Electronic Thesis and Dissertation Repository

7-23-2020 2:00 PM

Intra-field Canopy Nitrogen Retrieval from Unmanned Aerial Vehicle Imagery for Wheat and Corn Crops in Ontario, Canada

Hwang Lee, *The University of Western Ontario*

Supervisor: Wang, Jinfei, *The University of Western Ontario*
Co-Supervisor: Leblon, Brigitte, *University of New Brunswick*
A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Geography
© Hwang Lee 2020

Recommended Citation

Title:

Intra-field Canopy Nitrogen Retrieval from Unmanned Aerial Vehicle Imagery for Wheat and Corn Crops in Ontario, Canada

Abstract

The optimization of crop nitrogen fertilization to accurately predict and match the nitrogen (N) supply to the crop N demand is the subject of intense research due to the environmental and economic impact of N fertilization. Excess N could seep into the water supplies around the field and cause unnecessary spending by farmers. Understanding the detailed spatial information about a crop status is known as a farming management technique called precision agriculture, which allows farmers to maximize their yield and profit while reducing the inputs of fertilizers, pesticides, water, and insecticides.

The goal of this study is to document and test the applicability and feasibility of using Unmanned Aerial Vehicle (UAV) to predict nitrogen weight of wheat and corn fields in south-west Ontario. This is investigated using various statistical modelling techniques to achieve the best accuracy. Machine learning techniques such as Random Forests and Support Vector Regression are used, which provide more robust models than traditional linear regression models. The results demonstrate that most spectral indices have a non-linear relationship with canopy nitrogen weight and show high degree of multicollinearity among the variables. In this thesis, the final nitrogen prediction maps of wheat and corn fields using UAV images and the derived models are provided.

Keywords

Summary for Lay Audience

The analogy I like to give for the concept of precision agriculture is a person's desire for a cup of coffee from a coffee machine. Imagine trying to invent a coffee machine that can predict the quantities of coffee that the customers want. The coffee machine should be able to supply you with the right amount of coffee depending on the size of the cup, how tired you are, what time of the day it is, etc. All these factors contribute to how much coffee you need. For example, if you were very tired one day and inserted a large cup, you would be quite disappointed if the coffee machine supplied you with a little amount of coffee. This is how the phenomenon of nitrogen supply works with crops. The goal of precision agriculture is to accurately supply the agricultural crop's site-specific need, depending on various factors surrounding the crop (e.g. soil, precipitation, temperature, etc). If crops are deficient in nitrogen, their growth cycle is likely to be stunted, reducing its yield potential. On the contrary, if crops are supplied too much nitrogen, the excess supply can seep into the water supply, causing a negative environmental impact and unnecessarily uses up the farmer's nitrogen resources. Because crops cannot communicate their needs of nitrogen to us, researchers have performed extensive research using remote sensing techniques on various types of crops by estimating how much nitrogen they currently have. If we know how much nitrogen the crops have, we can add or reduce the nitrogen application for a particular area based on the guideline that the farmer has. This thesis dives into the statistical application of drone imagery and regression modelling to predict the quantification of nitrogen status within wheat and corn fields. Ultimately, when we predict the values of the nitrogen on a map, we are then able to supply it to the farmer for their next nitrogen fertilization application.

Co-Authorship Statement

This thesis was prepared in an integrated-article layout. The work in this thesis was carried out by the author and closely guided and overlooked under the supervision of Dr. Jinfei Wang and Dr. Brigitte Leblon. Both Drs. Wang and Leblon provided ideas for methodology and provided generous comments, edits and revisions of the articles. Both Drs. Wang and Leblon are co-authors of the two articles below:

Lee, H., Wang J., & Leblon, B. (2020). Intra-field canopy nitrogen retrieval from unmanned aerial vehicle imagery for wheat and corn fields. *Canadian Journal of Remote Sensing*, 1-19, doi: 10.1080/07038992.2020.1788384

Lee, H., Wang J., & Leblon, B. (2020). Using Linear Regression, Random Forests, and Support Vector Machine with Unmanned Aerial Vehicle Multispectral Images to Predict Canopy Nitrogen Weight in Corn. *Remote Sensing*, *12*(13), 2071, doi: https://doi.org/10.3390/rs12132071

Acknowledgments

First and foremost, I would like to thank my supervisors Dr. Jinfei Wang (Professor, University of Western Ontario and Dr. Brigitte Leblon (Professor, University of New Brunswick) for their gracious support and guidance throughout the past two years of my Masters. I am extremely grateful for the knowledge I have learned throughout the years by them. Dr Jinfei Wang has allowed me to continue my love and passion for geography and remote sensing after my undergraduate studies with her. Overall, I have developed and improved a tremendous amount of experience under their supervision and have ultimately set me up for the next chapter in my life.

I would like to thank the members of A and L Canada Laboratory and GITA lab for their help on retrieving and processing the data required for my thesis. This labour-intensive data collection in heat warning temperatures could have not been done alone.

I would like to thank the NSERC CRD project collaborated by Dr. Leblon, Dr. Wang and Dr. Sabarinathan for funding this thesis.

I would like to thank my girlfriend Sierra Steele for her love and support during my thesis project. Thank you for understanding the tremendous time commitment I have poured onto this thesis and even volunteered to cook meals when you knew I was stressed with work. I am very excited for the next chapter of our lives.

Last but not least, I would like to thank my family and friends for their support during my academic career. Thank you to my parents for allowing me to drive to Hamilton whenever I needed a nice home-cooked meal and provided countless quantity of frozen meat to take back to London. I would like to thank my friends Jonathan Lee, Alan Li, Mickey Wang and Jeremy Wong for their positive feedback and support on my projects through our Facebook group chats and well-needed gaming nights.

Table of Contents

List of Tables

List of Figures

List of Appendices

Glossary

ASD - Analytic Spectral Device; a handheld equipment used to retrieve spectral components of objects, measuring over visible and shortwave infrared wavelength region (325-1000nm).

Orthomosaic map- a detailed and accurate photo representation of an area, generated by many photos stitched together.

Pix4D - a photogrammetry software that is used to create orthomosaic images and 3D point cloud using images.

R and R Studio - a programming language and free to use open source integrated development environment (IDE) used for data cleaning, complex data analysis and data visualization.

Random Forests (RF)- an ensemble learning method using decision trees for classification or regression tasks. It builds trees for decision learning using random number of features and averages the output value of multiple trees built.

Regression- a statistical process for estimating the quantitative value of the relationship between the response and the explanatory variable(s) in interest.

Structure from Motion (SfM) – a technique of estimating the 3D structure of a scene from a set of 2D images.

Support Vector Machines (Regression) (SVM or SVR)- a supervised learning method used to create decision boundaries known as a hyperplane that help classify or predict data points. The dimension of a hyperplane depends on the number of input features.

Variation- a measurement of how far a set of numbers are spread out from their average value.

Chapter 1

1 Introduction

1.1 Background

1.1.1 Precision Agriculture

Agricultural resources are important to society because they are a renewable and dynamic natural resource. In Canada, agriculture and the agri-food industry employ approximately 2.3 million Canadians and contributes over $110 billion annually towards Canada's GDP, making Canada the 5th largest agriculture exporter in the world (Agriculture and Agri-Food Canada, 2018). In order for farmers to maximize their crop yield and profitability, they have been involved in an agricultural management technique called precision agriculture, which ensures the precise input of water, fertilizers, herbicides, and insecticides to the crops (Barbanti et al. 2018; Xie et al. 2018; Liu et al. 2016). A key component of this term is to use information technology that is retrieved from a variety of devices, such as: global positioning system (GPS), sensors, robotics, unmanned aerial vehicles, and autonomous vehicles (Schmaltz, 2017). One major aspect of precision agriculture is nitrogen fertilization management. Specifically, nitrogen is crucial for crops because a presence of nitrogen deficiency in the leaf can most likely affect the chlorophyll production. Furthermore, crop yield is directly affected by the plant nitrogen status (Loel et al. 2014; Munoz-Huerta et. al., 2013). Nitrogen is part of the chlorophyll molecule, (The chlorophyll molecule is partially composed of nitrogen) to help the plant obtain energy through the sun's rays. Plants with shortage of nitrogen present will also have a lower chlorophyll content, resulting in a non-optimal photosynthesis, greatly reducing the plant's growth. (Milford et al. 1985; Clevers & Kooistra, 2011). The deficiency of nitrogen could also be affected by the topography and drainage of the field, as the spatial variation of the topography and drainage could also affect the plant's nitrogen content. All these factors can ultimately reduce the plant's ability to grow adequately.

1.1.2   Remote Sensing Techniques

A prerequisite of precision agriculture is the extensive knowledge of the within-field information about the crop's nutrient status (Maes & Steppe, 2019; Zhang & Kovacs, 2012). To spatially examine the information about a crop, remote sensing technology such as unmanned aerial vehicle (UAV), satellite imagery, airborne imagery, and tractor-based sensors can determine the health of a certain crop, identify diseases, or potentially predict potential soil conditions. Different types of sensors: (i) red-green-blue (RGB), (ii) multispectral, (iii) hyperspectral, and (iv) thermal exist in remote sensing that contribute to a certain application in precision agriculture (Table 1-1).

Table 1-1. Overview of Applications and Suitability of Different Sensors. Adapted from Maes & Steppe, 2019

Application	Type of sensor/camera			
	RGB	Multispectral	Hyperspectral	Thermal
Drought Stress			S	HS
Pathogen detection	HS	HS	HS	S
Weed detection	HS	HS	HS	
Nutrient status	S	HS	HS	S
Growth vigour (growth stage, canopy height and biomass)	HS	HS		
Yield prediction	S	HS		

HS: highly suited; S: suited.

RGB cameras are relatively cheaper than the other sensors stated above, and have a high spatial resolution, although they are limited by its poor spectral resolution. RGB cameras

are generally used to generate digital elevation models (DEMs) and crop height maps. Multispectral cameras consist of a set of sensors with different lenses, with each sensor classifying a small region in the electromagnetic spectrum (ES) (such as: green region, red region, etc.). In contrast, hyperspectral sensors cover the full spectrum (400-2500 nm spectral region in the ES). Hyperspectral sensors also have the highest potential to quantify nutrient status due to the higher number of wavelengths. However, the trade-off is that hyperspectral sensors are very expensive along with a large volume of data, thus the data retrieved are typically too complex for technical interpretations (Furbank & Tester, 2011; White et al. 2012). Therefore, multispectral sensors have been favoured over hyperspectral sensors in practical applications (Prey & Schmidhalter, 2019). Lastly, thermal sensors are most often low-resolution cameras and consists of one band measured in the longwave infrared region (7000-12000nm). These cameras are typically used to extract canopy temperature for drought detection (Maes & Steppe, 2019).

Satellite and aerial images are currently used to monitor crop growth, crop stress, and predict yield. However, these images are often limited by weather conditions causing fewer cloud-free images over time and provides coarse spatial resolutions compared to UAV-based images (Zhang & Kovacs, 2012). Most UAV nitrogen studies have also built on the experience of tractor-based nitrogen sensors, due to their limitations on estimating the correct nitrogen status (Maes & Steppe, 2019) and the tractor's ability to drive on agricultural sites with difficult soil conditions (Gynp et al, 2016). Regardless of the image acquisition method, vegetation indices (VI) have been used extensively to monitor crop information from spectral sensor information. VIs are generated by using various spectral bands to create different equations to measure several properties of vegetation (Gutierrez-Rodriguez et al. 2005). The most simple and common VI is commonly known as the normalized difference vegetation index (NDVI) (Rousse, 1974).

$$NDVI = (NIR - RED) / (NIR + RED) \qquad [1]$$

NIR = near infrared region; RED = red region

The values of NDVI range between -1 to +1; a value closer to +1 indicates a presence of healthy green vegetation (Jones & Vaughan, 2010). These values are of great use since the areas of healthy green vegetation are associated with a sufficient amount of nitrogen. Due to nitrogen's direct relation to chlorophyll, healthy vegetation is shown in specific VI results considering the chlorophyll's activity present in the blue, red, green, and near infrared region (Inman, Khosla & Mayfield, 2005). However, a disadvantage to the popular NDVI is that its values begin to saturate with the response variable (e.g. nitrogen, leaf area index) once the canopy has become dense (Xie et al. 2018; Lee et al. 2020). Therefore, many other types of VIs have been developed in order to mitigate the saturation. For example, ratio vegetation index (RVI) has been used to estimate nitrogen status while simultaneously being insensitive to growth stages and crop type (Jordan, 1969; Muñoz-Huerta et al. 2013). Many researchers have also improved previously generated VIs to retrieve certain crop parameters. For example, modified triangular vegetation index 2 (MTVI2) is a reworked version of the triangular vegetation index (TVI), which was found to be most useful to estimate nitrogen content using multispectral images (Haboudance et al. 2004; Broge & Leblanc, 2001; Bagheri et al. 2013).

1.1.3   UAV Applications in Precision Agriculture

The development and application of UAV imagery has greatly expanded in the past decade. The application of UAVs is promising because image acquisition with UAV can be deployed both quickly and repeatedly. The operation of UAVs are generally at a low cost and can be used with greater flexibility than spaceborne and airborne platforms (Maresma et al. 2016; Raparelli & Bajocco, 2019; Zha et al. 2020). UAV images provide high spatial and temporal resolution for within-field crop monitoring. The flexibility of UAV data collection and processing allows the user to retrieve information and produce results regarding a particular field in real time. Flexible revisit times are necessary due to a quick response to unfavourable crop or field conditions (Zhang & Kovavs, 2012), whereas the user does not have the ability to control the revisit time of a satellite imagery.

A single UAV flight over a crop field can provide hundreds of images depending on the size of the field. A technique called Structure-from-Motion (SfM) is used to combine the

images into a single mosaic image. SfM is a method used to overlap multiple images as input and extracts features and 3D point clouds and generates image mosaics (Westoby et al. 2012). Fortunately, there are software programs like Pix4D (Pix4D SA, Lausanne, Switzerland) that utilize this technique in order to extract final products for the user. The outputs of the UAV images are produced at centimetre level resolutions. This fine level resolution also allows for differentiation of individual crops on the imagery.

1.2   Research Questions

Statistical approaches based on VIs are very popular in remote sensing due to their simplicity, robustness, and accuracy in retrieving targeted crop parameters (Jay et al. 2016). Certain VIs can be used to estimate structural crop properties, such as LAI (Darvishzadeh et al., 2011; Haboudane et al., 2004), green fraction (Comar et al. 2012), or biochemical properties such as leaf chlorophyll content (Zarco-Tejada et al. 2004) and leaf water content (Colombo et al. 2008). Due to the direct correlation between nitrogen content and chlorophyll, spectral VIs are able to quantify the nitrogen status of crops (Jay et al. 2016). Therefore, with the use of VIs to quantify nitrogen status of wheat and corn, the research questions of this thesis are:

(1) Is the relationship between the crop nitrogen weight and the spectral variable(s) linear? Can predictions of nitrogen levels be made using a parametric linear regression?

(2) Can the predictions of nitrogen levels be made using non-parametric machine learning models? Which machine learning model provides the best accuracy on predicting canopy nitrogen weight using spectral variables?

(3) By using the generated models, can the UAV prediction maps provide a good clear separation of nitrogen values in the map? What spatial resolution is fine enough to see the separation of nitrogen values throughout the map while considering the processing time?

1.3 Research Objectives

The objective of this thesis is to evaluate different statistical modelling techniques to predict nitrogen values in wheat and corn crops using UAV-based images. This thesis is an empirical driven research project where the crop parameter (nitrogen) is statistically related to the spectral variable(s) of interest. A multispectral sensor is attached to a UAV to retrieve reflectance values from various fields in south west Ontario. Therefore, the research objectives of this thesis to answer the research questions listed above are:

(1) Generate parametric regression models to predict crop nitrogen weight within wheat and corn fields using multispectral UAV imagery.

(2) Generate both parametric and non-parametric machine learning regression models to predict crop nitrogen weight in corn fields using multispectral UAV imagery.

 (3) Generate nitrogen prediction maps using UAV multispectral images and visually examine the spatial distribution of nitrogen within the fields.

Ultimately, these images should be given to farmers in highly accurate, quick and timely manner field information for their precision nitrogen fertilization management.

1.4 Thesis Structure

To answer the research questions, this thesis is divided into an introduction, two academic journal formatted papers, and a conclusion. The introduction (chapter 1) provides the background of this research, a brief review on remote sensing techniques used for crop monitoring and listed the objectives of the thesis. The next two chapters are separate published academic journal-formatted papers. The first paper (chapter 2) answers the research questions using a linear regression and attempts to derive conclusions without the use of non-linear transformation. This chapter uses ground spectral measurements of the canopy using an Analytic Spectral Device (ASD) and verified the reflectance values of the UAV images. Due to its similar values of ASD ground measurements and UAV derived measurements, chapter 3 directly used the reflectance values from the UAV for analysis. The second paper (chapter 3) answers the

research questions using nonparametric machine learning techniques such as Random Forests and support vector regression (SVR), while using parametric linear regression as a baseline comparison. Both studies have a similar data collection method, study area, and methodology. The fourth and final chapter of this thesis is the conclusion as it states/explains objective answers of this thesis. The 4th chapter also provides some limitations and suggestions for chapter 2 and 3.

1.5 Study Areas

The study area for both studies are location in Mt. Brydges and Melbourne, Ontario (30km west of London, Ontario) (Figure 1-1). The fields Bale, Century, Crandell, Hetzell, Jack North, and McColl are the study sites for chapter 2. The fields JJ and Susan are the study sites for chapter 3. Crop lands in this region are dominated by soybeans, grain corn, and winter wheat (Statistics Canada, 2016). The topography in southwest Ontario is generally flat with hot, humid summers. The study area is situated between Lake Huron and Lake Erie, which affects the tendency of frequent thunderstorms and heavy precipitation. Southern Ontario is known for its rich, Class 1 agricultural soil as a result of the glacial ice-age deposits (Leal, 2016).

Figure 1-1. Study areas located in Mt. Brydges/Melbourne, Ontario. White represents corn fields and red represents wheat fields

1.6 References

Agriculture and Agri-Food Canada (2018). We grow a lot more than you think. *Government of Canada. Retrieved from* http://www.agr.gc.ca/eng/about-us/publications/we-grow-a-lot-more-than-you-may-think/?id=1251899760841

Bagheri, N., Ahmadi, H., Alavipanah, S. K., & Omid, M. (2013). Multispectral remote sensing for site-specific nitrogen fertilizer management. *Pesquisa Agropecuária Brasileira*, *48*(10), 1394-1401.

Barbanti, L., Adroher, J., Damian, J. M., Di Virgilio, N., Falsone, G., Zucchelli, M., & Martelli, R. (2018). Assessing wheat spatial variation based on proximal and remote spectral vegetation indices and soil properties. *Italian Journal of Agronomy*, *13*(1).

Broge, N. H., & Leblanc, E. (2001). Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sensing of Environment*, *76*(2), 156-172.

Clevers, J. G., & Kooistra, L. (2011). Using hyperspectral remote sensing data for retrieving canopy chlorophyll and nitrogen content. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *5*(2), 574-583.

Colombo, R., Meroni, M., Marchesi, A., Busetto, L., Rossini, M., Giardino, C., & Panigada, C. (2008). Estimation of leaf and canopy water content in poplar plantations by means of hyperspectral indices and inverse modeling. *Remote sensing of environment*, *112*(4), 1820-1834.

Comar, A., Burger, P., de Solan, B., Baret, F., Daumard, F., & Hanocq, J. F. (2012). A semi-automatic system for high throughput phenotyping wheat cultivars in-field conditions: description and first results. *Functional Plant Biology*, *39*(11), 914-924.

Darvishzadeh, R., Atzberger, C., Skidmore, A., & Schlerf, M. (2011). Mapping grassland leaf area index with airborne hyperspectral imagery: A comparison study of statistical approaches and inversion of radiative transfer models. *ISPRS Journal of Photogrammetry and Remote Sensing*, *66*(6), 894-906.

Furbank, R. T., & Tester, M. (2011). Phenomics–technologies to relieve the phenotyping bottleneck. *Trends in plant science*, *16*(12), 635-644.

Gnyp, M. L., Panitzki, M., Reusch, S., Jasper, J., Bolten, A., & Bareth, G. (2016). Comparison between tractor-based and UAV-based spectrometer measurements in winter wheat. In *Proceedings of the 13th International Conference on Precision Agriculture, Monticello, IL, USA* (Vol. 31).

Gutierrez-Rodriguez, M., Escalante-Estrada, J. A., & Rodriguez-Gonzalez, M. T. (2005). Canopy reflectance, stomatal conductance, and yield of Phaseolus vulgaris L. and Phaseolus coccinues L. under saline field conditions. *Int. J. Agric. Biol*, *7*, 491-494.

Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J., & Strachan, I. B. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sensing of Environment*, *90*(3), 337-352.

Inman, D., Khosla, R., & Mayfield, T. (2005). On-the-go active remote sensing for efficient crop nitrogen management. *Sensor Review*, *25*(3), 209-214.

Jay, S., Maupas, F., Bendoula, R., & Gorretta, N. (2017). Retrieving LAI, chlorophyll and nitrogen contents in sugar beet crops from multi-angular optical remote sensing: Comparison of vegetation indices and PROSAIL inversion for field phenotyping. *Field Crops Research*, 210, 33-46.

Jones, H. G., & Vaughan, R. A. (2010). *Remote sensing of vegetation: principles, techniques, and applications*. Oxford university press.

Jordan, C. F. (1969). Derivation of leaf area index from quality of light on the forest floor. *Ecology*, *50*(4), 663-666.

Leal, J. (2016). Sustaining Ontario's agricultural soils. *Gov. of Ontario.* Retrieved from http://www.omafra.gov.on.ca/english/landuse/soil-paper.pdf

Lee, H., Wang J., & Leblon, B. (2020). Canopy nitrogen retrieval from unmanned aerial vehicle imagery for wheat and corn fields. (*accepted)*

Liu, Y., Cheng, T., Zhu, Y., Tian, Y., Cao, W., Yao, X., & Wang, N. (2016). Comparative analysis of vegetation indices, non-parametric and physical retrieval methods for monitoring nitrogen in wheat using UAV-based multispectral

imagery. In 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS'16), Beijing, China, July, pp. 7362-7365.

Loel, J., Kenter, C., Märländer, B., & Hoffmann, C. M. (2014). Assessment of breeding progress in sugar beet by testing old and new varieties under greenhouse and field conditions. *European journal of agronomy*, *52*, 146-156.

Maes, W. H., & Steppe, K. (2019). Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture. *Trends in plant science*, *24*(2), 152-164.

Maresma, Á., Ariza, M., Martínez, E., Lloveras, J., & Martínez-Casasnovas, J. A. (2016). Analysis of vegetation indices to determine nitrogen application and yield prediction in maize (*Zea mays* L.) from a standard UAV service. *Remote Sensing*, *8*(12), 973.

Milford, G. F. J., Pocock, T. O., & RILEY, J. (1985). An analysis of leaf growth in sugar beet. I. Leaf appearance and expansion in relation to temperature under controlled conditions. *Annals of Applied Biology*, *106*(1), 163-172.

Munoz-Huerta, R. F., Guervara-Gonzalez, R. G., Contreras-Medina, L. M., Torres-Pacheco, I., Prado-Olivarez, J. & Ocampo-Velazquez, R. V. (2013). *Sensors.* 13, 10823-10843.

Prey, L., & Schmidhalter, U. (2019). Simulation of satellite reflectance data using high-frequency ground based hyperspectral canopy measurements for in-season estimation of grain yield and grain nitrogen status in winter wheat. *ISPRS journal of photogrammetry and remote sensing*, *149*, 176-187.

Raparelli, E., & Bajocco, S. (2019). A bibliometric analysis on the use of unmanned aerial vehicles in agricultural and forestry studies. *International Journal of Remote Sensing*, *40*, 1-14.

Schmaltz, R. (2017). What is precision agriculture? *AFN.* Retrieved from https://agfundernews.com/what-is-precision-agriculture.html

Statistical summary of Ontario agriculture. (2016) *Ontario Ministry of Agriculture, Food and Rural Affairs.* Retrieved from http://www.omafra.gov.on.ca/english/stats/agriculture_summary.htm

Xie, Q., Dash, J., Huang, W., Peng, D., Qin, Q., Mortimer, H., Casa, R., Pignatti, S., Laneve, G., Pascucci, S. & Dong, Y. (2018). Vegetation indices combining the

red and red-edge spectral information for leaf area index retrieval. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*(5), 1482-1493.

Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J., & Reynolds, J. M. (2012). 'Structure-from-Motion'photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology*, *179*, 300-314.

White, J. W., Andrade-Sanchez, P., Gore, M. A., Bronson, K. F., Coffelt, T. A., Conley, M. M., ... & Jenks, M. A. (2012). Field-based phenomics for plant genetics research. *Field Crops Research*, *133*, 101-112.

Zarco-Tejada, P. J., Miller, J. R., Morales, A., Berjón, A., & Agüera, J. (2004). Hyperspectral indices and model simulation for chlorophyll estimation in open-canopy tree crops. *Remote sensing of environment*, *90*(4), 463-476.

Zha, H., Miao, Y., Wang, T., Li, Y., Zhang, J., Sun, W., Feng, Z. & Kusnierek, K. (2020). Improving Unmanned Aerial Vehicle Remote Sensing-Based Rice Nitrogen Nutrition Index Prediction with Machine Learning. *Remote Sensing,12*(2), 215.

Zhang, C., & Kovacs, J. M. (2012). The application of small unmanned aerial systems for precision agriculture: a review. *Precision agriculture*, *13*(6), 693-712.

Chapter 2

2 Canopy Nitrogen Retrieval from Unmanned Aerial Vehicle Imagery for Wheat and Corn Fields

2.1 Introduction

Precision agriculture or precision farming requires detailed spatial information of crop status in order to optimize crop inputs such as nutrients, pesticides, seeds or water as a function of the crop yield and reduce associated costs (Xie et al. 2018; Liu et al. 2016). Crop nitrogen content is one of the good crop status indicators. Indeed, nitrogen is the main plant nutrient needed for chlorophyll production, which has a direct impact on plant growth and yield. Therefore, the optimization of nitrogen fertilization has become a body of intense research due to its environmental and economic impact (Muñoz-Huerta et. al. 2013). There are two ways to measure crop nitrogen for precision agriculture purposes: ground-based and remote sensing methods. Ground-based methods require intensive field data collection, which can be time consuming, destructive, and limited to a small spatial area, making it impractical for fast and efficient results in most agricultural fields in Canada that can reach up to hundreds of acres in size. An alternative is to use remote sensing methods. Remote sensing methods are non-destructive and can cover much larger spatial areas than the ground-based method. However, remote sensing applications in agriculture often require high temporal and spatial resolution imagery. These images are often costly or difficult to obtain, if they are acquired by spaceborne or airborne platforms (Raparelli & Bajocco, 2019). Satellite imagery could be used to monitor nitrogen status across large areas. However, they cannot provide enough spatial and temporal accuracy because of their low spatio-temporal resolution (Zheng et al. 2018). Thus, this does not give the user much flexibility to determine their spatial or temporal resolutions requirements. Optical satellite imagery is also limited to data quality and accessibility due to weather conditions such as fog, haze and clouds. This makes them unreliable to monitor crop growths (Li et al. 2018). Synthetic aperture radar (SAR) imagery has the advantages to be acquired whatever the sky conditions, but SAR systems have poor revisiting periods and their images can get very expensive and complex to interpret (Zhang & Kovacs, 2012). Sentinel-1 SAR data are freely available but lack to

provide enough spatial resolution (10m) for precise small-scaled applications (Nasrallah et al. 2019).

The use of Unmanned Aerial Vehicle (UAV) imagery is an emerging technology, filling in gaps between satellite imagery, aerial photography and field samples (Kelcey & Lucieer, 2012). Most UAV studies have also built on the experience of tractor-based N sensors, due to their limitations on estimating the correct N status (Maes & Steppe, 2019) and its ability to drive on agricultural sites with difficult soil conditions (Gynp et al, 2016). Image acquisition with UAV can be deployed quickly and repeatedly, at a low cost, and with greater flexibility (Maresma, et. al. 2016). Since most agricultural fields in Canada have homogenous canopies, a clear advantage over airborne imagery is that UAVs can easily achieve high image overlapping of 90% or more, which is more useful for mosaicking uniform images (Song, 2016). UAVs can also be equipped with various types of equipment such as optical, radar and thermal sensors along with a georeferencing system.

The focus of this study is to test if UAV imagery can be used to retrieve crop nitrogen status in a perspective of precision nitrogen fertilization. The two most common remote sensing techniques to estimate nitrogen content at the canopy level are: Radiative Transfer Model (RTM), which estimates the chlorophyll or nitrogen content by describing the interaction between the sun's light and the crop canopy. An example of an RTM is the PROSAIL model, which uses various parameters at the leaf and canopy level and can be mathematically inverted to retrieve chlorophyll or nitrogen content from spectral data (Clevers & Kooistra, 2011; Clevers & Gitelson, 2013; Hansen & Schjoerring, 2003, Botha et al. 2010). The second technique uses empirical methods such as machine learning techniques, simple/multiple-linear regression methods to retrieve crop nitrogen from spectral data (Clevers & Kooistra, 2011). This project used the second technique, due to the complexity of using an RTM. Jay et al. (2017) has also shown that empirical methods using vegetation indices provide slightly more accurate estimations of chlorophyll and nitrogen content than PROSAIL inversion.

Currently, there are several cameras in the commercial market that are available for UAV-based precision agriculture. Deng et al. (2018) found that narrowband multispectral cameras acquire images with far better quality than broadband cameras. For this study, we used a MicaSense RedEdge narrowband multispectral camera attached to a fixed-wing UAV. Several studies tested linear regression to predict crop from MicaSense RedEdge imagery (Walsh et al. 2018; Olson et al. 2019; Sofonia et al. 2019). Almost all of these studies have focused on the use of NDVI (Maresma et al. 2016). Very few studies have tested the use of other vegetation indices such as the Ratio Vegetation Index (RVI). Many studies have avoided the use of RVI and instead, used VIs that perform well on single date/growth stages. This study wanted to incorporate multiple dates/growth stages and used ratio-based VIs, which was found to have a positive linear relationship with nitrogen throughout the growing stage in wheat (Muñoz-Huerta et. al. 2013).

In this study, we will use an empirical method to statistically relate spectral measurements and crop nitrogen contents. The resulting models will then be applied to UAV imagery for mapping crop nitrogen content. The objectives of this study are to: (i) generate a linear regression model that can be applied to each wheat and corn fields throughout the growing season to estimate nitrogen from MicaSense RedEdge imagery acquired over corn and winter wheat crops, (ii) determine which spectral variable(s) are the best predictor of crop nitrogen linearly; (iii) generate a nitrogen prediction map with the entire UAV image and analyse if the UAV images are able to detect any spatial variation of nitrogen within the fields. This study attempts to fill in research gaps on improving the accuracy of canopy nitrogen prediction throughout the entire growing season using UAV-based MicaSense multispectral imagery acquired over wheat and corn fields in Southwest Ontario. Figure 2-1 shows the flow chart of the procedure in this paper.
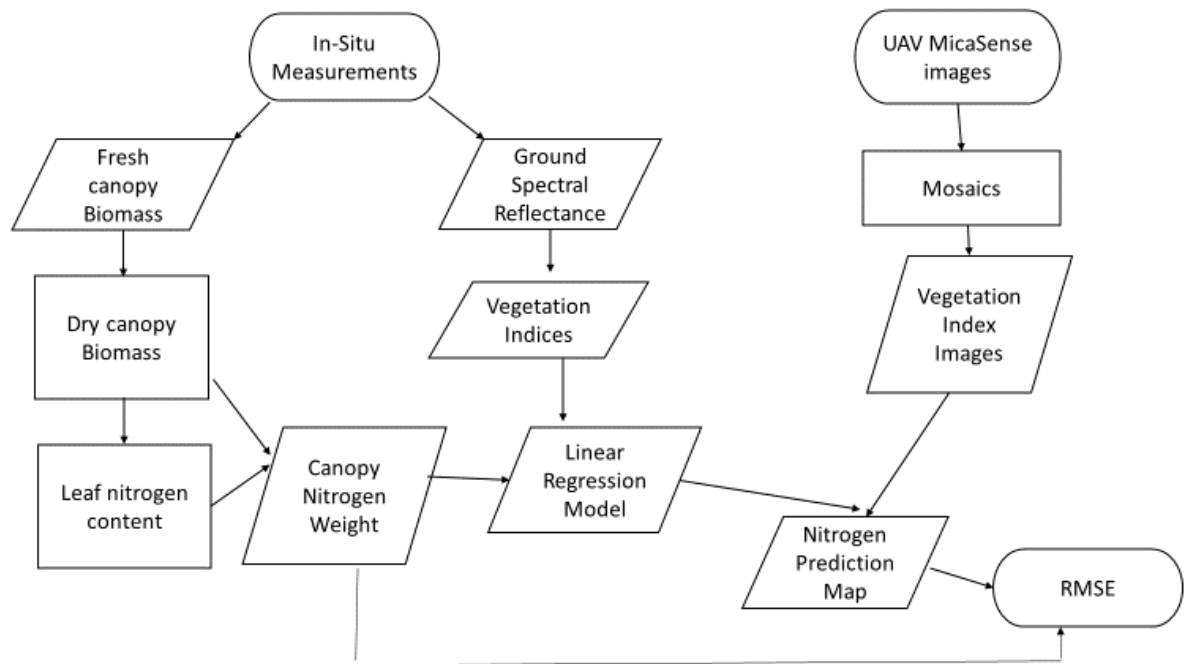
Figure 2-1. Flowchart of the methodology (chapter 2)

2.2 Materials and Methods

2.2.1 Study Area

The study site was located in Mt. Brydges, Ontario (42°54'N, 81°29'W; 20 km west of London, Ontario, Canada). This region is in the humid continental climate zone in Canada and the summers are typical hot and humid with an average temperature of 20° C and a high of 31° C in July, while experiencing harsh winters. The humidity in this region is around 75-80 % in July. The study site in Mt. Brydges is dominated by agricultural land with very little urban use. The common agricultural practice in this region is one crop harvested per annual cycle. The most dominant crops in this region of Southwest Ontario are wheat, corn and soybeans (Liao et al. 2019).

We collected field data from three hard winter wheat and three grain corn fields in the summer of 2018. Winter wheat seeds are usually sawn in October prior to the summer and harvested in September following the summer. Winter wheat is predominately harvested in south west Ontario and is favoured over spring wheat. Across Canada,

winter wheat is harvested on significantly less area than spring wheat, while producing considerably higher average yield at 4.3 tonnes per hectare (Larsen, 2012). Corn is generally planted in May just before the summer and typically harvested in late October/early November once the crop is dried and the starch content is high. Grain corn represents almost 95 % of the corn grown in Ontario and is used for livestock feed, ethanol fuel, and sweeteners (Hamel & Dorff, 2015).

Data was collected on a farmer's field, not an experimental field. A total of four sampling dates were collected for both wheat and corn. Fields on average were roughly 16-24 ha in size and were dispersed in a 3 km radius.

2.2.2   Field Data

In-situ data were collected over 8 sampling points on each field, with one wheat field holding 16 sample points from May-to-July 2018 (Figure 2-2). In-situ data included ground spectral reflectance and biomass. Ground spectral reflectance spectra between 325 nm and 1075 nm were acquired using an ASD FieldSpec Handheld v2. Field of view (FOV) was measured at 25°. Each sample point had an average of 8-10 readings and was measured at a height of roughly 1.5 m directly above the crop. The spectra were obtained on relatively cloud-free days between 9am-2pm. Calibration with a white board occurred before readings and were re-calibrated if sky conditions changed, such as in the case of cloud cover. Spectra reading occurred weekly during the growing season but was interrupted for corn when the height of the corn exceeded the user's height. This protocol follows Le Maire et al. (2008) who suggested that canopy spectral reflectance data should be retrieved weekly to avoid any nitrogen dilution, due to rapid biomass growth.
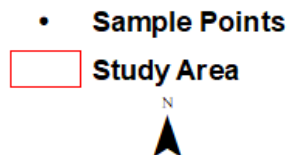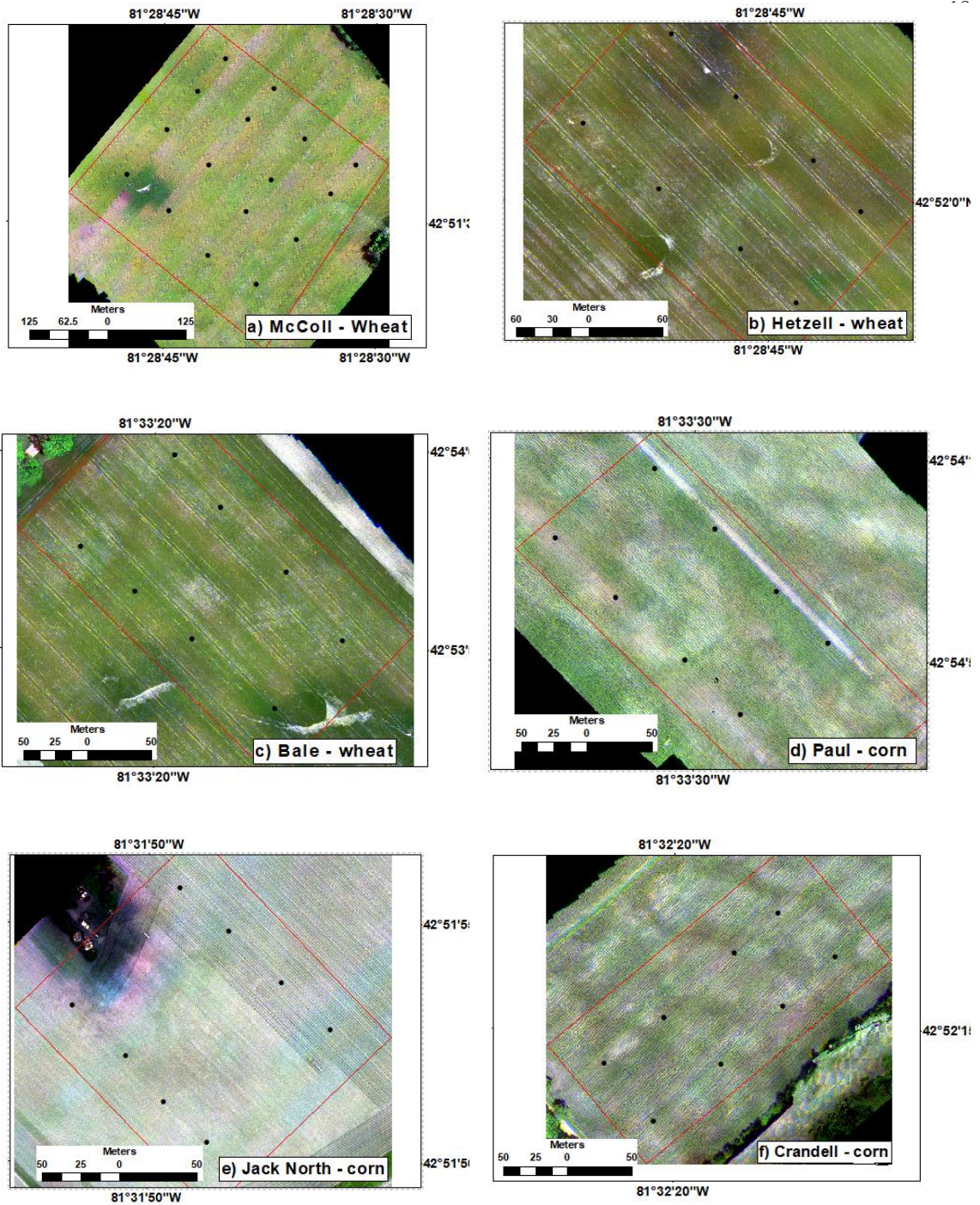
Figure 2-2 RGB mosaics made with the MicaSense images acquired on May 24th, 2018 over the wheat fields named at a) McColl, b) Hetzell, c) Bale) and on June 7th,

2018 of corn fields named d) Paul, e) Jack North, f) Crandell with ground sample point distribution. All fields are located in Mt. Brydges, Ontario

Crop biomass was measured destructively by removing the crop (at the bottom of the stem) immediately after collecting canopy reflectance on a 0.5 m^2 block for wheat and 1 m^2 block for corn. The average row distance for wheat was 18 cm and typically had three rows of wheat in a 0.5 m^2 block. The average row distance for corn was 75 cm and typically corn fields had an average of 12-14 plants per 1 m^2 area.

The plant biomass was lightly washed to remove any soil, dust, fertilizer, or spray residues on them. Biomass was then weighed at the fresh stage, then dried in 80° C oven for a 24-36 hours. Dried biomass weight (scaled at g/m^2) was measured then sent to A & L Canada Laboratory for plant tissue analysis. The leaves of the oven dried samples were grinded into powder form and passed through a 1mm sieve. The leaf nitrogen content (expressed as percentage) was then measured using the Laboratory Equipment Company (LECO) FP628 nitrogen/protein analyser that uses the total nitrogen combustion method at a temperature of 1050° C (AOAC 2006). Wheat leaf nitrogen content (%) throughout the growing season ranged between 3.4 % to 6.8 % and corn leaf nitrogen content (%) ranged between 2.3 % and 6 %. Both crops had a decrease trend in leaf nitrogen content, possibly due to the nitrogen contribution to the crop changes throughout the growing season.

2.2.3   UAV Imagery

As crop physiology and soil structure change over time, UAV collection is optimal when done during or immediately after or before field collection. UAV flights were performed the day of or within two days before collecting in-situ ground measurement data. UAV flights were performed on May 17th, May 24th, June 7th, and June 19th, 2018 by A & L Canada Labs Inc. over the entire fields (Table 2-1). The UAV imagery was acquired by a MicaSense Red edge camera using a Dà-Jiāng Innovations (DJI) Matrice 100 quadcopter, flown in a zigzag route at 40 m in height and 80 % overlap, with an average of 30-45 min of flight time depending on the size of the field. The software used to pre-program the flights was MicaSense Atlas Flight mission planner, a free app downloadable on a

mobile/tablet device on the Apple app store. The app allows the user to create a flight pattern with the desired speed, altitude and overlap parameters. The app also allows the ability to pause, stop and resume flights due to any circumstances such as battery change.

Table 2-1. Summary of UAV flight acquisition in the study (2018)

Crop	Field	Date of UAV Flight	Weather Condition	Growth Stage
Wheat	Bale	May 24th	27.3° C, Sunny	F6/F7
	Hetzell	May 24th	27.3° C, Sunny	F6/F7
		June 7th	23.2° C, Sunny	F10
	McColl	May 17th	26.3° C, Sunny	F5/F6
		May 24th	27.3° C, Sunny	F7/F8
Corn	Crandell	June 7th	23.2° C, Sunny	V2
		June 19th	26.0° C, Sunny	V5/V6
	Jack North	June 7th	23.2° C, Sunny	V2
	Paul	June 19th	26.0° C, Sunny	V5/V6

2.2.4 Field Canopy Reflectance Processing

ASD spectral reflectance data was processed using ViewSpecPro and converted to (.csv) files. ASD spectra were averaged at each sample point and divided by the white board spectra to compute canopy reflectance spectra. Because the purpose of the study is to test if UAV imagery acquired by the MicaSense Red edge camera can be used to monitor crop nitrogen, ASD reflectance spectra were converted to reflectance values that corresponded to the following 5 MicaSense Red edge bands (Table 2-2): (1) blue, (2) green, (3) red, (4) red edge, and (5) near-infrared. Figure 2-3 shows a typical reflectance spectrum of both wheat and corn canopies and the corresponding 5 MicaSense bands. Both band 1 (blue) and band 2 (green) have a bandwidth of 20 nm, while band 3 (red) and band 4 (red edge) have a narrow range of 10 nm. Lastly, band 5 (near infrared) has a bandwidth of 40 nm. These reflectances were then used to compute vegetation indices, which are algebraic transformation of reflectances in two or more bands. Vegetation indices are designed to enhance the contribution of the optical properties of the vegetation on the total spectral response of the canopy. Therefore, vegetation indices attempt to correct any confounding factors such as reflectance of soil backgrounds in a crop (Clevers & Kooistra, 2011). We used the most widely used vegetation indices (Table 2-4): RVI, NDVI, Green NDVI, MTVI2, and Red edge NDVI. RVI has been

already used to estimate nitrogen status while being insensitive to growth stages and crop type (Muñoz-Huerta et al. 2013). NDVI and Green NDVI have been used extensively in literature to study various biological parameters including nitrogen (Zhang & Kovacs, 2012). MTVI2, which is an improved vegetation index of the triangle vegetation index (TVI), was found to be useful to estimate nitrogen content with multispectral images (Bagheri et al. 2013). Red edge NDVI was common in other studies involving nitrogen management (Li et al. 2018; Olson et al. 2019) and was used to test the MicaSense's red edge band capability in predicting nitrogen status. Indeed, red edge-based vegetation indices were found to produce the highest R^2 with nitrogen in corn fields (Olson et al. 2019, Clevers & Gitelson, 2012).

Table 2-2. Spectral characteristics of the 5 MicaSense bands

Band #	Name	Band Range (nm)	Centre Wavelength (nm)	Bandwidth (nm)
1	Blue	465-485	475	20
2	Green	550-570	560	20
3	Red	663-673	668	10
4	Red edge	712-722	717	10
5	NIR	820-860	840	40

Table 2-3. Vegetation indices used in this study

Index ([1])	Formula ([2])	Authors
GNDVI	(NIR-GREEN)/(NIR+GREEN)	Gitelson & Merzlyak, 1996
MTVI2	$\dfrac{1.8(\text{NIR}-\text{GREEN})-3.75(\text{RED}-\text{GREEN})}{\sqrt{(2\text{NIR}+1)2\ -6(\text{NIR}-5\sqrt{\text{RED}})-0.5}}$	Bagheri et al. 2013
NDVI	(NIR-RED)/(NIR+RED)	Rouse et al. 1974
RE_NDVI	(NIR-REDEDGE)/(NIR+ REDEDGE)	Barnes et al. 2000
RVI	NIR/RED	Jordan, 1969

[1] GNDVI = green normalized difference vegetation index; MTVI2 = modified triangular vegetation index; NDVI = normalized difference vegetation index; RE_NDVI = red edge normalized difference vegetation index; RVI = ratio vegetation index [2] GREEN = green reflectance; RED = red reflectance; REDEDGE = red edge reflectance; NIR = near-infrared reflectance
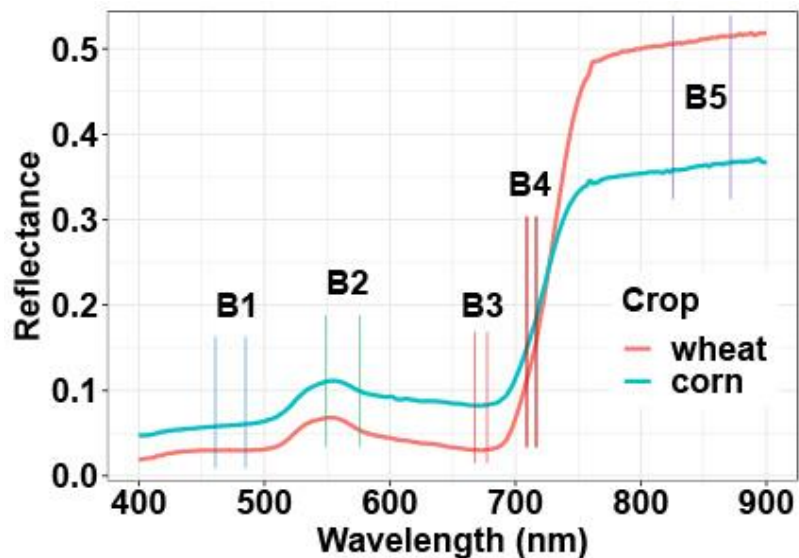
Figure 2-3. Typical reflectance spectrum of wheat and corn plants with the location of the 5 MicaSense bands. B1: band 1 (blue); B2: band 2 (green); B3: band 3 (red); B4: band 4 (red edge); B5: band 5 (near-infrared). Spectral reflectance of wheat taken at McColl field on May 14th, 2018 and corn at Crandell on June 12th, 2018

2.2.5   Nitrogen Estimation Modelling

All the modelling was performed in R Studio and figures were produced using the "ggplot2" package (Wickham, 2009). To describe the canopy nitrogen status, we used the canopy nitrogen weight that is defined by Hansen and Schjoerring (2003) as follows:

$$Canopy\ Nitrogen\ Weight\ = leaf\ nitrogen\ content\ * dry\ biomass\ weight \quad [2]$$

Equation 2 assumes that all the leaves from a sample gathered in the field contained the same amount of nitrogen. Canopy nitrogen weight (g/m^2) has the advantage of being a more absolute value, compared to plant or leaf nitrogen content (%), which is a more relative value. Absolute values allow the ability to compare the results among different fields and dates. Additionally, previous studies have shown that estimating biochemical concentrations at the leaf level is difficult. Therefore, focusing on the canopy level is optimal (Clevers & Kooistra, 2011).

The MicaSense Red edge vegetation index values derived from canopy reflectance spectra were used in simple linear regression models as independent variables to estimate the canopy nitrogen weight as follows:

$$Canopy\ Nitrogen\ Weight = a + b\ (VI_{ASD}) \qquad [3]$$

where VI_{ASD} represents the vegetation index values derived from the ASD-simulated MicaSense bands. We only considered in our study linear regressions because nonlinear regressions such as exponential, power or polynomial can lead to multiple nitrogen values for the same VI value. This is not practical as the developed method needs to be used by growers to precisely spray fertilizations and there is the need of having a unique nitrogen value for each VI value. The relationship between the vegetation index to the N status can be misleading if the best-fit function (R^2) is not linear because the sensitivity between the two will not be constant (Gitelson, 2013). After the analysis of simple linear regression of each vegetation index, we also used stepwise regression with backward selection of all the calculated VIs to develop a model that have more than one input vegetation index. In addition to R^2 and root mean square error (RMSE), the performance of the model was assessed as a function of possible significant multicollinearity between input variables using the "car" package (Fox & Weisberg, 2019) and VIF() function in R Studio. Multicollinearity is a statistical phenomenon in which two or more explanatory variables are correlated with each other and it reduces the accuracy of the estimates of the regression coefficients. Multicollinearity is often an issue and typically needs to be formally explored for any analysis using more than one explanatory variable using the variance inflation factor (VIF) equation:

$$VIF_j = \frac{1}{1-R_j^2} \qquad [4]$$

Where R^2_j is the R^2 from the regression of explanatory variable j to all the other predictors in the model. If R^2_j is close to one, then multicollinearity exists with that explanatory variable, as the VIF value will be large (James et al. 2013). A general rule of thumb is that VIF values > 5 present a high degree of multicollinearity and the explanatory variable should be removed from the model.

2.2.6   UAV Image Processing

The UAV images were processed using Pix4D (Pix4D SA, Lausanne, Switzerland) to generate an orthomosaic image of each field by stitching hundreds of different images captured during the same flight into one single 2D image or 3D point cloud and corrected for perspective. Because the camera on the UAV travels from point A to point B, Pix4D accounts for the differences in distance from when the camera changes its position during its route. Pix4D uses the technique called Structure from Motion (SfM). SfM is a technique that uses multiple overlapping images from various angles as inputs and extracts 3D surface information with matching points that were found in the overlapping space between the images (Harwin & Lucieer, 2012; Pricope et al. 2019; Westoby et al. 2012). The outputs of the mosaic UAV imagery can be produced at centimetre level resolutions, which allows the ability to easily differentiate between the soil and the canopy on the imagery. Keypoints in Pix4D was left at default 1 image scale. However, once the canopy became dense, Pix4D recommends changing the keypoints setting to ½ image scale for a higher number of calibrated images (Pix4D documentation, 2020). Typically, the mosaic imagery produced will show large areas of missing outputs if the number of the calibrated images are low. Therefore, we adjusted the keypoints parameter accordingly once the canopy became dense in the later growing stage, particularly in the imagery of Hetzell (wheat) on June 7th. The final output of the mosaic image was produced at 5 cm/pixel and five ground plates were used for georeferencing. For each flight acquisition, each field's vegetation index image was computed from the mosaic image and exported to a (.tif) file. Mosaic images were automatically radiometrically corrected in Pix4D.

The UAV vegetation index images were used to compute a prediction canopy nitrogen weight map of each field by applying the following equation to the imagery with the ArcMap Raster Calculator:

$$Canopy\ Nitrogen\ Weight = a_1 + b_1(VI_{UAV}) \qquad\qquad [5]$$

where VI_{UAV} is the vegetation index values derived from the UAV imagery, while *a* and *b* are the regression coefficients of Equation 3. After generating the nitrogen prediction

map of each field and at each date, root mean square error (RMSE) values were calculated by comparing the predicted and actual canopy nitrogen weight values around the sample point for each corn or wheat field using the following equation:

$$RMSE = \frac{\sum_{i=1}^{n}(P_i - O_i)^2}{n} \qquad [6]$$

where P_i represents the predicted canopy nitrogen weight value (g/m^2), while O_i represents the observed canopy nitrogen weight value (g/m^2). n is the number of observations in the calculation, and i is the index of summation in increments of 1. A polygon shapefile, roughly the size of the ground sample point was used to calculate the average value of the canopy nitrogen weight prediction values in ArcMap.

2.3 Results

2.3.1   Nitrogen Estimation Modelling

Canopy nitrogen weight shows an increase trend overtime throughout the growing season for both the wheat and corn crops (Figure 2-4). Figure 2-4 also shows that the later growing stage has a larger variation in canopy nitrogen weight for both the wheat (June 4th) and corn crops (June 26th and July 4th), being much more evident in the corn crops (Figure 2-4b). By contrast, the early growing stage of the corn crops (June 4th) has very little variation in canopy nitrogen weight (Figure 2-4b). This may be due to the small amount of biomass present at this early growing stage.
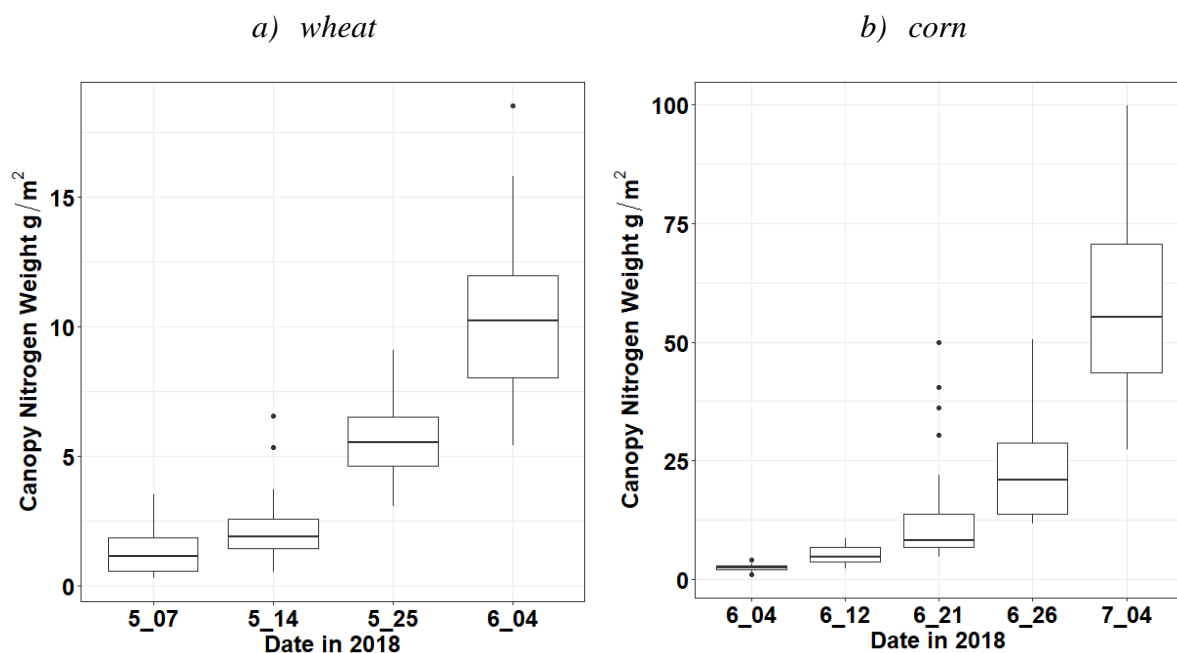
a) wheat
b) corn

Figure 2-4. Box plot showing the variation of canopy nitrogen weight (g/m²) as a function of the time in the studied wheat and corn fields during the 2018 growing season

In this study we found that there was almost no relationship between the vegetation indices and leaf nitrogen content (%) throughout the entire growing season for both wheat and corn crops (Table 2-4). The lack of relationship between leaf nitrogen content (%) and vegetation indices throughout the growing season was also observed in a study on sugar beet crops (Jay et al. 2017), on litchi orchards (Li et al. 2016) and on wheat crops (Song, 2016).

Table 2-4. R^2 between leaf nitrogen (%) and the vegetation indices used in the study

Crop	Vegetation Index	R^2
Wheat	RVI	0.39
	NDVI	0.38
	GNDVI	0.37
	MTVI2	0.41
	RE_NDVI	0.37
Corn	RVI	0.20
	NDVI	0.39
	GNDVI	0.28
	MTVI2	0.32
	RE_NDVI	0.31

Figures 2-5 and 2-6 presented the regression models between the reflectance of individual MicaSense simulated bands and the canopy nitrogen weight for the wheat and corn fields, respectively. For all bands except the near-infrared band (band 5), there is a decrease in the reflectance as canopy nitrogen weight (g/m^2) increases. Out of all the bands, the near-infrared (band 5) reflectance had the best correlation with canopy nitrogen weight for both the wheat and corn fields (R^2 = 0.61, R^2 = 0.54). Predicting nitrogen with a single near-infrared band could be useful. However, it showed signs of saturation after a reflectance value of 0.5 (Figures 2-5e and 2-6e). The red edge reflectance has already been shown to be sensitive to leaf chlorophyll content in plants (Jones & Vaughn, 2010), but also to be less sensitive at higher contents of chlorophyll (saturation effect) (Clevers & Kooistra, 2011). This could be also the case for nitrogen as nitrogen is the main component of chlorophyll and is well correlated to chlorophyll content (Clevers & Kooistra, 2011). The red edge reflectance has been found to be significantly related to corn nitrogen weight (Schlemmer et al. 2013; Li et al. 2014). However, in this study, the reflectance of the red edge band (band 4) consistently showed a weak correlation with canopy nitrogen weight for both the wheat and corn fields (R^2 = 0.24, R^2 = 0.001). This may be due to the red edge position effect that occurs with the sharp position of the sharp change in reflectance in the red edge region. This sharp change of reflectance is known to be a sensitive indicator of leaf chlorophyll (Jones & Vaughan, 2010; Zarco-Tejada et al. 2002; Curran et al. 1990). The MicaSense camera has a narrowband of 10 nm in the red

edge region, therefore this study may have had weak correlation with canopy nitrogen weight as the camera may have not fully captured the red edge position throughout the growing season. As shown in Figures 2-5c and 2-6c, green reflectance is also poorly related to canopy nitrogen weight, probably because we used data of the entire growing. The chlorophyll saturates in the middle to late growing stage, causing the crops reflecting the same amount of green wavelength. The green wavelength is closely related to the leaf chlorophyll *a* and *b* contents (Zhao et al. 2018). However, this result is not in agreement with Schlemmer et al. (2013) and Li et al. (2014) who observed a good relationship between green reflectances and corn nitrogen weight. We also observed weak relationships with either the blue or the red reflectance both crop types (Figures 2-5a, 2-5b, 2-6a, 2-6b). Similarly as the green reflectance, reflectances in the blue and red bands are highly related to chlorophyll contents because both bands correspond to chlorophyll absorption bands (Jones & Vaughn, 2010). However, the blue and red bands are reflecting less than the green band due to the chlorophyll reflecting more in the green wavelength. For the three visible bands (blue, green, and red), the relationship with the crop nitrogen weight is negative (Figures 2-5a-c, 2-6a-c). These results are similar to Yao et al. (2013) who observed that the reflectance in the visible bands decreases with increasing nitrogen (g/m^2) for wheat crops.
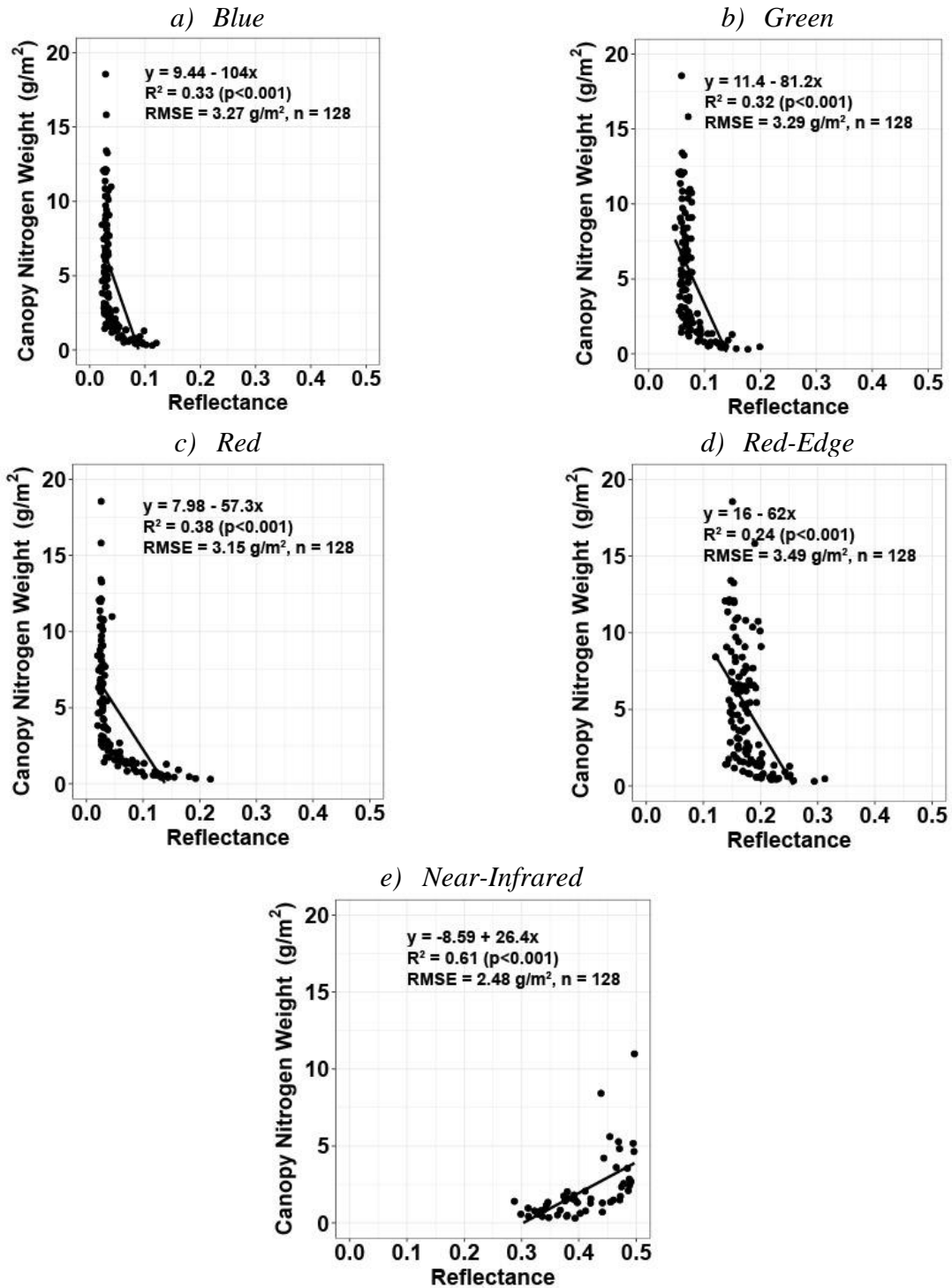
Figure 2-5. Relationship between canopy nitrogen weight and the MicaSense simulated reflectance in a) the blue MicaSense band, b) the green MicaSense band, c) the red MicaSense band, d) the red edge MicaSense band, and e) the near-infrared MicaSense band for the wheat fields

a) Blue

y = 26.9 - 296x
R^2 = 0.25 (p<0.001)
RMSE = 9.55 g/m^2, n = 176

b) Green

y = 24 - 110x
R^2 = 0.077 (p<0.01)
RMSE = 10.58 g/m^2, n = 176

c) Red

y = 26.3 - 206x
R^2 = 0.40 (p<0.001)
RMSE = 8.55 g/m^2, n = 176

d) Red-Edge

y = 13.7 -11.8x
R^2 = 0.001 (p=0.651)
RMSE = 11.01 g/m^2, n = 176

e) Near-Infrared

y = -18.3 + 80.8x
R^2 = 0.54 (p<0.001)
RMSE = 7.49 g/m^2, n = 176
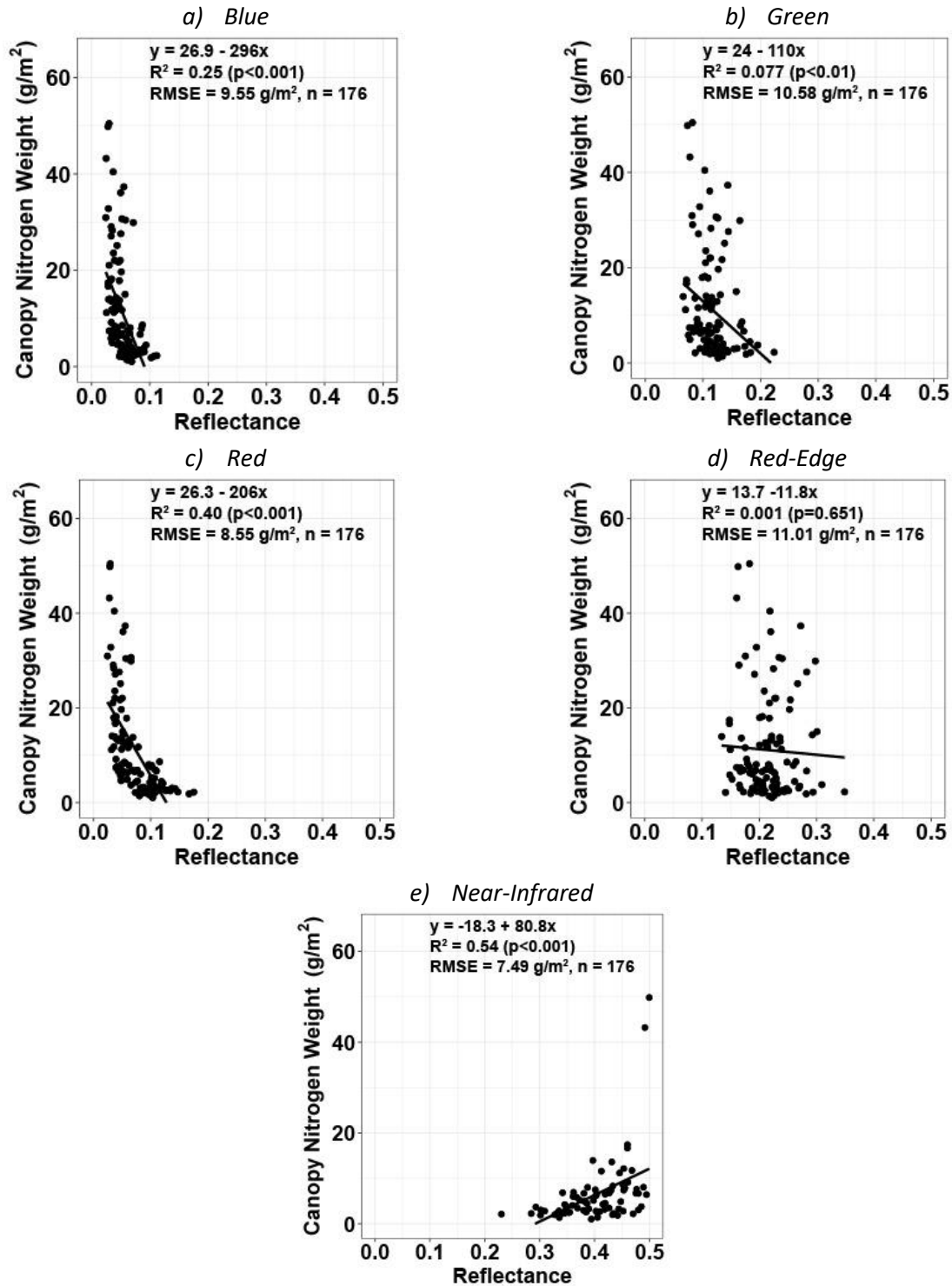
Figure 2-6. Relationship between canopy nitrogen weight and the MicaSense simulated reflectance in a) the blue MicaSense band, b) the green MicaSense band, c) the red MicaSense band, d) the red edge MicaSense band, and e) the near-infrared MicaSense band for the corn fields

As have already been showed in Liu et al. (2016), it is better to use reflectances in more than one band for estimating crop nitrogen. One way to do so is to use vegetation indices, which are algebraic combinations of multispectral reflectances. Among all the vegetation indices we tested, GNDVI produced a consistent lower R^2 value (around 0.50) than the other vegetation indices for both wheat and corn fields when considering the entire growing season (Figures 2-7 and 2-8). Song (2016) also found that GNDVI begins to saturate in the later growing season of wheat crops because of leaf senescence, while presenting a great relationship with nitrogen in the early/mid growing season. However, their relationship between GNDVI and nitrogen was evidently a non-linear exponential relationship. NDVI also presented low R^2 values because of the saturation of the relationship occurring in the later growing season for both the wheat and corn fields. This is not in agreement with Hansen and Schjoerring, (2003) who found that NDVI is the most useful index to predict nitrogen of wheat using a linear regression model, but these authors only considered the early growing stages of the wheat crops. Predicting crop nitrogen using NDVI was unreliable in this study because of the saturation of the relationship which leads to a large variation in the prediction of nitrogen in the later growing stage without increasing the NDVI values. This saturation effect was also present for the models with GNDVI, MTVI2 and Red Edge NDVI. The VIs that showed saturation in our study was also present in Frels et al. (2018), where the authors used a linear regression for each growing stage on wheat as a non-linear relationship was present when they combined all of the dates. However, the authors did not produce RVI in their list of VIs used in the study to estimate nitrogen status. Overall, most studies using VIs to estimate nitrogen have avoided the use of linear regression throughout the growing season to avoid the saturation in the later growing stages. On the other hand, the objective of this study aimed at achieving a linear relationship over the entire growing season and incorporating RVI to estimate nitrogen using linear regression.

Furthermore, RVI gave a consistent high R^2 value for both crops at the individual field level (Table 2-5) (around 0.93 for Hetzell) and when the fields are combined in the four sampling dates (around 0.66 for wheat and 0.64 for corn) (Figures 2-7 and 2-8). Li et al. (2018) working on rice crops found that ratio vegetation indices like RVI performed better in linear relationships with nitrogen than NDVI. However, even with RVI, our

relationship started to get weaker for the corn fields at the end of the growing stage (tasselling stage) because of a greater variation in the canopy nitrogen weight (Figure 2-8). This could be due to the leaf nitrogen content (%) remaining constant at the later growing stage, while the biomass increases. Figure 2-9 presents the best performing fields for both wheat and corn's RVI relationship with canopy nitrogen weight. Individually, both Hetzell and Crandell shows the best linear relationship with RVI and canopy nitrogen weight throughout the growing season.

Table 2-5. Statistics of the RVI-based regression models (*) as a function of the field and the number of dates sampled

Crop	Field	Dates Sampled	Regression Model	R^2	n
Wheat	Bale	3	y = -0.08014 + 0.22515(RVI)	0.80	24
		4	y = -0.701 + 0.351(RVI)	0.75	32
	Hetzell	3	y = -0.2222 + 0.13567(RVI)	0.93	24
		4	y = -0.625 + 0.368(RVI)	0.86	32
	McColl	3	y = -0.36529 + 0.26080(RVI)	0.71	48
		4	y = -2.0454 + 0.4218(RVI)	0.58	64
Corn	Crandell	3	y = 0.9639 + 0.7627(RVI)	0.70	24
		4	y = -2.583 + 1.54(RVI)	0.83	32
	Jack North	3	y = 0.1099 + 1.0457(RVI)	0.67	24
		4	y = -0.944 + 1.37(RVI)	0.81	32
	Paul	3	y = 1.8619 + 0.4708(RVI)	0.37	24
		4	y = -1.0225 + 1.0604(RVI)	0.71	32

(*) All the models are significant at p-value< 0.001
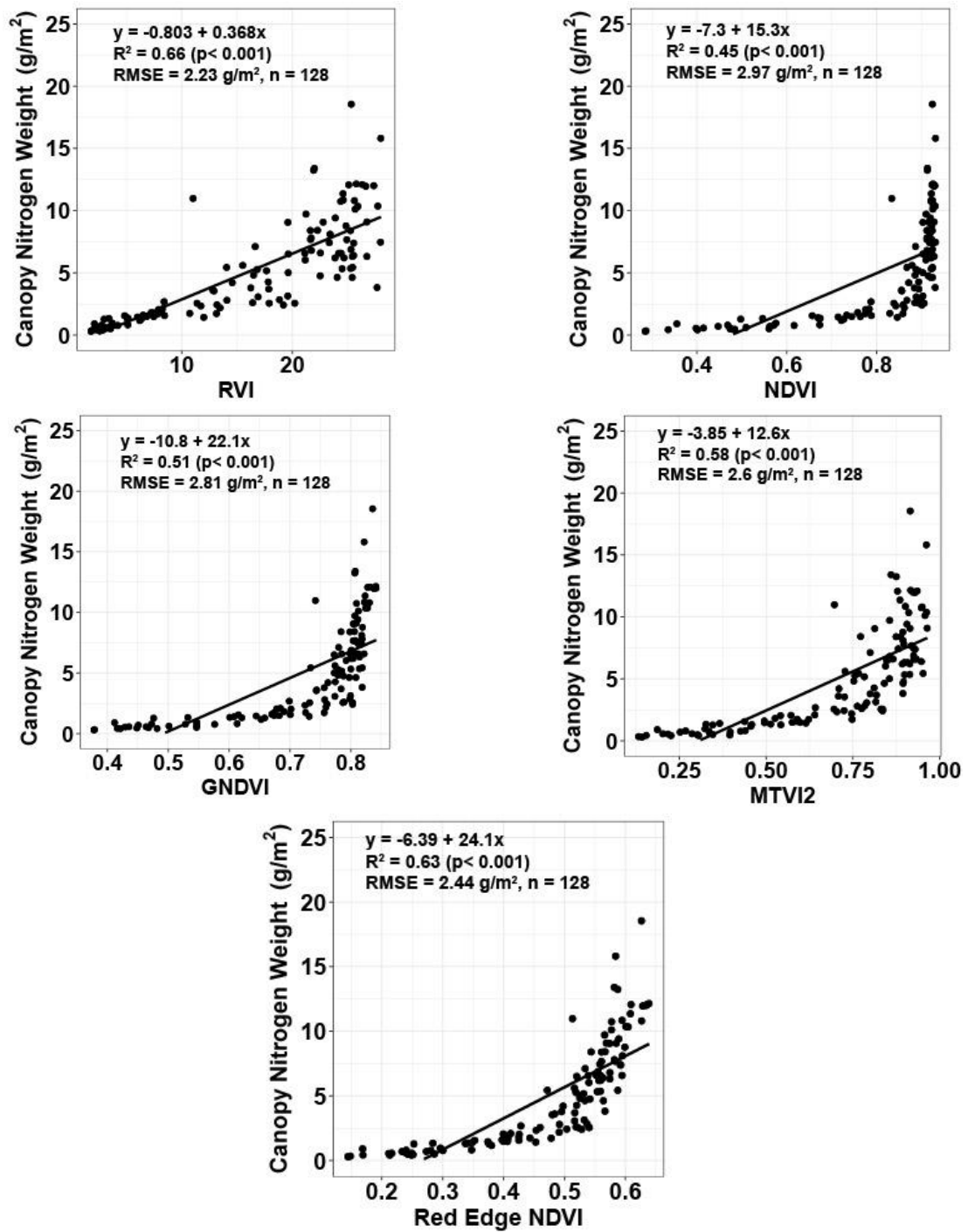
Figure 2-7. Relationship between canopy nitrogen weight and the five vegetation indices used in the study for the wheat fields combined
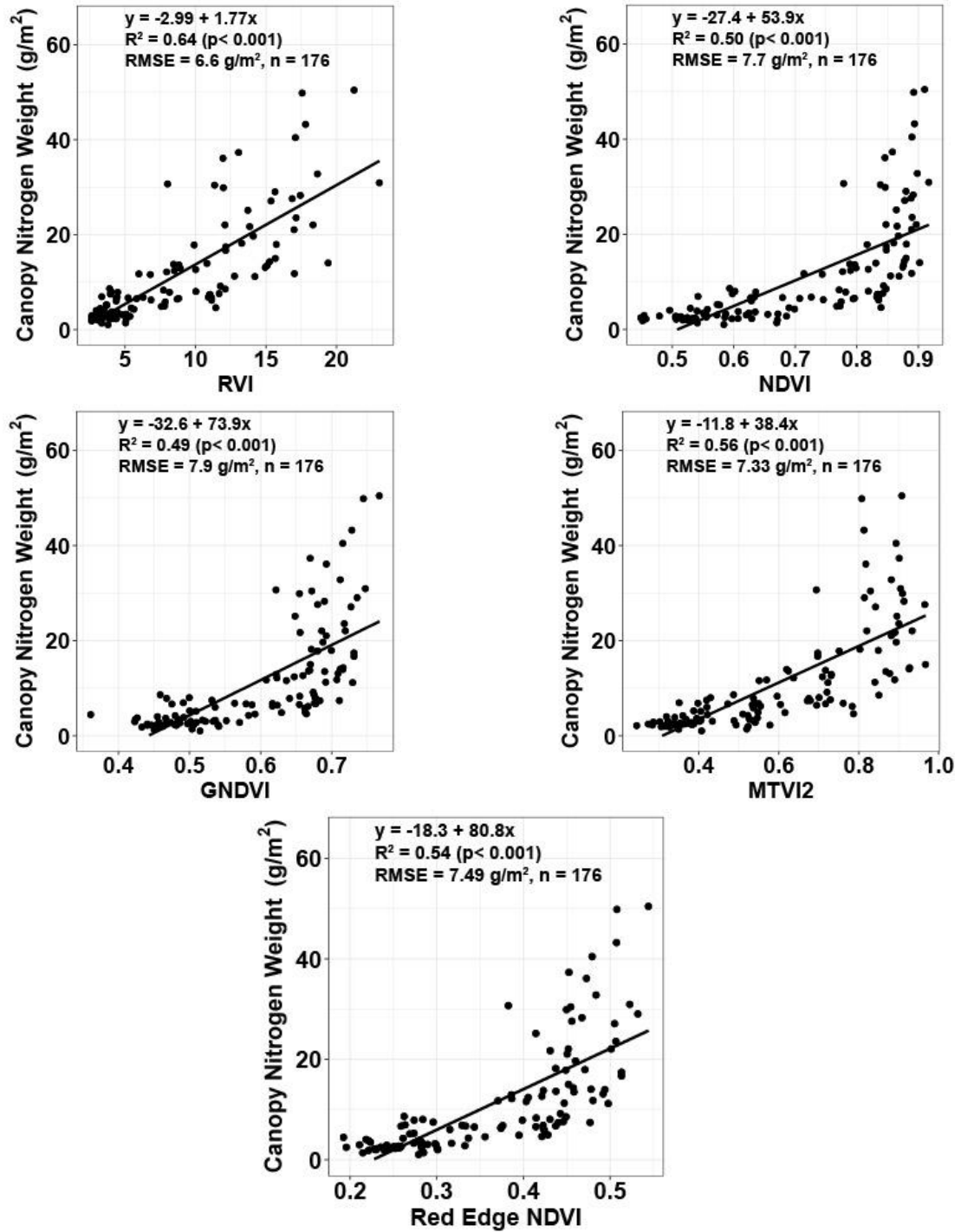
Figure 2-8. Relationship between canopy nitrogen weight and the five vegetation indices used in the study for the corn fields combined

a) Hetzell (wheat)

b) Crandell (corn)

y = -0.2222 + 1.3567x
R² = 0.93 (p<0.001)
RMSE = 1.31 g/m²
n = 24

y = -2.583 + 1.54x
R² = 0.83 (p<0.001)
RMSE = 4.13 g/m²
n = 32
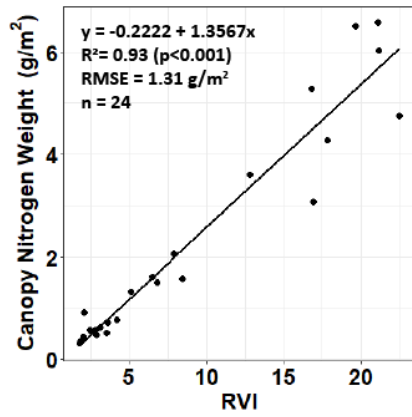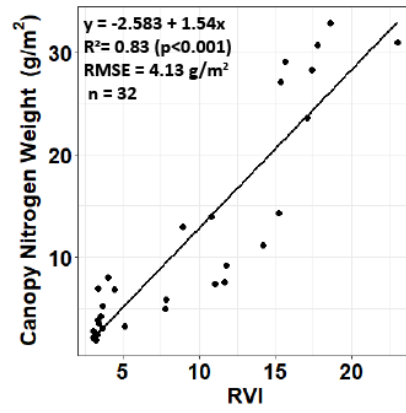
c) Bale (wheat)

d) Jack North (corn)

y = -0.08014 + 0.22515x
R² = 0.80 (p<0.001)
RMSE = 1.75 g/m²
n = 24

y = -0.944 + 0.1.37x
R² = 0.81 (p<0.001)
RMSE = 4.23 g/m²
n = 32
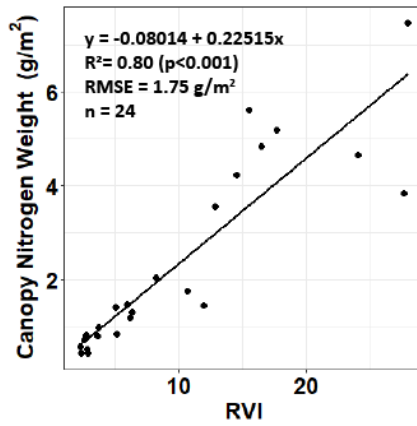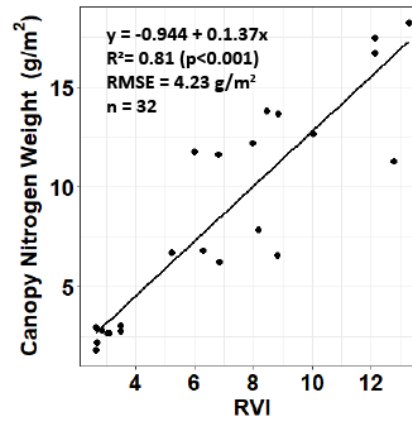
e) McColl (wheat)

f) Paul (corn)

y = -0.36529 + 0.2608x
R² = 0.71 (p<0.001)
RMSE = 1.98 g/m²
n = 48

y = -1.0225 + 1.0604x
R² = 0.71 (p<0.001)
RMSE = 5.31 g/m²
n = 32
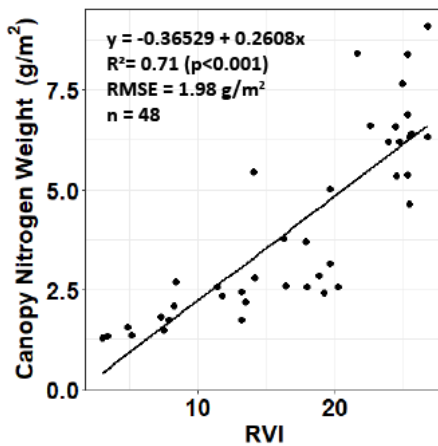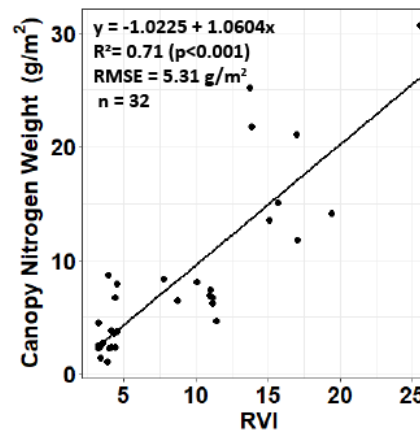
Figure 2-9. Relationship between canopy nitrogen weight and RVI of the best performing fields for (a) wheat and (b) corn

However, as shown in Table 2-5, when the data of the fourth sampling date was removed from the regression computation, the relationship between canopy nitrogen weight and RVI was not better for the corn fields. This could suggest that data of four sampling dates is needed for establishing a good relationship with canopy nitrogen weight for corn fields. Adding data from a fifth sampling date could be not useful given the saturation of the relationship. By contrast, for the wheat crops, Table 2-5 shows that removing the last sample date improved the linear model for all the fields, indicating that data from the first three sampling dates in the growing season is enough to model canopy nitrogen weight with vegetation indices. This may be due to the shift of nitrogen from the wheat's stem and leaves towards the grain, which is accompanied by the decline of leaf nitrogen content (%) at the last sampling date (Yang et al. 2019). Simultaneously due to the rapid biomass increase, the last sampling date may have caused a large variation of nitrogen, therefore weakening the relationship with vegetation indices as seen in Figure 2-4a.

A multiple linear regression was also established using stepwise regression on all the vegetation indices. However, the related VIF was higher than 10, which indicates a high degree of multicollinearity between the input vegetation indices, because some vegetation indices are well correlated along with each other, as shown in Table 2-6.

Table 2-6. Correlation matrix between the vegetation indices used in this study

	RVI	NDVI	GNDVI	MTVI2	RE_NDVI
RVI	1.0000				
NDVI	1.0000	1.0000			
GNDVI	0.9598	0.9598	1.0000		
MTVI2	0.9726	0.9726	0.9441	1.0000	
RE_NDVI	0.9478	0.9478	0.9868	0.9397	1.0000

With the simple linear regression method, since RVI is the best correlated with the canopy nitrogen weight among all the vegetation indices, it was used with the UAV imagery to derive a canopy nitrogen weight predictive map using Equation 5. However, before applying Equation 5, it is necessary to assess if the RVI values computed with the UAV imagery are similar as the RVI values computed with the ground-based reflectance data.

2.3.2 Comparison between the RVI computed with the UAV imagery and the ground measurements

In order to assess whether the RVI computed with the UAV imagery is similar as the RVI computed with the ground-based reflectance data, RVI values computed with the UAV imagery acquired with the MicaSense multispectral camera was correlated with the RVI values computed with the MicaSense-simulated reflectance data for all the fields combined (Figure 2-10). The resulting regression line has a R^2 of 0.94 and RMSE of 2.09. We can conclude that the RVI values derived from the UAV imagery are similar to those derived from the ASD data.
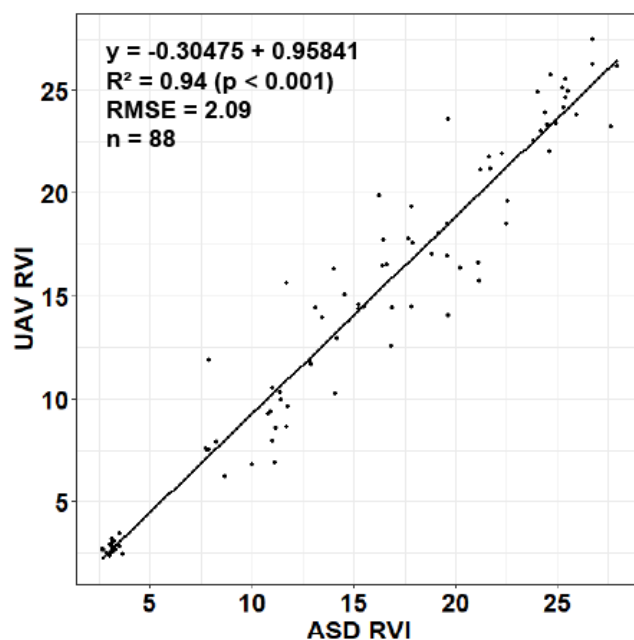
Figure 2-10. Comparison between ASD-measured and UAV-measured RVI for both crops combined

2.3.3 Crop Nitrogen Weight Predictive Map

RVI's linear regression equations of Table 2-5 were entered in the ArcMap Raster Calculator to compute canopy nitrogen weight for each image pixel and produce a canopy nitrogen weight map for each field and each UAV imagery date (Figure 2-11). From the resulting images, canopy nitrogen weight values were extracted around the sample points and compared to the measured ones. We achieved a RMSE ranging from

0.66 g/m^2 to 2.68 g/m^2 depending on the field, crop and the date of acquisition (Table 2-7). The UAV imagery acquired on June 7th over the corn field at Jack North achieved the best result with a RMSE of 0.66 g/m^2 using the linear regression model with four dates. The UAV imagery acquired on May 24th over the wheat field at McColl achieved the worst result with a RMSE of 2.68 g/m^2 using the linear regression model with four dates. The RMSE significantly increased for the wheat fields in Bale and McColl after removing data of the 4th sample date from the regression model. For all three corn fields in this study, the RMSE decreased when removing data of the 4th sample date from the regression model, indicating that data from a minimum of four sampling dates is needed for estimating canopy nitrogen weight in corn.
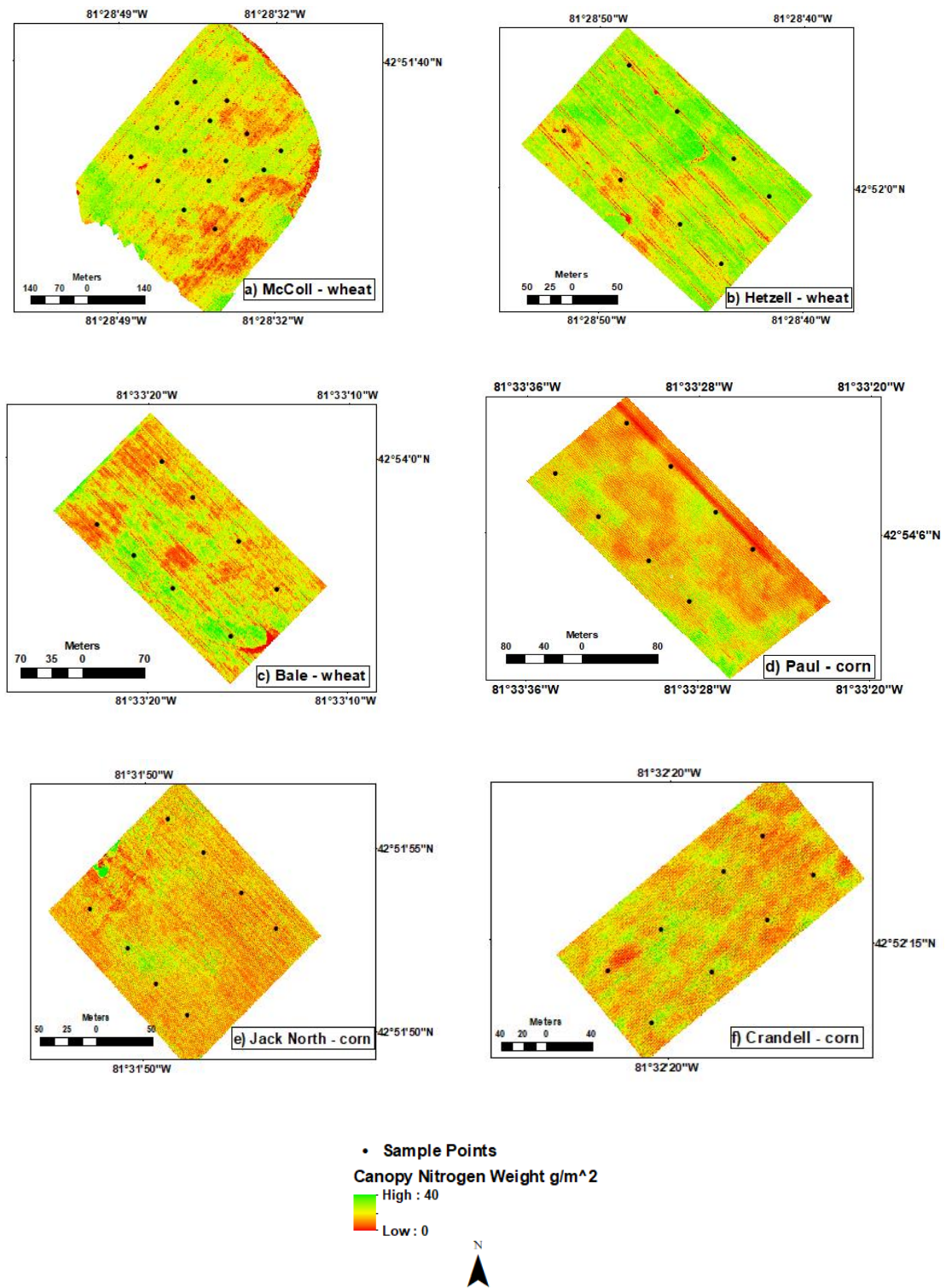
Figure 2-11. Canopy nitrogen weight prediction map with sample points derived from the RVI image on May 24th, 2018 for the following wheat fields a) McColl, b) Hetzell and c) Bale fields and on June 7th, 2018 for the following corn fields: d) Paul, e) Jack North, and f) Crandell fields

Table 2-7. RMSE for the RVI-based estimation model of canopy nitrogen weight (g/m²) when applied to the UAV imagery as a function of the crop, the field, the date of image acquisition, and the number of dates used for deriving the estimation model

Crop	Field	Date of the image	# of dates in the regression model	RMSE (g/m²)
Wheat	Bale	May 24	3	1.11
			4	2.36
	Hetzell	May 24	3	1.45
			4	1.50
		June 7	3	N/A[*]
			4	2.22
	McColl	May 17	3	1.09
			4	2.08
		May 24	3	0.95
			4	2.68
Corn	Crandell	June 7	3	1.21
			4	1.05
		June 19	3	3.01
			4	1.83
	Jack North	June 7	3	0.83
			4	0.66
	Paul	June 19	3	2.21
			4	2.09

*The UAV acquisition date was outside the dates of the linear regression model

On the resulting maps of Figure 2-11, high and low canopy nitrogen weight zones are displayed in green and in red respectively. This colour scheme allows distinguishing easily both zones, particularly in the case of the Hetzell field (Figure 2-11b). Similarly, the field of Figure 2-11d has a dark red strip zone that corresponds to bare soil as shown on Figure 2-1. Given that all the images have a spatial resolution of 5 cm per pixel, the images can easily allow detecting bare soil areas in both the wheat and corn fields and can easily separate the crop from the soil in the corn fields. Bagheri et al. (2013) found a very poor spatial distribution of nitrogen levels in their nitrogen prediction maps on corn, indicating that it is necessary to use variable-rate nitrogen application (applying certain amount of nitrogen fertilizer for each area) for differentiating nitrogen zones. However,

the authors used a 15 m resolution ASTER multispectral satellite imagery, making the spatial resolution inadequate for spatial differentiation of nitrogen status.

2.4   Conclusion

In this study, various vegetation indices were used in linear regression models to predict canopy nitrogen weight of wheat and corn fields using UAV MicaSense images. First, the models were established with MicaSense band-related vegetation indices derived from ASD spectra. The best vegetation index to estimate canopy nitrogen weight throughout the growing season is RVI for both the wheat and corn crops. RVI had a consistent linear relationship with canopy nitrogen weight throughout the entire growing season for the wheat crop, while the corn crop had a significant increase variation in nitrogen at the end of the growing season. However, the model was still able to predict nitrogen with high accuracy on the image gathered on June 7th (RMSE = 0.66 g/m^2). For the wheat fields, the best prediction was achieved with the image acquired May 24th (RMSE = 0.95 g/m^2).

The UAV nitrogen prediction maps can also detect spatial nitrogen variations within the fields. In practice, these results could be useful for farmers on retrieving quick information about the field's nitrogen status. Farmers will be able to know exactly which parts of their fields are deficient or excess in the amount of nitrogen present. Ultimately, this will lead to a much more efficient spraying program for the farmers as they will know precisely how much nitrogen is needed to use with their GPS-enabled nitrogen fertilizer. The study uses data acquired over Ontario fields in 2018 and there is the need to test the method over other datasets to see how general the developed methodology can be.

2.5 References

AOAC 2006. AOAC Official Method 972.43, Microchemical determination of carbon, hydrogen, and nitrogen, automated method, in *Official Methods of Analysis of AOAC International*, 18th edition, Revision 1, Chapter 12, pp. 5-6, AOAC International, Gaithersburg, MD.

Bagheri, N., Ahmadi, H., Alavipanah, S. K., & Omid, M. (2013). Multispectral remote sensing for site-specific nitrogen fertilizer management. *Pesquisa Agropecuária Brasileira*, 48(10), 1394-1401.

Barnes, E. M., Clarke, T. R., Richards, S. E., Colaizzi, P. D., Haberland, J., Kostrzewski, M., & Lascano, R. J. (2000). Coincident detection of crop water stress, nitrogen status, and canopy density using ground based multispectral data. In *Proceedings of the Fifth International Conference on Precision Agriculture, Bloomington, MN, USA* (Vol. 1619), 1-15.

Botha, E. J., Leblon, B., Zebarth, B. J., & Watmough, J. (2010). Non-destructive estimation of wheat leaf chlorophyll content from hyperspectral measurements through analytical model inversion. *International Journal of Remote Sensing,* 31(7), 1679-1697.

Clevers, J. G., & Gitelson, A. A. (2013). Remote estimation of crop and grass chlorophyll and nitrogen content using red edge bands on Sentinel-2 and-3. *International Journal of Applied Earth Observation and Geoinformation*, 23, 344-351.

Clevers, J. G., & Kooistra, L. (2011). Using hyperspectral remote sensing data for retrieving canopy chlorophyll and nitrogen content. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(2), 574-583.

Curran, P. J., Dungan, J. L., & Gholz, H. L. (1990). Exploring the relationship between reflectance red edge and chlorophyll content in slash pine. *Tree physiology*, 7(1-2-3-4), 33-48.

Deng, L., Mao, Z., Li, X., Hu, Z., Duan, F., & Yan, Y. (2018). UAV-based multispectral remote sensing for precision agriculture: A comparison between different cameras. *ISPRS Journal of Photogrammetry and Remote Sensing*, 146, 124-136.

Fox, J., & Weisberg, S. (2018). An R companion to applied regression. *Sage publications.*

Frels, K., Guttieri, M., Joyce, B., Leavitt, B., & Baenziger, P. S. (2018). Evaluating canopy spectral reflectance vegetation indices to estimate nitrogen use traits in hard winter wheat. *Field Crops Research*, 217, 82-92.

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

Gitelson, A. A. (2013). Remote estimation of crop fractional vegetation cover: the use of noise equivalent as an indicator of performance of vegetation indices. *International Journal of Remote Sensing*, *34*(17), 6054-6066.

Gnyp, M. L., Panitzki, M., Reusch, S., Jasper, J., Bolten, A., & Bareth, G. (2016). Comparison between tractor-based and UAV-based spectrometer measurements in winter wheat. In *Proceedings of the 13th International Conference on Precision Agriculture, Monticello, IL, USA* (Vol. 31).

Hamel, M. & Dorff, E. (2015). Corn: Canada's third most valuable crop. *Statistics Canada.* Retrieved from https://www150.statcan.gc.ca/n1/pub/96-325-x/2014001/article/11913-eng.htm

Hansen, P. M., & Schjoerring, J. K. (2003). Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sensing of Environment*, 86(4), 542-553.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning (Vol. 112, pp. 100-102). New York: springer.

Jay, S., Maupas, F., Bendoula, R., & Gorretta, N. (2017). Retrieving LAI, chlorophyll and nitrogen contents in sugar beet crops from multi-angular optical remote sensing: Comparison of vegetation indices and PROSAIL inversion for field phenotyping. *Field Crops Research*, 210, 33-46.

Jones, H. G., & Vaughan, R. A. (2010). *Remote sensing of vegetation: principles, techniques, and applications*. Oxford university press.

Jordan, C. F. (1969). Derivation of leaf area index from quality of light on the forest floor. *Ecology*, 50(4), 663-666.

Kelcey, J., & Lucieer, A. (2012). Sensor correction of a 6-band multispectral imaging sensor for UAV remote sensing. *Remote Sensing*, 4(5), 1462-1493.

Larsen, R. J. (2012). *Winter-Hardy Spring Wheat Breeding: Analysis of Winter x Spring Wheat Germplasm and the Development of Selection Tools* (Doctoral dissertation). University of Guelph.

Le Maire, G., François, C., Soudani, K., Berveiller, D., Pontailler, J. Y., Bréda, N., & Dufrêne, E. (2008). Calibration and validation of hyperspectral indices for the estimation of broadleaved forest leaf chlorophyll content, leaf mass per area, leaf area index and leaf canopy biomass. *Remote Sensing of Environment*, 112(10), 3846-3864.

Li, D., Wang, C., Liu, W., Peng, Z., Huang, S., Huang, J., & Chen, S. (2016). Estimation of litchi (Litchi chinensis Sonn.) leaf nitrogen content at different growth stages using canopy reflectance spectra. *European Journal of Agronomy*, 80, 182-194.

Li, F., Miao, Y., Feng, G., Yuan, F., Yue, S., Gao, X., & Chen, X. (2014). Improving estimation of summer maize nitrogen status with red edge-based spectral vegetation indices. *Field Crops Research*, 157, 111-123.

Li, S., Ding, X., Kuang, Q., Ata-Ul-Karim, S. T., Cheng, T., Liu, X., & Cao, Q. (2018). Potential of UAV-based active sensing for monitoring rice leaf nitrogen status. *Frontiers in Plant Science*, 9, 1834.

Liao, C., Wang, J., Dong, T., Shang, J., Liu, J., & Song, Y. (2019). Using spatio-temporal fusion of Landsat-8 and MODIS data to derive phenology, biomass and yield estimates for corn and soybean. *Science of The Total Environment*, 650, 1707-1721.

Liu, Y., Cheng, T., Zhu, Y., Tian, Y., Cao, W., Yao, X., & Wang, N. (2016). Comparative analysis of vegetation indices, non-parametric and physical retrieval methods for monitoring nitrogen in wheat using UAV-based multispectral imagery. In *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS'16),* Beijing, China, July, pp. 7362-7365.

Maes, W. H., & Steppe, K. (2019). Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture. *Trends in plant science*, *24*(2), 152-164.

Maresma, Á., Ariza, M., Martínez, E., Lloveras, J., & Martínez-Casasnovas, J. A. (2016). Analysis of vegetation indices to determine nitrogen application and yield prediction in maize (*Zea mays* L.) from a standard UAV service. *Remote Sensing*, 8(12), 973.

Muñoz-Huerta, R., Guevara-Gonzalez, R., Contreras-Medina, L., Torres-Pacheco, I., Prado-Olivarez, J., & Ocampo-Velazquez, R. (2013). A review of methods for sensing the nitrogen status in plants: advantages, disadvantages and recent advances. *Sensors*, 13(8), 10823-10843.

Nasrallah, A., Baghdadi, N., El Hajj, M., Darwish, T., Belhouchette, H., Faour, G., ... & Mhawej, M. (2019). Sentinel-1 Data for Winter Wheat Phenology Monitoring and Mapping. *Remote Sensing*, *11*(19), 2228.

Olson, D., Chatterjee, A., Franzen, D. W., & Day, S. S. (2019). Relationship of drone-based vegetation indices with corn and sugarbeet yields. *Agronomy Journal*, 111(5), 2545-2557.

Pix4D Documentation (2020). How to improve the outputs of dense vegetation areas? *Pix4D.* https://support.pix4d.com/hc/en-us/articles/202560159-How-to-improve-the-outputs-of-dense-vegetation-areas

Pricope, N. G., Mapes, K. L., Woodward, K. D., Olsen, S. F., & Baxley, J. B. (2019). Multi-Sensor Assessment of the Effects of Varying Processing Parameters on UAS Product Accuracy and Quality. *Drones*, *3*(3), 63.

Team, R. C. (2019). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. URL https://www.R-project.org/.

Raparelli, E., & Bajocco, S. (2019). A bibliometric analysis on the use of unmanned aerial vehicles in agricultural and forestry studies. *International Journal of Remote Sensing*, 40, 1-14.

Rouse, J.W., Jr., Haas, R.H., Schell, J.A., & Deering, D.W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. *NASA Goddard Space Flight Center 3d ERTS-1 Symposium*, Greenbelt, United States, January 1974.

Schlemmer, M., Gitelson, A., Schepers, J., Ferguson, R., Peng, Y., Shanahan, J., & Rundquist, D. (2013). Remote estimation of nitrogen and chlorophyll contents in

maize at leaf and canopy levels. *International Journal of Applied Earth Observation and Geoinformation*, 25, 47-54.

Sofonia, J., Shendryk, Y., Phinn, S., Roelfsema, C., Kendoul, F., & Skocaj, D. (2019). Monitoring sugarcane growth response to varying nitrogen application rates: A comparison of UAV SLAM LiDAR and photogrammetry. *International Journal of Applied Earth Observation and Geoinformation*, 82, 101878. (*in press*)

Song, Y. (2016). Evaluation of the UAV-based multispectral imagery and its application for crop intra-field nitrogen monitoring and yield prediction in Ontario. *Electronic Thesis and Dissertation Repository.* 4085. University of Western Ontario.

Walsh, O. S., Shafian, S., Marshall, J. M., Jackson, C., McClintick-Chess, J. R., Blanscet, S. M., & Walsh, W. L. (2018). Assessment of UAV based vegetation indices for nitrogen concentration estimation in spring wheat. *Advances in Remote Sensing*, *7*(02), 71-90.

Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J., & Reynolds, J. M. (2012). 'Structure-from-Motion'photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology*, *179*, 300-314.

Wickham, H. (2016). ggplot2: elegant graphics for data analysis. *Springer.*

Xie, Q., Dash, J., Huang, W., Peng, D., Qin, Q., Mortimer, H., & Dong, Y. (2018). Vegetation indices combining the red and red edge spectral information for leaf area index retrieval. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(5), 1482-1493.

Yang, B., Wang, M., Sha, Z., Wang, B., Chen, J., Yao, X., & Zhu, Y. (2019). Evaluation of aboveground nitrogen content of winter wheat using digital imagery of unmanned aerial vehicles. *Sensors,* 19(20), 4416.

Yao, X., Jia, W., Tian, Y., Ni, J., Cao, W., & Zhu, Y. (2013). Comparison and intercalibration of vegetation indices from different sensors for monitoring above-ground plant nitrogen uptake in winter wheat. *Sensors*, *13*(3), 3109-3130.

Zarco-Tejada, P. J., Miller, J. R., Mohammed, G. H., Noland, T. L., & Sampson, P. H. (2002). Vegetation stress detection through chlorophyll a+ b estimation and fluorescence effects on hyperspectral imagery. *Journal of environmental quality*, *31*(5), 1433-1441.

Zhang, C., & Kovacs, J. M. (2012). The application of small unmanned aerial systems for precision agriculture: a review. *Precision Agriculture*, 13(6), 693-712.

Zhao, B., Duan, A., Ata-Ul-Karim, S. T., Liu, Z., Chen, Z., Gong, Z., ... & Ning, D. (2018). Exploring new spectral bands and vegetation indices for estimating nitrogen nutrition index of summer maize. *European journal of agronomy*, *93*, 113-125.

Zheng, H., Li, W., Jiang, J., Liu, Y., Cheng, T., Tian, Y., & Yao, X. (2018). A comparative assessment of different modelling algorithms for estimating leaf nitrogen content in winter wheat using multispectral images from an Unmanned Aerial Vehicle. *Remote Sensing*, 10(12), 20-26

Chapter 3

3 Using Linear Regression, Random Forests, and Support Vector Machine with Unmanned Aerial Vehicle Multispectral Images to Predict Canopy Nitrogen Weight in Corn

3.1 Introduction

Precision agriculture (PA) is a farming management technique that requires detailed information on crop status. One of the important crop status indicators is the crop nitrogen (N) weight because nitrogen is the main plant nutrient needed for producing the chlorophyll, which has a direct impact on the plant photosynthesis and thus on crop growth and yield (Lemaire et al. 2004). Therefore, there is a need to understand the spatial distribution of crop nitrogen for better use of fertilizers. Ultimately, this information leads to a better yield among the crops and reduces costs for the farmer by matching the fertilizer supply to its demand (Xie et al. 2018; Liu et al. 2016). Traditionally, farmers used to rely on historical weather data, such as precipitation and temperature, and their past experiences, such as crop yields, to make decisions on their fertilizer operations for the upcoming season (Shahhosseini et al. 2019). Today, there have been many advances in technology, such as remote sensing data and machine learning algorithms, that can potentially aid farmers' decision making on fertilizer application. Remote sensing-based methods used to measure crop nitrogen are typically better than the traditional ground-based methods. Ground-based methods require intensive field data collection, which can be time-consuming, destructive, and limited to a small spatial area, making it impractical for fast and efficient results. Remote sensing-based methods are required for most agricultural fields in Canada, given that they can reach up to hundreds of ha in size. Remote sensing methods are non-destructive, can cover large spatial areas, and have been increasingly used for crop monitoring in precision agriculture. Crop monitoring based on remote sensing data can use spaceborne or airborne images, but these types of high-resolution imagery are either costly or difficult to obtain (Raparelli & Bajocco, 2019). Also, they have limited applicability in precision agriculture because of the too coarse temporal and spatial resolutions of the

imagery (Zheng et al. 2018; Zhang & Kovacs, 2012). Alternatively, free-of-charge Sentinel-1 Synthetic Aperture Radar (SAR) imagery could be used to monitor nitrogen status but does not provide enough spatial resolution (10 m) for precise small-scaled applications (Nasrallah et al. 2019). Bagheri et al. (2013) also found it difficult to differentiate levels of nitrogen status in corn fields using a 15 m Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) multispectral satellite imagery. An alternative is to use Unmanned Aerial Vehicle (UAV) imagery. The development and application of UAV imagery have increased in the past decade, filling in gaps between satellite imagery, aerial photography, and field samples (Kelcey & Lucieer, 2012; Yang et al. 2017). Image acquisition with UAV can be deployed quickly and repeatedly, at a low cost, and with greater flexibility (Maresma et al. 2016; Zha et al. 2020). The temporal resolution of UAVs is superior to the satellite and aerial photography platforms, which is easily defined by the user (Turner et al. 2013; Harwin & Lucieer, 2012). The low cost of UAVs could also be convinced without the use of purchasing ground control points (GCPs) (Turner et al. 2013).

The focus of this study is to test whether UAV multispectral imagery can be used to retrieve the crop nitrogen status over corn fields from a perspective of precision nitrogen fertilization. The spatial and temporal variations of the images we acquired were determined in order to match the crop requirements of nitrogen as closely as possible. One common remote sensing technique to estimate nitrogen content at the canopy level is the Radiative Transfer Model (RTM), which estimates the chlorophyll or nitrogen content by describing the interaction between the sun's light and the crop canopy. An example of an RTM is the PROSAIL model, which uses various parameters at the leaf and canopy level and can be mathematically inverted to retrieve both chlorophyll and nitrogen content from spectral data (Lemaire et al. 2004; Clevers & Kooistra, 2011; Clevers & Gitelson, 2013; Hansen & Schjoerring, 2003; Botha et al. 2010) . Other sets of remote sensing methods are empirical methods, such as machine learning algorithms, or simple/multiple-linear regression to retrieve crop nitrogen from canopy spectral data (Clevers & Kooistra, 2011). This paper used the second set of methods because empirical estimation has already been proven to be easier in estimating nitrogen than the PROSAIL model, given that an RTM requires the calibration of numerous parameters (Jay et al.

2017). In this study, we tested three empirical methods: (1) linear regression, (2) Random Forests, and (3) support vector regression (SVR) to statistically relate spectral measurements and canopy-level crop nitrogen weight in the case of two corn fields located in southwest Ontario, Canada. Many studies on predicting nitrogen were conducted in a controlled experimental condition and predicted nitrogen values at the leaf level (Chen et al. 2010; Tian et al. 2011; Mipkokasap et al. 2012; Fan et al. 2019), while few studies have been performed on real field conditions (Li et al. 2008). As already shown in Liu et al. (2016), it is better to use more than one spectral variable for estimating crop nitrogen. Therefore, multiple linear regression, Random Forests, and SVR was used in this paper to predict canopy nitrogen using a variety of spectral variables.

The first regression method is a traditional regression approach that has two major assumptions (multicollinearity and linear relationship). Most studies using UAV imagery have used linear regressions to estimate nitrogen status from vegetation indices (VIs) (Zha et al. 2020). The latter two approaches are machine learning techniques that have been advanced in recent years and are unaffected by the multicollinearity and linear relationship assumptions. In addition, they can handle overfitting (Zha et al. 2020; Kayad et al. 2019; Belgiu et al. 2016). Overfitting is a common problem in machine learning where the models produced perform poorly on unseen data. Random Forests has become popular recently within the remote sensing research community for classification and regression purposes. The variable importance plot provided by the Random Forests algorithm is very successful in identifying the most relevant input data in the model (Belgiu et al. 2016; Osco et al. 2019). Therefore, we used the Random Forests variable importance plots to identify the most important variables for canopy nitrogen estimation and adjusted the model parameters accordingly. Random Forests modelling has been found to perform very well out of all the non-parametric methods in various studies to monitor nitrogen content in wheat (Liu et al. 2016), rice (Zha et al. 2020), and citrus trees (Osco et al. 2019). Random Forests was also the most accurate machine learning approach for monitoring corn yield, which is related to crop nitrogen status (Kayad et al. 2019). Support vector modelling has also been popular in estimating nitrogen status due to its strong ability to decorrelate the input variables and can work with non-linear

relationships (Xiong et al. 2019; Wang et al. 2017; Karimi et al. 2008). SVR modelling has been found to predict nitrogen concentration with high accuracy in bokchoy (Xiong et al. 2019), wheat (Wang et al. 2017), and corn (Karimi et al. 2008).

Ideally, the relationship between the spectral data and the canopy nitrogen weight should be linear because there should be only one canopy nitrogen weight estimation for each level of input data, given that the relationship will be used to calibrate an N fertilizer sprayer that needs to have an exact determination of the crop N fertilization level. However, when the canopy becomes dense, the relationship can saturate and become non-linear. This is the case when using some VIs, such as the standard normalized difference vegetation index (NDVI), green NDVI, red edge NDVI, and modified transformed vegetation index 2 (MTVI2), which has been shown to saturate at high canopy densities (Lee et al. 2020). Therefore, this study tested the ability of Random Forests and SVR to work with non-linear saturated data, particularly when combining datasets throughout the growing cycle.

The best performing model was applied to the UAV imagery for mapping crop nitrogen content at the field level. Deng et al. (2018) found that mounting narrowband multispectral cameras on UAVs acquire images with far better quality than broadband multispectral cameras. Therefore, the objectives of this study are to (i) generate machine learning models to predict crop nitrogen weight in corn fields using UAV multispectral imagery, (ii) determine which individual MicaSense spectral bands and VIs have the most influence on the Random Forests decision tree when predicting nitrogen, (iii) generate nitrogen prediction maps by applying the best model to the entire UAV image and analyse whether the UAV images are able to detect any spatial variation of nitrogen within the fields. The study evaluates three different modelling approaches for predicting nitrogen in corn over different dates and growth stages using UAV multispectral imagery. The best algorithm not only fills the research gap between monitoring nitrogen and UAVs, but also has practical meaning for future modelling study designs. Ultimately, these images should be given to farmers in highly accurate, quick, and timely manner field information for their precision nitrogen fertilization management.

3.2   Materials and Methods

3.2.1   Study Area

The study site is located in Melbourne, Ontario, Canada (Figure 3-1). This region is in the humid continental climate zone in Canada and the summers are typically hot and humid, with an average temperature of 27 °C during the fieldwork month of July 2019. The study site is dominated by agricultural land with very little urban use. The closest large urban centre is London, Ontario, just 30 km east of the study site. Corn is generally planted in May just before the summer and typically harvested in late October/early November once the crop is dried and the starch content is high.

We collected field data from two corn (*Zea mays)* fields (labelled as JJ and Susan) in the summer of 2019. A total of six sampling dates were collected for corn, with at least a week in between each sampling date. Both corn fields together were roughly 60 ha in size and situated directly across from each other. The study fields are smaller than the average agricultural field sizes in Ontario (100 ha) (Mailvaganam, 2017). However, such field sizes are large enough to avoid weak relationships in using spectral data and related VIs to predict nitrogen because smaller fields can be more affected by bare soil surrounding the fields and the mixing effect of other nearby fields (Frels et al. 2018). Both fields were similar in topography and planted at the same time by the same grower, allowing the field data collected to be used in a single model.
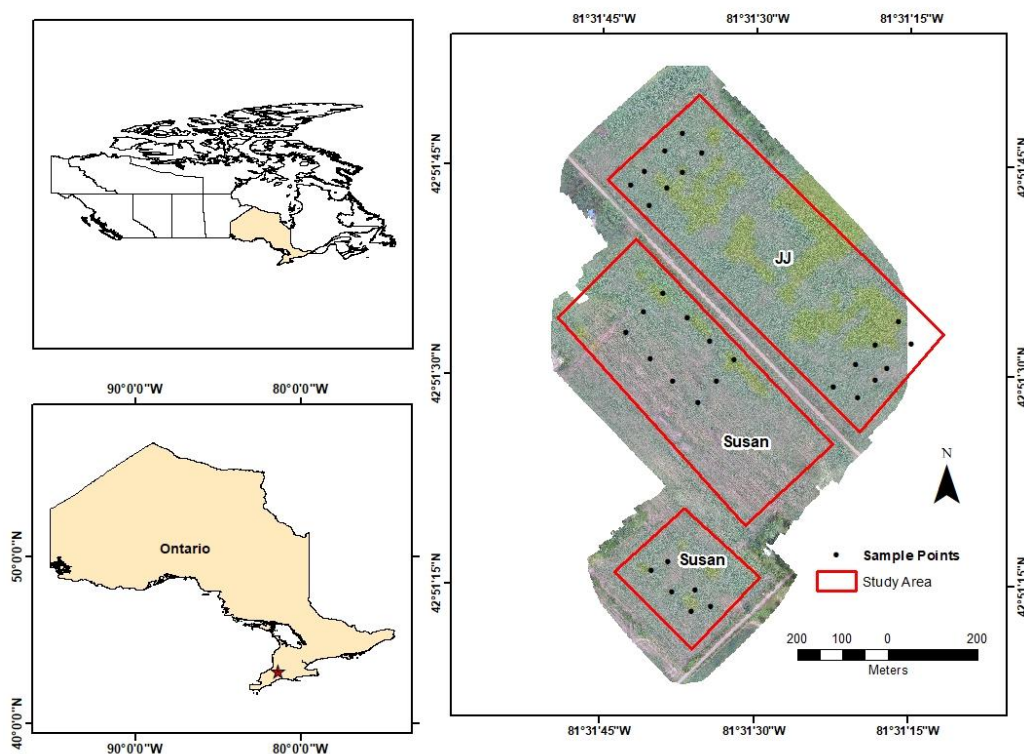
Figure 3-1. Study area of the corn fields (JJ and Susan) shown using red green blue (RGB) sensor mosaic imagery taken on July 18th, 2019 in Southwest Ontario, Canada. One corner of JJ is missing on the RGB image

3.2.2 Field Data

In-situ data were collected over 16 points on each corn field in June and July 2019. Whiteboard and red sticks were placed on each sample point, which was a ground control point to be identified on the UAV imagery for orthomosaicking. In-situ data included destructive biomass collection at each sampling point. Fresh biomass was measured in grams by gathering the fresh canopy in a 1 m^2 block around the sampling point and placed in large plastic bags for transport. Due to the intensive work and heavy weight of the fresh corn canopy, only two plants were taken per sample point and upscaled to the number of plants in the 1 m^2 block. The average row distance for corn was 75 cm and the fields typically had an average of 12–14 plants per 1 m^2 area. Following the fieldwork, biomass was weighed at a fresh stage in grams, then placed in an 80 °C oven for 36–48 hours. Dried biomass weight (scaled at g/m^2) was weighed then sent to A&L Canada

Laboratories for plant tissue analysis. The oven-dried samples were ground into a powder form and passed through a 1 mm sieve. The leaf nitrogen content (expressed as a percentage) was then measured using the Laboratory Equipment Company (LECO) FP628 nitrogen/protein analyser that uses the total nitrogen combustion method (AOAC, 2006).

3.2.3   UAV Imagery

UAV imagery collection is optimal when collected weekly and immediately before field data collection as crop physiology and soil structure change over time (Figure 3-2). UAV flights were performed before the field data collection to ensure that the biomass was present at the sampling points in the imagery. UAV flights were performed on June 26th, 3 July, July 10th, July 18th, and July 31st using a MicaSense RedEdge narrowband camera mounted on a Dà-Jiāng Innovations (DJI) Matrice 100 quadcopter (Table 3-1). The growth stages of the corn are also described in Table 3-1, where V(n) represents the vegetation stage and the amount of leaves present, excluding the initial emergence leaf. Flights were flown by pilots of A&L Canada Labs Inc. over the entire fields, flown in a zigzag route and 50 m in height, and 80–85% overlap. Our past studies (Lee et al. 2020) have indicated that a high overlap is required for corn as the canopy becomes very dense in the middle-late growing season.

Figure 3-2. Ground photos of corn (2019) taken on June 26th (left) and July 31st (right)

Table 3-1. Summary of UAV flight acquisition for JJ and Susan in the study (2019)

UAV Flight Date	Weather	Relative Humidity	Corn Growth Stage
26 June	27 °C, Sunny	69%	V1/V2
3 July	30.1 °C, Sunny	58%	V3/V4
10 July	30.8 °C, Sunny	45%	V5
18 July	29.3 °C, Sunny	70%	V6
23 July	24.3 °C, Sunny	73%	V7/V8
31 July	25.3 °C, Sunny	60%	V9

3.2.4 UAV Image Processing

Figure 3-3 shows the flowchart of the methodology in this study. The UAV images gathered in the summer of 2019 were processed using a photogrammetry software called Pix4Dmapper (Pix4D SA, Lausanne, Switzerland). Pix4Dmapper was used to generate an orthomosaic image of each field by stitching hundreds of different images captured during the same flight into one single 2D image and corrected for perspective. Pix4D uses the technique called Structure from Motion (SfM) and has been well-suited for UAV data as it combines images from multiple angles (Harwin & Lucieer, 2012). The mosaicked images were then exported to individual (.tif) files. The mosaic images were automatically radiometrically corrected in Pix4D with a spatial resolution of 5 cm/pixel. The mosaic images were scaled to 15 cm/pixel to reduce the computing time given the high number of pixels for a single field. For example, the Susan field scaled at 5cm/pixel is approximately 529 million pixels per each layer.
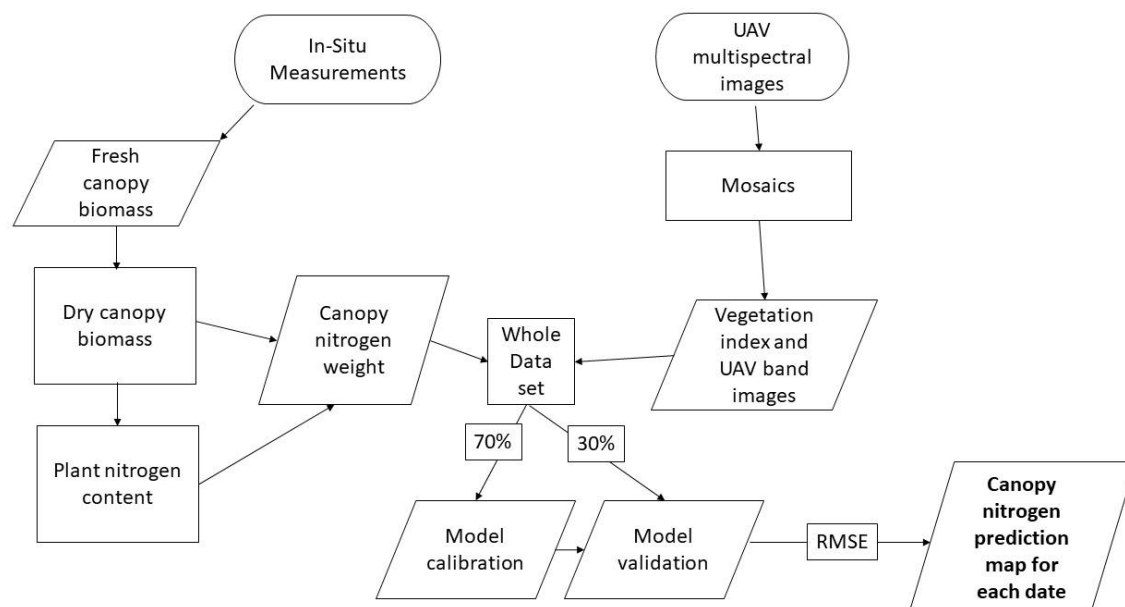
Figure 3-3. Flowchart of the methodology (Chapter 4)

3.2.5 Vegetation Indices

Reflectance values of the sample points were extracted from the MicaSense mosaic images. Five reflectance values of each sample point were acquired by the MicaSense RedEdge camera in the following bands: (1) blue, (2) green, (3) red, (4) red edge, and (5) near-infrared (Table 3-2) (Figure 3-4).

Table 3-2. Spectral characteristics of the 5 MicaSense bands

Band #	Name	Band Range (nm)	Centre Wavelength (nm)	Bandwidth (nm)
1	Blue	465-485	475	20
2	Green	550-570	560	20
3	Red	663-673	668	10
4	Red edge	712-722	717	10
5	Near-infrared (NIR)	820-860	840	40
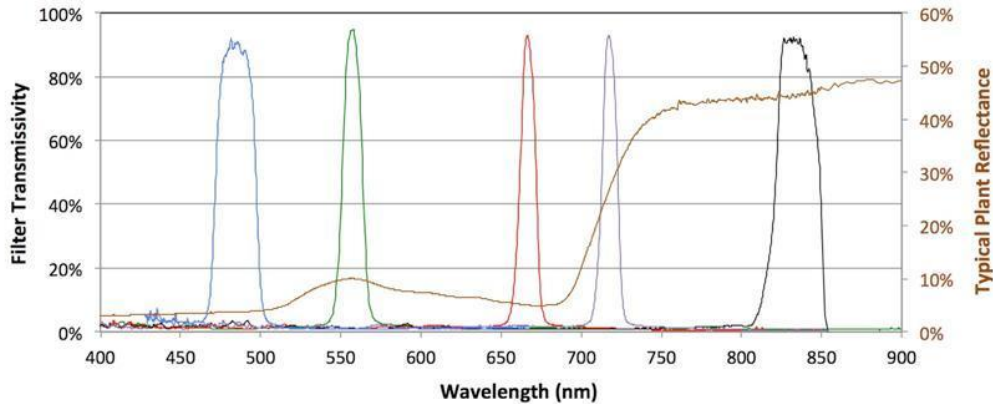
Figure 3-4. Spectral response curve for the MicaSense RedEdge Camera. The brown line is a typical reflectance profile of a green vegetation canopy. Figure derived from Tagle Casapia, 2017

These surface reflectance values were then used to compute 29 VIs that are commonly used to estimate canopy nitrogen variables (Table 3-3). These indices are intended to enhance the contribution of the optical properties of the vegetation on the total spectral response of the canopy. Therefore, VIs attempt to correct any confounding factors such as reflectance of soil backgrounds in a crop, particularly at the early stages of the growth cycle (Clevers & Kooistra, 2011).

Table 3-3. Vegetation indices used in the study

Index [1]	Formula [2]	Authors
BNDVI	$(NIR - BLUE)/(NIR + BLUE)$	Wang et al. 2007
CI_Green	$(NIR/GREEN) - 1$	Gitelson et al. 2003
CI_RE	$(NIR/REDEDGE) - 1$	Gitelson et al. 2003
EVI	$\dfrac{2.5(NIR - RED)}{(NIR + 6RED - 7.5BLUE) + 1}$	Huete et al. 2002
GARI	$\dfrac{NIR - [GREEN\text{-}1.7(BLUE\text{-}RED)]}{NIR + [GREEN - 1.7(BLUE - RED)]}$	Gitelson et al. 1996
GDVI	$NIR - GREEN$	Tucker et al. 1979
GNDVI	$(NIR - GREEN)/(NIR + GREEN)$	Gitelson; Merzlyak, 1998
GOSAVI	$(NIR - GREEN)/(NIR + GREEN + 0.16)$	Sripada et al. 2005
GRVI	$(GREEN - RED)/(GREEN + RED)$	Sripada et al. 2006
GSAVI	$1.5[(NIR - GREEN)/NIR + GREEN + 0.5)]$	Sripada et al. 2005
ISR	RED/NIR	Fernandes et al. 2003
MCARI	$[(REDEDGE\text{-}RED) - 0.2(REDEDGE\text{-}GREEN)] *(REDEDGE/RED)$	Daughtry et al. 2000
MCARI1	$1.2[2.5(NIR\text{-}RED) - 1.3(NIR - GREEN)]$	Haboudane et al. 2004
MCARI2	$\dfrac{3.75(NIR - RED) - 1.95(NIR - GREEN)}{\sqrt{(2NIR + 1^2)} - (6NIR - 5\sqrt{RED}) - 0.5}$	Haboudane et al. 2004
MSAVI	$\dfrac{2NIR + 1 - \sqrt{(2NIR + 1)^2 - 8(NIR - RED)}}{2}$	Qi et al. 1994
MSR	$\dfrac{(NIR/RED)\text{-}1}{\sqrt{\left(\dfrac{NIR}{RED}\right) + 1}}$	Chen, 1996
MTVI1	$1.2[1.2(NIR - GREEN) - 2.5(RED - GREEN)$	Haboudane et al. 2004
MTVI2	$\dfrac{1.8(NIR - GREEN)\text{-}3.75(RED - GREEN)}{\sqrt{(2NIR + 1)^2 - 6(NIR - 5\sqrt{RED}) - 0.5}}$	Haboudane et al. 2004
NDVI	$(NIR - RED)/(NIR + RED)$	Rouse et al. 1974
OSAVI	$1.6[(NIR - RED)/(NIR + RED + 0.16)]$	Rondeaux; Steven; Baret, 1996
RDVI	$\dfrac{(NIR\text{-}RED)}{\sqrt{(NIR + RED)}}$	Roujean; Breon, 1995
RE_NDVI	$(NIR - REDEDGE)/(NIR + REDEDGE)$	Gitelson; Merzyak, 1994
RGBVI	$(GREEN^2 - BLUE*RED)/(GREEN^2 + BLUE*RED)$	Bendig et al. 2015
RVI	NIR/RED	Jordan, 1969
SARVI	$\dfrac{(1 + 0.5)(NIR - y)}{(NIR + y + 0.5)}$	Rondeaux; Steven; Baret, 1996
SAVI	$(NIR - RED)(1 + 0.5)/(NIR + RED + 0.5)$	Huete, 1988
TVI	$\dfrac{120(NIR - GREEN) - 200(RED - GREEN)}{2}$	Broge; Leblanc, 2001
VARI	$(GREEN - RED)/(GREEN + RED - BLUE)$	Gitelson et al. 2002

| WDRVI | $(0.2*NIR - RED)/(0.2*NIR + RED)$ | Gitelson, 2013 |

[1] BNDVI = blue normalized difference vegetation index; CI_Green = chlorophyll index green; CI_RE = chlorophyll index red edge; EVI = enhanced vegetation index; GARI = green atmospherically resistant index; GDVI = green difference vegetation index; GNDVI = green normalized difference vegetation index; GOSAVI = green optimized soil adjusted vegetation index; GRVI = green ratio vegetation index; GSAVI = green soil adjusted vegetation index; ISR = infrared simple ratio; MCARI = modified chlorophyll absorption ratio index; MSAVI = modified soil adjusted vegetation index; MSR = modified simple ratio; MTVI = modified triangular vegetation index; NDVI = normalized difference vegetation index; OSAVI = optimized soil adjusted vegetation index; RDVI = renormalized difference vegetation index; RE_NDVI = red edge normalized difference vegetation index; RGBVI = red green blue vegetation index; RVI = ratio vegetation index; SARVI = soil atmospherically resistant vegetation index; SAVI = soil adjusted vegetation index; TVI = triangular vegetation index; VARI = visible atmospherically resistant index; WDRVI = wide dynamic range vegetation index; [2] BLUE = blue reflectance; GREEN = green reflectance; RED = red reflectance; REDEDGE = red edge reflectance; NIR = near-infrared reflectance; y = (RED−1)*(BLUE-RED).

3.2.6   Canopy Nitrogen Weight Estimation

To describe the canopy nitrogen status, we used the canopy nitrogen weight that is defined by Hansen and Schjoerring (2003) as follows:

$$CNW = \left(Nplants * \frac{Wd}{2}\right) * LNC \quad [7]$$

where CNW is the canopy nitrogen weight (g/m^2), N$_{plants}$ is the number of plants in the 1m^2 sampling point, Wd is the dry biomass weight (g/m^2) of two plants in the 1m^2 sampling point, and LNC is the leaf nitrogen content (%).

Equation (7) assumes that all the leaves from a sample gathered in the field contained the same amount of nitrogen. Canopy nitrogen weight (g/m^2) has the advantage of being a more absolute value, compared to plant or leaf nitrogen content (%), which is a relative value. Absolute values allow the ability to compare the results among fields and dates.

Previous studies have shown that estimating biochemical concentrations at the leaf level is difficult. Therefore, focusing on the canopy level is optimal (Clevers & Kooistra, 2011). Li et al. (2008) have also found that canopy nitrogen weight is more strongly correlated with spectral data than the other agronomic variables, such as Soil Plant

Analysis Development (SPAD) readings, plant nitrogen concentration%, nitrate content, and soil mineral N.

3.2.7   Canopy Nitrogen Weight Modelling

All modelling was performed using R programming language in R Studio (R Version 3.6.1) (R Core Team, 2019). The first modelling approach is the simple/multiple linear regression. We avoided using all the variables in the multiple linear regression to predict nitrogen as VIs are known to be highly intercorrelated with each other. Such multicollinearity between explanatory variables reduces the accuracy of the estimates of the regression coefficients (James et al. 2013). This makes the results of multiple linear regression difficult to interpret and unreliable with an increased number of variable inputs. Therefore, the linear regression model was established with the top six and top 12 most influential variables as determined by the Random Forests variable importance plot.

The second modelling approach used Random Forests. This is a decision tree nonparametric algorithm used for classification or regression. The algorithm selects a random number of samples from the training dataset chosen by the analyst. Afterward, the randomly chosen samples are used to develop a decision tree based on the most important variables. Trees are split at each node depending on the most contributing x_i (i^{th} explanatory variable) to y (response variable). For each prediction of $\hat{y}$ (predicted value of the response variable), it constructs a multitude of decisions trees and outputs the average value. Figure 3-5 shows an example of the decision tree modelling steps for the Random Forests model using the dataset. Hyper-tuning the parameters of Random Forests was unnecessary due to the results remaining unchanged after altering the number of trees of 500 and *mtry* parameter at default. The parameter *mtry* is the number of variables used for splitting at each tree node for decision tree learning. Random Forests in R defines the *mtry* by dividing the number of predictive variables divided by 3.
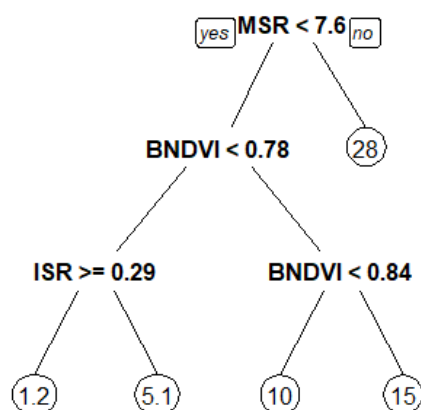
Figure 3-5. An example of a generalized decision tree model in Random Forests (R Studio) using all the spectral variables

The third modelling approach is SVR which is also a form of nonparametric modelling that defines boundaries in a high-dimensional space using a hyperplane. A hyperplane is a flat affine subspace of a dimension p-1, where p is the number of dimensions. In two dimensions, the hyperplane is a flat one-dimensional subspace (straight line) and splits the training data into different sections in a two-dimensional plot. The notion of (p-1) applies for any number of dimensions, while anything $p > 3$ is difficult to visualize (James et al. 2013). If the relationships of the data are nonlinear, SVR uses a nonlinear kernel function. We used the Radial Basis Kernel when performing SVR, which tricks the data into a higher-dimensional space to separate the data into different sections using the radial distance between the observations. Hyper-tuning the SVR model using a tune grid-search with the function tune() using different combinations of cost and gamma found that the best performing combinations were 2 and 0.5 for our dataset, respectively.

Random Forest decision trees and support vectors can work with non-linear relationships, whereas traditional linear regression models cannot. Most VIs on corn begin to saturate with nitrogen in the middle stage of the growing cycle, making the relationship more non-linear. Relationships that are non-linear between the dependent and independent variable(s) are not practical because they can lead to the prediction of multiple nitrogen

values for the same VI value. The relationship between the VI to the N status can be misleading if the best-fit function (R^2) is not linear because the sensitivity between the two will not be constant (Gitelson, 2013). Therefore, the information derived from the linear model can cause uncertainty to the growers to precisely spray fertilization due to the need of having unique nitrogen value for each VI value. Therefore, Random Forests and SVR were used to mitigate the uncertainty of non-linear VI values and canopy nitrogen weight, particularly with Random Forest's robustness to non-linear data (Louppe, 2014) and low variance in model prediction (Čeh et al. 2018).

Both Random Forest and SVR modelling were performed in R Studio using the "randomForest" (Liaw et al. 2002) and "e1071" (Meyer et al. 2018) packages, respectively. Linear regression modelling was performed in R Studio using the lm() function. The independent samples of 29 VIs and 5 individual MicaSense bands were then used to generate the linear, Random Forests and SVR models. The canopy nitrogen weight and VI values of July 3rd, July 10th, and July 18th were randomly split into a 70% calibration set and 30% for the validation set. The dates of July 3rd, July 10th, and July 18th were used for the modelling due to the availability of the entire dataset of both UAV and in-situ ground measurements for those dates. The calibration dataset was used to generate the models and the resulting models were compared with each other. The validation set was not used in the modelling but was used to test each modelling approach by using new datasets and avoiding overfitting. For both the calibration and validation datasets, the quality of the models was assessed using the R^2 and the Root Mean Square Error (RMSE). R^2 is measured from 0–1 and indicates how well the data fits the goodness of fit line and is calculated using the following equation:

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \qquad [8]$$

where y_i is the associated observed value in the dataset or formed in a vector as $y = [y_1, ..., y_n]^T$, $\hat{y}$ is the predicted value of the associated y_i, and $\bar{y}$ is the mean of the observed data.

RMSE measures how far on average the predicted values are from the measured ground truth values and is calculated using the following equation:

$$RMSE = \frac{\sum_{i=1}^{n}(P_i - O_i)^2}{n} \tag{9}$$

where P_i is the predicted canopy nitrogen weight value (g/m^2), O_i is the observed canopy nitrogen weight value (g/m^2), n is the number of observations, and i is the index of summation in increment of 1.

The model providing the lowest RMSE on the validation set was applied to the whole UAV images for predicting the spatial distribution of the canopy nitrogen weight in each field. To combine all the VI and MicaSense individual band images into a single data frame, the "raster" package (Hijmans, 2019) was used in R Studio. Individual VI and MicaSense band (.tif) files were imported into R Studio and combined using the stack() function. Afterward, the raster:predict() function was used to predict each pixel in the multi-layered (.tif) file using the best model. Finally, the writeRaster() function was used to generate the prediction map in (.tif) format, while the map characteristics were visualized using ArcMap.

3.3   Results

3.3.1   MicaSense Spectral Profile

Figure 3-6 shows the spectral profile of the MicaSense centre wavelengths for the dates of July 3rd, July 10th, and July 18th that were used in the modelling. The reflectance values of each point were averaged for both fields to represent the MicaSense band's reflectance for each date.
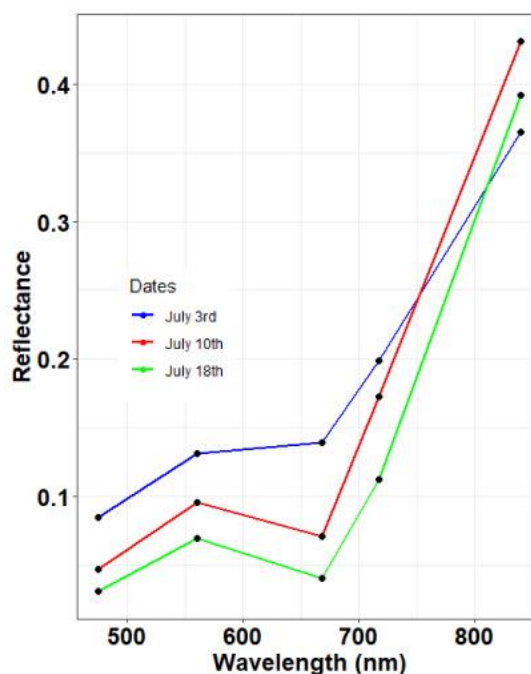
Figure 3-6. Spectral profile of the MicaSense reflectance using the centre wavelength for each band

3.3.2   Nitrogen Statistics

The leaf nitrogen content (%) in corn ranged between 2.24 % and 6.15 % throughout the growing season. The leaf nitrogen content had a decreasing trend in values, possibly due to the nitrogen contribution to the crop changes throughout the growing season (Table 3-4). Crop biomass is the cumulative production of plant photosynthesis throughout the growing season. Table 3-5 shows the summary statistics of the dried biomass weight (g/m^2).

Table 3-4. Summary statistics of the calibration and validation set for leaf nitrogen content %

Date	Growth Stage	Calibration Set			Validation Set		
		Min.	Max.	Average	Min.	Max.	Average
July 3rd	V3/V4	3.97	5.93	5.15	4.39	6.15	5.41
July 10th	V5	2.57	4.51	3.61	2.29	4.29	3.48
July 18th	V6	2.24	4.05	3.64	2.97	4.31	3.60

Table 3-5. Summary statistics of the calibration and validation set for dry biomass (g/m²)

		Calibration Set			Validation Set		
Date	Growth Stage	Min.	Max.	Average	Min.	Max.	Average
July 3rd	V3/V4	2.3	14.27	5.31	1.29	19.7	4.64
July 10th	V5	23.8	140.3	50.75	21.4	75.11	37.63
July 18th	V6	30.0	210.6	89.32	35.0	148.9	80.9

The canopy nitrogen weights (derived from Equation (7)) in corn presented a gradual increase of variation throughout the growing cycle (Figure 3-7). Figure 3-7 also shows that there is very little variation in canopy nitrogen weight in the early growing stage of the corn crops. By contrast, there is a larger variation in canopy nitrogen weights in the later growing season, due to the increase of biomass weight. One outlier is shown in Figure 3-7 but was not removed as it remained consistent throughout the growing cycle, indicating that this was not due to measurement error. Because the data collection and processing are the same for each sample point, the chances of measurement errors occurring on the same sample point three times are unlikely.
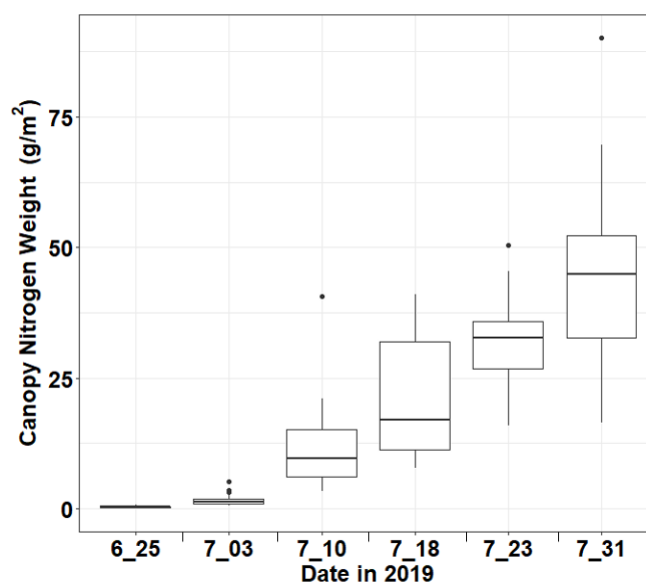
Figure 3-7. Box plot showing the variation of canopy nitrogen weight (g/m²) as a function of the date of field measurements in JJ and Susan corn field during the 2019 growing season. The dots on the graph represent outliers

3.3.3   Variable Importance Plot

Using the variable importance plot in R Studio, the best spectral variables that were important in the decision tree modelling are shown in (Figure 3-8). A large value of IncNodePurity indicates that the explanatory variables are an important predictor for canopy nitrogen. The red-edge band performed the worst out of all the individual MicaSense bands with canopy nitrogen. MSR performed the best out of all the VIs in the variable importance table, while GDVI and GARI had no weight.
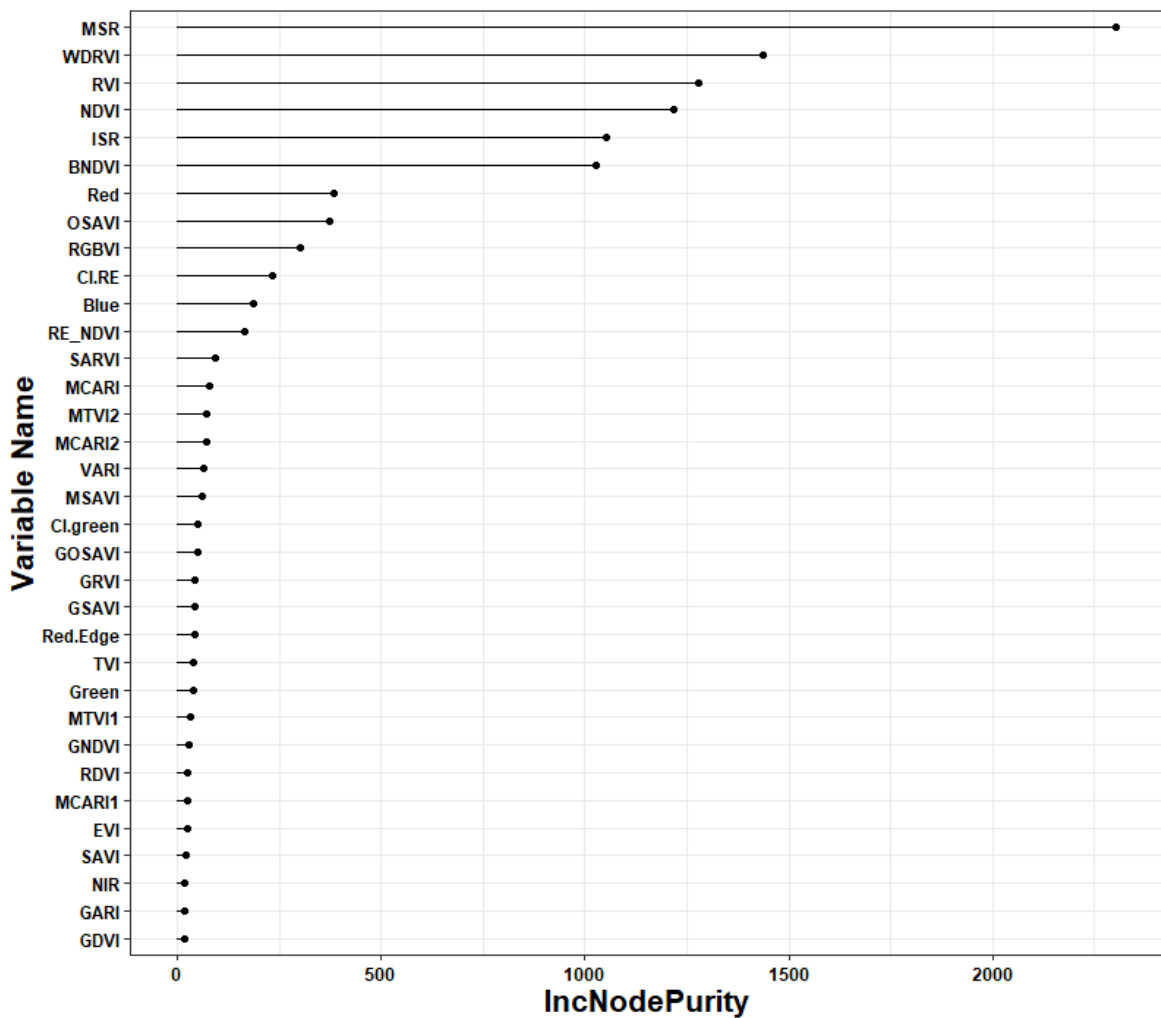
Figure 3-8. Variance Importance plot using the function varImpPlot() in R Studio. Higher IncNodePurity values indicate more impact on nitrogen

3.3.4   Calibration and Validation Models

Table 3-6 shows the statistics of the linear regression, Random Forests, and SVR applied to the calibration dataset. All computational processing times were quick (< 15 seconds), while computer hardware and larger sample sizes may factor processing speeds. Linear regression with 12 variables shows a great relationship with canopy nitrogen weight with ($R^2 = 0.87$) and an RMSE of 4.03 g/m^2. Multiple regressions of all the variable inputs were produced using the calibration data but the results are misleading due to the high degree of multicollinearity. None of the coefficients in the multiple linear regression using all the VIs were significant at $\alpha = 0.05$, indicating that the model was very sensitive to the multicollinearity present. Therefore, the model generated using multiple regression of all the variables was not applied to the validation dataset to avoid misleading/unreliable results. SVR performed well on the calibration set, while the 12-variable input model performed better than the 34-input variable model from the importance plot. This may be due to the higher degree of dimensions when adding more variables into the support vector model. Random Forests with all the variable combinations performed the best on the calibration set compared to the other two regression methods. An R^2 of 0.95 and RMSE of 2.25 g/m^2 were achieved (Figure 3-9a).

Table 3-6. Statistics for the calibration of canopy nitrogen model using various modelling approaches (n=63)[1]

Input Variables	Number of Variables	Model	RMSE (g/m²)	R^2
All VIs and 5 MicaSense bands	34	RF	2.25	0.95
MSR, WDRVI, RVI, NDVI, ISR, BNDVI, Red², OSAVI, RGBVI, CI_RE, Blue², RE_NDVI	12	RF	2.31	0.94
MSR, ISR, RVI, NDVI, BNDVI, WDRVI	6	RF	2.63	0.93
MSR, WDRVI, RVI, NDVI, ISR,	12	SVR	3.98	0.87
BNDVI, Red², OSAVI, RGBVI, CI_RE, Blue², RE_NDVI	12	Linear	4.03	0.87
All VIs and 5 MicaSense bands	34	SVR	4.08	0.87
MSR, ISR, RVI, NDVI, BNDVI, WDRVI	6	Linear	4.20	0.85
	6	SVR	4.28	0.85
MSR	1	Linear	4.41	0.83
All VIs and 5 MicaSense bands	34	Linear	9.14	0.58

[1] All models are significant at p-value < 0.001; [2] MicaSense individual band. RMSE, Root Mean Square Error.

R² = 0.94 (p < 0.001)
RMSE = 2.31 g/m²
n = 63

(a) Calibration

R² = 0.85 (p <0.001)
RMSE = 4.52 g/m²
n = 28
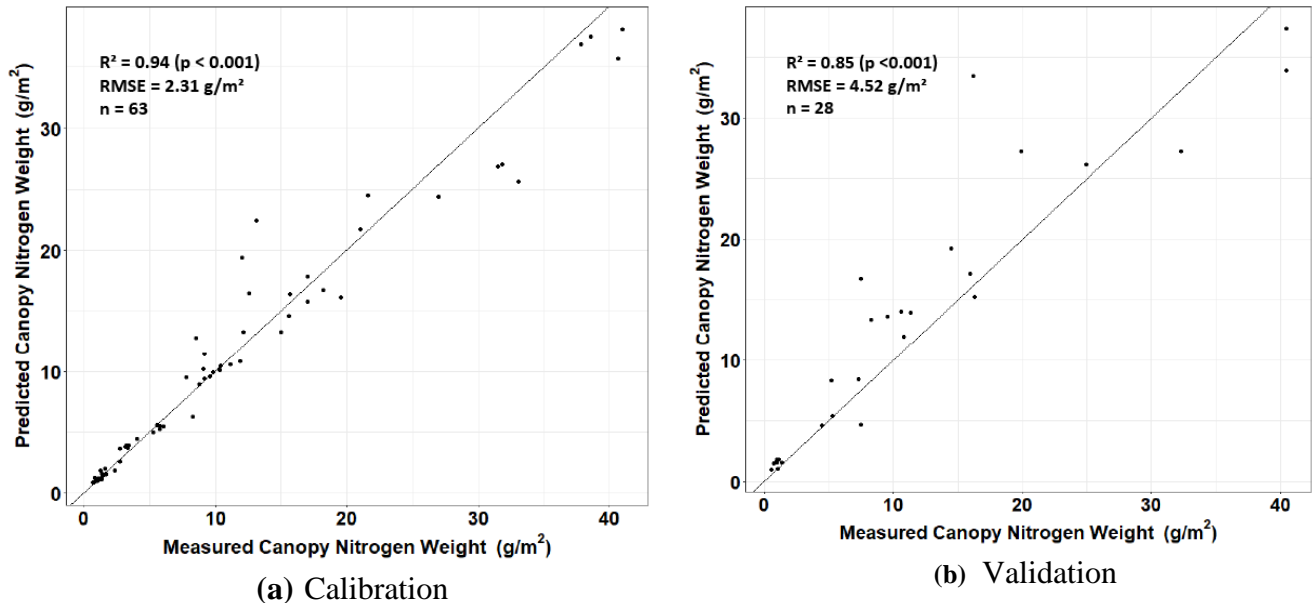
(b) Validation

Figure 3-9. Predicted versus measured canopy nitrogen weights when applying Random Forests model to the top 12 variables for (a) the calibration dataset and (b) validation dataset

When applying the models on the validation dataset, Table 3-7 also shows that Random Forest's top 12 and 34 variables performed the best out of all the other models. This is due to Random Forest's strong ability to avoid overfitting. Table 3-7 also shows that linear regression did not perform as well as the non-parametric Random Forests and SVR. This may be due to the nature of the non-linear relationship of several VIs on canopy nitrogen. The difference in RMSE between Random Forests models in Table 3-7 was only 0.01 g/m^2, making the model with 12 variables more realistic in terms of processing time. Figure 3-9b shows the predicted and measured canopy nitrogen weight on the validation dataset using the Random Forests model applied to the top 12 variables. The model with 12 variables was able to predict lower values of canopy nitrogen weight with very high accuracy but struggled to provide a high accuracy on predicting the higher values of canopy nitrogen weight. This may be due to the large variation of canopy nitrogen values on July 18th.

Linear regression performed the worst on the validation set compared to Random Forests and SVR, indicating that the calibration model showed some degree of overfitting and cannot work well compared to Random Forests. Interestingly, Table 3-7 shows that using all the variables performed marginally better than using all the top 12 variables in both the Random Forest and SVR models. The Random Forests importance plot was able to identify the variables with no or little effect on the model, reducing the processing time significantly when removing them. Removing the unused variables could also mitigate the errors caused in a higher dimensionality dataset. These results show that adding more independent variables does not necessarily mean that this will produce higher accuracy, and in fact might hurt the modelling performance.

Table 3-7. Statistics when applying various modelling approaches to the validation dataset (n = 28)[1]

Input Variables	Number of Variables	Model	RMSE (g/m^2)	R^2
All VIs and 5 MicaSense bands	34	RF	4.51	0.85
MSR, WDRVI, RVI, NDVI, ISR, BNDVI, Red[2] OSAVI, RGBVI, CI_RE, Blue[2], RE_NDVI	12	RF	4.52	0.85
All VIs and 5 MicaSense bands	34	SVR	4.58	0.84
MSR, WDRVI, RVI, NDVI, ISR, BNDVI, Red[2], OSAVI, RGBVI, CI_RE, Blue[2], RE_NDVI	12	SVR	4.74	0.84
MSR, WDRVI, RVI, NDVI, ISR, BNDVI, Red[2], OSAVI, RGBVI, CI_RE, Blue[2], RE_NDVI	12	Linear	4.78	0.85
MSR, ISR, RVI, NDVI, BNDVI, WDRVI	6	RF	5.21	0.83
MSR	1	Linear	5.47	0.82

[1] All models are significant at p-value < 0.001; [2] MicaSense individual band.

3.3.5 Crop Nitrogen Weight Predictive Map

As already shown in Figure 3-9b, there is a good agreement between the predicted and measured canopy nitrogen weight, particularly in the lower value ranges. Using our best model (Random Forests with 12 variables), we computed the canopy nitrogen weight for each image pixel with the 12 layers of VI and individual MicaSense bands in R Studio. Figure 3-10 shows the resulting canopy nitrogen prediction map of the UAV images of July 3rd, July 10th, and July 18th, 2019. The low and high canopy nitrogen weight zones are displayed in red and green, respectively. This color scheme allows the nitrogen level to be distinguished easily between the two colors, particularly on July 18th, in which the canopy nitrogen weight has a large variation. Given that all the images were downscaled to 15 cm per pixel, the images can still easily detect bare soil areas and can separate the crop from the soil in the corn fields. Most of the red pixels in the July 3rd image are mostly dominated by the soil, instead of the corn canopy (Figure 3-10a). However, as the plant height and density increases, it is harder to detect between the soil and the corn on July 10th, while showing variation in the canopy nitrogen level (Figure 3-10b). Both the low and high areas of canopy nitrogen weight on the field are also consistent throughout the three prediction images.

(a) July 3rd

(b) July 10th

(c) July 18th

Nitrogen Weight g/m^2
High : 33
Low: 0
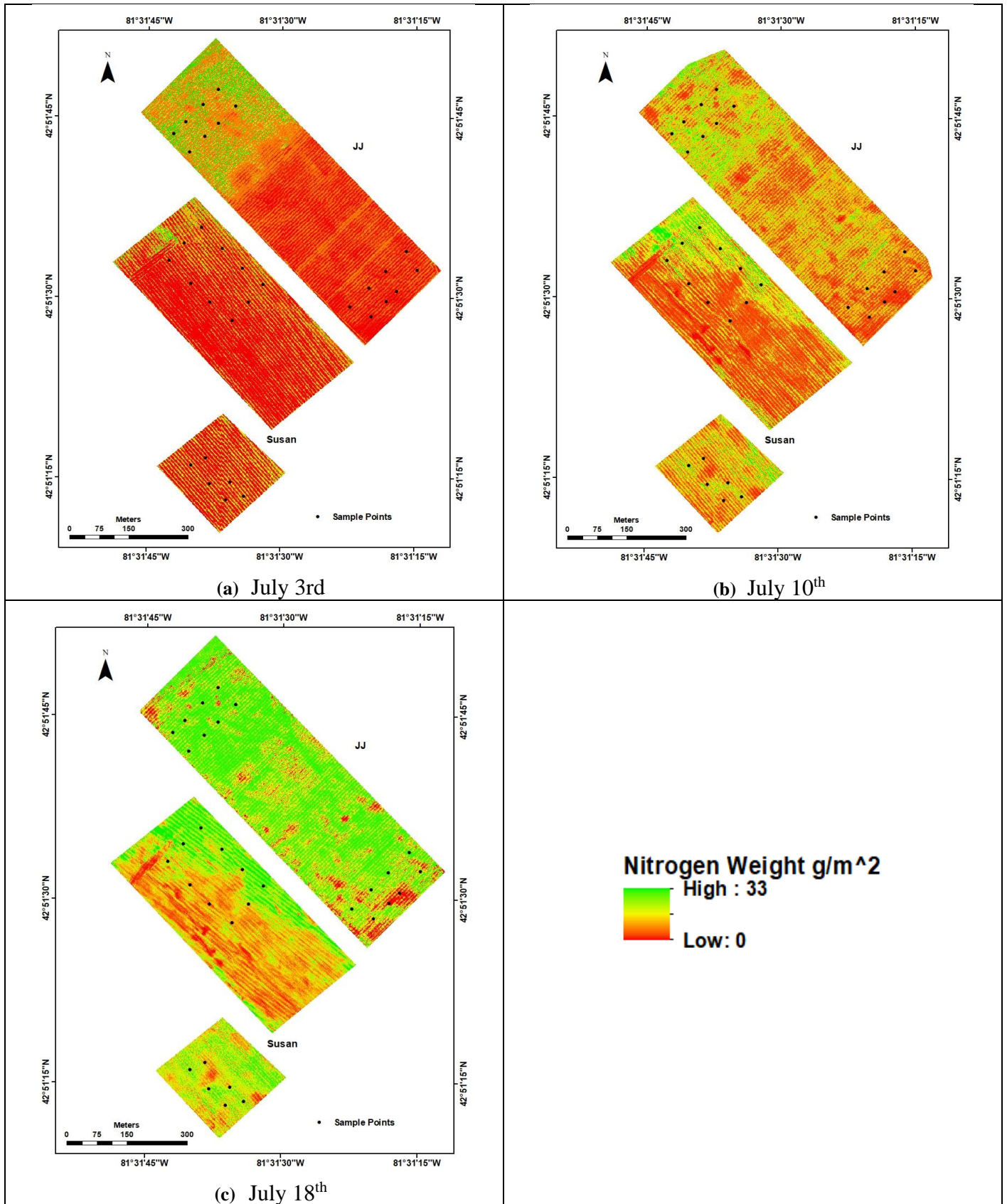
Figure 3-10. Canopy nitrogen prediction map derived when applying the RF model to the top 12 variable images for a) July 3rd, b) July 10th, and c) July 18th for the JJ and Susan fields. The locations of the ground sampling points are also given with a black dot

The RMSE of each nitrogen prediction map generated with the 12 variable Random Forests model were calculated using equation 9 (Table 3-8). The prediction of July 3rd canopy nitrogen values was very accurate at 0.62 g/m^2 RMSE. July 10th produced a high RMSE, possibly due to the single value of high canopy nitrogen weight from figure 3-7. In the later growing stage, July 18th produced an RMSE of 3.68 g/m^2.

Table 3-8. RMSE for the 12 variable Random Forests model applied to each UAV imagery

Date	Growth Stage	RMSE (g/m^2)
July 3rd	V3/V4	0.62
July 10th	V5	4.11
July 18th	V6	3.68

3.4 Discussion

Our study used five different MicaSense multispectral bands to derive various VIs to predict canopy nitrogen weight. Values of the in-situ canopy nitrogen weight saw an increase in variation on July 10th and July 18th. This is due to the rapid increase in biomass between the dates. This could be explained by the crop biomass variation increasing due to the factors that contribute to the crop's growth, such as the absorption, utilization, and transformation of solar energy; climate; and nutrient/water management (Yue et al. 2017; Campos et al. 2018).

Individual bands were tested to predict canopy nitrogen in all the models along with the VIs. Both the MicaSense green band and red-edge band performed poorly in the model compared to the other bands. The green wavelength is closely related to the leaf chlorophyll *a* and *b* contents, in which nitrogen is used for plant photosynthesis (Zhao et al. 2018). The poor performance in the model may be due to the chlorophyll saturating in the middle to the late growth stages, causing the crops to reflect the same amount of green wavelength. However, our results are not in agreement with Schlemmer et al. (2013) and Li et al. (2014), who observed a good relationship between the green reflectance and corn nitrogen weight. The red-edge spectral region is an interesting region, in which the position of the sharp change in reflectance (known as the red-edge position) is particularly known to be a sensitive indicator of leaf chlorophyll content (Jones & Vaughan, 2010; Zarco-Tejada et al. 2002; Curran et al. 1990). The red-edge position changes in the wavelength of 680–800 nm depending on the strength of the absorption of chlorophyll (Baranoski & Rokne, 2005). Therefore, the narrowband of 10nm in the MicaSense red-edge band may have not fully captured the red-edge position throughout the growing season, whereas another sensor with different band ranges could have captured it. A possible consideration in the future would be to fly two cameras simultaneously and compare the results. Furthermore, the red-edge reflectance has been also found to be significantly related to corn nitrogen weight in Schlemmer et al. (2013) and Li et al. (2014). Such difference in both green and red-edge reflectance can be explained by the fact that their study focused more on predicting nitrogen at individual growth stages, while our study considered all the growth stages in the model. All the top six VIs in the variable importance table use the near-infrared and red bands, indicating that they are both critical to the prediction of nitrogen. This is probably due to the chlorophyll absorption present in the red region and high reflectance of near-infrared energy for leaf development of healthy vegetation. However, the near-infrared band alone does not have a good relationship with canopy nitrogen weight, and therefore it must be included with other bands under the form of VIs. Interestingly, NDVI (the most commonly used VI in literature) did not perform as well as the top VIs in the variable importance plot. This is because NDVI is known to saturate with canopy nitrogen weight once the canopy of the crop becomes dense (Lee et al. 2020).

The use of the variable importance plot in Random Forests to eliminate spectral variable inputs was also performed in Osco et al. (2019) on predicting canopy nitrogen in citrus trees. The authors used the top five and ten variables of 33 spectral variables and found there was a slight decrease in the model performance. Similar to our results, Osco et al. (2019) found a decrease in performance relatively small, and the trade-off between the number of spectral indices used and obtained accuracy is something that should be considered. However, Osco et al. (2019) used a different list of VIs, but the application of the variable importance plot was the same. Random Forests performed poorly on corn in Fan et al. (2019) ($R^2 = 0.60$) compared to partial least squares regression ($R^2 = 0.80$) on the validation dataset. However, the authors used nitrogen content percentage at the leaf level and found weak correlation with most of the spectral variables, while our study incorporated our nitrogen values at the canopy level (g/m^2). Random Forests with spectral variables on the validation set performed much better at the canopy level ($R^2 = 0.85$), probably due to the spectral variables having a good correlation with nitrogen at the canopy level.

Our study on SVR modelling lines up with the results of Karimi et al. (2008). The authors found that SVR performed better and more consistent than its multiple linear regression counterpart in their study. The difference in results of Random Forests and SVR in our study are not too far apart in model performance. However, the concepts and outputs of Random Forests are a lot easier to interpret than the concepts and outputs of support vector machines.

Zha et al. (2020) found Random Forests performed better than SVR, multiple linear regression and artificial neural networks on predicting nitrogen content in rice using spectral indices. A model comparison study using spectral indices in Liu et al. (2016) also found Random Forests to perform better than the other non-parametric machine learning models in wheat. This could mean that the performance of Random Forests on nitrogen using spectral indices could be consistent on other types of crops and possibly give consistent results in different regions of Canada. Since different regions provide different climates, a comparison of the soil and nitrogen status in the crops could be studied. A consideration of future study could involve comparing Random Forests to other machine

learning models or deep learning using spectral indices on other types of crops. However, delving into deep learning requires a huge training dataset in order to be effective. Another drawback is the computational cost such as memory and processing power in order to tackle the datasets effectively with deep learning.

3.5   Conclusions

In this study, different regression methods were used to predict canopy nitrogen weight of corn using UAV MicaSense multispectral images. These models were established using the individual MicaSense bands and their associated VIs derived from the UAV reflectance values. Using the top 12 variables (in order: MSR, WDRVI, RVI, NDVI, ISR, BNDVI, Red band, OSAVI, RGBVI, CI Red edge, Blue band, and Red edge NDVI) derived with the Random Forests importance plot performed the best on estimating canopy nitrogen weight throughout the three dates in corn crops. Using the Random Forests model applied to the top 12 variables (RMSE = 4.52 g/m^2) on the validation performed marginally worse than the Random Forests model using all the variables (RMSE = 4.51 g/m^2), indicating that adding more variables into the model does not always improve its accuracy. However, because the difference of the accuracy is marginally different, this removes the unnecessary processing time of generating the 22 other VI images.

The UAV nitrogen prediction map can also detect spatial nitrogen variations within the field, especially in the July 18th image where the canopy nitrogen weight showed a large variation with the field data. In practice, these results could be useful for farmers in retrieving fast information about a field's nitrogen status, as they will know exactly which parts of their fields are in excess or deficient in the amount of nitrogen present. Practically, these results could be obtained on the day of the UAV flight, depending on the size of the field and the number of images acquired. Ultimately, this will lead to a much more efficient fertilizer application program for the farmers as they will know precisely how much nitrogen is needed in a particular spot with their GPS-enabled fertilizer spreader.

This study used data acquired over southwest Ontario fields in 2019 and there is the need to test the method over other datasets, such as different zones in Canada or a different crop. This will give an idea of how the developed method can be generalized and applied to different parts of Canada and whether it can be used on different crops. The study used MicaSense images with five spectral bands and there is a need to test different cameras that capture different wavelengths to understand which multispectral bands perform the best on predicting nitrogen using empirical regression techniques. Finally, another future consideration of this study can involve comparing the canopy nitrogen prediction map with other field spatial information, such as drainage and soil. This information can give a better idea on the contributions of nitrogen content that occur below the canopy.

3.6 References

AOAC 2006. AOAC Official Method 972.43, Microchemical determination of carbon, hydrogen, and nitrogen, automated method, in *Official Methods of Analysis of AOAC International*, 18th edition, Revision 1, Chapter 12, pp. 5-6, AOAC International, Gaithersburg, MD.

Bagheri, N., Ahmadi, H., Alavipanah, S. K., & Omid, M. (2013). Multispectral remote sensing for site-specific nitrogen fertilizer management. *Pesquisa Agropecuária Brasileira*, *48*(10), 1394-1401.

Baranoski, G. V. G., & Rokne, J. G. (2005). A practical approach for estimating the red edge position of plant leaf reflectance. *International Journal of Remote Sensing*, *26*(3), 503-521.

Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, *114*, 24-31.

Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., ... & Bareth, G. (2015). Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. *International Journal of Applied Earth Observation and Geoinformation*, *39*, 79-87.

Botha, E. J., Leblon, B., Zebarth, B. J., & Watmough, J. (2010). Non-destructive estimation of wheat leaf chlorophyll content from hyperspectral measurements through analytical model inversion. *International Journal of Remote Sensing*, *31*(7), 1679-1697.

Broge, N. H., & Leblanc, E. (2001). Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sensing of Environment*, 76(2), 156-172.

Campos, I., González-Gómez, L., Villodre, J., González-Piqueras, J., Suyker, A. E., & Calera, A. (2018). Remote sensing-based crop biomass with water or light-driven crop growth models in wheat commercial fields. *Field Crops Research*, 216, 175-188.

Čeh, M., Kilibarda, M., Lisec, A., & Bajat, B. (2018). Estimating the performance of random forest versus multiple regression for predicting prices of the apartments. *ISPRS International Journal of Geo-Information*, *7*(5), 168.

Chen, J. M. (1996). Evaluation of vegetation indices and a modified simple ratio for boreal applications. *Canadian Journal of Remote Sensing*, *22*(3), 229-242.

Chen, P., Haboudane, D., Tremblay, N., Wang, J., Vigneault, P., & Li, B. (2010). New spectral indicator assessing the efficiency of crop nitrogen treatment in corn and wheat. *Remote Sensing of Environment*, *114*(9), 1987-1997.

Clevers, J. G., & Kooistra, L. (2011). Using hyperspectral remote sensing data for retrieving canopy chlorophyll and nitrogen content. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *5*(2), 574-583.

Clevers, J. G., & Gitelson, A. A. (2013). Remote estimation of crop and grass chlorophyll and nitrogen content using red-edge bands on Sentinel-2 and-3. *International Journal of Applied Earth Observation and Geoinformation*, *23*, 344-351.

Curran, P. J., Dungan, J. L., & Gholz, H. L. (1990). Exploring the relationship between reflectance red edge and chlorophyll content in slash pine. *Tree physiology*, *7*(1-2-3-4), 33-48.

Daughtry, C. S. T., Walthall, C. L., Kim, M. S., De Colstoun, E. B., & McMurtrey Iii, J. E. (2000). Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sensing of Environment*, *74*(2), 229-239.

Deng, L., Mao, Z., Li, X., Hu, Z., Duan, F., & Yan, Y. (2018). UAV-based multispectral remote sensing for precision agriculture: A comparison between different cameras. *ISPRS Journal of Photogrammetry and Remote Sensing*, 146, 124-136.

Fan, L., Zhao, J., Xu, X., Liang, D., Yang, G., Feng, H., Yang, H., Chen, G. & Wei, P. (2019). Hyperspectral-based estimation of leaf nitrogen content in corn using optimal selection of multiple spectral variables. *Sensors*, *19*(13), 2898.

Fernandes, R., Butson, C., Leblanc, S., & Latifovic, R. (2003). Landsat-5 TM and Landsat-7 ETM+ based accuracy assessment of leaf area index products for Canada derived from SPOT-4 VEGETATION data. *Canadian Journal of Remote Sensing*, *29*(2), 241-258.

Frels, K., Guttieri, M., Joyce, B., Leavitt, B., & Baenziger, P. S. (2018). Evaluating canopy spectral reflectance vegetation indices to estimate nitrogen use traits in hard winter wheat. *Field Crops Research*, 217, 82-92.

Gitelson, A., & Merzlyak, M. N. (1994). Quantitative estimation of chlorophyll-a using reflectance spectra: experiments with autumn chestnut and maple leaves. *Journal of Photochemistry and Photobiology B: Biology*, *22*(3), 247-252.

Gitelson, A. A., Merzlyak, M. N., & Lichtenthaler, H. K. (1996). Detection of red edge position and chlorophyll content by reflectance measurements near 700 nm. *Journal of Plant Physiology*, *148*(3-4), 501-508.

Gitelson, A. A., & Merzlyak, M. N. (1998). Remote sensing of chlorophyll concentration in higher plant leaves. *Advances in Space Research*, *22*(5), 689-692.

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.

Gitelson, A. A., Gritz, Y., & Merzlyak, M. N. (2003). Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *Journal of Plant Physiology*, *160*(3), 271-282.

Gitelson, A. A. (2013). Remote estimation of crop fractional vegetation cover: the use of noise equivalent as an indicator of performance of vegetation indices. *International Journal of Remote Sensing*, *34*(17), 6054-6066.

Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J., & Strachan, I. B. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sensing of Environment*, *90*(3), 337-352.

Hansen, P. M., & Schjoerring, J. K. (2003). Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote sensing of environment*, *86*(4), 542-553.

Harwin, S., & Lucieer, A. (2012). Assessing the accuracy of georeferenced point clouds produced via multi-view stereopsis from unmanned aerial vehicle (UAV) imagery. *Remote Sensing*, *4*(6), 1573-1599.

Hijmans, R. J. (2019). raster: Geographic Data Analysis and Modelling. R package version 3.0-7.

Huete, A. (1988). A soil-adjusted vegetation index (SAVI). Remote Sensing of Environment. *Remote Sensing of Environment*, *25*, 295-309.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote sensing of environment*, *83*(1-2), 195-213.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112, pp. 99, 303-331). New York: Springer.

Jay, S., Maupas, F., Bendoula, R., & Gorretta, N. (2017). Retrieving LAI, chlorophyll and nitrogen contents in sugar beet crops from multi-angular optical remote sensing: Comparison of vegetation indices and PROSAIL inversion for field phenotyping. *Field Crops Research*, 210, 33-46.

Jones, H. G., & Vaughan, R. A. (2010). *Remote sensing of vegetation: principles, techniques, and applications*. Oxford University Press.

Jordan, C. F. (1969). Derivation of leaf-area index from quality of light on the forest floor. *Ecology*, *50*(4), 663-666.

Karimi, Y., Prasher, S. O., Madani, A., & Kim, S. (2008). Application of support vector machine technology for the estimation of crop biophysical parameters using aerial hyperspectral observations. *Canadian Biosystems Engineering*, *50*(7), 13-20.

Kayad, A., Sozzi, M., Gatto, S., Marinello, F., & Pirotti, F. (2019). Monitoring Within-Field Variability of Corn Yield using Sentinel-2 and Machine Learning Techniques. *Remote Sensing*, *11*(23), 2873.

Kelcey, J., & Lucieer, A. (2012). Sensor correction of a 6-band multispectral imaging sensor for UAV remote sensing. *Remote Sensing*, *4*(5), 1462-1493.

Le Maire, G., Francois, C., & Dufrene, E. (2004). Towards universal broad leaf chlorophyll indices using PROSPECT simulated database and hyperspectral reflectance measurements. *Remote sensing of environment*, *89*(1), 1-28.

Lee, H., Wang J., & Leblon, B. (2020). Canopy nitrogen retrieval from unmanned aerial vehicle imagery for wheat and corn fields. (*accepted)*

Liaw, A., & Wiener, M. (2002). Classification and Regression by RandomForest. *R News*, 2(3), 18-22.

Li, F., Gnyp, M. L., Jia, L., Miao, Y., Yu, Z., Koppe, W., ... & Zhang, F. (2008). Estimating N status of winter wheat using a handheld spectrometer in the North China Plain. *Field Crops Research*, 106(1), 77-85.

Li, F., Miao, Y., Feng, G., Yuan, F., Yue, S., Gao, X., & Chen, X. (2014). Improving estimation of summer maize nitrogen status with red edge-based spectral vegetation indices. *Field Crops Research*, *157*, 111-123.

Liu, Y., Cheng, T., Zhu, Y., Tian, Y., Cao, W., Yao, X., & Wang, N. (2016). Comparative analysis of vegetation indices, non-parametric and physical retrieval methods for monitoring nitrogen in wheat using UAV-based multispectral imagery. In 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS'16), Beijing, China, July, pp. 7362-7365.

Louppe, G. (2014). Understanding random forests. *Cornell University Library*.

Mailvaganam, S. (2017). Ontario farm data. *Ministry of Agriculture, Food and Rural Affairs 2017.* Retrieved from http://www.omafra.gov.on.ca/english/stats/census/summary.htm

Maresma, Á., Ariza, M., Martínez, E., Lloveras, J., & Martínez-Casasnovas, J. A. (2016). Analysis of vegetation indices to determine nitrogen application and yield prediction in maize (*Zea mays* L.) from a standard UAV service. *Remote Sensing*, *8*(12), 973.

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., & Leisch, F. (2018). e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien, 2017. R package version, 1(8).

Miphokasap, P., Honda, K., Vaiphasa, C., Souris, M., & Nagai, M. (2012). Estimating canopy nitrogen concentration in sugarcane using field imaging spectroscopy. *Remote Sensing*, *4*(6), 1651-1670.

Nasrallah, A., Baghdadi, N., El Hajj, M., Darwish, T., Belhouchette, H., Faour, G., ... & Mhawej, M. (2019). Sentinel-1 data for winter wheat phenology monitoring and mapping. *Remote Sensing*, *11*(19), 2228.

Osco, L.P., Marques Ramos, A. P., Roberto Pereira, D., Akemi Saito Moriya, É., Nobuhiro Imai, N., Takashi Matsubara, E., ... & Li, J. (2019). Predicting Canopy Nitrogen Content in Citrus-Trees Using Random Forest Algorithm Associated to Spectral Vegetation Indices from UAV-Imagery. *Remote Sensing*, *11*(24), 2925.

Pix4D Drone Mapping Software (2014). *Swiss Fed Inst Technol Lausanne, Route Cantonale, Switz*. Available online: http://pix4d.com

Qi, J., Chehbouni, A., Huete, A. R., Kerr, Y. H., & Sorooshian, S. (1994). A modified soil adjusted vegetation index. *Remote sensing of Environment, 48(2),* 119.

R Core Team (2019). *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria. Available online: https://www.r-project.org/index.html

Raparelli, E., & Bajocco, S. (2019). A bibliometric analysis on the use of unmanned aerial vehicles in agricultural and forestry studies. *International Journal of Remote Sensing*, 40, 1-14.

Rondeaux, G., Steven, M., & Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sensing of Environment, 55(2*), 95-107.

Roujean, J. L., & Breon, F. M. (1995). Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sensing of Environment*, *51*(3), 375-384.

Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. *NASA Special Publication*, *351*, 309.

Schlemmer, M., Gitelson, A., Schepers, J., Ferguson, R., Peng, Y., Shanahan, J., & Rundquist, D. (2013). Remote estimation of nitrogen and chlorophyll contents in maize at leaf and canopy levels. *International Journal of Applied Earth Observation and Geoinformation*, 25, 47-54.

Shahhosseini, M., Martinez-Feria, R. A., Hu, G., & Archontoulis, S. V. (2019). Maize yield and nitrate loss prediction with machine learning algorithms. *Environmental Research Letters*, *14*(12), 124026.

Sripada, R. P. (2005). Determining in-season nitrogen requirements for corn using aerial color-infrared photography. *NC State Theses and Dissertations.* North Carolina State University.

Sripada, R. P., Heiniger, R. W., White, J. G., & Meijer, A. D. (2006). Aerial color infrared photography for determining early in-season nitrogen requirements in corn. *Agronomy Journal*, *98*(4), 968-977.

Tagle Casapia, M. X. (2017). Study of radiometric variations in Unmanned Aerial Vehicle remote sensing imagery for vegetation mapping. *Lund University GEM thesis series*.

Tian, Y. C., Yao, X., Yang, J., Cao, W. X., Hannaway, D. B., & Zhu, Y. (2011). Assessing newly developed and published vegetation indices for estimating rice leaf nitrogen concentration with ground-and space-based hyperspectral reflectance. *Field Crops Research*, *120*(2), 299-310.

Tucker, C. J., Elgin Jr, J. H., McMurtrey Iii, J. E., & Fan, C. J. (1979). Monitoring corn and soybean crop development with hand-held radiometer spectral data. *Remote Sensing of Environment*, *8*(3), 237-248.

Turner, D., Lucieer, A., & Wallace, L. (2013). Direct georeferencing of ultrahigh-resolution UAV imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(5), 2738-2745.

Wang, F. M., Huang, J. F., Tang, Y. L., & Wang, X. Z. (2007). New vegetation index and its application in estimating leaf area index of rice. *Rice Science*, *14*(3), 195-203.

Wang, L., Zhou, X., Zhu, X., & Guo, W. (2017). Estimation of leaf nitrogen concentration in wheat using the MK-SVR algorithm and satellite remote sensing data. *Computers and Electronics in Agriculture*, *140*, 327-337.

Xie, Q., Dash, J., Huang, W., Peng, D., Qin, Q., Mortimer, H., Casa, R., Pignatti, S., Laneve, G., Pascucci, S. & Dong, Y. (2018). Vegetation indices combining the red and red-edge spectral information for leaf area index retrieval. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(5), 1482-1493.

Xiong, X., Zhang, J., Guo, D., Chang, L., & Huang, D. (2019). Non-Invasive Sensing of Nitrogen in Plant Using Digital Images and Machine Learning for Brassica Campestris ssp. Chinensis L. *Sensors*, *19*(11), 2448.

Yang, G., Liu, J., Zhao, C., Li, Z., Huang, Y., Yu, H., ... & Zhang, R. (2017). Unmanned aerial vehicle remote sensing for field-based crop phenotyping: current status and perspectives. *Frontiers in Plant Science*, *8*, 1111.

Yue, J., Yang, G., Li, C., Li, Z., Wang, Y., Feng, H., & Xu, B. (2017). Estimation of winter wheat above-ground biomass using unmanned aerial vehicle-based snapshot hyperspectral sensor and crop height improved models. *Remote Sensing*, 9(7), 708.

Zarco-Tejada, P. J., Miller, J. R., Mohammed, G. H., Noland, T. L., & Sampson, P. H. (2002). Vegetation stress detection through chlorophyll a+ b estimation and fluorescence effects on hyperspectral imagery. *Journal of environmental quality*, *31*(5), 1433-1441.

Zha, H., Miao, Y., Wang, T., Li, Y., Zhang, J., Sun, W., Feng, Z. & Kusnierek, K. (2020). Improving Unmanned Aerial Vehicle Remote Sensing-Based Rice Nitrogen Nutrition Index Prediction with Machine Learning. *Remote Sensing,*12(2), 215.

Zhang, C., & Kovacs, J. M. (2012). The application of small unmanned aerial systems for precision agriculture: a review. *Precision Agriculture*, *13*(6), 693-712.

Zhao, B., Duan, A., Ata-Ul-Karim, S. T., Liu, Z., Chen, Z., Gong, Z., ... & Ning, D. (2018). Exploring new spectral bands and vegetation indices for estimating nitrogen nutrition index of summer maize. *European Journal of Agronomy*, *93*, 113-125.

Zheng, H., Li, W., Jiang, J., Liu, Y., Cheng, T., Tian, Y., & Yao, X. (2018). A comparative assessment of different modelling algorithms for estimating leaf nitrogen content in winter wheat using multispectral images from an Unmanned Aerial Vehicle. *Remote Sensing*, 10(12), 20-26

4   Chapter 4

This chapter presents the responses of the objectives of this thesis, possible future studies and the limitations of this research.

4.1   Conclusion

The responses to the objectives of this thesis are presented:

(1) Linear regression was used to predict wheat and corn in chapter 2 using ground spectral reflectance calibrated to the MicaSense RedEdge camera bands. Ground ASD spectra were verified with the UAV reflectance with an R^2 of 0.94, indicating that the two measurements are well matched. The best performing linear regression with the VIs in the study was RVI. RVI performed the best on both wheat and corn and on all the fields in the study. The study found that the other VIs presented a saturation effect in the later growing stages, making predictions unreliable. Multiple regression was also evaluated and avoided due to the high degree of multicollinearity of all the spectral indices (VIF >10). This study has also found that three sampling dates of wheat performed better than using four sampling dates. Removing the fourth sampling date could mean that the later growing stage of wheat may not give much sensitive information in the overall model. Overall, the best performing wheat field had an RMSE of 0.95 g/m^2 on the date of May 24th imagery. The best performing corn field had an RMSE of 0.66 g/m^2 on the June 7th imagery.

(2) The methodology of chapter 3 learned from the results of chapter 2 and used different regression methods on predicting canopy nitrogen weight, using non-parametric tests for non-linear datasets. Chapter 2 showed that most vegetation indices presented a non-linear relationship and that multicollinearity was present when performing multiple linear regression. Therefore, chapter 3 used machine learning techniques like Random Forests and SVR that work well with non-linear datasets and mitigate the effect of multicollinearity among the explanatory variables. Random Forests' variable importance plot was used to determine the most contributing spectral variable on canopy nitrogen weight in corn. Random

Forests with all 34 explanatory variables was found to be the best performing model compared to SVR and linear regression on canopy nitrogen weight. However, in realistic practical terms, Random Forests with 12 variables was used to generate the final nitrogen prediction map as the difference of model's RMSE of 12 and 34 explanatory variables was 0.01 g/m^2, while the processing time was significantly reduced. SVR also produced good results; however, the concepts and interpretations of Random Forests are easier than support vector machines. Furthermore, the best performing nitrogen prediction map using the 12 variable Random Forests had an RMSE of 0.62 g/m^2 on the July 3rd imagery. The best result of chapter 3 (RMSE = 0.62 g/m^2) is an improvement over the best result of chapter 2 (RMSE = 0.66 g/m^2) on corn fields. Both Random Forests and SVR performed better than linear regression in chapter 3.

(3) The final products of chapter 2 and chapter 3 were nitrogen prediction maps using UAV multispectral imagery. Both chapters used the MicaSense RedEdge multispectral camera. In chapter 2, a 5cm/pixel spatial resolution imagery was used to produce the nitrogen map, while chapter 3 downscaled the resolution to 15cm/pixel due to the processing time of using multi-layered raster images. Both imageries were able to finely separate the soil from the crop in the early growth stages of the imagery. Most importantly, both chapters were able to identify the different areas of highs and lows of canopy nitrogen weight. This information is practical for farmers as they can identify the areas of their fields that are in excess or deficient in nitrogen.

4.2 Limitations

Several limitations exist in this thesis in both chapters 2 and 3. The first limitation is the distribution of the sample points throughout the entire field. Sample points were often weighted towards one side of the field. The issue with this is the enormous size of the fields. Evenly distributing the sample points throughout the entire field was not realistic in terms of labour. Temperatures during the fieldwork sometimes reached up to 35°C, while carrying equipment and the biomass of the crop. Therefore, the study's sample points were not able to fully capture the entire area of certain fields. A suggestion for

future research for this limitation is by having multiple teams retrieving the data in different sections of the field simultaneously. However, this would require more fieldwork members and a multiple set of equipment. This could get very costly, particularly with ASD equipment costing tens of thousands of dollars (depending on the model).

Another limitation was the limited number of cameras used in the study. Only one multispectral camera (MicaSense RedEdge) was used for both chapter 2 and chapter 3. Using a hyperspectral camera could have given a comparison of the results, as multispectral cameras capture certain wavelengths in the ES, whereas the hyperspectral would capture the entire spectrum of interest. However, limitation is the UAV being able to handle the weight of all sensors simultaneously. Two separate flights could have been flown; however, this would have to rely on the weather remaining the same for both of the flights.

4.3   Discussion and Future Work

Possible future studies could be proposed from the results of this thesis. As pointed out in the response from objective 1, removing the later growing stage of wheat from the overall model improved the relationship between canopy nitrogen weight and RVI. Therefore, specific study analysing the later growing stage of wheat separately or an insight on how to incorporate the later growing stage into the overall model could be proposed.

The results of chapter 2 indicated that the relationship between the canopy nitrogen weight and VIs presented a non-linear relationship. A linear fit was only shown in RVI in chapter 2, while NDVI, GNDVI, MTVI2 and RE_NDVI showed a saturated relationship. The linear fit for the latter VI was shown for visual purposes and not used for prediction. Therefore, this allowed a modification of the analysis of chapter 3 to include non-parametric modelling that can specifically work with non-linear data and highly intercorrelated variables. Chapter 3 proposed a set of machine learning methods to predict canopy nitrogen weight in corn: Random Forests and SVR to compare along with linear regression. The results have shown that non-parametric models perform better than traditional parametric regression models when using spectral indices due to its non-linear

relationship. Various non-parametric models such as: kernel ridge regression, principal components regression, and Gaussian process regression exist and could be used to compare with Random Forests and SVR on predicting canopy nitrogen weight.

Insights on other explanatory variables could be included in future models that capture the prediction of canopy nitrogen weight. Other spatial field information, such as drainage and soil information could be used in future models along with spectral VIs. This information can give a better idea on the contributions of nitrogen content that occur below the canopy and could possibly give another direction into analysing plant nitrogen stress.

5   Appendix

Appendix A - Field Nitrogen Summary Statistics

Table A-1. Summary statistics of the nitrogen measurements for wheat in the 2018 field campaign

Date	Field	Plant Nitrogen % (mean)	Plant Nitrogen % (sd)	Canopy Nitrogen Weight (mean) (g/m^2)	Canopy Nitrogen Weight (sd) (g/m^2)
May 7th	McColl	5.85	0.39	1.98	0.60
	Bale	6.23	0.41	0.64	0.19
	Hetzell	5.98	0.61	0.53	0.19
May 14th	McColl	5.30	0.39	2.95	1.29
	Bale	5.01	0.16	1.42	0.36
	Hetzell	5.67	0.36	1.25	0.53
May 25th	McColl	4.05	0.39	6.41	1.44
	Bale	4.80	0.48	4.91	0.23
	Hetzell	4.86	0.52	5.02	1.31
June 4th	McColl	4.69	0.33	11.25	2.83
	Bale	4.77	0.45	9.62	3.08
	Hetzell	4.56	0.57	8.61	1.65

Table A-2. Summary statistics of the nitrogen measurements for corn in the 2018 field campaign

Date	Field	Plant Nitrogen % (mean)	Plant Nitrogen % (sd)	Canopy Nitrogen Weight (mean) (g/m^2)	Canopy Nitrogen Weight (sd) (g/m^2)
June 4th	Crandell	5.29	0.32	2.35	0.39
	Jack North	5.65	0.17	2.65	0.49
	Paul	4.93	0.41	2.10	0.56
June 12th	Crandell	4.10	0.43	5.26	1.81
	Jack North	N/A	N/A	N/A	N/A
	Paul	3.56	0.54	5.12	2.31
June 20th	Crandell	3.79	0.23	9.31	3.54
	Jack North	3.01	0.36	8.88	2.93
	Paul	3.28	0.35	6.83	1.16
June 26th	Crandell	3.87	0.28	28.49	8.54
	Jack North	3.42	0.30	14.63	2.73
	Paul	3.57	0.44	19.13	6.63
July 4th	Crandell	3.78	0.42	69.39	20.72
	Jack North	3.60	0.53	45.86	15.65
	Paul	3.87	0.55	58.96	15.84

Table A-3. Summary statistics of the nitrogen measurements for corn in the 2019 field campaign

Date	Field	Plant Nitrogen % (mean)	Plant Nitrogen % (sd)	Canopy Nitrogen Weight (mean) (g/m²)	Canopy Nitrogen Weight (sd) (g/m²)
June 25th	JJ	5.55	0.35	0.28	0.21
	Susan	N/A	N/A	N/A	N/A
July 3rd	JJ	4.96	0.63	2.29	1.18
	Susan	5.48	0.31	1.02	0.29
July 10th	JJ	3.44	0.59	14.37	8.76
	Susan	3.75	0.55	7.93	2.58
July 18th	JJ	3.66	0.27	23.00	12.26
	Susan	3.71	0.31	19.18	10.54
July 23rd	JJ	3.66	0.54	32.45	8.96
	Susan	3.63	0.18	27.06	10.05
July 31st	JJ	3.76	0.25	61.28	17.27
	Susan	3.50	0.31	32.22	9.95

Appendix B - Field and Lab Photos

Figure B-1. Landscape photo of wheat field taken on May 2nd, 2019

Figure B-2. Bird's eye view of wheat taken on May 27th, 2019

Figure B-3. Close-up photo of wheat taken on July 2nd, 2019

Figure B-4. Landscape photo of corn taken on July 18th, 2019 with biomass collection

Figure B-5. Close-up photo of (dent) corn just before harvesting on October 25th, 2019

Figure B-6. Retrieving spectral reflectance data from corn using ASD (Robin Kwik, GITA lab)

Figure B-7. UAV (DJI Matrice 100) prior to take-off. Calibration panel is shown beside

Figure B-8. UAV (DJI Matrice 100) flying over corn field at early growth stage

Figure B-9. Using scale to weigh biomass of corn at A and L Canada Laboratory

Appendix C – R Code

Figure C-1. R code for modelling Random Forests, Support vector regression and Linear regression

```
https://gist.github.com/hwangl5/eec9653b840901d8475e6bb55f0e73b2
###############Splitting data into train and valid###########
set.seed(123);train <- sample(nrow(data), 0.7*nrow(data), replace = F)
trainset <- data[train,]
validset <- data[-train,]

####packages for RF and SVR modeling. Linear regression is defaulted in R####
library(randomForest); library(e1071)

###Random Forests model of training set#########   the "." indicates to use all the
variables in the dataframe########
RFmodelT<- randomForest(Nitrogen ~., data = trainset)

#### Apply the model on both calibration and validation set######
predictT <- predict(RFmodelT, newdata= trainset)
predictV <- predict(RFmodelT, newdata= validset)

###calculate RMSE of the model#####
rmsemodT <- sqrt(mean((predictT - trainset$Nitrogen)^2))
rmsemodV <- sqrt(mean((predictV - validset$Nitrogen)^2))

###### follow same steps for SVM and Linear regression###########
###### functions for SVM is "SVM() and Linear regresison is "lm()"
```

Figure C-2. R code for using the models generated to predict multi-layered raster images

```
https://gist.github.com/hwangl5/659b4ce5b7732c2487b7335d9caea4db
#### install packages prior to calling the library function #####
library(raster); library(sp); library(rgdal); library(raster); library(rasterVis)

####list.files functions loop through your folder and will look for all the images that
end with "tif" ####
fs<- list.files(path="D:/2019 thesis/Susanb_July18/", pattern = "tif$", full.names = T)

####create variable name rasterstack and use the function stack on the list of files that
were formed in the fs variable
rasterstack <- raster::stack(fs)

####predict function is used to call the rasterstack dataframe and predict each pixel
with the random forest model created or any model
prediction <-raster::predict(object=rasterstack,model =RFmodelT, fun=predict)

####this function will take in the prediction dataframe and create a raster file. The
second parametre is the location and name of file
##### you wish to name it #####
writeRaster(prediction, "D:/2019 thesis/Susanb_July18/PredictionJuly18sb.tif")
```

Curriculum Vitae

Name:	Hwang Lee
Post-secondary Education and Degrees:	University of Mary Bismarck, North Dakota, USA 2013-2015 B.Sc. (Transferred-not awarded) University of Western Ontario London, Ontario, Canada 2015-2018 B.A. (hons) University of Western Ontario London, Ontario, Canada 2018-2020 M.Sc.
Honours and Awards:	NSIC student-athlete award (Track and Field) 2013-2014 Social Science Honour Roll 2015-2017
Related Work Experience	Teaching Assistant University of Western Ontario 2018-2020 Research Assistant GITA lab, University of Western Ontario 2018-2020

Publications:

Lee, H., Wang, J., & Leblon, B. (2020). Canopy Nitrogen Retrieval from Unmanned

Aerial Vehicle Imagery for Wheat and Corn Fields. *Canadian Journal of Remote*

Sensing, 1-19

Lee, H., Wang J., & Leblon, B. (2020). Using Linear Regression, Random Forests, and

Support Vector Machine with Unmanned Aerial Vehicle Multispectral Images to Predict

Canopy Nitrogen Weight in Corn. *Remote Sensing*, *12*(13), 2071

Presentations:

Lee, H., Wang, J., & Leblon, B. (2019, June). Canopy Nitrogen Retrieval from Unmanned Aerial Vehicle Imagery for Wheat and Corn Fields. Paper presented at the 40th Canadian Symposium on Remote Sensing, Fredericton, New Brunswick.

Lee, H., Wang, J., & Leblon, B. (2020, July). Using Linear Regression, Random Forests, and Support Vector Machine with Unmanned Aerial Vehicle Multispectral Images to Predict Canopy Nitrogen Weight in Corn. Poster presented at the 41st Canadian Symposium on Remote Sensing, via Zoom.