SÃO PAULO STATE UNIVERSITY – UNESP

CAMPUS OF JABOTICABAL

IS IT POSSIBLE TO DETECT NEMATODES IN SOYBEAN (*Glycine max* L.) USING REMOTE SENSING?

Letícia Bernabé Santos

Agronomist

Jaboticabal – SP – Brazil

2020

SÃO PAULO STATE UNIVERSITY – UNESP

CAMPUS OF JABOTICABAL

IS IT POSSIBLE TO DETECT NEMATODES IN SOYBEAN (G*lycine max* L.) USING REMOTE SENSING?

Letícia Bernabé Santos
Advisor: Prof. Dr. Rouverson Pereira da Silva
Co-advisors: Prof. Dr. Pedro Luiz Martins Soares
Prof. Dr. José Eduardo Corá
Prof. Dr. Ignacio Antonio Ciampitti

Dissertation submitted to the Faculty of Agricultural and Veterinary Sciences – UNESP Campus of Jaboticabal, as requirement for the Master’s degree of Science in Agronomy (Crop Production).

2020

S237p Santos, Letícia Bernabé
É possível detectar sintomas de nematoides em soja (Glycine max l.) utilizando sensoriamento remoto? / Letícia Bernabé Santos. -- Jaboticabal, 2020
79 p.

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal
Orientador: Rouverson Pereira da Silva
Coorientador: Pedro Luiz Martins Soares

1. Imagens multiespectrais. 2. Detecção de doenças. 3. Nematoides. 4. Agricultura digital. 5. Aprendizado de máquina. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Jaboticabal

CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: "IS IT POSSIBLE TO DETECT NEMATODES IN SOYBEAN (*Glycine max* L.) USING REMOTE SENSING?"

AUTORA: LETÍCIA BERNABÉ SANTOS
ORIENTADOR: ROUVERSON PEREIRA DA SILVA
COORIENTADOR: IGNACIO ANTONIO CIAMPITTI
COORIENTADOR: PEDRO LUIZ MARTINS SOARES
COORIENTADOR: JOSÉ EDUARDO CORÁ

Aprovada como parte das exigências para obtenção do Título de Mestra em AGRONOMIA (PRODUÇÃO VEGETAL), pela Comissão Examinadora:

Prof. Dr. ROUVERSON PEREIRA DA SILVA
Departamento de Engenharia Rural / FCAV / UNESP - Jaboticabal

Dr. LEONARDO MENDES BASTOS (VIDEOCONFERÊNCIA)
KANSAS STATE UNIVERSITY / Manhattan/KS

Prof. Dr. LEONARDO CAMPOS DE ASSIS (VIDEOCONFERÊNCIA)
Instituto de Ciências e Tecnologia do Ambiente-Uniube / Uberaba/MG

Jaboticabal, 10 de agosto de 2020

CURRICULUM

LETÍCIA BERNABÉ SANTOS – Born in Franca, São Paulo, on February 18, 1995, André Neves Santos' and Andréia Bernabé's daughter, both also born in Franca, São Paulo. She was raised since the age of 5 in Rondonópolis, Mato Grosso where she attended Elementary School at Cenecista School "13 de Junho" and, High School at Pestalozzi School, in the city of Franca - SP, finished in 2012. Joined in Higher Education in 2013 in the Agronomic Engineering course at Universidade Estadual Paulista "Júlio de Mesquita Filho" (UNESP), Campus de Jaboticabal, obtaining the title of Agronomist, in February 2018. During her graduation she received a scholarship from the Tutorial Education (PET Agro) for three years, in research, teaching and extension. In addition, she participated in three scientific initiation projects, all aimed at the study of nematode control. To conclude the course, she did a curricular internship at the Bayer Company, in Sorriso - MT. In August 2018, she started her Masters course in Agronomy (Crop Production), in the Precision Agriculture area, at the São Paulo State University "Júlio de Mesquita Filho" - Campus de Jaboticabal, São Paulo, at the Agricultural Machinery and Mechanization Laboratory (LAMMA), where she conducted research using remote sensing to identify the occurrence of nematodes in soybean. In January 2020, she held a sandwich internship for 6 months at Kansas State University (KSU), Manhattan - KS, where she developed skills focused on statistics and writing.

EPIGRAPH

"Don't speak too quickly. Say only that the horse is not in the barn.

That is all we know; the rest is judgment.

If I've been cursed or not, how can you know? How can you judge?"

The Old Man and the White Horse
Max Lucado's Version (In the Eye of the Storm)

ACKNOWLEDGMENT

Always to God, who with His infinite wisdom full of mercy and love, accompanied and blessed me all the moments of my life.

To my parents, André Neves Santos and Andréia Bernabé, for their patience and education that have always been the pillar of my personality. Thank you for dedicating yourselves and believing in me! I'll also record a kiss for my little brother Ângelo.

To my advisor Dr. Rouverson Pereira da Silva, who welcomed me into his research team, taught me much more than just the profession. To the PhD Professors and also co-supervisors of this work Pedro Luiz Martins Soares, José Eduardo Corá and Ignacio Antonio Ciampitti for the suggestions and contributions in this thesis.

To São Paulo State University “Júlio de Mesquita Filho”, Faculty of Agricultural and Veterinary Sciences, Jaboticabal, in which I am proud to have been part of the institution's history since graduation. To the graduate program in Agronomy (Crop Production) for granting a scholarship and the opportunity to take the graduate course.

To Kansas State University (KSU), Manhattan, where I did the 6-month sandwich masters. Thanks to my coworkers, friends and teachers who contributed to my growth during this period.

To the technical team of LabNema, Lamma and KSUCrops Laboratories, for their friendship and support in carrying out the activities.

In order not forget anyone, I dedicate this paragraph to my friends, the old and the new ones, my eternal gratitude and affection for the moments lived. People I chose to share my life with and who certainly contributed to my development as a human being. And of course, Choppana will always have a space in my heart and an ankle mark to never forget this family that we are.

This work was carried out with the support of the Coordination for the Improvement of Higher Education Personnel in Brazil (CAPES) - Financing code 001.

SUMMARY

É POSSÍVEL DETECTAR SINTOMAS DE NEMATOIDES EM SOJA (*Glycine max* L.) UTILIZANDO SENSORIAMENTO REMOTO?

RESUMO

A utilização de técnicas de sensoriamento remoto teve expressivo aumento na agricultura nos últimos anos para avaliar diferentes estresses em culturas. Contudo, ainda são poucos trabalhos que avaliaram a viabilidade na identificação de nematoides. A atual amostragem de nematoides é intensiva em tempo e trabalho e não representa a variabilidade da infestação em todo o campo, dificultando a previsão precisa e o manejo. Portanto, foram desenvolvidos trabalhos visando verificar o potencial uso do sensoriamento aéreo na identificação de sintomas da ocorrência de nematoides na cultura da soja. No primeiro capítulo desta dissertação encontra-se a revisão de literatura contendo tópicos específicos relacionados ao tema central. No segundo, foi avaliada a variabilidade da ocorrência de nematoides em dois campos de produção da cultura da soja e o efeito da área de infestação em relação ao centro da reboleira (dentro, extremidade ou fora), assim como o comportamento multiespectral desses locais. Observou-se que as bandas do vermelho e NIR tiveram comportamento similar nas duas áreas avaliadas e foram capazes de se diferenciar em relação à localização na reboleira. No terceiro capítulo, o objetivo foi determinar a relação bivariada entre bandas espectrais individuais e índices de vegetação (IVs), todos relativos à condição da soja (plantas infectadas versus não infectadas). Utilizamos três algoritmos (Logistic Regression - LR, Random Forest – RF, Conditional Inference Tree – CIT) com três opções para entrada de dados: apenas bandas (modelo reduzido), apenas IVs e bandas mais IVs (modelo completo) para classificar as plantas. Os resultados demonstram a capacidade de se utilizar dados multiespectrais para distinguir plantas de soja infectadas por nematoides e não infectadas, utilizando sensoriamento aéreo combinado com aprendizado de máquina. No quarto capítulo, tem-se as considerações finais com os principais resultados, fatores limitantes, recomendações e os próximos passos da pesquisa.

PALAVRAS-CHAVE: imagens multiespectrais, detecção de doenças, nematoides, soja, aprendizado de máquina, agricultura digital.

IS IT POSSIBLE TO DETECT NEMATODES SYMPTOMS IN SOYBEAN (*Glycine max* L.) USING REMOTE SENSING?

Abstract

Remote sensing techniques have increased significantly in agriculture in recent years to assess different crop stresses. However, there are still few studies that have evaluated the viability to identify nematodes symptoms. The current nematode sampling approach is time and labor intensive and does not represent the variability of the infestation across the field, complicating its accurate prediction and management. Therefore, this work was carried out to verify the potential use of aerial sensing to identify nematodes symptoms occurring in soybean. The first chapter of this dissertation is a literature review containing specific topics. The second chapter is a variability analysis of nematode occurrence in soybean fields and the effect of the infestation area in relation to the center of the hotspot (inside, border or outside) were evaluated, as well as the multispectral behavior of these locations. It was observed that the red and near-infrared bands had similar behavior in the two areas and were able to differentiate different infestation areas in relation to location in the hotspot. In the third chapter, the objective was to determine the bivariate relationship between individual spectral bands and vegetation indices (VIs), all related to the condition of the soybean (infected versus non-infected plants). We used three algorithms (logistic regression, random forest, and conditional inference tree) with three options for data entry: only bands (reduced model), only VIs and bands plus VIs (complete model) to classify the plants. The results demonstrate the ability of multispectral data to distinguish nematode-infected and non-infected soybean plants, using aerial sensing combined with machine learning. In the fourth chapter, there are the final considerations with the main results, limiting factors, recommendations and the next steps of the research.

Keywords: multispectral imaging, disease detection, nematodes, soybean, machine learning, digital agriculture.

Chapter 1 – General Considerations

1. Introduction

This thesis is structured in four chapters with topics related to digital agriculture (DA), with focus on remote sensing at aerial level of data collection using an unmanned aerial vehicle (UAV) coupled with a multispectral sensor to detect nematode occurrence in soybean plants. Crop protection proved to be relevant as nematode attack can cause severe damage and advances in remote sensing equipment like UAV, development of small sensors, software, algorithms and digital imaging methodology enables the scalable assessment of field damage conditions.

Several plant pathogens can cause losses in soybean yield, but parasitic plant nematodes (PPN) are especially important given that the tropical conditions in Brazil are very favorable to their propagation. Due to nematodes ability to adapt to several agro-climatic zones, they can occur in any cropping system, being considered diverse and omnipresent (Nasu et al., 2018). Nematodes are round worms that can cause significant damage to crop plants affecting specially root growth development (inducing gall or feeding on the cortisol tissues), stunted growth due to the reduced nutrient uptake from soil and predisposition of roots to secondary pathogens invasion (Jones et al., 2013).

As the majority of its occurrence is in the soil, nematode damage is difficult to diagnose and manage. Currently, farmers around the world assess plant health through ground-based soil sampling to indicate the nematode occurrence. The method is destructive, highly subjective and does not represent the variability of the infestation throughout the field. Despite being widely used to detect different crop stresses, multispectral reflectance sensors for assessing nematode damage in soybeans has not been thoroughly evaluated. Therefore, in order to improve the nematode damage identification in the field, it is believed that remote sensing can be a potential tool capable of improving nematode management.

The first chapter presents the theorical framework to support the hypothesis that is possible to investigate nematode occurrence in soybean plants using remote sensing as a more efficient and non-destructive tool. The second- and third-chapters present results from application of remote sensing techniques in soybean plants infected with nematodes analyzed using analysis of variance (ANOVA) and machine learning algorithms which proved to be reasonable options for monitoring nematode infection. At last, the fourth chapter presents the final considerations of the presented master’s thesis, demonstrating the best results obtained in a cohesive manner, as well as some limitations and future next steps.

2. Objectives

In summary, this thesis aims to investigate the capability of multispectral data (remotely-sensed) for detection of nematode-infected soybean areas in two different approaches. The specific objectives of this study were to: i) assess the effect of nematode infestation area on soybean yield, ii) to assess soybean spectral reflectance at the blue, green, red, red-edge and near-infrared bands under different nematode infestation areas, iii) determine the response of individual bands and vegetation indices (VIs) to identify infected versus non-infected soybean areas; iv) compare the performance of three algorithms (logistic regression-LR, random forest-RF, and conditional inference tree-CIT) via the calculation of the accuracy, sensitivity, and specificity of each approach for classifying nematode infected areas, and once the algorithm is selected v) assess model performance when using only bands (reduced model), only VIs, and bands plus VIs (full model) as input variables.

3. Literature Review

3.1 Importance of soybean production and losses due to nematodes

Increased population and economic growth are two widely recognized challenges affecting food demand over the next years. Soybean (*Glycine* max L. Merrill) has been the most important commodity in the current international market due to its global consumption (Jia et al., 2020) as a major source of protein to humans and a high quality animal feed (Sugiyama et al., 2015). Among all the legumes, soybean has become one of the most important protein supplementation in developing countries due to the highest protein content (40-42%), oil content (18-22%), and relatively low cost production (Silva et al., 2017).

Soybean is the most planted oilseed in the world. In the season of 2018/2019, the United States, the largest soybean producer, planted 35.6 million hectares of soybeans and was responsible for 34.1% of the global production (123.66 million tons), followed by Brazil, with ~34.8 million hectares of planted soybeans and an estimated 32.5% of the total production (116.5 million tons). China is the largest importer of soybean in the world, being responsible for 64.56% of all world imports, followed by the European Union with 8.9%. As for exports, Brazil leads with 44.68%, followed by the United States (38.51%) and Argentina (4.94%), totaling 88.13% of all world exports (USDA: United States Department of Agriculture, 2019). The large area planted with the crop and high average yields are the main drivers for high level of production.

Improvements in soybean seed yield will be essential for matching the increased demand as all of the best farmland in the world is already been cultivated and facing yield stagnation (Barker & Koenning, 1998). Soybean production has been connected to concerns related to environmental sustainability and yield gap control (Silva et al., 2010). To achieve high soybean yields, site-specific crop management has to be implemented by producers, and monitoring plant disease is one key aspect. Plant diseases have a negative impact on the safety of food production, and can reduce quality and quantity of the harvested product. In severe cases, they can even induce to complete harvest loss.

Plant parasitic nematodes (PNN) represent an important group among microorganism that cause losses in soybean crops. Worldwide, yield losses can represent more than 12.6% (US$ 215.77 billion). Allen et al., 2017 reported that 12.5% of the losses on soybean yield are due to diseases with nematodes contributing to 31% of the total losses due to soybean diseases.

To sustain the increased food demand from population growth while mitigating environmental impacts, the use of innovate technologies in grain production have been evaluated. Among the most promising techniques, precision agriculture can provide efficient agricultural management of a farm by combining remote sensing (RS), geographic information systems and proximal data gathering (Ferrández-Pastor et al., 2018). These approaches are very useful to detect diseases and pests (Joalland et al., 2018; George Deroco Martins et al., 2017a; George Deroco Martins & Trindade Galo, 2015; Nutter et al., 2002).

Different strategies for reducing nematodes population and allowing their coexistence with cropping systems without causing significant yield losses have been evaluated. Integrated management has been extensively tested to control nematodes incidence in several crop systems and includes the use of chemical nematicides, resistant crop cultivars, biological control organisms and management practices such as crop rotation and cover crops (Ferraz et al., 2010).

3.2 Nematodes in soybeans

Multiple genus and species of nematodes infect most of the crops causing loss of yield potential (Nicol et al., 2011). Belonging to the phylum Nematoda, they evolved into more than 4000 species that can infect plants (Decraemer and Hunt, 2006). These soil borne pests are roundworms (vermiform), obligate parasites, that feed from the plant roots by using their protectable stylets to pierce the cell, release a secretion that can degrade the host cell wall, suppressing or avoiding host defenses, and establish a specific feeding site to extract nutrients from (Haegeman et al., 2012). Once the host is susceptible to the infection, feeding cells go through a severe morphological and metabolic modification to serve a new function as a nutrient sink to the developing nematodes (Hewezi & Baum, 2017).

Nematodes that infect roots are classified according to their location: the ones which pierce the cell and remain outside of the roots are called ectoparasites, while endoparasites have to enter into the roots and feed from the internal region. They can also be subclassified by their movement behavior within the roots (Akker et al.,2014). Nematodes that feed from the same site after the infection are called sedentary, while those that are mobile throughout their life cycle are called migratory nematodes that can cause massive tissue damage and necrosis, feeding along the way (Hewezi & Baum, 2017).

In Brazil, the most frequent and that have been causing most yield losses are the root knot nematode (RKN), (*Meloidogyne javanica* e *M. incognita*), the cyst nematode (*Heterodera glycines*), the root lesion nematode (*Pratylenchus brachyurus* Godfrey Filipjev e Schuurmans Stekhoven), and reniform (*Rotylenchulus reniformis* Linford & Oliveira).

Among these, *H. glycines* is the major suppressor of yield in soybean (Allen et al., 2017). Additionally, it is the hardest to control given that its resistance structure surrounding the eggs (the dead female body) can survive in the soil for several years until a new generation of juveniles can hatch again under favorable conditions (Dias et al., 2009). Furthermore, *Meloidogyne* spp., are as aggressive as the cyst nematode, and despite not having a resistance structure, is highly polyphagous which makes difficult to find non-host cover crops to use in crop rotation in infested areas (Ferraz & Brown, 2016).

Another polyphagous nematode is the root lesion nematode. The species behave as migratory endoparasite, creating lesions on the roots system that turn into necrotic areas, often followed by root rotting from secondary attack by soil fungi or bacteria (Jones et al., 2013). Moreover, the reniform nematode is a sedentary semi-endoparasite of several annual and perennial species that decreases root growth and creates many gelatinous egg sacks that may be seen on the root surface.

PPNs are considered the “unseen enemies” because the symptoms are non-specific and usually go unnoticed (Dutta et al., 2019). Unfortunately, the visible aboveground symptoms only become evident when the field reaches high population densities, and even though, it is unlikely that the farmer knows the extent of yield loss being caused by nematodes (Nutter et al., 2002). Those above ground symptoms arising from the infection include chlorosis and severely stunting (due to reduced nutrient uptake from soil and root predisposition to secondary invasions), which results from the reduction of chlorophyll content, plant growth, photosynthetic rate and plant nutrient concentration due to nematodes infection (Carneiro et al., 2002; Haseeb et al., 1990; Lu et al., 2014). The symptoms of nematodes occurrence can be misdiagnosed as typical drought or nutrient stress. Nonetheless, nematode symptoms usually appears in a patchy spatial pattern (K.R. Barker, 1985), and is controlled by environmental factors (T. Liu et al., 2019). The nematode attack extent can grow outwards away from the patch and still go undetected.

Currently, plant health and yield are assessed through ground-based soil sampling through costly testing methods that are labor and time intensive (Nutter et al., 2002). Remote sensing can offer a potential new method to detect subtle changes of crop canopies and can be crucial for effective disease management.

3.3 Remote sensing and the application in identifying diseases

Remote sensing can offer techniques and methods to obtain crop information with high repeatability and quality, without making physical contact with the object. These data are critical to build reliable statistics for agricultural management and decision-making (Mutanga et al., 2017).

Conventionally, plant disease has been detected through the visualization of the symptoms, followed by hand sampling and then submitting the samples to a laboratory.

However, this method requires expertise in plant pathology as well as time to complete the diagnosis. These gaps have been the major reasons for developing new methods that can assess the disease in the early stages with accuracy (Ali et al., 2019).

Plant disease and pests are severe threats to worldwide agriculture given the destructive power of its occurrence. Remote sensing techniques can be seen as a “radiodiagnosis” of plants, which deliver information of an entire field with no contact with it and can be repeated throughout the season. Considering that the occurrence of plant diseases can have heterogenous distribution, remote sensing techniques are useful to identify the primary location, extent and severity, improving the efficiency of plant protection.

Plants absorb light mostly to run photosynthesis and the amount of light radiation assimilated depends on the concentration of different pigments in the leaf, including chlorophyll a and b, carotenoids and anthocyanins (Chappelle et al., 1984). Several studies reported the reduction of chlorophyll content, plant growth, photosynthetic rate and plant nutrient concentrations due to nematode infection (Carneiro et al., 2002; Haseeb et al., 1990; Lu et al., 2014). Aboveground symptoms arising from the infection as chlorosis and severe stunting can be observed (Nutter et al., 2002). These plant physiological changes caused by nematodes can potentially be assessed via remote sensing.

The spectral data obtained through spectroscopy-based sensors arise from the interaction of electromagnetic spectrum and the object according to a specified wavelength and its behavior by radiation energy. In plants, the spectral reflectance is influenced by chemical conformation, physical formation and the spectral properties of the light source (Zwiggelaar, 1998). Reflectance in the visible range (400-700 nm) is associated with pigments, such as chlorophyll a and b, carotenoids and anthocyanins. Reflectance in the near-infrared (NIR) spectrum (700-1000 nm) is influenced by leaf morphology, structure and biomass. The red-edge spectral region has also been investigated recently. It is a narrow portion of the spectrum around 750 nm, between red and NIR regions, where the vegetation causes an unique reflectance, because the red light is absorbed mainly by chlorophyll and the NIR radiation is reflected (Scotford & Miller, 2005). As a result, capability of detecting plant disease infections has been studied by several researches and has been used as a rapid non-destructive and cost-effective method.

Grisham et al. (2010) found that the pigment data collected from sugarcane yellow leaf virus (SCYLV)-infected sugarcane plants was efficient in predicting the disease in 80% of the samples, prior to visual symptoms expression. Also, Cao et al., 2013 found that canopy hyperspectral reflectance can be applied to detect powdery mildew, specially before apparent symptoms were evident. In addition, certain soil borne diseases and root diseases that trigger systemic effects can also be detected (Zhang et al., 2019). Nutter et al. 2002, found significant relationships between image data, soybean yield and soybean cyst nematode (SCN) population densities, especially in the infrared region. Martins et al., 2017b reported that the spectral curve obtained from coffee canopy severely infected with root-knot nematodes had lower reflectance between 750 and 900 nm wavelengths when compared to healthy plants.

3.4 Vegetation Index

Vegetation indices (VIs) are mathematical combinations using single-band reflectance values. This information is simple and effective for quantitative and qualitative assessment condition of vegetation cover, among other approaches and applications in agriculture (Atzberger, 2013). Furthermore, VIs minimizes the effects of the soil, topography and viewing angle on spectral response of desired characteristic (Jackson et al., 2012).

For example, the normalized difference vegetation index (NDVI), the most widely used VI, uses the combination of red and NIR bands and is correlated with biophysical variables of the plant, like leaf area, leaf water content and total chlorophyll (Rouse et al., 1973). The index can vary from -1 to 1, given that very low values indicate bare soil and higher positive values are correlated with greater canopy density or greenness. Also, simple ratio indices (SR), which can also use the red and NIR bands, is an index used to indicate the relative biomass presented in the image (Baret & Guyot, 1991; Jordan, 1969).

However, red reflectance can be affected by other pigments such as anthocyanin and even when there is low content of chlorophyll, it saturates rapidly in higher biomass conditions (Gitelson et al., 2001). In order to increase the sensitivity of NDVI to chlorophyll content, researchers developed indices including green (green normalized difference vegetation index - GNDVI) (Gitelson, et al., 1996) and red-edge channels (normalized difference red Edge - NDRE) (Barnes et al., 2000) to overcome this limitations.

Several VIs were developed for assessing different vegetation parameters. In a nutshell, VIs combines different aspects of plant vigor information from the crops in a single value. It is common for researchers to test different VIs to evaluate the correlation with the studied variable. Joalland et al., (2018) used several VIs (NDVI,

SR, TCARI, TGI, ANTH, CHLG, PRI, NDWI, WI, HI) to validate the ability of aerial UAV-based and hyperspectral imaging to discriminate and rank susceptible and tolerant sugar beet cultivars and predict nematode population and final yield. They concluded that the selected VIs reflected plant traits such as leaf chlorophyll and water content, leaf area index, and biomass which are common severe symptoms from BCN (beet cyst nematode) infestation.

Thus, it is relevant to evaluate different VIs and their correlation with nematode occurrence, since the spectral signature of infected plants can change when they are under stress, with changes in the chlorophyll content and leaf structure.

3.5 Techniques for monitoring plant diseases and pests

Remote sensing data has been studied in combination with many algorithms for monitoring plant diseases and pests, including data classification. For example, non-parametric classification methods like random forest were used by Adelabu et al., (2014), and appeared promising to develop early warning systems for insect defoliation. Therefore, with appropriate classification techniques extracted from images, experts can implement reliable control prescriptions.

We studied two binomial machine learning classification models (random forest and conditional inference tree) and one method of statistical discriminant analyses (logistic regression). The main advantages of the algorithm techniques studied is the possibility of exploration of complex interactions and their robustness in relation to data distributional assumptions.

Random forest (RF) analysis is a powerful predictor for ensemble-learning methods (Breiman, 2001). When performed to predict classification problems, is considered the weighted majority of class votes (Kohestani et al., 2015), and it has been used in agricultural management (Pott et al., 2020). Conditional inference tree (CIT) is an algorithm used for recursive variable selection and binary partitioning on the explanatory variables according to a significance test and outputs an interpretable decision tree model (Hothorn et al., 2015). Logistic regression is a predictive analysis technique that provides a solution to the principle question of which variables are important in differentiating between binary responses (Neter et al., 1996).

Thus, to analyze remote sensing features obtained from different targets can be challenging, and to properly develop a robust method to classify the task, multivariate analyses have been a reasonable option.

References

Adelabu, S., Mutanga, O., & Adam, E. (2014). Evaluating the impact of red-edge band from Rapideye image for classifying insect defoliation levels. *ISPRS Journal of Photogrammetry and Remote Sensing*, *95*, pp. 34–41. https://doi.org/10.1016/j.isprsjprs.2014.05.013

Akker, S. E., Lilley, C. J., Jones, J. T., & Urwin, P. E. (2014). Identification and Characterisation of a Hyper-Variable Apoplastic Effector Gene Family of the Potato Cyst Nematodes. *PLoS Pathogens*, 10(9), e1004391. doi:10.1371/journal.ppat.1004391

Ali, M. M., Bachik, N. A., Muhadi, N. ‘Atirah, Tuan Yusof, T. N., & Gomes, C. (2019). Non-destructive techniques of detecting plant diseases: A review. In *Physiological and Molecular Plant Pathology* (Vol. 108, p. 101426). Academic Press. https://doi.org/10.1016/j.pmpp.2019.101426

Allen, T. W., Bradley, C. A., Sisson, A. J., Byamukama, E., Chilvers, M. I., Coker, C. M., Collins, A. A., Damicone, J. P., Dorrance, A. E., Dufault, N. S., Esker, P. D., Faske, T. R., Giesler, L. J., Grybauskas, A. P., Hershman, D. E., Hollier, C. A., Isakeit, T., Jardine, D. J., Kelly, H. M., … Allen Wrather, J. (2017). Soybean yield loss estimates due to diseases in the United States and Ontario, Canada, from 2010 to 2014. *Plant Health Progress*, *18*(1), 19–27. https://doi.org/10.1094/PHP-RS-16-0066

Atzberger, C. (2013). Advances in Remote Sensing of Agriculture: Context Description, Existing Operational Monitoring Systems and Major Information Needs. *Remote Sensing*, *5*(2), 949–981. https://doi.org/10.3390/rs5020949

Baret, F., & Guyot, G. (1991). Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sensing of Environment*, *35* (2–3), pp. 161–173. https://doi.org/10.1016/0034-4257(91)90009-U

Barker, K.R. (1985). *Sampling nematode communities*. Dept. of Plant Pathology. North Carolina State Univ.

Barker, Kenneth R., & Koenning, S. R. (1998). Developing sustainable systems for nematode management. *Annual Review of Phytopathology*, *36*(1), 165–205. https://doi.org/10.1146/annurev.phyto.36.1.165

Barnes, E. M., Clarke, T. R., Richards, S. E., Colaizzi, P. D., Haberland, J., Kostrzewski, M., Waller, P. M., Choi, C. Y., Riley, E., Thompson, T. E., Lascano, R. J., Li, H., Moran, M. S., Robert, P. C., Rust, R. H., & Larson, W. E. (2000). *Coincident detection of crop water stress, nitrogen status and canopy density using ground-based multispectral data*.

Breiman, L. (2001). *Random Forests* (Vol. 45).

Cao, X., Luo, Y., Zhou, Y., Duan, X., & Cheng, D. (2013). Detection of powdery mildew in two winter wheat cultivars using canopy hyperspectral reflectance. *Crop Protection*, *45*, 124–131. https://doi.org/10.1016/j.cropro.2012.12.002

Carneiro, R. G., Mazzafera, P., Ferraz, L. C. C. B., Muraoka, T., & Trivelin, P. C. O. (2002). Uptake and translocation of nitrogen, phosphorus and calcium in soybean infected with *Meloidogyne incognita* and *M. javanica*. *Fitopatologia Brasileira*, *27*(2), 141–150. https://doi.org/10.1590/s0100-41582002000200004

Chappelle, E. W., McMurtrey, J. E., Wood, F. M., & Newcomb, W. W. (1984). Laser-induced fluorescence of green plants 2: LIF caused by nutrient deficiencies in corn. *Applied Optics*, *23*(1), 139. https://doi.org/10.1364/ao.23.000139

Decraemer, W., & Hunt, D. J. (2006). Structure and classification. In ‘Plant nematology’.(Eds RN Perry, M Moens) pp. 3–32.

Dias, W. P., Silva, J. F., Carneiro, G. E., Garcia, A., & Arias, C. A. (2009). Nematoide de cisto da soja: biologia e manejo pelo uso da resistência genética. *Nematologia Brasileira*, 33(1), pp. 1-16.

Dutta, T. K., Khan, M. R., & Phani, V. (2019). Plant-parasitic nematode management via biofumigation using brassica and non-brassica plants: Current status and future prospects. In *Current Plant Biology* (Vol. 17, pp. 17–32). Elsevier B.V. https://doi.org/10.1016/j.cpb.2019.02.001

Ferrández-Pastor, F., García-Chamizo, J., Nieto-Hidalgo, M., & Mora-Martínez, J. (2018). Precision Agriculture Design Method Using a Distributed Computing Architecture on Internet of Things Context. *Sensors*, *18*(6), 1731. https://doi.org/10.3390/s18061731

Ferraz, L. C. C. B., & Brown, D. J. F. (2016). Nematologia de plantas: fundamentos e importância. Manaus: *Norma Editora*, 1, 251.

Ferraz, S., Freitas, L. G., Lopes, E. A., & Dias-Arieira, C. R. (2010). Manejo sustentável de fitonematoides, 306p. Editora UFV, Viçosa.

Gitelson, A. A., Merzlyak, M. N., & Chivkunova, O. B. (2001). Optical properties and nondestructive estimation of anthocyanin content in plant leaves. *Photochemistry and photobiology*, 74(1), pp. 38-45.

Grisham, M. P., Johnson, R. M., & Zimba, P. V. (2010). Detecting Sugarcane yellow leaf virus infection in asymptomatic leaves with hyperspectral remote sensing and associated leaf pigment changes. *Journal of virological methods*, 167(2), pp. 140-145.

Haegeman, A., Mantelin, S., Jones, J. T., & Gheysen, G. (2012). Functional roles of effectors of plant-parasitic nematodes. *Gene*, *492*(1), 19–31. https://doi.org/10.1016/J.GENE.2011.10.040

Haseeb, A., Srivastava, N. K., & Pandey, R. (1990). The influence of *Meloidogyne incognita* on growth, physiology, nutrient concentration and alkaloid yield of hyoscy amus niger. *Nematologia Mediterranea*, 18(2), pp.127–129.

Hewezi, T., & Baum, T. J. (2017). Communication of Sedentary Plant-Parasitic Nematodes With Their Host Plants. In *Advances in Botanical Research*, Vol. 82, pp. 305-324. https://doi.org/10.1016/bs.abr.2016.11.004

Hothorn, T., Hornik, K., & Zeileis, A. (2015). ctree: Conditional inference trees*. The Comprehensive R Archive Network*, 8.

Neter J., Kutner M.H., Nachtsheim C.J., & Wasserman W.. (1996). Logistic regression, Poisson regression and generalized linear models. *Applied Linear Statistical Models*, *4th ed.*, 567–627.

Jackson, L. E., Bowles, T. M., Hodson, A. K., & Lazcano, C. (2012). Soil microbial-root and microbial-rhizosphere processes to increase nitrogen availability and retention in agroecosystems. *Current Opinion in Environmental Sustainability*, *4*(5), pp. 517–522. https://doi.org/10.1016/J.COSUST.2012.08.003

Jia, F., Peng, S., Green, J., Koh, L., & Chen, X. (2020). Soybean supply chain management and sustainability: A systematic literature review. In *Journal of Cleaner Production* (Vol. 255, p. 120254). https://doi.org/10.1016/j.jclepro.2020.120254

Joalland, S., Screpanti, C., Varella, H., Reuther, M., Schwind, M., Lang, C., Walter, A., Liebisch, F., Joalland, S., Screpanti, C., Varella, H. V., Reuther, M., Schwind, M., Lang, C., Walter, A., & Liebisch, F. (2018). Aerial and Ground Based Sensing of Tolerance to Beet Cyst Nematode in Sugar Beet. *Remote Sensing*, *10*(5), 787. https://doi.org/10.3390/rs10050787

Jones, J. T., Haegeman, A., Danchin, E. G. J., Gaur, H. S., Helder, J., Jones, M. G. K., Kikuchi, T., Manzanilla-López, R., Palomares-Rius, J. E., Wesemael, W. M. L., & Perry, R. N. (2013). Top 10 plant-parasitic nematodes in molecular plant pathology. *Molecular Plant Pathology*, *14*(9), 946–961. https://doi.org/10.1111/mpp.12057

Jordan, C. F. (1969). Derivation of Leaf-Area Index from Quality of Light on the Forest Floor. *Ecology*, *50*(4), 663–666. https://doi.org/10.2307/1936256

Kohestani, V. R., Hassanlourad, M., & Ardakani, A. (2015). Evaluation of liquefaction potential based on CPT data using random forest. *Natural Hazards*, *79*(2), 1079–1089. https://doi.org/10.1007/s11069-015-1893-5

Liu, T., Hu, F., & Li, H. (2019). Spatial ecology of soil nematodes: Perspectives from global to micro scales. In *Soil Biology and Biochemistry* (Vol. 137, p. 107565). Elsevier Ltd. https://doi.org/10.1016/j.soilbio.2019.107565

Lu, P., Davis, R. F., Kemerait, R. C., Van Ierse, M. W., & Scherm, H. (2014). Physiological effects of meloidogyne incognita infection on cotton genotypes with differing levels of resistance in the greenhouse. *Journal of Nematology*, *46*(4), pp. 352–359.

Martins, G. D., Galo, M. D. L. B. T., & Vieira, B. S. (2017a). Detecting and Mapping Root-Knot Nematode Infection in Coffee Crop Using Remote Sensing Measurements. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(12), 5395–5403. https://doi.org/10.1109/JSTARS.2017.2737618

Martins, G. D., Galo, M. D. L. B. T., & Vieira, B. S. (2017b). Detecting and Mapping Root-Knot Nematode Infection in Coffee Crop Using Remote Sensing Measurements. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(12), 5395–5403. https://doi.org/10.1109/JSTARS.2017.2737618

Martins, G. D., & Trindade Galo, M. D. L. B. (2015). Caracterização espectral da cana-de-açúcar infectada por nematoides e migdolus fryanus por espectrorradiometria de campo. *Boletim de Ciencias Geodesicas*, *21*(4), 783–796. https://doi.org/10.1590/S1982-21702015000400046

Mutanga, O., Dube, T., & Galal, O. (2017). Remote sensing of crop health for food security in Africa: Potentials and constraints. *Remote Sensing Applications: Society and Environment*, *8*(October), 231–239. https://doi.org/10.1016/j.rsase.2017.10.004

Nasu, É. das G. C., Amora, D. X., Monteiro, T. S. A., Alves, P. S., de Podestá, G. S., Ferreira, F. C., & de Freitas, L. G. (2018). *Pochonia chlamydosporia* applied via seed treatment for nematode control in two soil types. *Crop Protection*, *114*, 106–112. https://doi.org/10.1016/j.cropro.2018.08.010

Nicol, J. M., Turner, S. J., Coyne, D. L., Nijs, L. den, Hockland, S., & Maafi, Z. T. (2011). Current Nematode Threats to World Agriculture. In *Genomics and Molecular Genetics of Plant-Nematode Interactions* (pp. 21–43). Springer Netherlands. https://doi.org/10.1007/978-94-007-0434-3_2

Nutter, F. W., Tylka, G. L., Guan, J., Moreira, A. J. D., Marett, C. C., Rosburg, T. R., Basart, J. P., & Chong, C. S. (2002). Use of Remote Sensing to Detect Soybean Cyst Nematode-Induced Plant Stress. *Journal of Nematology*, 34(3), pp. 222-231.

Pott, L. P., Amado, T. J., Schwalbert, R. A., Sebem, E., Jugulam, M., & Ciampitti, I. A. (2020). Pre-planting weed detection based on ground field spectral data. *Pest Management Science*, *76*(3), 1173–1182. https://doi.org/10.1002/ps.5630

Rouse, W., Haas H., Schell A., Deering W., 1973. Monitoring vegetation systems in the Great Plains with ERTS. In: Proc. Earth Res. Tech. Satellite-1 Symp., Goddard Space Flight Cent., Washington, DC., pp. 309–317.

Scotford, I. M., & Miller, P. C. H. (2005). Applications of spectral reflectance techniques in northern European cereal production: A review. In *Biosystems Engineering* (Vol. 90, Issue 3, pp. 235–250). https://doi.org/10.1016/j.biosystemseng.2004.11.010

Silva, F. C. S., Sediyama, T., de Cássia, R. T. O., Borém, A., da Silva, F. L., Bezerra, A. R. G., & da Silva, A. F. (2017). Economic importance and evolution of breeding. In *Soybean Breeding* (pp. 1–16). Springer International Publishing. https://doi.org/10.1007/978-3-319-57433-2_1

Silva, V. P., van der Werf, H. M. G., Spies, A., & Soares, S. R. (2010). Variability in environmental impacts of Brazilian soybean according to crop production and transport scenarios. *Journal of Environmental Management*, *91*(9), pp. 1831–1839. https://doi.org/10.1016/j.jenvman.2010.04.001

Sugiyama, A., Ueda, Y., Takase, H., & Yazaki, K. (2015). Do soybeans select specific species of *Bradyrhizobium* during growth? *Communicative & Integrative Biology*, *8*(1), e992734. https://doi.org/10.4161/19420889.2014.992734

Zhang, J., Huang, Y., Pu, R., Gonzalez-Moreno, P., Yuan, L., Wu, K., & Huang, W. (2019). Monitoring plant diseases and pests through remote sensing technology: A review. *Computers and Electronics in Agriculture*, *165*, 104943. https://doi.org/10.1016/j.compag.2019.104943

Zwiggelaar, R. (1998). A review of spectral properties of plants and their potential use for crop/weed discrimination in row-crops. In Crop Protection (Vol. 17, Issue 3, pp. 189–206). Elsevier. https://doi.org/10.1016/S0261-2194(98)00009-X

Chapter 2 - Detection of nematodes in soybean by multispectral images

Abstract

Plant parasitic nematodes can decrease soybean production by more than 30% without visible symptoms above ground. Multispectral reflectance sensors have been used to detect different stresses in crops, but their use in nematode soybean damage assessment has not been exhaustively evaluated. The objectives of this study were: i) to evaluate the effect of the nematode infestation area on soybean yield; and ii) evaluate the spectral reflectance of soybeans in the blue, green, red, red-edge and near infrared (NIR) bands at different nematode location within the patches. The experiment was carried out in two commercial soybean fields, showing different levels of infestation. Multispectral images were obtained at the R1 growth stage using a MicaSense RedEdge® sensor mounted on a drone. Sensor bands were used to identify critical nematode infestation points for targeted sampling. In each row, three nematode infestation areas were defined in relation to the patches' center: outside, border and inside. Grain yield was assessed at R8 by manual harvest at each sampling point. The nematode infestation area affected soybean productivity, with lower productivity observed inside the patches. The red, red-edge and NIR bands were able to differentiate nematode infestation areas. Outside the hotspot, there was significantly lower reflectance in the red band and greater reflectance in the red-edge and NIR bands, compared to the inside and border locations. Damage caused by nematodes affects soybean yield and can be detected by bands in the red, red-edge and NIR spectral regions.

Keywords: Disease identification, nematodes, remote sensing, multispectral images.

1. Introduction

Soybean (*Glycine max* L. Merr.) cultivation is considered the most important in the world economically (Chang et al., 2015), including in Brazil, where it has been cultivated throughout the country since 1980 (Meyer et al., 2017), being considered an extensive monoculture in the territory. However, this continuous production system results in decreased production and quality loss, which can be caused by biotic or abiotic factors (Eberlein et al., 2016), pests and diseases occurrence are serious crop threats (Oerke, 2006).

Among these limiting factors, nematodes are also an issue in agriculture even with several management options (Nasu et al., 2018). These are a diverse group of invertebrates, accounting for more than 15 thousand described species, and represent a small portion of Nematoda phylum (Barker & Koenning, 1998). Symptoms are similar to nutritional deficiency and lack of water (Blevins et al., 1995), as the plant exhibits symptoms leaf chlorosis, early senescence, reduced seed weight (Niblack, 2005) and, consequently, losses in grain yield.

Annual losses can reach 12.3% worldwide, equivalent to 157 billion dollars, in addition to 500 million dollars to control the disease (Hassan et al., 2013). Probably the damage caused by nematodes is underestimated. This is because the symptoms caused by plant parasitic nematodes (PPN) are nonspecific and in many regions, producers are unaware of their presence because they are microscopic (Siddique & Grundler, 2018).

For soybean crop, the soy cyst nematode (SCN), *Heterodera glycines* Ichinohe 1952, is considered the most devastating to decrease yield (Yan & Baidoo, 2018) and is also host to other species, such as *Pratylenchus brachyurus*, the root lesion nematode (RLN), *Meloidogyne* spp., gall nematode (GN), among others. In some areas, it is common to practice corn (*Zea mays* L.) rotation and soybeans, as this is considered the best system to maintain high grain yield (Grabau & Chen, 2016). However, this succession system supports important nematode reproduction, reaching very high population levels and causing big economic losses.

Symptoms caused due to the interaction of plants and parasite, including soil pathogens, are the basis for enabling remote sensing techniques in monitoring (Zhang et al., 2019a). Remote sensing for pests and diseases can be approached as plant “radiodiagnosis”, since it is possible to obtain continuous information about the disease without physical contact and, in an efficient manner (Zhang et al., 2019b). Production conditions attributes, such as water availability, soil properties, fertilizer management and irrigation, can significantly interfere in production potential (Campbell & Wynne, 2011). Therefore, remote sensing techniques can provide us with spectral, spatial and temporal information about objects trough land vehicles, remotely manned aircraft, satellites and manual radiometers to assist in decision making.

Reflectance sensors are coherent tools to detected pests and parasites in crops, registering its spectral changes and development in the field. Including differences in spectral reflectance caused by nematodes (Bajwa et al., 2017a; Christian Hillnhütter et al., 2012; Martins & Trindade Galo, 2015a).

Diseases detection and their spectral pattern behavior can be essential to provide effective disease control management, such as pesticides localized application, use of resistant cultivars and crop rotation in areas devastated by nematode (Bajwa et al ., 2017a). In this work, soybean plants parasitized with nematodes, showed higher reflectance in the visible band and lower reflectance in red-edge and NIR bands, which was also verified by (Bajwa et al., 2017a), in soybean plants infected with the nematode of cyst and sudden soybean death syndrome.

The objectives of this study were: i) to evaluate the nematode variability occurrence and the effect of the infestation area (outside, border and inside the patches) of nematodes on soybean grain yield; and ii) evaluate the spectral reflectance of soybeans in the blue, green, red, red-edge and near infrared (NIR) bands in different areas of nematode infestation.

2. Material and Methods

2.1 Experimental area description area and management

Two sites experiments were conducted in a commercial production field. It was located in Guaíra, São Paulo, Brazil (locations I and II, Figure 1) the soybean season of 2018/2019. The areas were chosen for their natural nematode infestation (Table 1), due to the intensification of crops that are host to the main nematodes such as soybeans, corn and beans. The previous harvest was common beans on both sites. Sites were managed under no-till practices. The maximum and minimum average air temperature was 30.0ºC and 21.2ºC, respectively and the cumulative rainfall was 822 mm.

Table 1. Site number, elevation, nematode infestation and total area for each field.

Site	Elevation(m)*	Nematode infestation genus	Area (ha)
I	540	*M. incognita*, *P. brachyurus*	8.2
II	502	*M. incognita*, *P. brachyurus* and *H. glycines*	16.1

**Above sea level.*

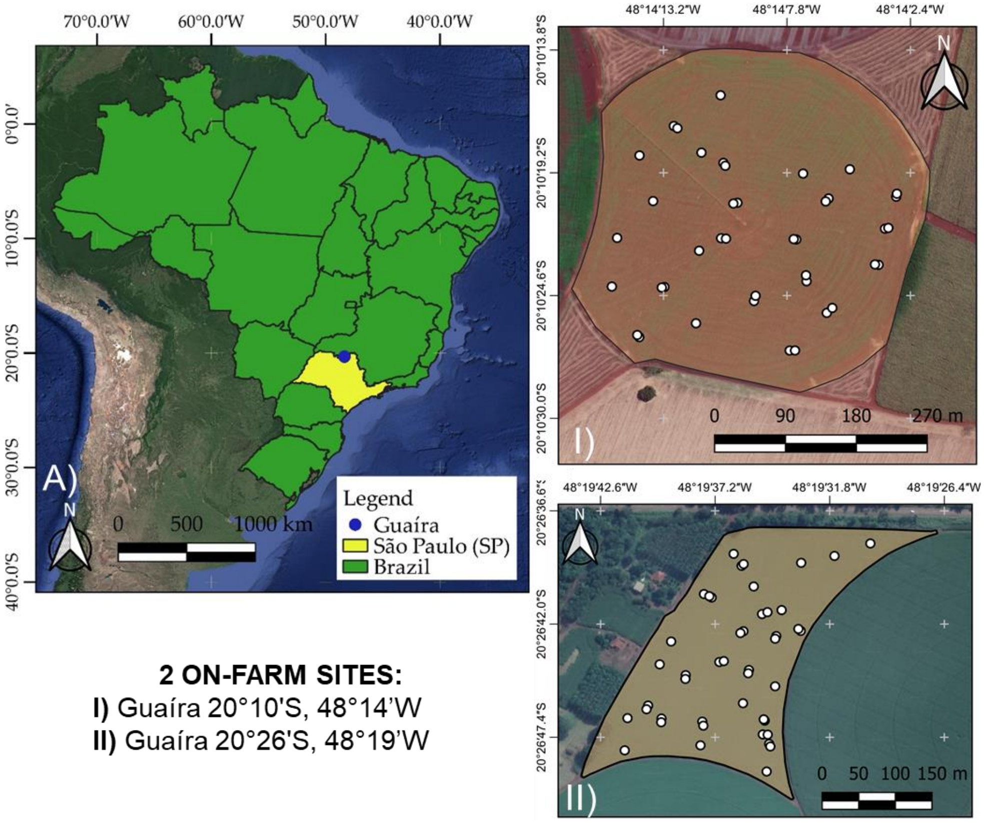

Figure 1. Experimental area characterization. A) Guaíra city highlighted in São Paulo state and Brazil maps. Sites I and II with white points representing the georeferenced sampling points.

Site I was sown on November 15th with cultivar DS 5916, resistant only to *M. incognita* counting 320,000 seeds ha $^{-1}$. Furthermore, Site II was sown on November 3rd with 290,000 seeds ha^{-1} with the cultivar BRS 7380 RR. This cultivar has low reproduction factor lesion root nematode (*P. brachyurus*), resistant to root-knot nematode, *M. incognita* e *M. javanica* and the cyst nematode (*H. glycines*) races 3, 4, 6, 9, 10 and 14.

Soybean seeds inoculated with a liquid inoculant containing *Bradyrhyzobium japonicum* SEMIA 5079 + SEMIA 5080 (Masterfix ®) and treated with Pyraclostrobin + Tiophanate-methyl + Fipronil (Standak Top ®). Site I was frequently irrigated by center-pivot to keep the moisture at field capacity while site II were in a rainfed condition. Fertilezers were used before sowing on the following rates 12 and 6 kg N ha^{-1}; 60 and 36 P_2O_5 ha^{-1}; and 60 kg K_2O ha^{-1} at each site, respectively.

2.2 Data collection

2.2.1 Pathogen and plant sampling

Sampling were made at the beginning of soybean seed filing rate (stage R5), 40 and 45 georeferenced points were selected on site I and II respectively. On each sampling point approximately 1L of soil around the rhizosphere and 30 g of roots from 5 consecutive soybean plants were collected from 0-20 cm depth. The soil and root samples were store in a cool and dark refrigerator. When the collect was done, we placed an EVA white plate on each point and a mapping was done with a Zenmuse RGB Camera for later georeferencing of the nematode collection points with the images of the MicaSense RedEdge (MicaSense Inc., Seattle, WA, USA; http://www.micasense.com/), using as control points physical features of the area perfectly identifiable in both images.

At harvest, yield was recorded after the physiological maturity of the grains. Two lines of 1 m were collected manually, for each point. Yield was estimated by correcting the water content to 0.13 kg kg^{-1} on a wet basis.

2.2.2 Nematodes extraction, identification and quantification

The procedures for nematode identification from soil and roots samples were taken at the Nematology Laboratory of São Paulo State University, Jaboticabal. Succeeding the entire sample homogenization, a volume of $100cm^3$ from each point were taken to extract nematodes by Jenkins (1964). Root samples were weighted in a scale, washed and cut into fragments of about 2 cm, crushed in a blender and the nematodes were extracted according to Coolen & D'herde (1972) proposed methodology.

Nematode identification were carried out using morphology and comparing with the original description of each species (Tihohod, 1989) with the assistance of a Peters chamber under an optical microscope. The estimated population was calculated by standardizing the result to 10 g of roots. For *H. glycines*, which is a species that just occurred in the Site II, was used Southey (1986) methodology to extract cysts from the soil.

After quantifying the nematodes, the number of individuals found in the soil (SNs) (Eq. 1), in the roots (SNr) (Eq. 2) and in the total (SNt) (Eq. 3) were calculated as follow the equations:

$$SNr = M.incognita + P.brachyurus + H.glycines + \text{eggs} \quad \text{(Eq. 1)}$$

$$SNs = M.incognita + P.brachyurus + H.glycines \quad \text{(Eq. 2)}$$

$$SNt = SNr + SNs \quad \text{(Eq. 3)}$$

2.2.3 Multispectral data

An unnamed aerial vehicle (UAV) DJI Matrice 100 (DJI, Shenzhen, China) was used as a platform to carry a RedEdge® (MicaSense Inc.®, Seattle, WA, USA) multispectral sensor, featuring five global shutter narrowband image bands: Blue (475±20 nm), Green (560±20 nm), Red (668±10 nm), Red edge (717±10 nm) and Near Infrared (NIR) (840±40 nm).

Flights were executed at R1 soybean growth stage, between 10 a.m. and 2 p.m. The UAV was regularly flown at 120m above the ground level, with 80% overlap between images that resulted in an 8cm spatial resolution in the final photogrammetric processing. The images were calibrated with assistance of a downwelling radiation sensor and photo from the calibration panel before the starting of the flight. The software Pix4D MapperPro version 3.3.13 (Pix4D SA, Lausanne, Switzerland) was used to calibrate the raw images captured into a reflectance orthomosaic. Yet, locational improvements were made implementing 5 control points in the software QGIS 3.12 (QGIS Development Team, 2017), Georeferencing plugin (Nearest Neighbor method). After that, it was possible to determine the center of the manual sampling in the processed imagery. For each point, a 1-m diameter buffer size was created from which mean band values were extracted using the ArcGIS (Redlards, US) software.

2.3 Data analysis

2.3.1 Nematode variability

Variability analysis of nematode occurrence was carried out using R software (R Code Team, 2020).

The control charts were separated by variables. Also, the model selected was the individual, in which control limits were established considering the variation of the data due to uncontrolled causes in the process (special causes), having been calculated based on the standard deviation of the variables, as shown in equations (4) and (5):

$$UCL = \bar{x} + 3O \quad \text{(Eq. 4)}$$

$$LCL = \bar{x} - 3O \quad \text{(Eq. 5)}$$

Where:

UCL: Upper limit control;

$\bar{x}$: Variable average;

O: Standard deviation;

LCL: Lower limit control.

2.3.2 Spectral band analysis and grain yield

Fixed effect linear models were conducted with each spectral band and grain yield as a responsible variable and as a result, the nematode sampling location (that is, plants inside, border and outside of the patches) as an explanatory variable using *lm* function of the stats package in R (R Core Team, 2017). Significant models were later analyzed, making paired comparisons between the results of the nematode sampling location at alpha = 0.05.

3. Results

3.1 Nematode population distribution in the areas

The main nematode species causing damage in soybean that occurred in site I were *M. incognita* and *P. brachyurus*. The SNs average (Figure 2a) was 1.8 and 1.9 times greater inside when compared to the border and outside of the patches, respectively. When observing the variability of the data, it appears that in the border of the patches was greater, followed by the inside and outside. We also noticed that the border region had a point out of control, in the same area where we observed high variability. Points out of control indicate instability in the sampling process, that is, the occurrence of special causes. For the other regions, the variability was random, indicating that there is natural variability in these regions.

Analyzing the nematode species in the soil individually (Figures 2b and 2c), we observed that *M. incognita* contributed more to cause damage, as its average was higher in all collection sites. The variability of the data grows along with the increase in the mean for both species and the most important point of instability occurred in the species *M. incognita*, as this was not diluted in the SNs.

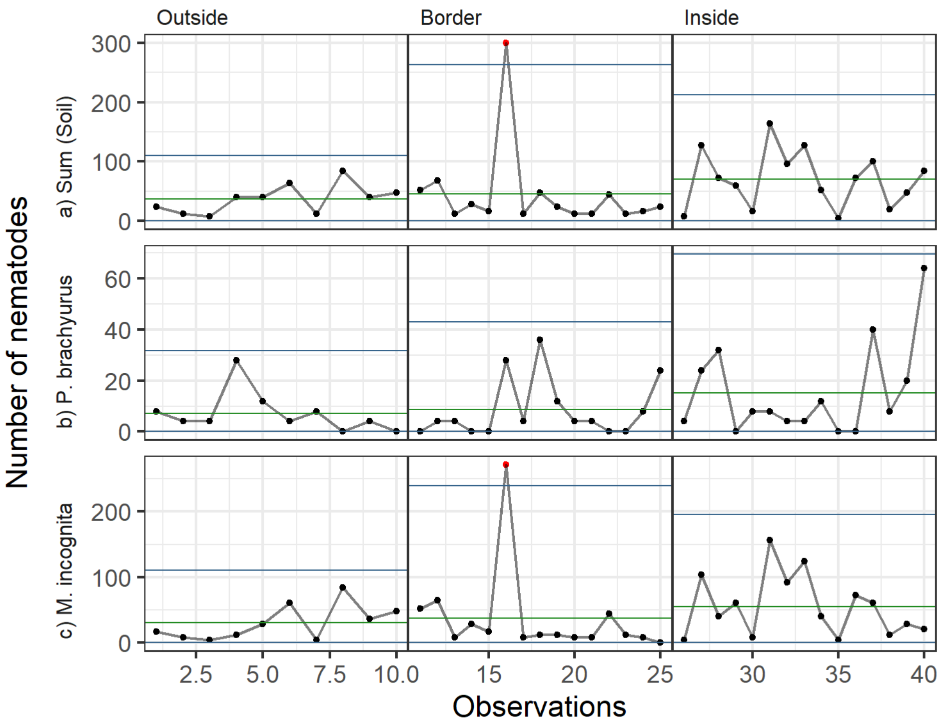

Figure 2. Control charts for (a) SNs, (b) *P. brachyurus* and (c) *M. incognita* counted in the soil (100 cm³) for site I.

The mean occurrence in SNr (Figure 3a) is higher in all sampling sites when compared to SNs (Figure 2a). In addition, the average of nematodes inside of the patches followed the same pattern as the SNs, remaining high.

When we unfold the SNr (Figures 3b and 3c), we observed that for both *P. brachyurus* and *M. incognita*, the variability inside the patches is greater and the unstable point that occurs in this region for both species is the same (observation 28). This sample point stood out for its high population density, being a source of inoculum that can spread throughout the area. This fact was also observed in SNr, being justified by this individual behavior of the species.

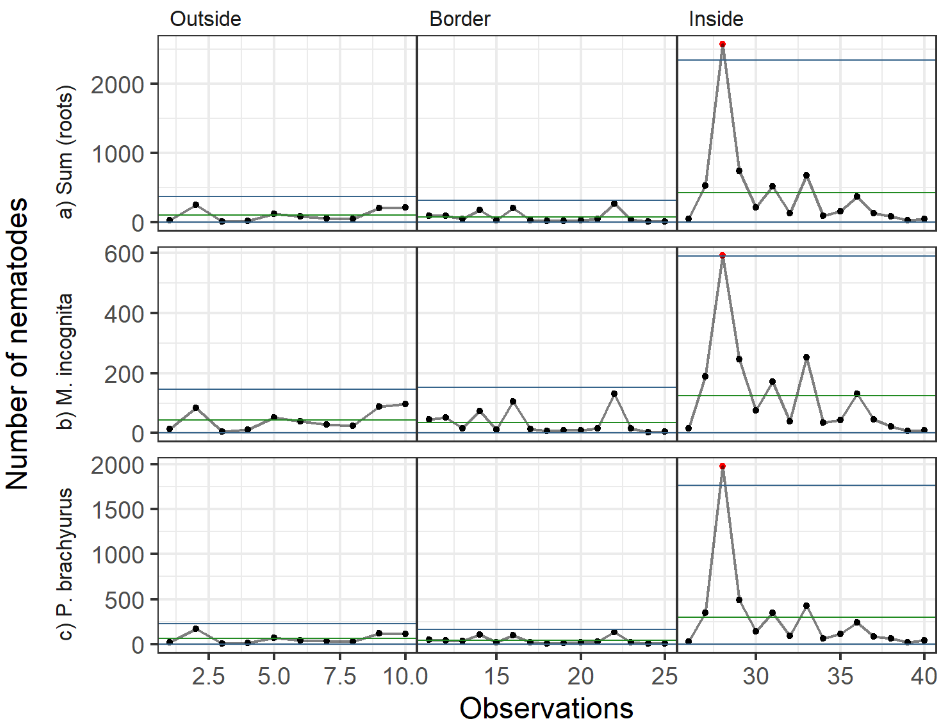

Figure 3. Control charts for (a) SNr, (b) *P. brachyurus* and (c) *M. incognita* counted at the root (10 g) of site I.

The average number of eggs at the border of the patches (Figure 4a) was 6.84 and 1.25 times greater than when compared to outside and at the end, respectively. The data variation was also greater in the region of the border of the patches.

In the SNt analysis (Figure 4b) we can see that the quantity of eggs was the one that most interfered, since the behavior of the data is similar in the higher points, averages, and in the variability.

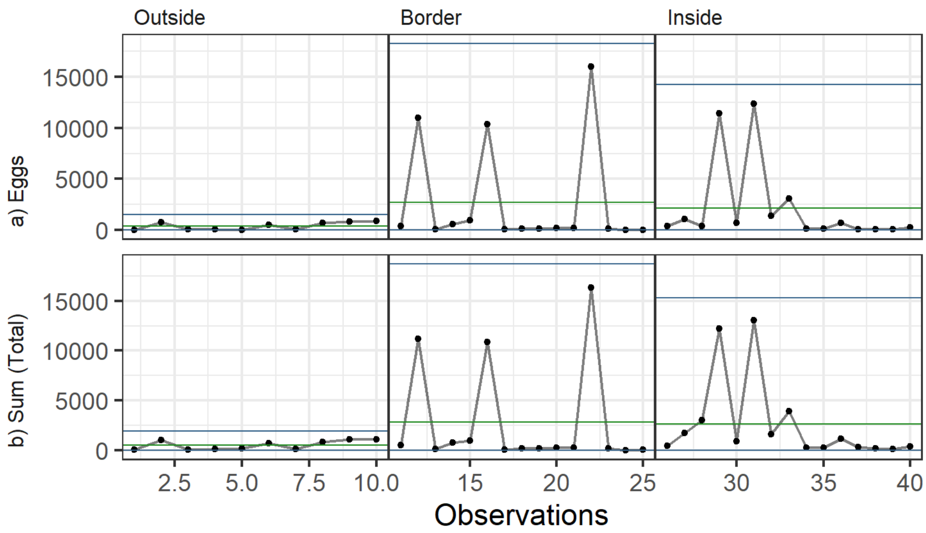

Figure 4. Control charts for (a) number of nematode eggs and (b) SNt.

As for site II, we observed that, in fact, the average of the SNs (Figure 5a) is the highest, as well as, we observed the great variability at the edge of the rocker and a point of instability. The other sampling regions have lower averages and more controlled variations, although both have points of instability.

Unfolding the SNs, we observed that *H. glycines* (Figure 5d), considered the most aggressive nematode for soybean (Jones et al., 2013), contributed 66.3% for the out of range point in the border of the patches, followed by *M. incognita* (Figure 5c) and *P. brachyurus* (Figure 5b). As it occurs in greater quantity, the SCN was responsible for the points of instability in SNs, the same happened with the behavior of the variability and average.

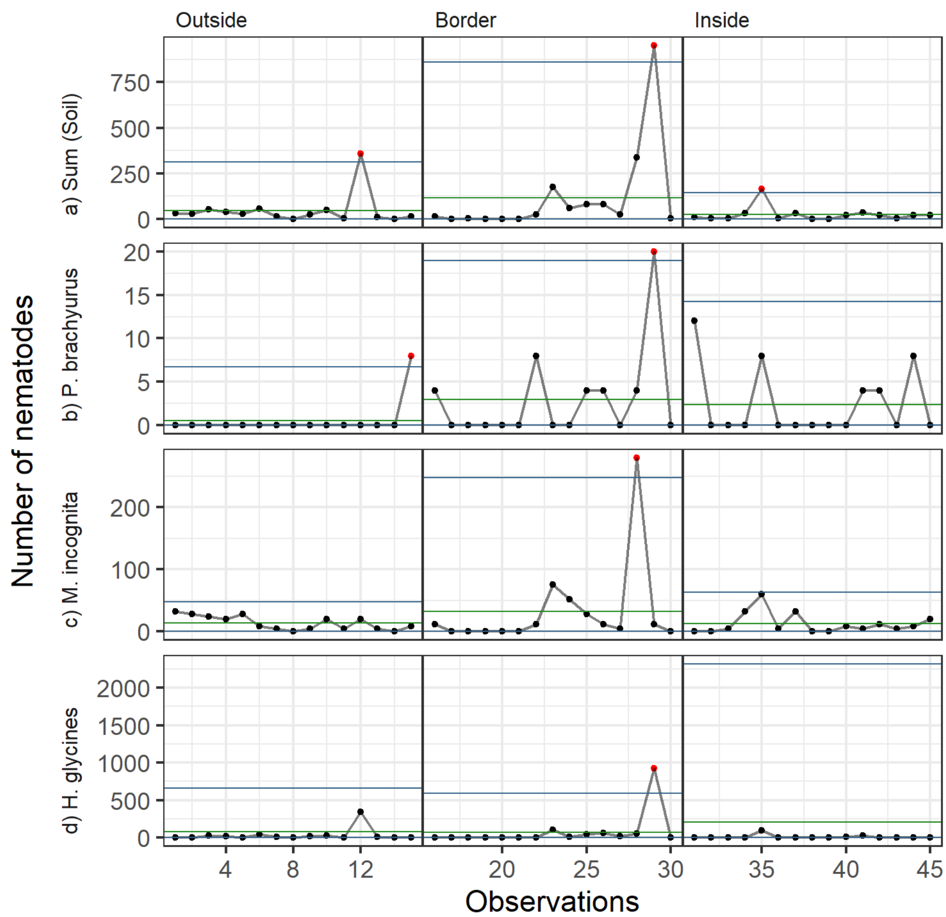

Figure 5. Control charts for (a) SNs, (b) *P. brachyurus*, (c) *M. incognita* and (d) *H. glycines* counted in the soil (100 cm³) in site II.

When the *H. glycines* females are fertilized and complete their life cycle, they detach from the roots forming a resistance structure called cyst and, within it, can contain up to 500 eggs. The non-viable cysts found in the area are those that already release J2 (second-stage juveniles), which are vermiform and their objective is to meet the roots. Viable cysts have not yet released J2s. We can see that viable cysts (Figure 6a) have a much lower average when compared to non-viable cysts (Figure 6b). This is due to the management adopted in past harvests, since the producer used only susceptible cultivars. The average of viable cysts was higher at the border, as well as the variability. These are the ones that remained dormant, knowing that they can last up to 40 years on the ground. The mean of non-viable cysts was higher at the border, however, with less variability.

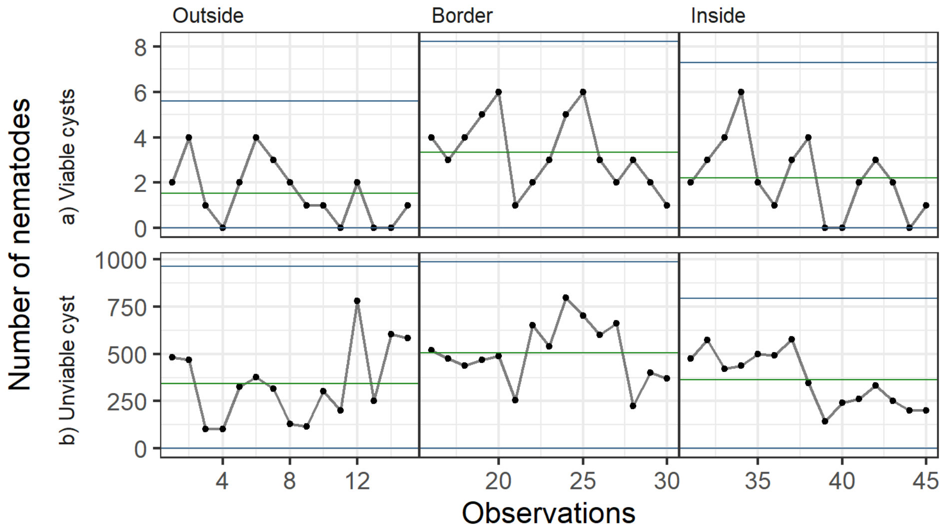

Figure 6. Control charts for number of (a) viable cysts and (b) unviable cysts counted on 100 cm^3 of soil at site II.

Regarding the mean and variability of nematodes occurrence at the root of site II, the RLN (Figure 7b) and SCN (Figure 7d) obtained higher values inside of the patches, while for gall nematode (Figure 7c) the highest mean and variability occurred outside of the patches.

We can observe that there are peaks in the occurrence of NG and eggs (Figures 7c and 8a), which allows us to infer that in the next harvest this region will have a higher occurrence of *M. incognita*, justified by the greater occurrence of eggs. In the same way, we observed the occurrence of SCN peaks and eggs in the region within the row (Figures 7d and 8a).

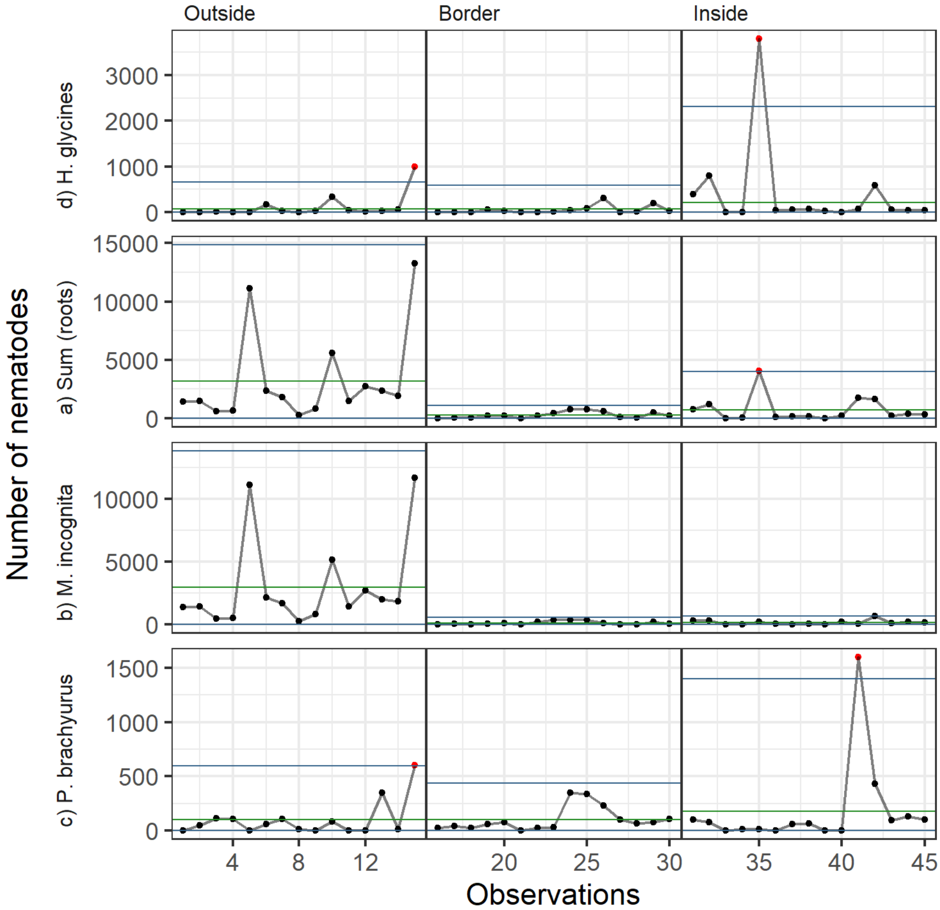

Figure7. Control charts for (a) SNr, (b) P. brachyurus, (c) *M. incognita* and (d) *H. glycines* counted in 10 g of root at site II.

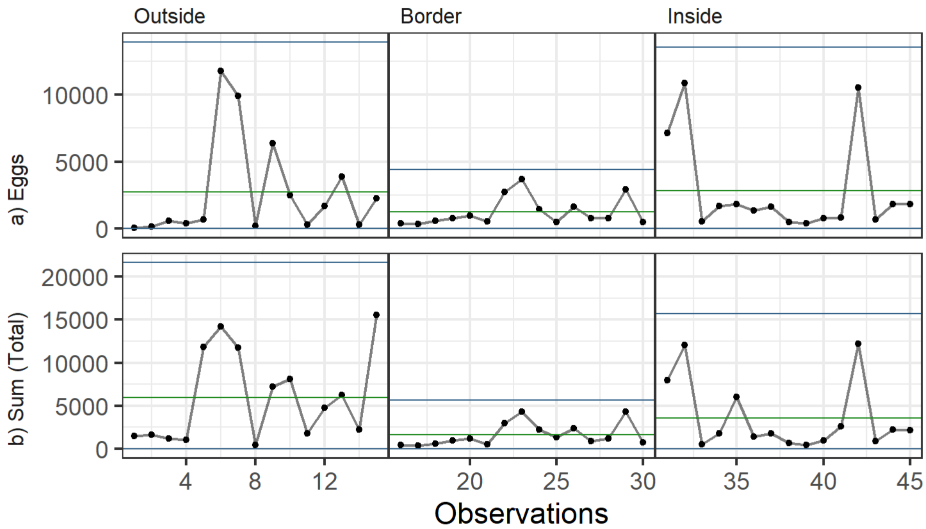

Figure 8. Control cards for (a) nematode's eggs in 10 g of root and (b) SNt counted in the area at site II.

The greatest variability and averages were found outside of the patches. However, no point was observed above the upper control limits, indicating that no collection region suffered from the action of special causes, that is, the detected variability is inherent in the process.

3.2 Multispectral bands and yield analysis

Soybeans yield average in site I was lower inside the patches when compared to the border, reaching a difference of 500 kg ha^{-1}. The outside mean did not differ from the means of inside and border of the patches. The visible region reflectance of the electromagnetic spectrum was different only in the red reflectance, and the others bands red-edge and NIR regions did not obtain any difference.

Soybean yield in site II also obtained a lower average on the inside of the patches when compared outside and at the border, which in turn obtained higher averages and did not differ. Visible reflectance spectrum obtained higher averages in the red band when located inside and at in the border of the patches. In the red-edge and NIR bands, the average was higher when located outside the patches.

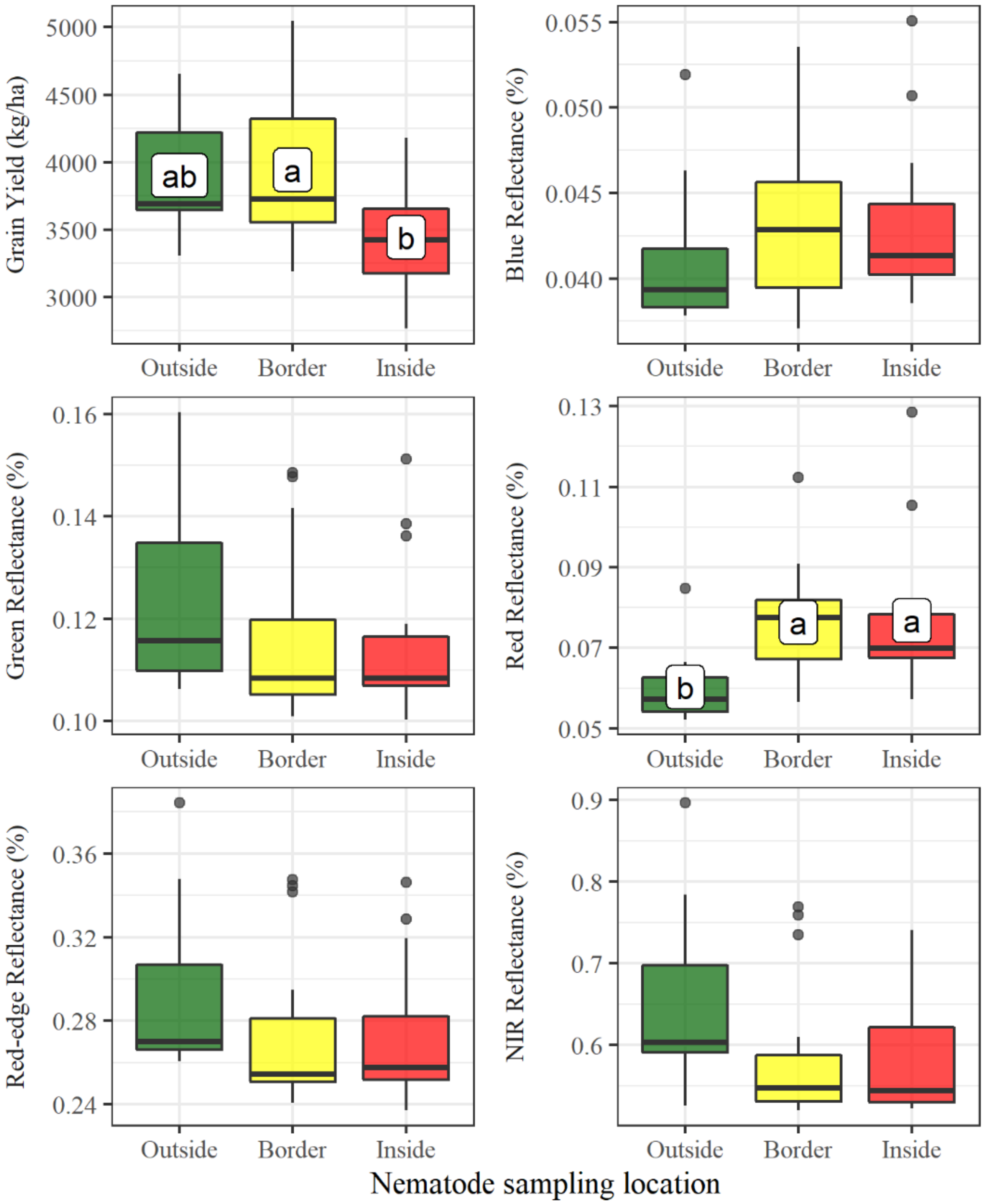

Figure 9. Grain yield (kg ha^{-1}) and multispectral reflectance (blue, green, red, red-edge and NIR) for each sampling site of the patches - inside (red), border (yellow) and outside (green) at site I. Letters are from the analysis of variance for the fixed-effect models and additional paired comparisons (important Tukey differences) were considered important at α = 0.05.

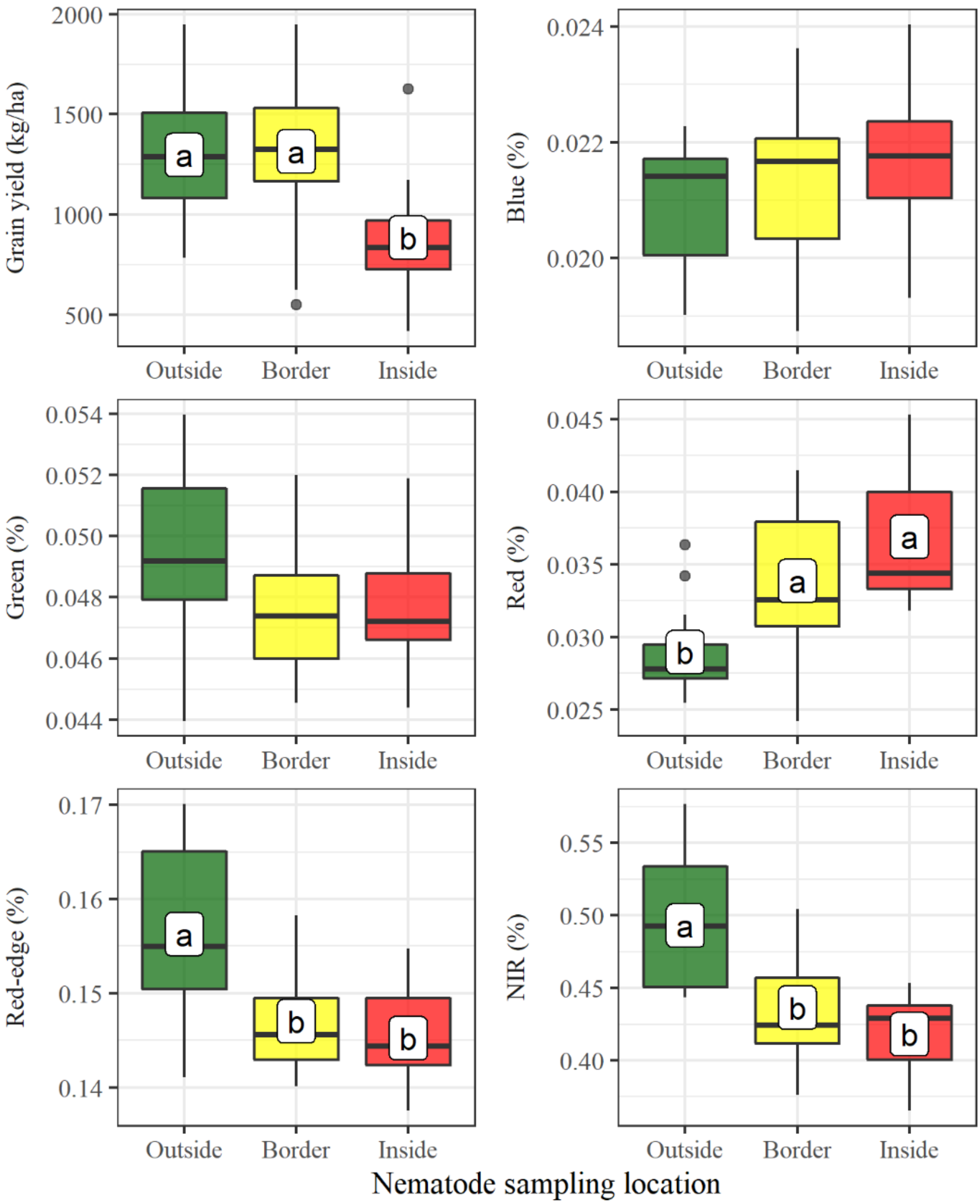

Figure 10. Grain yield (kg ha^{-1}) and multispectral reflectance (blue, green, red, red-edge and NIR) for each sampling site of the patches - inside (red), border (yellow) and outside (green) at site II. The letters are from the analysis of variance for the fixed-effect models and additional paired comparisons (significant Tukey differences) were considered significant at α = 0.05.

4. Discussion

This study addresses a perspective of a remote nematode detection in soybeans grown in the field to differentiate sites infected by nematodes, reducing collection time and intense work. In addition, it has the potential to be applied in the specific management of nematodes, use of resistant varieties and in the diversification of crop rotation in highly infested fields. The use of this tool for nematode detection has already been reported in the literature documenting that the visible and NIR bands are sensitive to the occurrence of nematodes (George Deroco Martins et al., 2017b; George Deroco Martins & Trindade Galo, 2015b; Nutter et al., 2002).

In the control charts analysis, points outside the upper or lower limits indicate the occurrence of special causes, caused by one or more factors (labor, machine, material, environment, method and measurement). The instability points indicate special causes in the nematode occurrence and are linked to the environment, mainly abiotic factors as soil that can vary the distribution of it in nature (Freckman & Virginia, 1989) and agriculture ecosystems (Duyck et al., 2012). Studies with environmental variables, such as topography, soil type and climate, were carried out to target nematode patterns assessing on a regional scale and showed contrasting results, making difficult to infer generalizations (Palomares-Rius et al., 2015).

Some plant parasitic nematodes, such as *M. incognita*, feed on the same plant cell for a very long period and lay their eggs in the same position. Such behavior results in a pattern of high spatial aggregation (Rossi et al., 1996), concentrating high populations at specific points, called patches (Pinheiro et al., 2008). Therefore, nematode’s spatial variability and migration also depends on external means such as agricultural machinery, contaminated seeds, animals, water and wind flow (Neher, 2010).

Nematode’s average found in roots was higher than the average found in the soil due to the parasitism habit of the main species found in the area*. P. brachyurus*, *M. incognita* are endoparasites, that is, they are contained within the root tissues, seeking to avoid encounters with predators and pathogens (Singh, 2015). In addition, obtaining higher averages at the border of the patches is also justified by the nematode's habit of parasitism. Nematodes are mandatory parasite, it needs the living tissue of the plant to reproduce and complete its life cycle (Vieira & Gleason, 2019). Plants that are located inside the patches have suffered from the infection since the beginning of its development and may lack living cells. Therefore, when the population increases rapidly more than the roots are capable of supporting, there is competition between nematodes and many die of starvation (Barbosa et al., 2013; Carvalho et al., 2013; Stirling, 1991).

Site II had a higher average of SNt, because it was infested with SCN, considered the most harmful parasite in soybean cultivation around the world (Kofsky et al., 2018) and underwent greater water stress because it did not have irrigation. Shekoofa & Sinclair (2018) reveals that the survival of annual crops still depends on the availability of water and the drought period is devastating for the producer's pocket. Studies show that environmental changes, such as lack of water, can influence the ability of plants to deal with biotic stresses, possibly increasing their susceptibility to parasitic infections (Newton et al., 2012). However, Snider et al., (2019), studied the water deficit and the infection of GN, *M. incognita*, in cotton, which negatively affected the production of cotton seeds. In other words, only the water supply was not compensatory for the negative effects of the nematode infection.

Plant parasitic nematodes have a specialized oral spear, called a stylus to penetrate plant cells, access the xylem and thus begin to ingest nutrients (Siddique & Grundler, 2018). Infections cause significant damage to the root system growth, predispose the roots to secondary pathogens invasion and jeopardize water and nutrients uptake (Jones et al., 2013). Several studies relate the chlorophyll content reduction, plant growth, photosynthesis rate and nutrient concentration to nematode infection (Carneiro et al., 2002; Haseeb et al. 1990; Nutter et al., 2002). Such factors were responsible for the lower productivity found inside the reel in both areas.

Due to aerial symptoms and the spatial growth in patches form, it is feasible to detect nematodes with the use of remote sensing (Hillnhütter et al., 2011b). The leaf visible reflectance is controlled by the presence of its constituents as nitrogen, chlorophyll, carotenoids and water (Ustin et al., 2004). The amount of solar radiation absorbed by the leaves is given by the photosynthetic components of the plant, like chlorophyll a and b and their concentration may vary. Consequently, low concentrations can limit photosynthetic potential and therefore grain yield (Richardson et al., 2002). In addition, plant's chlorophyll is related to the amount of nitrogen, and can be an indicator of physiological conditions, highlighting stress regions (Gitelson et al., 2002).

NIR is able to detect changes in the leaf structure of the canopy, just as the visible band can identify concentrations of photosynthetic pigments (Jensen, 2015; Sellers, 1987). These arguments explain the increase in reflectance at NIR and red-edge band and low reflectance in the red band at site II, mainly inside the patches (Figure 10), indicating physiological stress.

Soybean plants located inside the patches showed lower grain yield, greater reflectance in the red band in both areas. Site II, presented less reflectance in the red-edge and NIR bands, which was also verified by Bajwa et al. (2017b) in soybean plants infected with cyst nematode and sudden soybean death syndrome. The results obtained in this study corroborate other studies that used remote sensing to investigate plant stresses caused by soil pathogens. Martins et al. (2017a) also reported that the red, red-edge and NIR bands were used to define the spatial distribution of healthy, moderately infected and severely infected coffee plants, with an accuracy of 78%.

The main control method is the multiplication prevention and no propagation to uninfected areas, especially for pathogens with low soil mobility, such as nematodes. In addition, the high cost of sampling and the long period of time for conventional detection indicates the relevance of knowledge of spatial distribution in infested areas in the context of remote sensing.

5. Conclusion

In general, the greatest variability in the occurrence of nematodes was observed inside the patches for site I, both of the nematodes found in the soil and in the root. For site II, the greatest variability occurred in the region of the border (soil) and outside of the patches (roots). The lowest productivity was observed inside the patches for the two study sites.

It can be inferred that plants infected with nematodes will present higher values in the visible range and less reflectance in the RedEdge and NIR bands, when compared to less infested plants.

More detailed information on the behavior of the electromagnetic spectrum under nematode infection in cultivated plants can be further elaborated with the application of hyperspectral sensors. Future work should evaluate the behavior of the spectrum in each band and correlate with more data including abiotic factors (example: soil type, rainfall regime) and biotic factors (cultivars, severity levels).

Acknowledges

This work was carried out with the support of the Improvement Coordination of Higher Education Personnel in Brazil (CAPES) - Financing code 001.

References

Bajwa, S. G., Rupe, J. C., & Mason, J. (2017a). Soybean disease monitoring with leaf reflectance. *Remote Sensing*, *9*(2). https://doi.org/10.3390/rs9020127

Bajwa, S. G., Rupe, J. C., & Mason, J. (2017b). Soybean disease monitoring with leaf reflectance. *Remote Sensing*, *9*(2), 1–14. https://doi.org/10.3390/rs9020127

Barbosa, B. F. F., Dos Santos, J. M., Barbosa, J. C., Soares, P. L. M., Ruas, A. R., & De Carvalho, R. B. (2013). Aggressiveness of pratylenchus *brachyurus* to sugarcane, compared with key nematode P. zeae. *Nematropica*, *43*(1), 119–130.

Barker, K. R., & Koenning, S. R. (1998). Developing sustainable systems for nematode management. *Annual Review of Phytopathology*, *36*(1), 165–205. https://doi.org/10.1146/annurev.phyto.36.1.165

Blevins, D. G., Dropkin, V. H., & Luedders, V. D. (1995). Macronutrient uptake, translocation, and tissue concentration of soybeans infested with the soybean cyst nematode and elemental composition of cysts isolated from roots. *Journal of Plant Nutrition*, *18*(3), 579–591. https://doi.org/10.1080/01904169509364924

Campbell, J. B., & Wynne, R. H. (2011). Electromagnatic Radiation. In *Introduction to remote sensing* (5th ed, pp. 30–58). Guilford Press.

Carneiro, R. G., Mazzafera, P., Ferraz, L. C. C. B., Muraoka, T., & Trivelin, P. C. O. (2002). Uptake and translocation of nitrogen, phosphorus and calcium in soybean infected with *Meloidogyne incognita* and *M. javanica*. *Fitopatologia Brasileira*, *27*(2), 141–150. https://doi.org/10.1590/s0100-41582002000200004

Chang, K. F., Hwang, S. F., Conner, R. L., Ahmed, H. U., Zhou, Q., Turnbull, G. D., Strelkov, S. E., McLaren, D. L., & Gossen, B. D. (2015). First report of *Fusarium proliferatum* causing root rot in soybean (*Glycine max* L.) in Canada. *Crop Protection*, *67*, 52–58. https://doi.org/10.1016/J.CROPRO.2014.09.020

Carvalho, C., Fernandes, C. D., dos Santos, J. M., & Macedo, M. C. M. (2013). Densidade populacional de *pratylenchus* spp. em pastagens de *Brachiaria* spp. e sua influência na disponibilidade e na qualidade da forragem. *Revista Ceres*, *60*(1), 30–37. https://doi.org/10.1590/S0034-737X2013000100005

Coolen, W. A., & D'herde, C. J. (1972). A method for the quantitative extraction of nematodes from plant tissue. *A Method for the Quantitative Extraction of Nematodes from Plant Tissue*. https://www.cabdirect.org/cabdirect/abstract/19722001202

Duyck, P. F., Dortel, E., Tixier, P., Vinatier, F., Loubana, P. M., Chabrier, C., & Quénéhervé, P. (2012). Niche partitioning based on soil type and climate at the landscape scale in a community of plant-feeding nematodes. *Soil Biology and Biochemistry*, *44*(1), 49–55. https://doi.org/10.1016/j.soilbio.2011.09.014

Eberlein, C., Heuer, H., & Westphal, A. (2016). Impact of cropping sequences and production strategies on soil suppressiveness against cereal cyst nematodes. *Applied Soil Ecology*, *107*, 381–393. https://doi.org/10.1016/J.APSOIL.2016.07.012

Freckman, D. W., & Virginia, R. A. (1989). Plant-feeding nematodes in deep-rooting desert ecosystems. *Ecology*, *70*(6), 1665–1678. https://doi.org/10.2307/1938101

Gitelson, A. A., Kaufman, Y. J., Stark, R., & Rundquist, D. (2002). Novel algorithms for remote estimation of vegetation fraction. *Remote Sensing of Environment*, *80*(1), 76–87. https://doi.org/10.1016/S0034-4257(01)00289-9

Grabau, Z. J., & Chen, S. (2016). Influence of long-term corn-soybean crop sequences on soil ecology as indicated by the nematode community. *Applied Soil Ecology*, *100*, 172–185. https://doi.org/10.1016/j.apsoil.2015.12.016

Haseeb, A., Srivastava, N. K., & Pandey, R. (1990). The influence of *meloidogyne incognita* on growth, physiology, nutrient concentration and alkaloid yield of hyoscy amus niger. *Nematologia Mediterranea*, 127–129.

Hassan, M. A., Pham, T. H., Shi, H., & Zheng, J. (2013). Nematodes threats to global food security. *Acta Agriculturae Scandinavica Section B: Soil and Plant Science*, *63*(5), 420–425. https://doi.org/10.1080/09064710.2013.794858

Hillnhütter, C., Mahlein, A. K., Sikora, R. A., & Oerke, E. C. (2011). Remote sensing to detect plant stress induced by *Heterodera schachtii* and *Rhizoctonia solani* in sugar beet fields. *Field Crops Research*, *122*(1), 70–77. https://doi.org/10.1016/j.fcr.2011.02.007

Hillnhütter, Christian, Mahlein, A. K., Sikora, R. A., & Oerke, E. C. (2012). Use of imaging spectroscopy to discriminate symptoms caused by *Heterodera schachtii* and *Rhizoctonia solani* on sugar beet. *Precision Agriculture*, *13*(1), 17–32. https://doi.org/10.1007/s11119-011-9237-2

Jenkins, W. R. (1964). A rapid centrifugal-flotation technique for separating nematodes from soil. *Plant Disease Reporter*, 48(9). https://www.cabdirect.org/cabdirect/abstract/19650801105

Jensen, J. (2015). *Introductory digital image processing: a remote sensing perspective*. https://dl.acm.org/citation.cfm?id=2834479

Jones, J. T., Haegeman, A., Danchin, E. G. J., Gaur, H. S., Helder, J., Jones, M. G. K., Kikuchi, T., Manzanilla-López, R., Palomares-Rius, J. E., Wesemael, W. M. L., & Perry, R. N. (2013). Top 10 plant-parasitic nematodes in molecular plant pathology. *Molecular Plant Pathology*, *14*(9), 946–961. https://doi.org/10.1111/mpp.12057

Kofsky, J., Zhang, H., & Song, B. H. (2018). The untapped genetic reservoir: The past, current, and future applications of the wild soybean (*Glycine soja*). In *Frontiers in Plant Science* (Vol. 9). Frontiers Media S.A. https://doi.org/10.3389/fpls.2018.00949

Martins, G. D., Galo, M. D. L. B. T., & Vieira, B. S. (2017a). Detecting and Mapping Root-Knot Nematode Infection in Coffee Crop Using Remote Sensing Measurements. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(12), 5395–5403. https://doi.org/10.1109/JSTARS.2017.2737618

Martins, G. D., Galo, M. D. L. B. T., & Vieira, B. S. (2017b). Detecting and Mapping Root-Knot Nematode Infection in Coffee Crop Using Remote Sensing Measurements. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(12), 5395–5403. https://doi.org/10.1109/JSTARS.2017.2737618

Martins, G. D., & Trindade Galo, M. D. L. B. (2015a). Caracterização espectral da cana-de-açúcar infectada por nematoides e *Migdolus fryanus* por espectrorradiometria de campo. *Boletim de Ciencias Geodesicas*, *21*(4), 783–796. https://doi.org/10.1590/S1982-21702015000400046

Martins, G. D., & Trindade Galo, M. D. L. B. (2015b). Caracterização espectral da cana-de-açúcar infectada por nematoides e *Migdolus fryanus* por espectrorradiometria de campo. *Boletim de Ciencias Geodesicas*, *21*(4), 783–796. https://doi.org/10.1590/S1982-21702015000400046

Meyer, M. C., Favoreto, L., Klepker, D., & Marcelino-Guimarães, F. C. (2017). Soybean green stem and foliar retention syndrome caused by *Aphelenchoides besseyi*. *Tropical Plant Pathology*, *42*(5), 403–409. https://doi.org/10.1007/s40858-017-

0167-z

Nasu, É. das G. C., Amora, D. X., Monteiro, T. S. A., Alves, P. S., de Podestá, G. S., Ferreira, F. C., & de Freitas, L. G. (2018). *Pochonia chlamydosporia* applied via seed treatment for nematode control in two soil types. *Crop Protection*, *114*, 106–112. https://doi.org/10.1016/j.cropro.2018.08.010

Neher, D. A. (2010). Ecology of Plant and Free-Living Nematodes in Natural and Agricultural Soil. *Annual Review of Phytopathology*, *48*(1), 371–394. https://doi.org/10.1146/annurev-phyto-073009-114439

Newton, A. C., Torrance, L., Holden, N., Toth, I. K., Cooke, D. E. L., Blok, V., & Gilroy, E. M. (2012). Climate Change and Defense against Pathogens in Plants. In *Advances in Applied Microbiology* (Vol. 81, pp. 89–132). Academic Press Inc. https://doi.org/10.1016/B978-0-12-394382-8.00003-4

Niblack, T. L. (2005). Soybean cyst nematode management reconsidered. In *Plant Disease* (Vol. 89, Issue 10, pp. 1020–1026). https://doi.org/10.1094/PD-89-1020

Nutter, F. W., Tylka, G. L., Guan, J., Moreira, A. J. D., Marett, C. C., Rosburg, T. R., Basart, J. P., & Chong, C. S. (2002). *Use of Remote Sensing to Detect Soybean Cyst Nematode-Induced Plant Stress 1*.

Oerke, E. C. (2006). Crop losses to pests. In *Journal of Agricultural Science* (Vol. 144, Issue 1, pp. 31–43). https://doi.org/10.1017/S0021859605005708

Palomares-Rius, J. E., Castillo, P., Montes-Borrego, M., Navas-Cortés, J. A., & Landa, B. B. (2015). Soil Properties and Olive Cultivar Determine the Structure and Diversity of Plant-Parasitic Nematode Communities Infesting Olive Orchards Soils in Southern Spain. *PLOS ONE*, *10*(1), e0116890. https://doi.org/10.1371/journal.pone.0116890

Pinheiro, J. B., Pozza E. A., Pozza A. A. A., Moreira A. S., Alves M. C., & Campos V. P. (2008). *Influência da Nutrição Mineral na Distribuição Espacial do Nematóide de Cisto da Soja*. 270–278.

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. https://doi.org/10.1046/j.0028-646X.2001.00289.x

Rossi, J. P., Delaville, L., & Quénéhervé, P. (1996). Microspatial structure of a plant-parasitic nematode community in a sugarcane field in Martinique. *Applied Soil Ecology*, *3*(1), 17–26. https://doi.org/10.1016/0929-1393(95)00067-4

Sellers, P. J. (1987). Canopy reflectance, photosynthesis, and transpiration, II. The role of biophysics in the linearity of their interdependence. *Remote Sensing of Environment*, *21*(2), 143–183. https://doi.org/10.1016/0034-4257(87)90051-4

Shekoofa, A., & Sinclair, T. (2018). Aquaporin Activity to Improve Crop Drought Tolerance. *Cells*, *7*(9), 123. https://doi.org/10.3390/cells7090123

Siddique, S., & Grundler, F. M. (2018). Parasitic nematodes manipulate plant development to establish feeding sites. In *Current Opinion in Microbiology* (Vol. 46, pp. 102–108). Elsevier Ltd. https://doi.org/10.1016/j.mib.2018.09.004

Singh, R. . P. S. (2015). *Environmental Issues for Socio-ecological Development*. https://doi.org/10.13140/RG.2.1.4567.6405

Snider, J. L., Davis, R. F., Earl, H. J., & Timper, P. (2019). Water availability and root-knot nematode management alter seedcotton yield through similar effects on fruit distribution patterns. *Field Crops Research*, *233*, 88–95. https://doi.org/10.1016/j.fcr.2019.01.008

Southey, J. F. (1986). *Laboratory methods for work with plant and soil nematodes*. Technical bulletin 2. London: Her Majesty’s Stationery Office. http://agris.fao.org/agris-search/search.do?recordID=US201300438332

Stirling, G. R. (1991). Biological control of plant parasitic nematodes. *Biological Control of Plant Parasitic Nematodes.*

Tihohod, D. (1993). *Nematologia agrícola* (2nd edn). Funep.

Ustin, S. L., Roberts, D. A., Gamon, J. A., Asner, G. P., & Green, R. O. (2004). Using imaging spectroscopy to study ecosystem processes and properties. Bioscience, 54(6), 523-534.

Vieira, P., & Gleason, C. (2019). Plant-parasitic nematode effectors — insights into their diversity and new tools for their identification. In *Current Opinion in Plant Biology* (Vol. 50, pp. 37–43). Elsevier Ltd. https://doi.org/10.1016/j.pbi.2019.02.007

Yan, G., & Baidoo, R. (2018). Current Research Status of *Heterodera glycines* Resistance and Its Implication on Soybean Breeding. *Engineering*, *4*(4), 534–541. https://doi.org/10.1016/J.ENG.2018.07.009

Zhang, J., Huang, Y., Pu, R., Gonzalez-Moreno, P., Yuan, L., Wu, K., & Huang, W. (2019). Monitoring plant diseases and pests through remote sensing technology: A review. In *Computers and Electronics in Agriculture* (Vol. 165, p. 104943). Elsevier B.V. https://doi.org/10.1016/j.compag.2019.104943

Chapter 3 - Identifying nematode damage on soybean via remote sensing and machine learning techniques

Abstract

Nematode identification process in large soybean areas is cost-intensive and not always achievable in a practical way. Multispectral reflectance sensors have been tested to detect different crop stresses, but their use assessing nematode damage in soybeans (*Glycine max* L.) has not been thoroughly evaluated especially with the use of small unmanned aerial vehicle (sUAV). The main objectives of this study were to: i) determine the bivariate relationship between individual spectral bands and vegetation indices (VIs) all relative to the soybean condition (infected versus non-infected plants); and ii) select the best model testing for identifying plant condition using three algorithms (logistic regression - LR, random forest - RF, conditional inference tree - CIT) and three options for data input using only bands (reduced model), only VIs, and bands plus VIs (full model). The experiment was conducted in Brazil on three on-farm soybean fields presenting different species of nematode infestation. Multispectral imagery was obtained using a drone-mounted MicaSense RedEdge® sensor. At each sampling georeferenced point nematode infestation and spectral measurements of soybean plants were retrieved. Multispectral bands, plus VI image were assessed for a classification of infected and non-infected areas, according to the threshold level adopted. Bivariate analysis of variance (ANOVA), LR, RF, and CIT were used to select the multispectral bands/VIs that best discriminated among infected and non-infected soybean plants, assessing the best model via quantification of their respective parameters for accuracy, sensitivity, and specificity. The greatest classification accuracy (>0.70) was achieved when using the CIT algorithm with the spectral bands only, with green (560±20 nm) and near infrared (840±40 nm) included as the main spectral input variables in the model. These results demonstrate the capacity for remotely-sensed multi-spectral data collected via utilization of sUAS of distinguishing nematode-infected and non-infected soybean plants using aerial sensing combined with machine learning.

Keywords: Remote sensing, nematode, Multispectral mapping, disease detection, machine learning, Digital Agriculture.

1. Introduction

Globally, the estimated economic losses due to soilborne plant nematodes is roughly US$ 216 billion each year for about 20 crops considered essential in world subsistence (Askary T.H. & Martinelli P.P., 2015). Soybean (*Glycine max* L.) production can be significantly reduced up to 31% due to nematode damage (Allen et al., 2017). To prevent the spread and introduction of other invasive species early damage detection is important (Zhao et al., 2009). However, nematode occurrence is only evident via crop response, spatially concentrated but unstructured, and manual surveillance is both time- and labor-intensive, complicating its accurate and rapid prediction for timely management action.

Nematodes infect plant roots causing problems in plant water and nutrient absorption and translocation (Tihohod, 1993b). The aboveground symptoms include stunted growth and leaf chlorosis, challenging its correct diagnoses and typically confounding this symptom with other stresses such as drought and nutrient deficiencies (Blevins et al., 1995). Nonetheless, nematodes are spatially distributed in patches due to their low soil mobility, making its field-level detection prone to remote sensing use (Hillnhütter et al., 2011).

Nematode diagnosis and quantification of severity levels in large soybean areas is expensive and not always achievable in a practical way (George Deroco Martins et al., 2017b). Laboratory analysis of several soil and root samples distributed over the entire field is required for a detailed evaluation of potential damage (Wiesel et al., 2015), both labor- and cost-intensive. Therefore, a rapid and affordable method to diagnose its occurrence involves the development of rapid phenotyping (e.g. via remote sensing) detection methods in site-specific management.

Spectral reflectance bands have been studied as a useful tool in detecting plant stress due to the variance in the reflected light in the visible (VIS) and near infrared (NIR) range of the electromagnetic spectrum (Garcia-Ruiz et al., 2013; George Deroco Martins et al., 2017b). (Hillnhütter et al., 2011) Christian Hillnhutter et al. (2011) reported that hyperspectral imaging was useful to detect and discriminate the development of root diseases. (Bajwa et al., 2017) Bajwa et al., were able to discriminate between non-infected and infected plants based on the use of spectral reflectance and vegetation index (VI). Therefore, the use of spectral bands and VIs to identify nematode-infected and non-infected soybean plants should be investigated as an important strategy to improve nematode management and control.

The analysis of remote sensing to detect plants infected with diseases can be challenging and different statistical approaches have been evaluated for improving screening of this farming issue. Those include parametric methods such as logistic regression (LR) (Delalieux et al., 2007; Steddom et al., n.d.), and non-parametric classification methods such as decision tree-based algorithms (Adelabu et al., 2014). Two important decision tree methods are random forest (RF) and conditional inference tree (CIT). The RF algorithm has been used in multiple fields of remote sensing for imagery data extraction (Rodriguez-Galiano et al., 2012) with a high degree of accuracy but limited interpretability. The CIT is an algorithm used for recursive variable selection and binary partitioning on the explanatory variables according to a significance test and outputs of an interpretable decision tree model (Hothorn et al., 2006). These methods can be useful on identifying complex relationships between response and explanatory variables, and the choice of algorithm should be made based both on its performance and interpretability (Yuan et al., 2020). Remote sensing can create large amounts of data, with machine learning potentially assisting in untangling complex relationships and aid in plant protection management (Susič et al., 2018).

In summary, this paper aims to investigate the capability of remotely-sensed multispectral data for detection of nematode-infected soybean areas. The aims of this study were to: i) determine the bivariate relationship between individual spectral bands and VIs and soybean condition (infected versus non-infected); and ii) select the best model testing for identifying plant condition using three algorithms (logistic regression - LR, random forest - RF, conditional inference tree - CIT) and three options for data input using only bands (reduced model), only VIs, and bands plus VIs (full model).

2. Material and methods

2.1 Sites Description

Three different site experiments were conducted during the 2018/19 summer soybean growing season. The sites were located near Guaíra (site I and II) and Jaboticabal (site III) in the state of São Paulo, Brazil (Figure 1). Sites were chosen based on previous natural nematode occurrence (**Erro! Fonte de referência não encontrada.**).The previous crop was common beans at sites I and II, and corn at site III. The three sites have been managed under no-tillage practices. The growing season maximum and minimum average air temperature of site I and II was 30.0°C and 21.2°C, respectively and the cumulative rainfall was 822 mm, and for site III the averages air temperatures were 30.4 and 20.6°C, respectively, and cumulative rainfall of 551 mm.

Table 1. Coordinates, altitude, total area and nematode genus infestation for each site.

Site	Elevation (m)*	Nematode infestation genus	Area (ha)
I	502	*M. incognita*, *P. brachyurus* and *H. glycines*	16.1
II	540	*M. incognita*, *P. brachyurus*	8.2
III	630	*M. incognita*, *P. brachyurus* and *R. reniformis*	19.7

**Above sea level.*

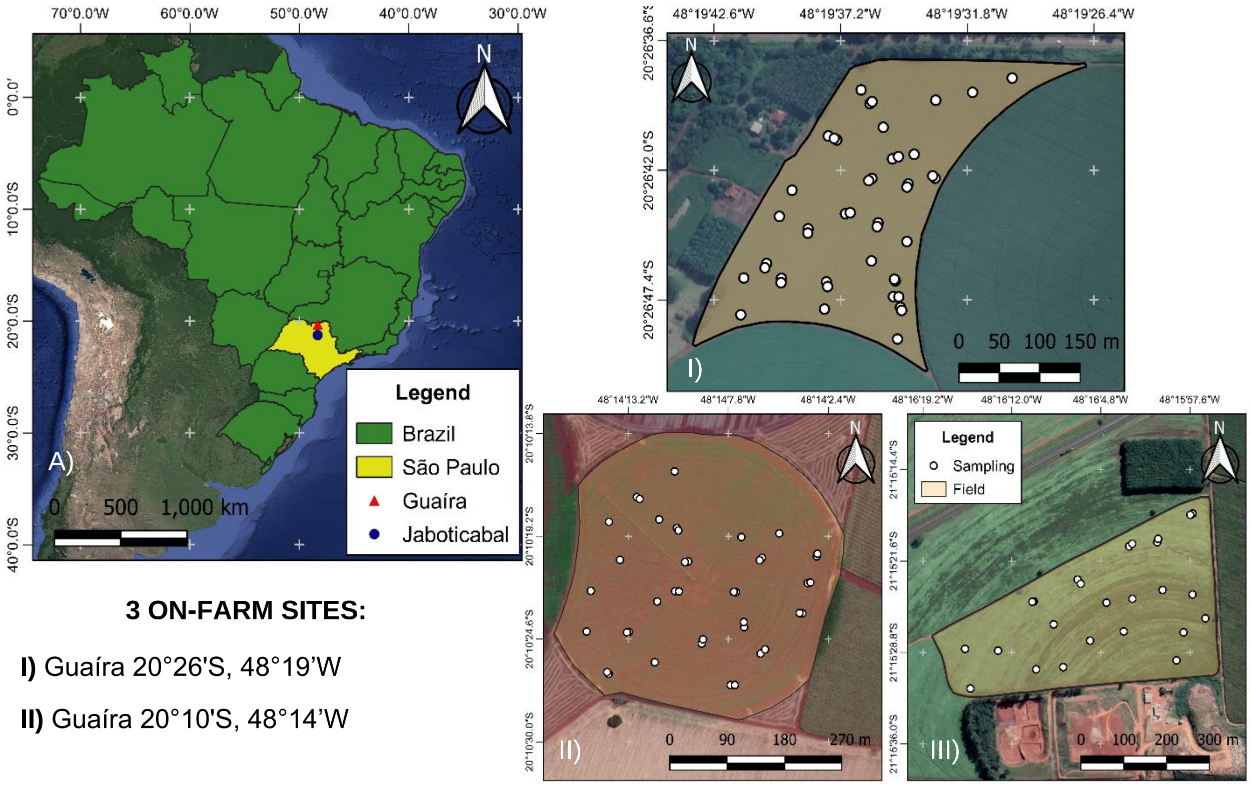

Figure 2. Characterization of the experimental area. A) Guaíra and Jaboticabal cities highlighted in the São Paulo and Brazil maps. I), II) and III) site experiments with white points representing the georeferenced sampling points.

Site I was sown at 290,000 seeds ha^{-1} on November 3rd with soybean cultivar BRS 7380 (Embrapa Ltd., Brazil), RR, resistant to cyst nematode (*H. glycines*) races 3, 4, 6, 9, 10 and 14; gall nematode, *M. incognita* e *M. javanica* and low reproduction factor to the lesion root nematode (*P. brachyurus*). Site II was sown at 320,000 seeds ha^{-1} on November 15th with soybean cultivar DS 5916 (Don Mario, Argentina), resistant only to *M. incognita*. Site III was sown at 290,000 seeds ha^{-1} on November 5th with soybean cultivar TMG 7063 (Tropical Melhoramento e Genética Ltda., Cambé-PR, Brazil), susceptible to nematodes.

Soybean seeds were treated with Pyraclostrobin + Tiophanate-methyl + Fipronil (Standak Top ®) and inoculated with *Bradyrhyzobium japonicum* SEMIA 5079 + SEMIA 5080 (Masterfix ® liquid inoculant). Sites I and III were rainfed and site II was center-pivot irrigated frequently to keep moisture at field capacity. Sites I, II and III were fertilized before sowing with a granular fertilizer at the rates of 12, 6, and 12 kg N ha^{-1}; 60, 36 and 60 kg P_2O_5 ha^{-1}; and 60 kg K_2O ha^{-1} at each site, respectively. Weeds, diseases and pests were controlled according to regional best management practices for soybean.

2.2 Data collection

2.2.1 Pathogen and plant sampling

At the R5 soybean growth stage (beginning of seed filling), 45, 40, and 30 geo-referenced and randomized points were selected on sites I, II and III, respectively. At each point, approximately 1L of soil and 30 g of soybean pivoting and radicle roots from 5 consecutive plants were collected from the 0-to-20 cm depth soil layer. Soil and root samples were stored in a dark refrigerator. After sample collection, an EVA white plate was placed at each point. A mapping was done with the Zenmuse RGB Camera for later georeferencing of the nematode collection points with the images of the MicaSense RedEdge (MicaSense Inc., Seattle, WA, USA; http://www.micasense.com/), using as control points physical features of the area perfectly identifiable in both images.

2.2.2 Nematodes extraction, identification and quantification

Nematode species were identified from soil and plant samples at the Nematology Laboratory of São Paulo State University, Jaboticabal. The samples were homogenized and a 100-cm³ soil volume representing each point was sampled, using the method from Jenkins (1964). For root samples, total volume was measured on a digital scale, the roots were washed and cut into pieces of about 2 cm, crushed in a blender, and the nematodes were extracted according to the methodology proposed by Coolen and D’herde (1972).

The identification and determination of the number of nematodes was performed morphologically following the method proposed by (Tihohod, 1993a) using 1 mL aliquots with the aid of the Peters chamber, under an optical microscope. From the results obtained, the nematode population was estimated by standardizing the result to 10 g of roots. For *H. glycines* species, which occurred only on site I, the cysts were extracted from the soil according to the methodology proposed by (Southey, 1986).

2.2.3 Multispectral data and VI calculation

A small unmanned aerial vehicle (sUAV) DJI Matrice 100 (DJI, Shenzhen, China) was used as the sensor-carrying platform. The sUAV was mounted with a RedEdge® (MicaSense Inc.®, Seattle, WA, USA) multispectral sensor, featuring five global shutter narrowband image bands: Blue (475±20 nm), Green (560±20 nm), Red (668±10 nm), Red edge (717±10 nm) and Near Infrared (NIR) (840±40 nm).

Flights were performed between 10 a.m. and 2 p.m., at the R1 soybean growth stage. The sUAV was constantly flown 120 m above the ground level, resulting in a spatial resolution of 8 cm with 80% overlap between adjacent images to avoid gaps and allow subsequent photogrammetric processing. The images were calibrated with assistance of a downwelling radiation sensor and photo from the calibration panel before the starting of the flight. The raw images were calibrated and collated into a reflectance orthomosaic using photogrammetric software Pix4D MapperPro version 3.3.13 (Pix4D SA, Lausanne, Switzerland). Further improvements in locational accuracy were obtained by establishing 5 control points using the software QGIS 3.12 (QGIS Development Team, 2017), Georeferencing plugin (Nearest Neighbor method). The center of the manual sampling location points was identified in the processed imagery. For each point, a 1-m diameter buffer size was created from which mean band values were extracted and vegetation indices were then calculated using the ArcGIS (Redlards, US) software (Table 3).

Table 2. Vegetation Index (VI), equation and citation used to analyze nematode infection.

VI	Description	Equation	Reference
NDVI	Normalized Difference Vegetation Index	(NIR - R) / (NIR+ R)	(Rouse et al., 1974)
GNDVI	Green Normalized Difference Vegetation Index	(NIR - Green) / (NIR + Green)	(Anatoly A. Gitelson et al., 1996)
SR	Simple Ratio	NIR / R	(Jordan, 1969)
RDVI	Renormalized Difference Vegetation Index	(NIR - R) / √ (NIR + R)	(Roujean & Breon, 1995)
SAVI	Soil Adjusted Vegetation Index	(NIR-R) / (NIR+R+L) * (1+L)	(Huete, 1988)
NDRE	Normalized Difference Red edge	(NIR-RE) / (NIR+RE)	(Barnes et al., 2000)
EVI	Enhanced Vegetation Index	2.5 * (NIR - R) / (NIR + 6 * R -7.5*B) +1	(H. Q. Liu & Huete, 1995)
VARI	Visible Atmospherically Resistant Index	(G – R) / (G + R – B)	(Gitelson et al., 2002)
NLI	Non-Linear Index	(NIR^2 - R) / (NIR^2 + R)	(Gong et al., 2003)

2.3 Data analysis

2.3.1 Nematode infection classification

Based on the nematode laboratory results, each sample were classified as infected and non-infected if any nematode count was above and below the threshold level, respectively (Tabela 4), adapted from (Koenning S.R., 2007) Koenning and (Dickerson et al., 2000) Dickson et al. (2000).

Table 3. Thresholds to classify non-infected and infected plants for the training model.

Nematode genus	Soil	Root
H. glycines (Cysts)	5	-
H. glycines	400	-
P. brachyurus	100	800
M. incognita	300	120
R. reniformis	300	-

2.3.2 Bivariate analysis of individual bands and VIs

Linear fixed-effect models were conducted with each individual band and VI as the response variable and nematode infection classification outcome (i.e., non-infected vs. infected plants) as the explanatory variable using the function *lm* from package stats in R (R Core Team, 2020). Significant models were further analyzed by performing pairwise comparisons between nematode infection classification outcome at alpha=0.05.

2.3.3 Nematode infection prediction

Nematode infection classification outcome was modelled as a function of bands, VIs, and band plus VIs using three different approaches: LR, RF and CIT. The explanatory variables (i.e., bands and VIs) were split in three sets in order to identify the spectral characteristics that allow more evident separation of the nematode classification. Those sets were: i) only bands (Blue, Green, Red, Red Edge and NIR); ii) only VIs (all VIs on); and iii) bands plus VIs.

The LR model was performed using the function *glm* (with binomial error distribution and link function) from package *stats* (R Core Team, 2020), RF models with function *randomForest* from package *randomForest* (Liaw & Wiener, 2002) , and CIT models with function *ctree* from package *partykit* (Hothorn et al., 2006). For RF and CIT, leave-one-out cross validation was used both to select the most appropriate hyperparameter values and to calculate unbiased model performance metrics. All analysis were performed using R software (R Code Team, 2020).

For the RF algorithm, all combinations between three hyperparameters were tested, including the number of trees (*ntree* = 50,100,300,600), the number of variables as predictors for each split (*mtry* = 1,2,4,6) and the minimum size in each terminal node (*nodesize* = 1,3,6,9). For the CIT algorithm, all combinations between two hyperparameters were tested, including the significance level for variable selection (*alpha* = 0.1, 0.05, 0.01) and the maximum depth of the tree (*maxdepth* = 2, 3, 4, 5). For both algorithms, the combination of hyperparameters that maximized classification accuracy was chosen for subsequent analysis. Therefore, after model optimization, RF selected hyperparameter values for *ntree*, *mtry*, and *nodesize* were 600, 4, and 1 for bands; 600, 1, and 1 for VIs; and 50, 2, and 9 for bans plus VIs, respectively. The CIT selected hyperparameter values for *alpha* and *maxdepth* were 0.01 and 2 for bands; 0.01 and 2 for VIs; and 0.05 and 2 for bands plus VIs, respectively.

2.3.4 Algorithm performance evaluation

LR models, RF and CIT models with their respective selected optimum hyper-parameters were evaluated for each of the explanatory variable sets (bands, VIs, bands plus VIs) based on overall classification accuracy (A), specificity (SP), and sensitivity (SN). These metrics are calculated based on the true-positive (TP), false-positive (FP), true-negative (TN), and false-negative (FN) rates. TP and TN represent the fundamental truth when the predicted class is infected or non-infected, respectively, and FP and FN illustrate the wrong classification for infected and non-infected occurrence. The metrics A, SP, and SN are then computed by the formulas:

$$A = \frac{TP + TN}{TP + TN + FP + FN} \qquad (1)$$

$$SP = \frac{TN}{TN + FP} \qquad (2)$$

$$SN = \frac{TP}{TP + FN} \qquad (3)$$

3. Results

3.1 Bivariate analysis of individual bands and VIs

The spectral reflectance curves for infected versus non-infected soybean condition presented large variability and overlap, displaying all three sites (Figure 2a). The curve behavior in the visible range appears traditional for chlorophyll absorption, and the percentage difference between non-infected and infected was less than 3%. The greatest discrepancies between infected and non-infected conditions were in the red edge (5.6%) and NIR (5%) spectral regions. Nonetheless, infected soybeans presented significantly greater reflectance in the regions of blue, green, red, and red-edge (Figure 2b).

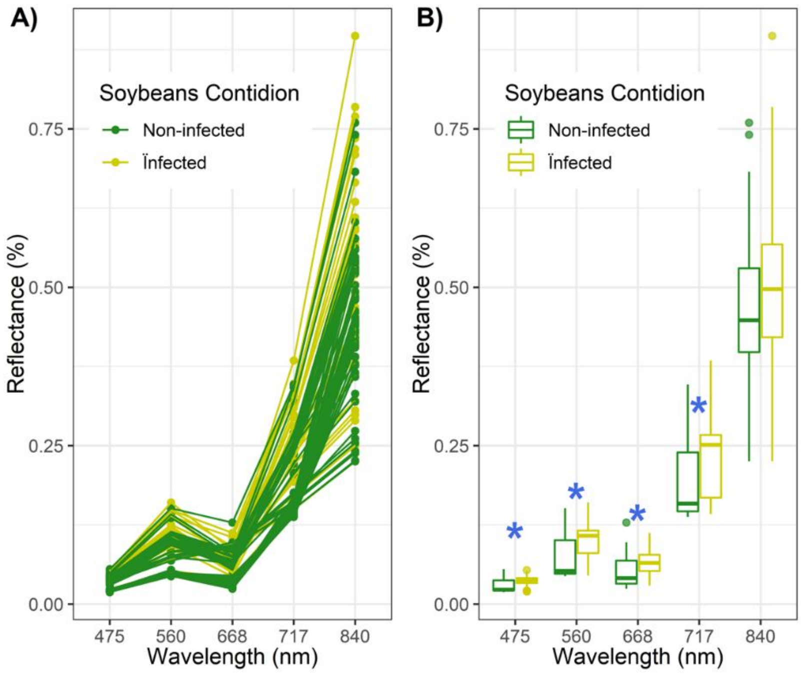

Figure 3. Spectral reflectance curves from non-infected (green) and infected (yellow) soybean plants with a) all replications (n=115) from all three sites, and b) summarized reflectance as boxplots obtained for each condition. On panel b, boxplots portray the 25^{th} (lower hinge), 50^{th} (solid black line), and 75^{th} (upper hinge) percentiles, largest value no further than 1.5×inter-quartile range (lower whisker), smallest value at most 1.5×inter-quartile range (upper whisker), and outlying observations (points). Different letters represent significant differences between conditions within a given band at alpha=0.05.

Over all tested VIs, only GNDVI, NDRE and SR were able to discriminate non-infected relative to infected soybean plants (Figure 3). In those cases, the VI values for the infected were significantly lower compared to non-infected condition. Those Indices are more sensitive to high-biomass conditions due to the replacement of the red band in NDVI by the green and red-edge bands in GNDVI and NDRE, respectively. The other indices were not able to statistically differentiate soybeans condition related to nematode infestation.

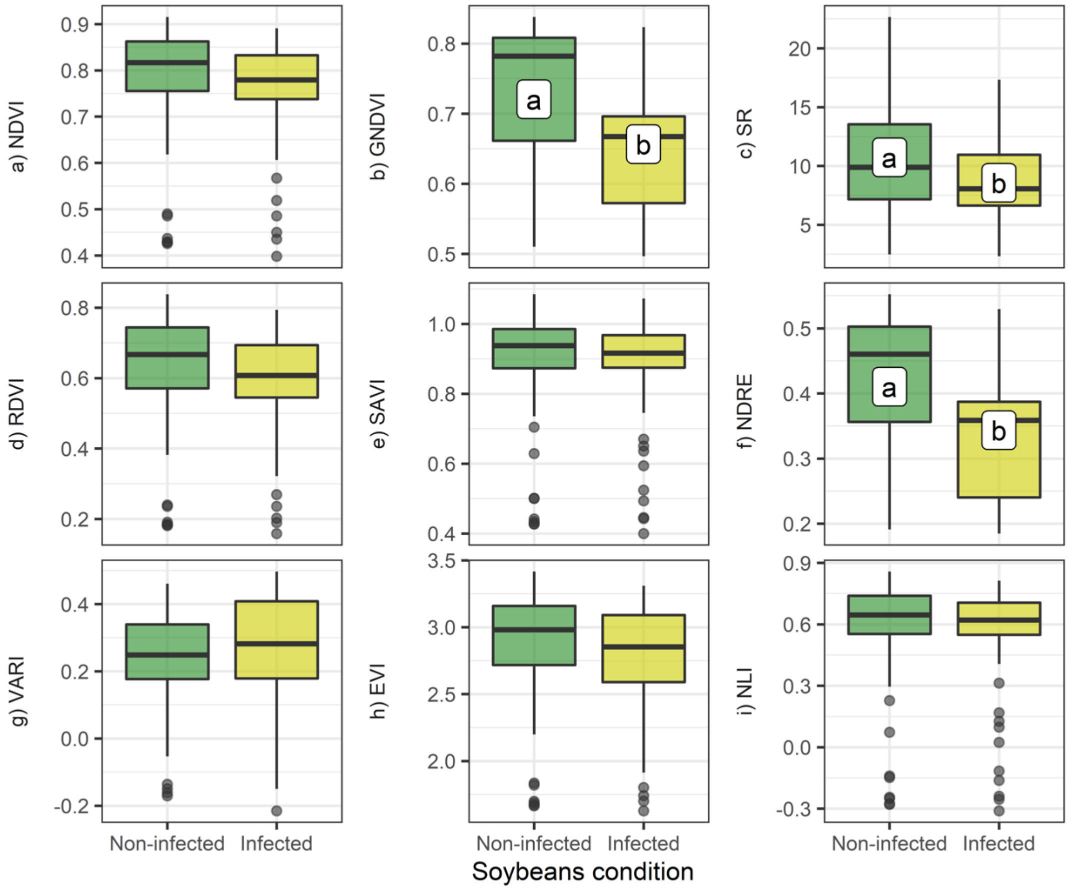

Figure 3. Boxplot representing the data distribution for VIs in different soybean conditions (non-infected in green, infected in yellow). Boxplots portray the 25th (lower hinge), 50th (solid black line), and 75th (upper hinge) percentiles, largest value no further than 1.5×inter-quartile range (lower whisker), smallest value at most 1.5×inter-quartile range (upper whisker), and outlying observations (points). Different letters represent significant differences between conditions at alpha=0.05.

3.2 Algorithms performance

Across all algorithms and input variable types, overall accuracy, sensitivity, and specificity ranged from 0.59-0.71, 0.60-0.72, and 0.55-0.78, respectively (Table 4). The greatest accuracy (0.71) was observed when bands were the main input variables and the algorithm used was CIT. Using bands plus VIs as the input variables did not improve the classification accuracy and sensitivity. The lowest accuracy, sensitivity and specificity were observed with the LR approach.

The difference in accuracy between the CIT and LR methods was 0.5, 0.8 and 0.12, respectively, when input variables included bands, VIs, and both bands and VIs, respectively, For sensitivity, CIT differed from LR by 0.01, -0.01, 0.10 and for specificity, by 0.10, 0.13, 0.06 for bands, VIs and bands plus VIs, respectively. The overall accuracy difference between CIT and RF was 0.01, 0.02 and 0.03, respectively and for sensitivity and specificity was 0.05, -0.02, 0.03 and 0, 0.11 and 0.02 for the input variables bands, VIs and bands plus VIs, respectively.

Table 4. Accuracy, sensitivity and specificity from the classification of soybean condition (infected vs. non-infected) from three algorithms (logistic regression - LR, random forest - RF, and conditional inference tree - CIT) and three data input (bands only, vegetation indices -VIs- only, bands plus VIs).

Metric	**Bands**			**VIs**			**Bands plus VIs**		
	LR	CIT	RF	LR	CIT	RF	LR	CIT	RF
Accuracy	0.66	0.71	0.70	0.59	0.67	0.65	0.59	0.71	0.68
Sensitivity	0.71	0.72	0.67	0.61	0.60	0.62	0.62	0.72	0.66
Specificity	0.61	0.71	0.71	0.55	0.78	0.67	0.55	0.71	0.69

3.2.1 Random Forest

For the RF algorithm, the most important variables as assessed by the mean decrease accuracy (MDA) were green and red-edge bands both when bands and bands plus VIs were used as input variables. When only bands were used as input variables, the MDA of green and red-edge bands was 9.4% and 6.5%, respectively, and the same bands were selected when the input type included VIs, with an MDA of 6.6% and 4.4%, respectively. When only VIs were used as the input variable, VARI and GNDVI were the most relevant variables, with an MDA of 5.4% and 3.5%, respectively.

3.2.2 Conditional inference tree

When using only bands as the data input type, the most important variables differentiating soybeans condition classes were green and NIR (Figure 4). Soybeans condition was classified as infected when green reflectance was greater than 0.105 (36.5% of the observations relative to all data), and non-infected when green reflectance was lower than 0.105 (63.5% of the observations relative to all data). Within the non-infected, when NIR reflectance was greater than 0.44 (41% of observations within the non-infected group), most of the observations within this node (90%) were non-infected soybean data. Lastly, when green reflectance was greater than 0.105 and NIR reflectance was lower than 0.44, soybean condition was classified as non-infected, although with less certainty due to both infected and non-infected classes existing in similar proportion at this node (41% and 59%, respectively). The same result was obtained when the input type was bands plus VIs, demonstrating a lack of improvement on discriminating soybean conditions. For VIs input variable type, soybean was considered infected when GNDVI was below 0.708, equivalent to 61% (n=70) from the total dataset (data not shown). This last model classified a greater proportion of the data as infected relative to the model using spectral bands only and full, spectral bands plus VI (same model as bands only; Figure 4).

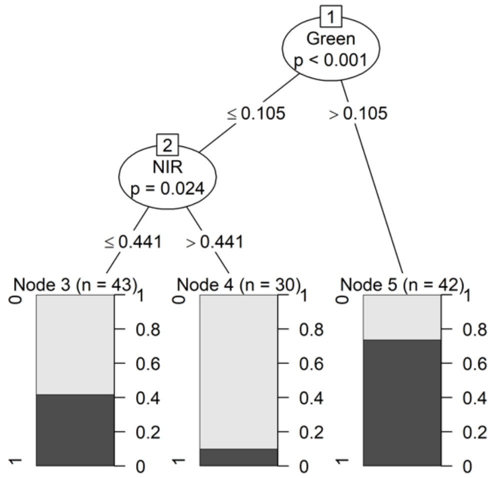

Figure 4. Conditional inference tree model using bands. Terminal node bars represent the proportion (right y-axis) of samples for each soybeans condition at that node. Left y-axis refers to plant condition classes with 0 as non-infected and 1 as infected.

4. Discussion

This study demonstrated a novel approach using remotely-sensed data and machine learning in field grown soybeans to differentiate between nematode-infected and non-infected areas, improving both labor- and time-intensity constraints relative to boots on the ground approach. This new approach has the potential to be implemented in site specific nematode management via the use of targeted pesticide treatments, utilization of tolerant soybean varieties, and for diversifying crop rotation in heavily infested fields. The application of remote sensing data to assess pest and diseases has been already reported in the literature documenting that the VIS-NIR reflectance is sensitive to nematode occurrence in several crops (Hillnhütter et al., 2011; George Deroco Martins et al., 2017b; Nutter et al., 2002).

From a physiological perspective, several studies reported the reduction of chlorophyll content, plant growth, photosynthetic rate and nutrient concentrations due to nematodes infection, with the aboveground canopy symptoms including chlorosis and severe stunted plants (Behmann et al., 2014; Carneiro et al., 2002; Haseeb et al.;

Nutter et al., 2002), symptoms that can be differentiated via utilization of remote sensing aerial imagery, spectral bands. Our results are in line with Sims & Gamon (2002), reporting that chlorophyll content and other plant pigments are strongly correlated with absorption in the visible spectral regions. Visible light (400-700 nm) is absorbed by the plant mostly for photosynthesis and lower concentration of pigments like chlorophyll a and b, carotenoids and anthocyanins can reduce photosynthetic rate (Chappelle et al., 1984). The red-edge band (centered in the 750 nm) has been identified as more sensitive to changes in chlorophyll content compared to blue, green and red bands under high-biomass conditions (Curran et al., 1990). Furthermore, leaf morphology and structure differences are retrieved by bands in the NIR region (800-1100 nm) (Behmann et al., 2014).

Considering the scientific literature on spectral reflectance information to distinguish nematode infected areas, an important outcome of this research was related to the capability of green and NIR bands on differentiating infected and non-infected soybeans condition. These findings show that multispectral data in the form of bands can be used as a rapid, non-destructive and cost-effective method for the detection of nematodes. Martins et al. (2017) reported that spectral curves obtained from coffee canopy severely infected with root-knot nematodes presented lower reflectance between 750 and 900 nm wavelength compared to non-infected plants. Our results agree with these authors, who stated that red-edge and NIR regions were sensitive to this type of plant stress. Therefore, presented stress characteristics caused by nematodes infection can be detected by leaf reflectance.

In addition to the investigation of the individual spectral bands, the VIs can estimate the relationship between crop physiological parameters and reflectance, with

NDVI considered as the most popular VI used in agricultural related topics (Bajwa et al., 2017; Hillnhütter et al., 2011; George Deroco; Martins & Galo, 2015). However, red reflectance can saturate rapidly even under low chlorophyll content (Anatoly A. Gitelson et al., 2001). In order to increase the sensitivity of NDVI to chlorophyll content, researchers developed indices including the green (Anatoly A. Gitelson et al., 1996) and red-edge bands (Barnes et al., 2000). Although NDVI was not an adequate VI to differentiate soybean condition, GNDVI and NDRE were capable to discriminate and isolate these conditions. The SR is used to indicate the relative biomass presented in the image (Baret & Guyot, 1991; Jordan, 1969), and in our study it was able to differentiate the lower-biomass infected plants. In the CIT models, GNDVI was selected as the only splitting variable and in the RF models it was identified as an important variable via MDA.

The statistical techniques demonstrated that similar variables were selected to differentiate the infected from non-infected condition regardless of the tested algorithm. The CIT model used green and NIR variable to split the tree. In contrast, NIR was not significant in explaining soybeans condition in the bivariate analysis. This demonstrates how the CIT model was able to expand the bivariate relationship in finding a significant NIR effect within a sub-population of the data. Also, the RF model had green and red-edge as the most important variables, agreeing with the bivariate analyses. Our results demonstrated that LR, RF and CIT were able to distinguish between soybean infected with nematodes from non-infected plants with reasonable accuracy, with RF and CIT presenting greater classification accuracy. Others have found LR to be successful in detecting rhizomania in sugar beets (Steddom et al., n.d.) and apple scab (Delalieux et al., 2007). Although not evaluated, based in the current results the outcomes presented by Delalieux et al. (2007) could have been improved if CIT and RF models had been tested in detecting biotic stress in apple trees. The CIT with spectral bands as input variables was the best model in regard to accuracy and sensitivity, but not specificity relative when only VIs were used as data input. Although the input variable VIs had the worst accuracy for all models, it had the best specificity when CIT was applied.

Future studies should evaluate both sources of algorithm and input data following a similar step process as presented in the current paper. Limitations of this study are related to the limited number of fields and the constraint on the single timing for the sUAV flight for data collection. For upscaling the study to a product, a more robust data set including a larger pool of abiotic (e.g. soil type, weather regime) and biotic (e.g. soybeans cultivars, nematode incidence and severity levels) variability is needed to retrain the models. Also, future studies should obtain spectral and ground-truthing information at various crop stages to determine the best moment to separate nematode-infected from non-infected soybean areas. Sensors with more spectral bands (hyperspectral sensor) available can also be useful to relate the occurrence in specific wavelength not observed in this study.

5. Conclusion

This research provided a useful nematode identification workflow for distinguishing soybean plants infected with nematodes using high spatial resolution aerial sensing data combined with machine learning. The findings can support further development of more precise soil-borne parasites identification models.

Statistical non-parametric techniques were able to identify spectral wavelengths differentiating between soybean plants infected or not with nematodes. Green and NIR spectral bands presented greater values of accuracy for model detection to separate infected versus non-infected plants, even when compared with different tested VIs. Spectral bands as an input applied to CIT were able to identify nematode-infected plants with reasonable classification accuracy and better performance than RF and LR methods.

Acknowledgments

Contribution no. 20-789 from the Kansas Agricultural Experiment Station.

Coordination for the Improvement of Higher Education Personnel Brazil (CAPES) for the scholarship granted - Financing code 001.

References

[1] Askary T.H. and Martinelli P.P., *Biocontrol Agents of Phytonematodes*. 2015.
[2] T. W. Allen *et al.*, “Soybean yield loss estimates due to diseases in the United States and Ontario, Canada, from 2010 to 2014,” *Plant Heal. Prog.*, vol. 18, no. 1, pp. 19–27, Apr. 2017, doi: 10.1094/PHP-RS-16-0066.
[3] L. Zhao, P. Jiang, L. M. Humble, and J. Sun, “Within-tree distribution and attractant sampling of propagative pinewood nematode, Bursaphelenchus xylophilus: An early diagnosis approach,” *For. Ecol. Manage.*, vol. 258, no. 9, pp. 1932–1937, Oct. 2009, doi: 10.1016/j.foreco.2009.07.040.
[4] D. Tihohod, “Nematologia agrícola aplicada,” 1993.
[5] D. G. Blevins, V. H. Dropkin, and V. D. Luedders, “Macronutrient uptake, translocation, and tissue concentration of soybeans infested with the soybean cyst nematode and elemental composition of cysts isolated from roots1,” *J. Plant Nutr.*, vol. 18, no. 3, pp. 579–591, 1995, doi: 10.1080/01904169509364924.
[6] C. Hillnhütter, A. K. Mahlein, R. A. Sikora, and E. C. Oerke, “Remote sensing to detect plant stress induced by Heterodera schachtii and Rhizoctonia solani in sugar beet fields,” *F. Crop. Res.*, vol. 122, no. 1, pp. 70–77, Apr. 2011, doi: 10.1016/j.fcr.2011.02.007.
[7] G. D. Martins, M. D. L. B. T. Galo, and B. S. Vieira, “Detecting and Mapping Root-Knot Nematode Infection in Coffee Crop Using Remote Sensing Measurements,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 10, no. 12, pp. 5395–5403, Dec. 2017, doi: 10.1109/JSTARS.2017.2737618.
[8] L. Wiesel, T. J. Daniell, D. King, and R. Neilson, “Determination of the optimal soil sample size to accurately characterise nematode communities in soil,” *Soil Biol. Biochem.*, vol. 80, pp. 89–91, Jan. 2015, doi: 10.1016/j.soilbio.2014.09.026.
[9] F. Garcia-Ruiz, S. Sankaran, J. M. Maja, W. S. Lee, J. Rasmussen, and R. Ehsani, “Comparison of two aerial imaging platforms for identification of Huanglongbing-infected citrus trees,” *Comput. Electron. Agric.*, vol. 91, pp. 106–115, Feb. 2013, doi: 10.1016/j.compag.2012.12.002.

[10] S. G. Bajwa, J. C. Rupe, and J. Mason, “Soybean disease monitoring with leaf reflectance,” *Remote Sens.*, vol. 9, no. 2, 2017, doi: 10.3390/rs9020127.
[11] S. Delalieux, J. van Aardt, W. Keulemans, E. Schrevens, and P. Coppin, “Detection of biotic stress (Venturia inaequalis) in apple trees using hyperspectral data: Non-parametric statistical approaches and physiological implications,” *Eur. J. Agron.*, vol. 27, no. 1, pp. 130–143, Jul. 2007, doi: 10.1016/j.eja.2007.02.005.
[12] K. Steddom, G. Heidel, D. Jones, and C. M. Rush, “Remote Detection of Rhizomania in Sugar Beets,” 2003.
[13] S. Adelabu, O. Mutanga, and E. Adam, “Evaluating the impact of red-edge band from Rapideye image for classifying insect defoliation levels,” *ISPRS J. Photogramm. Remote Sens.*, vol. 95, pp. 34–41, 2014, doi: 10.1016/j.isprsjprs.2014.05.013.
[14] V. F. Rodriguez-Galiano, B. Ghimire, J. Rogan, M. Chica-Olmo, and J. P. Rigol-Sanchez, “An assessment of the effectiveness of a random forest classifier for land-cover classification,” *ISPRS J. Photogramm. Remote Sens.*, vol. 67, no. 1, pp. 93–104, Jan. 2012, doi: 10.1016/j.isprsjprs.2011.11.002.
[15] T. Hothorn, K. Hornik, and A. Zeileis, “Journal of Computational and Graphical Statistics Unbiased Recursive Partitioning: A Conditional Inference Framework Unbiased Recursive Partitioning: A Conditional Inference Framework,” 2012, doi: 10.1198/106186006X133933.
[16] Q. Yuan *et al.*, “Deep learning in environmental remote sensing: Achievements and challenges,” *Remote Sens. Environ.*, vol. 241, p. 111716, May 2020, doi: 10.1016/j.rse.2020.111716.
[17] N. Susič *et al.*, “Discrimination between abiotic and biotic drought stress in tomatoes using hyperspectral imaging,” 2018, doi: 10.1016/j.snb.2018.06.121.
[18] W. R. Jenkins, “A rapid centrifugal-flotation technique for separating nematodes from soil.,” *Plant Dis. Report.*, vol. 48, no. 9, 1964, Accessed: Mar. 23, 2019. [Online]. Available: https://www.cabdirect.org/cabdirect/abstract/19650801105.
[19] W. A. Coolen and C. J. D’herde, “A method for the quantitative extraction of nematodes from plant tissue.,” *A method Quant. Extr. nematodes from plant tissue.*, 1972, Accessed: Aug. 01, 2019. [Online]. Available: https://www.cabdirect.org/cabdirect/abstract/19722001202.
[20] Tihohod, D. “*Nematologia agrícola aplicada*”. Funep, 1993.
[21] J. F. Southey, *Laboratory methods for work with plant and soil nematodes*. Technical bulletin 2. London: Her Majesty’s Stationery Office, 1986.
[22] R. W. H. Rouse, J. A. W. Haas, and D. W. Deering, “Monitoring vegetation systems in the great plains with erts,” Washington D.C., 1974. Accessed: Jul. 16, 2020. [Online]. Available: https://ntrs.nasa.gov/search.jsp?R=19740022614.
[23] A. A. Gitelson, Y. J. Kaufman, and M. N. Merzlyak, “Use of a green channel in remote sensing of global vegetation from EOS- MODIS,” *Remote Sens. Environ.*, vol. 58, no. 3, pp. 289–298, Dec. 1996, doi: 10.1016/S0034-4257(96)00072-7.
[24] C. F. Jordan, “Derivation of Leaf-Area Index from Quality of Light on the Forest Floor,” *Ecology*, vol. 50, no. 4, pp. 663–666, Jul. 1969, doi: 10.2307/1936256.
[25] J. Roujean and F. Breon, “Estimating PAR absorbed by vegetation from bidirectional reflectance measurements,” *Remote Sens. Environ.*, vol. 51, no. 3, pp. 375–384, 1995, doi: https://doi.org/10.1016/0034-4257(94)00114-3.
[26] A. R. Huete, “A soil-adjusted vegetation index (SAVI),” *Remote Sens. Environ.*,

vol. 25, no. 3, pp. 295–309, Aug. 1988, doi: 10.1016/0034-4257(88)90106-X.
[27] E. M. Barnes *et al.*, “Coincident detection of crop water stress, nitrogen status and canopy density using ground-based multispectral data.” 2000.
[28] H. Q. Liu and A. Huete, “A feedback based modification of the NDVI to minimize canopy background and atmospheric noise,” *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 2, pp. 457–465, Jun. 1995, doi: 10.1109/tgrs.1995.8746027.
[29] A. A. Gitelson *et al.*, “International Journal of Remote Sensing Vegetation and soil lines in visible spectral space: A concept and technique for remote estimation of vegetation fraction Vegetation and soil lines in visible spectral space: a concept and technique for remote estimation of vegetation fraction,” *int. j. Remote Sens.*, vol. 23, no. 13, pp. 2537–2562, Jul. 2002, doi: 10.1080/01431160110107806.
[30] P. Gong, R. Pu, G. S. Biging, and M. R. Larrieu, “Estimation of forest leaf area index using vegetation indices derived from Hyperion hyperspectral data,” *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 6 PART I, pp. 1355–1362, Jun. 2003, doi: 10.1109/TGRS.2003.812910.
[31] Koenning, S. R. 2007. Rotations for management of root-knot nematode in cotton. http://www.cottoninc.com/statesupport-program/State-Support-Program Projects/detail.asp? SelectedYear=2007&projectID=123. North Carolina State University/EUA.
[32] O. Dickerson, J. Blake, and S. Lewis, “Nematode guidelines for South Carolina. Clemson University Cooperative Extension Publication No. EC703. Clemson University, Clemson, SC,” 2000.
[33] A. Liaw and M. Wiener, “Classification and Regression by RandomForest,” 2002. Accessed: May 26, 2020. [Online]. Available: https://www.researchgate.net/publication/228451484.
[34] F. W. Nutter *et al.*, “Use of Remote Sensing to Detect Soybean Cyst Nematode-Induced Plant Stress 1,” 2002.
[35] R. G. Carneiro, P. Mazzafera, L. C. C. B. Ferraz, T. Muraoka, and P. C. O. Trivelin, “Uptake and translocation of nitrogen, phosphorus and calcium in soybean infected with *Meloidogyne incognita* and *M. javanica*,” *Fitopatol. Bras.*, vol. 27, no. 2, pp. 141–150, Apr. 2002, doi: 10.1590/s0100-41582002000200004.
[36] A. Haseeb, N. K. Srivastava, and R. Pandey, “The influence of *meloidogyne incognita* on growth, physiology, nutrient concentration and alkaloid yield of hyoscy amus niger,” *Nematol. Mediterr.*, pp. 127–129, Dec. 1990.
[37] J. Behmann, J. Steinrücken, and L. Plümer, “Detection of early plant stress responses in hyperspectral images,” *ISPRS J. Photogramm. Remote Sens.*, vol. 93, pp. 98–111, Jul. 2014, doi: 10.1016/j.isprsjprs.2014.03.016.
[38] D. A. Sims and J. A. Gamon, “Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages,” *Remote Sens. Environ.*, vol. 81, no. 2–3, pp. 337–354, 2002, doi: 10.1016/S0034-4257(02)00010-X.
[39] E. W. Chappelle, J. E. McMurtrey, F. M. Wood, and W. W. Newcomb, “Laser-induced fluorescence of green plants 2: LIF caused by nutrient deficiencies in corn,” *Appl. Opt.*, vol. 23, no. 1, p. 139, Jan. 1984, doi: 10.1364/ao.23.000139.
[40] P. J. Curran, J. L. Dungan, and H. L. Gholz, “Exploring the relationship between reflectance red edge and chlorophyll content in slash pine,” *Tree Physiol.*, vol. 7,

no. 1-2-3–4, pp. 33–48, Dec. 1990, doi: 10.1093/TREEPHYS/7.1-2-3-4.33.
[41] G. D. Martins, M. De, L. Bueno, and T. Galo, "Caracterização espectral da cana-de-açúcar infectada por nematoides e *Migdolus fryanus* por espectrorradiometria de campo In situ spectral characterization of sugar cane infected by nematodes and *Migdolus fryanus*," *Bol. Ciênc. Geod., sec. Artig.*, no. 4, pp. 783–796, 2015, doi: 10.1590/S1982-21702015000400046.
[42] A. A. Gitelson, M. N. Merzlyak, and O. B. Chivkunova, "Optical Properties and Nondestructive Estimation of Anthocyanin Content in Plant Leaves," *Photochemistry and photobiology*, vol. 74, no. 1, pp. 38-45, 2001.
[43] F. Baret and G. Guyot, "Potentials and limits of vegetation indices for LAI and APAR assessment," Remote Sens. Environ., vol. 35, no. 2–3, pp. 161–173, Feb. 1991, doi: 10.1016/0034-4257(91)90009-U.

Chapter 4 – Final considerations

Remote sensing techniques has increased significantly in agriculture in recent years and allows farmers to assess different crop stresses. Therefore, this work was important to evaluate the potential of remote sensing techniques in large soybean areas to detect the occurrence of soil parasites such as nematodes. These, in combination with non-parametric statistical techniques such as machine learning, were efficient in identifying the stress in question.

Regardless of the statistical approach adopted, it was possible to identify plants infected with nematodes using multispectral sensors and different statistical methods. To scale the techniques for a product, it is necessary to use more fields and further study at the time of image acquisition.

Future studies should consider phenological development. Crop structural and morphological growth changes dramatically from one phenological stage to another (Wang et al., 2019). Literature studies indicate that crop phenological growth can be inferred using remote sensing. It has the potential to provide spatial-temporal information over the entire crop growth respective to the size field (Canisius et al., 2018)

Sampling dates should be evaluated in others phenological period as well. Nematode population fluctuation can be inferred by plant growth and changes in the seasonal soil water content (Gomes et al., 2003). That includes sampling before any treatment or management decision is made, including planting. Also, it is important to sample nematodes when they are active and in high populations, which depends on the climatic varition.

Accurate information about the plant condition regard to nutrient necessity, water regime and surveillance diseases and pests are useful applications in order to optimize crop yield. For spectral data, consider the use of sensors with a greater number of spectral bands (hyperspectral sensor), for a deep study of which wavelength is related to the occurrence of nematodes. Wavelengths are useful for applications to optimize crop yield development and can be an important factor to analyze crop growth

Knowledge of this factor can assist in making agricultural decisions such as accurate crop management. Applications that can optimize crop yield. Others robust information including biotic and abiotic data could be used to train the models. The evaluation of different algorithms and input data is a workflow suitable for analysis, as was done in the present work.

References

Canisius, F., Shang, J., Liu, J., Huang, X., Ma, B., Jiao, X., Geng, X., Kovacs, J. M., & Walters, D. (2018). Tracking crop phenological development using multi-temporal polarimetric Radarsat-2 data. *Remote Sensing of Environment*, 210, 508–518. https://doi.org/10.1016/j.rse.2017.07.031

Gomes, G. S., Huang, S. P., & Cares, J. E. (2003). Nematode community, trophic structure and population fluctuation in soybean fields. *Fitopatologia Brasileira*, 28(3), 258–266. https://doi.org/10.1590/s0100-41582003000300006

Wang, H., Magagi, R., Goïta, K., Trudel, M., McNairn, H., & Powers, J. (2019). Crop phenology retrieval via polarimetric SAR decomposition and Random Forest algorithm. *Remote Sensing of Environment*, 231, 111234. https://doi.org/10.1016/j.rse.2019.111234