares and Regression by Main Components) have shown percentages in the classification accuracy greater than 80% in a wide cultivar of pathosystems, such as Wheat-Yellow Rust (90.0%), Cotton-Verticillium (82.4%), GLAVV-Vid-virus (81.0%), Tomato-*F. oxysporum* (85.0%-100%), Tomato-*Ralstonia solanacearum* (85.0%), Palm oil-*Ganoderma Boninense* (92%), Sugar beet-*Uromyces betae* (80.3%), Sugar beet-Cercospora (85%), Citron-*Candidatus liberibacter* americanus (Lam) (80%-90%), among others (Mahlein et al. 2013; Abu-Khalaf 2015; Alfadhi et al. 2017; Marin-Ortiz et al. 2018).

Even though of an appreciable amount of research has focused on the detection and classification of plant diseases using reflectance spectroscopy in the Vis/NIR as well as on the use of multivariate techniques for the analysis of high dimensionality data matrices, more detailed research is needed in the search for WSR. These subsequent specific indexes and analysis of data could be used for detection and early discrimination of systemic diseases in plants.

2.5 Conclusions

Plants of *S. lycopersicum* infected with *F. oxysporum* presented a clear spectral response compared to their respective controls, increasing their reflectance in the Vis and decreasing slowly in the NIRs measured (750nm-1000nm) during the incubation period of the disease. The tomato varieties evaluated (tolerant and susceptible) presented the same pattern of response in the the part of Vis/NIR range evaluated, but with a delay of the tolerant cultivar, mainly in regard to the decrease of reflectance on the measured infrared region. Traditionally vascular wilt has been related to the death of the plant due to hydric stress, caused by the plugging of vascular bundles, which prevents the flow of water in the

plant and causes its death in advanced stages of the disease. These results showed marked differences in the plants subjected to water stress in the Vis/NIR, which suggests that there are different physiological and structural response mechanisms to the two types of stress during the incubation period in which the symptoms are not visible.

The RSWs related to infection by F. oxysporum were found in the Vis range, which suggest changes in photosynthetic pigments in plants as a response to the pathogen, probably by variation in carotenoids (510nm-520nm range), Chlorophyll a (650nm-670nm range), and some groups of molecules that have oxygen, mainly water and free -OH alcohol (700-750nm range). Otherwise, the RSW related to water stress which were found (750nm, 900-960nm) were in the near-infrared range measured, in which the main information is generated regarding water absorption of the leaf (760-1100), suggesting high specificity and sensitivity to detect and discriminate *F. oxysporum* infection from hydric stress in tomato plants in the asymptomatic stage of the disease. However, it is important to highlight the importance of performing comparative studies with specific indices developed from RSW for different diseases and other indexes proposed in current literature, in order to evaluate the specificity and sensitivity of the wavelengths found in each type of infection.

The detection of the disease in tomato plants had a correct classification greater than 70%. Linear discriminant models on spectral reflectance data were able to classify plants infected with *F. oxysporum* from no infected ones with high precision (85% -93%), due to minor changes in the reflectance of diseased leaves at this stage. This study showed that the discrimination of systemic diseases in early infection stages is possible, but remains a challenge. Therefore, future research is required to provide additional information about factors that affect the spectral response in plants, such as differences between plant varieties, responses to various environmental conditions and nutritional considerations.

References

Abu-Khalaf N. 2015. Sensing tomato's pathogen using Visible/Near infrared (VIS/NIR) spectroscopy and multivariate data analysis (MVDA). Palestine Technical University Research Journal, 3(1): 12-22.

Baayen R.P., O`Donnell K., Bonants P.J.M., Cigelnik E., Kroon L.P.N.M., Roebroeck J.A., waalwijk C. 2000. Gene genealogies and AFLP analysis in the *Fusarium oxysporum* complex identify monophyletic and nonmonophyletic *formae especiales* causing wilt and rot disease. Phytopathology, 90(8): 891-900. DOI: 10.1094/PHYTO.2000.90.8.891

Berger S., Sinha A.K., Roitsch T. 2007. Plant physiology meets phytopathology: plant primary metabolism and plant–pathogen interactions. Journal of Experimental Botany, 58(15-16):4019-4026.

Berzaghi P., Riovanto R. 2010. Near infrared spectroscopy in animal science production: principles and applications. Ital.J.Anim.Sci. 8 (3). 39-62. Pharmaceutical and Medical Applications of Near-Infrared Spectroscopy. Impresión CRC, Segunda Edición. ISBN 9781420084146.

Carter G.A. 1994. Ratios of leaf reflectances in narrow wavebands as indicators of plant stress. International Journal of Remote Sensing, 15(3): 697-703. DOI: 10.1080/01431169408954109

Carter G.A., knapp A.K. 2001. Leaf optical properties in higher plants: linking spectral characteristics to stress and chlorophyll concentration. American Journal of Botany, 88(4): 677–684.

Cregeen S., Radisek S., Mandelc S., Turk B., Stajner N., Jakse J., Javornik B. 2015. Different Gene Expressions of Resistant and Susceptible Hop Cultivars in Response to Infection with a highly Aggressive Strain of Verticillium albo-atrum. Plant Molecular Biology Reporter, 33(3):689–704. DOI: 10.1007/s11105-014-0767-4

Demmig-Adams B., Gilmore A.M., Adams W.W. 1996. In vivo functions of carotenoids in higher plants. The FASEB Journal, 10(4):403-412. DOI: 10.1096/fasebj.10.4.8647339

Fradin E.F., Thomma B. 2006. Physiology and molecular aspects of Verticillium wilt diseases caused by V-dahliae and V-alboatrum. Molecular Plant Pathology, 7(2):71–86. DOI: 10.1111/j.1364-703.2006.00323.x

Genc L, Inalpulat M, Kizil U, Mirik M, Smith S, Mendes M. 2013. Determination of water stress with spectral reflectance on sweet corn (*Zea mays* L.) using classification tree (CT) analysis. Zemdirbyste-Agriculture, 100(1): 81–90. DOI: 10.13080/z-a.2013.100.011

Gitelson A.A., Zur Y., Chivkunova O.B., Merzlyak M.N. 2007. Assessing Carotenoid Content in Plant Leaves with Reflectance Spectroscopy. Photochemistry and Photobiology, 75(3): 272-281. DOI: 10.1562/0031-8655

Gitelson A.A, Gritz Y., Merzlyak M. 2003. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. J. Plant Physiol., 160:271-282. DOI: 10.1078/0176-1617-00887

Gitelson A.A., Merzlyak M.N. 1996. Remote estimation of chlorophyll content in higher plant leaves. Journal of Plant Physiology, 148(2): 494–500. DOI: https://doi.org/10.1080/014311697217558

Gitelson A.A., Merzlyak M.N. 1997. Signature analysis of leaf reflectance spectra: Algorithm development for remote sensing of chlorophyll. International Journal of Remote Sensing, 148(3-4): 2691–2697. DOI: https://doi.org/10.1016/S0176-1617(96)80284-7

Hatfield L.J., Gitelson A. A., Schepers S. J., Walthall L. C. 2008. Application of spectral remote sensing for agronomic decisions. Agronomy Journal, 100(3): 117-131. DOI: 10.2134/agronj2006.0370c

Huang MY, Huang WH, Liu LY, Huang YD, Wang JH, Zhao CH, Wan AM. 2004. Spectral reflectance feature of winter wheat single leaf infested with stripe rust and severity level inversion. Transactions of the CSAE, 20(1): 176-180.

Huang H, Yu H, Xu H, Ying Y. 2007. Near infrared spectroscopy for on/in-line monitoring of quality in foods and beverages: A review. Journal of Food Engineering, 87(3): 303-313. DOI: 10.1016/j.jfoodeng.2007.12.022

Jin X., Shi C., Yu Y., Yamada T., Sacks E.J. 2017. Determination of Leaf Water Content by Visible and Near-Infrared Spectrometry and Multivariate Calibration in *Miscanthus.* Front Plant Sci., Published online, 8-721. DOI: 10.3389/fpls.2017.00721

Khaled A.Y., Aziz S.A, Bejo S.K., Nawi N.M., Nawi N.M., Seman I.A. Onwude D.I.2018. Early detection of diseases in plant tissue using spectroscopy – applications and limitations. Applied Spectroscopy Reviews, 53(1): 36-64, DOI: 10.1080/05704928.2017.1352510.

Kira K., Rendel L. 1992. The Feature Selection Problem: Traditional Methods and a new algorithm. Proc. Tenth National Conference on Artificial Intelligence, MIT Press, 129-134.

Kononenko I., Simec E., Robnik-Sikonja, M. (1997). Overcoming the myopia of induction learning algorithms with RELIEFF. Applied Intelligence Vol7, 1, 39-55.

Larsolle A, Muhammed HH. 2007. Measuring crop status using multivariate analysis of hyperspectral field reflectance with application to disease severity and plant density. Precision Agriculture, 8 (1–2): 37–47. DOI: 10.1007/s11119-006-9027-4

Li M.Z. 2006. Spectroscopy Analysis Technology and Application, Science Press, Beijing, China.

Mahlein AK, Steiner U, Dehne HW, Oerke EC. 2010. Spectral signatures of sugar beet leaves for the detection and differentiation of diseases. Precision Agriculture, 11(4):413-431. DOI: 10.1007/s11119-010

Mahlein A.K., Rumpf T., Welke P., Dehne H.W., Plümer L., Steiner U., Oerke E.C. 2013. Development of spectral indices for detecting and identifying plant diseases. Remote Sensing of Environment, 128: 21-30. DOI: 10.1016/j.rse.2012.09.019

Marín-Ortiz J.C., Hoyos-Carvajal L.M., Botero-Fernández V. 2018. Detection of asymptomatic *Solanum lycopersicum* L. plants infected with *Fusarium oxysporum* using reflectance VIS spectroscopy. Colombian Journal of Horticultural Sciences, 12(2): 436-446. DOI: 10.17584/rcch.2018v12i2.7293

Merzlyak M.N., Solovchenko A.E., Gitelson A.A. 2003a. Reflectance spectral features and non-destructive estimation of chlorophyll, carotenoid and anthocyanin content in apple fruit. Postharvest Biology and Technology, 27(2):197-211. DOI: 10.1016/S0925-5214(02)00066-2

Merzlyak M.N., Gitelson A.A., Chivkunova O.B., Solovchenko A.E., Pogosyan S.I. 2003b. Application of Reflectance Spectroscopy for Analysis of Higher Plant Pigments. Russian Journal of Plant Physiology, 50(5):704-710.

Roberts D., Roth K., Perroy R. 2011. Hyperspectral Vegetation Indices. Hyperspectral Remote Sensing of Vegetation. DOI: 309-327. 10.1201/b11222-20.

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. DOI: https://doi.org/10.1016/j.compag.2010.02.007

Salman A, Lapidot I, Pomerantz A, Tsror L, Hammody Z, Moreh R, Huleihel M, Mordechai S. 2012. Detection of *Fusarium oxysporum* Fungal Isolates Using ATR Spectroscopy. Spectroscopy: An International Journal, 27(5-6): 551–556. DOI: 10.1155/2012/109708

Song S., Gong W., Zhu B., Huang Xi. 2011. Wavelength selection and spectral discrimination for paddy rice, with laboratory measurements of hyperspectral leaf reflectance. ISPRS Journal of Photogrammetry and Remote Sensing, 66(5):672–682. DOI: 10.1016/j.isprsjprs.2011.05.002

Storti E, Bogani P, Bettini P, Bonzi Morassi L, Pellegrini MG, Matteo M, Simeti C, Buiatti M. 1989. The pleiotropic phenotype of tomato cells selected for altered response to *Fusarium oxysporum* f. sp. *lycopersici* cell wall components. Theoretical and Applied Genetics, 78(5):689-695.

Storti E., Latil C., Salti S, Bettini P, Bogani P, Pellegrini M.G., Simeti C., Molnar A., Buiatti M. 1992. The in vitro physiological phenotype of tomato resistance to *Fusarium oxysporum* f. sp. *Lycopersici.* Theoretical and Applied Genetics, 84(1-2):123-128. DOI: 10.1007/BF00223991

Strzalka K., Kostecka-Gugala A., Latowski D. 2003. Carotenoids and Environmental Stress in Plants: Significance of Carotenoid-Mediated Modulation of Membrane Physical Properties. Russian Journal of Plant Physiology, 50(2):168-173. DOI: 10.1023/A:1022960828050

Teixeira C.A., LOPO M, PASCOA R, LOPES J. 2013. A Review on the Applications of Portable Near-Infrared Spectrometers in the Agro-Food Industry. Applied Spectroscopy, 67(11): 1215-1233. DOI: 10.1366/13

Weiner J., Ping-Tong H. 2003. Light-matter interaction. 1. Fundamentals and applications. Volume 1 edition. Hoboken, New Jersey : Wiley

Zhang M., Qin Z., Liu X., Ustin S.L. 2003. Detection of stress in tomatoes induced by late blight disease in California, USA, using hyperspectral remote sensing. International Journal of Applied Earth Observation and Geoinformation, 4(4): 295–310. DOI: 10.1016/S0303-2434(03)00008-4

Zhang Q, Li Q, Zhang G. 2012. Rapid Determination of Leaf Water Content Using VIS/NIR Spectroscopy Analysis with Wavelength Selection. Spectroscopy: An International Journal, 27(2): 93–105.

Zhang J., Pub R., Huanga W., Yuana L., Luoa J., Wanga J. 2012. Using in-situ hyperspectral data for detecting and discriminating yellow rust disease from nutrient stresses. Field Crops Research, 134:165-174. DOI: 10.1016/j.fcr.2012.05.011

Zur Y, Gitelson A.A., Chivkunova O.B., Merzlyak M.N. 2000. The spectral contribution of carotenoids to light absorption and reflectance in green leaves. Papers in Natural Resources. 272.

CHAPTER 3. LINKING PHYSIOLOGICAL PARAMETERS WITH VISIBLE/NEAR-INFRARED LEAF REFLECTANCE IN THE INCUBATION PERIOD OF VASCULAR WILT DISEASE

ABSTRACT

The photosynthetic pigments are mainly responsible for absorbing the light intended to promote photosynthesis on the chloroplast of the leaves. Different studies have related the spectral response in the leaves of plants with the biotic stress generated by pathogens. In general, maximum differences in reflectance have been found in the range of 380-750nm between plants subjected to biotic stress and healthy plants. In this study, it was possible to characterize and relate the spectral variance in leaves of *S. lycopersicum* infected with *F. oxysporum* with this physiological variation and pathogen concentration in tomato plants during the asymptomatic period of vascular wilt. Photosynthetic parameters derived from gaseous exchange analysis in the tomato leaves correlated related with four bands in the visible range (Vis). Additionally, five specific bands also present a high correlation with the increase in the concentration of *F. oxysporum* conidia measured at the root: 448-523nm, 624-696nm, 740-960nm, 973-976nm, and 992-995nm. These wavelengths allowed a 100% correct classification of the plants inoculated with *F. oxysporum* from the plants subjected to hydric stress and the control plants in the asymptomatic period of the disease. The spectral response to biotic and abiotic stress in the measured Vis/NIR range can be explained by the general tendency to change the concentration of chlorophyll and carotene in tomato leaves. These studies also highlight the importance of the implementation of robust multivariate analysis over the multiple univariate analysis used in the applied biological sciences and specifically in the agricultural sciences. These results demonstrate that specific wavelength responses are due to physiological
changes in plants subjected to stress, and can be used in indexes and algorithms applied to the early detection of diseases in plants on different pathosystems.

Keywords: plant diseases; vascular wilt; reflectance spectroscopy; multivariate analysis; early detection; plant physiology.

3.1 Introduction

Fusarium oxysporum (Schltdl. 1824) is a widely accepted model organism in studies for plants pathogenicity and one of the most morphologically diverse and widely distributed species of the genus *Fusarium* spp., currently considered as a complex of plant pathogenic fungal species (Baayen et al., 2000). This species has many pathogenic strains that parasitize more than 100 plant species named formae specialis, causing diseases in specific hosts that have economic importance such as tomato, banana, and cotton (Bosland, 1998).

This fungus is a hemibiotrophous pathogen (Chen et al., 2014). In the biotrophic phase. The fungus grows on its host tissues, with the recognition, deposition or contact, adaptation, and inoculation or penetration of the infective unit or inoculum as a preliminary stage. Subsequently, colonization occurs in a tissue recognized as a susceptible host. The time elapsed between inoculation and the moment of

symptomatic expression in the host is called the incubation period, this phase causes difficulties on detection of diseases caused by pathogens, since they are asymptomatic in its hosts for a period of time without causing visual changes, being foci of infection in natural ecosystems or crops. Once typical symptoms appear, observation is the traditional way to detect them. During the asymptomatic stage (incubation period) there are several methods that are currently used for the detection of diseases in plants, such as the Enzyme-Linked Immunoassay (ELISA) and the Polymerase Chain Reaction (PCR). These generally take time, resources and destroy the plant. Despite the availability of these techniques, a system of early detection of the disease can help reduce losses caused in crops and prevent a greater spread of the disease. A fast and reliable method is needed, with the sensitivity, selectivity and that does not require the destruction of samples.

Indicators of biotic and abiotic stress commonly use the spectral quality of the light absorbed, reflected and transmitted by the plant leaves (Sankaran et al., 2010). More specifically, in the past 20 years, important advances focused on the application of spectroscopy to the early detection of diseases in plants (Khaled et al., 2017). However, as a result of increased interest in remote sensing, the study of reflectance deepens more than the absorbance and transmittance in the last decades as responses to stress in plants (Gregory et al., 2001). Additionally, the spectral characteristics of the radiation reflected by the leaves can provide a frame to physiological responses on plants with different types of pathogens (Carter and Knapp, 2001).

In the latest decades, the development of new equipment and techniques in the application of reflectance spectroscopy in early detection of plant diseases also intensified in terms of searching for spectral characteristics related to physiological responses (Khaled et al., 2017). These techniques are based on measuring the amount of radiation reflected by a surface as a function of the wavelengths to produce a unique reflectance spectrum for each material, which can be used as a "fingerprint" (spectral signature) that allows to detect infected plants (Zhang et al 2003, Huang et al., 2004, Larsolle and Muhammed, 2007, Mahlein et al., 2010). Previous research has related some regions of the visible spectrum (400nm-750nm) and the physiological parameters obtained from the classical measurements, similar to those obtained from gaseous exchange analysis, and that these are determined for specific absorbance-reflectances patterns for the photoactive pigments, mainly chlorophylls, and carotenoids (Sims and Gamon, 2002).

These two types of pigments are mainly responsible for absorbing the light intended to promote photosynthesis in the chloroplast: chlorophylls (“a” and “b”) and carotenoids (alpha-carotene, beta-carotene and xanthophyll), with chlorophyll being the chromophore biomolecule that intervenes most directly in the absorption and conversion of light energy (Azcon, 2008). Consequently, changes in total chlorophyll foliar concentration and chlorophyll proportions “a” and “b” are indicators of physiological variation due to stress, leaf development, senescence and factors directly related to the primary production rate (Blackburn and Ferwerda, 2008). Reflectance in the leaf changes significantly due to the stress generated by pathogens at specific wavelengths in the visible range (Vis, 380nm-750nm), and a general variation in the far red range (690nm- 720nm) provides an indication of earlier or more consistent infection than the reflectance in other regions of the electromagnetic spectrum (Carter, 1994; Carter and Knapp, 2001). However, the level at which different pathogens can produce different spectral signatures in plants and the degree to which the spectral response to a particular stress factor can vary between species, are questions yet still unresolved on many pathosystems due to specificity.

Based on this, a large number of indexes and algorithms have been developed for the non-destructive estimation of chlorophylls and carotenes, which stand on the spectral reflectance in the Vis range in a wide cultivar of species, plants and organs (Gitelson et al. , 2002, Gitelson et al., 2003; Merzlyak et al 2003;. Anatoly et al, 2017. Solovchenko et al, 2005). The detection and discrimination of diseases in plants use these indices and algorithms based on specific wavelengths, related to physiological changes in plants (Naidu et al., 2009; Song et al., 2011; Zhang et al., 2012; Mahlein et al., 2013). However, the mechanisms responsible for close relationships between reflectance and plant physiology at early stages of infection, even when symptoms are not visible, need more detailed research.

The objective of this study was to relate the changes in specific spectral traces present due to the invasion of *F. oxysporum* in tomato, with the invasion of the pathogen in the tissue and associated physiological changes, which allow verifying that process of pathogenesis cause specific spectral variations. Additionally, since water stress in the plant is one of the consequences of *F. oxysporum* invasion in the tissue, hydric stress and its spectral fingerprint and other parameters were compared with infected plants, in order to verify the specificity of the spectral response found with the pathogen.

3.2Methodology

3.2.1 Biological material

The plants used in this study were maintained under semi-controlled greenhouse conditions, located at the Universidad Nacional de Colombia in Medellín (Antioquia, Colombia). The environmental conditions presented average temperatures between 18-24 °C, relative humidity between 60-70% and a photoperiod of 12 hours during the time the experiments lasted. In this study, the tomato cultivar Ponderosa was used. This cultivar is susceptible to all races of *F. oxysporum* (Reis and Boiteu, 2007). *F. oxysporum* Fo5 strain isolated from Passiflora edulis (passionfruit) was used because it is highly pathogenic on tomato plants. This strain showed an incubation period of 24 dpi on tomato plants. The maintenance protocol of the plants, the process of pathogen inoculation (Ortiz and Hoyos-Carvajal 2016), and the infectivity tests were described extensively in previous publications made by our work team (Marin et al. 2018).

3.2.2 Physiological parameters on foliar tissue

Some important photosynthetic parameters were measured on all plants (infected plants before the appearance of disease symptoms, those subjected to hydric stress and controls): the net assimilation rate of CO_2 (A), intercellular CO_2 concentration (Ci), stomatal conductance (gs) and transpiration rate (E), using an Infrared Gas Analyzer (ADC Scientific Ltd., model LCi, UK). It was also calculated the intrinsic water use efficiency (A/gs), transpiration efficiency (A/E) and the ratio of internal (Ci) and atmospheric (Ca) CO_2 concentration (Ci/Ca). The quantitative yield of PSII (ΦPSII) and continuous fluorescence (Ft) performance were measured with a modulated fluorometer (FluorPen 100 WP) to evaluate the efficiency of photosystem II in tomato leaves subjected to different treatments as indicators of biotic stress in plants. The measurements were made in 30 plants for each treatment, with five repetitions per plant, in the second developed leaf. Samples were taken on day 0 and 12 (the asymptomatic period for plants inoculated with *F. oxysporum*) and at the end of the experiment, at 24 dpi (the symptomatic period in plants inoculated with *F. oxysporum*).

3.2.3 Data Analysis

Five reflectance spectra of the adaxial face of the second leaf of each tomato plant were measured in the spectral range between 380-1000nm with a spectral resolution of ~0.5nm, using an Ocean Optics HR2000 spectroscope with a tungsten halogen light source HL-2000-HP (wavelength range of 360-2400

nm). A completely randomized design was carried out to compare two treatments: susceptible plants inoculated with *F. oxysporum*, and plants subjected to water stress at 60% field capacity; additionally, no infected plants were measured with 100% field capacity. Reflectance measurements were performed every three days after infection. Physiological parameters derived from the gas exchange and chlorophyll fluorescence analysis were performed at day 0 dpi, at day 12 dpi (when symptoms are not visible) and at day 24 dpi (with visible symptoms). Due to this design, we collected three groups of data with 150 spectra each (30 leaves per treatment), for each sampling day (450 spectra/day of sampling).

We performed a selection of the spectra and those with noise were removed, either because the spectra was deformed and/or because of a reading error. These spectra showed very different patterns. The differences were confirmed with an outlier analysis identified in a Principal Component Analysis (PCA) without prior data treatment. The SNV transformation was chosen as one of the best pre-treatments that allow a good grouping of the plants in the treatments carried out, according to the results of the analyses carried out in previous works (data not shown).

After performing the pretreatment, we built a matrix graph of the correlation coefficients initially calculated as a measure of the relationship between the growth of *F. oxysporum* in roots and leaves (CFU/mg) and the reflectance in the spectral range measured. The evaluated treatments were then compared using only the physiological parameters derived from the gas exchange analysis and the chlorophyll fluorescence, testing the difference between the means with a one-way ANOVA with multiple samples and represented with bar graphs. Then, the determination coefficients for linear regressions plotted with the objective of relating the spectral variance in leaves of *S. lycopersicum* submitted to the three treatments evaluated, with the variation in physiological parameters measured with the Infrared Gas Analyzer. Finally, we proceeded to make the adjustment of the regression lines with the largest R^2. Linear Discriminant Analysis (LDA) were made to carry out supervised classifications with the photosynthetic variables, and the specific wavelengths selected above. All analyses were done with the Software R (R Development Core Team, 2005).

3.3 Results

Plants inoculated with *F. oxysporum* developed visual symptoms between 21-24 dpi under moderate environmental conditions (temperature between 18-24 °C and relative humidity between 60-70%). The leaves suffered gradual chlorosis, starting from the lower layers upwards, but without loss of turgor

evident in most of the plants (Fig. 4). Plants subjected to water stress presented visual symptoms after 18 dpi (data not shown).

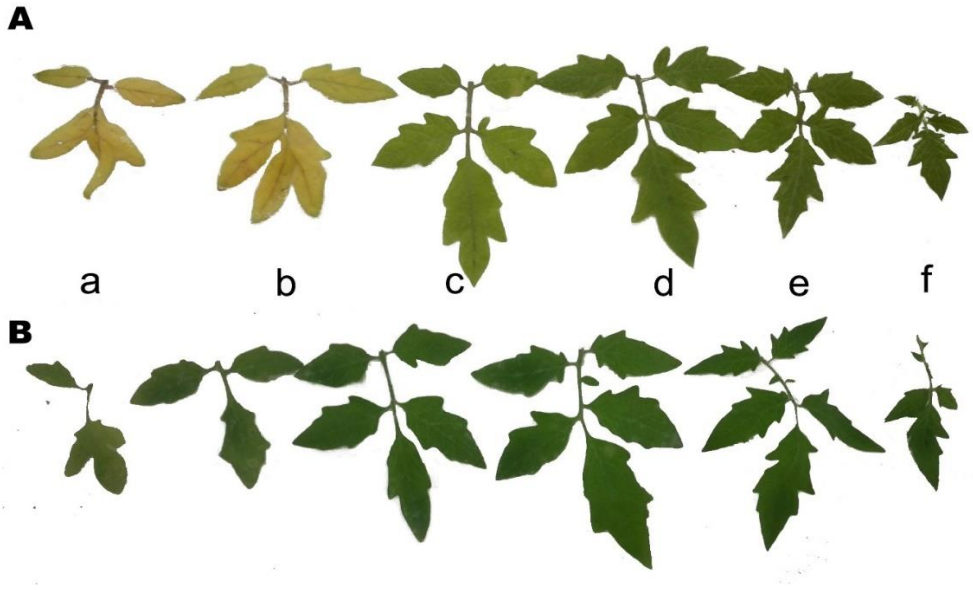

Fig. 4. The photo shows the first six leaves of a plant inoculated with *F. oxysporum* (A), and a no-infected plant (B): a) leaf 1, b) leaf 2, c) leaf 3, d) leaf 4, e) leaf 5, f) leaf 6.

3.3.1 Colonization of *F. oxysporum* in tomato plants and relationship with the spectral response in leaves during the incubation period of the disease

The growth of an isolate of *F. oxysporum* (Fo5) on root and stem, in a cultivar of susceptible tomato plants, was evaluated during the incubation period of vascular wilt. The increase in the concentration of *F. oxysporum* conidia in the course of the experiment describes the classical "J" type growth curves for both organs measured. During the first 12 dpi, a similar tendency, coincidently, was observed in the spectral response of the control plants, even presenting a slight decrease in the reflectance in the infected plants in the range Vis/NIR measured (Fig. 5A and 5B). After this period, there is a higher growth of fungi, up to 1.5 x 10^3 CFU/mg on inoculated roots(Fig. 5), and in terms of reflectance, after 12 dpi there is an increase of reflectancia in the Vis/NIR region in infected plants with respect to the non-infected plants (Fig. 5C and 5D).

Disease symptoms were clearly observed in most infected plants after 24 dpi. These symptoms coincide with a marked increase in reflectance in the VIS range and a decrease in the NIR range evaluated in the infected plants (Fig. 4e).

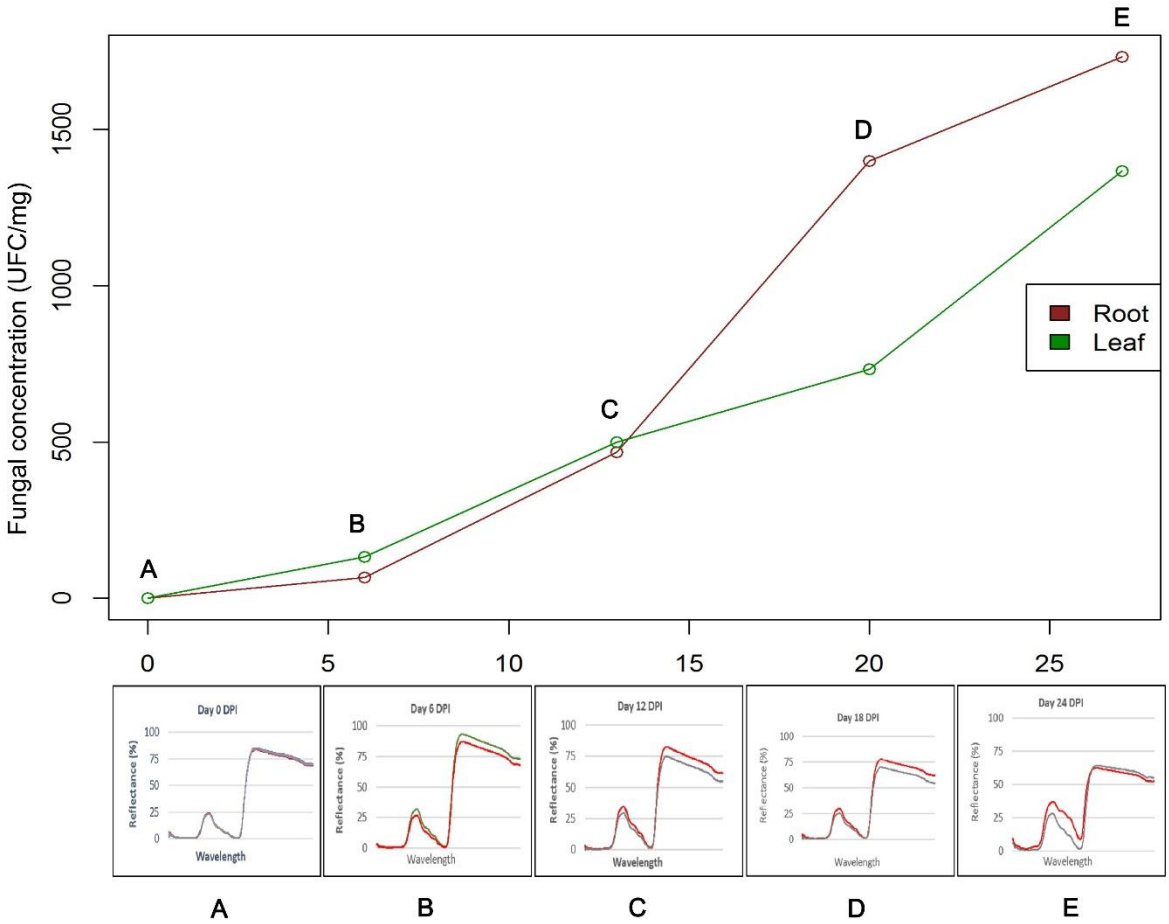

Fig. 5. Inoculum density of *F. oxysporum* (CFU) by mg in fresh plant tissue on the incubation period. A: day 0 dpi; B: day 6 dpi; C: day 12 dpi; D: day 18 dpi; E: day 24 dpi. Red series in boxes: Plants infected with *F. oxysporum*; Green series in boxes: control plants (without infection).

Five spectral bands most highly correlated with the growth of *F. oxysporum* in root and leaf ($r \geq 0.8$, $p \leq 0.05$) on infected tomato plants are shown (Fig. 6A). Two of them were positively correlated in the visible range, 448-523nm and 624-696nm, and three were negatively correlated in the near infrared (NIR) range measured, 740-960nm, 973-976nm, and 992-995nm. In general, higher values were observed in the correlation coefficients and lower values for p-value on the *F. oxysporum* measurements made on the leaves of the plants, in Figure 6B a specific example can be seen in the 992-995nm band. This figure shows a significant negative correlation between the concentration of conidia (CFU) and the percentage of reflectance at 950 nm measured in infected plants with F. oxysporum.

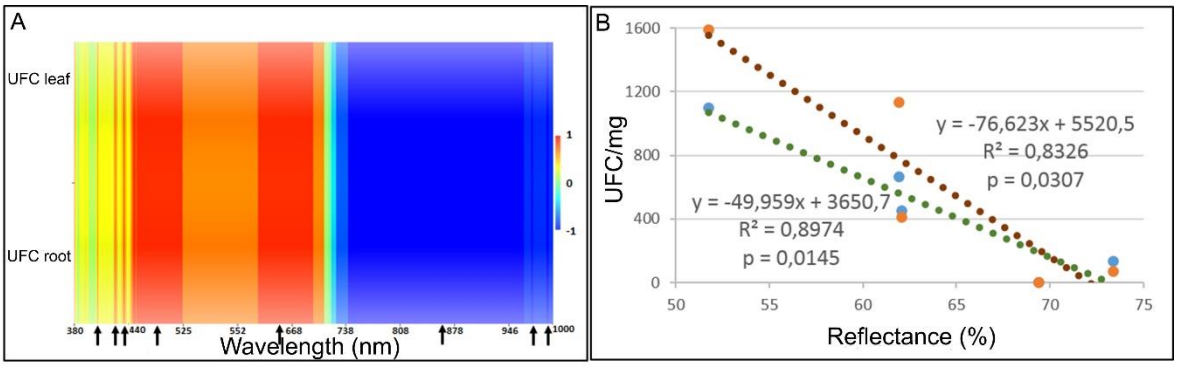

Fig. 6. Quantitative relationship between the growth of *F. oxysporum* on roots and leaves (CFU/mg) and the reflectance in tomato leaves. A) Matrix plot showing eight bands with high R^2 values ; B) Linear regression for the dependent variable "CFU" measured in roots (brown) and leaves (green) as a function of reflectance at 950 nm.

3.3.2 Physiological response of tomato plants inoculated with *F. oxysporum* during the incubation period and subjected to water stress

Regarding the physiological parameters evaluated by means of classical instruments, no significant differences were detected in the plants infected with *F. oxysporum* and subjected to water stress with respect to the control plants at 12 ppi (asymptomatic period, Fig. 7C, 7E), suggesting a minor differentiation between inoculation, water stress, and healthy plants. At 24 dpi, when the symptoms of the disease were already visible, the significant variation was observed in most of the physiological parameters evaluated in the plants subjected to the two types of stress with respect to the control plants. Specifically, in the plants infected with *F. oxysporum* there was a significant reduction with respect to the control plants for A, E, A/gs, A/E and Ft (Fig. 7A, 7C, 7E, 7F and 7H), and an increase in gs, Ci and Ci/Ca (Fig. 7B, 7D and 7I). Plants subjected to water stress presented the same pattern as infected plants, but there was no significant difference in gs (Fig. 7B).

3.3.3 Spectral variance in *S. lycopersicum* infected with *F. oxysporum* vs physiological responses

In all groups of plants, the determination of the coefficient resulting from each linear regression between reflectance vs physiological parameters (A, E, Ci and gs) result in high values on the range of 420-490nm (blue light), and had secondary peaks close to 560 nm (green), 680nm and 710nm (red) (Fig. 8).

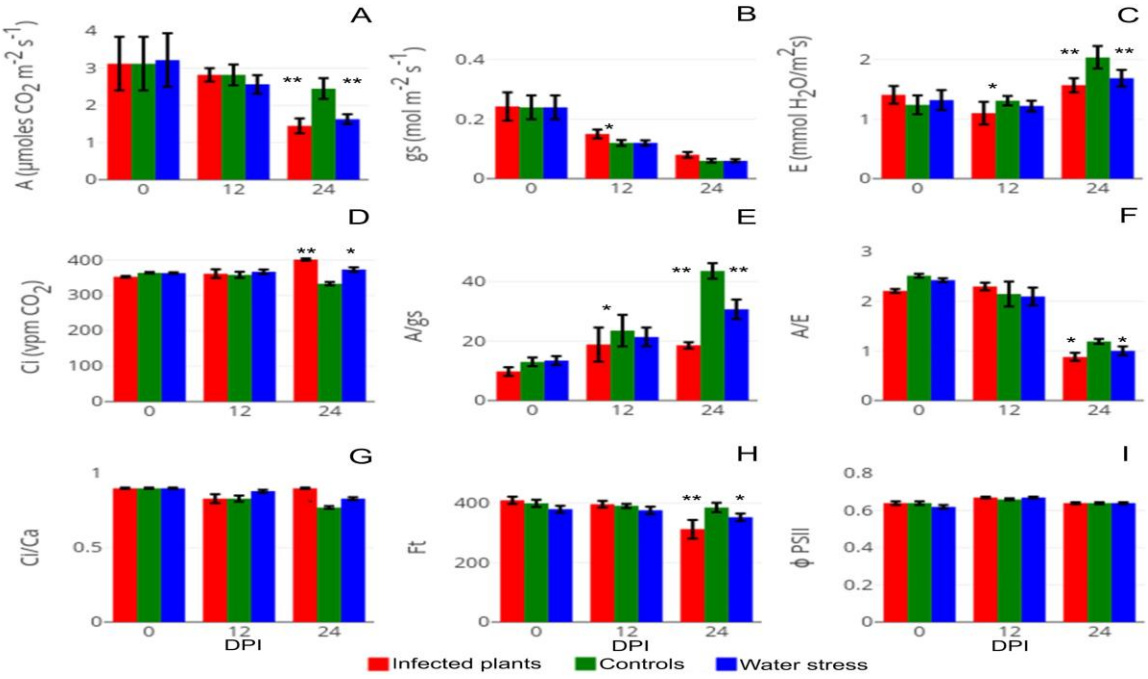

Figure 7. Bar plots showing the mean values (± standard error) of A, gs, E, Ci, A/gs, A/E, Ci/Ca (measured with an Infrared Gas Analyzer). The Ft' and Qy' (Φ PSII) under constant actinic light (measured with FluorPen 100 WP modulated fluorometer). Red: infected plants with *F. oxysporum*; blue: subjected plants to water stress; green: no-infected plants. Estimates indicated by * and ** were significant at the 5% and 1% levels, respectively.

Figure 8 shows the determination coefficients (R^2) calculated for each wavelength in the measured spectral range, with respect to each monitored physiological parameter. This R^2 are maximum on ranges 420-490 nm, 560 nm, and 680 nm, mainly in plants inoculated with *F. oxysporum* (Fig. 8A). On plants subjected to water stress, r was also slightly greater than 0.7 in the wavelengths of blue (Fig. 8C). In gs and E, the majority of wavelengths mentioned above had high correlation values (Fig. 8G-8L), except for the peak at 560nm that had low correlations with these physiological parameters in plants subjected to water stress (Fig. 9I, 9L), and in the particular case of E in the control plants (Fig. 8K). The graph referring to r for linear regressions between reflectance and intercellular CO_2 exhibits an inverse pattern to that observed in the other parameters (A, gs and E) (Fig. 8D-8F), in which they presented high peaks (R^2> 0.7) at 510 nm, 658 nm, 694 nm, and 750 nm only in plants subjected to water stress (Fig. 8F). The reasons why this inverse pattern occurs and other particularities will be discussed later in this manuscript.

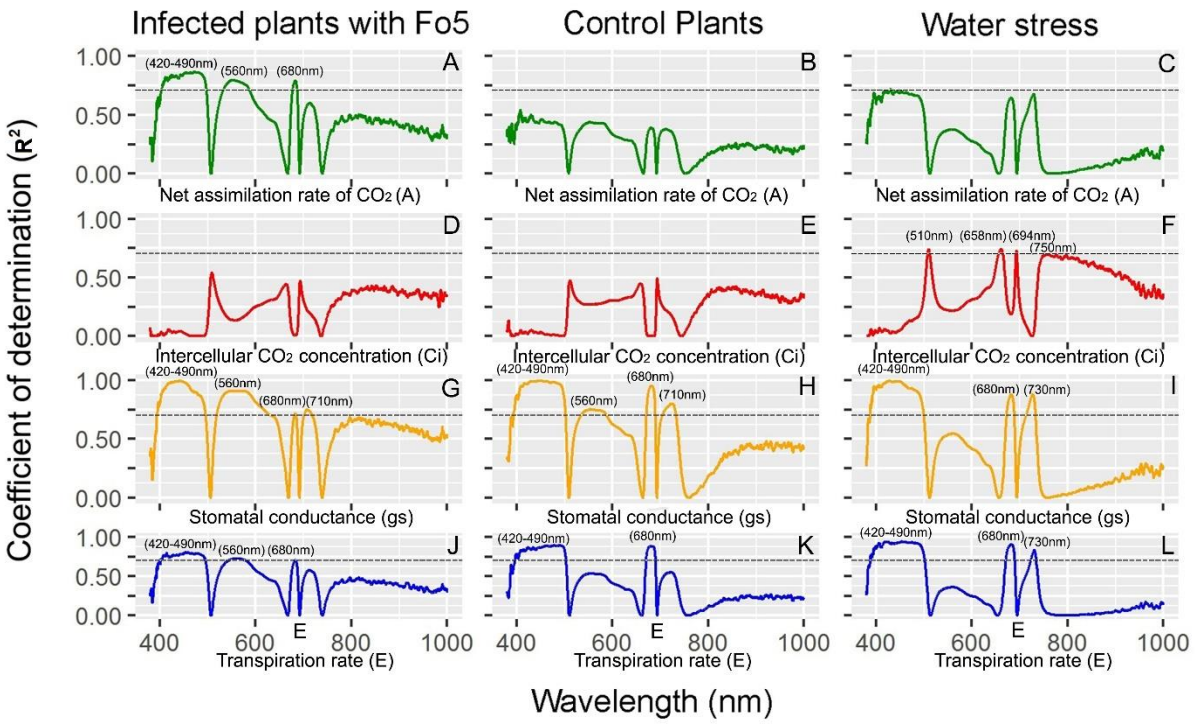

Figure 8. Coefficient of determination (R^2) vs wavelength for simple linear reflectance relationships in tomato leaves with photosynthetic parameters in plants infected with *F. oxysporum* (Fo5), subjected to water stress and controls. The wavelengths in which the best fit relationships were identified are indicated in parentheses (the black line indicates the $R^2 = 0.7$).

Once the wavelengths with a higher coefficient of determination are defined, adjusted regression lines with the highest R^2 (and r and p-value) were drawn and shown in fig. 9. The blue range is represented by a wavelength of 440 nm where some of the physiological parameters showed R^2 maxima. It is important to highlight the similarity between the linear regressions of the wavelengths in the blue (440 nm) and the first peak in the red (680nm), with positive relationships between predictor (X) and response variables with low values of p ($p << 0.05$). This suggests that fluctuations in the predictor variables (A, Ci, gs and E) are highly associated with the response variable (reflectance at 440nm and 680nm). This pattern is not equivalent on particular cases of intercellular CO_2 in plants inoculated with *F. oxysporum* and controls on which there is a negative relationship betwen reflectance (440nm and 680nm) and "Ci" (Fig. 5D and 5E); but since p values are high ($p >> 0.05$), predictor variable (Ci) and response are not associated. In the green wavelength (~ 560nm) the reflectance decreases with increasing A (Fig. 9A, 9B and 10C), gs (Fig. 10G, 10H and 10I) and E (Fig. 10J, 10K and 10L) with p values << 0.05, instead, a positive relationship between intercellular CO_2 was noticed with the explanatory variables. A similar pattern occurs in the red light wavelength (710nm) even with lower p values, except Ci on plants

subjected to water stress (p = 0.340) (Fig. 9F), revealing that changes in Ci are not linked with reflectance at this point.

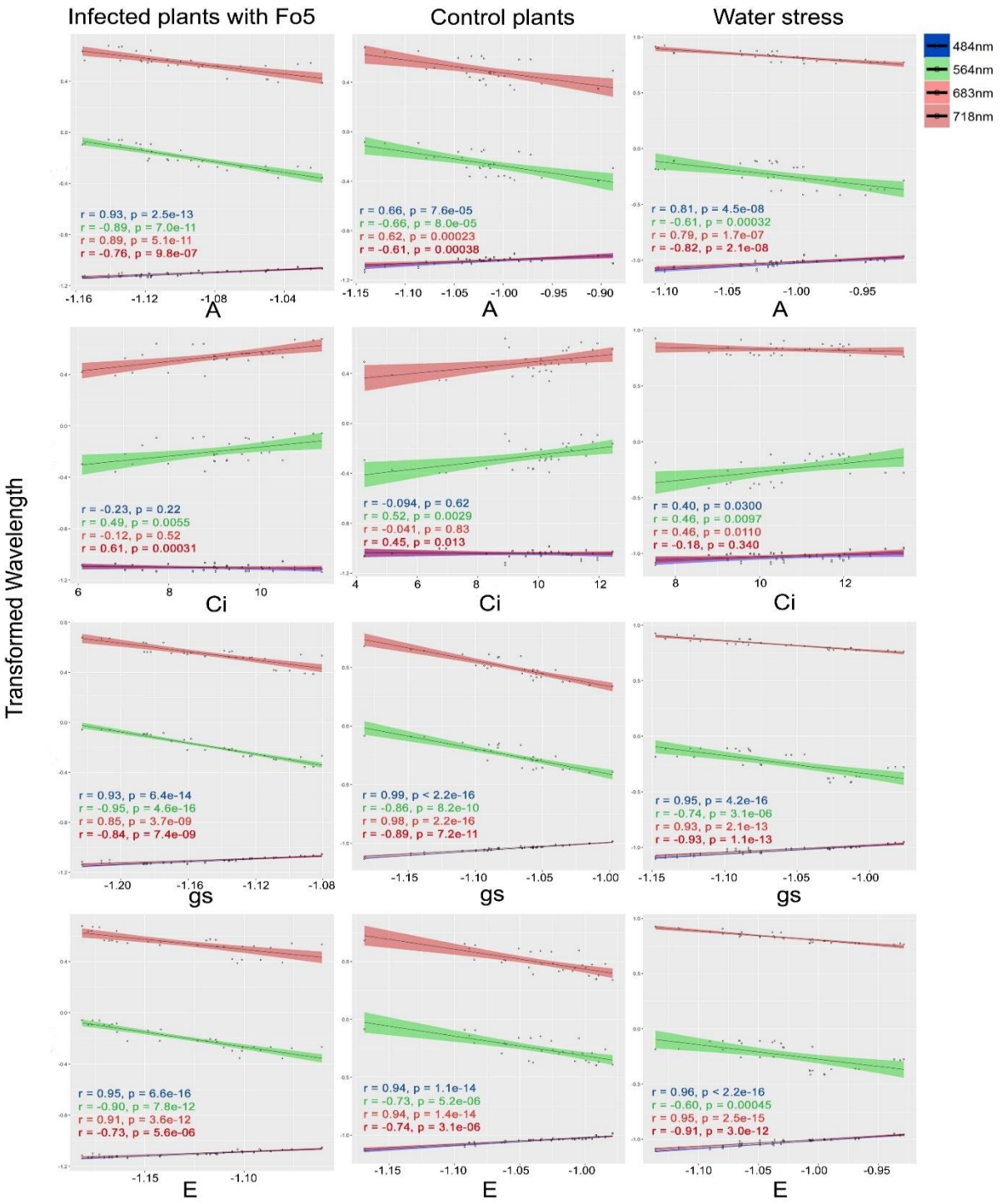

Fig. 9. Best adjusted linear reflectance near 440nm (blue), 564nm (green), 683nm (light red) and 718nm (dark red) with A, E, gs and Ci on tomato leaves inoculated with *F. oxysporum*, subjected to water stress and control plants on day 12 after inoculation (incubation period of the disease).

As a result of the Principal Components Analysis (PCA), the first two Principal Components (PC) explain 83.3% of the variance of the data (Fig. 10A), the quality of the variables is represented in the factor map by means of the square cosine or square coordinates (cos^2). Since high values of cos^2 indicate a satisfactory contribution of the variable in the PCs. Most variables are placed near the correlation circle on this study, indicating a good representation of the variables in the first PCs (red); others like Ci, Ft, and A/E present moderate correlation values (yellow), and A/gs, close to the center, have a reduced representation in the first PCs (blue).

To perform a dimensional reduction for the use of relevant variables in future prediction models, percentages of the contributions of the variables in each principal component were represented (Fig. 10A, 10B). Variables correlated with PC1 (Dim-1) are the physiological parameters gs, Qy, E, A, and the wavelengths at 718nm, 564nm, 484, 683, and 700nm (Fig. 6B); and with PC2 (Dim-2) are Ci, Ft and the wavelengths at 510nm, 800nm, 650nm, 750nm, 700nm, and 950nm. These are critical variables to explain the variability in the data set. Others that do not correlate with any PC or with last dimensions correspond to variables with poor contribution and could be eliminated to simplify the general analysis.

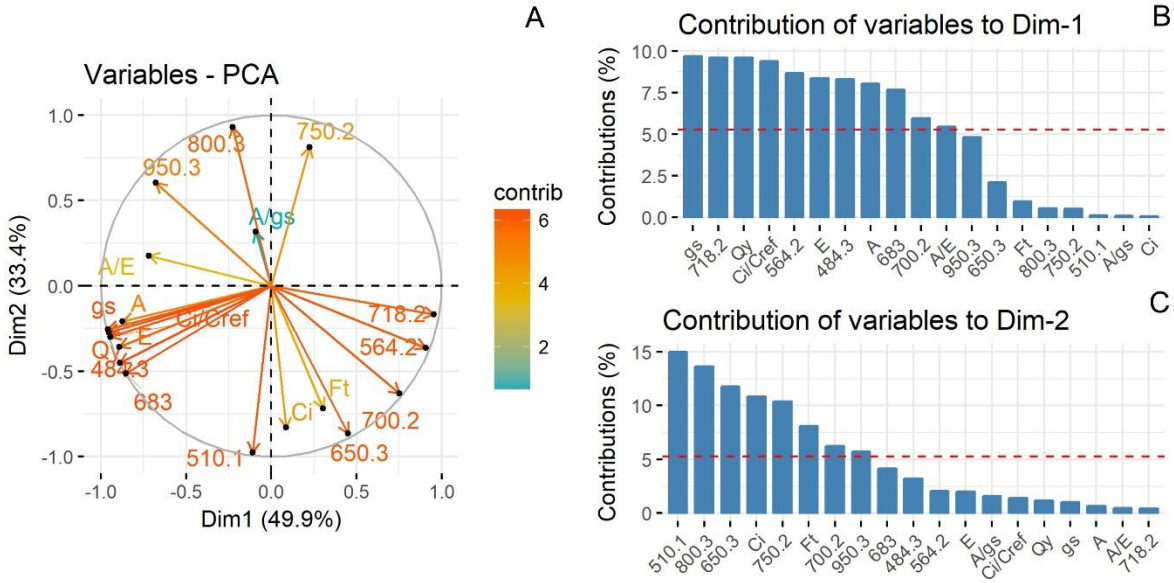

Fig. 10. A) Principal Components Analysis in which the contribution of representative physiological variables and wavelengths with high coefficients of determination to the first PCs by means of the square cosine is represented. B) Contribution of variables to PC1 (%). C) Contribution of variables to PC2 (%).

A statistical Linear Discriminant Analysis (LDA) with transformed data (SNV) was done to achieve a supervised classification with the photosynthetic variables and the specific wavelengths selected above. Three groups were defined a priori: i) plants inoculated *with F. oxysporum* (12 dpi, asymptomatic

period), ii) plants subjected to water stress and iii) controls (Fig. 11). The LDA with photosynthetic parameters obtained from the analysis of gaseous exchange and fluorescence of the chlorophyll achieved an acceptable classification percentage of 82%; plants inoculated with *F. oxysporum* are well separated from the other treatments, but the model does not differentiate plants subjected to water stress from the no-infected plants at 12 dpi (asymptomatic period) (Fig. 11A). Finally, the LDA done with the spectral variables achieved a correct classification of 100% (Fig. 11B), perfectly discriminating the three treatments at day 12 after the inoculation, when the symptoms are not visible yet.

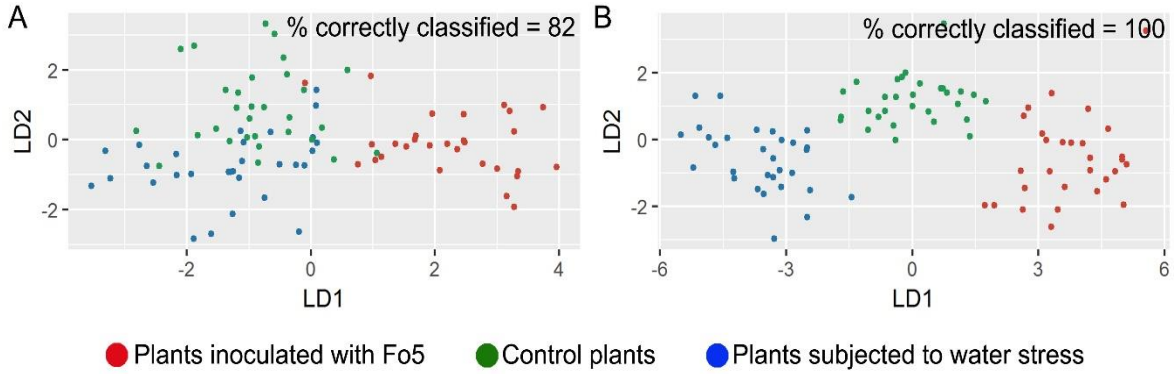

Fig. 11. Linear Discriminant Analysis (LDA) with transformed data (SNV) in plants inoculated with *F. oxysporum* (12 dpi), plants with water stress and controls. A) Physiological variables: A, gs, Qy, E, A, Ci, Ft, and Qy B) Reflectance data of the selected wavelengths: 484nm, 510nm, 564nm, 650nm, 683nm, 700nm, 750nm, 800 and 950nm.

3.4 Discussion

3.4.1 Relationship of *F. oxysporum* concentration with the spectral response in the incubation period of the disease

The measurement of conidial concentration of *F. oxysporum* in tomato and its relation with spectral response in leaves during the incubation period of the disease is a basic requirement to allow comparisons, repetitiveness, and analysis of disease tolerance and standardization of the results in many experiments on controlled and semi-controlled environments (Caligiore-Gei and Valdez, 2014). The inoculum concentration of *F. oxysporum* influenced the severity of the disease. Specifically, it has been demonstrated that the increase of inoculum of F. oxysporum showed an increase on the incidence, mortality (20%), the severity of the disease and number of lesions on gladiolus roots (Riaz et al., 2014). Additionally, it was found a high coefficient of determination (R^2 = 0.94) on the linear relationship between the severity of the disease generated by *F. solani* in bean plants and the concentration of chlamydospores in the substrate (Nicoli et al., 2013).

Results on this investigation suggest a high correlation between inoculum concentration and spectral response of tomato plants inoculated with *F. oxysporum* before the symptoms are visible, particularly in the blue (448nm-523nm) and red (624nm-696nm) regions of the visible spectrum and in the infrared plateau between 750nm-1100nm. The mechanism underlying this correlation is different for each region of the electromagnetic spectrum. Primarily the main photosynthetic pigments (chlorophylls and carotenoids) determine reflectance in the visible region, while in the infrared plateau the high reflectance is due to multiple dispersion within the leaf, in relation to air spaces within the tissue (internal structure) and its water content (Jacquemoud and Ustin, 2001). Maximum absorption peaks for chlorophyll and carotenoids in the blue range (chlorophyll a = 428nm, chlorophyll b = 453nm, carotenoids = 450nm) and red range (chlorophyll a = 661nm, chlorophyll b = 642nm) (Guidi et al., 2017) are coincident with wavelengths on the Vis spectrum correlated with the concentration of the pathogen (448nm-523nm and 624nm-696nm) under our experimental conditions. These results suggest a direct relationship between the inoculum concentration and the photosynthetic response in the plant that can be used for the design of models to quantify pathogen densities in asymptomatic periods of different vascular diseases. Spectral responses assessed on tomato plants for this research were complete at leaf scale, in order to observe specific changes on spectral characteristics in the course of the infection process. The inoculum concentration is indirectly measured by intermediate responses in tissue. However, to measure the direct concentration of spores or structures of the fungus in a particular tissue, the use of hyperspectral microscopes is necessary (Thomas et al, 2017).

3.4.2 Analysis of physiological changes in tomato plants

In this section, the effects generated in the exchange of gases from photosynthesis and the fluorescence of chlorophyll by the inoculation of *F. oxysporum* and the submission to water stress, on leaves, of a susceptible cultivar of tomato, in the incubation period of the disease are reported. Meanwhile during the asymptomatic period of the disease (12 dpi), a significant difference w found on the gs, E, A/gs and A/E, which suggests a water imbalance on plants inoculated with *F. oxysporum* respect to the control plants, thus corroborating water stress as an important factor for vascular wilt diseases (Nogués et al., 2002; Ghaemi et al., 2009; Ochola et al., 2015). Vessel obstruction is one of the relevant effects of *F. oxysporum* in the photosynthesis of leaves, but little is known about the effect of water stress induced by pathogens. However, the reduction of the diameter of the vascular bundles by species of the *Fusarium* spp., their metabolites, and enzymes or by inducing the accumulation of gummy and hairy

substances, generating resistance to the movement of water, has been widely reported (Aguirreolea et al., 1995, Jensen, 2002).

Under the experimental conditions here, non-significant differences were found on net photosynthesis at 12 dpi using univariate analysis with the Infrared Gas Analyzer (IRGA), neither with parameters of chlorophyll fluorescence (discussion of multivariate analysis will be addressed later), although it has been reported to be one of the first symptoms of vascular wilt. The often is reported indicating that the water imbalance generates a decrease in the rate of CO_2 fixation, on electron transport chain in chloroplasts, and lack of activity of ATP synthase (Yordanov et al., 1997, Flexas et al, 1999, Flexas et al, 1999b). However, different authors have also reported the absence of significant differences up to 17 dpi in gas exchange parameters measured with an infrared gas analyzer and changes in quantum yield of photosystem II only until 27 dpi in tomato plants infected with *F. oxysporum* (Lorenzini et al., 1997; Nogués et al., 2002).

The pattern development of the wilt by *Fusarium* reduces the photosynthetic activity depending on different mechanisms. Photosynthesis variation is difficult to detect on the asymptomatic period of the disease with these techniques. This can also be explained according to the type of vascular wilt (type I or II) described by Pshibytko *et al.* (2006), which depends on environmental conditions such as temperature, relative humidity and moisture content of the floor. In this research, the pattern of disease development on inoculated tomato plants was Type I, according to the classification of Pshibytko et al (2006). In this study, plants inoculated with Fo5 exhibited yellowing, and slow-wilting from lower to upper leaves (Fig. 1), on moderate environmental conditions (average relative humidity between 60% - 70%, temperature of 18-24 ° C), and gradual symptom display, only after 21 dpi. Pshibytko *et al.* (2006), suggest that on infection type I wilt, the mycelium of *F. oxysporum* partially obstructs xylem and grows extensively within the parenchyma, being able to absorb the nutrients and regulate the metabolic processes of the plant. Finally, in this type of wilting, plants die due to nutrient deficiency and poisoning by fungal toxins, and not due to water deficit. This hypothesis plausibly explains the chronology of the symptoms described above and why in this research the leaves of plants infected with *F. oxysporum* only showed loss of turgor shortly before wilting.

The difficulty in detecting responses associated with the disease in the asymptomatic period could be due to the lack of sensitivity in the equipment used for measuring gas exchange and fluorescence parameters of chlorophyll. Since in both cases few parameters are assessed on the plant, such as the

absorbance in the NIR to 1640cm-1 for the case of the IRGA, and a relation between the pulse of saturating light generated by the fluorometer (usually ultraviolet light) and its consequent response of lower energy generated by the sample (in the Vis or NIR). The conclusions obtained from the limited univariate analysis can also lead to erroneous conclusions. Therefore, it is necessary to implement more robust multivariate analyses (as in Fig. 7 and 8), which will be discussed later.

Finally, when the plants inoculated with *F. oxysporum* have finished the incubation period (24 dpi) and the symptoms of the disease on the leaf and on plants subjected to water stress can be observed, net photosynthesis decreases, as well transpiration rates, intrinsic efficiency in water use, transpiration efficiency (A/E), and the proportion between intercellular CO_2. Nevertheless, there is no evidence that this variation on parameters can lead to permanent damage on energy transduction phases, nor by stomatal or hydric limitations, coincident with results obtained in *P. edulis* with the same isolate of *F. oxysporum* (Cruz, 2012). Chlorophyll fluorescence parameters, especially the transient fluorescence (Ft), which significantly decreased on this study between evaluated treatments, remarks the importance to obtain information about the pigment complexes, the organization and the transfer of excitation energy between them, and different specific electron transfer reactions in photosystem II (Stirbet, 2012). Additionally, plant samples have a transient fluorescence that levels off in a new stable state before any change in the quality or quantity of the light to which they are exposed, regardless of whether the sample was adapted to darkness or light. Although Ft is an important tool to assess plant stress (Strasser et al., 2000), nevertheless in this study it is not an indicator of early stress on plant diseases.

3.4.3 Linking the physiological parameters with the leaf reflectance in the Vis/NIR during the incubation period of the disease

Chlorophyll (mainly “a” and “b”) and carotenoids are essential pigments for the conversion of light energy into stored chemical energy. The amount of solar radiation absorbed by a leaf is a function of the content of photosynthetic pigments; therefore, the chlorophyll content can determine directly photosynthetic potential and primary production (Filella et al., 1995, Gitelson et al., 2003). Higher coefficients of determination (R^2) between some regions of the Vis spectrum and the parameters obtained from the gas exchange measurements are determined by the specific absorbance-reflectance patterns of the photoactive pigments, mainly chlorophylls and carotenoids (Gitelson et al., 2002, Gitelson et al., 2003; Merzlyak et al. 2003; Anatoly et al., 2017; Solovchenko et al., 2005). These pigments have ranges of maximum absorption that coincide with the wavelengths linked with the *F. oxysporum* infection found in this work. In this research, plants inoculated with *F. oxysporum* showed

high correlations with A, gs and E on 4 bands of the Vis spectrum. The first high correlation zone is located between 420nm-490nm and coincides with the spectral region where the maximum absorbance of chlorophyll a (430nm), chlorophyll b (453nm) and carotenoids (450nm and 485nm) occurs, while secondary peaks of the chlorophylls in the far-red (642nm and 662nm) match with regions of low correlation (Fig. 12). Chlorophylls and carotenes absorb light on mesophyll mainly at wavelengths between 400-500nm (blue), its measurement is a signal of the physiological variation due to biotic and abiotic stress, leaf development, senescence and factors directly related to the photosynthetic primary production rate (Blackburn and Ferwerda, 2008). The results of this study support the idea that infection by *F. oxysporum* has correlation with photosynthetic pigments concentration on early stages of the disease. Additionally, the wavelengths with the highest coefficients of determination with the physiological parameters in the gas exchange analysis could be used in mathematical models for early detection of plant diseases based on spectral data in the Vis.

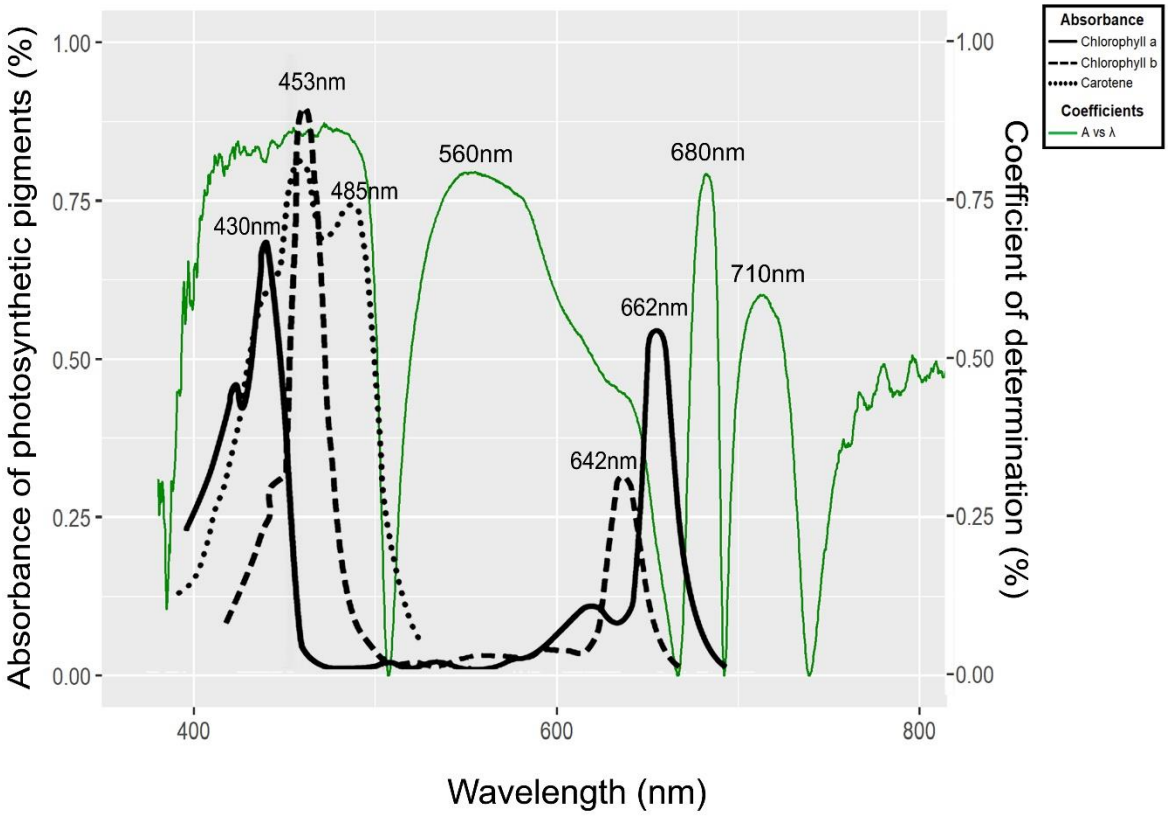

Fig. 12. Line plot illustrating the superposition of the relationship between the R^2 net photosynthesis vs wavelength in plants inoculated with *F. oxysporum* (green line) and the optimal absorption of light from the main photosynthetic pigments (lines black, absorption curves extracted from "The light-dependent reactions of photosynthesis: Figure 4," by OpenStax College, Biology [CC BY 3.0]).

Plants under water stress, show high correlations in the band of 420nm-490nm and the peak of 680nm, coincident with the absorbance of main photosynthetic pigments, due to the decrease in water potential on the leaf that suppresses the photosynthetic activity of the plant (discussed previously). However, on tomato, water stress has a high correlation between reflectances at 730nm and 750nm, which are not affected by the absorption of chlorophyll. These are linked with water stress or changes in tissues (Carter and Knapp, 2001). The control plants did not show high coefficients of determination (R^2> 0.7) between the A and Ci on none of the wavelengths in the measured spectral range, confirming the lack of significant variation of these parameters within the individuals for this treatment. The correlation found in these plants between gs and E, and the spectral bands 420nm-490nm, 560 nm, 680nm, may suggest differences in the concentration patterns of photosynthetic pigments due to gaseous exchange in the leaves (mainly water vapor) without significantly affecting net photosynthesis. This hypothesis is reinforced by the high correlation between stomatal conductance and reflectance at 710nm in this treatment (R^2> 0.75).

Reflectance spectroscopy is widely applied for the non-destructive estimation of leaf chlorophyll as an indirect measurement of the photosynthetic potential and primary production of the plant (Richardson et al., 2002). The bands of 420nm-490nm, 560nm and about 700nm (far red), which presented high values in the determination coefficients with gas exchange parameters in this research, have been widely used in the form of indexes and simple functions to relate the concentration of chlorophyll in a cultivar of plant species and in a wide range of photosynthetic pigment composition (Gitelson and Merzlyak, 1996; Gitelson and Merzlyak, 1997; Sims and Gamon, 2002; Gitelson et al., 2003; Merzlyak et al. ., 2003). Additionally, some bands in the Vis/NiR regions can be used to develop biotic stress indices during the incubation period and the moment that symptoms are already visible in the plant with vascular wilt disease (Sankaran et al., 2010; Khaled et al., 2017). Some of the bands identified in this study have been used independently and/or to develop indices, sensitive to plant disease processes, such as: the Photochemical Reflectance Index [PRI = (R570 - R531]/[R570 + R531)], Physiological Reflectance Indices [PhRI = (R550 - R531)/(R550 + R531)], Standard chlorophyll pigment rate index [NPCI = R680 - R430) / (R680 + R430)] and the Index of Anthocyanin Reflectance [ARI = (R550) -1 - (R700) -1] (Song et al., 2011; Zhang et al., 2012; Krishna et al., 2014). On plants exposed to water stress, main water absorption information in the spectral range measured in this research is in the range of short wave NIR (750-1000 nm) (Zhang et al. al., 2012; Genc et al., 2013). Due to the intention to relate

reflectance to the physiological parameters measured from a gaseous exchange analysis, it is expected that no specific bands are related to the water content in this spectral range.

The results obtained in the biplot of the PCA (Fig. 10) and the LDA (Fig. 11) reinforced the previous idea. On these analysis, some wavelengths that belong to the range of short wave NIR were added, which were highly correlated with the concentration of conidia in the plants inoculated with *F. oxysporum* (Fig. 5). In general, the wavelengths selected from the proposed methodology provided a high contribution to the first two main components, while only the parameters gs, Qy and E (derived from the gas exchange analysis) offered a contribution classified as high on main components.

The advantages of multivariate analysis over the multiple univariate analysis shown in Figure 10 should be highlighted, as a possible way of identifying system constructs of response variables, selecting subsets of variables and determining the relative value for each stand out variable (Huberty and Morris, 1989). Precisely, the addressing of this analysis with the multivariate approach, allowed to obtain high percentages of classification (82%) in the LDA, during the incubation period (12 dpi) with the variables derived from the analysis of gas exchange, while in the same period with multiple univariate analysis did not find significant difference for most parameters. Even with the multivariate analysis, the use of parameters derived from gaseous exchange analysis have limitations for the realization of applicable models in the early detection of plant disease. For example, the model applied in this study using these parameters did not allow to efficiently separate the plants subjected to water stress from the control plants (healthy and with 100% field capacity), while the model completed from the spectral reflectance with relevant specific wavelengths allows correct classification of 100% of the three treatments. The relevant specific wavelengths related to the physiological response of the plant to *F. oxysporum* can be used in the development of portable instruments for the early detection of vascular wilt, easily, quickly and non-destructively.

3.5 Conclusion

It is possible to relate, by means of an indirect methodology such as reflectance spectroscopy in the Vis/NIR, the concentration of conidia with the spectral response on the leaves of inoculated plants of *F. oxysporum* during the incubation period of the disease. The increase in the concentration of pathogen on the vascular system of the plant coincides with the increase of reflectance in the Vis region (380nm-750nm), but has an inverse relationship in the range of the infrared plateau (750nm-1000nm), at least in the last week of the incubation period. Five specific bands were found highly correlated with the

increase in the concentration of *F. oxysporum* conidia measured at root and leaves, two on the Vis range (448-523nm, 624-696nm), and three near the infrared measured (740- 960nm, 973-976nm, and 992-995nm).

Tomato plants inoculated with *F. oxysporum* maintained under greenhouse conditions did not present a significant difference in net photosynthesis with respect to control plants at 12 dpi (asymptomatic period), using only gas exchange and fluorescence analysis of chlorophyll. Multiple univariate analyzes only allowed to detect a significant decrease in this period in the transpiration, the transpiration efficiency and the proportion between the intercellular CO_2 and the environmental CO_2. However, there is no evidence of permanent damage in the energy transduction phase, nor by stomatal or water limitations. When the disease symptoms were observed (24 dpi) there was a marked decrease of the net photosynthesis, but the parameters derived from the chlorophyll fluorescence analysis showed no decrease in the electron transport efficiency of the photosystems in the plants infected with *F. oxysporum*.

Four bands in the Vis range correlated to the photosynthetic parameters derived from the gaseous exchange analysis in the tomato leaves subjected to biotic and abiotic stress. The sensitivity of the relationship between reflectance with net photosynthesis, stomatal conductivity and transpiration was higher on spectral ranges such as 420nm-490nm, 560nm, 680nm, and 710nm. Particularly, the far-red bands correlated more with those on net photosynthesis in the plants inoculated with *F. oxysporum*, allowing the estimation and detection of the infected plants during the early stages of infection.

The wavelengths selected whit this methodology allowed classifying correctly 100% of the plants inoculated with *F. oxysporum*, the plants subjected to water stress and the control plants in the asymptomatic period of the disease. These results allow a significant increase in the knowledge in the area of early detection of diseases in the specific pathosystem *S. lycopersicum-F. oxysporum*.

Based on the findings of this research, the use of robust multivariate analysis is recommended for the applied biological sciences and, specifically in the agricultural sciences. The efficacious use of these multivariate application tools can avoid falling on wrong conclusions on specific areas of interest, such as the identification of variable response systems, subset selection of variables and determination of the relative value for each variable in highly complex matrices.

References

Aguirreolea, J., Irigoyen, J., Sanchez-Diaz, M., Salaverri, J., 1995. Physiological alterations in pepper during wilt induced by Phytophthora capsici and soil water deficit. Plant Pathology, 44(3): 587-596.

Blackburn, G.A., Ferwerda, J.G., 2008. Retrieval of chlorophyll concentration from leaf reflectance spectra using wavelet analysis. Remote Sensing of Environment**,** 112(4): 1614-1632.

Bosland, PW., 1988. *Fusarium oxysporum* a pathogen of many plant species. Advances in plant pathology, 6(1): 281- 289.

Caligiore-Gei, F., Valdez J.G., 2015. Adjustment of a rapid method for quantification of *Fusarium* spp. spore suspensions in plant pathology. Revista Argentina de Microbiología, 47(2): 152-154.

Carter, G.A., 1994. Ratios of leaf reflectances in narrow wavebands as indicators of plant stress. International Journal of Remote Sensing, 15(3): 697-703.

Carter, G.A., knapp, A.K., 2001. Leaf optical properties in higher plants: linking spectral characteristics to stress and chlorophyll concentration. American Journal of Botany, 88(4): 677-684.

Filella, I., Serrano, I., Serra, J., Peñuelas, J., 1995. Evaluating wheat nitrogen status with canopy relfectance indices and discriminant analysis. Crop Science, 35(5): 1400-1405.

Flexas, J., Badger, M., Chow, W.S., 1999ª. Analysis of the Relative Increase in Photosynthetic O2 Uptake when Photosynthesis in Grapevine Leaves Is Inhibited Following Low Night Temperatures and/or Water Stress. Plant Physiology, 121(2): 675-684.

Flexas, J., Escalona, J.M., Medrano, H., 1999b. Water Stress Indices Different Leaves of Photosynthesis and Electron Transport Rate Regulations in Grapevines. Plant Cell Environment, 22: 39-48.

Ghaemi, A., Rahimi, A. Banihashemi, Z., 2009. Effects of Water Stress and *Fusarium oxysporum* f. sp. Lycoperseci on Growth (leaf area, plant height, shoot dry matter) and Shoot Nitrogen Content of Tomatoes Under Greenhouse Conditions. Iran Agricultural Research, 28(2): 51-62.

Gebrie, S.A., 2016. Biotrophic Fungi Infection and Plant Defense Mechanism. Journal of Plant Pathology and Microbiology, 7(9):1-6. DOI:10.4172/2157-7471.1000378

Genc, L., Inalpulat, M., Kizil, U., Mirik, M., Smith, S., Mendes, M., 2013. Determination of water stress with spectral reflectance on sweet corn (*Zea mays* L.) using classification tree (CT) analysis. Zemdirbyste-Agriculture, 100(1): 81-90.

Gitelson, A.A., Zur, Y., Chivkunova, O.B., Merzlyak, M.N., 2007. Assessing Carotenoid Content in Plant Leaves with Reflectance Spectroscopy. Photochemistry and Photobiology, 75(3): 272-281.

Gitelson, A.A, Gritz, Y., Merzlyak, M., 2003. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. Journal Plant Physiology, 160(3): 271-282. DOI: https://doi.org/10.1078/0176-1617-00887

Gitelson, A, Merzlyak, MN., 1996. Signature analysis of leaf reflectance spectra: algorithm development for remote sensing of chlorophyll. Journal Plant Physiology, 148(3-4): 494–500.

Gitelson, A, Merzlyak, MN., 1997. Remote estimation of chlorophyll content in higher plant leaves. International Journal of Remote Sensing, 18(12), 291-298. https://doi.org/10.1080/014311697217558

Gregory, P.J., Ingram, J.S.I., Andersson, R., Betts, R.A., Brovkin, V., Chase, T.N., Grace, P.R., Gray, A.J., Hamilton, N., Hardy, T.B., Howden, S.M., Jenkins, A., Meybeck, M., Olsson, M., Ortiz-Monasterio, I., Palm, C.A., Payne, T.W., Rummukainena, M., Schulze, R.E., Thiema, M., Valentin, A., Wilkinson, M.J., 2001. Environmental consequences of alternative practices for intensifying crop production. Agriculture, Ecosystems and Environment, 1853: 1-12.

Guidi, L., Tattini, M., Landi., M., 2017. How Does Chloroplast Protect Chlorophyll Against Excessive Light?. Cap. 3, 22-26. Published by INTECH, Colombia. isbn = 978-953-51-3107-6.

Huang, M.Y., Huang, W.H., Liu L.Y., Huang, Y.D., Wang, J.H., Zhao, C.H., Wan, A.M., 2004. Spectral reflectance feature of winter wheat single leaf infested with stripe rust and severity level inversion. Transactions of the CSAE, 20(1): 176-180.

Jacquemoud, S., Ustin, S., 2001. Modeling leaf optical properties. Proc. 8th International Symposium Physical Measurements & Signatures in Remote Sensing, Aussois (France), 8-12 January, CNES, 223-232.

Jensen, JR., 2002. Remote sensing of the environment – An earth resource perspective. [reprint.] edition. The MIT Press and MIT Press, Upper Saddle River, NJ, USA.

Khaled, A.Y., Aziz, S.A., Bejo, S.K., Nawi, N.M., Seman, I.A., Onwude, D.I., 2017. Early Detection of Diseases in Plant Tissue Using Spectroscopy – Applications and Limitations. Applied Spectroscopy Reviews, 53(1): 36-64.

Koeck, M., Hardhamb, A.R., Doddsa, P.N., 2011. The role of effectors of biotrophic and hemibiotrophic fungi in infection. Cell Microbiology, 13(12): 1849-1857.

Krishna, G., Sahoo, R.N., Pargal, S., Gupta, V.K., Sinha, P., Bhagat, S., Saharan, M.S., Singh, R., Chattopadhyay, C. 2104. Assessing Wheat Yellow Rust Disease through Hyperspectral Remote Sensing. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. XL-8, 1413-1416

Larsolle, A., Muhammed, H.H., 2007. Measuring crop status using multivariate analysis of hyperspectral field reflectance with application to disease severity and plant density. Precision Agriculture, 8 (1–2): 37–47. DOI: 10.1007/s11119-006-9027-4

Lorenzini, G., Guidi, L., Nali, C., Ciompi, S., Soldatini, G.F., 1997. Photosynthetic response of tomato plants to vascular wilt diseases. Plant Science, I24(2): 143-I52. https://doi.org/10.1016/S0168-9452(97)04600-1

Mahlein, AK, Steiner, U, Dehne, HW, Oerke, EC., 2010. Spectral signatures of sugar beet leaves for the detection and differentiation of diseases. Precision Agriculture, 11(4): 413–431.

Mahlein, A.K., Rumpf, T., Welke, P., Dehne, H.W., Plümer, L., Steiner, U., Oerke, E.C., 2013. Development of spectral indices for detecting and identifying plant diseases. Remote Sensing of Environment, 128: 21-30.

Marín-Ortiz J.C., Hoyos-Carvajal L.M., Botero-Fernández V. 2018. Detection of asymptomatic *Solanum lycopersicum* L. plants infected with *Fusarium oxysporum* using reflectance VIS spectroscopy. Revista Colombiana de Ciencias Hortícolas, 12(2): 436-446. DOI: 10.17584/rcch.2018v12i2.7293

Merzlyak, M.N., Solovchenko, A.E., Gitelson, A.A., 2003. Reflectance spectral features and non-destructive estimation of chlorophyll, carotenoid and anthocyanin content in apple fruit. Postharvest Biology and Technology, 27(2): 197-211. DOI: https://doi.org/10.1016/S0925-5214(02)00066-2

Merzlyak, M.N., Gitelson, A.A., Chivkunova, O.B., Solovchenko, A.E., Pogosyan, S.I., 2003. Application of Reflectance Spectroscopy for Analysis of Higher Plant Pigments. Russian Journal of Plant Physiology, 50(5): 704-710. DOI: https://1021-4437/03/5005

Naidu, R.A., Perry, E.M., Pierce, F.J. Mekuria, T., 2009. The potential of spectral reflectance technique for the detection of *Grapevine leafroll-associated virus-3* in two red-berried wine grape cultivars. Computers and Electronics in Agriculture, 66(1): 38-45. DOI: 10.1016/j.compag.2008.11.007

Nicoli, A., Zambolim, L., Trazilbo, P, Vieira, R.F., Teixeira H., Carneiro JE., 2013. Chlamydospore concentration for assessment of Fusarium root rot on common vean. Tropical Plant Pathology, 38(2): 149-151. DOI: 10.1590/S1982-56762013000200009

Nogués S., Cotxarrera L., Alegre L., Trillas M.I., 2002. Limitations to photosynthesis in tomato leaves induced by *Fusarium* wilt. New Phytologist. 154, 461-470. https://doi.org/10.1046/j.1469-8137.2002.00

Ochola, D., Ocimati, W., Tinzaara, W., Blomme, G., Karamura, E.B., 2015. Effects of water stress on the development of banana xanthomonas wilt disease. Plant Pathology, 64(3): 552-558.

Ortiz, E., Hoyos-Carvajal L., 2016. Standard methods for inoculations of *F. oxysporum* and *F. solani* in Passiflora. African Journal of Agricultural Research, 11(17): 1569-1575.

Pshibytko, N.L., Zenevich, L.A., Kabashnikova, L.F. 2006. Changes in the Photosynthetic Apparatus during Fusarium Wilt of Tomato. Russian Journal of Plant Physiology, 53(1): 25-31.

Riaz, T., Khan, S., Javaid, A., 2008. Effect of inoculum density on Fusarium corm rot disease of Gladiolus. Pak. J. Phytopathol, 20(2): 229-233.

Reis, A., Boiteux, L.S., 2007. Outbreak of Fusarium oxysporum f. sp. lycopersici race 3 in commercial fresh-market tomato fields in Rio de Janeiro State, Brazil. Horticultura Brasileira, 25(3): 451-454.

Richardson, A.D., Duigan, S.P., Berlyn, G.P., 2002. An evaluation of noninvasive methods to estimate foliar chlorophyll content. New Phytologist, 153(1): 185-194.

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.

Sims, D.A., Gamon, J.A., 2002. Relationship between leaf pigment content and spectral reflectance across a wide range species, leaf structures and development stages. Remote Sensing of Environment, 81(2-3): 337-354. https://doi.org/10.1016/S0034-4257(02)00010-X

Song, S., Gong, W., Zhu, B., Huang, X., 2011. Wavelength selection and spectral discrimination for paddy rice, with laboratory measurements of hyperspectral leaf reflectance. ISPRS Journal of Photogrammetry and Remote Sensing, 66(5): 672–682. DOI: https://doi.org/10.1016/j.isprsjprs.2011.05.002

Solovchenko, A.E., Chivkunova, O.B., Merzlyak, M.N., Gudkovsky, V.A., 2005. Relationships between chlorophyll and carotenoid pigments during on- and off-tree ripening of apple fruit as revealed non-destructively with reflectance spectroscopy. Postharvest Biology and Technology, 38(1): 9-17.

Strasser, R.J., Srivastava, A., Tsimilli-Michael, M., 2000. The fluorescence transient as a tool to characterize and screen photosynthetic samples. In: Yunus M, Pathre U, Mohanty P (eds), Probing Photosynthesis: Mechanisms, Regulation and Adaptation, 445-483. Taylor and Francis, London.

Stirbet A.· 2012. Chlorophyll a fluorescence induction: a personal perspective of the thermal phase, the J–I–P rise. Photosynth Res, 113(1-3):15-61. DOI: 10.1007/s11120-012-9754-5

Thomas, S., Thomas, K., Bohnenkamp D., Brugger, A., Alisaac E., Wahabzada, M., Behmann, J., Mahlein, A., 2017. Benefits of hyperspectral imaging for plant disease detection and plant protection: a technical perspective. Journal of Plant Diseases and Protection, 125(1): 5-20.

Yordanov, I., Tsonev, T., Goltsev, V., 1997. Gas Exchange and Chlorophyll Fluorescence during Water and High Temperature Stresses and Recovery. Probable Protective Effect of Carbamide Cytokinin 4-PU30. Photosynthetica, 33(3-4): 423-431.

Zhang, Q., Li, Q., Zhang, G., 2012. Rapid Determination of Leaf Water Content Using VIS/NIR Spectroscopy Analysis with Wavelength Selection. Spectroscopy: An International Journal, 27(2): 93-105.

Zhang, J., Pub, R., Huanga, W., Yuana, L., Luoa, J., Wanga, J., 2012. Using in-situ hyperspectral data for detecting and discriminating yellow rust disease from nutrient stresses. Field Crops Research, 134(1): 165-174. DOI: https://doi.org/10.1016/j.fcr.2012.05.011

Zhang, M., Qin, Z., Liu, X., Ustin, S.L., 2003. Detection of stress in tomatoes induced by late blight disease in California, USA, using hyperspectral remote sensing. International Journal of Applied Earth Observation and Geoinformation, 4(4): 295-310. DOI: 10.1016/S0303-2434(03)00008-4

CHAPTER 4. BINOMIAL LOGISTIC REGRESSION MODELS (BLRM) TO PREDICT *F. oxysporum* INFECTION IN TOMATO PLANTS DURING THE DISEASE INCUBATION PERIOD: A METHODOLOGICAL PROPOSAL

Abstract

The vascular wilt is a serious threat to a large number of economic crops. The evaluation of the disease incidence it is done visually, which makes it subjective and delayed, besides requiring the destruction of plants. The application of the Binomial Logistic Regression Models (BLRM) to predict Fusarium infection using reflectance data in the Visible (Vis) and near infrared (NIR) spectral range has not been attempted so far in any of its hosts. The field study was carried out during the asymptomatic period of the disease with two tomato varieties, one tolerant and one susceptible to all races of *F. oxysporum*, measuring the reflectance data every three dpi. We developed 16 BLRM, one model for each group data, which were highly significant ($p < 0.001$) and showed high goodness of fit after 6 dpi, represented by the Pseudo R^2 of McFadden > 0.5 and r2CU> 0.8. The assumptions that the BLRMs make about the data were verified: I) the dependent variable only has two categories: plant infected with *F. oxysporum* ("1") and healthy plants ("0"). II) Linear relationship between the logit function of the results and the three main predictor variables (R750, R550, and R430). III) Atypical values were selected and eliminated; IV) High correlations between predictors were decreased to the maximum using the generalized variance-inflation factors (vif). The variables chosen using the proposed methodology were significant in most models, except for R430 in I_{3dpi} and I_{9dpi}, R445 and R750 in I_{9dpi}, and R550 in I_{15dpi} on susceptible plants. In the models developed from reflectance data in tolerant plants only the variables R550 in I_{3dpi}, R970 in I_{3dpi}, and R704 in I_{9dpi} were not significant. According to the areas under the curve (AUC) obtained, the BLRMs generated from the reflectance in the tolerant plants have a higher prediction yield, surpassing AUCs of 0.9, and 0.8 in susceptible plants after 9 dpi. The BLRMs models developed in this study from reflectance data in the Vis/NIR range have a potential use for rapid detection and non-destructive estimation of vascular wilt incidence in tomato cultivars during the disease incubation period.

Keywords: Vascular wilt, spectral reflectancia, Binomial Logistic Regression Models, plant disease, infection prediction.

4.1 introduction

Vis/NIR applications are based on calibration models, which stablish a mathematical relationship between absorption or reflectance spectra and the interest factors. These models require measurements of sample's spectra of a population that includes all the possible variation for future prediction, considering the population as the set of all the measurements that cover the characteristics of the sample (Cao 2013). The potential of different statistical methods for the spectral data modeling in order to estimate physical and chemical properties can only be carried out if the properties studied are

dependent on molecular structure. This is because changes in molecular structure are reflect in the spectra, and linearly related to spectral intensities (Haaland and Thomas 1998).

Although different multivariate statistical tools have been used to perform calibration models applied to the quantitative analysis of ultraviolet, visible and near infrared, such as classical least squares modeling (CLS), inverse least squares (ILS) and Principal Components Regression (PCR). It is perhaps the methodology for Partial Least Square (PLS) one of the most widely used in prediction models with good levels of sensitivity and accuracy in detection, discrimination and quantification of plants disease (Wu et al. 2008; Song et al. 2011), furthermore the damage caused by insects (Couture et al. 2013; Ranjitha et al. 2014). This is because the PLS is related to the regression of principal components, but instead of finding hyperplanes of minimum variance between the response variable and the independent variables. A linear regression is found through the projection of the prediction variables and the observable variables to a new space, so it has some advantages over other methods for the analysis of spectral data. Although CLS modeling offers more qualitative information, PLS modeling can offer important information used to allocate bands and identify unexpected components in models.

The presence, discrimination and quantification of diseases in plants can be determined with good precision by means of multivariate calibration models of the spectral data based on plant-pathogen interaction (Martinelli et al. 2016). Most published studies report models performed from diseases in plants that develop local symptoms (Abu-Khalaf and Salman 2014; Krezhova et al. 2014; Lu et al. 2018). Local symptoms generated by physiological or structural changes within a limited area of the host tissue, such leaf spots, gill and cankers and for its measurement, the severity is calculated. This method of evaluation is a subjective visual estimate in which the level of infection in a given plant is established based on the amount of diseased tissue, so it refers to the percentage of the affected area of a specific organ of plant. In studies in which spectroscopy is applied in the modeling of infection of local diseases, the response variable is usually the percentage of severity, and the absorbance or reflectance at certain wavelengths are the explanatory variables (Wang et al. 2016).

However, studies focused on infection modeling in plants systemic diseases are less common, perhaps because their measurement is done by calculating the percentage of diseased plants throughout the crop (incidence) and should only be used for diseases that affect the entire plant. That is, if a specific measurement is made on affected organ with a systemic disease, that plant is diseased or healthy (a percentage of infection is not considered) (Leonberger et al. 2016). For this reason, in modeling

infection by systemic diseases in plants, the modeling techniques mentioned previously (CLS, ILS, PCR, PLS) are not recommended. For this purpose, the use of logistic regression methods is recommended, since it is one of the best-known techniques used to model a categorical response variable based on continuous or categorical predictor variables.

For the particular case of systemic diseases, Binomial Logistic Regression Models (BLRM) are more appropriate, since they predict the probability that an observation is assigned in one of the two categories of a dependent dichotomous variable, based on one or more independent variables that can be continuous or categorical, according to equation:

$$P(Y) = \frac{e^{b0+ b1x1+ b2x2+\ldots+ bnxn}}{1+ e^{b0+ b1x1+ b2x2+\ldots+ bnxn}}$$

Where P: probability of occurrence of Y, X_n: predictive variable, b_1: gradient line; b_n: Regression coefficient of X_n, *e*: logarithm in base *e*

Traditionally, nonlinear indices have been used in plants epidemiology, most often using linearized transformations or non-linear CLS (Macchiavelli et al. 2004). However, these results assume that the disease index has a normal distribution, that they are independent and that they have a constant variation, which cannot verified in the disease indexes. Otherwise, the BLRMs assume that the response variable must follow a binomial distribution, linear relationship between the independent variables and the link function (logit), absence of extreme values and low correlations between the predictors (Park 2013). However, the BLRMs do not assume normality of the data, although it is advisable to perform a normality analysis of the residuals of each model to avoid Type I error inflation (rejecting the null hypothesis when it is true in the population). Wherefore, a methodology is proposed to generate binary logistic regression models that allow predicting the incidence of Fusarium infection in plants based on reflectance spectra, this in order to perform a non-destructive diagnosis when the plants are in disease asymptomatic period.

4.2 Methodology

4.2.1 Study area and biological assays

The study was conducted between 2017-2018 with tomato plants of Ponderosa varieties, which is susceptible to all races of *F. oxysporum* (Reis and Boiteu 2007), and the Santa Cruz cultivar that is tolerant to races 1 and 2, maintained under semi-controlled greenhouse conditions at the National

University of Colombia, Medellín (Antioquia, Colombia). A completely randomized design carried out and used to compare two treatments: I) tomato plants var. Ponderosa inoculated with *F. oxysporum* (Fo5), II) tomato plants var. Santa Cruz inoculated. Additionally, uninfected control plants of each cultivar were used. More information about the experimental design, inoculation process and pathogenicity tests are widely described in chapters 2 and 3 of this manuscript.

4.2.2 Variable selection and model construction

In the previous chapters, a set of relevant specific wavelengths (RSW) were identified for Fo5 infection during the incubation period of the disease, using a classification algorithm (RELIEF) and relating changes in specific spectral bands with physiological changes associated with *F. oxysporum* infection in tomato during the incubation period. The selection of the wavelengths (explanatory variables) for Binomial Logistic Regression Models (BLRM) was done with the use of biplots and loading plots for Principal Components Analysis (PCA) for each data group without previous association. The data obtained from healthy and infected plants, with *F. oxysporum* in two tomato varieties on 27 variables (selected wavelengths). The training data for the construction of the models were obtained from the measurement of 5 reflectance spectra in each leaf of 25 plants per treatment, that is, 500 spectra in total. Using the methodology of selecting variables described above, 16 BLRMs were developed (Table 2), a model for each group of data that were taken every three days during the disease incubation. The Likelihood Ratio Test (LRTest) used to verify models significance and then effects of each predictor (explanatory variable) in the BLRMs were evaluated. Additionally, the Akaike Information Criterion (AIC) was calculated as a relative quality measure of models for a set of given data, and the pseudo R^2 of McFadden and Ragg and Uhler (r2CU), to evaluate the goodness of fit of models. All statistical analyses were performed using R Software, the main libraries and functions used are summarized in table 4.

Table 4. Description of main analyzes used in BLRMs development with R software

Analysis	Library	Function	Description
Pre processing	mdatools, prospectr	prep.snv, gapDer	Apply the transformations: standard normal variate (SNV) and derivatives of different orders to the rows of the data matrix
PCA	kazaam	Prcomp	Performs the principal component analysis on the data matrix by taking the singular value decomposition (SVD)
Models fit	stats	Glm, AIC	glm is used to fit generalized linear models; AIC calculating Akaike's Information Criterion

Models significance	lmtest	Lrtest	Generic function for comparisons of models via asymptotic likelihood ratio tests
goodness of fit	pscl	pR2	Compute various pseudo-R2 measures for various generalized linear models (GLMs)
LRGMs assumptions	car	Vif, residuals, kurtosis, plot	Functions to evaluate the assumptions for the data of GLMs
Make predictions	Raster	Predict	Generic function for predictions from the results of various model fitting functions
AUC calculation	ROCR	performance	Predictor evaluations are performed
Data visualization	mdatools, ggplot2	mdaplot, ggplot, plot	Functions used for data visualization (scatter plots, bars, histograms)

4.2.3 Verification of BLRM assumptions

BLRM makes several assumptions about the data and its verification is essential to build a good model. First, the response variable in this work has two possibilities: "1" (plants infected with *F. oxysporum*) and "0" (healthy plants) which realizes the first assumption, which requires to be the result of binary or dichotomous variable like “yes” vs “no”, “positive” vs “negative”, “1” vs “0”. To validate the second assumption, graphs were made of results logit function and three predictor variables that explained the pronounced variability of data in each model, since this requires linearity between them (Logit (p) = log (p / (1-p)), where p is the probability of result. To comply with the third assumption of BLRMs (values or atypical should not be taken), those values whose leverage was more than three times the critical value threshold in residual value exploration in an individualized manner were marked and eliminated (data not shown). Finally, the generalized variance-inflation factors (vif) were calculated at time of inserting each explanatory variable "one by one" to avoid high correlations between the predictors.

4.2.4 Evaluation of predictive capacity and model performance

To evaluate the predictive capacity of models, probabilities were generated in form of P (I = 1 | R) (I = variable response, R = explanatory variable), and a high threshold, which was 0.8. Therefore, if P (I = 1 | R)> 0.8 then I = 1, otherwise I = 0. To perform the corresponding validation in BLRMs, test data from a sample of susceptible and tolerant tomato plants (with their respective controls) taken independently (no cross-validation) in an experiment performed at different time from used to obtain the training data. The curves of Receiver Operating Characteristic (ROC) type were drawn, and the corresponding areas under the curve (AUC), and precisions were calculated, which are the typical measures performance for a binary classifier.

4.3 Results

The Figure 13 plots an example of variable selection process, in which a biplot (Fig. 13A) and the "loading plot" of the Main Component 1 (PC1) (Fig. 13B) to select the variables with highest loadings and incorporate them into the model. Once these variables are incorporated in PC1, the same procedure is use to variables select with the highest loadings in the first PCs (Fig. 13C and 13D) and incorporate them into the model “one by one”, as long as they meet the following criteria: A) the model remains significant. B) The new predictor variable is not highly correlated with the previously selected ones. C) Improve the model predictive capacity. In general, the reflectances at 550nm and 750nm (R550 and R750) had high positive charges in CP1 in all models performed, so this component may be related to the variables that determine plants infection. Other variables, such as R430, R445, R510, R704, and R970, were also selected using this method, although in general they explain a lower percentage of data variability in the PCAs. However, the relevance of model variables, their significance and predictive capacity of model with these variables will be discussed in detail below. To verify the second assumption, the logit function of the results Vs value of predictor variable was graphed.

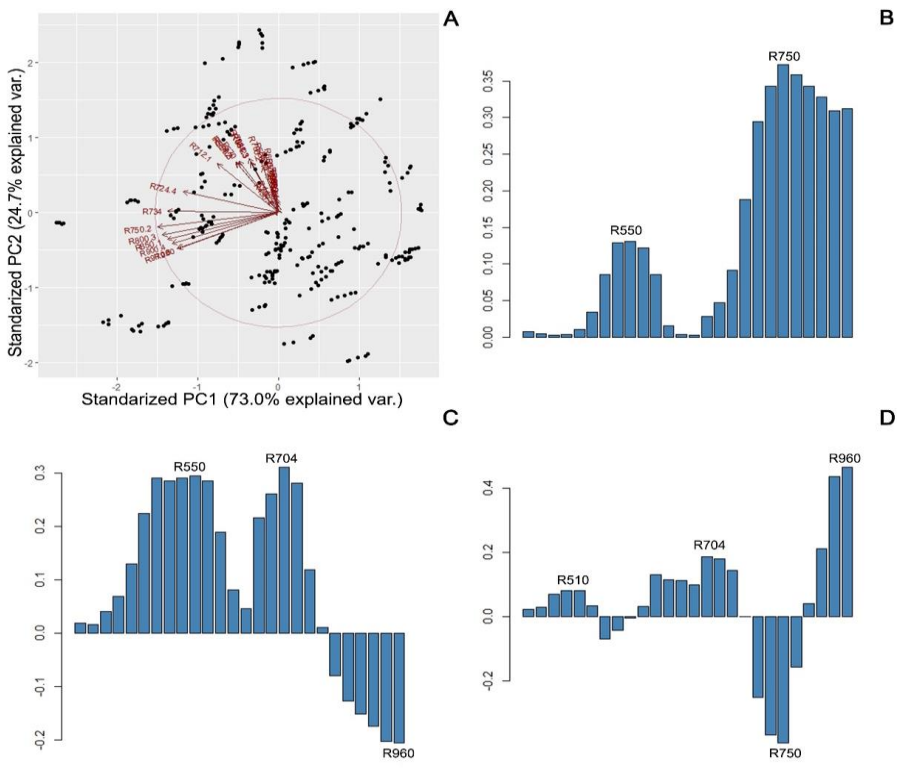

Figure 13. Biplot (A) and loading plots of Principal Component Analysis (PCA, A: PC1, B: PC2, C: PC3) with Vis/NIRs reflectance measures without grouping of infected (12 dpi) and healthy plants of a susceptible tomato cultivar.

4.3.1 BLRMs to predict *F. oxysporum* infection in tomato plants during the disease incubation period

Using the variable selecting methodology previously described, 16 BLRMs were developed, a model for each group of data, taken every three days during the incubation period of disease. Table 5 summarizes the developed models, denoting as "I" the response variable with an indicator of day after the inoculation (DPI) in which the reflectance data was taken, which has two possibilities: "1" (plants) infected with *F. oxysporum*) and "0" (healthy plants). The explanatory variables are denoted by the capital letter R followed by the corresponding wavelength value.

To analyze the goodness of fit of logistic regression models (LRM) there is no statistical equivalent to R^2. However, several pseudo R^2 have been developed, which receive the name since they are on a scale similar to R^2 (although some never reach "0" or "1"), but cannot be interpreted as an R^2 in the approach of the OLS. A clear example of this point is the pseudo R^2 of McFadden, which considered having good goodness of fit of LRM when it presents values between 0.2-0.4. Then we can consider that the models of both varieties in our study show good goodness of fit from the six dpi, which corroborated with pseudo R^2 of Cragg and Uhler (r2CU), which if it reaches values of 0-1. Additionally, the Likelihood Ratio Test (LRTest) used to verify the significance of models, which were highly significant (Table 5). It must be taken into account that this step is very important, since, if the global model were not significant, the effects of each predictor or explanatory variable would not be revised.

Table 5. Results of data adjustment to each binomial logistic regression model to predict *F. oxysporum* infection in tomato plants (susceptible and tolerant) using Vis/NIRs spectral data during the incubation period of the disease

	Equations	Pseudo R^2		Likelihood Ratio Test
DPI	**Susceptible plants**	**McFadden**	**r2CU**	**LRtest**
0	$I_{0dpi} = 3.86 + 6.42*R_{430} - 8.97*R_{484} + 0.25*R_{550} + -0.12*R_{750}$	0.24	0.66	***
3	$I_{3dpi} = -0.58 + 1.43*R_{430} - 3.70*R_{445} + 1.33*R_{510} - 0.55*R_{550} + 0.13*R_{750}$	0.16	0.80	***
6	$I_{6dpi} = -30.09 + 16.13*R_{430} - 4.69*R_{510} + 1.22*R_{564} - 0.19*R_{704} + 0.10*R_{750}$	0.52	0.69	***
9	$I_{9dpi} = -23.07 + 2.63*R_{430} + 4.45*R_{445} + -0.64*R_{510} + 0.30*R_{550} + 0.01*R_{750} + 0.15*R_{970}$	0.50	0.83	***
12	$I_{12dpi} = -94.34 - 2.47*R_{430} - 6.19*R_{445} + 1.82*R_{550} - 0.08*R_{750} + 0.85*R_{970}$	0.79	0.88	***
15	$I_{15dpi} = -103.75 + 15.01*R_{445} + 0.36*R_{550} + 1.09*R_{750}$	0.79	0.80	***
18	$I_{18dpi} = -18.32 + 12.19*R_{430} + 1.36*R_{550} - 1.28R_{704} +$	0.67	0.88	***

	$0.05*R_{750}$			
21	$I_{21dpi} = -20.71 - 11.25*R_{445} + 1.60*R_{550} + -0.04*R_{704} - 0.28*R_{750}$	0.67	0.77	***
	Tolerant plants			
0	$I_{0dpi} = -2.69 - 2.07*R_{430} + 0.01*R_{550} + 0.04*R_{750}$	0.04	0.08	**
3	$I_{3dpi} = -2.92 - 3.04*R_{430} - 0.01*R_{550} + 0.10*R_{750} - 0.04*R_{970}$	0.14	0.24	***
6	$I_{6dpi} = -8.36 + 0.87*R_{550} - 1.28*R_{650} - 0.09*R_{750}$	0.31	0.47	***
9	$I_{9dpi} = -20.43 - 5.75*R_{430} + 1.03*R_{550} -0.25*_{704} + 0.15*R_{750}$	0.62	0.77	***
12	$I_{12dpi} = -40.27 + 12.74*R_{430} + 3.35*R_{500} + 0.46*R_{750}$	0.87	0.94	***
15	$I_{15dpi} = -34.65 + 8.88*R_{430} + 0.66*R_{704} + 0.37*R_{750}$	0.87	0.94	***
18	$I_{18dpi} = -12.00 – 4.43*R_{430} + 4.85*R_{510} – 0.42*R_{704} + 0.20*R_{750}$	0.73	0.92	***
21	$I_{21dpi} = -66.57 + 5.07*R_{510} + 1.06*R_{550} + 0.51*R_{750}$	0.83	0.91	***

Significance codes: *p<0.05; **p<0.01; ***p<0.001

4.3.2 Verification of BLRM assumptions in reflectance data

The BLRMs makes several assumptions about the data. This chapter describes the main assumptions and provides a practical guide to verify if these assumptions are true for data used, which is essential to build a successful model. The first assumption requires that dependent variable be binary (dichotomous, "dummy"). In the particular case of this study, the dependent variable has only two categories: plant infected with *F. oxysporum* ("1") and healthy plants ("0"). To pass the second assumption there must be a linear relationship between the logit and the variables of each predictor. Figure 14 shows the relationship between the logit function of the results and the three-predictor variables that explained the greater variability of data in each model. In models made for days at 0 and 3 dpi for susceptible plants, the variables tend to move in the same relative direction, but not at a constant rate, that is, they have a "monotonous" relationship (the linear relationships are also monotonous). In addition, there are high dispersions in models made during this period. After 6 dpi, there are clear relationships between logit function and predictor variables, in addition to low dispersion of data during the incubation period of disease, except the R750 in model made for 21 dpi that presents a "curve pattern". The BLRM performed for tolerant plants follow the same pattern as the susceptible described above, with "monotone" relationships and high dispersion of data in relationships (logit-predictive variables) during the first week. In addition to clear linear relationships and low data dispersion in the models compared with the rest of incubation period, except for some particular cases (R750 in I_{6ddi}, R430 in I_{9dpi}, R430 and R750 in I_{18dpi}).

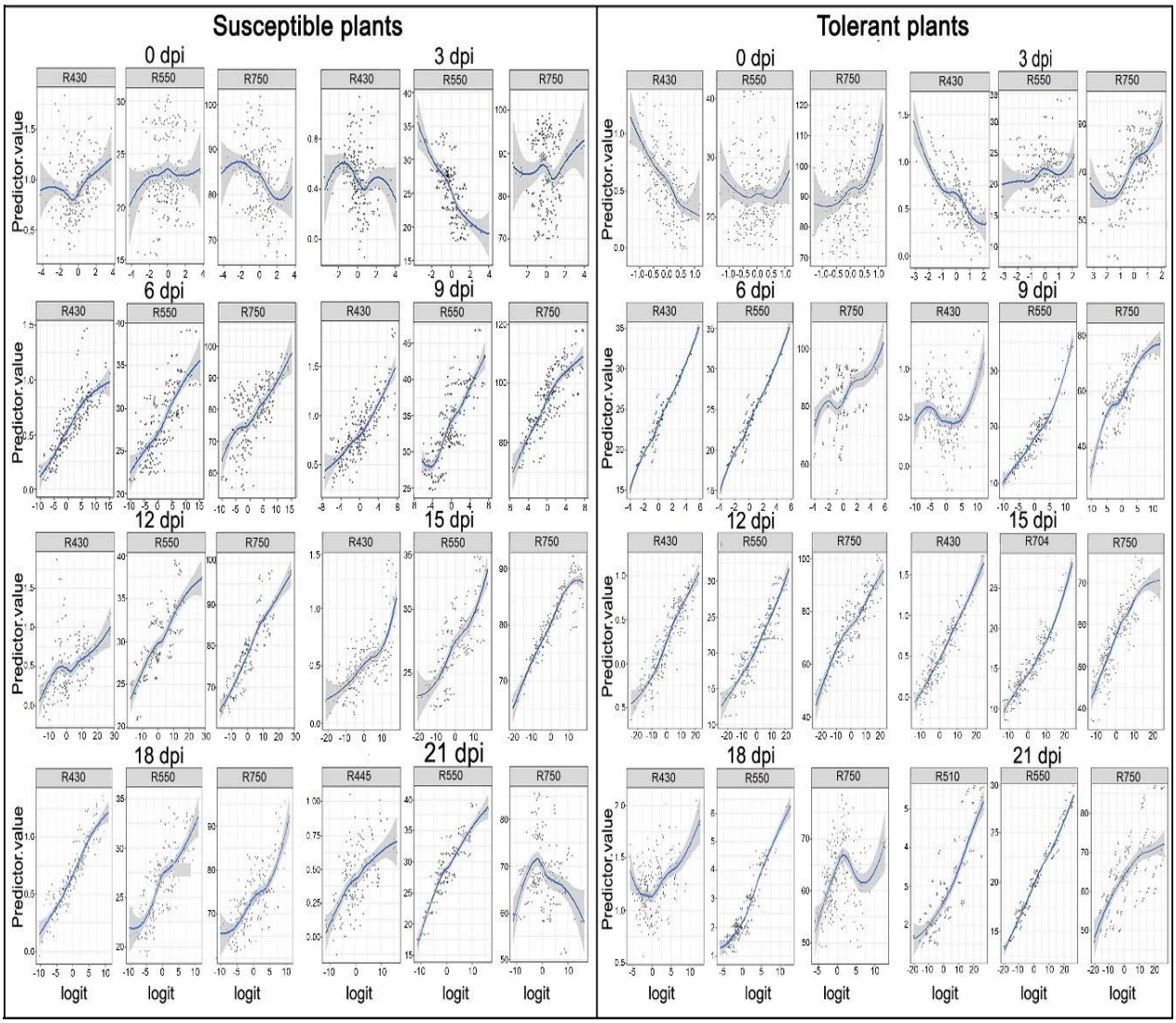

Figure 14. Relationship between the "logit" function of the result and three predictor variables that explain greater variability in each model

Before proceeding with the meanings of predictor variables, the overdispersion coefficients (ϕ) estimated, which were used to recalculate the significance (ϕ should be ~ 1). The final models for both varieties obtained ϕ values between 0.2-1.0 (Table 6). Finally, vifs calculated at time of inserting "one by one" each explanatory variable to avoid high correlations between the predictors. As a general criterion, variables were included in the different models when the magnitude of vif<10, with some point exceptions that showed high values of vif between two pairs of variables: R550-R704 for the BLRM in susceptible plants of 18 dpi, and R510 -R704 for BLRM in tolerant plants of the same day (Table 6).

Table 6. Overdispersion coefficients (ϕ) and generalized variance-inflation factors (vif)

DPI	ϕ	Susceptible plants Vif								
		R430	**R445**	**R484**	**R510**	**R550**	**R650**	**R704**	**R750**	**R970**
0	1.00	3.07		3.56		1.80			1.45	
3	0.99	3.31	3.83		5.80	7.51			2.97	
6	0.44	2.47	2.85		11.40	8.74		2.14	1.72	
9	0.69	2.30	2.75		9.30	6.93			8.04	7.46
12	0.37	7.11	11.33			8.13			4.49	7.11
15	0.37	1.46				1.02			1.48	
18	0.63	1.49				16.43		17.33	1.38	
21	0.66		3.90			5.94		2.06	2.61	
DPI	**ϕ**	**Tolerant plants Vif**								
		R430	**R445**	**R484**	**R510**	**R550**	**R650**	**R704**	**R750**	**R970**
0	1.02	2.13				3.25			2.62	
3	1.03	1.60				2.64			9.05	6.79
6	0.93					5.06	5.32		3.76	
9	0.61	2.01				6.57		5.65	1.18	
12	0.24	3.12				1.82			3.72	
15	0.24	1.20						1.02	1.21	
18	0.75	1.66			14.70			14.51	2.61	
21	0.23				4.26	1.22			3.85	

The logistic regression does not assume the assumption of normality of the data. Even so, a residuals normality analysis of each model was performed (data not shown), since if we do not have normality the Type I error can be inflating. It was confirmed that residuals of deviance (measure of the residual variation of a model) are centered at zero, and there is no predominance of negative or positive residual values, confirming the homoscedasticity of the models. The distribution of the residues is not biased, shows positive value and bias on the right, with low values of kurtosis (positive value close to 3); this slight bias does not substantially alter the critical alpha values of significance. Finally, the normal probability graph for each BLRM evaluated, in which the observed empirical data represented against the data that would obtained in a theoretical normal distribution, which confirmed that there is a slight deviation from the normal canonical assumptions.

4.3.3 Evaluation of explanatory variables significance

Because the models are highly significant at global level, we can proceed to evaluate significance of explanatory variables. First, we can see that the variables chosen are statistically significant in most models, except R430 (in I_3dpi and I_9dpi), R445 and R750 (in I_9dpi), and R550 (I_{15}dpi) in susceptible plants

(Table 7A). In the models developed from reflectance data in tolerant plants only the variables R550 (in I_0dpi, I_3dpi), R970 (in I_3dpi), R704 (in I9dpi) and R750 (in I_0dpi) were not significant (Table 6B). Regarding the statistically significant variables, in general R550 had the values of *p* lower in susceptible plants, whereas R750 had the lowest in the tolerant cultivar. his suggests a strong association between reflectance at 550nm and 750nm, and the probability that the plant is infected, but does not necessarily mean a high predictive capacity of models (this will be discussed later in this text). The positive coefficient for these two predictors in most models indicates infection in plants. At this point, it is important to remember that in the logit model the response variable is the probability of the record:

$$ln\left(\frac{p}{1-p}\right) = a * X1 + b * X2 + \cdots + Z * Xn$$

By performing a more parameters detailed analysis of each predictive variable, it can observed that an increase in one unit of reflectances at 550nm and 750nm increases the probability of record in small quantities compared to other variables. Conversely, an increase in one unit of reflectance to 430nm greatly increases the registration probabilities, as can be seen in the models for plants of the susceptible cultivar: 6.42 (0) dpi, 16.13 (6 dpi), 12.19 (12 dpi). Despite the model and highly significant variables, we cannot yet say anything about the models quality, compliance with assumptions of data used and their predictive capacity. Regarding the relative quality of the BLRM, we can see that the AICs are lower after the 12 DPI, which supposes a higher quality of the models generated from the explanatory variables selected in this stage of disease incubation period (Table 7).

Table 7. Coefficients summary of binomial logistic regression models to predict *F. oxysporum* infection in tomato plants (susceptible and tolerant) using Vis/NIRs spectral data during the incubation period of the disease

	Dependent variable (Infection)							
	A. Susceptible plants							
Wavelength	**0 DPI**	**3 DPI**	**6 DPI**	**9 DPI**	**12 DPI**	**15 DPI**	**18 DPI**	**21 DPI**
R430	6.42**	1.43	16.13**	2.63	-2.47		12.19**	
R445		-3.70*		4.45	-6.19	15.01**		-11.25**
R484	-8.97**							
R510		1.33**	-4.69**	-0.64				
R550	0.25	-0.55**		0.30*	1.82**	0.36*	1.36**	1.60**
R564			1.22**					
R704			-0.19				-1.28**	-0.04
R750	-0.12	0.13**	0.10*	0.01	-0.08	1.09**	0.05	-0.28**
R970				0.15	0.85***			

Constant	3.86*	-0.58	-30.09**	-23.07**	-94.34**	-103.75**	-18.32**	-20.71**
Observations	235	217	223	243	193	180	199	183
AIC	256.91	266.00	205.34	182.87	67.70	47.67	101.23	94.08
				B. Tolerant plants				
Wavelength	0 DPI	3 DPI	6 DPI	9 DPI	12 DPI	15 DPI	18 DPI	21 DPI
R430	-2.07**	-3.04**		-5.75**	12.74**	8.88**	-4.43**	
R500					3.35*			
R510							4.85**	5.07**
R550	0.01	-0.01	0.87**	1.03**				1.06*
R650			-1.28**					
R704				-0.25		0.66*	-0.42	
R750	0.04	0.10**	-0.09**	0.15**	0.46**	0.37**	0.20**	0.51**
R970		-0.04						
Constant	-2.69*	-2.92**	-8.36**	-20.43**	-40.27**	-34.65**	-12.00**	-66.57**
Observations	228	228	235	229	225	236	231	228
AIC	310.245	280.949	232.665	130.074	47.292	64.387	162.349	62.98

Significante codes: *p<0.05, **p<0.001

4.3.4 Evaluation of predictive capacity and models performance

Although so far it has been proven that BLRMs describe well the reflectance observations set in healthy plants and infected with *F. oxysporum*, it is clear that most researchers are more interested in accuracy of the predictions than in the goodness of fit (White 2013). In the previous steps, the adaptation of the BLRM was evaluated, now you can see how those models are performing when predicting on a new data set. The accuracies obtained after the sixth dpi were greater than 80% (except for I_{15} of the susceptible cultivar in which it was 0.79), even with values over 90% in tolerant cultivar. At this point, be aware of that these results depend to some extent on the test data origin and validation mechanism.

The ROCs constructed from True Positive Rate (TPR) Vs False Positive Rate (FPR) (Fig. 15). It should be remembered that the TPR defines how many correct positive results occur among all the positive samples available during the test (equivalent to sensitivity), while the FPR defines how many incorrect positive results occur among all negative samples available during test (1- specificity). The best possible prediction method would produce a point in the upper left corner or coordinate (0.1) of the ROC space, which represents 100% sensitivity (without false negatives) and 100% specificity (without false positives). A random assumption would give a point along a diagonal line (gray line in Fig. 16). The ROC graphs of models made from reflectance data at 0 dpi and 3 dpi are very close to the line of non-discrimination and have low values of AUC, which represents high randomness of the test and a Low

efficiency of models for correctly classify the observations during this period. On the contrary, the ROC curves of BLRMs for the tolerant cultivar showed maximum sensitivity (100%) after 6 dpi, with low FPRs (high specificity); in addition to AUC values greater than 0.9 after 9 dpi. The ROC curves of susceptible cultivar also reached high sensitivities, but with higher FPRs than in the tolerant cultivar (lower specificity), mainly between 12 dpi-21dpi. In general, their AUC were lower than previous cultivar, but with also high levels, greater than 0.8 (except ROC at 21 dpi).

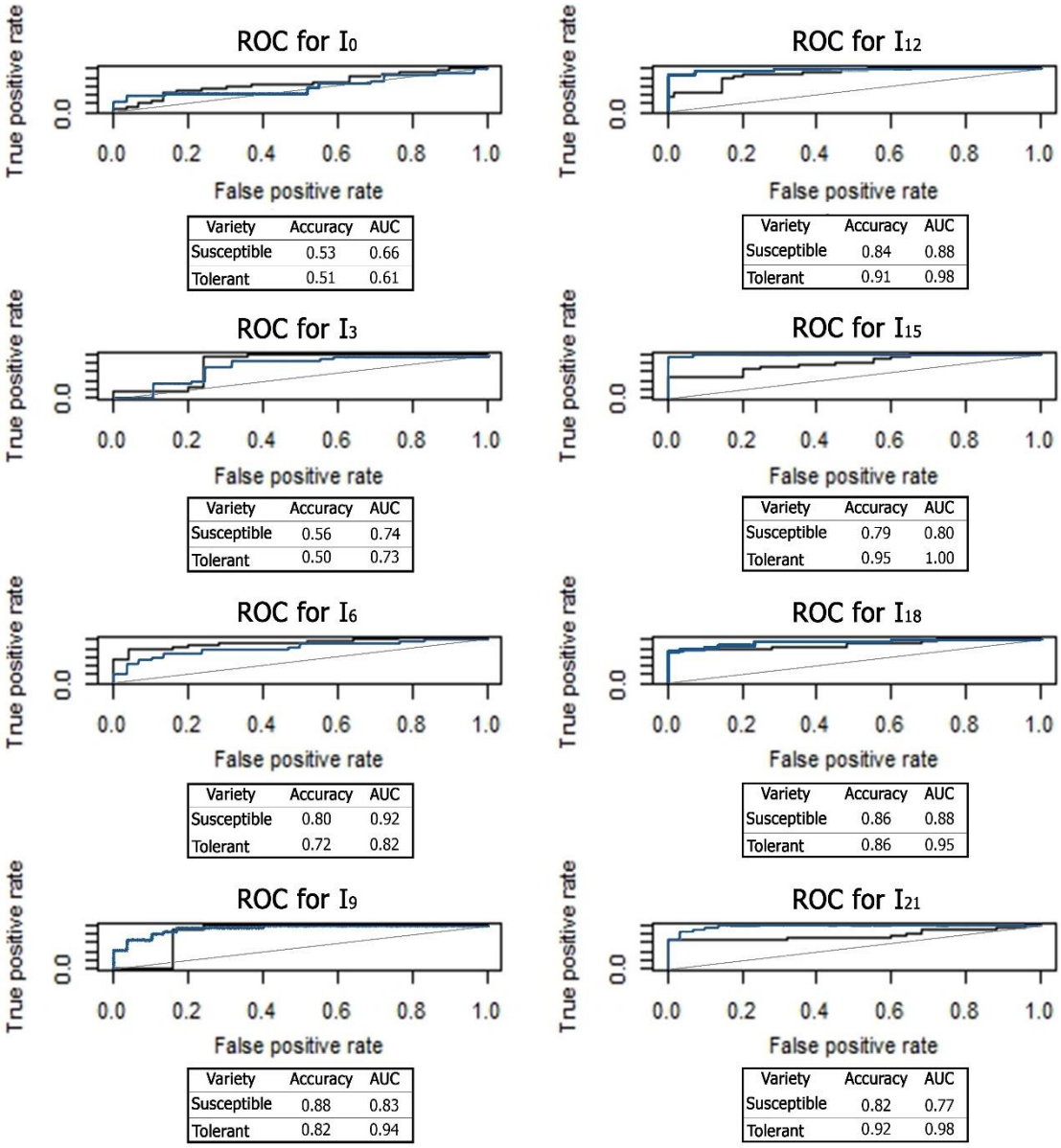

Figure 15. Receiver Operating Characteristic (ROC) curves of reflectance predictors to predict infection by *F. oxysporum* in two tomato varieties, susceptible (black line) and toleran (blue line). Diagonal line of no-discrimination (gray line)

4.4 Discussion

The absorbance and/or reflectance data in Vis/NIR ranges have used for models development in a wide range of plants applications. Such as the forage quality determination, chlorophylls and carotenes concentration, estimation of seeds oil content and total antioxidant capacity, among others (Hărmănescu et al. 2008; Ding and Fuchigami 2009; Elfadl et al. 2010; Asekova et al. 2015; Marin et al. 2018). The use of spectroscopy in plants diseases detection has carried out mainly in diseases with local symptoms from the so-called "Disease indices", which have the disadvantage of not being linear, adjusting most of the time with linearized transformations (White 2013). Specifically in infections with Fusarium, soft independent modeling of class analogy (SIMCA), neuronal and parametric classifiers have used for pathogen identification in corn grains, in addition to use LDA for detection of *F. oxysporum* isolates (Draganova et al. 2010; Salman et al. 2012). However, studies that describe the Vis/NIR application to predict systemic infections, and specifically vascular wilt in plants are very limited.

In this study, the Vis/NIR models developed to determine *F. oxysporum* infection during vascular wilt incubation period, when disease symptoms are not yet visible. The BLRMs constructed from described variables selection in the methodology were highly significant. The selected predictor variables that explained the highest percentage of data variability were R550 and R750, which were highly significant in most models. However, it noted that R550 was significant in most of BLRMs performed for susceptible cultivar during the incubation period, while R750 was significant mainly from the BLRMs for tolerant cultivar. The explanatory variable R430 was also significant in most models, but its percentage of data explained variation was very low. It should be taken into account that the test performed is a difference of null model residual deviances (the one that does not introduce any predictor variable, only the ordered one in the origin that coincides with the average of the answer) and interest model (Park 2013). Therefore, model significance does not imply that original data fulfill the assumptions for BLRM and/or that model has good predictive capacity. At this point, the question "which are the best models?" may arise, since each BLRM model developed in this study was made from a different data set, so it could not be compared with traditional indicators. So, in this work the Akaike Information Criterion (AIC) was used as relative quality measure of models for a set of given data, even with models made from different data groups. Thus, the AIC results in models of both varieties was quite consistent, as the BLRMs values were lower after 9 dpi, when the physiological response of the plant to the pathogen increased its intensity (see chapter 3).

In general, the data adjustment to the BLRM was good in both varieties, presenting values greater than 0.5 (McFadden) and 0.8 (r2CU) in the Pseudo R^2 of models made after 6 dpi; however, the Pseudo R^2 values tend to be slightly higher on tolerant cultivar models. Unfortunately, to date there are no other reports on the BLRMs performance to determine infection by *F. oxysporum*, but the Pseudo R^2 obtained are consistent (even superior) with R^2 homologs in multivariate models and traditional indexes applied to plants diseases detection (Song et al. 2011; Luo et al. 2013; Krishna et al. 2014; Zhang et al. 2014). It is important to note that the model estimates from a logistic regression are maximum likelihood estimates obtained through an iterative process that are not calculated to minimize the variance, so the ordinary least squares (OLS) approach is not applied (Czepiel 2019). For this reason, to logistic models goodness of fit evaluate several pseudo R^2 have been developed; they are called in this way since they are measured on similar scale (although some never reach "0" or "1"), but they cannot be interpreted as an R^2 in the strict approach of OLS.

The adjustment of reflectance data to BLRMs asks for verification of some assumptions, and it is important to discuss them for well-adjusted models development. The first assumption requires that dependent variable be binary, or in a practical way in this study, each reflectance spectrum measured on the sheet could only have two options as to its origin. Can be taken from a point in a sick plant that responds biochemically and physiologically to the pathogen (with or without visual symptoms), or be measured from a healthy plant. By definition, systemic diseases are those that alter the normal physiological function of an organ or whole organism, and specifically on infections caused by *F. oxysporum* it has been demonstrated that plant resistance to disease is based precisely on activation of plants systemic defense mechanisms (He et al. 2002). It is important to note that reflectance at selected wavelengths (explanatory variables) obtained linear relationships with link function (logit), that is, the second assumption (linearity) is fulfilled. This is very important, since we can discard some problems that can be generated by its non-compliance (specification error): omission of important independent variables, inclusion of irrelevant independent variables, an incorrect functional form, changing parameters and that dependent variable can be part of simultaneous equations system (Sapra 2005). Checking the third assumption requires the atypical and influential cases detection, and their subsequent treatment is a crucial task in any modeling exercise. It is common to find atypical data in the reflectance measurements used to develop BLRMs caused by environmental noise and human errors, which can involve large residues and often have marked effects on linear maximum likelihood predictor (Sarkar 2011). The "manual" process performed to select the variables comparing the vif between

explanatory variables of models is also important to reduce the collinearity that commonly exists in hyperspectral data (Maimaitiyiming et al. 2017), since the collinear predictor, variables cause unstable estimates and inaccurate variations that affect confidence intervals and hypothesis testing. Also inflating the variances of parameter estimates and, consequently, causing incorrect inferences about the relations between the explanatory and response variables (Midi et al. 2013).

The models developed in this study were restricted to short periods of time (every three days), while the symptoms of the disease were not visible. The lowest predictive capacity in models obtained with reflectance data in first week after infection is related to physiological low response of plant to *F. oxysporum* during this period, which has been supported in this study in previous chapter and by others authors (Michielse and Rep 2009; Zvirin et al. 2010). According to the areas under the AUC obtained, the BLRMs generated has a higher prediction yield, exceeding an AUC of 0.8 in susceptible plants and 0.9 for tolerant plants after 6 dpi. One possible reason for not obtaining higher values is that the wavelengths selected as predictors for BLRMs are in the spectral range of 400nm-1000nm and some important biomolecules in the plant-pathogen interaction have peaks of absorbance/reflectance NIR (Türker-Kaya and Huck 2017; Ozaki et al. 2018). The increase of the predictive variables in the models could also improve the values in the AUC, but the correlation between the multiple predictors used in this study would have caused problems in the adjustment of the model. A final point that draws attention is the higher prediction performance in the models developed from the reflectance data in tolerant plants. Since the plant-pathogen interaction is highly specific, each cultivar has particular physiological changes generated in the process of recognition of the pathogen and generates different polysaccharides important for its inhibition. These changes at different times of the incubation period may be causing differences in prediction yields (detailed discussion of this topic in Chapter 2).

The use of reflectance spectroscopy in the Vis / NIR in BLRM to predict the infection by *F. oxysporum* in tomato plants facilitates the objective, rapid and non-destructive estimation of the samples, which contrasts with other techniques for detecting diseases in plants (Sankaran et al. 2010). The methodology developed to generate the BLRM could be a viable technique to detect infection by *F. oxysporum* during the incubation period of the disease, when the symptoms are not yet visible. This research is the first step towards the application of BLRM based on reflectance data in the Vis/NIR range for the prediction of fungal diseases, which can be used as a basic input in the design of technological tools that allow the plant disease detection in real time.

4.5 Conclusion

In this chapter, we have succeeded applying the BLRM to the early detection of vascular wilt in tomato plants using reflectance data in the Vis/NIR range; additionally, a general methodology for calculating the adjustment has been provided. Several logistic regressions based on different combinations of predictive variables (reflectance at a specific wavelength) are shown in this study. By using the spectral data in BLRMs it was possible to predict incidence of wilt vascular in tomato plants with reasonable degree of accuracy (accuracy> 0.8, after 6 dpi). It was possible to create models with good predictive performance, mainly in the susceptible tomato cultivar evaluated, using nine identified variables (see table 5) with the selection method described above. However, the obtained results in this work suggest the infection in tomato plants can be predicted using three basic variables, R750, R550, and R430, wavelengths located in the upper limit of red, green and violet. Therefore, it is expected that by using some combinations of these three reflectances in BLRMs or in plant disease indices it should be possible to quickly examine a large number of tomato cultivars due to their reaction with *F. oxysporum* infection. The BLRMs committed fewer errors of identification when exceeding the 6 dpi, and those generated from the reflectance data in the tolerant cultivar were more efficient.

Finally, the results of this study provided valuable information for use of reflectance data for evaluation at organ and plant scales, which can be scaled for measurements made from remote sensors on aerial or satellite platforms for evaluation of large infected areas with vascular wilt. Subsequent studies should focus on the specific relationship of some important biomolecules in tomato-*F. oxysporum* interaction with reflectance in NIR ranges not covered in this project (100nm-2500nm).

References

Abu-Khalaf N., Salman M. Visible/Near infrared (VIS/NIR) spectroscopy and multivariate data analysis (MVDA) for identification and quantification of olive leaf spot (OLS) disease. Palestine Technical University Research Journal, 2(1):01-08

Asekova S., Han S., Choi H., Park S., Shin D., Kwon C., Shannon J.G., Lee J. 2015. Determination of forage quality by near-infrared reflectance spectroscopy in soybean. Turkish Journal of Agriculture and Forestry, 40(1):1-8. DOI: 10.3906/tar-1407-33

Czepiel, S. 2019. Maximum Likelihood Estimation of Logistic Regression Models: Theory and Implementation. https://czep.net/contact.htm

Ding P., Fuchigami L.H. 2009. Simple linear regression and reflectance sensitivity analysis used to determine the optimum wavelengths for the nondestructive assessment of chlorophyll in fresh leaves using spectral reflectance. Journal *of the* American Society for the Horticultural Sciences, 134(1):48–57.

Draganova T., Daskalov P., Tsonov R. 2010. An approach for identifying of Fusarium infected maize grains by spectral analysis in the visible and near infrared region, SIMCA models, parametric and neural classifiers. International Journal of Bioautomation, 14(2): 119–128.

Elfadl E., Reinbrechta C., Claupein W. 2010. Development of near infrared reflectance spectroscopy (NIRS) calibration model for estimation of oil content in a worldwide safflower germplasm collection. International Journal of Plant Production, 4 (4):1735-8043.

Koyshibayev M., Muminjanov H. 2016. Guidelines for monitoring diseases, pests and weeds in cereal crops. Food and Agriculture Organization of the United Nations (FAO). http:www.fao.org/3/a-i5550e.pdf

Hărmănescu M., Moisuc A., Drăgan S, Gergen I. 2008. The determination of total antioxidant capacity of medicinal plants from Romania by nir spectroscopy. Lucrări Ştiinţifice, 51(1):385:390.

He C.Y., Hsiang T., Woling D.J. 2002. Induction of systemic disease resistance and pathogen defence responses in *Asparagus officinalis* inoculated with nonpathogenic strains of *Fusarium oxysporum*. Plant Pathology, 51(2): 225-230. https://doi.org/10.1046/j.1365-3059.2002.00682.x.

Krezhova D., Dikova B., Maneva S. 2014. Ground based hyperspectral remote sensing for disease detection of tobacco plants. Bulgarian Journal of Agricultural Science, 20(5):1142-1150.

Krishnaa G., Sahooa R.N., Pargalb S., Guptaa V.K., Sinhac P., Bhagatd S., Saharane M.S., Singha R., Chattopadhyay C. 2014. Assessing wheat yellow rust disease through hyperspectral remote sensing. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information, XL(8): 9-12.

Leonberger K., Jackson K., Smith R., Gauthier N. 2016. Plant Diseases. Kentucky Master Gardener Manual Chapter 6. University of Kentucky College of Agriculture, Food and Environment, Lexington, KY, 40546.

Lu J., Ehsani R., Shi Y., de Castro A.I., Wang S. 2018. Detection of multi-tomato leaf diseases in different stages by using a spectral-based sensor. Scientific Reports**,** 8(1):1-8

Luo J., Huang W., Yuan L., Zhao C., Du S., Zhang J., Zhao J. 2013. Evaluation of spectral indices and continuous wavelet analysis to quantify aphid infestation in wheat. Precision Agriculture, 14(2):151-161.

Maimaitiyiming M., Vasit S., & Arianna B., Joseph W., Misha K. 2017. Early Detection of Plant Physiological Responses to Different Levels of Water Stress Using Reflectance Spectroscopy. Remote Sensing, 9(7):745-768. DOI: 10.3390/rs9070745

Marín-Ortiz J.C., Hoyos-Carvajal L.M., Botero-Fernández V. 2018. Detection of asymptomatic *Solanum lycopersicum* L. plants infected with *Fusarium oxysporum* using reflectance VIS spectroscopy. Colombian Journal of Horticultural Sciences, 12(2): 436-446. DOI: 10.17584/rcch.2018v12i2.7293

Martinelli F., Scalenghe R., Davino S., Panno S., Scuderi G., Ruisi P., Villa P., Stroppiana D., Boschetti M., Goulart L.R., et al. 2015. Advanced methods of plant disease detection. Agronomy for Sustainable Development, 35(1):1-25. DOI 10.1007/s13593-014-0246-1.

Michielse C., Rep M. 2009. Pathogen profile update: *Fusarium oxysporum*. Molecular Plant Pathology, 10(3), 311-324. DOI: 10.1111/J.1364-3703.2009.00538.X

Midi H., Saroje S., Sohel R. 2013. Collinearity diagnostics of binary logistic regression model. Journal of Interdisciplinary Mathematics. 13(3):253-267. 10.1080/09720502.2010.10700699.

Ozaki Y., Huck C. W., Ishigaki M., Ishikawa D., Ikehata A., Shinzawa H. 2018. Near-Infrared Spectroscopy in Biological Molecules and Tissues. Encyclopedia of Biophysics, 1-19. DOI: https://doi.org/10.1007/978-3-642-16712-6_138

Park H.A. 2013. An introduction to logistic regression: from basic concepts to interpretation with particular attention to nursing domain. Journal of Korean Academy of Nursing, 43(2):154-164.

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.

Salman A., Lapidot I., Pomerantz A., Tsror L., Shufan E., Moreh R., Mordechai S. 2012. Mahmoud Huleihel Identification of fungal hytopathogens using Fourier transform infraredattenuated total reflection spectroscopy and advanced statistical methods. Journal of Biomedical Optics, 17(1): 1-9

Sapra S. 2005. A regression error specification test (RESET) for generalized linear models. Economics Bulletin, AccessEcon, 3(1):1-6.

Sarkar S.K., Midi H., Rana S. 2011. Detection of Outliers and Influential Observations in Binary Logistic Regression: An Empirical Study. Journal of Applied Sciences, 11(1):26-35.

Song S., Bo Z., Huang X. 2011. Wavelength selection and spectral discrimination for paddy rice, with laboratory measurements of hyperspectral leaf reflectance. ISPRS Journal of Photogrammetry and Remote Sensing, 66(5): 672–682. https://doi.org/10.1016/j.isprsjprs.2011.05.002

Türker-Kaya S., Huck C. 2017. A Review of Mid-Infrared and Near-Infrared Imaging: Principles, Concepts and Applications in Plant Tissue Analysis. Molecules, 22(1):168. DOI:10.3390/molecules22010168

Wang H, Qin F, Ruan L, Wang R, Liu Q, et al. 2016. Identification and Severity Determination of Wheat Stripe Rust and Wheat Leaf Rust Based on Hyperspectral Data Acquired Using a Black-Paper-Based Measuring Method. PLOS ONE: 11(4):e0154648. DOI: https://doi.org/10.1371/journal.pone.0154648

White L. 2013. Logistic Regression Model Effectiveness: Proportional chance criteria and proportional reduction in error. Journal of Contemporary Research in Education, 2(1):4-10.

Zhang J., yuan L., Pu R., Loraamm R., Yang G., Wang J. 2014. Comparison between wavelet spectral features and conventional spectral features in detecting yellow rust for winter wheat. Computers and Electronics in Agriculture, 100(1):79-87. https://doi.org/10.1016/j.compag.2013.11.001

Zvirina T., Hermana R., Brotmanb Y., Denisovc Y., Belausovd E., Freemanc S., Perl-Treves R. 2010. Differential colonization and defense responses of resistant and susceptible melon lines infected by Fusarium oxysporum race 1 2. Plant Pathology, 59:576-585.

5. GENERAL CONCLUSIONS AND RECOMMENDATIONS

5.1 General conclusions

- Reflectance spectroscopy is a powerful and reliable technique to identify and discriminate plants infected with *F. oxysporum* from healthy plants and subjected to water stress. RSWs related to the disease were found in the Vis range (mainly the ranges 510nm-520nm, 650nm-670nm and 700-750nm), suggesting physiological changes in the plants in response to the pathogen. The diseased tomato plants were correctly classified using the RSWs obtained, with percentages greater than 70%, after 12 dpi in the varieties evaluated.

- In this work, it was possible to identify and relate some physiological and spectral responses before the appearance of visible symptoms caused by fungal infections in *S. lycopersicum*. The sensitivity for relationship between reflectance with net photosynthesis, stomatal conductivity and transpiration was greater in spectral ranges, such as 420nm-490nm, 560nm, 680nm and 710nm. In particular, distant red bands had greater correlation than that net photosynthesis in plants inoculated with *F. oxysporum*, which allows estimating and detecting infected plants during the early stages of infection. It is important to indicate that the tomato plants inoculated with *F. oxysporum* did not present a significant difference in the net photosynthesis with respect to the control plants at 12 dpi (asymptomatic period), using only univariate analyzes with data derived from the gases exchange and Chlorophyll fluorescence analysis. On the other hand, the analyzes derived from spectral data and even the multivariate analyzes with same photosynthetic parameters (measured with the IRGA) could discriminate the infected plants, healthy ones and those subjected to water stress with a 100% correct classification percentage.

- It was apply in this research the BLRMs with reflectance data in Vis/NIR range for prediction and early detection of vascular wilt in tomato plants with reasonable degree of accuracy, greater than 80% in susceptible cultivar and, 90% on tomato tolerant cultivar. The obtained results in this work suggest that is possible to achieve these degrees of precision using only three basic variables (R750, R550, and R430), which are wavelengths located in the upper limit of red, green and violet.

5.2 Recommendations

The results obtained in this thesis has highlighted a number of topics on which further research would be beneficial.

In general, there are few works focused on the early detection of systemic diseases using reflectance spectroscopy. Specifically, the scientific articles focused on *S. lycopersicum* - *F. oxysporum* using this technique have typically been limited to disease identification in plants and pathogen isolate classification in vivo cultures. Future studies could be focused on evaluating whether the wavelengths related to the infection are specific or not to the pathosystem studied, or may be important in other plant-pathogen associations (fungi, bacteria, viruses, and nematodes), nutritional deficiencies or even damage by abiotic environmental factors.

Additionally, it is important to use spectroscopes with a wider spectral range, especially that they measure in complete NIR (800nm-2500nm). In this range, it is possible to correlate changes in important biomolecules concentrations for specific plant-pathogen interactions, with spectral changes in plants leaves. These wavelengths obtained in this way tend to be more specific to each specific pathosystem, so they could be used in robust BLRMs to predict infections of different pathogens in their respective hosts.

Finally, it is important to emphasize to the data registration and publication describing the informative content of spectral data collection process (metadata), since they can be very important in general data analysis and discussion of patterns unexpected in study phenomenon. Detailing and reporting the methodology used in specific programs for data analysis (such as Software R) is also important to ensure its repeatability and facilitate its realization by new researchers in this research line.

ANNEXED

Annexed I. Publication of preliminary results (B Publindex Colciencias).

Doi: http://doi.org/10.17584/rcch.2018v12i2.7293

Detección de plantas asintomáticas de *Solanum lycopersicum* L. infectadas con *Fusarium oxysporum* usando espectroscopia de reflectancia VIS

Detection of asymptomatic *Solanum lycopersicum* L. plants infected with *Fusarium oxysporum* using reflectance VIS spectroscopy

JUAN CARLOS MARÍN-ORTIZ[1]
LILLIANA MARÍA HOYOS-CARVAJAL[1]
VERÓNICA BOTERO-FERNÁNDEZ[2]

Cultivo de tomate en producción.
Foto: L.M. Hoyos-Carvajal

RESUMEN

Las plantas asintomáticas son reservorios de patógenos, ya que pueden permanecer infectadas la mayor parte de su ciclo de desarrollo, convirtiéndose en fuente de contaminación para el resto del cultivo. El objetivo de este estudio fue evaluar un método de detección y discriminación de dos cepas de *Fusarium oxysporum* en tomate usando espectroscopia. La enfermedad en las plantas de tomate inoculadas con la cepa aislada de gulupa (F05) fue mayor a la observada en la cepa aislada de tomate (F07), presentando valores de 60,0% (11 días) y 81,8% (22 días); la cepa F07 presentó incidencias de 30,0 y 64,3% en ambas mediciones. La planta infectada con la cepa F05 fue mejor discriminada en el periodo de incubación de la enfermedad en ambos periodos de tiempo en los Análisis de Componentes Principales (PCA) y Análisis Discriminantes Lineales (LDA) realizados con los controles en comparación con la cepa F07. Estos resultados sugieren que la espectroscopia de reflectancia VIS es un método sensible y confiable que puede ser adecuado para el diagnóstico temprano de enfermedades en plantas.

Palabras clave adicionales: enfermedades, espectroscopia VIS/NIR, métodos de detección, reflectancia, análisis multivariado.

[1] Facultad de Ciencias Agrarias, Universidad Nacional de Colombia, Medellín (Colombia). ORCID Marín-Ortiz, J.C.: 0000-0002-1334-4721; ORCID Hoyos-Carvajal, L.M.: 0000-0003-0980-5035
[2] Facultad de Minas, Universidad Nacional de Colombia, Medellín (Colombia). ORCID Botero-Fernández, V.: 0000-0002-6964-239X
[3] Autor para correspondencia. juancarlosmo@gmail.com

Annexed II. Article certificate accepted for publication: first chapter in the Journal Of Plant Protection Research (A2 Publindex Colciencias).

JPPR-00293-2019-02

Journal of
Plant Protection Research

Authors:
Juan Marín Ortiz, Lilliana Hoyos Carvajal, Veronica Botero Fernandez

Decision letter:
Ref.: Ms. No. JPPR-00293-2019-02
Detection of significant wavelength for identifying and classification of Fusarium oxysporum during the incubation period and water stress in Solanum lycopersicum plants using reflectance spectroscopy
Journal of Plant Protection Research

Dear Juan Marín Ortiz,

I am pleased to tell you that your work has now been accepted for publication in Journal of Plant Protection Research.

It was accepted on 2019-03-05

Thank you for submitting your work to this journal.

Kindest regards,
Henryk Pospieszny
Editor-in-Chief
Journal of Plant Protection Research

Editorial System

Annexed III. Certificate of the article publication: third chapter in Saudi Journal of Biological Sciences (A1 Publindex Colciencias)

Contents lists available at ScienceDirect

Saudi Journal of Biological Sciences

journal homepage: www.sciencedirect.com

Original article

Linking physiological parameters with visible/near-infrared leaf reflectance in the incubation period of vascular wilt disease

Juan Carlos Marín-Ortiz [a,*], Nathalia Gutierrez-Toro [a], Verónica Botero-Fernández [b], Lilliana María Hoyos-Carvajal [a]

[a] Department of Agricultural Sciences, Universidad Nacional de Colombia, Medellín, Colombia
[b] Department of Geosciences and Environment, Universidad Nacional de Colombia, Medellín, Colombia

ARTICLE INFO

Article history:
Received 27 December 2018
Revised 19 May 2019
Accepted 23 May 2019
Available online xxxx

Keywords:
Plant diseases
Vascular wilt
Reflectance spectroscopy
Multivariate analysis
Early detection
Plant physiology

ABSTRACT

The photosynthetic pigments are mainly responsible for absorbing the light intended to promote photosynthesis on the chloroplast of the leaves. Different studies have related the spectral response in the leaves of plants with the biotic stress generated by pathogens. In general, maximum differences in reflectance have been found in the range of 380–750 nm between plants subjected to biotic stress and healthy plants. In this study, it was possible to characterize and relate the spectral variance in leaves of *S. lycopersicum* infected with *F. oxysporum* with this physiological variation and pathogen concentration in tomato plants during the asymptomatic period of vascular wilt. Photosynthetic parameters derived from gaseous exchange analysis in the tomato leaves correlated related with four bands in the visible range (Vis). Additionally, five specific bands also present a high correlation with the increase in the concentration of *F. oxysporum* conidia measured at the root: 448–523 nm, 624–696 nm, 740–960 nm, 973–976 nm, and 992–995 nm. These wavelengths allowed a 100% correct classification of the plants inoculated with *F. oxysporum* from the plants subjected to hydric stress and the control plants in the asymptomatic period of the disease. The spectral response to biotic and abiotic stress in the measured Vis/NIR range can be explained by the general tendency to change the concentration of chlorophyll and carotene in tomato leaves. These studies also highlight the importance of the implementation of robust multivariate analysis over the multiple univariate analysis used in the applied biological sciences and specifically in the agricultural sciences. These results demonstrate that specific wavelength responses are due to physiological changes in plants subjected to stress, and can be used in indexes and algorithms applied to the early detection of diseases in plants on different pathosystems.

Annexed IV. Nutrient solution for hydroponic tomato cultivation (Hort Americas)

Nutrient	**Concentration (ppm)**
Nitrate (NO_3)	200 ppm
Ammonium (NH_4)	7 ppm
Potassium (K)	240 ppm
Phosphate (PO_4)	50 ppm
Calcium (Ca)	220 ppm
Magnesium (Mg)	50 ppm
Iron (Fe)	1.5 ppm
Manganese (Mn)	0.55 ppm
Zinc (Zn)	0.33 ppm
Boron (B)	0.3 ppm
Copper (Cu)	0.05 ppm
Molybdenum (Mo)	0.05 ppm
Sulfates (SO_4)	20 ppm
Chloride (Cl)	<300 ppm
Sodium (Na)	<100 ppm