
Electronic Thesis and Dissertation Repository

11-19-2021 2:00 PM

Intra-field Nitrogen Estimation for Wheat and Corn using Unmanned Aerial Vehicle-based and Satellite Multispectral Imagery, Plant Biophysical Variables, Field Properties, and Machine Learning Methods

Jody Seymon Yu, *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

© Jody Seymon Yu 2021

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [Agricultural and Resource Economics Commons](#), [Geographic Information Sciences Commons](#), [Remote Sensing Commons](#), and the [Spatial Science Commons](#)

Recommended Citation

Yu, Jody Seymon, "Intra-field Nitrogen Estimation for Wheat and Corn using Unmanned Aerial Vehicle-based and Satellite Multispectral Imagery, Plant Biophysical Variables, Field Properties, and Machine Learning Methods" (2021). *Electronic Thesis and Dissertation Repository*. 8346.
<https://ir.lib.uwo.ca/etd/8346>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

Intra-field Nitrogen Estimation for Wheat and Corn using Unmanned Aerial Vehicle-based and Satellite Multispectral Imagery, Plant Biophysical Variables, Field Properties and Machine Learning Methods

Abstract

Management of nitrogen (N) fertilizers is an important agricultural practice and field of research to increase productivity, minimize environmental impacts and the cost of production. To apply N fertilizer at the right rate, time, and place depends on the crop type, desired yield, and field conditions. The objective of this study is to use Unmanned Aerial Vehicle (UAV) multispectral imagery, PlanetScope satellite imagery, vegetation indices (VI), crop height, leaf area index (LAI), field topographic metrics, and soil properties to predict canopy nitrogen weight (g/m^2) of corn and wheat fields in southwestern Ontario, Canada. Random Forests (RF) and Support Vector Regression (SVR) machine learning models were tested with combinations of variable datasets and evaluated for accuracy of canopy nitrogen weight prediction. The results demonstrate that UAV and satellite-based prediction models including spectral variables, crop biophysical parameters, and field conditions can provide accurate and useful information for fertilizer management.

Keywords

Unmanned Aerial Vehicle (UAV), satellite imagery, precision agriculture, nitrogen management, vegetation indices, machine learning, Random Forests, Support Vector Machine, corn, wheat

Summary for Lay Audience

The practice of agriculture first began thousands of years ago, when humans began building permanent settlements, growing their own crops, and raising livestock for reliable survival resources. Today, agriculture continues to be one of the most important industries around the world providing food and materials for over 7.7 billion people. As the world population is projected to continue growing, for agriculture to meet increasing demands requires sustainable production, adaptability to changing climates, and better methods of farming. Advances in technology have opened the way for precision agriculture, a management technique that gathers many types of data about a crop to improve resource use (e.g., water, fertilizer, pesticides), quality of production, profitability, and sustainability.

Remote sensing is the process of obtaining information about an object or phenomenon at a distance. In precision agriculture, remote sensing is especially useful without the need to make physical contact with plants to gather valuable crop data for analysis. With our eyes, we see colors such as blue, green, and red, through the way light is reflected or emitted from a surface. With special cameras designed for certain wavelengths of light, we can see beyond the visible spectrum and understand more about a plant's characteristics (i.e., healthy plants may reflect more infrared light than an unhealthy plant, but we cannot see that with our eyes). This thesis studied how we may estimate nitrogen levels in corn and wheat plants using both satellite and unmanned aerial vehicle (UAV; a.k.a. drone) remotely sensed data.

The growth of a plant depends on many different factors including light, water, nutrients, and more. Using machine learning in this thesis, we built computer models from gathered data of a crop's response to light, biophysical characteristics, and environmental conditions to improve nitrogen level estimation of crop fields. Nitrogen prediction maps of a field can be created from the best models showing areas of a field that need certain amounts of fertilizers. Following the map, a farmer can practice sustainable precision agriculture by applying resources at the right rate, time, and place for a bountiful crop.

Co-Authorship Statement

The thesis follows an integrated-article format. The work in this thesis was conducted by the author under supervision from Dr. Jinfei Wang and Dr. Brigitte Leblon. Ideas, guidance, reviews, comments, and edits of methodology and validation were generously provided by Drs. Wang and Leblon for the articles. Drs. Wang and Leblon are co-authors of the published article:

Yu, J., Wang, J., & Leblon, B. (2021). Evaluation of soil properties, topographic metrics, plant height, and unmanned aerial vehicle multispectral imagery using machine learning methods to estimate canopy nitrogen weight in corn. *Remote Sensing*, 13(16), 3105.

<https://doi.org/10.3390/rs13163105>

Drs. Wang, Leblon, and Yang Song are co-authors of the published article:

Yu, J., Wang, J., Leblon, B., & Song, Y. (2022). Nitrogen Estimation for Wheat Using UAV-Based and Satellite Multispectral Imagery, Topographic Metrics, Leaf Area Index, Plant Height, Soil Moisture, and Machine Learning Methods. *Nitrogen*, 3(1),1-25.

<https://doi.org/10.3390/nitrogen3010001>

Acknowledgments

I would first like to thank my supervisors Dr. Jinfei Wang (Professor, University of Western Ontario) and Dr. Brigitte Leblon (Professor, University of New Brunswick) for their continuous guidance, advice, and patience during my Master's studies. I am exceptionally grateful for their time and effort supervising my research, and I have learned much through the experience. Dr. Wang first introduced me to remote sensing during my undergraduate studies, teaching me much of what I know about this field of study and giving me many opportunities to pursue my academic fascinations. Her support and encouragement through the years were invaluable, and I will always be thankful to have had her as a professor.

I would like to sincerely thank the A&L Canada staff and the Geographic Information Technology and Applications (GITA) Lab for all their assistance with data collection and processing. Fieldwork by itself is tough work, not to mention the added challenges from the COVID-19 pandemic situation. Thank you to my colleagues for always being open to sharing and discussing ideas, teaching me new concepts, and helping me through problems over the years.

Lastly, thank you to my parents Paul and Elisa Yu, my family, my loved ones, and my friends for always being there for me. Thank you for believing in me and supporting me through thick and thin. I did not expect to be working remotely through most of my Masters, but it has taught me to be even more thankful for the time I get to spend with the people I love. Each shared meal, message, phone call, video call, games night – all these moments together make life all the better.

Table of Contents

Abstract.....	ii
Summary for Lay Audience.....	iii
Co-Authorship Statement.....	iv
Acknowledgments.....	v
Table of Contents.....	vi
List of Tables.....	ix
List of Figures.....	xi
List of Appendices.....	xiii
Glossary.....	xv
Chapter 1.....	1
1 Introduction.....	1
1.1 Background.....	2
1.1.1 Precision Agriculture & Nitrogen.....	2
1.1.2 Remote Sensing Techniques.....	3
1.1.3 Remote Sensing Platforms.....	6
1.2 Research Questions.....	7
1.3 Research Objectives.....	8
1.4 Thesis Structure.....	8
1.5 Study Areas.....	9
1.6 References.....	10
Chapter 2.....	14
2 Evaluation of Soil Properties, Topographic Metrics, Plant Height, and Unmanned Aerial Vehicle Multispectral Imagery using Machine Learning Methods to Estimate Canopy Nitrogen Weight in Corn.....	14
2.1 Introduction.....	14

2.2	Materials and Methods.....	17
2.2.1	Study Area & Data.....	17
2.2.2	UAV Imagery.....	20
2.2.3	UAV Image Processing.....	22
2.2.4	Vegetation Indices	24
2.2.5	Canopy Nitrogen Weight Estimation.....	25
2.2.6	Canopy Nitrogen Weight Modelling	26
2.3	Results.....	27
2.3.1	Nitrogen Statistics.....	27
2.3.2	Soil Statistics.....	28
2.3.3	Regression Models with all Parameters	28
2.3.4	Variable Importance Plot	30
2.3.5	Regression Models with Selected Variables.....	31
2.4	Discussion.....	33
2.5	Conclusion	35
2.6	References.....	37
	Chapter 3.....	43
3	Nitrogen Estimation for Wheat using UAV-based and Satellite Multispectral Imagery, Topographic Metrics, Leaf Area Index, Plant Height, Soil Moisture, and Machine Learning Methods	43
3.1	Introduction.....	43
3.2	Materials and Methods.....	47
3.2.1	Study Area & Data.....	47
3.2.2	UAV Imagery.....	51
3.2.3	UAV Image Processing.....	52
3.2.4	Satellite Imagery	54
3.2.5	Vegetation Indices	55

3.2.6	Canopy Nitrogen Weight Modelling	57
3.3	Results.....	59
3.3.1	Nitrogen Statistics.....	59
3.3.2	Soil Statistics.....	60
3.3.3	Regression Models with All Parameters.....	60
3.3.4	Variable Importance Plots.....	64
3.3.5	Regression Models with Selected Parameters	66
3.3.6	Crop Nitrogen Prediction Maps	70
3.4	Discussion.....	73
3.5	Conclusion	77
3.6	References.....	78
Chapter 4	85
4	Conclusion	85
4.1	Summary.....	85
4.2	Conclusions.....	86
4.3	Limitations & Future Work	88
5	Appendices.....	90
5.1	Appendix A – Field Photos.....	90
5.2	Appendix B – Pix4D Processing.....	98
5.3	Appendix C – Topographic Metrics Processing.....	100
5.4	Appendix D – LAI Layers	102
5.5	Appendix E – R Code	103
5.6	Appendix F - Copyrighted Material & Permissions	105
Curriculum Vitae	108

List of Tables

Table 2-1. Summary of Unmanned Aerial Vehicle (UAV) flight acquisitions with MicaSense RedEdge multispectral camera.	22
Table 2-2. Spectral characteristics of the five MicaSense bands.....	24
Table 2-3. Vegetation indices used in the study.	25
Table 2-4. Statistics for the calibration of the canopy nitrogen model with all 29 variables including vegetation indices, the 5 MicaSense band reflectances, plant physiology variables, soil metrics, and topographic metrics as a function of the date and the modelling approach (RF or SVR) ¹	29
Table 2-5. Statistics for the validation of the canopy nitrogen model with all 29 variables including selected vegetation indices, the 5 MicaSense bands, plant physiology variables, soil metrics, and topo-graphic metrics as a function of the date and the modelling approach (RF or SVR).	30
Table 2-6. Statistics for the calibration of the canopy nitrogen model with all dates (8, 15, 24 June), different combinations of variables (n = 70) ¹	32
Table 2-7. Statistics for the validation of the canopy nitrogen model with all dates (June 8, 15, 24), different combinations of variables (n = 30) ¹	33
Table 3-1. Summary of fieldwork data acquisitions.	48
Table 3-2. Spectral characteristics of the five MicaSense bands.....	51
Table 3-3. Spectral characteristics of the PlanetScope Dove-R sensors.....	54
Table 3-4. Vegetation indices used in the study.	56
Table 3-5. Statistics for the calibration of the UAV canopy nitrogen regression models with 28 variables as a function of the date and modelling approach (RF or SVR) ¹	61

Table 3-6. Statistics for the validation of the UAV canopy nitrogen models with 28 variables as a function of the date and modelling approach (RF or SVR).	62
Table 3-7. Statistics for the calibration of the PlanetScope canopy nitrogen models with 24 variables as a function of the date and the modelling approach (RF or SVR) ¹	63
Table 3-8. Statistics for the validation of PlanetScope canopy nitrogen models with 24 variables as a function of the date and the modelling approach (RF or SVR).....	64
Table 3-9. Statistics for the calibration of the canopy nitrogen model with 12, 20, 27 May MicaSense data and different combinations of variables (n = 89) ¹	67
Table 3-10. Statistics for the validation of the canopy nitrogen model with 12, 20, 27 May MicaSense data and different combinations of variables (n = 39) ¹	68
Table 3-11. Statistics for the calibration of the PlanetScope canopy nitrogen model with 12, 20, 27 May data and different combinations of variables (n = 100) ¹	69
Table 3-12. Statistics for the validation of the PlanetScope canopy nitrogen model with 12, 20, 27 May data and different combinations of variables (n = 44) ¹	69

List of Figures

Figure 1-1. Generalized spectra characteristics of common land surface types in the visible and infrared range. Adapted from Gosa, 2009..... 4

Figure 1-2. Study areas located near London, Ontario. The corn field is indicated in yellow, and the three wheat fields are indicated in red. Background image is from Google Earth. 9

Figure 2-1. Location of the corn field in Central Elgin, Ontario, Canada, over a Google Earth image..... 18

Figure 2-2. Digital elevation model showing the variation of the corn field's topography and the distribution of sample points..... 20

Figure 2-3. (a) DJI Matrice 100 UAV, with the MicaSense and RGB Gimbal cameras attached. (b) DJI Phantom 4 RTK UAV..... 22

Figure 2-4. Flowchart of the methodology used in the study. 23

Figure 2-5. Variation of canopy nitrogen weight (g/m^2) as a function of the date of field measurements during the 2020 growing season. Outliers are represented by dots on the graph. 28

Figure 2-6. Variance Importance plot produced by the Random Forest model of all three dates and 30 variables, using the function `varImpPlot()` in R Studio. Higher `IncNodePurity` values indicate more impact on nitrogen. ⁽¹⁾ Please refer to Table 1 for the full name of vegetation indices. `N_Weight`, plant nitrogen weight; `Soil_Nitrate_N`, soil nitrate nitrogen ($\text{NO}_3\text{-N}$); `Soil_Miner_N`, soil mineralizable nitrogen; `Soil_Total_N`, water extracted total soil nitrogen; `TWI_1`, total wetness index #1; `TWI_2`, total wetness index #2..... 31

Figure 3-1. Locations of the three 2020 wheat fields in southwestern Ontario, Canada over a Google Earth image. Field W1 is situated in the upper left corner, west of Mount Brydges. Fields W2 and W3 are in the lower right corner, south of Mount Brydges..... 47

Figure 3-2. Digital elevation model showing the variation of the wheat field's topography at each sample point for a) W1 wheat field, b) W2 wheat field, c) W3 wheat field. 50

Figure 3-3. Flowchart of the methodology used in the study. 53

Figure 3-4. Variation of canopy nitrogen weight (g/m^2) as a function of the date of field measurements during the 2020 growing season. An outlier is represented by a dot on the graph. 59

Figure 3-5. Variance Importance plot produced by the UAV RF regression model of data from 12, 20, and 27 May and 28 variables, using the function `varImpPlot()` in Rstudio. Higher `IncNodePurity` values indicate more impact on canopy nitrogen estimation. (1) Please refer to Table 1 for the full name of vegetation indices. TWI_1, total wetness index #1; TWI_2, total wetness index #2. 65

Figure 3-6. Variance Importance plot produced by the PlanetScope RF regression model of data from 12, 20, and 27 May and 24 variables, using the function `varImpPlot()` in Rstudio. Higher `IncNodePurity` values indicate more impact on canopy nitrogen estimation. (1) Please refer to Table 1 for the full name of vegetation indices. TWI_1, total wetness index #1; TWI_2, total wetness index #2. 66

Figure 3-7. Nitrogen prediction maps of wheat fields: from UAV-based models for a) W1, b) W2, and c) W3 fields; from satellite-based models for d) W1, e) W2, and f) W3 fields. 72

List of Appendices

Appendix A – Field Photos

Figure A 1. Example of ground control point marker. 90

Figure A 2. Piloting the unmanned aerial vehicles. The tall device on the left is the real time kinematic (RTK) navigation base station. 90

Figure A 3. Sensors attached to the underbody of the DJI Matrice 100 UAV. The RGB camera is on the top, and the red MicaSense RedEdge camera with five sensor bands is on the bottom. 91

Figure A 4. Photos of corn crop taken on 8 June 2020..... 92

Figure A 5. Photos of corn crop taken on 15 June 2020..... 93

Figure A 6. Photos of corn crop taken on 24 June 2020..... 94

Figure A 7. Photos of wheat crop taken on 12 May 2020. 95

Figure A 8. Photos of wheat crop taken on 20 May 2020. 96

Figure A 9. Photos of wheat crop taken on 27 May 2020. 97

Appendix B – Pix4D Processing

Figure B 1. Example of Pix4Dmapper flight pattern map view. Each red circle represents the UAV’s location with the sensor taking a picture. The diagonal flight patterns are due to the UAV returning to the starting point for a battery change midway through and returning upon completion of the specified flight area..... 98

Figure B 2. Example of the Pix4Dmapper rayCloud view of a field. Each blue-green dot is a representation of a UAV image tied to metadata (e.g., camera positions, geographic coordinate). The black and white layer beneath is an orthomosaic of all images taken..... 99

Appendix C - Topographic Metrics Processing

Figure C 1. Example of System of Automated Geoscientific Analyses (SAGA) interface and topographic metric processing results..... 100

Figure C 2. Example of the corn field aspect layer created in SAGA. 100

Figure C 3. Example of the corn field slope map created in SAGA..... 101

Figure C 4. Example of the corn field topographic wetness index map created in SAGA. . 101

Appendix D – LAI Layers

Figure D 1. Example of LAI layer for wheat field W2 on 16 May 2020. 102

Figure D 2. Example of LAI layer for wheat field W2 on 21 May 2020. 102

Appendix E – R Code

Figure E 1. R code of Random Forest and Support Vector regression models. 103

Figure E 2. R code for building a prediction image from multi-layered raster images. 104

Appendix F – Copyrighted Material and Permissions

Glossary

Ensemble learning: a machine learning method using multiple algorithms to improve predictive performance of models

Machine learning: the study and use of computer systems that can improve outcome predictions automatically by analyzing patterns in data with algorithms; a type of artificial intelligence.

Orthomosaic: a detailed, accurate photo representation of an area made by combining many smaller images mosaicked together

Precision agriculture: Management strategy that gathers, processes, and analyzes temporal, spatial and individual data and combines it with other information to support management decisions according to estimated variability for improved resource use efficiency, productivity, quality, profitability, and sustainability of agricultural production.

Pix4D: a suite of photogrammetry software designed for drone mapping applications including orthomosaicking images, generating 3D models, and more.

R programming language: a free programming language and software environment for statistical computing, graphics, and analysis.

RStudio: an Integrated Development Environment for R programming language; designed with a user-friendly interface for data cleaning, analysis, and visualization.

Random Forests (RF): an ensemble learning method for classification and regression using a multitude of decision trees to determine the most common or average of outputs.

Real-Time Kinematic (RTK) positioning: highly accurate spatial positioning system, up to centimeter-level, using global navigation satellite systems for real-time location correction. Commonly used application includes unmanned aerial vehicle navigation.

System for Automated Geoscientific Analyses (SAGA): free, geographic information system open-source software for spatial data processing and analysis.

Support Vector Machines / Regression (SVM / SVR): a supervised machine learning method for classification and regression using decision boundaries [hyperplanes] to classify or predict data points.

Unmanned Aerial Vehicle (UAV): an aircraft without any human pilot, crew, or passengers on board; part of an unmanned aircraft system with a ground-based controller and communication system.

Chapter 1

1 Introduction

Agriculture is the science of cultivating land, crops, and livestock for food and materials. Globally, agriculture is crucial to food security, economic development, and social structures (United Nations, 2015). In Canada, agriculture contributes over \$110 billion annually to the country's gross domestic product (GDP) (Agriculture and Agri-Food Canada, 2018). Over 2.3 million Canadians are employed in the industry, and Canada is the 5th largest agricultural exporter in the world. As the world population is projected to continue growing, for agriculture to meet increasing demands requires sustainable production, adaptability to changing climates, and better methods of farming. During the 2021 growing season, Western Canada experienced much drier conditions and higher-than-average temperatures, leading to very low soil moisture (Statistics Canada, 2021). In the Crop Condition Assessment Program (CCAP), results indicated many provinces had crops at peak health as much as four weeks earlier than normal before declining quickly with lack of moisture and high heat. Since Statistics Canada began monitoring crop conditions in 1987 with satellite imagery, these conditions have not been seen before. Canada's yield projections for canola, soybeans, barley, and oats are down substantially, with the greatest decrease in wheat production at 34.8% lower compared to the previous year (Statistics Canada, 2021). Of principal crops only corn yield is projected to increase slightly by 0.8% as Ontario, the major grain corn-producing province, had favorable growing conditions in 2021.

Research into agriculture practices have highlighted the severe environmental impact they can have; excess fertilizers contaminating water, greenhouse gas (GHG) emissions, insecticides, fungicides, and herbicides disrupting ecosystems (Montzka, Dlugokencky, Butler, 2011, Balfoutis et al., 2017). Plant genetics, land, water, and soil are key factors to crop production, and protecting these resources are vital. The farming management technique of precision agriculture can meet these needs as it aims to optimize inputs such as fertilizer, water, and pesticides to improve quality and yield of crops, reduce costs to farmers, and minimize negative environmental impacts (Das et al., 2015).

1.1 Background

1.1.1 Precision Agriculture & Nitrogen

From the International Society of Precision Agriculture (2019):

Precision agriculture is a management strategy that gathers, processes, and analyzes temporal, spatial and individual data and combines it with other information to support management decisions according to estimated variability for improved resource use efficiency, productivity, quality, profitability, and sustainability of agricultural production.

Within the field of precision agriculture, the efficient management of nitrogen fertilizers is one of the main goals (Good et al., 2004). Nitrogen is vital to plants as it is a component of amino acids for proteins of plant cells, enzymes, nucleic acids, chlorophyll, and more, allowing for plants to grow, reproduce, and absorb energy from light to perform photosynthesis (Marschner, 2011). As such, nitrogen content affecting above-ground plant tissue is an important indicator of crop health. Nitrogen is also one of the most expensive nutrients to supply and commercial fertilizers represent the major cost in plant production (Bryant et al. 2000; Statistics Canada, 2021). Out of the total fertilizers used in Canadian agriculture, 75% are nitrogen fertilizers including ammonia, urea, urea ammonium nitrate, ammonium nitrate, ammonium sulphate, monoammonium phosphate, and diammonium phosphate (Statistics Canada, 2021). Rates of nitrogen fertilizer application and type used are dependent on the crop variety, cultivar, desired yield, and nitrogen present in the soil.

In global studies, crops absorb an estimated <50% of nitrogen fertilizer applied (Bryant *et al.* 2000). This nitrogen can be leached from the soil and contaminate surface and ground water leading to harmful effects on human health (Wolfe and Patz, 2002), algal blooms and hypoxia in coastal waters (Rabalais, 2001), and denitrification that results in greenhouse gas emissions (Beaulieu et al., 2010). In general, excessive nutrients can reduce crop yield as well (Tan & Reynolds, 2003). Nutrients that have been added beyond the critical level of maximum growth continues to accumulate in the plant tissue without any further yield increase. Commonly in grain crops such as wheat, excessive nitrogen can cause plant stems to elongate leading to lodging – the stems bend over making it difficult to harvest, increases

chance of grain moisture, disease, and often reduces yield significantly (Foulkes et al., 2011). Crops such as corn can show little visual symptoms when above-optimal rates of nitrogen have been applied, which makes it costly for producers to recognize the issue (Marshner, 2011). Usually, nitrogen deficiency can be noted from chlorosis, the condition in which leaves yellow as the plant's chlorophyll content drops (Drew & Sisworo, 1977). With reduced photosynthetic activity, the plant will not reach peak health and yield will be low. Lack of sufficient nitrogen can also affect protein content in crops including wheat. As wheat protein content is a major component in Canada's wheat grading system, this can have great negative economic impact on farmers (Canadian Grain Commission, 2021).

With better management of nitrogen fertilizers, not only can costs and negative environmental impact be minimized but yield and quality can increase. An example would be fertilizer application on crops based on nutrient maps, such as that of nitrogen, to aid in technologically controlled input by farming machines. Input of the georeferenced crop nitrogen information will allow for only the necessary amounts of fertilizer to be distributed among the varying field conditions, preventing nitrogen over application or under application.

1.1.2 Remote Sensing Techniques

Information technology is key to precision agriculture, including the use of Global Positioning Systems (GPS), Geographic Information Systems (GIS), sensors, satellites, and Unmanned Aerial Vehicles (UAV) to gather precise data of fields (Zhu et al., 2008). Remote sensing is the process of obtaining information of the earth's surface, through data acquired by a device at a distance from the targeted area (Lillesand et al., 2015). In precision agriculture, remote sensing is especially useful without the need to make physical contact with plants to gather valuable crop data for analysis. Access to the technology, information, and interpretation allows for better agricultural data for government policies, industry management, producers to have data for variable-application methods, reducing inputs, and managing fewer crop acres while maximizing yield/acre.

Electromagnetic radiation of different wavelengths interacts with natural surfaces, and unique spectral properties of different land cover can be used for their identification and

classification in an image (Figure 1-1). In precision agriculture, spectral imaging provides an opportunity to derive useful information about the physical and biological characteristics of crops such as leaf pigment, cell structure, and water content to determine the health of the vegetation (Jones & Vaughn). There is much successful research on the use of spectral imaging for drought stress, pathogen detection, weed detection, nutrient status, phenology, and yield prediction (Sishodia et al. 2020).

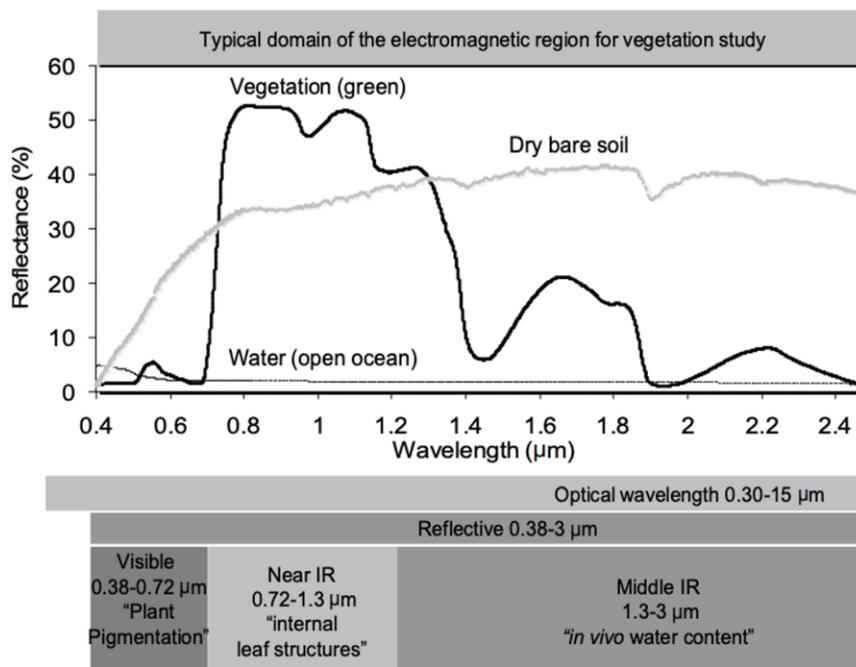


Figure 1-1. Generalized spectra characteristics of common land surface types in the visible and infrared range. Adapted from Gosa, 2009.

For precision agriculture, remote sensing systems are based on sensor platform and type of sensor. Common remote sensor types include red-green-blue (RGB), thermal, multispectral, and hyperspectral (Sishodia, 2020). From the study by Maes & Steppe (2019) on suitability of sensors for different precision agriculture applications, for nutrient status evaluation multispectral and hyperspectral imaging have highest performance. Multispectral and hyperspectral imaging is differentiated by the number and width of bands. A multispectral sensor has three to 10 bands on average, while a hyperspectral sensor is made of hundreds of bands. As differences in plant biophysical characteristics

can be within mere nanometers, hyperspectral sensors have a higher chance of capturing the spectral differences. However, hyperspectral sensors are expensive, generate large datasets requiring powerful computation and are complex to process and interpret (Jones & Vaughn, 2010). Multispectral sensors have lenses capturing distinct regions of the electromagnetic spectrum (e.g., blue, green, red, red-edge, near-infrared). When tailored for applications such as crop nutrient status, multispectral sensors are much more cost-efficient and can be just as effective in comparison to hyperspectral (Liu et al., 2018). RGB sensors have low spectral resolutions in comparison to multi and hyperspectral, but can have high spatial resolutions at low costs; applications include creating general orthomosaic maps, digital elevation models (DEMs), etc. Thermal sensors often only have one band covering the longwave infrared region (7000-12000 nm) and are most useful in applications such as drought stress and pathogen detection (Maes & Steppe, 2019).

From spectral data, most common methods of crop monitoring use vegetation indices. Vegetation indices (VIs) are mathematical calculations of canopy reflectance at specific visible and near-infrared wavelengths, and have been used in agricultural research to monitor and predict vegetation properties, such as nitrogen levels (Eitel *et al.*, 2007, Masclaux *et al.*, 2010, Frels, 2018). Many VIs are developed based on the inverse relationship between red and near-infrared (NIR) reflectance in healthy, green vegetation (Gomez, 2018). High spectral resolution data from remote sensing offers opportunities to select optimal wavelength bands to predict the chemical properties in plants. It should be kept in mind that VIs are developed from various remote sensing platforms (satellite, UAV, ground-based), and are mostly used on similar platform data (Lillesand et al., 2015). Commonly used VIs developed for agriculture nitrogen research include normalized difference vegetation index (NDVI), green NDVI (gNDVI), and Double-Peak Canopy Nitrogen Index (DCNI) (Gomez, 2018). VIs are also developed for chlorophyll estimation, as nitrogen is a key component of chlorophyll that absorbs more than 70% of blue and red incident spectral bands and reflects green and NIR bands (Daughtry et al., 2000). Examples of these VIs include the Modified Chlorophyll Absorption ratio index (MCARI), MCARI /Modified Triangular Vegetation Index-Improved (MCARI/MTVI2), and the Transformed Chlorophyll Absorption in Reflectance Index/Optimized Soil-Adjusted Vegetation Index (TCARI/OSAVI).

1.1.3 Remote Sensing Platforms

In precision agriculture, models for nitrogen monitoring and crop yield estimation has largely been through data collected by optical satellites (Agriculture and Agri-Food Canada, 2015). As demand for timely, accurate, and cost-effective data on the earth's surface has increased, numerous satellite systems have launched. Examples of current optical satellites in operation include Landsat 8, RapidEye, GeoEye-1, Sentinel-2, all of which have been used in studies on agriculture management (Chen et al., 2010).

Limitations in optical satellite imagery include the low spatial sensitivity, as the spatial resolution in the range of meters allows for analysis of large areas but not as well suited for individual fields. The temporal sensitivity can be rather low, as in the case of Landsat 8 - within a 16-day revisit time, crops will have changed significantly and valuable information on the different stages of growth would not be obtained. In addition to factors such as geometric distortion, atmospheric distortion, and cloud cover obscuring view of the land, processing to correct the images are often time consuming, costly, and the quality of images may not be sufficient for a user's purpose (Bolstad, 2016). However, with technological advances new satellite systems have higher spatial, temporal, and spectral resolutions. An example would be PlanetScope satellite constellation consisting of over 130 satellites, 3 to 5 m spatial resolutions with daily revisit.

Unmanned aerial vehicle (UAV) based remote sensing can provide low cost, high spatial and temporal resolution data for crop management. Individuals can operate a UAV using programmed routes specific to fields and collect images up to centimeter resolutions. They can be flown to capture more frequent image data and offers flexibility in operation for times when weather is ideal. In comparison to optical satellite images, UAV images can be more precise in capturing smaller areas such as crop fields. This makes it ideal for obtaining real time data, and field management can be conducted in a timely and accurate manner according to the needs of the crop (Hunt et al., 2010). The overall lower cost of a UAV, the system components, data collection and processing as opposed to satellites had made it a preferred method of precision agriculture remote sensing (Hunt et al., 2010).

1.2 Research Questions

Remote sensing is an established aspect of precision agriculture with the use of spectral imagery to retrieve crop characteristics including yield, biomass, and crop health (Breunig et al, 2020; Mulla, 2013; Sadeh et al., 2020). Most commonly, VIs are derived from spectral data to estimate crop parameters such as leaf area index (LAI), chlorophyll content, pathogen detection, and nutrient status (Sishodia et al., 2020). Studies have shown significant correlations between crop spectral variables derived from remotely sensed imagery and crop nitrogen content (Jiang et al., 2019; Zheng et al., 2018). Many studies are based on single or combinations of different spectral indices' relationships with crop nitrogen content, noting variation in the relationships at different stages of crop growth (Jiang et al., 2019). With the development of new remote sensing technologies, processing methods, and computing capabilities, estimation models for crop nitrogen can be improved. Machine learning is an area of research interest as they are capable of developing accurate crop monitoring models for large, nonparametric, nonlinear datasets (Chen et al., 2010; Lee et al., 2020.) However, most current literature in using remote sensing data and machine learning have only considered spectral information for crop nitrogen modelling (Schirrmann et al., 2016). As a crop's nitrogen status can be affected by many factors including fertilizer application methods, soil nutrient supply, water availability, field micro-topography, nitrogen prediction models may be improved if these plant physiological and environmental variables are considered. Thus, the research questions of this thesis are:

- (i) What is the importance of the relationships between the canopy nitrogen weight and plant physiology, soil properties, topographic factors, and spectral variables, respectively?
- (ii) Can the addition of non-spectral variables in machine learning models be used to improve the prediction of canopy nitrogen weight? Which machine learning model results in the highest accuracy for predicting canopy nitrogen weight, and with which combination of non-spectral and/or spectral variables?

- (iii) Using the UAV-based and satellite-based prediction models, can prediction maps be created, and will they provide useful indications of nitrogen values?

1.3 Research Objectives

The objective of this thesis is to evaluate machine learning modelling techniques to predict canopy nitrogen values of corn and wheat crops using UAV and satellite-based imagery. This empirical research is focused on the plant physiology, soil properties, topographic factors, and spectral variables in relation to the crop's canopy nitrogen. The objectives of this study include:

- (i) Generate machine learning regression models to predict canopy nitrogen weight of corn and wheat fields using plant height, topographic metrics, soil chemical properties, soil moisture conditions, multispectral UAV-based and satellite-based imagery.
- (ii) Determine the optimal combination(s) of spectral variable(s), crop plant physiology variables, and/or environmental conditions (soil, water, topographic data) for corn and wheat canopy nitrogen estimation and prediction.
- (iii) Create nitrogen prediction maps for corn and wheat canopy nitrogen weight prediction using UAV-based and satellite-based multispectral imagery.

1.4 Thesis Structure

This thesis is structured by an introduction, two academic journal papers, and a conclusion. Chapter 1 introduces the background information for this thesis, reviewing precision agriculture and associated remote sensing applications, research questions and objectives. Chapter 2 is a published journal paper on the use of plant, soil, topographic, and UAV-based spectral variables to estimate canopy nitrogen weight in corn crops. Chapter 3 is the second paper on the use of non-spectral variables, UAV-based and satellite based spectral variables to estimate canopy nitrogen weight in wheat crops. Chapter 4 is the conclusion of the completed objectives for this thesis, including limitations and suggestions for future research.

1.5 Study Areas

Southwestern Ontario is a fertile agricultural region, contributing significant portions of Canada's dominant crops including soybean, grain corn, and wheat (Ontario Ministry of Agriculture, Food, and Rural Affairs [OMAFRA], 2021b). Hence for this thesis, corn and wheat crops were chosen for the chapter studies. The study areas for this thesis are near London, Ontario (Figure 1-2). The corn field study site for chapter 2 is in Central Elgin, about 25 km south of London. The three wheat fields W1, W2, and W3 study sites for chapter 3 are near Mount Brydges, about 25 km southwest of London. In the continental climate zone of southwestern Ontario, growing seasons are typically from April to October with warm, humid summers (OMAFRA, 2021a). The topography of the area is overall flat and experiences frequent precipitation during summer months.

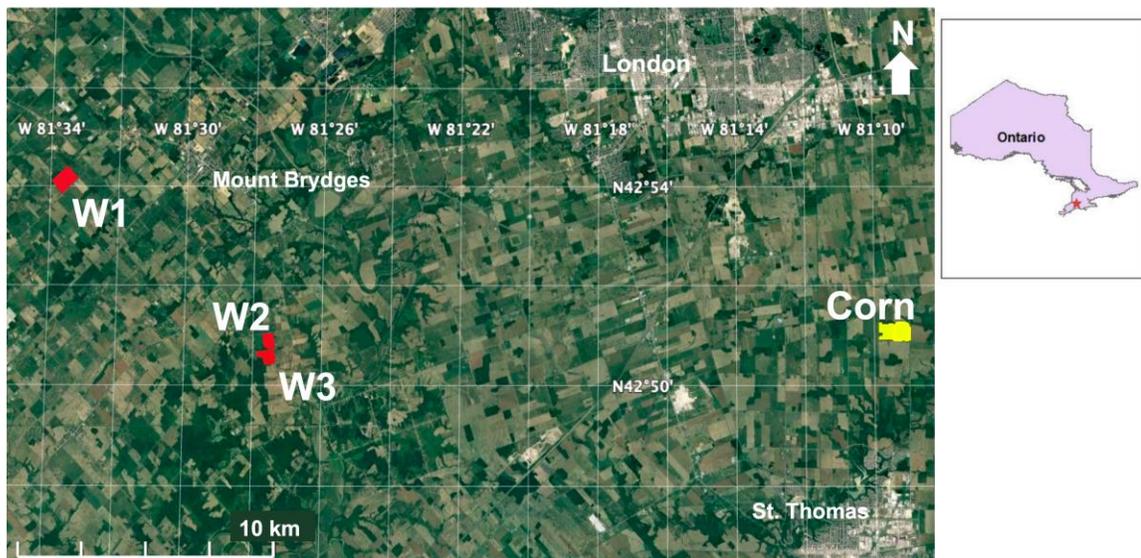


Figure 1-2. Study areas located near London, Ontario. The corn field is indicated in yellow, and the three wheat fields are indicated in red. Background image is from Google Earth.

1.6 References

- Agriculture and Agri-Food Canada. (2018). *Overview of the Canadian agriculture and agri-food sector 2018*. <https://agriculture.canada.ca/en/canadas-agriculture-sectors/sector-overviews-data-and-reports/overview-canadian-agriculture-and-agri-food-sector-2018>
- Balafoutis, A., Beck, B., Fountas, S., Vangeyte, J., Wal, T., Soto, I & Eory, V. (2017). Precision agriculture technologies positively contributing to GHG emissions mitigation, farm productivity and economics. *Sustainability*, 9(8), 1339. <https://doi.org/10.3390/su9081339>
- Breunig, F. M., Galvão, L. S., Dalagnol, R., Santi, A. L., Della Flora, D. P., & Chen, S. (2020). Assessing the effect of spatial resolution on the delineation of management zones for smallholder farming in southern Brazil. *Remote Sensing Applications: Society and Environment*, 19, 100325. <https://doi.org/10.1016/j.rsase.2020.100325>
- Bryant, C. R., Smit, B., Brklacich, M., Johnston, T. R., Smithers, J., Chiotti, Q., & Singh, B. (2000). Adaptation in Canadian agriculture to climatic variability and change. In *Societal adaptation to climate variability and change* (pp. 181-201). Springer, Dordrecht. https://doi.org/10.1007/978-94-017-3010-5_10
- Bolstad, P. (2016). *GIS fundamentals: A first text on geographic information systems*. Eider (PressMinnesota).
- Canadian Grain Commission. (2021). *Official Grain Grading Guide*. <https://www.grainscanada.gc.ca/en/grain-quality/official-grain-grading-guide/>
- 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. <https://doi.org/10.1016/j.rse.2010.04.006>
- Das, J., Cross, G., Qu, C., Makineni, A., Tokekar, P., Mulgaonkar, Y., & Kumar, V. (2015, August). Devices, systems, and methods for automated monitoring enabling precision agriculture. In *2015 IEEE International Conference on Automation Science and Engineering (CASE)* (pp. 462-469). IEEE. <https://doi.org/10.1109/CoASE.2015.7294123>
- Drew, M. C., & Sisworo, E. J. (1977). Early effects of flooding on nitrogen deficiency and leaf chlorosis in barley. *New Phytologist*, 79(3), 567-571. <https://doi.org/10.1111/j.1469-8137.1977.tb02241.x>
- Foulkes, M. J., Slafer, G. A., Davies, W. J., Berry, P. M., Sylvester-Bradley, R., Martre, P., Calderini, D., Griffiths, S., & Reynolds, M. P. (2011). Raising yield potential of wheat. III. Optimizing partitioning to grain while maintaining lodging resistance. *Journal of experimental botany*, 62(2), 469-486. <https://doi.org/10.1093/jxb/erq300>

- Good, A. G., Shrawat, A. K., & Muench, D. G. (2004). Can less yield more? Is reducing nutrient input into the environment compatible with maintaining crop production?. *Trends in plant science*, 9(12), 597-605.
<https://doi.org/10.1016/j.tplants.2004.10.008>
- Gosa, A.G. (2009). *Remote sensing of leaf area index: enhanced retrieval from close-range and remotely sensed optical observations*. [Doctoral dissertation, University of Helsinki] Helda Dissertation Repository.
<http://urn.fi/URN:ISBN:978-952-10-5873-8>
- International Society of Precision Agriculture. (2019). *Precision Ag Definition*.
<https://www.ispag.org/about/definition>
- Jiang, J., Cai, W., Zheng, H., Cheng, T., Tian, Y., Zhu, Y., ... & Yao, X. (2019). Using digital cameras on an unmanned aerial vehicle to derive optimum color vegetation indices for leaf nitrogen concentration monitoring in winter wheat. *Remote Sensing*, 11(22), 2667. <https://doi.org/10.3390/rs11222667>
- Jones, H.G.; Vaughan, R.A. (2010). *Remote sensing of vegetation: principles, techniques, and applications*, 1st ed.; University Press: New York, U.S.A.
- 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.
<https://doi.org/10.3390/rs12132071>
- Lillesand, T., Kiefer, R. W., & Chipman, J. (2015). *Remote sensing and image interpretation*. John Wiley & Sons.
- Liu, S., Li, L., Gao, W., Zhang, Y., Liu, Y., Wang, S., & Lu, J. (2018). Diagnosis of nitrogen status in winter oilseed rape (*Brassica napus* L.) using in-situ hyperspectral data and unmanned aerial vehicle (UAV) multispectral images. *Computers and Electronics in Agriculture*, 151, 185-195.
<https://doi.org/10.1016/j.compag.2018.05.026>
- 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.
<https://doi.org/10.1016/j.tplants.2018.11.007>
- Marschner, H. (2011) *Marschner's mineral nutrition of higher plants*, 2nd ed.; Academic Press: Adelaide, Australia.
- Montzka, S. A., Dlugokencky, E. J., & Butler, J. H. (2011). Non-CO₂ greenhouse gases and climate change. *Nature*, 476(7358), 43-50.
<https://doi.org/10.1038/nature10322>
- Mulla, D. J. (2013). Twenty-five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems engineering*, 114(4), 358-371. <https://doi.org/10.1016/j.biosystemseng.2012.08.009>
- Ontario Ministry of Agriculture, Food, and Rural Affairs. (2021a). *Climate zones and planting dates for vegetables in Ontario*. Retrieved June 29, 2021, from <http://www.omafra.gov.on.ca/english/crops/facts/climzoneveg.htm>

- Ontario Ministry of Agriculture, Food, and Rural Affairs. (2021b). *Statistical summary of Ontario agriculture*. Retrieved June 29, 2021, from http://www.omafra.gov.on.ca/english/stats/agriculture_summary.htm
- Rabalais, N. N., Turner, R. E., & Wiseman Jr, W. J. (2001). Hypoxia in the Gulf of Mexico. *Journal of environmental quality*, 30(2), 320-329. <https://doi.org/10.2134/jeq2001.302320x>
- Sadeh, Y., Zhu, X., Dunkerley, D., Walker, J. P., Zhang, Y., Rozenstein, O., Manivasagam V.S., & Chenu, K. (2021). Fusion of Sentinel-2 and PlanetScope time-series data into daily 3 m surface reflectance and wheat LAI monitoring. *International Journal of Applied Earth Observation and Geoinformation*, 96, 102260. <https://doi.org/10.1016/j.jag.2020.102260>
- Schirrmann, M., Giebel, A., Gleiniger, F., Pflanz, M., Lentschke, J., & Dammer, K. H. (2016). Monitoring agronomic parameters of winter wheat crops with low-cost UAV imagery. *Remote Sensing*, 8(9), 706. <https://doi.org/10.3390/rs8090706>
- Sishodia, R. P., Ray, R. L., & Singh, S. K. (2020). Applications of remote sensing in precision agriculture: A review. *Remote Sensing*, 12(19), 3136. <https://doi.org/10.3390/rs12193136>
- Statistics Canada. (2021). *Farm input price index, quarterly*. <https://www150.statcan.gc.ca/t1/tb1/en/tv.action?pid=1810025801&pickMembers%5B0%5D=1.1&cubeTimeFrame.startMonth=01&cubeTimeFrame.startYear=2016&cubeTimeFrame.endMonth=04&cubeTimeFrame.endYear=2021&referencePeriods=20160101%2C20210401>
- Statistics Canada. (2021). *Fertilizer shipments to Canadian agriculture and export markets, by product types and fertilizer year, cumulative data*. <https://www150.statcan.gc.ca/t1/tb1/en/tv.action?pid=3210003801>
- Statistics Canada. (2021). *Production of principal field crops, July 2021*. Retrieved August 30, 2021, from <https://www150.statcan.gc.ca/n1/daily-quotidien/210830/dq210830b-eng.htm>
- Tan, C. S., & Reynolds, W. D. (2003). Impacts of recent climate trends on agriculture in southwestern Ontario. *Canadian Water Resources Journal*, 28(1), 87-97. <https://doi.org/10.4296/cwrj2801087>
- United Nations. (2015). *Sustainable development goals knowledge platform: Food security and nutrition and sustainable agriculture*. <https://sustainabledevelopment.un.org/topics/foodagriculture>
- Wolfe, A. H., & Patz, J. A. (2002). Reactive nitrogen and human health: acute and long-term implications. *Ambio: A journal of the human environment*, 31(2), 120-125. <https://doi.org/10.1579/0044-7447-31.2.120>
- Zheng, H., Li, W., Jiang, J., Liu, Y., Cheng, T., Tian, Y., Zhu, Y., Cao, W., Zhang, Y., & Yao, X. (2018). A comparative assessment of different modeling algorithms for estimating leaf nitrogen content in winter wheat using multispectral images from an unmanned aerial vehicle. *Remote Sensing*, 10(12), 2026. <https://doi.org/10.3390/rs10122026>

Zhu, Y., Yao, X., Tian, Y., Liu, X., & Cao, W. (2008). Analysis of common canopy vegetation indices for indicating leaf nitrogen accumulations in wheat and rice. *International Journal of Applied Earth Observation and Geoinformation*, 10(1), 1-10. <https://doi.org/10.1016/j.jag.2007.02.006>

Chapter 2

2 Evaluation of Soil Properties, Topographic Metrics, Plant Height, and Unmanned Aerial Vehicle Multispectral Imagery using Machine Learning Methods to Estimate Canopy Nitrogen Weight in Corn

2.1 Introduction

Agriculture is an important industry as the basis of food security, and as a significant aspect of the world economy. However, factors such as rapidly increasing global demand, fluctuations in production due to climate change, and a greater awareness of the negative environmental impact of agriculture on surrounding ecosystems, contribute to an increasing need for more efficient and sustainable farming practices. Especially in Canada, where agriculture is a significant industry, developing agricultural methods to be adaptable and resilient is necessary (Tan & Reynolds, 2003). This is possible through precision agriculture (PA), a management technique that selectively applies crop farming resources such as fertilizer, water, pesticides, and herbicides based on the plant needs within a field (Bongiovanni & Lowenberg-Deboer, 2004; Mazaclaix-Daubresse et al., 2010; Sishodia et al., 2020).

One of the main fields of applications of precision agriculture is the management of nitrogen fertilizers (Good et al., 2004; Fageria et al., 2010). Nitrogen is an essential macronutrient to plants, as a major constituent of organic material, enzymic processes, and oxidation-reduction reactions (Marschner, 2011). As such, nitrogen content in above-ground plant tissue is an important indicator of crop health and yield potentials. Several global studies demonstrate that the mean nitrogen recovery efficiency by annual crops was less than 50% of the amount of fertilizer applied (Fageria, 2009; Fageria et al., 2010). Nitrogen is one of the most expensive nutrients to supply, and commercial fertilizers represent a major cost in plant production (Bryant et al., 2000). Rates of nitrogen fertilizer application depend on the crop type, desired yield, nitrogen present in the soil, and subsequently in the plants (Marschner, 2011). Excess nitrogen can reduce crop yield and can be leached from the soil, contaminating surface and groundwater, leading to harmful

effects on human health and ecosystem consequences such as algal blooms and hypoxia in water bodies (Wolfe et al., 2002). The United Nations Food and Agriculture Organization identifies classes of agricultural climate adaptation, one major class being management of field operation inputs including fertilizers (FAO, 2007). As such, optimizing the management of nitrogen fertilizers is an important field of research as new methods and technology are developed to improve nutrient use efficiency, quality, and crop yield while minimizing significantly negative environmental impacts and cost of production.

Literature does include much research on crop canopy nitrogen retrieval using UAVs, but model parameters are often focused on spectroscopy with the use of vegetation indices and spectral remote sensors (Lee et al., 2020). PA incorporates the use of many different types of spatial technologies such as geographic information systems (GIS), precision machinery, and remote sensing imagery to ground-based data collection (Zhu et al., 2008). In PA, remote sensing imagery is especially useful because it does not require physical or destructive contact with plants to gather valuable crop information. The spectral information provided by the imagery can be transformed into vegetation indices (VIs). VIs are mathematical combinations or transformations of spectral bands that have been widely used in agricultural research because they allow for deriving of specific plant properties such as chlorophyll or nutrient content by taking advantage of the differential spectral properties of plants in the visible and near-infrared wavelengths (Eitel et al., 2007, Sripada et al., 2008; Frels et al., 2018). The resulting VI data can then provide timely information for monitoring field conditions and crop health, allowing for the optimal number of resources to be placed where they are needed, when they are needed.

In PA, crop monitoring has largely been conducted using optical satellites (Agriculture & Agri-Food Canada, 2007). As demand for timely, accurate, and cost-effective data on the earth's surface increased in the last few decades, numerous satellite systems have been launched. Examples of optical satellites in operation include Landsat 8 (since 2013) and Sentinel-2 (since 2015), which have been used in studies on agriculture management (Sishodia et al., 2020). The Landsat program, which began in 1972 with the launch of Landsat 1, is the longest-running program for satellite imagery of the earth (USGS, 2021). Landsat 8 Operational Land Imager has nine spectral bands including visible, near-infrared,

and shortwave infrared, with varying spatial resolutions of 15 to 30 m. Taking more than 700 scenes a day, it has a 16-day revisit time to the same area. Sentinel-2 has 13 spectral bands in the visible, near-infrared, and shortwave infrared with varying spatial resolutions of 10 m, 20 m, and 60 m (ESA, 2020). With the constellation of twin satellites, the revisit cycle over an area is five days.

Limitations in optical satellite imagery include low spatial sensitivity, as the spatial resolution in the range of meters allows for analysis of larger-scale regional or national areas but is too coarse for small-scale crop fields. The temporal sensitivity can be rather low, as in the case of Landsat 8; within a 16-day revisit time, crops would have changed significantly and valuable information on the different stages of growth would not be obtained. In addition to factors such as geometric distortion, atmospheric distortion, and cloud cover obscuring view of the land, advanced processing expertise may be required to ensure sufficient image quality (Mulla, 2013). In comparison, UAV-based remote sensing can provide lower cost and higher spatial and temporal resolution data for crop management. Individuals with basic training can operate a UAV using programmed routes and collect images with <10 cm resolutions (Harwin & Lucieer, 2012). They can be flown to capture more frequent image data, including monitoring each significant stage of crop growth and offer flexibility in operation for times when weather is most suitable. This makes them ideal for field management conducted in a timely and accurate manner according to the needs of the crop (Hunt et al., 2010). Compared to satellites, overall UAV-based systems are often lower in cost for data collection and processing. As such, the use of UAV imagery in PA has become a research area of great interest due to its potential for larger environmental and economic impacts (Sishodia et al., 2020).

Corn was selected for this study because it is among the most grown crops in Ontario (Agriculture and Agri-Food Canada, 2007). Recent studies have tested the use of linear regression, Random Forest (RF), and Support Vector Regression (SVR) models in UAV-based canopy nitrogen weight prediction models (Lee et al., 2020). Although linear regression is a commonly used method to predict nitrogen, some VIs (e.g., NDVI) may saturate beyond the early growth crop stages and models may have reduced accuracy due to multicollinearity (Chen et al., 2010). By contrast, machine learning-based regression

methods such as RF and SVR have been found to produce more accurate models compared to classical linear regression methods, as they are unaffected by the assumptions of linear regression (Chen et al., 2010). However, Lee et al. (2020) considered only UAV spectral information and canopy nitrogen weight prediction in their study. The nitrogen prediction may be improved if plant physiology, topographic metrics, and soil variables are included in the analysis, given that crop nitrogen highly depends on these variables (Marschner, 2011).

To make well-informed fertilization management decisions, knowledge about the plant nutrient supply, health, and several environmental factors such as water availability, soil quality, and micro-topography of a field are key. The objectives of this study include, (i) studying the relationship between the spatial variation of canopy nitrogen weight and factors such as plant height, topographic metrics, soil chemical properties, and soil moisture conditions within a corn field in Southwestern Ontario using multispectral UAV-based imagery; (ii) determining the optimal combination(s) of spectral variable(s), crop plant physiology variables, and/or environmental conditions (soil, water, topographic data) for corn canopy nitrogen estimation and prediction; and (iii) evaluating the temporal variation of nitrogen estimation and prediction during early growth stages of corn using UAV images and select variables.

2.2 Materials and Methods

2.2.1 Study Area & Data

The study site is in Central Elgin, Ontario, Canada (Figure 2-1). Fieldwork was conducted during June–July 2020 with an average temperature of 25 °C and high humidity averaging 71%, typical of southwestern Ontario’s humid continental climate zone. The study area is situated in a predominantly agricultural area, about 25 km south of London, Ontario’s urban center. The corn was planted in early April, began sprouting in early June, and was harvested in late October once the plants were fully dried.

Data was collected from a corn field about 75 ha large, sown with cultivars “DKC48-56RI”. Beginning in early June, a total of five sampling dates were collected with at least seven days between each visit as the crop reached different significant growth stages. Corn

growth stages were recorded following the Biologische Bundesanstalt, Bundessortenamt und Chemische (BBCH) scale (Meier, 1997). We acquired and used data during the corn crops' early growth periods (BBCH 0–49) in June for nitrogen estimation, which is especially important because fertilizers used during this time can have the greatest impact on the final quantity and quality of yield (Marshner, 2011; Hansen et al., 2015). During July to August, corn crops reached middle to late growth stages (BBCH 50+) and application of nitrogen fertilizer is not recommended after plants begin tasseling (Fageria et al., 2010). The plant slows root nitrogen uptake, beginning to translocate nitrogen from vegetation to the grains, and excess fertilizer can leach from the field (Vetsch & Randall, 2004). Corn plants reached full maturity in early September, and the crop was left to dry in the field before harvest in October.



Figure 2-1. Location of the corn field in Central Elgin, Ontario, Canada, over a Google Earth image.

Before selecting sample points, a DJI Phantom 4 Real-Time Kinematics (RTK) UAV was flown over the bare soil of the field. With the UAV connected to an RTK global navigation satellite system (GNSS) base station, images have precise positioning metadata with 1 cm

horizontal positioning accuracy and 1.5 cm vertical positioning accuracy. These images were used to create a bare earth digital elevation model (DEM). Using Google Earth, 40 sample points were selected in the field based on the DEM and the UAV imagery (Figure 2-2). The sample points had to cover the variation of field topography sloping down from west to east. Considering corn row directions heading north-south, navigating along rows was more efficient compared to against rows. As well, the following factors were considered for sample point placement: the large dimensions of the field (1.2 km × 0.7 km), the intensive labour required, and the time-sensitive nature of in-situ data collection and processing. The sample points were placed in groups of ten spaced 60 m from one another, with a distance between groups at least 200 m apart to include a representative sample distribution of the field. A minimum distance of 50 m from roads was to reduce possible effects of transportation pollution. The sample points were exported from Google Earth to a KML file and downloaded onto mobile devices. During the first fieldwork date using the KML file for navigation, red flags were placed at the sample points for accurate positioning in the following weeks.

At each sampling point, fresh biomass samples were destructively collected within a 2 m radius at a different spot each week to obtain samples representative of the weekly growth stages. Biomass sampling for corn involve cutting the plant at the stem base above ground. The number of plants within a 1 m² block around the sample point were counted, and two plants were collected and placed in plastic bags. The average distance between rows was 80 cm. Following fieldwork collection on the same day, the fresh biomass was weighed in grams then fully dried in an oven at 60 °C for 48–72 h. Dry biomass was weighed and leaves at the top of the plant constituting the canopy layer were separated for A&L Canada Laboratories plant analysis using the Laboratory Equipment Company (LECO) FP628 nitrogen determinate combustion method (AOAC, 2006). The process involves grinding biomass leaves into a fine powder, which can be passed through a 1 mm sieve, and the combustion method obtains the leaf nitrogen content percentage.

On every fieldwork date, at each sampling point within a 1 m² block, six measurements of plant height in centimeters were taken to calculate an average height. Detailed plant phenology was recorded to determine growth stages according to the Biologische

Bundesanstalt, Bundessortenamt und Chemische (BBCH) scale at each sample point, as there can be variation in the field depending on growing conditions (Vetsch & Randall, 2004). Six measurements of soil moisture were collected using an ML3 ThetaProbe (Delta-T Devices Ltd., England) and averaged (Delta-T, 2017). On the first fieldwork date of 8 June, soil samples in the 0–30 cm surface layer were collected and sent to A&L Canada Laboratories for Vittellus soil health analysis (A&L, 2020). The test results include values of soil nitrate-nitrogen, mineralizable nitrogen, water extracted soil nitrate, water extracted total nitrogen, soil textural class, and A&L’s Soil Health Index rating.

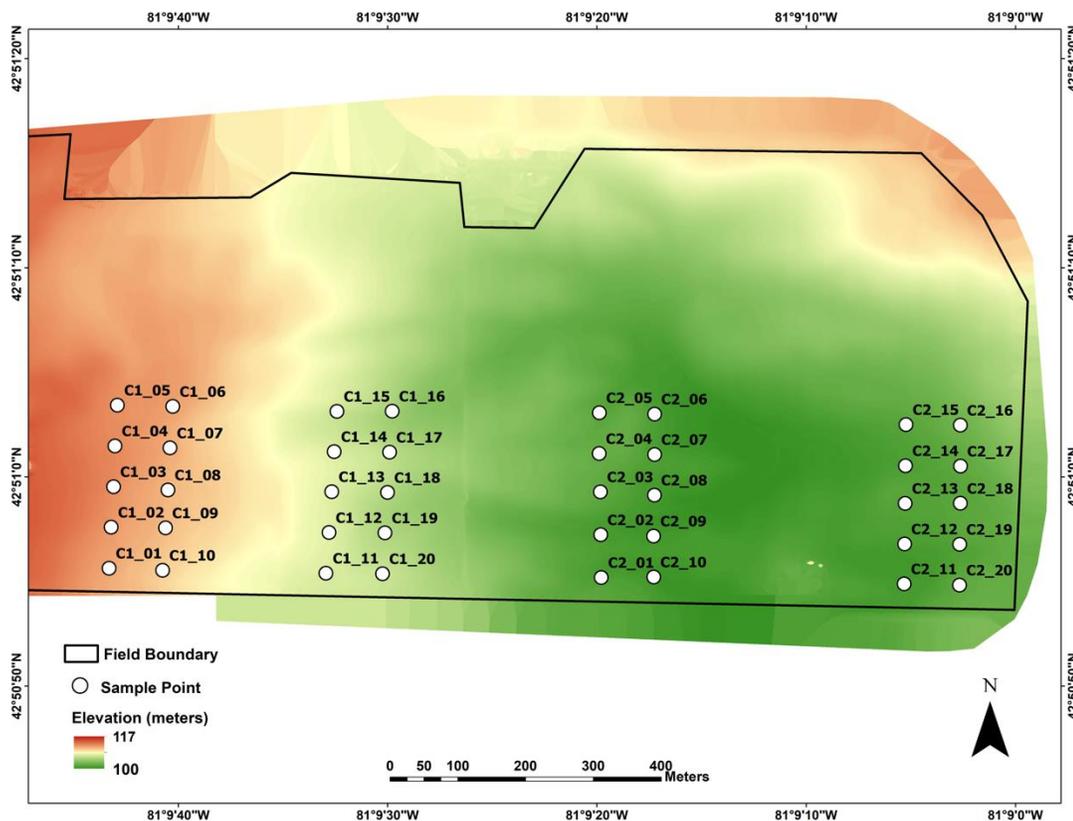


Figure 2-2. Digital elevation model showing the variation of the corn field's topography and the distribution of sample points.

2.2.2 UAV Imagery

For this study, two types of UAVs were used: a Da-Jiang Innovations (DJI) (DJI, China) Matrice 100 and a DJI Phantom 4 Real-Time Kinematics (RTK) (Figure 2-3). The DJI

Phantom 4 RTK was released in 2018, designed for centimeter-accurate horizontal and vertical positioning in images using a complementary metal-oxide-semiconductor (CMOS) sensor (DJI, 2018). Connected to an RTK global navigation satellite system (GNSS) base station, images with precise positioning metadata can be used to generate 3D point cloud datasets and digital elevation models (DEM). For this study, the DJI Phantom 4 RTK was flown at 30 m altitude as per manufacturer recommendations for optimal performance of the UAV's visual sensor system and RTK base station connection. The image resolution was 0.8×0.8 cm, and capture was set to 80% side and 80% front image overlap. Studies indicate that fine resolution (<10 cm) and high image overlap have higher success for mosaicking images together when crop canopy densifies through the season (Harwin & Lucieer, 2012).

First released in 2015, the DJI Matrice 100 is a model designed with a customizable aerial platform, ideal for research purposes of attaching small spectral sensors. Including batteries, it weighs 2431 g with a maximum take-off weight with a payload of 3600 g. DJI Matrice 100 carried a MicaSense RedEdge (MicaSense Inc., Seattle, WA, USA) narrowband multispectral camera (MicaSense, 2015). Multispectral imagery acquisition was aimed to be conducted on fieldwork dates before biomass collection. If the weather for the day was not ideal, flights were scheduled as soon as possible after that to maintain consistency with the plant physiology and field conditions (Table 2-1). The field flight plan was made in the "Pix4Dcapture" app, part of the Pix4D software suite (Pix4D S.A., Prilly, Switzerland), to cover the whole field in a zigzag pattern (Pix4D Documentation, 2020). Pix4Dcapture has the functions of adding custom UAV and sensor properties to calculate the flight plan's estimated total time, battery usage, and image resolution at selected altitudes. At the study field, wind and gust conditions > 60 m altitude were often greater than the UAV's manufacturer-recommended limits. Flight altitude was set at 60 m and based on the MicaSense camera's specifications (image width, sensor width, and focal length) the resulting image resolution was 4×4 cm, suitable for the scale of crop spectral analysis (Harwin & Lucieer, 2012). Image capture was set to 80% side and 80% front overlap. To streamline UAV flight patterns and maintain GPS connection with the controller, the study field was divided into two flight plans. Unfortunately, images from one flight on 26 June were corrupted and only data from half of the field were usable.

Images taken on and after 2 July could not be mosaicked in Pix4D due to the software limitations with recognizing and stitching the dense crop canopy at middle/late growth stages. Considering the nitrogen estimation for fertilizer management is most important in the early growth stages, data from 2 July onwards were omitted from the rest of this study.

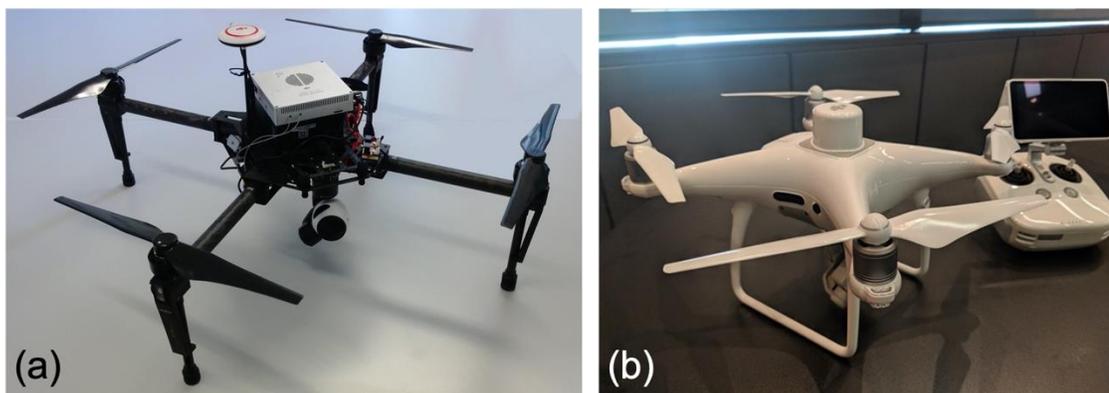


Figure 2-3. (a) DJI Matrice 100 UAV, with the MicaSense and RGB Gimbal cameras attached. (b) DJI Phantom 4 RTK UAV.

Table 2-1. Summary of Unmanned Aerial Vehicle (UAV) flight acquisitions with MicaSense RedEdge multispectral camera.

Flight Date	Air Temperature (°C)	BBCH* Growth Stage
8 June	26	10
16 June	25	15
26 June	28	31
2 July	32	34
9 July	33	45
15 July	29	53

*Biologische Bundesanstalt, Bundessortenamt und Chemische Industrie Scale [25]

2.2.3 UAV Image Processing

UAV images were processed according to the flowchart of Figure 2-4. DJI Phantom 4 RTK images, taken in April over the study site's bare soil, were inputted into Pix4Dmapper photogrammetry software to generate a continuous 3D point cloud dataset of the field. QGIS, an open-source geographic information software (GIS), was used to convert the point cloud dataset into a DEM in GeoTiff format (QGIS, 2021). The DEM enabled observation of topographic variation within the field, which can affect plant growth related

to landscape shape, soil structure, and water flow (Fageria et al., 2010). Topographic metrics were computed from the DEM with the System for Automated Geoscientific Analysis (SAGA), free, open-source software for spatial data analysis (Conrad et al., 2015). Topographic metrics exported in GeoTiff format included slope, aspect, profile curvature, plan curvature, and two topo-graphic wetness indices: (TWI) #1 using a Deterministic 8 algorithm, and TWI #2 using Multiple Flow Direction algorithm.

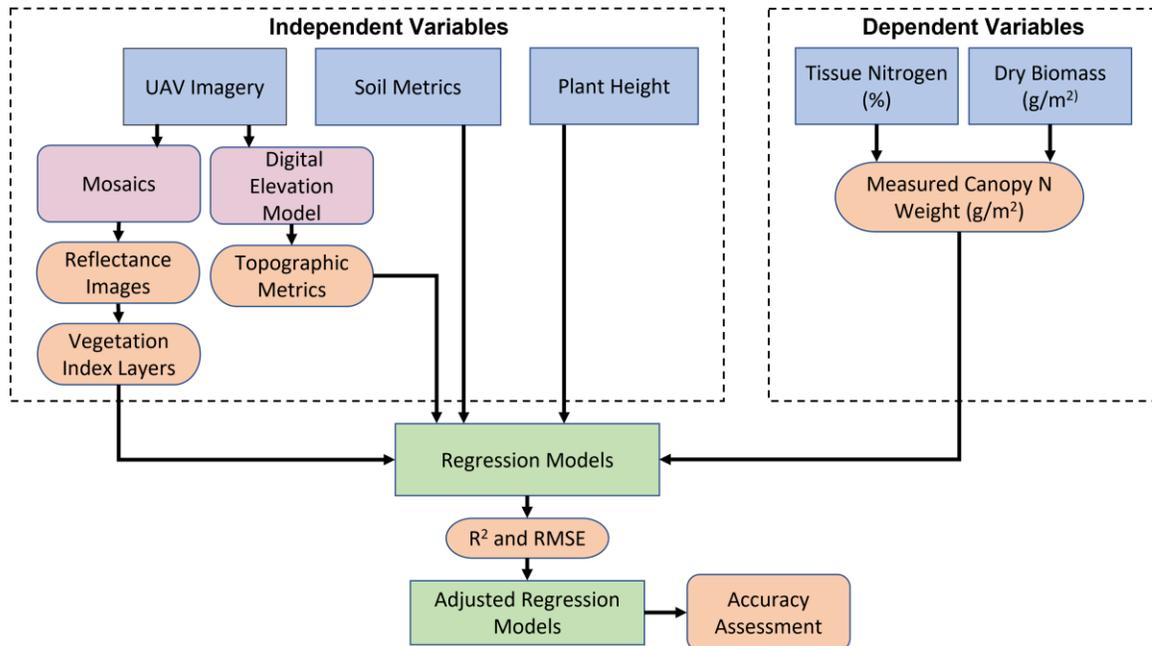


Figure 2-4. Flowchart of the methodology used in the study.

Multispectral images from the MicaSense camera were processed in Pix4Dmapper to create one orthomosaic image per band. Radiometric calibration of UAV images is important for the quality of image reflectance, taking into consideration the sensor influence and scene illumination. Before each flight, the MicaSense camera was positioned above a MicaSense Calibrated Reflectance Panel to acquire white reference images for each band. In Pix4Dmapper, the sensor settings, properties, and conditions can be obtained from the Exchangeable Image File Format (EXIF) metadata of the images. The white reference images and manufacturer-provided panel reflectance values were inputted in Pix4Dmapper processing options, enabling the software to calibrate and correct images' reflectance for each of the five bands. Pix4Dmapper then uses the Structure from Motion

(SfM) algorithm to correct image perspectives to stitch images together (Harwin & Lucieer, 2012). In addition to UAV image geolocation, Pix4Dmapper processing options include georeferencing orthomosaics with ground control points (GCP) to improve the absolute location accuracy. Five GCPs were positioned around the corn field area using black and white checkered boards, with coordinates obtained from a Global Positioning System (GPS) connected with the RTK. The output includes an orthomosaic GeoTiff image file with reflectance values of the entire flight area for each MicaSense band.

2.2.4 Vegetation Indices

The orthomosaics for each of the five MicaSense bands were exported into ArcGIS to extract crop canopy reflectance values at the sample points. The MicaSense RedEdge camera bands include the following bands: (1) blue, (2) green, (3) red, (4) red-edge, and (5) near-infrared (NIR) (Table 2-2).

Table 2-2. Spectral characteristics of the five 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

The orthomosaics were used to compute 11 VIs (Table 2-3). VIs were already found to be suitable for estimating canopy nitrogen in crops, such as the normalized difference vegetation index (NDVI), green NDVI (GNDVI), and Double-Peak Canopy Nitrogen Index (DCNI) (Jones & Vaughn, 2010). Some VIs are related to chlorophyll, which was found to be closely related to leaf nitrogen content as the photosynthetic enzyme, rubisco, comprises the largest proportion of nitrogen in leaves (Marschner, 2011). Chlorophyll absorbs more than 70% of blue and red radiation and reflects green and NIR radiation (Daughtry et al., 2000). The VIs were computed with the Raster calculator function of PCI Geomatica Banff and exported as GeoTiff files. Using ArcGIS, the VI values were extracted at each sample point.

Table 2-3. Vegetation indices used in the study.

Index ¹	Formula ²	Authors
BNDVI	$(\text{NIR} - \text{BLUE}) / (\text{NIR} + \text{BLUE})$	Wang et al., 2007
CI_RE	$(\text{NIR}/\text{REDEGE}) - 1$	Gitelson et al., 2003
ISR	RED/NIR	Fernandes et al. 2003
MSR	$\frac{(\text{NIR}/\text{RED}) - 1}{\sqrt{\frac{\text{NIR}}{\text{RED}} + 1}}$	Chen, 1996
NDVI	$(\text{NIR} - \text{RED}) / (\text{NIR} + \text{RED})$	Rouse et al., 1974
OSAVI	$1.6[(\text{NIR} - \text{RED}) / (\text{NIR} + \text{RED} + 0.16)]$	Rondeaux et al., 1996
NDRE	$(\text{NIR} - \text{REDEGE}) / (\text{NIR} + \text{REDEGE})$	Gitelson & Merzlyak, 1994
RGBVI	$(\text{GREEN}^2 - \text{BLUE} * \text{RED}) / (\text{GREEN}^2 + \text{BLUE} * \text{RED})$	Bendig et al., 2015
RVI	NIR/RED	Jordan, 1969
RVI2	NIR/REDEGE	Kanke et al., 2012
WDRVI	$(0.2 * \text{NIR} - \text{RED}) / (0.2 * \text{NIR} + \text{RED})$	Gitelson & Merzlyak, 1994

¹ Vegetation index abbreviation: BNDVI = blue normalized difference vegetation index; CI_RE = chlorophyll index red edge; ISR = infrared simple ratio; MSR = modified simple ratio; NDVI = normalized difference vegetation index; OSAVI = optimized soil adjusted vegetation index; NDRE = normalized difference red edge index; RGBVI = red green blue vegetation index; RVI = ratio vegetation index; WDRVI = wide dynamic range vegetation index

² Formula variable: BLUE = blue reflectance; GREEN = green reflectance; RED = red reflectance; REDEGE = red edge reflectance; NIR = near-infrared reflectance

2.2.5 Canopy Nitrogen Weight Estimation

Canopy nitrogen is defined by calculating canopy nitrogen weight using the following method (Hansen & Schjoerring, 2003):

$$CNW = \left(N_{plants} \times \frac{Wd}{N_{biomass}} \right) \times LNC \quad (1)$$

Where *CNW* is the canopy nitrogen weight (g/m^2), N_{plants} is the number of plants in the 1 m^2 area over the sampling point, *Wd* is the dry biomass weight (g/m^2), $N_{biomass}$ is the number of plants gathered for biomass at the sampling point, and *LNC* is the leaf nitrogen content (%). *CNW* assumes the plants collected for biomass within a 1 m^2 block around the sample

point have the same amount of nitrogen. For corn plants, the leaves constitute a majority of dry biomass weight, hence the use of total biomass per area in the formula. Compared to other agronomic parameters, including plant nitrogen concentration (%), plant nitrate content, and Soil Plant Analysis Development (SPAD) readings, canopy nitrogen weight has been found to have greater correlation with spectral data (James et al., 2013).

2.2.6 Canopy Nitrogen Weight Modelling

The modelling approaches in this study include Random Forest (RF) regression and Support Vector Regression (SVR). RF is an ensemble learning method that can be used for classification or regression models of large, nonparametric datasets. The user defines a percentage of the dataset to be randomly selected as training data; 70% is commonly used. Using the training data, the algorithm generates many decision trees to determine the importance of variables in the regression. Decision trees split at nodes depending on the most contributing independent variable to the dependent variable. The remainder of the dataset not used in training is used as validation data, and the output average of the individual trees is used to evaluate the regression model's performance. Advantages of RF include the algorithm not overfitting from the training data, it is quick to compute, and it has relatively high performance in studies (Lee et al, 2020).

SVR is part of the Support Vector Machine (SVM) learning algorithm, which uses a decision boundary in a hyperplane to split training data into classes based on the data characteristics. The support vectors that are closest to training samples are used to determine the optimal position of the decision boundary using the midpoint of the margin. SVR performs modelling in a high-dimensional space using the hyperplane. For non-linear data, SVR uses a kernel trick (i.e., Radial Basis Kernel), which places the data in a dimensional space to separate into groups using the radial distance between data points. Advantages of using SVR include its flexibility with nonparametric data, and it has been found to have better modelling capabilities compared to simple linear regression by capturing nonlinearity (James et al., 2013).

The modelling was written in R programming language using R Studio (Version 4.0.3), a free, open-source Integrated Development Environment (IDE) (RStudio, 2020). RF used

the “randomForest” package and SVR used the “e1071” package. For both models, the independent variables were the VIs, MicaSense bands, plant physiology variables, topographic metrics, and soil metrics, and the dependent variable was the canopy nitrogen weight. For variable measurements at each sample point, the average of the variable values within a 1 m² block was used. Data from 8 June, 15 June, and 24 June were randomly divided into 70% training set and 30% validation set. Dates were selected based on the availability of the dataset including UAV imagery and in-situ ground measurements. The quality of the models was assessed using the coefficient of determination (R^2) and Root Mean Square Error (RMSE), calculated using Equation (2) and Equation (3), respectively (Palaniswamy, 2006):

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \quad (2)$$

Where y_i is the observed value, \hat{y}_i is the predicted value, and \bar{y} is the mean of the observed values of the dataset; and:

$$RMSE = \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n} \quad (3)$$

Where \hat{y}_i is the predicted canopy nitrogen weight value (g/m²), y_i is the observed canopy nitrogen weight value (g/m²), n is the number of observations, and i is the index of summation in increments of one.

2.3 Results

2.3.1 Nitrogen Statistics

Canopy nitrogen weight gradually increased in variation from 8 June to 9 July during the early to middle growth stages, then decreased slightly on 15 July (Figure 2-5). As the crop developed into later growth stages, a decrease in canopy nitrogen results from the dilution effect, as discussed in (Justes et al., 1994; Li et al., 2008). An outlier is shown in Figure 5, but as it was consistent throughout the growing stages at the same sample point it is unlikely to be due to errors in measurements.

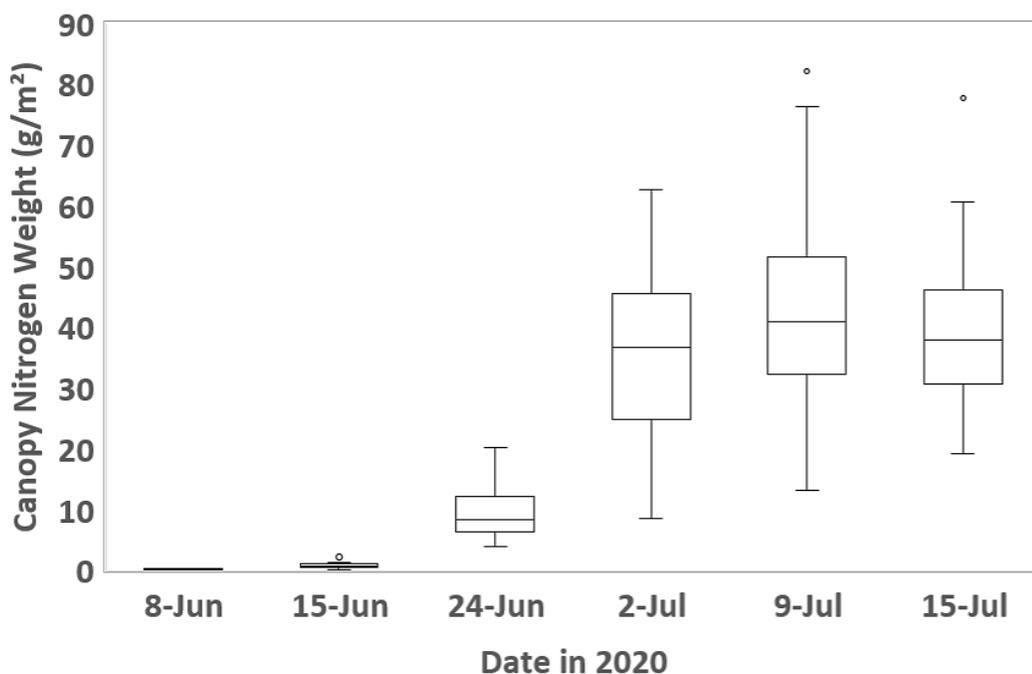


Figure 2-5. Variation of canopy nitrogen weight (g/m^2) as a function of the date of field measurements during the 2020 growing season. Outliers are represented by dots on the graph.

2.3.2 Soil Statistics

The Vittellus soil health analysis (A&L, 2020) produced the following mean values: soil nitrate-nitrogen of 72.75 ppm, mineralizable nitrogen of 30.75 ppm, water extracted soil nitrate of 71.38 ppm, and water extracted total nitrogen of 88.75 ppm. The A&L's Soil Health Index rating was in the "Good-High" category, and the soil textural class for the field was predominantly silt loam.

2.3.3 Regression Models with all Parameters

First, all 29 parameters including VIs, MicaSense bands, plant physiology variables, soil metrics, and topographic metrics were used in calibrating the RF and SVR model (Table 2-4). Single date datasets and combinations of the multi-date dataset were tested to evaluate the temporal effect on the models. From all calibrated models, RF had better performance in comparison to SVR. The best performing RF model was obtained with a combination of

all three dates, resulting in R^2 of 0.97 and RMSE of 0.71 g/m^2 . Other RF multi-date combinations had high performance close to the best model. Of all the RF models, the lowest R^2 was at 0.92 with an RMSE of 0.20 g/m^2 for 15 June. For SVR, the 15 June model also had the lowest performance. Although SVR RMSE values were low overall but close to the RMSE of the RF models', single date models of 8 June and 15 June had low R^2 values at 0.73 and 0.48 respectively. SVR multi-date models had much better performance compared to single-date models, but RMSE values were higher than with the RF multi-date models.

Table 2-4. Statistics for the calibration of the canopy nitrogen model with all 29 variables including vegetation indices, the 5 MicaSense band reflectances, plant physiology variables, soil metrics, and topographic metrics as a function of the date and the modelling approach (RF or SVR)¹.

Date	Model	R^2	RMSE (g/m^2)	(n)
8 June	RF	0.95	0.03	28
	SVR	0.73	0.04	28
15 June	RF	0.92	0.20	28
	SVR	0.48	0.31	28
24 June	RF	0.94	1.70	14
	SVR	0.94	1.53	14
8, 15 June	RF	0.92	0.15	56
	SVR	0.75	0.24	56
8, 24 June	RF	0.97	0.90	42
	SVR	0.97	1.10	42
15, 24 June	RF	0.97	0.96	42
	SVR	0.95	1.22	42
8, 15, 24 June	RF	0.97	0.71	70
	SVR	0.96	1.13	70

¹ All models are significant at p-value <0.001.

Next, the models were applied to the validation datasets (Table 2-5). RF with the combination of all three dates performed the best out of the models with R^2 of 0.75 and RMSE of 2.29 g/m^2 . Compared to the best model, there was only a small difference in the RF model of 8 June and 24 June with R^2 of 0.74 and RMSE of 2.48 g/m^2 . Single-date models for both RF and SVR had poor results overall. However, the small number of sample point data in single-date models may have resulted in the calibration model not

encompassing normal variation in the field data. Overall, multi-date models had better performance compared to single-date models.

Table 2-5. Statistics for the validation of the canopy nitrogen model with all 29 variables including selected vegetation indices, the 5 MicaSense bands, plant physiology variables, soil metrics, and topo-graphic metrics as a function of the date and the modelling approach (RF or SVR).

Date	Model	R ²	p-value ⁽¹⁾	RMSE (g/m ²)	(n)
8 June	RF	0.14	NS	0.09	12
	SVR	0.23	0.100	0.08	12
15 June	RF	0.04	NS	0.40	12
	SVR	0.01	NS	0.39	12
24 June	RF	0.30	NS	4.07	6
	SVR	0.02	NS	3.37	6
8, 15 June	RF	0.50	<0.001	0.31	24
	SVR	0.36	0.001	0.34	24
8, 24 June	RF	0.74	<0.001	2.48	18
	SVR	0.69	<0.001	2.28	18
15, 24 June	RF	0.63	<0.001	3.13	18
	SVR	0.53	<0.001	2.85	18
8, 15, 24 June	RF	0.75	<0.001	2.29	30
	SVR	0.58	<0.001	2.21	30

¹ NS = non-significant

2.3.4 Variable Importance Plot

RF modelling can be visualized with a variable importance plot in R Studio using the “varImpPlot()” function (Figure 2-6). The more important an explanatory variable is in the prediction of canopy nitrogen weight, the higher its IncNodePurity value is. Using the dataset containing all dates and the 29 variables used in the model, plant height was the most important predictor. The top ten variables were made by seven from the 11 VIs used and two MicaSense band reflectance mosaics (blue and red). Among all the topographic metrics, the profile curvature was the top-performing variable. From the soil metrics, soil moisture was the top-performing variable.

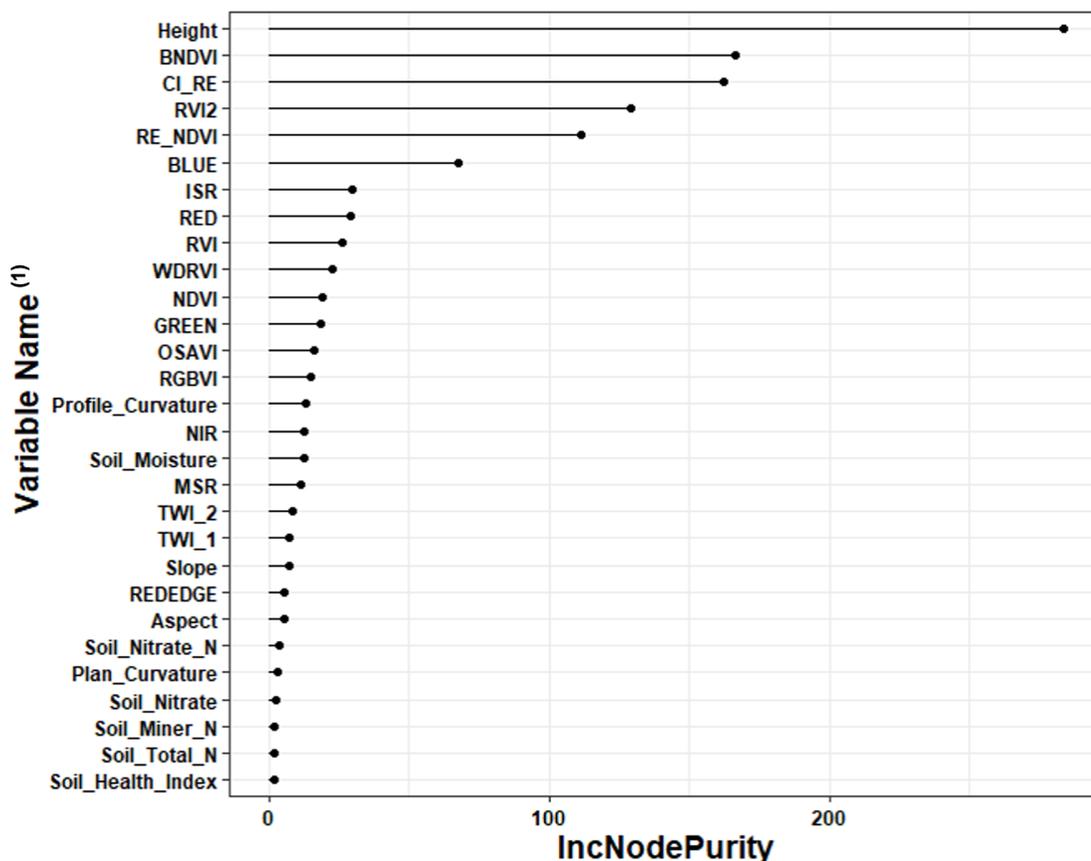


Figure 2-6. Variance Importance plot produced by the Random Forest model of all three dates and 30 variables, using the function varImpPlot() in R Studio. Higher IncNodePurity values indicate more impact on nitrogen. ⁽¹⁾ Please refer to Table 1 for the full name of vegetation indices. N_Weight, plant nitrogen weight; Soil_Nitrate_N, soil nitrate nitrogen (NO₃-N); Soil_Miner_N, soil mineralizable nitrogen; Soil_Total_N, water extracted total soil nitrogen; TWI_1, total wetness index #1; TWI_2, total wetness index #2.

2.3.5 Regression Models with Selected Variables

As the best performing model from the calibration and validation datasets was the combination of all three dates using RF, we tested numerous variable combinations with the data from all three dates. Based on Figure 6, we selected the top 6, 10, 15, 18, and 20 variables based on the evaluation of variable importance thresholds. In addition, we considered a separate group containing all spectral variables, because all the VIs and

MicaSense band reflectance mosaics, except the red-edge band mosaic, were in the top 20 variables. Table 2-6 displays the statistics of RF and SVR applied to the various combinations of variables from calibration datasets. Overall, RF had better performance than SVR, but there was not a large difference in the R^2 or RMSE values. The best model was RF with the combination of top 20 variables, resulting in a R^2 value of 0.97 and a RMSE of 0.70 g/m^2 . However, in comparison to the best model, there were only small differences in R^2 and RMSE for the RF models.

Table 2-6. Statistics for the calibration of the canopy nitrogen model with all dates (8, 15, 24 June), different combinations of variables (n = 70)¹.

Input Variables	Model	Number of Variables	R^2	RMSE (g/m^2)
Spectral-only: All VIs & 5 MicaSense bands	RF	13	0.95	0.93
	SVR	13	0.88	1.50
Top 6: Height, BNDVI, CI_RE, RVI2, RE_NDVI, BLUE	RF	6	0.97	0.73
	SVR	6	0.95	1.01
Top 10: Top 6 + ISR, RED, RVI, WDRVI	RF	10	0.96	0.81
	SVR	10	0.95	1.06
Top 15: Top 10 + NDVI, GREEN, OSAVI, RGBVI, Profile Curvature	RF	15	0.97	0.74
	SVR	15	0.96	0.96
Top 18: Top 15 + NIR, Soil Moisture, MSR	RF	18	0.97	0.73
	SVR	18	0.95	1.05
Top 20: Top 18 + TWI_2, TWI 1	RF	20	0.97	0.70
	SVR	20	0.95	1.10

¹All models are significant at p-value <0.001.

The models were then applied to the validation datasets (Table 2-7). The RF model using the top 15 variables had the best performance with an R^2 value of 0.73 and an RMSE of 2.21 g/m^2 . Compared to the top-performing model, the RF model using the top 18 variables had the same R^2 value with a slightly higher RMSE. Of the top 15 variables, only the plant height and the profile curvature were non-spectral variables. All RF models had higher R^2 values than SVM, but the RF model has RMSE values that were slightly higher as well. The RF variable importance plot allows identification of the variables that do not affect the model significantly. As found in the models with the top 20 variables, removing low importance variables from a model can improve the results.

Table 2-7. Statistics for the validation of the canopy nitrogen model with all dates (June 8, 15, 24), different combinations of variables (n = 30)¹.

Input Variables	Model	Number of Variables	R ²	RMSE (g/m ²)
Spectral-only: All VIs & 5 MicaSense bands	RF	13	0.61	2.89
	SVR	13	0.58	2.31
Top 6: Height, BNDVI, CI_RE, RVI2, RE_NDVI, BLUE	RF	6	0.67	2.50
	SVR	6	0.67	2.16
Top 10: Top 6 + ISR, RED, RVI, WDRVI	RF	10	0.68	2.47
	SVR	10	0.66	2.06
Top 15: Top 10 + NDVI, GREEN, OSAVI, RGBVI, Profile Curvature	RF	15	0.73	2.21
	SVR	15	0.60	2.23
Top 18: Top 15 + NIR, Soil Moisture, MSR	RF	18	0.73	2.25
	SVR	18	0.67	1.97
Top 20: Top 18 + TWI_2, TWI 1	RF	20	0.71	2.36
	SVR	20	0.59	2.17

¹All models are significant at p-value <0.001.

2.4 Discussion

In this study, RF and SVM regression methods were used to predict canopy nitrogen weight of corn using UAV MicaSense individual band reflectance, associated VIs, plant physiology variables, topographic metrics, and soil metrics. The variation of the in-situ canopy nitrogen weight measurements was very low in the earliest growth stage on 8 June and gradually increased until the latest sampling date of 15 July, with a marked decrease afterwards. The increase in canopy nitrogen variation during the early growth stages of BBCH 00-49 can be explained by the leaf growth and stem elongation because the crop biomass increases rapidly during that period. Then, as the plant reaches the BBCH 51 stage that corresponds to the inflorescence emergence and heading, the canopy nitrogen variation decreases because of the dilution effect (Fageria et al, 2010).

The RF and SVR models were first calibrated with all the 29 variables using single and multi-date datasets. With the validation datasets, single-date models had overall poor performance. Combinations of multi-date models led to better results, with the best performance obtained with the RF model. In the variance importance plot of the best RF model, the plant height was the most important predictor out of all variables used. Freeman et al. (2007) already found that plant height is a useful variable in identifying nitrogen

uptake in corn. Precision agriculture studies have used crop height for phenology, biomass, and yield prediction successfully, and this crop height can be derived from UAV point cloud datasets (Song & Wang, 2019). Among all the individual MicaSense band reflectances, the reflectance of the red-edge band has the poorest performance. The red-edge region (680–800 nm) represents a sharp change in the canopy reflectance and can provide important details about phenology (Kanke et al., 2012). Our result agrees with Lee et al.'s (2020) work that uses the same MicaSense camera. Likely, the narrow 10 nm band range of the red-edge band of the MicaSense camera did not capture the change in the region well. This could explain why our results are not in agreement with several other studies that find that the red-edge region is a sensitive indicator of leaf chlorophyll content, because of the high absorption in the red radiation and the high reflectance in the near-infrared region during plant growth stages (Schlemmer et al., 2013; Gahrouei et al., 2020). Overall, most of the soil metrics had little to no effect on the models, but the soil was sampled once at the beginning of the growing season. With the consideration of costs and historical farm operations where recommendations for soil tests are only once a year, this study emphasizes the limitations of the current soil testing practices. Soil metrics results from Mulvaney et al.'s (2006) and Tremblay et al.'s (2012) studies on a soil-based approach in corn nitrogen management, found that soil tests were useful in their models when field conditions were conducive to soil nitrogen mineralization, crop uptake, and utilization. With different sampling methods, soil metrics may still be useful in models. There is therefore the need to conduct soil tests at different dates to better characterize the soil condition changes, for example, because of fertilizer applications, precipitation patterns, and crop growth (Marschner, 2011).

The RF model's variable importance plot allowed selecting groups of top 6, 10, 15, 18, and 20 variables for developing new RF and SVM models. The top 20 variables included plant height, all the 11 VIs used in this study, all the MicaSense band reflectance mosaic but the red-edge one, the soil moisture, and the profile curvature, as well as the topo-graphic wetness indices #1 and #2. The group of top 15 variables that performed best has only the plant height and profile curvature as non-spectral parameters. Considering that topographic metrics were derived from the UAV Phantom 4 RTK imagery along with the possibility of deriving crop height across the field from point cloud data, all variables in the best model

can be measured from in-situ, non-destructive, UAV-based data collection (Freeman et al., 2007). Having all model data that can be collected by remote sensing could be a greater benefit, as common limitations of in-situ studies and subsequent application methods are the intensive labour and high costs required to obtain model input data.

In this study, the final validation of canopy nitrogen models with various combinations of variables indicated RF models had better performance than SVR in terms of R^2 values. This is consistent with results from Liu et al. (2016), Lee et al. (2020), and Zha et al. (2020) with RF yielding better nitrogen content prediction in wheat, rice, and corn crops compared to SVR models. Although SVR had lower RMSE values in comparison to RF, overall RF RMSE values were low as well in the context of nitrogen estimation for g/m^2 . In comparison to the study by Lee et al. (2020), the RMSE values of this study's models are much lower, which can be beneficial for fertilizer management recommendations to farmers in general. In the case of this study, the RF algorithmic method of using many decision trees may better suit the use of numerous variables in regression models. In comparison, using many variables in SVR requires user hyper-tuning and the kernel trick function to separate data into groups, relying on the radial distance between points to be meaningful in the model. Overall, the performance of SVR models was good, but RF models can be considered more useful in terms of ease of use and the quality of results.

2.5 Conclusion

This study tested machine learning regression methods to predict corn canopy nitrogen weight using UAV MicaSense band reflectance mosaics, associated VIs, plant physiology variables, topographic metrics, and soil metrics. With all 29 variables in RF and SVR models, the combination of all three dates with the RF model produced the best results: the validation model having an R^2 of 0.75 and an RMSE of 2.29 g/m^2 . From the multi-date RF model's variable importance plot, the top 6, 10, 15, 18, and 20 variables were tested in RF and SVR models. The best validation model was the RF model (R^2 value at 0.73 and RMSE at 2.21 g/m^2) with the top 15 variables, most of them being spectral variables.

We developed models for estimating canopy nitrogen weight from spectral, plant, soil, and topographic variables using machine learning algorithms, but the resulting models are still

empirical, and their applicability can be limited to the dataset on which they were built and validated. This is a common limitation in agricultural research as in-situ measurements often require intensive labour, costs, and variable conditions. Overall, many factors need to be considered to define plant growth conditions such as plant species, soil condition, environmental factors of field topology, moisture supply, weather, and more. There is a need to test the developed models on other datasets to determine their efficacy and to understand their applicability in precision agriculture. Future work can consider using a more deterministic modelling approach, for example, the PROSAIL model (Jay et al., 2017), as it is less empirical and applies to a high variety of conditions but requires more advanced parameter calibration. The PROSAIL model uses spectral data of leaf and canopy level parameters to retrieve chlorophyll and nitrogen content, with robust results from lab and field studies (Jay et al., 2017). Eventually, methods of crop height extraction from RTK UAVs (Song & Wang, 2019), in addition to UAV-derived topographic and spectral variables, can be used to develop a final map for a whole field. The model information can be used to improve field nitrogen prediction, leading to more effective and efficient N fertilizer management.

2.6 References

- A&L Canada Laboratories. (n.d.). *VitTellus soil health test & VitTellus Bio*. Retrieved June 14, 2020 from <https://www.alcanada.com/content/solutions/soil-health>
- Agriculture and Agri-Food Canada. (n.d.) *An overview of the Canadian agriculture and agri-food system*. Retrieved February 23, 2021 from <https://www.agr.gc.ca/eng/canadas-agriculture-sectors/an-overview-of-the-canadian-agriculture-and-agri-food-system-2017/?id=1510326669269>
- Association of Official Analytical Collaboration (AOAC) (2006) *International Official Method 972.43. Microchemical determination of carbon, hydrogen, and nitrogen, automated method*. AOAC
Int.http://www.aocofficialmethod.org/index.php?main_page=product_info&cPath=1&products_id=1492
- Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., Gnyp, M.L., & 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. <https://doi.org/10.1016/j.jag.2015.02.012>
- Bongiovanni, R., & Lowenberg-DeBoer, J. (2004). Precision agriculture and sustainability. *Precision agriculture*, 5(4), 359-387. <https://doi.org/10.1023/B:PRAG.0000040806.39604.aa>
- Bryant, C. R., Smit, B., Brklacich, M., Johnston, T. R., Smithers, J., Chiotti, Q., & Singh, B. (2000). Adaptation in Canadian agriculture to climatic variability and change. In *Societal adaptation to climate variability and change* (pp. 181-201). Springer, Dordrecht. https://doi.org/10.1007/978-94-017-3010-5_10
- 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. <https://doi.org/10.1080/07038992.1996.10855178>
- 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. <https://doi.org/10.1016/j.rse.2010.04.006>
- Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., & Böhner, J. (2015). System for automated geoscientific analyses (SAGA) v. 2.1. 4. *Geoscientific Model Development*, 8(7), 1991-2007. <https://doi.org/10.5194/gmd-8-1991-2015>
- Da-Jiang Innovations (DJI). (n.d.) Retrieved April 15, 2021 from <https://www.dji.com/ca>
- Daughtry, C. S., 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. [https://doi.org/10.1016/S0034-4257\(00\)00113-9](https://doi.org/10.1016/S0034-4257(00)00113-9)

- Delta-T Devices Ltd. (n.d.) User Manual for the ML3 ThetaProbe. Retrieved May 3, 2020 from <https://delta-t.co.uk/wp-content/uploads/2017/02/ML3-user-manual-version-2.1.pdf>
- Eitel, J. U. H., Long, D. S., Gessler, P. E., & Smith, A. M. S. (2007). Using in-situ measurements to evaluate the new RapidEye™ satellite series for prediction of wheat nitrogen status. *International Journal of Remote Sensing*, 28(18), 4183-4190. <https://doi.org/10.1080/01431160701422213>
- European Space Agency. (n.d.) *Sentinel-2*; Retrieved April 25, 2021 from www.sentinel.esa.int/web/sentinel/missions/sentinel-2
- Fageria, N. K. (2009). *The use of nutrients in crop plants*. CRC Press: New York, U.S.A.
- Fageria, N. K.; Baligar, V. C.; Jones, C. A. (2010) Growth and mineral nutrition of field crops, 3rd ed.; CRC Press: New York, U.S.A.
- 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. <https://doi.org/10.5589/m02-092>
- Food and Agriculture Organization (FAO) of the United Nations. (2007). *Adaptation to climate change in agriculture, forestry and fisheries: perspective, framework, and priorities*. Retrieved April 2, 2021 from www.fao.org/nr/climpag/pub/adaptation_to_climate_change_2007.pdf
- Freeman, K. W., Girma, K., Arnall, D. B., Mullen, R. W., Martin, K. L., Teal, R. K., & Raun, W. R. (2007). By-plant prediction of corn forage biomass and nitrogen uptake at various growth stages using remote sensing and plant height. *Agronomy Journal*, 99(2), 530-536. <https://doi.org/10.2134/agronj2006.0135>
- 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. <https://doi.org/10.1016/j.fcr.2017.12.004>
- Reisi Gahrouei, O., McNairn, H., Hosseini, M., & Homayouni, S. (2020). Estimation of crop biomass and leaf area index from multitemporal and multispectral imagery using machine learning approaches. *Canadian Journal of Remote Sensing*, 46(1), 84-99. <https://doi.org/10.1080/07038992.2020.1740584>
- 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. [https://doi.org/10.1016/1011-1344\(93\)06963-4](https://doi.org/10.1016/1011-1344(93)06963-4)
- 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. <https://doi.org/10.1080/01431161.2013.793868>

- 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. <https://doi.org/10.1078/0176-1617-00887>
- Good, A. G., Shrawat, A. K., & Muench, D. G. (2004). Can less yield more? Is reducing nutrient input into the environment compatible with maintaining crop production?. *Trends in plant science*, 9(12), 597-605. <https://doi.org/10.1016/j.tplants.2004.10.008>
- Hansen, N. C., Tubbs, S., Fernandez, F., Green, S., Hansen, N. E., & Stevens, W. B. (2015). Conservation agriculture in North America. In *Conservation Agriculture* (pp. 417-441). Springer, Cham. https://doi.org/10.1007/978-3-319-11620-4_17
- 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. [https://doi.org/10.1016/S0034-4257\(03\)00131-7](https://doi.org/10.1016/S0034-4257(03)00131-7)
- 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. <https://doi.org/10.3390/rs4061573>
- Hunt, E. R., Hively, W. D., Fujikawa, S. J., Linden, D. S., Daughtry, C. S., & McCarty, G. W. (2010). Acquisition of NIR-green-blue digital photographs from unmanned aircraft for crop monitoring. *Remote Sensing*, 2(1), 290-305. <https://doi.org/10.3390/rs2010290>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning: With applications in R*. (Vol. 112, p. 18). 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. <https://doi.org/10.1016/j.fcr.2017.05.005>
- Jones, H.G.; Vaughan, R.A. (2010). *Remote sensing of vegetation: principles, techniques, and applications*, 1st ed.; University Press: New York, U.S.A.
- Jordan, C. F. (1969). Derivation of leaf-area index from quality of light on the forest floor. *Ecology*, 50(4), 663-666. <https://doi.org/10.2307/1936256>
- Justes, E., Mary, B., Meynard, J. M., Machet, J. M., & Thelier-Huché, L. (1994). Determination of a critical nitrogen dilution curve for winter wheat crops. *Annals of botany*, 74(4), 397-407. <https://doi.org/10.1006/anbo.1994.1133>
- Kanke, Y., Raun, W., Solie, J., Stone, M., & Taylor, R. (2012). Red edge as a potential index for detecting differences in plant nitrogen status in winter wheat. *Journal of plant nutrition*, 35(10), 1526-1541. <https://doi.org/10.1080/01904167.2012.689912>

- 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*, 46(4), 454-472.
<https://doi.org/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.
<https://doi.org/10.3390/rs12132071>
- Li, F., Gnyp, M. L., Jia, L., Miao, Y., Yu, Z., Koppe, W., Bareth, G., Chen, X., & 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.
<https://doi.org/10.1016/j.fcr.2007.11.001>
- Liu, Y., Cheng, T., Zhu, Y., Tian, Y., Cao, W., Yao, X., & Wang, N. (2016, July). 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)* (pp. 7362-7365). IEEE, Beijing, China.
<https://doi.org/10.1109/IGARSS.2016.7730920>
- Marschner, H. (2011) *Marschner's mineral nutrition of higher plants*, 2nd ed.; Academic Press: Adelaide, Australia.
- Masclaux-Daubresse, C., Daniel-Vedele, F., Dechorgnat, J., Chardon, F., Gaufichon, L., & Suzuki, A. (2010). Nitrogen uptake, assimilation and remobilization in plants: challenges for sustainable and productive agriculture. *Annals of botany*, 105(7), 1141-1157. <https://doi.org/10.1093/aob/mcq028>
- Meier, U. (2001) Growth stages of mono- and dicotyledonous plants. Federal Biological Research Centre for Agriculture and Forestry. Blackwell Wissenschafts-Verlag.
- MicaSense. Retrieved April 15, 2021 from <https://micasense.com/about/>
- Mulla, D. J. (2013). Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems engineering*, 114(4), 358-371. <https://doi.org/10.1016/j.biosystemseng.2012.08.009>
- Mulvaney, R. L., Khan, S. A., & Ellsworth, T. R. (2006). Need for a soil-based approach in managing nitrogen fertilizers for profitable corn production. *Soil Science Society of America Journal*, 70(1), 172-182.
<https://doi.org/10.2136/sssaj2005.0034>
- Palaniswamy, U. (2006) *Handbook of Statistics for Teaching and Research in Plant and Crop Science*. The Haworth Reference Press: New York, U.S.A.
- Pix4D. (n.d.) *How to improve the outputs of dense vegetation areas*. Retrieved March 2, 2021 from <https://support.pix4d.com/hc/en-us/articles/202560159-How-to-improve-the-outputs-of-dense-vegetation-areas>

- QGIS (2021) 3.16. *Geographic information system developers manual*. Retrieved February 25, 2021 from www.docs.qgis.org/3.16/en/docs/developers_guide/index.html
- Rondeaux, G., Steven, M., & Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote sensing of environment*, 55(2), 95-107. [https://doi.org/10.1016/0034-4257\(95\)00186-7](https://doi.org/10.1016/0034-4257(95)00186-7)
- 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(1974), 309.
- RStudio Team. (n.d.) *RStudio: integrated development for R*. Retrieved March 2, 2021 from <http://www.rstudio.com/>
- 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. <https://doi.org/10.1016/j.jag.2013.04.003>
- Sishodia, R. P., Ray, R. L., & Singh, S. K. (2020). Applications of remote sensing in precision agriculture: A review. *Remote Sensing*, 12(19), 3136. <https://doi.org/10.3390/rs12193136>
- Song, Y., & Wang, J. (2019). Winter wheat canopy height extraction from UAV-based point cloud data with a moving cuboid filter. *Remote Sensing*, 11(10), 1239. <https://doi.org/10.3390/rs11101239>
- Sripada, R. P., Schmidt, J. P., Dellinger, A. E., & Beegle, D. B. (2008). Evaluating multiple indices from a canopy reflectance sensor to estimate corn N requirements. *Agronomy Journal*, 100(6), 1553-1561. <https://doi.org/10.2134/agronj2008.0017>
- Tan, C. S., & Reynolds, W. D. (2003). Impacts of recent climate trends on agriculture in southwestern Ontario. *Canadian Water Resources Journal*, 28(1), 87-97. <https://doi.org/10.4296/cwrj2801087>
- Tremblay, N., Bouroubi, Y. M., Bélec, C., Mullen, R. W., Kitchen, N. R., Thomason, W. E., & Ortiz-Monasterio, I. (2012). Corn response to nitrogen is influenced by soil texture and weather. *Agronomy Journal*, 104(6), 1658-1671. <https://doi.org/10.2134/agronj2012.0184>
- U.S. Geological Survey. (n.d.). *Landsat Satellite Missions*. Retrieved April 10, 2021 from www.usgs.gov/core-science-systems/nli/landsat/landsat-satellite-missions?qt-science_support_page_related_con=0#qt-science_support_page_related_con
- Vetsch, J. A., & Randall, G. W. (2004). Corn production as affected by nitrogen application timing and tillage. *Agronomy Journal*, 96(2), 502-509. <https://doi.org/10.2134/agronj2004.5020>

- 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. [https://doi.org/10.1016/S1672-6308\(07\)60027-4](https://doi.org/10.1016/S1672-6308(07)60027-4)
- Wolfe, A. H., & Patz, J. A. (2002). Reactive nitrogen and human health: acute and long-term implications. *Ambio: A journal of the human environment*, *31*(2), 120-125. <https://doi.org/10.1579/0044-7447-31.2.120>
- 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. <https://doi.org/10.3390/rs12020215>
- Zhu, Y., Yao, X., Tian, Y., Liu, X., & Cao, W. (2008). Analysis of common canopy vegetation indices for indicating leaf nitrogen accumulations in wheat and rice. *International Journal of Applied Earth Observation and Geoinformation*, *10*(1), 1-10. <https://doi.org/10.1016/j.jag.2007.02.006>

Chapter 3

3 Nitrogen Estimation for Wheat using UAV-based and Satellite Multispectral Imagery, Topographic Metrics, Leaf Area Index, Plant Height, Soil Moisture, and Machine Learning Methods

3.1 Introduction

Precision agriculture (PA) is a management technique that selectively applies crop farming resources such as fertilizer, water, pesticides, and herbicides based on the plant needs within a field (Bongiovanni & Lowenberg-Deboer, 2004; Masclaix-Daubresse et al., 2010; Sishodia et al., 2020). Nitrogen is an essential macronutrient to plants as a major constituent of organic material including enzymic processes, chlorophyll, and oxidation-reduction reactions; levels of nitrogen in plant tissue can indicate yield potential and crop health (Marschner, 2011). However, nitrogen is one of the most expensive nutrients to supply, and studies on nitrogen recovery efficiency by annual crops was on average less than 50% of the amount of fertilizer applied (Fageria, 2009; Fageria et al., 2010).

Excessive fertilizer can leach from the soil and contaminate waterways, disrupting local ecosystems, and denitrification that results in greenhouse gas emissions (Rabalais et al., 2001). Nutrients that have been added beyond the critical level of maximum growth can continue to accumulate in the plant tissue without any further yield increase (Marschner, 2011). Commonly in grain crops such as wheat, excessive nitrogen can cause plant stems to grow tall to the point of lodging – the stems bend over making it difficult to harvest, increasing chances of grain moisture, disease, and often reduces yield significantly (Foulkes, 2011). Usually, nitrogen deficiency can be noted from chlorosis, the condition in which leaves yellow as the plant's chlorophyll content drops (Drew & Sisworo, 1977). With reduced photosynthetic activity, the plant will not reach peak health and yield will be low. Water is also key to the transportation of nutrients from the soil to a plant. The availability of water to a plant depends on the weather conditions during the growing season, soil moisture, field micro-topography affecting water flow and accumulation (Walley et al., 2002; Si & Farrell, 2004). Understanding a field's characteristics as well as

monitoring plant biophysical characteristics including height, leaf area and colour of can provide useful information in nitrogen fertilizer applications.

In PA, remote sensing imagery is useful because it does not require physical or destructive contact with plants to gather valuable crop information (Jones & Vaughan, 2010). Vegetation indices (VIs) can be derived from the spectral information provided by the imager; VIs are mathematical combinations or transformations of spectral bands that have been widely used in agricultural research. VIs allow for deriving of specific plant properties such as chlorophyll or nutrient content by taking advantage of the differential spectral properties of plants in the visible and near-infrared (NIR) wavelengths (Eitel et al., 2007; Frels et al., 2018; Sripada et al., 2008). The VI information can then provide timely knowledge of crop conditions, allowing for suitable rate of application at the right time and location depending on the variations within a field.

Optical satellite imagery for crop monitoring has had several decades of research and application (Agriculture & Agri-Food Canada, 2007). Examples of recently launched optical satellites in operation include RapidEye since 2008, Landsat 8 since 2013, and Sentinel-2 since 2015, all frequently used in studies on crop nutrient, yield, and growth management (Sishodia et al., 2020). RapidEye has five spectral bands with 6.5 m resolution. Depending on the location, the five-satellite constellation revisit time is between one to five days. Landsat 8 Operational Land Imager has nine spectral bands with varying spatial resolutions of 15 to 30 m. It has a 16-day revisit time to the same area and takes over 700 scenes a day. Sentinel-2 has 13 spectral bands with 10 m, 20 m, and 60 m spatial resolutions depending on the band. Sentinel-2 constellation is composed of two satellites allowing for a five-day revisit time over the same area. Limitations in optical satellite imagery include low spatial sensitivity as the spatial resolution may be too coarse for small-scale crop fields (Jones & Vaughn, 2010). The temporal sensitivity can be rather low, such as Landsat 8 with a 16-day revisit time crops would have changed significantly and valuable information on the different stages of growth would not be obtained. Sentinel-2 and RapidEye have higher temporal resolution of one to five days, but it can vary by location and not all images may be useful due to cloud cover obscuring land. Additionally,

factors such as cloud cover, geometric distortion, and atmospheric distortion may require advanced processing expertise to ensure sufficient image quality (Mulla, 2013).

New satellite systems are improving in spatial and temporal sensitivity, such as the PlanetScope satellite constellation (Planet, 2020). Designed for collecting information for use in land-change detection, crop monitoring, climate monitoring and more, the PlanetScope satellite constellation is composed of over 130 satellites called Doves allowing for spatial resolutions of 3 to 5 m and daily revisit. Beginning with the first launch of a group of Doves in March 2016, over 10 more groups launched since to improve revisit time, spatial, and spectral resolutions. PlanetScope imagery products are also available in multiple asset forms with different radiometric processing and rectification, such as the “surface reflectance” product imagery downloaded for this study. Currently, a select portion of Planet data is available for free download under an open data access policy. PlanetScope imagery has been used in studies of wheat yield, biomass, and LAI monitoring and modelling with promising results (Breunig et al., 2020; Manivasagam et al., 2021; Sadeh et al., 2020). However, there are few studies focused specifically on nitrogen management using PlanetScope data which this study aims to fill.

With the rapid advancement in UAV technology in recent years, there is much research interest on UAV-based crop canopy nitrogen retrieval (Sishodia et al. 2020). UAV-based remote sensing can provide low cost and higher spatial and temporal resolution data for crop management. Individuals with basic training can operate a UAV using programmed routes and collect images with <10 cm resolutions (Harwin & Lucieer, 2012). They can be flown to capture more frequent image data and offer flexibility in operation for times when weather is most suitable (Hunt et al., 2012). Compared to satellites, overall UAV-based systems are often lower in cost for data collection and processing. Studies have shown significant correlations between crop spectral variables derived from UAV imagery and crop nitrogen content (Jiang et al., 2019; Liu et al., 2016; Zheng et al., 2018). Many studies are based on single or combinations of different spectral indices’ relationships with crop nitrogen content, noting variation in the relationships at different stages of crop growth (Asataoui et al., 2021; Jiang et al., 2019; Schirrmann et al., 2016). The spectral indices with the strongest relationship to crop nitrogen were noted to occur during early wheat growth

stages before and up to flowering. Often studies on estimating nitrogen were conducted in controlled experimental conditions, and more studies are needed on real field conditions.

Wheat was selected for this study because it is among the most grown crops in Ontario (Agriculture and Agri-Food Canada, 2021). With the development of new remote sensing technologies, processing methods, and computing capabilities, estimation models for crop nitrogen can be improved. Machine learning is an area of research interest as they can develop accurate crop monitoring models for large, nonparametric, nonlinear datasets (Lee et al., 2020). Recent studies have tested the use of linear regression, Random Forest (RF), and Support Vector Regression (SVR) models in UAV-based canopy nitrogen weight (CNW) prediction models (Astaoui et al., 2021; Lee et al., 2020; Liu et al., 2016; Zheng et al., 2018). Although linear regression is a commonly used method to predict nitrogen, some VIs (e.g., NDVI) may saturate beyond the early growth crop stages and models may have reduced accuracy due to multicollinearity (Chen et al., 2010; Zheng et al., 2018). By contrast, machine learning-based regression methods such as RF and SVR have been found to produce more accurate models compared to classical linear regression methods, as they are unaffected by the assumptions of linear regression (Chen et al., 2010). However, most current literature in using remote sensing data and machine learning have only considered spectral information for crop nitrogen modelling (Schirrmann, 2016). As a crop's nitrogen status can be affected by many factors including fertilizer application, soil characteristics, water availability, and field micro-topography, nitrogen prediction models may be improved if these plant physiological and environmental variables are considered (Yu et al., 2021).

With better management of nitrogen fertilizers, not only can costs and negative environmental impact be minimized but yield and quality can increase. This study aims to evaluate machine learning modelling methods with plant spectral, biophysical, and field environmental variables to predict CNW in wheat crops using UAV and satellite-based imagery. The objectives of this study include, (i) studying the relationship between the spatial variation of CNW and factors such as plant height, LAI, soil moisture, and topographic metrics within wheat fields in Southwestern Ontario using multispectral UAV- or PlanetScope - imagery; (ii) determining the optimal combination(s) of spectral

variable(s), crop variables, and/or environmental conditions (soil, water, topographic data) for wheat canopy nitrogen estimation and prediction; and (iii) evaluating the temporal variation of nitrogen estimation and prediction during early growth stages of wheat using UAV images or PlanetScope images, and related variables.

3.2 Materials and Methods

3.2.1 Study Area & Data

The study sites are in Strathroy-Caradoc, Ontario, Canada, nearby the community of Mount Brydges (Figure 3-1). Fieldwork was conducted during May-June 2020 with an average temperature of 22 °C and humidity averaging 73%, characteristic of southwestern Ontario's humid continental climate zone. The sites are in a predominantly agricultural area about 25 km southwest of London, Ontario's urban center.



Figure 3-1. Locations of the three 2020 wheat fields in southwestern Ontario, Canada over a Google Earth image. Field W1 is situated in the upper left corner, west of Mount Brydges. Fields W2 and W3 are in the lower right corner, south of Mount Brydges.

In southwestern Ontario, winter wheat is planted in autumn for germination, lying dormant during winter and resuming growth in spring (Ontario Ministry of Agriculture, Food, and Rural Affairs [OMFRA], 2019). This study’s wheat fields’ cultivar was “Soft Red Winter Wheat” planted in mid-October 2019, began sprouting mid-April 2020, and harvested in early August. The three fields labelled W1, W2, and W3 are sized at 48, 21, and 27 hectares respectively. Beginning in early May, sampling was conducted every 7-8 days for a total of five sample dates to capture significant growth stages of the crop (Table 3-1). Wheat growth stages were recorded following the Biologische Bundesanstalt, Bundessortenamt and Chemische Industrie (BBCH) scale (Meier, 2001). The data acquired for this study encompassed the wheat crop’s growth period from leaf development up to and including inflorescence emergence (BBCH 10 to 59) for nitrogen estimation. The early stages of growth before flowering are especially important as fertilizers applied then have better nitrogen use efficiency and yield response (Marshner, 2011; N. Hansen et al., 2015). Fertilizer can be applied during autumn planting, but lower probability of rainfall also decreases the amount of N from moving into the soil. Fertilizer left above soil during winter will not penetrate and may move during spring snowmelt. During later growth stages (BBCH 60+) from fruit development to ripening, root N uptake slows down as the plants translocate N from vegetation to grains and excess fertilizer can leach from the field (Vetsch & Randall, 2004).

Table 3-1. Summary of fieldwork data acquisitions.

Fieldwork Date	Air Temperature (°C)	BBCH* Growth Stage
May 5	17	22-24
May 12	12	23-26
May 20	14	30-32
May 27	21	39-41
June 4	23	52-58

*Biologische Bundesanstalt, Bundessortenamt und Chemische Industrie Scale

Prior to sample point selection, a DJI Phantom 4 Real-Time Kinematics (RTK) UAV was flown over the bare soil of the fields. The UAV connected to an RTK global navigation satellite system (GNSS) base station acquires images with positioning metadata at 1.5 cm vertical and 1 cm horizontal positioning accuracy. These images were mosaicked to create a digital elevation model (DEM) GeoTiff image. The DEM was imported into Google

Earth and 16 samples points were selected for each field based on the variation of elevation and coverage of representative areas (Figure 3-2). Considering the crop rows were planted northwest-southwest for field W1 and north-south for fields W2 and W3, the sampling pattern followed the row directions for navigation efficiency. The sample points were placed in a four-by-four grid, spaced 60m from one another. A minimum distance of 50m from roads was used to reduce possible effects of transportation pollution.

To measure biomass at each sampling point, a square guide made of plastic tubing 50 cm x 50 cm was placed around a patch of wheat. The plants within the guide were destructively collected by cutting at the stem base above ground, then placed in paper bags. The average distance between rows was 17 cm, and mostly three rows of wheat would be collected from the 0.5 m² biomass block. On the same day following fieldwork collection, the fresh biomass was weighed in grams then fully dried in an oven at 60°C for 48 h. Dry biomass was weighed before being sent to A&L Canada Laboratories for plant analysis. The biomass was ground into a fine powder able to pass through a 1 mm sieve before being used in the Laboratory Equipment Company (LECO) FP628 nitrogen combustion method to obtain leaf N content percentage (Association of Official Analytical Collaboration [AOAC], 2006).

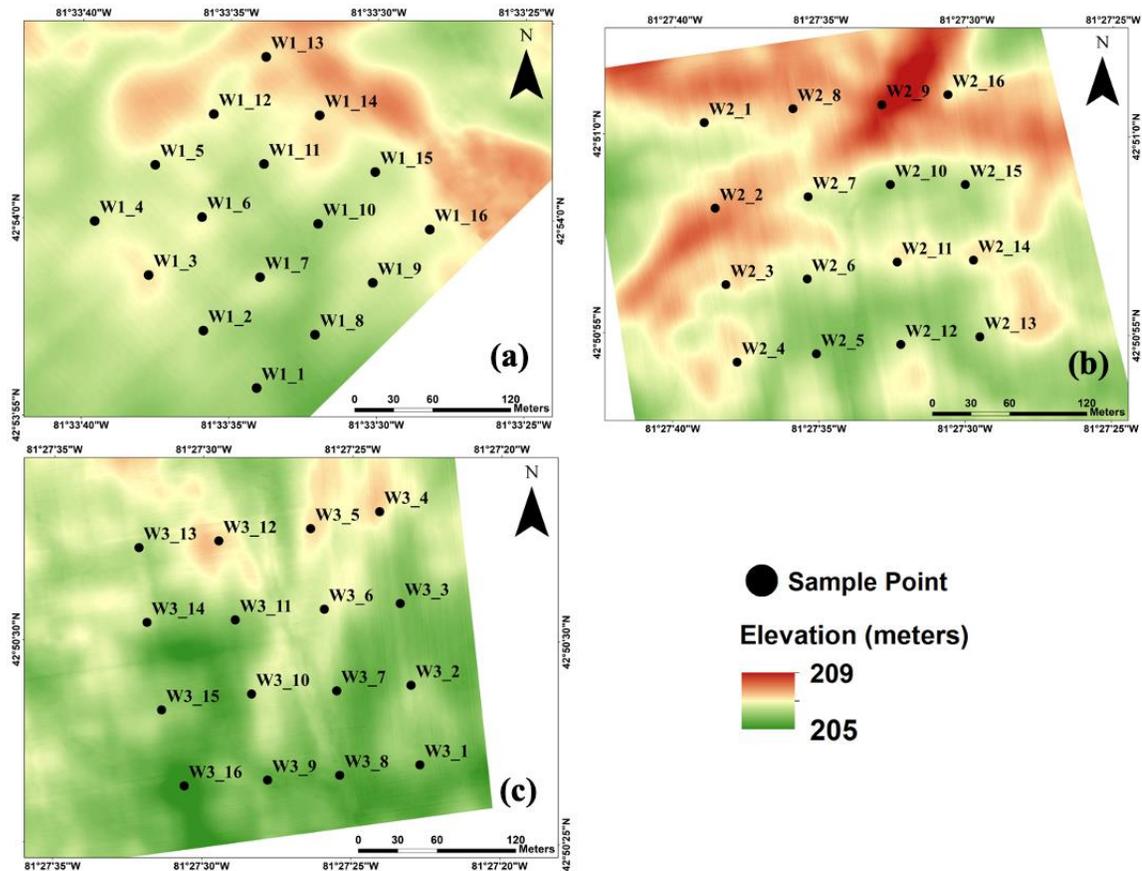


Figure 3-2. Digital elevation model showing the variation of the wheat field's topography at each sample point for a) W1 wheat field, b) W2 wheat field, c) W3 wheat field.

Six measurements of plant height in centimeters were taken around each sampling point within a 1 m² block on every fieldwork date, and an average height was calculated. Six measurements of soil moisture were collected within a 1 m² block around the sampling point using an ML3 ThetaProbe (Delta-T Devices Ltd., Burwell, Cambridge, UK) and averaged. Plant phenology was recorded at each sample point using the BBCH scale to determine growth stage during data collections.

Leaf area index (LAI) was measured non-destructively using a LAI-2200C Plant Canopy Analyzer (Li-Cor, Inc., Lincoln, Nebraska, USA). Following manufacturer recommendations, five measurements were taken along the row at each sample point during clear skies or uniform overcast. Files from the device were transferred to and

processed with the Li-Cor software File Viewer FV2200. Processing options include scattering correction based on the field conditions, and output text files with final LAI measurements.

3.2.2 UAV Imagery

In this study, the two UAVs used were a Da-Jiang Innovations (DJI) (DJI, China) Matrice 100 and a DJI Phantom 4 RTK. The DJI Phantom 4 was flown at 30m altitude per manufacturer recommendations of the visual sensor system's optimal performance and connection to the RTK GNSS base station. The RGB image resolution was 0.9 x 0.9 cm and capture set to 80% side and 80% front image overlap. A fine resolution typically < 10 cm and high image overlap ensures greater chance of successful image mosaicking as crop canopy densifies during a growing season (Harwin & Lucieer, 2012).

The DJI Matrice 100 carried a MicaSense RedEdge (MicaSense Inc., Seattle, WA, USA) narrowband multispectral camera including the following bands ordered by increasing wavelength: (1) blue, (2) green, (3), red, (4) red-edge, and (5) near-infrared (NIR) (Table 3-2). Image acquisition was conducted on the same dates as ground fieldwork before biomass collection. For the week of 27 May fieldwork, the weather conditions were characterized by strong winds and sudden rain showers, so the UAV was flown on different dates between 26 and 29 May for each field whenever the weather was suitable.

Table 3-2. Spectral characteristics of the five 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

The flight plans for each field were made in the Pix4D software suite “Pix4Dcapture” app (Pix4D S.A., Prilly, Switzerland) to cover entire fields in a zigzag pattern (Pix4D Documentation, 2020). Pix4Dcapture includes the function of adding custom UAV and sensor properties to calculate the flight plan's estimated total time, battery usage, and image

resolution at the plan's designated altitude. Flight altitude for each field's plan was set at 60m, as >60m wind and gust conditions were usually greater than UAV manufacturer recommendations. Based on the MicaSense camera's specifications, the resulting image resolution was 4 cm x 4 cm with capture set to 80% side and 80% front overlap. Unfortunately, the SD cards storing images for field W3 on May 5th and field W1 on May 27th were corrupted and data was unretrievable for further processing. The weather for subsequent days of those fieldwork weeks were also unsuitable for

3.2.3 UAV Image Processing

UAV image processing followed the flowchart of Figure 3-3. The DJI Phantom 4 RTK images taken in April over each field's bare soil were mosaicked in Pix4Dmapper software to generate a DEM GeoTiff. Using the System for Automated Geoscientific Analysis (SAGA), a free, open-source spatial data analysis software, topographic metrics were computed from the DEM (Conrad et al., 2015). Using the "Terrain Analysis – Morphometry" tool, metrics generated include slope, aspect, profile curvature, and plan curvature. Creating topographic wetness indices (TWI) required several steps. First, using the "Terrain Analysis – Hydrology" tool a flow accumulation layer of the field was created. Using the "Flow Width and Specific Catchment Areas (SCA)" tool, two SCAs were created using different algorithms: Deterministic 8 and Multiple Flow Direction. The "Topographic Wetness Index (TWI)" tool was then used to create two maps, TWI #1 and TWI #2 based on the respective algorithm SCAs. The final topographic metrics were exported as GeoTiff files.

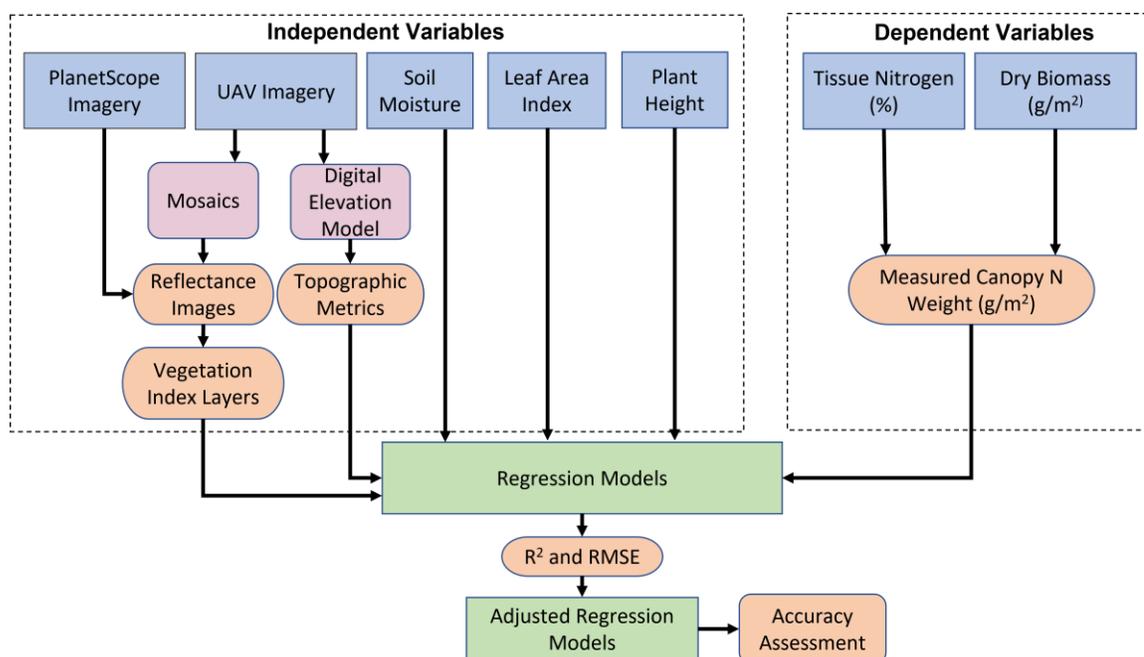


Figure 3-3. Flowchart of the methodology used in the study.

MicaSense camera multispectral images were processed in Pix4Dmapper to create an orthomosaic image per band with 4 cm x 4 cm resolution. An important step in producing a high-quality final image is radiometric calibration considering the sensor influence and scene illumination of the UAV flight. Prior to each flight over a field, the MicaSense camera was positioned above a MicaSense Calibrated Reflectance Panel to take a minimum of five white reference images for each band. From the Exchangeable Image File Format (EXIF) metadata of the images, Pix4Dmapper can read the sensor settings, properties, and geolocation at the time the images were taken. Before starting the mosaic processing, setting options include inputting the white reference images and manufacturer-provided panel reflectance values to calibrate and correct image reflectance for each of the bands. Then Pix4Dmapper uses the Structure from Motion (SfM) algorithm to correct image perspectives and recognize where to stitch images together (Harwin & Lucieer, 2012). The high image overlap parameters set in the flight plan enable the software to recognize greater, similar areas of each image for higher chance of mosaic success. Pix4Dmapper processing options include georeferencing with ground control points (GCP) to improve the absolute location accuracy. Following Pix4D's recommended number of five GCPs, black and white checkered boards were placed around the wheat fields and their

coordinates recorded using a Global Positioning System (GPS) connected to the RTK. The output orthomosaic images are GeoTiff files with reflectance values for each MicaSense band.

3.2.4 Satellite Imagery

In this study, PlanetScope satellite imagery was acquired from Planet Labs Inc. through submission and acceptance of the project proposal. Revisit times are almost daily worldwide with resampled spatial resolution of 3 m x 3 m. Image products available for this study are from the third-generation sensors, Dove-R, with four bands ordered by increasing wavelength: (1) blue, (2) green, (3) red, (4) NIR (Table 3-3). Products scenes are approximately 25 km x 23 km and in the case of the three study fields all were in one image. Five satellite images were downloaded, matching the capture dates with the ground data collection dates to maintain consistency with the plant physiology and field conditions. The images were available for download as GeoTiff surface reflectance assets, orthorectified and radiometrically corrected based on the atmospheric conditions of the specific ground locations. In ArcGIS, the large scenes were cropped to smaller images of each study field for subsequent processing

Table 3-3. Spectral characteristics of the PlanetScope Dove-R sensors.

Band	Name	Band Range (nm)	Centre Wavelength (nm)	Bandwidth (nm)
1	Blue	464-517	491	53
2	Green	547-585	566	38
3	Red	650-682	667	32
4	NIR	846-888	867	42

3.2.5 Vegetation Indices

MicaSense and PlanetScope reflectance images of each of their respective bands were exported into ArcGIS to extract crop canopy reflectance values at the sample points. With the “Raster Calculator” tool in PCI Geomatica Banff, the images were used to compute the VIs listed in Table 3-4 and exported as GeoTiff files. Using ArcGIS, the VI layer values were extracted at each sample point. As MicaSense has a red-edge band, the following three indices were included specifically for the UAV-data modelling: the chlorophyll index red-edge (CI_RE), normalized difference vegetation index (NDRE), and ratio vegetation index #2 (RVI2). VIs have been extensively studied for the purpose of crop monitoring and biophysical estimation, such as the normalized difference vegetation index (NDVI) (Jones & Vaughn, 2010; Zhu et al., 2008). VIs developed for chlorophyll estimation have been found to be related to plant nitrogen content as the photosynthetic enzyme, rubisco, encompasses the largest proportion of nitrogen in leaves (Marschner, 2011). Chlorophyll reflects green and NIR radiation and absorbs more than 70% of blue and red radiation (Daughtry et al., 2000).

Table 3-4. Vegetation indices used in the study.

Index ¹	Formula ²	Authors
BNDVI	$(\text{NIR} - \text{BLUE}) / (\text{NIR} + \text{BLUE})$	Wang et al. (2007)
CI_RE	$(\text{NIR}/\text{REDEGE}) - 1$	Gitelson et al. (2003)
ISR	RED/NIR	Fernandes et al. (2003)
MSR	$\frac{(\text{NIR}/\text{RED}) - 1}{\sqrt{\frac{\text{NIR}}{\text{RED}} + 1}}$	Chen (1996)
MTVI2	$\frac{1.8(\text{NIR} - \text{GREEN}) - 3.75(\text{RED} - \text{GREEN})}{\sqrt{(2\text{NIR} + 1)^2 - 6(\text{NIR} - 5\sqrt{\text{RED}})} - 0.5}$	Bagheri et al. (2013)
NDRE	$(\text{NIR} - \text{REDEGE}) / (\text{NIR} + \text{REDEGE})$	Gitelson and Merzyak (1994)
NDVI	$(\text{NIR} - \text{RED}) / (\text{NIR} + \text{RED})$	Rouse et al. (1974)
OSAVI	$1.6[(\text{NIR} - \text{RED}) / (\text{NIR} + \text{RED} + 0.16)]$	Rondeaux, Steven, and Baret (1996)
RDVI	$\frac{\text{NIR} - \text{RED}}{\sqrt{\text{NIR} + \text{RED}}}$	Roujean and Breon (1995)
RGBVI	$(\text{GREEN}^2 - \text{BLUE} * \text{RED}) / (\text{GREEN}^2 + \text{BLUE} * \text{RED})$	Bendig et al. (2015)
RVI	NIR/RED	Jordan (1969)
RVI2	NIR/REDEGE	Kanke et al. (2012)
WDRVI	$(0.2 * \text{NIR} - \text{RED}) / (0.2 * \text{NIR} + \text{RED})$	Gitelson (2013)
VDVI	$\frac{(2 * \text{GREEN} - \text{RED} - \text{BLUE})}{(2 * \text{GREEN} + \text{RED} + \text{BLUE})}$	Fuentes et al. (2012)

¹ Vegetation index abbreviations: BNDVI = blue normalized difference vegetation index; CI_RE = chlorophyll index red edge; ISR = infrared simple ratio; MSR = modified simple ratio; MTVI2 = modified triangular vegetation index; NDRE = normalized difference vegetation index; NDVI = normalized difference vegetation index; OSAVI = optimized soil adjusted vegetation index; RDVI = renormalized difference vegetation index; RGBVI = red green blue vegetation index; RVI = ratio vegetation index; WDRVI = wide dynamic range vegetation index; VDVI = visible band difference vegetation index

² Formula variables: BLUE = blue reflectance; GREEN = green reflectance; RED = red reflectance; REDEGE = red edge reflectance; NIR = near-infrared reflectance

3.2.6 Canopy Nitrogen Weight Modelling

Canopy nitrogen weight was calculated using the following method (Hansen & Schjoerring, 2003):

$$CNW = LNC \times W_d \quad (1)$$

where CNW is the canopy nitrogen weight (g/m^2), LNC is the leaf nitrogen content (%), and W_d is the dry biomass weight (g/m^2). CNW assumes the plants collected for biomass around the sample point, within the 0.5 m^2 block, have the same amount of nitrogen. The total biomass per area was used as dry biomass of wheat plants at early growth stages ($BBCH < 60$) since the leaves constitute most plant weight. Compared to other biophysical parameters such as plant nitrate content and plant nitrogen concentration (%), canopy nitrogen weight has been found to have greater correlation with spectral data (Li et al., 2008).

RF is an ensemble learning method suited for classification or regression of large, nonparametric datasets. Training data is randomly selected from the dataset based on a user-defined percentage; commonly 70% of a dataset is used to train a model. From the training data, many decision trees are generated by the algorithm to determine the variables' importance in the regression. Decision trees split at nodes depending on the independent variable that contributes most to the dependent variable. Validation data is the remaining dataset not used in training, and the average output from decision trees is used to evaluate the model's performance. Advantages of RF include quick computation, no overfitting from training data, and high performance in studies (Lee et al., 2020).

Support Vector Machines (SVM) are supervised learning algorithms used for classification and regression. SVR uses a decision boundary, known as a hyperplane, to split classes of training data based on data characteristics. The data points closest to either side of the hyperplane are known as support vectors, used as training samples to determine the optimal hyperplane position from the midpoint of the margin. SVR performs modelling in a high-dimensional space. SVR uses a kernel trick (i.e., Radial Basis Kernel) for nonlinear data, placing the data in a dimensional space to separate into groups based on radial distance between data points. SVR has advantages over simple linear regression models, as its

flexibility with nonparametric data has better modelling capabilities (James et al., 2013). Modelling for this study was written in R programming language using R Studio (Version 4.0.3), an open-source and free Integrated Development Environment (IDE) (Rstudio, 2020). For RF the “randomForest” package was used, and for SVR the “e1071” package.

Modelling was performed separately for MicaSense-based and PlanetScope-based sensor variables for comparison of UAV and PlanetScope data results. MicaSense models used all five MicaSense bands, all 14 Vis listed in Table 1, plant height, LAI, soil moisture, topographic metrics, and the dependent variable was the canopy nitrogen weight. PlanetScope models used all four PlanetScope bands, 11 Vis listed in Table 1 not including the red-edge based indices (CI_RE, NDRE, RVI2), plant height, LAI, soil moisture, topographic metrics, and the dependent variable was the canopy nitrogen weight. Measurements of variables at each sample point were averaged to a 1 m² scale. Datasets were randomly divided into 70% training set and 30% validation set. The quality of the models was assessed using the coefficient of determination (R²) and Root Mean Square Error (RMSE), calculated using Equation (2) and Equation (3), respectively (Palaniswamy, 2006):

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \quad (2)$$

where y_i is the observed value, \hat{y}_i is the predicted value, and \bar{y} is the mean of the observed values of the dataset; and:

$$RMSE = \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n} \quad (3)$$

where \hat{y}_i is the predicted canopy nitrogen weight value (g/m²), y_i is the observed canopy nitrogen weight value (g/m²), n is the number of observations, and I is the index of summation in increments of one.

3.3 Results

3.3.1 Nitrogen Statistics

Overall, canopy N weight for the wheat fields increased in variation during the fieldwork season (Figure 3-4). There was a slight decrease in canopy N weight from 6 May to 12 May, likely due to a several days of continuous rainfall between samplings. Rainfall can lead to leaching of soil N from areas around plant roots, reducing plant N utilization while the biomass continued to increase (Justes et al., 1994; Li et al., 2008). After 12 May, the farmer applied fertilizer to the wheat field once before 20 May and consequently the canopy N weight increased the following weeks. There is an outlier on the 20 May from a sample point and was removed from the dataset for modelling.

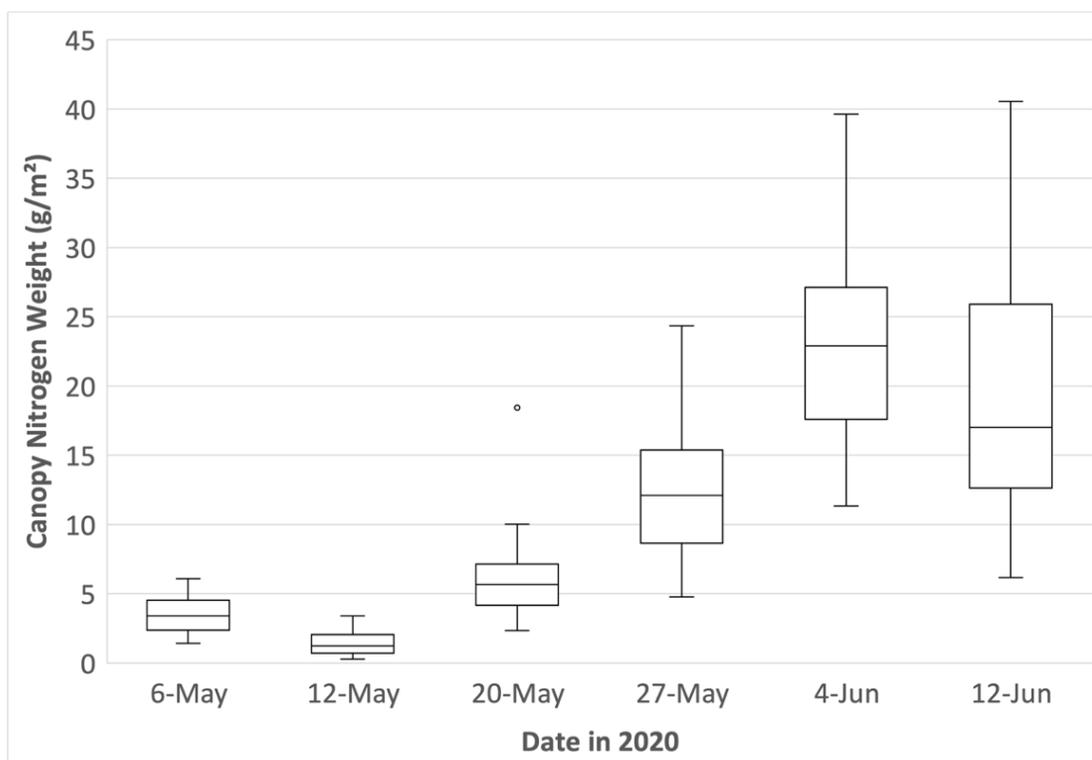


Figure 3-4. Variation of canopy nitrogen weight (g/m^2) as a function of the date of field measurements during the 2020 growing season. An outlier is represented by a dot on the graph.

3.3.2 Soil Statistics

A Vittellus soil health analysis (A&L, 2020) was conducted the study fields, producing the following mean values: soil nitrate-nitrogen of 46.81 ppm, mineralizable nitrogen of 29.17 ppm, and water-extracted soil nitrate of 44.38 ppm. A&L Canada Soil Health Index ratings were in the “Good-High” category, and soil textural class for the fields was predominantly sandy loam.

3.3.3 Regression Models with All Parameters

3.3.3.1 UAV Regression Models

For the UAV regression models, the first modelling step used all 28 variables including: the five MicaSense band reflectances, 14 Vis, plant height, soil moisture, LAI, topographic slope, aspect, profile curvature, plan curvature, TWI #1, and TWI #2. Single-date and combinations of multi-date datasets were tested to evaluate the temporal effect on models. 70% of each dataset was used to calibrate the RF and SVR models, before the remaining 30% was used to validate. From the calibrated models (Table 3-5), RF had better performance in comparison to SVR in terms of higher R^2 and lower RMSE values overall. The top model performance was RF with the combination of 12, 20, 27 May resulting in R^2 of 0.96 and RMSE of 1.07 g/m². Although 20, 27 May, 4 June model had a slightly higher R^2 of 0.97, RMSE was considerably greater at 1.76 g/m². Of the RF and SVR models, 20 May had the lowest performance. Of the SVR models, 12, 20, 27 May model had the best performance at R^2 of 0.87 and RMSE of 2.07 g/m².

Table 3-5. Statistics for the calibration of the UAV canopy nitrogen regression models with 28 variables as a function of the date and modelling approach (RF or SVR)¹.

Date	Model	R ²	RMSE (g/m ²)	(n)
5 May	RF	0.91	0.51	22
	SVR	0.78	0.70	22
12 May	RF	0.92	0.31	33
	SVR	0.80	0.46	33
20 May	RF	0.85	1.12	33
	SVR	0.58	1.89	33
27 May	RF	0.95	1.63	22
	SVR	0.84	2.43	22
4 June	RF	0.91	2.36	33
	SVR	0.81	3.00	33
5, 12 May	RF	0.95	0.55	56
	SVR	0.77	0.84	56
12, 20 May	RF	0.94	0.88	67
	SVR	0.80	1.53	67
20, 27 May	RF	0.95	1.37	56
	SVR	0.79	2.87	56
27 May, 4 June	RF	0.95	2.28	56
	SVR	0.76	3.86	56
5, 12, 20 May	RF	0.92	0.87	89
	SVR	0.79	1.40	89
12, 20, 27 May	RF	0.96	1.07	89
	SVR	0.87	2.07	89
20, 27 May, 4 June	RF	0.97	1.76	89
	SVR	0.85	3.79	89

¹ All models are significant at p-value <0.001.

The UAV models were then applied to the validation datasets (Table 3-6). All single date models for RF and SVR had low performance. This is likely due to the small number of sample point data from single date sets resulting in the calibration model not encompassing normal field data variation. The top performing model was RF and 12, 20, 27 May data resulting in R² of 0.74 and RMSE of 2.76 g/m². Overall, RF had better performance than SVR and multi-date models had better performance than single-date models.

Table 3-6. Statistics for the validation of the UAV canopy nitrogen models with 28 variables as a function of the date and modelling approach (RF or SVR).

Date	Model	R ²	p-value ¹	RMSE (g/m ²)	(n)
5 May	RF	0.18	NS	0.71	10
	SVR	0.06	NS	0.71	10
12 May	RF	0.07	NS	0.95	15
	SVR	0.14	0.1	0.79	15
20 May	RF	0.05	NS	2.28	15
	SVR	0.18	0.1	2.14	15
27 May	RF	0.25	0.1	4.35	10
	SVR	0.03	NS	5.05	10
4 June	RF	0.20	0.1	4.35	15
	SVR	0.06	NS	5.66	15
5, 12 May	RF	0.41	<0.001	0.99	24
	SVR	0.52	<0.001	0.81	24
12, 20 May	RF	0.66	<0.001	1.72	29
	SVR	0.71	<0.001	1.56	29
20, 27 May	RF	0.67	<0.001	2.26	24
	SVR	0.55	<0.001	2.38	24
27 May, 4 June	RF	0.45	<0.001	5.80	24
	SVR	0.30	0.001	6.44	24
5, 12, 20 May	RF	0.37	<0.001	1.58	39
	SVR	0.52	<0.001	1.32	39
12, 20, 27 May	RF	0.74	<0.001	2.76	39
	SVR	0.61	<0.001	3.50	39
20, 27 May, 4 June	RF	0.74	<0.001	4.12	39
	SVR	0.71	<0.001	4.46	39

¹ NS = non-significant

3.3.3.2 PlanetScope Regression Models

For the PlanetScope regression models, 24 variables were used in the first step including: the four PlanetScope band reflectances, 11 Vis, plant height, soil moisture, LAI, topographic slope, aspect, profile curvature, plan curvature, TWI #1, and TWI #2. To test temporal effect on the models, single-date and multi-date dataset combinations were used. Taking 70% of the dataset to calibrate RF and SVR models, overall RF had better performance to SVR with higher R² and lower RMSE values (Table 3-7). Of all models, RF with the combination of 12, 20, 27 May data had best performance with R² of 0.96 and

RMSE of 1.10 g/m². Of the SVR models alone, 12, 20, 27 May model had highest performance with R² of 0.87 and RMSE of 2.12 g/m².

Table 3-7. Statistics for the calibration of the PlanetScope canopy nitrogen models with 24 variables as a function of the date and the modelling approach (RF or SVR)¹.

Date	Model	R ²	RMSE (g/m ²)	(n)
5 May	RF	0.94	0.40	33
	SVR	0.82	0.60	33
12 May	RF	0.92	0.30	33
	SVR	0.71	0.49	33
20 May	RF	0.85	1.23	33
	SVR	0.52	2.03	33
27 May	RF	0.90	1.89	33
	SVR	0.60	3.41	33
4 June	RF	0.91	2.23	33
	SVR	0.70	3.80	33
5, 12 May	RF	0.95	0.55	67
	SVR	0.76	0.83	67
12, 20 May	RF	0.94	0.78	67
	SVR	0.78	1.48	67
20, 27 May	RF	0.94	1.27	67
	SVR	0.75	2.57	67
27 May, 4 June	RF	0.94	1.93	67
	SVR	0.74	3.87	67
5, 12, 20 May	RF	0.93	0.75	100
	SVR	0.74	1.34	100
12, 20, 27 May	RF	0.96	1.10	100
	SVR	0.87	2.12	100
20, 27 May, 4 June	RF	0.95	1.87	100
	SVR	0.83	3.68	100

¹ All models are significant at p-value <0.001.

Next, the PlanetScope models were applied to the remaining 30% of datasets for validation (Table 3-8). Of all the models, single-date 20 May and 4 June had very low performance for both RF and SVR. Comparing all RF to SVR models, RF usually had better performance than SVR except in the single date models of 5 May, 12 May, and 20 May and the multi-date model with the three dates combined. The top performing model was RF with 12, 20, 27 May data resulting in R² of 0.83 and RMSE of 1.77 g/m².

Table 3-8. Statistics for the validation of PlanetScope canopy nitrogen models with 24 variables as a function of the date and the modelling approach (RF or SVR).

Date	Model	R ²	p-value ¹	RMSE (g/m ²)	(n)
5 May	RF	0.44	0.001	0.92	15
	SVR	0.51	0.001	0.90	15
12 May	RF	0.45	0.001	0.62	15
	SVR	0.78	<0.001	0.53	15
20 May	RF	0.05	NS	2.24	15
	SVR	0.38	0.001	1.90	15
27 May	RF	0.63	<0.001	2.84	15
	SVR	0.50	0.001	3.40	15
4 June	RF	0.07	NS	6.24	15
	SVR	0.05	NS	5.54	15
5, 12 May	RF	0.44	<0.001	1.06	29
	SVR	0.29	0.001	1.08	29
12, 20 May	RF	0.79	<0.001	1.37	29
	SVR	0.71	<0.001	1.03	29
20, 27 May	RF	0.63	<0.001	3.24	29
	SVR	0.50	<0.001	3.79	29
27 May, 4 June	RF	0.63	<0.001	4.77	29
	SVR	0.63	<0.001	4.94	29
5, 12, 20 May	RF	0.50	<0.001	1.73	44
	SVR	0.68	<0.001	1.43	44
12, 20, 27 May	RF	0.83	<0.001	1.77	44
	SVR	0.67	<0.001	3.16	44
20, 27 May, 4 June	RF	0.75	<0.001	3.64	44
	SVR	0.74	<0.001	4.59	44

¹ NS = non-significant

3.3.4 Variable Importance Plots

The top performing UAV and PlanetScope regression models with all variables were both from RF and multi-date combination of 12, 20, 27 May datasets. Modelling of RF in Rstudio can be visualized with a variable importance plot using the “varImpPlot()” function. The higher a variable’s IncNodePurity value is, the more important the explanatory variable is in the prediction of canopy nitrogen weight. In the UAV model plot (Figure 3-5) with 28 variables, plant height was the most important predictor by far then followed by LAI. Red-edge Vis (NDRE, RVI2, CI_RE) were the third, fourth and sixth most important predictor variables respectively. Of the MicaSense band reflectances, NIR

was the most important. Soil moisture also appears to be among the top group of important variables with a greater difference in IncNodePurity compared to the other variables in the plot below. Of the topographic metrics, TWI #2 and TWI #1 held more importance than the rest.

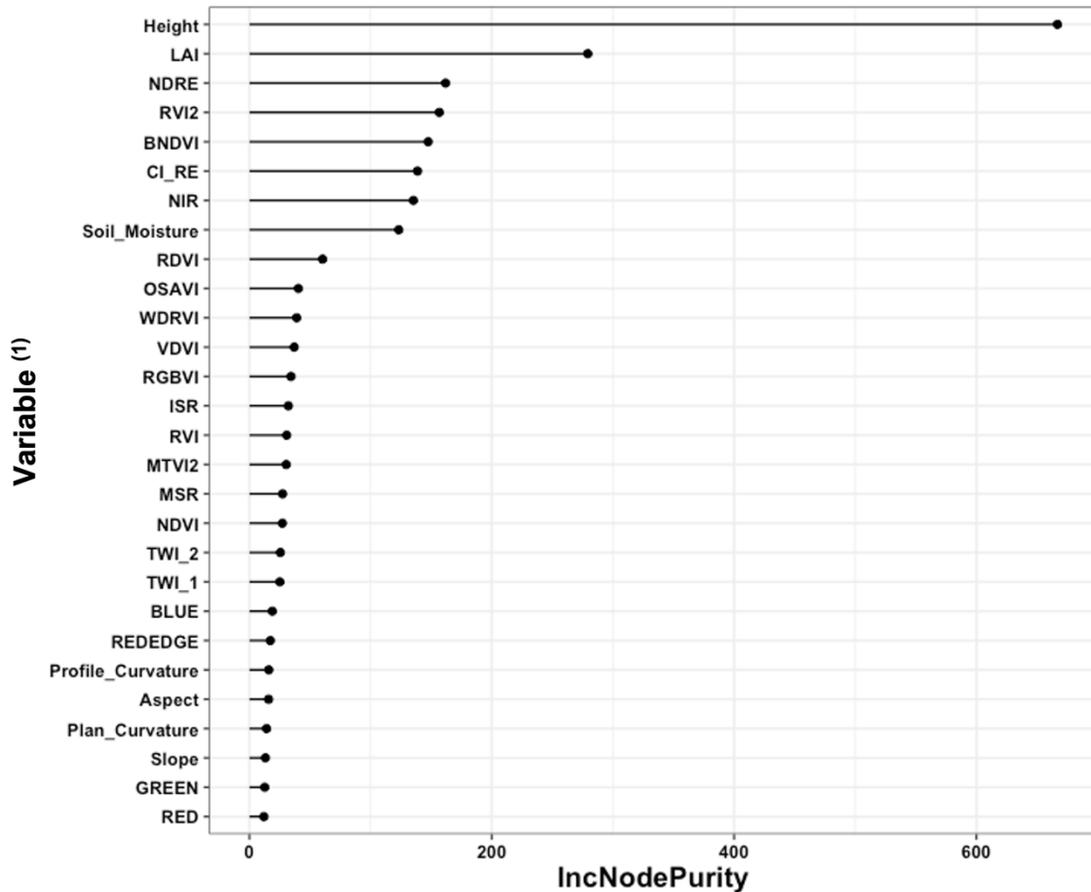


Figure 3-5. Variance Importance plot produced by the UAV RF regression model of data from 12, 20, and 27 May and 28 variables, using the function varImpPlot() in Rstudio. Higher IncNodePurity values indicate more impact on canopy nitrogen estimation. (1) Please refer to Table 1 for the full name of vegetation indices.

TWI_1, total wetness index #1; TWI_2, total wetness index #2.

The PlanetScope RF regression model plot with 24 variables indicates height as the most important predictor for canopy nitrogen weight (Figure 3-6). Of the four PlanetScope band reflectances, the blue band was most important and second-most important overall of all variables followed closely by LAI. Overall, Vis and PlanetScope bands were

among the most important variables while topographic metrics and soil moisture were the least important.

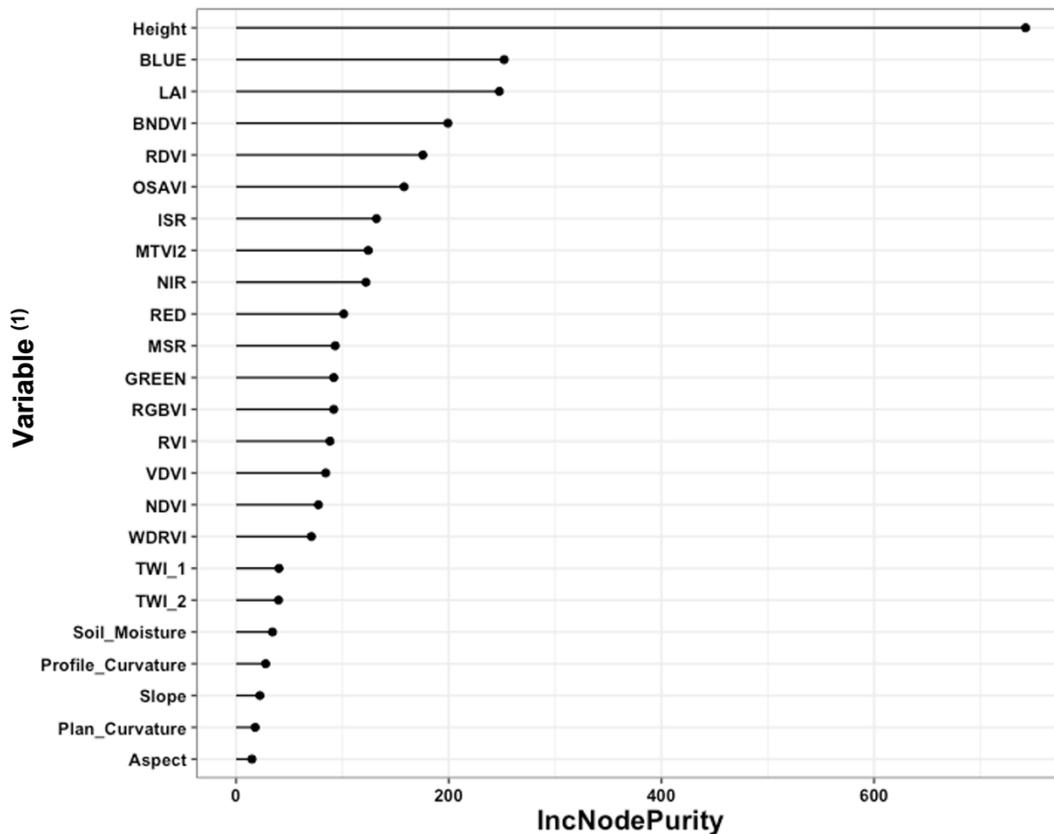


Figure 3-6. Variance Importance plot produced by the PlanetScope RF regression model of data from 12, 20, and 27 May and 24 variables, using the function `varImpPlot()` in Rstudio. Higher IncNodePurity values indicate more impact on canopy nitrogen estimation. (1) Please refer to Table 1 for the full name of vegetation indices. TWI_1, total wetness index #1; TWI_2, total wetness index #2.

3.3.5 Regression Models with Selected Parameters

3.3.5.1 UAV Regression Models

From the variable importance plot of the best performing UAV regression model using 12, 20, 27 May data, combinations of variables were tested in RF and SVR models. Evaluating based on thresholds of variable importance, as shown in Figure 3-5, we selected variable groups for the top 7, 8, 9, 13, and 16 variable groups. An additional group containing only

MicaSense band reflectances and Vis was included for comparison as common modelling in other studies use only spectral variables (Sishodia et al., 2020). 70% of each variable group dataset was used for model calibration (Table 3-9). Overall, RF had better performance than SVR in terms of higher R^2 and lower RMSE values. All variable group models using RF had a high R^2 of 0.96 except for the spectral-only variable group with lower R^2 . RF models' RMSE values were also very close within $<0.10 \text{ g/m}^2$ difference, the lowest being top 9 and top 13 variable groups at 1.07 g/m^2 . Again, the RF spectral-only variable group had poorer performance in comparison to other RF models, as its RMSE value is significantly higher. The SVR spectral-only variable group performed poorly compared to all the other SVR models. Of the SVR models, the Top 7 variable group had the best performance with R^2 of 0.86 and RMSE of 2.05 g/m^2 . Top 8, 9, 13, and 16 variable SVR models have only slightly lower R^2 and higher RMSE values compared to the best SVR model.

Table 3-9. Statistics for the calibration of the canopy nitrogen model with 12, 20, 27 May MicaSense data and different combinations of variables (n = 89)¹.

Input Variables	Model	Number of Variables	R^2	RMSE (g/m^2)
Spectral-only: All VIs & 5 MicaSense bands	RF	19	0.92	1.50
	SVR	19	0.66	3.04
Top 7: Height, LAI, NDRE, RVI2, BNDVI, CI_RE, NIR	RF	7	0.96	1.15
	SVR	7	0.86	2.05
Top 8: Top 7 + Soil Moisture	RF	8	0.96	1.12
	SVR	8	0.85	2.21
Top 9: Top 8 + RDVI	RF	9	0.96	1.07
	SVR	9	0.85	2.13
Top 13: Top 9 + OSAVI, WDRVI, VDVI, RGBVI	RF	13	0.96	1.07
	SVR	13	0.85	2.10
Top 16: Top 13 + ISR, RVI, MTVI2	RF	16	0.96	1.11
	SVR	16	0.85	2.11

¹ All models are significant at p-value <0.001 .

The calibrated models were applied to the remaining 30% of datasets for validation (Table 3-10). In the validation models, SVR had better performance than RF except for the spectral-only model. The best performing model is SVR with top 7 variables, resulting in

R^2 of 0.80 and RMSE of 2.62 g/m². Compared to the best model other SVR models had close performance, but as more variables were added the R^2 had lower values.

Table 3-10. Statistics for the validation of the canopy nitrogen model with 12, 20, 27 May MicaSense data and different combinations of variables (n = 39)¹.

Input Variables	Model	Number of Variables	R ²	RMSE (g/m ²)
Spectral-only: All VIs & 5 MicaSense bands	RF	19	0.50	3.83
	SVR	19	0.48	3.87
Top 7: Height, LAI, NDRE, RVI2, BNDVI, CI_RE, NIR	RF	7	0.73	2.81
	SVR	7	0.80	2.62
Top 8: Top 7 + Soil Moisture	RF	8	0.73	2.78
	SVR	8	0.78	2.82
Top 9: Top 8 + RDVI	RF	9	0.74	2.77
	SVR	9	0.77	2.80
Top 13: Top 9 + OSAVI, WDRVI, VDVI, RGBVI	RF	13	0.75	2.72
	SVR	13	0.77	2.76
Top 16: Top 13 + ISR, RVI, MTVI2	RF	16	0.74	2.73
	SVR	16	0.76	2.81

¹ All models are significant at p-value <0.001.

3.3.5.2 PlanetScope Regression Models

For the UAV regression models using single and multi-date datasets, the best performing combination was the model using 12, 20, 27 May data with RF. Evaluating by thresholds of variable importance from Figure 6, the selected variable groups for further model testing included top 6, 10, 13, 17, and a group containing only PlanetScope band reflectances and VIs. 70% of each variable group dataset was used for model calibration (Table 3-11). Of the calibration models, all RF models had the same high R^2 value with low RMSE compared to all SVR models. The best model performance was RF with top 6 variables with an R^2 at 0.96 with RMSE 1.18 g/m². For SVR models, the best performing was the combination of top 6 variables resulting in R^2 of 0.84 and RMSE of 2.31 g/m². From the calibration models, the spectral-only SVR model had the poorest performance at R^2 of 0.77 and RMSE of 2.81 g/m².

Table 3-11. Statistics for the calibration of the PlanetScope canopy nitrogen model with 12, 20, 27 May data and different combinations of variables (n = 100)¹.

Input Variables	Model	Number of Variables	R ²	RMSE (g/m ²)
Spectral-only: VIs & 4 PlanetScope bands	RF	15	0.96	1.21
	SVR	15	0.77	2.81
Top 6: Height, BLUE, LAI, BNDVI, RDVI, OSAVI	RF	6	0.96	1.18
	SVR	6	0.84	2.31
Top 10: Top 6 + ISR, MTVI2, NIR, RED	RF	10	0.96	1.20
	SVR	10	0.84	2.35
Top 13: Top 10 + MSR, GREEN, RGBVI	RF	13	0.96	1.21
	SVR	13	0.81	2.45
Top 17: Top 13 + RVI, VDVI, NDVI, WDRVI	RF	17	0.96	1.26
	SVR	17	0.82	2.63

¹ All models are significant at p -value <0.001.

The remaining 30% of variable datasets were used for validation models (Table 3-12). The poorest model performance for RF and SVR were from the spectral-only variable group, resulting in the lowest R² and highest RMSE values out of all tested. SVR had better performance compared to RF in the spectral-only, top 6, and top 10 groups, but for top 13 and top 17 RF was better. The best performing model is RF with top 17 variables with a R² of 0.92 and RMSE of 1.75 g/m².

Table 3-12. Statistics for the validation of the PlanetScope canopy nitrogen model with 12, 20, 27 May data and different combinations of variables (n = 44)¹.

Input Variables	Model	Number of Variables	R ²	RMSE (g/m ²)
Spectral-only: VIs & 4 PlanetScope bands	RF	15	0.66	3.20
	SVR	15	0.69	3.09
Top 6: Height, BLUE, LAI, BNDVI, RDVI, OSAVI	RF	6	0.72	2.90
	SVR	6	0.85	2.17
Top 10: Top 6 + ISR, MTVI2, NIR, RED	RF	10	0.83	2.16
	SVR	10	0.84	2.06
Top 13: Top 10 + MSR, GREEN, RGBVI	RF	13	0.87	2.20
	SVR	13	0.86	2.39
Top 17: Top 13 + RVI, VDVI, NDVI, WDRVI	RF	17	0.92	1.75
	SVR	17	0.91	1.85

¹ All models are significant at p -value <0.001.

3.3.6 Crop Nitrogen Prediction Maps

To create the crop nitrogen prediction maps, raster layers of each variable was required. The VI raster layers were created using the raster calculator in PCI Geomatica Banff from the MicaSense band rasters mosaicked from Pix4Dmapper. To extract raster layers for height, the Phantom 4 RTK flight processed in Pix4Dmapper include the option for a digital surface model (DSM) output. The DSM captures the natural and built features of the environment. By subtracting the DEM raster from the DSM raster pixel values using ArcGIS raster calculator, the output raster layer has the height data of the crop (Wu et al., 2017).

Generating an LAI raster layer is a more involved process. Based on methodology proposed by Song et al. (2020) using a simulated observation of point cloud Multiview angle (SOPC-M) designed to obtain 3D spatial distribution of vegetation and bare ground points to calculate gap fraction then crop effective LAI (LAI_e) from a UAV-based 3D point cloud dataset. In ArcGIS, the point data layer was converted to a raster layer using the “Point to Raster” conversion tool to a 2 m x 2 m resolution. From the linear regression equation, the digital hemispherical photography (DHP) LAI_e could be calculated using the ArcGIS raster layer to create the final LAI raster used in the model. For field W3, the LAI layer processing area error unfortunately only covered 12 of the 16 sample points, thus the final map produced does not cover all sample points.

Prior to running the regression models in R, the raster layers need to have the same resolution and extent for the functions to stack them. All raster layers were run through ArcGIS “Resample” tool to 1 m x 1 m resolution, as the canopy nitrogen weight variable is based on a 1 m² area. Resampling was conducting with the bilinear technique, which is a bilinear interpolation and determines the new value of a cell based on a weighted distance average of the four nearest input cell centers. As the LAI raster extent was the smallest of all rasters, it was used as the output extent reference layer for the ArcGIS “Clip Raster” tool with the selected option to maintain clipping extent. The columns and rows of the output raster are adjusted, and pixels resampled to match the reference layer exactly. Next, the ArcGIS “Extract by Mask” tool was used to extract the cell values specifically where

the LAI raster extent pixels held values greater than zero. Thus, each variable layer was prepared for the R modelling.

For the UAV best performing model, SVR with top 7 variables included height, LAI, NDRE, RVI2, BNDVI, CI_RE, and NIR. In R, the “raster”, “rgdal”, and “rasterVis” libraries were used to generate the final prediction map outputs. The variable raster layers were stacked using the “raster::stack” function, and “raster::predict” tool was used with the selected UAV-based SVR model using top 7 variables. The resulting prediction rasters for each field were exported as GeoTiff files into ArcGIS to create the final prediction maps with 1 m x 1 m resolution (Figure 3-7). The low and high canopy nitrogen weight areas are displayed in red and green respectively for distinct contrast between the nitrogen levels. From the prediction rasters, the predicted canopy nitrogen weight values around each sample point were extracted and compared to the measured values in-field. The resulting RMSE for W1, W2, and W3 were 4.27, 2.32, and 3.08 g/m² respectively.

The satellite best performing model was using RF with top 17 variables: height, LAI, 11 VIs, and the four PlanetScope bands. Following the same processing steps as the UAV-based model, the satellite variable raster images were processed to have the same extent and 1 m x 1 m resolution (Figure 3-7). As the satellite images began with 3 m x 3 m resolution, much larger compared to the UAV images, the resulting raster layers have a smoothed appearance. The final prediction rasters for each field are displayed as maps in Figure 8 with 1 m x 1 m resolution. From the prediction rasters, the predicted canopy nitrogen weight values around each sample point were extracted and compared to the measured values in-field. The resulting RMSE for W1, W2, and W3 were 3.12, 1.79, and 3.08 g/m² respectively.

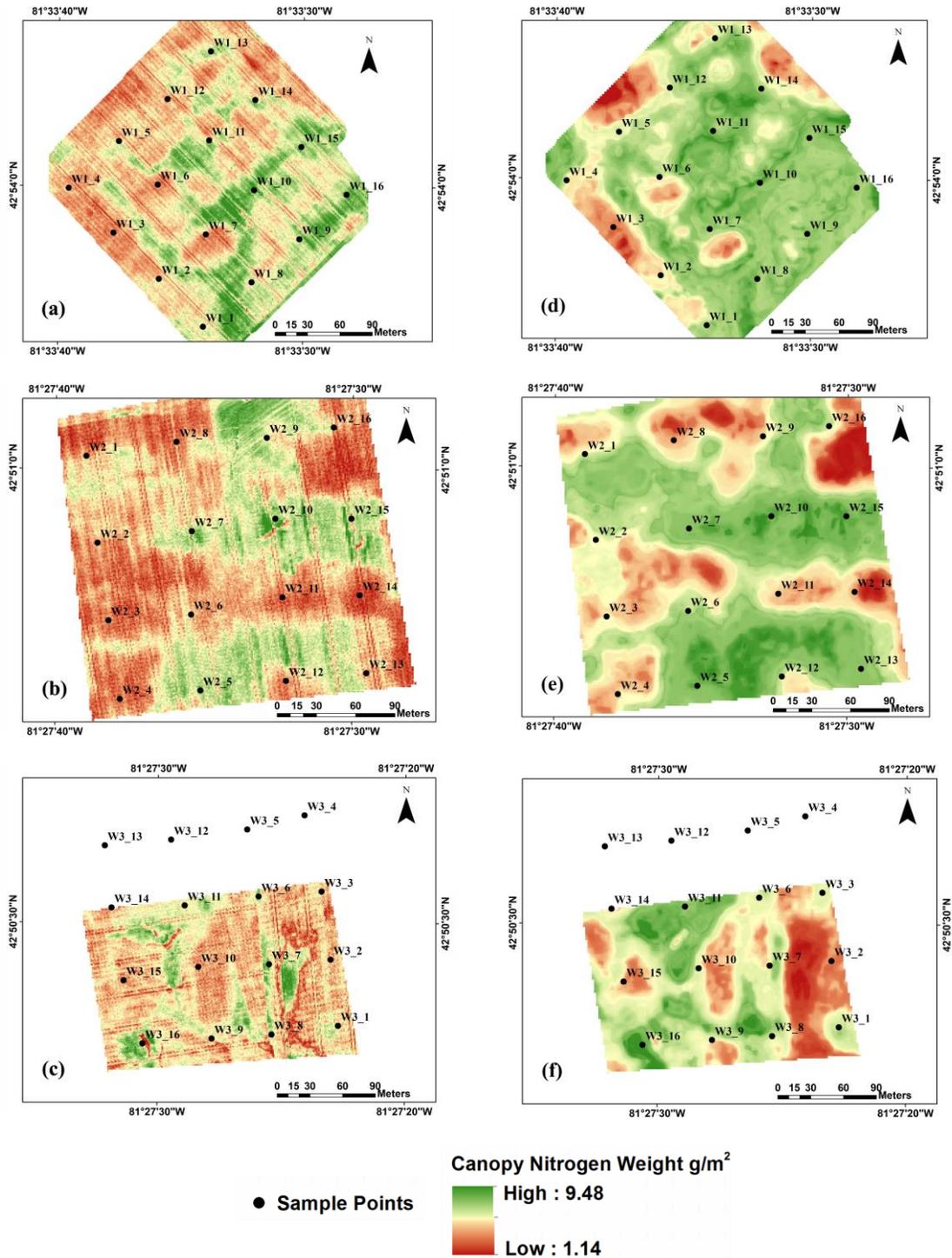


Figure 3-7. Nitrogen prediction maps of wheat fields: from UAV-based models for a) W1, b) W2, and c) W3 fields; from satellite-based models for d) W1, e) W2, and f) W3 fields.

From the maps in Figure 3-7, there are differences in the spatial distribution and values of CNW. It appears the UAV maps show more areas with low canopy nitrogen weight, while PlanetScope shows more areas with high CNW. This is likely due to the resampling method, as the UAV's centimeter-level image resolution includes spectral data of the visible soil between crop rows. While for PlanetScope images with coarser meter-level resolution, the pixel values average towards the spectral values of the crop rather than the soil.

The UAV-based nitrogen prediction maps may have lower performance compared to satellite-based maps because of the volatile weather conditions week of May 27th fieldwork. The UAV was flown on different dates between May 26th – May 29th for each field whenever the weather was suitable, and the crop spectral characteristics may have changed. Whereas the satellite images for the three fields for May 27th were all from the same date and present a more consistent relationship with the fieldwork data.

3.4 Discussion

For this study, RF and SVM regression methods were used to predict canopy nitrogen weight of wheat using UAV Micasense band reflectances, PlanetScope band reflectances, selected VIs, plant height, LAI, soil moisture, and topographic metrics. The models created were grouped by UAV-based and satellite-based data.

For the UAV RF and SVR regression models, calibration was conducted with 28 variables from single and multi-date datasets. Evaluating the validation models of each dataset, the performance of UAV single-date models was poor with R^2 values at most 0.25 and overall non-significant results. Combining UAV multi-date data yielded better results, with the best performance from the RF three-date model 12, 20, 27 May resulting in R^2 of 0.74 and RMSE of 2.76 g/m². For PlanetScope RF and SVR models, calibration of models used 24 variables for single and multi-date datasets. Of the validation models, single-date models 20 May and 4 June had the lowest performance. However, the other PlanetScope single-date models performances overall were much better compared to UAV single-date models. In general, PlanetScope multi-date models did not have significantly better results than its

single-date models. The best performing PlanetScope model was from three-date 12, 20, 27 May using RF at R^2 of 0.83 and RMSE 1.77 g/m². Both the UAV and satellite best models were from 12, 20, 27 May data, during which the wheat crops were in BBCH 23-41 growth stages mainly defined by tillering, stem elongation, and the beginning of the booting stage. As noted by Hawkesford (2017), the application of nitrogen fertilizer during these early growth stages before flowering is most conducive to efficient nitrogen use and yield response. To be able to accurately estimate nitrogen levels of the crop during early growth stages will be most beneficial for farmers.

In the RF variance importance plot of the best performing UAV model, of all variables plant height was the most important predictor of canopy nitrogen weight. Song & Wang (2019) have also noted that plant height is useful in estimating phenology, biomass, and yield in addition to nitrogen uptake in wheat. On the plot, LAI was the second most important predictor of canopy nitrogen weight. LAI has been used extensively in studies to predict crop chlorophyll content, biomass, and yield successfully (Reisi et al., 2020; Song et al., 2021). The study by Zhao et al. (2014) found a significant positive relationship between LAI and differences in crop nitrogen content across wheat growth stages. The gap fraction method of calculating LAI is more accurate during earlier growth stages of a crop when the canopy is not so dense, allowing for contrast between the vegetation and soil or vegetation and sky images (Song & Wang, 2019). Among the VI's used in the model, the red-edge VIs (NDRE, RVI2, CI_RE) were amongst the most important. The red-edge region (680-800 nm) has been shown to encompass sharp change in the canopy reflectance and can be used to identify important biophysical parameters of the crop. Nitrogen levels have shown the sensitivity of the red-edge region in estimating leaf chlorophyll content due to the high absorption of red radiation and high reflectance of NIR radiation (Reisi Gahrouei et al., 2020; Schlemmer et al., 2013). Of the MicaSense bands individually, the NIR band was of highest importance in the model while other individual bands had little effect. Soil moisture also appeared as a variable of high importance, and subsequent variables on the plot had noticeably lower importance. Studies have noted the importance of soil moisture in soil nitrogen mineralization, crop nitrogen uptake, and utilization (Mulvaney et al., 2006; Tremblay et al. 2012). Of the topographic metrics the topographic wetness indices were most important, while the remaining metrics had little effect.

In the RF variance importance plot of the best performing satellite model, similar to the UAV model, plant height was the most important predictor of canopy nitrogen weight with LAI following closely. Interestingly, the PlanetScope blue band was the second most important variable. As PlanetScope blue band has greater width compared to MicaSense, perhaps the wider bandwidth captured a change of canopy reflectance in the blue-edge region (480-517nm) which has been noted before in the study by Wei et al. (2008) to be related to crop nitrogen. Other PlanetScope bands in the model had varying levels of importance interspersed amongst the 11 VIs used. On the plot, other non-spectral variables of soil moisture and topographic metrics were of least importance to the model.

From the UAV-based RF variance importance plot, groups of variables were selected for testing in models. Groups of the top 7, 8, 9, 13, 16, and spectral-only variables were modeled, with the group of top 7 variables SVR model performing the best with R^2 of 0.80 and RMSE of 2.62 g/m². Top 7 variables included plant height, LAI, all three red-edge VIs, BNDVI, and MicaSense NIR band. Comparing to best models from studies by Asataoui et al. (2021), Jiang et al. (2019), and Zheng et al. (2018) using UAV-based spectral variables to estimate wheat nitrogen content, their models had lower R^2 values ranging from 0.76-0.63 and greater RMSE values. With the UAV-based best model in this study, significantly lower RMSE is a major advantage in reducing costs of nitrogen fertilizer.

The satellite-based variance importance plot was used to select variable groups for model testing including top 6, 10, 13, 17, and spectral-only variables. The RF model group of top 17 variables had the best performance with R^2 of 0.92 and RMSE of 1.75 g/m². Height, LAI, all four PlanetScope bands, and total 11 VIs were variables in the best performing model. For both UAV and satellite best performing selected variable models, plant height and LAI are the only non-spectral variables. With methods of deriving height and LAI of a wheat crop field from the UAV Phantom 4 RTK imagery, all variables in the top models can be obtained from in-situ, non-destructive, remote sensing data.

For both UAV and satellite spectral-only variable groups, results were poor with R^2 values <0.50 and significantly higher RMSE values compared to other tested variable groups. This is consistent with the studies by Astaoui et al. (2021) and Schirrmann et al. (2016) noting

within wheat crops, UAV imagery was limited for observing nitrogen status but had good performance in monitoring biophysical parameters. The UAV validation spectral-only model compared to the best performing top 7 variable model, RMSE dropped by 32%. In the satellite validation spectral-only model compared to its best performing top 17 variable model, RMSE dropped by 45%.

In the final validation of canopy nitrogen models with variable combinations, UAV SVR models mostly had greater R^2 values, but greater RMSE values as well compared to RF models. In the UAV spectral-only variable group models, RF had better results than SVR. Considering studies with spectral-only variables for crop nitrogen models, the results are consistent with RF yielding better nitrogen level prediction compared to SVR models (Liu et al. 2016; Lee et al. 2020; Zha et al. 2020). Only in the UAV best performing model of top 7 variables was SVR performance better in both a higher R^2 value and lower RMSE compared to RF. Of the satellite variable combination models, SVR had better performance than RF except for the top 13 and 17 variable groups. The best performing satellite model was RF with the top 17 variable group. Although it appears difficult to determine if RF or SVR models are better when built with non-spectral and spectral variables together, ultimately the ideal result is a model which can most accurately predict canopy nitrogen in wheat. In both the UAV and satellite models with different variable combinations, overall the top variable groups have good performances. In comparison to studies with spectral-only variable models, the variable combination models in this study all have lower RMSE values. In the context of nitrogen estimation and practical application, lower RMSE (g/m^2) in models is most beneficial for fertilizer management recommendations.

3.5 Conclusion

In this study, machine learning regression methods were tested to predict wheat canopy nitrogen weight using UAV MicaSense band reflectances, PlanetScope band reflectances, associated VIs, plant height, LAI, soil moisture, and topographic metrics. For UAV models using 28 variables, the combination of 12, 20, 27 May data with the RF validation model produced the best results with R^2 of 0.74 and RMSE of 2.76 g/m². From the model's variable importance plot, the top 7, 8, 9, 13, 16, and spectral-only variable groups were tested. The best validation model used SVR with the top 7 variables, which included plant height, LAI, four VIs, and the MicaSense NIR band. For the PlanetScope models using 24 variables, the best performing model was RF with 12, 20, 27 May data resulting in R^2 of 0.83 and RMSE of 1.77 g/m². Based on the model's variable importance plot the top 6, 10, 13, 17, and spectral-only variable groups tested. The validation model with the best performance was RF using top 17 variables including height, LAI, all four PlanetScope bands, and 11 VIs.

A common limitation of in-situ agricultural models including those developed in this study are the empirical nature, and applicability can be limited to the dataset they are built and validated upon. Each field and growing season has different conditions and factors that affect plant growth, so models will need further testing to determine their effectiveness in precision agriculture methods. PlanetScope satellite constellation has also launched and a third generation of sensors in 2020 known as SuperDove with the potential of capturing imagery with eight spectral bands including a red-edge band. Future work can consider further testing satellite-based nitrogen prediction models including red-edge VI variables.

3.6 References

- A&L Canada Laboratories. (n.d.). *VitTellus soil health test & VitTellus Bio*. Retrieved June 14, 2020 from <https://www.alcanada.com/content/solutions/soil-health>
- Agriculture and Agri-Food Canada. (n.d.) *An overview of the Canadian agriculture and agri-food system*. Retrieved February 23, 2021 from <https://www.agr.gc.ca/eng/canadas-agriculture-sectors/an-overview-of-the-canadian-agriculture-and-agri-food-system-2017/?id=1510326669269>
- Association of Official Analytical Collaboration (AOAC) (2006) *International Official Method 972.43. Microchemical determination of carbon, hydrogen, and nitrogen, automated method*. AOAC
Int.http://www.aocofficialmethod.org/index.php?main_page=product_info&cPath=1&products_id=1492
- Astaoui, G., Dadaiss, J. E., Sebari, I., Benmansour, S., & Mohamed, E. (2021). Mapping wheat dry matter and nitrogen content dynamics and estimation of wheat yield using UAV multispectral imagery machine learning and a variety-based approach: case study of Morocco. *AgriEngineering*, 3(1), 29-49.
<https://doi.org/10.3390/agriengineering3010003>
- Bagheri, N., Ahmadi, H., Alavipanah, S. K., & Omid, M. (2013). Multispectral remote sensing for site-specific nitrogen fertilizer management. *Pesquisa Agropecuária Brasileira*, 48, 1394-1401. <https://doi.org/10.1590/S0100-204X2013001000011>
- Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., Gnyp, M.L., & 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. <https://doi.org/10.1016/j.jag.2015.02.012>
- Bongiovanni, R., & Lowenberg-DeBoer, J. (2004). Precision agriculture and sustainability. *Precision agriculture*, 5(4), 359-387.
<https://doi.org/10.1023/B:PRAG.0000040806.39604.aa>
- Breunig, F. M., Galvão, L. S., Dalagnol, R., Santi, A. L., Della Flora, D. P., & Chen, S. (2020). Assessing the effect of spatial resolution on the delineation of management zones for smallholder farming in southern Brazil. *Remote Sensing Applications: Society and Environment*, 19, 100325.
<https://doi.org/10.1016/j.rsase.2020.100325>
- 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.
<https://doi.org/10.1080/07038992.1996.10855178>
- 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.
<https://doi.org/10.1016/j.rse.2010.04.006>

- Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., & Böhner, J. (2015). System for automated geoscientific analyses (SAGA) v. 2.1. 4. *Geoscientific Model Development*, 8(7), 1991-2007. <https://doi.org/10.5194/gmd-8-1991-2015>
- Da-Jiang Innovations (DJI). (n.d.) Retrieved April 15, 2021 from <https://www.dji.com/ca>
- Daughtry, C. S., 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. [https://doi.org/10.1016/S0034-4257\(00\)00113-9](https://doi.org/10.1016/S0034-4257(00)00113-9)
- Delta-T Devices Ltd. (n.d.) User Manual for the ML3 ThetaProbe. Retrieved May 3, 2020 from <https://delta-t.co.uk/wp-content/uploads/2017/02/ML3-user-manual-version-2.1.pdf>
- Drew, M. C., Sisworo, E. J. (1977) Early effects of flooding on nitrogen deficiency and leaf chlorosis in barley. *New Phytol.*, 79, 567-571. <https://doi.org/10.1111/j.1469-8137.1977.tb02241.x>
- Eitel, J. U. H., Long, D. S., Gessler, P. E., & Smith, A. M. S. (2007). Using in-situ measurements to evaluate the new RapidEye™ satellite series for prediction of wheat nitrogen status. *International Journal of Remote Sensing*, 28(18), 4183-4190. <https://doi.org/10.1080/01431160701422213>
- Fageria, N. K. (2009). *The use of nutrients in crop plants*. CRC Press: New York, U.S.A.
- Fageria, N. K.; Baligar, V. C.; Jones, C. A. (2010) Growth and mineral nutrition of field crops, 3rd ed.; CRC Press: New York, U.S.A.
- 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. <https://doi.org/10.5589/m02-092>
- Foulkes, M. J., Slafer, G. A., Davies, W. J., Berry, P. M., Sylvester-Bradley, R., Martre, P., Calderini, D., Griffiths, S., Reynolds, M. P. (2011) Raising yield potential of wheat. III. Optimizing partitioning to grain while maintaining lodging resistance. *Journal of Experimental Botany*, 62, 469-486. <https://doi.org/10.1093/jxb/erq300>
- 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. <https://doi.org/10.1016/j.fcr.2017.12.004>
- Fuentes, S., De Bei, R., Pech, J., & Tyerman, S. (2012). Computational water stress indices obtained from thermal image analysis of grapevine canopies. *Irrigation Science*, 30(6), 523-536. <https://doi.org/10.1007/s00271-012-0375-8>
- 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. [https://doi.org/10.1016/1011-1344\(93\)06963-4](https://doi.org/10.1016/1011-1344(93)06963-4)

- 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. <https://doi.org/10.1080/01431161.2013.793868>
- 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. <https://doi.org/10.1078/0176-1617-00887>
- Hansen, N. C., Tubbs, S., Fernandez, F., Green, S., Hansen, N. E., & Stevens, W. B. (2015). Conservation agriculture in North America. In *Conservation Agriculture* (pp. 417-441). Springer, Cham. https://doi.org/10.1007/978-3-319-11620-4_17
- 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. [https://doi.org/10.1016/S0034-4257\(03\)00131-7](https://doi.org/10.1016/S0034-4257(03)00131-7)
- 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. <https://doi.org/10.3390/rs4061573>
- Hawkesford, M. J. (2017). Genetic variation in traits for nitrogen use efficiency in wheat. *Journal of Experimental Botany*, 68(10), 2627-2632. <https://doi.org/10.1093/jxb/erx079>
- Hunt, E. R., Hively, W. D., Fujikawa, S. J., Linden, D. S., Daughtry, C. S., & McCarty, G. W. (2010). Acquisition of NIR-green-blue digital photographs from unmanned aircraft for crop monitoring. *Remote Sensing*, 2(1), 290-305. <https://doi.org/10.3390/rs2010290>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning: With applications in R*. (Vol. 112, p. 18). New York: Springer.
- Jiang, J., Cai, W., Zheng, H., Cheng, T., Tian, Y., Zhu, Y., ... & Yao, X. (2019). Using digital cameras on an unmanned aerial vehicle to derive optimum color vegetation indices for leaf nitrogen concentration monitoring in winter wheat. *Remote Sensing*, 11(22), 2667. <https://doi.org/10.3390/rs11222667>
- Jones, H.G.; Vaughan, R.A. (2010). *Remote sensing of vegetation: principles, techniques, and applications*, 1st ed.; University Press: New York, U.S.A.
- Jordan, C. F. (1969). Derivation of leaf-area index from quality of light on the forest floor. *Ecology*, 50(4), 663-666. <https://doi.org/10.2307/1936256>
- Justes, E., Mary, B., Meynard, J. M., Mached, J. M., & Thelier-Huché, L. (1994). Determination of a critical nitrogen dilution curve for winter wheat crops. *Annals of botany*, 74(4), 397-407. <https://doi.org/10.1006/anbo.1994.1133>
- Kanke, Y., Raun, W., Solie, J., Stone, M., & Taylor, R. (2012). Red edge as a potential index for detecting differences in plant nitrogen status in winter wheat. *Journal of*

plant nutrition, 35(10), 1526-1541.
<https://doi.org/10.1080/01904167.2012.689912>

- 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.
<https://doi.org/10.3390/rs12132071>
- Li-cor, *LAI-2200C Plant Canopy Analyzer Product*. Retrieved May 30, 2020 from
https://www.licor.com/env/products/leaf_area/LAI-2200C/
- Li, F., Gnyp, M. L., Jia, L., Miao, Y., Yu, Z., Koppe, W., Bareth, G., Chen, X., & 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.
<https://doi.org/10.1016/j.fcr.2007.11.001>
- Liu, Y., Cheng, T., Zhu, Y., Tian, Y., Cao, W., Yao, X., & Wang, N. (2016, July). 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)* (pp. 7362-7365). IEEE, Beijing, China.
<https://doi.org/10.1109/IGARSS.2016.7730920>
- Manivasagam, V. S., Sadeh, Y., Kaplan, G., Bonfil, D. J., & Rozenstein, O. (2021). Studying the Feasibility of Assimilating Sentinel-2 and PlanetScope Imagery into the SAFY Crop Model to Predict Within-Field Wheat Yield. *Remote Sensing*, 13(12), 2395. <https://doi.org/10.3390/rs13122395>
- Marschner, H. (2011) *Marschner's mineral nutrition of higher plants*, 2nd ed.; Academic Press: Adelaide, Australia.
- Masclaux-Daubresse, C., Daniel-Vedele, F., Dechorgnat, J., Chardon, F., Gaufichon, L., & Suzuki, A. (2010). Nitrogen uptake, assimilation and remobilization in plants: challenges for sustainable and productive agriculture. *Annals of botany*, 105(7), 1141-1157. <https://doi.org/10.1093/aob/mcq028>
- Meier, U. (2001) Growth stages of mono- and dicotyledonous plants. Federal Biological Research Centre for Agriculture and Forestry. Blackwell Wissenschafts-Verlag.
- MicaSense. Retrieved April 15, 2021 from <https://micasense.com/about/>
- Mulla, D. J. (2013). Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems engineering*, 114(4), 358-371. <https://doi.org/10.1016/j.biosystemseng.2012.08.009>
- Mulvaney, R. L., Khan, S. A., & Ellsworth, T. R. (2006). Need for a soil-based approach in managing nitrogen fertilizers for profitable corn production. *Soil Science Society of America Journal*, 70(1), 172-182.
<https://doi.org/10.2136/sssaj2005.0034>
- Ontario Ministry of Agriculture, Food, and Rural Affairs. (2019). *Optimum planting dates for winter wheat in Ontario*.

<http://www.omafra.gov.on.ca/english/crops/field/news/croptalk/2019/ct-0919a4.htm>

- Palaniswamy, U. (2006) *Handbook of Statistics for Teaching and Research in Plant and Crop Science*. The Haworth Reference Press: New York, U.S.A.
- Pix4D. (n.d.) *How to improve the outputs of dense vegetation areas*. Retrieved March 2, 2021 from <https://support.pix4d.com/hc/en-us/articles/202560159-How-to-improve-the-outputs-of-dense-vegetation-areas>
- Planet (n.d.) *Planet imagery product specifications*. Retrieved July 2, 2021 from https://assets.planet.com/docs/Planet_Combined_Imagery_Product_Specs_letter_screen.pdf
- Rabalais, N. N., Turner, R. E., Wiseman Jr, W. J. (2001) Hypoxia in the Gulf of Mexico. *Journal of Environment Quality*, 30, 320-329. <https://doi.org/10.2134/jeq2001.302320x>
- Reisi Gahrouei, O., McNairn, H., Hosseini, M., & Homayouni, S. (2020). Estimation of crop biomass and leaf area index from multitemporal and multispectral imagery using machine learning approaches. *Canadian Journal of Remote Sensing*, 46(1), 84-99. <https://doi.org/10.1080/07038992.2020.1740584>
- Rondeaux, G., Steven, M., & Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote sensing of environment*, 55(2), 95-107. [https://doi.org/10.1016/0034-4257\(95\)00186-7](https://doi.org/10.1016/0034-4257(95)00186-7)
- Roujean, J. L., & Breon, F. M. (1995). Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote sensing of Environment*, 51(3), 375-384. [https://doi.org/10.1016/0034-4257\(94\)00114-3](https://doi.org/10.1016/0034-4257(94)00114-3)
- 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(1974), 309.
- RStudio Team. (n.d.) *RStudio: integrated development for R*. Retrieved March 2, 2021 from <http://www.rstudio.com/>
- Sadeh, Y., Zhu, X., Dunkerley, D., Walker, J. P., Zhang, Y., Rozenstein, O., Manivasagam V.S., & Chenu, K. (2021). Fusion of Sentinel-2 and PlanetScope time-series data into daily 3 m surface reflectance and wheat LAI monitoring. *International Journal of Applied Earth Observation and Geoinformation*, 96, 102260. <https://doi.org/10.1016/j.jag.2020.102260>
- Schirrmann, M., Giebel, A., Gleiniger, F., Pflanz, M., Lentschke, J., & Dammer, K. H. (2016). Monitoring agronomic parameters of winter wheat crops with low-cost UAV imagery. *Remote Sensing*, 8(9), 706. <https://doi.org/10.3390/rs8090706>
- 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. <https://doi.org/10.1016/j.jag.2013.04.003>

- Si, B., Farrell, R. (2004) Scale-dependent relationship between wheat yield and topographic indices. *Soil Science Society of America Journal*, 68, 577-587. <https://doi.org/10.2136/sssaj2004.5770>
- Sishodia, R. P., Ray, R. L., & Singh, S. K. (2020). Applications of remote sensing in precision agriculture: A review. *Remote Sensing*, 12(19), 3136. <https://doi.org/10.3390/rs12193136>
- Song, Y., Wang, J., & Shan, B. (2021). Estimation of Winter Wheat Yield from UAV-Based Multi-Temporal Imagery Using Crop Allometric Relationship and SAFY Model. *Drones*, 5(3), 78. <https://doi.org/10.3390/drones5030078>
- Song, Y., Wang, J., & Shang, J. (2020). Estimating effective leaf area index of winter wheat using simulated observation on unmanned aerial vehicle-based point cloud data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 2874-2887. <https://doi.org/10.1109/JSTARS.2020.2995577>
- Song, Y., & Wang, J. (2019). Winter wheat canopy height extraction from UAV-based point cloud data with a moving cuboid filter. *Remote Sensing*, 11(10), 1239. <https://doi.org/10.3390/rs11101239>
- Sripada, R. P., Schmidt, J. P., Dellinger, A. E., & Beegle, D. B. (2008). Evaluating multiple indices from a canopy reflectance sensor to estimate corn N requirements. *Agronomy Journal*, 100(6), 1553-1561. <https://doi.org/10.2134/agronj2008.0017>
- Tremblay, N., Bouroubi, Y. M., Bélec, C., Mullen, R. W., Kitchen, N. R., Thomason, W. E., & Ortiz-Monasterio, I. (2012). Corn response to nitrogen is influenced by soil texture and weather. *Agronomy Journal*, 104(6), 1658-1671. <https://doi.org/10.2134/agronj2012.0184>
- Vetsch, J. A., & Randall, G. W. (2004). Corn production as affected by nitrogen application timing and tillage. *Agronomy Journal*, 96(2), 502-509. <https://doi.org/10.2134/agronj2004.5020>
- Walley, F., Yates, T., Van Groenigen, J.W., Kessel, C. (2002) Relationships between soil nitrogen availability indices, yield, and nitrogen accumulation of wheat. *Soil Science Society of America Journal*, 66, 1549-1561. <https://doi.org/10.2136/sssaj2002.1549>
- 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. [https://doi.org/10.1016/S1672-6308\(07\)60027-4](https://doi.org/10.1016/S1672-6308(07)60027-4)
- Wei, F., Yan, Z., Yongchao, T., Weixing, C., Xia, Y., & Yingxue, L. (2008). Monitoring leaf nitrogen accumulation in wheat with hyper-spectral remote sensing. *Acta Ecologica Sinica*, 28(1), 23-32. [https://doi.org/10.1016/S1872-2032\(08\)60018-9](https://doi.org/10.1016/S1872-2032(08)60018-9)
- Wu, M., Yang, C., Song, X., Hoffmann, W. C., Huang, W., Niu, Z., Wang, C., & Li, W. (2017). Evaluation of orthomosaics and digital surface models derived from aerial imagery for crop type mapping. *Remote Sensing*, 9(3), 239. <https://doi.org/10.3390/rs9030239>

- Yu, J., Wang, J., & Leblon, B. (2021). Evaluation of Soil Properties, Topographic Metrics, Plant Height, and Unmanned Aerial Vehicle Multispectral Imagery Using Machine Learning Methods to Estimate Canopy Nitrogen Weight in Corn. *Remote Sensing*, *13*(16), 3105. <https://doi.org/10.3390/rs13163105>
- 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. <https://doi.org/10.3390/rs12020215>
- Zhao, B., Yao, X., Tian, Y., Liu, X., Ata-Ul-Karim, S. T., Ni, J., Cao, W., & Zhu, Y. (2014). New critical nitrogen curve based on leaf area index for winter wheat. *Agronomy Journal*, *106*(2), 379-389. <https://doi.org/10.2134/agronj2013.0213>
- Zheng, H., Li, W., Jiang, J., Liu, Y., Cheng, T., Tian, Y., Zhu, Y., Cao, W., Zhang, Y., & Yao, X. (2018). A comparative assessment of different modeling algorithms for estimating leaf nitrogen content in winter wheat using multispectral images from an unmanned aerial vehicle. *Remote Sensing*, *10*(12), 2026. <https://doi.org/10.3390/rs10122026>
- Zhu, Y., Yao, X., Tian, Y., Liu, X., & Cao, W. (2008). Analysis of common canopy vegetation indices for indicating leaf nitrogen accumulations in wheat and rice. *International Journal of Applied Earth Observation and Geoinformation*, *10*(1), 1-10. <https://doi.org/10.1016/j.jag.2007.02.006>

Chapter 4

4 Conclusion

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

4.1 Summary

Precision agriculture is an important management technique and field of research to increase productivity, minimize environmental impacts and the cost of production. With the continuous advancement of remote sensing technologies, processing methods, and computing capabilities, new estimation models for crop nitrogen can be developed. It is well understood that the growth of a plant is affected by many factors such as nutrient supply, plant physiology, soil conditions, field topology, and water availability. Using the remotely sensed spectral properties, plant biophysical parameters, and environmental variables in canopy nitrogen estimation models can lead to more effective fertilizer management methods.

In Chapter 2, RF and SVM regression methods were used to predict canopy nitrogen weight of corn. Model variables included UAV MicaSense individual band reflectances, associated VIs, plant height, topographic metrics, and soil metrics. Single-date and multi-date data were tested, with the best model's variable importance plot used to identify groups of top variables related to corn canopy nitrogen weight. Further model development allowed for evaluating the benefit of using both spectral and non-spectral variables together, which types of variables are most important, and how they can be applied to precision agriculture methods feasibly.

In Chapter 3, RF and SVM regression methods were used to predict canopy nitrogen weight of wheat. Variables in the models included UAV Micasense band reflectances, PlanetScope band reflectances, selected VIs, plant height, LAI, soil moisture, and topographic metrics. The models created were grouped by UAV-based and satellite-based data. Combinations of multi-date data and single date data were used for model testing. The best performing UAV-based and satellite-based models' variable importance plots were used to select

groups of top variables related to wheat canopy nitrogen weight. Select spectral and non-spectral variable groups were evaluated, and the most important variables each for UAV-based and satellite-based models were determined. Nitrogen prediction maps were then created for each wheat field constructed solely from remotely sensed data.

4.2 Conclusions

The research objectives for this thesis were completed, and responses to each are as follows:

- (i) In both Chapters 2 and 3 for corn and wheat respectively, machine learning RF and SVM regression models were used to predict canopy nitrogen weight. First by testing single-date and multi-date data in RF and SVM models, the best temporal periods for nitrogen estimation were determined for each crop. In corn, the combination of three dates (8, 15, 24 June) had the best model performance of R^2 0.75 and RMSE 2.29 g/m² around BBCH 10-31. In wheat, both UAV-based and satellite-based best performing models were from the combination of three dates (12, 20, 27 May) at R^2 0.74, RMSE 2.76 g/m² and R^2 0.83, RMSE 1.77 g/m² respectively around BBCH 23-41. Using the variable importance plots from these multi-date data models, groups of spectral-only and both spectral/non-spectral variables together were tested in RF and SVM regression models.
- (ii) In Chapter 2 estimating canopy nitrogen weight of corn, the RF model using top 15 variables performed best with R^2 of 0.73 and RMSE of 2.21 g/m². Of the 15 variables, plant height and profile curvature were the only non-spectral, then spectral included 10 VIs and three MicaSense bands (blue, red, green). In Chapter 3 estimating canopy nitrogen weight of wheat, models were created separately from UAV-based and satellite-based data. For UAV-based, the group of top 7 variable SVR model performing the best with R^2 of 0.80 and RMSE of 2.62 g/m². Top 7 variables included plant height, LAI, all three red-edge VIs, BNDVI, and MicaSense NIR band. For satellite-based, the RF model group of top 17 variables had the best performance with R^2 of 0.92 and RMSE of 1.75 g/m². Height, LAI, all four PlanetScope bands, and total 11 VIs were

variables in the best performing model. For both UAV and satellite best performing selected variable models, plant height and LAI are the only non-spectral variables.

In the best performing models for predicting nitrogen weight in corn and wheat fields, all included non-spectral variables. The spectral-only models had significantly poorer performance with low R^2 values and high RMSE for both crop types in both RF and SVR methods. Of the top models' non-spectral variables for both corn and wheat, plant height was the most important predictor.

- (iii) For both Chapter 2 and Chapter 3, all the non-spectral variables in the top models (plant height, LAI, and topographic metrics) for both corn and wheat can be derived from in-situ UAV data, meaning practical application of the models is feasible. Algorithms for creating accurate maps of corn crop height still need improvement, thus nitrogen prediction maps were not created in Chapter 2. The difficulties in generating height maps of corn include the nature of the plant structure where the highest point is the tassel composed of narrow branches, layers of leaves are narrow and angled which could not be captured in detail with the UAVs used in this study. Creating accurate height maps for wheat is simpler as the plant structure and planting density create large, uniform areas of height. In Chapter 3, the UAV-based and satellite-based nitrogen prediction maps for wheat were created with 1 m x 1 m spatial resolution and able to differentiate areas of the field with highs and lows of canopy nitrogen weight. The low RMSE (g/m^2) of the top-performing models are useful to farmers for practical fertilizer management recommendations.

4.3 Limitations & Future Work

With in-situ fieldwork collection come many challenges and limitations. Sample point distribution in fields is dependent on many factors: the number of people in the fieldwork team, available equipment, size of field, intensity of labour, time constraints, weather, and associated operation costs. For both wheat and corn fields' sample point groups were selected with consideration of the field characteristics and feasibility of data collection, but ideally points would be placed across a whole field for full representation of conditions.

Other limitations involved the UAV data, in which a fieldwork week had consistently poor weather conditions including heavy rain, thunderstorms, and strong winds unsuitable for UAV flights. There was difficulty finding a day suitable for both fieldwork data collection and a UAV flight, and when conducted several days apart may affect results as plants reach new growth stages and field conditions (e.g., soil moisture) change. Although in Chapter 2 corn field soil metrics had little to no effect on the models and was subsequently omitted as variables in Chapter 3 wheat models, the soil for each field was only tested once at the beginning of the growing season. Considering the costs and typical recommendations for soil tests are only once a year, there is the limitation of soil testing practices. Future studies can consider more conducting soil tests throughout the growing season to improve evaluations of soil metrics in nitrogen prediction models, as it can be affected by fertilizer application, precipitation patterns, and crop growth.

As for the use of satellite-based spectral data in future model development, PlanetScope launched a new generation of sensors called SuperDove with five spectral bands: blue, green, red, NIR, and red-edge. As imagery from PlanetScope SuperDoves becomes available, future work can test its red-edge VIs in canopy nitrogen prediction models to evaluate the importance of a satellite-based red-edge band. For UAV-based model development, future work with hyperspectral sensors can explore more of the electromagnetic spectrum to improve prediction accuracies.

This thesis used machine learning Random Forest and Support Vector regression methods. Machine learning does have limitations in that models usually require very large training datasets in order to be effective, and computational costs of collecting, storing, and

processing power needed to handle big datasets. It should be noted that the regression analyses performed in this thesis are not explicitly spatial analyses. Spatial data was used in analyses, but there were assumptions of independence and homogeneity for the variables used. Explicit spatial analyses take into consideration spatial autocorrelation and heterogeneity, where relationships in data vary in space. For example, sample points which are closer to one another are likely to be more related than points further away. Different fields can have dissimilar conditions that may not have been fully captured in the methods used. Further work can try implementing spatial statistics, such as nearest neighbor analysis or Thiessen polygons, to discern nuances in data relationships for crop monitoring.

From the results of this thesis, the use of UAV and satellite spectral imagery in conjunction with non-spectral plant biophysical variables and field properties in wheat and corn crop nitrogen estimation models has promising results. The developed models need to be tested on other datasets to determine their efficacy, understand their applicability and feasibility in precision agriculture.

5 Appendices

5.1 Appendix A – Field Photos



Figure A 1. Example of ground control point marker.



Figure A 2. Piloting the unmanned aerial vehicles. The tall device on the left is the real time kinematic (RTK) navigation base station.



Figure A 3. Sensors attached to the underbody of the DJI Matrice 100 UAV. The RGB camera is on the top, and the red MicaSense RedEdge camera with five sensor bands is on the bottom.



Figure A 4. Photos of corn crop taken on 8 June 2020.

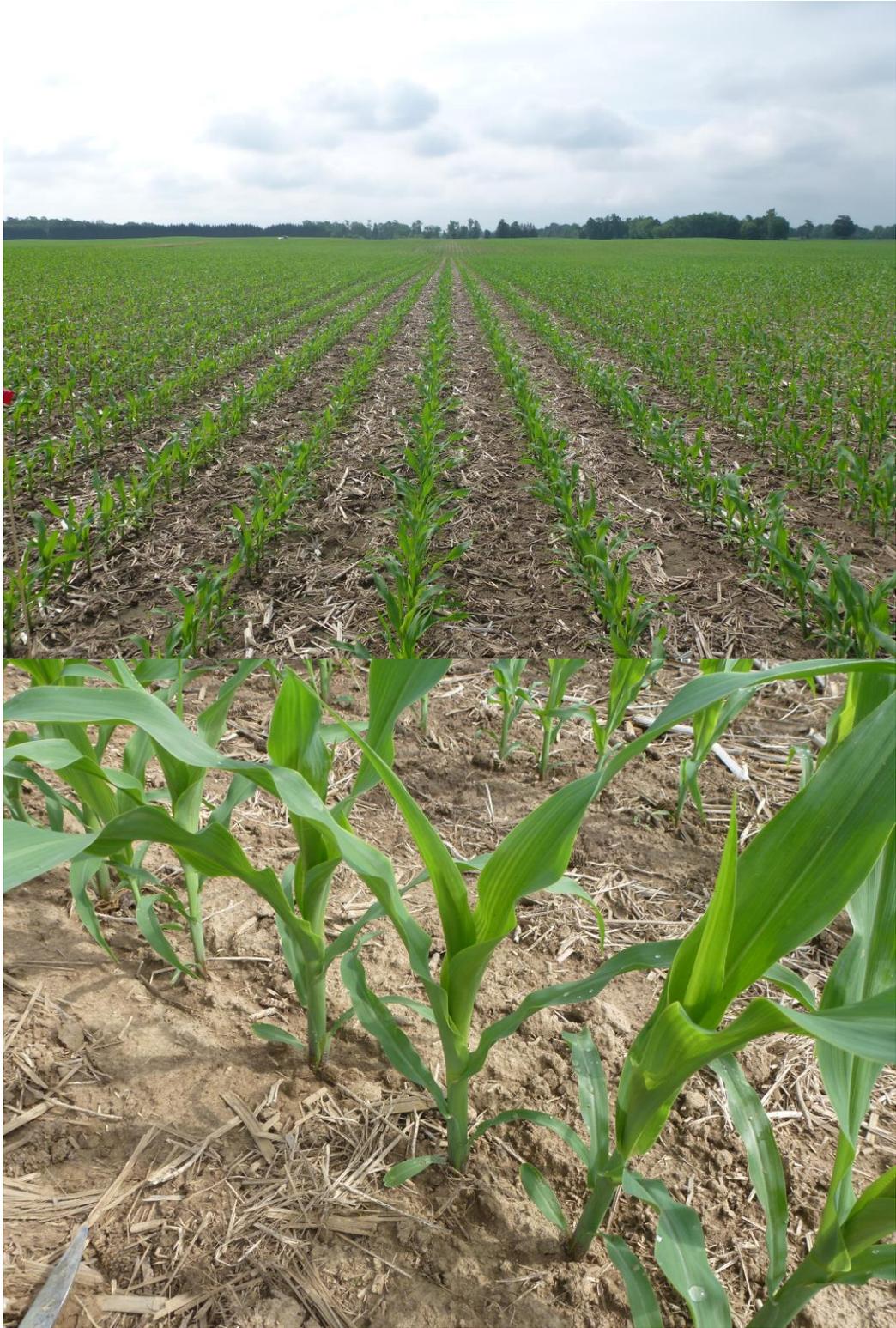


Figure A 5. Photos of corn crop taken on 15 June 2020.



Figure A 6. Photos of corn crop taken on 24 June 2020.

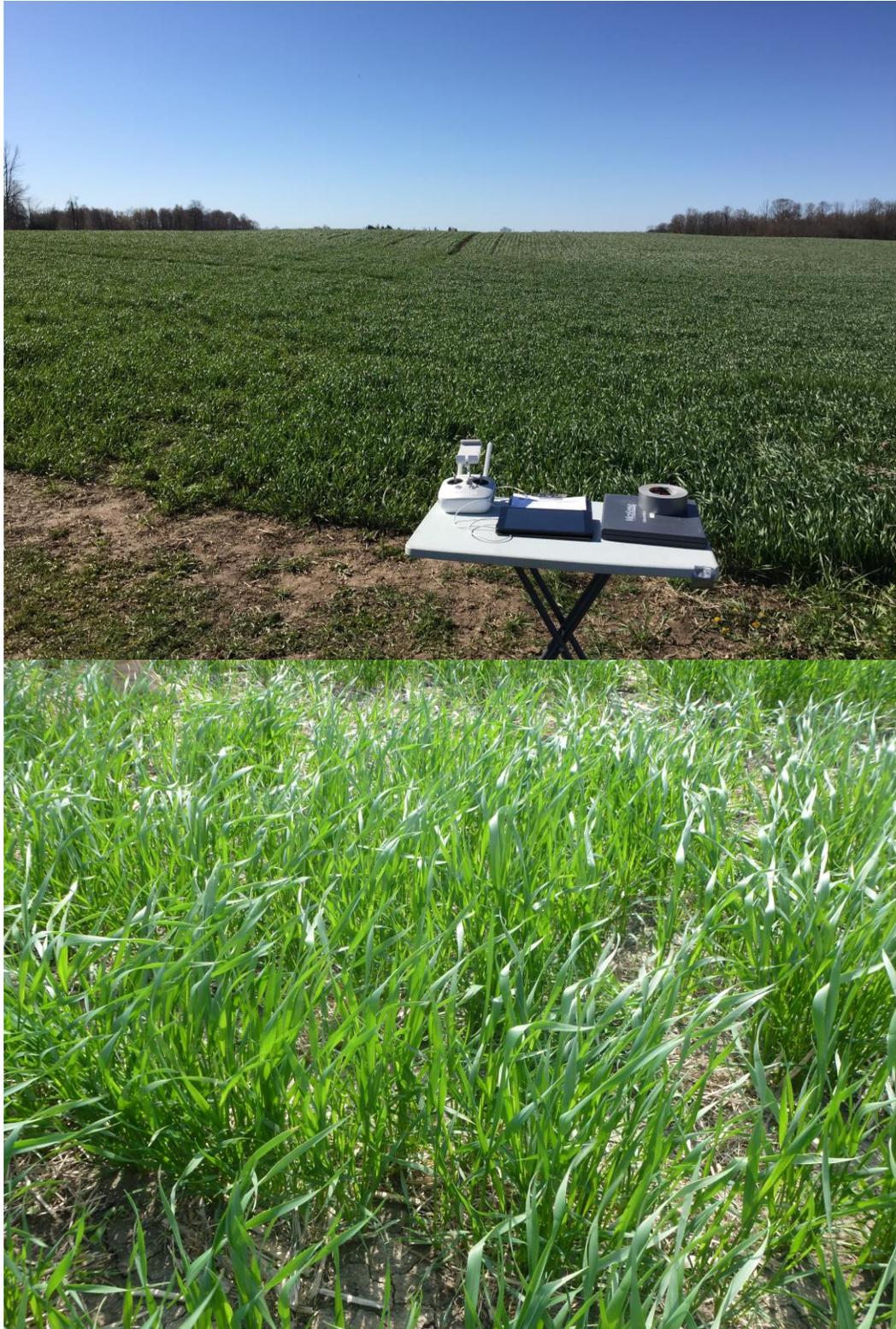


Figure A 7. Photos of wheat crop taken on 12 May 2020.



Figure A 8. Photos of wheat crop taken on 20 May 2020.



Figure A 9. Photos of wheat crop taken on 27 May 2020.

5.2 Appendix B – Pix4D Processing



Figure B 1. Example of Pix4Dmapper flight pattern map view. Each red circle represents the UAV's location with the sensor taking a picture. The diagonal flight patterns are due to the UAV returning to the starting point for a battery change midway through and returning upon completion of the specified flight area.

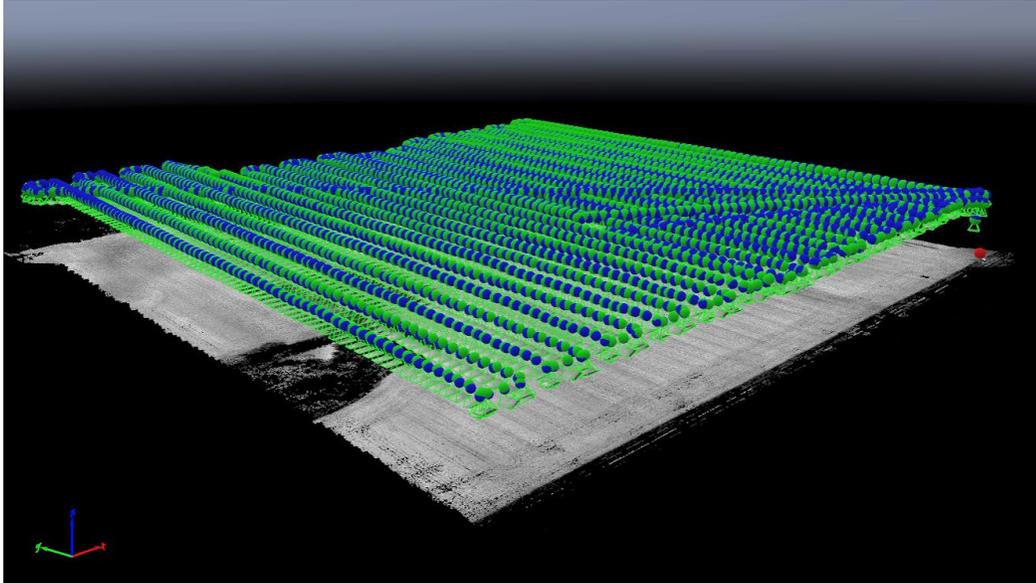


Figure B 2. Example of the Pix4Dmapper rayCloud view of a field. Each blue-green dot is a representation of a UAV image tied to metadata (e.g., camera positions, geographic coordinate). The black and white layer beneath is an orthomosaic of all images taken.

5.3 Appendix C – Topographic Metrics Processing

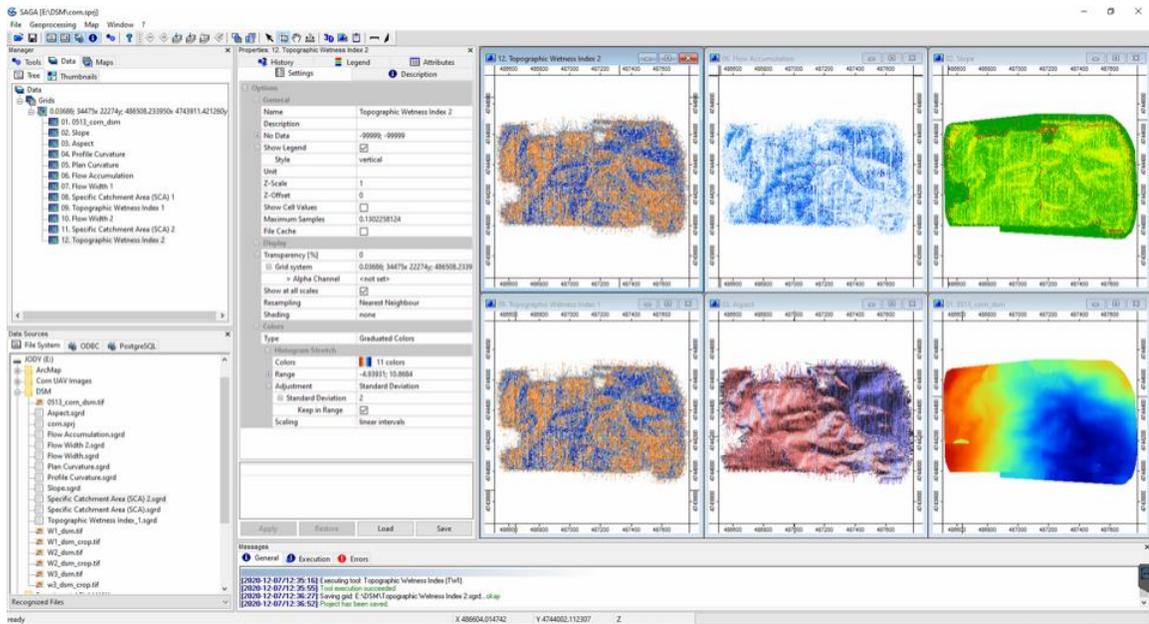


Figure C 1. Example of System of Automated Geoscientific Analyses (SAGA) interface and topographic metric processing results.

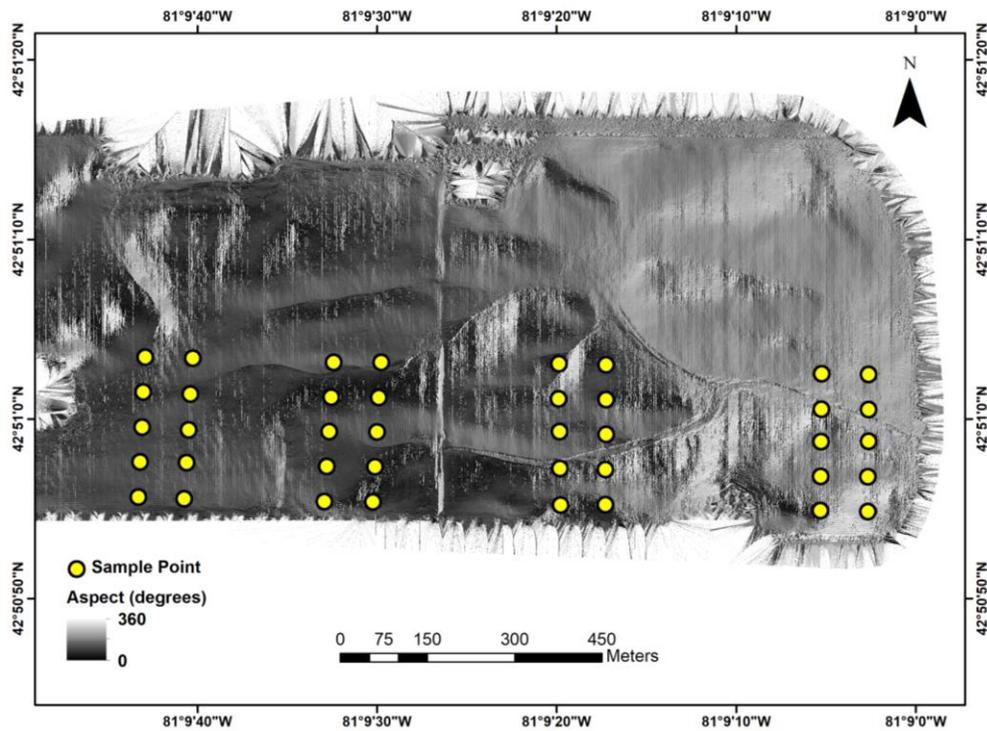


Figure C 2. Example of the corn field aspect layer created in SAGA.

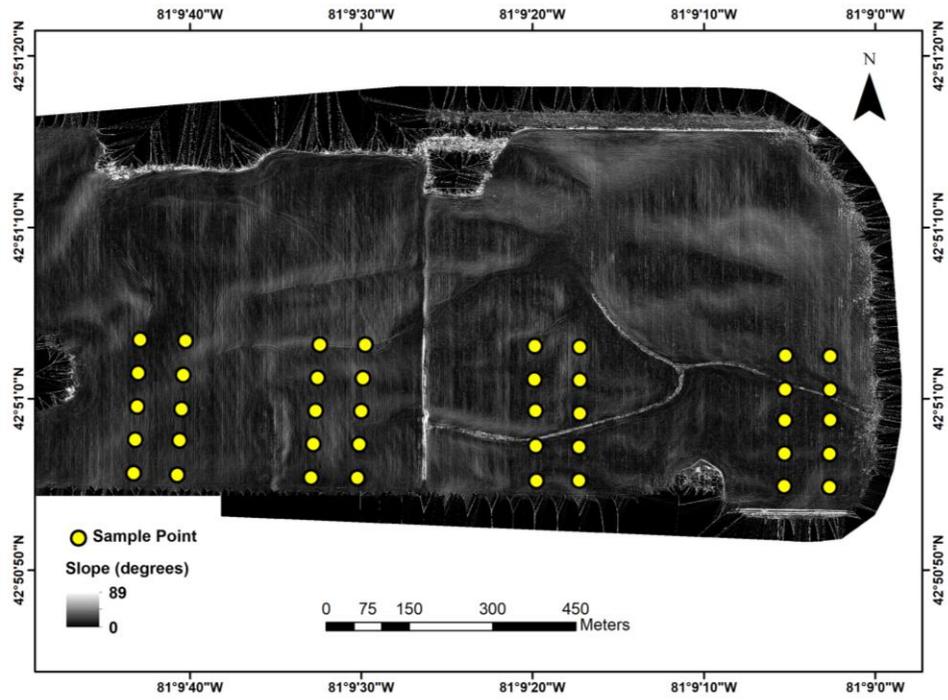


Figure C 3. Example of the corn field slope map created in SAGA.

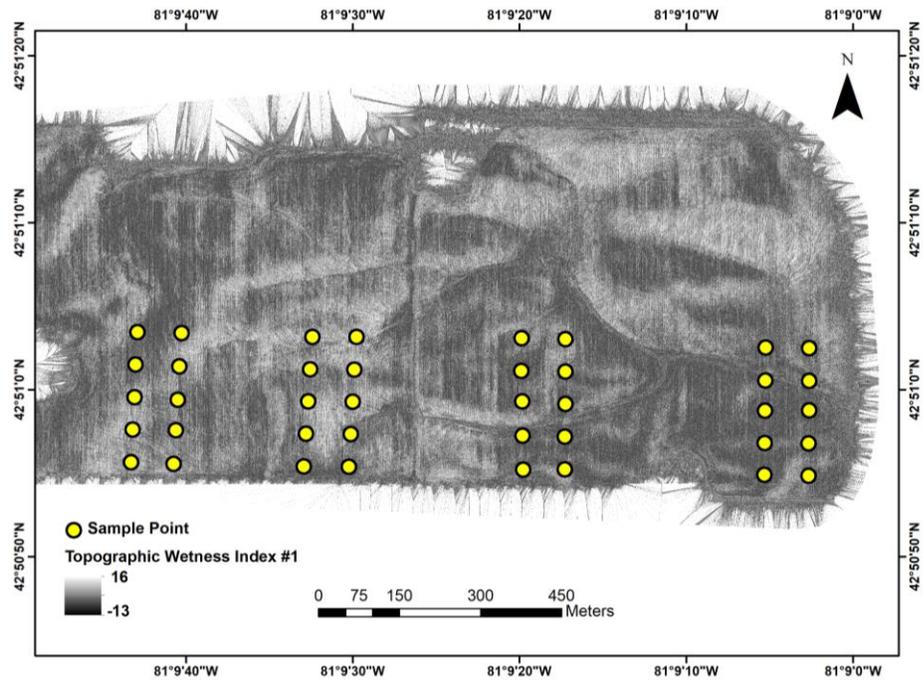


Figure C 4. Example of the corn field topographic wetness index map created in SAGA.

5.4 Appendix D – LAI Layers

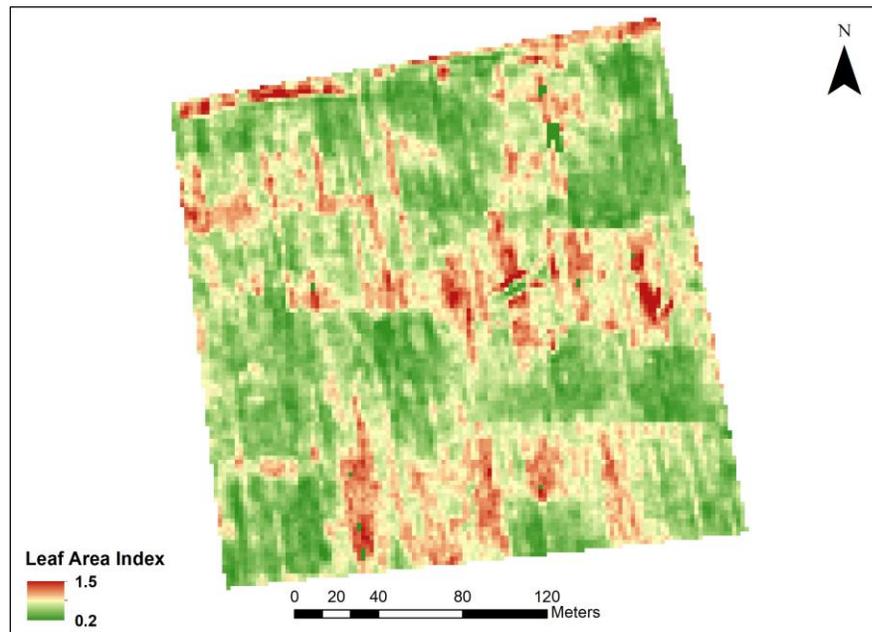


Figure D 1. Example of LAI layer for wheat field W2 on 16 May 2020.

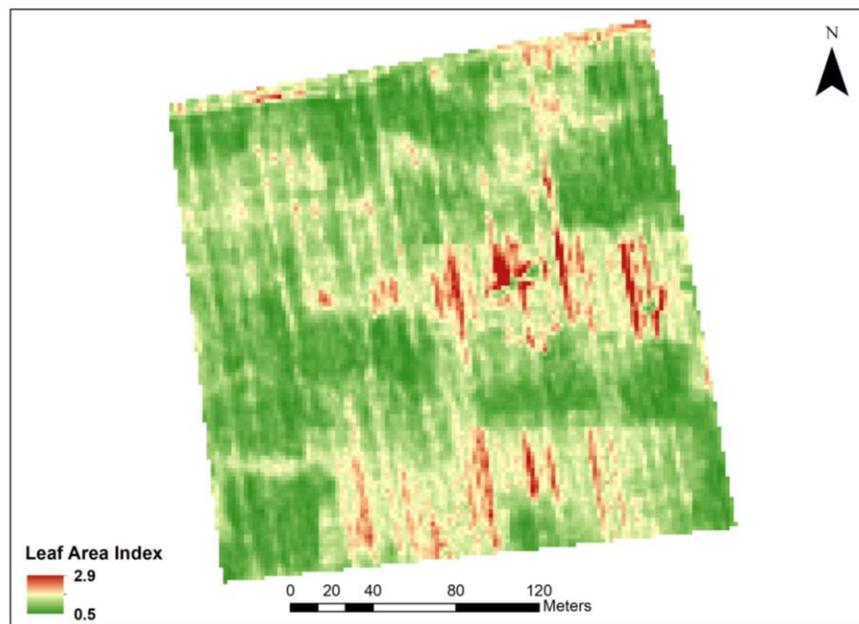


Figure D 2. Example of LAI layer for wheat field W2 on 21 May 2020.

5.5 Appendix E – R Code

```

#Install “randomForest” and “e1071” packages for Random Forest (RF) and Support
Vector Machine (SVM) before loading
#Load library packages
library (randomForest); library (e1071)

#Load data from Excel file into a variable and view to ensure it is correct
library (readxl)
data <- read_excel(“D:/data.xlsx”); View(data)

#Split dataset for models: 70% data for training and 30% for validation
#Set the starting number used to generate a sequence of random numbers
set.seed(123); train <- sample(nrow(data), 0.7*nrow(data), replace = F)
trainset <- data[train,]
validset <- data[-train,]

#Run RF model of training set, view results
RFmodT <- randomForest(N_weight ~., data = trainset); RFmodT

#Apply the model to both training and validation datasets
predictT <- predict(RFmodT, newdata = trainset)
predictV <- predict(RFmodT, newdata = validset)

#Calculate RMSE of the models, look at variable description for result
rmsemodT <- sqrt(mean((predict – trainset$N_weight)^2))
rmsemodV <- sqrt(mean((predict – trainset$N_weight)^2))

#Calculate R-squared of the models, view results
RsmodT <- lm(trainset$N_weight – predict, data = trainset); summary(RsmodT)
RsmodV <- lm(validset$N_weight – predict, data = validset); summary(RsmodV)

#Repeat steps above for SVM, starting with running model for training set
SVRmodT <- svm(N_weight ~., data = trainset); SVRmodT

```

Figure E 1. R code of Random Forest and Support Vector regression models.

```
#After creating regression models and selecting the best model, create a model
prediction image

#Install “raster”, “rgdal”, and “rasterVis” libraries before loading
library (raster); library (rgdal), library (rasterVis)

#Set up a folder containing the “.tif” images to use in the model
#Ensure the image names match the headers in your data for later steps
#The “list.files” function will read through the folder to look for files ending with “.tif”
images <- list.files(path = “D:/thesis_images”, pattern = “tif$”, full.names = T)

#The “raster::stack” function will stack the list of image files
#Ensure the images have the same resolution and extent for the function to work
imageStack <- raster::stack(images)

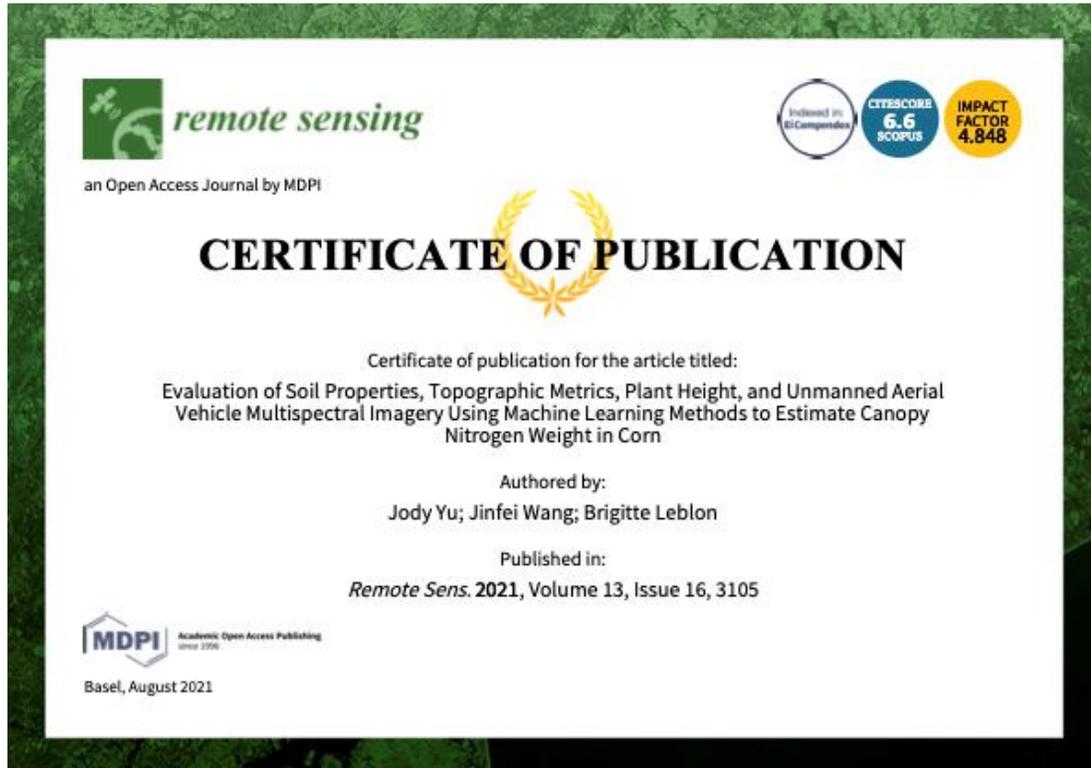
#The “raster::predict” function calls the image stack dataframe to predict values for
each pixel based on the model selected (see code for making regression models)
imagePredict <- raster::predict(object = imageStack, model = RFmodT, fun = predict)

#Create the final prediction raster using the “writeRaster” function
#Set the desired output location and file name; set file type to “.tif”
writeRaster(imagePredict, “D:/thesis_images/W1_predictN.tif”)
```

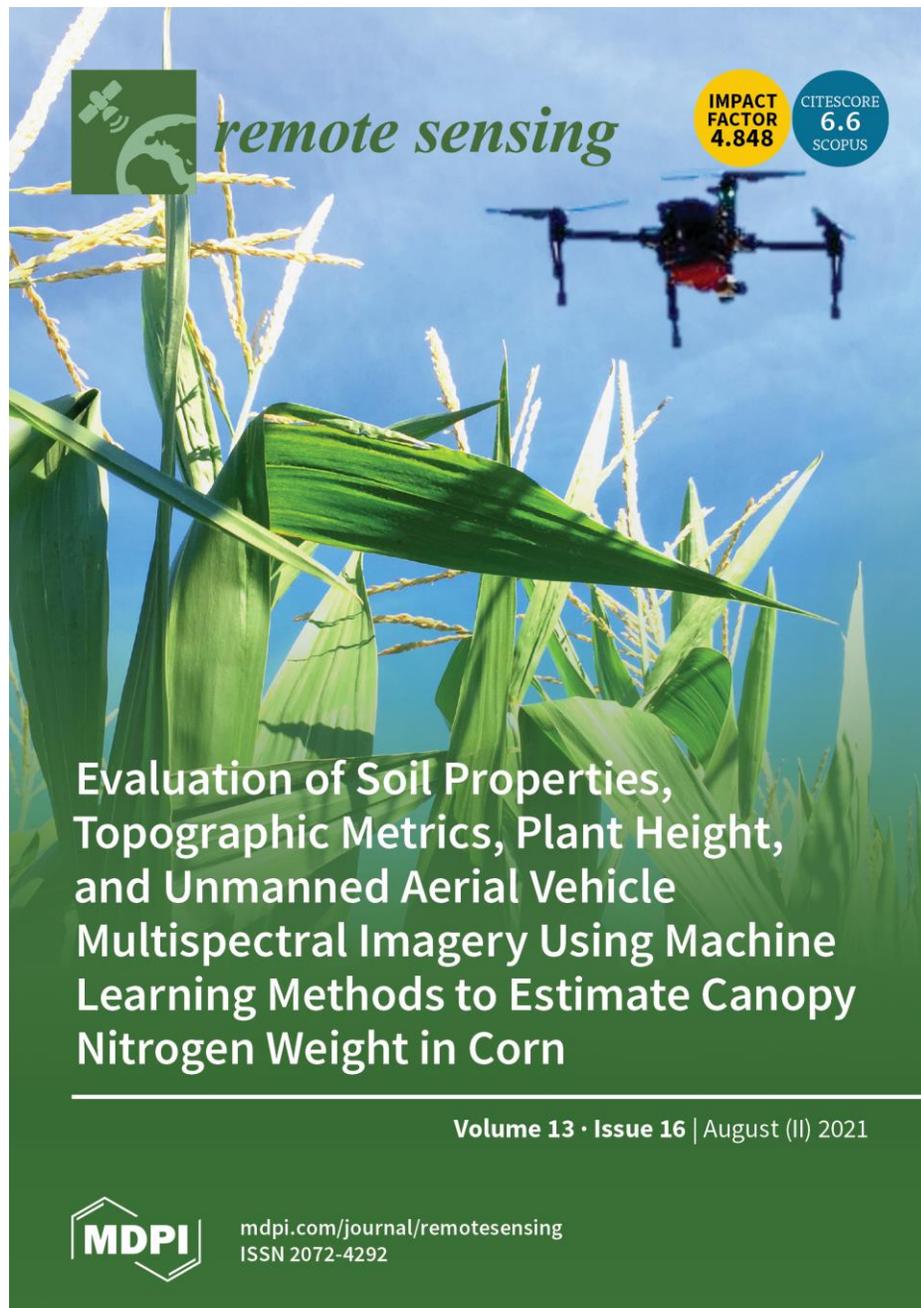
Figure E 2. R code for building a prediction image from multi-layered raster images.

5.6 Appendix F - Copyrighted Material & Permissions

Chapter 2, published under *MDPI Remote Sensing* requires no special permission to reuse all or part of an article under an open access Creative Common CC BY license, provided that the original article is clearly cited.



Our published paper was selected as the feature cover story in *MDPI Remote Sensing* Vol. 13 Issue 16. Available online at: <https://www.mdpi.com/2072-4292/13>



Chapter 3, published under *MDPI Nitrogen* requires no special permission to reuse all or part of an article under an open access Creative Common CC BY license, provided that the original article is clearly cited.



Curriculum Vitae

Name: Jody Yu

Post-secondary Education and Degrees: University of Western Ontario
London, Ontario, Canada
2015-2019 B.Sc. (Hons) Geographic Information Science

The University of Western Ontario
London, Ontario, Canada
2019-2021 M.Sc. Geography

Honours and Awards: Ontario Graduate Scholarship
2020

Western Gold Medal, BSc (Hons) GIS
2019

Canadian Association of Geographers Award, Top Thesis
2019

NSERC Undergraduate Student Research Awards
2018, 2019

Related Work Experience: Teaching Assistant
University of Western Ontario
2019-2021

UAV Pilot, Field & Lab Technician
A&L Canada Laboratories
2021

Research Assistant
GITA Lab, University of Western Ontario
2018-2021

Publications:

Yu, J., Wang, J., & Leblon, B. (2021). Evaluation of Soil Properties, Topographic Metrics, Plant Height, and Unmanned Aerial Vehicle Multispectral Imagery Using Machine Learning Methods to Estimate Canopy Nitrogen Weight in Corn. *Remote Sensing*, 13(16), 3105. <https://doi.org/10.3390/rs13163105>

Yu, J., Wang, J., Leblon, B., & Song, Y. (2022). Nitrogen Estimation for Wheat Using UAV-Based and Satellite Multispectral Imagery, Topographic Metrics, Leaf Area Index, Plant Height, Soil Moisture, and Machine Learning Methods. *Nitrogen*, 3(1), 1-25. <https://doi.org/10.3390/nitrogen3010001>