

Northumbria Research Link

Citation: Ahmed, Aishath Shara (2023) Use of unmanned aerial vehicle for environmental monitoring purposes and precision agriculture. Doctoral thesis, Northumbria University.

This version was downloaded from Northumbria Research Link:
<https://nrl.northumbria.ac.uk/id/eprint/51687/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

**Use of unmanned aerial vehicle for
environmental monitoring purposes and
precision agriculture**

A.S. Ahmed

PhD

2023

**Use of unmanned aerial vehicle for
environmental monitoring purposes and
precision agriculture**

Aishath Shara Ahmed

A thesis submitted in partial fulfilment of the
requirements of the University of Northumbria
at Newcastle for the degree of Doctor of
Philosophy

Research undertaken in the Faculty of Health
and Life Sciences

March 2023

Abstract

Despite numerous research studies on remote sensing applications in forestry and precision agriculture, there is a limited availability of image analysis techniques that are less complex, reproducible, and applicable to diverse locations and under a wide range of environmental conditions. Image analysis techniques currently in use employ complex machine learning approaches (regression-based models), for example, to identify tree species in forestry and estimate crop yield in precision agriculture. However, many challenges must be overcome before these modern machine learning approaches can potentially see widespread commercial and non-commercial implementation in agriculture and forestry. As a result, there is a need to investigate and develop simple, dependable, and reproducible image analysis methods by utilising remote sensing data applicable in forestry and precision agriculture.

Hence, the current study focuses on using a remote sensing platform of multispectral unmanned aerial vehicle (UAV) to monitor native and invasive tree species in an ancient semi-natural woodland and investigate the performance of a variety of crops for precision agriculture, including oilseed rape, winter beans, and winter oats. The multispectral UAV data were analysed using simple yet effective image analysis techniques such as principal component analysis (PCA), spectral vegetation indices combined with image classification methods of thresholding and clustering (k-means and iso-cluster). Also, the image analysis methods were performed with effective data manipulation software such as MATLAB and ArcGIS.

Identification and quantification of native and invasive tree species was achieved by PCA derived spectral vegetation indices, thresholding and k-means clustering. Additionally, the use of spectral vegetation indices and iso-cluster classification method in precision agriculture of crops assisted in estimation of crop yield three months before harvest. Also, strong correlation was observed between the estimate and actual crop yield. Furthermore, a pilot study using a multinomial logistic regression model with high sensitivity and accuracy enabled the identification of soil nutrient concentration and crop quality features for very high oats yield.

The simple and effective image analysis methods on multispectral UAV data for forestry and precision agriculture must be employed more frequently than complex machine learning approaches. Also, the estimated crop yield prior to harvesting aids farmers for precision agriculture of crops to maintain its performance. Whereas, in forestry these methods can be employed frequently to monitor the native tree species and emergence of new invasive tree species and remove them effectively to maintain a sustainable ecosystem.

Publications

Ahmed, S., Nicholson, C. E., Muto, P., Perry, J. J., & Dean, J. R. (2021). Applied aerial spectroscopy: A case study on remote sensing of an ancient and semi-natural woodland. *Plos one*, 16(11), e0260056.

Ahmed, S., Nicholson, C. E., Muto, P., Perry, J. J., & Dean, J. R. (2021). The use of an unmanned aerial vehicle for tree phenotyping studies. *Separations*, 8(9), 160.

Conference Presentations

Oral Presentation - Analytical Research Forum 2021 (15-16 June 2021 - Online event)

- Applied aerial spectroscopy- a case study using an unmanned aerial vehicle to assess our environment.

Poster Presentation - Analytical Research Forum 2022 (13 – 14 June 2022 – London UK)

- Use of unmanned aerial vehicle with a multi-spectral image camera for high throughput phenotyping in precision agriculture.

Table of Contents

Abstract	iii
Publications	iv
Conference Presentations	iv
Acknowledgement	10
Declaration	11
Chapter 1: History of aerial imaging by remote sensing platforms	12
1.1 Aerial imaging	12
1.2 History of aerial imaging using balloon platforms	12
1.3 History of aerial imaging using kite platforms	13
1.4 History of aerial imaging using manned aircraft.....	14
1.5 History of aerial imaging using satellites.....	14
1.6 History of aerial imaging using UAV platforms.....	16
1.7 Types of UAV.....	18
Chapter 2: UAV sensors and multispectral imaging	21
2.1 UAV sensors	21
2.2 RGB UAV sensors	22
2.3 Thermal UAV sensors.....	23
2.4 LiDAR UAV sensors	23
2.5 Hyperspectral UAV sensors.....	23
2.6 Multispectral UAV sensors.....	24
2.7 Concept of multispectral imaging (MSI)	24
2.8 Types of multispectral sensors for MSI.....	26
Chapter 3: Image processing by multispectral UAV	29
3.1 Image pre-processing	29
3.2 Feature extraction.....	30
3.2.1 Spectral features.....	30
3.2.2 Structural and texture features	30
3.3 Machine learning models	31
3.4 Research gap.....	32
3.5 Aims and Objectives	34
SECTION A	35
MULTISPECTRAL UAV IN FORESTRY	35
Chapter 4: Multispectral UAV monitoring of ancient and semi-natural woodland	36
4.1 Introduction.....	36

4.1.1 History of ancient and semi-natural woodland in UK	36
4.1.2 Significance of monitoring ASNW species	36
4.1.3 Methods to monitor native and invasive species in ASNW.....	37
4.1.4 Literature on multispectral UAV monitoring of native and invasive tree species	37
4.2 Research gap	38
Chapter 5: Multispectral UAV monitoring of native tree species in the central boundary of Priestclose ASNW	41
5.1 Aims and Objectives	41
5.1.1 Aims.....	41
5.1.2 Objectives	41
5.2 Experimental Design	41
5.2.1 Sampling site.....	41
5.2.2 Data collection by UAV.....	42
5.2.3 Field Data Collection	43
5.2.4 UAV Photogrammetric Processing.....	45
5.2.5 Image processing and Data Analysis	45
5.2.6 Normalised Difference Vegetation Index (NDVI).....	46
5.2.7 Mathematical background of PCA in multispectral images	46
5.2.8 PCA applied to data on Priestclose Wood	48
5.2.9 Calculation of pixels of native tree species and dead trees	54
5.3 Results and Discussion	54
5.3.1 Comparing the performance of classification methods by PCA	54
5.3.2 Quantitative information obtained by analysis of UAV MSI and field study data	59
5.4 Conclusion	62
Chapter 6: Multispectral UAV monitoring of invasive tree species in the southern boundary of Priestclose ASNW	63
6.1 Aims and Objective.....	63
6.2 Experimental Design	63
6.2.1 Sampling site.....	63
6.2.2 Data collection by UAV.....	64
6.2.3 UAV data analysis: Photogrammetric Processing	64
6.2.4 Field Data Analysis.....	65
6.2.5 Image processing and Data Analysis	65
6.2.6 Principal Component Analysis applied to obtained data.	65
6.2.7 Principle of K-means segmentation in image data.....	71
6.2.8 K-means segmentation and quantification of invasive tree species	73
6.3 Results and Discussion	73

6.3.1 Assessing the effectiveness of PCA and k-means segmentation in invasive tree species classification.	73
6.3.2 Quantitative information obtained by analysis of UAV MSI and field study data.	74
6.4 Conclusion	79
SECTION B	80
MULTISPECTRAL UAV IN PRECISION AGRICULTURE	80
Chapter 7: Multispectral UAV in precision agriculture of crops	81
7. 1 Introduction.....	81
7.1.1 Precision agriculture: Its Principles and Significance.....	81
7.1.2 Significant application of UAV sensing in precision agriculture	82
7.1.3 Crop varieties essential for precision agriculture in the UK	86
7.1.4 Literature review on precision agriculture of oilseed rape by multispectral-UAV	87
7.1.5 Literature review on precision agriculture of winter beans by multispectral-UAV.....	89
7.1.6 Literature review on precision agriculture of winter oats by multispectral-UAV	90
7.2 Research gap	92
7.3 Summary	94
Chapter 8: Multispectral UAV monitoring prediction of oilseed rape crop performance	95
8.1 Aims and Objective.....	95
8.2 Experimental Design	95
8.2.1 Sampling site Airy Holme Farm	95
8.2.2 Waskerley Edge agricultural field.....	96
8.2.3 Crop harvest	97
8.2.4 Unmanned Aerial Vehicle.....	97
8.2.5 UAV Photogrammetric Processing	98
8.2.6 Crop Height Estimation	99
8.2.7 Vegetation Index Calculation and Iso-cluster Classification	100
8.2.8 Oilseed rape seed yield estimation.....	101
8.2.9 Statistical Analysis.....	104
8.3 Results and Discussion	104
8.3.1 Estimation of OSR plant height by multispectral-UAV at the various phenological growth stages.....	104
8.3.2 Correlation study between estimated and actual seed yields	106
8.4 Conclusion	107
Chapter 9: Multispectral UAV monitoring prediction of winter bean crop performance	108
9.1 Aims and Objectives	108
9.2 Experimental Design	108
9.2.1 Sampling site Airy Holme Farm	108
9.2.2 High Dowell Agriculture field	109

9.2.3 Unmanned Aerial Vehicle.....	110
9.2.4 UAV Photogrammetric Processing.....	110
9.2.5 Crop height Estimation	111
9.2.6 Vegetation Index Calculation and Iso-cluster Classification	112
9.2.7 Winter bean seed yield estimation	113
9.2.8 Statistical Analysis.....	114
9.3 Results and Discussion	115
9.3.1 Estimation of winter bean development across the phenological growth stages.	115
9.3.2 Prediction of seed yield.....	117
9.4 Conclusion	122
Chapter 10: Multispectral UAV monitoring prediction of winter oats crop performance.....	123
10.1 Aims and Objectives	123
10.2 Experimental Design	124
10.2.1 Airy Holme Farm.....	124
10.2.2 Copse agricultural field.....	125
10.2.3 Unmanned Aerial Vehicle.....	125
10.2.4 Collection of crop phenotypic data: ground reference data	126
10.2.5 Comparison between UAV and ground reference data.....	127
10.2.6 Assessment of chemical soil health indicators.....	128
10.2.6.1 Soil sample collection	128
10.2.6.2 Analysis of NPK and pH by Palintest kit.....	129
10.2.7 Collection of crop yield and height data	130
10.2.7.1 Manual oat grain yield measurement	130
10.2.7.2 Oat β -glucan extraction.....	131
10.2.7.3 Analysis of glucose in oat β glucan by colorimetric assay	131
10.2.7.4 Calculation of oat β -glucan.....	131
10.2.7.5 Crop Height Estimation	132
10.2.7.6 Winter Oat yield Estimation by UAV	132
10.2.7.7 Thematic yield maps using Kernel density estimation.....	132
10.2.8 Statistical Analysis.....	133
10.3.1 Comparison of spectral profile between UAV and ground reference data	134
10.3.2 Statistical comparison of vegetation index from UAV and ground reference data.....	137
10.3.3 Soil nutrient data.....	142
10.3.4 Relationship between vegetation indices (NDVI and CI green) and chemical soil health indicators.....	154
10.3.5 Soil nitrate model development and validation.....	159
10.3.6 Estimation of oats plant height by multispectral-UAV at phenological growth stages.	163
10.3.7 Relationship between soil NPK and oat height in different zones.....	166

10.3.8 Correlation study between estimated and actual oats grain yield	173
10.3.9 Evaluation of thematic yield maps by kernel density estimation.....	175
10.3.10 Evaluating extraction and conversion efficiency of oat beta-glucan and analytical performance parameters of analysing glucose in oat beta-glucan.....	177
10.3.11 Hierarchical multinomial logistic regression model	178
10.4 Conclusion	191
Chapter 11: Overall conclusion and Future work	193
Chapter 12: Appendix	195
Chapter 13: References	235

Acknowledgement

Firstly, I want to express my sincere gratitude to my primary supervisor Professor John R. Dean for his unfailing support, encouragement, guidance, for helping me collect data and providing me with feedback on my work over the years. The assistance he provided has immensely made my PhD work shine. I could not have asked for a better supervisor to help me complete my PhD.

I would also like to thank my second supervisor, Dr Kate Nicholson, for sharing background knowledge and ideas on my data, as well as for generously lending me a computer to do all the heavy data interpretation. In addition, I would like to thank my third supervisor, Professor Justin Perry, for providing feedback on my work as well as funding for minor equipment/reagents required for my research. Also, I would like to thank Dr Nabanita Basu and Simon Rutter for helping me with data interpretation. I would also like to thank the lab technicians for providing me the instruments and equipment I needed to complete my research experiments, as well as for their encouraging words.

Furthermore, I would like to express my sincere gratitude to John Miller and his family for granting access to his agricultural fields for field sampling, granting permission to fly the UAV over his fields, and for providing agricultural data on his fields. I would also like to thank Jane Young, a former ecologist, and Paul Muto for providing background information on Priestclose wood, as well as Mr Geoff Dobbins, Senior Estates Officer (Reserves), Northumberland Wildlife Trust Ltd., for granting permission to fly the UAV over the wood. In addition, I would also like to thank Lynne Dean, Naomi Dean and John Marshall for assisting in field studies.

Finally, I would like to thank my family and friends for always encouraging me and providing me with the motivational words which always kept me in spirit to keep moving forward to complete my PhD. Also, my heartfelt gratitude goes to my Mom and Dad for supporting me financially without hesitation and for always wanting to see me shine and succeed. A special thanks also goes to my lovely sisters Sama and Shahama, who have been extremely supportive and encouraging throughout my PhD journey. In addition, I'd like to express my gratitude to my very close friends Maimoona and Shuaib for being there for me and supporting me. I also could not have done this without my best friends, Courtney Box, Courtney Olivo, Piegong, Sumaiya, and Sonali, who have listened to me, encouraged me, and pushed me to be the best version of myself. Finally, I would also like to thank my undergraduate supervisor, Dr. Mathi Kandiah, for her guidance, faith in me and being there for me always.

Declaration

I declare that the work contained in this thesis has not been submitted for any other award and that it is all my own work. I also confirm that this work fully acknowledges opinions, ideas and contributions from the work of others. Any ethical clearance for the research presented in this commentary has been approved. Approval has been sought and granted through the Researcher's submission to Northumbria University's Ethics Online System on 24.10.2020.

I declare that the Word Count of this Thesis is **45,237** words.

Name: Aishath Shara Ahmed

Signature:

Date: 27.03.2023

Chapter 1: History of aerial imaging by remote sensing platforms

1.1 Aerial imaging

Humans desire to see the Earth's surface for a multitude of practical and aesthetic reasons. This has been a success since the introduction of aerial imaging which helps to view and capture details of the respective Earth surface of interest above ground level. Aerial imaging captures photographs using film, digital, multispectral, hyperspectral, lidar, radar, image scanners or video cameras from remote sensing platform (Aber et al., 2010). The commonly used remote sensing manned aerial vehicle platforms are airplanes, helicopters, and satellites. However, since the beginning of mid-19th century when aerial imaging introduced, other remote sensing platforms such as balloons, tethered blimps, gliders, and kites were used, which have evolved over time to the remote sensing platforms that are commonly used today (Colwell, 1997). Aerial imaging was introduced in the beginning of mid-19th century, since then, the field has evolved through research, and the development of new equipment and techniques for use in a variety of applications.

1.2 History of aerial imaging using balloon platforms

Colonel Aime' Laussedat is credited with making the first attempt to take aerial photographs in 1849. He used kites and balloons, but his efforts were unsuccessful (Wolf and Dewitt, 2000). Later, in 1858, Gaspard Felix Tournachon, later known as "Nadar," took photographs of the French village of Petit Bicetre from a height of several hundred metres using a tethered balloon platform (Colwell, 1997). Later that year, after this successful attempt, he attempted to take aerial photographs using a glass-plate camera lifted by several kites, but it is uncertain if he was successful (Colwell, 1997). In this manner aerial imaging was used for other purposes such as during American Civil War between 1861 to 1865 to photograph the observation of enemy positions by hydrogen filled balloons as a platform, however non-of these photographs survived (Jensen, 2007). Samuel Archer King and James Wallace Black captured the earliest surviving image while levitating over Boston, Massachusetts, in a balloon (Jensen, 2007). In the meantime, Tournachon in France continued his experiments with balloons to take aerial pictures with only minimal success. Additionally, a German forester successfully identified and measured stands of forest trees in 1887 by taking aerial photographs with a balloon (Colwell, 1997). The improvement of 3D imaging and the experimenting of modern anastigmatic camera lenses by 1890 were both a result of the ongoing development of aerial photographic techniques in the 1850s (Zahorcak, 2007). Consequently, Frederic Eugene Ives

introduced and pioneered colour photography in 1890 (Kuehni, 2014). However, balloon aerial imaging was not widely used after the 19th century due to the method's inherent dangers and the high cost of maintenance (Aber et al., 2010). Due to its dependability and superiority to balloon aerial imaging, kite aerial imaging was developed and used for scientific, military, and general viewing of the Earth's surface (Beaufort and Dusariez, 1995). Edward Archibald, a British meteorologist, is thought to have taken the first aerial photograph using kites in the early 1880s after much debate on the subject (Colwell, 1997). In the meantime, Gaston and Albert known as the Tissandier brothers also carried out kite and balloon aerial imaging in France (Cohen, 2006). Also, according to other sources, Arthur Batut invented kite aerial imaging in France in 1888 (Beaufort and Dusariez, 1995).

1.3 History of aerial imaging using kite platforms

As a result, the early 20th century is regarded as the "golden age" of kite aerial imaging, during which kite platforms were widely available for hoisting a camera into the air to take aerial photographs (Aber et al., 2010). The 1906 earthquake and fires, which destroyed a large portion of San Francisco city, were successfully captured by George Lawrence in the 1890s using a train of nine kites that lifted a moving slit panoramic camera (Beaufort and Dusariez, 1995). He also photographed Pacific Grove, California, and Monterey Bay during the same trip. Thereafter, kite lifting methods began to develop during the era where Samuel Franklin Cody an English immigrant from the United States in the early 1900s began experimenting kites with the intention of designing kites to lift humans in the sky to capture aerial shots (Robinson, 2003b). To increase the lifting capacity to carry people, Cody created a double-celled Hargrave box kite with extended wings. He named this platform the "Cody kite" in 1901, and a successful demonstration of this manned kite platform took place at Whales Island in Portsmouth, England, in 1903 (Robinson, 2003b). Also, at different trials to capture aerial photographs, he further challenged at ascending the manned kites to heights of 60 m, 120 m, 240 m and 800 m, respectively. However, Cody's manned kite flying was practiced by few due to the cost and infinite risks involved in the method. Further, kite aerial imaging was practiced by René Desclée between 1910-1939 where his main interest was the Tournai city in France and its cathedral (Aber et al., 2010). He produced at least 100 outstanding aerial photographs using kites, amongst these the best-known aerial shots were portfolios prior to World War II (Beaufort and Dusariez, 1995). After his career, it was marked as the end to kite aerial imaging

golden age, as progress was made to more advanced manned lifting methods such as powered flight to capture aerial imaging.

1.4 History of aerial imaging using manned aircraft

Aerial imaging was also experimented with pigeons and rockets, as balloons and kites were too easy to shoot down during World War II (WWII) (Moore, 1979). In 1903, For carrier pigeons, Julius Neubranner created breast mounted cameras that weighed 70 grammes and had an automatic timing mechanism of 30 seconds for aerial photography (Baumann, 2014). However, photographs taken by pigeons were unpredictable as the flight paths taken by them were unreliable and the birds were occasionally shot down by hungry troops. Meanwhile, in 1897 Alfred Nobel experimented with aerial imaging on rocket platform where he mounted a camera on a rocket for aerial imaging. In 1906, Albert Maul in Germany was inspired by him to create a more dependable technique for taking aerial photographs at an altitude of 2,600 feet. However, all these platforms were not widely experimented with due to the rapid development of manned aircraft for aerial imaging.

In this way, L.P. Bonvillain, a passenger in a Wilbur Wright-piloted airplane, captured the first aerial photograph from a manned aircraft in 1909 while filming a military field at Centocelli, close to Rome (Crum, 2004). Aerial imaging from manned airplanes was used to a large extent during World War I (WWI) in which cameras were installed on manned airplanes to capture aerial shots up to 1000 images per day during the middle of WWI for the purpose of reconnaissance missions (Aber et al., 2010). Hence, in the beginning of WWII aerial imaging from manned airplanes were used for different applications such as glaciology, forestry, agriculture, and archeology (Konecny, 2014).

1.5 History of aerial imaging using satellites

Aerial images captured from balloons, kites and manned aircraft are equipped with a camera enabling to capture aerial photographs which are either panchromatic or coloured. It was in the beginning of the mid-20th century that new advancements in cameras for aerial imaging came into existence using satellite platforms (Aber et al., 2010). Space-based aerial imaging was first experimented with in 1960 when the United States launched TIROS-1 (Television and Infrared Observation Satellite) into space to capture aerial photographs for meteorological purpose. After this, in 1972, NASA initiated a remote sensing research program named Earth Resource

Technology (ERTS), which was later known as Landsat-1. This was in response to new advancements in cameras for aerial imaging (Read and Torrado, 2009). Landsat-1 was the first satellite to use multispectral scanners to provide non-military users with images of the Earth's surface in a predictable and repeatable manner. Since then, satellite platforms have seen further advancements in cameras, and in the 1980s, hyperspectral imagery systems were developed, enabling the capture of more concise images of the Earth's surface. As a result, satellites were launched in the 1990s, enabling the capture of the entire Earth's surface and in 1999 the first commercial satellites sensors were launched which provided improved details about the Earth's surface than previously available sensors (Read and Torrado, 2009). In this manner, satellite platforms in aerial imaging expanded in a wide range of application including urban mapping, precision agriculture, land use analysis, human environmental interactions and many more.

However, satellite-based aerial imaging platforms have a few drawbacks. Satellites have limited spatial and temporal resolution (Dubovik et al., 2005). The spatial resolution of a remote sensing sensor describes how many pixels are used to compose a digital image by the instantaneous field of view (IFOV) or ground sampling distance (GSD) (Figure 1) (Kislik et al., 2019). The temporal resolution is the time required to revisit and acquire data for the same location, and the shorter the revisiting cycle, the higher the temporal resolution. For example, Landsat sensor satellites have a minimum GSD of 30 m and 250 m spatial resolution respectively, and a 16-day revisit cycle (Wu et al., 2019). The increased revisiting cycle of satellites causes a large time-interval of re-acquiring satellite images causing a substantial gap for continuous analysis of land areas which can be a drawback in land use analysis applications.

Also, the spatial resolution in satellites is useful for capturing large areas, and unsuitable for capturing small areas such as rivers, lakes and farmlands, which only requires high spatial resolution to generate detail images with sufficient resolution. Additionally, weather conditions like cloud coverage have an impact on the quality of satellite images and must be rectified (Kislik et al., 2019). Due to these drawbacks of satellite platforms, remote sensing techniques using unmanned aerial vehicles (UAVs) were introduced. Also, UAVs have a spatial resolution of 0.1 - 0.5 m and can be flown multiple times per day to acquire sufficient images, making them ideal for continuous and long-term monitoring purposes.

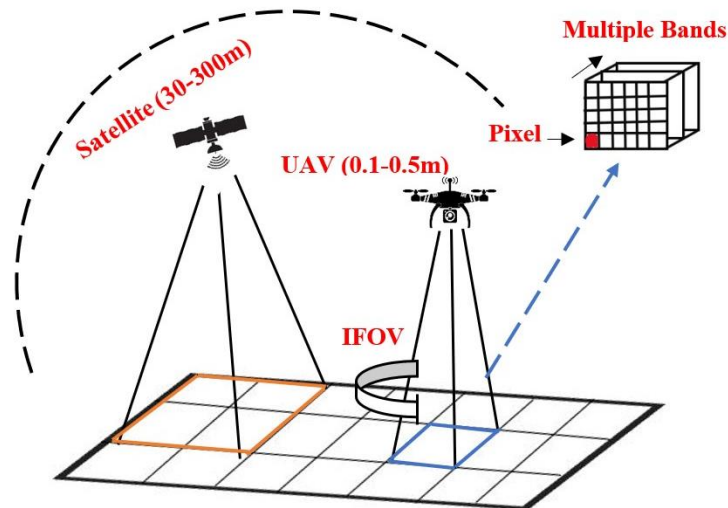


Figure 1: Spatial resolution of satellite and UAV.

1.6 History of aerial imaging using UAV platforms

UAVs are aircrafts controlled without a pilot or are completely programmed to automatically operate through software for flying. When Austria attacked Venice in 1849 using unmanned balloons loaded with explosives, the first attempt to fly a UAV was made (Cornelisse, 2003 and Scientific American, 1849). Afterwards, Nikola Tesla demonstrated his radio-controlled boat for the first time to an audience near New York Pond in Madison Square Garden (Posea, 2020). The craft responded to Tesla's directional signals by flashing its lights. This was a remarkable demonstration of what would later be known as radio-controlled aircraft. Also, the British forces produced one of the earliest examples of an orthomosaic of the German front during the Battle of Neuve Chapelle in the early 1900s, which was created by combining several individual aerial images (Daly, 2023). Additionally, the British Aerial Target, a small radio-controlled aircraft, made its first flight in March 1917, while the Kettering Bug, an American aerial torpedo, first flew in October 1918 (Imperial War Museum, 2023). In addition, the British used the term "drone" for the first time in 1935, which stands for Dynamic Remotely Operated Navigation Equipment (Posea, 2020). The Havilland DH.82B Queen Bee airplane, which was used for aerial target practice, served as the source of inspiration for the name drone. Consequently, Queen Bee is frequently cited as the very first modern drone. Afterwards, radio-controlled UAVs were also invented and used in the US for target practice and training. Also, in 1938, the first effective remote-control model plane known as "Big Guff" was spotted in the air (Daly, 2023). Additionally, the BQ-7 was the first instance of a first-person view UAV, developed by Boeing for the United States Air Force in 1943 (Posea 2023). The autopilot was

activated once the target was in sight, and the pilot bailed out of the plane and the BQ-7 would then autonomously fly to the objective. However, the BQ-7 was essentially ineffective in war, and pilots who bailed out frequently died or were captured. Meanwhile, the Predator, which was used in Afghanistan in 1996 to launch missiles and search for Osama Bin Laden, is the most well-known military UAV to this day (Imperial War Museum, 2023). In addition, the consumer UAV industry began in 2006 after the release of commercial UAV permit by the Federal Aviation Administration (FAA). In the same year Frank Wang, a student at Hong Kong University of Science and Technology, had the idea to drastically change the significance of quadcopters for the public, and he formally founded the DJI drone company (Haque, 2021) a company that is the world's largest seller of consumer UAVs today. Afterwards, in 2010, the French UAV manufacturer Parrot introduced the first UAV to be controlled directly from a smartphone application (Daly, 2023). In addition, in 2013, Amazon, Uber, Google, UPS and FedEx used UAVs to deliver their products in record time via airborne means (Posea, 2020). Even though the DJI company was formed in 2006, the well-known DJI Phantom 1 UAV was not released to the market until 2013 (Haque, 2021). This was the first UAV to hold a GoPro and be controlled from a distance. Following that, in 2015, the DJI Phantom 3 UAV was released, rapidly becoming one of the greatest consumer UAVs of all time (Crane, 2018). It had a stabilised 4k camera, the potential to view the flying view on a smartphone and had a battery life of 23 minutes. Furthermore, in 2016, the first concept of a foldable consumer UAV, the DJI Mavic was introduced (Haque, 2021). Also in 2019, DJI released the Mavic Mini, a 4k stabilised UAV weighing less than 250 g with extraordinary characteristics for such a small UAV (Crane, 2018). Further, the DJI Phantom 4 multispectral UAV was released in 2019 for specific applications of precision agriculture and environmental monitoring purposes (McNabb, 2019). This was the world's first UAV equipped with a multispectral sensor comprised of six different sensors. In recent years the history of aerial imaging using remote sensing platforms has shown an incredible advancement in technological innovation.

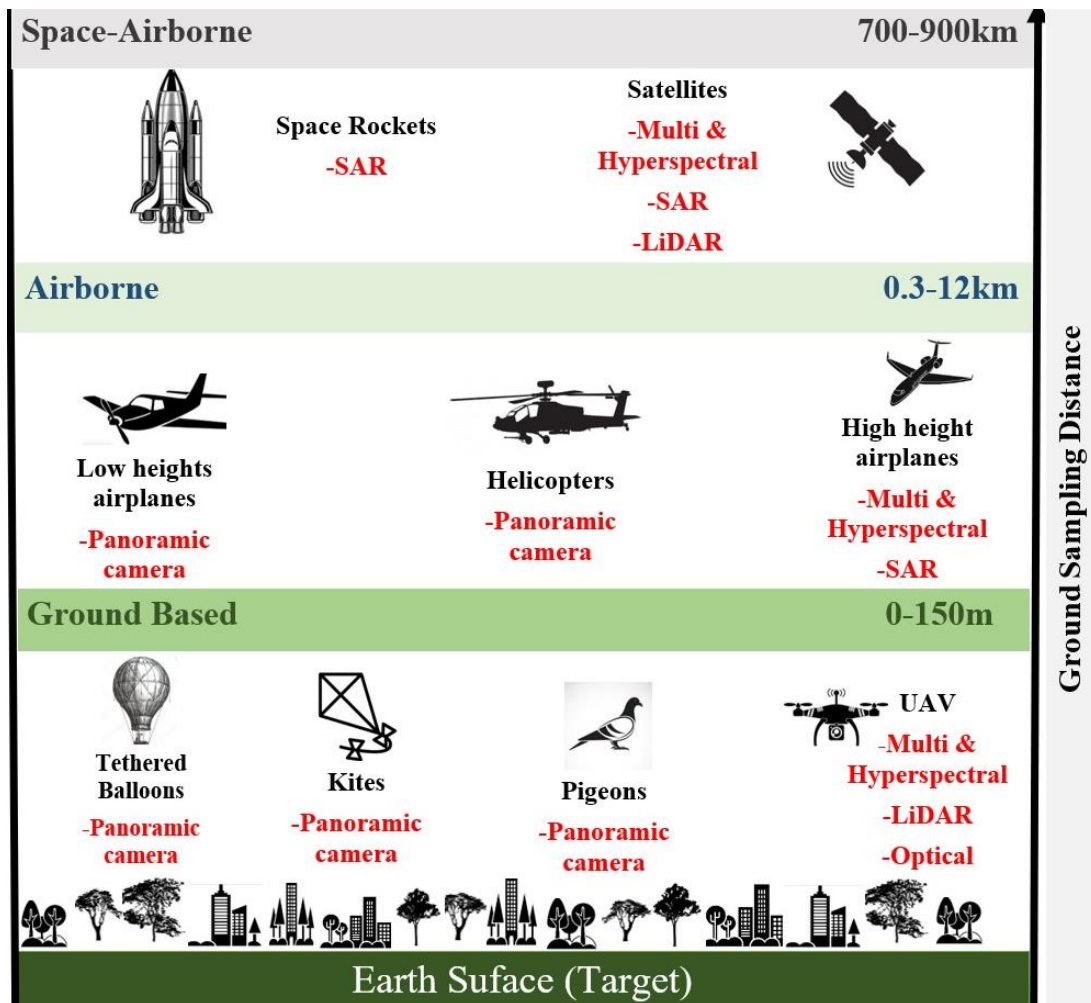





Figure 2: Evolution of remote sensing platforms for aerial imaging.

1.7 Types of UAV

UAVs vary in size, shape, and operational endurance (Ahmed et al., 2022). Hence, UAVs can be classified as either fixed wing and rotary wing. Fixed-wing UAVs have wings on both sides of the aircraft that enable them to lift themselves into the air and fly in only one direction at high speed (Laghari, 2023). The first miniaturised radio-controlled fixed wing UAVs were introduced in the 1960s (Daly, 2023). Further, rotary wing UAVs can be classified into single rotor and multirotor which are tricopters, quadcopters, hexacoverters or octocopters (Mohsan et al., 2022). Single rotor UAVs are helicopter-shaped which are expensive and require skilled training to operate (Laghari, 2023). Whereas tricopter UAVs have three propellers, quadcopter have four propellers, hexacoverters have six propellers and octocopters have eight propellers (Mohsan et al., 2022). Among these multirotor UAVs, quadcopters are the most widely used due to their simple structure, low cost, vertical landing, and high manoeuvrability (Mohsan et al., 2022). Etienne Omnichen also invented the first working quadcopter, the Omnichen 2,

which flew 360 metres, establishing a new world record. This was the first alternative to a single-rotor UAV (Posea, 2020). The benefits and drawbacks of fixed wing and rotor wing UAVs are summarised in Table 1.

Table 1: Benefits and drawback of different types of UAV

Types of UAV	Benefits	Drawbacks
Fixed wing 	<ul style="list-style-type: none"> -Operates at high speed -Large areas covered due to extended battery 	<ul style="list-style-type: none"> -Costly -Only moves forward and cannot hover in the air. -Horizontal take-off and landing necessitate open space. -Training is required to practice manual take-off
Single rotor 	<ul style="list-style-type: none"> -Long endurance and can hover in air. 	<ul style="list-style-type: none"> -Costly -Requires training to fly -Highly dangerous
Multirotor 	<ul style="list-style-type: none"> -Less costly -Vertical take-off and landing necessitate less open space -Easy to operate -Can move in forward direction and hover in air -Offers high stability 	<ul style="list-style-type: none"> -Short flight times due to limited battery capacity (30 minutes)

In comparison to other aerial imaging platforms, modern UAVs have very high spatial and low temporal resolution. As a result, aerial imaging by UAVs has a variety of applications, with forestry and precision agriculture being one of the most common. The use of UAV in forestry aids in the detection of tree diseases as well as the identification of native and invasive tree species. Precision agriculture employs UAVs to monitor crop fields and provide farmers with information on crop growth, health, and specific areas of nutrients that need to be applied to

crops to maximise yield. Due to the added advantage offered by the multirotor quadcopter UAV this PhD research uses a DJI phantom 4 multispectral UAV for analysis (Figure 3).



Figure 3: DJI Phantom 4 Multirotor UAV (DJI P4 Multispectral, 2023).

Chapter 2: UAV sensors and multispectral imaging

2.1 UAV sensors

The term ‘remote sensing’ was introduced by a researcher in the United States named as Evelyn Pruitt in the 1950s (Graham, 1999). This was since the traditional aerial imagery was inaccurate at describing the images gathered using electromagnetic radiation from outside the visible spectrum. Thus, remote sensing is defined as the gathering of data about objects on the Earth's surface without physical contact by spatially capturing reflected or emitted energy in any wavelength by sensors on remote sensing platforms (Jafarbiglu and Pourreza, 2022). All objects radiate electromagnetic energy through absorption, transmission, and reflection (Zhu et al., 2018). Frequently, the reflected energy from the objects is measured by the sensors on remote sensing platforms, which aids in the discovery of surface features important in forestry and precision agriculture. The contact between a sensor and the Earth's surface can be active or passive (Figure 4). Passive sensors use natural light (radiation from the sun), whereas active sensors use artificial light from their own source of energy and record the incoming reflected radiation (Pádua et al., 2017). Further, the passive sensors capture images either by scanning or non-scanning methods. Non-scanning methods are used by traditional remote sensing platforms of balloons, kites, pigeons, and airplanes to capture images by cameras. Whereas, scanning methods are used by modern UAV platforms, where the sensors scan the surface to collected multiple aerial images which then can be merged to form an orthomosaic. Further, the type of scanning sensors are thermal imaging sensors and optical imaging sensors of red, green, blue (RGB), multispectral and hyperspectral, while active sensors include radar imaging and light detecting and ranging (LiDAR) sensors (Pádua et al., 2017).

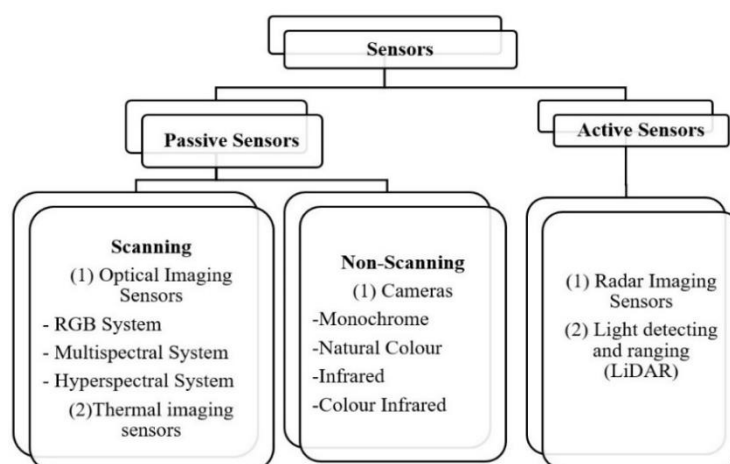


Figure 4: Classification of UAV sensors

The visible and infrared ranges of the electromagnetic spectrum as shown in Figure 5 are used by UAV sensors in forestry and precision agriculture.

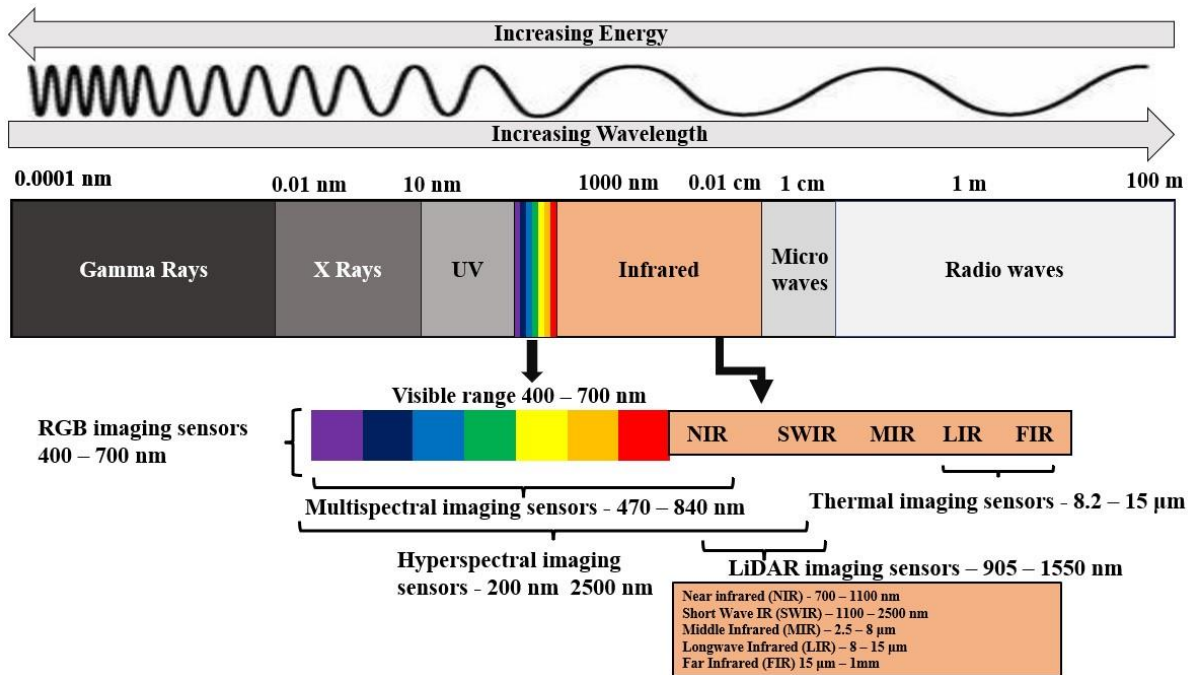


Figure 5 : Electromagnetic spectrum with UAV sensors applicable in forestry and precision agriculture

2.2 RGB UAV sensors

The red, green, and blue spectral bands in the visible portion of the electromagnetic spectrum are employed by RGB sensors to capture aerial images that are visible to the human eye. RGB operates in the wavelength range of 400 to 700 nm. Furthermore, RGB sensors installed on UAV platforms are widely available and capable of capturing high-resolution imagery for forestry and precision agricultural studies. RGB sensors have been effectively used to study visual symptoms of diseases in plant leaves (Padmavati and Thangadurai, 2016). Also, UAV RGB sensors have been used to monitor and identify canopy gaps to evaluate the floristic species diversity of the forest understory (Getzin et al., 2012). In addition, the images acquired by RGB sensors can be separated into red, green, and blue channels to work separately to reproduce meaningful features (Pádua et al., 2017). Hence, this has been utilised in vegetation segmentation of crops by retrieving information separately from individual RGB channels (Nolan et al., 2015). However, due to the limited spectral bands used, RGB sensors have limited application in forestry and precision agriculture, as information obtained in the near-infrared (NIR) spectral bands is significantly important for assessing vegetation properties.

2.3 Thermal UAV sensors

Thermal sensors operate approximately between the longwave infrared and far infrared of the electromagnetic spectrum between 8.2 – 15 μm (Jafarbiglu and Pourreza, 2022). The pixel intensity of an image captured by thermal sensors can be translated into a cooler or hot temperature. UAV thermal sensors have been used to generate thermal maps for land classification (Lagueta et al., 2015). Also, UAV thermal sensors have been used in soil salinity detection in irrigated agricultural lands (Abbas et al., 2013) and to assess water status assessment in different crop species (Gonzalez-Dugo et al., 2013). However, in comparison to RGB sensors, thermal sensors are costly and have low image resolution (Olson and Anderson, 2021). Also, thermal imaging is affected by high altitudes due to the atmospheric interference of light intensity, humidity, temperature, and wind (Costa et al., 2013).

2.4 LiDAR UAV sensors

To create 3D point clouds, LiDAR sensors emit an electromagnetic signal in the form of a laser beam in a specific direction to measure the distance between sensor and the object being measured (Debnath et al., 2023). The electromagnetic signal is produced in the near infrared range of the electromagnetic spectrum, with wavelengths ranging from 905 to 1550 nm. In forestry application UAV LiDAR has been used for mapping understory canopy (Chisholm et al., 2013) and forest canopy structure (Wallace et al., 2012). Furthermore, in agricultural studies, UAV LiDAR was used to monitor and manage agricultural lands (Tarolli et al., 2015), estimate biomass in maize crops (Jin et al., 2020), and estimate leaf area index (Arnó et al., 2012). The UAV LiDAR sensors generate data with millimetre spatial resolution, making it suitable for forestry and agricultural applications. However, processing LiDAR data points is computationally demanding and costly, and the software used to generate the data is even more costly (Jafarbiglu and Pourreza, 2022). Also, LiDAR data must be supplemented by images or videos for analysis.

2.5 Hyperspectral UAV sensors

Hyperspectral sensors utilise hundreds or thousands of spectral bands in the visible and the near infrared portion of the electromagnetic spectrum between 200 – 2500 nm. Detail features can be extracted to reveal significant information applicable in forestry and precision agriculture due to the high spectral resolution offered by hyperspectral sensors. Hyperspectral UAV sensors have been used to detect diseases in forest trees (Safonova et al., 2019), for obtaining accurate measurements of tree heights (Zou et al., 2019) and in tree species

identification (Cao et al., 2018). Further, hyperspectral UAV sensors have been applied in precision agriculture to estimate biomass and nitrogen content (Neale et al., 2013), for monitoring soil moisture content (Ge et al., 2019), for disease identification in crops (Oré et al., 2020) and for crop yield estimation (Tao et al., 2020). However, UAV hyperspectral sensors, suffer from the sensors being costly, the data being computationally heavy and time consuming to interpret, limiting its regular application for research purpose (Olson and Anderson, 2020).

2.6 Multispectral UAV sensors

Multispectral sensors utilise a limited number of spectral bands in the visible and the near infrared region of the electromagnetic spectrum between 470 – 840 nm. The significant advantages of multispectral UAV sensors in forestry and precision agriculture are the low-cost of the sensors and the easy manipulation of the multispectral data. Furthermore, data from the near-infrared spectral band is typically sufficient to extract significant information about the health and vegetation cover of trees and crop species. Hence, multispectral UAV sensors have been used to detect invasive and native tree species (Brovkina et al., 2018 and Liu et al.; 2018), and biomass estimation in tropical woodland species (Dominogo et al., 2019 and Di Gennaro et al., 2020). Further, multispectral UAVs have been used extensively in precision agricultural studies for biomass estimation in crops (Zheng et al., 2022), for nutrient assessment in crops (Noguera et al., 2021), for disease detection (Su et al., 2018) and yield estimation in crops (Zhou et al., 2021).

Due to the significant advantages provided by multispectral UAV sensors over other UAV sensors in forestry and precision agriculture, the concept of multispectral imaging (MSI) in relation to different types of multispectral sensors will be discussed further.

2.7 Concept of multispectral imaging (MSI)

Images are all around us in different forms, such as colourful photos in magazines, billboards, and paintings (Amigo, 2020). Also, digital images are seen on electronic devices such as computers and mobile phones which are generated by using digital cameras (Cromey, 2013). Digital cameras capture an image as light passes through the lens of the camera and reaches the digital sensor of the camera, where the light signal is converted into electrical signals. The electrical signals are stored as tiny dots which, when appeared on a camera screen or downloaded onto a computer screen, are turned into pixels (Note, 2011). Therefore, in short digital images are made up of small elements called pixels which when zoomed enough into

an image are observed like a mosaic arranged in a 2-dimensional grid by small, tiny squares (Figure 6).

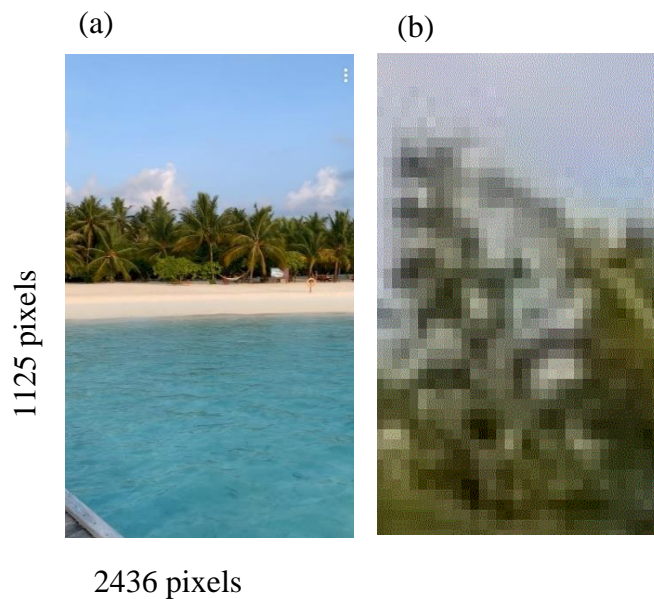


Figure 6: (a) A digital image captured from a digital camera (b) The digital image zoomed in to show the pixels in the image.

Hence, the quantity of pixels in an image is characterised by resolution which describes the amount of detail in an image expressed as the pixels from one side to the other 1125 pixels and from top to bottom 2436 pixels as shown in Figure 6. Thus, the image in Figure 6 has a pixel resolution of $1125 \times 2436 = 2,740,500$, or 2.74 megapixels. Pixel resolution describes the amount of detail in an image through pixels and insufficient pixels in an image leads to blurry or unclear images (Note, 2011). If staring closely at the screen of an old day TV the individual little red, blue, and green onscreen pixels can be seen. But, on today's LCD TVs this cannot be observed as it has millions of pixels (megapixels) in the same space leading to clearer images that appear more natural. Additionally, the smartphones used in today's world have cameras with up to 21 megapixels to capture very high-quality digital images. (Amigo, 2020). The amount of information a pixel contains is generated from a channel which corresponds to a range of wavelengths and contains spectroscopic information (Cromey, 2013). The colour images visible to the human eye generated from a camera are normally composed of three channels of spectral information, the red, green, and blue channels (Note, 2011). However, multispectral images are composed of more than three band channels incorporating the red edge and near infrared spectral bands (Amigo, 2020).

To understand multispectral images, it is important to initially understand how a colour image is generated by a digital colour camera. The digital camera consists of a sensor arranged of a colour filter array (CFA) also known as the ‘Bayer array’ consisting of three cells with a filter pattern of 50% green, 25% of red and 25% blue channels known as ‘RGGB’ pattern (Figure 7) (Bayer, 1976). The colour sensor works as light passes through the lens of the camera from the object to capture the colour image. As light reaches the sensor, it captures the image as red, green, and blue images using the colour filter array, and afterwards these three images are combined to form a colour image visible to the human eye (Amigo, 2020).

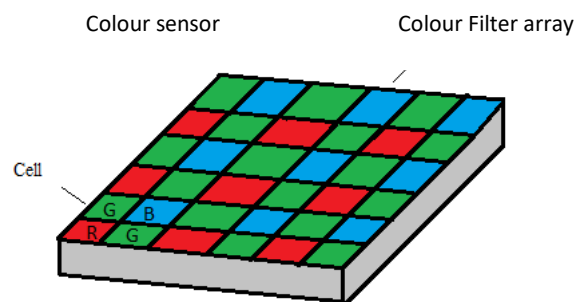


Figure 7: Colour filter array arrangement on a colour sensor found in digital camera to compose coloured images.

Whereas multispectral images reveal information that is neither observable to the human eye nor to the RGB camera. The colour images from the RGB camera are insufficient to draw meaningful interpretations; instead, a significant number of channels such as the NIR along with visible region of the electromagnetic spectrum are used in MSI. Therefore, multispectral images captured using the additional channels reveal hidden information through images (Note, 2011).

2.8 Types of multispectral sensors for MSI

Multispectral images can be generated by two main methods; scanning the area of interest by a spectroscopy technique and by snapshots in which images are captured using a camera/sensor. This PhD research is focused on using a UAV which has a sensor for MSI, therefore the development of snapshot method for MSI will be discussed in detail further.

The snapshots of the MSI technique captures images using a single sensor or multiple sensors by means of different channels or band filters (Lapray et al., 2014).

The beam splitting method uses several sensors and a setup of dichroic beam splitters to divide the incoming light into, for example, red, green, yellow, and blue beams which are subsequently displayed through lenses onto individual camera sensors to generate multispectral images (Lapray et al., 2014). The drawback of this technique in MSI is the limited number of channels, that requires calibration making it costly and limiting its application. Secondly, amongst the single sensor only few interferometer techniques have been adapted to MSI, mainly due to limited number of channels and to obtain more spectral channels it requires larger filters which is expensive to generate clear images (Kudenov et al., 2009). Thirdly, in the MSI method of filtered Lenslet array, the sensors responsible for producing multispectral images are smaller in size which results in a low resolution of picture (Shogenji et al., 2004). The tunable sensors in MSI cameras are photosensitive sensors capable of producing images at different spectral channels with a high sensitivity as spectral responses on each pixel can be tuned individually (Langfelder et al., 2009). However, the technique is again limited to few spectral channels thereby limiting the technique to few MSI application.

Finally, in comparison with other snapshot MSI approaches, the multispectral filter array (MSFA) has been of interest in the past few years. MSFA is simple and a compact system to capture images, as the earlier approaches are either bulky, have limited spectral channels, are expensive and require professional calibration (Lapray et al., 2014). MSFA incorporates pixelated channel filters in which the transmission characteristics can be precisely tuned during manufacturing. The MSFA consists of cells (eg; B1, B2, B3 and B4 Figure 8) which contains different channels/filters between 2-16 to separate colours in the image by a specified wavelength. These cells are arranged in a 2-dimensional array to form a MSFA mosaic which is aligned and integrated into a sensor in the camera (Wu et al., 2019). These features of MSFA captures multispectral images from a wider spectral channel which are easier to integrate, and the technique is relatively cost efficient. Hence, the sensor in the UAV (Figure 9) used in this PhD research uses a MSFA which consists of complementary metal oxide semi-conductor (CMOS) sensor of five different cells consisting of blue, green, red, red edge and NIR channels/filters applicable for multispectral imaging.

Hence, multispectral sensors were chosen for this PhD project due to the numerous benefits offered by the sensors to extract meaningful information in forestry and precision agricultural applications. Multispectral sensors can monitor subtle variations in canopy health as it captures the reflected light from the studying canopy within a critical part of the electromagnetic spectrum using visible and near-infrared spectral bands. Therefore, multispectral sensors with

fewer spectral bands are cost-effective and sufficiently effective to collect data on crop growth, yield and canopy classification in forestry and precision agriculture.

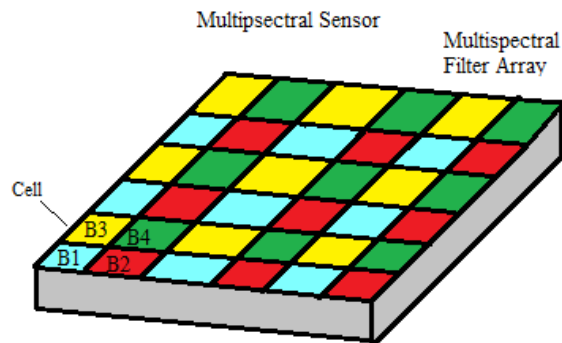


Figure 8: MSFA arrangement on a multispectral sensor to compose multiple images at specified channel of wavelength.

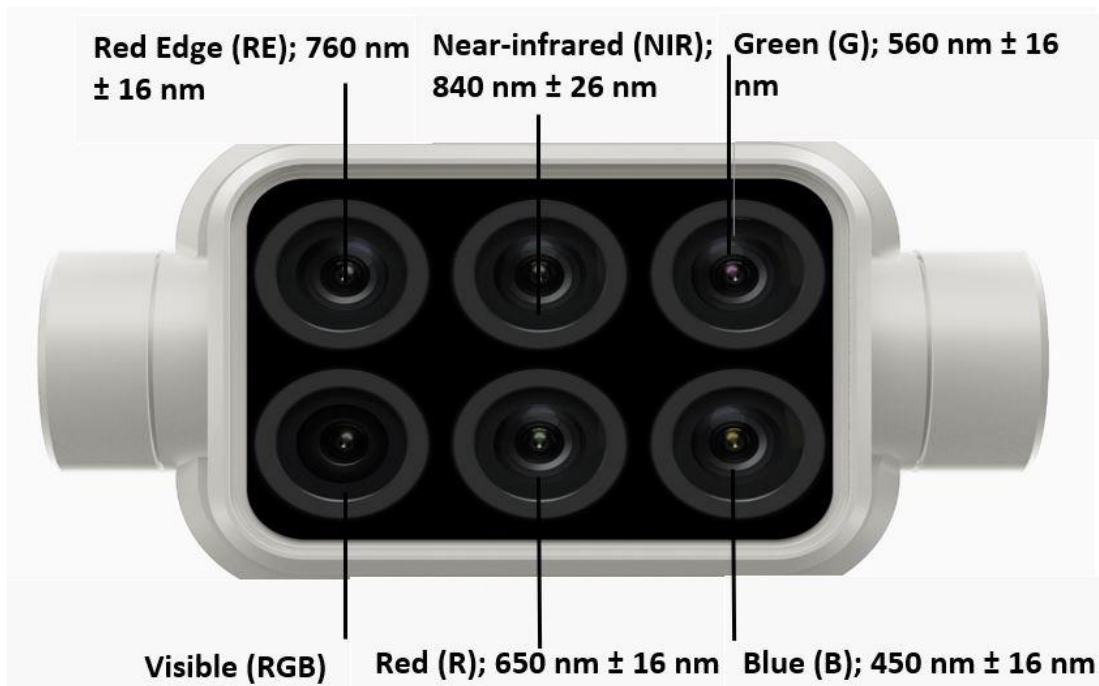


Figure 9: DJI Phantom 4 Multispectral UAV (DJI P4 Multispectral, 2023).

Chapter 3: Image processing by multispectral UAV

3.1 Image pre-processing

During UAV flights the multispectral sensors collect a large amount of raw data from each spectral band in the form of multiple sets of images which needs to be pre-processed to extract meaningful information. The first step of image pre-processing is atmospheric, geometric, and radiometric correction. Atmospheric correction is needed to account for image distortion caused by atmospheric phenomena like gaseous absorption, aerosol scattering, and absorption. However, unlike satellite platforms, UAV platforms fly at a lower altitude closer to the object of analysis, thus the atmospheric effect is minimal and does not require correction (Huang et al., 2021). To standardise the relationship between incident radiation and sensor output obtained at different times or locations, radiometric calibration is required (Shahi et al., 2022). The possible radiometric calibration methods are: (1) using reference reflectance panels with known reflectance factors placed in a well illuminated areas during UAV flight (2) having a static device on the ground to take ground readings during the UAV flight and (3) having an inflight irradiance sensor (Suomalainen et al., 2021). From these methods the most applied radiometric calibration in forestry and precision agriculture is using reference reflectance panels. Geometric calibration is performed when the sensors characteristics and external factors results in distorted images (Pádua et al., 2017). However, DJI UAV platforms require minimal geometric calibration due to complex handling and the sensors being more difficult to detach from the UAV (Cramer et al., 2017). Furthermore, the commercial UAV image pre-processing software Agisoft Metashape Professional and Pix4dMapper have been successful in performing calibration to produce reflectance orthomaps or orthomosaic images by stitching individual images which has been applicable in forestry and precision agriculture (Shahi et al., 2022). In addition, these software's uses a scale-invariant feature transform (SIFT) algorithm to perform UAV image stitching (Jia et al., 2015). It is divided into three steps; image pre-processing, image registration (feature extraction, matching, and model transformation), and image fusion.

3.2 Feature extraction

3.2.1 Spectral features

The orthomosaic generated consist of reflectance values of surface objects measured by the sensor on different spectral bands of the electromagnetic spectrum. As a result, orthomosaics are generated separately for each spectral band, and spectral indices can be calculated using two or more spectral bands to extract meaningful information applicable to forestry and precision agriculture. The surface of an object is affected by its materials, such as the soil, water, and greenery, on how much light it reflects. While soil is more reflective in the mid-infrared range and green vegetation is highly reflective in the near-infrared range and water, for instance, absorbs the most in these wavelength ranges. As a result, the reflectance and absorption of electromagnetic radiation by earth objects can be distinguished by spectral indices (Tsouros et al., 2019). The spectral indices calculate the intensity values for each pixel on the image, for instance representing the greenness of each pixel. Hence, the widely used spectral index of normalised difference vegetation index (NDVI) was a reliable indicator for estimating yield and observing the rates of fertiliser application in rice and wheat crops (Guan et al., 2019). Further, UAV derived NDVI orthomosaic was successful in distinguishing multiple tree species in a woodland (Fawcett et al., 2020). Although spectral features assist in the extraction of meaningful information from orthomosaics, additional feature extraction methods are required for effective data manipulation.

3.2.2 Structural and texture features

Structural feature extraction from orthomosaics include crop height and biomass estimation. By calculating the difference between the top of the crop (digital surface model) and the bottom of the crop (digital terrain model), orthomosaics can be used to create crop height models (CHM) As a result, UAV imagery was used to estimate faba bean crop height across different growth stages, as well as to estimate faba bean yield using the estimated crop height (Ji et al., 2022). Meanwhile, texture features from orthomosaic determines the spatial distribution of pixel intensities and multiple statistical approaches can be used to measure and extract texture features. Thus, the tasselling date of the summer maize crop was determined by using the grey level co-occurrence matrix (GLCM) to extract texture features from UAV images and by combining the extracted features with NDVI (Guo et al., 2021).

3.3 Machine learning models

Machine learning (ML) models are frequently used in forestry and precision agriculture to complement the extracted spectral, structural and texture features. For instance, ML models are used for yield estimation, disease detection and classification of crops in precision agriculture. Whereas, in forestry ML models are applicable in identification of native and invasive tree species. The ML models can be classified as supervised and unsupervised learning models covering both regression and classification algorithms (Figure 10) (Mehmood et al., 2022). The supervised models require a training data set to build the model, whereas the unsupervised models do not require a training data set.

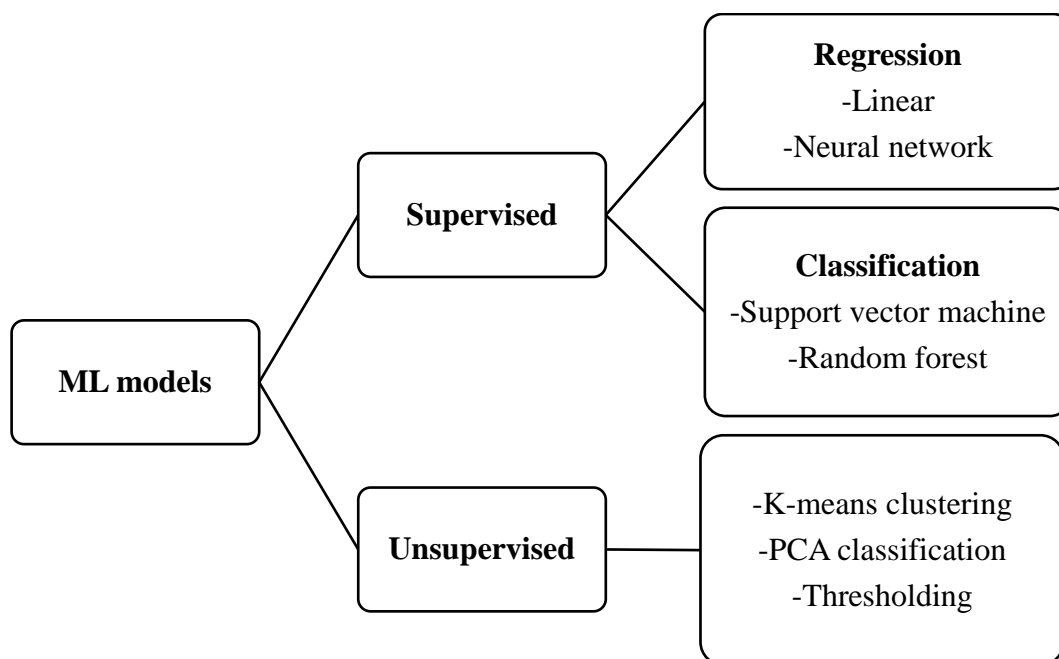


Figure 10: Classification of machine learning models

Hence, using multispectral UAV imagery spectral features were used to build multiple linear regression ML models to estimated rice crop yield (Zhou et al., 2017) and structural features of crop height were utilised to estimated corn yield by using a neural network ML model (García-Martínez et al., 2020). Further, northern blue light disease in maize crops were identified with an accuracy of 95% by neural network ML using UAV imagery (Wu et al., 2019). In addition, multiple crops were classified using GLCM texture features extracted from UAV images with an accuracy of 90% by random forest ML classification model (Kwark and Park, 2019). Also, an unsupervised ML K-means clustering algorithm was combined with a graph-cut algorithm on UAV RGB images to extract the pixel areas of rice crop to estimate rice yield (Reza et al., 2019). Meanwhile, multispectral UAV data was used for classification

of four native tree species by random forest classification model (Franklin and Ahmed, 2018) and multiple likelihood classifier was used for classification of eleven plant species in a nursery (Gini et al., 2018). Further, the workflow of utilising the UAV imagery data is summarised in Figure 11.

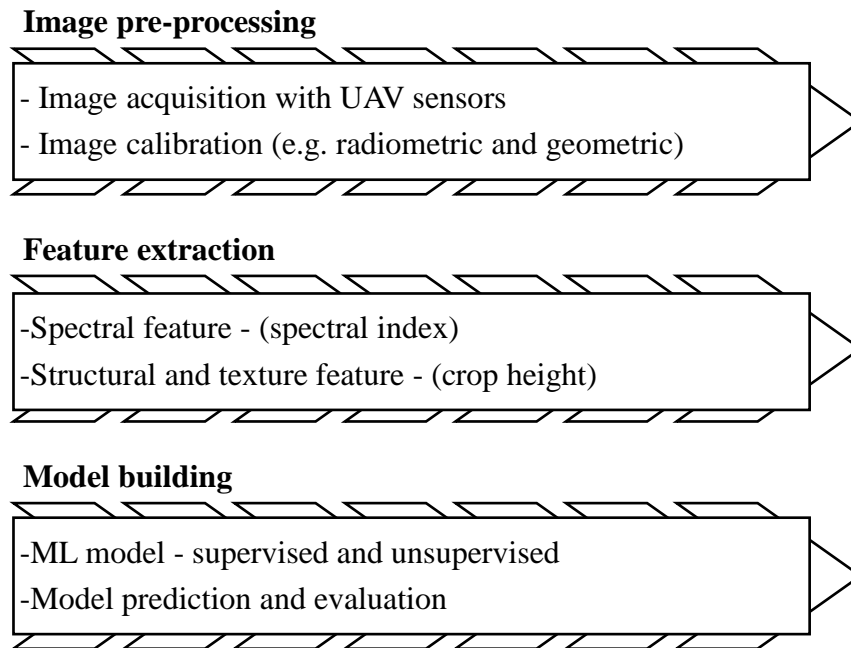


Figure 11: Workflow of using UAV imagery data.

3.4 Research gap

There has been a significant amount of study conducted with UAVs in forestry and precision agriculture over the last few years, but there are still some research gaps that need to be addressed. Firstly, the modern supervised ML models require thousands of training data set which can be computationally heavy and time consuming. Also, in most instances it is still unclear if these ML models are easily transferable for analysing other data sets as each model has its own limitation. Also, there are a limited number of studies which uses unsupervised ML models in both agroforestry and precision agriculture. The unsupervised classification ML models of k-means clustering, principal component analysis (PCA) and thresholding automatically learn the features within the image data set which makes the algorithm less time consuming. Further, to increase the accuracy of the unsupervised models the different classification models can be combined to retrieve information with high level of accuracy. Furthermore, PCA can be used to eliminate unwanted features from images and obtain the most

significant features by generating principal component (PC) score images. Information from the PC then can be used to generate spectral indices, which, when combined with unsupervised ML models, will aid in the classification of different tree species in forestry, which has previously not been performed. This can be a less time consuming and an alternative method to the already existing supervised ML models. Furthermore, there have been few to no studies on extracting structural features of crop height from UAV images at various growth stages in various agricultural crops of rapeseed, winter bean and winter oat grown in the UK. Also, extracted spectral features can be combined with an unsupervised clustering machine learning algorithm like k-means to segment pixel areas on these crops to estimate crop yield. Furthermore, a multinomial logistic regression supervised ML model can be used to predict the estimated yield using chemical soil health indicators (nitrogen, phosphorus, potassium, pH, and soil organic matter) and crop quality features (e.g., oats grain quality by β -glucan). This ML model, which has not previously been performed, would be useful for providing concentration range of soil health indicators at different crop yield levels (low yield and high yield). Finally, different software packages of eCognition, ENVI and R have mainly been used to perform the ML models in both precision agriculture and forestry. Limited number of studies have been performed using software packages of MATLAB and ArcGIS.

Filling these research gaps with appropriate data manipulation methodologies employing UAV could significantly benefit the forestry and precision agriculture fields in the long term. The simpler data manipulation strategies for instance would benefit in forestry application to identify native tree species and to map canopy gaps on a larger scale which is essential to maintain a sustainable ecosystem. Furthermore, decline in tree species can be reversed promptly when discovered by such highly accurate, short survey time, and cost-effective technologies. Also, these methodologies can be utilised for routine monitoring rather than the usual field survey, which is labour consuming and expensive. Additionally, due to a growing population, farmers typically find it difficult to meet the increased global demand for commodities. For improved scalability of farming these effective UAV data manipulation strategies will provide crop height and crop yield precisely across the growth stages over several months. This overcomes the manual crop evaluation method and enables farmers to accurately apply fertiliser to a specific area to increase yields of crops for precision agriculture. Furthermore, excess use and wastage of fertilisers are reduced, which might negatively impact crops and have a detrimental influence on the environment. Farmers will also be able to take

measures immediately and make essential plans to ensure high productivity of crops to meet the global demand.

3.5 Aims and Objectives

The goals of this PhD project were to use multispectral multirotor UAV imagery data in two applications: forestry (Section A) and precision agriculture (Section B):

The aims in forestry (Section A) were to use multispectral UAV imagery data and MATLAB software:

- To use PCA features to derive spectral vegetation indices to classify native and invasive tree species in a nature reserve woodland.
- To use an unsupervised ML algorithm of thresholding to classify the native tree species.
- To use an unsupervised ML models of k-means clustering to classify the invasive tree species.

The aims in precision agriculture (Section B) were to use multispectral UAV imagery data and ArcGIS software to analyse three different crops of oilseed rape, winter bean and winter oats:

- To analyse the most suitable spectral vegetation indices to distinguish crops.
- To use the features from the spectral indices to classify the crop area pixels using an unsupervised clustering algorithm inbuilt in ArcGIS software to estimate crop yield.
- To statistically evaluate the correlation of the estimated crop yield to the actual crop yield.
- To develop multinomial logistic regression ML models for determining the concentration range of chemical soil health indicators and crop quality features at various crop yield levels (e.g., low and high yield)
- To extract crop height across phenological growth stages of crops.

SECTION A

MULTISPECTRAL UAV IN FORESTRY

Chapter 4: Multispectral UAV monitoring of ancient and semi-natural woodland

4.1 Introduction

4.1.1 History of ancient and semi-natural woodland in UK

Nature reserve woodlands are sanctuaries for native tree and shrub species that provide habitat for a variety of fungi, invertebrates, birds, mammals, and reptiles, all of which contribute to a healthy ecosystem (Cooper, 2000). The history of woodlands in the UK can be traced back to the Ice Age (12,000 years ago) (Atkinson, 2011). By the Iron Age (750 BC- 40 AD), agricultural activity had cleared 50% of the UK's woodland (Goldberg et al., 2007). As a result, coniferous tree plantation was initiated in response to the need to conserve woodland, and the conservation movement began to protect ancient woodland as nature reserve sites in 1950 (Woodland Trust). The term "ancient woodland" refers to woodland that predates 1600 AD; however, some of these woodlands have been modified by human activity and are therefore often referred to as "ancient and semi-natural woodland" (ASNW) (Atkinson, 2011). In comparison to other European countries, the UK has one of the least wooded areas as they were slow to recognise the need to preserve nature reserves. It is thus an important cultural and social responsibility to protect the destruction of woodlands by urban development and to restore and create diverse and wildlife-rich natural ecosystems. In England, there are approximately 340,000 hectares of ancient woodland, defined as sites with continuous woodland cover since 1600 AD. Also, 200,000 hectares are considered semi-natural, implying they are of natural origin rather than artificially planted (Managing ancient and native woodland in England, 2010).

4.1.2 Significance of monitoring ASNW species

It is critical to monitor the ASNW species for the long-term viability of the ecosystem. Controlling the spread of invasive plant species that are not native to the area is also a major challenge in conserving ASNW (Papp et al., 2021). Invasive plant species in a woodland are trees and shrubs introduced by humans to a specific location, which can have a negative impact on the balance of a native ecosystem (Tobin, 2018). Invasive plant species can colonise large areas of land by competing for resources. This can eventually inhibit the growth of native species, affecting their development and the ecosystem (Papp et al., 2021). Furthermore, the spread of invasive plant species is ultimately expensive to remove after their occupation within a community (Drechsler et al., 2016).

4.1.3 Methods to monitor native and invasive species in ASNW.

Field-based assessment methods to assess trees below the canopy level have traditionally been used in ASNW to monitor native and invasive tree species. However, challenges with this approach may arise due to restricted access caused by ground cover plants and the physical location (Tehrany, 2017). Furthermore, field assessment methods are costly and time consuming, so the use of remote sensing technology has long been favoured as a tool for monitoring native and invasive species (Cooper, 2000).

The two main remote sensing approaches for mapping and locating invasive tree species, namely, low spatial resolution and high spectral resolution, via satellite, and UAV remote sensing platforms (Mafanya et al., 2017). As a result, several satellite platforms have been used to detect and delineate native trees (Chemura et al., 2015, Freudenberg et al., 2019, Strestasathiern et al., 2014, Satoso et al., 2016 and Li et al., 2014). Hyperspectral satellite platforms have also been used to differentiate invasive tree species based on their spectral signatures, structural and functional properties (Bradley and Mustard, 2006, Huang and Asner, 2009, Vilà et al., 2011, Lantz & Wang, 2013, Khar et al., 2017 and Ng et al., 2017).

Furthermore, conventional airborne multispectral and hyperspectral manned aircraft have been used to monitor invasive plant species (Mirik et al., 2013a, Mirik et al., 2013b, Dronova et al., 2017 and Skowronek et al., 2017). However, even if these traditional remote sensing platforms provide relevant spatial resolution, they are expensive and have increased temporal resolution in comparison to UAVs, which delays regular time series monitoring. Furthermore, the low spatial resolution of 30-300 m/pixel provided by satellite platforms makes monitoring and isolating a small population of invasive tree species and shrubs ineffective (Lehmann et al., 2017). UAV platforms, on the other hand, can capture images with exceptional spatial resolution (centimetre scale) and high temporal resolution, allowing for multiple images to be captured per day.

4.1.4 Literature on multispectral UAV monitoring of native and invasive tree species

Several machine learning methods of object based-image analysis (OBIA) were used to determine target species in recent research studies using multispectral UAV imaging for species recognition in areas with high biodiversity of tree species such as Canada, Italy, Nepal, and Costa Rica. Multi resolution and random forest methods were used to classify target plant species of Aspen, white birch, sugar maple, and red maple in Canada, as well as trees and shrub species along the Himalayan ectone in Nepal (Franklin and Ahmed, 2017 and Mishra et al., 2018).

Furthermore, maximum likelihood classifiers were used to classify various plant species in an Italian plant nursery (Gini et al., 2018). Yaney-Keller et al., 2019 used NDVI with a spatial resolution of 100cm/pixel to distinguish seven abundant mangrove species in Costa Rica, which were then classified using a support vector machine. In addition, UAV studies for monitoring invasive plant species have emerged globally from Africa, North America, South America, Europe, and New Zealand with promising results (Dvořák et al., 2015, Lehmann et al., 2017, Lishawa et al., 2017, Mafanya et al., 2017, Perroy et al., 2017, Martin et al., 2018, Lopatin et al., 2019 and Papp et al., 2021). Furthermore, studies using multispectral UAVs to detect invasive tree species have been used for mapping Sydney golden wattles plants in Portugal using random forest classification models (De Sá et al., 2019). Also, multispectral UAV allowed detection of invasive exotic Scots pine and ponderosa pine in New Zealand using random forest and logistic regression models (Dash et al., 2019).

4.2 Research gap

Firstly, there has been no multispectral UAV monitoring of native and invasive tree species in ASNW in the UK. Given that the UK is one of Europe's least wooded countries, monitoring these species is essential. Secondly, the commonly used methods for native and invasive tree species in woodlands are performed by supervised ML models. Less time-consuming algorithms that are user-friendly, accurate, and less complicated are still required to monitor tree species, particularly in situations where tree species analysis is required on a regular basis. As a result, image processing methods that use algorithms like normalised difference spectral indices (NDSI) can be derived by utilising information from multispectral images, which improves the ability to distinguish spectral features in tree species classification. The most used spectral index or vegetation index of NDVI is used to monitor crop growth and health (Rouse et al., 1974). Other vegetation indices, such as the optimised soil adjusted vegetation index (OSAVI), are used for crop management and biomass estimation, while the enhanced vegetation index (EVI) is used to quantify the greenness of vegetation (Pádua et al., 2017). So far, each vegetation index has been used for a specific purpose. As a result, since each tree species contains varying amounts of pigments such as chlorophyll and carotenoids, it will be more reasonable to use information from multispectral bands to derive a suitable NDSI specifically to determine a tree species. Furthermore, these pigments react differently to the spectral bands.

PCA can improve the sensitivity of individual multispectral images before deriving a NDSI. The discriminated tree species derived from a specific NDSI can be further segmented using

simpler thresholding segmentation methods. Thresholding is a pixel-based method for segmenting image regions based on intensity values. In contrast to OBIA methods, the thresholding method only considers an image's spectral features. If NDSI can distinguish between tree species well enough, it can be a useful segmentation method for classifying limited native tree species. Furthermore, multiple invasive tree species can be classified using k-means clustering, an image segmentation technique that divides the number of trees in an image into distinct tree clusters to be quantified. As a result, these practical classification models have not been used to classify native and invasive tree species in ASNW in the UK.

4.3 Summary

Due to the importance of preserving ASNW in the UK, the site known as Priestclose Wood an ASNW in the UK was selected for multispectral UAV monitoring of native and invasive tree species. Priestclose Wood is a 19.35-hectare woodland area near the centre of Prudhoe, of which 15.26 hectares are managed as a Local Nature Reserve by the Northumberland Wildlife Trust (NWT), and of which 17.44 hectares are classified as ASNW (Priestclose Wood). (Figure 12). Silver birch (*Betula pendula*), pedunculate and sessile oak (*Quercus robur* and *Quercus petraea*), rowan (*Sorbus aucuparia*), holly (*Ilex aquifolium*), willow elm (*Ulmus glabra*), sycamore (*Acer pseudoplatanus*), and some ash (*Fraxinus excelsior*) are also present in the wood. When evaluated with the ground vegetation, the woodland most closely resembles the British National Vegetation Classification (NVC) community W10 (*Quercus robur* - *Pteridium aquilinum* - *Rubus fruticosus* woodland). Additionally, there is a section of wet woodland on the northern boundary. While rhododendron bushes predominate at ground level, oak trees at canopy level occupy the eastern edge of the woods, which is adjacent to Prudhoe Hall Drive. The ancient woodland's southern boundary, which borders the Cottier Grange housing development, was planted in the 1890s. It includes Norway maple (*Acer platanoides*), Scots (*Pinus sylvestris*), and Austrian Pine (*Pinus nigra*) trees. Along the edges of the woods, sycamore (*Acer pseudoplatanus*) is also widely distributed. A portion of the woodland was burned in the 1930s, and the oak trees that were damaged were removed and taken to a nearby sawmill. Additionally, during the 1970s, trees were cut down for firewood, which led to additional areas of natural regeneration. Standing dead tree patches in the woodland's centre can serve as an important habitat for biodiversity and as a space for new tree growth (Priestclose Wood). Additionally, it is crucial for the survival of this ASNW to recognise and confirm the native tree species that are present in the woodland, especially the oak and silver birch trees, as they contribute to the local biosphere by serving as a place to live. Also, they are a source of

food, and a component of the localised water cycle that allows various native tree species to survive and maintain a healthy ecosystem (Bargali, 2015). As a result, and in accordance with ASNW, native trees must be planted correctly.






-  Priority Habitat Inventory – Deciduous Woodland (England)
-  Ancient and Semi-Natural Woodland
-  Local Nature Reserves (England)

Figure 12: Priestclose Wood's location and identifier boundaries.

Two distinct multispectral UAV data manipulation methods were used to identify native and invasive tree species in Priestclose woodland using MATLAB software, the outcomes of each approach will be discussed in two separate chapters.

Chapter 5: Multispectral UAV monitoring of native tree species in the central boundary of Priestclose ASNW

The central boundary of the Priestclose ASNW consists native trees species of oak and Silver birch along with some amount of dead trees. Hence, the multispectral UAV approach for identifying native tree species and dead trees in Priestclose Woodland's central boundary will be covered in Chapter 5.

5.1 Aims and Objectives

5.1.1 Aims

The aim of this chapter is to identify native tree species of oak and silver birch in the central boundary of the Priestclose ASNW using PCA derived vegetation indices and image classification method.

5.1.2 Objectives

1. To perform PCA classification on multispectral images to help derive a suitable NDSI which allows in discrimination of native species richness.
2. To segment the discriminated native tree species by thresholding method
3. To compare the performance of NDVI and NDSI algorithms in discriminating the native species.
4. To investigate woodland coverage at canopy level with respect to the overall health of the woodland and contribute to the field of precision woodland management.

5.2 Experimental Design

5.2.1 Sampling site

From the total area of the Priestclose Woodland (19.35 hectares), approximately 11.98 hectares of the central area of the woodland consist of the native tree species of oak and silver birch with an understory of holly and hazel. Hence, the chosen area from the central region for data manipulation has a rectangular border drawn which approximates to 9.25 hectares (Figure 13).

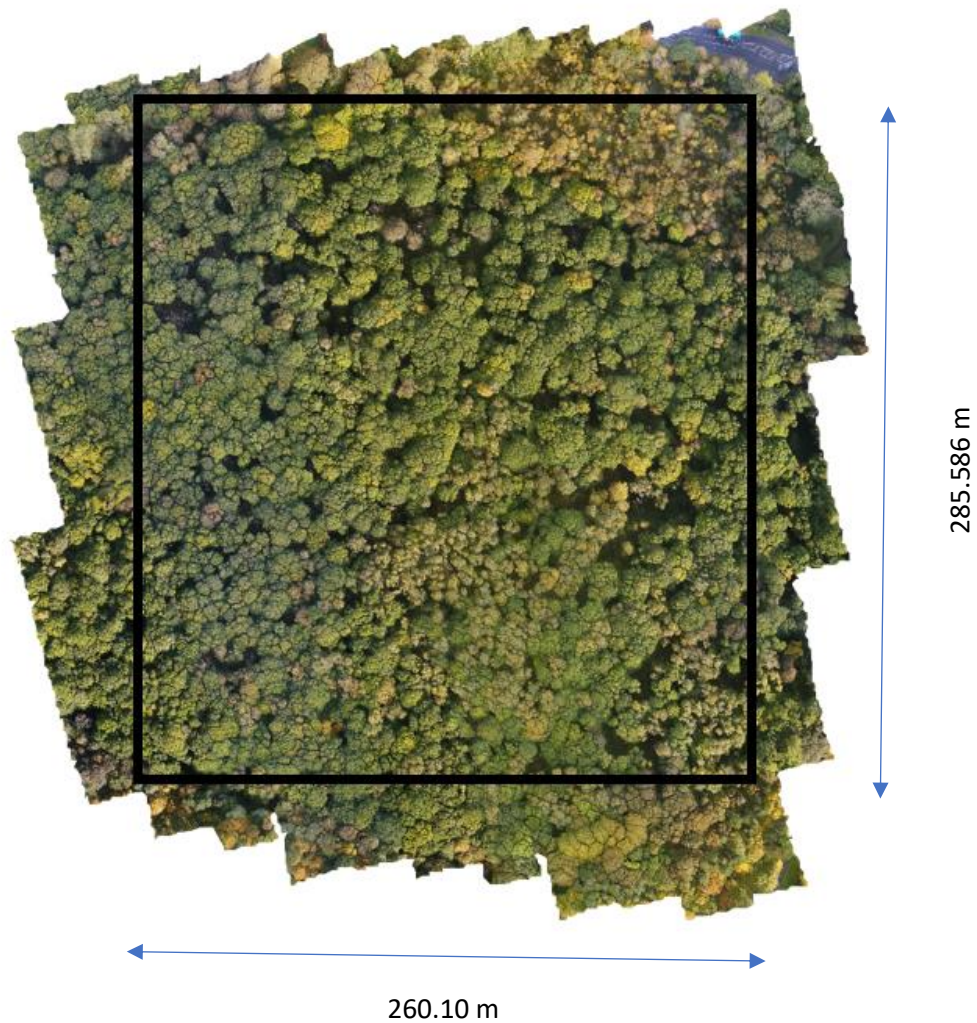


Figure 13: an image of the woodland that was stitched in visible mode and included the actual area (insert) that was being analysed.

5.2.2 Data collection by UAV

The Northumberland Wildlife Trust Mr. Geoff Dobbins offered permission to fly the multispectral UAV (DJI Phantom 4, North Shields, UK) over the Priestclose wood to collect data. The multispectral image camera uses a five-camera array to capture images in the blue (450–16 nm), green (560–16 nm), red (650–16 nm), red edge (760–16 nm), and near infrared (840–26 nm) spectra. A sixth camera can also capture live images in RGB (visible) mode. A 3-axis gimbal stabilises all cameras. Data was captured in hover & capture mode, with the camera always perpendicular to the ground.

Averaging 100 metres in height, the UAV travelled at a speed of 5.0 m/s. With a resolution of 5.3 cm/px, a front overlap ratio of 75%, a side overlap ratio of 60%, and a course angle of 90°, all flights were captured. A hand-held anemometer (Benetech® GM816, available on Amazon UK) was used to record the wind speed and direction, as well as the UAV pilot's anecdotal observation of cloud coverage. These specific weather conditions were identified with specific dated data. Flight information was gathered between the hours of 10:30 and 17:10 from September 17 through October 14, 2020, with an average flight time of about 22 minutes. Per flight, 1076 images were collected in total.

5.2.3 Field Data Collection

Field data was obtained at ground level over the period of late 2020 / early 2021. Photographic documentation of the typical canopy coverage (Figure 14(a)), the visual identification of dead trees (Figure 14(b)), and other low level plant growth, such as holly bushes (Figure 14(c)), were collected from the field. On the weekend of January 30 and 31, 2021, field sampling was conducted to count the number of oak and silver birch trees (based on their bark) that were present at canopy level as well as the number of dead trees that were present above and below canopy level. Two independent tree surveyors covered a 10,800 m² field site on foot. Afterward, on February 16, 2021, two independent tree surveyors used a hand-held clinometer and basic algebra to determine the average tree height of silver birch and oak trees in the field survey area.

(a)



(b)



(c)



Figure 14: Ground level photographs (a) central region (canopy coverage), (b) dead tree, and (c) holly bush. [camera: a Nikon D3500 body with an autofocus Nikon DX VR AF-P Nikkor 18-55 mm 1:3.5-5.6G lens; images (a) taken on October 1, 2020, and (b) and (c) taken on January 31, 2021].

5.2.4 UAV Photogrammetric Processing

Using Agisoft Metashape Professional (64 bit) software version 1.7.1 (Agisoft LLC, St. Petersburg, Russia), an orthomosaic image was created from the multispectral UAV's captured images. The following Agisoft settings were used for photogrammetric processing. The individual images were aligned with medium accuracy setting (key point limit: 40,000 and tie point limit: 4,000). After that, a dense point cloud was created using aggressive depth filtering and low-quality scanning. The next step was to create a mesh model with the following settings: surface type: height field, source data: sparse cloud, polygon count: high, advanced interpolation: enabled, and calculated vertex colours: checked. The orthomosaic image was then constructed and saved as a tif file. The programme offers an automated image processing cycle to align numerous separate images that can be stitched together to build an orthomosaic image, also known as an aerial image. Figure 13 indicates the area used for UAV data analysis.

5.2.5 Image processing and Data Analysis

MATLAB v.R2020b (MathWorks Inc, USA), a programming language software, was used to perform additional image processing and the implementation of algorithms, such as PCA and spectral indices, on the multispectral UAV images from Priestclose Wood (MATLAB code in Appendix 1.1). Figure 15 summarises the workflow for processing data from Priestclose Wood.

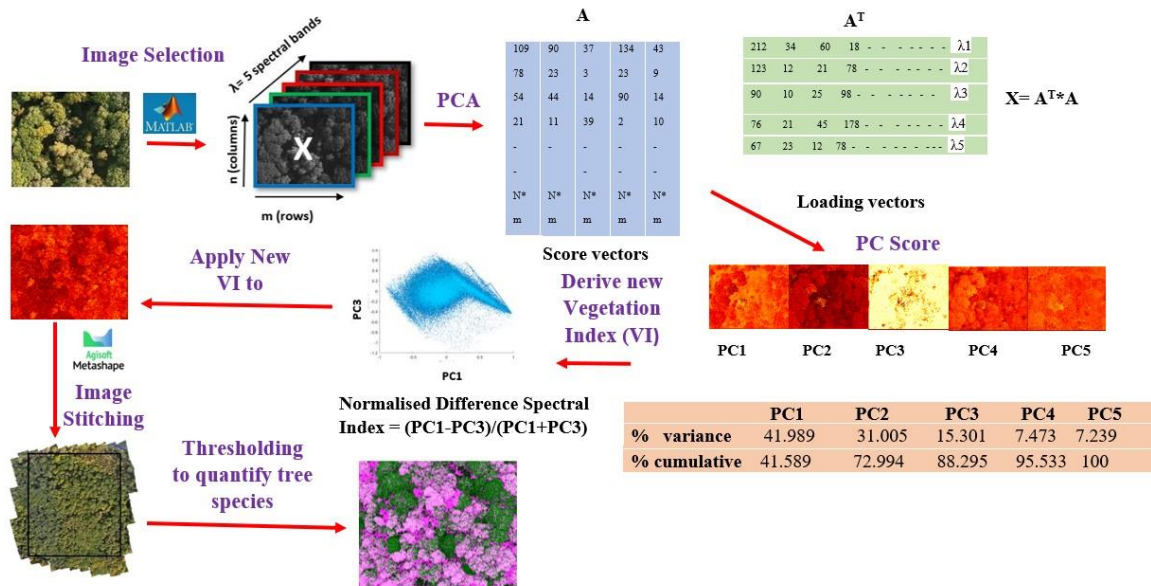


Figure 15: Data processing workflow of identifying native trees in central area of Priestclose Wood.

5.2.6 Normalised Difference Vegetation Index (NDVI)

NDVI is a mathematical formula to determine the growth of vegetation in a particular area and more specifically the health of the vegetation (Santo, 2012). The NDVI is calculated by the difference between reflected energy in the NIR and red spectral bands normalised by the sum of both NIR and red spectral bands using the equation below:

$$NDVI = \frac{(NIR-Red)}{(NIR+Red)} \quad (1)$$

The NDVI values range between $-1 + 1$. A pixel value close to 1, indicates a large increase in the reflectance energy in the NIR spectral band and low reflectance energy in the red spectral band; this corresponds to a greater spectral response of vegetation development and density (Xue and Su, 2017). Values closer to -1 have a greater reflectance energy in the red spectral band than the NIR spectral band and are considered as water saturated areas (Xue and Su, 2017). Ideally, NDVI might generate different spectral index values for different types of tree species as the reflectance energy between NIR and red spectral band might be different for different types of tree species in an area. Therefore, NDVI images were generated by MATLAB using the red and NIR multispectral images to test its effectiveness in classification of native tree species in Priestclose Wood.

5.2.7 Mathematical background of PCA in multispectral images

An individual pixel of an image has a vector x whose dimension is equal to the number of spectral bands N . In this case it is multispectral images composed by 5 spectral bands of red, green, blue, red edge and NIR.

$$x = [x_1 + x_2 + x_3 \dots x_N] \quad (2)$$

An image from each spectral band is a matrix X with m rows and n columns composing $M = m \times n$ pixels, namely $i = 1, \dots, M$. The mean vector μ of an image from each spectral band can therefore be denoted as:

$$\mu = \frac{1}{M} \sum_{i=1}^M [x_1 \ x_2 \ x_3 \ \dots \ x_N] \quad (3)$$

Hence, the covariance matrix Cx can be denoted as:

$$Cx = \frac{1}{M} \sum_{i=1}^M (x_i - \mu_i)(x_i - \mu_i)^T \quad (4)$$

Where T is the superscript the transpose matrix.

The covariance matrix can take the form of:

$$Cx = ADA^T \quad (5)$$

Where D is the diagonal matrix composed of the eigenvalues $\lambda = (\lambda_1, \lambda_2 \dots \lambda_n)$ of the covariance matrix, and A is an orthogonal matrix composed of eigenvectors in the form of:

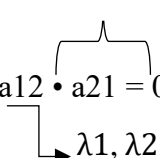
$$A = \begin{pmatrix} a_{11} \\ \vdots \\ a_{nn} \end{pmatrix}^T = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{nn} \end{pmatrix} \begin{matrix} \leftarrow 1^{\text{st}} \text{ eigenvector} \\ \\ \\ \leftarrow n^{\text{th}} \text{ eigenvector} \end{matrix} \quad (6)$$

And,

$$D = |C_x - \lambda \cdot \mathbf{I}| = 0 \quad (7)$$

$$D = \begin{pmatrix} [a_{11} - \lambda_1] & a_{12} & a_{13} \\ a_{21} & [a_{22} - \lambda_2] & a_{23} \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ a_{n1} & a_{n2} & [a_{nn} - \lambda_k] \end{pmatrix} = 0 \quad (8)$$

$$D = (a_{11} - \lambda_1) \cdot (a_{22} - \lambda_2) - a_{12} \cdot a_{21} = 0 \quad (9)$$

\mathbf{I}


Equation 7 can be used to derive eigenvalues λ where \mathbf{I} is the identity matrix. The eigenvalues can be used to further derive the percentage of total variance explained by each principal component (PC) by computing the ratio of each eigenvalue to the total sum of all the eigenvalues. Therefore, eigenvalues of λ_1 and λ_2 are eigenvalues of PC1 and PC2, respectively. The first PC contains usually about 40-50% of the total variance implying the proportion of original information mainly retains in PC1 (Chitroub, 2001). On the other hand, A is the matrix of eigenvectors C_x . Eigenvectors indicate the proportion that each input spectral band contributes to each individual PC. Hence, the eigenvector with the highest eigenvalue contributes to PC1 and results in most of the variation in an image. Likewise, PC2 has the second larger variation to PC1 and so on. Therefore, usually the first 3-4 PCs are sufficient at describing the entire dataset. Finally, a transformed data set (p_i) is created using the eigenvector matrix A to form the new PC images depending on the new variables:

$$p_i = A^T (x_i - \mu_i) \quad (10)$$

$$\begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_i \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{nn} \end{pmatrix} * \begin{pmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \\ \vdots \\ x_i - \mu_i \end{pmatrix} \quad (11)$$

For example, p_1 is $a_{11} (x_1 - \mu_1) + a_{12} (x_2 - \mu_2)$, and p_2 is $a_{21} (x_1 - \mu_1) + a_{22} (x_2 - \mu_2)$.

The terms p_1 and p_2 represent PC1 and PC2, respectively. Thus, the images composed from PC1 contain most information residing in the original multispectral images as the number of spectral bands and the amount of image noise are reduced retaining the most important information for data analysis (Pal and Neifeld, 2003).

5.2.8 PCA applied to data on Priestclose Wood

To categorise the native and dead tree species in the woodland, PCA was used, and the results were compared to NDVI. The set of images displaying dead trees (Figure 16(a)) and native tree species (Figure 16(b)) were selected from the image data set of the woodland using RGB (visible) images.

(a)



(b)

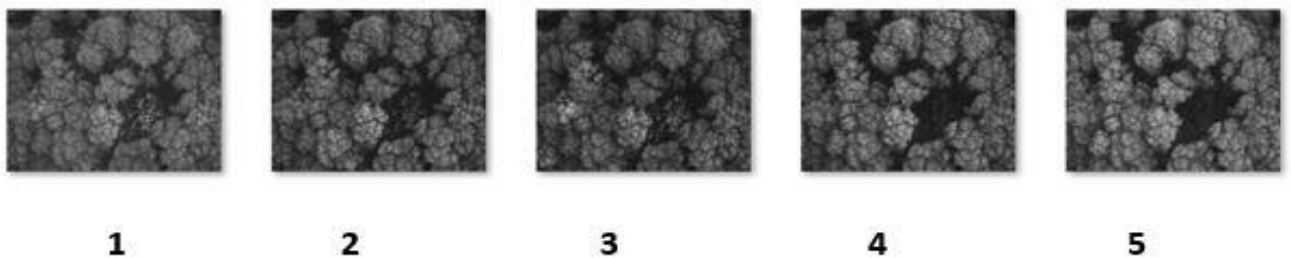


Figure 16: RGB images from the data set of the central area of Priestclose Wood (a) Dead trees with a canopy coverage of oak trees (b) Oak and silver birch trees.

Visual observation of the images from the data set was crucial. All the images, from the data set, do not contain dead trees and some images do not have both silver birch and oak trees; hence by randomly selecting an image from the data set would not serve the purpose of this study. These methods led to the selection of two multispectral image data sets that represented dead trees (Figure 17(a)) and silver birch and oak trees (Figure 17(b)) respectively from 14 October 2020 data set. PCA was performed using MATLAB on the red, green, blue, red edge, and NIR multispectral images (Figure 17a and 17b) to generate the eigenvectors in principle component (PC) images and % variance (Table 2) to extract features for distinguishing dead trees and native trees. After PCA, the % variance obtained of combining PC1, PC2 and PC3 images for the dead tree was 86.0 (Table 2(a)), and for oak and silver birch trees, was 88.3 (Table 2(b)). This highlights most of the information about the images retain within these principal components. Additionally, a new spectral index as below can be created using the eigenvectors contained in these PC images to identify different tree species.

$$\text{Normalised Difference Spectral Index (NDSI)} = (x_1 - x_2) / (x_1 + x_2) \quad (12)$$

(a)



(b)



Figure 17: Multispectral images used from the central area of Priestclose Wood to perform PCA (a) multispectral image data set representing dead trees (b) multispectral image data set representing oak and silver birch trees [1 = Blue, 2 = Green, 3 = Red, 4 = red edge and 5 = NIR].

Table 2: Percentage Variance of PC1 – PC5 resulting from PCA applied to multispectral images to classify (a) dead trees, and (b) oak and silver birch trees.

(a)	PC1	PC2	PC3	PC4	PC5
% variance	41.470	30.758	13.777	7.473	6.522
% cumulative	41.470	72.228	86.005	93.478	100.000
(b)					
% variance	41.989	31.005	15.301	7.473	7.239
% cumulative	41.989	72.994	88.295	95.533	100.000

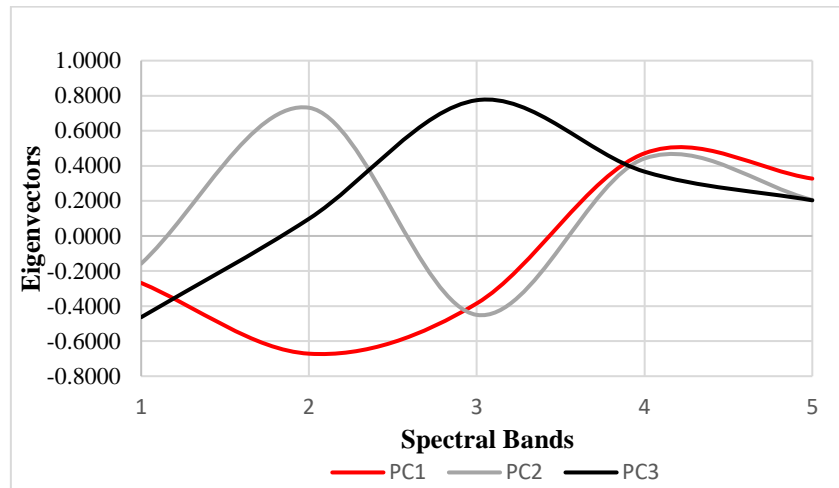
NDSI is the difference between the two spectral bands (x1 and x2) normalised by their sum; it can be used to differentiate scene elements and enhance spectral features that are not visible to human eye. Additionally, PCA offers the choice to select and exclude the appropriate spectral bands. To distinguish the dead trees from oak and silver birch trees, the proper PCs were used to derive an NDSI (Figure 18). In both dead trees and oak and silver birch trees, PC1 and PC3 exhibit strikingly different spectral responses at various spectral bands, as shown in Figure 18. Green and red spectral bands in PC1 and PC3 showed a difference in spectral response for dead trees (Figure 18a). For oak and silver birch trees (Figure 18b) a difference in spectral response was observed in all five spectral bands, with a pronounced difference observed in the blue, red and NIR spectral bands in PC1. However, comparing the spectral bands in PC1 and PC2 for dead, silver birch and oak trees only, a difference was observed in the blue and green spectral bands which is insufficient at extracting useful features as the spectral response of vegetation mainly uses spectral bands from red to NIR. Hence, a new NDSI was derived using PC1 and PC3 (Figure 18 c,e,h. and j), as below separately to identify dead trees using the multispectral images in Figure 17a and oak and silver birch using the multispectral images in Figure 17b:

$$\text{Spectral index} = (\text{PC1}-\text{PC3})/(\text{PC1}+\text{PC3}) \quad (13)$$

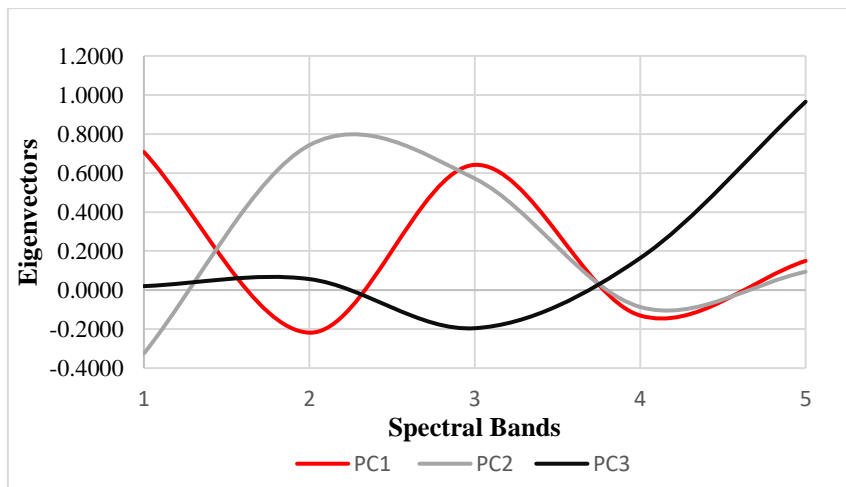
Further, to observe the variability a spectral index using PC1 and PC2 was derived.

This new spectral index was then applied on Priestclose Wood data over the period of 17 September – 14 October 2020 to generate multispectral images which was used for further data analysis.

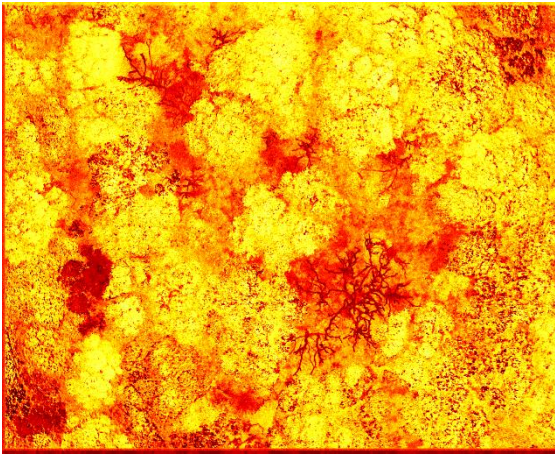
(a)



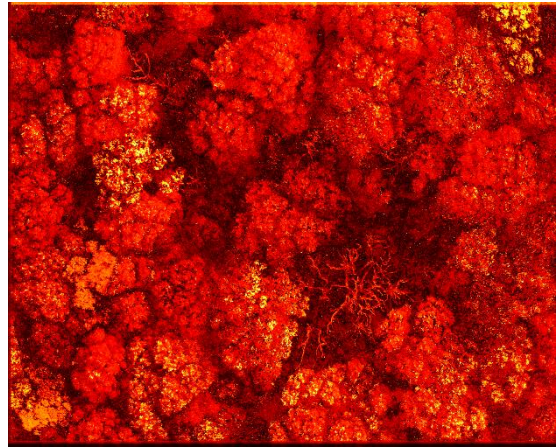
(b)



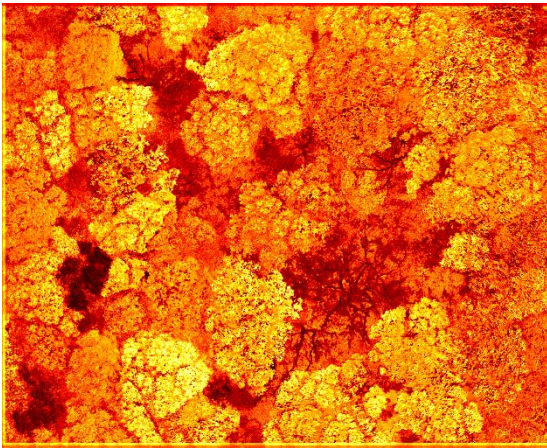
(c)



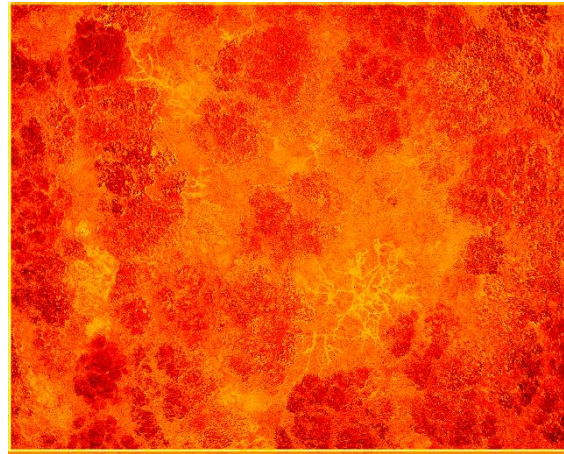
(d)



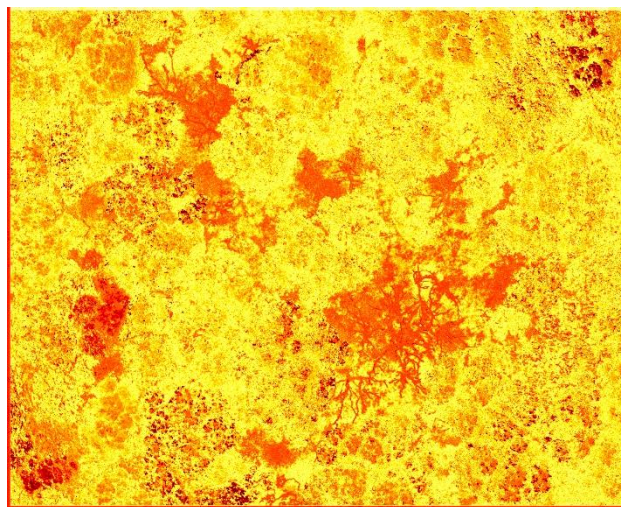
(e)



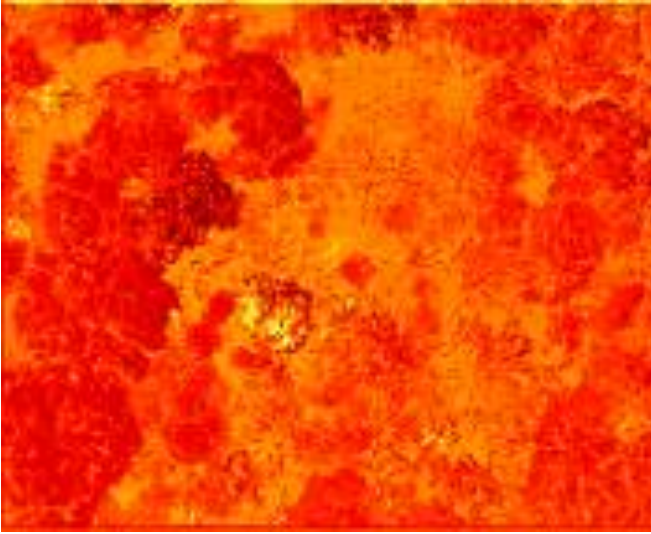
(f)



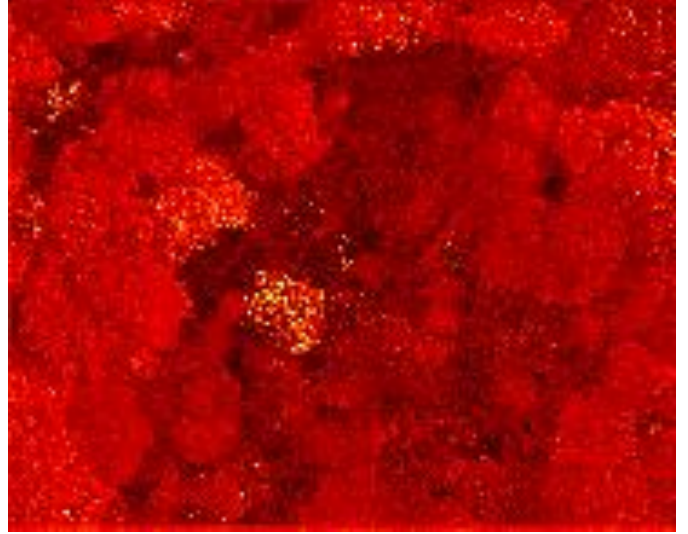
(g)



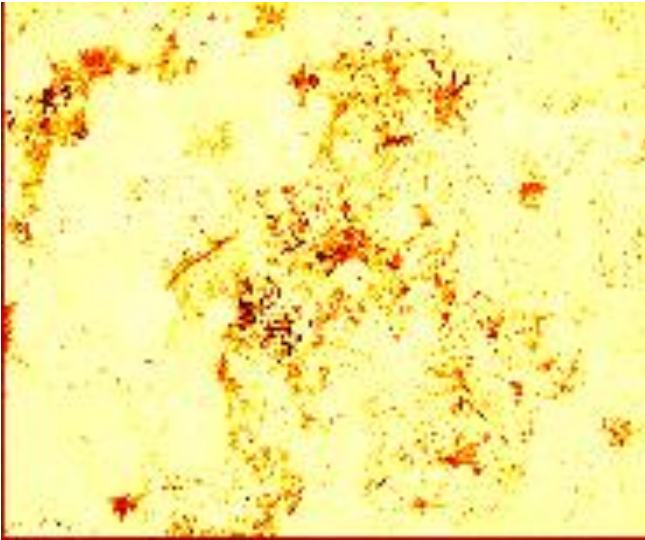
(h)



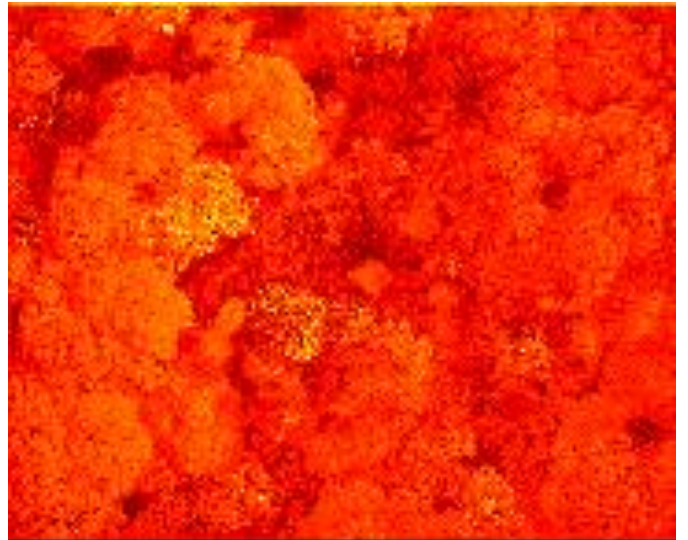
(i)



(j)



(k)



(l)

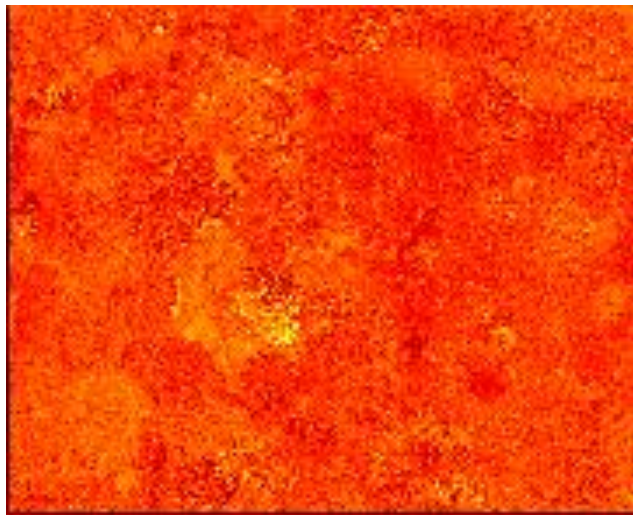


Figure 18: Eigenvectors indicating the proportion of each spectral band contributing to form each individual PC1, PC2 and PC3 image (a) for dead trees (b) oak and silver birch trees. [1 = Blue, 2 = Green, 3 = Red, 4 = DR and 5 = NIR]

PC for dead trees (c) PC1 (d) PC2 (e) PC3 (f) PC4 (g) PC5; PC for Oak and Silver birch trees (h) PC1 (i) PC2 (j) PC3 (k) PC4 (l) PC5

5.2.9 Calculation of pixels of native tree species and dead trees

An orthomosaic image was created by stitching together images derived from the new NDSI = $(PC1-PC3)/(PC1+PC3)$ using Agisoft Metashape Professional software. The orthomosaic image of Priestclose Wood in Figure 13(b) shows native oak and silver birch trees as well as dead trees. The different spectral responses from the native tree species and dead trees were highlighted by the new NDSI, which resulted in different pixel values being assigned to oak, silver birch, and dead trees. The total pixel numbers for each area representing oak tree, silver birch tree and dead tree were counted using MATLAB. A grayscale image was created by stitching together the individual new NDSI images (Appendix 1.2). By analysing this grayscale image, specific threshold values for oak (0.35 to 0.38), silver birch (0.44 to 0.48), and dead trees could be ascertained (0.20). Due to the trees' normal phenological cycle, it is important to note that the threshold values for oak and silver birch trees changed slightly over the course of the 5 different days. To represent the area of woodland under investigation, the total number of pixels in each area was multiplied by the resolution of the drone images that were taken ($0.053 \text{ m/pixel} * 0.053 \text{ m/pixel}$). By dividing each area by the total area of the woodland, the percentage coverage of dead trees and the canopy coverage of oak and silver birch trees were calculated. The pixel calculation for native tree species and dead trees was examined over the course of five different days to evaluate the accuracy of the data.

5.3 Results and Discussion

5.3.1 Comparing the performance of classification methods by PCA

The results from the study demonstrates the performance of classification methods of NDVI, PCA applied NDVI, and newly derived NDSI from PCA to identify the native tree species of oak and silver birch trees along with dead trees from Priestclose ASNW. Initially, oak, and silver birch trees were identified by classification methods of NDVI and the PCA applied

NDVI method. The NDVI images were composed by using the original NIR and red multispectral UAV images. Whereas PCA applied NDVI images were composed by using NIR and red eigen vectors from PC1 (Table 2). PC1 consisted of 41% of the total variance implying the proportion of original information mainly retained in PC1. The NDVI images after the application of the eigen vectors from PC1 (Figure 19(f)) showed a better classification of both oak and silver birch trees than NDVI without weightings from PC1 eigen vectors (Figure 19(e)), as the empty areas surrounded by the silver birch trees were not clearly identified in comparison to the PCA applied NDVI image (Figure 19(f)). This demonstrates the significance of the use of PCA classification as the noise from the original multispectral images have been reduced retaining the most important information in each spectral band. While the classification using NDVI with PCA showed possibilities in identifying oak and silver birch trees, its main use has been in its application to represent the vegetation cover between the reflectance energy in red and NIR spectral band (Xu and Su, 2017).

In addition, the NDSI was derived using PC1 and PC3 (equation 13) to classify the native tree species. This was as the spectral bands of blue, green, red, red edge and NIR showed a huge difference in spectral response between PC1 and PC3 (Figure 18b). Also, NDSI derived using PC1 and PC3 was more suitable at classifying both dead trees and native tree species in comparison to using PC1 and PC2 (Figure 20). In PC1 and PC2 (Fig 18a and 18b) there is an observable difference in spectral bands between green and red, and the remaining spectral bands consist of negative eigenvector components implying that these spectral bands do not hold significant information to help discriminate native tree species or dead trees.

Also, combining the information from additional spectral bands shows a greater sensitivity at differentiating the variation of chlorophyll in different tree species (Figure 19(b) and 19(c)). This new approach has enabled the identification of the oak trees and the individual standing silver birch trees (Figure 19(b) and 19(c)). This is due to the enhanced spectral resolution observed by inclusion of the additional spectral bands in new NDSI. This effect has also been reported by (Heikkinen et al., 2010), who showed that the addition of the red edge band to the existing VIS-NIR band sensors improved the classification of single trees i.e., spruce, pine and silver birch. Additional benefits of the new NDSI are that it has allowed the discrimination of the oak (green region) and silver birch tree (pink region) in Figure 19(c); this was difficult to observe in the original RGB image (Figure 19(a)). While the new NDSI using PC1 and PC3 (Figure 18b) could identify the oak and silver birch trees it was not a particularly useful index to identify the dead trees. Therefore, a separate NDSI was derived by using PC1 and PC3

(Figure 18a) to specifically identify dead trees (Figure 21). The newly derived NDSI for identification of dead trees (Figure 21(b)) enhanced the identification of individual branches of the dead trees in comparison to NDVI and NDVI applied PCA Figure 21(c) & 21(d).

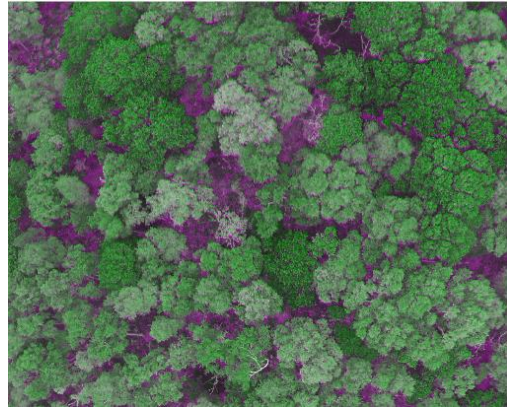
Along with the classification methods of NDVI and PCA, a further useful trait for species classification of the phenology was considered. Phenology is defined as the study of changes in the leaf colour of tree species in deciduous temperate forests in relation to seasonal changes such as in autumn; the phenology can also vary amongst different tree species (Gärtner et al., 2016). Due to leaf senescence the decomposition of chlorophyll pigments in different tree species are faster in comparison to anthocyanins (Fassnacht et al., 2016). As the images in this study were captured between the 17 September and 14 October 2020, i.e., the onset of autumn, the application of the new NDSI was investigated further. The importance of considering the phenology is exemplified by comparing the data (derived from the data set from 14 October 2020) (Figure 19(c)) with the data set from the 17 September 2020 (Figure 19(d)) where the 14 October 2020 (Figure 19(c)) allows better classification of trees. Further, the spectral response of chlorophyll and carotenoid composition present in trees during phenology indicates a higher spectral response in the blue, red, red edge and NIR spectral bands (Morley et al., 2020). Therefore, the new NDSI is composed using these spectral bands, further confirming that UAV multispectral cameras used in this study provide a significant advantage at classifying trees. Hence, the new NDSI using PC (equation 13) was derived using the multispectral images from 14 October 2020 and applied to the remaining dates as repeats.

The results demonstrate the importance of feature extraction by PCA which in turn allows selection and elimination of spectral bands according to their spectral response to derive a new NDSI for tree species identification. Also, the results highlight the importance of considering the phenological cycle of tree species to aid with their identification using multispectral imagery.

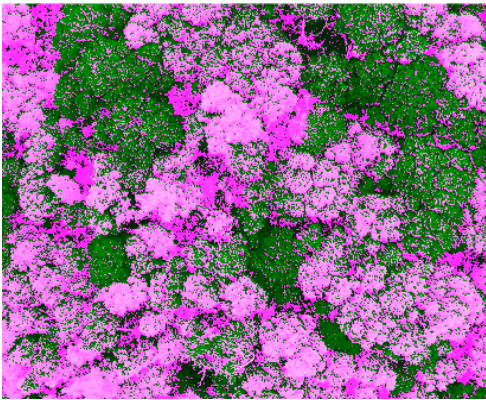
(a)



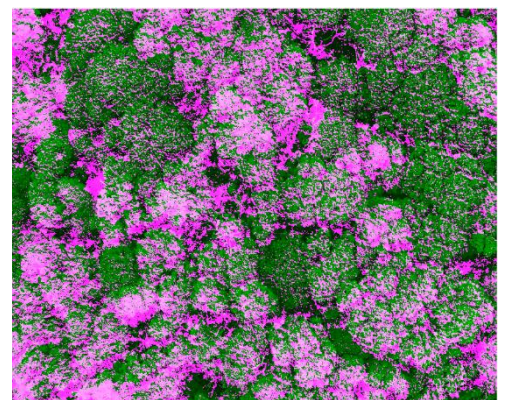
(b)



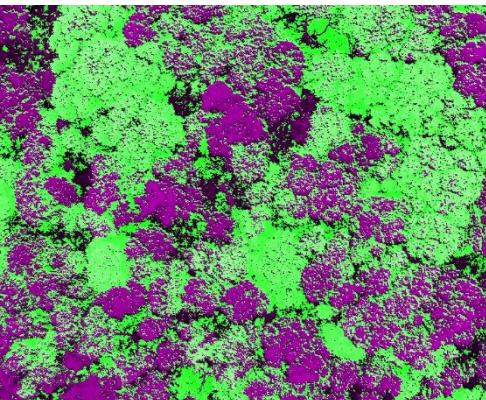
(c)



(d)



(e)



(f)

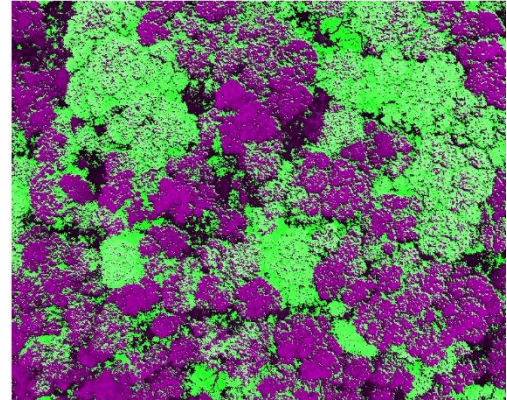


Figure 19: (a) RGB image (b) new NDSI image differentiating oak (dark green) and silver birch trees (light green) (c) Image (14 October 2020) derived from new NDSI thresholding to segment oak (pink) and silver birch trees (green) (d) Image (from 17 September 2020) derived from new NDSI thresholding to segment oak (green) and silver birch trees (pink) (e) NDVI image after thresholding (green as oak trees and purple as silver birch) (f) PCA derived NDVI image after thresholding (green as oak trees and purple as silver birch).

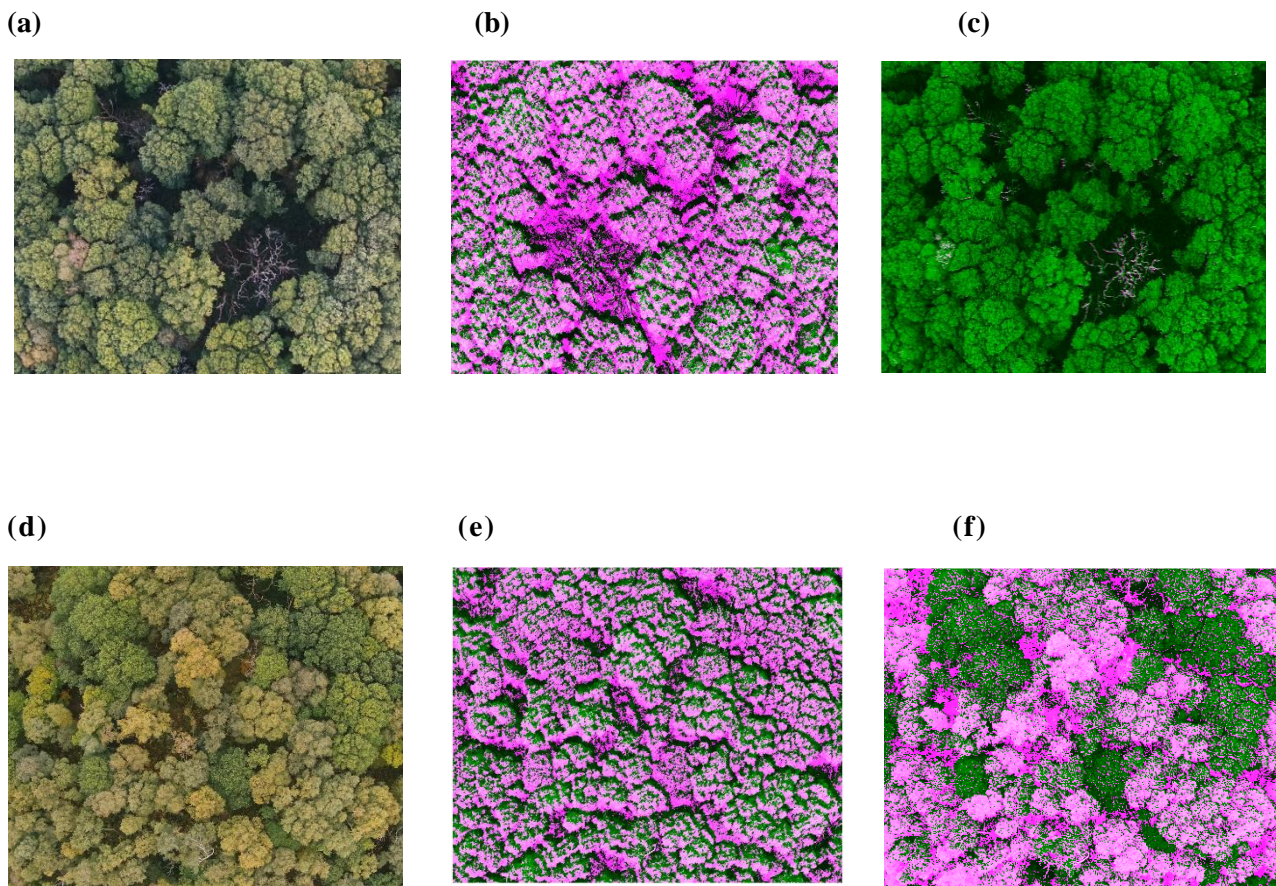


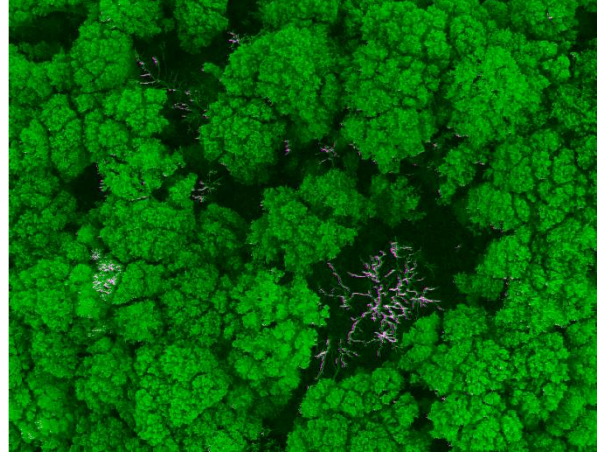
Figure 20: Thresholding Images derived from NDSI (a) RGB image, (b) image derived from $(PC1-PC2)/(PC1+PC2)$ (c) image derived from $(PC1-PC3)/(PC1+PC3)$ for dead trees.

(d) RGB image (e) image derived from $(PC1-PC2)/(PC1+PC2)$ (f) image derived from $(PC1-PC3)/(PC1+PC3)$ for oak and silver birch trees.





(c)



(d)

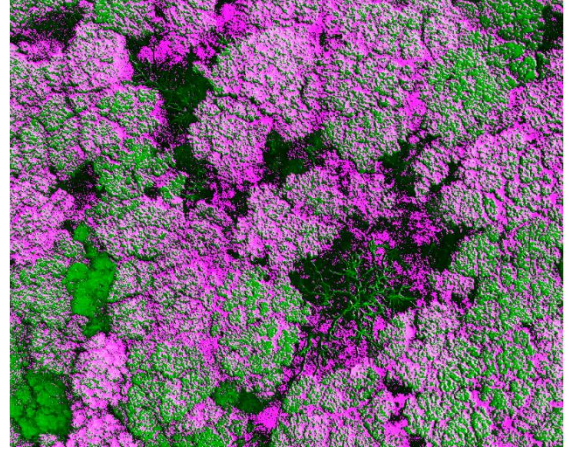
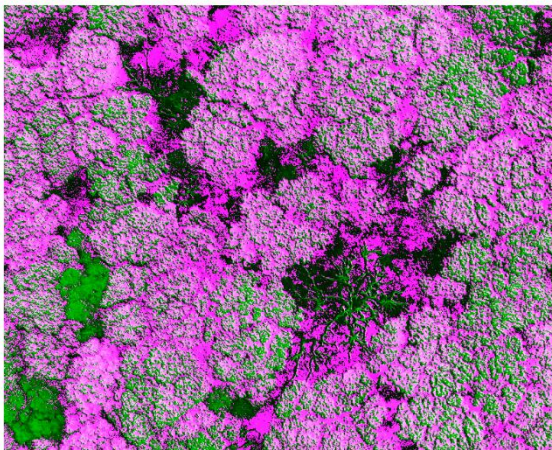


Figure 21: RGB image outlining dead trees. (b) Image derived from new NDSI to differentiate dead trees after thresholding (14 October 2020) (c) NDVI image after thresholding (d) PCA NDVI image after thresholding

5.3.2 Quantitative information obtained by analysis of UAV MSI and field study data

On five separate occasions, the pixels in the orthomosaic images were counted, and percentages of dead trees' coverage and the canopy coverage of oak and silver birch trees were calculated (days). The average results of this quantitative study using multispectral UAV identified that there were $37 \pm 1\%$ silver birch trees and $53 \pm 1\%$ oak trees. Additionally, the combined

percentage of dead trees and empty space was $10.3 \pm 1\%$ empty space and $1 \pm 0.15\%$ dead trees, as identified at canopy level (Table 3). It should be noted that a single UAV run's data capture took place over a $73,934 \pm 1,264 \text{ m}^2$ area in 22 minutes. On the other hand, a thorough ground-level field study was also carried out over a $10,800 \text{ m}^2$ area, and the results showed that 47% of oak trees and 45.6% of silver birch trees were identified by manual counting (Table 3). Additionally, manual data collection was carried out to calculate the percentage of dead trees at canopy level (7.4%); this value is comparable to the data obtained by UAV (10.3%). At ground level, it was observed that a significant portion of the open space at canopy level was made up of smaller dead trees (Figure 14b) and vegetation (such as holly bushes, Figure 14c). Most dead trees found in the field study data were below the canopy level, where UAV data analysis revealed that they were blurry or difficult to distinguish. Previous studies report this occurrence with remote sensing of species below the canopy level (Hernandez-Santin et al., 2019). They stated that the cloud cover, the sun's angle at the time of image acquisition, and the impact of shadows cast by species above canopy level on species below canopy level all had an impact on the data analysis from a UAV platform (Hernandez-Santin et al., 2019). At ground level, it was determined that the 9% of empty spaces reported in this study by the UAV were caused due to holly bushes and small dead trees (Figure 14). Using a hand-held clinometer, the canopy's height was calculated as part of the field study to be between 18 and 22 metres. Two people worked on the ground-level field study for at least eight hours, manually counting, identifying trees, and marking the ground with cordoned-off areas. Both methods noted that oak trees have wider leaves that radiate to branches to form multiple crowns, which contribute significantly to a larger canopy of oak in the woodland (Figure 19(a)).

Table 3: Quantitative information obtained by analysis of UAV MSI and field study data.

UAV flight [#]	From analysed UAV data			Calculated data				Field study data [@]					
	Total woodland area (m ²)	Area of oak tree coverage (m ²)	Area of silverbirch tree coverage (m ²)	% oak trees	% silverbirch trees	% empty space with lower lying canopy	% dead trees at canopy level	% oak trees	% silver birch trees	% dead trees at canopy level		Height of canopy (m) ^{&}	
										Oak	Silver birch	Oak ^{\$}	Silverbirch ^{\$}
1	74,281	39,048	28,222	53	38	8	1.36	47.0	45.6	5.3	2.1	22 ± 3 22 ± 3	18 ± 4 18 ± 3
2	73,896	38,870	26,441	53	36	10	1.49						
3	71,849	38,536	26,335	54	37	9	1.09						
4	74,409	38,549	27,404	52	37	10	1.27						
5	75,236	39,737	27,930	53	37	9	1.24						
Average	73,934	38,948	27,266	53	37	9	1.29						
SD	1,264	492	855	1	1	1	0.15						

Notes

[#] UAV flights: 1. 17 September 2020; 2. 18 September 2020; 3. 28 September 2020; 4. 8 October 2020; 5. 14 October 2020

[@] Field data collected on 30 January 2021 over an area of 10,800 m². The sample size was estimated, at the 95% confidence level, with a 5% confidence interval (\pm margin of error) by assuming every tree could occupy a space of 1 m² (i.e. an assumption that 73,934 trees were present) that the sample size should be 382 trees. In total 453 were manually counted. The field study identified a total of 377 trees as being at canopy level with 7.4% identified as dead trees. Also, an additional 76 trees were identified as dead that were present below canopy level i.e. not visible via the UAV. Additionally, 32 clumps (single or multiple trunk) holly bushes were identified at ground level.

[&] Determined using a hand-held clinometer and calculated using algebra, height was based on two independent people each making repeat measurements (n = 3) on 10 different trees of the same type, around the field survey site. In addition, the reproducibility was assessed by making repeat measurements on the same tree, by two independent people. Mean height of silver birch tree (n = 10) as determined by person 1 was 20.35 \pm 2.0 m and person 2 was 20.01 \pm 2.0 m whereas Mean height of oak tree (n = 10) as determined by person 1 was 18.47 \pm 2.0 m and person 2 was 19.83 \pm 2.1 m. ^{\$} mean \pm SD, based on the results from two independent field workers.

5.4 Conclusion

Although there are currently many studies reporting the use of UAVs to identify and quantify tree species, this study has demonstrated some additional benefits. The results have demonstrated the benefits of a PCA classification approach which allows the selection of additional spectral bands to derive a new NDSI to classify tree species. The greater number of spectral bands increases the probability to identify species having different species-specific spectral signatures. However, the PCA threshold classification method requires subsequent interpretation of data to derive a new and suitable NDSI. The UAV data was compared with a ground level field study that confirmed the precision of the newly developed method. Further research will apply this new approach to investigate the transferability of the developed approach to areas of woodland with multiple tree species,

Chapter 6: Multispectral UAV monitoring of invasive tree species in the southern boundary of Priestclose ASNW

The southern boundary of the Priestclose ASNW consist of Norway maple, Sycamore, Scots Pine and Black Pine invasive tree species. Hence, the multispectral UAV approach for invasive species identification in Priestclose Woodland's southern boundary will be covered in Chapter 6.

6.1 Aims and Objective

6.1.1 Aims

The aim of this chapter is to identify invasive tree species of Scots pine, Black pine, Sycamore and Norway Maple in the southern boundary of the Priestclose ASNW using PCA-derived vegetation indices and image classification method.

6.1.2 Objectives

1. To perform PCA classification on the multispectral images to derive an NDSI algorithm which allows determination of the invasive 20th century species.
2. To segment the discriminated invasive tree species by a k-means clustering method
3. To quantify the invasive species and validate with ground reference.

6.2 Experimental Design

6.2.1 Sampling site

The Priestclose wood's southern boundary (Figure 22) is a planted 20th-century design landscape that was added between the 17th and mid-18th centuries and consist of invasive Scots pine, Black pine, Sycamore, and Norway Maple trees. The area of data analysis is outlined in Figure 22.



Figure 22: Stitched image (insert), in RGB mode of the southern boundary of Priestclose Wood.

6.2.2 Data collection by UAV

The Northumberland Wildlife Trust Mr. Geoff Dobbins offered permission to fly the UAV over the Priestclose wood in order to collect data. A multispectral multirotor UAV (DJI Phantom 4, Leeds, UK) was employed. A 3-axis gimbal was used to stabilise every camera. Data was captured in hover and capture mode, with the camera always perpendicular to the ground. Images were recorded as 16-bit TIF file. The UAV travelled at a speed of 5.0 m/s and had a height of 100 m on average. All flights were captured with 5.3 cm/px resolution, a course angle of 90°, a front overlap ratio of 75%, and a side overlap ratio of 60%. A handheld anemometer (Benetech® GM816, available on Amazon UK) was used to record the wind speed and direction, as well as the anecdotal cloud coverage observations of the UAV pilot. These weather conditions are identified with specific dates in the data.

6.2.3 UAV data analysis: Photogrammetric Processing

The multispectral UAV data was used to produce orthomosaic images for the spectral bands using Agisoft Metashape Professional software. Few steps were performed to implement Agisoft parameters for photogrammetric processing. The initial alignment of each image was done with medium accuracy, with a key point limit of 40,000 and a tie point limit of 4,000. After that, a dense point cloud was created using aggressive depth filtering and low-quality scanning. A mesh model was then formed using the following parameters: surface type: height field, source data: sparse cloud, high polygon count, advanced interpolation enabled, and calculated vertex colours checked. The orthomosaic was then developed and saved as a tiff file.

6.2.4 Field Data Analysis

On March 13, 2021, ground level tree identification and mapping using a portable GPS (Garmin Oregon, 600) were performed. Visual identification was carried out by examining the leaf buds, canopy shape, and tree bark. Black pine, Norway maple, oak, Scots pine, silver birch, and sycamore were the main tree species identified. An independent tree surveyor visited the area on foot while being accompanied by a recorder to conduct the survey. On the orthomosaic image (Figure 22) the GPS coordinates of the trees that were identified in the field study were assigned. ArcGIS Pro v.2.8.0 was used to input the x,y coordinates for each tree location (Esri Inc, West Redlands, CA, USA).

6.2.5 Image processing and Data Analysis

On the multispectral UAV images, additional image processing and implementation of algorithms, including PCA and vegetation indices (VI) and image segmentation by k-means, have been carried out using the programming language MATLAB v.R2020b (MathWorks Inc, USA) (Appendix 1.1). Figure 23 summarises the data processing workflow.

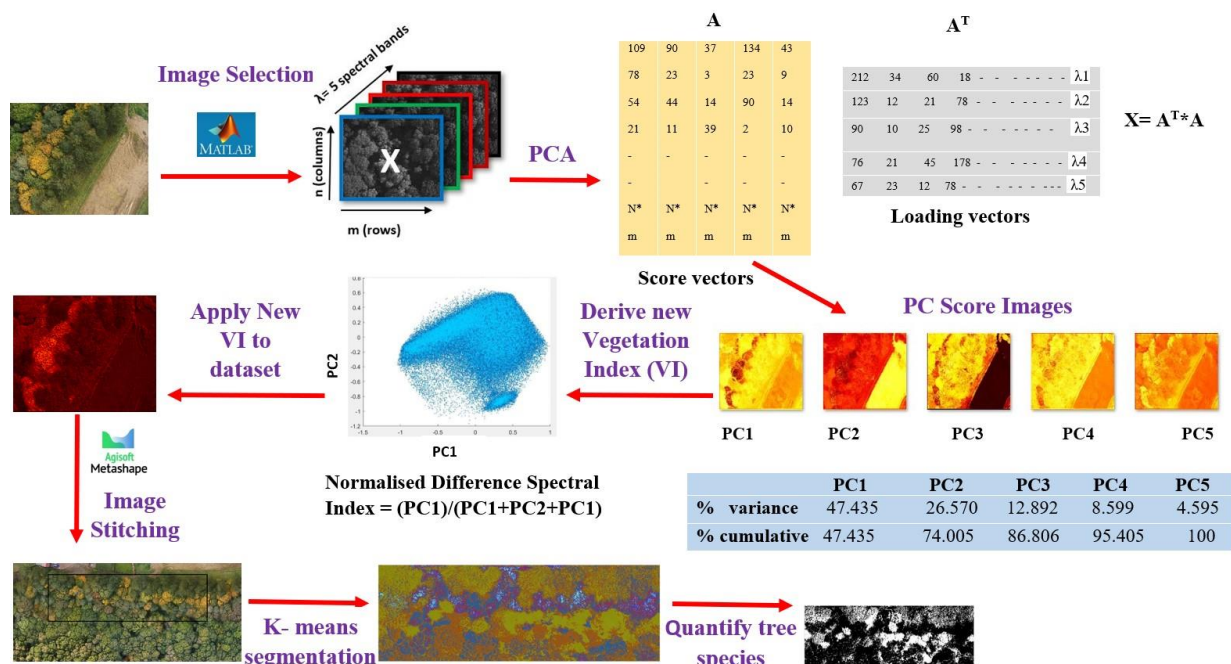


Figure 23: Workflow of processing UAV data from Southern boundary in Priestclose wood.

6.2.6 Principal Component Analysis applied to obtained data.

RGB (visible) images were used to identify the sets of images that display the various tree species from the image data set of the southern boundary (Figure 24). Visual analysis of the data set images was essential since not all the images in the data set contained all the invasive

tree species, and randomly choosing an image from the data set would not have served the objectives of this research. Afterwards, multispectral image data sets depicting the local tree species were chosen (Figure 25(a)). The red, green, blue, red edge, and NIR multispectral images were used, and the PCA was then carried out using MATLAB. The eigenvectors and % variance (Table 4), which represent the features extracted from the multispectral image data set, were produced for this data interrogation. Table 4(a) shows that the variance obtained for pine, Norway maple, and sycamore was 86.81%. The first three principal component (PC) score images retain the most crucial information that can be used to analyse the data efficiently. To distinguish the invasive tree species, the proper PCs were used to derive an NDSI. Figure 26(a) shows that PC1 and PC2 exhibit different spectral responses in the green, red, red edge, and NIR bands, with the red spectral band exhibiting the most pronounced difference. Since the spectral response of vegetation primarily uses the red and NIR spectral bands, these spectral bands would be useful for extracting useful features. As a result, a new NDSI was created using PC1 and PC2 as follows (Figure 26 c and d);

$$\text{Spectral index} = (\text{PC1}) / (\text{PC2} + \text{PC2} + \text{PC1}) \quad (14)$$

The new NDSI (equation 14), however, did not adequately distinguish between Scots and black pine as two distinct species of trees. Hence, to perform PCA to categorise the pine trees, an RGB image and the multispectral data set (Figure 25(b)) were chosen. For the first three PCs, the percent variance obtained was 87.552 (Table 4(b)). In accordance with (Figure 26(b)), PC1, PC2, and PC3 exhibit varying spectral responses in five spectral bands, with the NIR spectral band exhibiting a notable increase in spectral response for all three PCs. As a result, combining the data from the first three PC will improve the spectral response of some trees, which then aids in correctly classifying them. By combining the eigenvectors from the first three PC score images, a new spectral index was generated to extract features to distinguish the Scots pine trees as below combining PC1, PC2 and PC3 (Figure 26 h, i and j):

$$\text{Spectral Index} = (\text{PC1} + \text{PC2} + \text{PC3}) \quad (15)$$

The PC for the new indices in equation 14 and 15 was derived from 14 October 2020 data set (Figure 25) and applied to the remaining date of 19 October 2020.

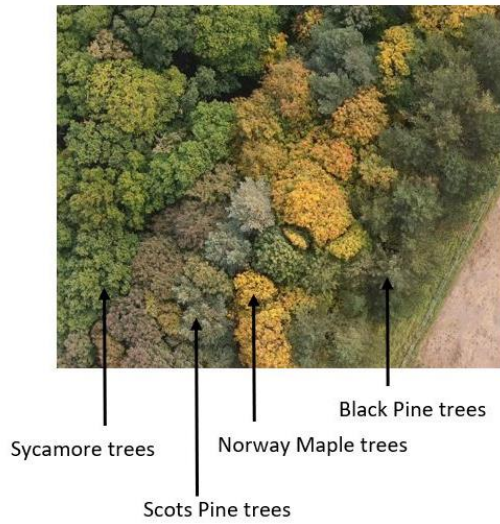
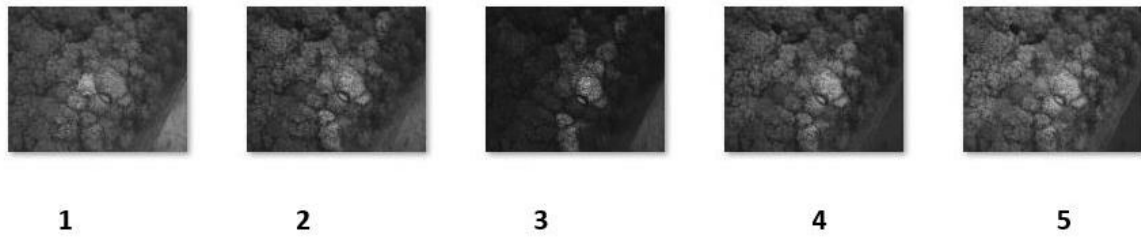


Figure 24: RGB (visible) which represents the invasive tree species: black pine, Norway maple, Scots pine and sycamore.

(a)



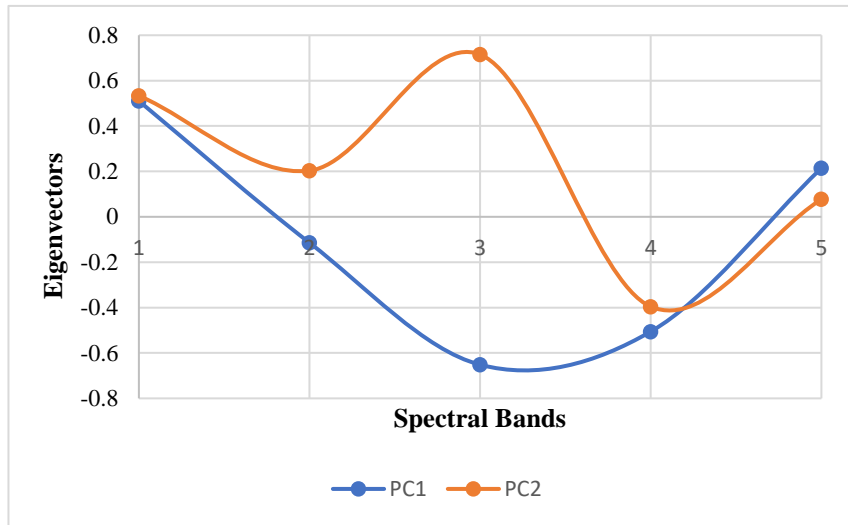
(b)



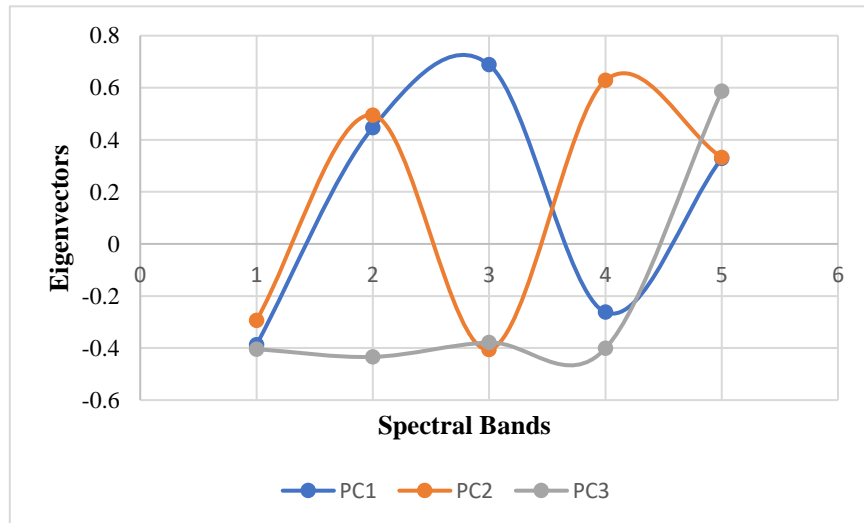
Figure 25: Multispectral image datasets used to perform PCA for (a) pine trees (Scots and black), Norway maple and sycamore trees (b) Scots Pine trees only.

[1 = Blue, 2 = Green, 3 = Red, 4 = Red Edge and 5 = NIR].

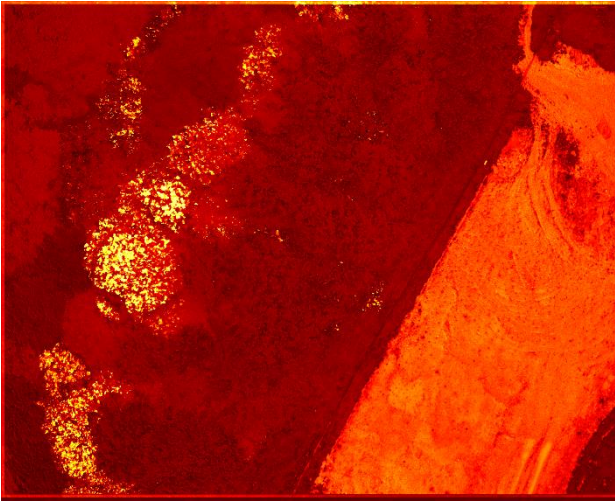
(a)



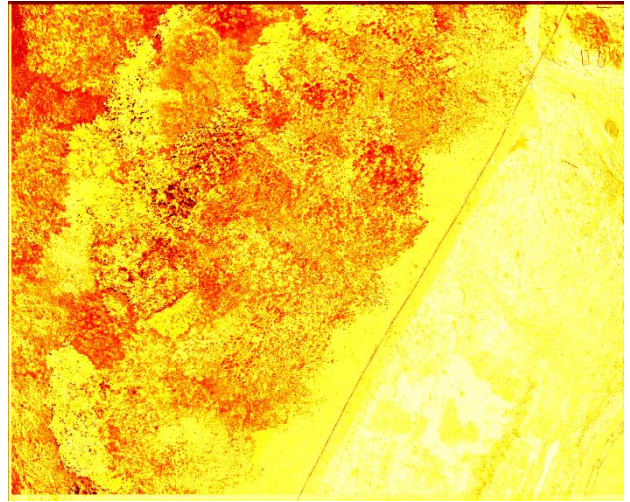
(b)



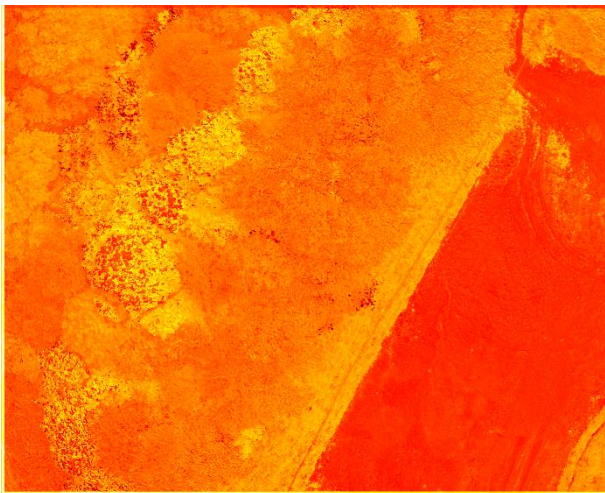
(c)



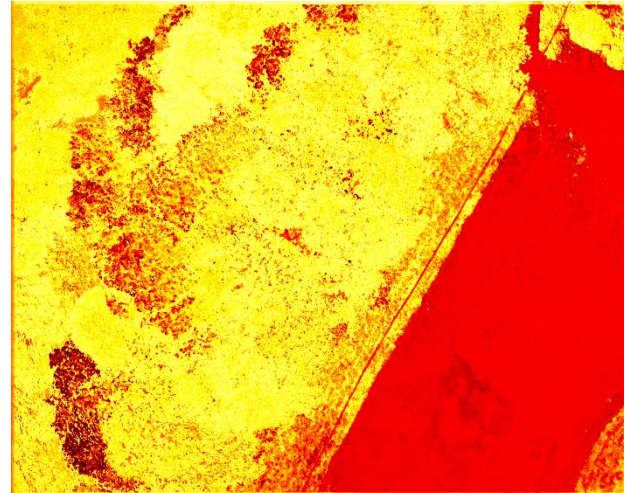
(d)



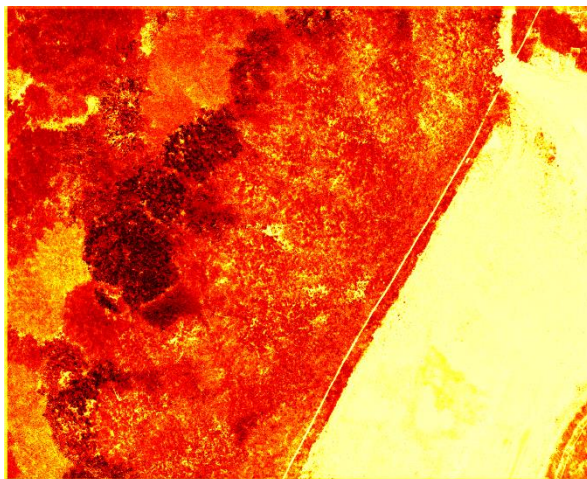
(e)



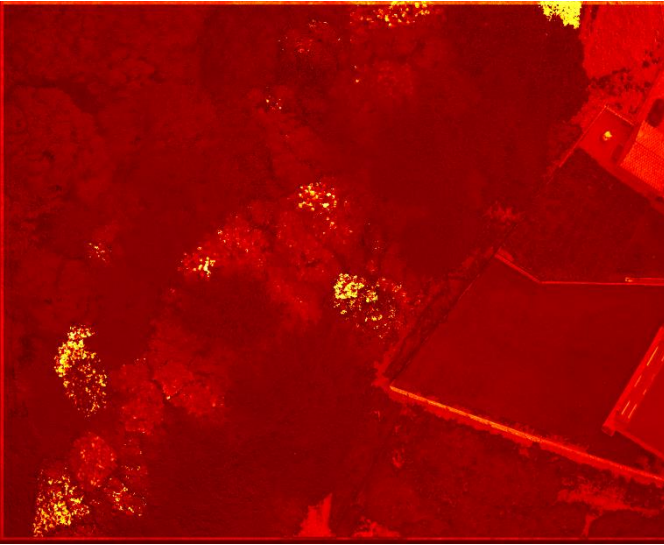
(f)



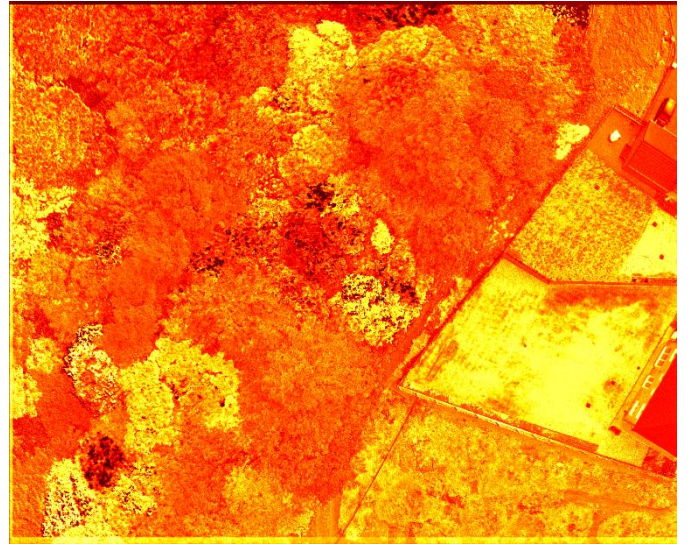
(g)



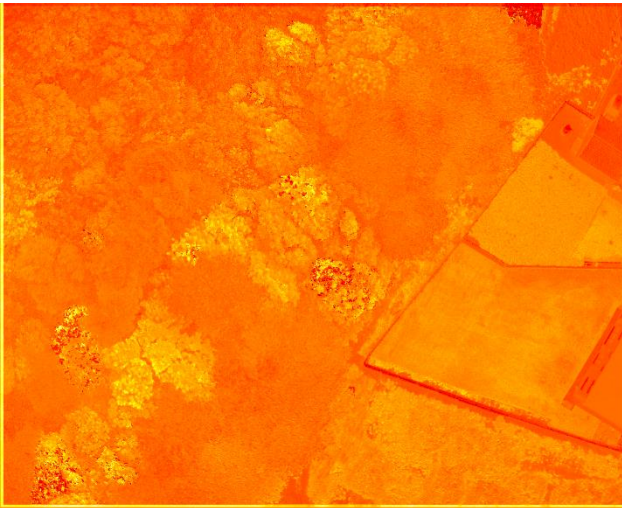
(h)



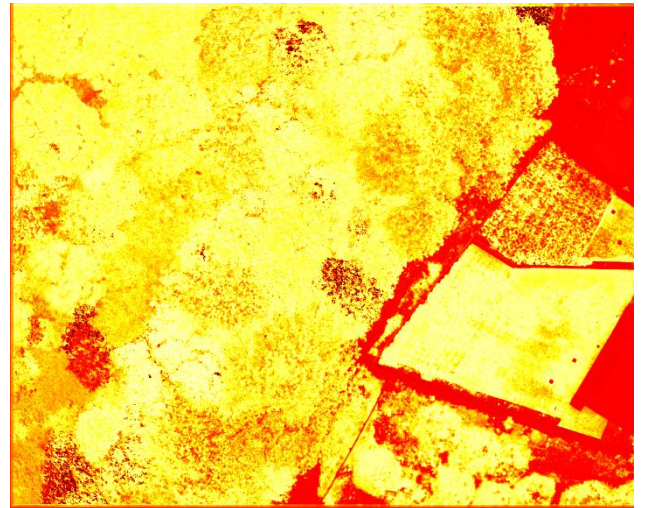
(i)



(j)



(k)



(l)

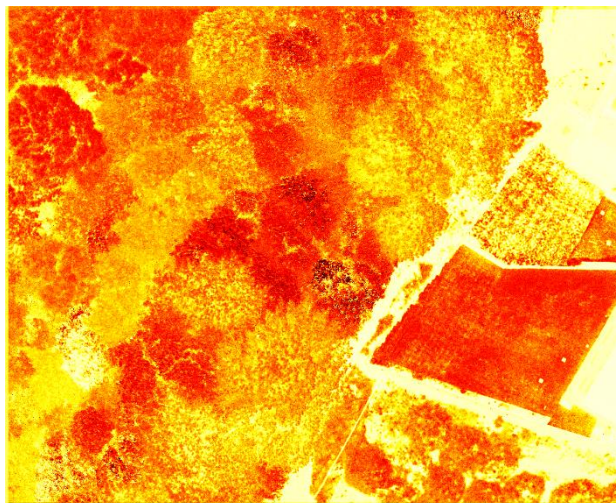


Figure 26: The eigenvectors of spectral band contributing to form each individual PC image for (a) pine (Scots and Black) trees, Norway maple and sycamore trees (b) Scots pine trees only.

[1 = Blue, 2 = Green, 3 = Red, 4 = Red Edge and 5 = NIR].

PC for identifying pine (Scots and Black) Norway maple and sycamore (c) PC1 (d) PC2 (e) PC3 (f) PC4 (g) PC5; PC for identifying Scots Pine tree (h) PC1 (i) PC2 (j) PC3 (k) PC4 (l) PC5

Table 4: Percentage Variance of PC1 – PC5 resulting from PCA applied to multispectral images to classify (a) maple, pine (Black and Scots) and sycamore trees and (b) Scots pine trees.

(a)	PC1	PC2	PC3	PC4	PC5
% variance	47.435	26.570	12.801	8.599	4.595
% cumulative	47.435	74.005	86.806	95.405	100.000
(b)					
% variance	56.222	21.738	9.592	7.540	4.908
% cumulative	56.222	77.960	87.552	95.092	100.000

6.2.7 Principle of K-means segmentation in image data

Image segmentation is separating meaningful information existing in an image to segments. The segments are oriented by dividing extremely similar pixels into groups using image segmentation techniques (Burney et al., 2014). There are numerous image segmentation techniques including edge-based, threshold-based, cluster-based, and neural network-based (Dhanachandra et al., 2012). However, one of the most effective techniques from the existing techniques is cluster-based methods of K-means clustering, Fuzzy C means clustering, mountain clustering and subtractive clustering methods. In simple words cluster-based methods, groups data (pixels) in an image with similar characteristics enabling them to be identified as groups. Amongst these cluster-based methods one of the simplest cluster-based method is K-means clustering. K-means clustering is an unsupervised machine learning algorithm used to segment different groups of samples (e.g., tree species) from the background into different clusters representing a similar pixel value. The clusters are obtained by observing the similarity in the data represented by the assigned K-number of clusters, which then

calculates and places centroids according to the k-number of clusters (Shan et al., 2018). Afterwards, the Euclidean distance for each pixel is calculated to the centroid which classifies the pixel into groups based on the distance. The longer the distance between pixels the smaller the similarity and chances of pixels being in separate groups. The closer the distance between the pixels the greater the similarity of them being clustered as one group.

K-means algorithm for image segmentation is composed of several steps:

Step 1: Determine K-clusters or K-random pixel values, which are the initial assumed centroids. The pixel in a region is similar with characteristic features of colour, intensity, and texture. (e.g., if an image has two distinct colours K-cluster = 2).

Step 2: Randomly place the centroids on the data set of pixels and Euclidean distance is calculated. There are two concepts of calculating distance in K-means clustering. Within Cluster Sums of Squares (WSS) and Between Cluster Sum of Squares (BSS). WSS is the sum of distance between the pixels and the corresponding centroids for each cluster and BSS is the sum of distance between the centroids and the total sample mean multiplied by the number of data points within each cluster.

The BSS distance is calculated marking the regions of clusters, and the pixel values closer to each centroid is grouped into one cluster.

$$BSS = \sum_{i=1}^{Kc} |Ci|. d(\mu_{Ci}, \mu)^2 \quad (16)$$

C_i = Cluster; Kc = number of clusters; μ_{Ci} = Cluster centroid; μ = sample mean; d = distance

Step 3: The WSS distance is calculated for each pixel value and centroids assigning the pixel value to the nearest cluster.

$$WSS = \sum_{i=1}^{Kc} \sum_{x \in Ci} d(x, \mu_{Ci})^2 \quad (17)$$

X = data point

Step 4: The centroids is moved towards the centre of each cluster using the mean of the distances between pixel values.

Step 5: Repeat the iteration measure the distance using the mean of each cluster until the centroids stabilize and there is no change in clustering.

6.2.8 K-means segmentation and quantification of invasive tree species

After k-means clustering, each area's total pixel count was multiplied by the resolution of the drone images that were taken (0.053 m / pixel * 0.053 m / pixel) to determine the total woodland under investigation as well as the different tree species.

6.3 Results and Discussion

6.3.1 Assessing the effectiveness of PCA and k-means segmentation in invasive tree species classification.

Using the eigenvectors from PC1 and PC2, the PCA results were used to derive an NDSI (Figure 26(a)), which displayed a strong spectral response in the green, red, red edge, and NIR spectral bands between PC1 and PC2. The spectral features that allowed for the identification of pine (Scots and black), Norway maple, and sycamore trees were improved by this difference in spectral bands. The segmentation method of thresholding or k-means clustering is used to further categorise the clusters. By giving each pixel a threshold value between 0 and 1, which separates the image's light and dark regions, thresholding transforms a greyscale image into a binary image. When there are only a few species of trees present, thresholding for image segmentation is effective. Thresholding is less accurate at segmenting numerous trees since it solely considers intensity and ignores the relationship between pixels, which could result in the inclusion or exclusion of pixels for a desired region of a particular tree. Figure 27(a) illustrates this effect by depicting Norway maple trees in white areas, which suggests that they occupy a larger area in the woodland than in reality. The different intensities of the regions are challenging to differentiate as some appear to have a similar intensity level which allows them to be misclassified by the threshold segmentation approach. Due to this effect k-means segmentation was applied to the PCA derived NDSI image. K-means segmentation is an unsupervised algorithm with no labelled data used to identify clusters in in the data set. Initially, the NDSI derived PCA (Figure 27(b)) was used for k-means segmentation. The k-number of clusters assigned was six which allowed segmentation of the different clusters of invasive tree species into pine (Scots and black), Norway maple and sycamore (Figure 27(c)). The mustard green coloured areas representing the pine (Scot and black) trees as one cluster, and the brown coloured areas representing the sycamore trees as the second cluster (Figure 27(c)). The third clusters were regions with light blue, pinkish red, and dark blue coloured areas in the centre representing the Norway maple trees (Figure 27(c)). The final cluster at the bottom of the

image, with a different shade of blue colouration, was identified to be a mixture of oak and silver birch trees (Figure 27(c)).

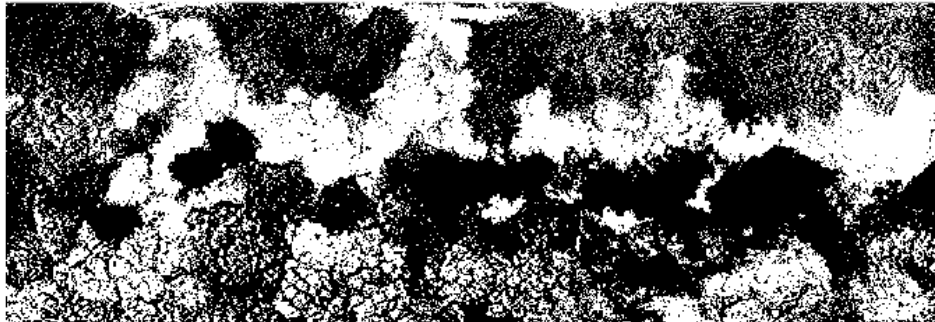
However, despite numerous interpretations, no spectral index was able to distinguish the tree species (oak and silver birch) as distinct tree species due to their poor phenology. Figure 27(d), which segments the Norway maple trees more precisely than threshold segmentation (Figure 27(a)), demonstrates the effectiveness of k-means segmentation. Figure 27(e), in which k-means segmentation is carried out directly on the RGB image without PCA, also demonstrates the significance of multispectral imaging and the process of developing an NDSI algorithm prior to k-means segmentation. The segmentation of invasive tree species is extremely poor in the image (Figure 27(e)). However, the current NDSI classified the two types of pine trees (Scots and black) as one tree type. Hence, eigenvectors from the first three PC score images were combined which retained 87.56% (Table 4b) of information to extract features enabling to classify the Scots pine tree (Figure 28b). The newly derived spectral index (equation 15) was segmented by k-means clustering where k-number of clusters for the image was 4 representing the sky-blue regions in a single row as one cluster of Scots Pine trees (Figure 28b). Finally, when compared to the data from the field study, the k-means segmentation results showed good correlation with the locations of the invasive tree species. (Figure 29).

6.3.2 Quantitative information obtained by analysis of UAV MSI and field study data.

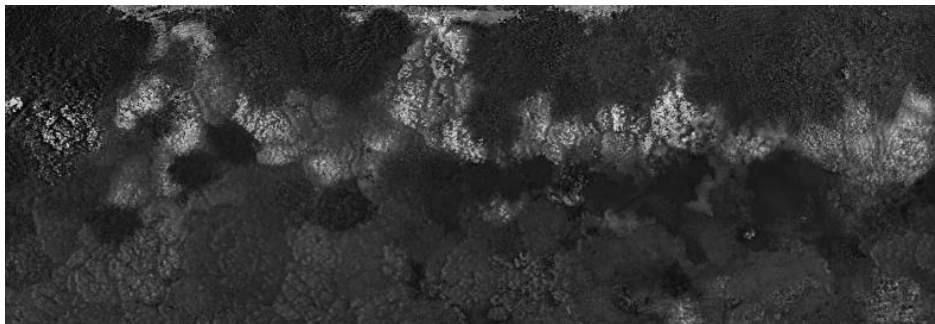
The invasive tree clusters of Scots pine, Black pine, Norway maple and Sycamore trees obtained from the k-means segmentation images (Figure 27c and Figure 28b) were quantified by calculating the number of pixels occupied by each cluster of trees divided by the area of the woodland on 2 separate days. The average results of the quantitative results using UAV MSI identified 19% Norway maple trees, 12% Scots pine trees, 23% Black pine trees and 19% of Sycamore trees (Table 5). Further, the UAV data capture was 22 minutes over an area of 8,052.28 m². In contrast, a detailed ground-level field study was undertaken around 4 hours over an area of 6,785 m² which resulted in the manual counting, identification and GPS mapping of 30% Norway Maple trees, 10% Scots Pine trees, 26% Black Pine trees, and 14% Sycamore trees (Table 5). A correspondence can be observed in the quantification of invasive trees between UAV MSI and field data study. For example, the % Scots pine in both the studies was around 20%. However, field study data is counting and identifying tree species based on ground-level observations, whereas the UAV MSI data phenotypes tree type from the canopy based on its shape and colouration. Also, the data based on scaling of the tree canopy versus the tree trunk has not been attempted. Further, the other trees which was oak and silver birch

at the bottom edge of the southern boundary of the woodland were identified by UAV MSI (27%) and field study (20%).

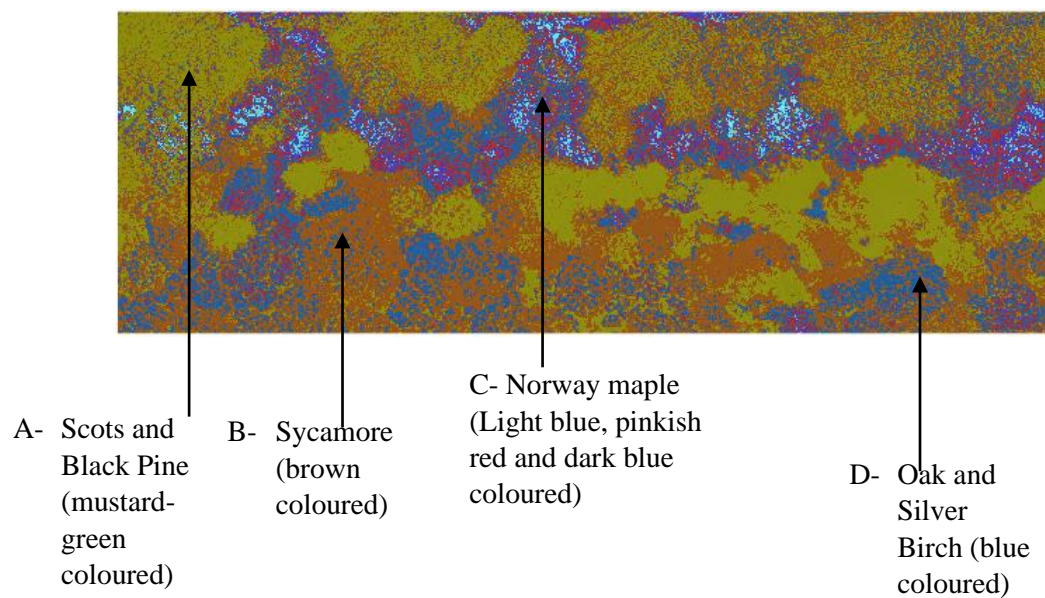
(a)



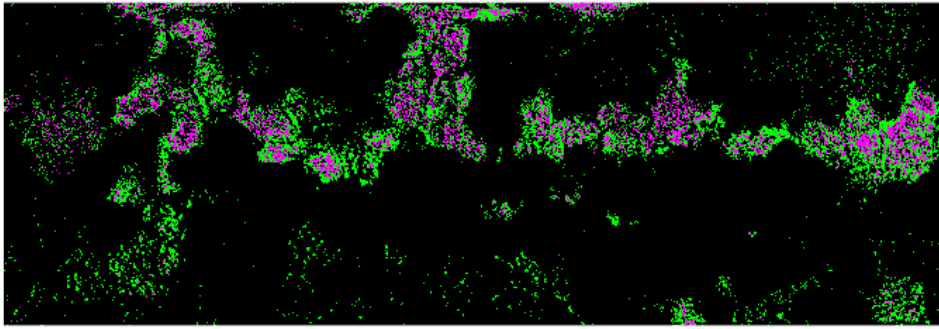
(b)



(c)



(d)



(e)

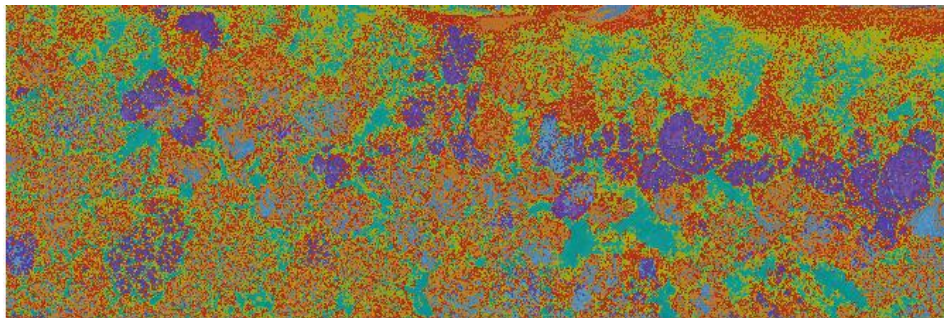


Figure 27: PCA and k-means segmented images for classifying pine (Scots and Black), Norway maple and sycamore trees (a) threshold segmentation representing Norway maple trees from PCA derived NDSI, (b) PCA derived $NDSI = (PC1) / (PC2 + PC2 + PC1)$, (c) k-means segmented image from PCA derived NDSI, (d) k-means cluster representing the Norway maple trees, and (e) k-means segmented image directly on RGB image

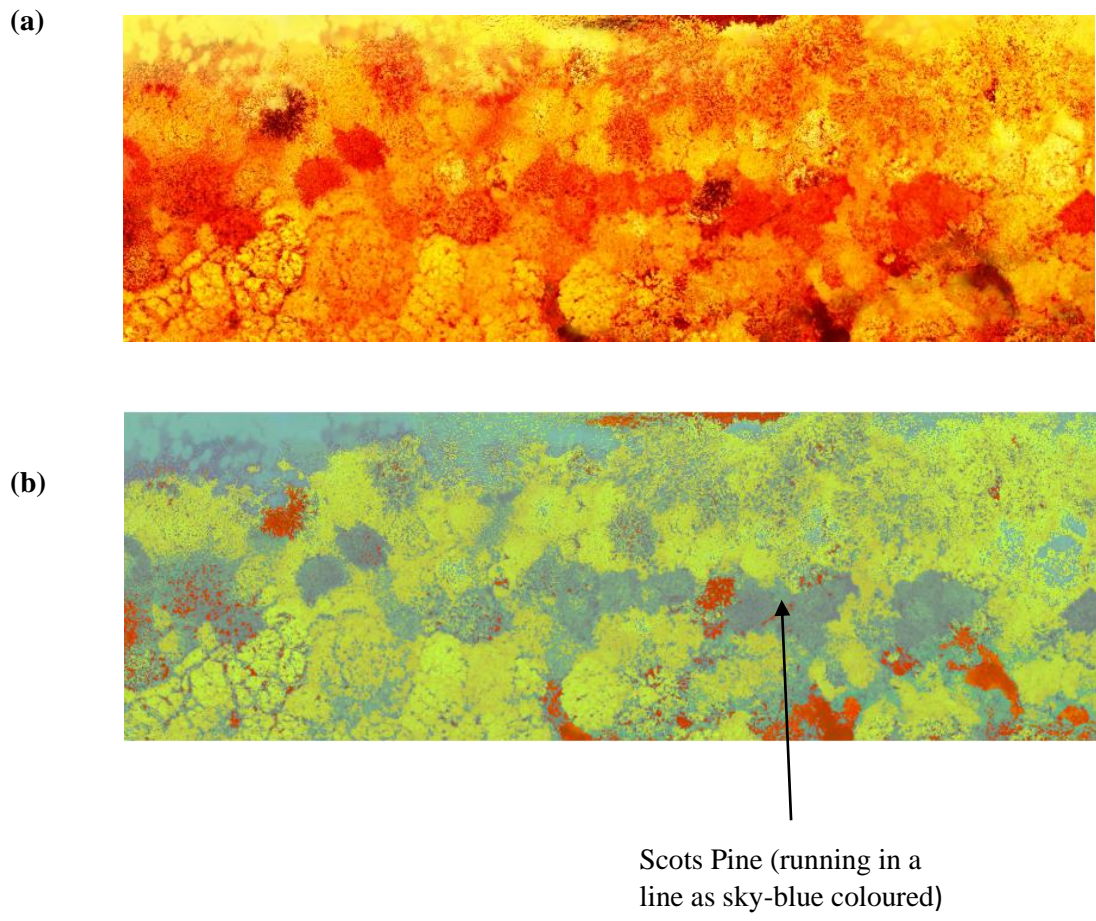


Figure 28: PCA and k-means segmented images for classifying Scots pine (a) PCA derived (PC1 + PC2 + PC3) colourmap image, and (b) k-means segmented image.



Figure 29: GPS coordinates for the tree species in the woodland boundary

Table 5: Quantitative information obtained by analysis of UAV MSI and field study data.

UAV flight [#]	From analysed UAV data										Field study data				
	Total Woodland Area (m ²)	Area of Norway Maple coverage (m ²)	Area of Scots Pine coverage (m ²)	Area of Black Pine coverage (m ²)	Area of Sycamore coverage (m ²)	% Norway Maple trees	% Scots Pine trees	% Black Pine trees	% Sycamore trees	% Other trees (oak and silver birch)	% Norway Maple trees	% Scots Pine trees	% Black Pine trees	% Sycamore trees	% Other trees (oak and silver birch)
Value*	8052	1565	934	1805	1493	19	12	23	19	27	30	10	26	14	20
Range ^{&}	7866-8239	1485-1644	903-964	1794-1817	1453-1533	NA					NA				

[#] UAV flights: 14 October 2020 and 19 October 2020.

* Value based on n = 2 determinations for the tree area.

[&] Range of values, based on individual analysis.

[@] Field data collected on 13 March 2021 over an area of 6,785 m².

NA = not applicab

6.4 Conclusion

This study has demonstrated a simple approach to classifying multiple invasive tree species with some additional benefits. The results have demonstrated the benefits of using PCA in the classification process. The use of PCA allows the selection of the most appropriate spectral bands to classify tree species into clusters and consequently the derivation of a new spectral index. Current spectral indices, such as NDVI, have limited application. Building up new spectral indices for a specific purpose is thus a straightforward method that has demonstrated to offer promising outcomes for classifying tree species in an ASNW. Furthermore, in this approach, k-means segmentation has been determined to be the most effective segmentation method for identifying multiple trees. K-means allows for further refinement of the spectral index, which leads to quantification of tree species types. Data from a ground-level field study confirmed the accuracy of this simple approach. Current methods for classifying tree species rely on object-based classification methods, which necessitate extensive and time-consuming supervised machine learning and deep learning methods to construct classification models. The current approach, on the other hand, is pixel-based and allows for the selection of spectral bands to derive spectral indices suitable for tree species classification. It is a powerful tool for identifying different tree clusters when combined with the k-means segmentation approach. Furthermore, our approach does not require a training data set, making it simple to implement, adaptable, and ultimately less time-consuming.

SECTION B

MULTISPECTRAL UAV IN PRECISION AGRICULTURE

Chapter 7: Multispectral UAV in precision agriculture of crops

7.1 Introduction

7.1.1 Precision agriculture: Its Principles and Significance

In recent years, agriculture has become growingly crucial for the global economy. According to statistics, agricultural production must rise by 60–100% to satisfy the increasing demand of food for a future global population of 9 -10 billion people (Monteiro et al., 2021). To satisfy this demand, precision agriculture of crops is essential. Precision agriculture entails crop management using low-cost sensing or observational technological methods to respond promptly with management strategies to increase environmental performance and crop yield. The technologies used in precision agriculture are global positioning systems (GPS) and geographic information systems (GIS). After the late 1990s, GPS systems for precision agriculture emerged. These systems use ground-based sensors mounted on, for instance, tractors or farm equipment to enable variable rate application (VRA) of seeds and fertiliser on to the field (Vrchota et al., 2022). Whereas GIS systems employ remote sensing technology to create survey maps, monitor crop growth, and identify fields that need fertilisers to increase crop yield and prevent overuse of fertilisers that could lead to environmental issues. Hence, to achieve optimal outcomes in crop precision agriculture, it is recommended that both GPS and GIS technologies must be used in combination. Precision agriculture by remote sensing is also known as high throughput phenotyping of crops, in which important information about crops are discovered from cell to the canopy level at various growth stages (Tao et al., 2022). These remote sensing image-based crop phenotyping techniques enable rapid and unbiased evaluation in precision agriculture for nutrient management, disease management, weed management, and crop yield estimation. Typically, remote sensing requires a higher spatial resolution of 1-3 m for crop yield estimation and a spatial resolution of 5 - 10 m for nutrient management (Mulla, 2013). Whereas weed mapping requires a finer spatial resolution of 5 - 50 cm to effectively identify weed patches around crops. Even though satellites can provide high spatial (5 m) and temporal (daily) resolution images, most publicly available satellite products have a coarse resolution for many precision agricultural applications (Sishodia et al., 2020). In comparison, UAV platforms can typically provide low spatial resolution (<5 m), making them suitable for effective precision agricultural applications.

7.1.2 Significant application of UAV sensing in precision agriculture

Remote sensing in precision agriculture provides the tools and technologies for data assessment to improve agricultural practises and increase crop yield. Among the most common uses of remote sensing in precision agriculture are irrigation water management, nutrient management, disease management, weed management and crop yield monitoring (Sishodia et al., 2020). Throughout the growth period of crops remote sensing imagery are gathered to identify numerous agricultural water demand indicators, including evaporation rate, soil moisture, and crop water stress (Evans et al., 2013 and McDowell, 2017). These indicators are used to precisely plan irrigation and assess crop water requirements. Vegetation indices derived from remote sensing data have shown strong correlation with crop chlorophyll contents and growth which helps the farmer understand the crop nutrient status across the field for precision application of fertilisers (Franzen et al., 2016 and Scharf et al., 2011). Further, remote sensing offers the capability to monitor disease in crops at early stages of disease development by vegetation indices and machine learning approaches (Di Gennaro et al., 2016). Also, weeds surrounding crops can be differentiated from remote sensing data by image classification approaches based on their spectral, phenological and morphological attributes (Partel et al., 2019 and Huang et al., 2020). Additionally, remote sensing data enables crop yield monitoring through vegetation indices, estimation of the leaf area index, and biomass detection, enabling the farmer to estimate the yield months before harvesting (Ali et al., 2020).

Traditional ground-based precision agriculture relies on hand-held spectrometers to collect phenotypic data in the field. Remote sensing methods, on the other hand, are non-destructive in gathering phenotypic crop data in the form of an image (Yang et al., 2017). Both hand-held spectrometers and remote sensing platforms are incorporated with spectral sensors with varying spectral bands to record the reflected light from a sampling surface of interest. As the sensors can detect the presence of chlorophyll in crops, they are useful in analysing the phenological growth stages of a crop by deriving vegetation indices. Different sensors may use different conversion methods to generate a reflectance value from a desired spectral band, which can influence the vegetation index calculation (Di Gennaro et al., 2022). Hence, few studies compare the performance of different UAV-based cameras with hand-held spectrometer in crop monitoring under the same environmental conditions.

A study conducted in New Zealand compared the spectral response of a variety of ryegrass pastures using four different types of UAV based sensors to a ground based hyperspectral hand-held spectrometer (von Buren et al., 2016). The UAV multispectral camera ($R^2 = 0.92$) and the

UAV high spectral resolution spectrometer ($R^2 = 0.98$) were found to have a higher correlation to the hand-held spectrometer data. However, a lower correlation was found in the spectral data of RGB ($R^2 = 0.63$) and converted near-infrared camera ($R^2 = 0.65$) in comparison to the hand-held spectrometer. Similarly, spectral data from a wide range of barley crops using two lightweight hyperspectral UAV sensors had similar spectral profiles to the data from a portable spectrometer (Bareth et al., 2015). Also, a study performed on oats crop revealed a very strong positive correlation ($R^2 = 0.89-1.00$) between the NDVI values measured using a hand-held system and the UAV-installed camera (Csajabok et al., 2022). Further, multispectral UAV camera compared the performance of different vegetation indices on a wide variety of crops from rapeseed, barley, onion, and potato in comparison to the data obtained using a hyperspectral field spectrometer (Nebiker et al., 2016). High correlation was found between the UAV and ground based spectral data. Additionally, the vegetation cover of maize crop was tested using the vegetation index NDVI and red-edge NDVI by a Mini-MCA6 multispectral camera and Sequoia multispectral camera and compared to the soil-plant analysis development (SPAD) meter (Deng et al., 2018). The results from the study highlighted that the Mini-MCA6 camera had a higher accuracy on the NDVI values and Sequoia had a higher accuracy on red-edge NDVI. However, regardless of which camera was used, SPAD accuracy on red-edge NDVI was always higher than NDVI. In addition, recent research examined the ability of two multispectral UAV sensors of Sequoia and DJI Phantom 4 Multispectral (P4M) to accurately produce spectral profiles and vegetation indices when identifying a large area of grassland (Lu et al., 2020). The study revealed the sensors to be highly correlated with portable spectroradiometers data ($R^2 > 0.90$) and to provide good accuracy (Sequoia RMSE 0.07; P4M RMSE 0.09). Finally, UAV sensors of SENOP HSC-2 hyperspectral and DJI P4M multispectral were used to characterise six common targets in vineyards: bare soil, bare-stony soil, stony soil, soil with dry grass, partially grass covered soil, and canopy (Di Gennaro et al., 2022). The accuracy of the vegetation index calculation using UAV sensor datasets was then evaluated using the percentage error with respect to the true value measured with the GER 3700 reference spectroradiometer. The percentage error for vegetation index values was found to be relatively high for images that were not radiometrically calibrated. This concludes the fundamental importance of radiometric calibrated UAV images in accurately matching ground truth data to obtain precise vegetation index calculation. Given the widespread use of UAVs by researchers in crop phenotyping in precision agriculture there is a need for additional performance evaluation of UAV sensors to ground based field spectrometer in different crops. Little attention is paid to evaluate the behaviour of multiple vegetation indices across different

phenological growth stages of crops. Also, as NDVI is the most widely used vegetation index in crop phenotyping, it is critical to investigate the applicability of various other spectral indices to develop phenology detection with remote sensing in crops.

Another important application in precision agriculture is the evaluation of soil nutrient content for effective growth of crops. In traditional precision agricultural methods soil nutrient content are analysed manually in laboratory analysis which are extremely time consuming and laborious. An alternative approach for assessing soil health indicators in crops have been possible through remote sensing technologies. Soil health indicators are a set of measurable physical, chemical, and biological characteristics used to evaluate soil health (Raghavendra et al., 2020). The most well-established chemical indicators are pH, soil organic matter and three major soil nutrients of nitrogen (N), phosphorus (P) and potassium (K). NPK are primary soil nutrients that crops require for optimal growth and development (Rütting et al., 2018). Nitrogen is essential for photosynthesis and flower bud differentiation during the vegetative growth phase of crops, which increases crop yield (Anas, 2020). Insufficient amounts of N can hinder the growth and development of crops. Phosphorus stimulates root growth and flowering in crops by enhancing cell division (Rütting et al., 2018). As a result, P deficiency can cause plant maturity to be delayed. Potassium promotes the movement of water, carbohydrates, and nutrients such as nitrogen in plant tissues, resulting in optimal plant growth; thus, K deficiency can postpone optimal plant growth (Xu et al., 2020). K also acts as an activator of key enzymes involved in protein synthesis, sugar transport, and photosynthesis. Further, the quality and quantity of soil organic matter is significant in the management of soil fertility, nutrient supply, and carbon dynamics (Paul, 2016). Also, the loss of organic carbon content in soil can limit the soil's ability to provide nutrients for long-term crop development. Finally, soil pH is a primary factor in soil since it regulates the amount of soil nutrients and chemicals that are soluble in water, and the amount of nutrients available to crops (Neina, 2019). Most agricultural crops thrive in soil pH ranges of 5.5 to 7.5, and at lower pH levels, most essential nutrients become unavailable to crops. Furthermore, excessive usage of these chemical indicators can result in economic loss, fertiliser waste, resulting costly crop management (Hossen et al., 2021). Whereas overuse of NPK fertilisers causes the release of harmful greenhouse gases into the atmosphere and the eutrophication of waterways (Anas, 2020). As a result, these chemical soil health indicators must be preserved and used appropriately to ensure favourable nutrient conditions to maximise crop yield contributing to food security (Bascietto, 2021).

An important aspect of soil health management for farmers is the timely application of fertilisers for precision agriculture of crops (Maes and Steppe, 2019). Determination of spatial patterns across a crop field is critical for fertiliser application. In many instances crop nitrogen status is estimated indirectly from variables of chlorophyll content and leaf area index using chlorophyll meters (Maes and Steppe, 2019) For instance, N application rates for corn and hardwood species were determined using SPAD chlorophyll meter (Chand and Robinson, 2002 and Hawkins et al., 2007). However, chlorophyll meter methods fail to capture spatial variability often present within a crop field. Also, these ground-based methods necessitate the collection of many samples, which is deeply destructive to crops, labour-intensive, and time-consuming (Maes and Steppe, 2019). Remote sensing techniques are a viable alternative to ground-based measurements as they provide crop reflectance and diagnostic information on crop nutrient concentration in a timely and spatially contextualised manner (Walsh et al., 2018). The two main remote sensing technologies used in precision agriculture for soil mapping are satellites and UAV. For instance, vegetation indices derived using remote sensing data from multispectral sensors can be used to detect N statuses in crops (Schelmera et al., 2013). This is due to the strong correlation between N concentration and chlorophyll content at canopy scale, as majority of N is localised within chlorophyll molecules of plant leaves (Schelmera et al., 2013). Hence, several chlorophyll sensitive VIs have been successfully employed to estimate chlorophyll in crops and correlate to N concentrations. A study found a strong relationship between UAV-based vegetation indices of normalised difference red-edge index (NDRE) and chlorophyll index (CI) green and plant N concentration in spring wheat (Walsh et al., 2018). Additionally, a recent study identified the simple ratio of NIR, and red spectral band (SR_{red}) derived from UAV data as a suitable VI to estimate N uptake in various winter cover crops (Holzhauser et al., 2022). Furthermore, research has been conducted to assess the relationship between remote sensing VI and various nitrogen treatments, as well as the strengths and weaknesses of VI in detecting nitrogen fertiliser treatments. Amongst recent studies NIR/VIS indices were adequate for detecting lower N status without sacrificing precision, while the red edge spectral band was critical for detecting higher N status in winter wheat crop (Prey and Schmidhalter, 2019). Also, low N management using simple linear NDVI, and variable rate application of N fertilisation are likely to reduce the economic and environmental impact of N fertilisation in winter wheat (Vizzari et al., 2019). Furthermore, NDVI and leaf chlorophyll index yielded promising results in determining whether nitrogen can be applied for a second round, known as topdressing nitrogen, during late dormancy to promote early greening of white oats crops (Barbosa et al., 2020). Also, green NDVI, NDRE, and CI green remote sensing VI

showed a very strong relationship with nitrogen fertiliser treatment in maize crop. (Burn et al., 2022). These indices were able to detect severe N deficiency in maize crops at various vegetative growth stages. Also, machine learning estimative models were used to estimate total nitrogen in crop soil from multispectral-UAV images with a root mean square percent error (RMSPE) of 10.8% (Hossen et al., 2021). In this study the multispectral-UAV based imaging data used red, NIR, green spectral bands and NDVI indices to develop the machine learning models. A few recent studies have also used satellite data to construct VI to identify a possible relationship with soil nutrient indicators. Hence, a strong positive correlation was found between NDVI derived from satellite imagery time-series data and soil nitrogen content in a rice crop field (Trigunasif and Saifollah, 2022). Similarly, a study discovered that nitrogen fertilisation had a much stronger effect on UAV derived multiple vegetation indices than potassium and phosphorus fertilisation on rye crops (Wójcik-Gront et al., 2022). Furthermore, significant positive correlations were found between NDVI for bare soil and pH, and all soil nutrients studied on multiple maize fields (Mazur et al., 2022). However, there was no relationship found between NDVI derived from satellite data and soil macronutrients of NPK in wheat crops, indicating that NDVI cannot be used as a proxy to determine soil NPK status in crops (Bascietto et al., 2021). Also, most of these studies were subjected with treatments having large ranges in (most often) N rates which are not necessarily representative for field conditions (Maes and Steppe, 2019). Furthermore, since nitrogen is by far the most studied soil nutrient, there is less research identifying the direct association between remote sensing data of crops performance and soil chemical health indicators. The recent studies which were conducted to identify direct relationship between soil nutrients and VI (Bascietto et al., 2022 and Trigunasif and Saifolla, 2022), used remote sensing platform of satellites.

7.1.3 Crop varieties essential for precision agriculture in the UK

In the UK, agriculture and farming are crucial as they constitute a significant source of the food that the nation needs to consume (Statista Research Department, 2023). Around 71 percent of the UK's land area was used for agriculture in 2021, and the agricultural industry generated more than 9 billion British pounds. Approximately one-third of the agricultural area, or 6.1 million hectares, is used for agricultural crops, and the most common crops grown in these land areas are 19% cereal, 2% oilseed rape (OSR), 1% potatoes, and 4 % of other crops. (Department for Environment, Food & Rural Affairs, 2022).

Further, according to the UK's Agriculture and Horticulture Development Board (AHDB) anticipated an OSR production of 1.04 Mt (Speight, 2021) which is the lowest production this

century. This decreased OSR production was identified to be due to extreme weather conditions, linked to climate change, and the cost of effective crop management (Searby, 2020). Meanwhile, the total UK bean production varies between 500,000 to 600,000 tonnes per year (Senova, 2021). Typically, around 150,000 tonnes of beans are exported globally for human consumption and the remainder is sold locally (Senova, 2021). Annual demand of beans for animal feed in the UK ranges between 270,000 to 350,000 tonnes (Senova, 2021). In addition, among the cereal crop in the UK, the oats production fell by 8 % to 825 thousand tonnes in 2022, due to a 12 % decrease in area (National Statistics, 2022). This however caused a 4.5 % increase in yield in comparison to 2021. Also, in 2021/22, the UK exported approximately 675 metric tonnes of oats to non-EU countries (Statista, 2022). However, as for all crops, disease, weed growth, and adverse weather conditions can significantly influence the crop yield (Jones, 2016). Hence, according to the crop statistics in the UK, efficient crop management plans and monitoring systems through precision agriculture are required to stabilise the increasing demand for these crops to establish higher yields.

7.1.4 Literature review on precision agriculture of oilseed rape by multispectral-UAV

Oilseed rape (*Brassica napus* L.) is a widely cultivated crop worldwide due to its oil-rich seeds (Gardy et al., 2014). After harvesting, the seeds are crushed to liberate the highly desirable oil which has been consumed at 28-29 million tonnes per year worldwide for the past seven years (Shahbandeh, 2023). In addition, the by-products are used for animal feed, biofuel, and medicine (Chong et al., 2017). The phenological growth stages of OSR have been codified (Figure 30) (Meier, 1997). Stages 1 and 2 incorporate both seed germination and plant development. Flowering of the OSR occurs in stage 3, while in stages 4 and 5 OSR pod development and ripening take place.



Figure 30: Phenological growth stages of oilseed rape (OSR).

Monitoring the flowering stage of OSR through to the maturation of the OSR pod allows a more accurate estimation of the seed yield. Traditionally, the collection and use of OSR flower, plant and pod data from field investigation is subjective, labour intensive and potentially destructive to crops (Fang et al., 2016). Hence, the deployment of remote sensing techniques has become an area of interest in acquiring objective and accurate crop information. Specifically, the use of remote sensing approaches to predict OSR seed yield has been a growing area of interest over the last few years for precision agriculture (Wan et al., 2018). The distinctive visibility of the OSR flowers (Stage 3) (Figure 30), can be distinguished using characteristic spectral bands by remote sensing platforms. Recent studies using satellite remote sensing platforms have proposed spectral indices which have been used to estimate the OSR seed yield (Zang et al., 2020, Han et al., 2021 and Tang et al., 2022).

However, low spatial resolution (e.g., 500 m/pixel) obtainable using satellite images diminishes the accuracy for smaller agricultural sites (e.g., 5-10 ha). Hence, UAV remote sensing platforms have become the dominant option to estimate the OSR seed yield. An estimate of the biomass of winter OSR using vegetation indices coupled with a random forest regression model based on gathered UAV multispectral images has been reported (Liu et al., 2019). This study evaluated the potential of UAV to retrieve plot level information about a

specific OSR crop profile. Further, UAV multispectral data has also shown promising results by predicting OSR seed yield from correlation with the leaf area index during different phenological growth stages (Peng et al., 2019 and Hussain et al., 2020). However, these studies require field crop samples to be analysed in the laboratory for validation, and as they need datapoints from different phenological stages are time consuming. Meanwhile, the OSR flowers provide spectral features which can be interpreted by the sensors used in remote sensing platforms. The yellowness of OSR flowers is potentially due to the carotenoid absorption of blue light and the reflectance of green and red spectral bands (Sulik and Long, 2015 and Fang et al., 2016). Sulik and Long, 2015 reported that the ratio of green and blue spectral bands were strongly correlated ($R^2 = 0.87$) to determine the number of OSR flowers. They then proposed that a new normalised difference yellowness index (NDYI) based on the green and blue spectral bands was a significant OSR seed yield predictor (Sulik and Long, 2016).

Other studies which have combined multiple vegetation indices and image classification methods (e.g., k-means clustering and pixel level mixture analysis) to improve the identification of OSR flowers and hence improved seed yield prediction (Gong et al., 2018 and Wan et al., 2018). Similarly, Zang et al., (2020) used image-based phenotyping of thresholding to classify winter ($R^2 = 0.84$) and spring OSR flowers ($R^2 = 0.72$). Additionally, a simplified approach using multi-series data, generated by NDYI-based flowering pixel, and classified by thresholding successfully estimated OSR seed yield (Zhang et al., 2021). However, the most relevant of these studies estimated OSR seed yield based on a single data set on a relatively small field area by the classification of thresholding. In addition, the unsupervised k-means image processing studies were carried out on RGB image datasets rather than NDYI derived maps (Gong et al., 2018 and Wan et al., 2018). This has limited the understanding of the performance of combining an unsupervised classification model with NDYI to estimate OSR seed yield.

7.1.5 Literature review on precision agriculture of winter beans by multispectral-UAV

Winter beans are a widely cultivated crop worldwide consumed as both food and animal feed (Limagrains UK, 2019). According to data from the Processors Growers Research Organisation (PGRO) in the UK the leading varieties of winter bean are Vincent, Vespa, Bumble, Norton, Wizard, Honey, and Tundra (Nickerson). Tundra is the highest yielding variety on the PGRO winter bean recommended list as it has a high protein content making it suitable for export market, human consumption, and animal feed (Nickerson). The phenological growth stages of

winter bean is shown in Figure 31. Stage 1 and 2 involves seed germination and stem elongation. Stage 3 and 4 involves flowering and pod development. Finally, stage 5 incorporates pod ripening and senescence. Furthermore, no studies on winter beans have been conducted, and only a few studies using remote sensing platforms have focused on precision agriculture of other bean varieties. Hence, a recent study, using a UAV, monitored both the growth of the faba bean across the phenological growth stages and using machine learning algorithms estimated the yield of faba bean (Ji et al., 2022). Soybeans have been the focus of a few studies. A UAV-MSI has been deployed to predict soybean yield based on vegetation indices coupled with machine learning regression models (Alabi et al., 2022). While soybean yield has been predicted, from UAV-MSI data using a deep learning method (Maimatijiang et al., 2020).

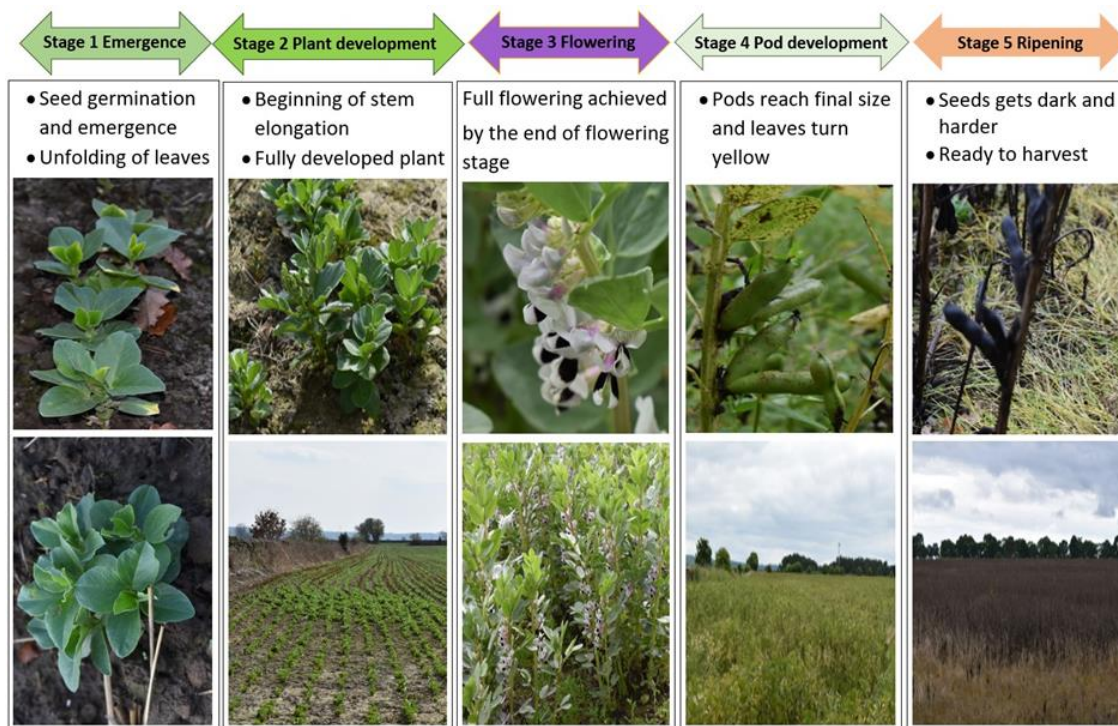


Figure 31: Phenological growth stages of winter beans

7.1.6 Literature review on precision agriculture of winter oats by multispectral-UAV

Oats (*Avena sativa* L.) are Europe’s fifth largest crop and the sixth most grown cereal worldwide (Buerstmayr et al., 2007). Oats contain a variety of nutrients, allowing them to be used primarily for human food and animal feed (Paudel et al., 2021). Also, the majority of oats grown are used for animal feed and approximately 25% of the crop is milled for human consumption (Tosh and Miller, 2016). Due to their high fibre and protein content, oats are commonly found in breakfast cereals, beverages, bread, and in infant foods (Paudel et al.,

2021). Based on the decimal code for cereal growth stages developed by Zadoks et al., 1974, the phenological growth stages of oats can be classified into five critical stages (Figure 32). Stage 1 incorporates leaf emergence and tillering. Stage 2 involves stem elongation and panicle formation. Flowering occurs in stage 3 and grain filling occurs in stage 4. Finally, stage 5 involves ripening of the oats crop. Hence, using the phenological growth stages white oats grain yield was estimated by spectral models previously calibrated using NDVI and Inverse ratio vegetation index (IRVI) values retrieved from a remote sensing handheld spectrometer (Coelho et al., 2020). The models were found to have a strong positive correlation ($R^2 = 0.73-0.90$) for estimating white oats grain yield, with the flowering stage of oats having the lowest error for estimating oats yields. Similarly, NDVI in comparison to Falker chlorophyll index (FCI) was found to be suitable at estimating white oats grain yield at later phenological growth stages (Coelho et al., 2019). The methods used in these studies to estimate yield are traditional machine learning models of linear, exponential, and multiple regression models. Meanwhile, modern machine learning methods of artificial neural network, partial least square, random forest, support vector machine, and convolutional neural network (CNN) have also been recently used to estimated yield and biomass. As a result, four machine-learning algorithms were used to build oat biomass estimation models using a variety of vegetation indices derived from UAV-based multispectral imagery (Sharma et al., 2022). The machine learning models demonstrated promising results at estimating biomass in oat breeding nurseries. However, the main drawback was inaccurate model estimation across different locations. Further, automated machine learning model using hyperspectral UAV data accurately estimated oats yield at panicle formation and at flowering stage (Li et al., 2022). In addition, a study used CNN models to estimate cereal crop yield (wheat, barley, and oats) from time series UAV data collected over a 15-week period (Nevavuori et al., 2022). Also, using UAV imagery data, a study successfully classified winter oat varieties and foliar fertilisation treatments by silicon and sulphur (Csjabok et al., 2022). By using discriminant statistical analysis, the percentage accuracy for correctly classified oats varieties were 88.9% and treatments were 87.5%. However, no studies have been conducted to estimate oat yield by combining information from vegetation indices and image-based classification methods.



Figure 32: Phenological growth stages of winter oats.

7.2 Research gap

An important agronomic trait which can be used through every phenological growth stage to assess development is the crop height. The traditional, field-based method for estimating general crop height involves measuring the height of two or three plants from a single plot using a ruler which are laborious, and time consuming (Zhang(b) et al., 2021). However, the sensors mounted on remote sensing UAV platform allows images to be collected with centimetre-resolution which can be processed through structure from motion-based softwares to extract 3D point clouds and ultra-high aerial images (Verhoeven, 2011). Exploring 3D point clouds, crop height models can be generated with centimetre resolution by subtracting digital surface models from a digital terrain model. Retrieving crop height models are a non-destructive method and has been successfully implemented in several agricultural crop studies (Acorsi et al., 2019). Hence, crop heights in maize crops were successfully estimated using hundreds of ultra-high resolution stereo image data from a low-cost UAV platform (Li et al., 2016). Also, remote estimation of canopy height in different cereal crops have also been reported in literature. Where wheat crop height was calculated using ultra-high UAV resolution canopy surface models from booting to grain filling growth stages (Schirmaan et al., 2016). Also, using a moving cuboid filter to remove noise from the UAV based 3D point cloud, winter wheat crop height was successfully estimated (Song and Wang, 2019). In addition, ultra-high canopy height models for barley crop derived from UAV showed strong positive correlation ($R^2 = 0.92$) with the ground reference crop height data collected across different growth stages

(Bendig et al., 2014). Furthermore, summer barley crop height was estimated using UAV data-derived models that were suitable for estimating barley biomass production (Bendig et al., 2016). Another recent study demonstrated UAV RGB derived canopy height models for black oats crop to have positive correlations ($R^2 = 0.68 - 0.92$) with ground reference data across three different growth periods (Acorsi et al., 2019). Recent studies have evaluated remote sensing techniques to estimate OSR crop height using a UAV integrated with a RGB and LiDAR sensor (Xie et al., 2021 and Zhang(b) et al, 2021). However, in these studies, data was gathered over a few days, limiting their flexibility and hence utility.

Even though most studies have demonstrated the competence of UAV imaging combined with structure from motion algorithms to accurately estimate canopy height, no study has investigated crop height estimation throughout the five critical phenological stages of OSR, winter bean and winter oats. Furthermore, the relationship between crop height across growth stages and chemical soil health indicators has not been investigated.

Also, according to the literature review on precision agriculture of OSR, winter bean and winter oats, most studies focus on estimating yield of the crop by using supervised machine learning models. Although the supervised ML models performed well in terms of generalisation for estimating oat yield, the data required to build these models are time consuming and require extensive training and automation to implement. Also, it is extremely improbable that these models can be easily transferred to estimate yield from distinct locations. Limited studies have focused on estimating crop yield by vegetation indices and image segmentation classification models, which are easier to interpret in comparison to ML models. Also, it is significant to estimate the crop yield prior to harvesting and correlate with the actual yield, which would be helpful to provide information at regional and national scales. This helps to define crop prices and production destinations, for instance whether it should be used for animal feed or as human food products. Currently, in several countries crop production estimate ability is performed after harvest resulting in high expenses and delays in the process (Food and Agriculture Organization of the United Nations Rome, 2017). This also aids in generating thematic maps to identify specific management areas, and high-yielding areas can receive less resources, such as water and fertiliser, than low-yielding area. Further, the relationship between estimated oats yield by remote sensing data and oats grain quality have not been previously studied. Determining this relationship will help determine whether areas with high oat yield also have high grain quality oats. Hence, the quality of oat grain can be determined by extracting and analysing oat β -glucan, which is composed of a linear branched chain of D-glucose

monosaccharides bonded by a mixture of β (1 \rightarrow 3) and β (1 \rightarrow 4) linkages (Boyaci et al., 2002). Furthermore, rather than manually measuring soil properties in the field, predicting soil parameters for an estimated crop yield using UAV data assists farmers in precision agriculture by providing an estimated range of fertiliser in the region. This can be performed by developing ML models for predicting the concentration range of chemical soil health indicators and crop quality features at various crop yield levels (low yield and high yield) estimated by simpler UAV image classification models.

7.3 Summary

Hence, ArcGIS software (Esri Inc., West Redlands, CA, USA) was used for the multispectral UAV data manipulation of OSR, winter bean and winter oats.

The focus research area on OSR and winter bean was estimating the crop height across phenological growth stages and estimate yield at the vegetative growth stage prior to harvesting of the crop. The estimated yield was determined by NDYI for OSR and multiple vegetation indices for winter bean crop and an image classification model in ArcGIS software. Therefore, the outcome of the approaches on OSR crop will be discussed in Chapter 8 and winter bean crop will be discussed in Chapter 9.

However, for the winter oats crop at various phenological growth stages, multiple ground reference data of soil health indicators, yield, oats grain quality, and spectral profiles were collected. As a result, the relationship between ground reference data and multispectral UAV yield data for the oats crop will be thoroughly discussed in Chapter 10.

Chapter 8: Multispectral UAV monitoring prediction of oilseed rape crop performance

Oilseed rape is one of the most widely cultivated crops worldwide and estimating the seed yield before harvest predominately helps the farmer to precisely manage the crop. Hence, the focus research area on Chapter 8 for OSR crop would be estimating the crop height across phenological growth stages and estimating seed yield at the vegetative growth stage prior to harvesting.

8.1 Aims and Objective

8.1.1 Aims

The aim of this chapter is to estimate OSR crop height throughout the phenological growth stages and OSR yield in the flowering stage using a vegetation index.

8.1.2 Objectives

1. To evaluate the performance of a UAV with a MSI camera to estimate OSR canopy height across the different phenological growth stages over an 11-month growth cycle.
2. To predict the final OSR seed yield (in mid-August) using NDYI, with a simplified pixel based iso cluster classification method using ArcGIS software, during the flowering stage (May).
3. To assess the accuracy of the predicted yield.

8.2 Experimental Design

8.2.1 Sampling site Airy Holme Farm

Access to visit and monitor the farm at Kiln Pitt Hill, Consett DH8 9SL, was granted by John Miller and family (T & AE Miller) (Figure 33a). Winter oil seed rape (codex) was planted at a rate of 76 seeds / m² using a Claydon Hybrid T4 trailed drill (Rickerby, Hexham, UK) pulled by a Claas ARES 836 RZ tractor (Rickerby, Hexham, UK) to a depth of 10 mm on the 1 August 2020 in an 8.1-hectare field, known locally as Waskerley Edge. The field was treated as follows in general terms, for protection of the oilseed rape: initially at the sowing stage (August) a specific herbicide for oilseed rape was applied to control invasive broadleaved and grass weeds, Banastar® (BASF Agricultural Solutions UK, Cheadle, Cheshire, UK) along with an adjuvant Grounded AD® (Helena Agri-Enterprises, LLC, Collierville, USA). Adjuvants are added as they can enhance the efficacy and adsorption of soil applied herbicides (and other types of pesticides) in several ways including acting as a wetting agent; by improving the weatherability; enhancing the penetration and translocation; and adjusting pH. At this time also

(August) slug pellets were applied to the field to protect the recently deployed (sown) seed. To prevent damage from the cabbage stem flea, beetle and oilseed rape insecticide (Kung Fu®, Syngenta UK Ltd., Fulbourn, UK) was applied in September along with an adjuvant (Activator® 90, Nutrien Ag Solutions, Loveland, USA). Subsequently, and additionally a post emergence herbicide was applied during the winter (Falcon®, Adama Agricultural Solutions UK Ltd., Reading, UK) and spring (Korvetto®, Corteva Agriscience, Cambridge, UK). To restrict (i.e. reduce) winter growth, and prevent excessive winter growth, a growth regulator was applied in December (Caryx®, BASF Agricultural Solutions UK, Cheadle, Cheshire, UK). In early spring (March) a wide-ranging fungicide (Protefin, Clayton Plant Protection, Dublin, Ireland) is applied followed in late spring (May) with a further fungicide (Recital®, Bayer Crop Science Ltd., Cambridge, UK) with an adjuvant that reduces spray droplet surface tension and allow even coverage of the fungicide (Roller, Agrovista UK Ltd., Nottingham, UK) for the treatment and prevention of diseases in oil seed rape e.g. light leaf spot, mildew, mould and phoma stem canker. Finally, 21 days prior to harvesting of the oil seed rape a pre-harvest desiccate was applied (Roundup Vista plus, Bayer Crop Science, Cambridge, UK). The addition of a desiccate assists in producing an evenly ripe crop (and eliminates perennial weeds) to aid harvesting.

To promote growth of the oil seed rape fertiliser, NPK(S), 20-8-12(7SO₃) with the composition: total nitrogen (N) of which nitric nitrogen (9.2%) and ammoniacal nitrogen (10.8%); phosphorus pentoxide soluble in neutral ammonium citrate and in water (P₂O₅) 8% (3.5% P) of which phosphorus pentoxide soluble in water (P₂O₅) 7.7% (3.3% P); potassium oxide soluble in water (K₂O) 12% (10% K); sulfur trioxide soluble in water (SO₃) 7% (2.8% S) (CF Fertilisers UK Ltd., Billingham, UK) was applied on the field in spring 2021 (March and April). Additional nutrients were also added in spring (April – May) and included supplements for boron and molybdenum (Lebosol® MoBo, Lebosol® Dunger GmbH, Elmstein, Germany), phosphate, potassium and magnesium (Yaravita Magphos K, Yara Ireland, Grimsby, UK) and plant based amino acids and trace elements (Terra-Sorb Foliar, Agrovista UK Ltd.). An adjuvant, Boost (Dow Agro Sciences, King's Lynn, UK) was added to assist with ground coverage, soil penetration, crop uptake and rain fastness.

8.2.2 Waskerley Edge agricultural field

Soil analysis was done on this agricultural field in March 2021 (by Lancrop Laboratories, Pocklington, UK in association with Agrovista UK Ltd., Nottingham, UK). Soil analysis indicated that the sandy silt loam with a pH of 7.1, organic matter (5.4%) and a cation exchange

capacity of 12.3 meq/100 g had normal soil levels of the major nutrients i.e. phosphorus (30 ppm), potassium (185 ppm) and magnesium (69 ppm) alongside secondary and micronutrients i.e. calcium (2120 ppm), copper (3.3 ppm), iron (735 ppm) and zinc (5.2 ppm). However, deficiencies were noted in term of secondary and micronutrients, specifically, sulfur (3 ppm) against a guideline of 15 ppm; sodium (14 ppm) against a guideline of 90 ppm; boron (1.3 ppm) against a guideline of 2.1 ppm; manganese (58 ppm) against a guideline of 75 ppm; and, molybdenum (0.01 ppm) against a guideline of 0.3 ppm. Therefore, supplements were added to the field to rectify these deficiencies.

8.2.3 Crop harvest

The crop was harvested on the 15 August 2021 using a Claas Lexion 570, Terra-Trac combine harvester with header attachment (Rickerby, Hexham, UK). To begin harvesting the header of the combine harvester is hydraulically lowered. The oilseed rape is gathered at the front, as the combine moves forward, by the header. A slowly rotating pickup wheel pushes the crop down towards the cutter. The cutter bar, which runs the entire length of the header, cuts the crop at its base. As the oilseed rape is cut, spinning augers (or screws) guide the crop up a conveyor, for processing within the combine harvester. A threshing drum beats the cut crop to release the seeds which then pass-through sieves whose hole size is optimised to maximise seed recovery. The unwanted other plant material is passed along a conveyor, called a straw walker, and discarded out of the back of the combine. The recovered oilseed rape seed is collected in the grain tank. The oilseed rape seed then travels from the grain tank up an elevator where it is shot out of a side pipe into a waiting trailer, towed by a tractor. Sensors (FarmTRX, Troo Corp., Ottawa, Canada), mounted on the upside of the elevator, measure the amount of seeds recovered, by weight. The weight of seed is recorded on a data logger which transmits location and yield (amount of seed) using GPS technology using on-board data processing software.

8.2.4 Unmanned Aerial Vehicle

A multirotor UAV (DJI Phantom 4, Leeds, UK) was used with a multispectral camera with a 5 camera-array for data analysis. Images were captured as 16-bit TIF files corrected for ambient radiance values. The UAV speed was 5.0 m/s and had an average height of 50.6 m. All flights were recorded with a resolution of 2.7 cm/px, a front overlap ratio of 75%, a side overlap ratio of 60% and a course angle of 90°. Specific weather conditions relating to daytime temperature during flight, wind speed and direction (recorded using a handheld anemometer (Benetech®

GM816, Amazon UK), and UAV pilot anecdotal observations on cloud coverage were identified with specific dated data.

8.2.5 UAV Photogrammetric Processing

The multispectral UAV images were used to build an orthomosaic image using Agisoft Metashape Professional (64 bit) software v.1.7.1 (Agisoft LLC, St. Petersburg, Russia). The UAV photogrammetric processing steps were as follows: the aerial images were merged and aligned to construct a sparse point cloud by matching similar image attributes. The images were then accurately positioned to generate a 3D point cloud based on the GPS coordinates of each image and, a solid mesh model was created. Finally, an orthomosaic image (Figure 33b) was built using the WGS 1984 Web Mercator coordinate system.

(a)



(b)

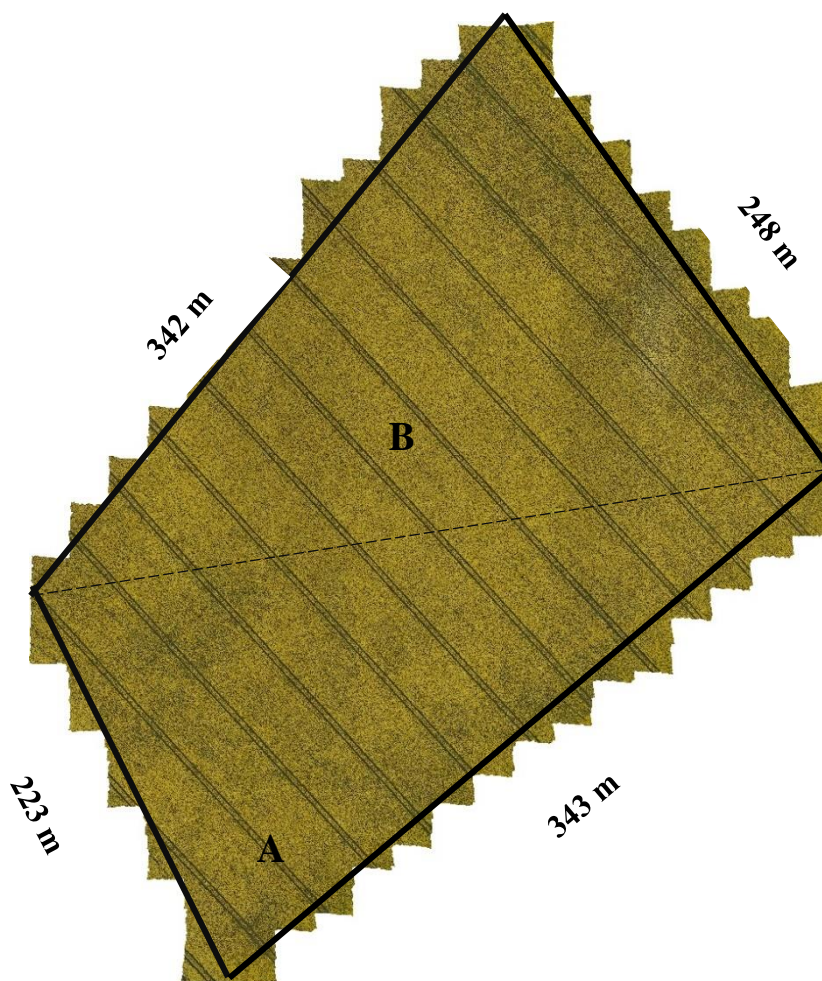


Figure 33: (a) Airy Holme Farm site map (b) Orthomosaic image of oilseed rape field.

[Note: The chosen area to investigate is outlined with an approximate area of 8.1 ha (80,653 m²). Area of A = 3.82 ha (38,245 m²) and Area of B = 4.24 ha (42,408 m²).

8.2.6 Crop Height Estimation

The OSR height estimation was performed using ArcGIS Pro v.2.8.0 software (Esri Inc., West Redlands, CA, USA). ArcGIS Pro software can construct a canopy height model (CHM) based on the principle of structure from motion (SfM). The SfM photogrammetry creates a rigid 3D model by matching similar features of several overlapping images (Xie et al., 2021). Hence, the 3D point clouds built by Agisoft were extracted and used by ArcGIS Pro to create CHMs based on time-series data on rapeseed phenological growth stages.

Initially, the 3D points were used to generate a digital surface model (DSM) and a digital terrain model (DTM). The DSM was calculated by including features which are elevated from above ground and the DTM was calculated by interpolation of features from ground or soils surfaces (Figure 34). The DSM and DTM generated for each spectral band were merged to produce a multispectral raster image. The CHM, which is the height between the ground and the top of the rapeseed crop, was then calculated using the raster calculator tool in ArcGIS (CHM = DSM - DTM). Finally, a polygon was drawn to extract OSR height from a specific region.

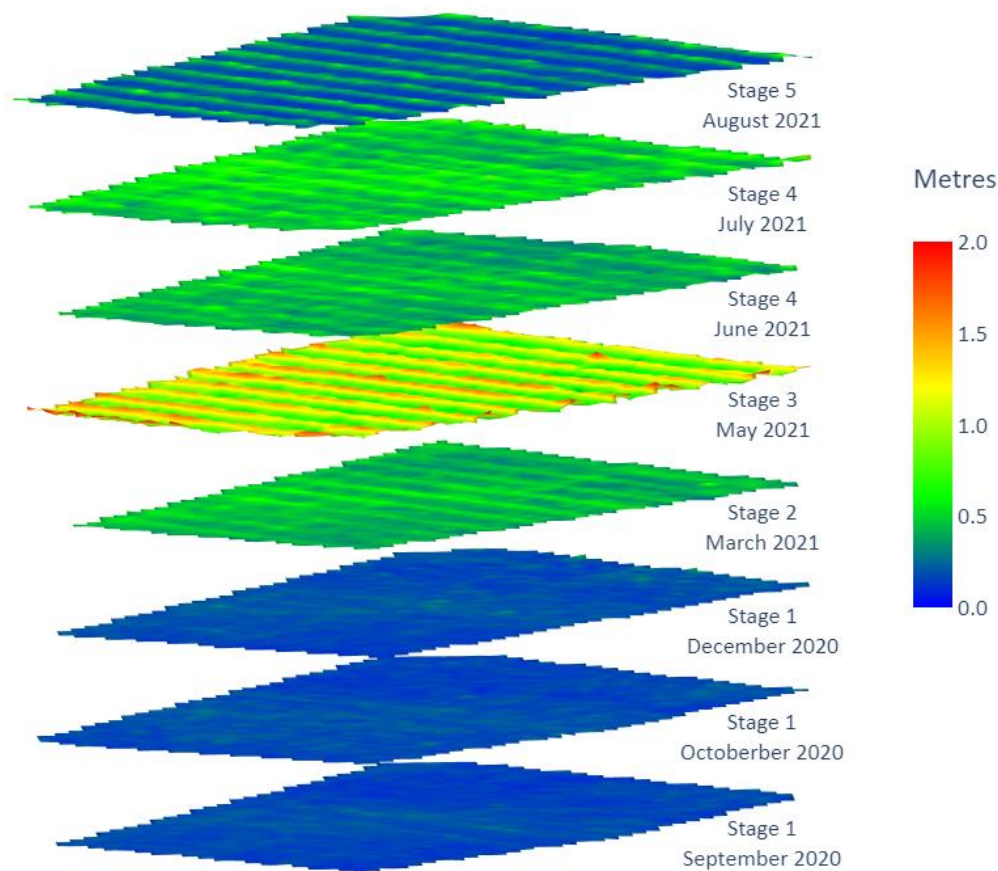


Figure 34: 3-Dimensional canopy height model of oilseed rape at the five phenological growth stages.

8.2.7 Vegetation Index Calculation and Iso-cluster Classification

ArcGIS Pro software was used for vegetation index calculation and iso-cluster classification (Figure 35). The NDYI (equation 18) raster was derived by calculation of the reflectance of green and blue orthomosaic image using the raster calculator tool.

$$\text{NDYI} = \frac{(\text{Green}-\text{Blue})}{(\text{Green}+\text{Blue})} \quad (18)$$

The NDYI captures the yellowness of the OSR flowers utilising the green and blue spectral bands. Therefore, the NDYI was calculated on two image datasets of OSR flowering in May (14 and 27 May 2021). The NDYI raster was further classified by iso-cluster classification. Iso-cluster classification is an unsupervised classification tool in ArcGIS pro which automatically groups similar clusters to output a classified image. It uses a similar principal to k-means clustering where centroids are placed according to the number of clusters assigned. The Euclidian distance for each pixel with the respective centroid is calculated. The pixels for each cluster are classified on its closest Euclidian distance into separate clusters which each cluster having a similar value. Five clusters were assigned to classify the NDYI raster. The first cluster had pixels representing the soils surface, while the second cluster include pixels representing green vegetation. The remaining three clusters represented rapeseed flower pixels of varying yellowness.

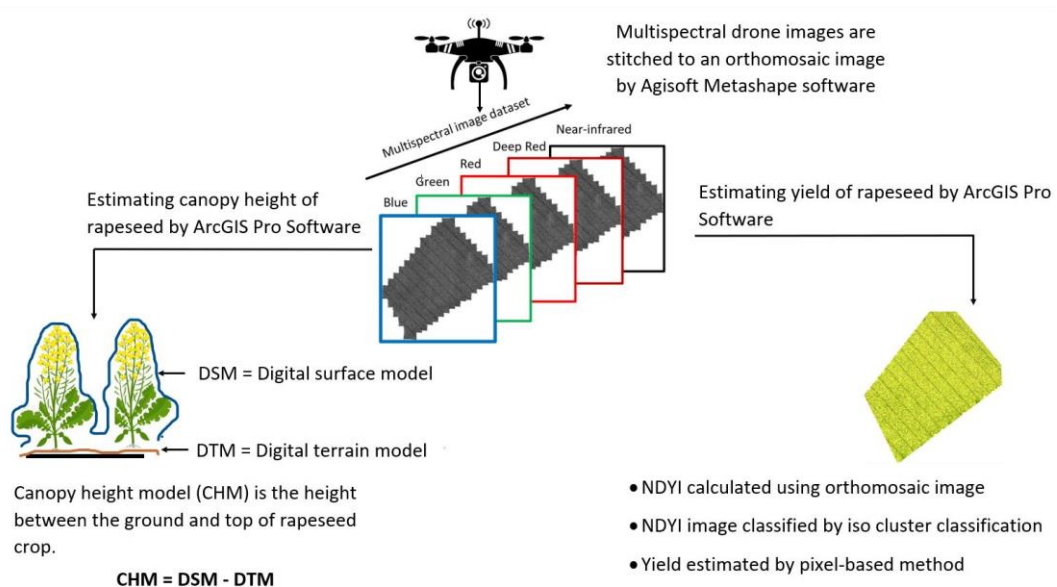


Figure 35:

Workflow of data processing to derive the canopy height model (CHM) and normalised difference yellow index (NDYI) maps.

8.2.8 Oilseed rape seed yield estimation

Around 70% of OSR flowers will convert into ripened pods, and this can be assessed against the final seed yield derived from the combine harvester processing. Therefore, pixels representing the OSR flowers by NDYI iso-cluster classification can be used to estimate seed

yield. Hence, a simplified pixel-based approach (as originally proposed by Shirsath et al., 2020) was used to segregate OSR flower pixels to estimate the seed yield, in units of t/ha) using the following equation:

$$\text{Estimated seed yield (t/ha)} = \frac{F \times P(\text{OSR})}{A} \quad (19)$$

Where, F is the weighting factor to scale the relationship. The higher the F value the higher the crop yield will be. The value for F is determined as follows: $F = \text{rapeseed flower pixels} / \sum \text{pixels}$; $P(\text{OSR}) = \text{The area (m}^2\text{) of OSR flowers in the field is calculated by multiplying the number of pixels of OSR flowers by the resolution of the drone images (i.e. (0.027 m / pixel)}^2\text{)}$; and $A = \text{Area of the field in ha}$.

According to the actual corrected seed yield map (Figure 36), supplied by FarmTRX (Troo Corp., Ottawa, Canada), some variation of yield is noted across the 8.15 ha. The yield map has variation across its area, highlighted as follows: low yield (red = 3.41t/ha), medium (orange = 3.89 t/ha), high (yellow = 4.22 t/ha) and very high (green > 4.61 t/ha) yield areas (Figure 36). As a result, and with reference to the corrected yield map (FarmTRX), the predicted seed yield based on the NDYI iso-cluster image was determined in 25 distinct places for the four identified yield sites (i.e. low, medium, high and very high yield) (Figure 37).

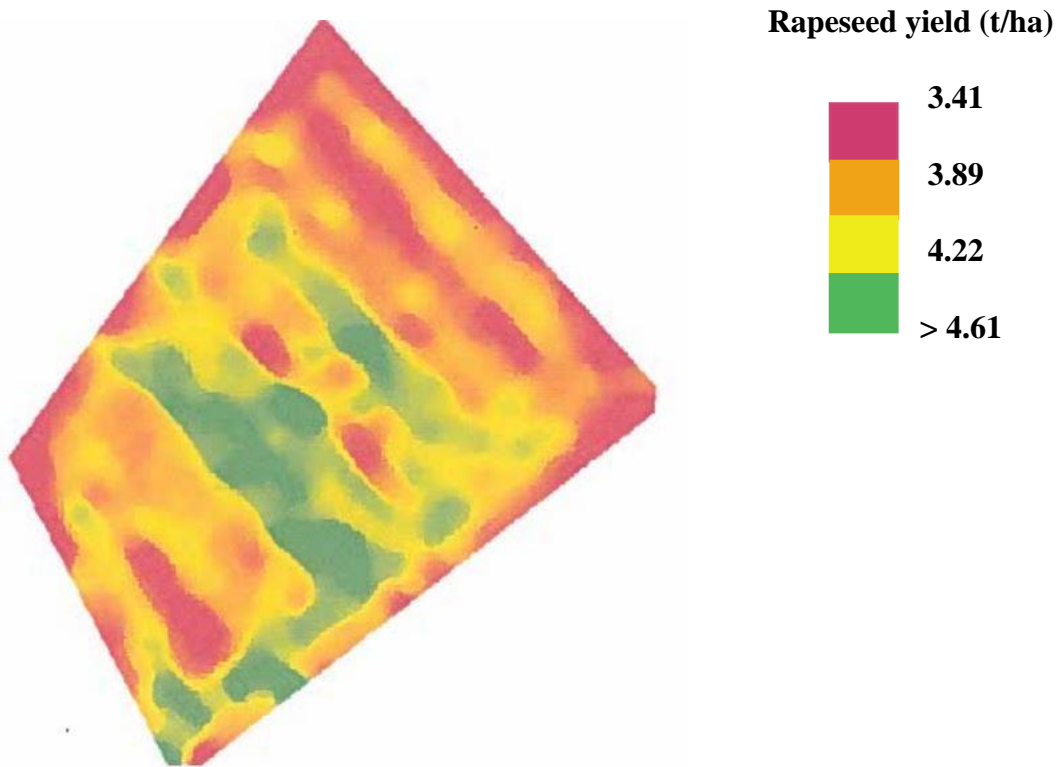


Figure 36: Actual seed yield map of oilseed rape

[Note: The reported field area is 8.15 ha (81,500 m²)].

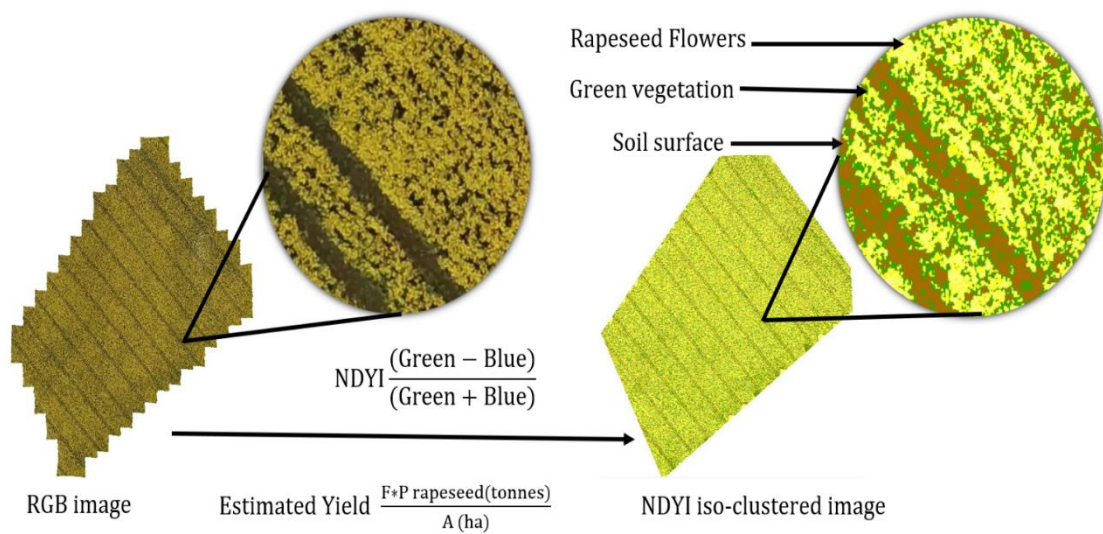


Figure 37: Normalised difference yellow index (NDYI) iso-clustered image representing the rapeseed flowers, green vegetation, and soil surface.

8.2.9 Statistical Analysis

To analyse the link between the estimated seed yield and the actual yield, Pearson correlation and root mean square error (RMSE) calculations were made. The larger R^2 and low RMSE will indicate the higher precision and accuracy of the estimated yield model.

$$\text{RMSE} = \frac{\sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2}}{(n)} \quad (20)$$

where y_i = estimated yield; \hat{y}_i = actual yield; and n = number of observations.

8.3 Results and Discussion

8.3.1 Estimation of OSR plant height by multispectral-UAV at the various phenological growth stages.

Multispectral-UAV estimates of plant height from the CHM were conducted on an 8.1 ha of OSR field (Figure 33). This study has no ground-truth data of OSR crop height, over the five stages, to validate the UAV data. Therefore, the UAV images were validated by measuring the height of the wall adjacent to the OSR field. Statistical analysis (t-test) was performed on manually measured wall heights against the UAV-MSI estimated wall height (Table 6). It was found that no statistical difference was noted between the two measurements (at a p-value of 0.17, 95% confidence interval), and the estimated OSR plant height measurements, as estimated from UAV-MSI images are used as accurate for the rest of the study.

A boxplot statistically representing the distribution of numerical data with minimum, median, mean and maximum OSR plant height across the field is shown in Figure 38. In stage 1 (Rosette), the OSR height has values of 0.092 m, 0.21 m and 0.35 m (minimum, mean and maximum). In stage 2 (Bolting), the OSR height has values of 0.29 m, 0.53 m and 0.83 m (minimum, mean and maximum). In stage 3 (Flowering), the OSR height has values of 0.39 m, 0.91 m and 1.35 m (minimum, mean and maximum). In stage 4 (Podding), the OSR height has values of 0.26 m, 0.58 m and 0.96 m (minimum, mean and maximum). In stage 5 (Ripening), the OSR height has values of 0.24 m, 0.62 m and 1.07 m (minimum, mean and maximum). It can clearly be seen (Figure 38) that the highest OSR plant height was achieved at stage 3, the flowering stage. Further, the decline of OSR height after flowering (stage 3) can be due to the challenging canopy architecture of the OSR plant. At the flowering stage, the canopy height is measured precisely due to the inflorescence which can be visible on aerial images. However, after the flowering stage, when the OSR starts podding, the pods become heavier, and the plants slouch whilst accommodating the extra weight. Therefore, it becomes challenging to estimate

the plant dimensions (rather than their height above ground level) accurately using the CHM generated from UAV images.

As reported in other studies, and in agreement with our data, the main stem of the OSR plant reaches between 30-60 % of its maximum length at stage 2 (Bolting) (Searby, 2020). Additionally, the OSR plant achieves its full height in the flowering stage (stage 3), with an average plant height of between 0.75 m to 1.75 m. Similarly, Zhang(a) et al., 2021, used a UAV with RGB sensors to estimate winter OSR crop height which they found varied between 0.80 m and 1.6 m. However, it should be noted that differences in OSR plant height data will vary, based on its geographical location along with variation in soil type and fertiliser and other treatment variations, as well as climatic conditions will all influence OSR growth and development. It is recommended therefore that localised geographical area phenotyping is done in relation to precision agriculture practices.

Table 6: Statistical analysis, using the student’s t-test, on actual and estimated wall height.

Location of wall height	Actual measured wall height (m)	Estimated wall height using UAV-MSI generated data (m)
1	0.96	0.96
2	1.17	1.04
3	1.09	1
Average	1.07	1
SD	0.04	0.11
P value (0.05) = 0.17		

Note: For each location the UAV-MSI data is calculated on three different days and represented as a mean value.

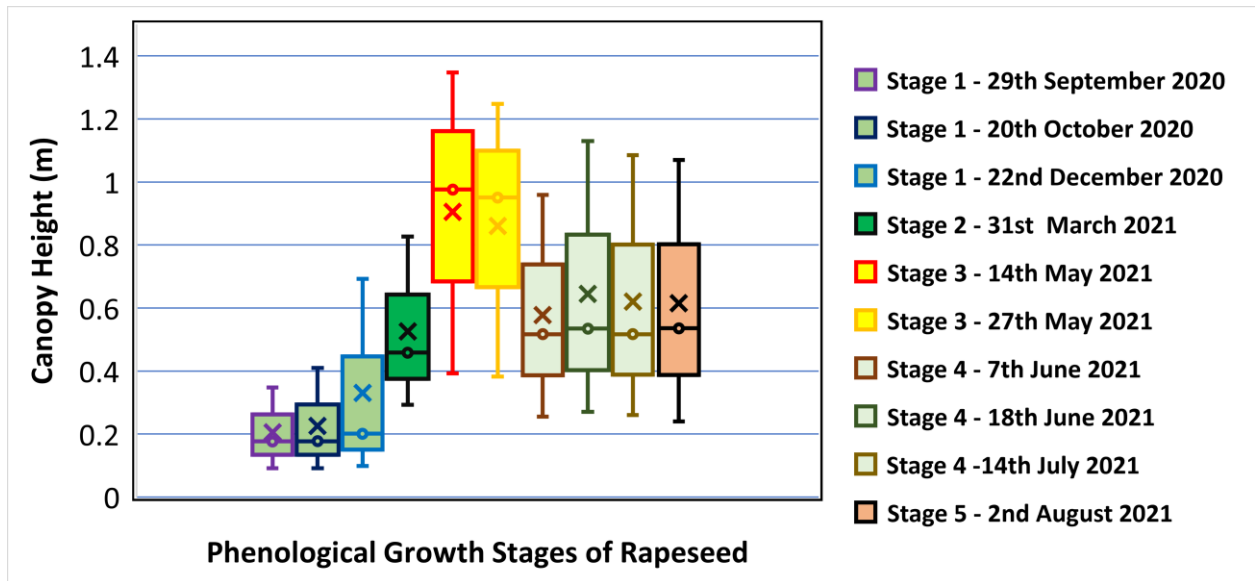


Figure 38: Determined oilseed rape canopy height at the five phenological growth stages.

8.3.2 Correlation study between estimated and actual seed yields

Initially, two dates in May 2021 (14 and 27 May) were chosen by visual inspection of the OSR inflorescence in the RGB images to determine the estimated seed yield (Figure 39). The NDYI maps were classified into five clusters by iso-cluster classification. The classified clusters were grouped as soil surface, green vegetation, and OSR flowers (Figure 37). If the NDYI value of a cluster pixel was more than 0.12 reflectance it was classified as OSR flowers. In this manner, three clusters were identified to have NDYI pixel values greater than 0.12. All three clusters were grouped as rapeseed flowers. Similarly, the NDYI iso-clustered pixel values for green vegetation was in the range between -0.15 to -0.45 and soil surface was in the range between -0.45 to -1. Using a similar approach on winter OSR (Tian et al., 2021), this method identified flowers with NDYI pixel values greater than 0.28. In contrast, their NDYI values were significantly lower for green vegetation and soil surface regions. The classified OSR flower pixels, by NDYI iso-cluster classification, was incorporated in equation 19, to estimate the OSR seed yield. The results (Figure 39) from our study demonstrate the strong correlation ($R^2 = 0.86$ and $RMSE = 0.49$) between the estimated seed yield and the actual OSR seed yield.

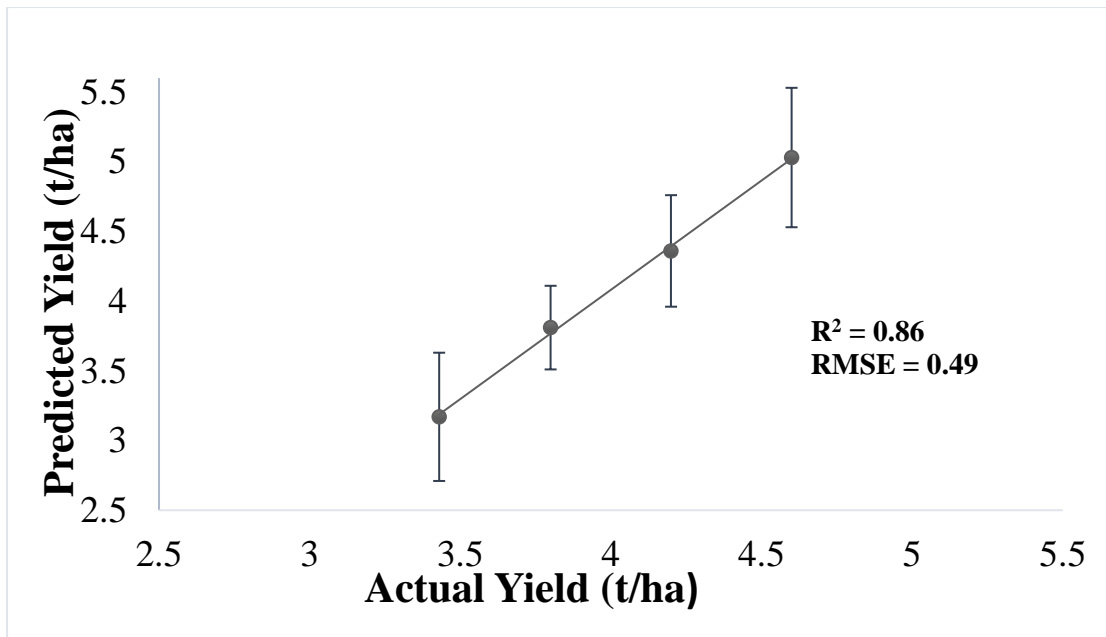


Figure 39: Correlation study between the predicted and actual seed yields.

8.4 Conclusion

The results of this work provide an effective and simple approach for classification of OSR flowers as an estimate of seed yield. Our new approach, using an automated iso-cluster classification model, has proven to be an effective estimator of seed yield and with good accuracy and precision. Future work will seek to evidence the robustness of the iso-cluster classification approach for assessing crop yield based on the flowering stage, for both the same and other widely grown crops. In addition, the effective use of a calibrated UAV-MSI system can estimate crop height to not only follow, over an 11-month period, the development of OSR, but also assess the point of optimum flowering for estimation of seed yield 3 months ahead of harvesting. The deployment of this approach in precision agriculture, without the need for challenging computer programming, could be a convenient and effective approach to be used by agronomists, as well as provide farmers with insight into estimated OSR seed yield 3 months ahead of harvesting. This could be an asset to farmers as it would allow the sale of the OSR crop in advance of harvesting.

Chapter 9: Multispectral UAV monitoring prediction of winter bean crop performance

Winter beans are a widely cultivated crop worldwide consumed as both food and animal feed. Limited research has been performed using multispectral UAV to monitor the winter beans crops. Hence, the focus research area on Chapter 9 for winter bean crop would be estimating the crop height across phenological growth stages and estimating yield at the vegetative growth stage prior to harvesting.

9.1 Aims and Objectives

9.1.1 Aims

The aim of this chapter is to estimate winter bean crop height throughout the phenological growth stages and yield in the flowering stage using three different vegetation indices.

9.1.2 Objectives

1. To evaluate the performance of a UAV-MSI camera to estimate winter bean crop development across the different phenological growth stages over an 11-month growth cycle.
2. To predict the final winter bean seed yield (in mid-August) using NDVI, GNDVI and NDRE coupled with an iso-cluster classification method, during the flowering stage (May),
3. To assess the accuracy of each vegetation index to predict the seed yield.

9.2 Experimental Design

9.2.1 Sampling site Airy Holme Farm

Winter bean seed (variety, Tundra) was planted at a rate of 220 kg / hectare using a Claydon Hybrid T4 trailed drill (Rickerby, Hexham, UK) pulled by a Claas ARES 836 RZ tractor (Rickerby, Hexham, UK) to a depth of 70 mm on the 29 September 2020 in a 5.29-hectare field, known locally as High Dowell (Figure 40a). The field was treated as follows in general terms, for protection of the winter bean crop: Within a week the sowing stage (early October), a specific herbicide was applied to the soil prior to crop and weed emergence (Clomate®, Albaugh Europe, Lausanne, Switzerland) alongside Most Micro, a pendimethalin herbicide, used for the control of annual grass and broadleaved weeds (Most Micro, Sipcam (UK) Ltd., Royston, Herts). In (early December) a foliar acting selective herbicide with systemic activity on a wide range of grass weeds and volunteer cereals (Falcon®, Adama Agricultural Solutions UK Ltd, Reading) was applied. A broad-spectrum systemic fungicide (Custodia, Adama

Agricultural Solutions UK Ltd, Reading) was applied in early June. Finally, 21 days prior to harvesting of the winter bean crop a pre-harvest desiccant was applied (Roundup Vista plus, Bayer Crop Science, Cambridge, UK). The addition of a desiccant assists in producing an evenly ripe crop (and eliminates perennial weeds) to aid harvesting.

To promote growth of winter beans fertiliser, NPK(S), 20-8-12(7SO₃) with the composition: total nitrogen (N) of which nitric nitrogen (9.2%) and ammoniacal nitrogen (10.8%); phosphorus pentoxide soluble in neutral ammonium citrate and in water (P₂O₅) 8% (3.5% P) of which phosphorus pentoxide soluble in water (P₂O₅) 7.7% (3.3% P); potassium oxide soluble in water (K₂O) 12% (10% K); sulfur trioxide soluble in water (SO₃) 7% (2.8% S) (CF Fertilisers UK Ltd., Billingham, UK) was applied at the time of sowing (29 September 2020). Additional nutrients were also added in spring (April) and included supplements for boron, manganese and molybdenum (Lebosol® Rapsmix SC, Lebosol® Dunger GmbH, Elmstein, Germany), phosphate, potassium and magnesium (Yaravita Magphos K, Yara Ireland, Grimsby, UK) and, in June, phosphorus and calcium (Calfite extra, Agrovista UK Ltd., Nottingham, UK). In addition, an adjuvant, Boost (Dow Agro Sciences, King's Lynn, UK) was added to assist with ground coverage, soil penetration, crop uptake and rainfastness.

9.2.2 High Dowell Agriculture field

Soil analysis was done on this agricultural field in March 2021 (by Lancrop Laboratories, Pocklington, UK in association with Agrovista UK Ltd., Nottingham, UK). Soil analysis indicated that the sandy loam with a pH of 7.6, organic matter (4.0%) and a cation exchange capacity of 15.1 meq/100 g had normal soil levels of the major nutrients i.e. phosphorus (33 ppm), potassium (164 ppm), magnesium (44 ppm) and sodium (17 ppm) alongside secondary and micronutrients i.e. calcium (3006 ppm), copper (3.3 ppm), iron (687 ppm) and zinc (6.2 ppm). However, deficiencies were noted in terms of secondary and micronutrients, specifically, sulfur (2 ppm) against a guideline of 10 ppm; boron (1.2 ppm) against a guideline of 1.6 ppm; manganese (32 ppm) against a guideline of 100 ppm; and molybdenum (0.07 ppm) against a guideline of 0.2 ppm. Therefore, supplements were added to the field to rectify these deficiencies. The desiccated crop was harvested on the 1 September 2021 using a Claas Lexion 570, Terra-Trac combine harvester (Rickerby, Hexham, UK). The seed is monitored by sensors (FarmTRX, Troo Corp., Ottawa, Canada), which record both the yield, and its location using GPS technology, into an on-board data logger.

9.2.3 Unmanned Aerial Vehicle

A multirotor UAV (DJI Phantom 4, Leeds, UK) was used with a multispectral camera with a 5 camera- for data collection. In all cases, the camera was angled perpendicular to the ground, with data capture occurring in hover and capture mode. Images (1548 image files per flight gathered over 256 waypoints) were captured as 16-bit TIF files corrected for ambient radiance values. The UAV speed was 5.0 m/s and had an average height of 50.6 m for the 2901 m flight distance. All flights were recorded with a resolution of 2.7 cm/px, a front overlap ratio of 75%, a side overlap ratio of 60% and a course angle of 90°. Specific weather conditions relating to daytime temperature during flight, wind speed and direction (recorded using a handheld anemometer (Benetech® GM816, Amazon UK)), and UAV pilot anecdotal observations on cloud coverage are identified with specific dated data.

9.2.4 UAV Photogrammetric Processing

The multispectral UAV images were used to build an orthomosaic image using Agisoft Metashape Professional (64 bit) software v.1.7.1 (Agisoft LLC, St. Petersburg, Russia). The aerial images were merged and aligned to create a sparse point cloud by matching similar image attributes during the UAV photogrammetric processing steps. The images were then precisely positioned to generate a 3D point cloud based on each image's GPS coordinate to create a solid mesh model. Finally, using the WGS 1984 Web Mercator coordinate system, an orthomosaic image (Figure 40b) was created.

(a)



(b)

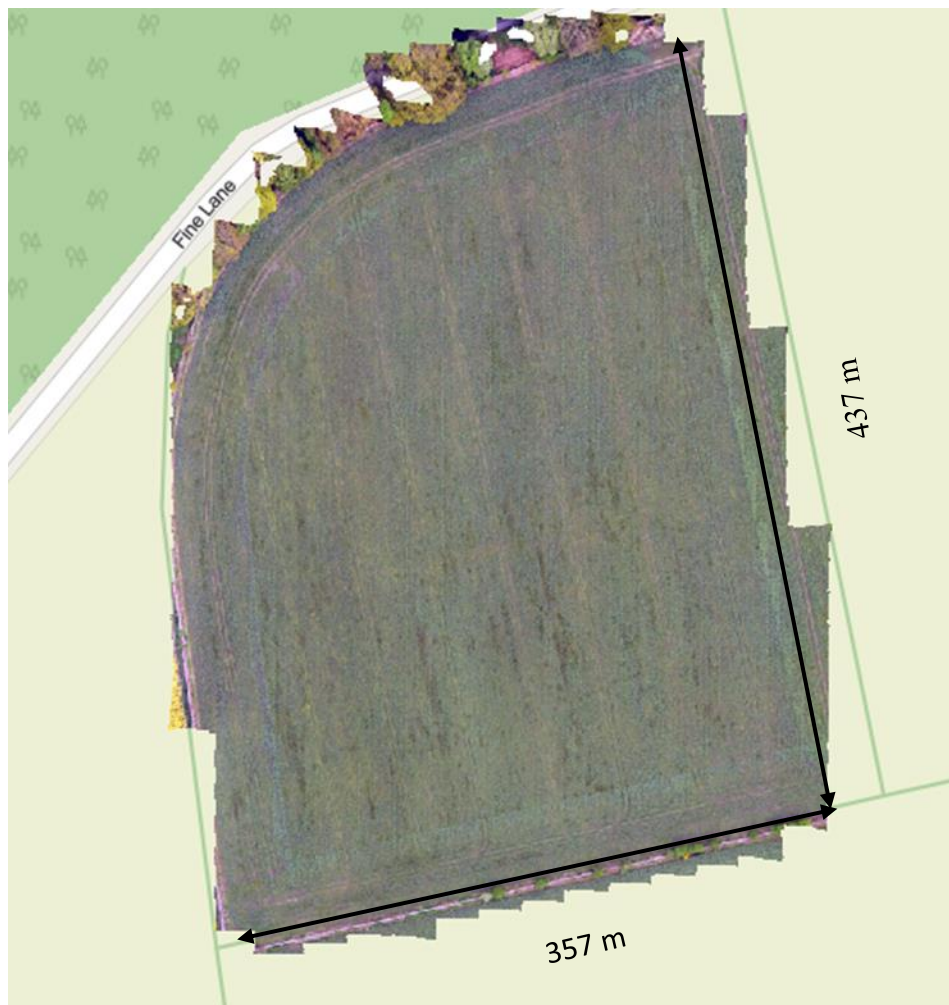


Figure 40: (a) Airy Holme site winter beans (b) Orthomosaic image of High Dowell: winter bean field, with an area of 5.29 ha.

9.2.5 Crop height Estimation

An estimation of the height of the winter bean crop, across the phenological growth stages, was performed on ArcGIS Pro v.2.8.0 software (Esri Inc., West Redlands, CA, USA). ArcGIS Pro software generates a CHM (Figure 41) based on the concept of structure from motion principle. The DSM and DTM for each spectral band were combined to create a multispectral raster image. The CHM, which is the height between the ground and the top of the winter bean crop was then calculated in ArcGIS using the raster calculator tool ($CHM = DSM - DTM$). Finally, a polygon was drawn to extract the winter bean height across the winter bean field (Figure 40).

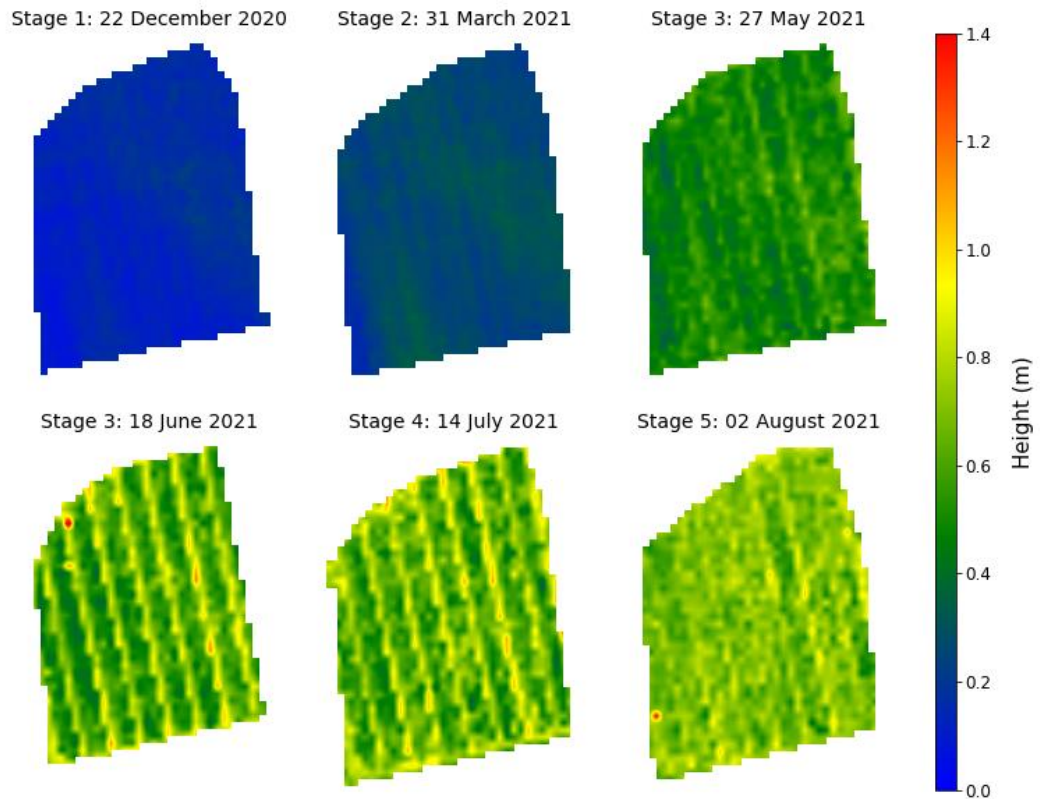


Figure 41: 3D Canopy height model of winter bean at different phenological growth stages.

9.2.6 Vegetation Index Calculation and Iso-cluster Classification

ArcGIS Pro software was used for vegetation index calculation and iso-cluster classification. The raster calculator tool in ArcGIS was used to generate specific vegetation index maps by applying three vegetation indices: NDVI, GNDVI, and NDRE to multispectral data. The NDVI spectral index was calculated by measuring the difference between the NIR and red spectral bands (Rouse et al., 1974). NDVI is a standardised graphical index which helps to quantify healthy vegetation with higher NDVI values and unhealthy vegetation with lower NDVI values. Whereas GNDVI was calculated by measuring the difference between NIR and green spectral bands (Gitelson and Merzlyak, 1996). GNDVI has a greater saturation point and is more sensitive to chlorophyll changes in the crop than NDVI and is useful in evaluating crop vigour in the early stages in dense canopies. While NDRE index was calculated by measuring the difference between NIR and red edge spectral bands (Gitelson and Merzlyak, 1994). Furthermore, NDRE can quantify higher amounts of chlorophyll within dense canopy layers more effectively than NDVI. NDRE is most suited for use during the mid and late growing season of crops.

Winter beans canopy were classified as NDVI values between 0.67 – 1, GNDVI values between 0.63 – 1 and NDRE values between 0.64 – 1. Grasslands were classified as NDVI values between 0.66 – 0.24, GNDVI values between 0.62 – 0.13 and NDRE values between 0.63 – 0.16. Soil surfaces were classified as NDVI values less than 0.24, GNDVI values less than 0.13 and NDRE values less than 0.16.

$$\text{NDVI} = \frac{(\text{NIR}-\text{Red})}{(\text{NIR}+\text{Red})} \quad (1)$$

$$\text{GNDVI} = \frac{(\text{NIR}-\text{Green})}{(\text{NIR}+\text{Green})} \quad (21)$$

$$\text{NDRE} = \frac{(\text{NIR}-\text{Deep Red})}{(\text{NIR}+\text{Deep Red})} \quad (22)$$

The vegetation indices were calculated on two gathered image datasets during flowering in May (14 and 27 May) 2021. Vegetation index maps were further classified by iso-cluster classification in ArcGIS Pro. The first cluster had pixels representing the soils surface, while the second cluster include pixels representing grasslands or shrubs. The remaining four clusters represented healthy vegetation pixels in winter bean crop.

9.2.7 Winter bean seed yield estimation

By calculating vegetation pixels from various locations in High Dowell can be used to estimate the final seed yield. The pixels representing the vegetation, determined by vegetation index maps, are combined with iso-cluster classification and used to estimate seed yield. Hence, a simplified pixel-based approach (as originally proposed by Shrishath et al;2022) was used to segregate healthy vegetation pixels in the winter bean crop to estimate the seed yield, in units of t/ha, using the following equation:

$$\text{Estimated seed yield (t/ha)} = \frac{F \times P}{A} \quad (19)$$

Where, F is the weighting factor to scale the relationship. The higher the F value the higher the crop yield. The value for F is determined as follows: $F = \text{healthy vegetation pixels} / \sum \text{pixels}$; $P = \text{area (m}^2\text{) of healthy vegetation pixels in the field}$. This is calculated by multiplying the number of pixels of healthy vegetation by the resolution of the drone images (i.e. $(0.027 \text{ m} / \text{pixel})^2$); and, $A = \text{Area of the field (ha)}$. According to the actual corrected seed yield map (Figure 42), supplied by FarmTRX (Troo Corp., Ottawa, Canada), some variation of yield is noted across the 5.29 ha. The yield map has variation across its area, highlighted as follows:

A. low yield (red with minimum = 0 t/ha, mean = 1.15 t/ha and maximum = 1.21t/h); B. medium (orange with minimum = 1.21 t/ha, mean = 1.28 t/ha and maximum = 1.33 t/h); C. high (yellow with minimum = 1.33 t/ha, mean = 1.38 t/ha and maximum = 1.42 t/h); and, D. very high (green with minimum = 1.4 t/ha, mean = 1.57 t/ha and maximum = 1.9 t/h) yield areas (Figure 42). As a result, and with reference to the corrected yield map (FarmTRX), the predicted seed yield based on the vegetation index map by iso-cluster image was determined in 25 distinct places within High Dowell for the four identified yield sites (i.e., low, medium, high and very high).

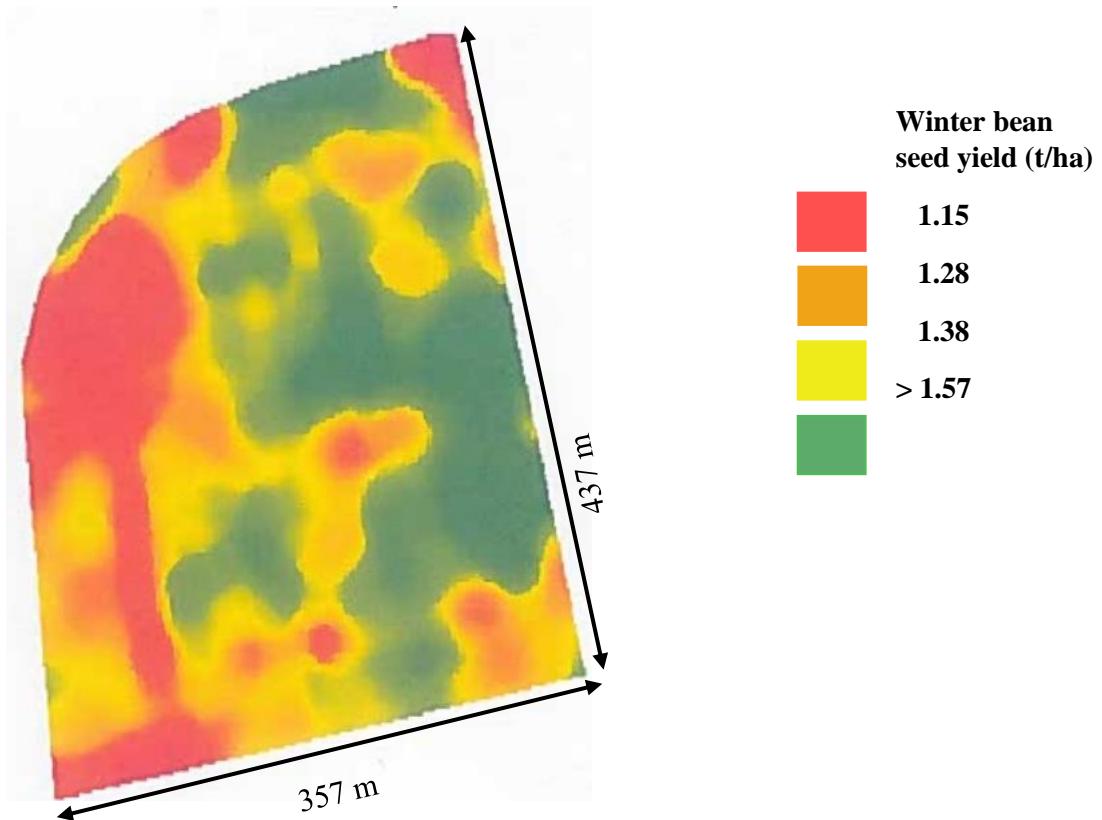


Figure 42: Mean actual yield map of winter bean seed yield

(Note: The reported field area is 5.29 ha (or 52,900 m²))

9.2.8 Statistical Analysis

To analyse the link between the estimated seed yield and the actual yield, Pearson correlation and root mean square error (RMSE) calculations were made. The larger R² and low RMSE will indicate the higher precision and accuracy of the estimated yield model.

$$RMSE = \frac{\sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2}}{(n)} \quad (20)$$

where y_i = estimated yield; \hat{y}_i = actual yield; and n = number of observations.

9.3 Results and Discussion

9.3.1 Estimation of winter bean development across the phenological growth stages.

The height of the winter bean crop in High Dowell (5.29 ha) across the phenological growth stages was estimated using the CHM. To corroborate the canopy height data of winter beans using UAV-MSI data treatment, ground truth data was obtained by measuring the height of the wall used as a boundary to the east of the field and compared with UAV-MSI estimated data. The manually measured wall heights (1.07 ± 0.04 m, $n = 3$) compared favourably with the UAV-MSI CHM estimated wall height (1.00 ± 0.11 , $n = 3$, as determined on three different days) using statistical analysis (t-test). The t-test results indicated a p-value of 0.17 at the 95% confidence interval. It was concluded that there was no statistically significant difference between the two measurements (ground truth versus UAV-MSI CHM data). Therefore, the estimated winter bean plant height measurements, estimated from UAV-MSI CHM, were used as obtained for the rest of the study (Table 7).

A boxplot illustrating the phenological growth stages of winter beans from December 2020 to August 2021 is shown in Figure 43. In stage 1 (emergence), the mean winter bean height is 0.14 m (minimum 0.046 m and maximum 0.24 m). In stage 2 (plant development), the mean winter bean height is 0.27 m (minimum 0.11 m and maximum 0.37 m). In stage 3 (flowering), the mean winter bean height is 0.60 m (minimum 0.30 m and maximum 0.82 m). In stage 4 (pod development), the mean winter bean height is 0.60 m (minimum 0.27 m and maximum 1.0 m). Finally, in stage 5 (ripening), the mean winter bean height is 0.71 m (minimum 0.37 m and maximum 1.2 m). The results from the study (Figure 43) indicate a steady increase in winter bean plant height until the final stage 5 of ripening. The biggest growth in plant height occurs between stages 2 (plant development) and 3 (flowering) where the average plant height has increased by nearly 33% from 0.27 m to 0.6 m (Figure 43). This growth is facilitated by the addition of nutrients, in the form of boron, magnesium, manganese, molybdenum, phosphate and potassium, to the soil. Further, between stage 3 (Flowering) until stage 5 (Ripening) there has only been a slight increase of average winter bean plant height by 10% from 0.6 m to 0.7 m (Figure 43). Generally, plant growth slows down after the flowering stage as the plant prepares itself for pod development and ripening. Furthermore, in this study, the maximum winter bean height by stage 5 (ripening) is 1.2 m (Figure 43) with a similar plant height of 1.12 m for Tundra winter bean reported by PGRO (Limagrains UK, 2019). However,

differences in winter bean plant height will vary depending on geographical location, climate conditions, soil type and fertiliser and other treatment variations. It is therefore recommended that localised geographical area phenotyping is carried out in relation to precision agriculture practises.

Location of wall height	Actual measured wall height (m)	Estimated wall height using UAV-MSI generated data (m)
1	0.96	0.96
2	1.17	1.04
3	1.09	1
Average	1.07	1
SD	0.04	0.11
P value (0.05) = 0.17		

Table 7: Statistical analysis, using the student’s t-test, on actual and estimated wall heights.

Note: For each location the UAV-MSI data is calculated on three different days and represented as a mean value.

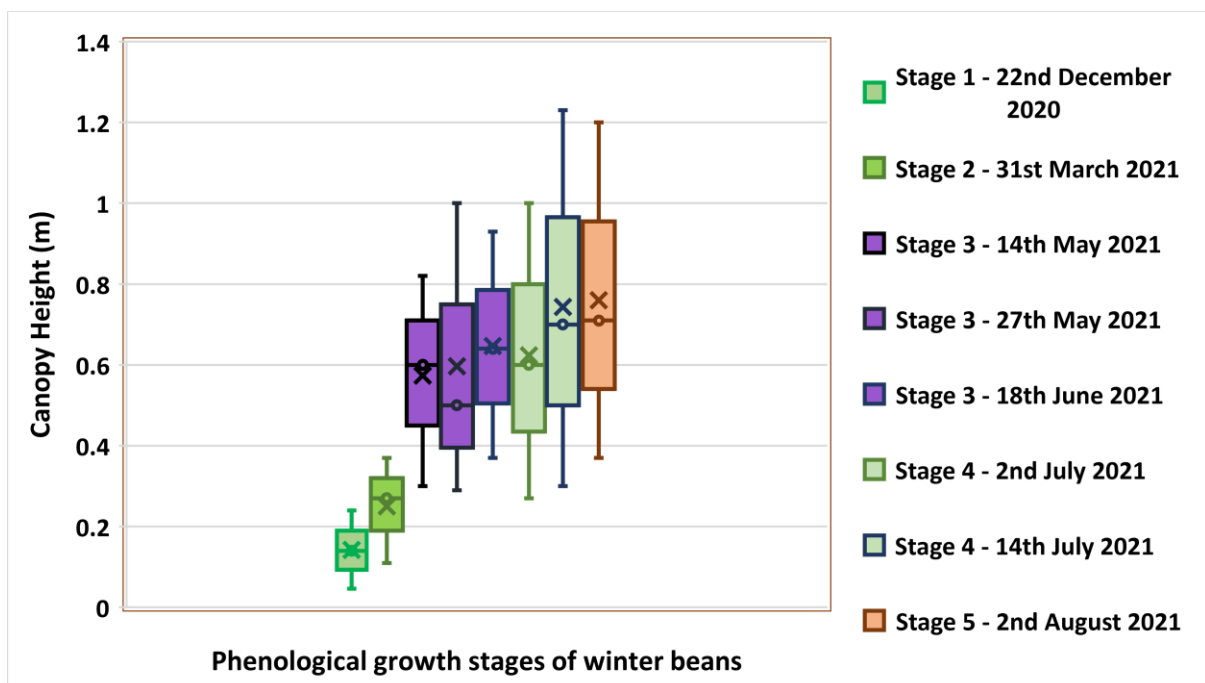


Figure 43: Winter bean canopy height at phenological growth stages.

9.3.2 Prediction of seed yield

Initially, two dates in May 2021 (14 and 27 May) were chosen to estimate seed yield during the growing peak period (Stage 3, flowering) of the winter bean crop. The vegetation index maps were classified into six clusters by iso-cluster classification. The classified clusters were grouped as soil surface, grassland or shrub, and a healthy winter bean crop (Figure 44). The healthy winter bean crop had an average NDVI value of 0.5, a GNDVI value of 0.41 and an NDRE value of 0.45. These are considered moderate values due to the sparse nature of the vegetation of winter bean crop across the High Dowell field (Figure 44(a)). Generally, higher values of NDVI, GNDVI and NDRE (typically between 0.6 to 1) are found in dense vegetation due to the high levels of chlorophyll. The final seed yield of the winter bean crop was estimated by calculating the pixel areas of healthy vegetation within the crop classified by iso-cluster classification using NDVI, GNDVI and NDRE vegetation index maps (Figure 44). The NDVI iso-cluster classified image (Figure 44(b)) shows a better classification for the winter bean crop vegetation from the surrounding grassland or shrub and soils surfaces in comparison to the GNDVI and NDRE (Figure 44(c) and 44(d)) iso-cluster classified images.

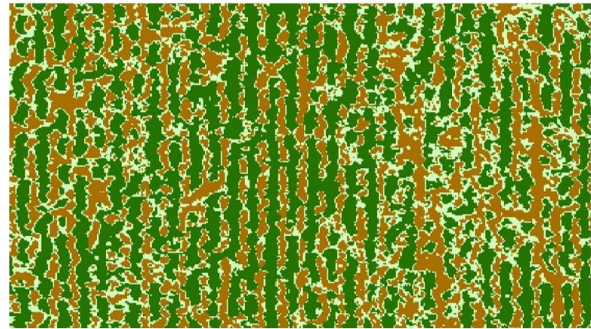
A comparison between the estimated yield by NDVI has been done with the actual seed yields within the High Dowell site (Figure 42) i.e. A = low yield (mean 1.15 t/ha), B = medium yield (mean 1.28 t/ha), C = high yield (mean 1.38 t/ha), and D = very high yield (mean >1.57 t/ha). A strong correlation is noted between the estimated and actual seed yields across the two days investigated in May 2021 (Figure 45). The correlation coefficient and RMSE by NDVI for the 14 May was an R^2 of 0.84 with an RMSE of 0.32 (Figure 45(a)) and for the 27 May was an R^2 of 0.87 with an RMSE of 0.53 (Figure 45(b)). Whereas the estimated yield by GNDVI and NDRE was significantly lower and showed weak correlation between the estimated and actual seed yield (Figure 46 and 47). The correlation coefficient and RMSE from GNDVI for the 14 May was $R^2 = 0.66$ with an RMSE of 0.92 (Figure 46(a)), whilst from the 27 May was $R^2 = 0.66$ with an RMSE of 0.57 (Figure 46(b)). For the NDRE, the correlation coefficient for the 14 May was $R^2 = 0.55$ with an RMSE of 0.42 (Figure 47(a)) and for the 27 May the R^2 was 0.7 with an RMSE of 0.62 (Figure 47(b)).

However, for very high yield area (D in Figure 42) the estimated yield by NDRE was significantly better than the estimated NDVI (Figure 47(a) and 47(b)). A similar observation

was observed for the high yield area (D in Figure 42) estimated by GNDVI on 27 May (Figure 46(b)). This is because light in the green and red edge spectral band can deeply penetrate the leaf layers compared to blue or red light, in both cases underestimating the yields. It is worth noting that both NDRE and GNDVI will be more sensitive and hence provide better indicators of yield for crops with dense vegetation in their later growth stages that have accumulated high levels of chlorophyll in their leaves. In contrast, NDVI has shown to be a better indicator of vegetation for crops with sparse vegetation i.e. winter beans.

(a)

(b)



(c)

(d)

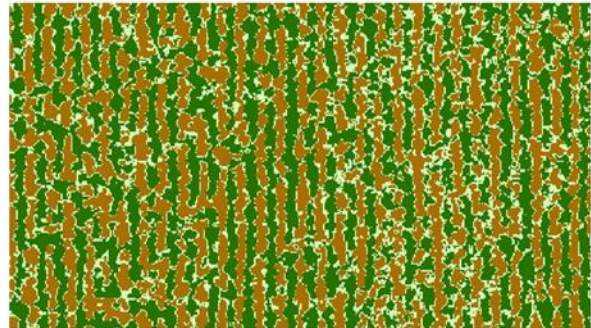
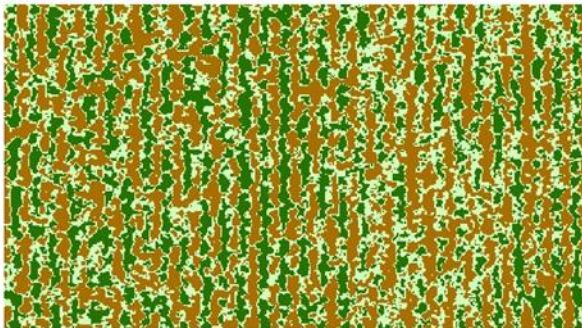
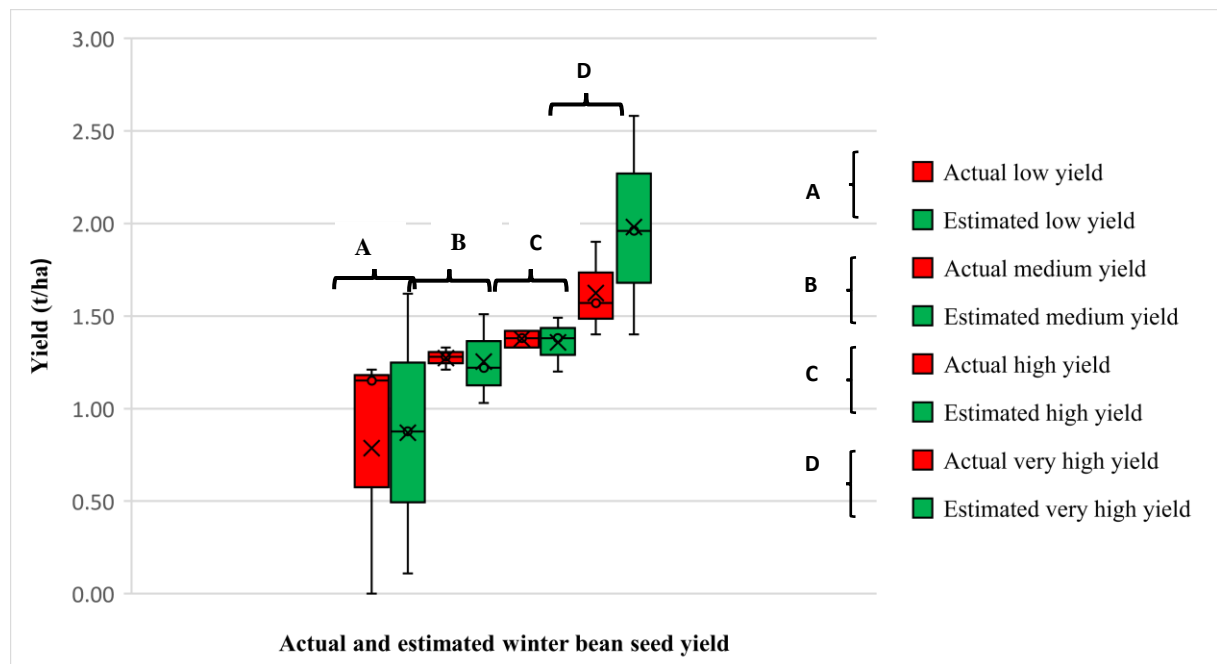


Figure 44: (a) RGB image (b) NDVI iso-clustered image (c) GNDVI iso-clustered image (d) NDRE iso-clustered image. Note for (b)-(d): Dark green represents healthy winter bean crop. Light green represents grasslands or shrubs. Brown represents soil surface.

(a)



(b)

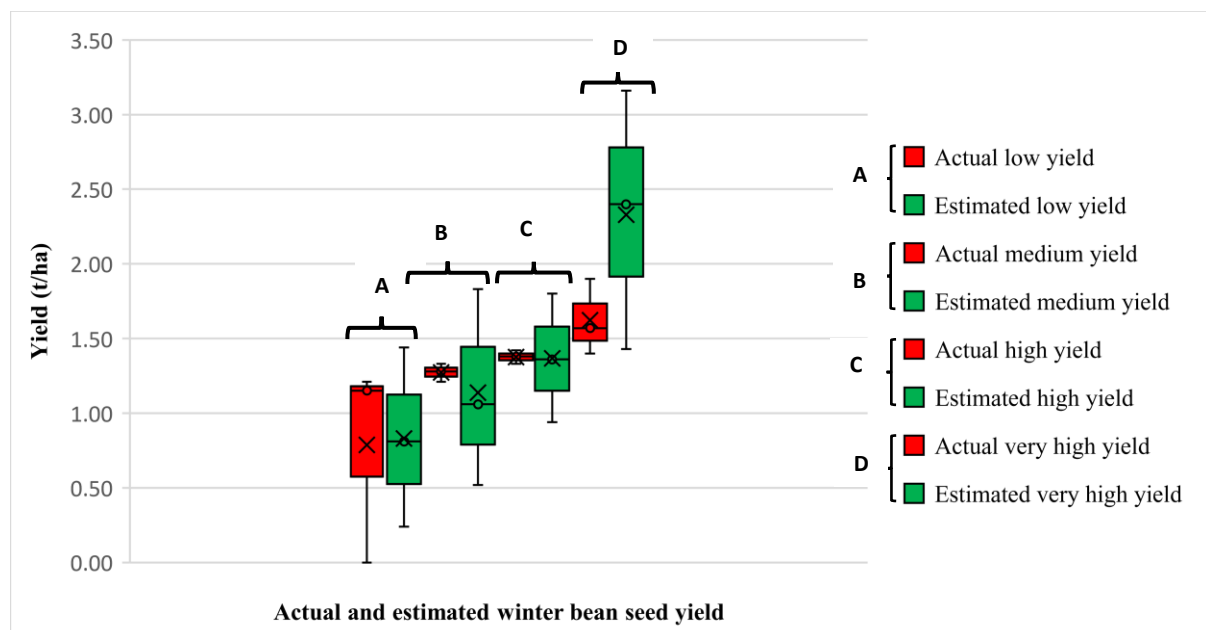
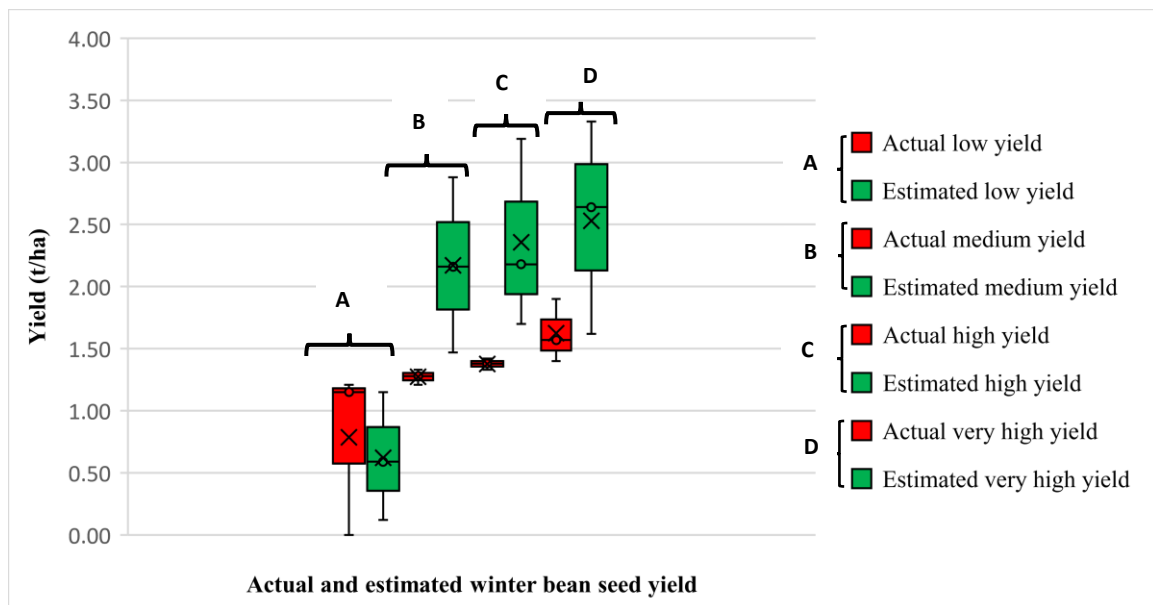


Figure 45: (a) Comparison between actual and estimated winter bean yield by NDVI on (a) 14 May 2021, and (b) 27 May 2021.

[Correlation coefficient and RMSE for 14 May 2021 ($R^2 = 0.84$ and $RMSE = 0.32$) and 27 May 2021 ($R^2 = 0.87$ and $RMSE = 0.53$)]

(a)



(b)

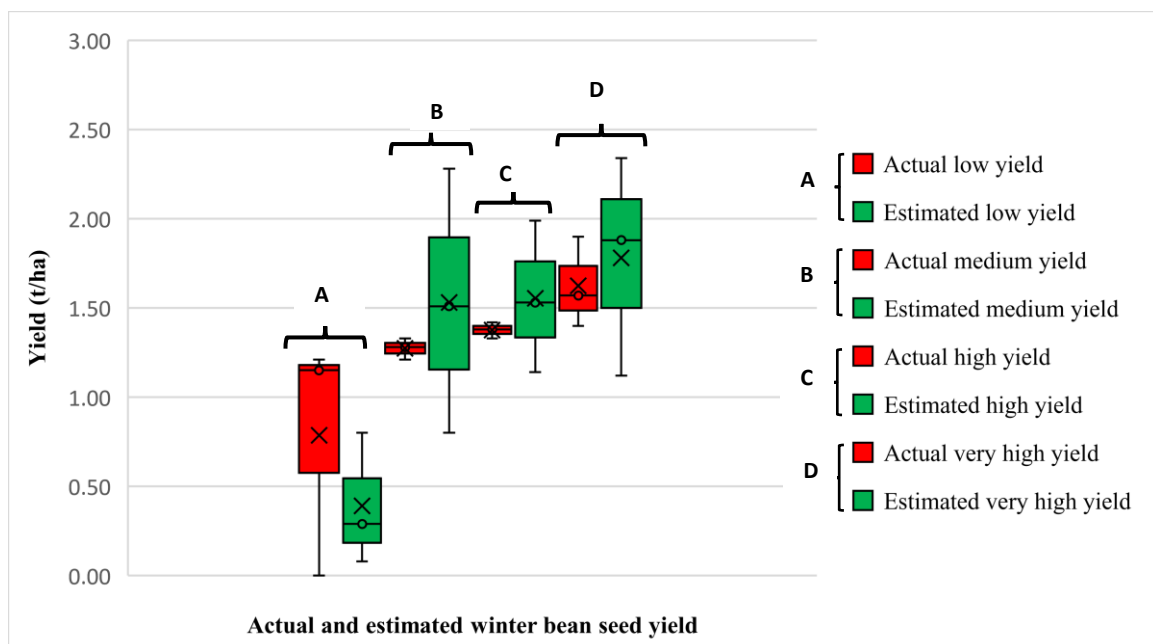
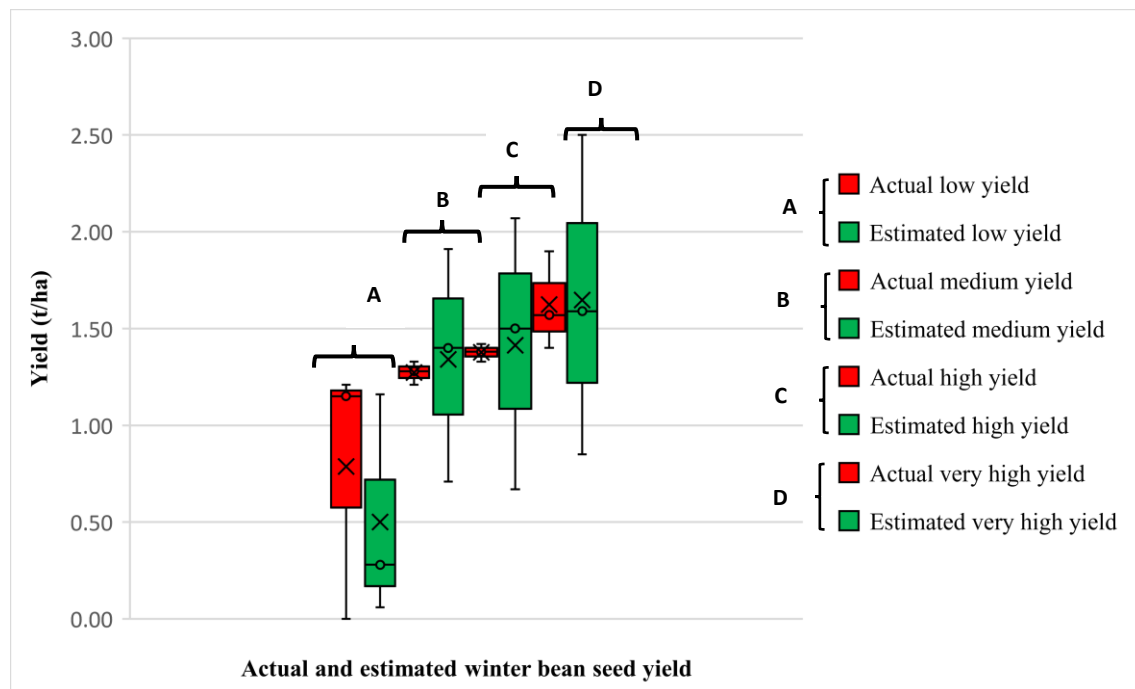


Figure 46: Comparison between actual and estimated winter bean seed yield by GNDVI on (a) 14 May 2021, and (b) 27 May 2021.

[Correlation coefficient and RMSE for 14 May 2021 ($R^2 = 0.66$ and $RMSE = 0.92$) and 27 May 2021 ($R^2 = 0.66$ and $RMSE = 0.57$)]

(a)



(b)

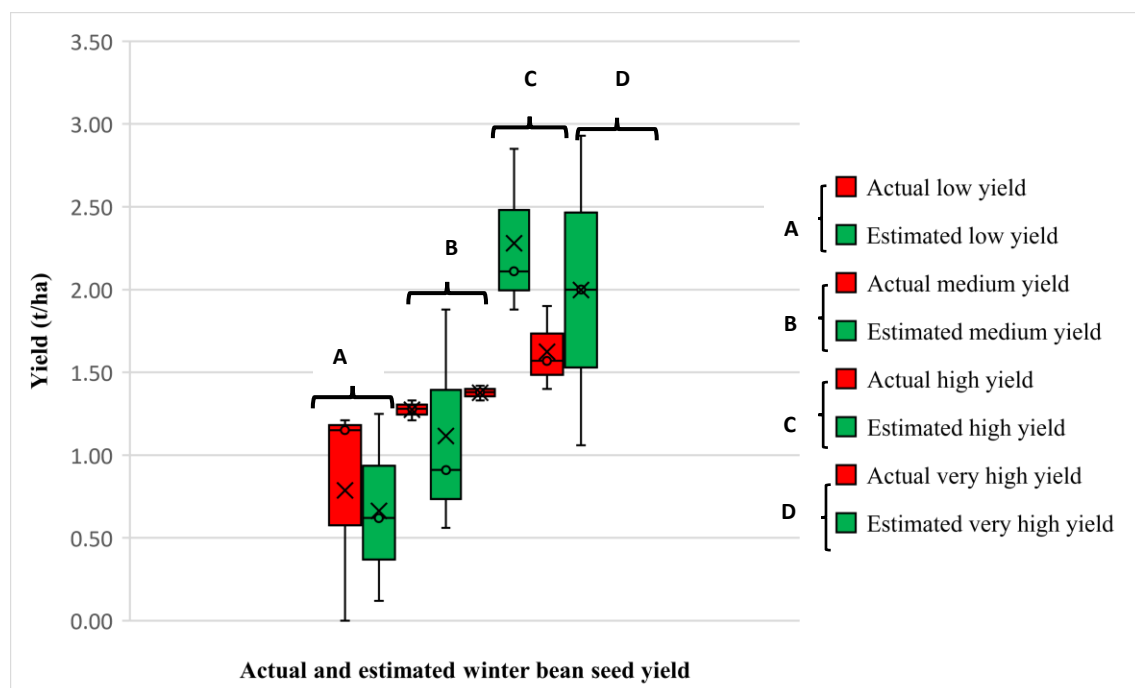


Figure 47: Comparison between actual and estimated winter bean seed yield by NDRE on (a) 14 May 2021, and (b) 27 May 2021.

[Correlation coefficient and RMSE for 14 May 2021 ($R^2 = 0.55$ and $RMSE = 0.42$) and 27 May 2021 ($R^2 = 0.70$ and $RMSE = 0.62$)]

9.4 Conclusion

This research has shown that a canopy height model, resultant from UAV-MSI data, can effectively predict the crop height of winter beans over the phenological growth stages (9 months). Secondly, the study demonstrates a simple and effective method for estimating winter bean seed yield at the flowering stage (Stage 3 in the phenological growth stages) 3 months prior to harvesting, using three different vegetation indices. Finally, NDVI has been shown to be a good estimator of winter bean seed yield, due to its sparse vegetation. Therefore, this new approach for winter beans can provide insights to the farmer, and agronomist, in terms of crop development, and inform application of soil nutrients, as well as estimate the final seed yield in advance of its ripening and harvesting using a direct, automated approach with commercially available software that does not require challenging coding.

Chapter 10: Multispectral UAV monitoring prediction of winter oats crop performance

10.1 Aims and Objectives

Winter oats is a widely cultivated crop used for both animal and human consumption playing a vital role in ensuring world food security. Hence, remote sensing technology is required for the precision agriculture of crops to increase the crop yield and the meet the global need. The remote sensing technique of multispectral UAV for monitoring winter oats crop would be discussed in Chapter 10.

10.1.1 Aims

The aim of this chapter is to estimate winter oats crop height, oats grain yield by vegetation indices and identify the relationship between oats grain yield and chemical soil health indicators at different growth stages.

10.1.2 Objectives

1. To investigate and statistically compare the performance of multiple vegetation indices in analysing winter oats crop at three different phenological growth stages (flowering, grain filling, and ripening) using a portable hand-held spectrometer and multispectral UAV.
2. To identify the relationship between chlorophyll sensitive VI of NDVI and CI green from multispectral-UAV and chemical soil health indicators (NPK, pH and soil organic matter) across different growth stages of winter oats crop.
3. To evaluate the performance of UAV multispectral imaging to estimate winter oats canopy height across the different phenological growth stages and identify the effect of chemical soil health indicators on canopy height.
4. To estimate the final oats grain yield (in late August) using vegetation indices (NDVI and CI green), coupled with a simplified pixel based iso cluster classification method, during the flowering stage (early June).
5. To compare the estimated winter oats grain yield with the actual winter oats grain yield (late August) measured in the laboratory, and to generate thematic maps of the estimated winter oats grain yield by a statistical method of kernel density estimator.
6. To use hierarchical logistic regression supervised machine learning model to identify the relationship of oats grain yield with chemical soil health indicators (NPK, pH and soil organic matter) and oats quality (beta-glucan) at different winter oats growth stages.

10.2 Experimental Design

10.2.1 Airy Holme Farm

Winter oats seed (variety, Mascani) was planted at a rate of 155 kg / hectare using a Claydon Hybrid T4 trailed drill (Rickerby, Hexham, UK) pulled by a Claas ARES 836 RZ tractor (Rickerby, Hexham, UK) to a depth of 25-30 mm on the 28 September 2020 in a 3.58-hectare field (Lat. 54.880690; Long. -1.915923), known locally as Copse (Figure 49a). The field was treated as follows in general terms, for protection of the winter oats crop: Within a week of the sowing date (early October), a specific herbicide (containing flufenacet and diflufenican) was applied to the soil prior to crop and weed emergence (Herold®, Adama Agricultural Solutions UK Ltd, Reading) alongside Grounded AD (Agrovista UK Ltd., Nottingham, UK) a paraffin oil-based concentrate designed to minimise drift of ground applied pesticides. In spring 2021 (late April), further herbicides were applied simultaneously to control broadleaved weeds i.e. Duplosan KV (containing mecoprop-P), Paramount Max (containing florasulam and tribenuron-methyl) (both from NuFarm UK Ltd., Bradford, Yorkshire) and Gal-Gone (containing fluroxypyr) (Belchim Crop Protection, St Neots, Cambridgeshire). In mid-May the fungicide Protefin (containing prothioconazole and tebuconazole) (Clayton Plant Protection, Dublin, Ireland) was applied. In mid-June, additional fungicides were applied as Aderya® XE (containing fluxapyroxad and mefentrifluconazole) (BASF Agricultural Solutions UK, Littlehampton, Sussex) and Tebucur 250 (contains tebuconazole) (Belchim Crop Protection, St Neots, Cambridgeshire) alongside the growth regulator Canopy, which contains mepiquat chloride and prohexadione calcium) (BASF Agricultural Solutions UK, Littlehampton, Sussex). Finally, 21 days prior to harvesting of the winter oats crop a pre-harvest desiccant was applied (Roundup Vista plus, Bayer Crop Science, Cambridge, UK). The addition of a desiccant assists in producing an evenly ripe crop (and eliminates perennial weeds) to aid harvesting.

To promote growth of winter oats fertiliser, NPK(S), 20-8-12(7SO₃) (CF Fertilisers UK Ltd., Ince, Cheshire) with the composition: total nitrogen (N) of which nitric nitrogen (9.2%) and ammoniacal nitrogen (10.8%); phosphorus pentoxide soluble in neutral ammonium citrate and in water (P₂O₅) 8% (3.5% P) of which phosphorus pentoxide soluble in water (P₂O₅) 7.7% (3.3% P); potassium oxide soluble in water (K₂O) 12% (10% K); sulfur trioxide soluble in water (SO₃) 7% (2.8% S) (CF Fertilisers UK Ltd., Billingham, UK) was applied at the time of sowing (28 September 2020) at a rate of 104 kg / hectare. Additional fertiliser, NPK(S), 20-8-12(7SO₃) was also applied on 6 March 2021 (rate: 225 kg / hectare), 30 March 2021 (rate: 200

kg / hectare); and 13 April 2021 (rate: 200 kg / hectare). A final addition of muriate of potash (Origin Fertilisers, Royston, Hertfordshire) was applied in July 2021 (rate: 60 kg / hectare). In mid-June, 3 ALO t6p, which contains a precursor of the sugar trehalose-6-phosphate was applied to regulate metabolic and development processes within the crop (Agrovista UK Ltd., Nottingham, UK). In addition, an adjuvant was applied in mid-May and mid-June, Boost (Dow Agro Sciences, King's Lynn, UK) to assist with ground coverage, soil penetration, crop uptake and rainfastness. The desiccated crop was harvested on the 13 August 2021 using a Claas Lexion 570, Terra-Trac combine harvester (Rickerby, Hexham, UK). The seed is monitored by sensors (FarmTRX, Troo Corp., Ottawa, Canada), which record both the yield, and its location using GPS technology, into an on-board data logger.

10.2.2 Copse agricultural field

Soil analysis was done on this agricultural field in March 2021 and data reported on the 15 March 2021 (by Lancrop Laboratories, Pocklington, UK in association with Agrovista UK Ltd., Nottingham, UK). Soil analysis indicated the following characteristics, a sandy silt loam (sand 41.3%; silt 47.3%, and clay 11.4%) with a pH of 7.5, organic matter (4.7%) and a cation exchange capacity of 19.0 meq/100 g. The field had an adequate level of the major nutrient, phosphorus (34 ppm), against a guideline value of 16 ppm. A normal level of potassium (176 ppm) against a guideline value of 121 ppm, and magnesium (48 ppm) against a guideline value of 50 ppm. For secondary and micronutrients the following levels were noted (with guidance) for calcium (3761 ppm) against a guideline value of 1600 ppm (adequate), sulfur (3 ppm) against a guideline value of 10 ppm (consider treatment), sodium (13 ppm) against a guideline value of 90 ppm (not an issue for oats), boron (1.16 ppm) against a guideline value of 1.60 ppm (consider treatment), copper (3.5 ppm) against a guideline value of 4.1 ppm (priority for treatment), iron (614 ppm) against a guideline value of 50 ppm (adequate), manganese (69 ppm) against a guideline value of 95 ppm (priority for treatment), molybdenum (0.01 ppm) against a guideline value of 0.20 ppm (low priority for oats), and zinc (6.2 ppm) against a guideline value of 4.1 ppm (adequate).

10.2.3 Unmanned Aerial Vehicle

For data collection, a multispectral camera with five sensors was used with a multirotor UAV (DJI Phantom 4, Leeds, UK). Images were recorded as 16-bit TIF files that were adjusted for ambient radiance values (1554 image files per flight collected over 256 waypoints). The UAV travelled 2901 metres at a speed of 5.0 m/s and an average height of 50.6 m. With a resolution

of 2.7 cm/px, a front overlap ratio of 75%, a side overlap ratio of 60%, and a course angle of 90°, all flights were captured. Specific weather conditions including daytime temperature during flight, wind direction and speed (recorded using a handheld anemometer; Benetech® GM816, Amazon UK), and anecdotal cloud coverage observations from the UAV pilot were recorded on the specific date.

10.2.4 Collection of crop phenotypic data: ground reference data

For ground truth measurements of oats crop, a Spectro 1 (Variable, Inc., Chattanooga, TN, USA; (Figure 48a) pocket-size, portable hand-held spectrometer was used. Spectro 1, has an 8 mm measurement aperture which allows spectrophotometric measurements in the visible region from 400 to 700 nm at 10 nm intervals. For this experiment 48 location in the oats field was selected for spectrophotometric measurements over three different growth stages of oats crops (flowering (June), grain filling (July) and full ripening (August)). However, due to the greater number of data points, 9 locations were used for data analysis. Five replicate scans of oats crops were collected from each location. Prior to each measurement, the spectrometer was calibrated on a white plate according to the manufacturer's instruction (Figure 48b). Further, since the device is easily operated with the Spectro application (Variable, Inc., Chattanooga, TN, USA) on the user's smartphone, the collected reflectance data was automatically uploaded and stored in the manufacturer's cloud storage service.

(a)



(b)



Figure 48: (a) Spectro 1 hand-held spectrometer (b) Calibration plates

10.2.5 Comparison between UAV and ground reference data

For ground reference data GRVI was calculated by using the reflectance data of green and red provided by the Spectro 1 spectrometer. The orthomosaics generated by Agisoft Metashape software for UAV data are not in reflectance but as "reflectivity" represented by digital numbers normalised by the information provided by the irradiance sensor mounted on top of the multispectral UAV. Hence, to normalise the data to get the reflectance values between 0 to 1 a pseudo correction was performed. Based on the UAV image histogram the maximum DN value was divided with each spectral band by using the raster calculator tool in ArcGIS Pro software for vegetation index calculation (Appendix 2 Figure A 2.1).

The GRVI vegetation index raster was derived by calculation of the reflectance of green and red orthomosaic image. GNDVI was derived by calculation of the reflectance of green and NIR orthomosaic image. NDVI was derived by calculation of the reflectance of red and NIR orthomosaic image. NDRE was derived by calculation of the reflectance of red-edge and NIR orthomosaic image. SAVI was derived by calculation of the reflectance of red, NIR orthomosaic image and a soil brightness correction factor of (L) defined as 0.5. CI green was derived by calculation of the reflectance ratio of NIR and green orthomosaic image as below:

$$\text{NDVI} = \frac{(\text{NIR} - \text{Red})}{(\text{NIR} + \text{Red})} \quad (1)$$

$$\text{GNDVI} = \frac{(\text{NIR} - \text{Green})}{(\text{NIR} + \text{Green})} \quad (21)$$

$$\text{NDRE} = \frac{(\text{NIR} - \text{Rededge})}{(\text{NIR} + \text{Rededge})} \quad (22)$$

$$\text{GRVI} = \frac{(\text{Green} - \text{Red})}{(\text{Green} + \text{Red})} \quad (23)$$

$$\text{SAVI} = \frac{(\text{NIR} - \text{Red})}{(\text{NIR} + \text{Red} + \text{L})} * (1 + \text{L}) \quad (24)$$

$$\text{CI green} = \frac{(\text{NIR})}{(\text{Green})} - 1 \quad (25)$$

10.2.6 Assessment of chemical soil health indicators

10.2.6.1 Soil sample collection

Soil samples from oats field were collected from 6 rows naming from A-F (figure 49). Using a measuring tape, 8 locations from each row were marked with a 30 m distance between each data location. From each location surface soil between 0-10 cm were collected using a small spade. The spade was cleaned with antibacterial wipes in between sample collection. A total of 48 samples were collected from the 6 rows (figure 49b). The collected soil samples were stored in brown paper bag and labelled for soil analysis by Palintest kit (Palintest UK, Kingsway, Team Valley, England). Furthermore, the samples were collected during the first week of June, July, and August 2021.

(a)



(b)



Figure 49: (a) Airy Holme Site winter oats field (b) Soil sample collection location in winter oat Copse agricultural field.

10.2.6.2 Analysis of NPK and pH by Palintest kit

The soil Palintest kit consists of the Soiltest 10 Bluetooth photometer (Figure 50), multiparameter pH probe, sample collection and preparation equipment's and soil test reagents of soil nutrient extraction powders and tablets. Appendix 2.2 presents a detailed method for measuring soil nitrogen, phosphorus, potassium, and pH using the Palintest kit, as well as an analysis of soil organic matter (SOM).



Figure 50: Palintest soil test kit.

10.2.7 Collection of crop yield and height data

10.2.7.1 Manual oat grain yield measurement

The yield of oats grain was calculated by harvesting a 1m² area of oats on 4 August 2021 from the 48 locations when the oats were fully ripe and ready to harvest (Figure 49). Since, the oats grain yield calculation from the entire 1m² would be time consuming a subsample of 20 g (three replicates) from each 48 sample was used. Hence, the initial total weight of the oats (husk + oats grain) separated from the stalk were weighed for 48 samples. Followed by this, the husks

from the oats were separated for the 20 g subsamples manually by hand and weighed to determine the oats grain weight. Then the measured subsample oats grain weight was used to estimate the final oats grain weight for 48 samples in the 1m² (Appendix 2 Table 2.3). Finally, the oats grain yield was calculated and represented as t/ha.

10.2.7.2 Oat β -glucan extraction

The detail method of oats β -glucan extraction is in Appendix 2.2.

10.2.7.3 Analysis of glucose in oat β glucan by colorimetric assay

Analysis of glucose was performed by using a glucose colorimetric assay kit named as Randox (Randox Laboratories Ltd, County Antrim, UK) and the detail analysis is in Appendix 2.2

10.2.7.4 Calculation of oat β -glucan

Six standard glucose curves were developed by using a glucose concentration ranging from 0.1 mM to 40 mM. The absorbance of the glucose standards was measured after the colorimetric assay by randox kit. Hence, glucose concentration as x in the sample was measured by incorporating the absorbance values for the y in the linear regression $Y = 0.05973x - 0.00645$ ($R^2 = 0.99$) developed using the mean y-intercept and slope of six glucose calibration curves (Appendix 2.4). The free glucose ($M_r = 180$) determined are adjusted by multiplying by a conversion factor of 0.9, to account for the difference in molecular weight of glucose monomers bound in a polysaccharide of β -glucan ($M_r = 162$) (Danielson et al., 2010). Therefore, the percentage β -glucan in oat was determined by dividing the amount of glucose measured in each sample by the fresh weight of oat sample as below:

Initially, the mass of glucose in 66 μ L of sample was calculated using the measured glucose concentration.

Number of moles = Glucose concentration * volume in 66 μ L

Mass of glucose in 66 μ L (g) = moles * molecular weight of glucose ($M_r = 180$)

$$\% \beta - \text{glucan} = \frac{\text{Mass of glucose in 66 } \mu\text{L (g)} * \text{DF} * \text{CF}}{\text{Fresh weight of oats sample (g) (0.2g)}} * 100$$

$$\text{Dilution factor (DF)} = \frac{2000 \mu\text{L}}{66 \mu\text{L}} = 30$$

$$\text{Conversion factor (CF)} = \frac{162 (\text{Beta-glucan polysaccharide})}{180 (\text{Glucose})} = 0.9$$

10.2.7.5 Crop Height Estimation

Agisoft's 3D point clouds were extracted and used by ArcGIS Pro to generate CHMs based on time-series data on oat phenological growth stages. The height between the ground and the top of the oat crop, CHM, was calculated in ArcGIS using the raster calculator tool (CHM = DSM - DTM). Finally, to extract plant height from a specific region, a polygon was drawn.

10.2.7.6 Winter Oat yield Estimation by UAV

The oat crop pixel extracted by NDVI and CI green iso-cluster classification by ArcGIS pro can be used to estimate oat yield. As a result, Shirsath et al., 2020 proposed a simplified pixel-based approach to segregate rapeseed flower pixels to estimate rapeseed yield using the following equation:

$$\text{Estimated Yield} = \frac{F * P \text{ oats } (t)}{A \text{ (ha)}} \quad (19)$$

F is the weighting factor to scale the relation. The higher the F value the higher the crop yield will be.

$F = \text{extracted oats pixel} / \sum \text{pixels}$

P oats = The area of oats in the field is calculated by multiplying the pixel of oats by the resolution of drone images (0.027m/pixel*0.027m/pixel). Divide the oats area by 1000 to estimate the oats production in tonnes.

A = Area of the field in ha.

As a result, the F value was multiplied by the amount of oats produced and divided by the field area to estimate the yield value at the pixel level in t/ha. As a result, the oat yield for 1 m² was calculated using the VI iso-cluster classified image on the 48 location points.

$n = \text{number of observations}$

10. 2.7.7 Thematic yield maps using Kernel density estimation.

KDE is a non-parametric spatial analysis method of estimating probability function or regression function (Chen, 2017). Hence, KDE mathematical function can be defined as kernel density estimator \hat{f} at location \mathbf{x} taken from a set of data (e.g. crop yield) $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2 \dots \mathbf{x}_n)$ of N samples with an unknown probability function $\mathbf{f}(\mathbf{x})$ and defined as a kernel estimate $\hat{\mathbf{f}}(\mathbf{x})$:

$$\hat{f}(x) = \frac{1}{Nb} \sum_{N=i}^N K(di) \quad (26)$$

$$\hat{f}(x) = \frac{1}{Nb} \sum_{N=i}^N K\left(\frac{X_i - x}{b}\right) \quad (27)$$

N = number of samples, di = distance between two points (e.g., $X_i - x$) b = bandwidth ($b > 0$, a positive number that defines the smoothness of a density plot and K = denotes kernel function

Among the various types of symmetric kernel functions found in the literature include Epanechnikov, Quartic, Triangular, Gaussian, and Rectangular (Wglarczyk, 2018). ArcGIS software employs the Quartic (Biweight) kernel function, which reveals more detailed information about a data point than the other symmetric kernel function.

Hence, the Quartic kernel function can be defined as:

$$K(u) = \int_0^{\frac{15}{16}} (1 - di^2)^2, \quad \text{for } |di| \leq 1 \text{ and } |di| \geq 1 \quad (28)$$

Hence, thematic maps of the estimated oats grain yield measured in the laboratory, by NDVI and CI green for 48 locations in $1m^2$ were created using the KDE function in ArcGIS pro software.

10.2.8 Statistical Analysis

To compare the accuracy of the vegetation indices of UAV data ($n = 5$) to ground reference data ($n = 5$), statistical t-test was performed using Excel software (version 365). If $p < 0.05$ the VI means are statistically significant indicating there is a difference between the VI means from UAV data and ground reference data. Likewise, if $p > 0.05$ the VI means are statistically insignificant indicating there is no difference between the VI means from UAV data and ground reference data. The Pearson correlation coefficients between UAV-derived vegetation indices (NDVI and CI green) and chemical soil health indicators were calculated (NPK, pH and soil organic matter using Excel software (version 365). The Pearson correlation coefficients of the results were tabulated. To estimate soil nitrate, simple regression models were developed. The training data sets, which comprised 80% of the total data, were used to develop regression models and determine the best fit between soil nitrate concentration and UAV-based VI. The performance of regression models was cross validated further using a 20% test data. This was done by incorporating the UAV-based VI (from the remaining 20 % dataset) into the regression

model equation to estimate the soil nitrate concentration. Afterwards, the estimated and measured soil nitrate level was correlated using Pearson correlation. Pearson correlation and root mean square error (RMSE) calculations were performed to comprehend the connection link between the estimated yield and the actual yield. The greater the R^2 and lower the RMSE, the greater the precision and accuracy of the estimated yield model.

$$\text{RMSE} = \frac{\sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2}}{n} \quad (20)$$

y_i = estimated yield

\hat{y}_i = actual yield

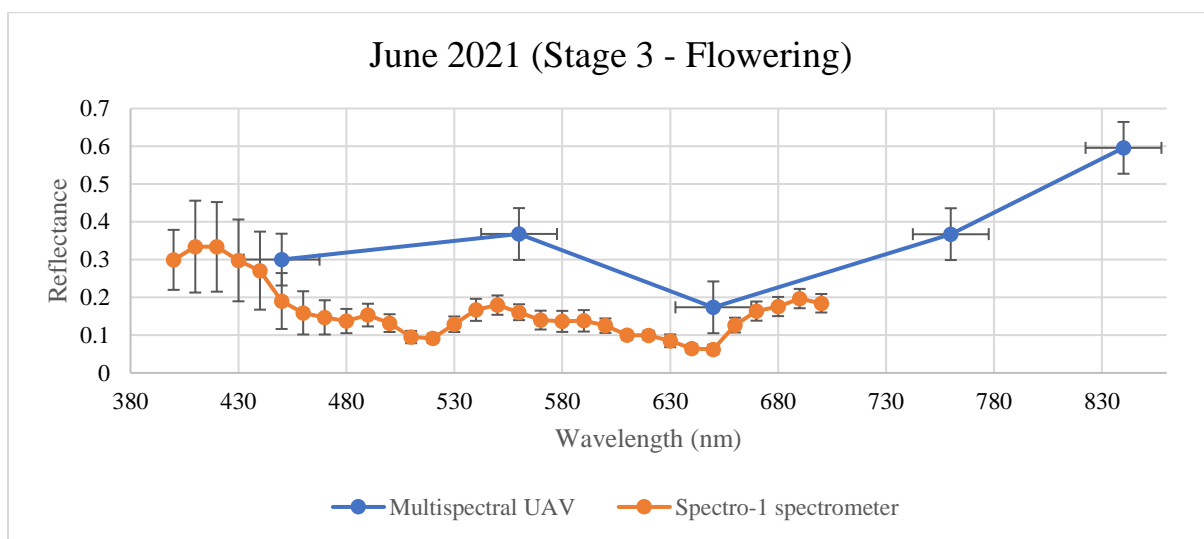
In addition, the measured and UAV oats grain yield data was used to build a hierarchical multinomial logistic regression model. The model was developed on Matlab R2021a. The dependent variable in this study is the yield of oat. Soil concentration of nitrate, P, K, pH, SOM and β -glucan concentration were used as independent variables in the model.

10.3 Results and Discussion

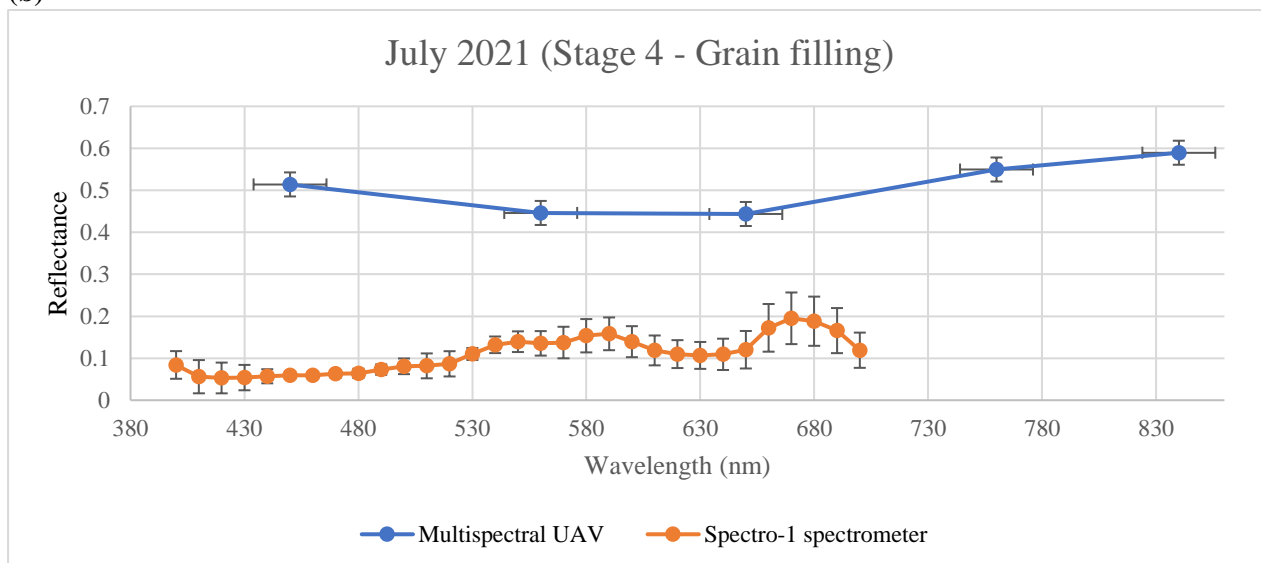
10.3.1 Comparison of spectral profile between UAV and ground reference data

The UAV reflectance spectral profile data shows variation with the hand-held spectrometer data over the three-month growth stages (Figure 51). Initially, there is an overestimation of reflectance values in the green (560 nm) and red (650 nm) spectral bands, with a high estimate of reflectance in the green band in UAV data for the month of June in comparison to spectrometer reflectance values (Figure 51a). Secondly, for July and August data the overestimation of reflectance values is higher in all spectral bands in comparison to the spectrometer reflectance values (Figure 51b and 51c).

(a)



(b)



(c)

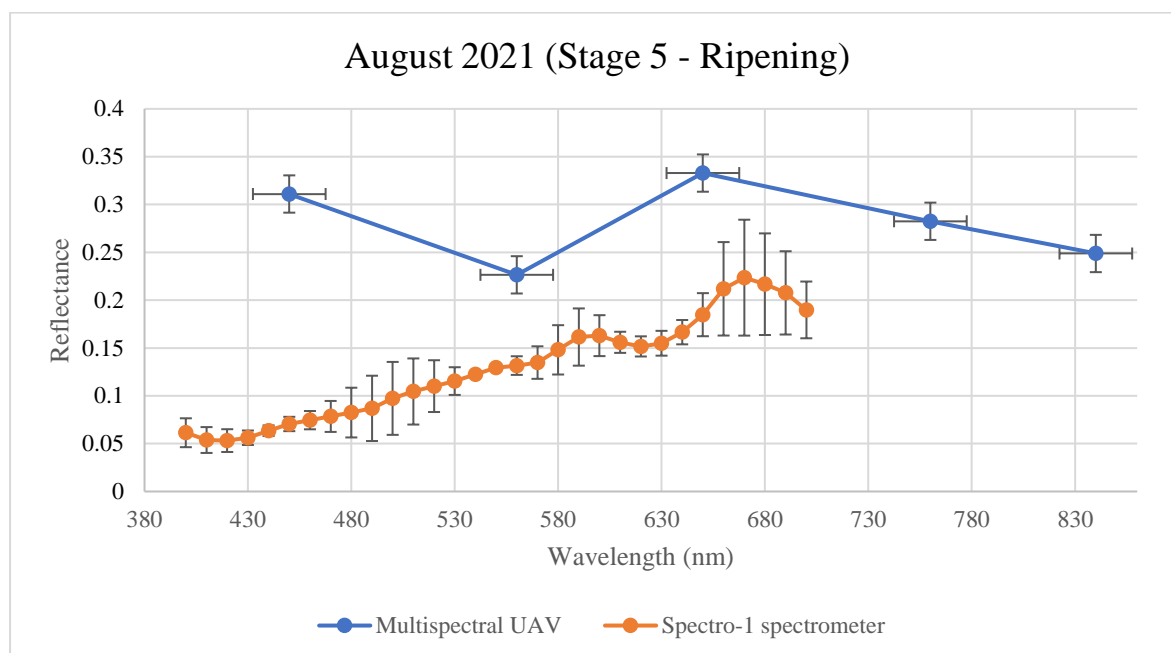


Figure 51: Comparison of spectral profile in reflectance between multispectral UAV (blue) and ground reference data by Spectro-1 spectrometer (orange) across three growth stages of winter oats. (a) June (Stage 3- flowering) (b) July (Stage 4 – Grain filling) (c) August (Stage 5 - Ripening).

Regarding, the higher reflectance values by UAV in comparison to the ground truth measurements performed with spectrometer have been previously reported to cause by the flight height in which the UAV data are collected (Mamaghani and Salvaggio, 2019). Higher altitude between the sampling location and the sensor of the remote sensing platform could have an impact on the quality of the data. Also, the variability can be justified due to the opposite data acquisition approach by the two methods (Di Gennaro et al., 2022). Data collected by spectrometer method is proximal and static, whereas UAV method is remote and dynamic. Further, Stow et al., 2019 reported illumination geometry had a visible effect on the reflectance values in spectral bands. Accuracy of reflectance values in spectral bands was dependent on the time during the day in which the data was collected. The most suitable time for data collection during the day was during morning as data collected between 14:50 and 16:00 influenced the reflectance values due to appearance of shadows (Stow et al., 2019). This

could explain the higher reflectance values in all spectral bands in the July (Figure 51b) month as the data on this day was collected between 15.44 to 16.01. Whereas the data collected in June and August were during morning and such higher reflectance values are not observed in these two data sets in comparison to July.

10.3.2 Statistical comparison of vegetation index from UAV and ground reference data

The Spectro 1 spectrometer can provide spectral profile only in the visible region. Therefore, only GRVI was calculated using Spectro 1 data. However, the GRVI from the spectrometer was compared not only to the GRVI measurements from the UAV, but also to all the VI measured by the UAV. The statistical comparison of UAV and ground reference data with respect to some vegetation index is represented in Table 8. For the month of June and July there is a statistical difference ($p < 0.05$) between the means of GRVI values from UAV and ground reference data, except for location (D3) in June and (A6 and C3) in July. However, in August data there is no statistical difference ($p > 0.05$) between the means of GRVI values from UAV and ground reference data, except for only 3 locations (B6, D5 and D3).

Table 8: Statistical comparison of vegetation indices between multispectral UAV and ground reference data by Spectro-1 spectrometer in (a) June (Stage 3- Flowering), (b) July (Stage 4 – Grain filling), and (c) August (Stage 5 - Ripening)

(a)

Sample	Spectro-1 spectrometer (GRVI)	MSI-UAV (GRVI)	P-value	MSI-UAV (NDVI)	P-value	MSI-UAV (GNDVI)	P-value	MSI-UAV (SAVI)	P-value	MSI-UAV (NDRE)	P-value	MSI-UAV (CI green)	P-value	Repeats
A6	0.5262	0.3300	0.0001	0.6295	0.0000	0.5983	0.0137	0.6544	0.0001	0.4926	0.0473	2.7459	0.0000	5
A4	0.4602	0.3512	0.0015	0.6673	0.0002	0.4855	0.1196	0.6545	0.0002	0.4457	0.2429	1.8112	0.0000	5
A2	0.4567	0.3561	0.0074	0.7385	0.0000	0.4674	0.3098	0.5077	0.0101	0.4706	0.2171	3.5470	0.0000	5
B6	0.5210	0.3957	0.0176	0.6296	0.0208	0.4821	0.1866	0.4984	0.2845	0.4303	0.0288	1.8456	0.0000	5
C6	0.4501	0.2873	0.0076	0.6668	0.0041	0.4803	0.2314	0.5644	0.0215	0.4791	0.2351	2.4125	0.0000	5
C3	0.5235	0.2486	0.0009	0.7139	0.0078	0.4901	0.2358	0.5556	0.2372	0.4081	0.0288	3.0511	0.0000	5
C1	0.4497	0.3548	0.0313	0.8138	0.0022	0.6409	0.0033	0.5743	0.0145	0.4149	0.1989	4.6234	0.0000	5
D5	0.5142	0.3300	0.0000	0.7257	0.0000	0.6511	0.0003	0.5872	0.0036	0.4196	0.0076	2.1216	0.0000	5
D3	0.3311	0.4509	0.0984	0.6722	0.0492	0.5448	0.0773	0.5358	0.0816	0.3941	0.2214	2.5926	0.0000	5

(b)

Sample	Spectro-1 spectrometer (GRVI)	MSI-UAV (GRVI)	P-value	MSI-UAV (NDVI)	P-value	MSI-UAV (GNDVI)	P-value	MSI-UAV (SAVI)	P-value	MSI-UAV (NDRE)	P-value	MSI-UAV (CI green)	P-value	Repeats
A6	0.1637	0.0021	0.1364	0.5698	0.0574	0.3822	0.1047	0.3872	0.1014	0.3518	0.1185	1.5022	0.0191	5
A4	0.2665	0.0075	0.0000	0.4209	0.0010	0.3694	0.0030	0.3367	0.0183	0.3646	0.0095	0.9730	0.0000	5
A2	0.2070	-0.0721	0.0110	0.3693	0.0306	0.2553	0.1858	0.3027	0.0770	0.4029	0.0212	0.6909	0.0001	5
B6	0.3189	-0.0164	0.0228	0.5469	0.0453	0.4431	0.1183	0.4378	0.1241	0.4482	0.1120	1.1627	0.0009	5
C6	0.2621	-0.0615	0.0011	0.3201	0.0891	0.2898	0.2306	0.2065	0.0932	0.3149	0.1016	0.6959	0.0000	5
C3	0.2893	0.1140	0.0567	0.3870	0.1682	0.2908	0.4916	0.1071	0.0965	0.3690	0.1977	0.4952	0.0889	5
C1	0.3472	0.0338	0.0417	0.2430	0.1901	0.1457	0.0807	0.1525	0.0850	0.1708	0.0983	0.3541	0.4738	5

D5	0.1838	-0.0207	0.0300	0.1836	0.1094	0.0868	0.0532	0.0841	0.0523	0.1701	0.0978	0.2741	0.2611	5
D3	0.2789	-0.0103	0.0012	0.3021	0.2844	0.2346	0.1587	0.2210	0.1173	0.2795	0.4945	0.3299	0.1405	5

(c)

Sample	Spectro-1 spectrometer (GRVI)	MSI- UAV (GRVI)	P-value	MSI- UAV (NDVI)	P- value	MSI- UAV (GNDVI)	P- value	MSI- UAV (SAVI)	P- value	MSI- UAV (NDRE)	P- value	MSI- UAV (CI green)	P-value	Repeats
A6	-0.1356	-0.2793	0.0640	0.3902	0.0050	0.2032	0.0115	0.3207	0.0065	0.1613	0.0148	0.7405	0.0018	5
A4	-0.1707	-0.1890	0.3272	0.3299	0.0024	0.1336	0.0063	0.2640	0.0030	0.1502	0.0055	0.5618	0.0011	5
A2	-0.0714	-0.0796	0.4779	0.2607	0.1062	0.1311	0.1637	0.2696	0.1031	0.0886	0.1976	0.4235	0.0725	5
B6	-0.1064	-0.3540	0.0012	0.2994	0.0006	0.2635	0.0217	0.2441	0.0236	0.2205	0.0255	0.8685	0.0000	5
C6	-0.0011	-0.0915	0.1816	0.2620	0.0360	0.2635	0.0352	0.2441	0.0405	0.1695	0.0739	0.6388	0.0068	5
C3	-0.1507	-0.0806	0.1296	0.1563	0.0067	0.1775	0.0057	0.1693	0.0058	0.2002	0.0049	0.4812	0.0016	5
C1	-0.0941	-0.1833	0.1678	0.1516	0.0429	0.0887	0.0703	0.1573	0.0409	0.1092	0.0584	0.3131	0.0168	5
D5	-0.1944	-0.1575	0.0125	0.1544	0.0000	0.1537	0.0000	0.1599	0.0000	0.1304	0.0000	0.3456	0.0000	5
D3	-0.1197	-0.1002	0.0125	0.2360	0.0000	0.2175	0.0000	0.1443	0.0000	0.2213	0.0000	0.7521	0.0000	5

Note: For each VI 5 repeats were taken for each sampling point using the hand-held spectrometer and UAV to perform the t-test and obtain the p-value.

Firstly, the statistical difference between the means can be explained by the spectral profiles from (Figure 51). The red and green spectral bands are used to calculate GRVI, and the higher spectral response in the green and red bands in June and July compared to August may explain the larger difference in means from UAV in these two months. Further, even a minimal over or under estimation of reflectance values will have a greater impact on the vegetation index calculation. Secondly, GRVI index is evaluated to be useful as a phenological indicator (Motokha et al., 2010). Positive GRVI values greater than 0 indicates green vegetation and negative values indicated the change in green vegetation to autumn colouring. During June oats crop begins the flowering stage and are at the peak of green vegetation and in July the oats crop starts grain filling and by August the crops are fully ripe and ready to harvest. In our research, although there is a statistical difference between the GRVI values in June, the values appear to be positive between (0.33 to 0.53) for handheld and (0.25 to 0.45) for UAV data (Table 8a). Further, when the oats are fully ripe the GRVI values (Table 8b) appear to be negative between (-0.001 to -0.19) for handheld and (-0.07 to -0.28) for UAV data. Similarly, Motokha et al., 2010 found during the peak of green vegetation for coniferous forest canopy a positive GRVI value between (0.2 to 0.24) and a negative GRVI value between (-0.14 to -0.17) during autumn. Also, a study performed over a peanut field proved that GRVI outperformed the NDVI in detecting phenological crop changes especially during ripening (Chen et al., 2019). This could support the means of the UAV and ground reference data in our study being statistically similar for most locations in August for oats crop during ripening. (Table 8). In addition, another possible factor for the statistical difference in means between UAV and ground reference data could be the method of radiometric correction used in our study. The importance of radiometric calibration methods has been demonstrated using reference reflectance panels. This performs a correction in the entire spectrum of the signal detectable during a UAV monitoring, from underlying signals (soil and grass) to high signals (canopy) (Di Gennaro et al., 2022). In comparison to the radiometric correction performed using reference reflectance panels, the radiometric correction based on irradiance provided higher errors in UAV images, particularly for vegetation indices related to the green spectral band such as GNDVI than NDVI or NDRE. (Di Gennaro et al., 2022). Henceforth, in our study the correction of the UAV images is based on irradiance sensors instead of reference reflectance panels which could further explain the difference of GRVI values. Also, the GRVI index is calculated by using the green spectral band. Oppositely, a study evaluated the radiometric correction based on irradiance sensor

provided higher accuracy in comparison to the correction performed using reference reflectance panels (Cao et al., 2020). Therefore, despite the radiometric calibration method used the true potential of the vegetation index of UAV and ground reference data cannot be fully evaluated by only comparing the performance of GRVI for phenology of oats crop. Therefore, to fully comprehend the potential, a spectrometer with additional spectral bands of red-edge and NIR are required. Since, vegetative crops highly reflect light in these bands and could be a useful indicator for crop phenology.

Finally, GRVI values compared from ground reference data to several other vegetation indices (NDVI, GNDVI, SAVI, NDRE and CI green values) from UAV data (Table 8) provides a better understanding on the index value between the phenological growth stages of winter oats. Generally, NDVI, GNDVI SAVI and NDRE values tends to be between 0.6 to 0.8 for healthy green crops and 0.2 to 0.3 for early developing and senescence crops. Similarly, CI green values between 0 to 5 represents crops with less chlorophyll and higher CI green values greater than 5 will indicate crops with high levels of chlorophyll. According to the June data NDVI values have a higher value between 0.63 to 0.81 and CI green have a value between 1.81 to 4.42 and are statistically significant ($p < 0.05$) to GRVI ground reference data (Table 8a). In a recent study, similar NDVI values of 0.74 were obtained for oats during the peak growth stage and NDVI values of 0.30 during the final growth stage (Stepanov et al., 2020). However, most of GNDVI, SAVI and NDRE values are statistically insignificant ($p > 0.05$) to GRVI ground reference data in June (Table 8a). As June was the flowering stage of oats, the values for these vegetation indices are expected to be higher, comparable to the NDVI values from this study. Further, the July NDVI data have values between 0.18 to 0.56 and CI green have a value between 0.27 to 1.50 and are statistically significant ($p < 0.05$) to GRVI ground reference data (Table 8b). On the other hand, most of the July GNDVI, SAVI and NDRE values are statistically insignificant ($p > 0.05$) to GRVI ground reference data. Finally, August NDVI data have values between 0.15 to 0.39 and CI green have a value between 0.34 to 3.15. Along with NDVI values, GNDVI, SAVI and NDRE are also statistically significant ($p < 0.05$) to GRVI ground reference data (Table 8c).

Hence, the results from the study highlight the most suitable vegetation index as a phenology indicator of oats crop could be the GRVI, NDVI and CI green index. Firstly, this is since the index values of the GNDVI, SAVI, and NDRE do not vary significantly across phenological growth stages. Hence, it is challenging to differentiate the oats crops in different phenological growth stages over different months. On the other hand, GRVI, NDVI and CI green values

have shown to vary significantly across phenological growth stages. Particularly, positive GRVI values were obtained when oats were green and negative values were recorded when the oats were fully ripe and ready to harvest. Similarly, when the oats were green and fully ripe, the NDVI and CI green values were significantly higher and lower, indicating that these indices could be used to differentiate oats at different phenological growth stages. As a result, the NDVI and CI green indices were used to examine the relationship between soil health indicators and yield estimation in the phenological growth stages of oats.

10.3.3 Soil nutrient data

To validate the results from Palintest soil testing kits, a correlation study between standard laboratory methods for soil analysis and Palintest soil testing kits was conducted (Eddy and Johnston, 2009). The research concluded a very strong correlation between Palintest methods and external laboratory analysis for nitrate ($R^2 = 0.96$), phosphate ($R^2 = 0.95$), potassium ($R^2 = 0.96$), and pH meter ($R^2 = 0.98$) (Eddy and Johnston, 2009). As a result, the Palintest soil testing kits are a viable method for soil analysis.

Furthermore, the soil nutrient test for the oats field was performed in March 2021 by Agrovista company UK using laboratory soil test methods according to AHDB (Agriculture and Horticulture Development Board, 2023). As a result, the change in soil nutrient content can be compared using this data in conjunction with the Palintest soil data from June, July, and August 21. Agrovista's only comparable soil nutrient data from March 21 were potassium, phosphorus, and pH. As a result, these soil nutrient data were compared to Palintest data from June, July and August 21. The full results for analysis of the 48 soil sub-samples, with respect to nitrate, phosphorus, potassium, pH and soil organic matter for the months of June, July and August are reported in Table 10.

Initially, soil nutrient data analysis by Agrovista Company UK, was reported on 16 March 21, before the two rounds of soil fertiliser application (28/09/21 and 06/03/21). Hence, according to the March 21 Agrovista data (Table 9) the mean phosphorus, potassium, and pH values across the 4 zones were 27 mg/l, 173.8 mg/l and 7.4 respectively. The minimum phosphorus, potassium and pH values across the field were 17mg/l, 112mg/l and 7.3 respectively Table 9. Further, the maximum phosphorus, potassium and pH values across the field were 38mg/l, 214mg/l and 7.5 Table 9.

Further, the Palintest phosphorus levels for June 21 were recorded as minimum = 8mg/l, mean= 21.7mg/l, and maximum= 38 mg/l values Table 10. The phosphorus levels for July 21 were recorded as minimum = 9 mg/l, mean= 16.5 mg/l, and maximum= 36 mg/l values Table 10. Final, phosphorus levels for August 21 were recorded as minimum = 7 mg/l, mean = 15.4 mg/l, and maximum= 27 mg/l values Table 10. The mean and maximum phosphorus levels recorded by Palintest (mean = 21.7 mg/l and maximum 38 mg/l Table 10) in June 21 and those measured by Agrovista in March 21 (mean = 27 mg/l and maximum = 38 mg/l Table 9) do not differ significantly in terms of levels. This can be explained due to the addition of fertiliser in two rounds (30/03/21 and 13/04/2021). Furthermore, no additional phosphorus fertiliser was applied to the field following this fertiliser application. As a result, phosphorus levels in the field were noticeably low from June 21 to August 21 (Figure 52).

Meanwhile, the Palintest potassium levels for June 21 were recorded as minimum = 80 mg/l, mean = 106.9 mg/l, and maximum= 180 mg/l values Table 10. The potassium levels for July 21 were recorded as minimum = 100 mg/l, mean = 157 mg/l, and maximum= 245 mg/l values Table 10. Final, potassium levels for August 21 were recorded as minimum = 95 mg/l, mean= 161.2 mg/l, and maximum= 255 mg/l values Table 10. Despite the addition of fertiliser in two rounds (30/03/21 and 13/04/2021) the mean potassium levels across the field remained comparatively low in June 21 by Palintest (mean =106.9mg/l Table 10) in comparison to the March 21 data by Agrovista data (mean = 173.8 mg/l Table 9). This could be explained by potassium being depleted faster than phosphorus in the field. As a result, an additional potassium fertiliser application was performed on July 21, resulting in a significant spike in potassium levels in the July 21 Palintest data (Table 10) which could be observed throughout the field (Figure 53).

The nitrate levels for June 2021 were recorded as mean (minimum – maximum) as 26.5 mg/l (1.3 mg/l – 80.6 mg/l), for July 2021, 17.9 mg/l (0.9 mg/l – 48.3 mg/l) and August, 14.4 mg/l (0 mg/l – 39.4 mg/l) (Table 10). The addition of two rounds of fertiliser (30/03/2021 and 13/04/2021) contribute to the highest nitrate levels in June followed by a steady depletion, by crop uptake, until August (Figure 54).

Finally, the Palintest pH levels for June 21 were recorded as minimum = 7.04 mean = 7.62 and maximum= 7.89 values Table 10. The pH levels for July 21 were recorded as minimum = 7.21 mean = 7.61 and maximum= 7.86 values Table 10. Final, potassium levels for August 21 were recorded as minimum = 7.63 mean = 7.79 and maximum= 7.92 values Table 10. Since March

21, the pH levels in the field have been increasing until August 21, making the soil highly alkaline in nature. A high alkaline soil environment will reduce the available of certain nutrients particularly phosphorus and potassium (Neina, 2019). Particularly, the higher pH levels were observed at the north end of the field which also has the poorest oat yield in comparison to the southern end of the field (Figure 55). Inversely, Palintest analysis found lower levels of phosphorus and potassium at the north end of the field as opposed to the southern end of the field in June, July, and August 21 soil maps (Figure 52 and Figure 53). Similarly, Agrovista soil nutrient maps generated in March 21 revealed that the north end of the field had significantly lower levels of phosphorus and potassium than the southern end of the field (Figure 56). This could be since there are fewer crops at north end the field than at southern end of the field.

Table 9: Agrovista UK chemical soil health indicator data for March 2021.

March-21			
Sample	P (mg/l)	K (mg/l)	pH
Zone 1	38	214	7.5
Zone 2	31.1	199.9	7.4
Zone 3	21.9	169.3	7.3
Zone 4	17	112	7.5
Mean	27	173.8	7.4
Minimum	17	112	7.3
Maximum	38	214	7.5

Table 10: Chemical soil health indicator data determined using standard methods, for June, July, and August 2021 across the 3 phenological growth stages and oat grain quality (assessed by determination of β -glucan in oats).

Sample	June 2021 (Stage 3 – Flowering)					July 2021 (Stage 4 – Grain filling)					August 2021 (Stage 5 -Ripening)					Oat Harvest
	NO ₃ - (mg/l)	P (mg/l)	K (mg/l)	pH	SOM (%)	NO ₃ - (mg/l)	P (mg/l)	K (mg/l)	pH	SOM (%)	NO ₃ - (mg/l)	P (mg/l)	K (mg/l)	pH	SOM (%)	β - glucan (%)
A1	44.2	21	100	7.73	9.76	37.7	27	235	7.59	9.78	17.3	16	190	7.8	9.86	0.36
A2	50.45	23	105	7.74	11.22	9.7	13	205	7.57	9.31	0.9	10	150	7.79	8.18	0.21
A3	45.15	20	100	7.62	12.12	7.1	11	150	7.48	10.76	5.3	15	120	7.79	9.72	0.14
A4	9.73	20	110	7.84	12.69	7.5	18	150	7.6	10.45	23	11	200	7.92	10.14	0.74
A5	18.14	22	100	7.76	11.03	25.7	20	205	7.34	9.72	7.1	11	245	7.9	11.13	0.42
A6	23.02	13	80	7.78	11.06	31.9	9	105	7.51	11.19	15.9	11	130	7.79	10.08	1.08
A7	29.21	19	135	7.8	11.53	9.7	12	165	7.66	11.55	39.4	11	145	7.83	11.86	0.56
A8	30.98	20	145	7.89	11.04	28.4	22	170	7.69	9.96	25.7	20	190	7.84	9.75	0.70
B1	13.2	18	80	7.6	8.34	7.1	17	165	7.56	10.21	3.1	19	95	7.64	8.54	0.50

B2	27	20	100	7.63	10.55	14.2	17	200	7.56	10.01	8.9	14	125	7.75	8.17	1.18
B3	69.6	22	105	7.38	10.42	37.2	12	245	7.43	10.70	1.8	13	120	7.63	9.32	3.39
B4	23.4	20	100	7.65	11.87	33.2	11	175	7.69	10.83	4.4	16	185	7.85	11.82	1.34
B5	17.6	23	105	7.59	12.76	20.8	20	120	7.7	13.10	26.1	13	180	7.78	15.78	2.33
B6	8.9	13	90	7.58	9.60	3.5	25	100	7.84	11.87	2.7	16	100	7.84	9.57	0.51
B7	8.9	14	80	7.56	10.52	6.6	10	150	7.86	10.30	7.1	10	135	7.8	9.30	0.09
B8	4.4	11	115	7.18	6.61	4.4	10	175	7.8	10.39	4.4	9	155	7.64	9.57	0.15
C1	80.6	22	165	7.52	9.22	16.8	13	205	7.31	9.09	3.5	7	160	7.65	7.77	0.75
C2	58.9	25	150	7.64	9.89	12	16	215	7.43	9.00	2.7	12	210	7.73	5.68	1.41
C3	34.1	20	100	7.04	9.50	7.5	12	125	7.21	10.00	33.2	16	130	7.77	16.07	3.94
C4	18.6	20	90	7.2	10.75	34.1	10	215	7.28	10.78	7.5	18	135	7.84	12.03	0.96
C5	36.3	20	100	7.68	12.61	26.6	17	130	7.36	12.56	20.8	11	140	7.78	11.57	1.71
C6	13.3	23	90	7.52	10.78	9.3	9	150	7.57	9.77	17.7	16	200	7.63	9.52	1.68
C7	12	8	100	7.65	9.19	19	11	120	7.78	10.25	8.9	12	200	7.77	8.41	0.32
C8	33.7	9	110	7.59	9.92	16.4	18	150	7.82	10.43	7.1	12	95	7.82	10.96	2.19

D1	47.8	31	100	7.38	9.23	40.3	16	100	7.51	8.24	4.1	19	145	7.79	7.29	0.22
D2	23	33	132	7.61	10.13	8.4	20	105	7.63	11.68	7.9	19	185	7.77	8.97	0.95
D3	19.5	33	95	7.46	10.48	18.6	26	170	7.48	9.47	13	20	190	7.76	10.51	1.26
D4	20.4	27	90	7.7	11.50	14.6	11	110	7.49	9.78	38.5	15	200	7.82	10.34	0.79
D5	22.2	22	105	7.57	10.33	6.2	12	180	7.42	13.78	13.7	17	130	7.83	14.99	0.58
D6	46.1	18	95	7.66	12.76	25.7	16	155	7.57	13.27	22.6	16	195	7.85	10.01	2.11
D7	16.8	13	105	7.75	9.58	48.3	20	125	7.68	4.85	20.8	13	130	7.88	9.35	2.93
D8	36.2	17	105	7.78	9.98	23.9	20	130	7.78	10.98	13.7	13	165	7.87	10.32	1.86
E1	14.2	28	95	7.76	10.10	22.6	19	215	7.81	9.37	21.3	20	225	7.69	8.19	0.62
E2	22.2	20	90	7.77	9.52	6.6	20	135	7.73	9.63	4.9	14	125	7.72	7.88	0.37
E3	54.4	24	115	7.64	9.23	2.7	13	110	7.83	9.01	18.2	14	155	7.72	9.33	0.71
E4	1.3	22	100	7.67	14.79	12.8	13	120	7.78	10.60	8.8	9	110	7.72	9.62	0.74
E5	40.8	38	120	7.48	9.99	1.8	13	105	7.45	10.47	35.4	13	120	7.72	10.12	0.44
E6	19	19	90	7.51	11.02	8.4	16	100	7.72	10.84	15.9	16	160	7.86	9.48	0.47
E7	1.8	26	95	7.87	11.55	12	15	165	7.83	10.63	7.1	12	140	7.89	9.09	0.98

E8	15.5	21	140	7.84	10.80	23.9	16	140	7.83	10.31	35.4	12	165	7.87	9.97	0.80
F1	48.3	25	110	7.11	11.31	40.3	36	240	7.48	10.47	19	23	230	7.73	7.58	1.00
F2	29.2	25	110	7.52	9.98	15.1	24	130	7.57	9.94	23	21	255	7.64	9.29	1.22
F3	5.8	29	180	7.42	8.95	38.5	21	215	7.58	8.80	10.6	20	125	7.82	9.03	0.43
F4	6.6	35	120	7.79	10.58	36.3	23	235	7.69	7.64	4.4	27	225	7.83	10.14	0.71
F5	8.9	29	105	7.75	11.30	12.8	20	135	7.68	17.61	15.4	25	215	7.8	9.82	0.92
F6	11.5	21	85	7.78	13.02	4.9	13	165	7.69	11.39	26.5	26	185	7.86	12.90	1.34
F7	47.8	17	80	7.89	13.15	9.3	13	108	7.76	10.96	15.4	14	100	7.88	11.09	0.84
F8	2.7	20	115	7.76	10.10	0.9	18	120	7.86	12.17	0	20	125	7.86	9.99	1.06
Mean	26.5	21.7	106.9	7.62	10.7	17.9	16.5	157	7.61	10.5	14.4	15.4	161.2	7.79	10.0	1.04
Minimum	1.3	8	80	7.04	6.6	0.9	9	100	7.21	4.9	0	7	95	7.63	5.7	0.09
Maximum	80.6	38	180	7.89	14.8	48.3	36	245	7.86	17.6	39.4	27	255	7.92	16.1	3.94

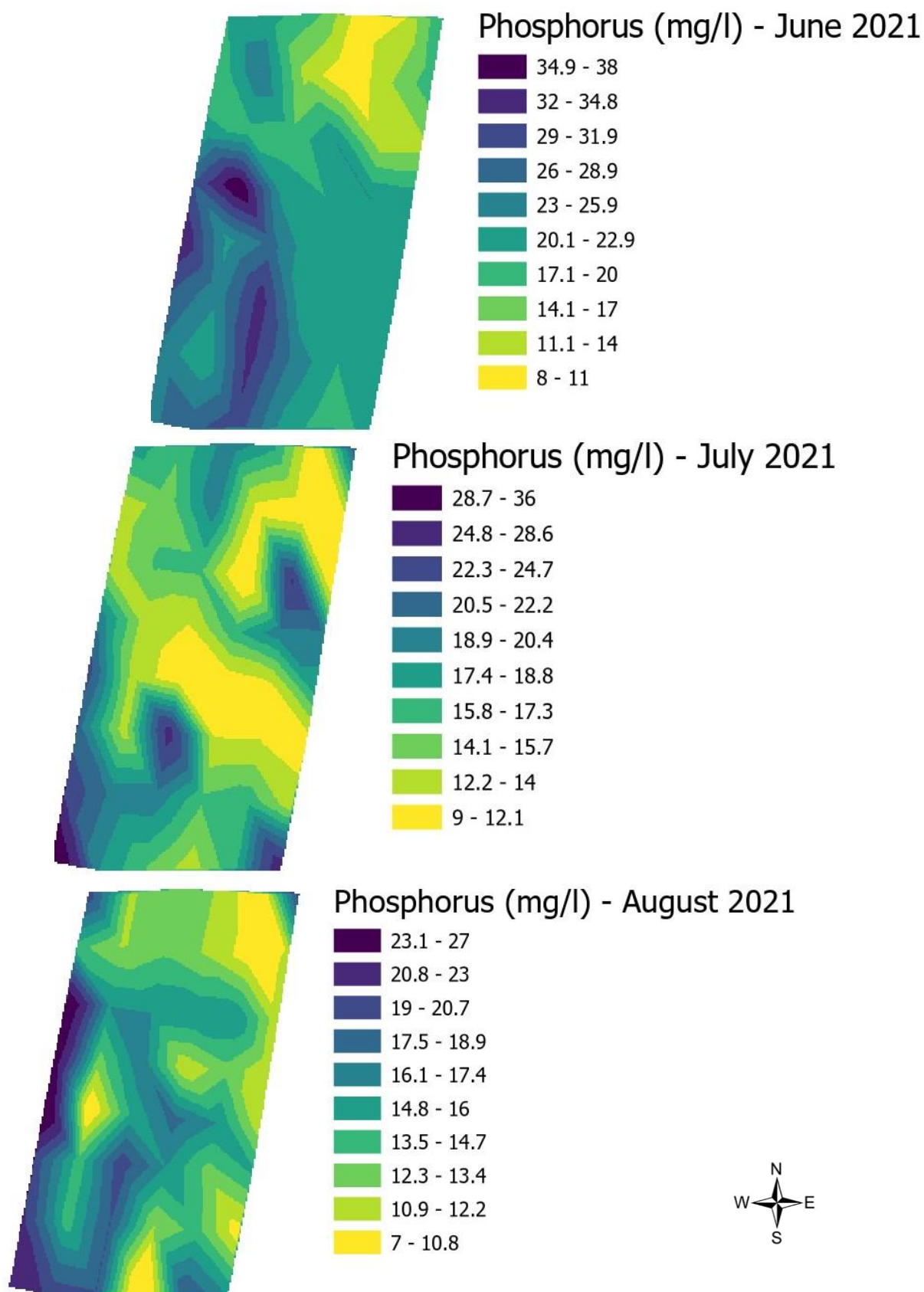


Figure 52: Palintest of phosphorus maps across phenological growth stages

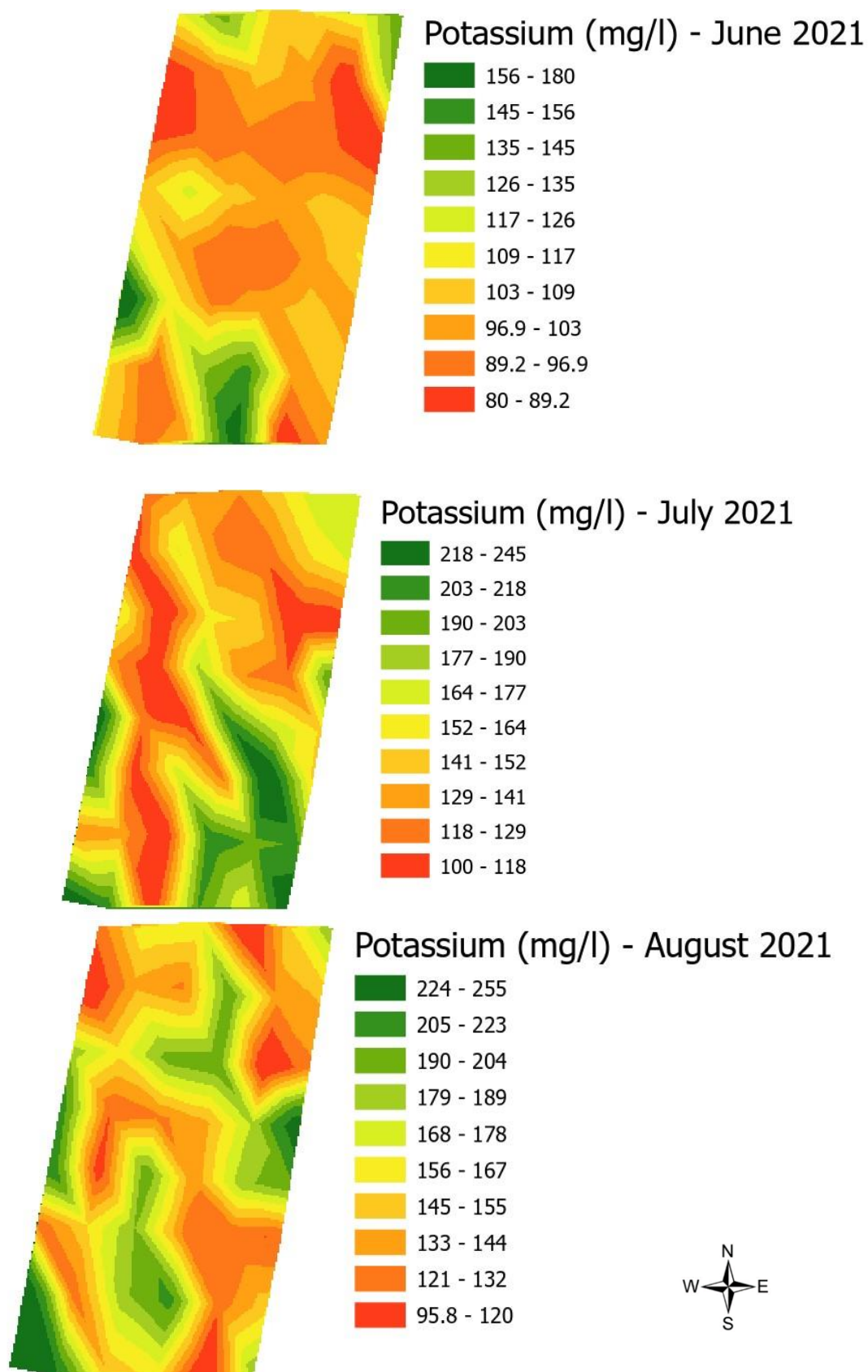


Figure 53: Palintest of potassium maps across phenological growth stages.

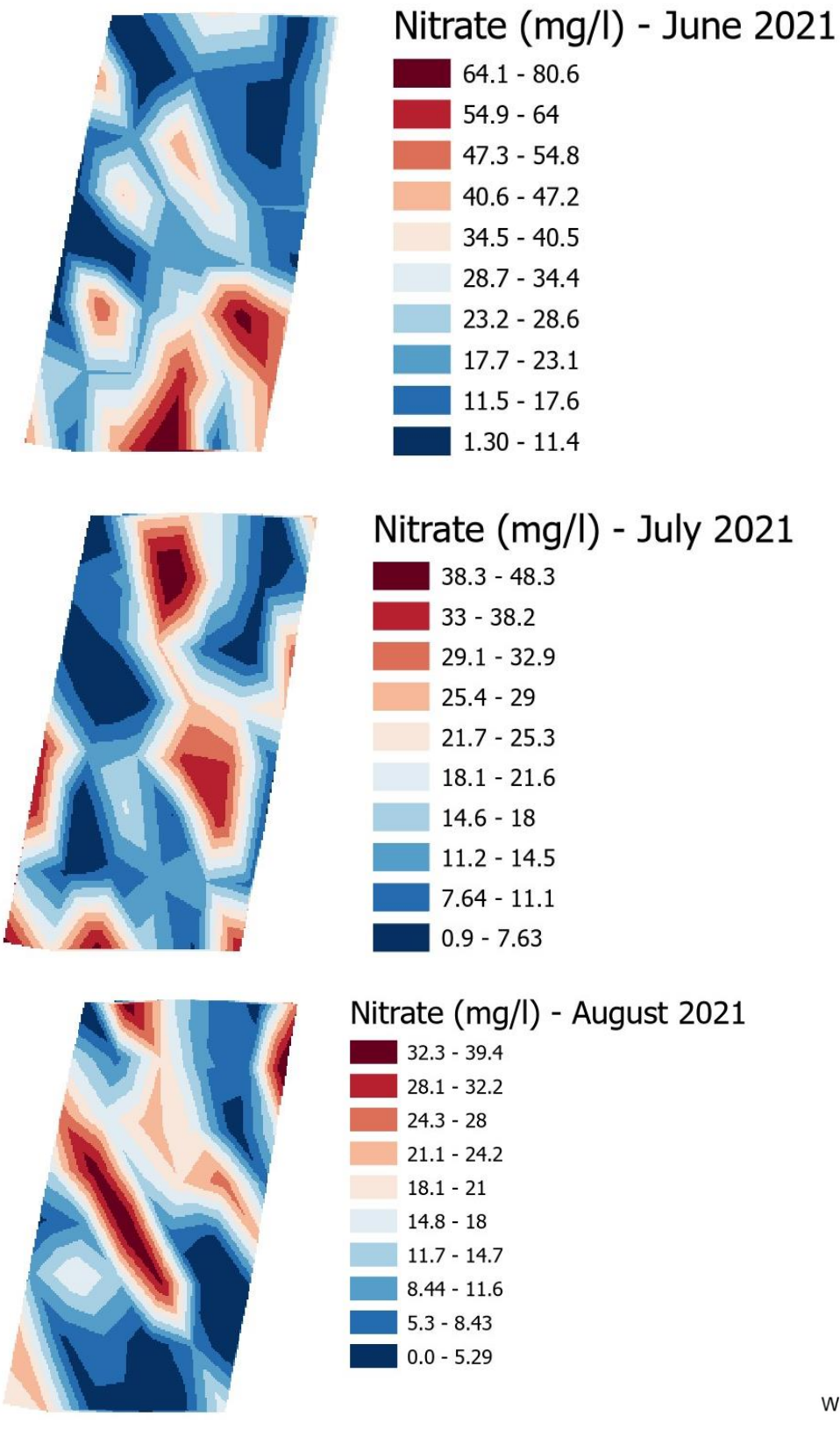
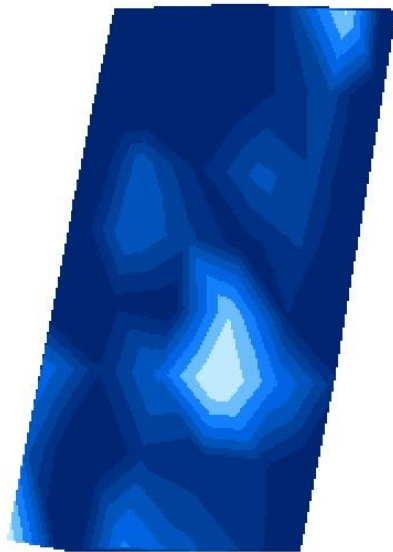
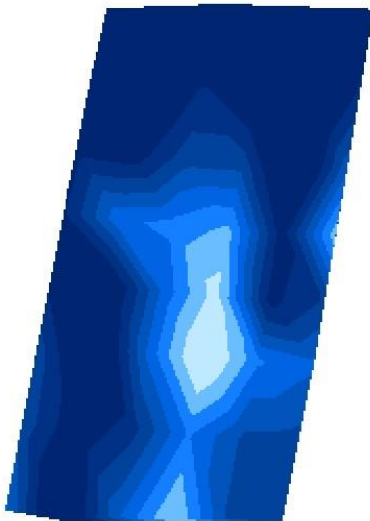


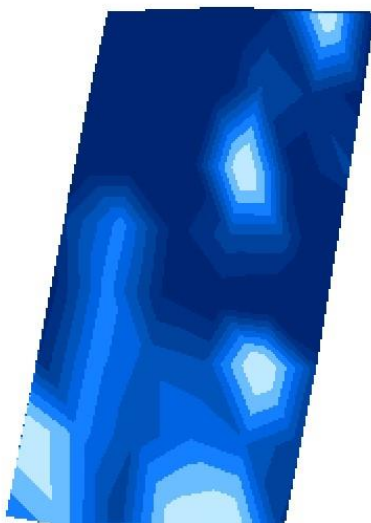
Figure 54: Palintest of soil nitrate maps across phenological growth stages.



pH - June 2021



pH - July 2021



pH - August 2021

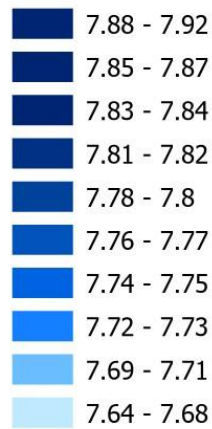
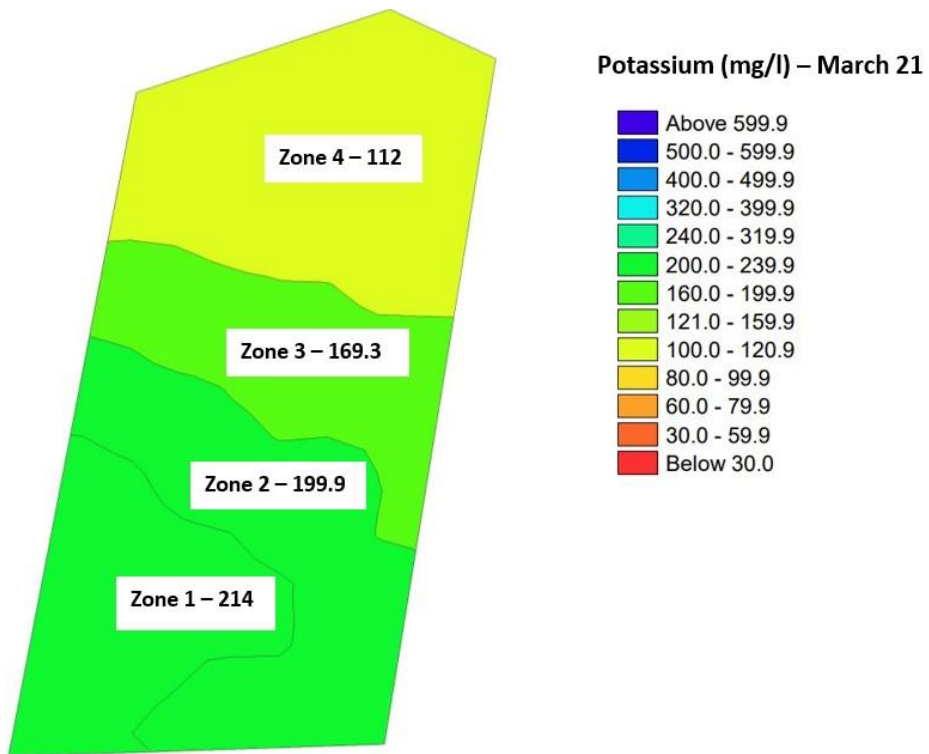
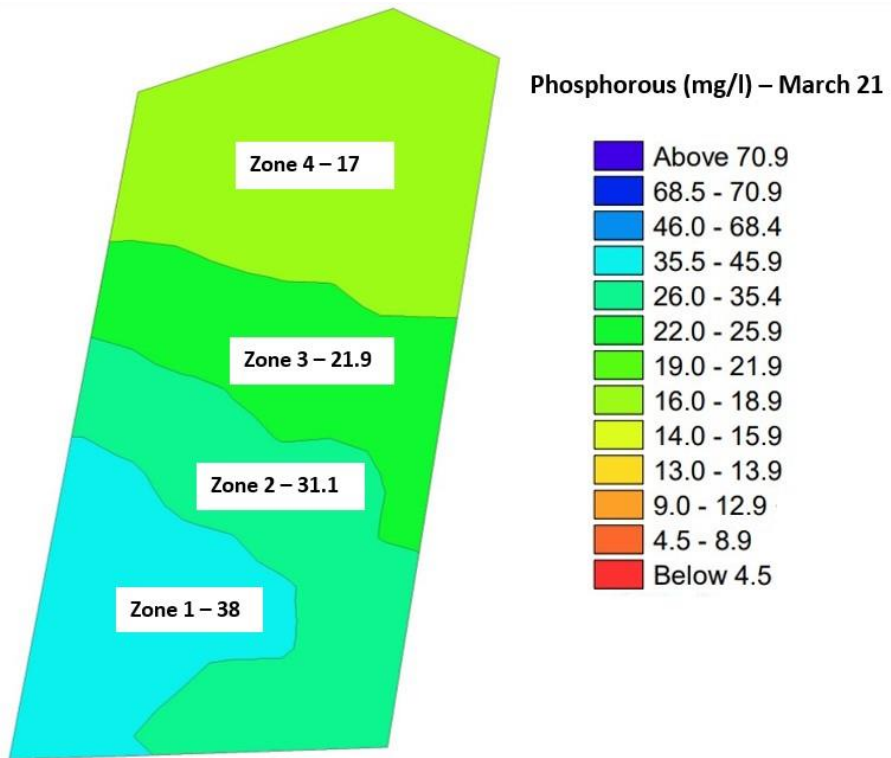


Figure 55: Palintest of pH maps across phenological growth stages.



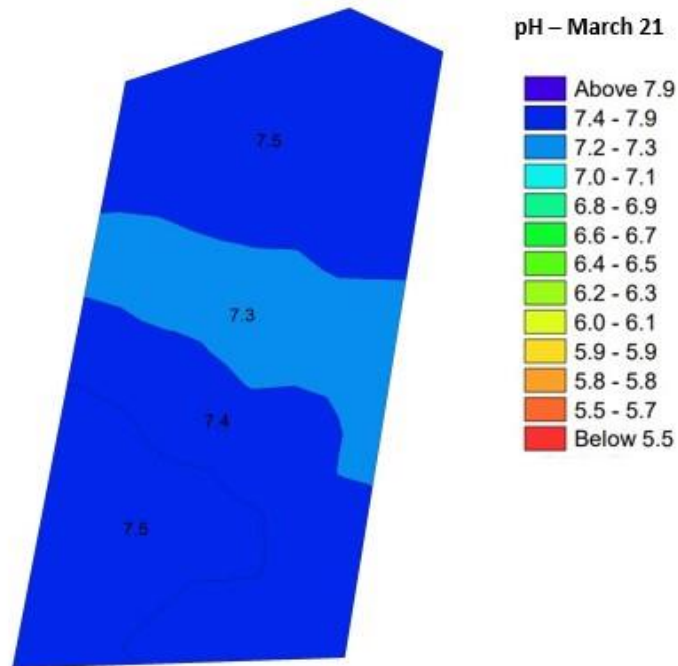


Figure 56: Chemical soil health indicator maps for March 2021 by Arovista Company UK.

10.3.4 Relationship between vegetation indices (NDVI and CI green) and chemical soil health indicators

In this research study, NPK fertilization plan for the oats field were conducted in four rounds. The first round of NPK fertiliser treatment occurred after crop sowing in September 2020, followed by two rounds in March 2021 and a single round in April 2021 during crop emergence. However, only K fertiliser were supplied to the crops during the grain filling stage 4 in July 2021. As a result, evaluating the relationship between vegetation indices across the field and soil health indicators will provide an idea of whether the crops have taken up the supplied fertilisers and whether the crops require additional fertiliser treatment. Generally, the vegetation index values are anticipated to vary at different phenological growth stages, assisting in identifying the health and growth stage of crops. For instance, during the vegetative flowering growth stage 3 (June), the oats are at its most vigorous stage, resulting in higher green biomass and a higher VI values for NDVI (minimum = 0.46, mean = 0.64 and maximum = 0.81) and CI green (minimum = 1.01, mean = 2.63 and maximum = 4.62) table 11. During

the grain filling stage 4 (July), the oats begin to change colour from green to a golden yellowish colour, resulting in a lower value of NDVI (minimum = 0.18, mean = 0.42 and maximum = 0.64) and CI green (minimum = 0.27, mean = 1.64 and maximum = 2.91) table 11. Finally, during the oats ripening stage 5, all the oats should change to golden yellow, implying that the crops are fully ripe and ready to harvest, resulting in lower NDVI values (minimum = 0.11, mean = 0.34 and maximum = 0.57) and CI green (minimum = 0.31, mean = 1.55 and maximum = 3.15) table 11.

In theory, retaining fertiliser in the soil indicates a slower uptake of fertilisers by crops, delaying the growth rate during the respective phenological growth stage, and thus varying the VI values. For instance, if the soil NPK levels are higher in some places in the field, NDVI and CI green values will be higher in the senescence stage indicating that the crops are ripening at a delayed rate. Therefore, evaluating the relationship between vegetation indices and soil health indicators will determine whether the VI can be used as a substitute to determine chemical soil health indicators. Hence, allowing farmers to determine whether these nutrients can be supplied to the desired area to maximise the growth rate of crops. Therefore, the relationship between NDVI and CI green values and soil health indicator in oats were determined across 48 soil samples collected across different phenological growth stages (Figure 49).

The results of the Pearson correlation analysis presented in Table 12 proves weak, moderate, strong, and very strong relationships between the NDVI, CI green and soil health indicators (NPK, organic matter and pH).

Table 11: NDVI and CI green basic statistical parameters across phenological growth stages at 48 location points

Phenological growth stage	NDVI				CI green			
	Mean ± SD	Min	Max	N	Mean ± SD	Min	Max	N
Stage 3 – Flowering (June)	0.64 ± 0.083	0.46	0.81	48	2.63 ± 0.82	1.01	4.62	48
Stage 4 – Grain filling (July)	0.41 ± 0.13	0.18	0.64	48	1.64 ± 0.71	0.27	2.91	48
Stage 5 – Ripening (August)	0.34 ± 0.11	0.11	0.57	48	1.55 ± 0.73	0.31	3.15	48
June, July and August	0.46 ± 0.17	0.11	0.81	144	1.94 ± 0.88	0.27	4.62	144

Table 12: Pearson correlation coefficient (r) between chemical soil health indicators and NDVI and CI green across the 48 locations

Phenological growth stage	NDVI					CI green				
	Nitrate (NO ₃ ⁻)	P	K	pH	% Organic matter	Nitrate (NO ₃ ⁻)	P	K	pH	% Organic matter
Stage 3 – Flowering (June)	0.81	0.14	0.37	-0.18	0.020	0.93	0.04	0.27	-0.12	-0.033
Stage 4 – Grain filling (July)	0.70	0.34	0.17	0.007	-0.19	0.74	0.42	0.22	-0.003	-0.22
Stage 5 – Ripening (August)	0.73	0.13	0.21	0.26	0.23	0.77	0.15	0.23	0.12	0.37
June, July and August	0.68	0.47	-0.34	-0.24	0.11	0.84	0.38	-0.17	-0.17	0.11

Initially, the correlation between NDVI and CI green, evaluated in terms of phenological growth stages, and the pH and soil organic matter were very weaker, and inconclusive (Table 12). Meanwhile, correlation between NDVI and CI green with phosphorus and potassium were weak and moderate across phenological growth stages. At flowering stage very, weaker correlation was observed between NDVI and phosphorus ($r = 0.14$) and CI green and phosphorus ($r = 0.04$) (Table 12). At grain filling stage weak correlation was observed between NDVI and phosphorus ($r = 0.34$) and a moderate correlation for CI green and phosphorus ($r = 0.42$) (Table 12). During ripening stage, a very weak weaker correlation was observed between NDVI and phosphorus ($r = 0.13$) and CI green and phosphorus ($r = 0.15$) (Table 12). However, combining all the data from the phenological growth stages, the correlation was identified to be moderate between NDVI and phosphorus ($r = 0.47$) and a weak correlation for CI green and phosphorus ($r = 0.38$) (Table 12). Whilst, at flowering stage a weak correlation was observed between NDVI and potassium ($r = 0.37$) and CI green and potassium ($r = 0.27$) (Table 12). At grain filling stage very, weak correlation was observed between NDVI and potassium ($r = 0.17$) and CI green and potassium ($r = 0.22$) (Table 12). Similarly, during ripening stage weak correlation was observed between NDVI and potassium ($r = 0.21$) and CI green and potassium

($r = 0.23$) (Table 12). Also, combining all the data from the phenological growth stages, the correlation was identified to be negatively weak between NDVI and phosphorus ($r = -0.34$) and CI green and phosphorus ($r = 0.17$) (Table 12). Interestingly, correlation between NDVI and CI green with soil nitrate were strong and very strong across phenological growth stages. At flowering stage, a very strong positive correlation was observed between NDVI and nitrate ($r = 0.81$) and CI green and nitrate ($r = 0.93$) (Table 12). At grain filling stage a strong positive correlation was observed between NDVI and nitrate ($r = 0.70$) and CI green and nitrate ($r = 0.74$) (Table 12). Similarly, during ripening stage strong positive correlation was observed between NDVI and nitrate ($r = 0.73$) and CI green and nitrate ($r = 0.77$) (Table 12). Also, combining all the data from the phenological growth stages, the correlation was identified to be moderately positive between NDVI and nitrates ($r = 0.68$) and very strongly positive for CI green and nitrates ($r = 0.84$) (Table 12).

Hence, according to the findings of this study, the relationships between the values of vegetation indices and the chemical soil health indicators of pH, soil organic matter, potassium, and phosphorus were typically weak. Only strong correlations were observed with soil nitrate and vegetation indices. This could be due to the mobility of soil nitrogen through the soil matrix in comparison to soil phosphorus and potassium (Bascietto et al., 2022). However, a study conducted on *Carex cinerascens* a wetland grass, there was a strong correlation between various three-band vegetation indices (TBVI) and foliar phosphorus content (Wang et al., 2016). Similarly, strong relationships were found between potassium content in rice leaves and various vegetation indices (Lu et al., 2022). Also, soil phosphorus content in subtropical wetland were strongly correlated with satellite derived NDVI values (Rivero et al., 2009). Meanwhile, both phosphorus and potassium contents in legume-based pastures were strongly related to the photochemical reflectance index (PRI) (Kawamura et al., 2011). Additionally, in a wheat study, there were significant and insignificant correlations between selected vegetation indices and potassium and phosphorus content in various locations (Pimstein et al., 2011).

Hence, these studies indicate the relationships between soil health indicators and vegetation indices differ depending on soil and crop conditions. Also, most strongly correlated studies between VI and phosphorus and potassium were analysed from leaf contents rather than soil samples. Therefore, it might be worth identifying the relationship between VI and phosphorus and potassium content in oats leaves. Further, the soil studies that analysed phosphorus and potassium contents which showed positive correlation to VI were performed on wet soil texture type. In our study the soil type for oats field is sandy silt loam, which could potentially cause

for a weaker correlation between VI and soil phosphorus and potassium. Also, the soil pH in the studied oats field is highly alkaline (pH = 7.5), which is likely to have a negative impact on the essential nutrients such as phosphate and potassium. Further, combining the VI data from all the phenological stages improved the correlation between VI to phosphorus rather than analysing the data separately at each phenological growth stage. However, this was not observed for potassium and resulted in a weak negative correlation between VI and potassium across the data from all phenological growth stage. This effect could be explained by addition of potassium fertiliser to the field during the grain filling stage (July 21) which caused a spike in potassium contents in the soil samples, further affecting the relationship between VI and potassium.

Therefore, phosphorus, potassium, organic matter, and pH in oats soil do not show a strong relationship with VI. While soil nitrate does show a strong relationship with VI

10.3.5 Soil nitrate model development and validation

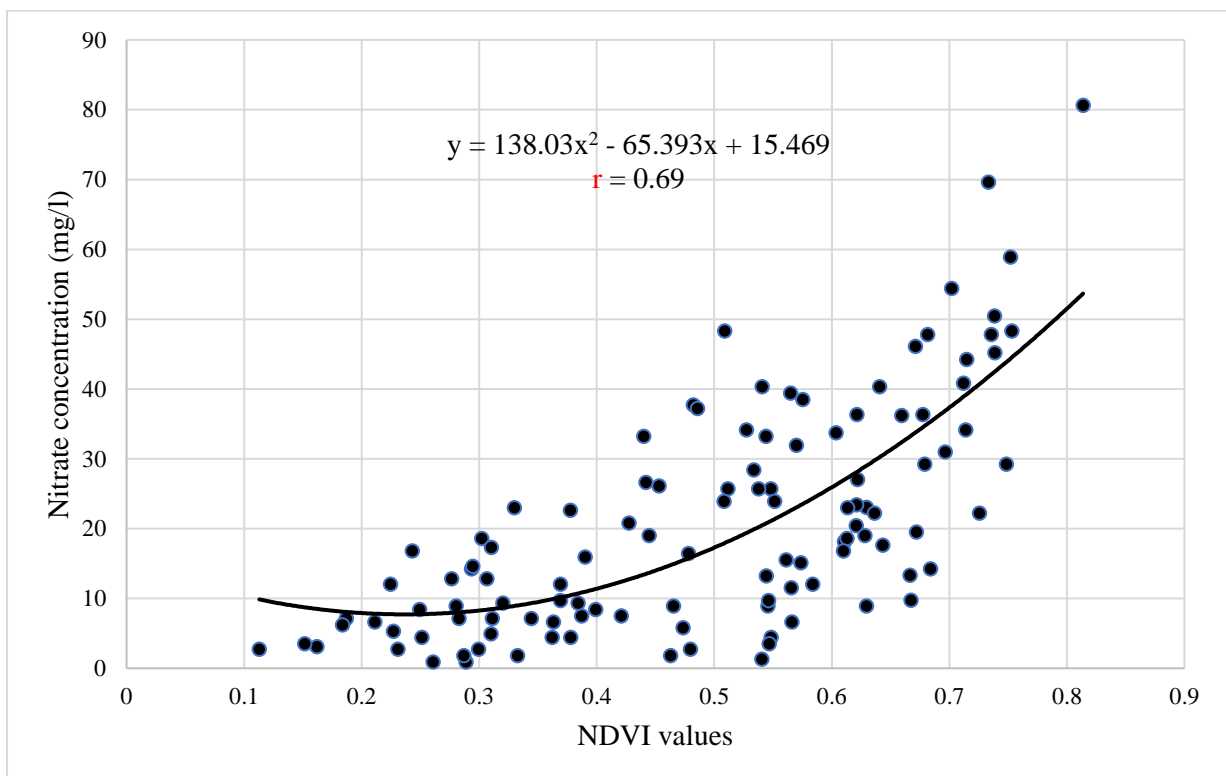
Since, a strong positive correlation was observed between VI and soil nitrate, regression models were developed to predict soil nitrate. To do this data must be split into training and test data set. Around 80% of the data were used in training data set to build soil nitrate predictive regression model using NDVI and CI green values. The remaining 20% of the dataset was known as the test data set which was used to predict soil nitrate from the developed regression models using the NDVI and CI green values. If the soil nitrate predictive regression models are efficient it can accurately predict soil nitrate on the test data sets. Further, a strong relationship between the measured soil nitrate with the predicted soil nitrate must be observed.

As a result, two models were developed and used to predict nitrate concentrations based on NDVI and CI green. The best relationship fit between VI and soil nitrate concentration were found to be polynomial regression models (Figure 57). Pearson correlation coefficient between NDVI and soil nitrate using the developed model were $r = 0.69$ across all phenological growth stages (Figure 57a). Thus, Pearson correlation coefficient between CI green and soil nitrate using the developed model were $r = 0.84$ across all phenological growth stages (Figure 57b). To validate the developed polynomial regression models (Figure 57a and 57b) cross validation was performed to predict soil nitrate using the test data set. Hence, Figure 58a and 58b displays the results of comparison between measured soil nitrate concentration and predicted soil nitrate concentration retrieved from the developed polynomial regression models. Pearson correlation between measured nitrate concentration and predicted nitrate concentration using NDVI developed model across all phenological growth stages was $r = 0.68$ (Figure 58a). Further, the Pearson correlation between measured nitrate concentration and predicted nitrate concentration using CI green developed model across all phenological growth stages was $r = 0.78$ (Figure 58b). The results highlight CI green regression models developed has a stronger correlation at predicting soil nitrate in comparison to NDVI regression models.

Furthermore, the NDVI and CI green regression models (Figure 57a and 57b) were tested to predict soil nitrate concentration in early May, when oats were in the early stages of vegetative growth. Figure 59 displays the results of comparing the soil nitrate concentration measured in June (flowering stage) vs the predicted soil nitrate in May for NDVI and CI green. According to Figure 59a the Pearson correlation for NDVI between measured soil nitrate in June and predicted soil nitrate in May displayed very weak and insignificant correlation $r = -0.02$. However, the Pearson correlation for CI green between measured soil nitrate in June and predicted soil nitrate in May displayed moderate correlation $r = 0.28$ (Figure 59b). Also, in

comparison to NDVI in CI green most of the predicted soil Nitrate data for May lies between the 95% confidence interval range. However, it may be worthwhile to run this data set through modern machine learning models (e.g., random forest) to see if soil nitrate can be predicted much more accurately than simple regression models. Also, a limitation of chlorophyll sensitive vegetation index such as NDVI and CI green is that even if a suitable soil nutrient model is developed using these indices it can only be used to predict soil nitrate at peak growth stage of crops. This is due to a positive linear relationship observed between VI and soil nutrient contents. As a result, when crops are immature and at an early growth stage, they have a lower VI value and thus a lower prediction of soil nitrate. This prediction will be inaccurate at the early growth stage of crops since realistically, there is supposed to be a higher soil nutrient concentration at the early growth stage. As a result, the temporal variation of VI models will exhibit a wide range of inaccuracy, especially during the initial stages of crop growth.

(a)



(b)

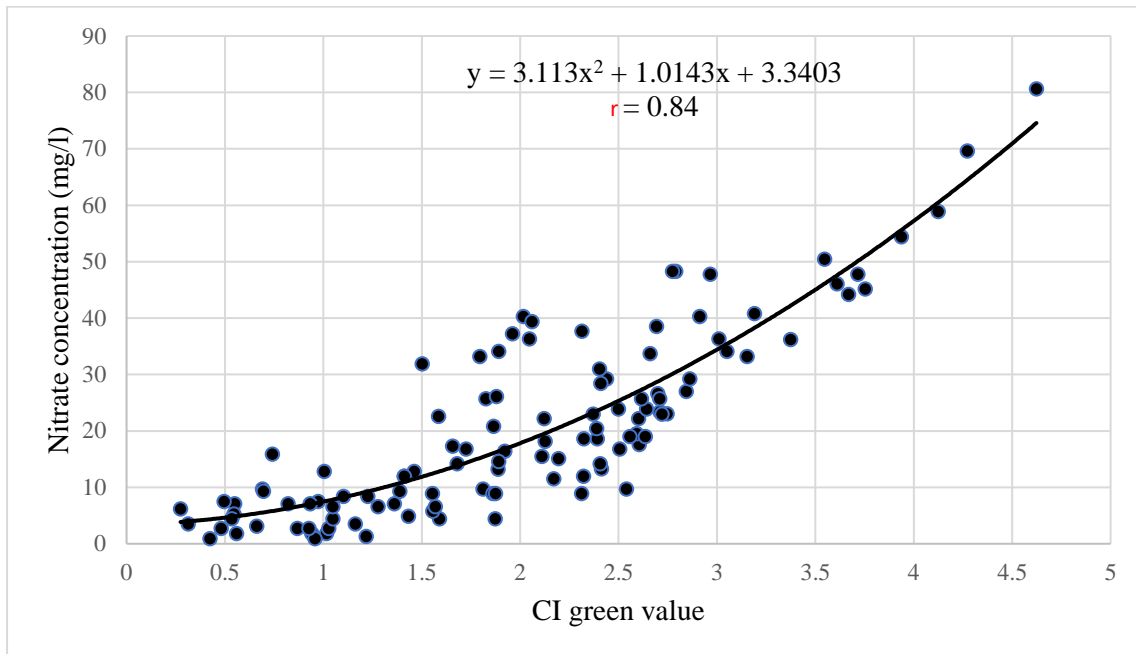
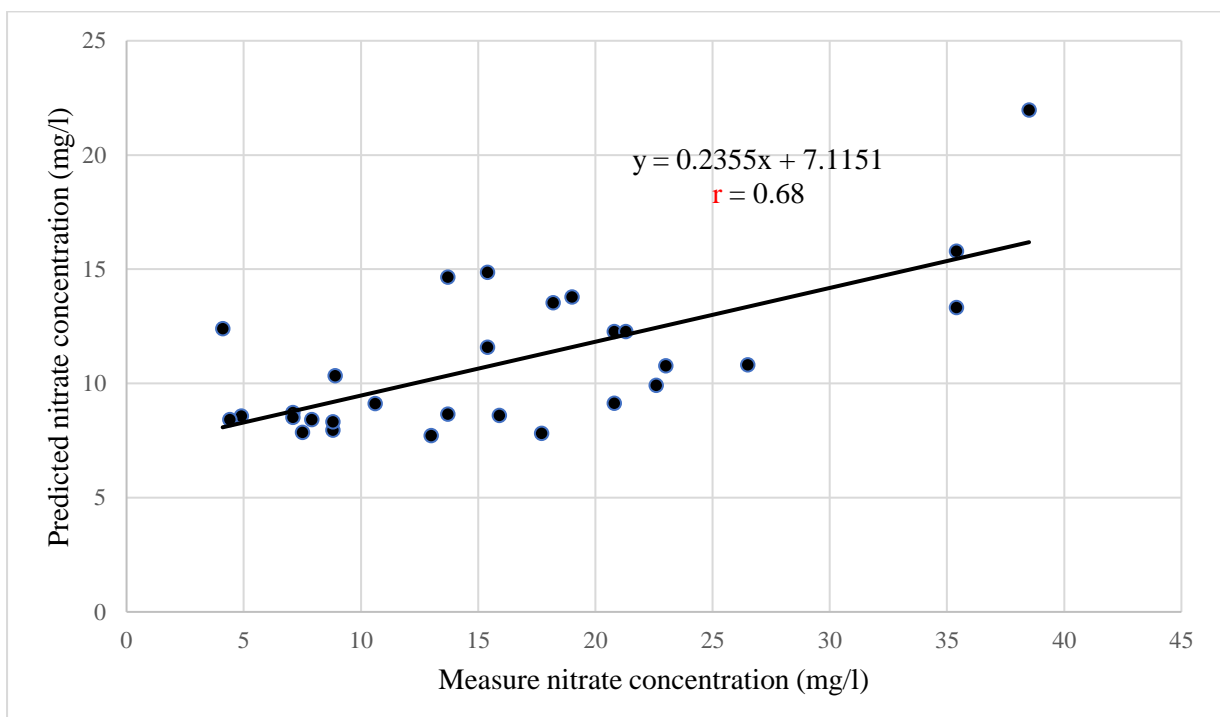


Figure 57: Relationship between measured soil nitrate concentration from training dataset vs (a) NDVI (b) CI green across all phenological growth stages.

(a)



(b)

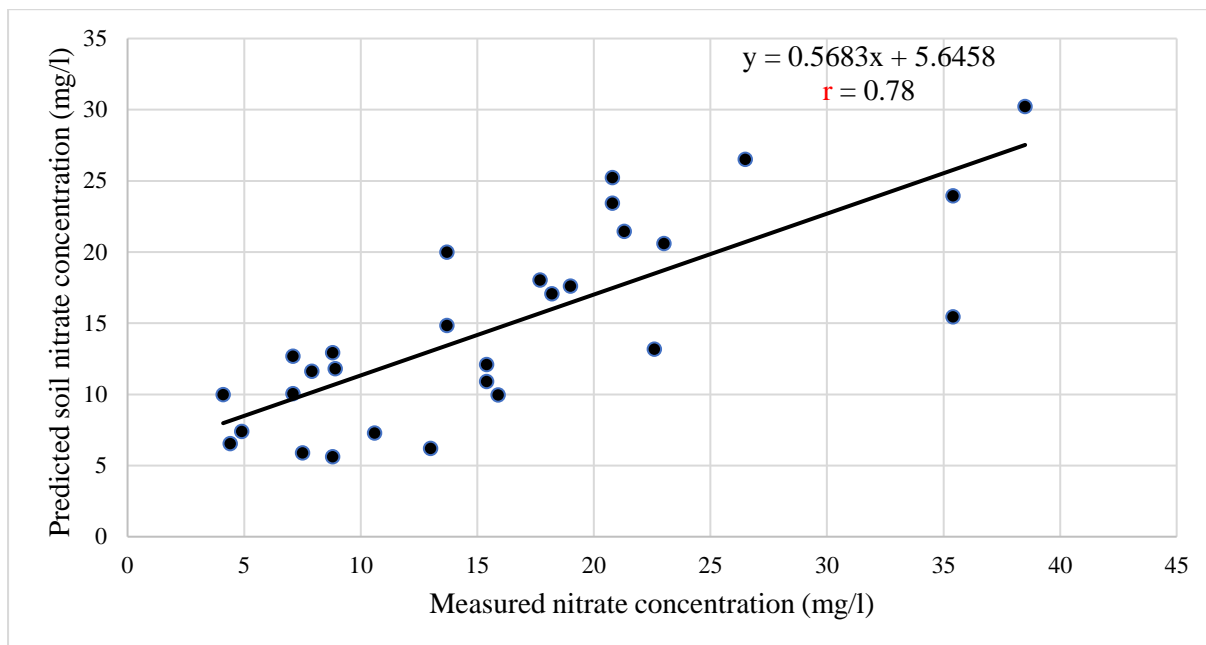
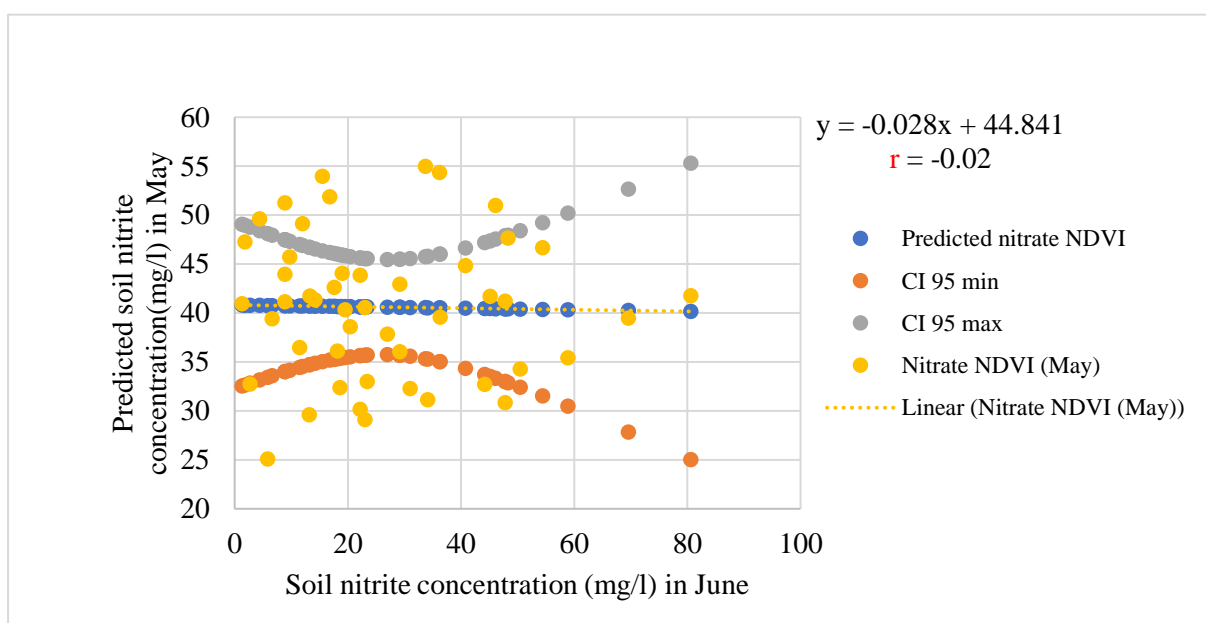


Figure 58: Cross validation using test data set to predict soil nitrate from developed predictive models on (a) NDVI (b) CI green across all phenological growth stages.

(a)



(b)

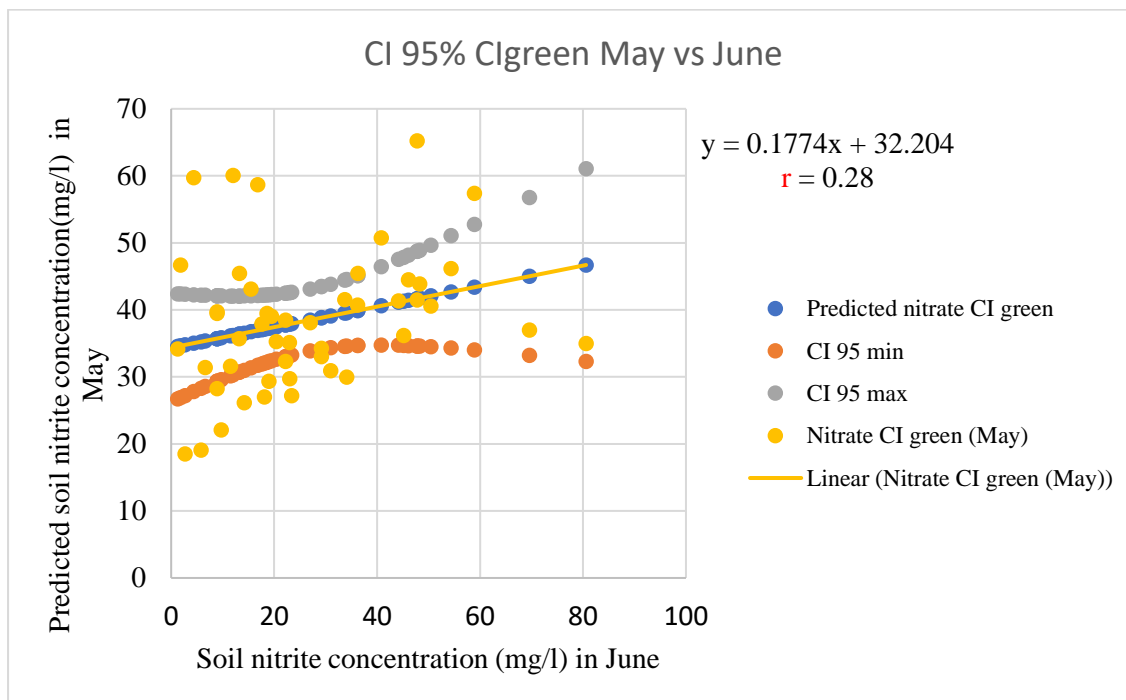


Figure 59: Comparison of soil nitrate predicted in May vs June (Flowering stage) by (a) NDVI (b) CI green.

10.3.6 Estimation of oats plant height by multispectral-UAV at phenological growth stages.

On 4.2 ha of winter oats field, multispectral-UAV estimates of plant height were performed using the CHM (Figure 60). To validate the UAV data, there is no ground-truth data of oat crop height over the five stages in this study. As a result, the height of the wall adjacent to the oats field was measured to validate the UAV images. The manually measured wall heights were statistically analysed (t-test) against the UAV-MSI estimated wall height (Table 13). There was no statistical difference between the two measurements (at a p-value of 0.17, 95% confidence interval). Hence, the estimated oats plant height measurements from UAV-MSI images are accurate for the study.

Table 13: Statistical t-test on estimated and actual wall height

Location of wall height	Actual Wall Height (m)	Calculated using UAV Data Wall Height (m)
1	0.96	0.96
2	1.17	1.04
3	1.09	1
Average	1.07	1
SD	0.04	0.11
P value (0.05) = 0.17		

Note: For each location the UAV data is calculated on three different days and represented as a mean value.

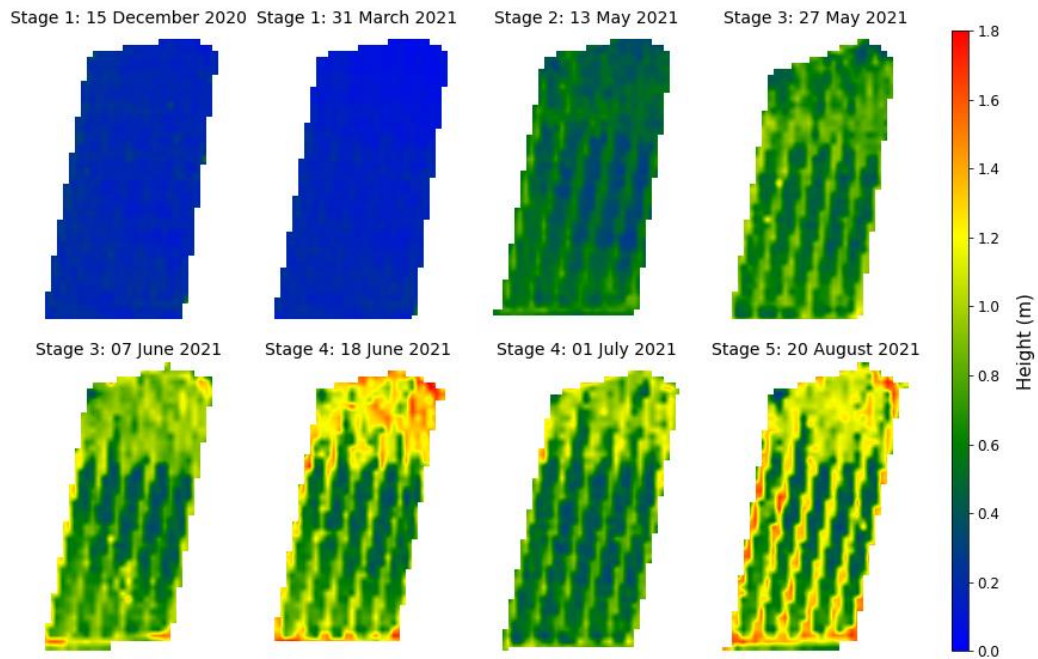


Figure 60: 3D CHM for oat at different phenological growth stages

Figure 61 depicts a statistical boxplot of the distribution of numerical data with minimum, median, mean, and maximum oats plant height across the field. The oats height in stage 1 (Tillering) has values of 0.077 m, 0.14 m, and 0.37 m. (minimum, mean and maximum). The oats height in stage 2 (Panicle formation) has values of 0.18 m, 0.44 m and 0.63 m (minimum, mean and maximum). The oats height in stage 3 (Flowering), has values of 0.25 m, 0.79 m and 1.43 m (minimum, mean and maximum). The oats height in stage 4 (Grain filling), has values of 0.32 m, 0.95 m and 1.52 m (minimum, mean and maximum). The oats height in stage 5 (Ripening), has values of 0.43 m, 1.18 m and 1.75 m (minimum, mean and maximum).

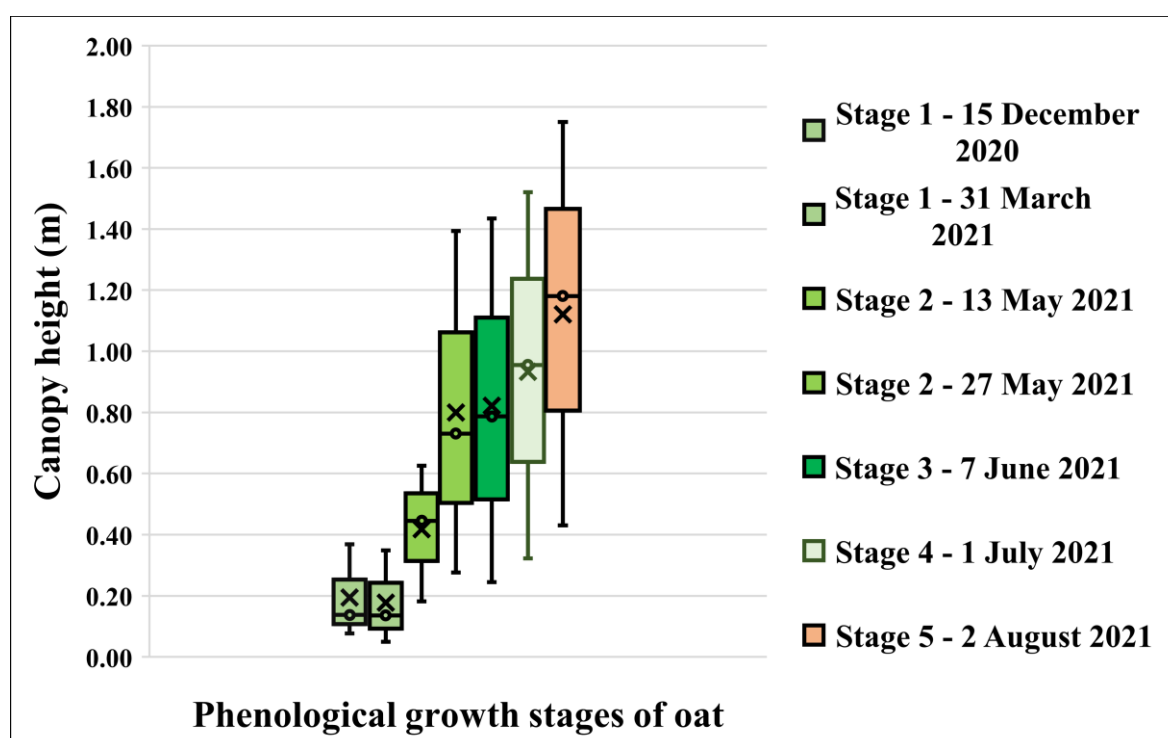


Figure 61: Oat canopy height at phenological growth stages.

Note: The limits of the box represent the upper and lower quartile of the data as assessed at the 95% confidence limit while the whiskers show the minimum and maximum heights determined. The horizontal line within the box represents the median height while the cross is the mean height.

According to winter oat growth guide, the height of the oats crop grown in the North of UK at stage 5 is recorded to be 1.00 m (Oat growth guide). This mean height complements the data from this research study, as the mean oat height at stage 5 is recorded to be 1.18 m. However, it should be noted that oat height can vary by environmental and management factors.

10.3.7 Relationship between soil NPK and oat height in different zones

The effect of fertilisers (NPK) on oat height across phenological growth stages in different zones was studied. In theory, oat crop height and fertilisers should have an inverse relationship; as the crops use up the fertilisers, the crop height will increase. The zones were divided into four zones similar to Agrovista data (Figure 56), as shown in Figure 62. The soil parameters for the zones are recorded in ((Table 4.1) in Appendix 4).

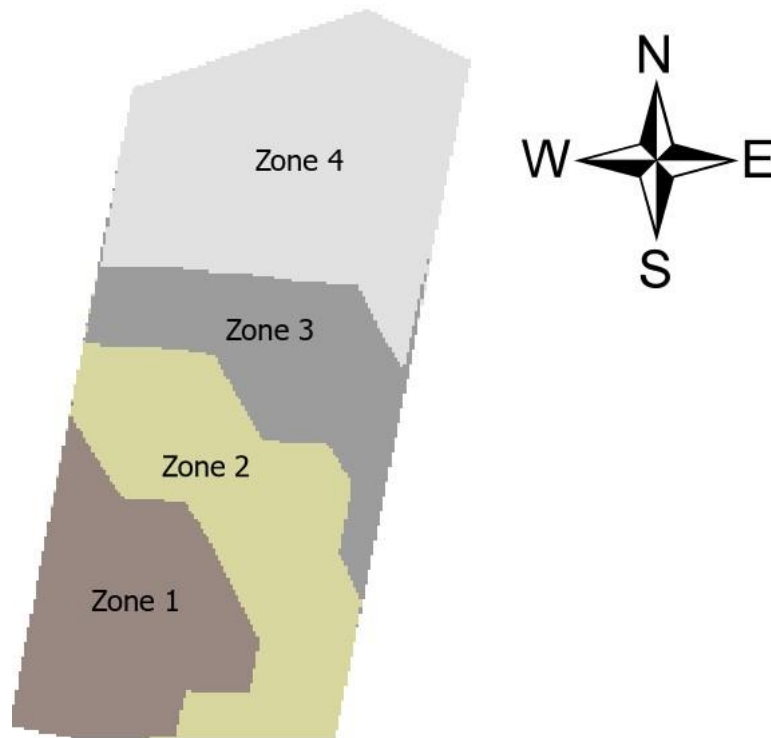


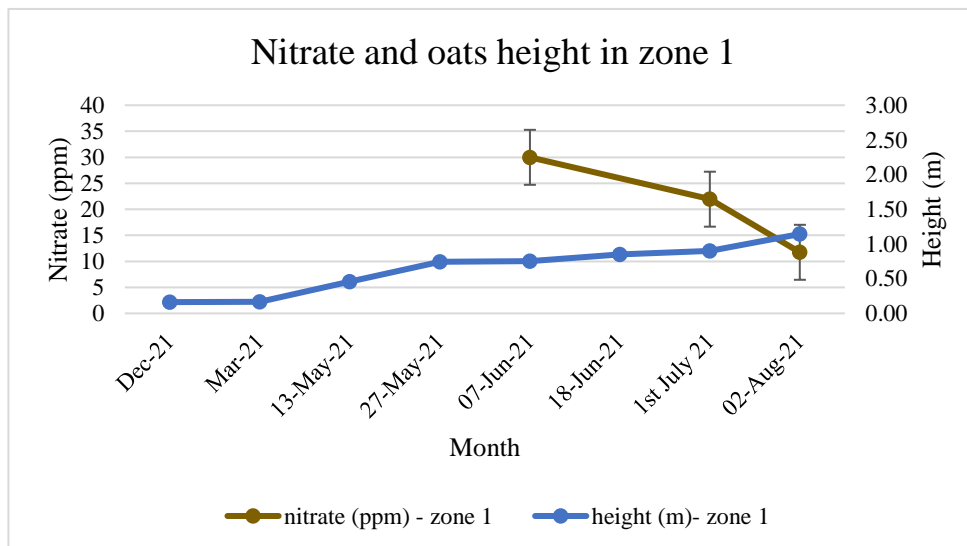
Figure 62: Zones used to compare oats height and soil NPK.

The oat height data for the zones were extracted across phenological growth stages from December 20 to August 21. The oats height data was compared to soil NPK results from the March 21 Agrovista data set and the June, July, and August 21 Palintest data set. Hence, the relationship between soil nitrate, phosphorus and oats height across all the zones shows an inverse relationship (Figure 63 and 64). The soil nitrate levels significantly decrease after June 21 across all the zones (Figure 63). This could be as the rate of nitrogen uptake by winter oats crop is greatest between the month of May and June which is also the panicle formation and flowering stage (Oats growth guide). Similarly, phosphorus levels significantly decrease after June 21 for zones 1, 2, and 3 (Figure 64 a, b and c), and there is a consistent phosphorus depletion across the months for zone 4 (Figure 64d). However, soil potassium decreases significantly from March 21 till June and after June there is an increase in potassium levels across all zones (Figure 65). The spike in potassium levels after June 21 is due to the addition

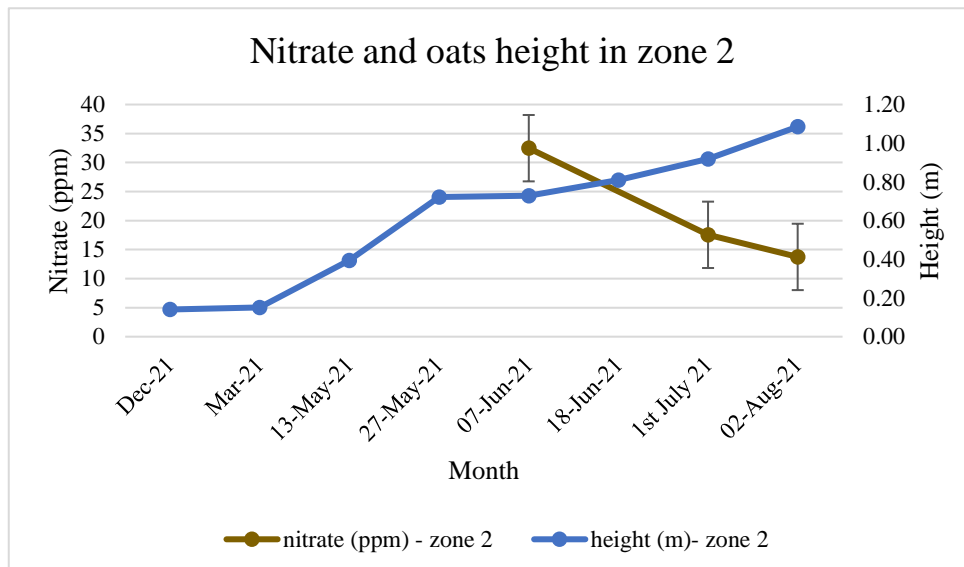
of potassium fertiliser by the farmers during the month of July 21. Therefore, the effect of soil potassium on oat height across phenological growth stages in different zones remains unclear.

Meanwhile, the height across zones increases significantly towards the end of May 21 (beginning of flowering stage), indicating that the height increase could be due to the oat crop's uptake of nitrate and phosphorus fertiliser. Therefore, oats height extracted by UAV data can be a potential indicator to understand if the crops have sufficiently utilised soil nitrate and phosphorus. However, UAV height data can only be used to understand fertiliser usage by crops assuming fertilisers were not administered over the essential growth period. For instance, the effect of potassium levels and oats height across the growth stages are unclear since fertiliser was added in July 21. (Figure 65).

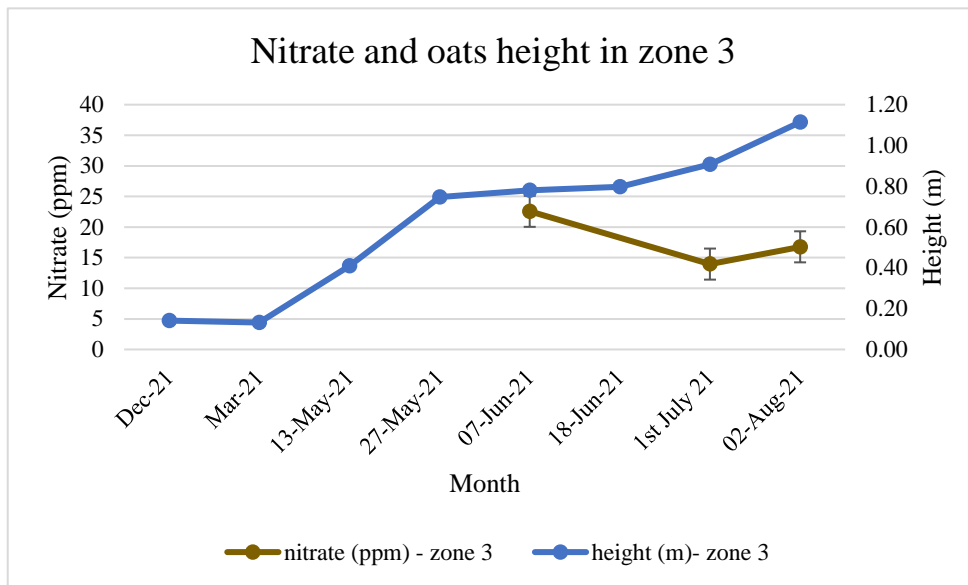
(a)



(b)



(c)



(d)

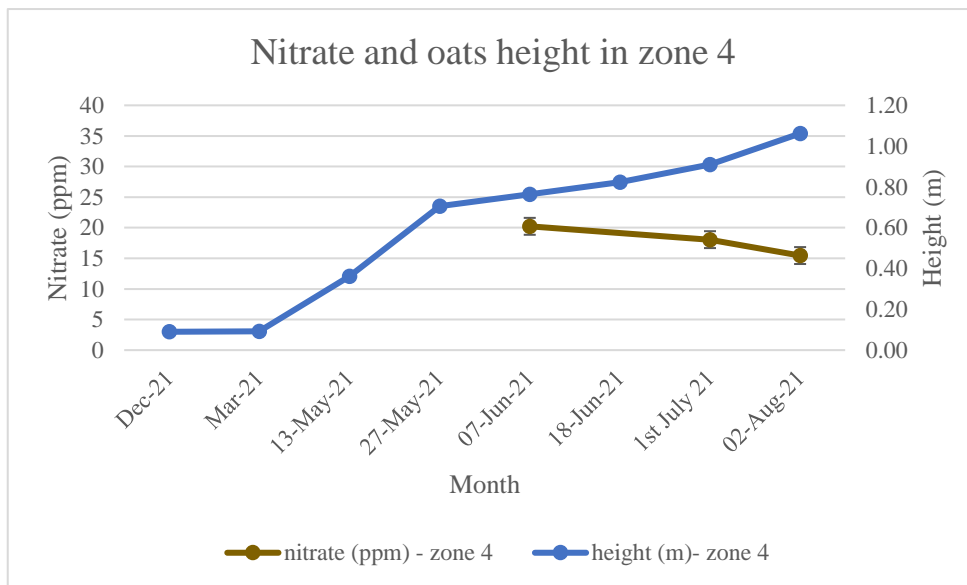
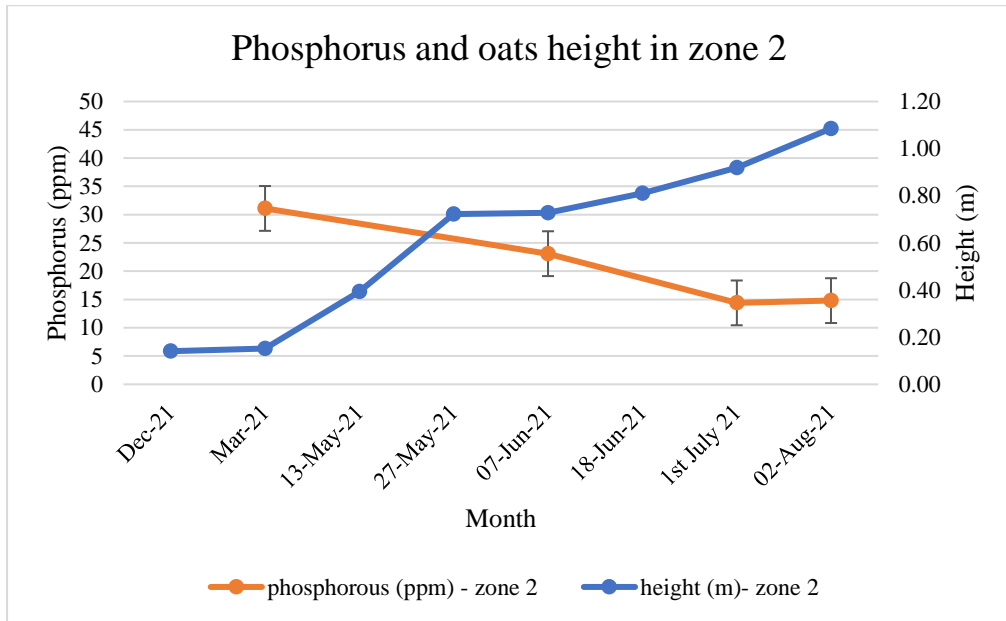
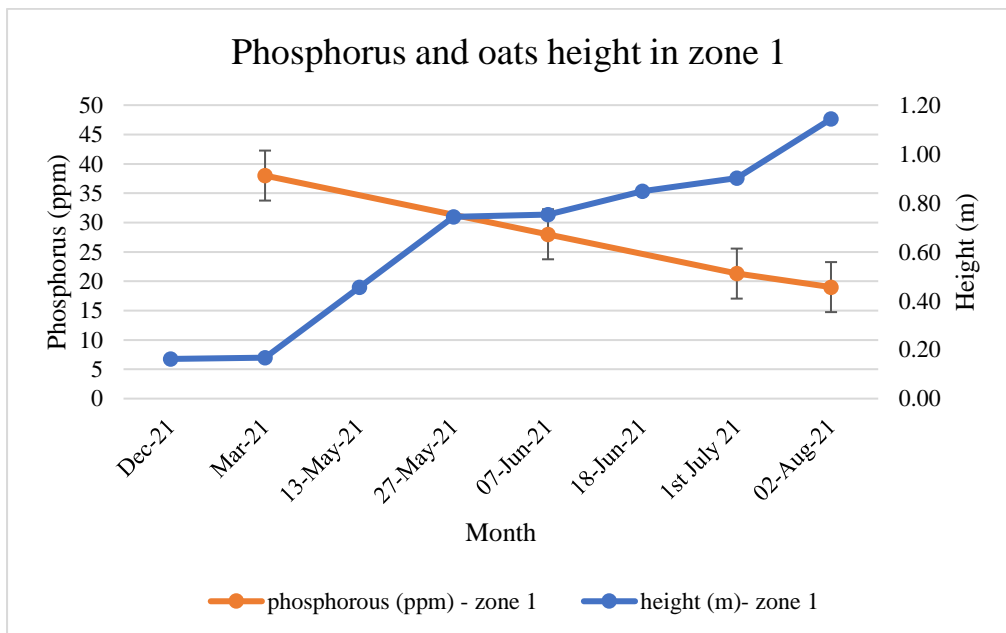


Figure 63: Soil nitrate and oats height in (a) zone 1 (b) zone 2 (c) zone 3 (d) zone 4.

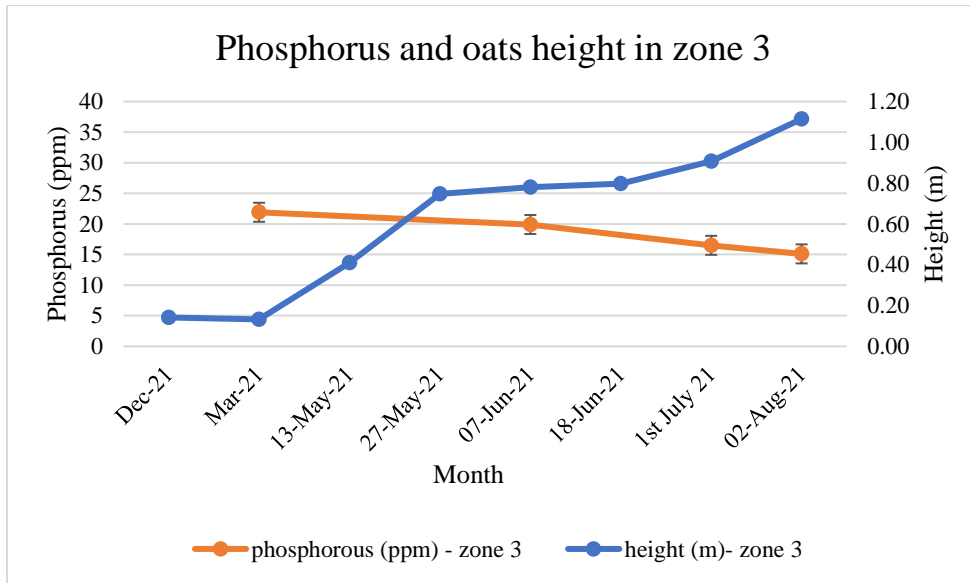
(a)



(b)



(c)



(d)

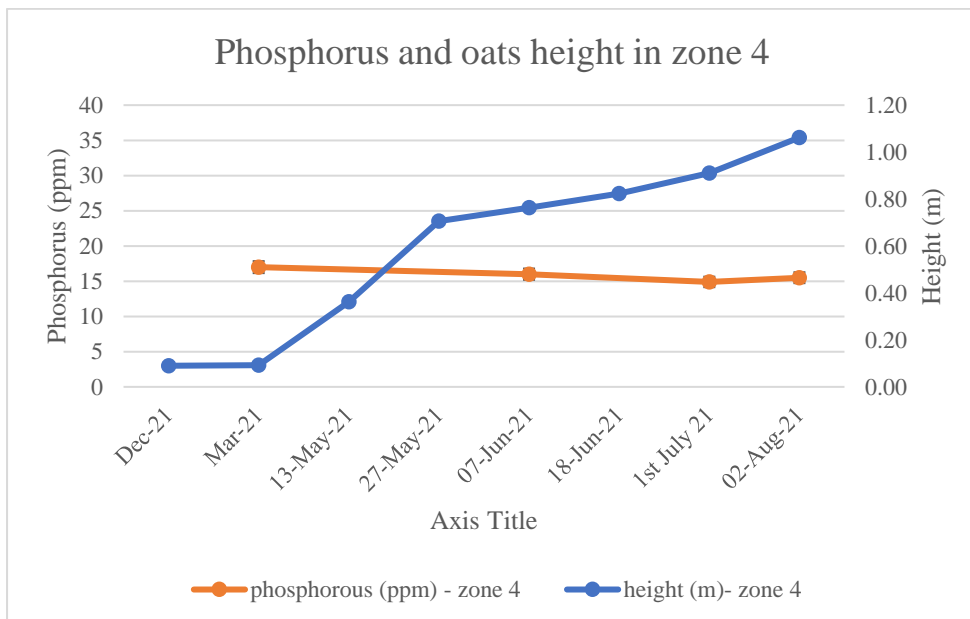
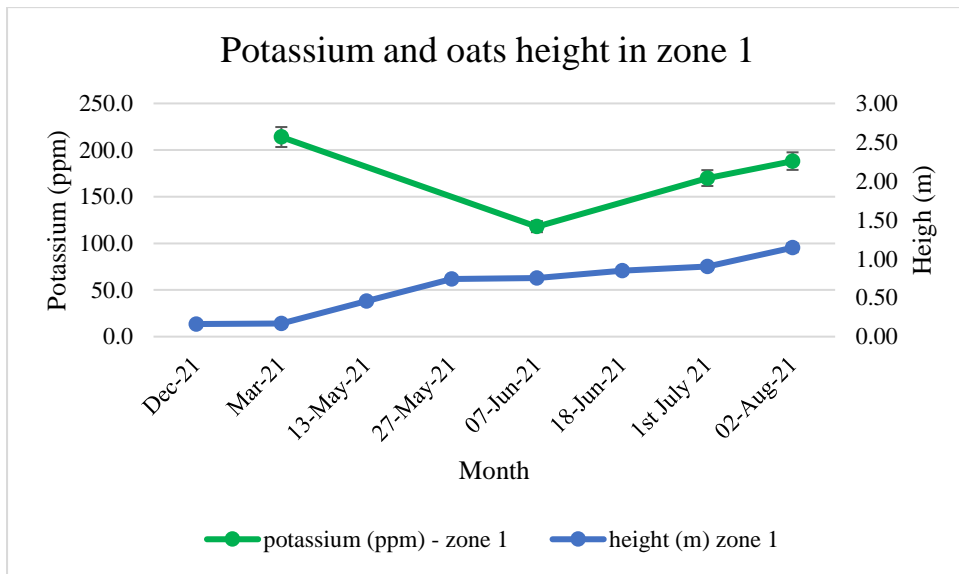
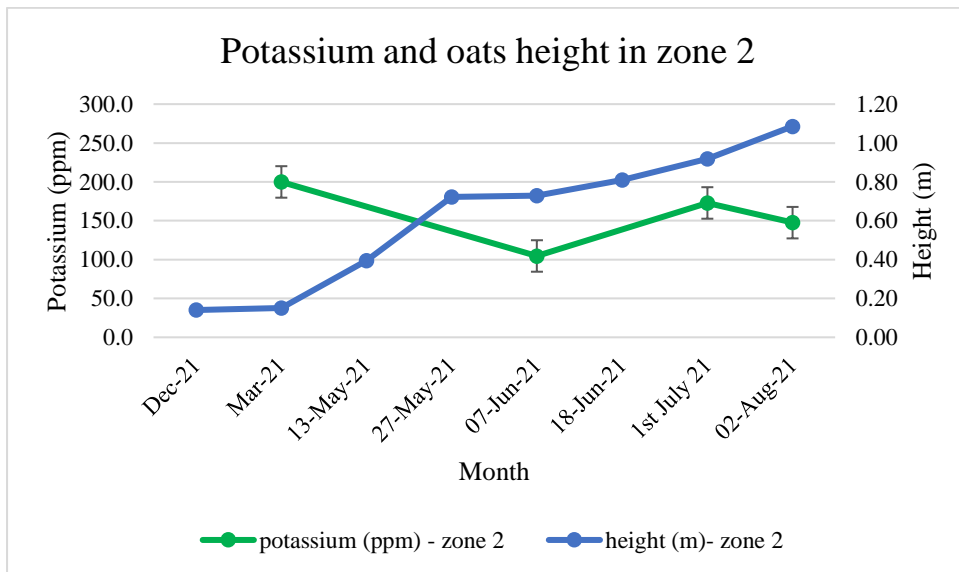


Figure 64: Soil phosphorus and oats height in (a) zone 1 (b) zone 2 (c) zone 3 (d) zone 4.

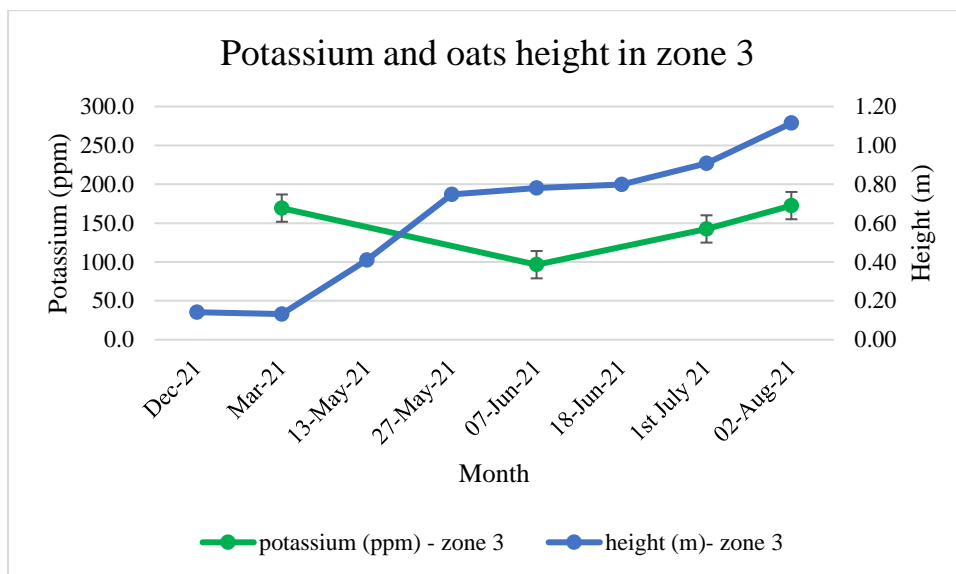
(a)



(b)



(c)



(d)

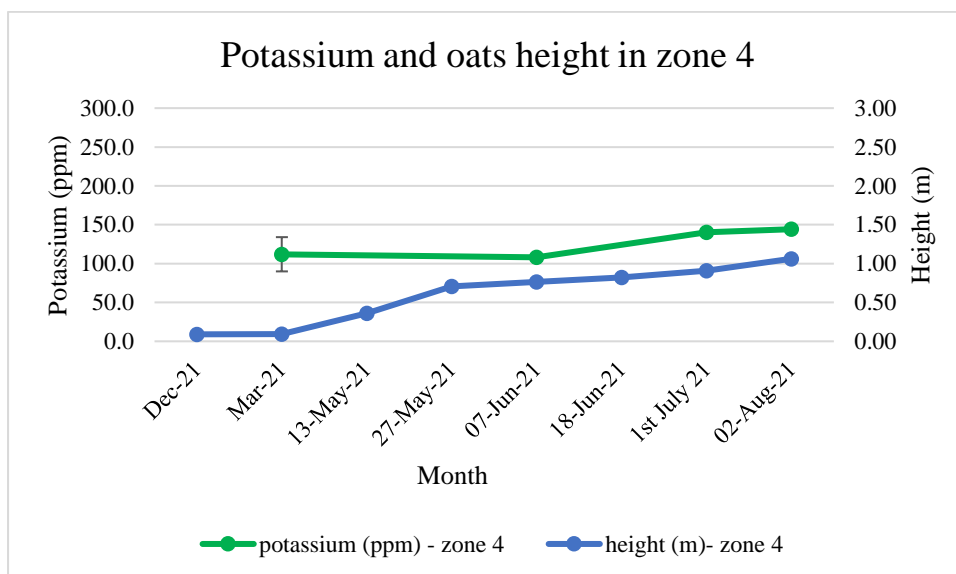


Figure 65: Soil potassium and oats height in (a) zone 1 (b) zone 2 (c) zone 3 (d) zone 4.

10.3.8 Correlation study between estimated and actual oats grain yield

Initially, the date of 7 June 2021 was chosen to determine the estimated oats grain yield which was the peak period of flowering of oats crop. Iso-cluster classification was used to classify the vegetation index maps into three clusters. The clusters were classified as soil surface, grasslands, and oats crop (Figure 66). Oats were classified as NDVI values between 0.6 - 0.8 and CI green values between 2 - 7. Grasslands were classified as NDVI values between 0.59 - 0.12 and CI green values between 1.8 - 0.5. Finally, soil surfaces were classified as NDVI values less than 0.12 and CI green values less than 0.5. As a result, the final oats grain yield was estimated by calculating the pixel areas of the oats crop visible at the top of the panicle classified by iso-cluster classification for the NDVI and CI green vegetation index maps.

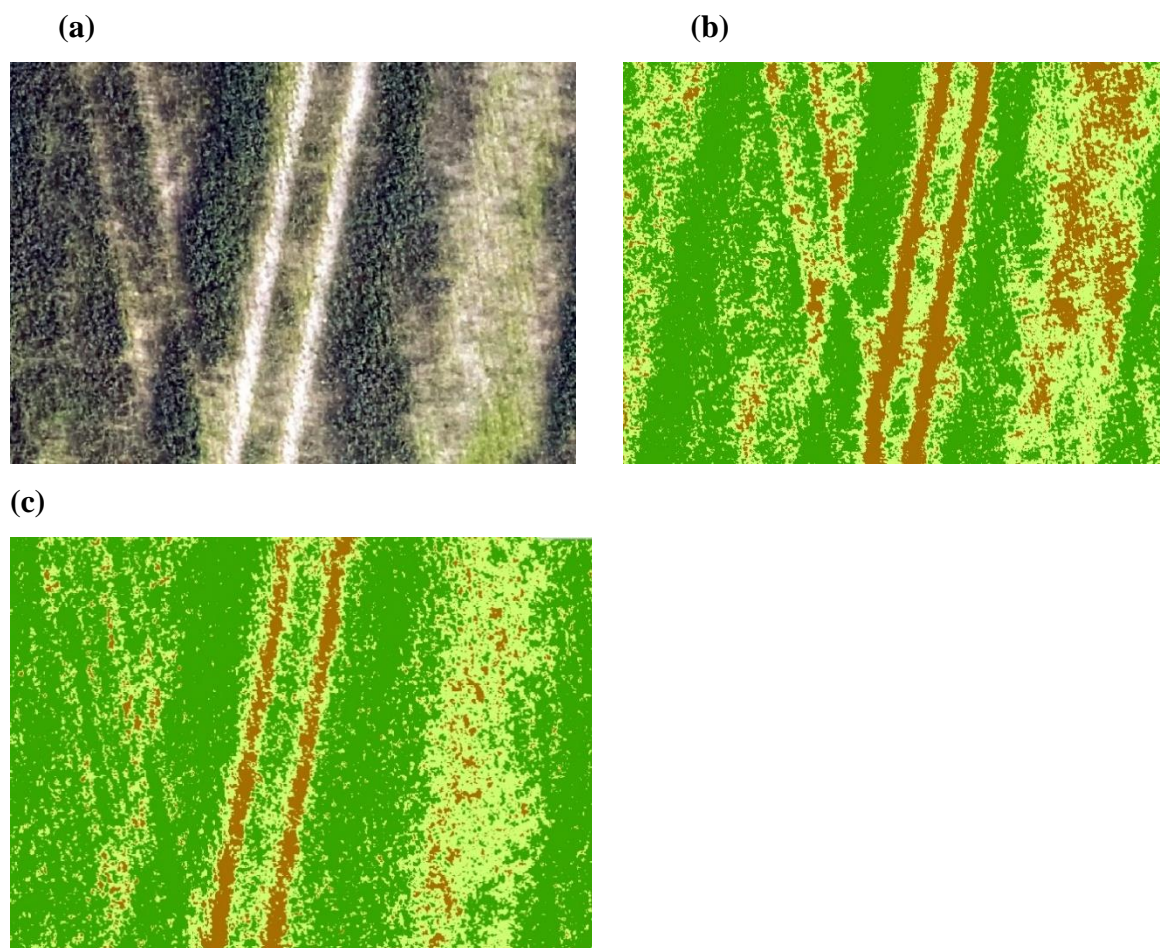
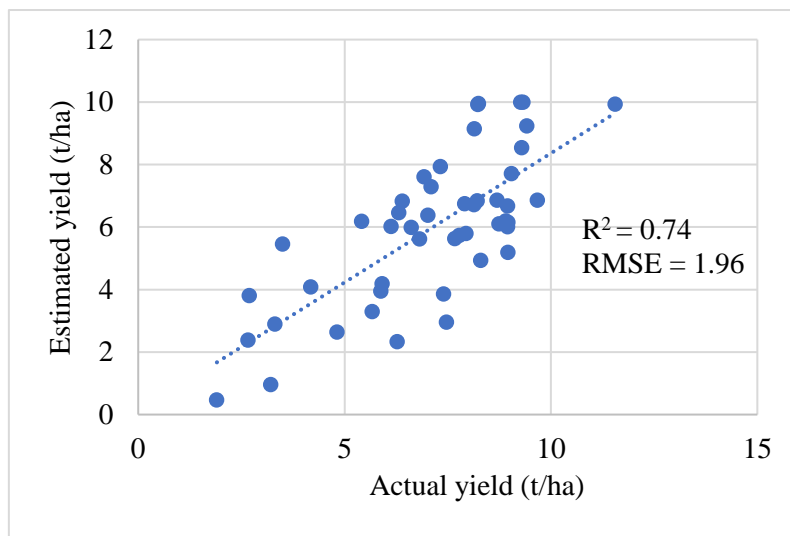


Figure 66: (a) RGB image (b) NDVI iso-clustered image (c) CI green iso-clustered image (d).

Note: Dark green represents oats crop. Light green represents grasslands. Brown represents soil surface.

The results (Figure 67) from the research study demonstrates a moderate correlation between the estimated oats grain yield and the actual oats grain yield for NDVI ($R^2 = 0.74$ and $RMSE = 1.96$) and CI green ($R^2 = 0.70$ and $RMSE = 1.83$). NDVI iso-cluster classified image (Figure 66) demonstrated better classification of oats crop from surrounding grasslands and soil surfaces than CI green. Therefore, this could explain the slightly better correlation between estimated and actual oats grain yield by NDVI in comparison to CI green. Also, research studies reveal that NDVI to have stronger correlation in estimating white oat grain yield in comparison to other vegetation indices (Coelho et al., 2019 and Coelho et al., 2020).

(a)



(b)

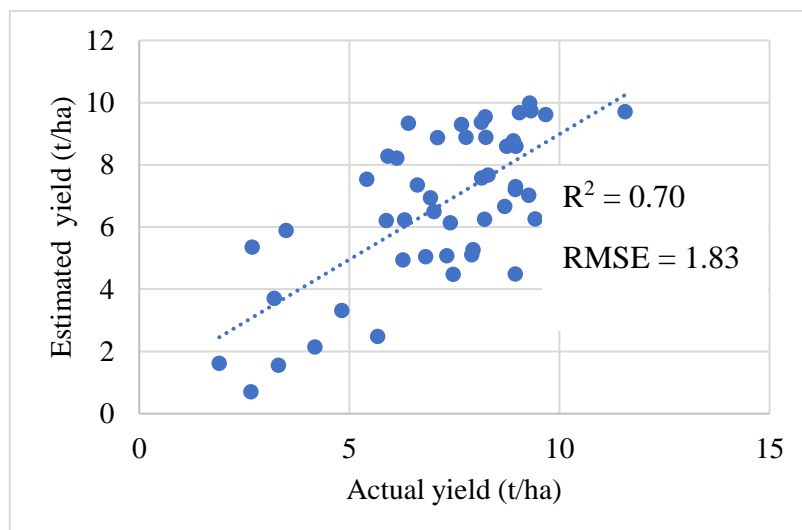


Figure 67: Correlation between estimated and actual yield (a) NDVI (b) CI green

10.3.9 Evaluation of thematic yield maps by kernel density estimation

Yield maps using estimated and actual oats grain yield were generated by a statistical pattern analysis method of KDE. The oats grain yield map variation was highlighted as follows: very low (red = 0 – 4.97 t/ha), low (orange = 4.97 – 6.18 t/ha), medium (yellow = 6.18 – 7.11 t/ha), high (light green = 7.11 – 7.62 t/ha) and very high (dark green > 7.62 t/ha) yield areas (Figure 69). According to Figure 68 the estimated oats grain yield maps by NDVI and CI green represent visually similar pattern to the actual oats grain yield measured in laboratory. Also, significant proportion of the low yield areas identified in both the estimated and actual yield maps are at the north end of the field. Further, according to the actual corrected oats grain yield map supplied by FarmTRX (Troo Corp., Ottawa, Canada), during the harvesting of oats is illustrated in Figure 69. In comparison to the corrected oats grain yield by FarmTRX, the yield maps generated by KDE method has shown for an improved sensitivity in spatially representing the yield patterns across the field. This is since, KDE yield maps provide a better visual representation of high and low yield clusters at respective locations, as well as indicating potential tractor line areas across the field (Figure 68). Furthermore, KDE examines the likelihood of yield in neighbouring areas from a defined yield in a location which improves the sensitivity of the yield maps. As a result, oats grain yield estimated by vegetation index iso-cluster classification can successfully be used to generate estimated oats grain yield map by KDE method across the field two months prior to harvesting of the oats crop. Furthermore, analysing these yield patterns prior to harvesting benefits farmers since specific crop management plans can be implemented to supply additional fertilisers to low yielding areas. Farmers can also specify crop prices and production destinations, such as whether it will be used for animal feed or as a human food product.

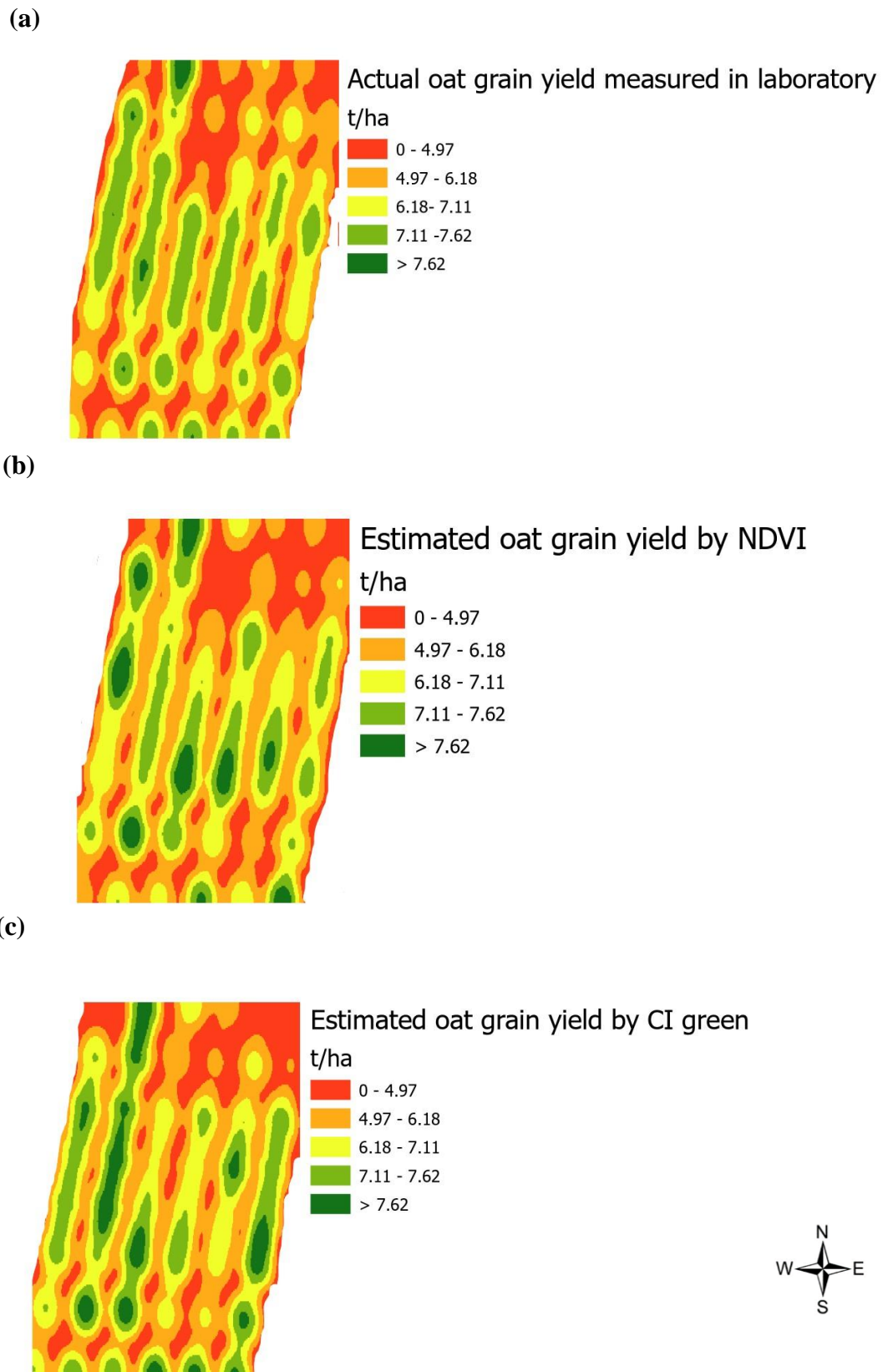


Figure 68: Oats grain yield maps by KDE (a) actual yield (b) estimated yield by NDVI (c) estimated yield by CI green.

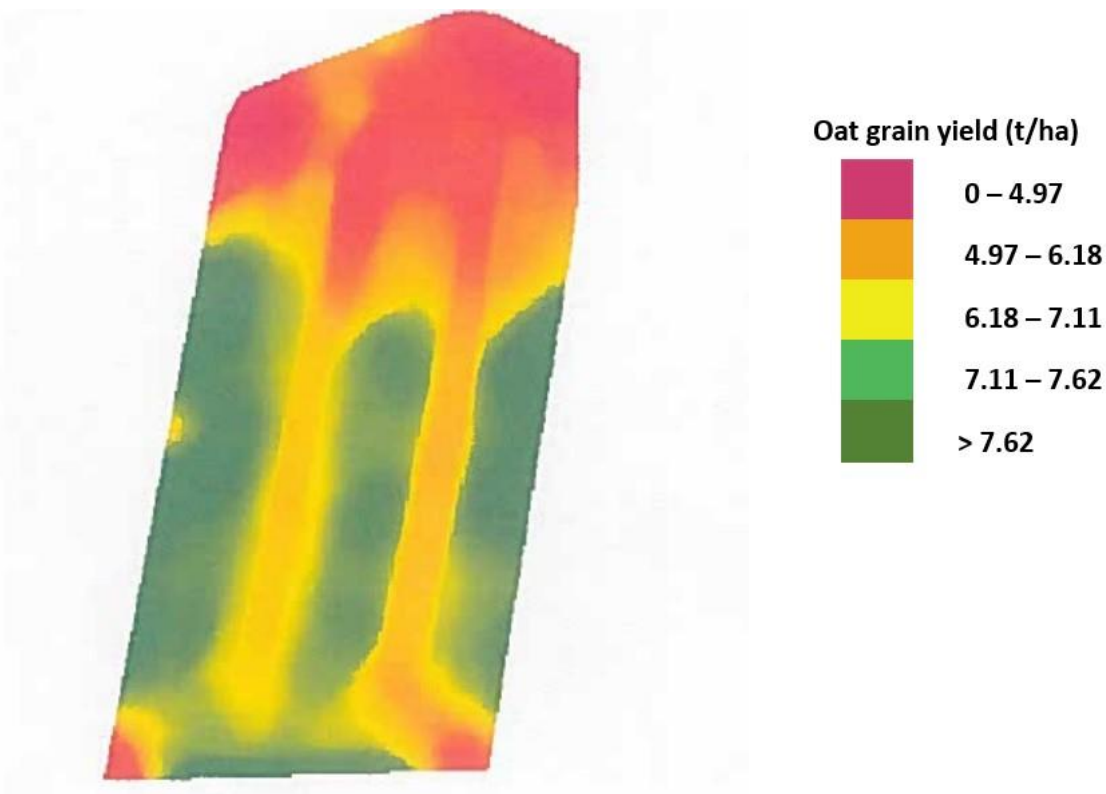


Figure 69: Actual yield map of oat grain yield by FarmTRX Note: The reported field area is 4.2 ha.

10.3.10 Evaluating extraction and conversion efficiency of oat beta-glucan and analytical performance parameters of analysing glucose in oat beta-glucan

Validation of the β -glucan determination, determined as glucose, was assessed across its four extraction stages, and reported (Appendix 2.4 a). The mean extraction efficiency, assessed by spiking glucose (15 mM) on oat extracts, was as follows: Stage 1 (before alkali extraction): mean 84.4% (85.2%; 83.5%); Stage 2 (before acid neutralisation): mean 90% \pm 5% (n = 6); Stage 3 (before freeze drying): 95% \pm 6% (n = 6); and, Stage 4 (before acid hydrolysis): mean 105% \pm 12% (n = 6). In addition, the conversion efficiency of β -glucan to glucan was assessed by spiking a sample with β -glucan (15 mM); this was determined to be 102% \pm 8% (N = 6) (Appendix 2.4 b)). The analytical performance data, using this colorimetric assay method, was determined as follows: for glucose, a lower limit of detection (LLOD) of 34 mg/l and a limit of quantitation (LOQ) of 102 mg/l. Precision was determined to be 9.5% at the low concentration (0.1 mM glucose) and 1.9 % at the higher concentration (40 mM glucose). The equivalent LLOD for β -glucan was 30 mg/l with an LOD of 92 mg/l (Appendix 2.4 c). The

LLOD and LOQ were determined using the standard curve method: $LLOD = (3.3\sigma)/S$ and $LOQ = (10\sigma)/S$, where σ is the standard deviation and S is the slope of the curve (ICH guideline Validation of Analytical Procedures – test and methodology). The full results for analysis of the 48-oat grain sub-samples for their β -glucan content in August 2021 are reported in Table 10.

10.3.11 Hierarchical multinomial logistic regression model

Matlab code used to develop the model is available in Appendix 5. With a given set of independent variables, a hierarchical multinomial logistic regression model can be built to predict the probabilities of the various possible outcomes of a categorical dependent variable. The dependent variable in this study is the yield of oat grain (either measured in the laboratory, or estimated via interpretation of UAV-MSI data as NDVI and CI green) which have previously been categorised into five yield classes: very low (red = 0 – 4.97 t/ha), low (orange = 4.97 – 6.18 t/ha), medium (yellow = 6.18 – 7.11 t/ha), high (light green = 7.11 – 7.62 t/ha) and very high (dark green > 7.62 t/ha) across the agricultural field (Figure 69). In contrast, the soil concentration of nitrate, P, K, as well as the soil pH and SOM as determined in June, July, and August 2021 and the β -glucan concentration, as determined in the harvested crop in August 2021, were used as independent variables in the model. The significance of building a model using the measured yield, and the UAV-MSI estimated yields using NDVI and CI green alongside soil indicators is that it helps identify the range of soil properties that are required to be maintained to obtain the desired yield. This in turn provides the users (farmers) who are operating a Precision Agriculture management strategy the information required to maintain the soil nutrient concentration by appropriate intervention by application of fertilisers. A flow chart (Figure 70) outlines the key steps required to build this machine learning model. Firstly, the measured yield, as well as the NDVI and CI green estimated yield by UAV data calculated for the 48 locations were recoded into the five yield classes, as previously indicated. After recoding, only 40-50 % similarity was evidenced in the measured yield, NDVI and CI green estimated yield by UAV data across all 48 datapoints (Table 14).

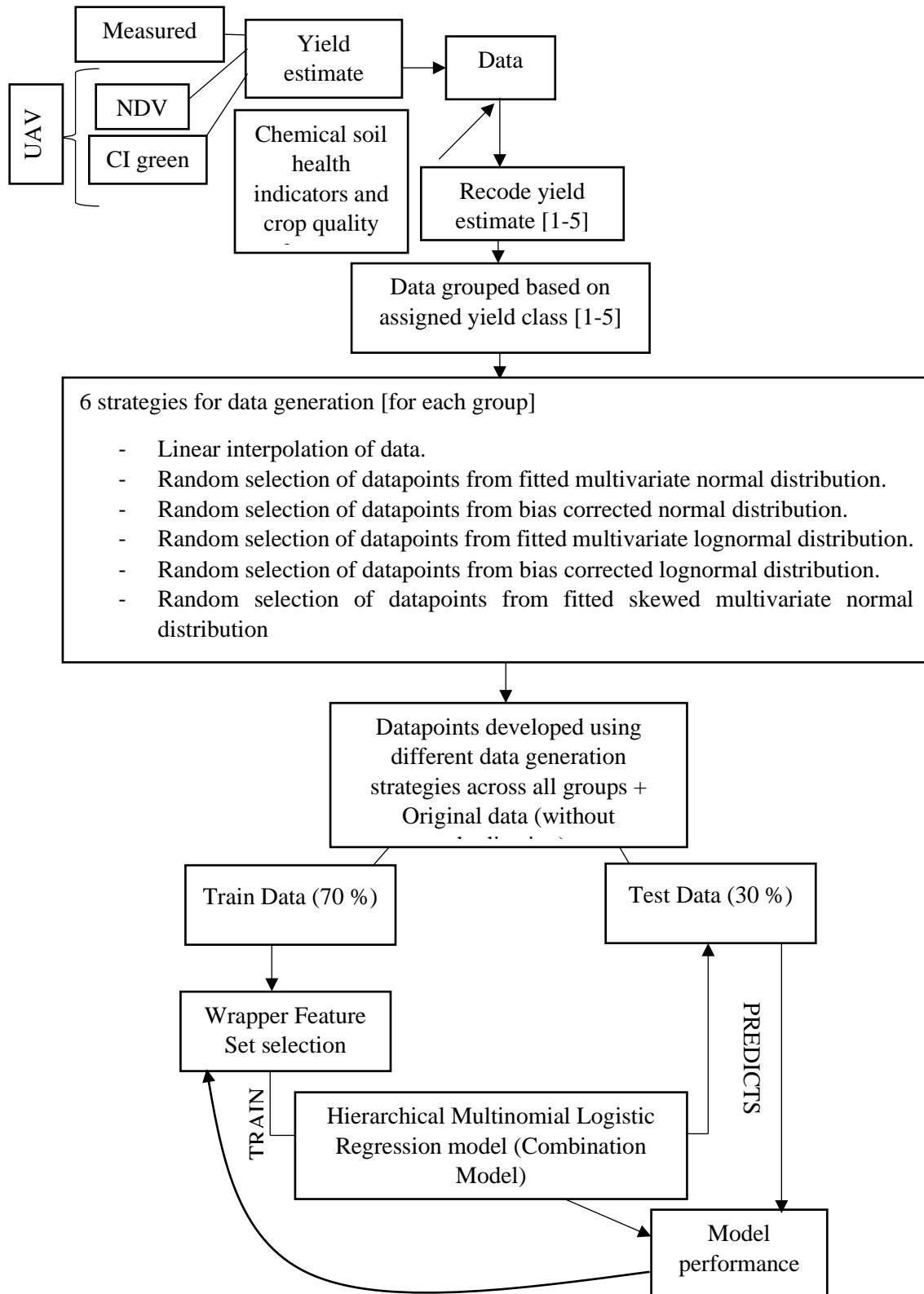


Figure 70: Flow chart representing data generation, model training and performance evaluation.

Table 14: Similarity in the recoded class assigned to the measured yield estimate, NDVI estimated yield and CI green estimated yield.

Recoded yield classes	Measured yield	Estimated yield by NDVI on UAV data	Estimated yield by CI green on UAV data
Measured yield	1	0.4167	0.4792
Estimated yield by NDVI on UAV data	0.4167	1	0.4583
Estimated yield by CI green on UAV data	0.4792	0.4583	1

Since 48 data points are insufficient for developing an effective model and validating its accuracy, the dataset was augmented using different approaches to generate sufficient data.

The 6 strategies used to generate data were:

- Multivariate linear interpolation of datapoints for each yield class.
- Random selection of datapoints from a normal distribution fitted to the sample dataset (consisting of original and interpolated datapoints for each yield class).
- Random selection of datapoints from a bias corrected normal distribution fitted to the sample dataset.
- Random selection of datapoints from a lognormal distribution fitted to the sample dataset.
- Random selection of datapoints from a bias corrected lognormal distribution fitted to the sample dataset.
- Random selection of datapoints from a skewed normal distribution fitted to the sample dataset.

Given the differences in the yield classes by measured yield and estimated yield by NDVI and CI green (Table 14), the data was generated separately for each of the three yield estimation methods. Firstly, for each of the yield estimation methods, data was grouped into the 5 yield classes. The multivariate datapoints assigned to each class were then linearly interpolated to have substantial number of datapoints in each class to train and test a classification model and to fit a multivariate probability distribution (Figure 71).

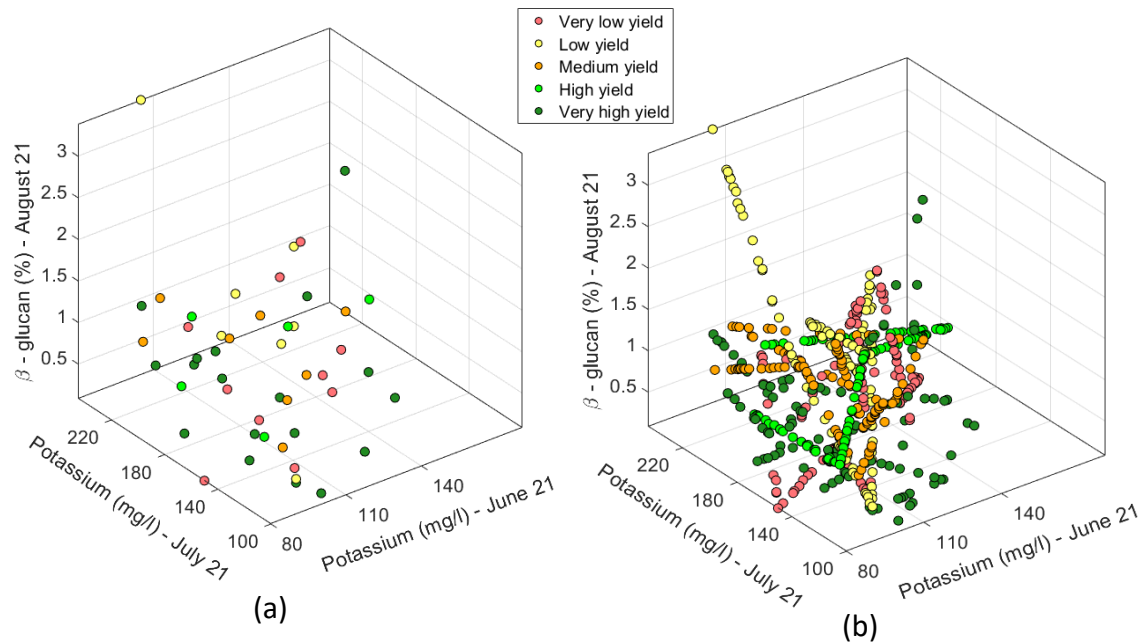


Figure 71: Scatter plot of CI green yield for selected features of potassium (mg/l) - June 21, potassium (mg/l)- July 21 and % β -glucan – August 21. (a) 48 datapoints. (b) linearly interpolated datapoints along with the original 48 datapoints.

[Note : Datapoints defined in terms of the soil nutrients were grouped based on the recorded five yield classes. Datapoints pertaining to each yield class were separately interpolated]

The linearly interpolated data assumes a constant rate of change between variables and is limited to the sample dataset's range of feature values. As a result, the linearly interpolated data across the five yield classes does not accurately reflect the population from the data points derived. Hence, in the absence of a representative sample dataset, the linearly interpolated dataset along with the original datapoints for each yield class was used to fit parametric distributions. Multivariate parametric distributions were fitted separately to the sample dataset of each yield class. It is hypothesized, that if sufficient datapoints were collected for each yield class then the data would follow a normal distribution.

Therefore, a multivariate normal distribution was fitted to datapoints for each yield class and a total of five multivariate normal distributions were developed for each yield estimated method (i.e., measured, NDVI and CI green estimated yield) (Figure 72). By neglecting datapoints consisting of any negative values, a set of 100 datapoints were randomly selected from each multivariate distribution.

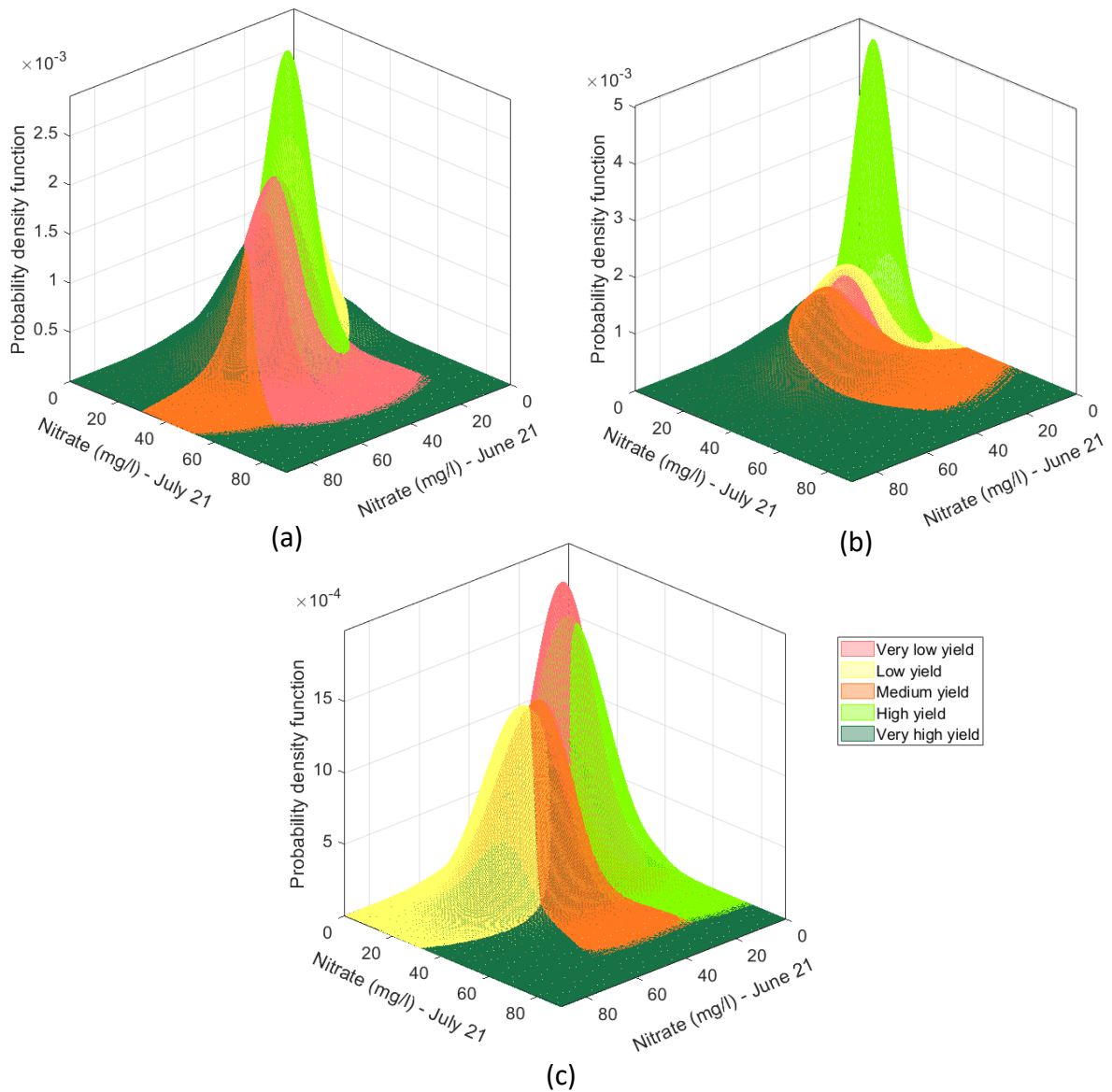


Figure 72: Multivariate normal distribution for nitrate (mg/l) – June 21 and nitrate (mg/l) – July 21 fitted to the datapoints grouped using the (a) recoded measured yield estimate, (b) recoded NDVI estimated yield (c) recoded CI green estimated yield.

Due to the presence of some extreme values (i.e., outliers) in the data, the covariance and mean of the fitted distribution are frequently overestimated and/or biased (Figure 73a). The Orthogonalized Gnanadesikan Kettenring algorithm (Maronna and Zamar, 2022) was therefore applied to the dataset since this corrects for the overestimation in the sample dataset. Using the overestimation corrected mean and covariance parameters a multivariate normal distribution

was generated for each of the 5 yield classes (Figure 73b). This allowed a set of 100 random points, with no negative values in any of the 16 features, to be selected for each yield class.

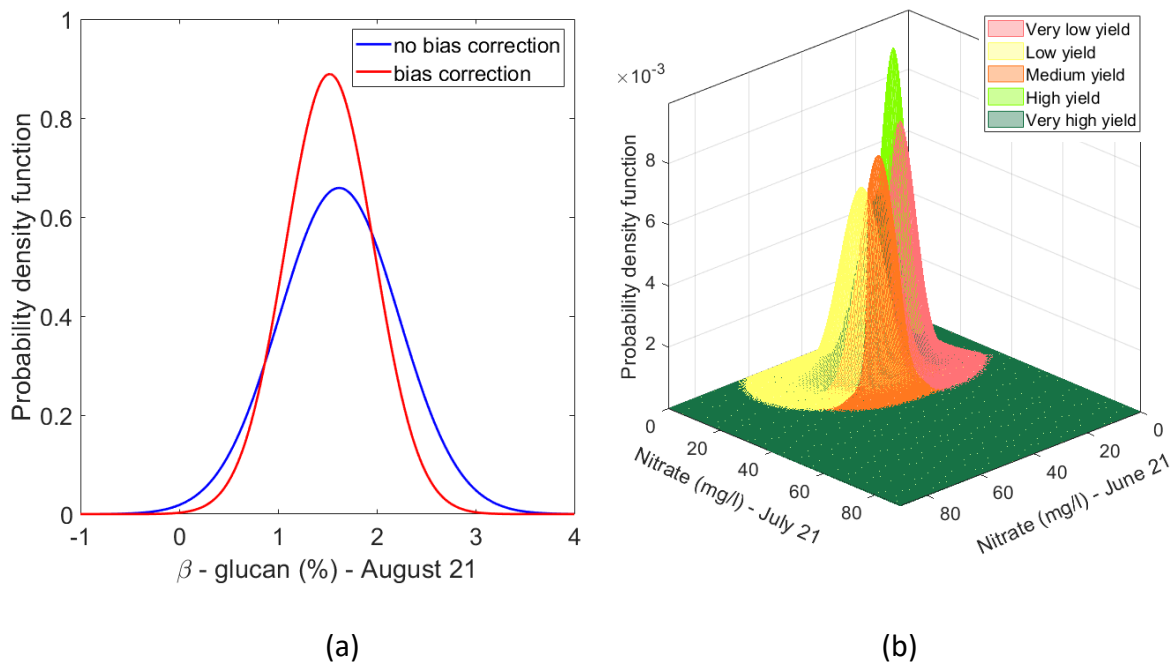


Figure 73: (a) Example of univariate distribution of β -glucan (%) - August 21 for low yield estimate using the CI green method (b) Example of Multivariate overestimation/bias corrected normal distributions of nitrate (mg/l) – June 21 and Nitrate (mg/l) –July 21 fitted to datapoints pertaining to each yield class based on the CI green yield estimate.

[Note: The univariate representation shows the effect of orthogonalized Gnanadesikan Kettenring method on the distribution of the normal distribution fitted to the variable. The original distribution is represented in blue. Overestimation of variance owing to presence of outliers is compensated by the orthogonalized Gnanadesikan Kettenring method]

The 16 features (nitrate, potassium, phosphate, pH, SOM in June 21, July 21, Aug 21 and % β -glucan in August 21) representing the soil characteristics and crop quality does not contain any negative values (Table 10). Due to this generated characteristic of the sample dataset, a multivariate lognormal distribution was fitted to the sample data for each of the 5 yield classes, in the three datasets pertaining to the yield estimation method (Figure 74a). The fitted multivariate lognormal distributions were used to randomly select 100 datapoints for each yield class. Further, after correction for overestimation and/or bias the mean and covariance, were used to fit lognormal distributions (Figure 74b) and 100 datapoints were randomly selected for each yield estimation method. Furthermore, a closer examination of the 48 datapoints reveals

that the variables can have absolute zero values (for example, absolute absence of nitrate in August Table 10). Hence, the dataset was adjusted by adding a value of 1.001 to fit a lognormal distribution. In view of the sample data characteristics, a skewed normal distribution was fitted to the sample dataset for each yield class (Figure 74c). From the skewed normal distribution, a set of 100 datapoints with no negative variable values were selected.

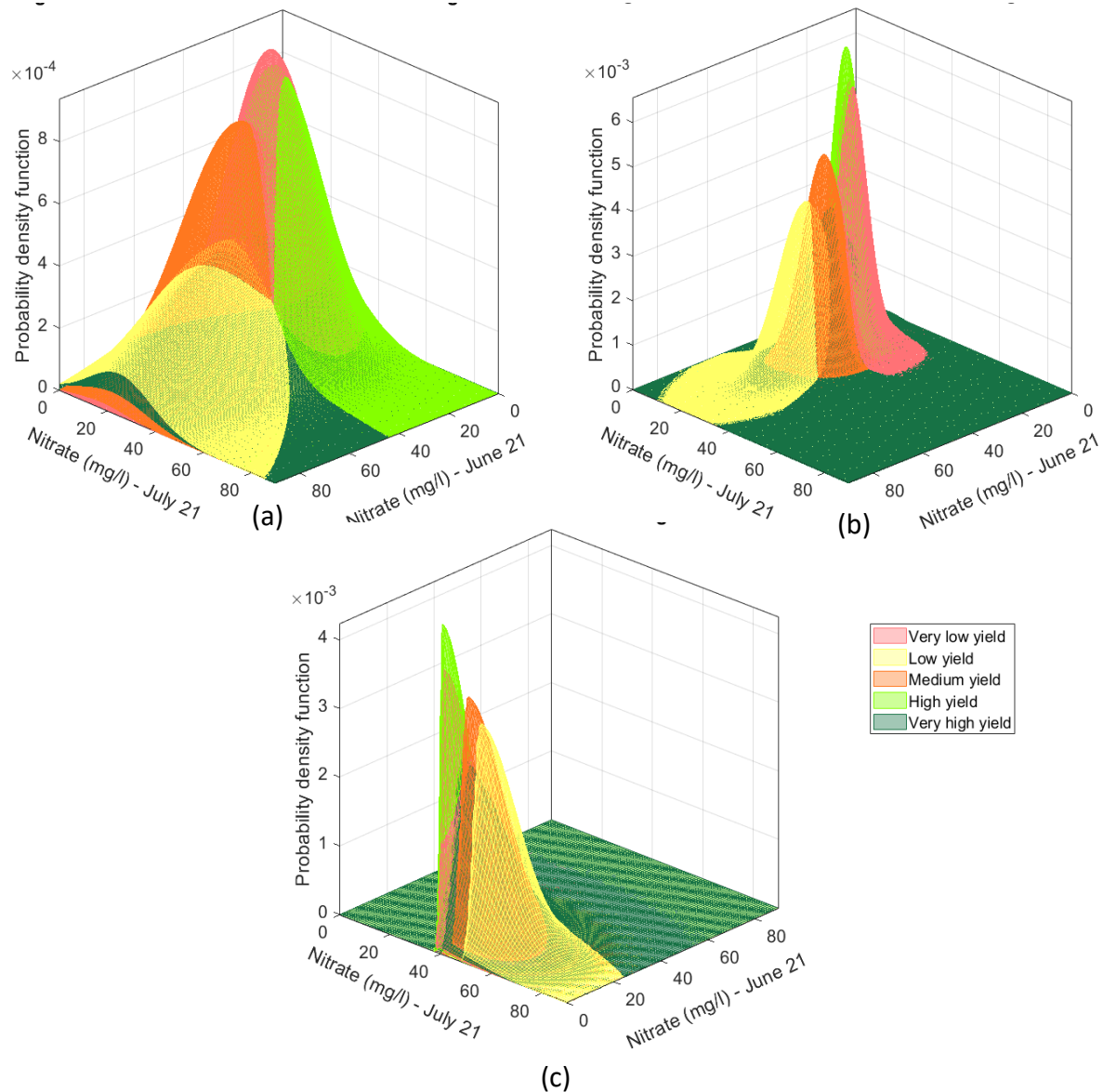


Figure 74: Multivariate distribution for CI green yield estimate representing the feature nitrate (mg/l) –June 21 and nitrate (mg/l) –July 21 (a) Multivariate log normal distributions fitted to datapoints grouped to 5 classes of CI green yield estimate (b) Multivariate lognormal distribution corrected for overestimation due to outliers by the Orthogonalized Gnanadesikan Kettenring method (c) Skewed normal distributions fitted to datapoints grouped based on the CI green yield estimate .

Hence, the six data generated strategies (e.g. the multivariate linear interpolation of datapoints, random selection of datapoints from a normal distribution fitted to the sample dataset, random selection of datapoints from a bias corrected normal distribution fitted to the sample dataset, random selection of datapoints from a lognormal distribution fitted to the sample dataset, random selection of datapoints from a bias corrected lognormal distribution fitted to the sample dataset and random selection of datapoints from a skewed normal distribution fitted to the sample dataset) for each of the 5 yield classes were used separately to train and test hierarchical multinomial logistic regression models.

Performance of each model was estimated by:

- Trained model deviance over an intercept only model at the 0.05 significance level.
- McFadden’s pseudo R^2 , and
- Accuracy, Sensitivity and Specificity of the model (as assessed on the data test set).

The variability of the yield can be explained by application of the McFadden's pseudo- R^2 approach (McFadden, 1973), and if the given model accurately fits the data, the McFadden's R^2 value will be close to 1 (Smith & McKenna, 2013). Specificity, on the other hand, indicates whether the model correctly identifies a datapoint as not belonging to a specific class. Whereas sensitivity indicates whether the model correctly identifies whether a datapoint belongs to a specific class. Finally, accuracy identifies the correctly predicted total number of datapoints by the model.

Hence, all generated datapoints and original samples were used to train and test the performance of a single hierarchical multinomial logistic regression model. The equations of the hierarchical multinomial logistic model in view of the five yield classes identified are:

$$\log\left(\frac{p(0 < y \leq 4.97)}{p(y > 4.97)}\right) = \beta_0 + \beta_{01} X_1 + \beta_{02} X_2 \dots \dots \beta_{016} X_{16} \dots \dots \dots (29)$$

$$\log\left(\frac{p(4.97 < y \leq 6.18)}{p(y > 6.18)}\right) = \beta_1 + \beta_{11} X_1 + \beta_{12} X_2 \dots \dots \beta_{116} X_{16} \dots \dots \dots (30)$$

$$\log\left(\frac{p(6.18 < y \leq 7.11)}{p(y > 7.11)}\right) = \beta_2 + \beta_{21} X_1 + \beta_{22} X_2 \dots \dots \beta_{216} X_{16} \dots \dots \dots (31)$$

$$\log\left(\frac{p(7.11 < y \leq 7.62)}{p(y > 7.62)}\right) = \beta_3 + \beta_{31} X_1 + \beta_{32} X_2 \dots \dots \beta_{316} X_{16} \dots \dots \dots (32)$$

To prevent localized clustering of datapoints that have the same yield class assigned or were generated from the same distribution, the datapoints were randomly shuffled. 70% of the unique datapoints in the dataset were used to train the model, and the remaining 30% were used to test the model. Since the recoded yield class assigned to the manually measured yield and the estimated yields using NDVI and CI green represented about 40-50% similarity, three hierarchical multinomial logistic regression models were separately trained and tested. Finally, the hierarchical multinomial logistic regression models trained on the data developed using the six different strategies will hereafter be referred to as a 'combination model'.

From the training data, sets of variables/ features (i.e., soil nutrient eg. Nitrate, P, K etc. in this case) that have a correlation lower than 0.4 among themselves were selected and the logistic regression model was trained and tested using a dataset defined by only these sets of variables. The prediction accuracy of the models trained on the selected feature sets was compared among themselves and with a model that is trained on a dataset that takes into consideration all the 16 variables/features. This was done to estimate the effect of multicollinearity (i.e., features bringing in similar/overlapping information) on classifier performance. The threshold value of 0.4 was arbitrarily selected.

For the datasets generated based on the estimated yield using NDVI and CI green, the combination model using all 16 features (nitrate, P, K, SOM, pH in June 21, nitrate, P, K, SOM, pH in July 21 and nitrate, P, K, SOM, pH, and β -glucan in August 21) was found to have the best prediction accuracy. However, this soil parametric feature set selection is dependent on the training set and the prediction performance of the training model based on the test set data used in the study. While for the measured yield data a subset of 9 features (K, P in June 21, nitrate, K, P, SOM in July 21 and nitrate, K, and β -glucan in August 21) was found to perform best (Table 15). The 9 features are found to be relevant for all the three yield models, separately trained and tested, using the multinomial logistic regression model (i.e. K, P in June 21, nitrate, K, P, SOM in July 21 and nitrate, K, and β -glucan in August 21).

Table 15: Performance of hierarchical multinomial logistic regression model (i.e. Combination model) trained on measured yield data, estimated yield data using the NDVI and CI green data. Performance is defined based on the goodness of fit of the model to the training data and the model's ability to make predictions on unseen/test data.

Combination model - Measured yield			
Significance of Model developed over Intercept only model	$\chi^2 (36) = 4.6701e+03, p=0 <0.05$ [Based only on training data]		
McFadden's R^2	0.9476 [Based only on training data]		
test_data	Specificity	Sensitivity	Accuracy
Class 1 - very low yield (0 – 4.97 t/ha)	0.56	0.98	0.89
Class 2 - low yield (4.97 – 6.18 t/ha)	0.75	0.94	0.91
Class 3 - medium yield (6.18 – 7.11 t/ha)	0.76	0.97	0.93
Class 4 - high yield (7.11 – 7.62 t/ha)	0.89	0.88	0.88
Class 5 - very high yield (>7.62t/ha)	0.72	0.90	0.86
Overall			0.74
Combination model – NDVI yield			
Significance of Model developed over Intercept only model	$\chi^2 (64) = 4.9506e+03, p=0 <0.05$ [Based only on training data]		
McFadden's R^2	0.5594 [Based only on training data]		
test_data	Specificity	Sensitivity	Accuracy
Class 1 - very low yield (0 – 4.97 t/ha)	0.13	1.00	0.83
Class 2 - low yield (4.97 – 6.18 t/ha)	0.36	0.89	0.78
Class 3 - medium yield (6.18 – 7.11 t/ha)	0.81	0.80	0.80
Class 4 - high yield (7.11 – 7.62 t/ha)	1.00	0.86	0.89
Class 5 - very high yield (>7.62t/ha)	0.73	0.97	0.92
Overall			0.61
Combination model – CI green yield			
Significance of Model developed over Intercept only model	$\chi^2 (64) = 5.7487e+03, p=0 <0.05$ [Based only on training data]		
McFadden's R^2	1.00 [Based only on training data]		
(test data)	Specificity	Sensitivity	Accuracy
Class 1 - very low yield (0 – 4.97 t/ha)	0.58	0.99	0.91
Class 2 - low yield (4.97 – 6.18 t/ha)	0.47	0.94	0.84
Class 3 - medium yield (6.18 – 7.11 t/ha)	0.95	0.86	0.88
Class 4 - high yield (7.11 – 7.62 t/ha)	1.00	0.96	0.97
Class 5 - very high yield (>7.62t/ha)	0.87	0.95	0.93
Overall			0.76

As shown in Table 15, the hierarchical multinomial logistic regression model for the CI green estimated yield data provides the best performance in terms of specificity, sensitivity, and As

As evidenced in Table 15, the hierarchical multinomial logistic regression model for the CI green estimated yield data gives the best performance. The trained hierarchical multinomial logistic regression model equation for the CI green estimated yield are:

$$\log\left(\frac{p(0 < y \leq 4.97)}{p(y > 4.97)}\right) = -4.54 + 0.22X_1 + 1.44X_2 \dots \dots - 0.19X_{16} \dots \dots \dots (33)$$

$$\log\left(\frac{p(4.97 < y \leq 6.18)}{p(y > 6.18)}\right) = -5.79 - 0.5X_1 + 2.77X_2 \dots \dots + 3.6X_{16} \dots \dots \dots (34)$$

$$\log\left(\frac{p(6.18 < y \leq 7.11)}{p(y > 7.11)}\right) = -4.38 - 0.63X_1 - 3.15X_2 \dots \dots + 8.32X_{16} \dots \dots \dots (35)$$

$$\log\left(\frac{p(7.11 < y \leq 7.62)}{p(y > 7.62)}\right) = -146.66 - 37.93X_1 + 12.69X_2 \dots \dots - 60.42X_{16} \dots (36)$$

The coefficients $\beta_0, \beta_{01}, \beta_{02}$ and β_{016} in Eq. (29), $\beta_1, \beta_{11}, \beta_{12}$ and β_{116} in Eq. (30), $\beta_2, \beta_{21}, \beta_{22}$ and β_{216} in Eq. (31) and $\beta_3, \beta_{31}, \beta_{32}$ and β_{316} in Eq. (32) were replaced with their values on the trained CI green data model in Eq. (33) - (36). Also, the coefficients of β for the 16 features in the CI green combination model (Table 15) for the equation 33 – 36 which for the five yield classes is shown in Table 16. The specificity, sensitivity and accuracy on the trained measured yield data model, yield estimated using NDVI and CI green data models are provided in Appendix 6 (Table 6.1).

The performance of the models separately fitted to the linearly interpolated datapoints and the randomly selected points from different fitted parametric distributions for the measured yield data, yield estimated using NDVI and CI green method on UAV data are also presented in Appendix 6 (Table 6.2).

Table 16: The variables of hierarchical multinomial logistic regression model equation for the CI green yield which satisfies all the five yield classes.

Equation variable	Feature/Variable	B variable of eq. 33	Coefficient of β for eq. 33	β variable of eq. 34	Coefficient of β for eq. 34	β variable of eq. 35	Coefficient of β for eq. 35	β variable of eq. 36	Coefficient of β for eq. 36
Constant	Constant	β_0	-4.54	β_1	-5.79	β_2	-4.38	β_3	-146.66
X ₁	Nitrate (mg/l) - June 21	β_{01}	0.22	β_{11}	-0.50	β_{21}	-0.63	β_{31}	-37.93
X ₂	Potassium (mg/l) - June 21	β_{02}	1.44	β_{12}	2.77	β_{22}	-3.15	β_{32}	12.69
X ₃	Phosphorus (mg/l) - June 21	β_{03}	-4.56	β_{13}	-4.40	β_{23}	-0.01	β_{33}	4.30
X ₄	soil organic matter (mg/l) - June 21	β_{04}	-0.58	β_{14}	1.33	β_{24}	1.84	β_{34}	-44.82
X ₅	pH - June 21	β_{05}	4.21	β_{15}	-6.54	β_{25}	-9.91	β_{35}	179.75
X ₆	Nitrate (mg/l) - July 21	β_{06}	-0.57	β_{16}	0.67	β_{26}	-1.16	β_{36}	53.20
X ₇	Potassium (mg/l) - July 21	β_{07}	2.03	β_{17}	-0.54	β_{27}	0.03	β_{37}	-26.29
X ₈	Phosphorus (mg/l) - July 21	β_{08}	-0.75	β_{18}	0.14	β_{28}	-1.73	β_{38}	-18.33
X ₉	Soil organic matter (mg/l) - July 21	β_{09}	-0.61	β_{19}	3.19	β_{29}	1.55	β_{39}	-31.02
X ₁₀	pH - July 21	β_{010}	2.71	β_{110}	0.30	β_{210}	-2.18	β_{310}	-272.26
X ₁₁	Nitrate (mg/l) - August 21	β_{011}	0.67	β_{111}	-3.13	β_{211}	1.70	β_{311}	-10.56
X ₁₂	Potassium (mg/l) - August 21	β_{012}	-1.66	β_{112}	0.62	β_{212}	2.38	β_{312}	-27.49
X ₁₃	Phosphorus (mg/l) - August 21	β_{013}	1.30	β_{113}	-2.04	β_{213}	1.28	β_{313}	16.87
X ₁₄	soil organic matter (mg/l) - August 21	β_{014}	0.65	β_{114}	-2.14	β_{214}	-8.72	β_{314}	108.45
X ₁₅	pH - August 21	β_{015}	-6.12	β_{115}	3.24	β_{215}	7.34	β_{315}	130.97
X ₁₆	β - glucan (%) - August 21	β_{016}	-0.19	β_{116}	3.60	β_{216}	8.32	β_{316}	-60.42

In terms of the usability of the model as part of a Precision Agriculture management strategy, based on the nutrient quantity of the soil in June 21, July 21 and August 21 and crop quality feature of β -glucan in August 21, the ‘combination model’ by CI green (Table 16) can assist farmers, and agronomists, to estimate the oat crop yield, across the 3 phenological growing seasons (months). This is since the chosen 16 soil feature parameters show good, specificity, sensitivity, and accuracy at estimating the oat crop yield over the five different classes (Table 15). However, this probabilistic machine learning model only holds if the crops aren’t affected by any natural or adverse calamities in the months before harvest.

But, using the original 48 datapoints a proportional range of concentration of the different soil nutrients in June 21, July 21, August 21 and β -glucan in August 21 can be used to classify the oat yield by all the three methods yield estimation. This can be performed by taking into consideration only the datapoints that have been identically grouped by all three yield estimate methods of measured yield, NDVI and CI green yield. However, in coherence with the similarity outlined in Table 14, only 6 datapoints were identically marked as ‘very low yield’ by all the three yield estimate models; no agreement on any datapoint was found for ‘low yield’ and ‘high yield’ classes among the three methods used for yield estimation. Again, only 2 datapoints were found to be marked out as ‘medium yield’ all the three yield estimation methods, Also, 6 datapoints were found to be marked out as ‘very high yield’ (>7.62t/ha). by all the three methods yield estimation method and the range of the soil nutrient composition for the identified 6 datapoints has been represented in Table 17

Table 17: Proportional concentration range of soil nutrients that resulted in very high yield across all 3 methods

All yield estimates	Nitrate(mg/l) - June 21 (Flowering)	Nitrate (mg/l) - July 21 (Grain filling)	Nitrate (mg/l) - August 21 (Ripening)
Class 5 - very high yield (>7.62t/ha)	15.50 - 80.60	6.60 - 37.70	3.50 - 35.40
	Phosphorus (mg/l) - June 21	Phosphorus (mg/l)- July 21	Phosphorus (mg/l) - August 21
	20 - 33	13 - 27	7 - 20
	Potassium (mg/l) - June 21	Potassium (mg/l) - July 21	Potassium (mg/l) - August 21
	90 - 165	105 - 235	125 - 190
	pH - June 21	pH - July 21	pH - August 21
	7.46 - 7.84	7.31 - 7.83	7.65 - 7.87
	SOM - June 21	SOM - July 21	SOM - August 21
	9.22 - 10.80	9.09 - 11.68	7.77 - 10.51
			β -glucan (%) - August 21
		0.37 - 2.26	

Maintaining the range of soil nutrients and their action leading to crop quality (i.e. % β -glucan) are required as part of a Precision Agricultural management strategy for winter oats. This is attributable to the farmer's knowledge that the soil nutrient levels in the winter oats crop must be kept within ranges to maintain very high yields at the three different growth stages of oats in flowering (June 21), grain filling (July 21), and ripening (August 21) (Table 17). Also, it provides an idea to the farmer on the quality of oats (% β -glucan) which would be expected for very high yield by the time of harvest (Table 17). Also, according to the winter growth guide, the β -glucan content of winter oat Mascani grown in the North of UK ranges from 3.7 % to 4.2%, with a mean of 3.9 % (Oat growth guide, 2019). In our study the β -glucan for very high yield ranges between 0.37 – 2.26 %. Variation in the oat β -glucan reported in this study can be since oat β -glucan significantly differ among cultivator, and are affected by growing locations, storage, and processing condition (Paudley et al., 2021).

Further, it must be noted that these ranges for soil nutrients and β -glucan (Table 17) for the very high yield are represented using the original data set and not the simulated data. This is since the main purpose of this pilot study elucidates that it is possible to predict yield based on soil nutrients and crop quality features using the 'combination model'. Hence, the concentration ranges of soil nutrients and β -glucan cannot simply be provided to the farmer for very high yield (Table 17) without this model since it must be initially evaluated that soil nutrients and β -glucan can be used to estimate oat yield.

However, to further improve the proposed yield prediction model and the proportional soil nutrient concentrations, more data collection and further experimentation would be required. Further, this 'combination model' particularly CI green (Table 16) can be used for future studies in alternate crops to identify if the chemical soil health indicators and crop quality features aid in estimation of crop yield. Hence, if the model can estimate yield effectively using soil nutrients and crop quality features, the farmer can be provided with ranges for soil nutrients and crop quality features for very high yield at different growth stages of the crop. This would be beneficial for the precision agriculture of crops.

10.4 Conclusion

Finally, across the phenological growth stages of oats, UAV GRVI data are very similar to ground reference data. Positive GRVI values were obtained when oats were green and negative values were recorded when the oats were ready to harvest. Also, the negative GRVI values for most data locations during senescence were statistically insignificant ($p < 0.05$) to ground reference data. Furthermore, NDVI and CI green values appear promising for differentiating

oats across different phenological growth stages. However, it is worthwhile to validate these indices using ground reference data. In addition, the validity of the indices to ground reference data must be determined by comparing the calibration methods of reflectance calibration method to irradiance calibration method.

With reference to VI and chemical soil health indicators a positive strong relationship was observed between NDVI, CI green and soil nitrate. However, no potential relationship was observed between NDVI, CI green and other soil health indicators (potassium, phosphorus, pH and organic matter). Hence, the strong relationship between UAV-derived VI and soil nitrate can help farmers understand how crops use nitrate efficiently until harvest. This is only assumed if no additional fertiliser treatment is applied between phenological growth stages. Meanwhile, the maximum crop height of oats plant by multispectral-UAV can be effectively monitored over a 11-month period. Also, oats height extracted by UAV data can be a potential indicator to understand if the crops have sufficiently utilised soil nitrate and phosphorus. Finally, a simplified effective method for estimating oat grain yield at the flowering stage using NDVI and CI green coupled with iso-cluster classification has shown positive correlation with actual yield. NDVI was a better predictor of oat grain yield estimation in comparison to CI green. Also, KDE shows to be an effective method to map estimated oat grain yield across the field. Further, it has been identified that oats yield can be estimated using chemical soil health indicators and crop quality features. This has enabled to provide concentration range of soil nutrients and crop quality features expected for very high yield across three phenological growth stages of oats crop.

Chapter 11: Overall conclusion and Future work

Firstly, the use of simple and effective image classification algorithms on multispectral UAV images have been applied in forestry for the identification of native and invasive tree species. The extraction of PCA features from the UAV multispectral images helps derive spectral vegetation indices for the identification of native and invasive tree species residing in an ASNW. Furthermore, the use of an unsupervised ML thresholding algorithm has proven to be an adequate pixel-based image classification approach for classifying a limited number of tree species. The unsupervised ML algorithm of k-means clustering has been an appropriate image classification approach to classify multiple invasive tree species.

Secondly, using the multispectral UAV imagery data the most appropriate spectral vegetation indices to distinguish a variety of crops were identified. NDYI was best suited for enhancing spectral features for the identification of oilseed rape, whereas NDVI was for winter beans and winter oats. Using these spectral vegetation indices, the unsupervised iso-cluster algorithm built into the ArcGIS software was demonstrated to be a viable image classification approach to classify the crop area pixels. The estimated yield of oilseed rape by NDYI-iso cluster approach showed a strong positive correlation ($R^2 = 0.86$) with the actual oilseed rape. Further, the estimated winter beans and winter oats yield by NDVI iso-cluster approach showed a strong positive correlation with actual winter bean ($R^2 = 0.84$) and winter oats ($R^2 = 0.74$) yield. Furthermore, crop height of oilseed rape, winter beans, and winter oats was successfully monitored over an 11-month growth period using multispectral UAV imagery data. In addition, a highly sensitive and accurate multinomial logistic regression ML model for predicting the concentration ranges of chemical soil health indicators and crop quality features at various winter oats crop yield levels (e.g., low, and high yield) have been developed.

Thirdly, additional methods for collecting ground reference data can be implemented to validate the UAV data and strengthen the research conducted in this PhD. Initially, reflectance reference panels can be used for radiometric calibration. This corrects the images reflectance according to the values given by the reference panels for each spectral band whilst considering the illumination condition of the UAV flight date, and location (Suomalainen et al., 2021). Also, data acquired over multiple flights can be compared accurately with this approach rather than solely using the irradiance sensor radiometric calibration method. In addition, UAV vegetation index values and height obtained for crop varieties must be validated with ground reference data. This would provide a better understanding on the accuracy of the UAV data in this PhD. Furthermore, a larger number of data (VI and chemical soil health indicators) must

be employed to strengthen the multinomial regression model used to estimate oats crop yield, since this would precisely account for variation observed across the field.

Also, for future studies, PCA derived spectral vegetation indices coupled with image segmentation methods can be employed to classify a wide range of other tree species located in woodlands. Furthermore, the iso-cluster classification approach can be examined to classify and estimate yield in a wide range of crops. Additionally, the multinomial regression model built in this study can be tested on other crop types to identify if crop yield can be estimated by soil nutrients and crop quality features. This would evaluate the model's replicability and provide concentration ranges of soil nutrients belonging to a class of yield for precision agriculture.

Chapter 12: Appendix

Appendix 1

Appendix 1.1

The steps below are performed using MATLAB and uses the function in the folder named as 'PCA on Multispectral images' in the supplementary materials.

To perform PCA on multispectral images

- 1) To run the function below select the Npcatif script, copy the function line, paste it in the command window, and press Enter. Ensure all the other function scripts specified in the folder of 'PCA Multispectral images' are present in the folder.
- 2) This will allow to select the multispectral images from the spectral band of blue, green, red, red edge and NIR to perform PCA.
- 3) The output will consist of principle components images, variance, eigenvectors labelled as 'p' contributing to each spectral band in the principal components.
- 4) The comment on function is written in bold.

```
function [ bil,d,p,score,l] = Npcatif(numberofimages)
%original indirectbil( numberofimages,fc,fr,lc,lr,scl )
%Reads, rangescales and registers n images, crops and rescales all
%identically and applies PCA.
%To ensure functionality, ensure your current folder contains your images.
%OUTPUT: [.bil multispectral image,data,pc,score]
%INPUT: (number of images to use,topleft x,topleft y,btm right x,btm right y,scale)
global nrows
bil=tifget3(numberofimages);
bil=registerbil(bil);
%bil=cropbil(bil,fr,lr,fc,lc,scl);
[bil]=pixelrange(bil);
[d,p,score,l,e]=makepcabil(bil);
scoreimages=reshape(score,nrows,[],numberofimages);
putvar(scoreimages);
explain=e;
latent=l;
putvar(nrows,latent,explain) end
```

To derive and apply the new spectral index by PCA on images.

The steps below are performed using MATLAB and uses the function in the folder named as 'Spectral index by PCA' in the supplementary materials.

- 1) To run the function as below select the spectralindex script, copy the function line, paste it in the command window, and press Enter. Ensure all the other function scripts specified in the folder of 'spectral index by PCA' are present in the folder.
- 2) The comment on function is written in bold.

function spectralindex

**%select folder, get path name as string variable and number of tif files in
%folder named as 'out' and hence the number of loops 'nummaps is out/5'%
%as well as this script you will also need the grs2rgb folder active on the
%same path as this matlab script. if it is greyed out in the lefthand
%directory you can right click on it and add selected folder to path.**

```
mypath = uigetdir;  
a=dir([mypath '/* .tif']);  
out=size(a,1);  
nummaps=out./5;  
nm=1:nummaps
```

**%make image datastore to address image files. NB this will include the RGB
%image as index 1,6,etc so address each one of the multispectral images in
%positions 2-5 within the for loop.**

```
ds=imageDatastore(mypath);
```

%create filenames for maps to be written to

```
name6 = mypath + "NEW1" + nm + ".tif"  
name7 = mypath + "NEW2" + nm + ".tif"  
name8 = mypath + "NEW3" + nm + ".tif"  
name9 = mypath + "NEW4" + nm + ".tif"
```

**%assigns the colourmap to overlay on the calculated images - hot = standard
%heatmap or you can switch to any other you may wish to create for the RGB
%intensities to write.**

map=hot(256)

%idx sets the start of the loop and indexes on each run through it

idx=0

**%this loop will cycle for the number of tif files in the selected folder,
%and extract multiples of position 2,3,4,5,6 for the selected multispectral
%images.**

for n = 1:nummaps

idx=idx+1

k(idx)= n*1

l(idx) = (6*k(idx))-6

**% Maths required for image processing drone files. Note files are saved as
% DJI_XXXY.TIF where Y=1 for B, 2 for G, 3 for R, 4 for DR, and 5 for NIR.
% These read in multiples from the image datastore and create the temporary
% variables B,G,R,DR and NIR that are used in the calculations**

B= readimage(ds,(l(idx)+2));

G= readimage(ds,(l(idx)+3));

R= readimage(ds,(l(idx)+4));

DR= readimage(ds,(l(idx)+5));

NIR= readimage(ds,(l(idx)+6));#

**% gets rid of the blurry effect from the houses (you should have register
% images code in your folder)**

[MOVINGREG]=registerImages(G,B);

G = MOVINGREG

```
[MOVINGREG]=registerImages(R,B);
R = MOVINGREG
[MOVINGREG]=registerImages(DR,B);
DR = MOVINGREG
[MOVINGREG]=registerImages(NIR,B);
NIR = MOVINGREG
```

```
G=G.RegisteredImage;
R=R.RegisteredImage;
DR=DR.RegisteredImage;
NIR=NIR.RegisteredImage;
```

**%register the bil images to get rid of the blurry effect (you should have
%registerbil and pixel range in your folder).**

```
bil=zeros(1300,1600,5);
bil(:,:,1)=B;
bil(:,:,2)=G;
bil(:,:,3)=R;
bil(:,:,4)=DR;
bil(:,:,5)=NIR;
bil=registerbil(bil);
[ bil ] = pixelrange( bil );
```

**%do the maths required to calculate the relevant indexes, note the images must be first
%converted to greyscale from the matrices read in the step above to achieve this
%read in the variable p from the PCA outputs (eigenvector value for each spectral band
in PC1, PC2, PC3, PC4 and PC5) and put it in an excel file to calculate PC images
%an example of excel file named as 'p.mat' is attached in the folder the rows represents
eigenvectors in each spectral band and the five column represents the PC1, PC2, PC3
PC4 and PC5**

```
a='p.mat.xlsx';
p=readmatrix(a);
```

```

PC1 = mat2gray(((p(1,1)*double(B))) + (p(1,2)*double(G)) + (p(1,3)*double(R)) +
(p(1,4)*double(DR)) + (p(1,5)*double(NIR)));
PC2 = mat2gray(((p(2,1)*double(B))) + (p(2,2)*double(G)) + (p(2,3)*double(R)) +
(p(2,4)*double(DR)) + (p(2,5)*double(NIR)));
PC3 = mat2gray(((p(3,1)*double(B))) + (p(3,2)*double(G)) + (p(3,3)*double(R)) +
(p(3,4)*double(DR)) + (p(3,5)*double(NIR)));
PC4 = mat2gray(((p(4,1)*double(B))) + (p(4,2)*double(G)) + (p(4,3)*double(R)) +
(p(4,4)*double(DR)) + (p(4,5)*double(NIR)));
PC5 = mat2gray(((p(5,1)*double(B))) + (p(5,2)*double(G)) + (p(5,3)*double(R)) +
(p(5,4)*double(DR)) + (p(5,5)*double(NIR)));

```

```

NEW1= mat2gray(PC1-PC2)./(PC1+PC2);

```

**%add heatmap this applies the selected colourmap to the calculated images,
 %giving the heatmap version as an output that is more familiar when using
 %these indexes**

```

hNEW1=grs2rgb(NEW1,map);

```

% write calculated images to file appended with loop number to filename,
 %this will write the files as TIF format to the parent folder of the one
 %selected. Note that if you run the DJI matlab script a second time on a
 %folder it will append the images to anything bearing the same filename.
 %So it's a good idea to keep your folder names varied with date and site
 %identifiers.

```

imwrite(hNEW1,name6(k(idx)),'tif');

```

```

end

```


To perform thresholding and k means segmentation.

For image thresholding first read the image on MATLAB, convert it to grayscale and specify a threshold to get the pixels for the specified threshold

Example

```
Im = a; (read image)  
aa= mat2gray(Im) (convert image to grayscale)  
aa2=nnz(aa<0.45) (get the number of pixel)
```

For k-means segmentation was performed used MATLAB and uses the function in the folder named as 'k-means segmentation' in the supplementary materials

- 1) To run the function as below select the seg script, copy the function line, paste it in the command window, and press Enter.
- 2) The comment on function is written in bold.

function seg

```
% This example demonstrates intensit based image segmentation using
```

```
% kmeans
```

```
% Read in image
```

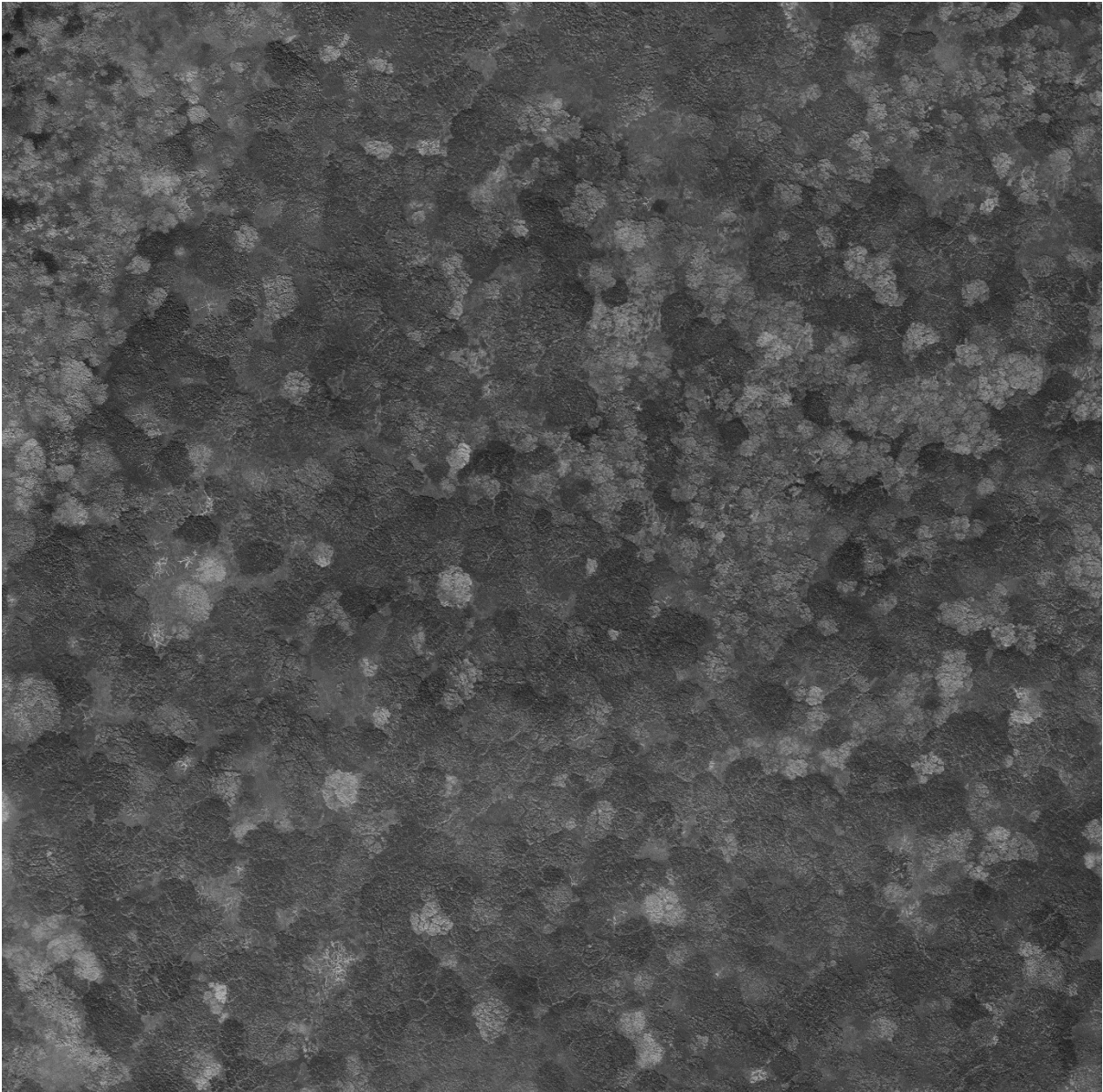
```
Im = a;  
subplot(1,2,1)  
imshow(Im)  
title('Original Image');
```

```
specify the number clusters in this case its 3
```

```
[L,Centers] = imsegkmeans(Im,3);  
B = labeloverlay(Im,L);  
subplot(1,2,2)  
imshow(B)  
title('Labeled Image');
```

```
end
```

Appendix 1.2 Priestclose wood threshold image of oak and silver birch trees (light regions are silver birch and dark regions are oak).



Appendix 2

2.1 Pseudo corrected figure

Figure A2.1 Pseudo corrected raster NIR band. The NIR spectral band was divided by the maximum DN value of 65535 to convert reflectivity (a) to reflectance by performing to normalise the data and obtain values between 0 to 1.

(a)



(b)



Appendix 2.2 Analysis of NPK and pH by Palintest soil kit, organic matter and beta-glucan

The Soiltest 10 Bluetooth photometer in Palintest kit was used to analyse the chemical soil health indicators of NPK. At the beginning of analysis, the photometer was calibrated using certified Palintest check standards. The Soiltest 10 Bluetooth photometer analyses soil nutrients accurately and reliably. It employs the photometric testing principle, which is based on the absorption or scattering of the measured intensity of the incident light in comparison to the light intensity reaching the detector (Palintest, 2020). The light intensity is calculated as a percentage of transmittance (%T) or absorbance (A) and compared to calibration tables stored in the Soiltest 10 Bluetooth photometer. The calibration table that is saved converts the %T or A to results in a variety of units (mg/l or ppm) (Palintest, 2020). Further, Palintest soil analysis methods can be carried out on fresh or dried samples. In this study the collected 48 samples were tested for soil NPK and pH within one week after sample collection. However, nitrogen was analysed first among soil NPK, as it is recommended to analyse nitrogen on damp soil samples to prevent nitrogen loss due to biological activity.

Soil nitrogen analysis

Firstly, to a sample container filled with 50ml of deionised water, 2.5 ml of 1 M ammonium chloride powder (Extract N) was added and shaken to dissolve the contents. Afterwards, 2ml soil sample was added to the container and shaken for 1 minute to extract the nitrate and nitrate from soil samples. Then, one level spoonful of Nitratest™ powder was added to the container and shaken for 1 minute to reduce nitrate to nitrate. Afterwards, a filter paper was folded into quarters and inserted into a funnel and placed in a clean container where the extract solution was poured to the funnel to collect the extraction filtrate. 10ml filtrate was then transferred to a clean cuvette and used to blank the Soiltest 10 photometer. After blanking the photometer, one Nitricol tablet was added to the same cuvette and crushed to dissolve, and the cuvette was left at room temperature for 10 minutes to allow full colour development. Nitricol helps to produce a range of colours from pale pink to deep violet depending on the nitrogen concentration in soil sample. After 10 minutes the cuvette was inserted into the photometer and the result was recorded as nitrate in mg/l. The Soiltest 10 Bluetooth photometer can record soil nitrate in the range 0-25 mg/l.

Soil phosphorus analysis

Firstly, to a sample container filled with 50ml of deionised water, 5 tablets of 0.5 M sodium bicarbonate (Extract P) was added and shaken gently to dissolve the contents. Afterwards, 2ml soil sample was added to the container and shaken for 1 minute to extract the phosphorus from soil samples. Then, a filter paper was folded into quarters and inserted into a funnel and placed in a clean container where the extract solution was poured to the funnel to collect the extraction filtrate. 2ml filtrate was then transferred to a clean cuvette and filled until 10ml using deionised water and one Acidifying S tablet was crushed and mixed gently to dissolve the contents. This solution was used to blank the photometer and afterwards one Phosphate P tablet was crushed and mixed to dissolve. Then the cuvette was left at room temperature for 10 minutes to allow full colour development. Phosphate P tablet helps to form a bright blue complex solution depending on the phosphorus concentration in soil sample. After 10 minutes the cuvette was inserted into the photometer and the result was recorded as phosphorus in mg/l. The Soiltest 10 Bluetooth photometer can record soil phosphorus in the range 0-150 mg/l.

Soil Potassium analysis

Firstly, to a sample container filled with 50ml of deionised water, 2.5 ml of 0.1 M magnesium acetate (Extract K) was added and shaken gently to dissolve the contents. Afterwards, 2ml soil sample was added to the container and shaken for 1 minute to extract the potassium from soil samples. Then, a filter paper was folded into quarters and inserted into a funnel and placed in a clean container where the extract solution was poured to the funnel to collect the extraction filtrate. 10ml filtrate was then transferred to a clean cuvette and used to blank the Soiltest 10 photometer. After blanking the photometer, one Potassium K tablet was added to the same cuvette and crushed to dissolve, and the cuvette was left at room temperature for 2 minutes to allow full colour development. The Potassium K tablet consists of sodium tetraphenylboron to form an insoluble white complex depending on the potassium concentration in soil sample. After 2 minutes the cuvette was inserted into the photometer and the result was recorded as potassium in mg/l. The Soiltest 10 Bluetooth photometer can record soil potassium in the range 0-450 mg/l.

Soil pH analysis using the multiparameter pocket sensor.

Prior to analysis the pH probe was calibrated using the pH 4, 7 and 10 buffers provided by the Palintest kit. Afterwards, two levels of 2 ml soil sample and 10ml deionised water was added to a sample container and shaken gently for 1 minute. Then, 10ml sample was transferred to the multiparameter pocket sensor cap. The pH measurement mode was selected on the multiparameter pocket sensor and then inserted into the cap allowing the reading to stabilise and the soil pH was recorded.

Analysis of soil organic matter using ‘loss on ignition’ method

Firstly, 5 g of soil sample in a crucible was oven dried overnight at 105°C and weighed afterwards. The samples were then re-heated to 800°C and weighed afterwards. This process is known as ‘loss on ignition, essentially to burn the organic matter in the soil. The percentage organic matter was calculated as below.

$$\% \text{ soil organic matter} = \left(\frac{\text{Weight } 105^{\circ}\text{C} - \text{Weight } 800^{\circ}\text{C}}{\text{Weight } 105^{\circ}\text{C}} \right) * 100$$

Oat beta-glucan extraction

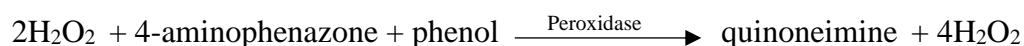
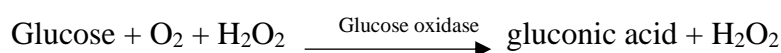
Initially, 0.2 g of weighed oats separated from the husk was crushed using a mortar and pestle to make a fine powder. Oat beta-glucan for 48 samples was extracted according to Bzducha-Wróbel at al., 2012) with few modifications. The beta-glucan from the samples were extracted in three stages. Stage 1 was alkali extraction where the oats samples were dissolved in 15 ml of 3% NaOH in a water bath at 75°C for 1 hour. The samples were then centrifuged at 8,000 rpm for 30 minutes. The alkali extraction was performed three times for each sample to obtain supernatant with beta-glucan. Hence, in stage 2 acid neutralization was performed where the supernatant obtained from stage 1 was mixed with 20 ml of 0.1 M Tris-HCl (pH = 7.4) and centrifuged at 8,000 rpm for 30 minutes. Further, 20 ml of 0.01 M Tris-HCl (pH = 7.4) was added and centrifuged at 8,000 rpm for 30 minutes. The pellet obtained after acid neutralisation in stage 2 consists of oat beta-glucan extracted in gel deposit form. Hence, in stage 3 the excess solvents in the beta-glucan gel deposit were removed by freeze drying where the samples were left in a freeze dryer overnight.

After extraction of beta-glucan acid hydrolysis can be performed to determine the total glucose released by breaking the bonds in beta-glucan polysaccharide chains which would aid in

quantifying the beta-glucan contents in oats sample. Hence, extracted beta -glucan samples were acid hydrolysed in the final stage 4, where 2 ml of de-ionised water was added to the extracted beta-glucan sample and vortexed. Acid hydrolysis was performed by converting beta-glucan to glucose monomers where to 66 µL of beta-glucan sample a 33 µL of trifluoroacetic acid was added and heated on a heating block for 2 h at 90 °C. And for the control, 66 µL of beta-glucan sample and 33 µL of de-ionised water was added and heated for 2 h at 90 °C. The control was used to check for contamination of samples. The acid hydrolysed glucose samples were stored in a freezer at -4 °C until further analysis.

Analysis of glucose in oat beta glucan by colorimetric assay

Analysis of glucose was performed by using a glucose colorimetric assay kit named as Randox (Randox Laboratories Ltd, County Antrim, UK). The kit employs GOD-PAP reagent and a buffer that performs enzymatic oxidation in the presence of glucose oxidase to determine glucose. Under the catalysis of peroxidase, the formed hydrogen peroxide reacts with phenol and 4-aminophenazone to form a red - violet quinoneimine dye as an indicator. The absorbance of the sample can be measured using a spectrophotometer (Randox, 2023)



Hence, 10 µL of sample and 1000 µL of reagent (GOD-PAP reagent + Buffer) was mixed in a cuvette and incubated for 25 min at room temperature. As a standard, a glucose standard solution from the Randox kit was used, and deionised water was used as a reagent blank. Afterwards, the absorbance of the sample, standard and the reagent blank was measured at 500 nm using a spectrophotometer (Jenway 7300 and 7305 spectrometer).

Appendix 2.3 Manually measured oats data.

Table 2.3: Oats data to estimate oats grain yield.

Sample	Recovered oats + husk from sampled 1m ² (g)	Average oats + husk (20 g)	± SD (n = 3)	Average oats (20 g)	± SD (n = 3)	Calculated oats 1 m ² (g)	Unit conversion for oats 1 m ² (t)	Final Oats yield in 1m ² (t/ha)
A-1	896.4	20.6272	0.0948	18.9745	0.2653	824.6	0.0008246	8.25
A-2	939.4	20.5003	0.3909	19.4250	0.2287	890.1	0.0008901	8.90
A-3	703.1	20.7281	0.2285	18.8708	0.2900	640.1	0.0006401	6.40
A-4	630.6	20.6553	0.2455	19.3659	0.2080	591.2	0.0005912	5.91
A-5	971.1	20.3844	0.4785	18.7863	0.5741	8954.0	0.0008950	8.95
A-6	564.9	20.7074	0.2475	19.8530	0.1209	541.6	0.0005416	5.42
A-7	527.2	20.6318	0.2494	16.3601	5.2574	418.0	0.0004180	4.18
A-8	201.6	20.7656	0.1505	19.6161	0.2395	190.4	0.0001904	1.90
B-1	800.6	20.1838	0.1120	19.3314	0.1387	766.8	0.0007668	7.67
B-2	802.4	20.4741	0.3359	19.0529	0.0275	746.7	0.0007467	7.47
B-3	843.1	20.4068	0.0508	19.1494	0.1256	791.2	0.0007912	7.91
B-4	781.7	20.6551	0.3221	19.3389	0.3419	731.9	0.0007319	7.32
B-5	810.2	20.4540	0.0976	19.6350	0.2051	777.8	0.0007778	7.78
B-6	788.0	20.6843	0.2568	19.4324	0.3017	740.3	0.0007403	7.40
B-7	664.2	20.5087	0.0773	19.3786	0.0886	627.6	0.0006276	6.28
B-8	346.8	20.3913	0.2377	19.4864	0.3055	331.4	0.0003314	3.31
C-1	1016.3	20.3864	0.2306	18.6343	0.2807	929.0	0.0009290	9.29
C-2	724.8	20.3786	0.3946	19.1671	0.4289	681.7	0.0006817	6.82
C-3	863.5	20.5385	0.1612	19.3801	0.2140	814.8	0.0008148	8.15
C-4	891.3	20.5954	0.2449	18.9850	0.5568	821.6	0.0008216	8.22
C-5	843.1	20.5335	0.2343	19.3484	0.2024	794.4	0.0007944	7.94
C-6	728.4	20.4974	0.3359	19.4950	0.3146	692.8	0.0006928	6.93
C-7	597.6	20.5716	0.1293	19.5135	0.2610	566.9	0.0005669	5.67
C-8	517.9	20.5719	0.0557	19.1451	0.5683	482.0	0.0004820	4.82
D-1	885.0	20.7355	0.2281	19.4518	0.1913	830.2	0.0008302	8.30
D-2	961.3	20.6760	0.3833	19.4563	0.5084	904.6	0.0009046	9.05
D-3	875.4	20.6096	0.2941	19.3750	0.4638	823.0	0.000823	8.23
D-4	929.0	20.7926	0.0870	19.4672	0.0425	869.8	0.0008698	8.70
D-5	955.9	20.4049	0.4865	19.1066	0.5327	895.1	0.0008951	8.95

D-6	284.9	20.3851	0.3311	19.2472	0.7382	269.0	0.0002690	2.69
D-7	339.6	20.4036	0.2748	19.3010	0.6562	321.2	0.0003212	3.21
D-8	370.8	20.5214	0.2770	19.3437	0.2983	349.5	0.0003495	3.50
E-1	658.3	20.2823	0.1332	18.8869	0.1570	613.0	0.0006130	6.13
E-2	996.4	20.4655	0.1798	19.1481	0.0491	932.3	0.0009323	9.32
E-3	748.1	20.5399	0.0204	19.2648	0.0901	701.7	0.0007017	7.02
E-4	1034.3	20.1673	0.1497	18.8590	0.3007	967.2	0.0009672	9.67
E-5	881.5	20.4418	0.3609	18.8695	0.3651	813.7	0.0008137	8.14
E-6	935.6	20.3488	0.2582	19.0134	0.2760	874.2	0.0008742	8.74
E-7	767.5	20.3836	0.1661	18.8579	0.1733	710.1	0.0007101	7.10
E-8	1219.4	20.3724	0.1201	19.3083	0.0960	1155.7	0.0011557	11.56
F-1	622.3	20.7099	0.2865	19.5649	0.2903	587.9	0.0005879	5.88
F-2	765.8	20.4470	0.2105	16.8513	4.0714	631.1	0.0006311	6.31
F-3	718.8	20.5860	0.2804	18.9438	0.1124	661.5	0.0006615	6.61
F-4	970.3	20.5449	0.3873	18.9623	0.3907	895.6	0.0008956	8.96
F-5	991.1	20.6072	0.2988	19.2725	0.4014	926.9	0.0009269	9.27
F-6	965.5	20.5251	0.3233	19.0468	0.2350	896.0	0.0008960	8.96
F-7	1007.8	20.6408	0.3678	19.2869	0.2164	941.7	0.0009417	9.42
F-8	284.9	20.6081	0.3317	19.2244	0.6082	265.8	0.0002658	2.66

Appendix 2.4

Appendix 2.4 Investigation of β -glucan analysis (a) Extraction efficiency of β -glucan in different extraction stages, (b) Conversion efficiency of β -glucan to glucose by acid hydrolysis, and (c) Analytical performance parameters for analysing glucose in oat beta-glucan by colorimetric assay and (d) Glucose calibration curve by Randox kit colorimetric assay

(a)

β-glucan extraction in different stages	Extraction efficiency (%) (N = 6), mean \pm SD (individual values)
Base level - no spike (mM)	-
Stage 1 - Before alkali extraction	84.4* (85.2, 83.5)
Stage 2 - Before acid neutralization	90 \pm 5 (97.4, 85.1, 90.7, 90.2, 92.5, 83.9)
Stage 3 - Before freeze drying	95 \pm 6 (97.3, 97.9, 89.4, 88, 105.2, 92.4)
Stage 4 - Before acid hydrolysis	105 \pm 12 (114.3, 113, 96.6, 89.7, 98.5, 119.7)

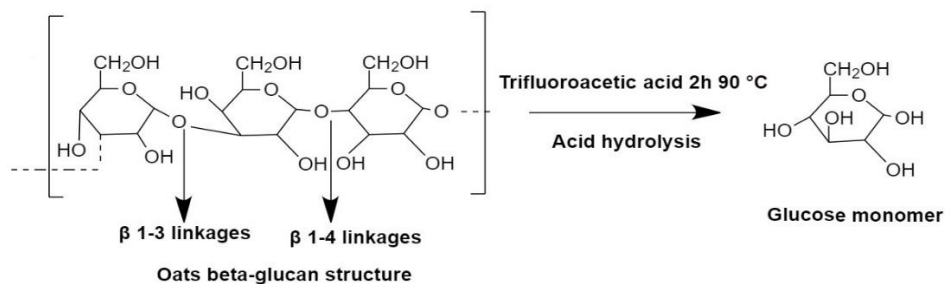
*N = 2

Note: The extraction efficiency was determined by spiking 15 mM of glucose to the oat samples prior to the four stages of the extraction method. This provides information on the amount of β -glucan extracted from oats at each stage, indicating the efficiency of the extraction method. Stage 1 - Alkali extraction by 3% NaOH, 75 °C (1 h x 3) to obtain supernatant fraction consisting of β -glucan. Stage 2 - Acid neutralisation by 0.1 M and 0.01 M Tris-HCl to obtain gel deposit of β -glucan. Stage 3 – Freeze drying to remove excess solvents in the gel deposit of β -glucan. Stage 4 - acid hydrolysis converts β -glucan to glucose monomers.

(b)

β-glucan to glucose by acid hydrolysis	Conversion efficiency (%) n = 6, mean \pm SD (individual values)
Stage 4 - Acid hydrolysis	102 \pm 8 (114.4, 98.6, 103.6, 95.7, 108.2, 93.0)

Note: The conversion efficiency of beta-glucan was determined by spiking 15 mM of β -glucan to the samples prior to acid hydrolysis. This allows to determine if the β -glucan linkages in the oats sample are efficiently broken to glucose monomers by acid hydrolysis method.



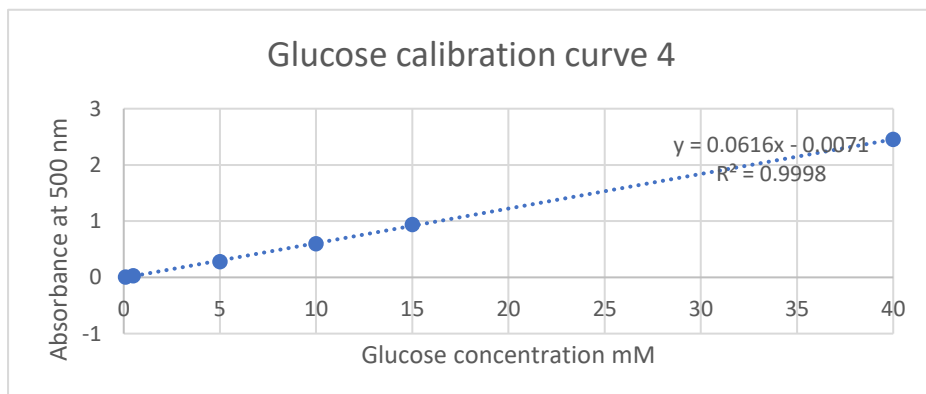
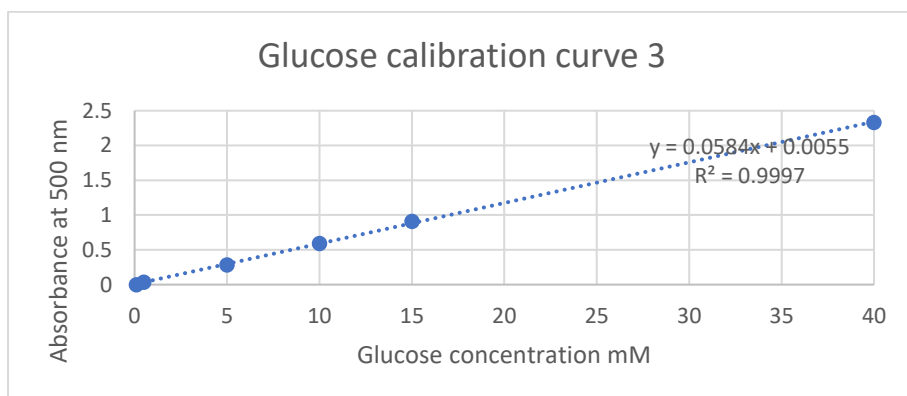
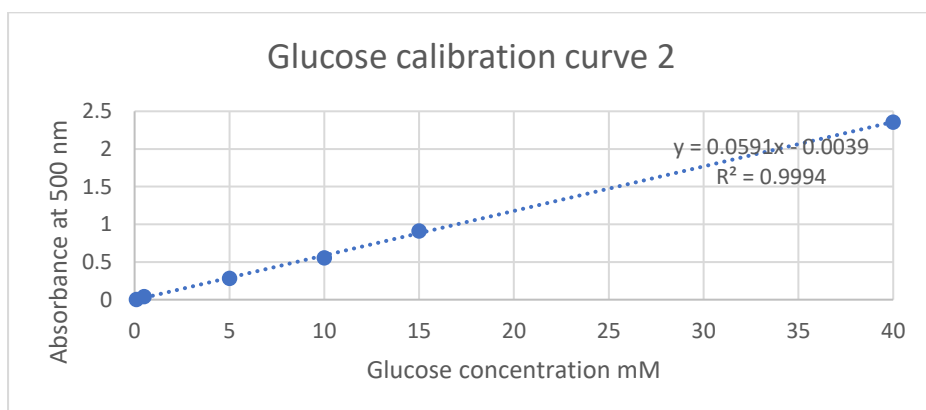
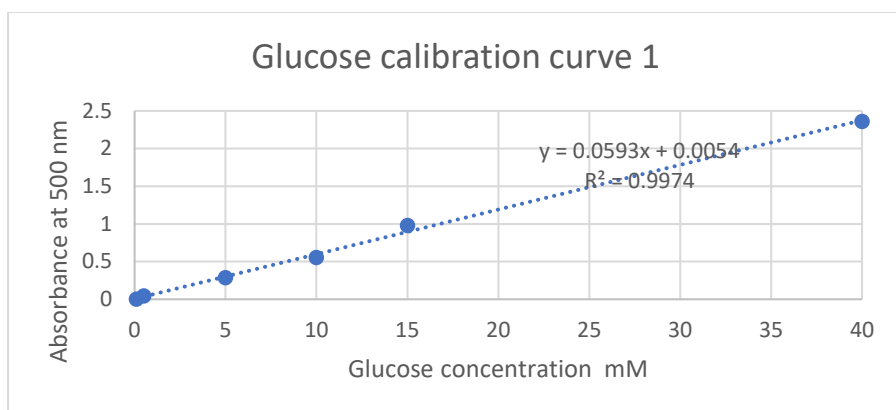
The limit of detection (LOD), limit of quantification (LOQ) and % RSDs, by the colorimetric Randox assay for analysing glucose were calculated based on the information of six standard calibration glucose curves (d). The LOD and LOQ for determine glucose was 33.6 mg/L and 101.9 mg/L respectively (c). The % RSDs for low concentration was 9.5 and high concentration was 1.9 (c). The glucose content measured in samples did not fall below the LOD and LOQ (c).

(c)

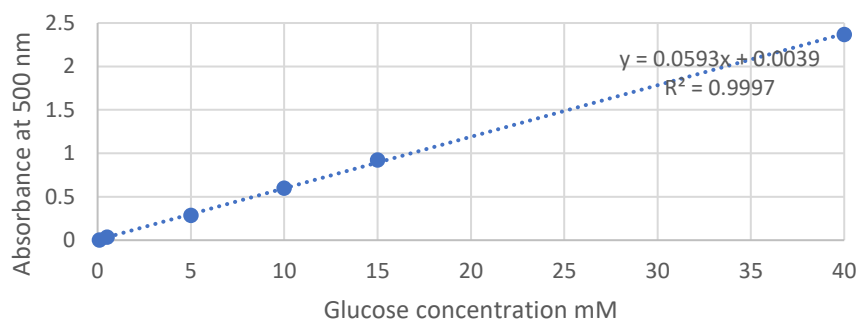
Analyte	N	Linear regression ($Y = mx + c$)	Correlation coefficient (R^2)	Limit of detection (LOD) mg/L	Limit of quantification (LOQ) mg/L	Precision (RSDs %) low concentration	Precision (RSDs %) high concentration
Glucose	6	$Y = 0.05973x - 0.00645$	0.9992	33.6	101.9	9.5	1.9
β -glucan	6			30.3	91.7		

Note: The free glucose ($M_r = 180$) determined was adjusted by multiplying by a conversion factor of 0.9, to account for the difference in molecular weight of glucose monomers bound in a polysaccharide of β -glucan ($M_r = 162$).

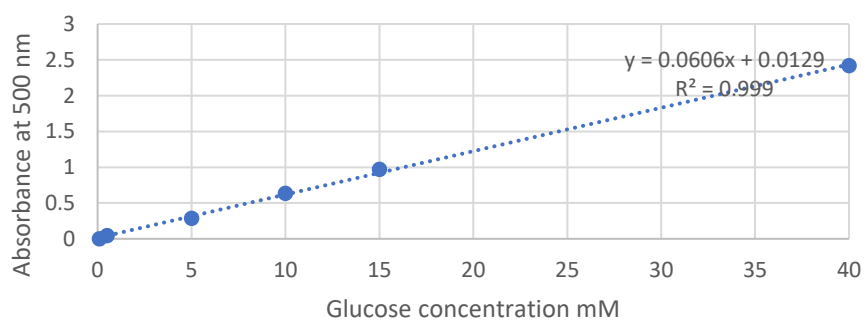
(d)



Glucose calibration curve 5



Glucose calibration curve 6



Appendix 3

Raw data of Comparison of spectral profile between UAV and ground reference data (Chapter 10)

Table 3.1: Ground reference data using the Spectro-1 spectrometer to generate a spectral profile for sample location A4 sample over 3 months (as displayed in Figure 51 in Chapter 10 of thesis).

Wavelength (nm)	June 2021		July 2021		August 2021	
	Average (n = 6)	SD	Average (n = 2)	SD	Average (n = 2)	SD
400	0.2993	0.0794	0.0653	0.0074	0.0614	0.0206
410	0.3344	0.1216	0.0333	0.0059	0.0538	0.0136
420	0.3337	0.1186	0.0322	0.0089	0.0532	0.0097
430	0.2978	0.1082	0.0368	0.0082	0.0562	0.0072
440	0.2708	0.1033	0.0479	0.0071	0.0636	0.0078
450	0.1904	0.0740	0.0562	0.0090	0.0706	0.0105
460	0.1589	0.0573	0.0580	0.0103	0.0746	0.0135
470	0.1470	0.0453	0.0641	0.0122	0.0785	0.0221
480	0.1373	0.0321	0.0669	0.0140	0.0825	0.0328
490	0.1531	0.0300	0.0767	0.0153	0.0869	0.0421
500	0.1319	0.0235	0.0898	0.0149	0.0974	0.0489
510	0.0950	0.0164	0.0980	0.0139	0.1046	0.0446
520	0.0915	0.0124	0.1031	0.0144	0.1102	0.0356
530	0.1289	0.0205	0.1137	0.0184	0.1154	0.0196
540	0.1670	0.0292	0.1280	0.0261	0.1225	0.0039
550	0.1795	0.0256	0.1375	0.0346	0.1295	0.0022
560	0.1607	0.0210	0.1401	0.0396	0.1316	0.0055
570	0.1398	0.0250	0.1482	0.0461	0.1348	0.0045
580	0.1365	0.0278	0.1618	0.0524	0.1481	0.0260
590	0.1379	0.0285	0.1584	0.0552	0.1615	0.0367
600	0.1251	0.0190	0.1338	0.0504	0.1629	0.0253
610	0.0998	0.0117	0.1097	0.0452	0.1560	0.0081
620	0.0997	0.0119	0.1001	0.0404	0.1517	0.0041
630	0.0850	0.0168	0.0975	0.0392	0.1550	0.0097
640	0.0647	0.0109	0.0973	0.0438	0.1666	0.0005
650	0.0623	0.0126	0.1047	0.0503	0.1849	0.0230
660	0.1266	0.0196	0.1527	0.0641	0.2119	0.0660
670	0.1637	0.0254	0.1740	0.0700	0.2236	0.0836
680	0.1758	0.0254	0.1710	0.0714	0.2167	0.0712
690	0.1969	0.0256	0.1523	0.0683	0.2076	0.0548
700	0.1844	0.0245	0.1118	0.0566	0.1898	0.0210

Table 3.2: Multispectral UAV data to generate a spectral profile for sample location A4 sample over 3 months (as displayed in Figure 51 in Chapter 10 of thesis).

Wavelength (nm)	June 2021		July 2021		August 2021	
	Average (n = 5)	SD	Average (n = 5)	SD	Average (n = 5)	SD
450	0.3000	0.0371	0.5140	0.0439	0.3116	0.0420
560	0.3677	0.0394	0.4461	0.0083	0.2265	0.0326
650	0.1764	0.0192	0.4444	0.0157	0.3329	0.0534
760	0.3674	0.0528	0.5495	0.0275	0.2825	0.0191
840	0.5958	0.0455	0.5895	0.0325	0.2489	0.0332

Table 3.3: Ground reference GRVI data from handheld Spectro-1 spectrometer for samples in June 21, July 21 and August 21 (Table 8 of Chapter 10 in thesis)

Sample	June 2021		July 2021		August 2021	
	Average (n = 6)	SD	Average (n = 3)	SD	Average (n = 3)	SD
A2	0.4567	0.0271	0.2070	0.0722	-0.0714	0.2285
A4	0.4602	0.0397	0.2665	0.0227	-0.1707	0.0578
A6	0.5262	0.0315	0.1637	0.1038*	-0.1356	0.0904
B6	0.5210	0.0716	0.3189	0.1274	-0.1064	0.1274
C1	0.4497	0.0604	0.3472	0.1601	-0.0941	0.1322
C3	0.5235	0.0659	0.2893	0.0805*	-0.1507	0.0601
C6	0.4501	0.1038	0.2621	0.0792	-0.0011	0.1283
D3	0.3311	0.0519	0.2789	0.0839*	-0.1196	0.1015
D5	0.5142	0.0159	0.1836	0.0393*	-0.1944	0.0067

Table 3.4: Vegetation indices data from multispectral-UAV for samples in (a) June 2021, (b) July 2021 and (c) August 2021 (as displayed in table 8 of chapter 10 in thesis)

(a)

	A2		A4		A6		B6	
VI	Average (n = 5)	SD	Average (n = 5)	SD	Average (n = 5)	SD	Average (n = 5)	SD
GRVI	0.3561	0.0592	0.3512	0.0400	0.3300	0.0512	0.3957	0.0765
NDVI	0.7385	0.0080	0.6673	0.0113	0.6295	0.0192	0.6296	0.0182
GNDVI	0.4674	0.0348	0.4855	0.0155	0.5983	0.0462	0.4821	0.0421
SAVI	0.5077	0.0231	0.6545	0.0148	0.6544	0.0252	0.4984	0.0141
NDRE	0.4706	0.0218	0.4457	0.0206	0.4926	0.0097	0.4303	0.0202
CI green	3.5470	0.1341	1.8112	0.1918	2.7459	0.2784	2.7857	0.2063

C1		C3		C6		D3		D5	
Average (n = 5)	SD	Average (n = 5)	SD	Average (n = 5)	SD	Average (n = 5)	SD	Average (n = 5)	SD
0.35489	0.0279	0.2486	0.0566	0.2873	0.0429	0.4509	0.0614	0.3300	0.0311
0.81385	0.0145	0.7139	0.0067	0.6668	0.0331	0.6722	0.0248	0.7257	0.0232
0.6409	0.0275	0.4901	0.0402	0.4803	0.0448	0.5448	0.0228	0.6511	0.0191
0.5743	0.0302	0.5556	0.0221	0.5644	0.0283	0.5358	0.0301	0.5872	0.0324
0.4149	0.0132	0.4081	0.0120	0.4791	0.0221	0.3941	0.0185	0.4196	0.0083
4.6234	0.1805	3.0511	0.0616	1.8456	0.2331	2.5926	0.1875	2.1216	0.0629

(b)

	A2		A4		A6		B6	
VI	Average (n = 5)	SD	Average (n = 5)	SD	Average (n = 5)	SD	Average (n = 5)	SD
GRVI	-0.0721	0.0145	0.0075	0.0355	0.0021	0.0162	-0.0164	0.0211
NDVI	0.3693	0.0135	0.4209	0.0173	0.5698	0.0211	0.5469	0.0134
GNDVI	0.2553	0.0157	0.3694	0.0150	0.3822	0.0300	0.4431	0.0240
SAVI	0.3027	0.0203	0.3367	0.0098	0.3872	0.0152	0.4378	0.0140
NDRE	0.4029	0.0014	0.3646	0.0092	0.3518	0.0066	0.4482	0.0248
CI green	0.6909	0.0757	0.9730	0.0395	1.5022	0.0727	1.1627	0.0699

C1		C3		C6		D3		D5	
Average (n = 5)	SD	Average (n = 5)	SD	Average (n = 5)	SD	Average (n = 5)	SD	Average (n = 5)	SD
0.03389	0.0639	0.1140	0.0686	-0.0615	0.0154	-0.0103	0.0559	-0.0207	0.0438
0.2430	0.0284	0.3870	0.0105	0.3201	0.0152	0.3021	0.0045	0.1836	0.0050

0.1457	0.011 2	0.2908	0.0033	0.2898	0.023 6	0.2346	0.031 6	0.0868	0.008 4
0.1525	0.010 0	0.1071	0.0064	0.2065	0.005 7	0.2210	0.008 6	0.0841	0.009 2
0.1708	0.004 0	0.3690	0.0096	0.3149	0.007 3	0.2795	0.018 1	0.1701	0.009 9
0.3541	0.023 6	0.4952	0.0357	0.6959	0.053 8	0.3299	0.014 6	0.2741	0.049 4

(c)

	A2		A4		A6		B6	
VI	Average (n = 5)	SD	Average (n = 5)	SD	Average (n = 5)	SD	Average (n = 5)	SD
GRVI	-0.0796	0.0626	-0.1890	0.0347	-0.2793	0.0541	-0.3540	0.0815
NDVI	0.2607	0.0300	0.3299	0.0216	0.3902	0.0167	0.2994	0.0474
GNDVI	0.1311	0.0114	0.1336	0.0186	0.2032	0.0087	0.2635	0.0095
SAVI	0.2696	0.0097	0.2640	0.0097	0.3207	0.0128	0.2441	0.0166
NDRE	0.0886	0.0063	0.1502	0.0144	0.1613	0.0015	0.2205	0.0181
CI green	0.4235	0.0214	0.5618	0.0214	0.7405	0.0232	0.8685	0.0477

C1		C3		C6		D3		D5	
Average (n = 5)	SD	Average (n = 5)	SD	Average (n = 5)	SD	Average (n = 5)	SD	Average (n = 5)	SD
-0.1833	0.1043	-0.0806	0.0990	-0.0915	0.050 1	-0.1002	0.059 7	-0.1575	0.074 9
0.1516	0.0236	0.1563	0.0202	0.2620	0.021 9	0.2360	0.006 7	0.1544	0.016 4
0.0887	0.0206	0.1775	0.0139	0.2635	0.009 5	0.2175	0.019 0	0.1537	0.010 7
0.1573	0.0158	0.1693	0.0061	0.2441	0.016 6	0.1443	0.008 7	0.1599	0.007 5
0.1092	0.0023	0.2002	0.0071	0.1695	0.004 3	0.2213	0.007 1	0.1304	0.007 1
0.3131	0.0165	0.4812	0.0156	0.6368	0.029 8	0.7521	0.024 8	0.3456	0.024 6

Appendix 4

Table 4.1 Palintest chemical soil health indicator data for June, July, and August 2021 across phenological growth stages in different zones (For Figure 63- 65 in chapter 10)

ZONE 1												
June 2021				July 2021					August 2021			
Sample	NO3- (mg/l)	P (mg/l)	K (mg/l)	pH	NO3- (mg/l)	P (mg/l)	K (mg/l)	pH	NO3- (mg/l)	P (mg/l)	K (mg/l)	pH
C2	58.9	25	150	7.64	12	16	215	7.43	2.7	12	210	7.73
D1	47.8	31	100	7.38	40.3	16	100	7.51	4.1	19	145	7.79
D2	23	33	132	7.61	8.4	20	105	7.63	7.9	19	185	7.77
D3	19.5	33	95	7.46	18.6	26	170	7.48	13	20	190	7.76
E1	14.2	28	95	7.76	22.6	19	215	7.81	21.3	20	225	7.69
E2	22.2	20	90	7.77	6.6	20	135	7.73	4.9	14	125	7.72
E3	54.4	24	115	7.64	2.7	13	110	7.83	18.2	14	155	7.72
F1	48.3	25	110	7.11	40.3	36	240	7.48	19	23	230	7.73
F2	29.2	25	110	7.52	15.1	24	130	7.57	23	21	255	7.64
F3	5.8	29	180	7.42	38.5	21	215	7.58	10.6	20	125	7.82
F4	6.6	35	120	7.79	36.3	23	235	7.69	4.4	27	225	7.83
Mean	29.99	28.00	117.91	7.55	21.95	21.27	170.00	7.61	11.74	19.00	188.18	7.75
Min	5.8	20	90	7.11	2.7	13	100	7.43	2.7	12	125	7.64
Max	58.9	35	180	7.79	40.3	36	240	7.83	23	27	255	7.83
ZONE 2												
June 2021				July 2021					August 2021			
Sample	NO3- (mg/l)	P (mg/l)	K (mg/l)	pH	NO3- (mg/l)	P (mg/l)	K (mg/l)	pH	NO3- (mg/l)	P (mg/l)	K (mg/l)	pH
A1	44.2	21	100	7.73	37.7	27	235	7.59	17.3	16	190	7.8
A2	50.45	23	105	7.74	9.7	13	205	7.57	0.9	10	150	7.79

B1	13.2	18	80	7.6	7.1	17	165	7.56	3.1	19	95	7.64
B2	27	20	100	7.63	14.2	17	200	7.56	8.9	14	125	7.75
B3	69.6	22	105	7.38	37.2	12	245	7.43	1.8	13	120	7.63
B4	23.4	20	100	7.65	33.2	11	175	7.69	4.4	16	185	7.85
C1	80.6	22	165	7.52	16.8	13	205	7.31	3.5	7	160	7.65
C3	34.1	20	100	7.04	7.5	12	125	7.21	33.2	16	130	7.77
C4	18.6	20	90	7.2	34.1	10	215	7.28	7.5	18	135	7.84
D4	20.4	27	90	7.7	14.6	11	110	7.49	38.5	15	200	7.82
D5	22.2	22	105	7.57	6.2	12	180	7.42	13.7	17	130	7.83
E4	1.3	22	100	7.67	12.8	13	120	7.78	8.8	9	110	7.72
E5	40.8	38	120	7.48	1.8	13	105	7.45	35.4	13	120	7.72
F5	8.9	29	105	7.75	12.8	20	135	7.68	15.4	25	215	7.8
Mean	32.48	23.14	104.64	7.55	17.55	14.36	172.86	7.50	13.74	14.86	147.50	7.76
Min	1.3	18	80	7.04	1.8	10	105	7.21	0.9	7	95	7.63
Max	80.6	38	165	7.75	37.7	27	245	7.78	38.5	25	215	7.85
ZONE 3												
June 2021					July 2021				August 2021			
Sample	NO3- (mg/l)	P (mg/l)	K (mg/l)	pH	NO3- (mg/l)	P (mg/l)	K (mg/l)	pH	NO3- (mg/l)	P (mg/l)	K (mg/l)	pH
A3	45.15	20	100	7.62	7.1	11	150	7.48	5.3	15	120	7.79
A4	9.73	20	110	7.84	7.5	18	150	7.6	23	11	200	7.92
A5	18.14	22	100	7.76	25.7	20	205	7.34	7.1	11	245	7.9
B5	17.6	23	105	7.59	20.8	20	120	7.7	26.1	13	180	7.78
B6	8.9	13	90	7.58	3.5	25	100	7.84	2.7	16	100	7.84
C5	36.3	20	100	7.68	26.6	17	130	7.36	20.8	11	140	7.78
C6	13.3	23	90	7.52	9.3	9	150	7.57	17.7	16	200	7.63
D6	46.1	18	95	7.66	25.7	16	155	7.57	22.6	16	195	7.85
E6	19	19	90	7.51	8.4	16	100	7.72	15.9	16	160	7.86
F6	11.5	21	85	7.78	4.9	13	165	7.69	26.5	26	185	7.86
mean	22.57	19.90	96.50	7.65	13.95	16.50	142.50	7.59	16.77	15.10	172.50	7.82

min	8.9	13	85	7.51	3.5	9	100	7.34	2.7	11	100	7.63
max	46.1	23	110	7.84	26.6	25	205	7.84	26.5	26	245	7.92
Zone 4												
June 2021					July 2021				August 2021			
Sample	NO3- (mg/l)	P (mg/l)	K (mg/l)	pH	NO3- (mg/l)	P (mg/l)	K (mg/l)	pH	NO3- (mg/l)	P (mg/l)	K (mg/l)	pH
A6	23.02	13	80	7.78	31.9	9	105	7.51	15.9	11	130	7.79
A7	29.21	19	135	7.8	9.7	12	165	7.66	39.4	11	145	7.83
A8	30.98	20	145	7.89	28.4	22	170	7.69	25.7	20	190	7.84
B7	8.9	14	80	7.56	6.6	10	150	7.86	7.1	10	135	7.8
B8	4.4	11	115	7.18	4.4	10	175	7.8	4.4	9	155	7.64
C7	12	8	100	7.65	19	11	120	7.78	8.9	12	200	7.77
C8	33.7	9	110	7.59	16.4	18	150	7.82	7.1	12	95	7.82
D7	16.8	13	105	7.75	48.3	20	125	7.68	20.8	13	130	7.88
D8	36.2	17	105	7.78	23.9	20	130	7.78	13.7	13	165	7.87
E7	1.8	26	95	7.87	12	15	165	7.83	7.1	12	140	7.89
E8	15.5	21	140	7.84	23.9	16	140	7.83	35.4	12	165	7.87
F7	47.8	17	80	7.89	9.3	13	108	7.76	15.4	14	100	7.88
F8	2.7	20	115	7.76	0.9	18	120	7.86	0	20	125	7.86
mean	20.23	16.00	108.08	7.72	18.05	14.92	140.23	7.76	15.45	13.00	144.23	7.83
min	1.8	8	80	7.18	0.9	9	105	7.51	0	9	95	7.64
max	47.8	26	145	7.89	48.3	22	175	7.86	39.4	20	200	7.89

Appendix 5 Hierarchical multinomial logistic regression model

The steps below are performed using MATLAB and uses the function in the folder named as 'hierarchical multinomial logistic regression model' in the supplementary materials.

- 1) To run the function below select 'Oats_yield_production.m' script hit run it. Ensure all the other function scripts specified in the folder of 'hierarchical multinomial logistic regression model' and the 'data' file from excel are present in the folder.
- 2) The data file consist of the measured yield, NDVI and CI green along with the soil nutrients and beta-glucan levels collected across three months of June, July and August 21.
- 3) The output after running the scripts will show the labels as below;
 - r_corr_mat : correlation matrix between all pairs of features
 - pV_mat : associated p-value for correlation values in r_corr_mat
 - B_cell : B coefficient values across all partitions in Leave-one-out
 - cross validation
 - dev_cell : fitted logistic regression model deviance across all partitions in Leave-one-out cross validation
 - pihat_dp : probability estimates for each of the unique labels in the
 - data
 - MI : labels assigned by the trained logistic regression model to each test datapoint
 - pp : p-value representing model significance against an intercept only model
 - pseudo_R2 - McFadden's pseudo R2 explains variability in dependant variable explained by independent predictors over an intercept only model
 - accuracy - model prediction accuracy across all labels
 - Val_all - Val_all(:,5)- Sensitivity; Val_all(:,6)- Specificity; Val_all(:,7)- Accuracy for each label.

Script Oats_yield_production.m'

% Read in Data from excel file

```
addpath('./Data/');  
addpath('./functions/');
```

```

Sh1=xlsread('Data.xlsx',1);
Sh2=xlsread('Data.xlsx',2);

data=Sh1(:,[3:7,12:16,21:25]); % data consisting of only soil composition values in June,
July and August

% Dependant variable
dep_var=Sh2(2:end,1:3);

B_glucan_aug=Sh2(2:end,4);

reco_labels= zeros(size(dep_var,1),size(dep_var,2));

% recoding variables in dep_var - 3 dependent variables
for i=1:size(dep_var,1)
    for j=1:size(dep_var,2)
        if dep_var(i,j)>7.62
            reco_labels(i,j)=5;
        elseif dep_var(i,j)>7.11
            reco_labels(i,j)=4;
        elseif dep_var(i,j)>6.18
            reco_labels(i,j)=3;
        elseif dep_var(i,j)>4.97
            reco_labels(i,j)=2;
        else
            reco_labels(i,j)=1;
        end
    end
end

% data matrix consisting of soil composition over the 3 months
data=[data B_glucan_aug];

% variable cols and names
col=1:16;
names={'Nitrate (mg/l) - June 21','Potassium (mg/l) - June 21','Phosphorus (mg/l) - June
21','Organic matter (mg/l) - June 21','pH - June 21','Nitrate (mg/l) - July 21','Potassium (mg/l)
- July 21','Phosphorus (mg/l) - July 21','Organic matter (mg/l) - July 21','pH - July 21','Nitrate
(mg/l) - August 21','Potassium (mg/l) - August 21','Phosphorus (mg/l) - August 21','Organic
matter (mg/l) - August 21','pH - August 21','\beta - glucan (%) - August 21'};

%% for 3 datasets and 3 different models

%% Model 1 - using mean and covariance of multivariate normal distribution
%% Model 2 - using robust mean and covariance of multivariate normal distribution
%% Model 3 - using mean and covariance of multivariate lognormal distribution
%% Model 4 - using robust mean and covariance to remove skew in the multivariate
lognormal distribution data
%% Model 5 - using the robust cov +mean for skewed multivariate normal distribution

```

```
%% Model 6 - using only linearly interpolated datapoints
```

```
for i=1:6
```

```
    tt=strcat('Case ',num2str(i), ' being processed.');
```

```
    disp(tt);
```

```
    % disp('Model 1 completed');
```

```
    % model development #1
```

```
    [r_corr_mat1,pV_mat1,B_cell1, dev_cell1, stats_cell1, pihat_dp1, MI1, pp1, pseudo_R21,  
    accuracy1, Val_all1, selected_fs1, mu_grp1, std_grp1, Data_gen1,labels1] =  
    model_development(data,reco_labels,1,100,1,1,i);  
    disp('Model 1 completed');
```

```
    % model development #2
```

```
    [r_corr_mat2,pV_mat2,B_cell2, dev_cell2, stats_cell2, pihat_dp2, MI2, pp2, pseudo_R22,  
    accuracy2, Val_all2,selected_fs2, mu_grp2, std_grp2, Data_gen2,labels2] =  
    model_development(data,reco_labels,2,100,1,1,i);  
    disp('Model 2 completed');
```

```
    % model development #3
```

```
    [r_corr_mat3,pV_mat3,B_cell3, dev_cell3, stats_cell3, pihat_dp3, MI3, pp3, pseudo_R23,  
    accuracy3, Val_all3,selected_fs3, mu_grp3, std_grp3,Data_gen3,labels3] =  
    model_development(data,reco_labels,3,100,1,1,i);  
    disp('Model 3 completed');
```

```
    Model{i}.corrMat1=r_corr_mat1;
```

```
    Model{i}.corrMat2=r_corr_mat2;
```

```
    Model{i}.corrMat3=r_corr_mat3;
```

```
    Model{i}.pV_mat1=pV_mat1;
```

```
    Model{i}.pV_mat2=pV_mat2;
```

```
    Model{i}.pV_mat3=pV_mat3;
```

```
    Model{i}.B_cell1=B_cell1;
```

```
    Model{i}.B_cell2=B_cell2;
```

```
    Model{i}.B_cell3=B_cell3;
```

```
    Model{i}.dev_cell1=dev_cell1;
```

```
    Model{i}.dev_cell2=dev_cell2;
```

```
    Model{i}.dev_cell3=dev_cell3;
```

```
    Model{i}.stats_cell1=stats_cell1;
```

```
    Model{i}.stats_cell2=stats_cell2;
```

```
    Model{i}.stats_cell3=stats_cell3;
```

```
    Model{i}.pihat_dp1=pihat_dp1;
```

```
    Model{i}.pihat_dp2=pihat_dp2;
```

```

Model{i}.pihat_dp3=pihat_dp3;

Model{i}.MI1=MI1;
Model{i}.MI2=MI2;
Model{i}.MI3=MI3;

Model{i}.pp1=pp1;
Model{i}.pp2=pp2;
Model{i}.pp3=pp3;

Model{i}.pseudo_R21=pseudo_R21;
Model{i}.pseudo_R22=pseudo_R22;
Model{i}.pseudo_R23=pseudo_R23;

Model{i}.accuracy1=accuracy1;
Model{i}.accuracy2=accuracy2;
Model{i}.accuracy3=accuracy3;

Model{i}.Val_all1=Val_all1;
Model{i}.Val_all2=Val_all2;
Model{i}.Val_all3=Val_all3;

Model{i}.selected_fs1=selected_fs1;
Model{i}.selected_fs2=selected_fs2;
Model{i}.selected_fs3=selected_fs3;

Model{i}.mu_grp1=mu_grp1;
Model{i}.mu_grp2=mu_grp2;
Model{i}.mu_grp3=mu_grp3;

Model{i}.std_grp1=std_grp1;
Model{i}.std_grp2=std_grp2;
Model{i}.std_grp3=std_grp3;

Model{i}.Data_gen1=Data_gen1;
Model{i}.Data_gen2=Data_gen2;
Model{i}.Data_gen3=Data_gen3;

Model{i}.labels1=labels1;
Model{i}.labels2=labels2;
Model{i}.labels3=labels3;

end

% plot the correlation matrix change for CI green dataset for change in the
% covariance
% Feature set you want to look at in the plot
featureS=[1:5 16];

```



```
figure(1)
Data_m1d1=Model{1,1}.Data_gen1(labels1==1,:);
plotmatrix(Data_m1d1(featureS,featureS));
```

```
figure(2)
Data_m2d1=Model{1,2}.Data_gen1(labels1==1,:);
plotmatrix(Data_m2d1(featureS,featureS));
```

```
figure(3)
Data_m3d1=Model{1,3}.Data_gen1(labels1==1,:);
plotmatrix(Data_m3d1(featureS,featureS));
```

```
figure(4)
Data_m4d1=Model{1,4}.Data_gen1(labels1==1,:);
plotmatrix(Data_m4d1(featureS,featureS));
```

```
figure(5)
Data_m5d1=Model{1,5}.Data_gen1(labels1==1,:);
plotmatrix(Data_m5d1(featureS,featureS));
```

```
figure(6)
Data_m6d1=Model{1,6}.Data_gen1(labels1==1,:);
plotmatrix(Data_m6d1(featureS,featureS));
```

%% Mismatch in the labels assigned to the 3 dependant variables

% between 1, 2 and 3

```
mismatch_12=(sum(reco_labels(:,1)~=reco_labels(:,2)))/size(reco_labels,1);
mismatch_13=(sum(reco_labels(:,1)~=reco_labels(:,3)))/size(reco_labels,1);
mismatch_23=(sum(reco_labels(:,2)~=reco_labels(:,3)))/size(reco_labels,1);
```

%% outlining the concentration range for different selected fetures(i.e. nutrients) for each unique yield label

```
unique_labels=unique(reco_labels);
```

```
Range1=zeros(length(unique_labels),2,size(data,2));
Range2=zeros(length(unique_labels),2,size(data,2));
Range3=zeros(length(unique_labels),2,size(data,2));
```

```
Xaxis={};
```

```
for j=1:size(Model,2)
    for i=1:length(unique_labels)
        % model #1 all feature range
        Model{j}.Range1(i,1,:)=min(Model{j}.Data_gen1(labels1==i,:));
        Model{j}.Range1(i,2,:)=max(Model{j}.Data_gen1(labels1==i,:));
```

```

    Model{j}.Xaxis{1,i}=Model{j}.Data_gen1(labels1==i,:);
%end

% model #2 all feature range
% for i=1:length(unique_labels)
    Model{j}.Range2(i,1,:)=min(Model{j}.Data_gen2(labels2==i,:));
    Model{j}.Range2(i,2,:)=max(Model{j}.Data_gen2(labels2==i,:));
    Model{j}.Xaxis{2,i}=Model{j}.Data_gen2(labels2==i,:);

%end

% model #3 all feature range
%for i=1:length(unique_labels)
    Model{j}.Range3(i,1,:)=min(Model{j}.Data_gen3(labels3==i,:));
    Model{j}.Range3(i,2,:)=max(Model{j}.Data_gen3(labels3==i,:));
    Model{j}.Xaxis{3,i}=Model{j}.Data_gen3(labels3==i,:);
end
end

all_Data1=[];
all_Data2=[];
all_Data3=[];

all_labs1=[];
all_labs2=[];
all_labs3=[];

for i=1:size(Model,2)
    all_Data1=[all_Data1; Model{i}.Data_gen1];
    all_labs1=[all_labs1; Model{i}.labels1];

    all_Data2=[all_Data2; Model{i}.Data_gen2];
    all_labs2=[all_labs2; Model{i}.labels2];

    all_Data3=[all_Data3; Model{i}.Data_gen3];
    all_labs3=[all_labs3; Model{i}.labels3];
end

[r_corr_mat_A1, pV_mat_A1,B_A1, dev_A1,devD_A1,dfed_A1, stats_A1, pihat_A1,
MI_A1, pp_A1, pseudo_R2_A1, accuracy_A1, Val_A1, selected_features_A1,
train_data_A1, test_data_A1] = full_model_development(all_Data1,all_labs1,1);
[r_corr_mat_A2, pV_mat_A2,B_A2, dev_A2, devD_A2,dfed_A2, stats_A2, pihat_A2,
MI_A2, pp_A2, pseudo_R2_A2, accuracy_A2, Val_A2, selected_features_A2,
train_data_A2, test_data_A2] = full_model_development(all_Data2,all_labs2,2);
[r_corr_mat_A3, pV_mat_A3,B_A3, dev_A3,devD_A3,dfed_A3, stats_A3, pihat_A3,
MI_A3, pp_A3, pseudo_R2_A3, accuracy_A3, Val_A3, selected_features_A3,
train_data_A3, test_data_A3] = full_model_development(all_Data3,all_labs3,3);

% intersection of the features marked significant by 3 different data

```

```

% models
fs_intersect=intersect(intersect(selected_features_A1,selected_features_A2,'stable'),selected_
features_A3,'stable');

% Features significant across all 3 models
F_sig=[];
for i=1:length(fs_intersect)
    F_sig=[F_sig; string(names{fs_intersect(i)})];
end

for j=1:(size(Model,2)-1)
    %size(Model,2)
    %figure;
    if j==1 | j==2
        plot_cl_pdfs(Model{j}.Xaxis,[5 5 5],1,6, 200,names,'mvn',j);
    elseif j==3 | j==4
        plot_cl_pdfs(Model{j}.Xaxis,[5 5 5],1,6, 200,names,'logn',j);
    else
        plot_cl_pdfs(Model{j}.Xaxis,[5 5 5],1,6, 100,names,'mvsn',j);
    end
end

colorsS=[255,114,118;255,255,102;255, 165, 0;0, 255, 0;34,139,34];
colorsS=colorsS./255;
nn=get(gcf,'Number');
Model_6_title=[{'Linearly interpolated yield plot - Measured Data'},{'Linearly interpolated
yield plot - NDVI Data'},{'Linearly interpolated yield plot - CI green Data'}];
%scatter plot of last model
for i=1:size(Model{6}.Xaxis,1)
    figure(nn+i);
    for j=1:size(Model{6}.Xaxis,2)
        A=Model{6}.Xaxis{i,j};

scatter3(A(:,2),A(:,7),A(:,16),60,colorsS(j,:),'MarkerEdgeColor','k','MarkerFaceColor',colors
S(j,:));
        hold on;
        xticks([min(A(:,2)):30:max(A(:,2))]);
        yticks([min(A(:,7)):40:max(A(:,7))]);
        xlabel(names{2});
        ylabel(names{7});
        zlabel(names{16});
        box on;
        axis square;
        axis tight;
    end
    legend(['Very low yield'],'Low yield'],'Medium yield'],'High yield'],'Very high
yield'])
    %title(Model_6_title{1,i});
    set(gca,"FontSize",16)
    hYLabel = get(gca,'YLabel');

```

```

    set(hYLabel,'rotation',-
35,'Units','normalized','VerticalAlignment','baseline','HorizontalAlignment','right','Position',[0
.32673749997845,-0.118479586512629,0])
    hXLabel = get(gca,'XLabel');

set(hXLabel,'rotation',20,'Units','normalized','VerticalAlignment','baseline','HorizontalAlignm
ent','left','Position',[0.556156794063963,-0.0978,0])

end

%% Original data scatter - CI green
nn=get(gcf,'Number');
F={};
E=unique(reco_labels(:,3));
for i=1:length(E)
    F{i}=data(reco_labels(:,3)==E(i),:);
end

figure(nn+1);
for i=1:size(F,2)
    AA=F{i};

scatter3(AA(:,2),AA(:,7),AA(:,16),60,colorsS(i,:),'MarkerEdgeColor','k','MarkerFaceColor',c
olorsS(i,:));
    hold on;
    xticks([min(AA(:,2)):30:max(AA(:,2))]);
    yticks([min(AA(:,7)):40:max(AA(:,7))]);
    xlabel(names{2});
    ylabel(names{7});
    zlabel(names{16});
    box on;
    axis square;
    axis tight;
end
legend(['Very low yield'],'Low yield'],'Medium yield'],'High yield'],'Very high
yield'])
set(gca,"FontSize",16)
hYLabel = get(gca,'YLabel');
set(hYLabel,'rotation',-
35,'Units','normalized','VerticalAlignment','baseline','HorizontalAlignment','right','Position',[0
.32673749997845,-0.118479586512629,0])
hXLabel = get(gca,'XLabel');
set(hXLabel,'rotation',20,'Units','normalized','VerticalAlignment','baseline','HorizontalAlignm
ent','left','Position',[0.556156794063963,-0.0978,0])

%% Univariate plot showing normal and 'ogk' mu and cov for same distribution
%% for Beta glucan variable class 2 in CI green

```

```

AAE=Model{1}.mu_grp3.mu_F2(1,16);
AAEc=sqrt(Model{1}.std_grp3.cov_F2(16,16));
AAE2=Model{2}.mu_grp3.mu_F2(1,16);
AAE2c=sqrt(Model{2}.std_grp3.cov_F2(16,16));

tt=normpdf([-1:0.01:4],AAE,AAEc);
tt1=normpdf([-1:0.01:4],AAE2,AAE2c);

nn=get(gcf,'Number');

figure(nn+1);
plot([-1:0.01:4],tt,'b-','LineWidth',2);
hold on;
plot([-1:0.01:4],tt1,'r-','LineWidth',2);
box on;
axis tight;
axis square;
xlabel(names{16});
ylabel('Probability density function');
ylim([0 1]);
legend(['no bias correction'],'bias correction'],'Location','northeast');
set(gca,"FontSize",20)

%%% range Values for very high and very low yield from the original 48 data
data_hy=data(reco_labels(:,1)==5 & reco_labels(:,2)==5 & reco_labels(:,3)==5,:);

min_data_hy=min(data_hy);
max_data_hy=max(data_hy);

Range_valvH=num2cell([min_data_hy' max_data_hy']);
Range_finalvHigh=[names' Range_valvH];

data_ly=data(reco_labels(:,1)==1 & reco_labels(:,2)==1 & reco_labels(:,3)==1,:);

min_data_ly=min(data_ly);
max_data_ly=max(data_ly);

Range_valvL=num2cell([min_data_ly' max_data_ly']);
Range_finalvLow=[names' Range_valvL];

```

Appendix 6

Table 6.1 Model Accuracy and McFadden pseudo R².

Model 1 - Multivariate normal distribution without bias correction	Accuracy	McFadden pseudo R²
Measured yield	0.81	0.79
NDVI yield	0.6	0.58
CI green yield	0.88	0.76
Model 2 - Multivariate normal distribution with bias correction	Accuracy	McFadden pseudo R²
Measured yield	0.78	0.84
NDVI yield	0.64	0.57
CI green yield	0.82	0.8
Model 3 - Multivariate lognormal distribution without bias correction	Accuracy	McFadden pseudo R²
Measured yield	0.79	0.61
NDVI yield	0.52	0.51
CI green yield	0.78	0.7
Model 4 - Multivariate lognormal distribution with bias correction	Accuracy	McFadden pseudo R²
Measured yield	0.86	0.79
NDVI yield	0.55	0.58
CI green yield	0.82	0.76
Model 5 - Multivariate skewed normal distribution with bias correction	Accuracy	McFadden pseudo R²
Measured yield	0.76	0.86
NDVI yield	0.7	0.63
CI green yield	0.94	0.72
Model 6 - Linearly interpolated datapoint model	Accuracy	McFadden pseudo R²
Measured yield	0.94	0.78
NDVI yield	0.56	0.58
CI green yield	1	0.83
Model 7 - Combination model	Accuracy	McFadden pseudo R²
Measured yield	0.68	0.74

NDVI yield	0.72	0.61
CI green yield	0.84	0.76

Table 6. 2 Specificity, sensitivity, and accuracy of model 1, model 2, model 3, model 4, model 5, model 6 and model 7 for measured, NDVI and CI green yield.

	Model 1 - Multivariate normal distribution without bias correction			Model 2 - Multivariate normal distribution with bias correction		
	Measure yield			Measured yield		
	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy
Class 1 - very low yield (0 – 4.97 t/ha)	0.91	0.98	0.97	0.91	0.98	0.97
Class 2 - low yield (4.97 – 6.18 t/ha)	0.84	0.98	0.95	0.84	0.98	0.95
Class 3 - medium yield (6.18 – 7.11 t/ha)	0.81	0.93	0.91	0.81	0.93	0.91
Class 4 - high yield (7.11 – 7.62 t/ha)	0.99	0.95	0.96	0.99	0.95	0.96
Class 5 - very high yield (>7.62t/ha)	0.67	0.95	0.89	0.67	0.95	0.89
	Model 1 - Multivariate normal distribution without bias correction			Model 2 - Multivariate normal distribution with bias correction		
	NDVI yield			NDVI yield		
	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy
Class 1 - very low yield (0 – 4.97 t/ha)	0.48	0.97	0.87	0.15	0.98	0.81
Class 2 - low yield (4.97 – 6.18 t/ha)	0.42	0.87	0.78	0.27	0.86	0.74
Class 3 - medium yield (6.18 – 7.11 t/ha)	0.37	0.89	0.78	0.84	0.86	0.86
Class 4 - high yield (7.11 – 7.62 t/ha)	1.00	0.83	0.86	1.00	0.83	0.86
Class 5 - very high yield (>7.62t/ha)	0.66	0.91	0.86	0.66	0.93	0.88
	Model 1 - Multivariate normal distribution without bias correction			Model 2 - Multivariate normal distribution with bias correction		
	CI green yield			CI green yield		
	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy
Class 1 - very low yield (0 – 4.97 t/ha)	0.57	0.98	0.90	0.77	0.97	0.93
Class 2 - low yield (4.97 – 6.18 t/ha)	0.44	0.98	0.87	0.70	0.97	0.91
Class 3 - medium yield (6.18 – 7.11 t/ha)	0.95	0.84	0.86	0.74	0.98	0.93
Class 4 - high yield (7.11 – 7.62 t/ha)	1.00	0.92	0.94	0.99	0.87	0.89
Class 5 - very high yield (>7.62t/ha)	0.82	0.97	0.94	0.78	0.96	0.92

	Model 3 - Multivariate lognormal without bias correction			Model 4 - Multivariate lognormal with bias correction		
	Measured yield			Measured yield		
	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy
Class 1 - very low yield (0 – 4.97 t/ha)	0.45	0.99	0.88	0.59	0.98	0.90
Class 2 - low yield (4.97 – 6.18 t/ha)	0.40	0.96	0.86	0.95	0.94	0.94
Class 3 - medium yield (6.18 – 7.11 t/ha)	0.50	0.96	0.87	0.74	0.98	0.93
Class 4 - high yield (7.11 – 7.62 t/ha)	0.96	0.67	0.73	0.96	0.88	0.89
Class 5 - very high yield (>7.62t/ha)	0.72	0.93	0.88	0.72	0.96	0.91
	Model 3 - Multivariate lognormal without bias correction			Model 4 - Multivariate lognormal with bias correction		
	NDVI yield			NDVI yield		
	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy
Class 1 - very low yield (0 – 4.97 t/ha)	0.46	0.97	0.87	0.42	0.96	0.85
Class 2 - low yield (4.97 – 6.18 t/ha)	0.28	0.87	0.75	0.34	0.89	0.78
Class 3 - medium yield (6.18 – 7.11 t/ha)	0.24	0.86	0.74	0.59	0.82	0.77
Class 4 - high yield (7.11 – 7.62 t/ha)	1.00	0.84	0.87	1.00	0.93	0.94
Class 5 - very high yield (>7.62t/ha)	0.61	0.84	0.80	0.57	0.87	0.81
	Model 3 - Multivariate lognormal without bias correction			Model 4 - Multivariate lognormal with bias correction		
	CI green yield			CI green yield		
	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy
Class 1 - very low yield (0 – 4.97 t/ha)	0.47	0.97	0.87	0.56	0.96	0.88
Class 2 - low yield (4.97 – 6.18 t/ha)	0.44	0.93	0.83	0.59	0.95	0.88
Class 3 - medium yield (6.18 – 7.11 t/ha)	0.85	0.87	0.86	0.93	0.88	0.89
Class 4 - high yield (7.11 – 7.62 t/ha)	1.00	0.93	0.94	0.99	0.96	0.96

Class 5 - very high yield (>7.62t/ha)	0.72	0.93	0.89	0.75	0.95	0.91
---------------------------------------	------	------	------	------	------	------

	Model 5 - Skewed multivariate normal distribution with bias correction			Model 6 - Linearly interpolated datapoint		
	Measured yield			Measured yield		
	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy
Class 1 - very low yield (0 – 4.97 t/ha)	0.78	0.98	0.94	0.62	1.00	0.92
Class 2 - low yield (4.97 – 6.18 t/ha)	0.96	0.99	0.99	0.72	0.97	0.92
Class 3 - medium yield (6.18 – 7.11 t/ha)	0.89	0.96	0.95	0.72	0.97	0.92
Class 4 - high yield (7.11 – 7.62 t/ha)	0.96	0.98	0.98	1.00	0.82	0.85
Class 5 - very high yield (>7.62t/ha)	0.75	0.91	0.88	0.86	0.98	0.95
	Model 5 - Skewed multivariate normal distribution with bias correction			Model 6 - Linearly interpolated datapoint		
	NDVI yield			NDVI yield		
	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy
Class 1 - very low yield (0 – 4.97 t/ha)	0.49	0.98	0.88	0.25	0.96	0.81
Class 2 - low yield (4.97 – 6.18 t/ha)	0.27	0.91	0.78	0.50	0.89	0.82
Class 3 - medium yield (6.18 – 7.11 t/ha)	0.79	0.87	0.86	0.50	0.89	0.81
Class 4 - high yield (7.11 – 7.62 t/ha)	1.00	0.82	0.85	1.00	0.79	0.83
Class 5 - very high yield (>7.62t/ha)	0.65	0.96	0.90	0.69	0.94	0.89
	Model 5 - Skewed multivariate normal distribution with bias correction			Model 6 – Linearly interpolated datapoint		
	CI green yield			CI green yield		
	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy
Class 1 - very low yield (0 – 4.97 t/ha)	0.57	0.97	0.89	0.75	1.00	0.95
Class 2 - low yield (4.97 – 6.18 t/ha)	0.21	0.97	0.82	0.51	0.95	0.86
Class 3 - medium yield (6.18 – 7.11 t/ha)	0.98	0.81	0.84	0.92	0.89	0.90
Class 4 - high yield (7.11 – 7.62 t/ha)	1.00	0.92	0.94	1.00	0.96	0.97

Class 5 - very high yield (>7.62t/ha)	0.85	0.98	0.95	0.98	1.00	1.00
-------------------------------------------------	------	------	------	------	------	------

Chapter 13: References

1. Abbas, A., Khan, S., Hussain, N., Hanjra, M. A., & Akbar, S. (2013). Characterizing soil salinity in irrigated agriculture using a remote sensing approach. *Physics and Chemistry of the Earth, Parts A/B/C*, 55-57, 43-52. doi:10.1016/j.pce.2010.12.004
2. Aber, J.S., Marzloff, I. & Ries, J.B. (2010). Chapter 1- to small format-aerial photography, Small Format Aerial Photography. *Principles, Techniques and Geoscience Applications*, 1-13.
3. Acorsi, M. G., das Dores Abati Miranda, F., Martello, M., Smaniotto, D. A., & Sartor, L. R. (2019). Estimating Biomass of Black Oat Using UAV-Based RGB Imaging. *Agronomy*, 9(7). doi:10.3390/agronomy9070344
4. AHDB. 2023. *Measuring soil nutrients, pH and organic matter*. <https://archive.ahdb.org.uk/measuring-soil-nutrients-ph-and-organic-matter>
5. Ahmed, F., Mohanta, J. C., Keshari, A., & Yadav, P. S. (2022). Recent Advances in Unmanned Aerial Vehicles: A Review. *Arab J Sci Eng*, 47(7), 7963-7984. doi:10.1007/s13369-022-06738-0.
6. Alabi, T. R., Abebe, A. T., Chigeza, G., & Fowobaje, K. R. (2022). Estimation of soybean grain yield from multispectral high-resolution UAV data with machine learning models in West Africa. *Remote Sensing Applications: Society and Environment*, 27. doi:10.1016/j.rsase.2022.100782
7. Ali, A., & Imran, M. M. Evaluating the potential of red edge position (REP) of hyperspectral remote sensing data for real time estimation of LAI & chlorophyll content of kinnow mandarin (*Citrus reticulata*) fruit orchards. (2020). *Sci. Hortic. Amst*, 267, 109326. doi: 10.1016/j.scienta.2020.109326
8. Amigo, J.M. (2020). Hyperspectral and multispectral imaging: setting the scene. *Department of Food Science*, 32, 3-16. <https://doi.org/10.1016/B978-0-444-63977-6.00001-8>.
9. Anas, M., Liao, F., Verma, K. K., Sarwar, M. A., Mahmood, A., Chen, Z. L., ... & Li, Y. R. (2020). Fate of nitrogen in agriculture and environment: agronomic, eco-physiological and molecular approaches to improve nitrogen use efficiency. *Biological Research*, 53(1), 1-20. [10.1186/s40659-020-00312-4](https://doi.org/10.1186/s40659-020-00312-4)

10. Arnó, J., Escolà, A., Vallès, J. M., Llorens, J., Sanz, R., Masip, J., . . . Rosell-Polo, J. R. (2012). Leaf area index estimation in vineyards using a ground-based LiDAR scanner. *Precision Agriculture*, *14*(3), 290-306. doi:10.1007/s11119-012-9295-0
11. Atkinson, S., & Townsend, M. (2011). The state of the UK's forests, woods and trees: perspectives from the sector. *Woodland Trust, Grantham, Lincolnshire*.
12. Barbosa, J. d. A., Faria, R. T. d., Coelho, A. P., Dalri, A. B., & Palaretti, L. F. (2020). Nitrogen fertilization management in white oat using spectral indices. *Pesquisa Agropecuária Tropical*, *50*. doi:10.1590/1983-40632020v5064924
13. Bareth, G., Aasen, H., Bendig, J., Gnyp, M. L., Bolten, A., Jung, A., . . . Soukkamäki, J. (2015). Leichte und UAV-getragene hyperspektrale, bildgebende Kameras zur Beobachtung von landwirtschaftlichen Pflanzenbeständen: spektraler Vergleich mit einem tragbaren Feldspektrometer. *Photogrammetrie - Fernerkundung - Geoinformation*, *2015*(1), 69-79. doi:10.1127/pfg/2015/0256
14. Bargali, K., Joshi, B., Bargali, S. S., & Singh, S. P. (2015). Oaks and the biodiversity they sustain. *International Oaks*, *26*, 65-76.
15. Bascietto, M., Santangelo, E., & Beni, C. (2021). Spatial Variations of Vegetation Index from Remote Sensing Linked to Soil Colloidal Status. *Land*, *10*(1). doi:10.3390/land10010080
16. Baumann, P.R. (2014). *History of remote sensing, aerial photography*. <http://employees.oneonta.edu/baumanpr/geosat2/rs%20history%20i/rs-history-part-1.html>
17. Bayer, B.E. (1976). Color imaging array. *United States Patent*, 1-10.
18. Beaufort, G. D. & Dusariez. M. (1995). *Aerial photographs taken from a kite: yesterday and today*. KAPWA-Foundation Publishing.
19. Bendig, J., Bolten, A., Bennertz, S., Broscheit, J., Eichfuss, S., & Bareth, G. (2014). Estimating Biomass of Barley Using Crop Surface Models (CSMs) Derived from UAV-Based RGB Imaging. *Remote Sensing*, *6*(11), 10395-10412. doi:10.3390/rs61110395
20. Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., . . . Bareth, G. (2015). Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. *International Journal of Applied Earth Observation and Geoinformation*, *39*, 79-87. doi:10.1016/j.jag.2015.02.012

21. Boyacı, İ., Şeker, U., & Mutlu, M. (2002). Determination of β -glucan content of cereals with an amperometric glucose electrode. *European Food Research and Technology*, 215(6), 538-541. doi:10.1007/s00217-002-0609-8
22. Bradley, B. A., & Mustard, J. F. (2006). Characterizing The Landscape Dynamics Of An Invasive Plant And Risk Of Invasion Using Remote Sensing. *Ecological Applications*, 16(3), 1132-1147. doi:10.1890/1051-0761(2006)016[1132:Ctldoa]2.0.Co;2
23. Brovkina, O., Cienciala, E., Surový, P., & Janata, P. (2018). Unmanned aerial vehicles (UAV) for assessment of qualitative classification of Norway spruce in temperate forest stands. *Geo-spatial Information Science*, 21(1), 12-20. doi:10.1080/10095020.2017.1416994
24. Buerstmayr, H., Krenn, N., Stephan, U., Grausgruber, H., & Zechner, E. (2007). Agronomic performance and quality of oat (*Avena sativa* L.) genotypes of worldwide origin produced under Central European growing conditions. *Field Crops Research*, 101(3), 343-351.
25. Burney, S. A., & Tariq, H. (2014). K-means cluster analysis for image segmentation. *International Journal of Computer Applications*, 96(4).
26. Burns, B. W., Green, V. S., Hashem, A. A., Massey, J. H., Shew, A. M., Adviento-Borbe, M. A. A., & Milad, M. (2022). Determining nitrogen deficiencies for maize using various remote sensing indices. *Precision Agriculture*, 23(3), 791-811. doi:10.1007/s11119-021-09861-4
27. Bzducha-Wróbel, A., Błażej, S., & Tkacz, K. (2012). Cell wall structure of selected yeast species as a factor of magnesium binding ability. *European Food Research and Technology*, 235, 355-366.
28. Cao, H., Gu, X., Wei, X., Yu, T., & Zhang, H. (2020). Lookup Table Approach for Radiometric Calibration of Miniaturized Multispectral Camera Mounted on an Unmanned Aerial Vehicle. *Remote Sensing*, 12(24). doi:10.3390/rs12244012
29. Cao, J., Leng, W., Liu, K., Liu, L., He, Z., & Zhu, Y. (2018). Object-Based Mangrove Species Classification Using Unmanned Aerial Vehicle Hyperspectral Images and Digital Surface Models. *Remote Sensing*, 10(2). doi:10.3390/rs10010089
30. Chang, S. X., & Robison, D. J. (2003). Nondestructive and rapid estimation of hardwood foliar nitrogen status using the SPAD-502 chlorophyll meter. *Forest Ecology and Management*, 181(3), 331-338. doi:10.1016/s0378-1127(03)00004-5

31. Chemura, A., van Duren, I., & van Leeuwen, L. M. (2015). Determination of the age of oil palm from crown projection area detected from WorldView-2 multispectral remote sensing data: The case of Ejisu-Juaben district, Ghana. *ISPRS journal of photogrammetry and remote sensing*, *100*, 118-127.
32. Chen, A., Orlov-Levin, V., & Meron, M. (2019). Applying high-resolution visible-channel aerial imaging of crop canopy to precision irrigation management. *Agricultural Water Management*, *216*, 196-205. doi:10.1016/j.agwat.2019.02.017
33. Chen, Y. C. (2017). A tutorial on kernel density estimation and recent advances. *Biostatistics & Epidemiology*, *1*(1), 161-187. <https://doi.org/10.1080/24709360.2017.1396742>
34. Chisholm, R. A., Cui, J., Lum, S. K. Y., & Chen, B. M. (2013). UAV LiDAR for below-canopy forest surveys. *Journal of Unmanned Vehicle Systems*, *01*(01), 61-68. doi:10.1139/juvs-2013-0017
35. Chitroub, S., Houacine, A., & Sansal, B. (2001, June). Principal component analysis of multispectral images using neural network. In *Proceedings ACS/IEEE International Conference on Computer Systems and Applications* (pp. 89-95). IEEE.
36. Chong, K. L., Kanniah, K. D., Pohl, C., & Tan, K. P. (2017). A review of remote sensing applications for oil palm studies. *Geo-spatial Information Science*, *20*(2), 184-200. doi:10.1080/10095020.2017.1337317
37. Coelho, A. P., Faria, R. T. d., Leal, F. T., Barbosa, J. d. A., & Rosalen, D. L. (2020). Validation of white oat yield estimation models using vegetation indices. *Bragantia*, *79*(2), 236-241. doi:10.1590/1678-4499.20190387
38. Coelho, A. P., Faria, R. T. d., Leal, F. T., Barbosa, J. d. A., Dalri, A. B., & Rosalen, D. L. (2019). Estimation of irrigated oats yield using spectral indices. *Agricultural Water Management*, *223*. doi:10.1016/j.agwat.2019.105700
39. Cohen, J.L. (2006). *Above Paris: The Aerial Survey of Roger Henrard by Jean-Louis Cphen (2006-09/28)*. Princeton Architectural Press.
40. Colwell, R.N. (1997). *Manual of Photographic Interpretation*. American Society for Photogrammetry and Remote Sensing.
41. Cooper, N. S. (2000). How natural is a nature reserve?: an ideological study of British nature conservation landscapes. *Biodiversity & Conservation*, *9*, 1131-1152.
42. Cornelisse, D. G. (2003). *Splendid vision, unswerving purpose: Developing air power for the United States Air Force during the first century of powered flight*. Department of the Air Force.

43. Costa, J. M., Grant, O. M., & Chaves, M. M. (2013). Thermography to explore plant-environment interactions. *J Exp Bot*, *64*(13), 3937-3949. doi:10.1093/jxb/ert029
44. Cramer, M., Przybilla, H. J., & Zurhorst, A. (2017). Uav Cameras: Overview and Geometric Calibration Benchmark. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLII-2/W6*, 85-92. doi:10.5194/isprs-archives-XLII-2-W6-85-2017
45. Crane, A. (2018, May 14). *From startup to empire: the evolution of DJI drones*. <https://dronelife.com/2018/05/14/from-startup-to-empire-the-evolution-of-dji-drones/>
46. Cromey, D. W. (2013). Digital images are data: and should be treated as such. *Methods Mol Biol*, *931*, 1-27. doi:10.1007/978-1-62703-056-4_1
47. Crum, R.J. (2004) Constructing the image of flight: The Wright Brothers, Photography, and their Visual Heritage. *History of Photography*, *28*, 10–24.
48. Csajbók, J., Buday-Bódi, E., Nagy, A., Fehér, Z. Z., Tamás, A., Virág, I. C., . . . Kutasy, E. (2022). Multispectral Analysis of Small Plots Based on Field and Remote Sensing Surveys—A Comparative Evaluation. *Sustainability*, *14*(6). doi:10.3390/su14063339
49. Daly, D. (2023). *A not so short history of unmanned aerial vehicle (UAV)*. <https://consortiq.com/uas-resources/short-history-unmanned-aerial-vehicles-uavs>
50. Danielson, M. E., Dauth, R., Elmasry, N. A., Langeslay, R. R., Magee, A. S., & Will, P. M. (2010). Enzymatic method to measure β -1, 3- β -1, 6-glucan content in extracts and formulated products (GEM Assay). *Journal of agricultural and food chemistry*, *58*(19), 10305-10308. doi: 10.1021/jf102003m
51. Dash, J. P., Watt, M. S., Paul, T. S. H., Morgenroth, J., & Pearse, G. D. (2019). Early Detection of Invasive Exotic Trees Using UAV and Manned Aircraft Multispectral and LiDAR Data. *Remote Sensing*, *11*(15). doi:10.3390/rs11151812
52. De Sá, N.C., Castro, P., Carvalho, S., Marchante, E., López-Núñez, F.A. & Marchante, H. (2018). Mapping the Flowering of an Invasive Plant Using Unmanned Aerial Vehicles: Is There Potential for Biocontrol Monitoring? *Frontier Plant Science*, *9*, 293
53. Debnath, S., Paul, M., & Debnath, T. (2023). Applications of LiDAR in Agriculture and Future Research Directions. *Journal of Imaging*, *9*(3). doi:10.3390/jimaging9030057
54. Deng, L., Mao, Z., Li, X., Hu, Z., Duan, F., & Yan, Y. (2018). UAV-based multispectral remote sensing for precision agriculture: A comparison between different cameras.

- ISPRS Journal of Photogrammetry and Remote Sensing*, 146, 124-136.
doi:10.1016/j.isprsjprs.2018.09.008
55. Department for Environment, Food & Rural Affairs. (2022, September 26). *Agriculture in the UK Evidence Pack*. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1106562/AUK Evidence Pack 2021 Sept22.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1106562/AUK_Evidence_Pack_2021_Sept22.pdf)
 56. Dhanachandra, N., Manglem, K., & Chanu, Y. J. (2015). Image segmentation using K-means clustering algorithm and subtractive clustering algorithm. *Procedia Computer Science*, 54, 764-771.
 57. Di Gennaro, S. F., Nati, C., Dainelli, R., Pastonchi, L., Berton, A., Toscano, P., & Matese, A. (2020). An Automatic UAV Based Segmentation Approach for Pruning Biomass Estimation in Irregularly Spaced Chestnut Orchards. *Forests*, 11(3). doi:10.3390/f11030308
 58. Di Gennaro, S. F., Toscano, P., Gatti, M., Poni, S., Berton, A., & Matese, A. (2022). Spectral Comparison of UAV-Based Hyper and Multispectral Cameras for Precision Viticulture. *Remote Sensing*, 14(3). doi:10.3390/rs14030449
 59. Di Gennaro, S.F., Battiston, E., Di Marco, S., Facini, O., Matese, A., Nocentini, M., Palliotti, A., & Mugnai, L. . (2016). Unmanned Aerial Vehicle (UAV)-based remote sensing to monitor grapevine leaf stripe disease within a vineyard affected by esca complex. *Phytopathol. Mediterr*, 55, 262–275.
 60. DJI (2023). *P4 Multispectral*. <https://www.dji.com/uk/p4-multispectral?site=brandsite&from=nav>
 61. DJI (2023). *P4 Multispectral*. <https://www.dji.com/uk/p4-multispectral?site=brandsite&from=nav>
 62. Domingo, Ørka, Næsset, Kachamba, & Gobakken. (2019). Effects of UAV Image Resolution, Camera Type, and Image Overlap on Accuracy of Biomass Predictions in a Tropical Woodland. *Remote Sensing*, 11(8). doi:10.3390/rs11080948
 63. Drechsler, M., Touza, J., White, P. C., & Jones, G. (2016). Agricultural landscape structure and invasive species: the cost-effective level of crop field clustering. *Food Security*, 8, 111-121.
 64. Dronova, I., Spotswood, E. N., & Suding, K. N. (2017). Opportunities and constraints in characterizing landscape distribution of an invasive grass from very high resolution multi-spectral imagery. *Frontiers in Plant Science*, 8, 890.

65. Dubovik, O., Schuster, G. L., Xu, F., Hu, Y., Bösch, H., Landgraf, J., & Li, Z. (2021). Grand Challenges in Satellite Remote Sensing. *Frontiers in Remote Sensing*, 2. doi:10.3389/frsen.2021.619818
66. Dvořák, P., Müllerová, J., Bartaloš, T., & Brůna, J. (2015). Unmanned Aerial Vehicles for Alien Plant Species Detection and Monitoring. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-1/W4*, 83-90. doi:10.5194/isprsarchives-XL-1-W4-83-2015
67. Eddy, S., & Johnston, S. R. (2009). Comparison of Palintest Soil Analysis to External Laboratory Analysis. *Palintest Ltd.*
68. Evans, R. G., LaRue, J., Stone, K. C., & King, B. A. (2013). Adoption of site-specific variable rate sprinkler irrigation systems. *Irrig. Sci*, 31, 871–887.
69. Fang, S., Tang, W., Peng, Y., Gong, Y., Dai, C., Chai, R., & Liu, K. (2016). Remote Estimation of Vegetation Fraction and Flower Fraction in Oilseed Rape with Unmanned Aerial Vehicle Data. *Remote Sensing*, 8(5). doi:10.3390/rs8050416
70. Fassnacht, F. E., Latifi, H., Stereńczak, K., Modzelewska, A., Lefsky, M., Waser, L. T., ... & Ghosh, A. (2016). Review of studies on tree species classification from remotely sensed data. *Remote Sensing of Environment*, 186, 64-87.
71. Fawcett, D., Bennie, J., & Anderson, K. (2020). Monitoring spring phenology of individual tree crowns using drone-acquired NDVI data. *Remote Sensing in Ecology and Conservation*, 7(2), 227-244. doi:10.1002/rse2.184
72. Food and Agriculture Organization of the United Nations Rome. (2017). *The future of food and agriculture trends and challenges*. <https://www.fao.org/3/i6583e/i6583e.pdf>
73. Franklin, S. E., & Ahmed, O. S. (2017). Deciduous tree species classification using object-based analysis and machine learning with unmanned aerial vehicle multispectral data. *International Journal of Remote Sensing*, 39(15-16), 5236-5245. doi:10.1080/01431161.2017.1363442
74. Franklin, S. E., & Ahmed, O. S. (2018). Deciduous tree species classification using object-based analysis and machine learning with unmanned aerial vehicle multispectral data. *International Journal of Remote Sensing*, 39(15-16), 5236-5245.
75. Franzen, D., Kitchen, N., Holland, K., Schepers, J., & Raun, W. (2016). Algorithms for in-season nutrient management in cereals. *Agron. J*, 108, 1775–1781. doi:10.2134/agronj2016.01.0041

76. Freudenberg, M., Nölke, N., Agostini, A., Urban, K., Wörgötter, F., & Kleinn, C. (2019). Large scale palm tree detection in high resolution satellite images using U-Net. *Remote Sensing*, *11*(3), 312.
77. García-Martínez, H., Flores-Magdaleno, H., Ascencio-Hernández, R., Khalil-Gardezi, A., Tijerina-Chávez, L., Mancilla-Villa, O. R., & Vázquez-Peña, M. A. (2020). Corn Grain Yield Estimation from Vegetation Indices, Canopy Cover, Plant Density, and a Neural Network Using Multispectral and RGB Images Acquired with Unmanned Aerial Vehicles. *Agriculture*, *10*(7). doi:10.3390/agriculture10070277
78. Gardy, J., Hassanpour, A., Lai, X., & Rehan, M. (2014). The influence of blending process on the quality of rapeseed oil-used cooking oil biodiesels. *Int. Sci. J. Environ. Sci*, *3*, 233-240.
79. Gärtner, P., Förster, M., & Kleinschmit, B. (2016). The benefit of synthetically generated RapidEye and Landsat 8 data fusion time series for riparian forest disturbance monitoring. *Remote Sensing of Environment*, *177*, 237-247.
80. Ge, X., Wang, J., Ding, J., Cao, X., Zhang, Z., Liu, J., & Li, X. (2019). Combining UAV-based hyperspectral imagery and machine learning algorithms for soil moisture content monitoring. *PeerJ*, *7*, e6926. doi:10.7717/peerj.6926
81. Getzin, S., Wiegand, K., & Schöning, I. (2012). Assessing biodiversity in forests using very high-resolution images and unmanned aerial vehicles. *Methods in Ecology and Evolution*, *3*(2), 397-404. doi:10.1111/j.2041-210X.2011.00158.x
82. Gini, R., Sona, G., Ronchetti, G., Passoni, D., & Pinto, L. (2018). Improving Tree Species Classification Using UAS Multispectral Images and Texture Measures. *ISPRS International Journal of Geo-Information*, *7*(8). doi:10.3390/ijgi7080315
83. Gitelson, A. A., Kaufman, Y. J., & Merzlyak, M. N. (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote sensing of Environment*, *58*(3), 289-298.
84. Gitelson, A., & Merzlyak, M. N. (1994). Spectral reflectance changes associated with autumn senescence of *Aesculus hippocastanum* L. and *Acer platanoides* L. leaves. Spectral features and relation to chlorophyll estimation. *Journal of plant physiology*, *143*(3), 286-292.
85. Goldberg, E., Kirby, K., Hall, J., & Latham, J. (2007). The ancient woodland concept as a practical conservation tool in Great Britain. *Journal for Nature Conservation*, *15*(2), 109-119.

86. Gong, Y., Duan, B., Fang, S., Zhu, R., Wu, X., Ma, Y., & Peng, Y. (2018). Remote estimation of rapeseed yield with unmanned aerial vehicle (UAV) imaging and spectral mixture analysis. *Plant Methods*, *14*, 70. doi:10.1186/s13007-018-0338-z
87. Gonzalez-Dugo, V., Zarco-Tejada, P., Nicolás, E., Nortes, P. A., Alarcón, J. J., Intrigliolo, D. S., & Fereres, E. (2013). Using high resolution UAV thermal imagery to assess the variability in the water status of five fruit tree species within a commercial orchard. *Precision Agriculture*, *14*(6), 660-678. doi:10.1007/s11119-013-9322-9
88. Graham, S. (1999, September 17). *Introduction and history*. <https://earthobservatory.nasa.gov/features/RemoteSensing/remote.php>
89. Guan, S., Fukami, K., Matsunaka, H., Okami, M., Tanaka, R., Nakano, H., . . . Takahashi, K. (2019). Assessing Correlation of High-Resolution NDVI with Fertilizer Application Level and Yield of Rice and Wheat Crops using Small UAVs. *Remote Sensing*, *11*(2). doi:10.3390/rs11020112
90. Guo, Y., Fu, Y. H., Chen, S., Robin Bryant, C., Li, X., Senthilnath, J., . . . de Beurs, K. (2021). Integrating spectral and textural information for identifying the tasseling date of summer maize using UAV based RGB images. *International Journal of Applied Earth Observation and Geoinformation*, *102*. doi:10.1016/j.jag.2021.102435
91. Han, J., Zhang, Z., & Cao, J. (2020). Developing a New Method to Identify Flowering Dynamics of Rapeseed Using Landsat 8 and Sentinel-1/2. *Remote Sensing*, *13*(1). doi:10.3390/rs13010105
92. Haque, N. (2021, August 29). *The history and dominance of drone pioneer DJI*. https://techinspection.net/history-of-dji-drones/#google_vignette
93. Hawkins, J. A., Sawyer, J. E., Barker, D. W., & Lundvall, J. P. (2007). Using Relative Chlorophyll Meter Values to Determine Nitrogen Application Rates for Corn. *Agronomy Journal*, *99*(4), 1034-1040. doi:10.2134/agronj2006.0309
94. Hernandez-Santin, L., Rudge, M. L., Bartolo, R. E., & Erskine, P. D. (2019). Identifying Species and Monitoring Understorey from UAS-Derived Data: A Literature Review and Future Directions. *Drones*, *3*(1). doi:10.3390/drones3010009
95. Holzhauser, K., Rübiger, T., Rose, T., Kage, H., & Kühling, I. (2022). Estimation of Biomass and N Uptake in Different Winter Cover Crops from UAV-Based Multispectral Canopy Reflectance Data. *Remote Sensing*, *14*(18). doi:10.3390/rs14184525

96. Hossen, M. A., Diwakar, P. K., & Ragi, S. (2021). Total nitrogen estimation in agricultural soils via aerial multispectral imaging and LIBS. *Sci Rep*, *11*(1), 12693. doi:10.1038/s41598-021-90624-6
97. Huang, C. Y., & Asner, G. P. (2009). Applications of remote sensing to alien invasive plant studies. *Sensors (Basel)*, *9*(6), 4869-4889. doi:10.3390/s90604869
98. Huang, H., Lan, Y., Yang, A., Zhang, Y., Wen, S., & Deng, J. (2020). Deep learning versus Object-based Image Analysis (OBIA) in weed mapping of UAV imagery. *Int. J. Remote Sens*, *41*, 3446–3479. doi: [10.1080/01431161.2019.1706112](https://doi.org/10.1080/01431161.2019.1706112)
99. Huang, S., Tang, L., Hupy, J. P., Wang, Y., & Shao, G. (2021). A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing. *Journal of Forestry Research*, *32*(1), 1-6.
100. Hussain, S., Gao, K., Din, M., Gao, Y., Shi, Z., & Wang, S. (2020). Assessment of UAV-Onboard Multispectral Sensor for Non-Destructive Site-Specific Rapeseed Crop Phenotype Variable at Different Phenological Stages and Resolutions. *Remote Sensing*, *12*(3). doi:10.3390/rs12030397
101. ICH. *ICH guideline Validation of Analytical Procedures—Test and Methodology*. <http://www.ich.org/products/guidelines/quality/quality-single/article/validation-of-analytical-procedures-text-and-methodology.html>.
102. Imperial War Museum. (2023). *A brief history of drones*. <https://www.iwm.org.uk/history/a-brief-history-of-drones>
103. Jafarbiglu, H., & Pourreza, A. (2022). A comprehensive review of remote sensing platforms, sensors, and applications in nut crops. *Computers and Electronics in Agriculture*, *197*. doi:10.1016/j.compag.2022.106844
104. Jensen, J.R. (2000). *Remote Sensing of the Environment: An Earth Resource Perspective*. Prentice-Hall: New Jersey.
105. Ji, Y., Chen, Z., Cheng, Q., Liu, R., Li, M., Yan, X., . . . Yang, T. (2022). Estimation of plant height and yield based on UAV imagery in faba bean (*Vicia faba* L.). *Plant Methods*, *18*(1), 26. doi:10.1186/s13007-022-00861-7
106. Jia, Y., Su, Z., Zhang, Q., Zhang, Y., Gu, Y., & Chen, Z. (2015). Research on UAV Remote Sensing Image Mosaic Method Based on SIFT. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, *8*(11), 365-374. doi:10.14257/ijsp.2015.8.11.33

107. Jin, S., Su, Y., Song, S., Xu, K., Hu, T., Yang, Q., . . . Guo, Q. (2020). Non-destructive estimation of field maize biomass using terrestrial lidar: an evaluation from plot level to individual leaf level. *Plant Methods*, 16, 69. doi:10.1186/s13007-020-00613-5
108. Jones, D. (2016, April 1). *Disease alerts for winter beans after mild winter*. <https://www.fwi.co.uk/arable/pulses/disease-alert-winter-beans-mild-winter>
109. Kawamura, K., Mackay, A. D., Tuohy, M. P., Betteridge, K., Sanches, I. D., & Inoue, Y. (2011). Potential for spectral indices to remotely sense phosphorus and potassium content of legume-based pasture as a means of assessing soil phosphorus and potassium fertility status. *International Journal of Remote Sensing*, 32(1), 103-124. doi:10.1080/01431160903439908
110. Khare, S., Latifi, H., & Ghosh, S. K. (2018). Multi-scale assessment of invasive plant species diversity using Pléiades 1A, RapidEye and Landsat-8 data. *Geocarto international*, 33(7), 681-698.
111. Kislik, C., Dronova, I., & Kelly, M. (2018). UAVs in Support of Algal Bloom Research: A Review of Current Applications and Future Opportunities. *Drones*, 2(4). doi:10.3390/drones2040035
112. Konecny, G. (2014). *Geoinformation: Remote Sensing, Photogrammetry and Geographic Information Systems*, 2nd ed.; CRC Press, Taylor & Francis Group: Boca Raton: USA.
113. Kudenov, M.W., Jungwirth, M.E.L., Dereniak, E.L. & Gerhart, G.R. (2009). White light Sagnac interferometer for snapshot linear polarimetric imaging. *Optic Express*, 17(25), 22520-22534. doi:<https://doi.org/10.1364/OE.17.022520>
114. Kuehni, R. (2014). Ives, Frederic Eugene. In *Encyclopedia of Color Science and Technology* (pp. 1-3).
115. Kwak, G.-H., & Park, N.-W. (2019). Impact of Texture Information on Crop Classification with Machine Learning and UAV Images. *Applied Sciences*, 9(4). doi:10.3390/app9040643
116. Laghari, A. A., Jumani, A. K., Laghari, R. A., & Nawaz, H. (2023). Unmanned aerial vehicles: A review. *Cognitive Robotics*, 3, 8-22. doi:10.1016/j.cogr.2022.12.004.

117. Lagüela, S., Díaz, L., Roca, D., & Lorenzo, H. (2015). Aerial thermography from low-cost UAV for the generation of thermographic digital terrain models. *Opto-Electronics Review*, 23(1), 78-84.
118. Langfelder, G., Longoni, A. & Zaraga, F. (2009). The Transverse Field Detector: A CMOS active pixel sensor capable of on-line tuning of the spectral response. In *Proceedings of 2009 IEEE Sensors, Christchurch, New Zealand*, 1652–1657. doi:[0.1109/ICSENS.2009.5398512](https://doi.org/0.1109/ICSENS.2009.5398512)
119. Lantz, N. J., & Wang, J. (2013). Object-based classification of Worldview-2 imagery for mapping invasive common reed, *Phragmites australis*. *Canadian Journal of Remote Sensing*, 39(4), 328-340.
120. Lapray, P. J., Wang, X., Thomas, J. B., & Gouton, P. (2014). Multispectral filter arrays: recent advances and practical implementation. *Sensors (Basel)*, 14(11), 21626-21659. doi:[10.3390/s141121626](https://doi.org/10.3390/s141121626)
121. Lehmann, J. R. K., Prinz, T., Ziller, S. R., Thiele, J., Heringer, G., Meira-Neto, J. A. A., & Buttschardt, T. K. (2017). Open-Source Processing and Analysis of Aerial Imagery Acquired with a Low-Cost Unmanned Aerial System to Support Invasive Plant Management. *Frontiers in Environmental Science*, 5. doi:[10.3389/fenvs.2017.00044](https://doi.org/10.3389/fenvs.2017.00044)
122. Li, K.-Y., Sampaio de Lima, R., Burnside, N. G., Vahtmäe, E., Kutser, T., Sepp, K., . . . Sepp, K. (2022). Toward Automated Machine Learning-Based Hyperspectral Image Analysis in Crop Yield and Biomass Estimation. *Remote Sensing*, 14(5). doi:[10.3390/rs14051114](https://doi.org/10.3390/rs14051114)
123. Li, W., Niu, Z., Chen, H., Li, D., Wu, M., & Zhao, W. (2016). Remote estimation of canopy height and aboveground biomass of maize using high-resolution stereo images from a low-cost unmanned aerial vehicle system. *Ecological Indicators*, 67, 637-648. doi:[10.1016/j.ecolind.2016.03.036](https://doi.org/10.1016/j.ecolind.2016.03.036)
124. Li, Xinghua, Huanfeng Shen, Liangpei Zhang, Hongyan Zhang, Qiangqiang Yuan, & Gang Yang. (2014): "Recovering quantitative remote sensing products contaminated by thick clouds and shadows using multitemporal dictionary learning." *IEEE Transactions on Geoscience and Remote Sensing* 52, 11, 7086-7098.
125. Limagrain UK. (2019, May). *Tundra winter bean*. <https://www.arablescotland.org.uk/sites/www.arablescotland.org.uk/files/2020-06/Tundra%20Winter%20Bean%20Technical%20Summary.pdf>

126. Lishawa, S.C., Carson, B.D., Brandt, J.S., Tallant, J.M., Reo, N.J., Albert, D.A., Monks, A.M., Lautenbach, J.M. & Clark, E. (2017). Mechanical Harvesting Effectively Controls Young *Typha* spp. Invasion and Unmanned Aerial Vehicle Data Enhances Post-treatment Monitoring. *Frontier Plant Science*, 8.
127. Liu, C., Ai, M., Chen, Z., Zhou, Y., & Wu, H. (2018). Detection of *Firmiana danxiaensis* Canopies by a Customized Imaging System Mounted on an UAV Platform. *Journal of Sensors*, 2018, 1-12. doi:10.1155/2018/6869807
128. Liu, Y., Liu, S., Li, J., Guo, X., Wang, S., & Lu, J. (2019). Estimating biomass of winter oilseed rape using vegetation indices and texture metrics derived from UAV multispectral images. *Computers and Electronics in Agriculture*, 166. doi:10.1016/j.compag.2019.105026
129. Lopatin, J., Dolos, K., Kattenborn, T., & Fassnacht, F. E. (2019). How canopy shadow affects invasive plant species classification in high spatial resolution remote sensing. *Remote Sensing in Ecology and Conservation*, 5(4), 302-317.,
130. Lu, H., Fan, T., Ghimire, P., & Deng, L. (2020). Experimental Design Design Evaluation and Consistency Comparison of UAV Multispectral Minisensors. *Remote Sensing*, 12(16). doi:10.3390/rs12162542
131. Lu, J., Eitel, J. U. H., Engels, M., Zhu, J., Ma, Y., Liao, F., . . . Tian, Y. (2021). Improving Unmanned Aerial Vehicle (UAV) remote sensing of rice plant potassium accumulation by fusing spectral and textural information. *International Journal of Applied Earth Observation and Geoinformation*, 104. doi:10.1016/j.jag.2021.102592
132. 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. doi: 10.1016/j.tplants.2018.11.007.
133. Mafanya, M., Tsele, P., Botai, J., Manyama, P., Swart, B., & Monate, T. (2017). Evaluating pixel and object based image classification techniques for mapping plant invasions from UAV derived aerial imagery: *Harrisia pomanensis* as a case study. *ISPRS Journal of Photogrammetry and Remote Sensing*, 129, 1-11. doi:10.1016/j.isprsjprs.2017.04.009
134. Maimaitijiang, M., Sagan, V., Sidike, P., Hartling, S., Esposito, F., & Fritschi, F. B. (2020). Soybean yield prediction from UAV using multimodal data fusion and deep learning. *Remote Sensing of Environment*, 237. doi:10.1016/j.rse.2019.111599

135. Mamaghani, B., & Salvaggio, C. (2019). Multispectral Sensor Calibration and Characterization for sUAS Remote Sensing. *Sensors (Basel)*, 19(20). doi:10.3390/s19204453
136. Managing ancient and native woodland in England. (2010). *Forestry Commission (England)*. <https://www.forestresearch.gov.uk/research/managing-ancient-and-native-woodland-in-england/>.
137. Maronna, R. A., & Zamar, R. H. (2002). Robust estimates of location and dispersion for high-dimensional datasets. *Technometrics*, 44(4), 307-317.
138. Martin, F.M., Müllerová, J., Borgniet, L., Dommagnet, F., Breton, V. & Evette, A. (2018). Using Single- and Multi-Date UAV and Satellite Imagery to Accurately Monitor Invasive Knotweed Species. *Remote Sensing*, 10, 1662.
139. Mazur, P., Gozdowski, D., & Wójcik-Gront, E. (2022). Soil Electrical Conductivity and Satellite-Derived Vegetation Indices for Evaluation of Phosphorus, Potassium and Magnesium Content, pH, and Delineation of Within-Field Management Zones. *Agriculture*, 12(6). doi:10.3390/agriculture12060883
140. McDowell, R.W. (2017). Does variable rate irrigation decrease nutrient leaching losses from grazed dairy farming? *Soil Use Manag*, 33, 530–537. doi: [10.1111/sum.12363](https://doi.org/10.1111/sum.12363)
141. McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in econometrics* (pp. 104-142). New York: Academic Press.
142. McNabb, M. (2019, September 24). *Enterprise Scale: DJI introduces new P4 multispectral for precision agriculture and land management*. <https://dronelife.com/2019/09/24/enterprise-scale-dji-introduces-new-p4-multispectral-for-precision-agriculture-and-land-management/>
143. Mehmood, M., Shahzad, A., Zafar, B., Shabbir, A., Ali, N., & Ahmad, A. (2022). Remote Sensing Image Classification: A Comprehensive Review and Applications. *Mathematical Problems in Engineering*, 2022, 1-24. doi:10.1155/2022/5880959
144. Meier, U. (1997). *Growth stages of mono-and dicotyledonous plants*. Blackwell Wissenschafts-Verlag.
145. Mirik, M., Ansley, R.J., Steddom, K., Jones, D.C., Rush, C.M., Michels, G.J. & Elliott, N.C. (2013b). Remote Distinction of A Noxious Weed (Musk Thistle:

- CarduusNutans) Using Airborne Hyperspectral Imagery and the Support Vector Machine Classifier. *Remote Sensing*, 5, 612–630
146. Mirik, M., Chaudhuri, S., Surber, B., Ale, S., & James Ansley, R. (2013a). Detection of two intermixed invasive woody species using color infrared aerial imagery and the support vector machine classifier. *Journal of Applied Remote Sensing*, 7(1), 073588-073588.
 147. Mishra, N. B., Mainali, K. P., Shrestha, B. B., Radenz, J., & Karki, D. (2018). Species-level vegetation mapping in a Himalayan treeline ecotone using unmanned aerial system (UAS) imagery. *ISPRS International Journal of Geo-Information*, 7(11), 445.
 148. Mohsan, S. A. H., Khan, M. A., Noor, F., Ullah, I., & Alsharif, M. H. (2022). Towards the Unmanned Aerial Vehicles (UAVs): A Comprehensive Review. *Drones*, 6(6). doi:10.3390/drones6060147.
 149. Monteiro, A., Santos, S., & Gonçalves, P. (2021). Precision agriculture for crop and livestock farming—Brief review. *Animals*, 11(8), 2345. <https://doi.org/10.3390/ani11082345>
 150. Moore, G.K. (1979). What is a picture worth? A history of remote sensing. *Hydrological Sciences Bulletin*, 24, 2-10.
 151. Morley, P. J., Jump, A. S., West, M. D., & Donoghue, D. N. (2020). Spectral response of chlorophyll content during leaf senescence in European beech trees. *Environmental Research Communications*, 2(7), 071002.
 152. Motohka, T., Nasahara, K. N., Oguma, H., & Tsuchida, S. (2010). Applicability of Green-Red Vegetation Index for Remote Sensing of Vegetation Phenology. *Remote Sensing*, 2(10), 2369-2387. doi:10.3390/rs2102369
 153. 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>
 154. National statistics. (2022, October 21). *Chapter 7: Crops*. <https://www.gov.uk/government/statistics/agriculture-in-the-united-kingdom-2021/chapter-7-crops>
 155. Neale, C. M. U., Pölönen, I., Saari, H., Kaivosoja, J., Honkavaara, E., Pesonen, L., & Maltese, A. (2013). *Hyperspectral imaging based biomass and nitrogen content estimations from light-weight UAV*. Paper presented at the Remote Sensing for Agriculture, Ecosystems, and Hydrology XV.

156. Nebiker, S., Lack, N., Abächerli, M., & Läderach, S. (2016). Light-Weight Multispectral Uav Sensors and Their Capabilities for Predicting Grain Yield and Detecting Plant Diseases. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLI-B1*, 963-970. doi:10.5194/isprsarchives-XLI-B1-963-2016
157. Neina, D. (2019). The role of soil pH in plant nutrition and soil remediation. *Applied and environmental soil science*, 2019, 1-9. <https://doi.org/10.1155/2019/5794869>
158. Nevavuori, P., Narra, N., Linna, P., & Lipping, T. (2020). Crop Yield Prediction Using Multitemporal UAV Data and Spatio-Temporal Deep Learning Models. *Remote Sensing*, 12(23). doi:10.3390/rs12234000
159. Ng, W. T., Rima, P., Einzmann, K., Immitzer, M., Atzberger, C., & Eckert, S. (2017). Assessing the Potential of Sentinel-2 and Pléiades Data for the Detection of *Prosopis* and *Vachellia* spp. in Kenya. *Remote sensing*, 9(1), 74.
160. Nickerson. Tundra. <https://www.nickersonseeds.co.uk/products/pulses/tundra/data-sheet/>
161. Noguera, M., Aquino, A., Ponce, J. M., Cordeiro, A., Silvestre, J., Arias-Calderón, R., . . . Andújar, J. M. (2021). Nutritional status assessment of olive crops by means of the analysis and modelling of multispectral images taken with UAVs. *Biosystems Engineering*, 211, 1-18. doi:10.1016/j.biosystemseng.2021.08.035
162. Nolan, A. P., Park, S., O'Connell, M., Fuentes, S., Ryu, D., & Chung, H. (2015). Automated detection and segmentation of vine rows using high resolution UAS imagery in a commercial vineyard. In *International Congress on Modelling and Simulation 2015: Partnering with industry and the community for innovation and impact through modelling* (pp. 1406-1412). Modelling and Simulation Society of Australia and New Zealand Inc.(MSSANZ).
163. Note, M. (2011). Digital image basics. In *Managing Image Collections* (pp. 39-59).
164. Olson, D., & Anderson, J. (2021). Review on unmanned aerial vehicles, remote sensors, imagery processing, and their applications in agriculture. *Agronomy Journal*, 113(2), 971-992. doi:10.1002/agj2.20595
165. Oré, G., Alcântara, M. S., Góes, J. A., Oliveira, L. P., Yepes, J., Teruel, B., . . . Hernandez-Figueroa, H. E. (2020). Crop Growth Monitoring with Drone-Borne DInSAR. *Remote Sensing*, 12(4). doi:10.3390/rs12040615

166. Padmavathi, K., & Thangadurai, K. (2016). Implementation of RGB and Grayscale Images in Plant Leaves Disease Detection – Comparative Study. *Indian Journal of Science and Technology*, 9(6). doi:10.17485/ijst/2016/v9i6/77739
167. Pádua, L., Vanko, J., Hruška, J., Adão, T., Sousa, J. J., Peres, E., & Morais, R. (2017). UAS, sensors, and data processing in agroforestry: A review towards practical applications. *International journal of remote sensing*, 38(8-10), 2349-2391.
168. Pal, H. S., & Neifeld, M. A. (2003). Multispectral principal component imaging. *Optics Express*, 11(18), 2118-2125.
169. Palintest. (2020). *SKW 500 complete soil kit operation manual*. https://www.palintest.com/wp-content/uploads/2019/04/SKW500_User-Manual_Web.pdf
170. Papp, L., van Leeuwen, B., Szilassi, P., Tobak, Z., Szatmári, J., Árvai, M., . . . Pásztor, L. (2021). Monitoring Invasive Plant Species Using Hyperspectral Remote Sensing Data. *Land*, 10(1). doi:10.3390/land10010029
171. Partel, V., Kakarla, S.C., & Ampatzidis, Y. . (2016). Development and evaluation of a low-cost and smart technology for precision weed management utilizing artificial intelligence. *Comput. Electron. Agric*, 157, 339–350. doi: [10.1016/j.compag.2018.12.048](https://doi.org/10.1016/j.compag.2018.12.048)
172. Paudel, D., Dhungana, B., Caffè, M., & Krishnan, P. (2021). A Review of Health-Beneficial Properties of Oats. *Foods*, 10(11). doi:10.3390/foods10112591
173. Paul, E. A. (2016). The nature and dynamics of soil organic matter: Plant inputs, microbial transformations, and organic matter stabilization. *Soil Biology and Biochemistry*, 98, 109-126. doi:10.1016/j.soilbio.2016.04.001
174. Peng, Y., Zhu, T. e., Li, Y., Dai, C., Fang, S., Gong, Y., . . . Liu, K. (2019). Remote prediction of yield based on LAI estimation in oilseed rape under different planting methods and nitrogen fertilizerfertiliser applications. *Agricultural and Forest Meteorology*, 271, 116-125. doi:10.1016/j.agrformet.2019.02.032
175. Perroy, R.L., Sullivan, T. & Stephenson, N. (2017). Assessing the impacts of canopy openness and flight parameters on detecting a sub-canopy tropical invasive plant using a small unmanned aerial system. *ISPRS Journal of Photogrammetry of Remote Sensing*, 125, 174–183.
176. Pimstein, A., Karnieli, A., Bansal, S. K., & Bonfil, D. J. (2011). Exploring remotely sensed technologies for monitoring wheat potassium and phosphorus using

- field spectroscopy. *Field Crops Research*, 121(1), 125-135.
doi:10.1016/j.fcr.2010.12.001
177. Posea, P. (2020). *The complete history of drones (1898 – 2021)*.
<https://dronesgator.com/the-history-of-drones/>
178. Prey, L., & Schmidhalter, U. (2019). Sensitivity of Vegetation Indices for Estimating Vegetative N Status in Winter Wheat. *Sensors (Basel)*, 19(17).
doi:10.3390/s19173712
179. Priestclose wood. *Northumberland Wildlife Trust*.
<https://www.nwt.org.uk/nature-reserves/priestclose-wood>
180. Quaker. 2019. *Oat growth guide*.
<https://www.hutton.ac.uk/sites/default/files/files/publications/Oat-Growth-Guide.pdf>
181. Raghavendra, M., Sharma, M. P., Ramesh, A., Richa, A., Billore, S. D., & Verma, R. K. (2020). Soil Health Indicators: Methods and Applications. *Soil Analysis: Recent Trends and Applications*, 221-253.
182. Randox. (2023). *Reagent glucose*. <https://www.randox.com/glucose/>
183. Read, J.M. & Torrado, M. (2009) *International Encyclopedia of Human Geography*. Remote Sensing, Maxwell School of Syracuse University, Syracuse: NY,USA
184. Reza, M. N., Na, I. S., Baek, S. W., & Lee, K.-H. (2019). Rice yield estimation based on K-means clustering with graph-cut segmentation using low-altitude UAV images. *Biosystems Engineering*, 177, 109-121.
doi:10.1016/j.biosystemseng.2018.09.014
185. Rivero, R. G., Grunwald, S., Binford, M. W., & Osborne, T. Z. (2009). Integrating spectral indices into prediction models of soil phosphorus in a subtropical wetland. *Remote Sensing of Environment*, 113(11), 2389-2402.
doi:10.1016/j.rse.2009.07.015
186. Robinson, M. (2003b). The flying cowboy. *Kiting*. 25(3), 34-35.
187. Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. *NASA Spec. Publ*, 351(1), 309.
188. Rütting, T., Aronsson, H., & Delin, S. (2018). Efficient use of nitrogen in agriculture. *Nutrient Cycling in Agroecosystems*, 110, 1-5.
189. Safonova, A., Tabik, S., Alcaraz-Segura, D., Rubtsov, A., Maglinets, Y., & Herrera, F. (2019). Detection of Fir Trees (*Abies sibirica*) Damaged by the Bark Beetle

- in Unmanned Aerial Vehicle Images with Deep Learning. *Remote Sensing*, 11(6). doi:10.3390/rs11060643
190. Santo, R. D. E. (2012). Principal Component Analysis applied to digital image compression. *Einstein (São Paulo)*, 10, 135-139.
191. Santoso, H., Tani, H., & Wang, X. (2016). A simple method for detection and counting of oil palm trees using high-resolution multispectral satellite imagery. *International journal of remote sensing*, 37(21), 5122-5134.
192. Scharf, P. C., Shannon, D. K., Palm, H. L., Sudduth, K. A., & Drummond, S. T., Kitchen, N. R., Mueller, L. J., Hubbard, V. C., & Oliveira, L. F. (2011). Sensor-based nitrogen applications out-performed producer-chosen rates for corn in on-farm demonstrations. *Agron. J*, 103, 1684–1691. doi: [10.2134/agronj2011.0164](https://doi.org/10.2134/agronj2011.0164)
193. 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). doi:10.3390/rs8090706
194. 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. doi:10.1016/j.jag.2013.04.003
195. Scientific American. (1849) More About Balloons. *Scientific American*, 4, 205.
196. Searby, L. (2020, August 20). *Declining UK rapeseed production: not just a little crush issue*. <https://www.feednavigator.com/Article/2020/08/20/Declining-UK-rapeseed-production-not-just-a-little-crush-issue>
197. Senova. (2021, September). *Winter bean growers guide*. <https://irp.cdn-website.com/55da3d94/files/uploaded/Winter%20Bean%20Growers%20Guide%205th%20Edition%209.21.pdf>
198. Shahbandeh, M. (2023, Feb 9). Rapeseed oil usage worldwide 2015/16-2022/23. <https://www.statista.com/statistics/1086226/world-rapeseed-oil-usage-distribution/>
199. Shahi, T. B., Xu, C.-Y., Neupane, A., & Guo, W. (2022). Machine learning methods for precision agriculture with UAV imagery: a review. *Electronic Research Archive*, 30(12), 4277-4317. doi:10.3934/era.2022218
200. Shan, P. (2018). Image segmentation method based on K-mean algorithm. *EURASIP Journal on Image and Video Processing*, 2018(1), 1-9.

201. Sharma, P., Leigh, L., Chang, J., Maimaitijiang, M., & Caffè, M. (2022). Above-Ground Biomass Estimation in Oats Using UAV Remote Sensing and Machine Learning. *Sensors (Basel)*, 22(2). doi:10.3390/s22020601
202. Shirsath, P. B., Sehgal, V. K., & Aggarwal, P. K. (2020). Downscaling Regional Crop Yields to Local Scale Using Remote Sensing. *Agriculture*, 10(3). doi:10.3390/agriculture10030058
203. Shogenji, R., Kitamura, Y., Yamada, K., Miyatake, S. & Tanida, J. (2004). Multispectral imaging using compact compound optics. *Optic Express*, 12(8), 1643–1655. doi:[10.1364/opex.12.001643](https://doi.org/10.1364/opex.12.001643).
204. 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>
205. Skowronek, S., Ewald, M., Isermann, M., Van De Kerchove, R., Lenoir, J., Aerts, R., ... & Feilhauer, H. (2017). Mapping an invasive bryophyte species using hyperspectral remote sensing data. *Biological Invasions*, 19, 239-254.
206. Smith, T. J., & McKenna, C. M. (2013). A comparison of logistic regression pseudo R2 indices. *Multiple Linear Regression Viewpoints*, 39(2), 17-26.
207. 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). doi:10.3390/rs11101239
208. Speight, A. (2019, December 4). *A poor forecast for UK oilseed rape: Grain Market Daily*. <https://ahdb.org.uk/news/a-poor-forecast-for-uk-oilseed-rape-grain-market-daily>
209. Srestasathiern, P., & Rakwatin, P. (2014). Oil palm tree detection with high resolution multi-spectral satellite imagery. *Remote Sensing*, 6(10), 9749-9774.
210. Statista research department. (2023, Jan 30). *Agriculture industry in the UK - statistics & facts*. <https://www.statista.com/topics/4150/agriculture-industry-in-europe/#topicOverview>
211. Statista. (2022). *Volume of oats exported to non-EU countries from the United Kingdom (UK) from season 2002/03 to 2021/22*. <https://www.statista.com/statistics/299472/oat-uk-exports-non-eu-united-kingdom/>
212. Stepanov, A., Dubrovin, K., Sorokin, A., & Aseeva, T. (2020). Predicting Soybean Yield at the Regional Scale Using Remote Sensing and Climatic Data. *Remote Sensing*, 12(12). doi:10.3390/rs12121936

213. Stow, D., Nichol, C., Wade, T., Assmann, J., Simpson, G., & Helfter, C. (2019). Illumination Geometry and Flying Height Influence Surface Reflectance and NDVI Derived from Multispectral UAS Imagery. *Drones*, 3(3). doi:10.3390/drones3030055
214. Su, J., Liu, C., Coombes, M., Hu, X., Wang, C., Xu, X., . . . Chen, W.-H. (2018). Wheat yellow rust monitoring by learning from multispectral UAV aerial imagery. *Computers and Electronics in Agriculture*, 155, 157-166. doi:10.1016/j.compag.2018.10.017
215. Sulik, J. J., & Long, D. S. (2015). Spectral indices for yellow canola flowers. *International Journal of Remote Sensing*, 36(10), 2751-2765. doi:10.1080/01431161.2015.1047994
216. Sulik, J. J., & Long, D. S. (2016). Spectral considerations for modeling yield of canola. *Remote Sensing of Environment*, 184, 161-174. <https://doi.org/10.1016/j.rse.2016.06.016>
217. Suomalainen, J., Oliveira, R. A., Hakala, T., Koivumäki, N., Markelin, L., Näsi, R., & Honkavaara, E. (2021). Direct reflectance transformation methodology for drone-based hyperspectral imaging. *Remote Sensing of Environment*, 266. doi:10.1016/j.rse.2021.112691
218. Tang, W., Tang, R., Guo, T., & Wei, J. (2022). Remote Prediction of Oilseed Rape Yield via Gaofen-1 Images and a Crop Model. *Remote Sensing*, 14(9). doi:10.3390/rs14092041
219. Tao, H., Feng, H., Xu, L., Miao, M., Yang, G., Yang, X., & Fan, L. (2020). Estimation of the Yield and Plant Height of Winter Wheat Using UAV-Based Hyperspectral Images. *Sensors (Basel)*, 20(4). doi:10.3390/s20041231
220. Tao, H., Xu, S., Tian, Y., Li, Z., Ge, Y., Zhang, J., ... & Jin, S. (2022). Proximal and remote sensing in plant phenomics: Twenty years of progress, challenges and perspectives. *Plant Communications*, 100344.
221. Tarolli, P., Sofia, G., Calligaro, S., Prodocimi, M., Preti, F., & Dalla Fontana, G. (2015). Vineyards in Terraced Landscapes: New Opportunities from Lidar Data. *Land Degradation & Development*, 26(1), 92-102. doi:10.1002/ldr.2311
222. Tehrany, M. S., Kumar, L., & Drielsma, M. J. (2017). Review of native vegetation condition assessment concepts, methods and future trends. *Journal for Nature Conservation*, 40, 12-23. doi:10.1016/j.jnc.2017.08.004

223. Tian, H., Wang, Y., Chen, T., Zhang, L., & Qin, Y. (2021). Early-Season Mapping of Winter Crops Using Sentinel-2 Optical Imagery. *Remote Sensing*, *13*(19). doi:10.3390/rs13193822
224. Tobin, P. C. (2018). Managing invasive species. *F1000Res*, *7*. doi:10.12688/f1000research.15414.1
225. Tosh, S. M., & Miller, S. S. (2016). Oats. In *Encyclopedia of Food and Health* (pp. 119-125).
226. Trigunasih, N. M., & Saifulloh, M. (2022). Correlation Between Soil Nitrogen Content and NDVI Derived from Sentinel-2A Satellite Imagery. *Jurnal Lahan Suboptimal : Journal of Suboptimal Lands*, *11*(2), 112-119. doi:10.36706/jlso.11.2.2022.574
227. Tsouros, D. C., Bibi, S., & Sarigiannidis, P. G. (2019). A Review on UAV-Based Applications for Precision Agriculture. *Information*, *10*(11). doi:10.3390/info10110349
228. Verhoeven, G. (2011). Taking computer vision aloft - archaeological three-dimensional reconstructions from aerial photographs with photostan. *Archaeological Prospection*, *18*(1), 67-73. doi:10.1002/arp.399
229. Vila, M., Espinar, J. L., Hejda, M., Hulme, P. E., Jarosik, V., Maron, J. L., . . . Pysek, P. (2011). Ecological impacts of invasive alien plants: a meta-analysis of their effects on species, communities and ecosystems. *Ecol Lett*, *14*(7), 702-708. doi:10.1111/j.1461-0248.2011.01628.x
230. Vizzari, M., Santaga, F., & Benincasa, P. (2019). Sentinel 2-Based Nitrogen VRT Fertilization in Wheat: Comparison between Traditional and Simple Precision Practices. *Agronomy*, *9*(6). doi:10.3390/agronomy9060278
231. von Bueren, S. K., Burkart, A., Hueni, A., Rascher, U., Tuohy, M. P., & Yule, I. J. (2015). Deploying four optical UAV-based sensors over grassland: challenges and limitations. *Biogeosciences*, *12*(1), 163-175. doi:10.5194/bg-12-163-2015
- a. Vrchota, J., Pech, M., & Švepešová, I. (2022). Precision Agriculture Technologies for Crop and Livestock Production in the Czech Republic. *Agriculture*, *12*(8), 1080. <https://doi.org/10.3390/agriculture12081080>
232. Wallace, L., Lucieer, A., Watson, C., & Turner, D. (2012). Development of a UAV-LiDAR system with application to forest inventory. *Remote sensing*, *4*(6), 1519-1543.

233. Walsh, O. S., Shafian, S., Marshall, J. M., Jackson, C., McClintick-Chess, J. R., Blanscet, S. M., . . . Walsh, W. L. (2018). Assessment of UAV Based Vegetation Indices for Nitrogen Concentration Estimation in Spring Wheat. *Advances in Remote Sensing*, *07*(02), 71-90. doi:10.4236/ars.2018.72006
234. Wan, L., Li, Y., Cen, H., Zhu, J., Yin, W., Wu, W., . . . He, Y. (2018). Combining UAV-Based Vegetation Indices and Image Classification to Estimate Flower Number in Oilseed Rape. *Remote Sensing*, *10*(9). doi:10.3390/rs10091484
235. Wang, J., Shi, T., Liu, H., & Wu, G. (2016). Successive projections algorithm-based three-band vegetation index for foliar phosphorus estimation. *Ecological Indicators*, *67*, 12-20. doi:10.1016/j.ecolind.2016.02.033
236. Węglarczyk, S. (2018). Kernel density estimation and its application. In *ITM Web of Conferences* (Vol. 23, p. 00037). EDP Sciences. <https://doi.org/10.1051/itmconf/20182300037>
237. Wójcik-Gront, E., Gozdowski, D., & Stępień, W. (2022). UAV-Derived Spectral Indices for the Evaluation of the Condition of Rye in Long-Term Field Experiments. *Agriculture*, *12*(10), 1671. <https://doi.org/10.3390/agriculture12101671>
238. Wolf, P.R. & Dewitt, B.A. (2016). *Elements of photogrammetry: with application is GIS*. McGraw-Hill,2000.
239. Woodland Trust. *Ancient Woodland*. <https://www.woodlandtrust.org.uk/trees-woods-and-wildlife/habitats/ancient-woodland/>.
240. Wu, D., Li, R., Zhang, F., & Liu, J. (2019). A review on drone-based harmful algae blooms monitoring. *Environ Monit Assess*, *191*(4), 211. doi:10.1007/s10661-019-7365-8.
241. Wu, H., Wiesner-Hanks, T., Stewart, E. L., DeChant, C., Kaczmar, N., Gore, M. A., . . . Lipson, H. (2019). Autonomous Detection of Plant Disease Symptoms Directly from Aerial Imagery. *The Plant Phenome Journal*, *2*(1), 1-9. doi:10.2135/tppj2019.03.0006
242. Xie, T., Li, J., Yang, C., Jiang, Z., Chen, Y., Guo, L., & Zhang, J. (2021). Crop height estimation based on UAV images: Methods, errors, and strategies. *Computers and Electronics in Agriculture*, *185*. doi:10.1016/j.compag.2021.106155
243. Xu, X., Du, X., Wang, F., Sha, J., Chen, Q., Tian, G., ... & Jiang, Y. (2020). Effects of potassium levels on plant growth, accumulation and distribution of carbon, and nitrate metabolism in apple dwarf rootstock seedlings. *Frontiers in Plant Science*, *11*, 904.

244. Xue, J., & Su, B. (2017). Significant remote sensing vegetation indices: A review of developments and applications. *Journal of sensors*, 2017.
245. Yaney-Keller, A., Santidrian Tomillo, P., Marshall, J. M., & Paladino, F. V. (2019). Using Unmanned Aerial Systems (UAS) to assay mangrove estuaries on the Pacific coast of Costa Rica. *PLoS One*, 14(6), e0217310. doi:10.1371/journal.pone.0217310
246. Yang, G., Liu, J., Zhao, C., Li, Z., Huang, Y., Yu, H., . . . Yang, H. (2017). Unmanned Aerial Vehicle Remote Sensing for Field-Based Crop Phenotyping: Current Status and Perspectives. *Front Plant Sci*, 8, 1111. doi:10.3389/fpls.2017.01111
247. Zadoks, J. C., Chang, T. T., & Konzak, C. F. (1974). A decimal code for the growth stages of cereals. *Weed research*, 14(6), 415-421.
248. Zahorcak, M. (2007). *The Focal Encyclopedia of Photography*, 4th edn, Michael Peres: Burlington.
249. Zang, Y., Chen, X., Chen, J., Tian, Y., Shi, Y., Cao, X., & Cui, X. (2020). Remote Sensing Index for Mapping Canola Flowers Using MODIS Data. *Remote Sensing*, 12(23). doi:10.3390/rs12233912
250. Zhang(a), T., Vail, S., Duddu, H. S. N., Parkin, I. A. P., Guo, X., Johnson, E. N., & Shirtliffe, S. J. (2021). Phenotyping Flowering in Canola (*Brassica napus* L.) and Estimating Seed Yield Using an Unmanned Aerial Vehicle-Based Imagery. *Front Plant Sci*, 12, 686332. doi:10.3389/fpls.2021.686332
251. Zhang(b), C., Craine, W. A., McGee, R. J., Vandemark, G. J., Davis, J. B., Brown, J., . . . Sankaran, S. (2021). High-throughput phenotyping of canopy height in cool-season crops using sensing techniques. *Agronomy Journal*, 113(4), 3269-3280. doi:10.1002/agj2.20632
252. Zheng, C., Abd-Elrahman, A., Whitaker, V., & Dalid, C. (2022). Prediction of Strawberry Dry Biomass from UAV Multispectral Imagery Using Multiple Machine Learning Methods. *Remote Sensing*, 14(18). doi:10.3390/rs14184511
253. Zhou, J., Zhou, J., Ye, H., Ali, M. L., Chen, P., & Nguyen, H. T. (2021). Yield estimation of soybean breeding lines under drought stress using unmanned aerial vehicle-based imagery and convolutional neural network. *Biosystems Engineering*, 204, 90-103. doi:10.1016/j.biosystemseng.2021.01.017
254. Zhou, X., Zheng, H. B., Xu, X. Q., He, J. Y., Ge, X. K., Yao, X., . . . Tian, Y. C. (2017). Predicting grain yield in rice using multi-temporal vegetation indices from

UAV-based multispectral and digital imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 130, 246-255. doi:10.1016/j.isprsjprs.2017.05.003

255. Zhu, L., Suomalainen, J., Liu, J., Hyypä, J., Kaartinen, H., & Haggren, H. (2018). A Review: Remote Sensing Sensors. In *Multi-purposeful Application of Geospatial Data*.
256. Zou, X., Liang, A., Wu, B., Su, J., Zheng, R., & Li, J. (2019). UAV-Based High-Throughput Approach for Fast Growing *Cunninghamia lanceolata* (Lamb.) Cultivar Screening by Machine Learning. *Forests*, 10(9). doi:10.3390/f10090815