Advancing Crop Monitoring and Disease Detection in Potatoes (*Solanum tuberosum* L.) Through High-Throughput Phenotyping Utilizing Unmanned Aerial Vehicle and Remote Sensing Technology

Thesis submitted for the degree of Doctor of Philosophy by

Phatchareeya Waiphara

School of Natural and Environmental Sciences

Newcastle University, Newcastle-upon-Tyne

United Kingdom

March 2024

ABSTRACT

Due to increasing food demands and changing dietary preferences, the potato (*Solanum tuberosum* L.) has emerged as a key crop in various regions. This thesis, titled "Advancing Crop Monitoring and Disease Detection in Potatoes (*Solanum tuberosum* L.) through High-Throughput Phenotyping Utilizing Unmanned Aerial Vehicles and Remote Sensing Technology," introduces an innovative approach to address the intricacies of potato crop physiology and phenotyping through the integration of unmanned aerial vehicles (UAVs), remote sensing, and machine learning technologies.

This study examined the interaction among genotype, environmental factors, crop physiology, and yield prediction using high-throughput phenotyping and UAV-based remote sensing across various potato varieties over three growing seasons (2020-2022). The primary objective is to enhance our understanding of potato growth variations and how different potato varieties respond to and adapt to different field conditions. Multispectral imaging data were analysed using structure-from-motion (SfM) algorithms to generate canopy height, ground cover, and vegetation indices. These canopy parameters were then used to establish the relationship between UAV-based data and proximal data, providing the foundation for developing a predictive model for both yield estimation and disease detection. A strong correlation was observed between canopy height obtained from UAV and proximal measurement (R^2=0.93). UAV data collection during the tuber initiation stage (UAV flight 2) showed the strongest correlation with ground-based measurements. As potato plants reached the tuber bulking stage (UAV flight 3), the correlation remained strong but became slightly less pronounced. However, as the plant reached the maturity stage, the correlation decreased dramatically. The results emphasise the critical importance of selecting appropriate approaches and timing for data collection to enhance the efficacy of plant phenotyping and genotyping studies.

Maturation and yield predictive models have incorporated various crop parameters and machine learning techniques to improve predictive accuracy. The Random Forest (RF) approach demonstrated promising results, achieving an R^2 of 78.31 in 2022 by using canopy volume, canopy area, canopy height, NDVI, and NDRE extracted from UAV flight 2 (tuber initiation stage). However, yield prediction accuracy varied across seasons due to environmental variability. Furthermore, the evaluation of potato maturity showed that the RF model outperformed the Partial Least Square Regression (PLSR) and Decision Tree models.

In terms of tuber quality, this study investigated the accumulation of potato glycoalkaloids (PGAs) in relation to greening phenomena in potato tubers. Analysis from nine potato varieties showed significant variations in total PGA concentrations across different greening scores (0-5 scale), with the ‘Craigs Royal’ variety demonstrating the highest concentration (2635 ± 638 mg/kg FW) at a greening score of 5. Concerningly, some varieties, such as ‘Dundrod’ and ‘Anna’, also exceed the food safety limit (200 mg/kg FW) even at low greening levels. In contrast, non-green tubers remained within the safety limits (21.8-189.5 mg/kg FW).

In summary, this research not only advances the domain of potato phenotyping and disease detection but also enhances the broader agricultural sector through the development of the pipeline for enhancing crop monitoring and disease detection through innovative image sensing and data analysis technologies. The thesis showcases the potential of UAV-based remote sensing for crop monitoring and enhancing disease detection and yield prediction by exploiting spectral responses and canopy characteristics obtained through high-throughput techniques. The integration of UAV-based phenotyping with machine learning methodologies represents a leap forward in our capacity to monitor crop health and boost productivity.

DECLARATION

I hereby declare that the work presented in this thesis is my own original research, with the exception of some of the unmanned aerial vehicle flights and subsequent image stitching and the point cloud construction that were performed by Jock Souter (Survey Solution Scotland) as part of a research collaboration, all subsequent analysis and interpretation of the data were my own work.

I certify that thesis is my own work and that it has not been submitted previously to this or any other University. Where other sources of information have been used, they have been acknowledged.

ACKNOWLEDGEMENT

I would like to express my gratitude to the individuals and organisations who contributed to this research. Without their assistance, the field campaign would have been significantly more challenging. First and foremost, I extend my sincere appreciation to my supervisor, Dr Ankush Prashar for his invaluable guidance and support throughput this endeavour. Special thanks to Jock Souter from Survey Solutions Scotland for his expertise in conducting the fixed-wing UAV flights, which greatly enriched the data collection process. I am also grateful to Gavin Hall, Racheal Chapman, and Michale Botha for their agronomic insights and technical support, which provide instrumental in navigating various challenges encountered. I extend my appreciation to Dr Avinash Agarwal and Dr Enas Sufar for their invaluable contributed to the collection of ground truth measurement throughout the growing season. To my family and friends for their support and encouragement throughout this journey, I offer my heartfelt gratitude. I would like to acknowledge all Thai taxpayers, without whom the government would not be able to fund my PhD and this research. Your contributions have facilitated the advancement of scientific research in agriculture. Lastly, I would like to thank Me.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

3D Model – Three-Dimensional Model

AHDB – Agriculture Horticulture Development Board

ANOVA – Analysis of Variance

BGI – Blue Green Pigment Index

BI – Bare Soil Index

CHM – Canopy Height Model

CRP – Calibrated Reflectance Panel

CVI Chlorophyll Vegetation Index

DAP – Day After Planting

DAPC – Discriminant Analysis of Principle Components

DCA –Discriminant Component Analysis

DMF – N, N – Dimethylformamide

DN – Digital Number

DSM – Digital Surface Model

DTM – Digital Terrain Model

EVI – Enhance Vegetation Index

FW – Fresh Weight

GA – Glycoalkaloid

GIS – Geographical Information System

GLI – Green Leaf Index

GNDVI – Green Normalised Difference Vegetation Index

GPC – Ground Point Control

GPS – Global Positioning System

HPLC – High-performance Liquid Chromatography

HTFP – High-throughput Field Phenotyping

HTP – High-throughput Phenotyping

Kg – Kilogram

LAI – Leaf Area Index

MAE – Mean Absolute Error

MLR – Multiple Linear Regression

MSE – Mean Square Error

NDRE – Normalised Difference Red Edge

NDVI – Normalised Difference Vegetation Index

NIR – Near-Infrared

PCA – Principal Component Analysis

PGA – Potato Glycoalkaloid

PLRV – Potato Leaf Roll Virus

PLSR – Partial Least Square Regression

PVY – Potato Virus Y

QGIS – Quantum Geographic Information System

RF – Random Forest

RGB – Red, Green, Blue

SAVI – Soil-adjusted Vegetation Index

SfM – Structure from Motion

RMSE – Root Mean Square Error

R^2 – R square

RS – Remote sensing

RVI – Ratio Vegetation Index

SV – Stereo Vision

TBC – Trimble Business Centre

TVI – Triangular Vegetation Index

TIFF – Tagged Image Field Format

UV/VIS – Ultraviolet/Visible

VARI – Visible Atmospherically Resistant Index

VI – Vegetation Index

UAV – Unmanned Aerial Vehicle

CHAPTER 1. INTRODUCTION AND LITERATURE REVIEW

1.1 GENERAL INTRODUCTION

Potato (*Solanum tuberosum* L.) is the most important non-grain food commodity in the world (Campos, 2020). Due to an increasing global population and changing dietary preferences, the potato has become a key crop in many regions to meet rising food demands. Commercial potato production is vital in ensuring global food security by providing nourishment, generating revenue, and creating employments (Lutaladio and Castaldi, 2009; Devaux *et al.*, 2021; Adekanmbi *et al.*, 2023). Trends in climate change and extreme weather conditions significantly affect potato crop production, making it essential to develop new potato varieties and planting strategies to mitigate the impacts of climate variability (Adekanmbi *et al.*, 2023). Utilising advanced techniques and methodologies to evaluate crop performance is essential to address these challenges and improve potato productivity.

Field-based phenotyping (FBP) involves the utilisation of tools and methodologies to evaluate crop performance, study environmental interactions, and analyse specific traits such as crop performance and development through the use of vegetation indices (White *et al.*, 2012; Aakash *et al.*, 2019). The challenges in FBP originated from the need to improve the efficiency and accuracy of plant breeding while concurrently monitoring crop performance to mitigate potential economic losses and ensure food security. FBP serves as a critical tool for both plant breeding and precision agriculture, albeit with distinct requirements contingent upon field size, plot numbers, and the trait of interest. In plant breeding, FBP facilitates the assessment of diverse crop varieties across varying environmental conditions, often necessitating the subdivision of fields into several small plots to evaluate a myriad of traits with sufficient replication. This data serves as a foundation for selecting superior varieties tailored for breeding purposes, leading to the development of crops with desirable traits and characteristics. Conversely, in precision agriculture, FBP assumes a pivotal role in monitoring crop health and productivity across expansive fields. Exploiting remote sensing technologies like unmanned aerial vehicles (UAVs) or satellite imagery, FBP gathers critical data on plant growth, nutrient levels, weed populations, and pest or disease infestations. This data informs judicious decision-making about field interventions such as irrigation, pest controls, and fertilisation applications, ultimately optimising crop yields and minimising resource wastages. Historically, field inspectors, farmers, breeders, and agronomists have manually carried out field phenotyping and plant monitoring. Challenges associated with conventional ways of field-based

phenotyping include environmental variability, which can impede the accuracy of the assessment of crop performance, while the conventional methods are effective for qualitative traits but time-consuming, labour-intensive, and prone to limitations such as human error or restricted sensor use (Campos *et al.*, 2019). Thus, there is a need for a shift towards more flexible, rapid, accurate, and robust crop monitoring techniques to address these challenges and advance agriculture to meet the demands of modern agriculture.

Remote sensing is a valuable tool in agricultural technology, with various advanced high-throughput plant phenotyping (HTP) approaches demonstrating significant achievements. However, the predominant focus of HTPs on controlled environment systems, which restricts their applicability, leaving FBP as a critical bottleneck in crop breeding due to its adoptability to diverse environments (David *et al.*, 2014; Virlet *et al.*, 2016). To overcome this challenge, high-throughput field phenotyping with a remote sensing platform (HTFP-RS) has emerged as an alternative solution, offering rapid and non-invasive measurements of crop traits to address the limitations of traditional methods. This approach offers a robust way to evaluate and comprehend crop variability, facilitated by non-destructive growth measurements and the identification of critical growth stages for dataset collection with high spatial and temporal resolutions. For instance, multispectral data acquired from unmanned aerial vehicles (UAVs) unveil field variability invisible to the human eye, thereby facilitating early disease detection and empowering decision-making to enhance crop productivity.

Modern high-throughput field phenotyping has broadened its focus from initial commercially relevant characteristics to complex physiological traits linked to crop heath and performance, reflecting a shift towards more comprehensive trait assessments. Although many approaches have been developed to translate remote sensing data in the field of phenotyping, they are still under rapid development, with the main aim to reduce the cost and complexity of data handling and management. This literature review aims to highlight the progression in HTFP in potato production using remote sensing methods, with a focus on the evolution of techniques and their impact on research in this field. Furthermore, the literature review will examine the challenges and limitations associated with using high-throughput field phenotyping tools in potato production. However, while high-throughput field phenotyping has demonstrated potential in research settings, it is essential to test and validate these tools to ensure their practicality in commercial potato production.

Considering the significance of potatoes as a staple crop, their sensitivity to environmental variations, their complex growth stages, coupled with their susceptible to both biotic and abiotic stresses, make them suitable for evaluating and validating the efficacy of remote sensing technology in improving crop productivity. The following sections will explore potato taxonomy and growth stages to enhance the understanding of the applicability of these phenotyping methodologies.

1.2 POTATO TAXONOMY

Potato taxonomy refers to the systematic classification of potatoes within the plant kingdom. As a member of the nightshade family (Solanaceae), potatoes share botanical lineage with other crucial food crops such as tomatoes, eggplants, and peppers (Datiles and Acevedo-Rodríguez, 2022). Within the Solanaceae family, potatoes are classified under the Solanum genus, specifically in the section Petota. The species name *Solanum tuberosum* L. denotes the widely cultivated potato species prevalent across Europe and many other regions worldwide (Machida-Hirano, 2015). Originating from the Bolivian-Peruvian Andes region, potatoes have undergone extensive domestication, resulting in the proliferation of both cultivated and wild varieties (Britannica, 2022).

The taxonomic classification of potatoes has undergone several revisions, leading to ongoing debate and controversy among botanists and taxonomists regarding the precise categorization of different potato varieties (Gopal and Khurana, 2006). S. *tuberosum* is a complex species with diploid, triploid, and tetraploid variations, with tetraploid being the most important commercially (Datiles and Acevedo-Rodríguez, 2022). Traditionally, potatoes have been divided into two main groups: the Andigena group and the Chilotanum group. Broadly, the taxonomy of cultivated potatoes follows the delineation as below:

Domain: Eukaryota

Kingdom: Plantae

Phylum: Spermatophyta

Subphylum: Angiospermae

Class: Dicotyledonae

Order: Solanales

Family Solanaceae

Genus: Solanum

Species: *Solanum tuberosum*

Physical description

The potato, which is one of approximately 150 tuber-bearing species in the *Solanum* genus, has a tuber that is the swollen end of an underground stem. The first description of potatoes was given in Herball, or General Histories of Plantes (1597) by Gerard (1995). The tuber's size, shape, and colour vary widely depending on the variety, with some being round or oblong, ranging from small to large. The skin can be smooth or rough, and the flesh can be white, yellow, or purple. Potato plants can grow up to 1.5 metres (5 feet) tall, and the underground tubers can be found at varied depths, ranging from just beneath the soil surface to several feet deep in some varieties. The compound leaves are spirally arranged; each leaf is 20-30 cm (approximately 8-12 inches) long and consists of a terminal leaflet and two to four pairs of leaflets. The flower colour can be white, lavender, or purple with five united petals and yellow stamens. Potatoes also produce a fruit that contains viable seeds, and each seed can be used in propagation to potentially produce a new variety (Dean, 2018; Thornton, 2020).

1.3 POTATO GROWTH AND DEVELOPMENT

Potato tubers are modified stems that serve as the vegetative "seed" (asexual reproduction for reproducing a specific variety) (Bonierbale *et al.*, 2020). In order to ensure the best possible growth and development of potatoes, it is crucial to consider the plant shallow root system, and the type of cultivation method used, such as ridges or beds. Deep cultivation can aid in drainage and make harvesting easier, especially when growing on a large scale (Jackson *et al.*, 1997). Ideal conditions for potato growth include a soil temperature of 15-21 degrees Celsius with adequate moisture and nutrients as well as appropriate lighting conditions (day length) (Dean, 2018; Garcia-Gonzalez *et al.*, 2022). It is also important for growers to understand the different growth stages of potatoes in order to monitor and make informed decisions throughout the growing season.

The growth stages of a potato plant can be separated into five phenological stages, including sprout development, plant/vegetative establishment, tuber initiation, tuber bulking, and maturation (Griffin *et al.*, 1993; Thornton, 2020) (**Figure 1**).

- Sprout development/initial: The initial stage of potato development involves the emergence of sprouts from the seed's eyes. The emergence typically occurs around 2-4 weeks after planting, with soil temperature and moisture playing a crucial role during this period. At this stage, the seed serves as a source of energy for initial development. Unlike other major food crops, the potato is propagated vegetative using whole or cut tubers (known as seed pieces) rather than true seeds (Thornton, 2020).
- Vegetation establishment: The vegetative growth is distinguished by rapid leaves, branches, roots, and stolon formation. This stage begins when the sprouts emerge from the soil and lasts until the tuber initiation stage (from 30-70 days, depending on the planting date, soil temperature, and other environmental conditions) (Haverkort, 1990). During this stage, the plant concentrates on growing a strong root system and foliage to maximise photosynthesis.
- Tuber initiation: The tuber initiation stage marks the transition from vegetative to reproductive development, which could take around 6-8 weeks after planting depends on day length and temperature (Griffin *et al.*, 1993). During this stage, the plant starts producing tubers at the end of the stolon but not enlarging. The plant may produce 20-30 little tubers about the size of 10 mm, but only 5-15 tubers normally achieve maturity.
- Tuber filling: Tuber filling or bulking is the stage when the tuber develops in size. The leaf area of the canopy affects its capacity to capture radiation and tuber growth rate during this period (Ingram and McCloud, 1984). The number and size of tubers that reach this stage are proportional to the amount of energy that plant produces, as they require starch and other nutrients (Griffin *et al.*, 1993). Variable weather and moisture levels are crucial for even tuber set and growth. This has a significant impact on the final yield (Zebarth *et al.*, 2022). This stage lasts until the above ground starts to die (45-60 days).
- Maturation and senescence: Maturation and senescence occur when the potato has reached its maximum size, weight, and dry matter content and is nearly ready to harvest. The main visual indicator that the plant has entered this stage is when the haulm/vines (above ground vegetation) have turned yellow and lost their leaves (Heltoft *et al.*, 2017). While some varieties may naturally undergo senescence (yellowing and defoliation of the vines), in some cases, farmers may need to induce vine death to initiate the maturation process. This can be achieved through herbicide application or flailing. This process allows the plant to transfer nutrients from the foliage to the tubers,

allowing it to continue growing and reach maturity before harvest (Schnabel *et al.*, 2010).

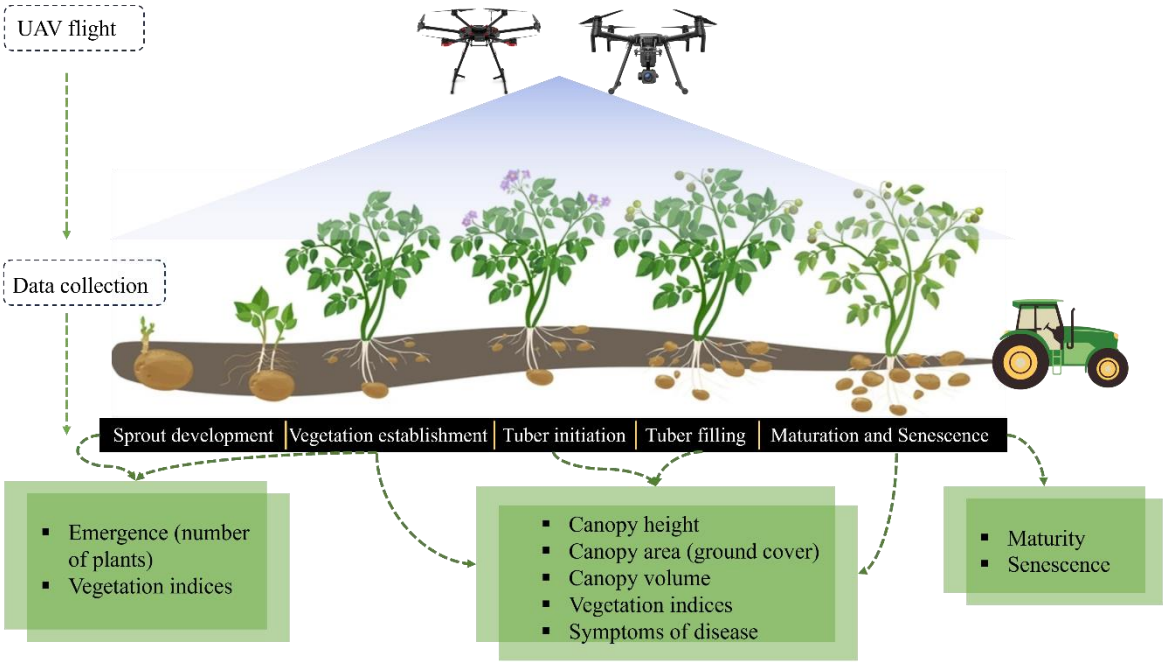

Figure 1 From seed to harvest, modern agriculture's symbiosis with technology. Utilising the use of unmanned aerial vehicles to capture the potato growth cycle.

Understanding these growth stages is crucial for potato production to monitor plant performance, make informed decisions, and optimise yield and quality throughout the growing season. Especially the tuber initiation and maturity stages, as these stages have a significant impact on the final yield. In addition, mature potato tubers are better to store compared to immature tubers, as they are less susceptible to injury and disease and have low respiration rates (Heltoft *et al.*, 2017). This contributes to longer a shelf-life and greater resistance to pathogens, thereby reducing economic loss (Thornton, 2020).

1.4 POTATO PRODUCTION

Potato (*Solanum tuberosum* L.) is one of the most essential food commodities in the world, with the production of more than 380 million metric tonnes annually. The potato was brought to Europe by Spaniards around 400 years ago and progressively replaced cereals in northern Europe (Haverkort *et al.*, 2009). Nowadays, potatoes are widely cultivated in over 150 countries across the world, where they are serving as a major staple food and a table vegetable. One-third of the world's potatoes are produced in China and India (FAO (Food and Agriculture Organisation), 2022). According to AHDB, potato production in the UK for 2020

is estimated at around 5.37 million metric tonnes and is ranked the seventh largest in the world since 2010. Potatoes are an essential crop, not only an important source of nutrients for humans but also serving as a feedstock and alcohol for industrial products (Bradshaw and Bonierbale, 2010; Koch *et al.*, 2020). With over 5000 varieties, potatoes have a high level of variation in tuber colour and shape as a result of their widespread cultivation throughout the world (Lutaladio and Castaldi, 2009; Datiles and Acevedo-Rodríguez, 2022). Potato cultivation ranges from small to large scale worldwide, however, in temperate zones, they are primarily grown in monocultures.

The United Kingdom (UK) is among the top five potato producers in Europe, benefiting from highly favourable soil and climate conditions for potato growth in this region (Goffart *et al.*, 2022). In the UK, potato production is growing throughout various locations. Two primary locations account for 46 percent of the UK's potato cultivation area: the East of England and Scotland, which produce the majority of UK's seed potatoes (Goffart *et al.*, 2022). The potato planting season starts from March to May and is influenced by factors such as soil type, variety, and region (each with different climate conditions). More than 170 potato cultivars are commercially produced and classified into 'early' and 'main crop' depending on their planting and harvesting dates (Daccache *et al.*, 2012). The term 'early' refers to all potatoes that are typically planted in mid-March (southern England) or early April (northern England) and are ready to harvest approximately 10-15 weeks from the planting date. 'Main crop' potatoes are sown in the first half of April (southern England) or late April or early May (northern England) and harvested after 15-20 weeks (Daccache *et al.*, 2012).

The growing season for potatoes varies depending on a range of factors, including variety, geographical area, and climate conditions that determine the optimal timing for planting and harvesting. For example, highland areas in Bolivia, Peru, and Mexico are favoured as summer crops, whereas China, South America, East Africa, and lowland subtropics produce potatoes all year round (Bradshaw and Bonierbale, 2010). Potatoes are often grown throughout spring and autumn in Mediterranean regions, while they are grown in the summer in lowland temperate areas such as North America, Eastern Europe, Australia, and New Zealand (Bradshaw and Bonierbale, 2010). When deciding on the planting time, it is essential to consider not only the climate conditions but also the duration of the growing season to ensure optimal growth and yield. Weather conditions play a crucial role in determining the planting date, particularly the early cultivars that have a shorter growing season, which can lead to a lower yield compared to the main crop cultivars (Garcia-Gonzalez *et al.*, 2022). Effective

management practices, including crop rotation, soil management, and pest and disease management, are essential for ensuring a successful potato crop during the growing season.

1.4.1 ORGANIC PRODUCTION

Conventional farming methods are very productive but are now considered as harmful to the environments and unsustainable (Boschiero *et al.*, 2023; Kumar *et al.*, 2023). In order to achieve international and European sustainability growth, an alternative farming method such as organic production, has been a primary strategy. Organic farming in Europe has experienced significant expansion, but this varies across different countries (Acs *et al.*, 2005; Goffart *et al.*, 2022). The potato production industry in the United Kingdom (UK) is well established and successful, as indicate by the significant increase in market value (McGregor, 2007). However, one major challenge facing the potato industry in this region is the need to maintain soil quality, nutrient management, and the increasing restriction on the use of chemical pesticides and herbicides (Goffart *et al.*, 2022). Organic farming of potatoes in the UK is defined by the absence of artificial pesticides, fertilisers, and genetically modified organisms (GMOs) (Barber-Rowe, 2021).

However, organic production can present some challenges. Although organic production provides environmental benefits, research indicates that when produced on a large-scale, it can be more expensive and results in lower yield compared to conventional production due to the higher costs associated with disease and nutrient management (Hagman *et al.*, 2009; Kazimierczak *et al.*, 2019). For instance, Maggio *et al.* (2008) found that organic farming led to a 25% reduction in marketable yield compared to conventional farming, with a higher percentage of larger tubers. Seufert *et al.* (2012) found that organic yields are typically lower than conventional yields by 5 to 34 % depending on system and site characteristics. Whereas, Brazinskiene *et al.* (2014) and Kazimierczak *et al.* (2019) found that organic potato yield are on average 50% lower than conventional yields and vary within the variety. These findings indicate that organic farming may have some constraints leading to yield reduction, especially when the use of chemical pesticides and fertiliser is limited by organic regulations.

To address these challenges, organic farmers need to adopt more sustainable methods that can promote soil health through the use of organic fertiliser, crop rotation, reduced tillage, and real-time crop management, to manage disease and nutrients effectively (Watson *et al.*, 2002; Gattinger *et al.*, 2012). It is also advisable to use early to moderately maturing cultivars that have a high tolerance to late blight to minimise yield losses (Pawelzik and Möller, 2014).

Although there are various existing potato varieties that are suitable for organic production, there is a demand for new varieties that produce high yields with low inputs (Pandey *et al.*, 2021). Selecting cultivars that are more suitable for organic production is an effective strategy to overcome some of the challenges in organic potato production. Organic potato varieties are often bred to have better resistance to diseases and pests which might replace susceptible varieties which are currently grown in organic farming (Speiser *et al.*, 2006). These cultivars may also exhibit improved nutrient uptake efficiency, which can help to improve yields (Pandey *et al.*, 2021).

Since most of the available varieties have been developed for conventional farming, conventional breeding often overlooks several key traits essential for organic farming such as resistance to soil pathogens and seedborne diseases, nutrient efficiency, good weed suppression, and qualitative traits (Hagman *et al.*, 2009). Therefore, selecting potato varieties that are well-suited to local environmental conditions and resistant to disease is crucial for increasing the competitiveness of the organic sector (Meglic *et al.*, 2020). The study aims to develop methods, strategies, and infrastructure for breeding organic potato varieties with enhanced stress resistance, exemplified by the Ecobreed project. Ecobreed project tests numerous potato varieties to identify those that require low inputs, have higher yield potential, and are well suited to the region to increase the competitiveness of organic sector (Ecobreed, 2020). A collection of potato genotypes was evaluated across a range of environment in four countries (Slovenia, Hungary, Poland, and the United Kingdom). A list of organic potato varieties in provided in Chapter 5 (**Table 19**). These potato varieties were selected from across Europe, including traditional cultivars, breeding materials, research materials, commercial cultivars, and improved cultivars. They have been tested, bred, and evaluated for their suitability for organic or low-input production systems.

Nevertheless, the choice of varieties may vary based on different factors and can change from year to year. Testing and selecting potato varieties that require fewer inputs while maintaining yield potential and disease resistance will help farmers overcome the limitations of disease management and resource use. This will ultimately improve the sustainability and commercial viability of organic potato production.

1.4.2 POTATO VARIETY DEVELOPMENT

Potato domestication (cultivated potatoes) from wild species traces its origin to Andean and Chilean landraces developed by pre-Colombian cultivators (Gopal and Khurana, 2006).

Within the species, there are many different cultivars or varieties that have been bred for desired underground traits, such as yield, disease resistance, and tuber quality (Spooner *et al.*, 2005). These cultivars can be further classified into different groups based on characteristics (e.g., skin and flesh colour, cooking quality, and maturity). High yield is not the only priority, but also a tuber dormancy period has become an important goal as it relates to a longer storage period (Mori *et al.*, 2015). However, the fundamental goal of potato variety development is to create improved varieties that are suitable for varied growing conditions and different climates, as well as to meet the demands of farmers and consumers while simultaneously addressing the rise in global food demand and climate change (Andrivon, 2017; Bonierbale *et al.*, 2020; Devaux *et al.*, 2021).

Understanding the genetic diversity and population structure of breeding material is crucial for potato variety development (Vales *et al.*, 2022). Although the progress has been achieved in genotyping and molecular-based breeding techniques, the approach of physiological phenotyping remains limited to a small number of genotypes. For over a half century, this has been a critical limitation for applying physiological information to breeding programmes. While yield is often considered the most important commercial traits, other traits frequently determine the marketability of potato tubers. When selecting the commercial variety, it is important to consider many quality characteristics, such as resistance to common as well as region specific, defects, diseases, pests, and stresses (Stark *et al.*, 2020b). To achieve future progress in potato breeding, there is a need to renew the emphasis on potato physiological phenotyping for specific and well-defined traits. For example, to evaluate the pest and disease responses of a new potato variety, a replicated field trial must be conducted for a minimum two years (Bonierbale *et al.*, 2020; Vales *et al.*, 2022).

Despite the large number of potato cultivars currently available, there is a clear need for the development of new cultivars to meet the growing demand for high-yielding and biotic/abiotic-tolerant varieties while reducing the demand for nutrients and water input (Pandey *et al.*, 2021). This demand is driven by population growth and changes in climate conditions around the globe. To achieve higher yields and good trait quality, breeders and biotechnologists have focused on improving traits related to biotic and abiotic stresses as well as the ability of cultivars to adapt to different production environments. In the European Union (EU), the potato industry is trying to increase potato usage in an economically and environmentally sustainable manner, with the new cultivar required to deliver a higher yield of a saleable product under low inputs, possess inbuilt resistances to pests, diseases, and

environmental stresses, and offer improved nutritional, flavour, and health benefits (Bradshaw, 2007; Lutaladio and Castaldi, 2009). In addition, to develop a new variety, it required several years of field trials to ensure a high fresh market rating score. For example, the most widely grown UK variety (cv Maris Piper) has been replaced by a selection of alternative varieties that are more suitable for specific end uses, such as 'cv Markies' which is ideal for frying, 'cv Innovator' and 'cv Royal' are suitable for French fry processing, and 'cv Melody' for baking (Goffart *et al.*, 2022). This indicates that there is a need for a new cultivar that meets consumer or market demands.

1.4.3 PRECISION AGRICULTURE

Sustainably increasing productivity in a changing climate is one of the most important challenges when conducting research potatoes worldwide. The yield of potatoes is affected by many factors, including varieties, weather conditions, production systems, supplies, and available resources. To improve quality and yield, to fulfil the growing demand, suitable cultivars are required along with a good decision support system tailored to local growing conditions. Climate change poses a significant challenge to potato production, as it can lead to unpredictable weather conditions that affect crops. Additionally, the increasing price of fertilisers adds to the production cost, making it necessary to find more efficient methods to increase yield and reduce expenses. These challenges highlight the need for innovative approaches like precision agriculture to sustainably increase productivity in the face of changing conditions.

Precision agriculture (PA) is a farming approach that utilises advanced technologies such as remote sensing, geographical information systems (GIS), and data analysis to manage field variability for the purpose of optimising crop productivity and reducing environmental impact (Robert, 1993; Radoglou-Grammatikis *et al.*, 2020; Wu *et al.*, 2022). The concept of precision agriculture was first introduced by Dr Pierre Robert in the 1980s, but it has continuously evolved alongside technological advancements and our growing understanding of its potential (McBratney *et al.*, 2005). Today, precision agriculture involves the use of real-time data and sensors to monitor various factors that can affect crop growth and development, allowing farmers to make informed decisions and adjust input resources accordingly (Moysiadis *et al.*, 2021). Precision agriculture has transformed potato farming by providing farmers with real-time data and remote monitoring capabilities. For instance, farmers can use sensors and other tools to monitor soil moisture levels and other factors that impact crop

growth. With this information, farmers can adjust input resources, accordingly, leading to a reduction in costs and an increase in productivity and quality (Adamchuk *et al.*, 2010). Precision agriculture can also help improve soil variability management, allowing farmers to optimise soil health and fertility in the long term (McBratney *et al.*, 2005).

While precision agriculture offers many advantages, it does require the use of specialised equipment and software to gather and analyse the information. This can be a potential challenge for farmers, as it may involve additional investments and training given the complexity of the process. Furthermore, the effectiveness of precision techniques may vary depending on the specific growing conditions and the availability of resources. For example, a disease detection approach found effective in a greenhouse setting might not produce the same result in field conditions (Sishodia *et al.*, 2020). An in-depth examination of PA covering all the associated technologies, and reference of good practices around UK and Europe is crucial (Moysiadis *et al.*, 2021). It is important for farmers to carefully evaluate the feasibility and cost-effectiveness of implementing precision agriculture practices. Future research in PA should concentrate on applying this knowledge to real world or farm condition practices (Maes and Steppe, 2019).

Overall, precision agriculture technology can play a critical role in helping potato farmers increase productivity sustainably, in the face of climate change conditions, while also reducing costs and improving crop quality. The use of precision agriculture in combination with organic practices and suitable cultivars, can significantly contribute to a more sustainable potato sector in the United Kingdom. By adopting precision agriculture, farmers can preserve long-term soil health and fertility while lowering environmental impact and boosting profitability by reducing the input needed.

1.5 POTATO TUBER QUALITY

One of the most important aspects of potato production is tuber quality, that include external quality traits (such as marketable tuber size, shape, skin appearance, absence of diseases or defects) and internal quality such as flavour, colour, dry matter, starch content, and texture (Carputo *et al.*, 2004; Naumann *et al.*, 2020; Stark *et al.*, 2020a). Consumer preferences for fresh market potatoes commonly refer to visual properties, including tuber's shape and appearance with an absence of defects and disorders (Storey, 2007; Jemison Jr *et al.*, 2008). According to the study by McGregor (2007) reveals key areas identified by consumers including packaging, convenience, usage marketing, product innovation, health, and product

quality, with more than 60% wanted potatoes that are quick and easy to cook. As these tuber quality traits are genetically controlled, selecting a potato variety to grow is a key to meet market demands and consumer preferences (D'hoop *et al.*, 2008; Stark *et al.*, 2020a; Thornton *et al.*, 2020).

1.5.1 TUBER SIZE AND SHAPE

Fresh market and processing potato have specific demands regarding tuber size and shape. A tuber size threshold of more than 35 mm of diameter was used to determine marketable fraction of tuber yield (Zarzyńska and Boguszewska-Mańkowska, 2024). Consumers in the fresh market often prefer medium-sized tubers due to their versatility in several culinary purposes (Haverkort *et al.*, 2023). While the processing sector prefers larger and uniform tubers due to higher yield and consistency in products such as French fries, chips (crisps), and frozen and chilled products (Storey, 2007; Haverkort *et al.*, 2023). For the UK and Europe market, the size range 40-95 mm is acceptable for processing (Kirkman, 2007). For tuber shape, the industry standard for processing product is round to short-oval and should be uniform (Kirkman, 2007). Long and round varieties are suitable for fresh market (Stark *et al.*, 2020a).

The shape and marketable size of potato tubers are influenced by several key factors such as production system, variety, weather condition, and nutrient availability (Bussan *et al.*, 2007; Stark *et al.*, 2020a). Different potato varieties have specific genetic traits that influence tuber size, shape and fresh weight (Gondwe *et al.*, 2020). Some varieties naturally produce larger tubers while others are bred for smaller sizes. Although tuber shape has a relatively high heritability and is generally a stable trait, it is not always a clear defined trait as they are potentially influenced by different physiological and environmental processes during tuber development (Neilson *et al.*, 2021).

Optimal temperatures for potato growth are between 14 - 22-degree Celsius, high temperature can reduce tuber size, while low temperatures may slow growth (Struik, 2007; Ávila-Valdés *et al.*, 2020). Adequate and consistent moisture is crucial during tuber development. Drought stress can significantly reduce the fresh weight of the marketable tuber size (Aliche *et al.*, 2019). Nutrient levels are necessary for vegetative growth, root development, tuber initiation and starch accumulation which contributed to final yield and marketable size (Rosen and Bierman, 2008; Rens *et al.*, 2015).

1.5.2 PHYSIOLOGICAL DISORDERS

Potatoes are prone to various physiological disorders that are not directly caused by infectious pest but they result from unfavourable environmental conditions, improper management practices, high temperature, and inadequate moisture or nutrient levels (Sowokinos, 2007; Mikitzel, 2014; Thornton *et al.*, 2020). These disorders can lead to abnormal growth, external and internal appearance on tubers such as tuber malformations, growth cracks, bruises, enlarged lenticels, brown spot (IBS), and greening (Fiers *et al.*, 2010; Mikitzel, 2014; Bamberg *et al.*, 2015; Koch *et al.*, 2019; Thornton *et al.*, 2020; Pentangelo *et al.*, 2021).

Common physiological disorder such as greening is cause by the exposure to light during growth and storage. The greening occurs in the periderm which cause the accumulation of chlorophyll and the toxic glycoalkaloids toxic glycoalkaloids (Bishop *et al.*, 2012; Tanios *et al.*, 2018). Whereas the internal shattering or cracking of tuber tissue is a result in a combination of problems including disease, severe defects created by discolouration and the development of scars from the wound (Lulai, 2007; Mikitzel, 2014). Bruise and shatter bruise are damaged on the skin and the underlying tissue directly below occurs when potatoes are damaged during harvest and handling process (Hussein *et al.*, 2020; Hendricks *et al.*, 2022). As these physiological disorders can significantly affect the tuber appearance and mostly identified through post-harvest symptom evaluation, hence, lead to a significant reduction in yield and quality of marketable tubers. Losses form many physiological disorders can be minimised using good agricultural practices that promote uniform plant and tuber during the growing season.

1.5.3 TUBER SPROUTING

Tuber sprouting is associated with further quality loss during storage (Bishop *et al.*, 2012). The commercial value of sprouted potatoes decreases significantly due to undesirable changes in their morphological characteristics and taste. Sprouting alters potato quality by causing softening, shrinkage, and alkaloid buildup, and consequently leads to changes in texture and taste (Dourado *et al.*, 2019; Visse-Mansiaux *et al.*, 2022).

When harvested, potato tubers are in a dormancy state, which lasts for weeks or months depending on the variety and the pre- and post-harvest conditions (Sonnewald and Sonnewald, 2014; Thornton, 2020). As temperature, humidity, and light exposure can break dormancy and stimulate sprout initiation (Sonnewald and Sonnewald, 2014; Gong *et al.*, 2021a; Hu *et al.*,

2023). Therefore, adequate storage conditions are crucial to control sprouting and reduce quality loss. The preferred storage condition to suppress sprouting include maintain temperature between 3 and 4 degrees Celsius, high relative humidity (80-90%), and good air circulation (Suttle, 2007; Gong *et al.*, 2021a). Additionally, sprout control can be achieved through the application of radiation, chemical compounds (i.e., maleic hydrazide) or natural substances (such as mint essential oil) (Sonnewald and Sonnewald, 2014; Hu *et al.*, 2023). Research by Mølmann and Johansen (2020) also found that using red light-emitting diodes (LEDs) during storage can effectively inhibit sprouting.

1.5.4 POTATO DISEASES

The tuber quality decreases by between 5% and 40% due to potato pest and diseases (Lulai, 2007). Potato diseases are caused by various pathogens, including viruses, bacteria, phytoplasmas, nematodes, and fungi (Kirkman, 2007; Fiers *et al.*, 2010; Dourado *et al.*, 2019). These diseases can impact potatoes at any growth stages from seed to storage which lead to reductions in tuber quality (Nolte *et al.*, 2020). Diseases, such as potato leafroll virus (PLRV), potato virus Y (PVY), early blight (*Alternaria solani*), and late blight (*Phytophthora infestans*), which shorten the length of the tuber growth period and damage foliage can also affect production and tuber quality (Kreuze *et al.*, 2020; Stark *et al.*, 2020a). Tuber yield losses are caused by either of them in single infections and can reach more than 80% in combination with ither viruses (Kreuze *et al.*, 2020).

Disease development in potato plants is linked to plant nutrition as nutrition affects both plant physiology and disease resistance or tolerance (Dordas, 2008; Tein *et al.*, 2015). Plants demonstrates highest disease resistance when their nutritional status is optimal, as balanced nutritional promotes healthy growth and enhance the plant's defences against pathogen. For example, severity of leaf blight is enhanced by potassium and calcium deficiency (Huber *et al.*, 2012). Thus, maintaining adequate nutrition is essential for minimising disease susceptibility and enhancing overall plant vitality.

Reducing the chances of disease development can be achieved through practice measures which include crop rotation, resistant varieties, and the use of healthy and certified seed tubers, while regular monitoring and integrated pest management (IPM) are employed (Vilvert *et al.*, 2022; AHDB, 2024b).

1.5.5 GREENING IN POTATO TUBERS

While potato is a member of the nightshade family (Solanaceae), the edible part is the tuber, not the fruit or leaves, which contain toxic substances. The greening of potato tubers is one change that happens when the tubers are exposed to light. This greening is the single biggest cause of wasted potatoes due to its association with the accumulation of toxic alkaloids known as potato glycoalkaloids (PGAs). To assure the safety and quality of market potatoes, it is important to understand the mechanisms causing potato greening. To prevent poisoning, the standard PGA (potato glycoalkaloids) contents in fresh market potatoes have been suggested. Although the formation of chlorophyll has no effect on taste, it can still result in consumer rejection, resulting in economic loss (Bamberg *et al.*, 2015; Tanios *et al.*, 2020). Furthermore, consuming green or sprouted potatoes can be harmful due to the higher levels of toxins they may contain. Therefore, it is necessary to prevent potato greening to avoid the buildup of PGAs and maintain the safety of potato consumption. In addition to the health risk associated with the consumption of green potatoes, it also affects the quality and marketability of potatoes, which ultimately leads to economic loss.

THE GREENING PROCESS

Chlorophyll is a green pigment found in the chloroplasts of green plants and algae, which is important in the photosynthesis process. The greening process in potato tubers is the result of the formation of chlorophyll when exposed to light. Potato tubers are non-photosynthetic tissues, however, after exposure to light, it results in greening in potato tubers, due to the accumulation of chlorophyll in the peripheral layers (Pavlista, 2001). Green potato tubers are easily identified by their surface coloration. However, chlorophyll is not the only pigment that causes potato tubers to become green. Other pigments, like xanthophyll and carotenoids, also contribute to the tubers' green hue. The carotenoid pigment is naturally found in red, yellow, and orange crops and can also be found in dark green vegetables (Maoka, 2020; Sousa, 2022). Important factors influencing the formation of chlorophyll and the greening process are influenced by many factors, including varieties, growing conditions, and post-harvest management, with genetic variation being an important factor (Tanios *et al.*, 2020). However, the greening process is also associated with the formation of potato glycoalkaloids (PGAs) that can cause allergic reactions and illness. Unlike chlorophyll, light is not needed for glycoalkaloid formation, but with light, glycoalkaloid formation increases.

1.5.6 POTATO GLYCOALKALOIDS

Glycoalkaloids (GAs) are natural plant metabolic compounds found in the Solanaceae plant family, including potatoes (Friedman, 2006). Glycoalkaloids (GAs) are natural secondary plant metabolic compounds that contain nitrogen, in a steroidal structure, and carbohydrate side chains (aglycons) at the 3-OH position (Barceloux, 2009). There are other glycoalkaloids found in potatoes, with 95 percent of the total glycoalkaloids primarily comprised of α-chaconine and α-solanine, in cultivated and wild potatoes (Friedman *et al.*, 1997; Friedman *et al.*, 2003; Duke Gekonge *et al.*, 2016). These two are trisaccharide steroidal glycoalkaloids but differ in the composition of the sugar side chain; α-chaconine consists of two rhamnose molecules and one glucose, whereas α-solanine is formed up of galactose, glucose and rhamnose (Friedman *et al.*, 2003). The significance of potato glycoalkaloids (PGAs) in humans is that a high concentration in potatoes may cause poisoning and could be fatal. The recommended safety limit for new potato varieties is 200 mg per kilogramme of fresh weight. Reducing the PGA levels in fresh market potato will provide numerous advantages from the cultivation stage to the processing, post-harvest, transportation, and consumption of potato and potato-based products (Friedman and McDonald, 1999).

The PGAs are found in all parts of the potato plant, with the highest concentrations found in the flowers and the lowest in the tubers; however, variation can be observed both between and within varieties and tuber maturity (Grunenfelder *et al.*, 2006a; Tanios *et al.*, 2020). The analysis of PGAs has been the recognition and separation of individual glycoalkaloids, to which many techniques have been applied, such as gas chromatography, thin layer chromatography (TLC), high pressure liquid chromatography (HPLC), radioimmunoassay (RIA), and enzyme linked immunosorbent assay (ELISA) (Coxon, 1984; Morgan *et al.*, 1985; Okamoto *et al.*, 2020). The use of HPLC with UV/Visible spectrometer detection is the most common approach due to the speed of analysis, high reliability of separation, repeatability, and ability to distinguish between the two glycoalkaloids in different plant species without derivatization (Matsuda *et al.*, 2004; Friedman, 2006; Nema *et al.*, 2008).

1.5.7 POST-HARVEST AND STORAGE

Maintaining high-quality tuber is crucial for fresh market potato, processing, and economic losses. Tuber quality is established in the field, with final qualities are influencing by tuber maturity, varieties, and field variability (Alamar *et al.*, 2017). Fresh potatoes are typically available for two to three months after harvest; therefore, potatoes need to be stored

to maintain a year-round supply (Gottschalk and Ezekiel, 2006). To ensure the tuber quality during storage, they need to be kept under controlled conditions that prevent sprouting and damage.

A healthy periderm serves as a protective barrier for potato tubers from pests, diseases, infection, and other factors that lead to quality deterioration and financial losses (Lulai, 2007). Understanding the physiological changes that occur post-harvest is also critical for maintaining tuber quality. Since potato tubers remain biologically active after harvest, they require appropriate storage conditions to preserve their quality (Olsen and Kleinkopf, 2020). Optimal storage conditions depend on the purpose use of potatoes (table stock, processing, and seed) and varieties, for example, varieties for table use are typically stored at 4-6 degree Celsius and for processing at 7-12 degree Celsius (Gottschalk and Ezekiel, 2006; Bradshaw, 2021). Light also increase the chlorophyll, glycoalkaloids and other substances in potato tubers that may have positive effects on resistance against pathogens (Johansen and Mølmann, 2018).

The industry has always innovated and invested in improved post-harvest storage. However, there is concern about traditional chemical post-harvest treatments for sprout suppression and disease control. Therefore, there is a need in creating alternative post-harvest treatment. Employing molecular and advanced phenotyping techniques to enhance the understanding of mechanisms which mediate physiological responses during pre-harvest production, post-harvest management, storage, and processing will improve tuber quality (Alamar *et al.*, 2017; Dourado *et al.*, 2019).

1.6 **PHENOTYPING**

Phenotype refers to the concepts of observing and measuring the physical properties or characteristic traits of an organism that are observable and measurable which developed from the interaction with its environment (Fiorani and Schurr, 2013; Pieruschka and Schurr, 2019). In plant research, these traits can range from simple features such as plant height, canopy cover, canopy volume, and disease resistance or tolerance, to more complex features like root architecture and chlorophyll contents in both greenhouses and field environments (Cendrero-Mateo *et al.*, 2017). The challenges originate from the need for a deep understanding of the interaction between plant performance and its growing conditions (Pieruschka and Schurr, 2019).

In potato production, plant phenotyping is an important aspect of genotype studies and potato breeding. It helps to enhance the efficiency and accuracy of plant breeding and monitor crop performance to reduce economic losses for future food security, which has led to challenges in field-based phenotyping. However, conventional phenotyping methods pose significant challenges due to their being labour-intensive, time-consuming, expensive, and lacking consistency in characterising the effects of environment and genotype on plant traits (Watanabe *et al.*, 2017). While significant progress has been achieved in the efficient examination of genetic data and non-invasive assessment of physical traits (Großkinsky *et al.*, 2015), large-scale analysis is still lacking.

To overcome these challenges, over the last three decades or so, novel opportunities such as sensors, automation, computer vision, and quantitative data analysis have been developed for quantifying plant traits with increasing throughput and accuracy (Fiorani and Schurr, 2013; Pieruschka and Schurr, 2019). One of the novels is the high-throughput phenotyping pipeline (HTPP), which has been developed as a non-invasive plant trait measurement (Watt *et al.*, 2020). This approach allows the researchers to study the form and function of many genotypes subjected to different growing conditions. With HTPPs, research can collect data on multiple traits simultaneously, making progress much faster and more efficient. This approach also reduces the potential for human error and increases the accuracy of data collection.

1.6.1 HIGH-THROUGHPUT FIELD PHENOTYPING TECHNIQUES

Field phenotyping is an essential practice in plant breeding and crop monitoring that involves the measurement and analysis of the physical and biochemical characteristics of plants in their natural growing environment (Chawade *et al.*, 2019). This process helps in identifying and selecting plant varieties that are best adapted to a specific growing environment. It includes assessing various key plant traits such as growth, yield, disease resistance, and response to environmental stress such as temperature and water. However, current field phenotyping techniques can be challenging due to the variability of field conditions and the difficulty of collecting accurate and consistent data in large-scale field trials. Environmental variability can influence crop performance and may pose a challenge in field phenotyping as it limits the potential for genetic improvement (Araus *et al.*, 2018).

In recent years, high-throughput phenotyping (HTP) has emerged as a new technique in plant science research, with the development of automated systems for studying plant traits

in both greenhouse and field settings (Wang *et al.*, 2020). One of the key advancements in using high-throughput techniques for field-based phenotyping is the use of cameras to capture the raw image of a large number of plants. Image collection and processing pipelines have also been developed to automate this process for micro-plots in the field and for individual plants (Daviet *et al.*, 2022). While HTP has shown promise in the research environment, applications for the field environment may need to be tested and validated to ensure their practicality.

1.6.2 POTATO PHENOTYPING

Phenotyping in potato breeding research involves the characterisation and measurement of various traits related to the growth, development, and performance of potato plants. These traits can include plant morphology, yield, tuber quality, disease resistance, and response to environmental stress such as drought or heat. By evaluating these and other traits in potato field phenotyping studies, researchers and breeders can better understand how different genotypes respond to environmental factors and identify traits that can be targeted for improvement through breeding or management practice (Watt *et al.*, 2020).

Field-phenotyping plays an important role in potato breeding as it enables breeders to identify and select potato varieties with desirable traits that can improve crop productivity, reduce economic losses, and enhance food security. Traditionally, potato phenotyping has been carried out through manual measurements or field inspection. Potato breeding involves targeted crossbreeding, followed by repeated selection, based on field observation over multiple generations to identify an improved phenotype (Slater *et al.*, 2017). However, effective field phenotyping requires more than just trait selection, there is also a need to consider various factors, such as environmental limitations (Araus *et al.*, 2018). A traditional approach is time consuming, labour-intensive, and subject to human error. Therefore, the recent advance in technology has led to the development of high-throughput field phenotyping.

The development of effective field phenotyping tools is a key bottleneck in deciphering and acquiring a better understanding of the morphological changes on potato growth and their impact on yield. Non-invasive technologies are used to analyse experimental and breeding populations and promise to speed up the breeding process (Watt *et al.*, 2020). With advances in high-throughput phenotyping techniques, including proximal and remote sensing of plants in the field, potato breeders are now able to efficiently and accurately phenotype a large number of potato genotypes (varieties) over a large study area. This accelerates the selection of superior potato varieties for commercial cultivation. While HTP has shown promise in the research

environment, applications for the field environment may need to be tested and validated to ensure their practicality.

1.7 PLANT PHENOTYPING TECHNOLOGY

The interest in the application of remote sensing technologies among the plant phenotyping community has been increasing, which leads to the development of a number of techniques and applications able to characterise key physiological traits (Tripodi *et al.*, 2018). Remote sensing is used for on-site observation or data acquisition; its aim is to characterise the objective or feature of interest from its neighbour non-destructively and non-invasively. Extracting crop information from remote sensing can be achieved by measuring the vegetation spectral properties, which vary within the wavelengths, also known as a spectral signature (Das *et al.*, 2019). The spectral properties of vegetation make it possible to monitor vegetation dynamics and their spatial and temporal variability using various remote sensing platforms.

1.7.1 REMOTE SENSING IN AGRICULTURE APPLICATION

Remote sensing is a key technology in plant breeding and precision agriculture for spectral and spatial data acquisition. The rapid development of remote sensing has taken place in the field of non-destructive measurements that allow the assessment of plant traits with high throughput (Walter *et al.*, 2015). The method of remotely sensed field phenotyping differs based on plants, traits of interest, growth stages, and available resources (Aakash *et al.*, 2019). For plant breeding, field phenotyping is done on many small plots of hundreds or even thousands of breeding lines under different environmental conditions (Aakash *et al.*, 2019). The main aims of HTP with remote sensing for plant breeding are to accelerate phenotypic gains in order to develop new cultivars that perform well under diverse field conditions and management systems. In this aspect, remote sensing has been introduced for plant breeding to improve selection efficiency through informed decision-making by evaluating the quantitative traits such as yield and disease resistance. Whereas, in precision agriculture production, evaluation is mainly done to support decision making such as monitoring fertiliser requirements, disease occurrence, weed detection, and yield prediction (Son *et al.*, 2014; Gago *et al.*, 2015). For future developments, early disease detection holds a promise to revolutionise precision agriculture.

1.7.2 PHENOTYPING PLATFORMS

To implement information technologies for field-scale phenotyping; several platforms are being used, from proximal to remote sensing with a wide range of sensors, each with their own constraints and advantages. A good platform should be flexible to handle an assorted range of sensors and mounted with an accurate and precise positioning system. The plant phenotyping community increases interest in the application of remote sensing technologies, which leads to the development of several approaches and applications able to characterise key physiological traits (Tripodi *et al.*, 2018).

There are three broad categories of remote sensing platforms available: ground based, airborne, and satellite. The first developments in remote sensing applications using imaging spectrometry started with the development of satellite-based imagery resources (Adao *et al.*, 2017). Spectral satellite imaging has been being used for large-scale field evaluation in agriculture for a long time (Chapman *et al.*, 2014). It has been successfully used to monitor crop health and performance (Rembold *et al.*, 2013). Currently, there are increased capabilities available in satellite-based approaches for detecting vegetation health and stress with wide spatial and spectral resolutions (Yuan *et al.*, 2016; Hu *et al.*, 2019). However, the limitations of satellite-based systems are that they are not easily affordable, there is a lack of data acquisition due to the climate conditions, and in some cases, there are long revisit times (Zaman-Allah *et al.*, 2015).

The alternative solution is the use of unmanned aerial vehicles (UAVs)-based platforms. The use of UAVs as a remote sensing tool for aerial image phenotyping has become critical in advanced crop breeding and precision agriculture (Zaman-Allah *et al.*, 2015), as it delivers accurate and precise physiological crop information more rapidly, effortlessly, and non-destructively (Adamchuk *et al.*, 2010; Shakoor *et al.*, 2017). These are possible due to the development of light-weight UAVs with better flight controls and the possibility of mounting high spatial and spectral sensors that effectively increase temporal resolution. The UAV equipped with sensors designed for phenotyping studies has demonstrated the capabilities of large-scale crop monitoring, field phenotyping, yield and quality forecasting (Zaman-Allah *et al.*, 2015), and early disease detection (Sugiura *et al.*, 2016; Franceschini *et al.*, 2019) with high throughput (Richard *et al.*, 2018).

Beside the UAV and satellite platforms, proximal remote sensing (PRS) platforms have been used for high precision field phenotyping and disease detection (Virlet *et al.*, 2016;

Tripodi *et al.*, 2018). Field-phenotyping using proximal platforms refers to phenotyping done with ground-based vehicles coupled with sensors, which can be handheld or mounted on platforms such as ground vehicles, stationary towers and field scanners (Aakash *et al.*, 2019). Although PRS is fast, has fewer errors, and has the possibility to carry more sensors, it is relatively expensive to build and operate. Three types of remote sensing were directly compared (satellite-based imagery, UAV, and PRS). UAV-based remote sensing was found to be the most precise, efficient, and cost-effective for field phenotyping (Tattaris *et al.*, 2016).

1.7.3 SENSOR TYPES AND CAPABILITIES IN REMOTE SENSING PLATFORMS

Remote sensing platforms that are available for field phenotyping include UAV-based, ground-based, and satellite platforms, can be mounted with a wide range of sensors, such as RGB, thermal infrared, LIDAR, multispectral, and hyperspectral (Yang *et al.*, 2017; Richard *et al.*, 2018). The number and type of sensors vary from one to multiple sensors, and their applicability depends upon the type of crops, parameters to be evaluated, resolution levels, field or plot sizes, and local conditions or areas to be exploited. The choice of sensor and its capabilities will depend on the specific research objectives and the characteristics of the vegetation and surrounding environment.

THERMAL SENSING

Thermography or thermal imaging is one of the techniques that works on the emittance and can be used for detecting crop stresses, canopy temperature, water deficit, and disease occurrence over a large field by visualising the differences in canopy temperature (Shakoor *et al.*, 2017). Several authors have been discussed and reviewed the viability of infrared thermography as a high-throughput field phenotyping (HTFP) tool (Prashar *et al.*, 2013; Prashar and Jones, 2014; Deery *et al.*, 2016). Recently, studies have utilised the UAV-RS for thermal imaging acquisition for early disease detection (Prashar *et al.*, 2013; Zhu *et al.*, 2018). Thermal cameras can also be used to predict the final yield. For example, Chapman *et al.* (2014) used thermal data to compare the yield and sugar content of sugarcane clones under different irrigation treatments (drought, half irrigation, and full irrigation). The result showed that there was a variation among clones within the irrigation environment, and a higher canopy temperature is associated with a lower transpiration rate. Rutkoski *et al.* (2016) used UAV platforms to acquire multispectral data coupled with a thermal camera to predict grain yield in wheat.

REFLECTANCE SENSING

Sensors measure the radiance, which represents the intensity of light in a certain direction towards the sensor. Reflectance or spectral sensors assess the spectral information of vegetation by measuring the spectral reflectance in the visible (400-700 nm), near infrared (700-1300 nm), and shortwave infrared (1300-2700 nm) (Burger and Gowen, 2011). Spectral reflectance measurement of vegetation can be related to crop structure and chemical characteristics; it acts like a signature that enables the estimation of crop traits based on the spectral properties of plants (Tripodi *et al.*, 2018). There are three broad reflectance sensors, including RGB, multispectral, and hyperspectral, which are categorised based on the number and width of measured wavebands, also known as spectral resolution (Mahlein, 2016a).

Images from an RGB digital camera, which records spatial data in the visible region (red, green, and blue bands: RGB) (Tripodi *et al.*, 2018), have been widely used for non-invasive phenotyping and disease detection as they are suitable for the use under natural illumination (Shakoor *et al.*, 2017). RGB imagery has been extensively used as a primary tool for assessing field phenotyping and disease screening. For instance, Sugiura *et al.* (2016) developed a potato late blight severity estimation technique in a field using RGB imagery from unmanned aerial vehicles. This study showed that disease severity estimated from RGB imagery is more efficient compared to visual disease assessment. Moreover, it has been established to be sufficient for the estimation of leaf area and LAI (Leaf Area Index) (Tripodi *et al.*, 2018).

Reflectance datasets from the hyperspectral and multispectral have the potential to discriminate between healthy and diseased crops (Franceschini *et al.*, 2019) and plant architecture by relating spectral signature and crop architecture information (Shakoor *et al.*, 2017). Focusing on spectral information related to disease occurrence, multispectral imaging is an initial tool for assessing genetic variation and early disease detection on the field scale at the canopy level (Franceschini *et al.*, 2019). Many studies have reported successful discrimination of healthy plants from diseased plants in various crops, such as stripe rust infection and nitrogen deficiency in wheat (Devadas *et al.*, 2015), mosaic virus in cassava (Raji *et al.*, 2016), leaf spot in maize (Adam *et al.*, 2017), verticillium wilt in olive trees (Calderón *et al.*, 2015), and late blight in potato (Sugiura *et al.*, 2016). Higher spectral resolution allows higher detection capability, but this comes with the possible limitation of data storage and probably acquisition time. Multispectral is limited by the number of spectral bands (two to ten),

whereas hyperspectral provides ten to a hundred spectral bands in visible, near infrared, and shortwave infrared. Therefore, hyperspectral offers a more targeted approach to spectral acquisition and is richer in data than multispectral.

1.7.4 ADVANCING CROP MONITORING

Crop monitoring and disease observation have traditionally relied on manual methods conducted by the field inspector or agronomist. The conventional methods perform well for qualitative traits; however, they are time-consuming, prone to human error, and labour-intensive, or permit the use of only a few sensors at a time (Campos *et al.*, 2019). With the climate change situation is expected to lead to a yield reduction due to a shorter season length (Haverkort and Struik, 2015). Nonetheless, the effects of climate change on potato yield are anticipated to vary across different regions. To overcome these challenges with the anticipated impacts of future climate changes on plant growth and field management practices, there is a growing need for flexible, fast, accurate, and precise crop monitoring methods. These advancements in crop monitoring are crucial to effectively responding to changing climatic conditions and optimising land management strategies.

Producers worldwide are exploring the adoption of precision agriculture technologies, as real-time crop monitoring can be highly beneficial in the context of climate change situations (Adamchuk *et al.*, 2010). With the occurrence of erratic weather patterns and extreme events like droughts, heatwaves, and storms, real-time monitoring allows for timely interventions, empowering farmers to mitigate risks and maximise yields amidst changing conditions. Furthermore, effective crop monitoring enables precise resource management, including water, fertilisers, and pesticides. By tailoring inputs to crop requirements based on monitoring data, farmers can minimise waste and reduce environmental impact. Additionally, crop monitoring provides valuable insights into plant responses to climate stressors, facilitating the development, and implementation of adaptation strategies (Sugiura *et al.*, 2016). This encompasses the selection of resilient crop varieties, the adjustment of planting dates, and the optimisation of cultivation techniques. Through these measures, real-time crop monitoring plays a crucial role in enhancing agricultural resilience and sustainability in the face of climate change.

Advancements in crop monitoring techniques, including remote sensing, sensor networks and unmanned aerial vehicles (UAVs) equipped with sensors, offer significant improvement in wide-area coverage and the frequency of data collection at a localised level.

These technologies enable close monitoring of individual fields or crop areas. Numerous plant phenomics' studies have developed useful tools for assessing plant phenotypes and information in the field, such as biomass (Golzarian *et al.*, 2011), leaf area index (Liang *et al.*, 2015), canopy height (Friedli *et al.*, 2016; Han *et al.*, 2018), growth rates (Holman *et al.*, 2016; Liu *et al.*, 2020), yield potential (Gómez *et al.*, 2019), diseases (Mahlein, 2016b), and stress tolerances (Sugiura *et al.*, 2016; Franceschini *et al.*, 2017). It is clear that this information is crucial for evaluating crop development, health, and environmental responses. However, device variability and plant population structure can influence data collection (Liebisch *et al.*, 2015). Despite these challenges, the advancement in crop monitoring techniques is complemented by data analytics, which enhance the interpretation of monitoring data, identify patterns, and generate predictive models. This helps breeders and farmers make precise selections and crop management decisions (Watt *et al.*, 2020).

In addition, in the face of future climate change challenges, flexible, fast, accurate, and precise crop monitoring is essential for climate-responsive crop management. With advancements in remote sensing, UAVs, sensor networks, and data analytics, breeders and farmers can make informed decisions, adapt to changing conditions, and optimise resources. Future research and development in crop monitoring techniques will continue to improve their efficacy, accessibility, and integration into agricultural practices, ultimately contributing to sustainable and climate-resilient food production systems.

1.7.5 SENSING APPROACH FOR FIELD PHENOTYPING

Cutting-edge field phenotyping technologies, especially in close-range and remote sensing, hold the potential to expedite early-stage breeding procedures (Tripodi *et al.*, 2018; Watt *et al.*, 2020). The concept of high-throughput field phenotyping is the practice of utilising autonomous vehicle quipped with several sensors to gather extensive data on the variability of plants (Shi *et al.*, 2016). As remote sensing technologies offer wide-area coverage and frequent data acquisition, by analysing the spectral signatures and patterns captured by remote sensing sensors, valuable insights into plant characteristics and phenotypic traits can be obtained.

The integration of unmanned aerial vehicles (UAVs) into agriculture remote sensing provides improved data collection with high-resolution imagery and detailed data capture over a large area (Shi *et al.*, 2016). They enable on-point monitoring of individual fields or specific crop areas, facilitating targeted interventions and precision agriculture practices. UAVs can play a significant role in field phenotyping by capturing high-quality images of crops, allowing

for the extraction of plant traits and the assessment of phenotypic variations across different areas of the field (Shi *et al.*, 2016). Although there is various research in multispectral remote sensing, there is a heavily focus on monitoring over large field areas. Guo *et al.* (2018) research highlights the importance of integrating spatial remote sensing data with crop growth models to evaluate crop growth, yield prediction, and study climate change at the spatial scale. This research also highlights that in order to achieve high accuracy and efficacy, there is a need to manage the computational challenges of this integration.

By integrating these sensing approaches with field phenotyping techniques, such as the extraction of plant traits from imaging sensors, the field of crop monitoring is advancing rapidly. These technologies offer opportunities for real-time monitoring, precise resource management, adaptation strategies, and an enhanced understanding of plant responses to climate stressors. As the agricultural sector faces the challenges of climate change and the need for sustainable practices, the development and utilisation of these sensing approaches in field phenotyping provide valuable tools for farmers, researchers, and agronomists to optimise crop production and ensure food security in a changing world.

1.7.6 DATA PROCESSING AND ANALYSIS APPROACH

Data processing and analysis are critical components to extract useful information from complex raw data from phenotyping platforms. High-throughput phenotyping platforms are increasingly being used to collect data on plant growth and development over a large area. However, they can create massive amounts of data. The raw data obtained from phenotyping platforms can be complex and huge in volume. The evaluation of data is crucial for maximising the potential of phenotyping in research and breeding (Watt *et al.*, 2020). Manual analysis of such data is not only time-consuming but also prone to errors and limitations due to human subjectivity.

To overcome these challenges, sophisticated data processing and analysis techniques are employed. These techniques involve employing statistical methods, machine learning algorithms, and data visualisation tools to extract meaningful information and patterns from the raw data (Yu *et al.*, 2006; Watt *et al.*, 2020). By leveraging these techniques, researchers and scientists can uncover relationships, identify trends, and gain insights into plant performance and responses to environmental factors. Data processing involves pre-processing the raw data, which may include data cleaning, normalisation, and transformation to ensure the data is suitable for analysis (Al-jabery *et al.*, 2020). Subsequently, various statistical and

machine learning algorithms can be applied to analyse the data and derive meaningful information. These algorithms can help identify correlations between different plant traits, predict plant responses under different conditions, and classify plant phenotypes based on specific criteria (Rossi *et al.*, 2022). Furthermore, data visualisation techniques are utilised to present the analysed data in a more interpretable and accessible manner. Visualisation tools enable researchers to explore and interpret complex data sets, facilitating a better understanding of plant performance and aiding in decision-making processes (Canonico *et al.*, 2021).

Several approaches can be used to extract plant traits (phenotyping) from remote sensing-based images, providing valuable insights for field phenotyping studies. One of the common approaches is the use of (deep) machine learning algorithms (Pound *et al.*, 2017; Watt *et al.*, 2020). Machine learning techniques such as the Random Forest, which can be trained to classify different crop features based on their spectral signatures. Machine learning algorithms offer a supervised approach to analysing field images and can effectively identify and classify various crop characteristics from remote sensing data (Zhang *et al.*, 2017; Khan *et al.*, 2018). Through the process of training the algorithm using labelled data, they acquire the ability to identify patterns that are linked to particular plant characteristics (Pound *et al.*, 2017). This empowers researchers to efficiently analyse and integrate phenotypic data into genomic prediction through image annotation, presenting a solution to the human annotation bottleneck (Zhang *et al.*, 2017; Watt *et al.*, 2020).

Another approach is object-based image analysis (OBIA) techniques, where the multispectral image is segmented into smaller, homogenous regions, and each region is analysed separately to extract crop features (Benz *et al.*, 2004). OBIA offers a more detailed and precise identification of specific crop characteristics and can be particularly useful for complex field conditions (Nuijten *et al.*, 2019). Image processing techniques, such as image thresholding and filtering, can further enhance the extraction of crop features from images. These techniques help to isolate and enhance specific characteristics of interest, contributing to a more accurate and detailed analysis. In addition to machine learning and OBIA, other approaches that can be used to extract crop features from remote sensing-based images include spectral indices. Spectral indices, such as the Normalised Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI), these indices utilise specific spectral bands to calculate vegetation indices, providing information on crop health, biomass, and photosynthetic activity (Bannari *et al.*, 1995; Haboudane *et al.*, 2004; Bendig *et al.*, 2015).

The choice of approach will depend on the specific crop features of interest, the type, resolution and quality of the imagery, as well as the overall research objectives. Overall, effective data processing and analysis are essential for extracting useful information from the complex raw data obtained from phenotyping platforms. This strategy not only simplifies the retrieval of valuable data for phenotyping studies but also adds to the advancement of precision agriculture and crop monitoring approaches employing advanced techniques. Researchers, breeders, and farmers may gain a comprehensive understanding of their crop and optimise input for higher efficiency and sustainability, improve crop management practices, and accelerate the development of new varieties that are better adapted to changing environmental conditions.

1.7.7 FUTURE DIRECTIONS FOR UAV-BASED PHENOTYPING

To further advance, phenotyping techniques are required for improving the selection of efficiency, being fully automated for monitoring plants, and being cost effective. The future direction of the application of UAV-based phenotyping is expected to experience substantial progress, driven by novel computational tools. As technology continues to advance, data processing and analysis will become more efficient through the development of more advanced and integrated systems that can provide real-time data on various plant features. UAV-based phenotyping is expected to follow several key directions.

First, the development and integration of more sophisticated sensors, such as hyperspectral, thermal, and LiDAR sensors, into UAV platforms will enhance the precision and range of traits that can be monitored and detected. The presence of numerous stresses in field crops poses challenges for optical, automated stress phenotyping by utilising multispectral or hyperspectral technologies, which offer potential solutions (Ray *et al.*, 2011; Van De Vijver *et al.*, 2020; Sarić *et al.*, 2022). Over the years, hyperspectral imaging has gained popularity as a tool for plant phenotyping. This technique offers a more detailed and accurate view of the potato plant's biochemical and physiological properties (Moghimi *et al.*, 2018). When integrated with a UAV remote sensing platform to obtain high spectral resolution data on multiple parameters simultaneously, leading to a more comprehensive understanding of the potato plant's growth and development. This can be done in various locations, such as laboratories, greenhouses, experimental fields, and large farm areas (Žibrat *et al.*, 2019; Sarić *et al.*, 2022).

Second, considering the amount of data that needs to be processed and analysed, various statistical techniques have been developed to help identify factors that influence the trait of

interest. The advancement in computational power is also benefiting research in terms of managing collections such as phenotypic data (e.g., plant height, yield, and disease resistance), genotypic data (e.g., DNA markers), environmental and weather condition data, soil quality, and market data (Slater *et al.*, 2017). This information is then used to make informed breeding decisions and detect the best individual for seed production. There are a few papers that address the potential and challenges of managing big data, underscoring the pivotal role of advanced technologies in enhancing plant phenotyping's effectiveness and applicability (Tripodi *et al.*, 2018).

Third, future UAV-based phenotyping will likely focus on integrating phenotypic data with genomic, environmental, and management data to provide more in-depth information on plant performance. Over the last few years, there has been an increased discussion about the use of novel computational tools such as machine learning, and artificial intelligence (AI) has the potential to enable breeders to analyse vast amounts of genetic and phenotypic data in a more efficient and accurate manner. For example, in order to identify complex relationships between traits and genotype data that may be difficult to detect by using traditional statistical methods. A machine learning algorithm can be used to select desirable traits, resulting in more precise and faster development of a new variety that is better adapted to changing environmental conditions. Additionally, artificial intelligence can optimise breeding programmes by automating tasks such as seed selection, disease identification, and phenotyping.

Next, while remote sensing technologies are becoming more sophisticated and affordable, UAV systems will become more tailored to specific crops and environments. In the case of potato production, the use of remote sensing methods in high-throughput field phenotyping has already shown promising results in research settings (de Jesus Colwell *et al.*, 2021). UAVs are equipped with adaptable, small-sized sensors to gather various sorts of data and allow for the quantification of both spectral and morphological data related to plants (Tripodi *et al.*, 2018). However, their practical implementation in commercial potato production still requires further testing and validation. The ongoing development of remote sensing technologies and field-based phenotyping will continue to improve the ability of plant breeders and farmers to optimise crop yields and produce more efficient and sustainable agricultural practices.

Overall, the use of UAV-based phenotyping in conjunction with novel computational tools can certainly accelerate the development of new varieties that meet the demands of a changing world. By enabling more efficient and accurate data processing and analysis, leading to more informed breeding decisions, we may effectively shorten the breeding cycle in a way that is more precise and better adapted to changing environmental conditions. However, efficient and accurate data processing and analysis methods are key challenges to overcome. Continued research and development in this field will unlock the full potential of UAV-based phenotyping for potato production, leading to improved crop yields, more efficient agricultural practices, and sustainable food production.

1.8 THE THESIS AIMS

The study acknowledges the current constraints of remote sensing technologies, primarily the cost of sensors and the complexity of imaging processing, handling, and acquisition. These factors limit the accessibility of advanced phenotyping platforms, which offer detailed canopy development descriptions with high accuracy and reproducibility. Although advanced phenotyping platforms are performing well, they are not easily affordable for many research and practical applications. Studies focusing on what is a suitable resolution to describe the relationship between canopy characteristics and disease infection at the field scale are still lacking. To address this gap, the study aims to investigate and understand the variability between different systems and the resolution dependence of crop growth and disease detection characteristics in potatoes (*Solanum tuberosum* L.). While assessing this variation, the research will also focus on identifying a suitable resolution to describe the relationship between canopy characteristics, disease infection, and crop variability at the field scale under natural conditions. The structured research consists of methodologies (Chapter 2), field phenotyping utilising multispectral imagery and 3D models (Chapter 3), maturation and yield prediction models (Chapter 4), UAV-based disease detection and classification (Chapter 5), and the analysis of greening tuber and glycoalkaloid content (Chapter 6). Chapter 7 integrates these findings into a comprehensive discussion on advancing crop monitoring.

CHAPTER 2. METHODOLOGY AND DATA ANALYSIS

In the rapid development field of agriculture technology, the advancement of unmanned aerial vehicles (UAVs) in the agriculture sector is revolutionising the way we farm and opening a new era of precision agriculture. Robust tools and techniques coupled with comprehensive data analytic farmwork is the backbone of precision agriculture and recently, effective precision agriculture is heavily dependent on the UAV-based approaches and the analysis of the data it provides. This approach is not only enhancing crop productivity but also paving the way towards sustainable farming practices. A well-defined methodology is important to ensure that the data captured by the UAV-based platform in the field environment is both accurate and reliable. Therefore, a methodology that includes sensor calibration processes and user-friendly applications, is essential for tasks such as identifying crop productivity and health, also known as crop monitoring. A clear methodology allows for the reproducibility of studies and is essential to validating the data. This chapter's focus is on UAV- based multispectral image data analysis and its application in potato phenotyping. The approach for methodology and data analysis in this chapter includes data collection, data processing, and comprehensive computational data analysis. The practicability, repeatability, and robustness of two developed approaches are compared and demonstrated in this chapter including multispectral based imagery in R and three-dimensional canopy models. Using a structured approach, images from UAVs makes it possible to devise the crop phenotypes of small potato plots in the field, even with uncertainties of environment and weather conditions. By adapting the presented methodologies, farmers and researchers can compare the data across different fields which will enable insights into crop monitoring and precision agriculture. One of the aspects of this chapter is its adaptability for comprehensive analysis by standardising data collection and processing techniques. This allows for meaningful comparisons across the field environment. As a brief overview this chapter focuses on UAV-based multispectral image data and its application in potato phenotyping. Each results chapter refers to the methodologies detailed in this chapter, providing explanation for specific employing UAVs for data analysis with the research paving the way for future innovations in agriculture by utilising the power of remote sensing technology for precision agriculture. We are stepping into a future where farming is not supposed to be increasingly productive but also environmentally sustainable. In summary, this chapter presents approaches to potato phenotyping using UAV-based technologies. It stands as a testament to the transformative power of precision agriculture, offering an

alternative tool for future farming methodologies that are efficient, sustainable, transferable, and globally applicable.

2.1 INTRODUCTION

Plant phenotyping is the description and characterisation of plants, which has been a long-understanding practice of plant research in both greenhouse and field environments (Cendrero-Mateo *et al.*, 2017). Although optical approaches for assessing crop performance and quantifying plant traits were developed between the 1950s and 1970s, they were not specifically referred to as "plant phenotyping" at that time (Moran *et al.*, 2003). Such methods have contributed significantly to our understanding of plant biology and agricultural practices. However, to collect plant traits from the early stages to the end of their cycle, poses challenges for quantitative analysis, due to traditional phenotyping methods often relying on manual measurements or field inspection, which are time-consuming, labour-intensive, and limited in scale (McBratney *et al.*, 2005; Martinelli *et al.*, 2015; Mahlein, 2016b; Pieruschka and Schurr, 2019). Therefore, there's a call for high-throughput, non-invasive phenotyping techniques that integrate detailed characterisation, phenotyping, and genotyping for underlying physiological processes (Großkinsky *et al.*, 2015).

Aerial imagery, when combined with multispectral sensors, offers advantages in the field of phenotyping for agriculture research and crop monitoring (Guo *et al.*, 2020). Aerial imagery captured small-scale spatial variation within the field. All aerial images are recoded in RAW data format, and in order to interpret the spectral data captured from the UAV-based imagery, they need to be converted to orthomosaic or the tagged image field format (TIFF) (Michael *et al.*, 2016). There are three main types of spectral data that play a crucial role in aerial imagery. Firstly, according to Adao *et al.* (2017), a digital number (DN) is a value that sensors record but has no actual physical meaningful units. Secondly, radiance is an amount of radiance coming from object and how much light the sensor receives. The amount of light received by the sensor is affected by multiple factors, such as the source of radiation, scattering, and sensor characteristics. Lastly, reflectance is the ratio of amount of radiation leaving an object to the amount of incident radiation. It is a property of the object being observed, also known as a “spectral signature” (Huete, 2004). The “spectral signature” helps in understanding spatial variation by capturing the large-scale multispectral data and offering a bird’s-eye view of the field, and aerial imagery captures field spatial variation at a fine scale. These spectral signatures can be used to distinguish between various vegetation types, health status, or even

detecting the early signs of stress or disease. Field surface reflectance measurements obtained in the field can also be utilised in several remote sensing analytical methods, including the modelling of the canopy (Peddle *et al.*, 2001). By combining aerial imagery benefits with the depth of insights provided by spectral data, agriculture practice can be refined, and research can be developed deeper into plant dynamics and environmental interaction (Jung *et al.*, 2021). However, the development of such an approach focused on phenotyping, and little emphasis has been placed on environmental monitoring and reducing error variances (Araus and Cairns, 2014). Environmental variation masks important generic variation for key traits and has the potential to reduce repeatability, thus reducing cost and increasing the precision of the phenotyping technique.

There are two primary challenges to phenotyping at plot or canopy levels (Guo *et al.*, 2020). First, the plant segmentation from a weedy background. This challenge involves accurately distinguishing the target plant from the background and surrounding weeds using imaging techniques. In a field environment, different plants may grow closely together, posing a challenge for automated systems like UAV technology to distinguish and separate the targeted plant for evaluation (Jung *et al.*, 2021). Accurate phenotyping relies on effective segmentation to ensure measurements are captured from the plant without interference from nearby plants or weeds. Secondly, the estimation of plant traits that are difficult to measure manually. Several plant traits, particularly structural or physiological, can be complex and difficult to measure precisely using manual techniques, such as root architecture, leaf surface area, or unique physiological responses. It is important to address these problems to advance plant phenotyping, especially for research and breeding purposes. Overcoming these challenges will allow for more accurate and effective evaluation of plant characteristics, and bridge the gap between genetic differences, and enhance crop selection and breeding methods.

Overview of the 3D model of the crop canopy and its application

In the last decades, common trends have been demonstrated in the extraction of relevant phenotypic information from plant three-dimensional (3D) models acquired by Structure from Motion (SfM) imaging, stereo vision (SV), LiDAR, and RGB-D cameras (Rossi *et al.*, 2022; Liu *et al.*, 2023). The images that contain an extra depth component (RGB-D) can be used in conjunction with SfM techniques to generate point clouds or mesh representing the plant canopies (Ziamtsov and Navlakha, 2019). This approach not only captures visual (RGB) information but also depth data (D), offering a comprehensive view of plant physiology. The

point cloud data is then filtered to remove noise, aligned, and merged into a cohesive 3D model representing the plant's structure (Liu *et al.*, 2023). Several studies have already explored the feasibility of SfM in conjunction with an unmanned aerial system (UAS) for estimating plant height across various crops, including maize (Malambo *et al.*, 2018), sorghum (Tunca *et al.*, 2024), barley (Bendig *et al.*, 2014), wheat (Holman *et al.*, 2016; Zhou *et al.*, 2023), and rice (Bendig *et al.*, 2013b). For example, Tunca *et al.* (2024) proposed the technique for estimating canopy height from the Digital Surface Model (DSM), Digital Terrain Model (DTM), and vegetation indices (VIs) generated from 3D point cloud data and the SfM algorithm. Malambo *et al.* (2018) proposed the use of UAV and SfM techniques to generate 3D point clouds to collect crop height in maize and sorghum plots. In some studies, point clouds of crop canopies at plot-level have been segmented into quantiles based on height, with 99 percentiles representing the top of the plant and 1 percentile representing the soil and above ground (Anderson II *et al.*, 2019). This method allows researchers to differentiate between various sections of the plant for detailed analysis. Despite the advances in capturing 3D data of plant structure, developing methods to efficiently analyse 3D point cloud data of plant structure throughout the crop growing cycle remains challenging for many phenotyping purposes (Malambo *et al.*, 2018; Ziamtsov and Navlakha, 2019).

The use of crop 3D models in plant phenotyping research represents a significant advancement in agricultural research, offering detailed insights of plant structure and function that were difficult to obtain with traditional methods. These models facilitate a deeper understanding of how different phenotypic traits contribute to plant growth, health, and productivity. By employing these technologies, researchers can improve crop breeding strategies, enhance precision agriculture practices, and make better predictions of plant performance under field environmental conditions.

Overview of the Image R package and its application in multispectral image analysis

Using remote sensing effectively in agriculture requires various software tools for image processing. Software such as Agisoft, PhotoScan, and Pix4D Mapper, are being used to combine overlapping aerial photographs into a unified orthomosaic image or TIFF image file. While generating an orthomosaic is mostly automated, identifying experimental units (plots) and extracting plot-level features are more complex and involve numerous processes (Matias *et al.*, 2020). Breeders or agriculturists often rely on customised and specialised GIS (geographic information system) software, such as ArcGIS, QGIS, and GRASS GIS. To tackle

this issue, the proposal of using packages in a user-friendly software such as the R programming language was suggested to simplify these processes (Bivand *et al.*, 2008; Matias *et al.*, 2020).

R is a programming language and open-source software environment widely used for statistical analysis, data visualisation, and data manipulation (Team, 2021). R provides powerful libraries and packages such as raster, sp, GeoXp, and ggplot2 that offer the user the ability to integrate code, analyse, visualise, and prepare UAV data for further spatial data analysis (Bivand *et al.*, 2008; Laurent *et al.*, 2008). R also has a large community, which means there are resources, tutorials, and an online forum available to help with the analysis workflow and extract valuable insight from the drone data. Its open-source nature and wide availability make it accessible to users worldwide and well suited for analysing UAV datasets.

One of the challenges for small plot field phenotyping is the segmentation of the individual plots from the background in image analysis. The image R package is a powerful tool for spatial data analysis and processing of images in the R programming language. However, packages that require using them for drone image analysis need to be designed for the specific format. For example, in order to analyse five-band multispectral images, a package has been developed to accommodate this, known as FieldimageR (Matias *et al.*, 2020). In this chapter, we delve into the capabilities of the R programming language, with a special focus on the image package. In this chapter, we provide an overview of how the image package can be employed for image analysis and processing tasks. This will include data extraction and analysis for potato canopy traits.

2.2 RESEARCH APPROACHES

This research is abductive by combining a theory that is already available and building it based on the observed data. In the context of using unmanned aerial vehicles and remote sensing for agriculture, a deductive approach involves testing hypotheses about the effectiveness of UAVs in improving the way we study plant phenotyping and plant monitoring.

This study employs the research approach using real time monitoring across fields. In this chapter, we will explore the approaches that have been used or employed with the optical imaging sensor technique for the potato production or breeding programme. The proposed method for assessment of different potato genotypes (phenotypic differences) for different varieties of potato in the field condition is to detect and monitor the changes in potato canopy

characteristics and stresses (both biotic and abiotic) using UAV-based multispectral imagery combined with field inspection (morphological proximal using handheld proximal devices). With the supervised classification and machine learning with the threshold techniques to remove the background and the unwanted objects in the background (e.g., soil and weed). The main stage of the proposed study is to explore the possibility of employing the low-cost and low altitude UAV as an alternative way to monitor potato growth in field conditions.

2.3 RESEARCH DESIGN

2.3.1 EXPERIMENTAL DESIGN

The experiment was conducted using the potato (*Solanum tuberosum* L.) association panel available in our group in the form of a large-replicated study at Nafferton Farm, Newcastle University. A total of 290 varieties are available in this panel and were planted in May 2020, 2021, and 2022. The field plan was designed to minimise the possible impact of external factors that could potentially affect the data collection process. To address this issue, the experimental design consisted of two replicate blocks and had a total of 300 plots (linked to each variety) per replicate block. Each block consisted of six rows with 50 plots per row (**Figure 2**), and each plot was comprised of four tuber seeds of a given variety. To minimise edge effect, a row of guard plants was planted surrounding each block, and for imaging evaluation, the planting distance between rows was kept at 90 cm and between plots at 150 cm. The trials were conducted using standard conventional management practices. The variation in the elevation of the field is small, with a height difference of around 10 cm going from northeast to southwest. An aerial view of the experimental setup presented in **Figure 3**, provides a distinct visual representation of the plot rows and spacing capture with an UAV at 30-metre altitude.

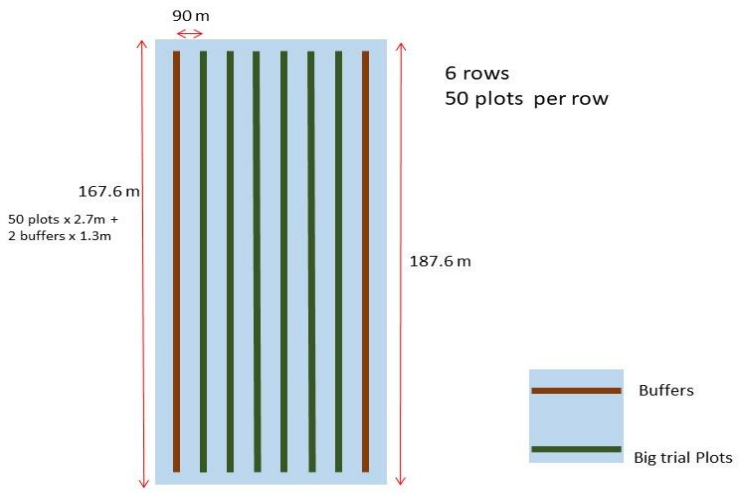

Figure 2 A diagrammatic representation of field layout.

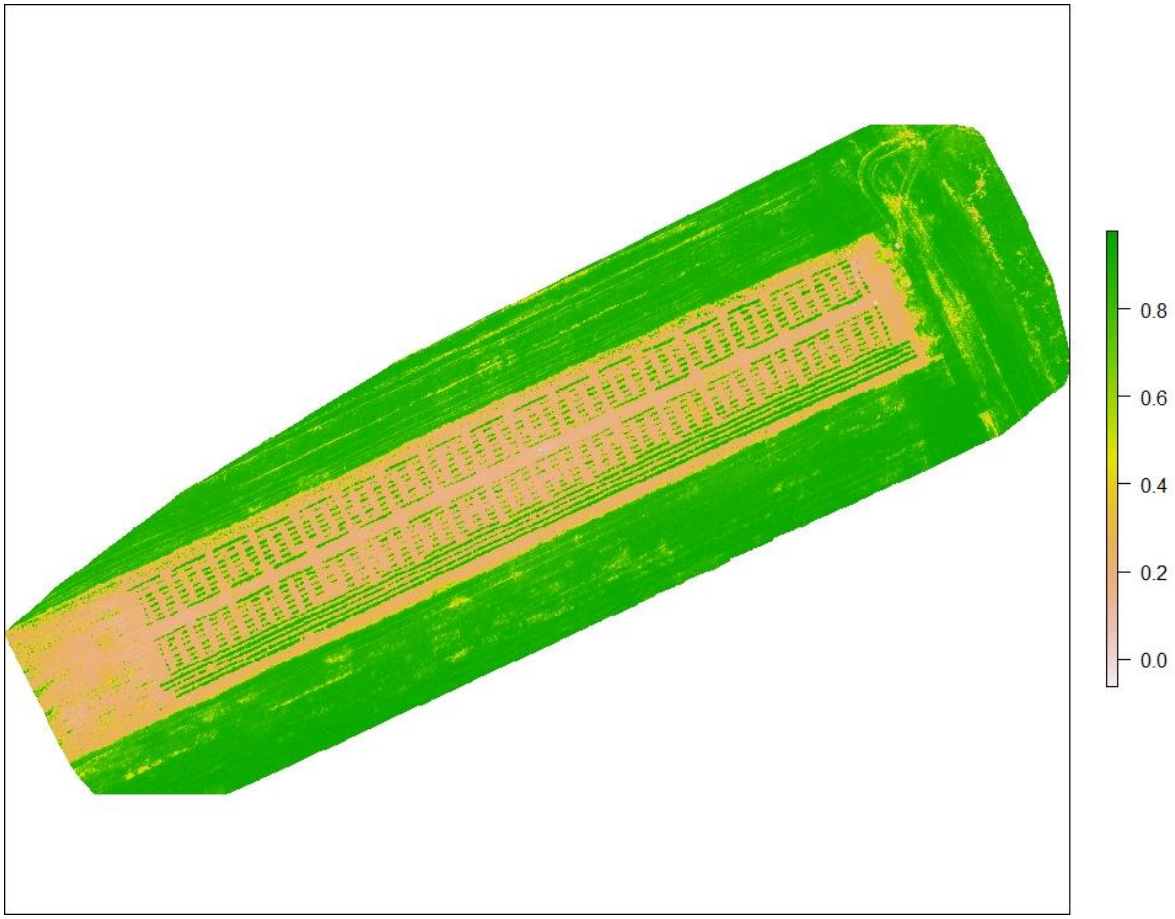

Figure 3 Aerial view of the potato field layout with digital surface model elevation (m).

2.3.2 STUDY AREA

The study was conducted at Nafferton Farm, Newcastle University, Northumberland, United Kingdom. The location of Nafferton Farm is indicated with latitude and longitude

coordinates in **Figure 4**. The farm has been followed the soil association standard for organic field to ensure that the field is certified as organic.

Figure 4 Map of Nafferton Farm, Newcastle University, UK, showing the boundary and layout of the farm. The legend shows specific field sites used for potato cultivation in three different years: 2020 (green), 2021 (yellow), and 2022 (blue).

As a result of the COVID-19 pandemic limiting physical access and availability of support personnel, trials in 2020 were carried out in conventional fields only. Potatoes were planted in both conventional and organic fields from 2021.

2.3.3 FIELD MANAGEMENT

The organic blocks were fertilised 3 weeks before planting with 180 N Kg/ha of farmyard manure, with weeds being controlled mechanically. Throughout the experiment, a consistent fertilisation application was employed for conventional fields**.** In the experimental plots cultivated, fertilisers were applied as follows:

- Nitrogen (N): 180 kg N/ha
- Phosphate (P): 134 kg P/ha

- Potassium (K): 200 kg K/ha

Fertilisers were applied in a split application to optimise nutrient intake. The initial application is at planting, where a basal dose containing a mixture of N, P, and K was applied. A second application as a top-dressing of nitrogen was applied at the tuber initiation stage to promote optimal tuber growth and development.

A detailed list of pesticides and herbicides used in this experiment field for 2020, 2021, and 2022 is described in **Table 1**. The application was timed and designed based on potato growth stages, pest, disease and weed present, and environmental conditions. However, regular field monitoring was conducted to adjust the time of application as needed.

Table 1 A detailed pesticide and herbicide application schedule was followed each year.

Year	Application date	Chemical used	Description	Application
2020	18-05-2020	Wicket	Herbicide	5 L/ha
	18-05-2020	Prosulfocarb	Herbicide	3 L/ha
	18-05-2020	Praxim	Herbicide	4 L/ha
	18-05-2020	Metobromuron	Pesticide	3 L/ha
	30-06-2020, 16-07-2020, 06-08-2020, 20-08-2020	Shirlan Fluazinam	Fungicide	300 mL/ha
	08-07-2020, 31-07-2020, 13-08-2020	Mancozeb	Fungicide	1.7 kg/ha
2021	20-05-2021	Wicket	Herbicide	5 L/ha
	20-05-2021	Prosulfocarb	Herbicide	3 L/ha
	20-05-2021	Praxim	Herbicide	4 L/ha
	20-05-2021	Metobromuron	Pesticide	3 L/ha
	10-06-2021, 22-06-2021, 13-07-2021	Shirlan Fluazinam	Fungicide	300 ml/ha

	30-06-2021	Panarex	Herbicide	2.25 L/ha
	30-06-2021, 22-07-2021	Fubol Gold	Fungicide	1.9 kg/ha
2022	26-05-2022	Wicket	Herbicide	5 L/ha
	26-05-2022	Prosulfocarb	Herbicide	4 L/ha
	26-05-2022	Praxim	Herbicide	4 L/ha
	26-05-2022	Metobromuron	Pesticide	0.5 L/ha
	10-06-2022, 23-06-2022, 15-07-2022, 22-07-2022	Shirlan Fluazinam	Fungicide	300 ml/ha
	17-06-2022, 07-07-2022, 01-08-2022	Fubol Gold	Fungicide	1.9 kg/ha

In preparation for final harvest, both organic and conventional potato blocks were flailed two to three weeks prior to the scheduled harvest date. This practice allows the hardening of the tuber skin, which is crucial for minimising damage during harvesting and enhancing the storage quality of the tubers.

2.4 RESEARCH STRATEGY

Developing an effective and comprehensive research strategy for early and accurate detection of plant diseases and plant monitoring involves considering various aspects of the problem, including technological advancements, interdisciplinary approaches, and practical implementation.

The structure of the research strategy includes, first, conducting a thorough review of existing literature to understand the current stage of plant monitoring and detection technologies. Then, identify research gaps and challenges in current methodologies and explore the advancements in sensors, imaging approaches, data analytics, and other relevant fields. Second, identify the specific objectives of the trend in the research, such as improving early detection, enhancing accuracy, or developing real-time monitoring systems. Next, determine the scope of your search, including the types of potato disease to be addressed and the differences in potato varieties.

Given that UAV-based multispectral imaging generates quantitative data on plant health and growth, we believe in objective measurement and the scientific method to interpret this data. The core proposal of this study is both exploratory and explanatory. The methodology aims to explore the potentials of the use of UAV-based imagery in potato phenotyping and explain the correlations between the imagery data and field measurement of potato growth parameters and important traits. For representativeness, stratified random sampling techniques are employed. Potato fields are designed with the aim of achieving an in-depth analysis of the potato varieties by sensing and ensuring external factors and variations are accounted for and minimised.

2.5 GROUND TRUTH AND PROXIMAL DATA COLLECTION

The objective of manual or proximal measurement, including visual disease assessment and morphological ground truthing data, was to validate the spectral data and determine the potential relationship between the estimated data obtained from a UAV-based aerial imagery approach and the data obtained from the proximal method. Phenotyping with proximal sensors allow high-precision measurements of plant traits both in controlled conditions and in the field (Odilbekov *et al.*, 2018).

The data collection to broadly understand the pattern of plant growth and disease incidence in 300 different potato varieties was done simultaneously using different approaches, including visual rating for ground-truthing and spectral datasets using UAV-based aerial imagery. This was done throughout the growing season (May – September) and aerial imagery and field measurement were coincided wherever possible in the 2020 and 2022 trials. Using their multifactored approaches ensured that we obtained a robust and comprehensive dataset for analysis, giving insights into various nuances of potato growth and dynamics. For plant and canopy growth assessment, the conventional method of assessing plant performance involves physically walking through the plots and performing the visual rating. Field visual assessments were employed for the evaluation of crop emergence, plant maturity, and disease diseases in potato crops, as detailed below.

2.5.1 WEATHER DATA

The weather data recorded at Nafferton Farm includes average wind direction (in degrees), maximum wind speed (in metres per second), average wind speed (in metres per second), average temperature (in degrees Celsius), average atmospheric pressure (in

hectopascals), average relative humidity (in percent), average dew point (in degrees Celsius), average soil temperature (in degrees Celsius), total rainfall (in millimetres), and average solar radiation (in Watts per square metre), obtained from the weather station (**Table 2**).

Table 2 Weather data recorded at Nafferton Farm from 2020 to 2022 (average).

Parameter	year	month					
		April	May	June	July	August	September
Relative humidity (percentage)	2020	95.64	92.1	89.08	92.65	92.62	96.71
	2021	86.96	94.16	87.43	87.87	85.82	89.77
	2022	94.39	91.5	84.4	94.97	88.89	86.79
Temperature (degree Celsius)	2020	8.63	11.52	13.44	14.2	15.17	12.56
	2021	5.36	8.91	14.15	16.28	14.74	14.48
	2022	7.84	11.73	14.09	16.89	16.33	13.89
Soil temperature (degree Celsius)	2020	8.04	9.71	11.92	12.68	14.07	12.41
	2021	6.77	9.35	13.37	15.39	14.72	14.01
	2022	8.26	10.73	13.08	15.77	15.86	14.26
Total rainfall (millimetres)	2020	0.001	0.012	0.017	0.016	0.03	0.02
	2021	0.006	0.025	0.006	0.027	0.023	0.018
	2022	0.007	0.018	0.009	0.01	0.01	0.036
solar radiation (Watts per square meter)	2020	164.63	219.71	162.96	181.39	128.54	101.21
	2021	176.61	167.8	217.76	184.75	134.75	94.87
	2022	160	177.18	220.79	195.11	170.79	118.1

The weather data for 2020, 2021, and 2022 needed to be structured in a format that could be easily correlated with the crop parameter data. This can be done by aligning the

weather data with the corresponding growth stages of the potato crop. Then calculate the average of each weather parameter for each growth stage.

- April and May (initial or establishment stage)
- June (tuber initiation stage)
- August (maturation stage)
- September (senescence)

Finally, combine the structured weather data with the crop data for further analysis.

2.5.2 CROP EMERGENCE/ NUMBER OF PLANTS PER PLOT

Potato plant emergence involves visually counting the number of individual plants within the plot by physically inspecting the field to determine how many plants have emerged from the ground.

2.5.3 DISEASES

Potatoes can be infected by many different biotic and abiotic stresses that can reduce yield and tuber quality. This assessment focuses on the crop tolerance or resistance of stresses ensuing the selection of resilient varieties. This included fungal pathogens: including *Phytophthora infestans* (casing late blight) and *Alternaria* spp., and Viruses: including potato virus Y (PVY) and potato leaf roll virus (PLRV).

For field evaluation of late blight resistance during vegetation growth, the visual observation of late blight symptoms was carried out for each plot using the degree of infection (1–10-degree scale). Late blight symptoms on leaves generally appear as irregular dark spots that enlarge as the disease develops (**Figure 5**). On the upper surface, a lighter green halo often surrounds the necrotic area, and on the lower surface, whitish sore-bearing mould develops around the lesions under moist conditions. Temperature and relative humidity are critical in determining the risk of infection (AHDB, 2023). The estimated scale value in this study was adopted from Hansen *et al.* (2005) as detailed in **Table 3**.

Figure 5 Late blight symptoms observed in the field.

Table 3 Rating system for assessing late blight symptoms during field trials.

Rating Score	Plant health indicator
10	Overall excellent plant health; no disease symptoms visible
9	Less than 5% of leaves developed the irregular dark spots; overall healthy plant
8	Around 10 % of leaves developed the irregular dark spots
7	Around 20 % of leaves developed the irregular dark spots
6	Around 30 % of leaves developed the irregular dark spots
5	Around 40 % of leaves developed the irregular dark spots
4	Around 50 % of leaves developed the irregular dark spots
3	Around 60 % of leaves developed the irregular dark spots
2	Around 80 % of leaves developed the irregular dark spots
1	Completely dead

Table 4 Estimation of the resistance of potato to *P. infestans* according to 10-degree scale (modified from Sieczka (2001)).

Degree of resistance	Symptoms of infection	Infection %	
		range	mean
10	No symptoms	0.0 – 0.0	0.0
9	No symptoms, occasional necrotic spots	0.0 – 0.5	0.2
8	Occasional spots on individual plants, 2 leaves infected	0.6 – 2.3	1.1
7	Slight infection on 9 leaves	2.4 – 9.5	4.7
6	All plants infected, about 20% leaves blighted	9.6 – 32.1	18.3
5	50% of leaves blighted, petioles infected	32.2 – 67.9	50.0
4	80% of leaves blighted, petioles and stems infected	68.0 – 90.4	81.7
3	Heavy infection, about 9 leaves healthy	90.5 – 97.7	95.3
2	Very heavy infection, individual leaves green	97.8 – 99.5	98.9
1	Plants completely blighted; occasional parts of steams non infected	99.6 – 100.0	99.8

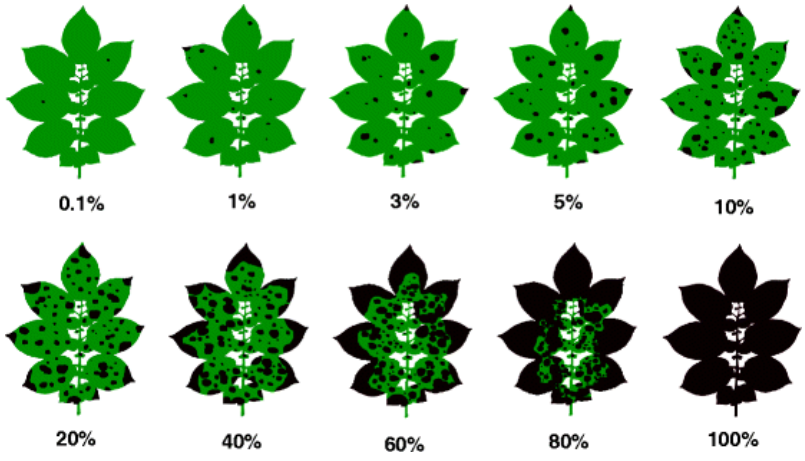

Figure 6 Area diagrams for early blight severity scoring (Duarte *et al.*, 2013).

Field evaluation for Ecobreed project (page number 140) will be using **Figure 6** and **Table 4** Estimation of the resistance of potato to *P. infestans* according to 10-degree scale (modified from Sieczka (2001)).

2.5.4 SENESCENCE

Senescence was carried out by field visual assessment using the below rating system (**Table 5**).

Table 5 Senescence rating system for potato field trials

Senescence rating score	Plant health indicator
1	Overall excellent plant health; no visible browning or yellowing of leaves
2	Some browning and yellowing of leaves; overall healthy plants
3	Around 75 % of green material remains, with some yellowing of the leaves
4	Around 50 % of green material remains, leaving open gaps in rows
5	Around 25 % of green material remains, with a significant gap in rows
6	Around 10 % of green material remains, most of the row is exposed
7	Completely dead plant (no visible green material, including stems)

2.5.5 VIRUSES

Potatoes are susceptible to many different viruses that can negatively impact on both yield and tuber quality. Common indicators of viral infection include mosaic patterns on the foliage, stunting of the plant, and the malformation of both leaves and tubers. However, symptoms are not always expressed due to the interactions between the virus and the potato plant, growing conditions such as soil fertility and weather, or the stage of the plant when it is infected. Two of the most important viruses affecting potatoes are potato virus Y (PVY) and potato leafroll virus (PLRV). During this study, the specific type of virus present in the field was not identified: rather, observations were made whenever virus symptoms appeared in any plots. This observation will be instrumental for future disease management strategies.

2.5.6 PLANT HEIGHT

Plant height is an essential trait for assessing the growth and health of most crops, making it a key factor in crop screening processes. Measuring plant height is straightforward and can be done manually with simple tool like a ruler (Perez-Harguindeguy *et al.*, 2016). In the context of this study, plant heights were determined by using a measuring ruler on each

plot, with measurements taken from the top of the ridge to the apex of the plants. To estimate the average height for each plot (variety), measurements were taken from three different samples within the plot, ensuring a more accurate representation of the plant height across the entire plot.

2.5.7 GROUND COVER OR LEAF AREA INDEX (LAI)

The density of the potato canopy was estimated using the Leaf Area Index (LAI), which is a measure of leaf area per unit of ground area. To assess light interception, the Photosynthetic Active Radiation (PAR) was measured. Both LAI and PAR were recorded on a monthly basis starting from the tuber initiation stage, using Decagon's AccuPAR model LP-80 LAI/PAR Ceptometer (Pullman, WA, USA) as illustrated in **Figure 7**. Additionally, PAR measurements can also be used in conjunction with climate data to estimate the biomass production of the canopy.

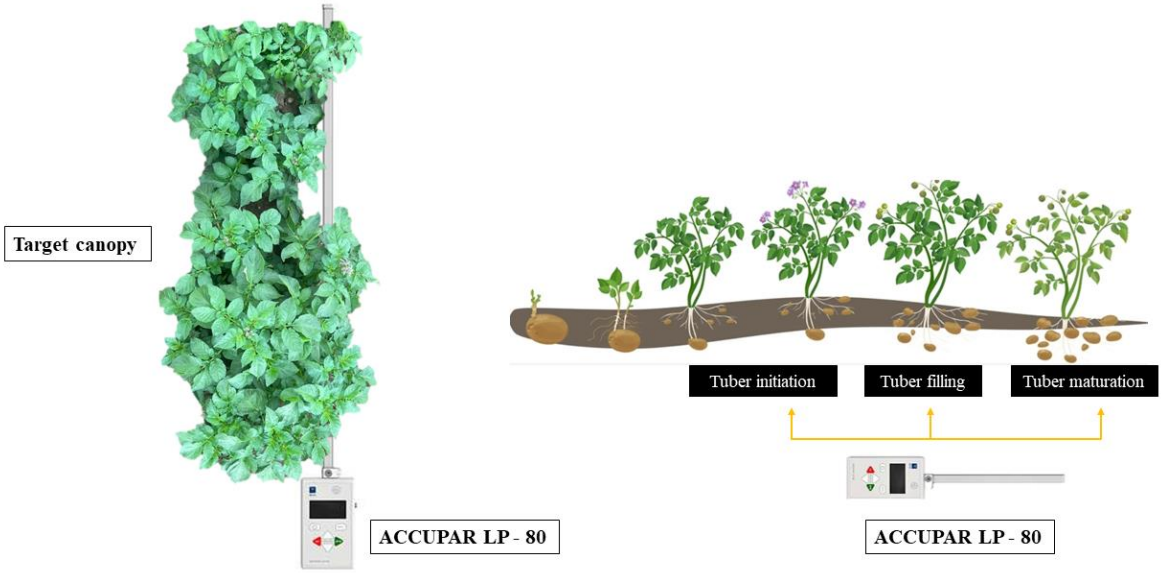

Figure 7 Measurement of LAI using the ACCUPAR LP-80 (ceptometer). The left image shows the sensor aligned with a potato canopy to measure the light interception via photosynthetic active radiation (PAR) within the canopy. The right image shows that the ceptometer is used to measure the potato canopy throughout the growth stages.

2.5.8 VEGETATION INDICES (VIS)

Vegetation indices (VIs), such as the Normalised Difference Vegetation Index (NDVI) and the Normalised Difference Red Edge Index (NDRE), are commonly used to assess plant health and vigour, namely extracted from handheld sensor and remote sensing data. These indices utilise specific spectral bands to calculate vegetation indices, providing information on crop health, biomass, vigour, and photosynthetic activity. For this study, NDVI and NDRE were measured using the handheld three-channel photodetector, RapidSCAN CS-45 device (Holland Scientific, Lincoln, the Netherlands) (**Figure 8**). This sensor measures crop reflectance in the 670 nm, 730 nm, and 780 nm wavelengths using an integrated polychromatic light source. This sensor is capable not only of capturing NDVI and NDRE but also recording latitude and longitude coordinates, along with basic reflectance information about the target plant.

Figure 8 Greenseeker and RapidSCAN handheld devices for ground-based vegetation indices.

Measurements were taken at each plot within the potato canopy, starting from the tuber filling stage through to the maturity stage of the potato growth cycle. To optimise accuracy and consistency in data collection, the sensor was maintained at a level and positioned around 60 - 75 cm above the canopy, with all reading conducted by the same person. This standardisation of the measurement process (illustrated in **Figure 9**) ensures the reliability of the VIs in reflecting the actual health and vigour of the potato crops.

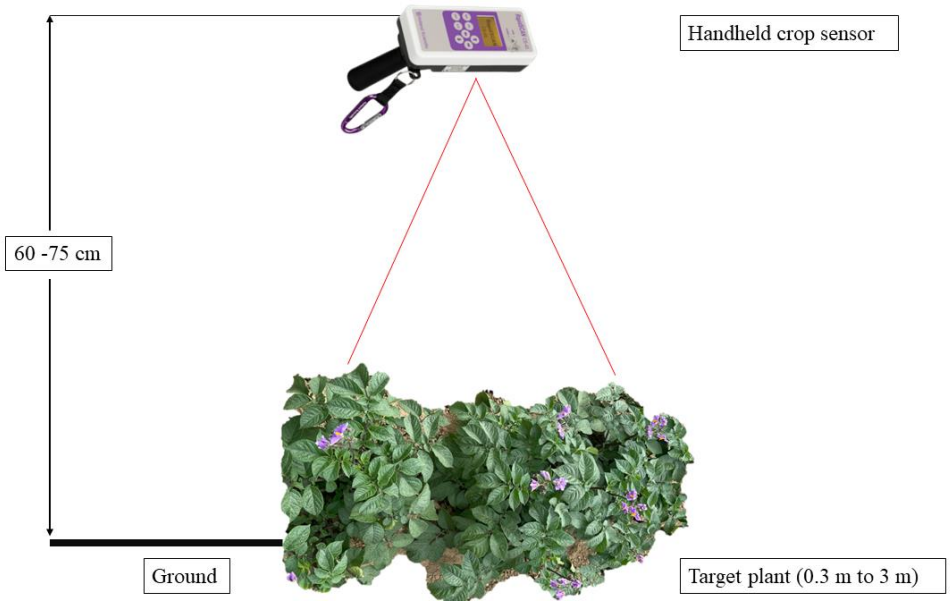

Figure 9 The use of handheld crop sensors to collect vegetation indices. Red lines indicate the field of vision of the sensor and outline the target dimension for optimal sensing distances. It highlights the significance of maintaining a uniform distance and angle to ensure the precision of the measurement.

2.5.9 POTATO HARVEST

Total tuber yield was determined at each plot (for each potato variety). At each plot, all potato tubers were harvested and stored in a net bag by hand. The harvested tubers were then weighted (kg) and recorded. The total tuber yield of each variety and each production system was then divided into three categories: high, medium, and low yield. Then the yield from each variety will be graded and kept for next year growing season.

2.6 UAV-BASED DATA COLLECTION

The planning document for a UAV mission with a scheduled flight for the purpose of studying potato phenotyping using UAV-based imagery was done in ScaleFlyt environment (**Figure 10**). Where flight missions were conducted through the MicaSense application (**Figure 11**). Regarding the sensor equipped with the UAV platform, the planning and collection approaches are similar but depend on the flight mission.

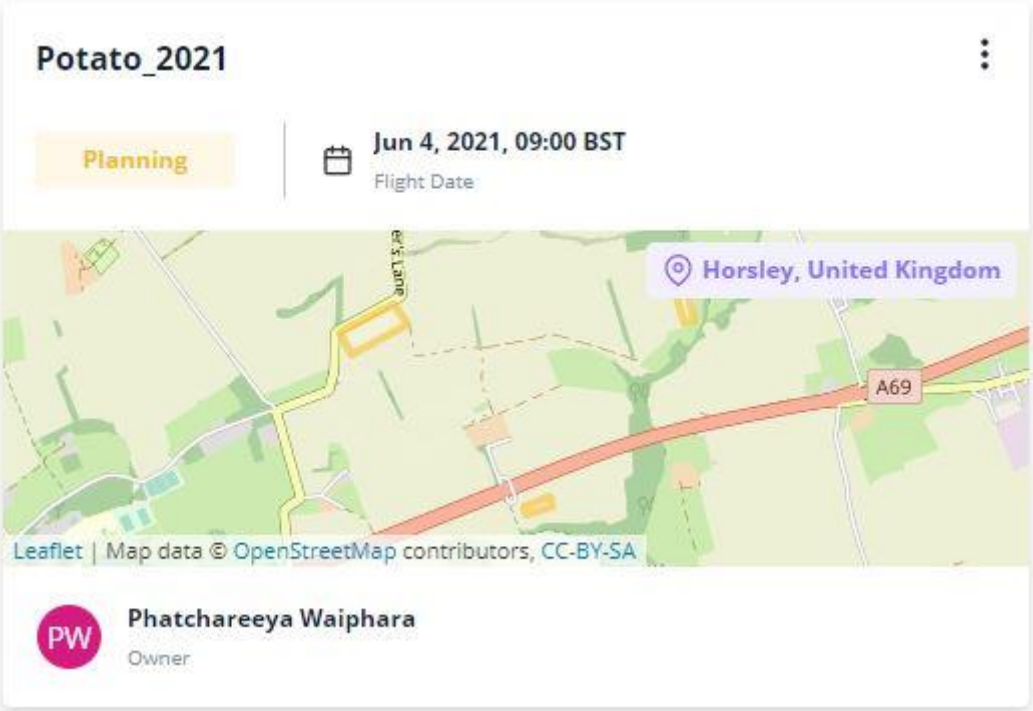

Figure 10 The interface of UAV flight planning prior to the flight date.

2.6.1 UAV FLIGHT PARAMETERS

Aerial images were captured using an unmanned aerial vehicle. Two platforms were used: a multi-rotary wing and a fixed-wing UAV. The multi rotary-wing UAV, a multi-rotor DJI M200 and DJI M600, is equipped with a MicaseseRedEdge-M five-band multispectral sensor (MicaSense, Seattle, USA). The multispectral sensor captures images in blue, green, red, NIR, and red-edge bands. The UAV with a remote sensing system was operated at an altitude of 30 metres with 85 percent image overlap. The camera was set to capture images automatically in a given space of one image every second. The corresponding spatial resolution for this data is 1 cm/pixel. The irradiance sensor has filters to detect the amount of radiation from direct sunlight or diffuse light in cloudy weather conditions. The first UAV flight was performed when over 50 percent of the plots had emerged, approximately 40 days after planting (sprout development stage). Fixed-wing UAV equipped with RGB camera was collected aerial images at a 75-metre altitude with 80 percent image overlap. Details of flight dates and their correlation with potato growth stages are shown in **Table 6**.

Table 6 Detailed flight schedules in correlation with potato growth in the 2020, 2021, and 2022 trials.

Planting date	Flight No	Field	Flight date	Day after planting (DAP)	Potato growth stage
05-05-2020	1	Conventional	25-06-2020	50	Initial/ Establishment
	2	Conventional	02-07-2020	57	Tuber initiation
	3	Conventional	16-07-2020	72	Tuber Initiation
	4	Conventional	24-07-2020	80	Tuber Filling
	5	Conventional	07-08-2020	95	Tuber Filling
	6	Conventional	10-08-2020	98	Maturation
	7	Conventional	20-08-2020	108	Maturation
	8	Conventional	08-09-2020	119	Maturation
30-04-2021	1	Conventional and Organic	08-06-2021	39	Initial/ Establishment
	2	Conventional and Organic	07-07-2021	68	Tuber initiation
	3	Conventional and Organic	23-07-2021	84	Tuber filling
	4	Conventional and Organic	10-08-2021	102	Maturation
04-05-2022	1	Conventional and Organic	30-05-2022	25	Initial/ Establishment
	2	Conventional and Organic	20-06-2022	46	Tuber initiation
	3	Conventional and Organic	11-07-2022	67	Tuber filling

	4	Conventional and Organic	05-08-2022	92	Maturation

The data for each year shows the progression through different stages of potato growth from initial establishment to tuber maturation (**Figure 12**). The flight dates provide the observation at key stages of potato crop development as this will be used to analysing the growth pattern of the potato crop correlating it with other data such as yield and weather condition to gain a deeper insight into the factors that affecting each potato variety for the production.

The variation in flight number in **Table 6** is primarily due to differences in field management and operational constraints. In 2020, we conducted 8 UAV flights as we are only a conventional field to manage (COVID-19 disruption). This allows for more flexibility to schedule additional flights across all critical growth stages of potatoes. In 2021 and 2022, the study expanded to include both organic and conventional fields, which increased workload and the need to balance operations between the two sites (**Figure 4**). As a result, we were able to conduct 4 flights each field (8 flight in total). These four flight were timed to cover all key potato growth stages and the UAV flights was conducting for both fields on the same day.

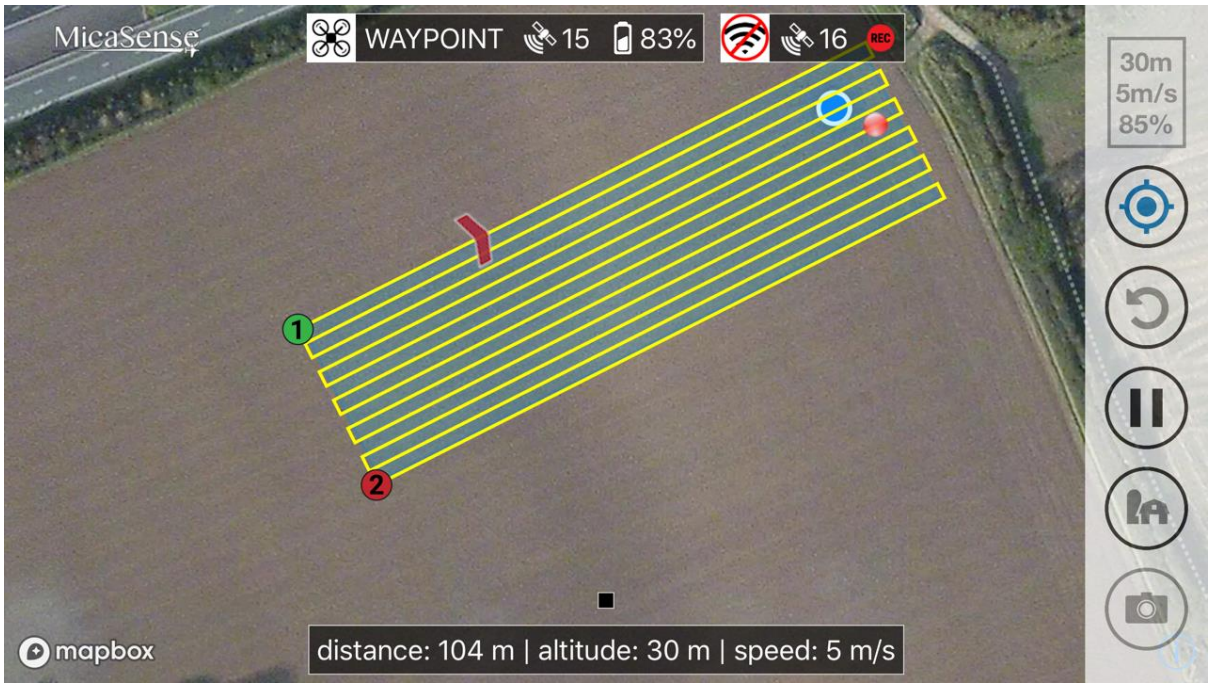

Figure 11 A user interface for a drone flight planning application connected with a MicaSense multispectral camera. Additional information on the interface includes a flight parameter.

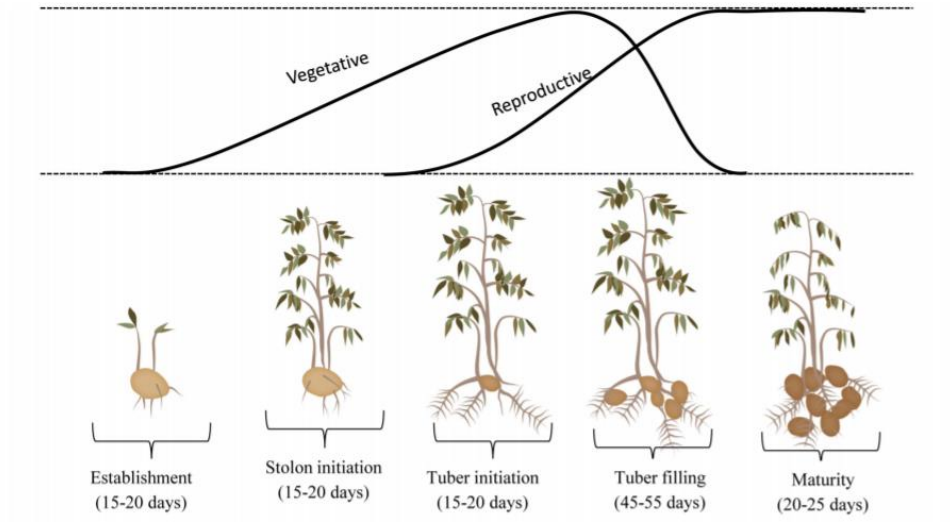

Figure 12 Potato growth stages modified from Obidiegwu *et al.* (2015)

2.7 IMAGE PROCESSING

Regarding the platform from which multispectral data is acquired, the processing steps and methods are similar, except for the pre-processing stage (Adao *et al.*, 2017). In this study, the computation analysis methods developed for processing raw field spectra to calibrated reflectance measurements for potato field. This includes (i) calibration of reflectance panels and radiometric calibration; (ii) computing the ground point control (GPC); (iii) image alignment and export the final output files.

2.7.1 PHOTOGRAMMETRIC PROCESSING

This step is aimed at the radiometric calibration and conversion of raw sensor readings into reflectance for Micasense Rededge images by following the procedure indicated by the original manufacturer (Micasense Inc., Seattle, USA). The certified reflectance data of our Micasense Rededge Calibrated Reflectance Panel (CRP) consists of five reflectance values, including red, green, blue, red edge, and near-infrared (**Figure 13**). In this process, the goal is to obtain calibrated reflectance values for further detailed analysis. It is important that the reflectance panel measurements be calibrated. This is because the KGCs may not fully reflect all incident radiation, and their reflectance might vary significantly depending on the illumination angle (solar zenith angle) (Peddle *et al.*, 2001).

2.7.2 RADIOMETRIC CALIBRATION

Radiometric calibration is the process of converting raw sensor readings (digital numbers, DN) into physical units using radiance values. This process takes into account sensor characteristics, such as sensor sensitivity, noise, and other factors that might affect the recorded values (Lillesand *et al.*, 2015). The calibration helps ensure that the data collected has an accurate, reliable, and consistent relationship between the recorded DN and the actual radiance of the images captured. Radiance is the amount of light that comes from a scene and strikes the sensor. It is a quantifiable physical property. In essence, radiometric ensures that the digital images captured by sensors are accurate and consistent and represent field scenes.

2.7.3 CONVERSION TO REFLECTANCE

Where reflectance is the ratio of the amount of light that is reflected by a surface to the amount of incident light. This step involves converting the radiance values obtained from the calibration of the calibrated sensor data into reflectance values. Reflectance values are most meaningful for analysis because they account for variations in illumination and atmospheric conditions. The Micasense Rededge calibrated reflectance panel (CRP) is a known reflectance target with calibrated reflectance values (**Figure 13**). By capturing images of this panel under consistent lighting conditions, the Micasense Rededge camera can calibrate the sensor data it has gathered. This process can be used to establish a relationship between the recorded sensor values (**Table 7**) and known reflectance values with sensor responses. Then, the calibrated reflectance values are used to apply a correction factor to the images taken in the field, ensuring the data's accuracy and comparability over time and across different conditions.

Figure 13 The calibrated reflectance panel, serial number RP04-1826003-SC

Table 7 The five reflectance values (albedo values) correspond to different wavelengths captured by the Micasense RedEdge-M camera.

RedEdge-M band	Centre wavelength (nm)	Band width (nm)	Albedo values	Percentage
Blue	475	20	0.54	54.08%
Green	560	20	0.54	54.08%
Red	668	10	0.54	53.93%
Red-Edge	717	10	0.54	53.82%
Near infrared (NIR)	840	40	0.53	53.47%

Note that a surface's albedo values represent the percentage of sunlight it reflects. In the context of multispectral imaging, these values can help in understanding the reflective properties of surfaces often used in vegetation health monitoring or other land cover types. The

albedo value is the calibrated reflectance value for specific wavelengths of the electromagnetic spectrum. These reflectance values are measured relative to the amount of incident solar radiation at those wavelengths, specifically for the Micasense RedEdge-M camera. These albedo values allow for more accurate calibration and correction of the remote sensing data collected by the sensor, which converts raw sensor readings into accurate reflectance values. By using a known reference panel, the final reflectance image has been adjusted for variations in illumination, atmospheric conditions, and sensor sensitivity to obtain more comparable reflectance values from the target area.

2.7.4 PIX4D PROCESSING WITH CORRECTED IMAGES

The corrected digital images collected from the UAV were pre-processed using the Pix4DMapper software version 4.4.12 (Pix4D, SA). Pix4D Mapper is a commercial image processing software. This preliminary phase is aimed at obtaining calibrated data from the Pix4D structure from motion bundle adjustment capabilities. This is for the generation of multispectral orthomosaic and point clouds. During the pre-processing step, a calibrated reflectance panel was used for radiometric calibration, and ground control points (GCPs) were manually added. Four geo-referenced GCPs, providing coordinates using Network Real-Time Kinetic (RTK), were used in all flights to improve the accuracy of the project and reduce the noise. This enabled us to create the final products, including a geo-referenced orthomosaic image, the reflectance map, the Digital Surface Models (DSM) and the Digital Terrain Model (DTMs), which were stored in TIFF format. This format is commonly used for storing raster data and is suitable for further analysis and visualisation (**Figure 14**).

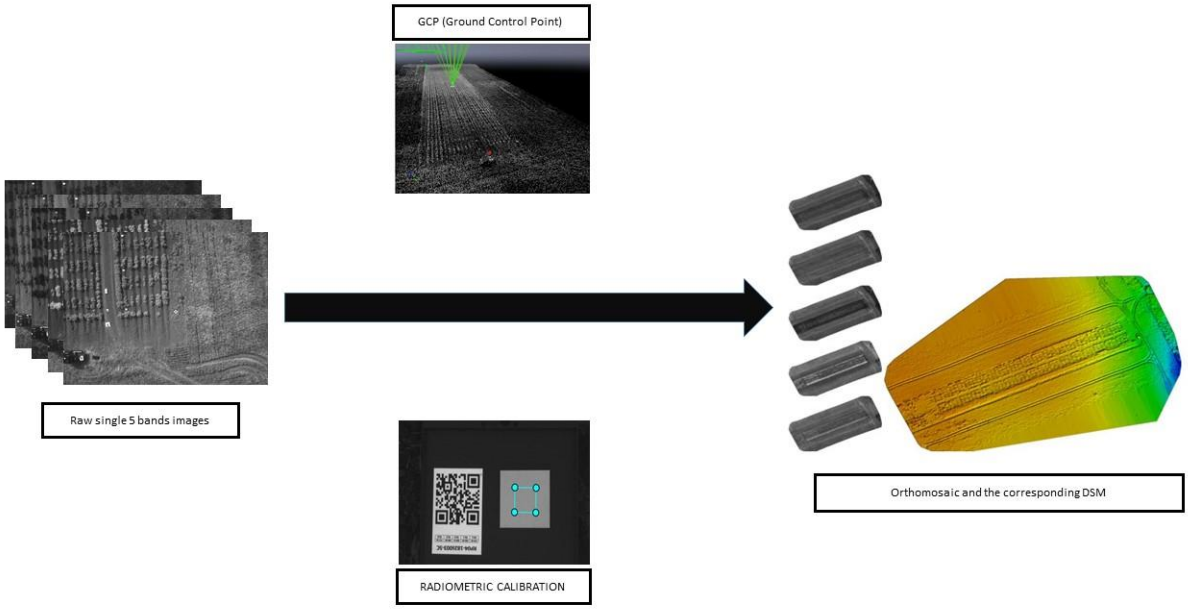

Figure 14 Pre-processing flowchart in Pix4D software involves radiometric calibration, the input of ground control points (GCPs), the final orthomosaic image, the corresponding DSM and DTM.

2.8 DATA EXTRACTION

After successfully calibrating and processing of multispectral images using Pix4D Mapper, the focus shifted towards in-depth analysis and extraction of specific plant-related features, including plant counts, canopy cover area percentage, vegetation indices, and plant heights. This comprehensive analysis was carried out in the R programming language using R Studio software (Team, 2021) using the following packages: Raster, SP, and FIELDImageR (Matias *et al.*, 2020) to handle various aspects of the imaging process, spatial data manipulation, and field image analysis. Running the script and loading the function in R from the GitHub repository. The FIELDImageR package in R is specifically designed for high-throughput field phenotyping image analysis. The plant feature extraction processing used the imported multispectral images, which were cropped to reduce the size and processing time. This merged image was used to generate the plot shape file, which essential for computing vegetation indices and extracting plant feature values by segmenting the soil background. This approach underscores the integration of advanced image processing techniques and spatial analysis in extracting data from multispectral images.

2.8.1 IMAGE SEGMENTATION

The image segmentation approach was carried out to extract the pixels defining the plant or crop area from the orthomosaic image, by removing the background that included soil. An accurate vegetation index is required to identify plant biomass versus soil and residue background (Haverkort and Struik, 2015). This was done using the overall HUE index, which was used to mask the soil from the green plant. The segmentation approach using the HUE index transformed the RGB image into a binary image, where positive values were assigned to soil pixels and zero values to plant pixels. This demonstration is in **Figure 15**, where the result image shows only green plants. In order to extract the values from individual plots, a shape file was created and overplayed on the binary image file to distinguish individual plots.

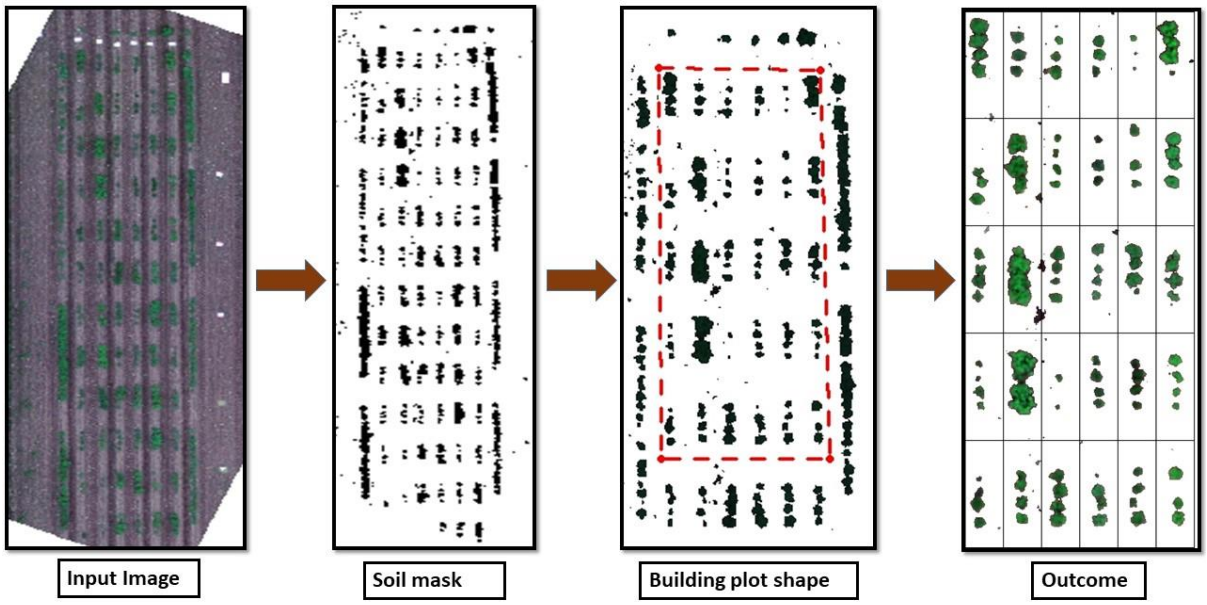

Figure 15 A step-by-step approach to segmenting and processing a potato field image to distinguish plants. The outcome involved segmenting the field into small plots, and each plot has a collection of plants that are clearly distinguishable from the soil background.

The process visually illustrates how imaging data can be separated into individual plots for further plot data extraction.

2.8.2 VEGETATION INDICES

Vegetation indices (VIs) serve as quantitative tools to measure various crop traits or characteristics, including chlorophyll content, leaf area, and overall health. VI images were computed for each UAV flight in R language software. The computation was based on the reflectance reading captured from different spectral bands from multispectral images. The VIs, formular, and their description used in this study as shown in **Table 8**. The binary images of potato fields were generated using the overall HUE threshold computed with the FIELDImageR package in R Studio (Matias *et al.*, 2020).

Table 8 Describes the vegetation indices and formular used in this thesis study.

Index	Formula	Description	Related trait	Reference
HUE	Atan(2*(B-G-R)/30.5*(G-R))	Overall Hue index	Soil colour	Matias *et al.* (2020)
BGI	B/G	Blue green pigment index	Chlorophyll and LAI	Richardson and Wiegand (1977)
BI	Sqrt((R^2+G^2+B^2)/3)	Bare soil index	Vegetation coverage and water content	Richardson and Wiegand (1977)
CVI	(NIR*R)/(G^2)	Chlorophyll vegetation index	Chlorophyll	Vincini *et al.* (2008)
DVI	NIR-RE	Difference vegetation index	Chlorophyll and nitrogen	Jordan (1969)
EVI	2.5*(NIR-R)/NIR+(2.4*R) +1	Enhance vegetation index	Chlorophyll and biomass	Huete *et al.* (2002)
GLI	(2*G-R-B)/(2*G+R+B)	Green leaf index	Chlorophyll	Louhaichi *et al.* (2001)
GNDVI	(NIR-G)/(NIR+G)	Green normalised difference vegetation index	Chlorophyll, biomass, yield and LAI	Gitelson *et al.* (1996)

NDVI	(NIR-G)/(NIR+G)	Normalised difference vegetation index	Chlorophyll, nitrogen, and maturity	Tucker *et al.* (1980)
NDRE	(NIR-RE)/(NIR+RE)	Normalised difference red edge index	Chlorophyll and maturity	Gitelson and Merzlyak (1994)
SAVI	(NIR-R) (1+L)/(NIR+R+L)	Soil-adjusted vegetation index	Soil colour	Huete (1988)
RVI	NIR/R	Ratio vegetation index	Biomass, water content, and nitrogen	Bannari *et al.* (1995)
TVI	0.5*(120*(NIR-G)-200*(R-G)	Triangular vegetation index	Green LAI, chlorophyll, and canopy	Broge and Leblanc (2001)
VARI	(G-R)/((G+R-B)	Visible atmospherically resistant index	Canopy, biomass, and chlorophyll	Gitelson *et al.* (2002)

After VIs were computed (**Figure 16**), VI values were subtracted from the binary image for each individual plot using the plot shape file.

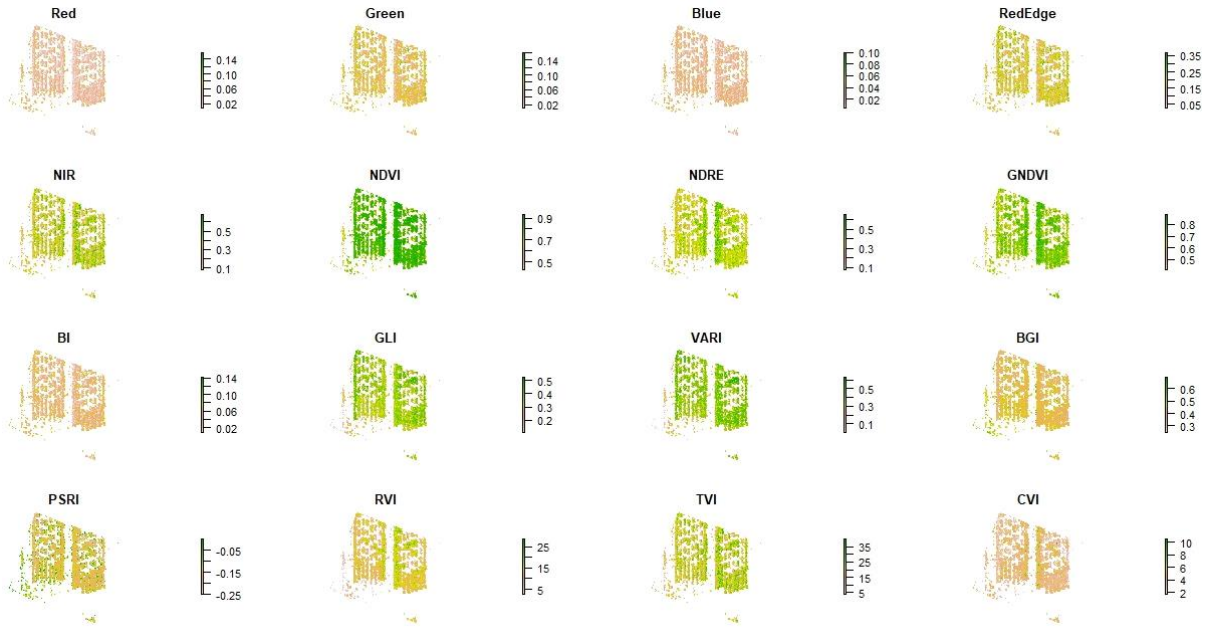

Figure 16 Spatial representations of various vegetation indices and reflectance values used in this study. Each map corresponds to a different spectral index including, Red, Green, Blue, Red Edge, NIR, NDVI, NDRE, GNDVI, BI, GLI, VARI, BGI, PSRI, RVI, TVI, and CVI. The colour scales on the right-hand side of each map represent the range of values for each index and reflectance band, with different colours corresponding to different levels of index values.

2.8.3 NUMBER OF PLANT PER PLOT

UAV-based images are used to determine the number of plants per plot, commonly known as crop emergence. The aerial images are taken during the early stage of potato growth (sprout development or vegetative growth). After aerial image collection, the images are then processed to determine and count the number of plants that have emerged in each plot. This activity is usually carried out while the potato plants are still small, which aligns with the operational capabilities of the software used for this task. The early-stage imaging ensures that each plant is clearly identifiable and separable for an accurate counting process, before the plants grow larger and their canopy begin to overlap.

2.8.4 ESTIMATED PLANT HEIGHT

In this study, the plant height data, derived from multispectral images using R and Quantum Geographical Information System (QGIS) software, is consistent in its numeric form, yet the methodologies and tools for data extraction and analysis may vary between the software environments. A Crop Height Model (CHM) was developed through a pixel-wise subtraction,

where altitudes from a Digital Terrain Model (DTM) were subtracted from a Digital Surface Model (DSM) (**Figure 17**). This method facilitated the extraction of estimated height values for each plot, which are quantified in meters. These height estimations are crucial for assessing plant growth stages, uniformity across the field, and identifying potential stress factors affecting crop development. To validate these estimates, ground-truth measurements were taken directly from the field and used as benchmarks to ensure the accuracy and reliability of the height data. Ground-truth measurements taken in the field served as reference points.

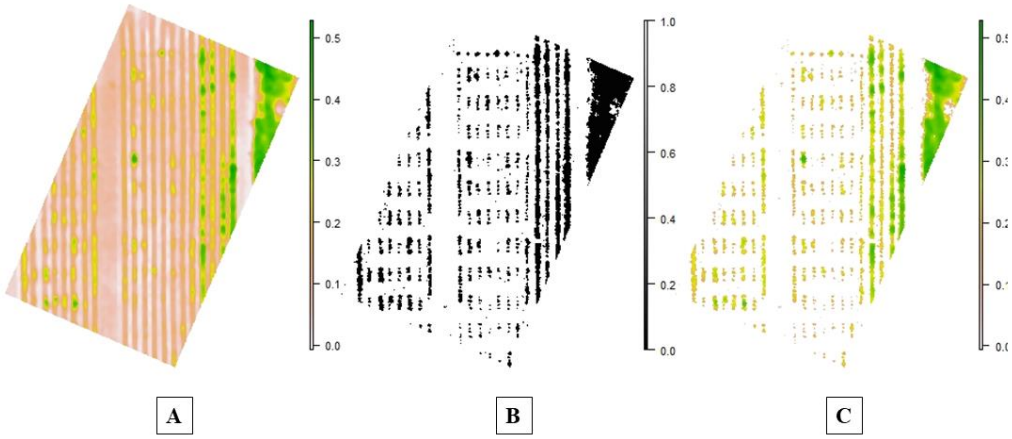

Figure 17 Key components of the crop height estimation process. (A) The elevation of the field's surface model, including objects on it (DSM). (B) The image describes how the canopy height model (CHM) for individual plots is created by pixel-wise subtraction of ground elevation (DTM) from the elevation surface model (DSM), which allows for the extraction of plant heights. (C) The final CHM map presents the plant canopy after the background removal process, with different colours corresponding to different heights, and the height values are presented in metres.

2.8.5 PLOT DATA EXTRACTION

The height data extracted from multispectral images in R and QGIS may not be inherently different in terms of the numeric values themselves. However, the tools and processes used to extract and process the data may differ. A crucial step in estimating above ground canopy height from a UAV is the identification of ground points and accurate

reconstruction of the DTM and DSM to a produce canopy height model (CHM) from the digital elevation model (Anderson II *et al.*, 2019). In this study, a crop height model (CHM) was performed using a pixel-wise subtraction of the digital terrain model (DTM) altitudes from the digital surface model (DSM). From this model, we are able to extract estimated height values for individual plots. These values are presented in meters and provide insights into plant growth stages, uniformity, and potential stresses. To ensure the accuracy and reliability of the height data, a validation process was undertaken. Ground-truth measurements taken in the field served as reference points.

Canopy height, canopy area, and VIs for each individual plot are extracted and presented in the Excel spreadsheet. The unit of measurement is in metres. Then these data points correlate with the plant growth stages according to **Table 6.** For example, **Figure 17** demonstrates the crucial step involved in separating the canopy height model (CHM) from soil in order to extract the height values. This visual representation provides an intuitive understanding of the spatial distribution of plant height across the field. The map uses a colour gradient to indicate varying plant heights, with each colour representing a different height level, quantified in meters.

2.9 THREE-DIMENTIONAL MODEL OF POTATO CANOPY PROCESSING

In order to construct a three-dimensional (3D) model of potato plants or canopy, a different unmanned aerial vehicle (UAV) equipped with a camera suitable for reconstructing the 3D point cloud data is required. The details of the processing and UAV flight parameters are provided below.

2.9.1 UAV FLIGHT PARAMETERS

Aerial images were captured in collaboration with Survey Solution Scotland using a fixed wing UX5 HP UAV (Trimble, Sunnyvale, California, United States). The UX5 HP used Global Navigation Satellite Systems (GNSS) post processed kinematic techniques to determine the UAV trajectory. The images were taken using a Sony Alpha7R 36MP full frame 35 mm RGB camera with a custom made Voigtlander 35 mm lens. The 35 mm lens was selected to deliver a 1.0 cm GSC at 75 m AGL. Given the importance of the canopy volume in this research a UAV sensor with a global shutter was chosen rather that a sliding shutter as this greatly reduce noise in the images which leads to clearer and more precise images. The data was

collected at 75 m altitude with overlaps of 85% (both front and side) between neighbouring images.

The irradiance sensor has filters to detect the amount of radiation from direct sunlight or diffuse light in cloudy weather conditions. The first UAV flight was performed when over 50 percent of the plots had emerged, approximately 40 days after planting (sprout development stage). Fixed-wing UAV images were collected with an RGB camera quipped. Fix-wing surveys were conducted at a 75-metre altitude with 80 percent image overlapping.

2.9.2 UAV DATA PRE-PROCESSING

Raw GNSS data is recorded using the UX5 HP UAV by the on-board 366-channel multi-constellation GNSS receiver, which is downloaded at the end of the flight and processed against a local base station situated within our flight area. The local based position was established by processing the local base against Ordnance Survey CORS stations. Post-processed kinetic (PKK) was used to create the trajectory as it is more robust than alternative methods which may rely on radio or other communications.

2.9.3 IMAGE PROCESSING

The UAV trajectory was processed and analysed using Trimble Business Centre (TBC) software versions 5.9 for 2021 and 2022 datasets. This involves the following activities: utilising Ground Point Control (GCP) referencing, generating point cloud data, creating digital surface models (DSM), creating digital terrain models (DTM), and manually marking plot boundaries. The point cloud data was then used to generate an individual 3D model for each plot canopy, calculating difference models and extracting canopy characteristic values such as canopy cover, canopy height and canopy area.

Alternatively, the point cloud method identifies ground points within each DSM and creates an independent DTM for each plot (Anderson II *et al.*, 2019). **Figure 18** demonstrates that the point cloud method produced a 3D model of each potato plot using SfM photogrammetry from high-resolution, high-altitude UAV images of a breeding potato field trial.

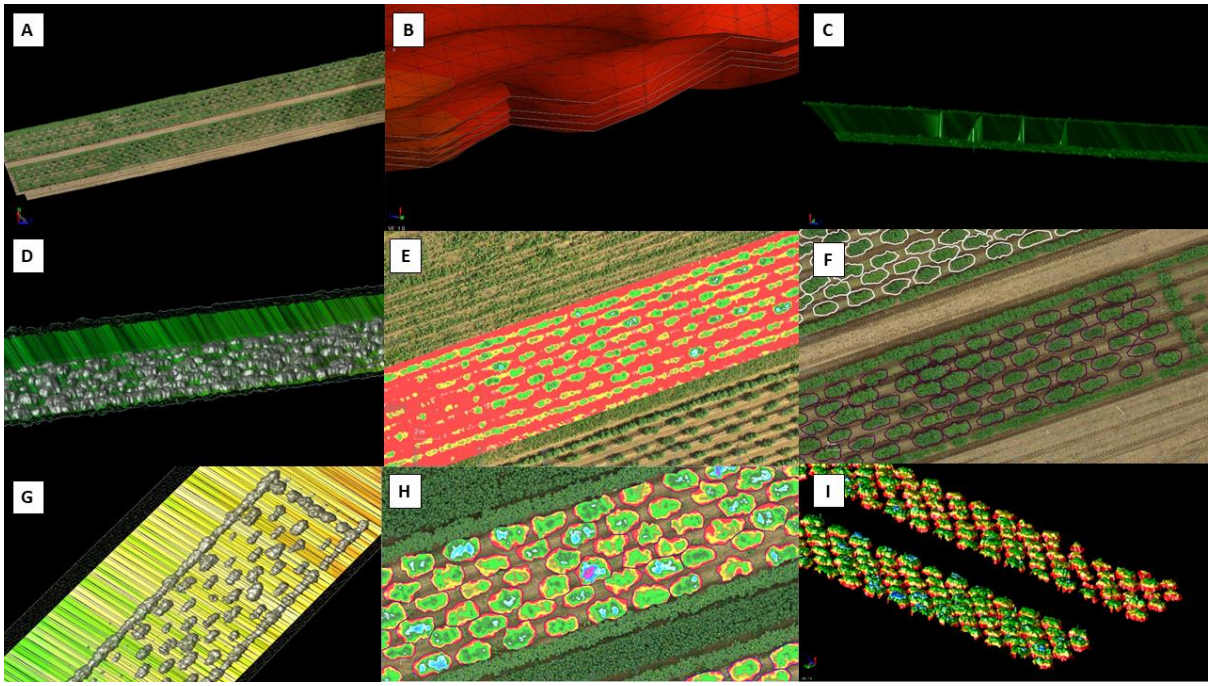

Figure 18 Visual presentation of a 3D model of a potato canopy. Each panel (A-I) shows each step-in order to extract potato 3D model. (A) 3D point clouds of the potato field. (B) Generating of soil reference surfaces. (C) 3D representation of the point cloud before error removing process. (D) 3D representation of an overlay canopy and soil estimation model. (E) a color-coded heatmap indicating the differences in height from the group point. (F) the output of a process identifying individual potato plots. (G) 3D representation of the final overlap between soil estimation model and canopy model. (H) the output of plot segmentation, the colour differences indicate the height of canopy. (I) final output of 3D model of potato canopy for each plot.

2.9.4 DATA EXTRACTION

Reconstructed results contained a scanning trajectory and raw point clouds for scanned area. In trajectory, the lowest position indicated the starting and ending points of each data collection session. Raw point clouds were rendered by colour using point height information, with rad to purple representing low to hight values.

For plant height calculation, the difference model allows the determination the height data at any point within the plot area at a resolution of 1 cm, permitting only the determination of the highest point within the demarcated area (representing the maximum height of the 4 plants within the plot). Canopy ground cover and canopy volume are defined as the sum

planimetric area and total canopy volume. There is above the level of the soil reference surface of each plot difference model.

Subsequent statistical analysis and other data processing was carries out in R (R language software).

2.10 DATA ANALYSIS METHODS

Quantitative analysis is employed using R language programme and Quantum Geographical Information System (QGIS) software to process and interpret multispectral data. Statistical methods, such as correlation and regression analysis, are used to compare UAV data with ground-truth data computed in R Studio, R language software.

2.10.1 QUANTITATIVE ANALYSIS

Quantitative analysis is employed using R language software and QGIS to process and interpret multispectral data. Due to the complexity of the dataset (a large number of groups) that includes over 290 varieties of potato and different management practices, when visualising the data, we need to select only those varieties that show significant differences or patterns.

2.10.2 STATISTICAL ANALYSIS

The basic workflow outlines the steps to perform correlation on data extracted from multispectral images in R Studio software (Team, 2021). The null hypothesis is that there is no difference between data extracted from UAV-based imagery and proximal data. A low p-value shows significant support to the finding, therefore rejecting the null hypothesis. Additionally, to explore more advanced statistical techniques and visualisation to gain deeper insights into the relationship between variables, the least squares means were calculated for all proximal measurements, including height, LAI, disease, and senescence scores. Furthermore, the least square means for the vegetation indices obtained by remote sensing were calculated, and the correlation between plant health was established. Regression analysis was performed to test the strength and trends in relationship between vegetation indices and other crop growth parameters.

ANOVA (analysis of variance) was employed to compare the means of measurements across different methods and potato varieties. Subsequent *post-hoc* tests, such as Tukey's HSD, were conducted to determine differences between groups. The 'tidyr' package in R was used

to reshape the dataset, allowing us to filter and analyse differences across measurement methods over time using Linear models for consistency.

Discriminant analysis of principle components (DAPC) is a multivariate method ideal for identifying and describing clusters of genetically related individuals. It is particularly useful when analysing the data that consists of various measurements or genotypes for different potato varieties. The process includes two main processing steps: principal component analysis (PCA) and discriminant component analysis (DCA). The first step is to understand the structure or pattern of the data and then use the principal components to distinguish between different varieties.

2.10.3 CLASSIFICATION

The objective of the classification was to categorised data into classes based on their characteristics using advanced statistical and machine learning techniques. Partial least squares regression (PLSR) and partial least squares discriminant analysis (PLS-DA) were used to quantify and identify the relationship between predictive variables and responsive variables. K-mean clustering was used to group data points into clusters to identify patterns in the data set. The number of clusters is based on elbow plots.

2.10.4 HANDLE THE MISSING VALUE

Before using the data for modelling, it's important to clean and pre-process the dataset. To enhance the clarity of the clustering and potentially reduce the impact of outliners, there are several approaches that need to be considered.

1. Outliner detection and removal: identify and remove outliners from the dataset before performing any analysis using statistical methods.
2. Focused visualisation: adjust the scale of the plot in specific regions of the plot where most data points are clustered.
3. Reduce the number of potato varieties: as the data contained over 290 potato varieties, visualising all of them together can be challenging. By reducing the number of varieties or divided them into sub-groups, will improve the graph.
4. Simplify the plot: instead of using advanced visualisation techniques, use the basic scatter plot with a limited number of varieties.

Given the nature of data values, the dataset contains a number of non-numeric values, such as, different management systems, potato variety names and plot numbers. To process, we need to clean the data by either removing or imputing, depending on the quantity and distribution of such values. With datasets such this, we proceed with imputation using the median of the crop parameters columns. After imputation, we then run PCA and DCA. This is done by standardising the imputed data to ensure all parameter contribute equally to the analysis. Then, we reduced the dimension of the data using PCA and then used the DCA to distinguish between different potato varieties.

2.10.5 MODEL DEVELOPMENT

Before developing the model, the following steps were considered:

- identify which crop parameter or traits are most relevant for predicting potato maturation and yield. Identifying the correlation between yield and weather parameters (i.e., temperature, rainfall, and soil temperature) to determine the feature importance of each parameter on yield production.
- Then using the selected features to perform correlation analysis to understand the relationships between different parameters and yield or maturity and determine which parameters are strongly linked with yield or maturity.
- Building the predictive model for yield based through statistical and machine learning techniques including regression, random forests, and decision tree methods.

In this study the partial least square discriminant analysis (PLS-DA) was used to classify the correlated plant parameters with the actual yield data and to combine the yield data with corresponding weather data and other crop parameters for the yield estimation model.

Multiple linear regression models

Multiple linear regression (MLR) models were constructed using stepwise variable selection based on the correlation in R programme using ‘lm’ functions for benchmarking the performance of Random Forest regression models (Guo *et al.*, 2021).

The general form of a multiple linear regression is:

$$Y=\beta 0+\beta 1X1+\beta 2X2+...+\beta nXn+\epsilon \quad (1)$$

where:

- Y is the crop yield.

- β0is the y-intercept.

- β1, β2, ..., βn are the coefficients for each parameter.

- X1, X2, ..., Xn are the crop parameters.

- ϵ is the error term.

Decision tree model

A decision tree is a model used in machine learning that partitions the data into subsets based on decision rules inferred from the input features.

$$\mathrm{RSS} = \sum_{m=1}^{M} \sum_{i \in R_m} (y_i - \hat{y}_{R_m})^2 \quad (2)$$

The Random Forest models (RF).

Random Forests can perform both classification and regression tasks based on the crop parameters (Breiman, 2001). This study trained and used RF regression models, binary-bias machine-learning method to predict potato yield for 2020, 2021 and 2022 trials. In this study RF was also used as a classification tool for potato maturity and senescence. The RF package in the R programme produced a measure of importance of the predictor variables and a measure of the internal structure of the data using 'randomForest' function (Jurado *et al.*, 2022).

2.10.6 PREDICTIVE MODELLING

In order to model the weather data and predict the potato yield, the combination of datasets for proximal and remote sensing is used Certain weather conditions may increase the risk of disease incidents. Modelling the risks are based on the weather data and field observation. This has been done after exploring the data analysis of the combined data to find patterns, trends, and correlations within the parameters. The tool used in the data analysis is carried out in R Studio, R language software (Team, 2021).

1. Then using the model building, including linear regression, partial least squares regression (PLSR), and machine learning such as decision trees and Random Forest models.
2. Then training the model by split the data into training and test set (80/20).
3. Model evaluation: validate the model using the test set with
4. Interpretation: account for different potato varieties and management practices (organic and conventional practices).

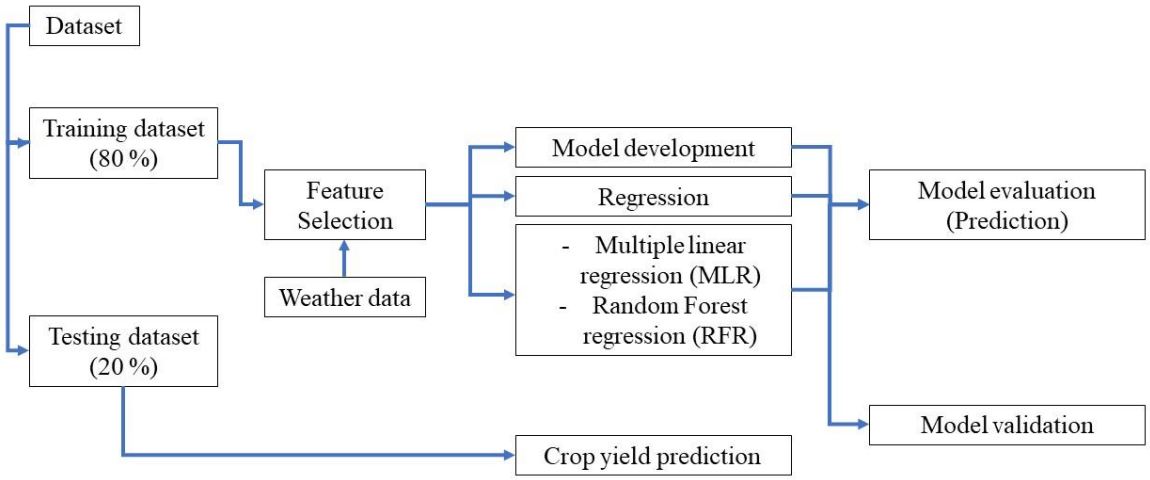

Figure 19 The development of the yield prediction model workflow.

By following the workflow showed in **Figure 19**, we can create a yield prediction model that leverages both weather data and crop parameters to help predict the potato yields with greater accuracy. Regression models and coefficients are provided for all regression in this study. Then, a package in R, such as the 'ggplot2', was used for data visualisation.

2.10.7 VISUALISATION

Created visualisation that overlay the weather data with field sensing data. This will help explain how the data points correlate with different plant growth stages, uniformity, and potential stress indicators. Then, the ground-truth measurements that are taken and used as a reference point to ensure accuracy and reliability. Creating a visitation that overlays weather data with field sensing data involves integrating datasets to provide comprehensive insights

into how each variety behaves cross years and weather conditions. This process typically involves the following steps:

(i) Data integration: integrated the weather data with the field sensing data and perform statistical analyses that representing in the visualisation, for example, calculating correlations between weather conditions and plant health indicators, and creating a time series to visualise how each crop parameters change over the growing season.

(ii) Visualisation development: by selecting visualisation tool that support the overlay of spatial data, such as GIS software (e.g., QGIS, ArcGIS) or programming libraries (e.g., R, Matplolib)

(iii) Presentation: display the information using colour scheme, adjusting scale, labels, and annotations to make the data easy to interpret. Then looks for patterns or trends that informative and present.

2.11 VALIDITY AND RELIABILITY

To ensure validity, we calibrate our UAV sensors before each flight. The reliability is enhanced by repeated flights over the same plots, ensuring consistent data capture. For model evaluation we employed the metrics for evaluating regression models include mean absolute error (MAE), mean square error (MSE), root mean square error (RMSE), and R-squared value.

Mean absolute error (MAE) is the average of the absolute difference in the original to predicted value.

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}\,| \tag{3}$$

mean squared error (MSE)

$$MSE = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y})^2 \tag{4}$$

root mean squared error (RMSE)

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y})^2} \quad (5)$$

Coefficient of determination or R-squared is a measure of how well the values fit together in relation to the starting values. The percentage represented by the value between 0 and 1. The closer the value to 1, the better the model fit.

$$R^2 = 1 - \frac{\sum(y_i - \hat{y})^2}{\sum(y_i - \hat{y})^2} \quad (6)$$

2.12 ETHICAL CONSIDERATIONS

The study respects the privacy of the farm. No imagery that identifies specific individuals or infringes on private structures will be published. All data is anonymized and stored securely. University ethical number 11-445487.

2.13 RESEARCH LIMITATIONS

This chapter focuses on the application of multispectral imaging for potato phenotyping. This study also intends to investigate and understand the variability between different systems and how resolution dependent are the crop growth and disease detection parameters in potato. While assessing this variation, we will evaluate different imaging sensors to depict the best approach and resolution suitable for disease monitoring and accessing crop variability in potatoes (*Solanum tuberosum* L.) at field-scale under natural conditions. By combining the proximal measurements with the data obtained from UAV-based aerial imagery, we can establish correlations and evaluate the accuracy of the remote sensing approach. As with any remote sensing activity, the generation of accurate and quantitative data using UAV-based sensors require calibration and validation. This process involves characterising the UAV data and validating their measurements against reference data obtained from ground-based or proximal measurement. While this study offers promising into the potential of multispectral imaging for potato phenotyping, it is crucial to acknowledge its limitation. This research is limited to potato crops grown under field conditions. As there are many different platforms and

sensors available for crop phenotyping, in this project evaluate these resources using a 290-potato genotype panel with the aim to investigate crop variability and non-invasive disease monitoring. The findings may not be directly applicable to other crop species or greenhouse settings. Moreover, this chapter primarily utilized the R programming language to conduct the analyses. R, renowned for its statistical and data analysis capabilities, offered the flexibility and robustness required for this research. This limitation, however, pave the way for further research opportunities, pushing the boundaries of what can be achieved in crop phenotyping and remote sensing.

2.14 SUMMARY

This chapter focus into the advanced applications of unmanned aerial vehicles (UAVs) equipped with multispectral imaging capabilities, emphasizing their application in potato phenotyping. We explore how UAV-based multispectral imaging serves as a transformative tool, allowing for high-resolution, non-destructive assessment of potato crops. By capturing detailed spectral data, these UAVs enable the detect plant parameters related to plant health, growth, and trend. The chapter further elucidates the methods of data acquisition, processing, and analysis, underscoring their importance in generating actionable insights for potato crop management. Through real-world applications, we demonstrate the profound impact of this technology in advancing precision agriculture, particularly in enhancing our understanding of potato crop dynamics and optimizing yield outcomes, which will take about in the following chapters. With a structured research strategy, this chapter sets the foundation for data collection and analysis. The subsequent chapters will delve into the findings, interpretations, and implications of using UAV-based multispectral imaging in potato phenotyping.

2.15 CODE AVAILABILITY

The research analysis performed for this study leveraged pre-existing packages sourced from GitHub, the popular open-source platform. These packages were not merely used as-is; they were adapted and tailored to meet the specific requirements and objectives of this study. For transparency, reproducibility, and to aid interested parties in understanding our computational approach, the modified R code used for our data analyses is display in the appendix B. For more detailed code will be made upon request.

CHAPTER 3. DEVELOPMENT AND VALIDATION FOR HIGH-THROUGHPUT FIELD PHENOTYPING AND DISEASE DETECTION OF POTATOES USING MULTISPECTRAL IMAGERY AND PLANT 3D-MODEL

Potato (*Solanum tuberosum* L.) is one of the world's most important staple crops, providing a significant portion of the global food commodity. Accurate and efficient monitoring of potato growth and development is crucial for optimising yield and quality for marketable produce. Traditional field phenotyping methods are labour-intensive, time-consuming, and often limited in capturing detailed information about plant health and stress response. The advent of multispectral imaging technology offers a promising solution to overcome these challenges by enabling high-throughput and non-destructive collection. New technology considerably reduces the costs of sensors and autonomous vehicles. Low investment in sensor, and platforms, or pipelines presents a trade off with labour costs, as plant handling and labour represent the major proportion of costs in phenotyping experiments. The cost of high-throughput phenotyping field experiments and automated platforms is similar regardless of platforms or vehicles. The development of software applications, e.g., imaging, processing, modelling, and decision support systems, is also a major part of the cost.

This chapter focuses on crop variability detection and its dependency on varieties and crop management systems. Precision agriculture and plant phenotyping required new approaches to improve traditional plant monitoring to keep up with the challenges. When used together, high-throughput field phenotyping (HTFP) and remote sensing platforms show detailed patterns and relationships between various crop characteristics and disease incidence. The multispectral images captured were processed into a crop surface model by using the structure-from-motion (SfM) software Pix4DMapper. Generated plant height, ground cover, and vegetation indices were used to establish the linear relationship between UAV-based data and proximal data. High differences between cultivars were observed and increased during the growing season. The proposed UAV-based multispectral imaging showed a correlation coefficient of 0.93 between proximal height and canopy height model (CHM) with a mean difference of 0.01 m in the z-direction. The method and the use of UAV-based multispectral image data enable high-throughput phenotyping of small potato plots at field scale with uncertainties of field conditions.

The strong positive correlations observed between proximal height and UAV height indicate a consistent relationship between the two methods, regardless of the growth stage. This consistency suggests that both proximal and UAV methods can effectively capture plant height variations during different potato growth phases. The practicability, repeatability, and robustness of two developed approaches were demonstrated in this study.

3.1 INTRODUCTION

Plant phenotyping is the detailed description and characterization of plants, which has been a long-standing practice of plant research in both greenhouse and field environment settings (Cendrero-Mateo *et al.*, 2017). Although optical approaches for assessing crop performance and quantifying plant traits emerged between the 1950s and 1970s, they were not initially referred to as "plant phenotyping" at that time (Moran *et al.*, 2003). These methods have contributed significantly to our understanding of plant biology and agricultural practices. In the context of potato production, plant phenotyping is an important aspect of genotype studies and breeding. It helps to enhance the efficiency and accuracy of plant breeding and monitor crop performance to reduce economic losses for future food security. However, traditional phenotyping methods often rely on manual measurements or field inspection, which pose significant challenges due to their being labour-intensive, time-consuming, expensive, and lacking consistency in capturing the effects of environment and genotype on plant traits (Watanabe *et al.*, 2017).

Not only did plant phenotyping require a fast approach to help accelerate breeding lines, but also to understand disease resistance. Late blight, caused by *Phytopthora infestans* (Mont.) de Bary and various potato viruses, is among the most severe potato diseases that damage potato production worldwide (Akino *et al.*, 2014; Lin *et al.*, 2014). Since the Irish famine, the development of late blight resistant cultivars has been a priority in potato breeding (Akino *et al.*, 2014). Infection by potato viruses or different stains of a virus has been reported to increase late blight susceptibility after virus infection (Lin *et al.*, 2014). Although there are various potato cultivars that have been bred to resist the late blight, the latest change in the population of P. *infestans* may not respond to traditional disease management (Fry, 2020). Due to the complexity of disease symptoms that interact with the environment and plant health themselves, when the symptom becomes visible to the naked eye, usually it is already too late or hard to control. Therefore, if breeders or farmers have a tool that helps them to detect it early, it will help prevent the spread of disease. Accurate assessment of the resistance and

response of cultivars to diseases under field growing conditions is required to develop cost-effective approaches (Hansen *et al.*, 2005), especially as inadequate crop monitoring can result in the inefficient use of resources such as water and fertiliser (Valente *et al.*, 2020).

The traditional method of assessing crop performance or monitoring crops involves physically walking through the plots and performing visual and handheld instrument measurements. Proximal phenotyping is a type of phenotyping that uses ground-based instruments and vehicles for plant trait measurements (Chawade *et al.*, 2019). Such proximal data is important and plays a crucial role in validating spectral data obtained from UAV-based aerial imagery. The key objective of these measurements is to establish relationships between estimated data and ground truth measurements, visual disease assessment, and morphological ground truthing data. To validate the data derived from UAV-based aerial imagery, visual disease assessment and morphological ground truthing are conducted. For instance, crop emergence evaluation is obtained through field visual assessment, where the emergence and establishment of crops are visually inspected and recorded. This ground truth information serves as a benchmark to validate and calibrate the estimates derived from the aerial imagery. The assessment of plant maturity and diseases in potato varieties is also carried out using visual ratings. This involves visually observing the crops throughout their growth stages and assigning ratings to indicate the level of maturity or disease presence. Assessment of diseases and maturity using remote sensing platforms will allow consistent monitoring and real-time detection, which will help to improve productivity and make it more sustainable (Rodríguez *et al.*, 2021).

Recent technological advancements have led to the emergence of various phenotyping techniques aimed at enhancing plant phenotyping in both laboratory and field settings (Bucksch *et al.*, 2014). Alongside this, sensor techniques are increasingly being used for detecting, identifying, and quantifying plant diseases, promising more accurate disease control due to their sensitivity and high availability (Haverkort and Struik, 2015). High-throughput phenotyping (HTP) platforms were created due to the need to quickly and accurately measure many plant traits on a large scale in order to speed up breeding programmes and learn more about how plants interact with their environments (Wang *et al.*, 2020). HTP platforms address these needs by using a variety of technologies, including remote sensing, imaging, and aerial vehicles equipped with multiple sensors and sensor networks, to automate data collection and analysis. By collecting reflected light, HTPs can identify many things about a plant, such as its structure (e.g., height and leaf area index), its nutrient status, and whether it is under abiotic or

biotic stress. To enhance plant growth and yield through engineering, a deep understanding of the genetic factors involved is essential (de Jesus Colwell *et al.*, 2021). Precision measurement of phenotyping assessments can be linked to this genetic knowledge, which not only improves our understanding of plant phenotyping studies but also facilitates the selection of plant varieties. This integration of information leads to accelerated breeding cycles and the development of improved cultivars.

However, despite the significant advantages offered by UAV-based and proximal sensing, there are challenges regarding the precision required for specific agricultural research applications (Hu *et al.*, 2018). For example, in a field setting, environmental conditions are inherently variable. It is essential to characterise and understand plant phenotypes under these dynamic conditions. High-throughput field phenotyping (HTFP) is a technique that integrates HTP approaches to measuring plant traits on a large scale under field conditions and settings (Chawade *et al.*, 2019). When HTPs, imaging sensors, and GIS technologies are used together, they make phenotyping quick and require less effort. This helps breeding programmes, makes farming more efficient, and makes crops more resistant to damage. While significant progress has been made, the current constraints of remote sensing technologies are the cost of the sensors and the complexity of imaging processing, handling, and acquisition. Although advanced phenotyping platforms are performing well, they are not easily affordable in many research and practice settings. While unmanned aerial vehicles (UAVs) have been extensively used to measure plant traits at the group or plot level, their application for individual plants has been limited (Guo *et al.*, 2020).

Despite these limitations, various researchers have demonstrated and reported successful discrimination between healthy and diseased crop patches and plants based on high resolution imagery at canopy level by sensors mounted on the UAV using HTFP approaches. Not only are current field phenotyping endeavours primarily centred around quantifying crop yield and related characteristics (Watt *et al.*, 2020), but they also have the ability to evaluate the healthy and diseased potato canopy for diseases such as late blight (Van De Vijver *et al.*, 2020; Rodríguez *et al.*, 2021), early blight (Chakraborty *et al.*, 2022), and other diseases (Appeltans *et al.*, 2020). Not only for the early detection but also for the severity of the diseases or stresses, as seen in the research from Franceschini *et al.* (2019). However, effective field phenotyping demands more than trait selection or using the best tools and data management practices (Araus *et al.*, 2018). Therefore, this chapter aims to explore the different approaches

used for potato canopy trait extraction and evaluate the high-throughput field phenotyping in potato varietal trials.

Research questions and hypotheses

In the previous chapter (Methodology and Data Analysis), we explored the methodology of data collection and different approaches on potato canopy trait extraction from the UAV-based imagery. In this chapter, the primary objective is to develop and evaluate a high-throughput field phenotyping approach using multispectral images for potatoes. The specific goals are as follows:

- design and implement a multispectral imaging system suitable for potato field phenotyping.
- assess the effectiveness of multispectral images in capturing diverse phenotype traits of potato crops.
- explore the potential of multispectral data for early detection of stress-related and diseases traits in potato crops.
- compare the efficiency and accuracy of multispectral field phenotyping with traditional methods.
- provide insights into optimising potato production practices based on the acquired phenotypic data.

It was possible to establish a correlation between what we have observed on the ground and what the UAV had captured from the air by aligning aerial imagery collection with proximal field measurements. This integration offered both a micro and macro perspective on potato growth and development and disease dynamics, ensuring data accuracy and validation.

Scope and limitations: This chapter focuses on the application of multispectral imaging for potato phenotyping. This study also intends to investigate and understand the variability between different systems and how resolution dependent the crop growth and disease detection parameters in potatoes are. While assessing this variation, we will evaluate different imaging sensors to depict the best approach and resolution suitable for disease monitoring and accessing crop variability in potatoes (*Solanum tuberosum* L.) at field-scale under natural conditions. By combining the proximal measurements with the data obtained from UAV-based aerial imagery, we can establish correlations and evaluate the accuracy of the remote sensing approach. As with any remote sensing activity, the generation of accurate and quantitative data using UAV-

based sensors requires calibration and validation. This process involves characterising the UAV data and validating their measurements against reference data obtained from ground-based or proximal measurements. While this study offers promising insights into the potential of multispectral imaging for potato phenotyping, it is crucial to acknowledge its limitations. This research is limited to potato crops grown under field conditions. As there are many different platforms and sensors available for crop phenotyping, this project evaluates these resources using a 290-potato genotype panel with the aim of investigating crop variability and non-invasive disease monitoring. The findings may not be directly applicable to other crop species or greenhouse settings. This limitation, however, paves the way for further research opportunities, pushing the boundaries of what can be achieved in crop phenotyping and remote sensing.

3.2 STATISTICAL ANALYSIS

The basic workflow involves performing correlation analysis on data extracted from multispectral images using R studio software (Team, 2021). The null hypothesis is that there is no difference between UAV-based imagery data and proximal sensor data. A low p-value indicates significant findings, leading to the rejection of the null-hypothesis. Additionally, to explore more advanced statistical techniques and visualisation to gain deeper insights into the relationship between variables, the least square means were calculated for all proximal measurements, including canopy height, LAI, disease, and senescence scores. Similarly, we calculated least square means for vegetation indices obtained from remote sensing and established correlations with plant health.

ANOVA (analysis of variance) was employed to compare the means of measurements across different methods and potato varieties. Subsequent *post-hoc* tests, such as Tukey's HSD, were conducted to determine differences between groups. The 'tidyr' package in R was used to reshape the dataset, allowing us to filter and analyse differences across measurement methods over time using Linear models for consistency. For visual representation, the 'ggplot2' package was utilised for a comparative view across varieties and observational consistency over years. We also applied robust least squares regression to evaluate model accuracies, with root mean squared error (RMSE) and coefficient of determination (R^2) serving as the criteria for model performance. Person's correlation coefficient (r) was also used to assess the linear relationship between observed and the predicted values.

Finally, to analyse the variation in blight infection across different years and locations, we used a regression model that incorporated field assessment and vegetation indices to estimate the severity. Following this, principal component analysis (PCA) was conducted to reduce the dimension of the data, which was then followed by discriminant component analysis (DCA) to distinguish and group potato varieties. K-means clustering was performed afterward to cluster data into groups based on elbow plots.

3.3 RESULTS AND DISCUSSION

3.3.1 THE ESTIMATION OF CROP EMERGENCE VERSUS STAND COUNT

Estimating crop emergence (number of plants per plot) accurately is crucial for optimising resource use in agriculture. The conventional method involves physically walking through fields and counting plants that have emerged from the ground. The primary aims of this step is to adjust the ground-level plant estimation method for UAVs and assess its accuracy. This study evaluates the method's performance, focusing on its repeatability and potential limitation posed by spatial resolution for UAV-based observations. The emergence rate and emergence uniformity were obtained from all plots from proximal data collection were compared with the data obtained from the UAV-based imaging. The findings suggest suitability for field phenotyping, with **Figure 20** illustrating the counting process using UAV imagery for an automated plant counting process. The top panel of the figure illustrates the example of RGB (red, green, and blue) images for individual plots. This represented a reference for the automated counting process. The bottom panel shows the automated counting process executed in the R software environment using UAV-derived imagery. The black shapes represent the original plants, maintaining their shape and relative position from the RGB section, with the markers indicating the plant that has been detected and identified within the plot. This ensured that the image processing algorithm had not altered the physical characteristics of the plants. Once the number of plants is obtained from the UAV-based imagery, it can be compared with the proximal data.

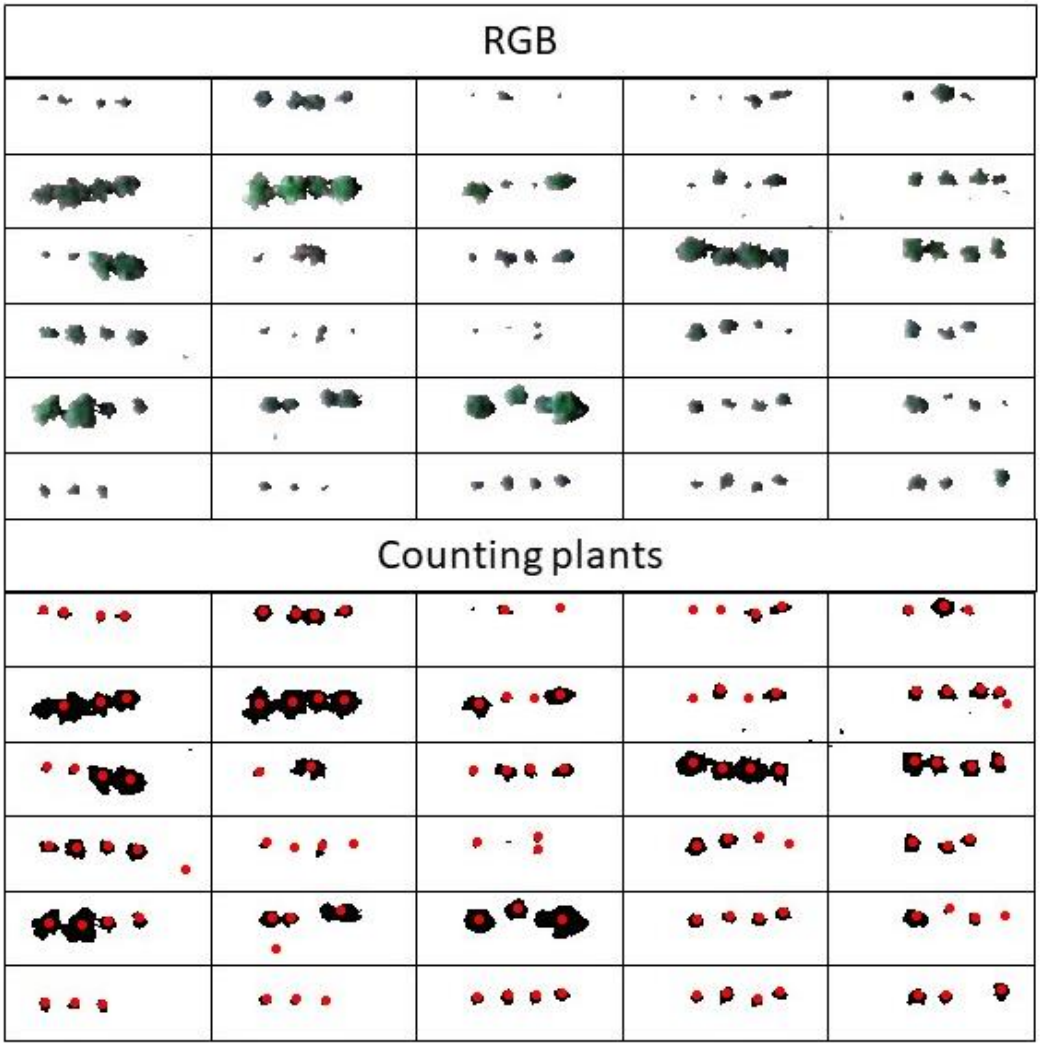

Figure 20 RGB section represents the original visual representation captured by the UAV. Top: RGB image after soil removal process with grid lines to identify individual plots. Bottom: Plant counting process, where the black shape represents the original plants while the red dots indicated each individual plant that the algorithm has successfully identified.

The emergence count obtained from both proximal (stand counting) and UAV-based imaging are compared to assess how well the two approaches align in estimating crop emergence and stand count. The automated counting process mimics the manual method's accuracy, showing a strong correlation with approximately 75.54% of the variance observed in the automated counting (R^2=0.7545, p<0.001) (**Figure 21**, **A**). Despite expected deviations due to the inherent methodological differences, the 2021 data indicated the highest match between manual and automated counting with an R^2 of 0.937 (**Figure 21, B**). However, automated methods may overcount compared to manual methods, potentially due to algorithm over-detection. Why was the best match/ correlation in 2021.

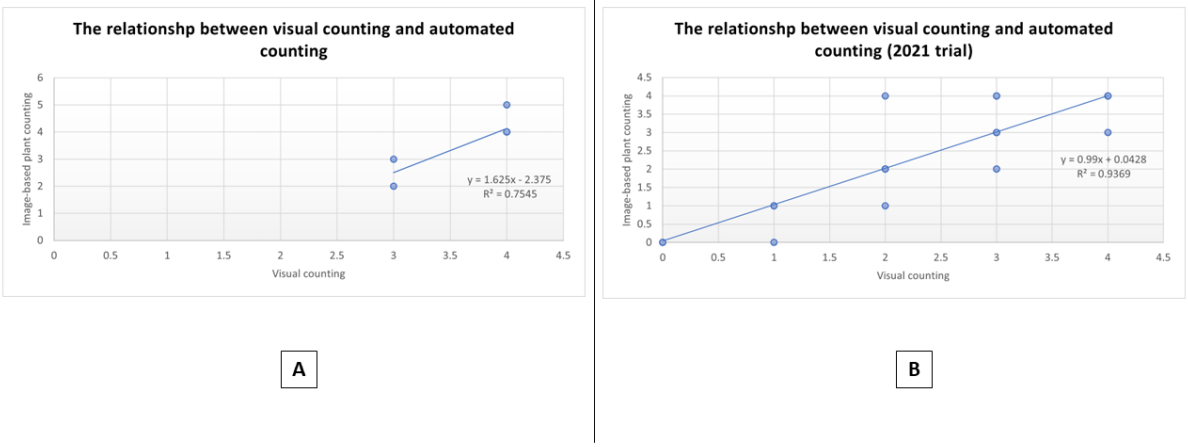

Figure 21 Relationship between visual stand count and automated counting where (A) coefficient of determination (R-square) = 0.7545 (three-year field trials combined). (B) coefficient of determination (R-square) = 0.9369 (2021 potato field trial).

Comparing the overall data, the R-squared value indicates a strong positive correlation, though not as robust as a one-to-one relationship might show. This result suggests that the automated method may tend to count more plants that the visual method, potentially due to over counting by the algorithm. By comparison between flights, we found that the accuracy of this technique is higher when the potato plants are smaller. This might be due to the less obstructive canopy size and reduced weed interference affecting the UAV-based imagery data extraction.

Studies using a similar technique to identify and count the number of plants in the field indicate that lower-altitude UAV flights can provide a much better resolution, hence greater accuracy. For instance, Jin *et al.* (2017) achieved an R-squared of 0.81 using a Support Vector Machine (SVM) for analysis of very low-altitude UAV imagery (3-7 m altitude above ground), which allows capturing images at a ground resolution of 0.2 mm to 0.45 mm. Whereas this study maintained UAV flight parameter at a 30 m altitude, resulting in a ground resolution of approximately 1 cm. While lower altitudes of UAV flight offer improved ground solution, they also pose challenges for photogrammetric techniques to retrieve the camera's precise position at the time image is acquired (Jin *et al.*, 2017).

This study along with others in the field of plant counting, primarily utilised the Red, Green, and Blue (RGB) spectral bands. Li *et al.* (2019) reached a 0.96 correlation coefficient for crop emergence estimation using an unsupervised Random Forest-based method with high-resolution RGB ortho-images. Valente *et al.* (2020) achieved a 95 percent accuracy using machine vision with excess green index and Otsu's method and transfer leaning using convolutional neural networks, and Lin *et al.* (2023) improved this to 99.53 percent by employing a unique combination of Red, Green, and Near-infrared (NIR) spectral bands. These findings underscore the effectiveness of the used of specific spectral bands to enhance plant counting accuracy.

Thus, the further study from this chapter need to explore other spectral bands and different approaches to improve the accuracy of automated crop stand counting. Considering lower UAV flight altitudes may be necessary for achieving the desired ground resolution without overly complicating the method.

3.3.2 PLANT HEIGHT MEASUREMENT

In this study, the plant height of 290 potato varieties distributed over 1200 plots was examined, with each variety having four replicates. Manual measurements were taken at growth stages corresponding with UAV flights in order to establish a correlation between proximal measurements and estimating plant height from the UAV-based imagery. This study utilised the Canopy Height Model (CHM) to estimate plant height from the difference between the Digital Surface Model (DSM) and Digital Terrain Model (DTM) as suggested by Anderson II *et al.* (2019). This method was applied using the FIELDimageR tool proposed by Matias *et al.* (2020) and point cloud data processed using Structure from Motion technique (SfM) through Trimble Business Centre (TBC) software as proposed by de Jesus Colwell *et al.* (2021) to estimate plant heights in three years of potato trials.

Finding indicates a high overall Pearson's correlation coefficient (R^2=0.7985, $p<0.001$) between UAV-based model estimation and ground truth across three years of research trials (**Figure 22**). Among the methods used in this study, the point cloud method showed the best accuracy and repeatability, followed by the R package method.

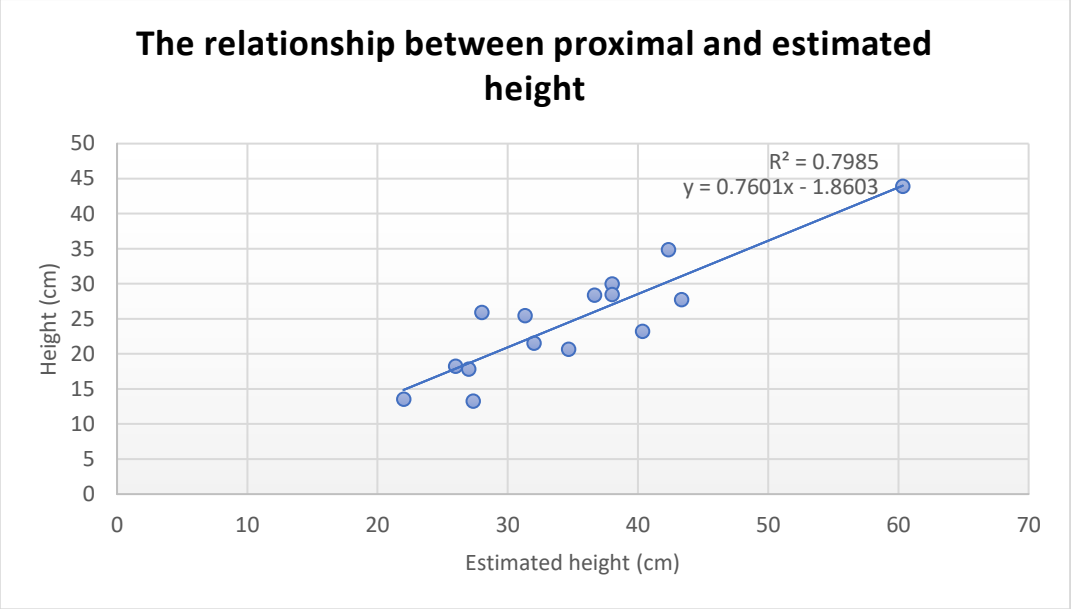

Figure 22 Correlation between proximal measurements and the estimating of potato height (in metres) over three years of research trials.

However, detailed data analysis for 2022 on the canopy height and potato crop growth stage showed the strongest correlation during the vegetative growth and tuber initiation stages. During the vegetative growth and tuber initiation stage (UAV flight 2), the correlation with proximal data was the strongest at R^2 of 0.8131 ($p<0.001$). When potatoes plants were closer to the maturity stage (UAV flight 3), the correlation remained strong but was less pronounced ($r=0.806$, $p<0.001$). When the plant reached the maturity stage (UAV flight 4) the correlation reduced dramatically ($r=0.5864$, $p<0.001$) (**Figure 23**). Canopy height increased exponentially until reaching the maximum height when the plant entered the flowering stages. Then, canopy height started to decrease as the plant has entered the maturity stage. The overall mean and median of canopy height were calculated based on the field height measurement.

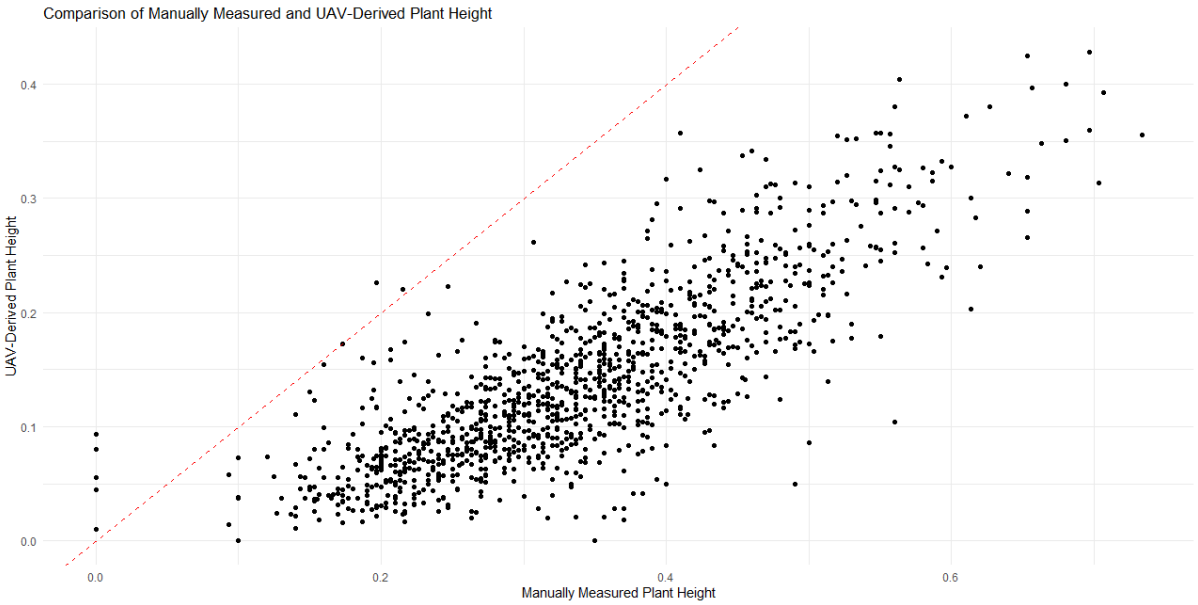

Figure 23 The comparison between manual plant height measurement and UAV-derived plant height (from CHM) when the plant reaches the maturity stage (UAV flight 4).

Among the three methods used to estimate plant heights (proximal, CHMs, and 3D model), ANOVA analysis suggested that all the measurement method and the potato variety significantly affects average plant height ($F_{(2,3309)} = 2344.23$, $P<0.001$). The method of measurement has a pronounced effect, with a high F-value and extremely low p-value. The variety also plays a significant role, but the differences between varieties, while statistically significant, are not as strong as the differences between measurement methods ($F_{(288,3309)} = 10.91$, $P<0.001$). **Figure 23** contrasts manual plant height measurements with UAV-derived data at the maturity stage (UAV flight 4), reinforcing the differences observed between the two measurement approaches and highlighting the need for careful consideration of methodology in plant height assessment. The method used to measure the plant height can have a substantial impact on the results obtained, which is important to consider for research and practical applications.

Comparative analysis with previous studies showed variability in the effectiveness of UAV-based height estimation techniques across different crops. The self-calibration method showed the best performance ($R^2 = 0.63$) and outperformed 'point cloud' and 'reference ground' methods in terms of repeatability (0.34 and 0.38, respectively) in sorghum and maize (Hu *et al.*, 2018). Similarly, de Jesus Colwell *et al.* (2021) found that the image based plant height estimation, particularly after data cleaning, showed the best performance with an R^2 of 0.52 in potatoes. Pugh *et al.* (2018) observed variability in correlation across different crops, with

correlation ranging from of 0.4 to 0.5 in sorghum and 0.7-0.9 in maize. The study by Shi *et al.* (2016) also highlighted a stronger correlation between UAV estimated and ground truth height in sorghum plots ($R^2 = 0.55$). ten Harkel *et al.* (2020) utilised UAV-LiDAR technology with the 3DPI algorithm, achieving accurate height estimations for sugar beet and wheat ($R^2 = 0.70$, RMSE = 7.4 cm and $R^2 = 0.78$, RMSE = 3.4 cm, respectively) but encountering challenges with potatoes ($R^2 = 0.50$, RMSE = 12 cm). Unlike crop with uniform and vertically structured canopy, such as wheat, rice, and maize, potato plants have an irregular and non-uniform canopy. The complex canopy structure and the ridges characteristic of potato cultivation posed difficulties, resulting in lower reliability for plant height estimation.

These findings emphasise the significance of choosing suitable UAV-based estimating technique that are customised to the unique characteristics of the crop. Although UAV technology along with data analysis is valuable for plant study, the complex characteristics of specific crops, such as potatoes, require further methodological adjustments to improve accuracy and repeatability in plant height estimation.

3.3.3 LEAF AREA INDEX ESTIMATION

Several techniques have been developed to measure Leaf area index (LAI) from both the ground truth and through remote sensing techniques (Bréda, 2003). These methods utilised spectral-based vegetation indices (VIs), plant height data (obtained manually or through sensors), along with machine learning algorithms for enhanced accuracy. Recent advancements in methodologies for measuring LAI, as detailed in the studies by Luo *et al.* (2020), Sun *et al.* (2022) and Yu *et al.* (2023), emphasise the effort to evaluate plant health and productivity through non-destructive approaches. This effort combines traditional ground-based proximal measurements with spectral-based VIs and plant height data.

Table 9 Model performance summary for leaf area index (LAI) estimation using vegetation indices (VIs) and canopy height.

Model	Parameters	R^2	MSE
Random forest	VIs	0.253	0.0151
Random forest	VIs + canopy height	0.294	0.0143
PLS regression	VIs	0.286	0.0145

PLS regression	VIs + canopy height	0.351	0.0132

Data collection from potato canopies was conducted using a handheld device (ceptometer), as well as multispectral cameras, which served as input variables in regression models. Feature selection was guided by correlation analysis, identifying Biomass Index (BI) and Plant Senescence Reflectance Index (PSRI) as having a positive correlation with LAI among other variables. However, other VIs showed weak correlations throughout the growing season, similar to the result from Gong *et al.* (2021b). In order to improve LAI estimation accuracy, we integrated VI with canopy height data across potato growth stages.

The Random Forest (RF) models showed modest correlations between estimated and actual LAI measurements, with R^2 values ranging from 0.25 to 0.29. In order to improve the estimation accuracy further, the Partial Least Square Regression (PLSR) method was also employed. The results shows that the PLSR outperforms the RF approach for estimating LAI, showing R^2 values ranging from 0.28 to 0.35 (**Table 9**). The same result was found in the study from Li *et al.* (2020) and Liu *et al.* (2022). Despite these advancements, the current predictive modelling effort has relied on a dataset with a singular ground measurement of LAI. While the models deployed, including Random Forest and PLSR, have demonstrated varying degrees of predictive capability, the limited ground truth data for LAI poses a significant constraint. This limitation inherently restricts the model's learning scope, potentially impacting the accuracy and reliability of the LAI predictions.

This summary highlights the importance of feature selection in model performance and suggests that combining different types of ecological and phenological data can significantly improve predictive accuracy. Nonetheless, this approach demonstrates a significant step forward in precision agriculture for potato varietal study. The ability to accurately estimate LAI with non-destructive offers a more efficient and repeatable method.

3.3.4 FIELD DISEASE DETECTION

In term of disease detection, this study does not represent a true genotype x environment analysis due to the small plot size and used of seed across years, despite testing potato varieties in both organic and conventional management systems over three consecutive years (2020-2022). Additionally, the trials were subject to natural infections from the local pathogen population, but the structures of the pathogen populations remained unknown. It is presumed

that the varieties were exposed to a variable range of pathogen genotypes during the growing season. Variable weather conditions and timing of disease occurrence led to fluctuations of the blight season across the different year trials.

Potato crops can be infected by numerous viruses that can affect yield and tuber quality. Commonly, viruses can be diagnosed through mosaic patterns on the leaves, stunting of the plant, and the deformation of leaves and tubers (Lacomme and Jacquot, 2017). However, symptom expression can be inconsistent, influenced by the interactions between the virus and plant, growing conditions such as soil fertility and weather, or the stage of the plant when it is infected. The most important potato viruses include potato virus Y (PVY) and potato leafroll virus (PLRV) (Chatzivassiliou *et al.*, 2008). In this study, while specific viruses were not identified, observations of virus symptoms that occurred in a particular plot were recorded for future disease management (**Figure 24** Stunning symptoms caused by potato viruses on potato plants observed in the field. **Figure 25** Leafroll symptoms observed in the field.). It was noted that viruses could be found in 'healthy plants' which tolerate infection without showing any symptoms. Confirmation of specific virus presence, however, requires plant pathogen analysis, such as enzyme-linked immunosorbent essay (ELISA), for detecting the presence of PVY and PLRV (El-Amin *et al.*, 1994; Chatzivassiliou *et al.*, 2008).

Figure 24 Stunning symptoms caused by potato viruses on potato plants observed in the field.

Figure 25 Leafroll symptoms observed in the field.

The annual variation in disease incidence across the potato variety trials offers valuable insights into the susceptibility and resilience of different varieties to virus infections. Over three years, a distinct set of varieties each year displayed sign of disease illustrated in **Table 10**. Disease detection varied annually, with infections identified in 11 potato varieties in 2020 trial, 16 varieties in 2021, and 13 varieties in 2022.

Table 10 Potato varieties affected by disease across three years trials.

Variety Name	**2020**	**2021**	**2022**
12601AB1	No	Yes	No
Alpha	No	Yes	No
Anna	No	No	Yes
BF 15	No	No	Yes
Bzura	Yes	No	No
Celine	No	Yes	No
Ditta	No	Yes	No
Dr McIntosh	No	No	Yes
Duke of York	No	No	Yes
Dunbar Standard	No	Yes	No
Dundrod	Yes	No	No
Eginhiemer	No	Yes	No
Fontaine	Yes	No	No
Gardena	Yes	No	No
GL 78/79	No	Yes	Yes

Gold Marie	Yes	No	No
Harlequin	No	No	Yes
Juliette	No	No	Yes
Kerr’s Pink	No	No	Yes
Kingston	No	Yes	No
Lady Christl	Yes	No	No
Linton	No	Yes	No
Marabel	Yes	No	No
Maris Bard	No	Yes	Yes
Markies	No	Yes	No
Mimi	No	Yes	Yes
Mira	Yes	No	No
NDTX 4271 5-R	No	Yes	No
Pimpernel	No	Yes	Yes
Rubesse	Yes	No	No
Safari	No	Yes	No
Saphire	No	No	Yes
Slaney	Yes	No	No
Spey	No	Yes	Yes
Timate	Yes	No	No

The summarised results from **Table 10** reveal a pattern of disease occurrence among different potato varieties. In 2021, there was an increase in disease presence with 16 varieties impacted, and among these, 'Pimpernel', 'Spey', 'Maris Bard', 'Mimi', and 'GL 78/79' were also affected in the 2022 trials. This indicates that these varieties more prone to disease. The impact of disease infections led to the loss of variety ‘Fontaine’ in 2020, ‘Maris Bard’ and ‘GL 79/42’ in 2022. ‘Saphire’ and ‘Mimi’ produced significant reduction in yield, casing an average yield of less than half a kilogramme. This observation highlights that diseases severity can greatly vary between varieties and the need for robust disease monitoring techniques.

In response to the detection of viral symptoms within the plot, a crucial management decision was made to exclude tubers from affected plots from future seeding to prevent the potential virus spread and protect future cop health, given that viruses can be transmit from infected plants to seeds or tubers, hence potentially affecting the next crop generation (Lacomme and Jacquot, 2017). Moreover, the rapid spread of disease under favourable weather conditions underscores the importance of immediate disease mapping (Dammer *et al.*, 2016). In instances where plants are severely infected, it may be necessary to remove them to prevent the spread of disease to neighbouring plots. In conclusion, this preventive strategy is crucial for controlling and managing crop health.

3.3.5 VEGETATION INDICES FOR DISEASE PREDICTION

Partial least squares-discriminant analysis (PLS-DA) is recognised for its robust capabilities in both predictive and descriptive modelling but requires optimization of parameters for accurate results (Lee *et al.*, 2018). Despite its proven effectiveness with high-dimensional datasets, many researchers encounter difficulties in constructing valid models, highlighting a gap between the algorithm's potential and its practical application. Furthermore, while PLS-DA has been widely utilised in food and plant chemistry research (Almeida *et al.*, 2013; Uarrota *et al.*, 2014; Grasel and Ferrão, 2016; Sampaio *et al.*, 2020; Valderrama *et al.*, 2022; Birenboim *et al.*, 2023), its adoption in plant phenotyping remains limited. For example, Hao *et al.* (2010) and Spoladore *et al.* (2021) explored its utility in differentiating plant varieties using Visible/Near Infrared (Vis/NIR) spectroscopy with modelling, demonstrating its potential on a small number of varieties despite the challenges.

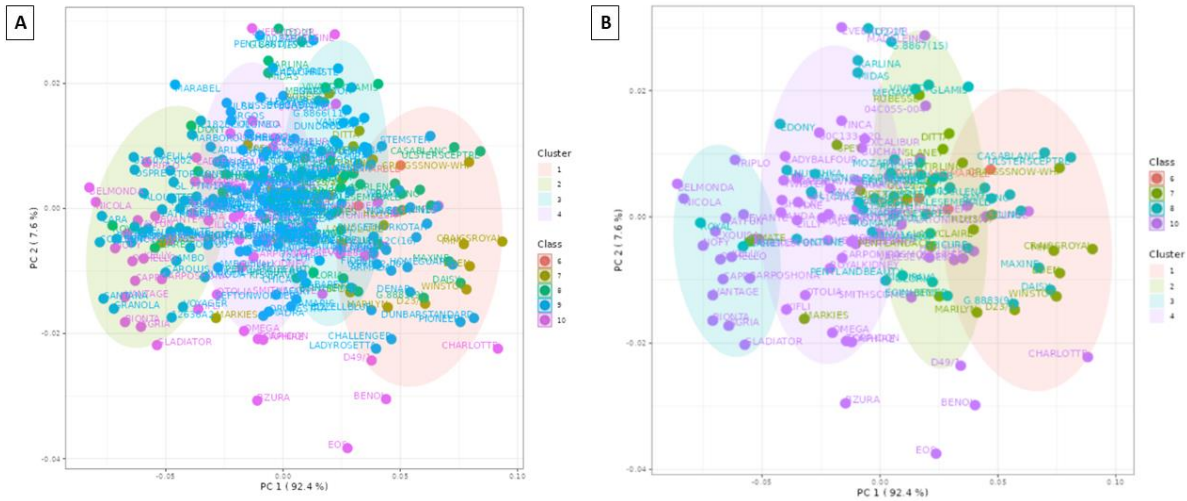

Figure 26 PCA scatter plots and clustering analysis of potato varieties based on vegetation indices (NDVI and NDRE) and field disease severity score class (1-10 scale). Each dot represents a potato variety, and cluster is indicated by coloured shading. (A) visual summary of the relationship and similarities or differences between potato varieties based on vegetation indices and field disease evaluation score. (B) the distribution of varieties with the removal of disease class 9 (healthy plant).

This study utilised multispectral imaging to evaluate potato plant canopy reflectance across growth stages. By calculating vegetation indices (NDVI and NDRE) and developing statistical models, disease severity was estimated and comparted with field disease severity

classes (1-10 score scale) (**Table 3**). The PCA scatter plots with K-means clustering presented in **Figure 26 A** shows a densely packed clustering with noticeable overlap among potato varieties, indicating the challenges of differentiating closely related varieties. Therefore, **Figure 26 B** represents a PCA plot where disease class 9 (representing healthy plants) was excluded, allowing for a clearer visualisation of the relationship and differences among varieties. From the K-means clustering analysis, four distinct clusters were identified, with each point in a cluster representing varieties that share similar traits based on vegetation indices and disease classes. Cluster 1 (red) and 2 (yellow) shows overlap, suggesting similarities between certain varieties in these group. Cluster 1 and 2 contain varieties with low vegetation index value and low disease class (showing disease symptom), indicating significant disease susceptibility and lower overall plant health performance. Cluster 3 (blue) and 4 (purple) also showed overlap, these two clusters contain potato varieties that display high vegetation index value and high disease classes (healthy plant). Cluster 3 and 4 suggest disease-resistant varieties that have a robust response to environmental stress and may be suitable for farmers in the region where disease pressure is moderate and disease management practices are feasible.

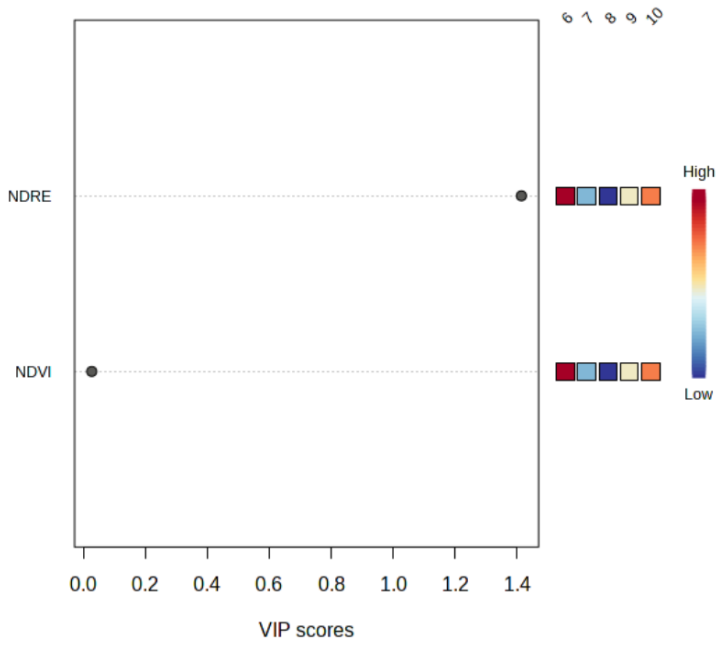

Figure 27 Variable importance in projection (VIP) score for NDVI and NDRE in multivariate partial least square regression model (PLSR). The x-ais plot displays VIP scores, which quantify the importance of both vegetation indices in the model.

In the partial least square analysis, the variable importance plot graph shows the variable importance in projection (VIP) values or scores. A VIP score is a measure of variable's importance as a predictor in modelling (Eriksson *et al.*, 2013). If the variable importance is greater than 1.5, it indicates better vegetation index to discriminate the variety of potatoes according to disease score. The PLS-DA analysis revealed the variable importance in projection (VIP) score ranging from 0 to over 1.4 (**Figure 27**). The variable importance plot highlighted the significance of NDVI and NDRE indices in the regression model, with the NDRE score closer to 1.5 indicating higher potential for predicting the disease severity than NDVI.

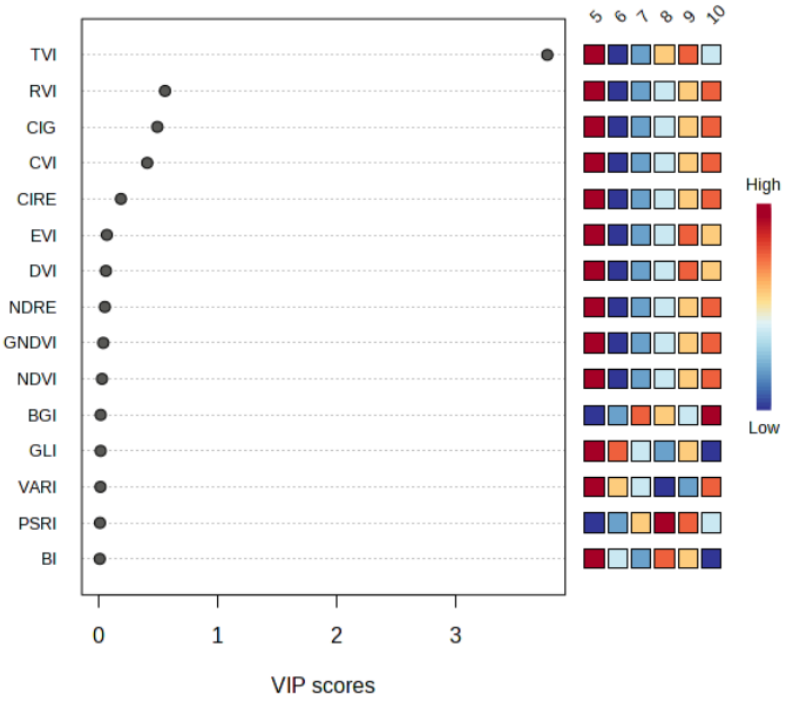

Figure 28 Variable importance in projection (VIP) score plot from a partial least square regression (PLSR) model, listing various vegetation indices extracted from multispectral based imagery on the y-axis and their corresponding VIP scored on the X-axis.

A higher VIP score indicates a greater ability to discriminate between different disease severity classes within the dataset. In the context of plant phenotyping, especially concerning disease incidence or severity in potato crops, the selection and assessment of vegetation indices are crucial. The PLS-DA analysis, visualised in **Figure 28**, shows VIP scores ranging from 0 to over 3. While many of vegetation indices display relatively low VIP scores, the Transformed

Vegetation Index (TVI) stands out with the highest score. The indices like TVI, Ratio Vegetation Index (RVI), Chlorophyll Index-Green (CIG), and Chlorophyll Vegetation Index (CVI) have VIP scores near or beyond the threshold of 1, indicating that they are also significant predictors in the model. Specially, the high VIP score for TVI indicates a strong correlation with the disease scores obtained from field evaluations, which are integrated into the regression model.

In addition to identifying significant indices, the study also utilised the K-means algorithm to categorise observation into separate cluster based on TVI and RVI. This clustering analysis reveals the underlying trends in the data (**Figure 29**). Cluster 1 (red) primarily consists of observations from class 6 and 7. Cluster 2 (green) is characterised by a combination of classes, with a significant representation of classes 5 and 8. Cluster 3 (purple) generally comprises observation from classes 9 and 10. The arrangement of these cluster in space enhances the analysis as the observation for classes 9 and 10 show a greater spread across the plot. This indicates a higher level of variability within these classes. In contrast, class 5 observations exhibit a higher level of centralisation indicating a greater degree of homogeneity. Classes 6 and 7 are closely clustered in the lower left quadrant, suggesting that they may have similar characteristics or reactions to disease.

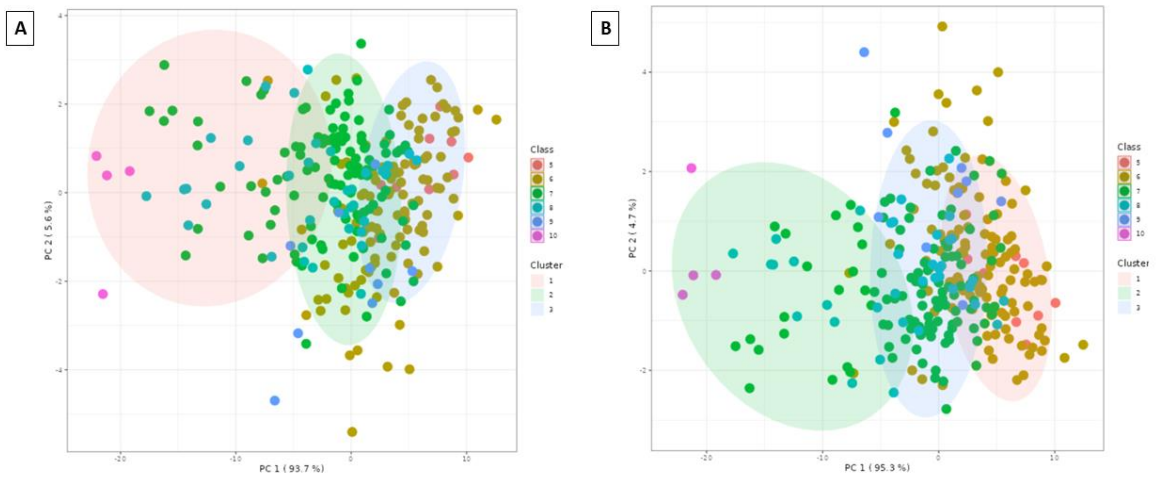

Figure 29 PCA scatter plots with K-means clustering for potato varieties based on vegetation index values and field disease severity classes (1-10 scale). The data points are grouped into clusters with each cluster representing the relationship between vegetation indices and disease severity across all varieties. (A) Clustering using all vegetation indices and disease classes. (B)

Clustering using points based on Transformed Vegetation Index (TVI) and Ratio Vegetation Index (RVI).

By calculating five vegetation indices and developing a regression model, we estimated late blight severity. These estimations were then validated against field visual disease assessments, scoring from 1 to 10 (**Table 3**), across the field. The PLS-DA model was then applied to distinguish between healthy and diseased plants. The PLS-DA score plot revealed distinct clustering of potato varieties based on different vegetation indices. This suggests the capability of each index to identify the presence and severity of diseases that affecting photosynthesis and overall plant health. The study also employed K-means clustering demonstrates the variable nature of disease impact, offering insights into the homogeneity and variability within classes. The comparison with previous studies, such as Hao *et al.* (2010) and Gold *et al.* (2020) confirmed that the early physiological characteristics of the disease obtained from PLS-regression trait models, although among different cultivars responses to the disease differently, they shared some spectral properties that made it possible to distinguish between them.

In summary, in this study, the strategic application of VIP scores and K-means clustering in this PLS-DA analysis not only offers a window into the relationship between vegetation indices and disease incidence in potatoes but also sets the stage for targeted disease management strategies. By identifying which vegetation indices serve as reliable indicators of plant health, researchers and farmers can monitor crop vitality and disease progression. Due to the complexity and crop-specific characteristics of plant phenotypic data, it is crucial for future research to concentrate on improving PLS-DA modelling methodologies that are tailored to this field. Additionally, efforts should be made to address the knowledge gaps that now limit its widespread application. This would require in-depth investigations into the unique features of phenotypic data, which include factors such as environmental adaptability and genetic expression patterns.

3.3.6 THE CONSISTENCY AND ACCURACY BETWEEN PROXIMAL AND UAV-BASED MEASUREMENT

To analyse the obtained results in relation to research objective and hypotheses that the remote sensing-based approach is highly correlated with the direct ground measurements, we compared each methodology used to extract the same crop parameters including remote sensing based and direct measurement (**Table 11**). For validation, we first validate the aerially

observed vegetation indices by comparison with proximal ground measurements for all varieties (**Figure 30**). Comparing the data extraction method across different models or software and evaluating the visual observation against these methodologies in term of accuracy and precision, involves a structured approach. Additionally, identifying the method that provides consistent results over multiple years requires a careful analysis.

1. Accuracy: measures how close the results are the true values.
2. Precision: measures how repeatable or reproducible the results are when the method is applied under the same conditions.
3. Consistency: assesses how stable the method's performance is across different datasets or over time.

Table 11 Different approaches and their evaluation metrics (Correlation Coefficient, R-squared, RMSE, MAE) when tested against the proximal approach. The MAE for the 'R package' and 'TBC' methods is calculated separately using their respective predicted values.

Approach	**Crop parameter**	**Correlation**	**R-squared**	**Root Mean Square Error (RMSE)**	**Mean Absolute Error (MAE)**
R package	Average plant height	0.68	0.4575	0.1448	0.8068
TBC (Trimble Business Centre)	Average plant height	0.78	0.5116	0.0498	0.1984
R package	NDVI	0.72	0.529	0.0002	0.0111
R package	NDRE	0.69	0.229	0.027	0.022

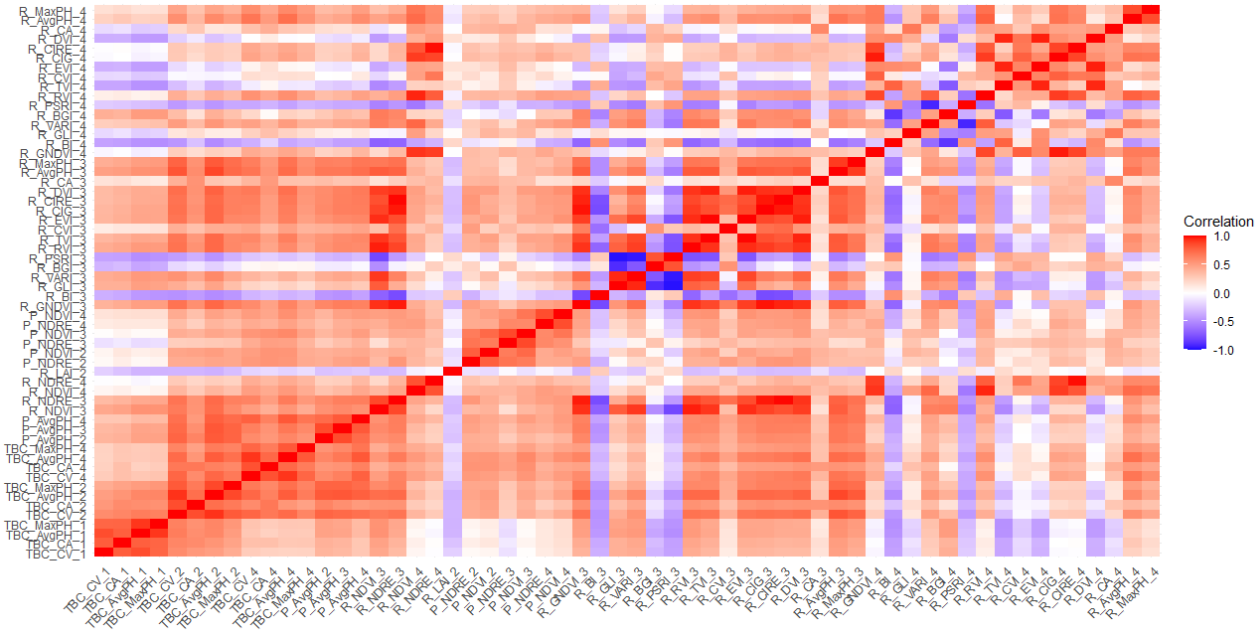

Figure 30 A heatmap visualising the correlation between each measurement method for the 2022 potato trial, with colour intensities reflecting the strength and direction of the correlations.

CANOPY HEIGHT MEASUREMENT

In this study we found that the canopy crop model (CHM) extracted from both R and TBC approach is highly correlated with ground measurement. However, when we further investigate different stages of potato crop and corresponding CHM (**Figure 31**) and 3D canopy model (**Figure 32**), we found that when plants reach the maturity stage, the correlation is weaker compared to vegetative growth and tuber filling stage.

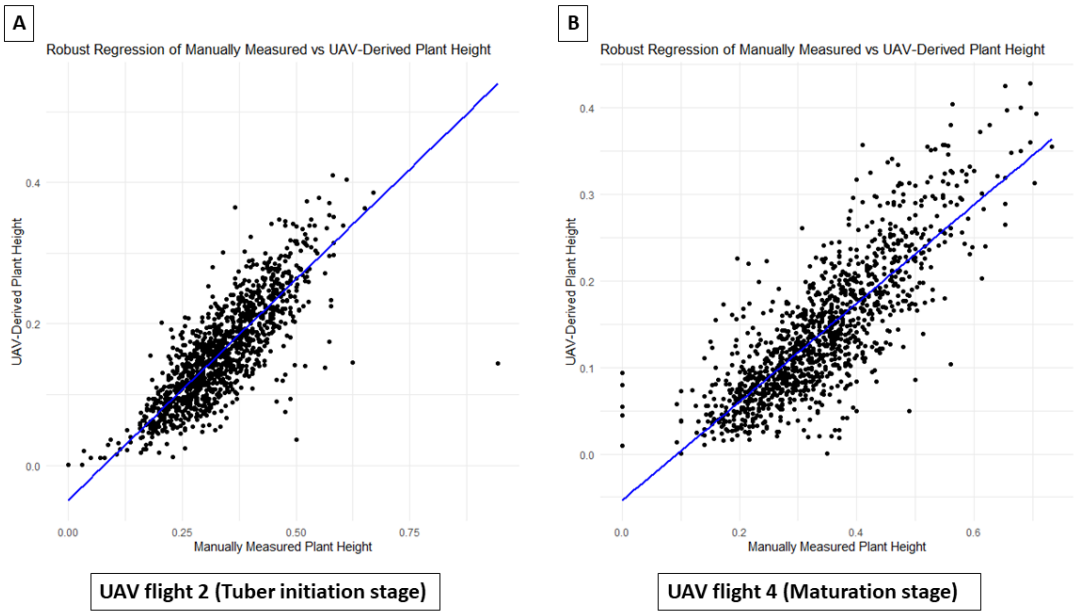

Figure 31 A different stage of potato crop development extracted from UAV-based imagery compare with manual height measurement. The blue line indicates the line of best fit through the data point. (A) Data from UAV flight 3 corresponding to tuber filling stage. (B) Data from UAV flight 4 corresponding to maturation stage.

A robust regression analysis was carried out to evaluate the relationship between the plant height estimated from UAV-derived data (R and TBC approaches) and the plant height measured manually (proximal approach). The 'R package' approach, when applied to average plant height, shows a moderate correlation (0.68) with a corresponding R^2 of 0.4575, indicating that about 45.75% of the variance in plant height is explained by the model. The 'TBC' method scores higher in both correlation (0.78) and R-squared (0.5116), suggesting a stronger and more accurate relationship with average plant height. It has a significantly lower RMSE (0.0498) and MAE (0.1984), indicating more precise predictions. The residuals of the model, representing the differences between observed and predicted values, ranged from -0.1755 to 0.1773, with the median close to zero (-0.0012). This findings mirror similar result reported in previous studies conducted on wheat (Holman *et al.*, 2016; Jimenez-Berni *et al.*, 2018), potato (Njane *et al.*, 2023), and maize Han *et al.* (2018).

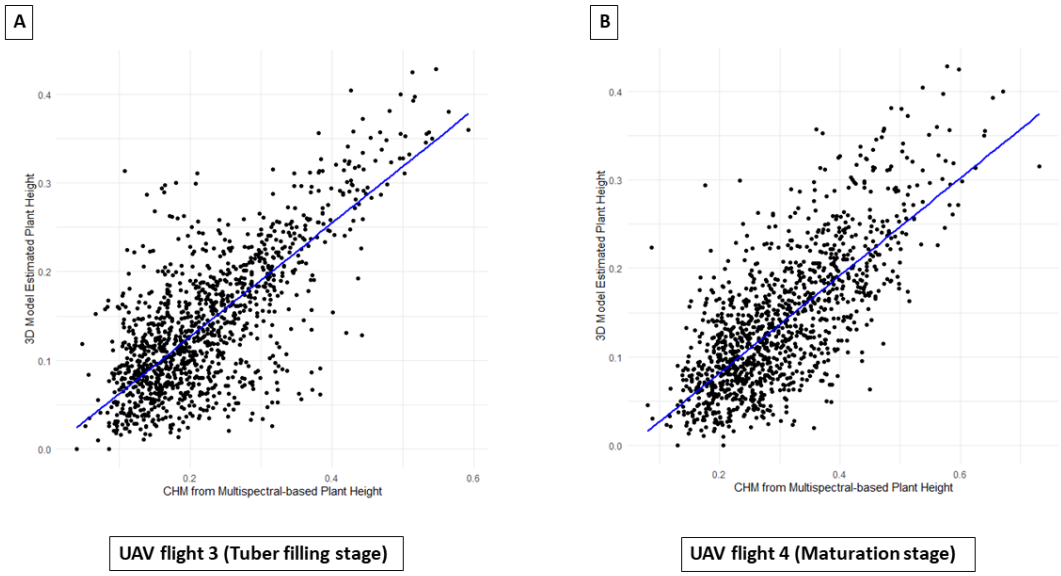

Figure 32 A different stage of potato crop development extracted from UAV-based multispectral imagery (CHM) and 3D point cloud data. The blue line indicates a line of best fit. (A) shows data from UAV flight 3 corresponding to tuber filling stage. (B) the scatter plot represents UAV flight 4 corresponding to maturation stage.

The intercept of the model, which represents the estimated UAV-derived plant height when the manually measured plant height is zero, was not significantly different from zero (coefficient = -0.0535, standard error = 0.0038, t-value = -13.9839). However, the slope of the model was significantly positive (coefficient = 0.5691, standard error = 0.0107, t-value = 52.9735), indicating a strong and positive relationship between manually measured plant height and UAV-derived estimates. For every one unit increase in manually measured plant height, there is an associated increase of 0.5691 units in the UAV-derived plant height estimate.

The robust nature of the regression model helps to mitigate the influence of potential outliers and provides a more reliable estimate of the relationship between these two methods of measuring plant height. The strong positive relationship and significant t-value for the slope suggest that manually measured plant height is a good predictor of UAV-derived plant height for these potato crops.

VEGETATION INDICES

Our analysis revealed a substantial positive correlation between proximal NDVI and UAV-NDVI during the tuber initiation stage, suggesting that both sensor types are consistent

in capturing vegetative health indicators. This correlation diminishes during subsequent stages, possibly due to the differing dynamics of plant growth and changing reflectance patterns as the plant matures. The correlation between proximal NDRE and UAV-NDRE also varies across stages, with a stronger correlation during tuber filling, indicating a more consistent relationship for this metric. Our analysis further identified potential factors contributing to weaker correlations, particularly during the maturation stage for both NDVI ($r=0.581$, $p<.001$) and NDRE ($r=0.562$, $p<0.001$). The difference in data collection methods, with proximal sensors offering a static view and UAV sensors providing a dynamic remote acquisition, likely contributes to variations in correlation strength. Additionally, illumination geometry and shadow effects from potato canopy observed during maturation can impact the accuracy of UAV-derived data, thereby affecting correlations with proximal data. For NDRE, the highest correlation is observed in tuber initiation stage (flight 2) while other stages show weaker correlation especially at the maturation stage (coefficient = 0.431, $R^2 = 0.229$) (**Figure 33**).

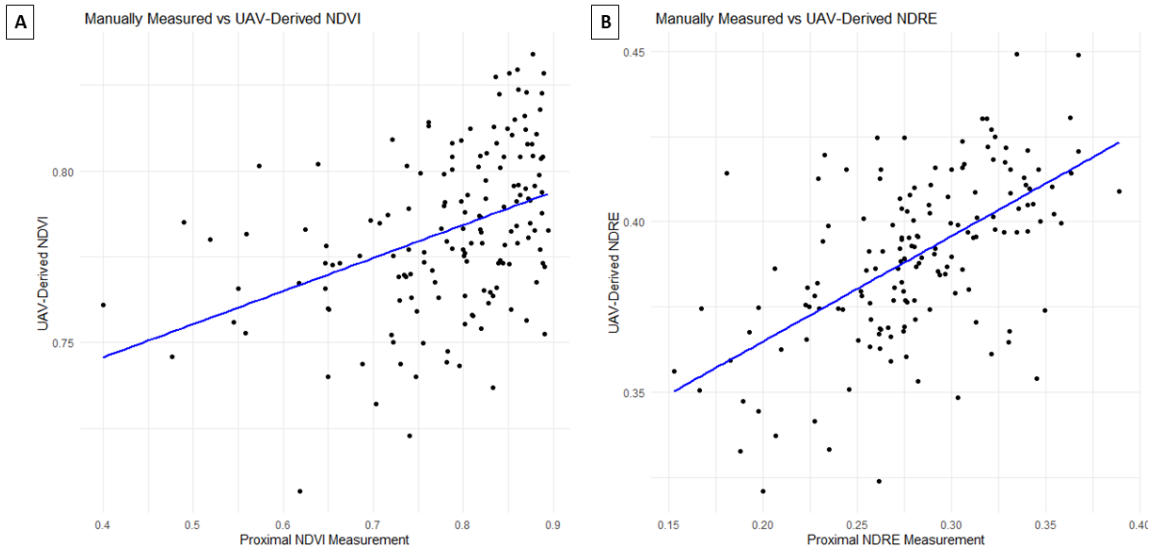

Figure 33 The correlation between the data derived from UAV-based and proximal measurement for NDVI (A) and NDRE (B) vegetation indices from maturation stage (UAV flight 4).

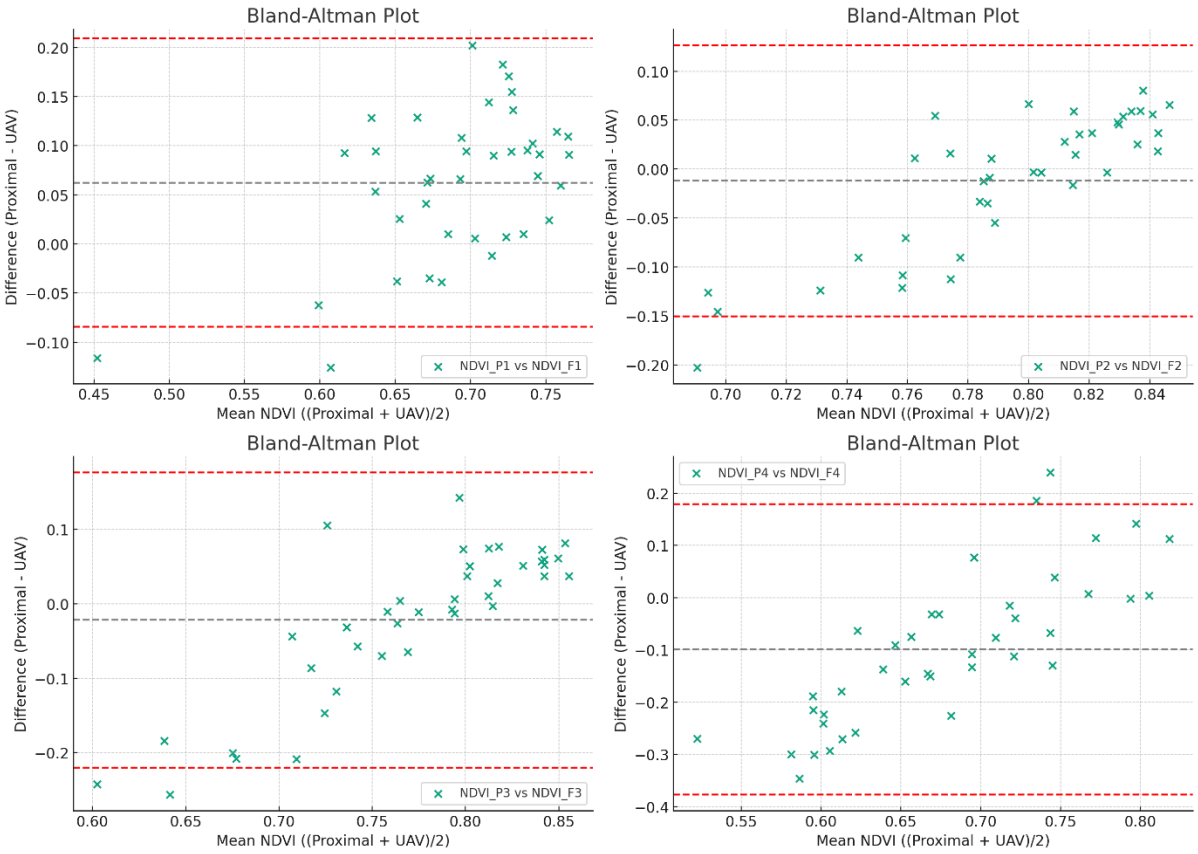

Figure 34 Bland-Altman plots visually evaluate the correlation between proximal and UAV-derived NDVI data for each potato growth stages. The central dashed line in each plot represents the mean difference between the proximal and UAV-derived measurements. The outer dashed lines represent the limits of correlation, calculated as the mean difference ± 1.96 standard deviations of the differences.

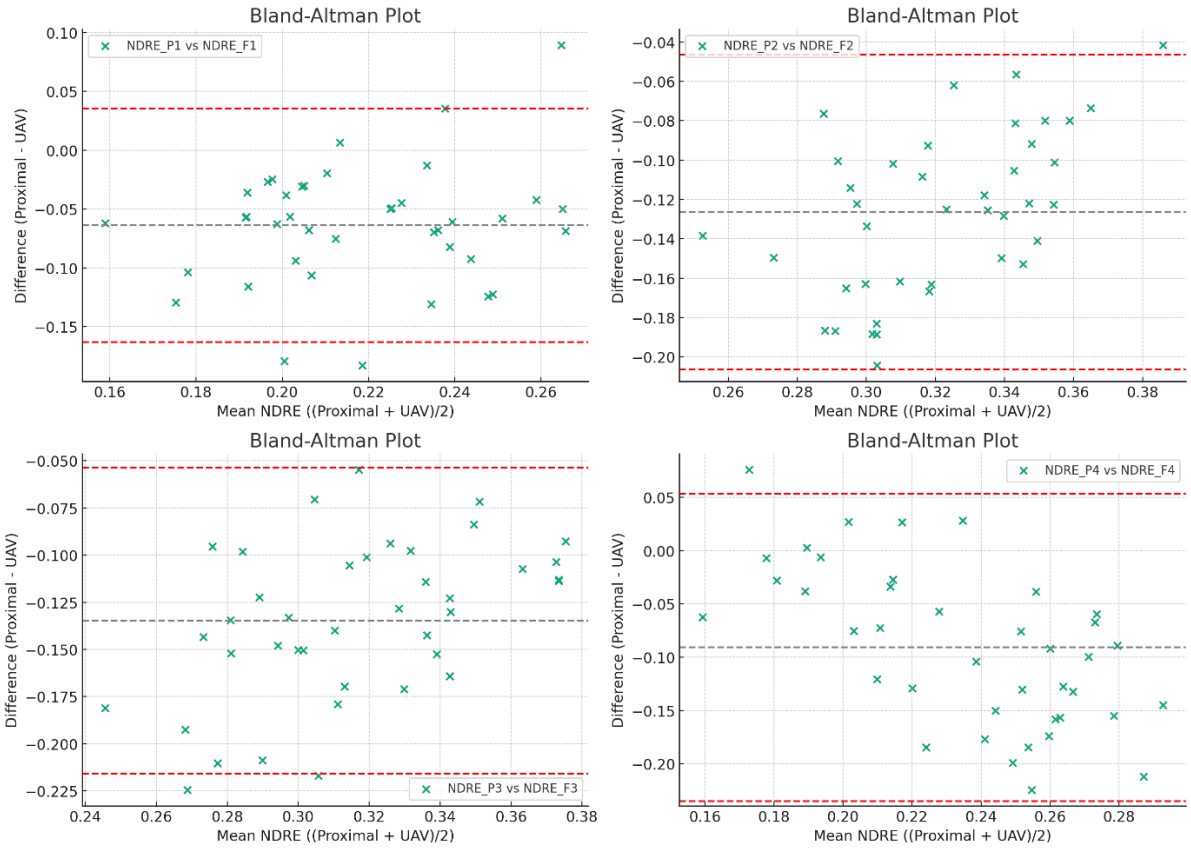

Figure 35 Bland-Altman plots visually evaluate the correlation between proximal and UAV-derived NDRE data for each potato growth stages. The central dashed line in each plot represents the mean difference between the proximal and UAV-derived measurements. The outer dashed lines represent the limits of correlation, calculated as the mean difference ± 1.96 standard deviations of the differences.

To complete our assessment with a Bland-Altman analysis (**Figure 34** and **Figure 35**), we plot the differences against the means of the proximal and UAV-derived NDVI and NDRE measurements for each potato growth stages (four growth stages). For the first UAV flight or tuber initiation stage, there appears to be a moderate spread of differences around the mean, indicating some level of correlation but with variability. Similar observations can be made for flight 2,3, and 4 (tuber initiation, tuber filling, and maturation stages) with differences spread around the mean. The extent of variability and the number of points outside the limits of correlation vary among different growth stages. In summary while there is high correlation between the proximal and UAV-derived data, the Bland-Altman plots reveal variability in this correlation across different potato growth stages. This analysis underscores the importance of considering both systematic differences and the consistency of these differences when

assessing the accuracy and reliability of UAV-derived data against ground-based measurements.

3.3.7 YIELD PERFORMANCE

Potato varieties can be grouped based on yield performance. This process typically involves classifying varieties into distinct groups based on specific yield-related parameters. In this study we employ the potato variety categorised based on yield threshold and clustering analysis techniques.

CATEGORISATION BASED ON YIELD THRESHOLDS

The detailed observation of diseases and growth pattern across different varieties and years provide the dynamics of variety performance under varying environmental conditions. This is essential for effective disease management and field planning. Among over 290 different potato varieties, we categorised them based on their yield performance. Yield performance groups were defined based on quantiles of average yield, dividing yields into "Hight", "Medium" and "Low" categorises (**Figure 36**).

- "High" ~ average yield ≥ quantile (average yield, 0.75)
- "Medium" ~ average yield < quantile (average yield, 0.75) & average yield ≥ quantile (average yield, 0.25)
- "Low" ~ average yield ≤ quantile (average yield, 0.25)

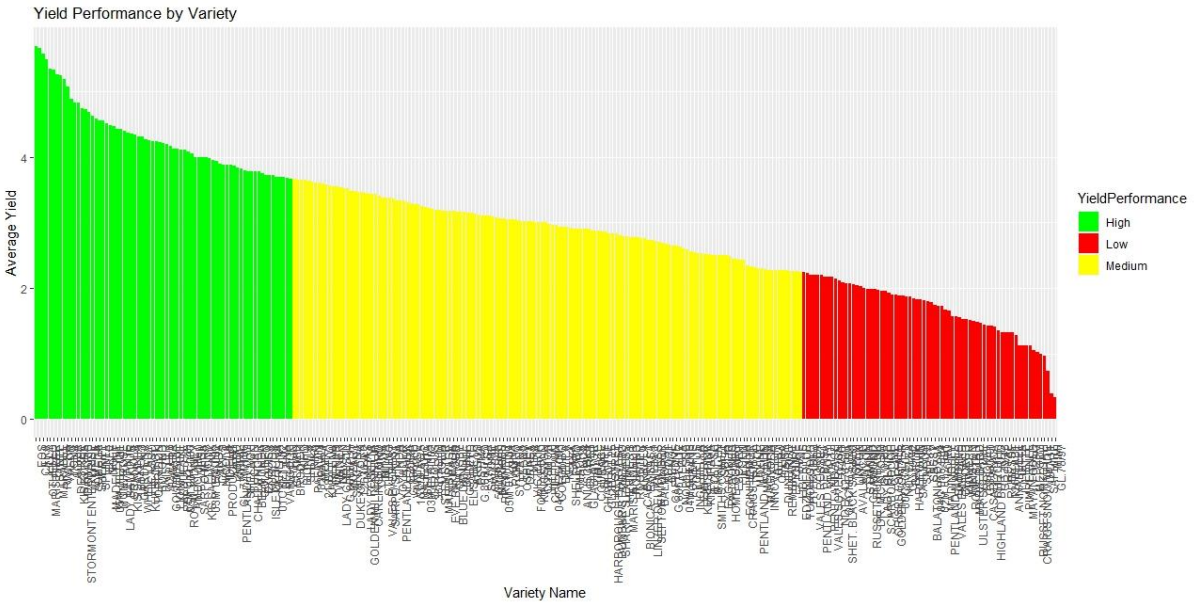

Figure 36 The distribution of yield performance (in kg of fresh weight) by variety for three-year trials, categorised into three groups: High, medium, and low.

CLSTERING ANALYSIS

PCA can be used to reduce the dimensionality of yield-related traits and identify principal components that explain the most variation in yield. By plotting the varieties in PCA space, we can group them based on yield performance and identify key traits driving differences between groups.

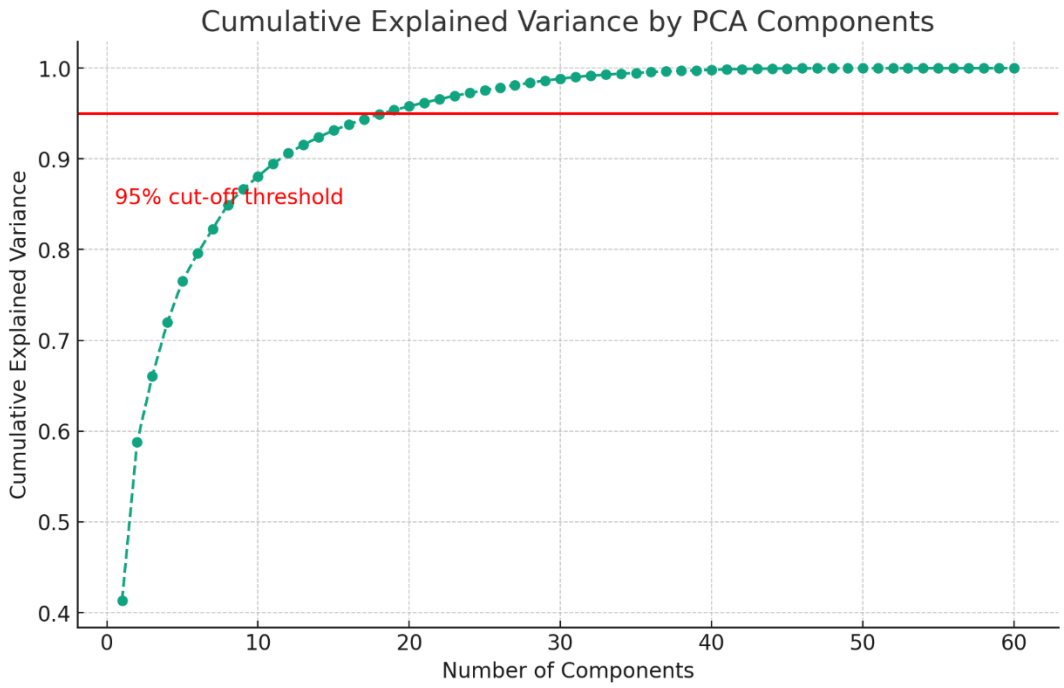

Figure 37 Cumulative explained variance plot shows that 19 principal components are explain around 95% of the variance.

Combining three years trial yield data, we performed Principal Component Analysis (PCA) to reduce the dimensionality of the dataset, which consists of quantitative variables for different potato varieties. It was found that the first 19 principal components explain approximately 95.37 percent of the variance (**Figure 37**). These components were then utilised in a Discriminant Content Analysis (DCA) to differentiate the potato varieties (**Figure 38**). Despite a high number of varieties casing crowding in the middle, K-means clustering allowed us to group potato varieties based on yield performance into 10 distinct groups.

Figure 38 K-mean clustering plot (10 clusters) of potato varieties based on yield performance. Each point represents an individual potato variety, with distinct symbols and colours indicating the distinct clusters identified.

3.3.8 CROP MONITORING AND TIME-SERIES

There are distinct spectral signatures for healthy and diseased plants, which can be detected using different vegetation indices. Different potato varieties have different levels of susceptibility or resistance to specific diseases which is reflected in their spectral profiles. We performed ANOVA for each variety across these different measurement methods for average height and other crop parameters. However, since these measurements are taken at different times (corresponding to different UAV flights), we cannot directly compare the measurements across different flights as they are not repeated measures on the same subject. Consequently, we conducted separate ANOVAs per flight date/growth stage, using a mixed model approach to account for these variations.

The analysis revealed significant variability in phenology and growth traits across potato varieties. The data showed that the varieties 'Markies' and 'Sarpo Mira' achieved the maximum yield in both organic and conventional fields, with 'Sarpo Mira' consistently ranking among the top 10 for yield across three separate trials. In contrast, 'Mimi' consistently produced the lowest yields.

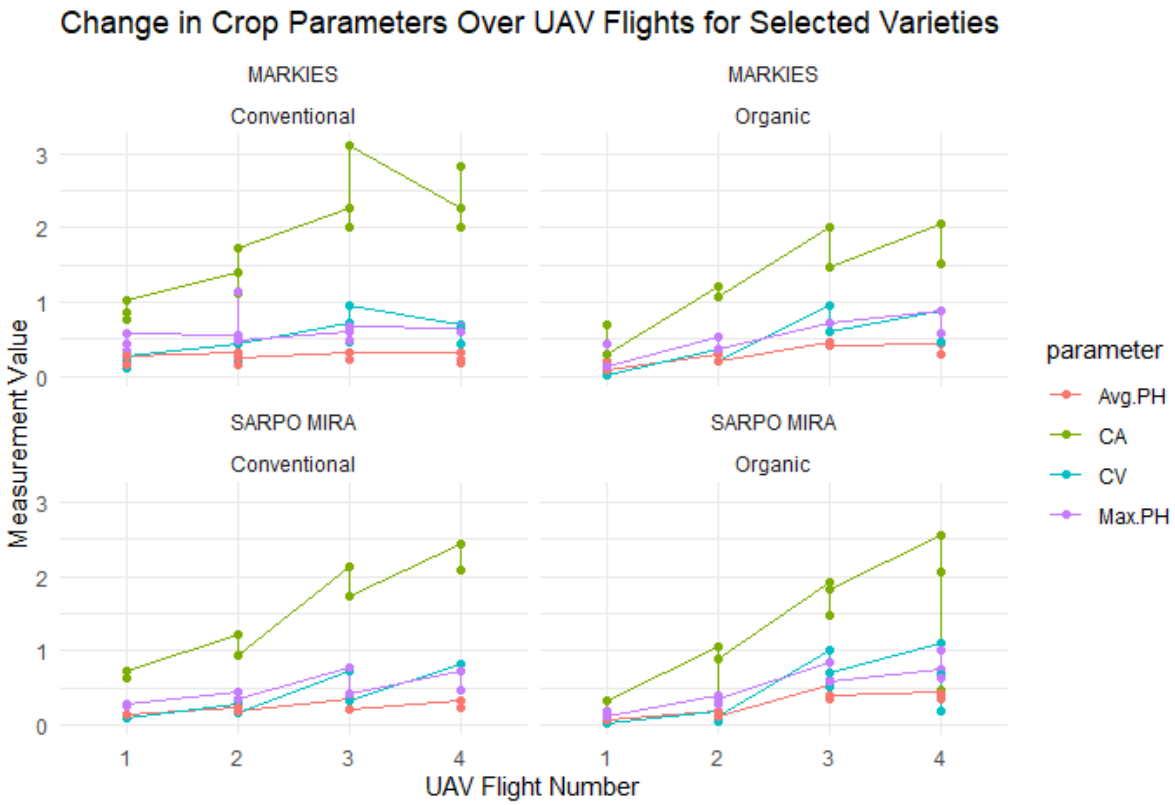

Figure 39 Changes in crop growth parameters including average canopy height (Avg.PH), canopy area (CA), canopy volume (CV) and maximum canopy height (Max.PH) over-time measured over multiple UAV flight for varieties Markies and Sarpo Mira under conventional and organic management (2022 trial).

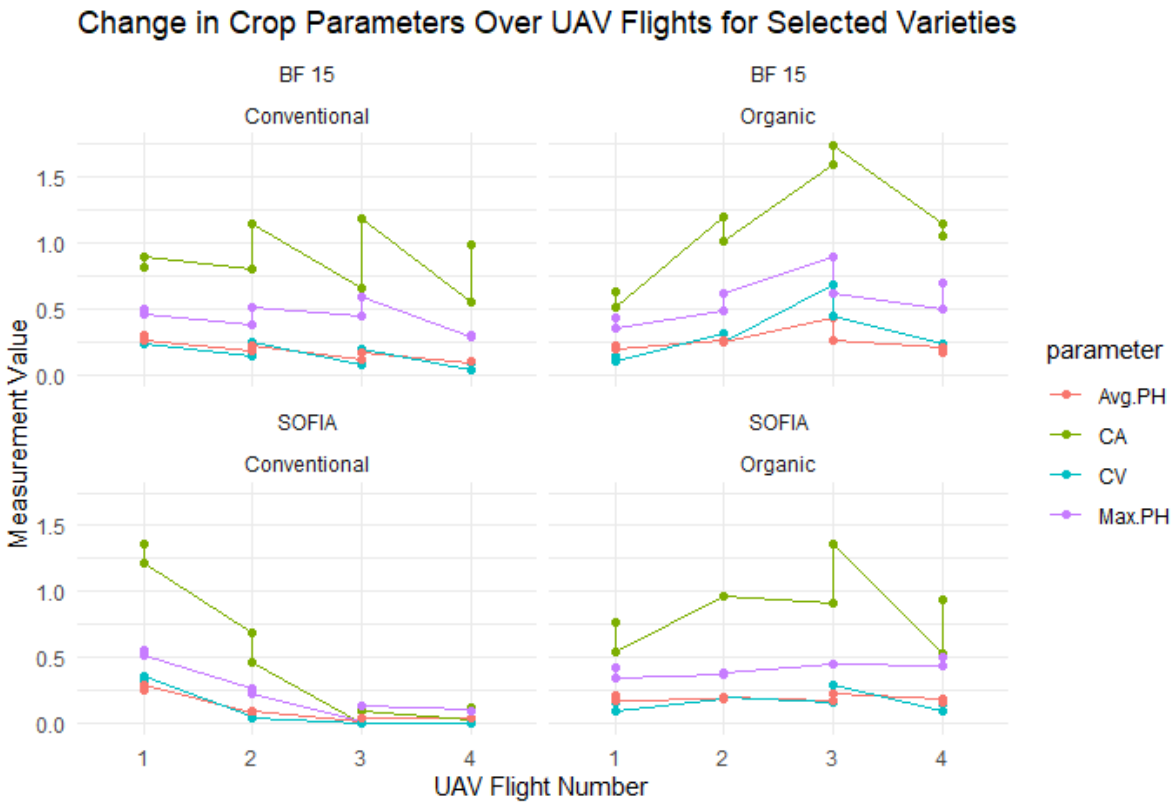

Figure 40 Changes in crop growth parameters including average canopy height (Avg.PH), canopy area (CA), canopy volume (CV) and maximum canopy height (Max.PH) over time for varieties BF 15 and Sofia under conventional and organic management (2022 trial).

The crop monitoring graphs provide insights into how each variety responds to field conditions, revealing patterns of growth that are consistent across both conventional and organic farming practices. As illustrated in **Figure 39**, the parameters including average canopy height (Avg.PH), canopy area (CA), canopy volume (CV), and maximum canopy height (Max.PH) for the varieties 'Markies' and 'Sarpo Mira' under both conventional and organic management show a positive progression over time, which correlates with high yield production. This upward trend demonstrates the robustness and adaptability of these varieties in different farming systems.

In contrast, **Figure 40** examines the varieties 'BF 15' and 'Sofia', showing a decline in the same crop parameters over time under both conventional and organic management. This downward trend indicates less vigorous growth, which subsequently results in lower yields. These comparisons highlight the importance of selecting potato varieties that not only adapt well to specific farming practices but also maintain consistent growth and productivity over the growing season. Tessema *et al.* (2020) also emphasise the importance of acknowledging

the inherent variability among potato varieties in terms of tuber yield and other important growth traits. Such insights are invaluable for farmers and agronomists in making informed decisions about variety selection to optimize yield outcomes in various agricultural management systems.

3.4 CONCLUSION

This comprehensive study demonstrates the possibilities and the challenges of implementing high-throughput field phenotyping (HTFP) using multispectral imagery in potato varietal study trials. The research findings indicate a strong correlation between unmanned aerial vehicle (UAV)-based measurements and traditional proximal methods across various growth stages for diverse phenotype traits. This highlights the effectiveness of UAVs in capturing detailed phenotypic information.

Key findings include the successful application of multispectral imaging to differentiate between healthy and diseased plants, which is critical due to the varying susceptibilities of potato varieties to specific diseases. The study utilised advanced image processing and machine learning techniques to facilitate the extraction of valuable data, such as plant height, vegetation indices, and disease indicators, which are pivotal for informed decision-making in crop management and breeding. However, the study also encountered limitations, primarily due to the vast diversity of potato varieties and the consequent data volume required for robust statistical analysis and model accuracy. The research emphasises the need for more extensive data collection, including both replicates and longitudinal studies, in order to enhance the accuracy of yield prediction models and disease detection methodologies.

Combining technological advancement with traditional farming knowledge offers a hopeful approach to precision agriculture, which improves production, disease control, and sustainable farming methods. There is a potential for a revolutionary impact on potato cultivation, including the possibility of achieving larger yields, enhanced disease resistance, and increased resource use efficiency. However, small-scale farms face significant obstacles, especially due to economic and technological restrictions. The high costs associated with sensor and the complexity of data analysis could potentially restrict the availability of advantages offered by precision agriculture.

This study establishes the foundation for future research by recommending the implementation of larger-scale studies, more thorough investigation of resistance specific to

individual pathogens, and the creation of focused management strategies. By furthering our comprehension of HTFP's capacities and constraints, we can approach the realisation of its complete potential in enhancing worldwide food security and agricultural sustainability.

CHAPTER 4. THE DEVELOPMENT AND VALIDATION OF METHODOLOGY FOR ESTIMATING POTATO MATURATION AND YIELD FORECASTING BASED ON UAV IMAGERY

Studies focusing on what is a suitable resolution to describe the relationship between canopy characteristics to estimate potato maturity and yield estimation at field scale are still lacking. In consideration of this, the study aims to explore the variability across different data collection systems and the dependency of crop growth and disease detection parameters on resolution in potato crops. Through evaluating various imaging sensors and machine learning approaches, this research seeks to identify the most suitable approach and resolution for effective assessment of potato canopy maturity, senescence, and yield prediction.

Potato maturation and yield estimation models have considered a comprehensive array of crop parameters, integrating them with machine leaning models to enhance predictive accuracy. Among the models explored, the Random Forest (RF) algorithm appears to be the most effective for predicting potato plant senescence and yield estimation. In 2022 yield estimation, the RF model, which accounts for canopy volume, canopy area, canopy height, NDVI and NDRE, achieved an R^2 of 78.31. However, in 2020 and 2021 the accuracy of the model was lower, highlight the significant impact field conditions and annual weather variations have on yield. For maturity estimation, RF also best performed model, explaining approximately 53.4% of the variance in the senescence score, with the PLSR having a lower predictive value (50.8%) whilst Decision Tree gave the lowest R^2 score (29.3%).

In conclusion, this chapter presents a perspective on the application of advanced 3D modelling, remote sensing technologies, and machine leaning to accurately evaluate maturity, senescence, and yield prediction of potatoes. By utilising these technologies and addressing existing research gaps, there is a possibility to improve our understanding of crop growth dynamics, disease detection, and ultimately, agricultural productivity and sustainability.

4.1 INTRODUCTION

Potato breeding has evolved over the past 150 years with the goal of developing cultivars that meet the changing demands of consumers and the potato industry. Early potato breeding efforts focused on increasing yield and disease resistance, however, as consumer preferences and market demands changed, breeders began to focus on other traits such as earlier maturity, compact plant size, smooth tubers, skin finish, improved processing quality, and

improved nutritional quality (Jansky, 2009). Growth and development of the potato plant can be categorised into distinct growth stages based on the development of the plant above ground (canopy/haulm) and below ground (tuber), these are influenced by both genetic and environmental factors (Norman and Campbell, 1989; Celis-Gamboa *et al.*, 2003). The information about the canopy and environmental interaction helps determine the impact on vital processes such as photosynthesis, transpiration, growth and infestation of pests and pathogens (Norman and Campbell, 1989). In addition to monitoring these physiological indicators, measuring total leaf area, total dry matter, and light interception can be used to indicate the size and effectiveness of the crop canopy and the crop growth rate (Mackerron and Davies, 1986).

Early yield estimation is crucial in both potato production and breeding process. This allows breeders to identify and select variety that produce higher yield under various environmental conditions, and farmer to optimise the input and plan for market demands (Al-Gaadi *et al.*, 2016; Luo *et al.*, 2020). In potato production, one important factor that affects the quality of harvested potatoes is the physiological status of the crop at harvesting time. It is crucial to distinguish between maturity and senescence in potato production because harvesting at the appropriate maturity stage is critical in order to achieve high-quality tubers while delaying harvest beyond the maturity stage can result in tuber deterioration and loss of yield (Caldiz *et al.*, 2001; Kempenaar and Struik, 2007). Tuber maturity is a predominant factor influencing quality at harvest and throughout storage, mature potatoes tend to store better than immature ones, given their reduced susceptible to skin injury, lower respiration rates, and are decreased susceptible to disease (Heltoft *et al.*, 2017). On the other hand, monitoring and managing senescence is vital in maintaining plant health and productivity and may influence the harvesting time.

Achieving **crop maturity** is complicated and is linked to the maturation of the entire plant, which in potato refers to four distinct processes that occur as the crop approaches harvest (Bussan, 2009). The maturity indicators included haulm/vine greenness (haulm/vine maturity), tuber skin set (physical maturity), tuber dry matter content (physiological maturity) and tuber contents of sucrose, glucose and fructose (chemical maturity) (Heltoft *et al.*, 2017). Tubers must undergo each process; therefore, we can describe the maturity of potato crop as the stage in potato growth when the potato plant has stopped growing, the tubers have reached their full size and have developed the proper dry matter content, starch content, skit set and is physiologically ready for harvest (Bussan, 2009). Maturity can vary depending on the potato

variety/cultivar and environmental conditions, but it is generally between 90 to 120 days after planting. However, weather conditions are a key factor affecting growth and development of the plant, as well as the spread of pest and diseases affecting potato crops (Kooman *et al.*, 1996; Economou *et al.*, 2023). Consequently, these factors can impact the timing of plant maturity.

Senescence in potato plants is a natural ageing-dependent plant process that occurs after maturity or 14-28 days after flowering (Meyer and Neto, 2008). Meyer and Neto (2008) explained that during senescence, the above ground/ haulm begins to deteriorate, leading to plant death on a discrete visual scale. In breeding, plants that senesce more slowly may reach higher yield as plants retain higher photosynthetic capacity (Kooman *et al.*, 1996). However, senescence is not only influenced by internal factors (genetic) but also multiple environmental factors such as temperature, light and nutrient availability and it also can be accelerated by stress factors such as drought and pathogen infection (Bussan, 2009; Guo *et al.*, 2021; Economou *et al.*, 2023).

Conventionally, potato canopy maturity and senescence can be assessed using a combination of visual observation, physiological measurements, and laboratory analysis. Visual observation is one of the simplest, changes during maturation and senescence typically includes degradation of chlorophyll in the leaf (turn yellow). Destructive laboratory analysis, such as measuring leaf chlorophyll content, sugar, and nitrate levels in petiole sap, can provide valuable information about the physiological status of the potato (Mackerron and Davies, 1986). A sudden decline in vegetative traits such as height and canopy area can also be used as an appropriate indicator of the onset of maturity and then senescence (Mackerron and Davies, 1986). These indicators can give insights into the plant's nutrient status, photosynthetic efficiency, and overall health. Therefore, we can describe the maturity of crop as the stage in potato growth when the potato plant has stopped growing and is ready to harvest.

4.1.1 BRIEF OVERVIEW OF THE 3D-MODEL APPROACH AND ITS POTENTIAL BENEFITS

In recent years, the deployment of unmanned aerial vehicles (UAVs) and a variety of imaging technologies have been used to collect data for quantitative studies of complex plant traits related to growth, yield, and adaptation to biotic and abiotic stress (Li *et al.*, 2014; Comba *et al.*, 2018). The technique that are being used in plant phenotyping platforms include visible imaging (machine vision), imaging spectroscopy (multispectral and hyperspectral remote

sensing), thermal infrared imaging, fluorescence imaging, 3D imaging, and tomographic imaging (MRT, PET and CT) (Li *et al.*, 2014).

Image-based modelling technique can combine high-quality photogrammetry and efficient computer vision algorithms using 2-dimention (2D) images to obtain 3-dimention (3D) models (Herrero-Huerta *et al.*, 2015). Image mapping on three-dimensional models aims to define the texture or colour information on the surface, which can be represented as point-clouds or triangle meshes (Jurado *et al.*, 2022). Point-cloud refers to a 3D coordinate system dataset that represents external surface of visible objects where light is reflected (Comba *et al.*, 2018). 3D point cloud can be obtained directly from laser scanners (such as light detection and ranging systems—LiDAR) (Lumme *et al.*, 2008) derived from RGB, multispectral and thermal imagery by point cloud generating method such as Structure from Motion (SfM) (Akca and Grün, 2007; Comba *et al.*, 2018). SfM is a common application of remote sensing methods, stereo photogrammetry or air born scanning are used to extract plant height and crop surface model (CSM) in agriculture. (Bendig *et al.*, 2013a). 3D model has been utilised for studying crop development extensively in vineyard (Herrero-Huerta *et al.*, 2015; Mack *et al.*, 2017; Comba *et al.*, 2018; De Castro *et al.*, 2018). However, there is a potential to utilise the 3D model to study the variability of potato canopy characteristics.

4.1.2 UTILISING PLANT 3D MODELS

Utilising a 3D model of the potato canopy is a useful tool to study the variability of potato canopy characteristics and monitor the changes in canopy over time (de Jesus Colwell *et al.*, 2021). A 3D model of the potato canopy can be used to determine when the plant is approaching maturity and senescence, firstly, changes in the measured height of the canopy throughout the growing season may provide indications of plant approaching maturity and senescence (Ji *et al.*, 2022). Secondly, the 3D model can be used to estimate and monitor the changes in leaf area index (LAI) or canopy cover, the LAI may decrease as the plant mature and senesce. LAI is a valuable measurement in helping to assess canopy density and biomass. Thirdly, changes in density/volume of the canopy at different heights can provide insight into the distribution of biomass within the canopy and how it changes over the growing season. Additionally, as the potato plant senesce, the leaves may turn yellow and die off (Heltoft *et al.*, 2017). By using a 3D model, the spatial patterns of senescence can be evaluated within the canopy. Overall, the use of a 3D model can provide a comprehensive understanding of how

senescence progresses through the canopy and how it affects the overall health and productivity of the plants (yield performance).

4.1.3 RELEVANT THEORETICAL CONCEPTS AND METHODOLOGIES

The integration of unmanned aerial vehicle (UAV)-remote sensing and 3D modelling at the field scale has marked a significant advancement in agricultural research and application. Mack *et al.* (2017) proposed the automated approach for accurately phenotyping grape bunches through complete 3D reconstruction from high-resolution sensor data. This approach's high precision and recall rates highlight its potential for practical application such as quality control and yield estimation in grape production. Comba *et al.* (2018) introduced an unsupervised algorithm for vineyard detection and vine-row features evaluation using 3D point-cloud maps. This algorithm was found to be efficient and robust, even with missing plants and steep terrain slopes. Li *et al.* (2022) proposed the use of 3D point clouds generated from UAV oblique imagery coupled with an adaptive micro terrain model, to map the leaf area index (LAI) and height of maize canopies, showing strong correlations with manual measurements.

To develop a methodology for estimating potato maturation and yield forecasting based on crop parameters extracted from the UAV imagery approaches includes process-based modelling and statistical modelling (Meyer and Neto, 2008). Statistical modelling such as Simple or multiple linear regression (MLR) models are the most commonly used to quantifying the effect of each variable (Guo *et al.*, 2021), other common supervised machine-learning algorithm for predictive model such as random forest, regression tree and neural systems are distinct from classical statistical modelling as they account for complex non-linear relationships in datasets (Jurado *et al.*, 2022). Statistical modelling is able to reproduce many of the key features of process-based model response to weather change (Guo *et al.*, 2021). Process-based modelling is more focused on the underlying processes and interaction in the environment (Guo *et al.*, 2021).

The literature review reveals that the majority of the previous research has focused on identifying and characterising gene associated with plant maturity and senescence, such as the studies by Celis-Gamboa (2002) and Hurtado *et al.* (2012). This is due to the recognition that chemical maturity is the most related to above ground plant maturity. For example, Hurtado *et al.* (2012) introduced a smoothed generalised linear model with a Quantitative Trait Loci (QTL) analysis to decode the genetic underpinnings of potato senescence.

To investigate the potential of potato maturity indicators measured in the field, this step involves identifying and extracting relevant features from the 3D model that are indicative of potato maturity. These features are then used to classify the maturity stage of the potatoes through machine learning algorithm, such as convolutional neural networks (CNNs), support vector machines (SVMs), Random Forest (RF) or decision trees, and can be used to train and annotate datasets to recognise different maturity stages (Teodoro *et al.*, 2021; Njane *et al.*, 2023). Further contribution to this field include Heltoft *et al.* (2017), who developed a liner regression model to categorise potato cultivars and maturity levels using indicators measured 1-3 weeks before harvest and at harvest. Ávila-Valdés *et al.* (2020) implemented principal component analysis (PCA) to differentiate between genotypes. In a related methodology, Teodoro *et al.* (2021) explored the application of machine learning, including RF, SVM and linear regression to predict days to maturity based on plant height in soybeans. Given that potato maturity significantly impacts tuber quality management and yield prediction, yet it is complex and involves a range of process including haulm senescence, tuber maturation and physiological maturation (chemical maturation), this chapter aims to investigate the feasibility of using plant traits extract from 3D model for maturity prediction.

These methods mentioned above can be used to monitor the development of potato over the growing season and estimate the best harvesting period for maximum yield and quality. Due to the nature of conventional methods which are destructive, time-consuming, laborious, and infeasible for a large-scale crop monitoring. Therefore, the use of remote sensing has been introduced to accurately monitoring crop and estimate the best harvesting time to achieve best yield performance and quality. However, the effectiveness of UAV-based maturity and yield estimation depends on several factors, including the quality and resolution of the imagery, the algorithm used for image processing and classification, and the variability of the potato varieties and growth conditions.

4.2 DATA ANALYSIS

The data analysis method employed tin this chapter correspond with those outlines in Chapter 2. These include key analysis such as correlation and model evaluation. For yield estimation, the model's performance was evaluated by calculating the Mean Absolute Error (MAE), mean squared error (MSE), root mean squared error (RMSE) and R-squared (Coefficient of Determination) between the predicted yield values and the actual fresh yield values in the test set.

Furthermore, model calibration and validation were performed to evaluate the precision of predictions obtained from the remote sensing platform. The model's performance was assessed to determine if the data if the data extracted from the remote sensing platform overpredicts or underpredicts compared to yield reference. The residuals of the prediction model were analysed to assess the differences between the observed values (actual yield) and the predicted values. A positive residual indicates an underprediction (the actual value is higher than the predicted value), while negative residual indicates an over prediction (the actual value is lower than the predicted value).

$$R = Y_{actual} - Y_{predicted} \quad (4)$$

4.3 RESULTS AND DISCUSSION

4.3.1 CROP MATURITY

Canopy maturity indicators included haulm/vine greenness (haulm/vine maturity), tuber skin set (physical maturity), tuber dry matter content (physiological maturity), and tuber contents of sucrose, glucose, and fructose (chemical maturity) (Heltoft *et al.*, 2017). A potato crop reaches maturity when the plant stops growing, the tubers have fully developed in size, and they have the proper dry matter content, starch content, skin set and are physiologically ready for harvest (Bussan, 2009). Maturity can vary depending on the potato variety/cultivar and environmental conditions, but it typically occurs between 90 and 120 days after planting, with some variations. Moreover, weather conditions significantly influence phenological growth and development and regulating the spread of pest and diseases of potato crops (Kooman *et al.*, 1996; Economou *et al.*, 2023). Therefore, the is a need for field inspection to evaluate the stage of plant maturity and make a harvesting decision.

Estimating crop maturity in the field involves walking through the field to assess plant stages, this helps farmers decide when to make a final decision on the harvesting time. This method is cost-effective and does not require expensive equipment, and provides immediate results, making it a practical and efficient approach. However, in large-scale fields with diverse plant verities, this method poses a challenge in term of accurately recording the plant growth stage. The UAV imagery has made predicting potato plant senescence feasible. Although studies of potato senescence are limited, methodologies from other crops, such as soybeans (Teodoro *et al.*, 2021), have been adapted to this study using machine learning algorithms.

For each plot inspected and recorded the senescence score based on the condition of the plant haulms using the rating system from **Table 5** Senescence rating system for potato field trials to assign a score ranging from 1 to 7, where 1 indicates excellent health and 7 indicates complete death of the plant. If the majority of the plants are in the range of scores 4-5, it indicates that the crop is nearing or at the ideal maturity for harvesting.

To build a model for predicting the potato plant maturity, we need to understand how potato plants change overtime. The sudden decrease in the plant height during the growth period was a result of senescence where the leaves begin to wilt thus leading to weaker stems and decreased height. It was observed that from UAV flight 3 (tuber filling/bulking stage) there was a sudden slight decreased in height in most varieties. However, there are some varieties such as ‘Cara’ and ‘Sarpo Mira’, that maintained height although it is approaching the maturation stage. The correlation tests were performed to identify which plant features in the dataset are most relevant to predicting maturity. The Biomass index (BI) had the highest correlation with the Senescence scores, follow by transformed vegetation index (TVI). Therefore, we build the model based on the vegetation indices extracted from UAV flight 4 (maturation stage).

Before developing the model, we utilised the partial lease square discriminant analysis (PLS-DA). From the PLS-DA score plot showing a clustering of potato varieties regarding the different vegetation indices. PLS-DA is particularly useful in situation where the goal is to predict or classify observations into groups based on predictor variable and can also be used to visualise the data in a low-dimensional space.

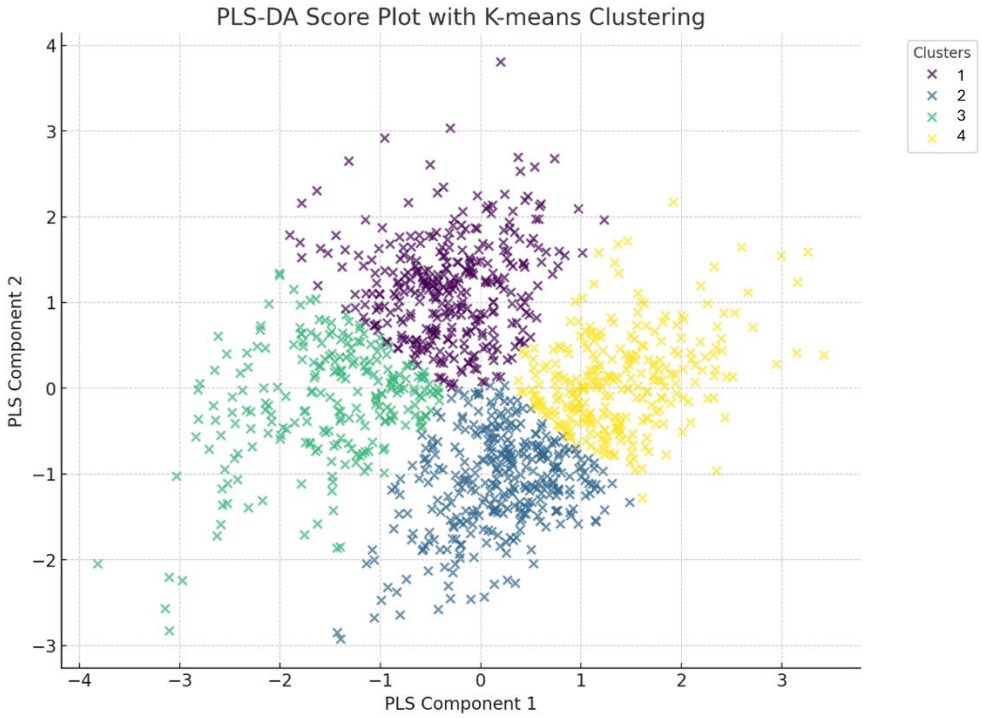

Figure 41 PLS-DA score plot with K-means clustering (k=4) The clusters show four group of potato varieties based on their PLS-DA scores which have similar vegetation indices and senescence scores.

PLS-DA score plot with K-means clustering (k=4) suggest some variation in the senescence across clusters, with cluster 3 showing the lowest average senescence score (**Table 30**), indicating potentially healthier or less green plants compared to the other clusters (**Figure 41**). Clusters that are closer to each other in the plot represent varieties with similarity in the measured spectral properties, whereas clusters that are further apart indicate more significant differences. These grouping also reflect how different varieties respond to environmental conditions. The clustering thus provides a framework for exploring the relationships among potato varieties based on spectral data.

Model development

In this step, we split the data into training (80%) and testing (20%) sets. Three different models, including Random Forest (RF), Decision Tree, partial lease square regression (PLSR) was used to predict potato plant maturity using field Senescence scores as the target variable.

Table 12 Statistical results of the performance of three predictive models (Random Forest, Decision Tree and Partial Least Squared Regression (PLSR) based on their R-squared and Mean absolute Error (MAE).

Model	R-squared (R^2)	Mean Absolute Error (MAE)
Random Forest	0.534	0.636
Decision Tree	0.293	0.675
PLSR	0.508	0.657

The Random Forest has the lowest MAE, suggesting it makes predictions closest to the actual values, followed closely by PLSR, with the Decision Tree having a slightly higher MAE. In terms of the R^2 score, which indicates the proportion of variance in the dependent variable that is predictable from the independent variables, the RF performs the best, explaining approximately 53.4% of the variance. The PLSR is slightly behind, while the Decision Tree has the lowest R^2 score, explaining about 29.3% of the variance.

The Random Forest model appears to be the most effective for predicting potato plant senescence among the models tested (**Table 12**), offering a good balance between complexity and prediction accuracy. This suggests that leveraging the ensemble method provided by Random Forest helps in capturing more complex patterns in the data without fitting too closely to the training set, a common limitation of the simpler Decision Tree model.

From the results of this study, potato maturity estimation shows that the Random Forest (RF) model performs the best, explaining approximately 53.4% of the variance in "Senescence score". The PLSR is slightly behind at 50.8%, while the Decision Tree regression has the lowest R^2 score at 29.3%. This suggests that the statistical method provided by Random Forest is better at capturing complex patterns in the data without overfitting, a common limitation with the simpler like Decision Tree.

Although the R^2 score from these models are not exceptionally high, the RF model still appear to be the most effective for predicting potato plant senescence among the tested models. This opens a way for further research into senescence estimation in potatoes. Future studies could involve analysing the average or distribution of senescence scores with cluster analysis to directly relate these cluster to senescence scores. This could potentially reveal how different senescence stages correspond to the spectral characteristics, and when compared with known senescence stages or other physiological data, it would help validate interpretations and understand the practical significance of the clustering in terms of potato plant health and

management. Nevertheless, this approach offers a data-driven way to classify and understand variability in potato plants based on spectral indices, contributing to more informed agricultural practices and studies on plant health.

4.3.2 YIELD CORRELATION

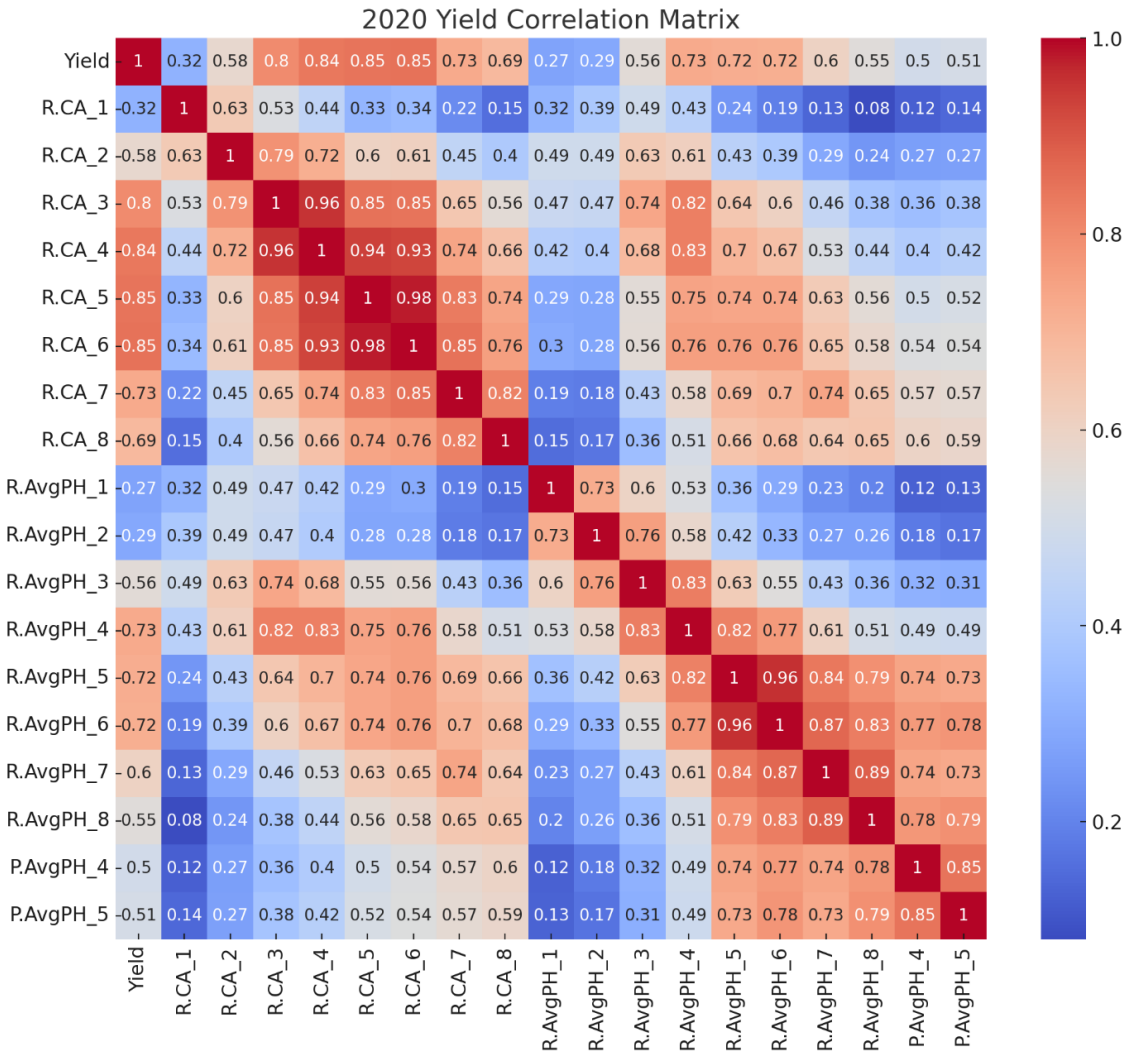

Figure 42 Heat map shows correlation between yield and other crop parameters for the 2020 potato trial. The colour scale indicates the strength of the correlation: red (value closer to 1.0) indicate strong positive correlation, blue (closer to -1.0) indicate strong negative correlation, and white represents weak or no correlation (close to 0).

Figure 42 shows a correlation heat map between yield and other crop parameters for 2020 potato trial. In this figure, CA refers to canopy area and AvgPH refers to average canopy height. The prefix 'R' refers to data extracted from aerial images using R software, while 'P' indicates data obtained using proximal sensing methods. The number following each parameter

(e.g., 1-8) correspond to UAV flight numbers, which are associated with different potato plant growth stages, as detailed in **Table 6**. The heat map highlights the average plant height from flight number 4, 5 and 6 (tuber filling stage) showed the highest correlation with the fresh yield as during these UAV flights the potato crop was during the tuber filling stage.

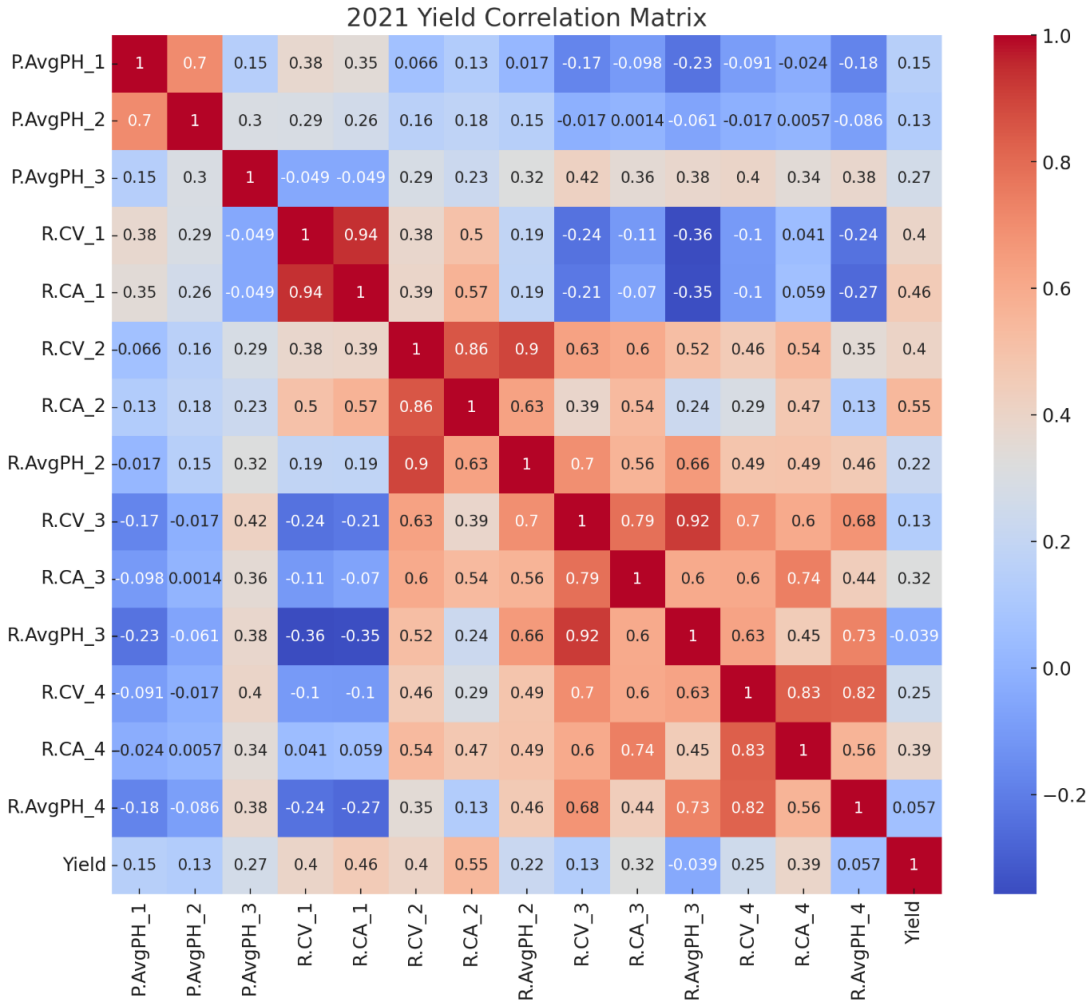

Figure 43 Heat map shows correlation between yield and other crop parameters for 2021 potato trial. The colour scale indicates the strength of the correlation: red (value closer to 1.0) indicate strong positive correlation, blue (closer to -1.0) indicate strong negative correlation, and white represents weak or no correlation (close to 0).

Figure 43 shows a correlation heat map between yield and other crop parameters for 2020 potato trial. In this figure, CA refers to canopy area, CV represents canopy volume, and AvgPH refers to average canopy height. The prefix 'R' refers to data extracted from aerial images using R software, while 'P' indicates data obtained using proximal sensing methods. The number following each parameter (e.g., 1-4) correspond to UAV flight numbers, which are associated with different potato plant growth stages, as detailed in **Table 6**.

As seen in the heatmap (**Figure 42**, **Figure 43**, and **Figure 44**) the correlation is different when look at the different years due to differences in weather pattern. The correlation between yield and crop parameters are relatively low compared to 2021 and 2022 trial, with the canopy area from flight number 2 (tuber initiation) showed the highest correlation.

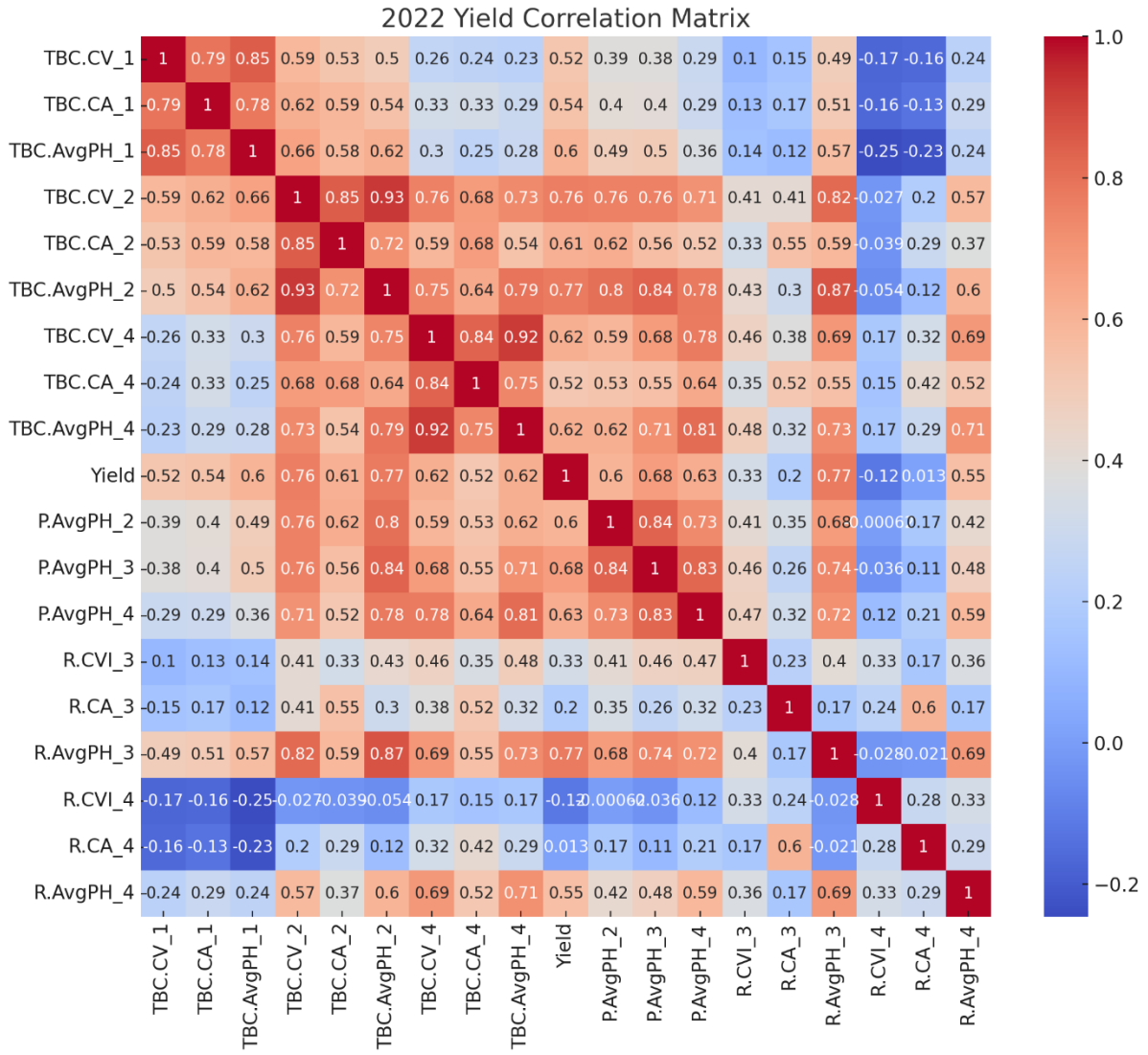

Figure 44 Heat map shows correlation between yield and other crop parameters for 2022 potato trial. The colour scale indicates the strength of the correlation: red (value closer to 1.0) indicate strong positive correlation, blue (closer to -1.0) indicate strong negative correlation, and white represents weak or no correlation (close to 0).

Figure 44 shows a correlation heat map between yield and other crop parameters for 2020 potato trial. In this figure, CA refers to canopy area, CV represents canopy volume, and AvgPH refers to average canopy height. The prefix 'R' refers to data extracted from aerial images using R software, while 'P' indicates data obtained using proximal sensing methods.

The number following each parameter (e.g., 1-4) correspond to UAV flight numbers, which are associated with different potato plant growth stages, as detailed in **Table 6**.

In the 2022 trial, the average plant height from flight number 3 (tuber filling stage) showed the highest correlation with the fresh yield. Follow by average height and canopy volume from flight number 2 (tuber initiation stage).

4.3.3 YIELD ESTIMATION

For yield estimation, the actual yield data is needed to calibrate the estimation method. In this study, we performed the correlation map between yield, different vegetation indices and plant height at different growth stages. The data was split into a training set (e.g., 70 or 80 percent) and a testing set (e.g., 20 or 30 percent) for model evaluation. To perform a linear regression model for yield estimation, average height and different vegetation indices data were extracted from the multispectral data collected during the UAV flights. Height measurement is commonly used to represent plant growth including potato. We have shown the compared with the average height provides more consistent measure of plant height in Chapter 3. As demonstrated by using temporal data on the canopy, the average height is a better represent of plant growth.

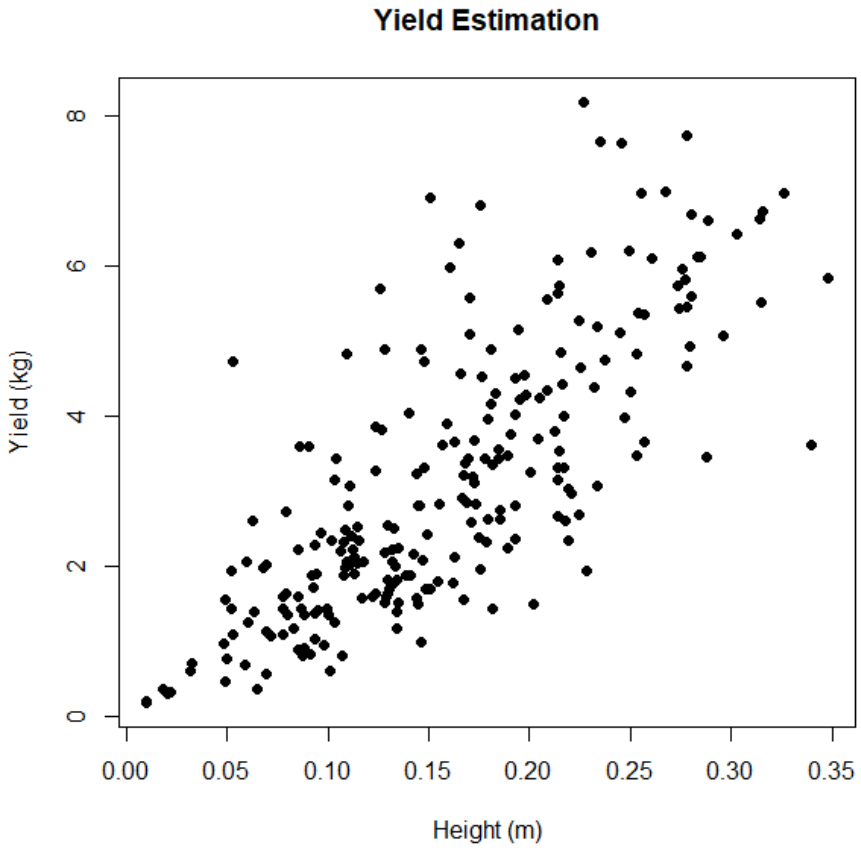

Figure 45 Result extracting potato height and yield estimation model (potato trial 2022).

2020

Before running the yield prediction model, the correlation test has been tested to see what flight number and what crop parameter gives the highest correlation value to the fresh yield weight (kg). From the correlation test, flight number 5 and 6 were highly correlated with the yield. The canopy area (flight number 5 and 6, 0.8453 and 0.8478) is the most correlated with the yield, follow by mean height (flight 6, 0.7136) and NDVI (flight 6, 0.6387) (**Figure 42** Heat map shows correlation between yield and other crop parameters for the 2020 potato trial.).

Table 13 Statistical results of fit linear regression model with difference predictors and their evaluation metrics (R-squared, P-value, MAE, MSE, and RMSE) for 2020 potato trial.

Predictor	Model fit	P-value	MAE	MSE	RMSE	Fitting equation

Mean Height	R-square =0.511	P<0.001	0.788	1.112	1.054	Y=-0.734+15.279*Mean height
NDVI +Mean Height	R-square =0.567	P<0.001	0.740	0.981	0.990	Y=-4.338+6.296*NDVI+11.1847*Mean height
Canopy area (5)	R-square =0.714	P<0.001	0.5941	0.6381	0.8356	Y=0.231+10.19*Canopy area
Canopy area (6)	R-square =0.719	P<0.001	0.5998	0.665	0.8155	Y=0.115+9.816*Canopy area
Canopy area (6) + NDVI	R-square = 0.725	P<0.001	0.586	0.655	0.809	Y=-1.699+8.941*Canopy area+2.67*NDVI
Canopy area (6) + Mean Height	R-square = 0.734	P<0.001	0.586	0.655	0.809	Y=-0.3168+8.211*Canopy area + 3.8585*Mean height
Canopy area (6) + NDVI +Mean height	R-square = 0.736	P<0.001	0.586	0.651	0.807	Y=0.116+9.817*Canopy area

The examination of yield predictors through linear regression models unveiled the importance of both spectral and structural crop parameters (**Table 13**). The model utilising canopy area from the tuber filling stage (UAV flights 5 and 6) demonstrated strong predictive capabilities, with R^2 values of 0.714 and 0.719 respectively. Adding NDVI to canopy area from

flight 6 slightly elevated the R^2 to 0.725, while a model combining canopy area from the same flight with average height achieved an R^2 of 0.734. The most predictive model included canopy area from flight 6, NDVI, and average height, together achieving the highest R^2 value of 0.736.

Table 14 The performance of random forest model applied with difference set of predictors and their evaluation metrics (R-squared, MAE, MSE, RMSE, and % variance explained) for 2020 potato trial.

Variable	R-squared	MAE	MSE	RMSE	% Var explained
Canopy area (5)	0.892	0.344	0.236	0.486	59.58
Canopy area (6)	0.898	0.349	0.223	0.472	61.35
Canopy area + NDRE	0.93	0.271	0.154	0.393	71.13
Canopy area + NDVI	0.945	0.252	0.12	0.346	71.34
Canopy area + NDRE + average height	0.938	0.255	0.136	0.37	73.39
Canopy area + NDVI + average height	0.95	0.24	0.113	0.337	75.88

The Random Forest models across various predictor combinations for yield prediction demonstrated significant insights into the predictive power of integrating different crop parameters (**Table 14Error! Reference source not found.**). Canopy area alone from UAV flight 5 and 6 already demonstrated high predictive accuracy, with R^2 values nearing 0.9. With the model using canopy area from flight 6, achieving an R^2 of 0.898 and explaining over 61% of the yield variance. The integration of NDRE and NDVI into models with canopy area significantly enhanced prediction capabilities, NDRE reached an R^2 of 0.93, NDVI further improving this to R^2 of 0.945, both models explaining more than 71% of the yield variance.

Moreover, when added average plant height as a predictor alongside canopy area and spectral indices led to even more robust models. The combination of canopy area, NDVI, and average height outperform all model with R^2 of 0.95 and explaining approximately 76% of the yield variance.

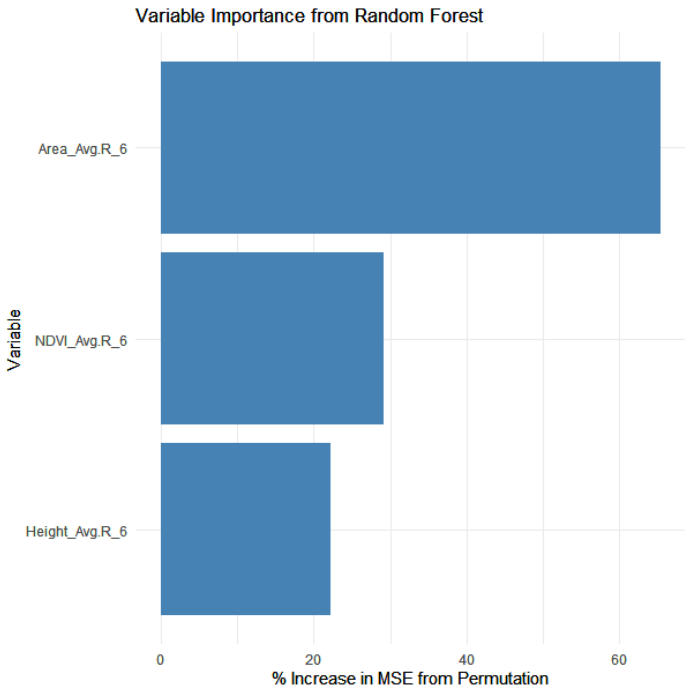

Figure 46 The variable importance from a random forest model. The percentage increased in Mean Squared Error (MSE) from permutation from three variables (canopy area, NDVI and height).

Visualising the random forest regression model (**Figure 46**) indicates that canopy area is the most important variable with the highest percentage increase in MSE when this variable is permuted. This suggest that changes in canopy area significantly affect the model's accuracy and its strong predictive power regarding the fresh yield.

Decision Trees

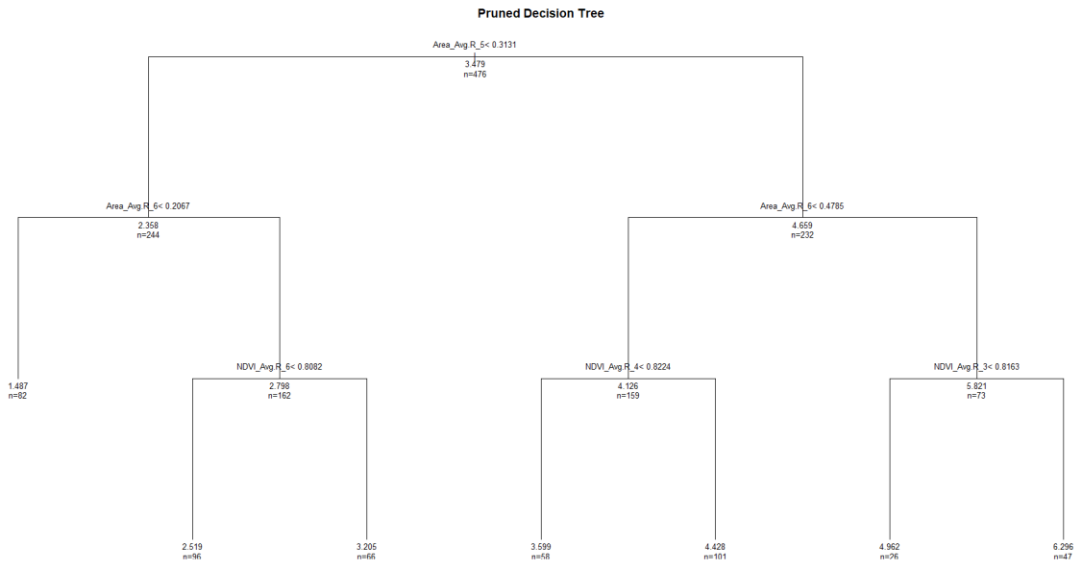

Figure 47 A pruned decision tree used for regression analysis. In this tree, the decision nodes are based on the variable canopy area and NDVI which are the same variable we discussed from the random forest model.

The pruned decision tree (**Figure 47**) is a regression model used to predict outcomes based on canopy area and NDVI, which align with the important variables identified in the Random Forest model. The nodes of this tree split the data into branches based on thresholds of the canopy area and NDVI. The result from the decision tree confirmed the result form the RF model that predicted values are used to estimate the yield based on the canopy area and NDVI.

2021

Following the correlation analysis for 2021 data, crop parameters measured during tuber filling stage (UAV flight 2) has the highest correlation with yield, with canopy has the greatest influence, followed by canopy volume from the same flight and canopy area from the maturation stage (flight 4) (**Figure 43**).

Table 15 Statistical results of fit linear regression model with difference predictors and their evaluation metrics (R-squared, P-value, MAE, MSE, and RMSE) for 2021 potato trial.

Training	Predictor	Model fit	P-value	MAE	MSE	RMSE	Fitting equation

80%	Canopy area	R-square = 0.374	p<0.001	1.233	2.53	0.804	Y=1.11+12.644* Canopy area
80%	Canopy volume	R-square = 0.315	P<0.001	1.42	3.12	0.804	Y=0.368+6.909* Canopy volume
80%	Canopy area + Canopy volume	R-square = 0.402	p<0.001	1.244	2.574	0.804	Y=0.4272+ (3.36730*Canopy area) -(5.361*Canopy volume)

The linear regression analysis (**Table 15**) revealed that models incorporating canopy area and canopy volume separately provided moderate fits to the yield data, with R^2 values of 0.374 and 0.315 respectively, both achieving statistical significance. Combining these two predictors into a single model improved the fit slightly, achieving an R^2 of 0.402.

Random forest

Table 16 The performance of random forest model applied with difference set of predictors and their evaluation metrics (R-squared, MAE, MSE, RMSE, and % variance explained) for 2021 potato trial

Variable	R-squared	MAE	MSE	RMSE	% Var explained
Canopy area	0.82	0.659	0.722	0.849	21.47
Canopy volume	0.722	0.831	1.139	1.067	19.74
Canopy area + Canopy volume	0.84	0.616	0.635	0.78	30.74
Canopy area + Canopy volume + Maximum height	0.84	0.624	0.658	0.811	31.69

Canopy area + Canopy volume + average height + Maximum height	0.841	0.618	0.647	0.804	45.52

The Random Forest models showed the predictive strength of canopy area, which on its own accounted for 21.47% of the yield variance with an R^2 of 0.82 (**Table 16**). However, when canopy area was combined with canopy volume, the model's performance was further enhanced, explaining 30.74% of the variance. The addition of maximum height as a predictor increased the variance explained marginally to 31.69%. The most comprehensive model included canopy area, canopy volume, average height, and maximum height, achieving the highest predictive accuracy with an R^2 of 0.841 and explaining 45.52% of the yield variance.

2022

Following the correlation test (**Figure 44**), UAV flight 3 (tuber filling stage) has the highest correlation with yield follow by UAV flight 4 (maturation stage). From crop parameters, canopy volume extracted from UAV flight 3 has the highest correlation with yield (r = 0.808), follow by canopy volume from UAV flight 2 (tuber initiation stage), NDRE from UAV flight 3 (r = 0.76) and maximum height from UAV flight 2 (r = 0.69). The results from linear regression models offer an insight into the factors affecting fresh yield in potato. A strong positive relationship among plant height, NDVI canopy area and yield, as indicated by the significant p-value and the positive coefficient (model fit).

Table 17 Statistical results of fit linear regression model with difference predictors and their evaluation metrics (Correlation Coefficient, R-squared, RMSE, MAE) for 2022 potato trial.

Training	Predictor	Model fit	P-value	MAE	MSE	RMSE	Fitting equation
80%	Canopy volume	R-square = 0.57	p<0.001	0.898	1.273	1.128	Y=1.11+12.644* Canopy volume

80%	Maximum height	R-square = 0.408	p<0.001	1.067	1.87	1.368	Y=0.043+9.975* Maximum height
80%	NDVI	R-square= 0.646	p<0.001	0.8	1.15	1.072	Y=-23.253+32.723* NDVI
80%	NDRE	R-square = 0.572	p<0.001	0.9	1.415	1.189	Y=-10.105+33.211* NDRE
80%	NDVI + NDRE	R-square = 0.66	p<0.001	0.775	1.096	1.047	Y=-20.62 + (24.26*NDVI) + (10.531*NDRE)
80%	Canopy volume + NDVI	R square = 0.736	p<0.001	0.683	0.793	0.89	Y=-15.707+ (6.658 *Canopy volume) + (22.053*NDVI)
80%	Canopy volume + Maximum height + NDVI	R square = 0.736	p<0.001	0.68	0.8	0.892	Y=-15.707+ (6.887*Canopy volume) + (-0.2908* Maximum height) + (22.137* NDVI)
80%	Canopy + NDVI + NDRE	R square = 0.736	p<0.001	0.68	0.8	0.892	Y=-15.29 + (6.457* Canopy volume) + (20.352 * NDVI) + (2.516 * NDRE)
80%	Canopy volume + Maximum height + NDVI + NDRE	R square = 0.736	p<0.001	0.68	0.8	0.891	Y=-15.28 + (6.76*Canopy area) + (-0.4*Maximum height) + (20.334*NDVI) + (2.715* NDRE)

The statistical results for the linear regression models with different predictor for the year 2022 (**Table 17**) indicate how each model fits the yield data. The model that included both canopy volume and NDVI shows the best fit for yield prediction with an R^2 of 0.736. Adding additional predictors such as maximum height and NDRE did not significantly improve the model, although there was a minor improvement in RMSE, suggesting a slight increase in the precision of the yield predictions.

Random forest model

Table 18 The performance of random forest model applied with difference set of predictors and their evaluation metrics (R-squared, MAE, MSE, RMSE, and % variance explained) for 2022 potato trial.

Variable	R-squared	MAE	MSE	RMSE	% Var explained
Canopy volume	0.835	0.561	0.537	0.733	44.23
NDRE	0.873	0.505	0.412	0.642	44.22
NDVI	0.905	0.431	0.308	0.555	62.67
NDVI+NDRE	0.928	0.374	0.236	0.486	68.64
Canopy volume + NDVI	0.942	0.331	0.189	0.434	73.79
Canopy volume + NDVI+NDRE	0.946	0.315	0.176	0.419	75.27
Canopy volume + Average height + NDVI+NDRE	0.943	0.324	0.184	0.429	76
Canopy volume + Maximum height + NDVI+NDRE	0.946	0.31	0.177	0.421	76.02

Among the Random Forest models for predicting potato yield in 2022, the R^2 value indicate the proportion of variance in yield explained by the models, ranged from 0.835 for canopy volume alone to 0.946 when combining canopy volume with NDVI, NDRE, and

maximum height (**Table 18**). The addition of NDVI and NDRE to canopy volume progressively increased the R-squared values, with NDVI alone achieving 0.905 and both NDVI and NDRE together reaching 0.928. This illustrates the importance of spectral data from NDVI and NDRE in understanding yield variations. The combination of canopy volume, maximum plant height, NDVI, and NDRE variables demonstrated the best performance model achieved R^2 value of 0.946 and its ability to explain 76.02% of the yield variance, indicating a robust predictive capability.

Decision Tree model

Based on the result from the Random Forest model, a decision tree model was used to calculate with variables that best explain the variability in the yield.

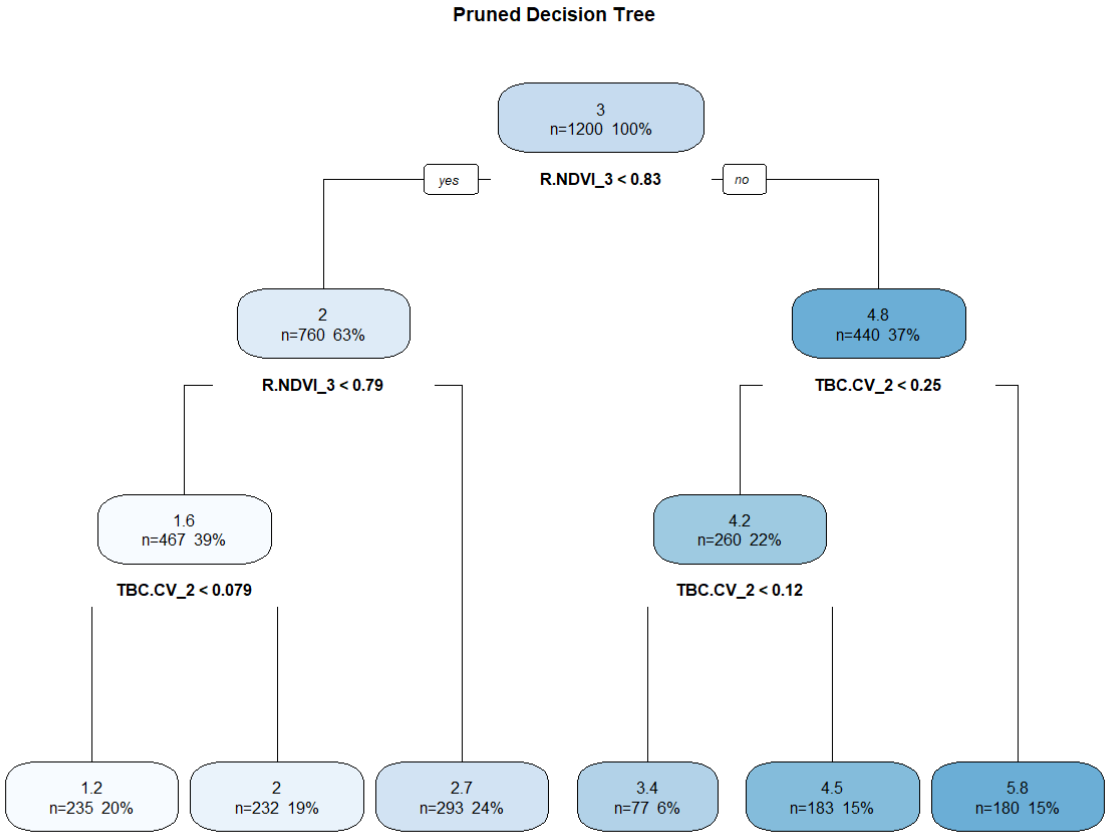

Figure 48 A pruned decision tree used for regression analysis. In this tree, the decision nodes are based on the variable canopy volume and NDVI which are the same variable we discussed from the random forest model.

From the decision tree model (**Figure 48**) through the tree and explain a prediction by the contributions added at each decision node. The accuracy of all models varied significantly

with change in split ratio of the dataset into training and testing data. Each internode represents a test on an attributed and each branch represents the outcome of the test. Each leaf node represents a class label and decision taken after computing all attributes. The tree diagram indicates the flow of decisions that leads to different outcome. The tree starts with a root node splitting on the NDVI extracted from UAV flight 3 (tuber filling stage) at a threshold of 0.83, then further splits into other nodes based on additional thresholds of CV (canopy volume) extracted from UAV flight 2 (tuber initiation stage). This pruning strategy has led to a tree structure that is concise and likely more generalizable to unseen data.

OVERALL

The use of UAVs at different time points and growth stages of potato crops allows for the analysis of vegetation health over time and its relationship to yield. As the crop approaches harvest or the tuber bulking stage (the end of vegetative growth), vegetation indices, plant height, canopy area, and canopy volume tend to decrease compared to the mid-growth stage, making them fewer effective predictors. This observation aligns with our correlation analysis, which flight that captured the potato plant during tuber filling stage give the highest correlation with yield.

Linear models often struggle to achieve high prediction accuracy as the linear models are quantify effect of each variable but they are unable to capture the complex, non-linear interactions with the environment (Meyer and Neto, 2008). Research has focused on identifying optimal conditions for potato crops to maximize yield. The Random Forest (RF) algorithm, known for its analytical and predictive capabilities, achieved an accuracy of 99.74% in one study (Guo *et al.*, 2021). Another study by Bussan (2009) reported the accuracy around 96, 76, 75, and 85 percent for wheat, maize, and potato predictions, respectively, using Random Forest classification.

This observation aligns with our yield prediction model, which showed a moderate R^2 value when performed linear regression model, with the best performed model only achieved the highest R^2 of 0.736, 0.402, and 0.736, for 2020, 2021, and 2022 respectively. Compare to the RF model achieved R^2 value of 0.95, 0.946, and 0.841, for 2020, 2021 and 2022, respectively. The results from RF for yield prediction indicates a robust predictive capability. Despite this success, the model requires further refinement due to data variability from field conditions.

Many studies have confirmed that crop parameters are closely interrelated and dependent on growth stages. The growth and yield of individual stems largely depend on tuber size. During our experiment, we selected seeds of a uniform size to ensure the gap between row, which facilitate easy assessment with machinery. Which conflicts with the suggestion that large seed are necessary for the highest yields. Although the number of sprouts and stems per tuber is asymptotically related to tuber size over the seed treatment used.

A consideration of incident radiation and leaf area would suggest that because tuber form a constant fraction of total weight, bulking rates would increase to a maximum and then gradually decrease as the amount of light intercepted radiation increased and then decreased (Guo *et al.*, 2021). A study by Ji *et al.* (2022) found that using 2D-RGB images could accurately estimate maximum height, correlating strongly with yield (R^2= 0.99). For yield estimation, Support Vector Machines (SVM) performed best (R^2= 0.72), followed by Random Forest and Decision Tree. Li *et al.* (2020) found that random forest regression model demonstrated high prediction accuracy for crop yield using vegetation indices and crop height (R^2>0.90) with the imagery data obtained 90 days after planting. Whereas in this study, we found then using canopy height, canopy area with vegetation index (NDVI) is best to use in our prediction model (R^2 of 0.95) using imagery obtained 98 days after planting (tuber filling stage). Comparatively, Zhou *et al.* (2017) confirmed that the most effective for predicting grain yield in wheat using NDVI (R^2= 0.75). Salvador *et al.* (2020) predicted potato tuber yield using random forest model trained with NDVI (R^2=0.76 and RMSE = 0.149), finding that images collected at earlier growth stages such as tuber initiation, vegetative growth, tuber bulking has better prediction performances that those at the later stage when close to the end of season (Tanabe *et al.*, 2019; Li *et al.*, 2020; Li *et al.*, 2021). Sun *et al.* (2020) reported a positive prediction with an R^2 of 0.63 when using hyperspectral image features at tuber maturation stage. A more accurate estimation of crop parameters derived from UAV could be benefit to yield prediction model (Hu *et al.*, 2018). This research highlights the potential and challenges of using UAV imagery and machine learning for crop yield prediction, indicating the importance of selecting appropriate growth stages and parameters for analysis.

4.3.4 LIMITATION OF THE STUDY AND FUTURE RESEARCH DIRECTION

The complexity of integrating high-dimensional UAV imagery data under field conditions poses a challenge in data processing and analysis. Developing robust analytical method will be crucial for extracting more accurate and precise data. While this study focuses

on a large set of potato genotypes, in order to ensure that the finding is applicable across different potato varieties and growing conditions it is essential for the broader applicability of the result from this chapter. Some crops, such as cereals, have more uniform and vertically structure canopies,

The most significant contribution of the new prediction model is its capability to produce accurate prediction and explainable insights simultaneously. This was achieved through training algorithm to select features and interactions that are spatially and temporally robust to balance prediction accuracy for the training data and generalisability of the test data (Kooman *et al.*, 1996). For efficient crop production management, crop yield prediction models should incorporate with remote sensing imagery, soil information, weather data. Weather conditions are a key factor phenological growth and development and affecting regulating the spread of pest and diseases of potato crops. Potato that emerge quickly are able to capture sunlight at the beginning of the season, which is crucial for maximising the amount and quality of the potato tuber (Pavek and Thornton, 2009). However, the complexity of combining data with weather information across the different year presents challenges in this chapter, leading the use of advanced models for improved accuracy. This results in difficulties in creating direct non-linear or linear mappings between raw data and crop yield values (Elavarasan and Vincent, 2020).

Given the complexity of agricultural systems, more advanced models, including regression trees or conception neural networks (CNN), may offer better accuracy as they can capture non-linear relationships and interactions between predictors. However, for simplicity and interpretability, multiple linear regression is a good starting point, especially if the relationships between the parameters and yield are linear. It is crucial to select the right statistical model based on the specific environmental factor and scale of analysis. Future research, the widespread adoption of UAV technology and 3D modelling in agriculture will depends on reducing costs, improving a more user-friendly interface, and demonstrating clear benefits to farmers and breeders.

4.4 CONCLUSION

The correlation tests and subsequent predictive models established a clear linkage between specific crop parameters, such as canopy area, average height, and vegetation indices with fresh yield of potatoes. UAV flights that corresponding to the tuber filling stage, emerged as critical periods where UAV imagery data showed high correlation with yield outcomes. This

finding suggests that monitoring crop parameters during specific growth stages can provide actionable insights for optimizing harvest times and improving yield predictions. While linear regression models confirmed the positive correlation between plant height and yield, indicating that taller plants usually yield more. However, the lack of significance in NDVI's effect indicates the complexity of factors influencing yield and the need for a multifaceted approach to yield prediction. The Random Forest model appears to be the most effective and outperformed linear regression method, with a notable accuracy of up 76.02% for yield prediction and 53.4% for maturity prediction.

In conclusion, this comprehensive investigation into yield estimation and potato maturity estimation through UAV-remote sensing and machine learning methodologies illuminates the transformative potential of technology in agriculture. The findings advocate for a multidisciplinary approach, combining traditional agricultural knowledge with cutting-edge technology and data analysis techniques, to achieve more sustainable and productive farming practices. Future research should aim to refine these models, explore additional predictors of yield and maturity, and validate the methodologies across different environmental conditions and crop varieties to universalize their application. By continuing to bridge the gap between technology and traditional farming, we can look forward to achieving higher precision in crop management, leading to enhanced food security and agricultural sustainability.

CHAPTER 5. POTATO BREEDING AND DISEASE MONITORING: INTEGRATING UAV IMAGERY AND MACHINE LEARNING

This chapter delves into the intricate relationship between environmental factors and potato yield across diverse varieties over three consecutive growing seasons (2020-2022). By harnessing vegetation indices (VIs) extracted from multispectral UAV imagery alongside traditional crop monitoring methods, we aim to predict disease incidence and forecast yield in organic potato (Solanum tuberosum L.) under field conditions. Preliminary findings from a Random Forest classification model underscored the significance of growth stage-specific monitoring, pinpointing the tuber filling stage as particularly crucial for disease management based on vegetation indices data. Subsequent analyses delved into the impacts of temperature and relative humidity (RH) during the growing season, revealing nuanced, variety-specific responses to these environmental factors.

Our results unveil a complex, multifactored, and non-linear relationship between weather variables and potato yields, shedding light on the dynamic influence of environmental changes on the growth patterns of different potato genotypes. Notably, potato varieties exhibiting high yields have displayed an adaptive response to elevated temperatures, while those with lower yields have thrived under cooler conditions.

In essence, our integrated approach, combining UAV-based imagery with traditional screening methods, lays the groundwork for a streamlined breeding strategy. By swiftly identifying varieties with desired traits, we aim to expedite the breeding process in a cost-effective manner, paving the way for enhanced potato cultivation practices.

5.1 INTRODUCTION

Potato breeding has evolved over the past 150 years with the goal of developing cultivars that meet the changing demands of consumers and the potato industry. Early potato breeding efforts focused on increasing yield and disease resistance, however, as consumer preferences and market demands changed, breeders began to focus on other traits such as earlier maturity and compact tubers. A collection of potato phenotypes is forming the basis of a core collection. By assessing variety characteristics in field trials, researchers can identify varieties that perform well, resilience, high yields, and good quality traits. The increased yield is attributed to enhance agronomy and crop management skill that have evolved alongside the

growing concentration and specialisation among UK farmers (McGregor, 2007). There are shifting trends in potato consumption at home in the UK, the variety or type of potato purchased depends on the meal occasion (McGregor, 2007). This information is vital for breeding programmes, helping make decision in variety selection and developing cultivation practices that enhance productivity and sustainability while meeting the market demands.

While breeding and selection of variety can be done without on-depth genetic information, understanding and analysing of the target breeding traits helps enhance and accelerate breeding processes (Slater *et al.*, 2017). Morphological and phenological traits: these traits are fundamental indicators of potato growth and development. This included canopy height, canopy cover, number of days to flowering, vegetation indices, and maturity. Tolerance or resistance to biotic and abiotic stresses: this assessment focuses on the crop health ensuing the selection of resilient varieties. This included fungal pathogens: including *Phytophthora infestans* (casing late blight) and *Alternaria* spp., Viruses: including potato virus Y (PVY) and potato leafroll virus (PLRV), and environmental stresses: including drought and temperature extreme. Yield and yield component parameters: yield assessments provide quantitative measurement of potato variety performance and capacity. These traits have become central to understanding ecological and evolutionary patterns across various scales, driving a need for standardized methods to measure them accurately (Perez-Harguindeguy *et al.*, 2016). This research is crucial for building predictive models of plant-environment and vegetation-atmosphere dynamics. Recognising the importance of these issues, there is an urgent call for improved and more extensive data collection, underlining the value of standardized protocols for trait measurement.

A study on potato yield variation and tuber in the UK by Taylor *et al.* (2018) revealed that yield, tuber size and stem density are more structured than seed production field where it showed more variability in yield attributes. According to the study by AHDB (2024a) on climate change impacts on UK potato production, the result showed that with the climate change impact could end up in earlier planting and harvest dates, and there is a called for a change to better adapted varieties that less dependence on soil with low water holding capacities, crop movement to regions with suitable agroclimatic and water availability. This has also been confirmed by the study from Garcia-Gonzalez *et al.* (2022) planting date effects were likely driven by humid, warm weather later in the season that was conductive to disease and detrimental to crop development. This study also indicated that early planting date and variety selection are an effective management approach to reduce to disease incident and

maximise tuber yield. With the aims of the project Ecobreed is increasing the efficiency and competitiveness of organic crop breeding (Ecobreed, 2020).

Potatoes are prone to several diseases brought on by viruses, fungus, bacteria, and other pathogens. Among these, early and late blight diseases are particularly common. Early blight, caused by two species of the *Alternaria* genus (*Alternaria solani and Alternaria alternata*), manifests as lesions on the leaves, often exhibiting a target spot appearance with concentric rings (Tsedaley, 2014; AHDB, 2024b). This disease can lead to yield losses in potato crops ranging from 20 percent to 50 percent (Horsfield *et al.*, 2010). On the other hand, late blight, primarily caused by the pathogen *Phytophthora infestans*, is the most destructive disease affecting potato crops by infecting foliage, stems, and tubers which spreads rapidly through the air (Scheufele, 2022; AHDB, 2024b). Infestation of late blight can result in yield losses in potato crop ranging from 30 to 100 percent (Hirut *et al.*, 2017; Namugga *et al.*, 2018).

The existing techniques for identifying these diseases involve field-based approaches that rely on manual inspection of vast fields which are time consuming, demands significant human effort, and outcomes may vary subjectively (Shi *et al.*, 2022). Hence, these factors contribute to challenges in managing potato disease through the utilization of field-based approaches. The advancement of high-throughput precise phenotyping platforms is offering a new and fascinating tool for precise screening (Obidiegwu *et al.*, 2015). By integrating innovative screening techniques, it can assist in selecting variety that suitable for particular settings. Nevertheless, employing remote sensing techniques using inexpensive sensors attached to UAVs offers an alternative method for swiftly and reliably monitoring and managing potato diseases. Early detection of the potato diseases enables minimal fungicide application for efficient disease control, thereby substantially mitigating the environmental repercussions associated with excessive pesticide usage.

Hence, Van De Vijver *et al.* (2022) used ultra-high-resolution images from UAV to detect early blight disease on potato leaves by utilising deep learning models. In addition, Duarte-Carvajalino *et al.* (2018) used machine learning methods with multispectral UAV images to evaluate late blight in 14 potato genotypes. Their results suggest machine learning (ML) algorithms could replace visual late blight severity estimation. Further, Franceschini *et al.* (2019) and Rodríguez *et al.* (2021) assessed the viability of employing high resolution UAV multispectral imagery for early detection and severity evaluation of late blight in potato crops. Additionally, numerous other studies have explored the potential use of hyperspectral UAV

imagery with various machine learning models such as support vector machine (SVM), PLS-DA, and Random Forest (RF). These studies investigated spectral and spatial characteristics at both canopy and leaf levels to identify late blight disease in potato crops, achieving classification accuracies of over 80 percent by the machine learning models (Patil *et al.*, 2017; Gao *et al.*, 2021; Shi *et al.*, 2022).

While machine learning models trained on UAV hyperspectral imagery data yield impressive accuracy, they are hindered by computationally demanding datasets and lengthy training periods. As a result, their implementation for routine monitoring aimed at detecting potato diseases becomes ineffective and time-consuming. Another cost-effective option involves utilizing multispectral UAV data. Limited research has been undertaken to thoroughly explore the effectiveness of using multispectral UAV data for detecting the most suitable growth stage to monitor and manage potato diseases. Additionally, its potential for monitoring disease progression over time with data collected over a timeframe has not been conducted. Achieving this entails leveraging spectral features, by constructing vegetation indices from multispectral UAV imagery, to ascertain whether these indices can serve as an indirect means of detecting potato disease. It is hypothesized that healthy crops in the canopy would yield higher spectral index values from vegetation indices, while disease-infected areas would exhibit the opposite trend. In addition, these spectral characteristics can be combined with k-means clustering and random forest classification models to assess the ability to accurately group disease-prone potato crops based on their spectral features.

Hence, the aim of this chapter is to (1) To understand the most suitable classification approach to accurately group disease-prone potato crops based on multiple vegetation indices derived from multispectral UAV (2) To understand the most significant growth stage of the potato to monitor the disease (3) To understand whether the classification method demonstrates accuracy when applied across various years.

5.2 STUDY AREA AND PLANT MATERIAL

The field experiment was performed at Nafferton Farm, Newcastle University, the United Kingdom as part of the multi-environment organic potato testing trials called the Ecobreed project, where field trials of a wide range of potato and wheat varieties are on-going (Ecobreed, 2020; Meglic *et al.*, 2020). Ecobreed project tests numerous potato varieties to identify those that require low inputs, have higher yield potential, and are well suited to the region to increase the competitiveness of organic sector (Ecobreed, 2020)

Annual experiments were arranged in the randomised complete block (RCB) designs with three blocks (replications). 64 potato varieties were selected from across Europe by Ecobreed committee including traditional cultivars, breeding or research materials, commercial cultivars, and improved cultivars (listed in **Table 19**). All 64 verities were evaluated for their suitability for breeding organic potato varieties across four counties including Hungary, Poland, Slovenia, and the United Kingdom.

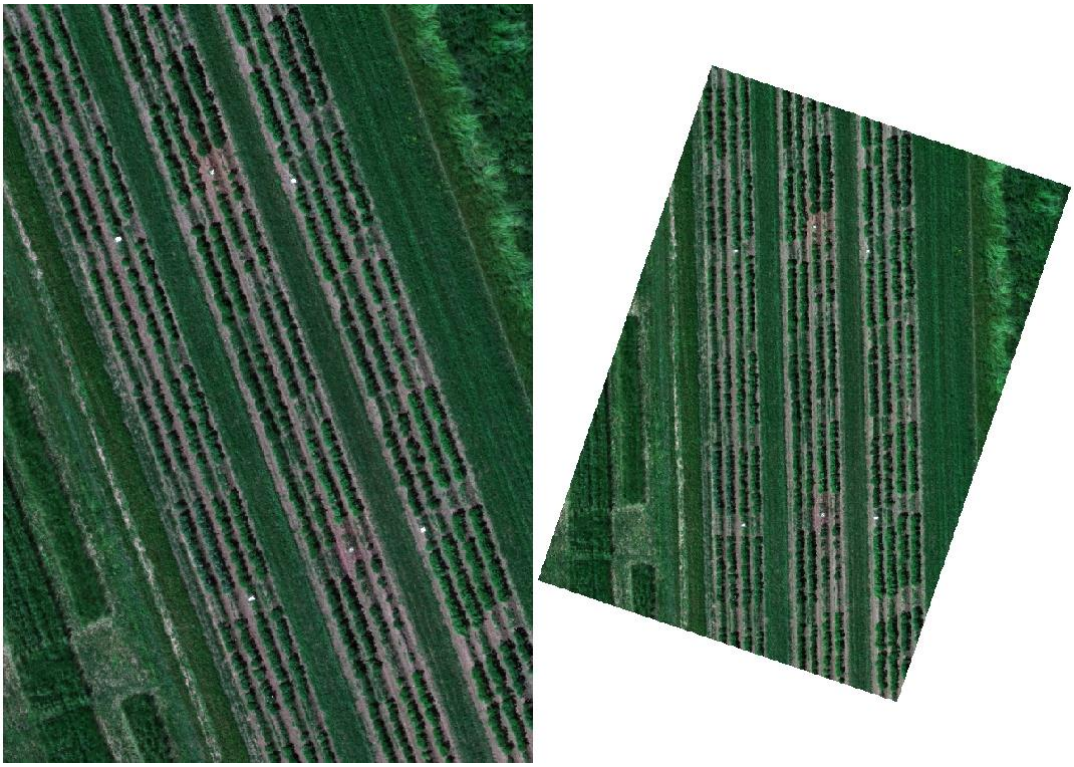

Figure 49 Field layout, each block consisting of 64 plots each growing a different potato variety. This layout was repeated three times.

The experimental design for the Ecobreed study different slightly from the main field trial if this study. The experimental design consisted of three replicate blocks, each block containing four rows with 16 plots per row. Each plot consisted of two row, each 4.5 meters (m) long and was planted with 15 tuber seeds of a given variety. The trial followed a solid row configuration, with a row spacing of 90 cm while the distance between plot was 1.2 metres (**Figure 49**).

Table 19 Potato variety list used in Ecobreed project (Ecobreed, 2020).

VARIETY NAME (ACCENAME)	COUNTRY OF ORIGIN (ORIGCTY)	ORIGINAL BREEDER (BREDNAME)	BREEDING INSTITUTE CODE (BREDCODE)	YEAR OF RELEASE	Maturity
AGRIA	DEU	EUROPLANT Pflanzenzucht GmbH	DEU383	1985	LATE
ALOUETTE	NLD	Agrico	NLD022	2014	MEDIUM LATE
AMBO	IRL	IPM Potato Group Limited	N/A	1993	MEDIUM LATE
ANUSCHKA	DEU	EUROPLANT Pflanzenzucht GmbH	DEU384	2004	EARLY
BALATON ROSZA	HUN	MATE (Hungarian University of Agriculture and Life Sciences) /Pannon University of Agriculture	HUN007		EARLY
BASA	HUN	MATE (Hungarian University of Agriculture and Life Sciences) /Pannon University of Agriculture	HUN007		LATE
BELANA	DEU	EUROPLANT Pflanzenzucht GmbH	DEU385	2000	EARLY
BELMONDA	DEU	Solana Deutschland GmbH & Co. KG	N/A	2010	MEDIUM LATE
BIONTA	AUT	NÖ. Saatbaugenossenschaft	N/A	1993	LATE
BOTOND	HUN	MATE (Hungarian University of Agriculture and Life Sciences) /Pannon University of Agriculture	HUN007		EARLY
BZURA	POL	Hodowla Ziemniaka Zamarte	N/A	1983	LATE
CAPRICE	DEU	Norika	N/A	2010	MEDIUM LATE
CAPUCUNE	FRA	Germicopa	N/A	2018	MEDIUM LATE
CARA	IRL	IPM Potato Group Limited	N/A	1973	LATE
CAROLUS	NLD	Agrico	NLD022	2012	LATE

CASABLANCA	GBR	Cygnet PB Ltd.	N/A	2010	EARLY
CHARLOTTE	FRA	Germicopa	N/A	1981	MEDIUM LATE
COLEEN	IRL	IPM Potato Group Limited	N/A		EARLY
COLOMBA	NLD	HZPC	N/A	2011	EARLY
DAMARIS OO1	DEU	Norika	N/A		MEDIUM LATE
DELILA	FRA	Germicopa	N/A	2017	LATE
DENAR	POL	Hodowla Ziemniaka Zamarte	N/A	1999	EARLY
DITTA	AUT	NÖ. Saatbaugenossenschaft	N/A	1989	LATE
EDONY	FRA	Germicopa	N/A	2012	LATE
ELFFE	DEU	EUROPLANT Pflanzenzucht GmbH	DEU386	2003	EARLY
ERIKA	AUT	NÖ. Saatbaugenossenschaft	N/A	2009	EARLY
FIDELIA	DEU	Norika	N/A	2011	MEDIUM LATE
FORTUS	NLD	HZPC	N/A	2015	MEDIUM LATE
GARDENA	POL	Hodowla Ziemniaka Zamarte	N/A		EARLY
GATSBY	GBR	Cygnet PB Ltd.	N/A	2013	MEDIUM LATE
GOLDMARIE NN	DEU	Norika	N/A		MEDIUM LATE
GRANOLA	DEU	Solana Deutschland GmbH & Co. KG	N/A	1975	MEDIUM LATE
KARLENA	DEU	Norika	N/A	1988	EARLY
KELLY	FRA	Germicopa	N/A	2018	LATE
KIS KOKRA	SVN	Agricultural Institute of Slovenia	SVN019	2010	MEDIUM LATE
KIS SAVINJA	SVN	Agricultural Institute of Slovenia	SVN019	2016	EARLY
KIS SLAVNIK	SVN	Agricultural Institute of Slovenia	SVN019	2015	EARLY
KIS VIPAVA	SVN	Agricultural Institute of Slovenia	SVN019	2012	EARLY
LEVANTE	NLD	Agrico	NLD022	2018	EARLY

LILLY	DEU	Solana Deutschland GmbH & Co. KG	N/A		MEDIUM LATE
LORD	POL	Hodowla Ziemniaka Zamarte	N/A	1999	EARLY
MAGNOLIA	POL	Pomorsko Mazurska Hodowla Ziemniaka Sp. z o.o.	N/A		EARLY
MAYAN GOLD	GBR	The James Hutton Institute	GBR048	2001	LATE
MICHALINA	POL	Hodowla Ziemniaka Zamarte	N/A		EARLY
NOBLESSE	NLD	HZPC	N/A	2015	MEDIUM LATE
NOFY	NLD	Agrico	NLD022	2017	EARLY
OMEGA	DEU	EUROPLANT Pflanzenzucht GmbH	DEU387	2004	MEDIUM LATE
OTOLIA	DEU	EUROPLANT Pflanzenzucht GmbH	DEU388	2015	MEDIUM LATE
OWACJA	POL	Pomorsko Mazurska Hodowla Ziemniaka Sp. z o.o.	N/A		EARLY
PREMIERE	NLD	Agrico	NLD022	1979	EARLY
RIVIERA	NLD	Agrico	NLD022	1998	EARLY
SALOME	DEU	Norika	N/A	2001	EARLY
SARPO MIRA	GBR	Sarpo Potatoes Ltd.	N/A	2003	LATE
SARPO SHONA	GBR	Sarpo Potatoes Ltd.	N/A	2010	LATE
TAJFUN	POL	Pomorsko Mazurska Hodowla Ziemniaka Sp. z o.o.	N/A	2004	MEDIUM LATE
TINCA	DEN	DANESPO A/S	N/A	2018	EARLY
TRIPLO	NLD	HZPC	N/A	2000	MEDIUM LATE
TWINNER	NLD	Agrico	NLD022	2016	EARLY
TWISTER	NLD	Agrico	NLD022	2017	MEDIUM LATE
VALOR	GBR	Caledonia Potatoes	N/A	1993	LATE
VOYAGER	NLD	HZPC	N/A	2003	MEDIUM LATE
WEGA	DEU	Norika	N/A	2010	MEDIUM LATE

WHITE LADY	HUN	MATE (Hungarian University of Agriculture and Life Sciences) /Pannon University of Agriculture	HUN007	1994	LATE
YONA	FRA	Germicopa	N/A	2008	LATE
12-LHI-6	DEN	DANESPO A/S	N/A		MEDIUM LATE
KIS BLEGOŠ	SVN	Agricultural Institute of Slovenia	SVN019	2021	EARLY
KIS RAZOR	SVN	Agricultural Institute of Slovenia	SVN019	2019	MEDIUM LATE
KIS TAMAR	SVN	Agricultural Institute of Slovenia	SVN019	2022	LATE
OSPREY	GBR	Dunnett Dr. Jack	N/A	2000	EARLY
INCA BELLA	GBR	The James Hutton Institute	GBR048	2010	LATE

5.3 GROUND TRUTH DATA COLLECTION

The methodology for field evaluation described in this chapter follows the procedures outlined in Chapter 2 (Ground truth and proximal data collection page 42).

5.3.1 CANOPY TRAITS

The following characteristics were assessed in the field trials. Morphological and phenological traits: these traits are fundamental indicators of potato growth and development. This included canopy height, canopy cover, number of days to flowering, vegetation indices, and maturity.

5.3.2 DISEASE ASSESSMENT

For disease evaluation, this assessment focuses on the crop tolerance or resistance of stresses ensuing the selection of resilient varieties. This included fungal pathogens: including *Phytophthora infestans* (casing late blight) (**Table 4**), and *Alternaria* spp. (**Figure 6**) and Viruses: including potato virus Y (PVY) and potato leaf roll virus (PLRV).

5.3.3 FINAL YIELD

Yield assessments provide quantitative measurement of potato variety performance and capacity. Yield data (weight of tubers per row) was averaged for each variety and plot in kilogram (kg).

5.4 AERIAL DATA COLLECTION

5.4.1 UAV FLIGHT PARAMETERS

In this chapter, aerial images were captured using an unmanned aerial vehicle. Two platforms were used: a multi-rotary wing and a fixed-wing UAV. The multi rotary-wing UAV, a multi-rotor DJI M200 and DJI M600, is equipped with a MicaseseRedEdge-M five-band multispectral sensor (MicaSense, Seattle, USA). The multispectral sensor captures images in blue, green, red, NIR, and red-edge bands. The UAV with a remote sensing system was operated at an altitude of 30 metres with 85 percent image overlap. The camera was set to capture images automatically in a given space of one image every second. The corresponding spatial resolution for this data is 1 cm/pixel. The irradiance sensor has filters to detect the

amount of radiation from direct sunlight or diffuse light in cloudy weather conditions. The first UAV flight was performed when over 50 percent of the plots had emerged, approximately 40 days after planting. Fixed-wing UAV images were collected with an RGB camera quipped. Fix-wing surveys were conducted at a 75-metre altitude with 80 percent image overlap. Details of flight dates and their correlation with potato growth stages are shown in **Table 20**.

Table 20 UAV flight detailed for Ecobreed fields (2020, 2021, and 2022)

Planting date	Flight No	Flight date	Potato growth stage
21-04-2020	1	10-08-2020	Tuber filling
	2	20-08-2020	Tuber filling
	3	08-09-2020	Maturation
13-04-2021	1	08-06-2021	Initial/ Establishment
	2	07-07-2021	Tuber filling
	3	10-08-2021	Maturation
12-04-2022	1	30-05-2022	Initial/ Establishment
	2	20-06-2022	Tuber initiation
	3	11-07-2022	Tuber filling
	4	05-08-2022	Maturation

5.5 DATA PROCESSING

We follow the step explained in **chapter 2 Methodology and Data Analysis**. From the multispectral UAV-based imagery of the potato field, extract canopy parameters from the drone-based imagery using R Studio in R language software (Team, 2021). Data extraction steps include soil/background removing, plot boundary identification, calculating vegetation

indices, calculating canopy height and area from DSM and DTM, and exportation of crop parameters from each plot (**Figure 50** Background removal process and plot boundary identification.).

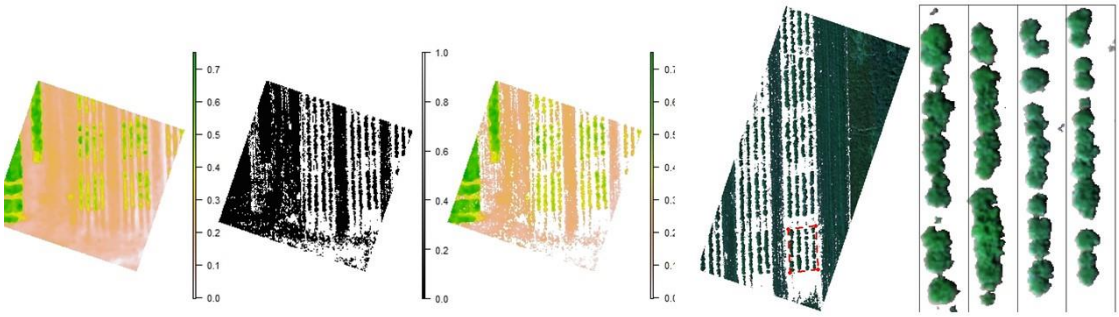

Figure 50 Background removal process and plot boundary identification. The labels on the y-axis represent the pixel height in the image, which was used as the threshold for background removal, as explained in **Chapter 2**.

The vegetation indices were computed and extracted for each plot for further analysis (**Figure 51** Example of vegetation indices extracted from the multispectral UAV-based imagery in R environment). The vegetation indices and their description to related plant traits was explained in **Table 8** Describes the vegetation indices and formular used in this thesis study**.,** Methodology and Data Analysis).

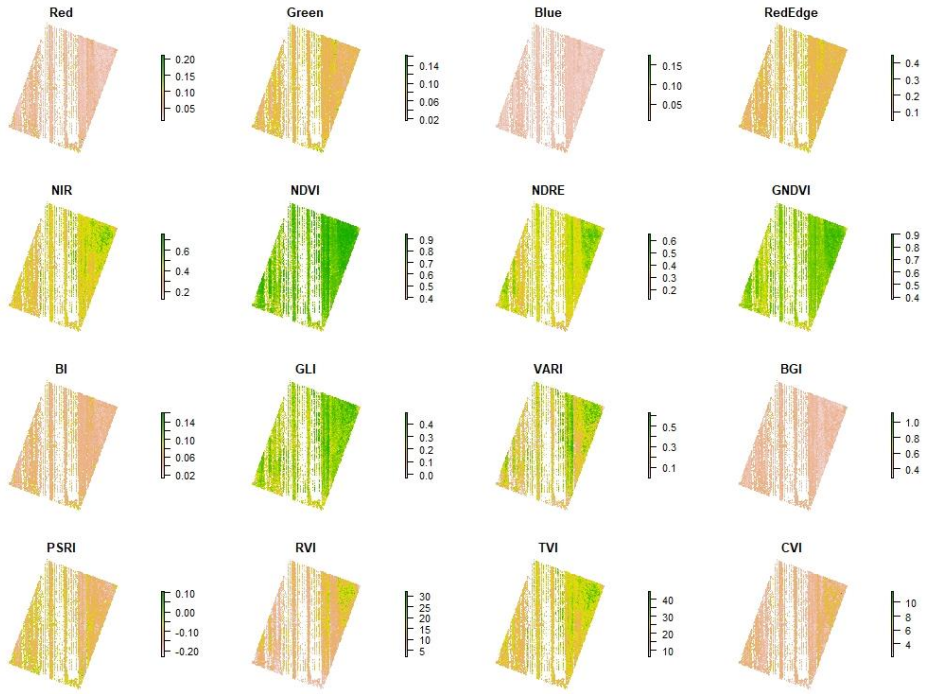

Figure 51 Example of vegetation indices extracted from the multispectral UAV-based imagery in R environment. The map displayed separate maps with unique colour palettes with its own scale bar on the y-axis indicate the value ranges, which differ according on the index.

5.6 DATA ANALYSIS

Data preparation includes aligning the datasets temporally to ensure the weather data matches the time frames of the potato observation. Performing correlation analysis to identify the relationships between weather parameters and potato parameters. For potato crop parameters, we were focused on analysing the height, vegetation indices and disease incident in the field. Next, visualising the relationship between crop parameters and yield. Performed statistical analysis to understand the relationship between different crop parameters. No comparison of total yield per plot was performed due to the differences in yield potential and maturity between varieties. ANOVA and correlation analysis were performed in R Studio, R language software (Team, 2021).

For predictive modelling, Random Forest (RF) algorithms were employed for both regression and classification tasks (Breiman, 2001; Biau, 2012). Regression models aimed at yield prediction, while classification models focused on disease prediction. All data was divided into training and test sets for both yield and disease prediction model. The accuracy of these predictive models was evaluated using a Root mean squared error (RMSE) and

coefficient of determination (R^2). Person's correlation coefficient (R^2) was also considered to assess the linear correlation between the observations and the predictions. Furthermore, K-means clustering were performed to segment the data into *K*-clusters as suggested by the analysis of elbow plot (Huang, 1998).

Conceptual framework for analysis

1. Identify weather conditions for each growth stage: For each potato growth stage in 2020, 2021, and 2022, identify the weather condition in the corresponding month.

2. Correlate weather conditions with potato traits: for each growth stage, correlate the weather conditions (temperature, soil temperature, rainfall, and relative humidity) of the corresponding month with the potato traits. This involved comparing the data directly.
3. Statistical analysis: correlation coefficients and regression analysis were performed to determine the relationship between weather conditions and crop parameters.

5.7 EXPLORE DATA ANALYSIS

Comparing the data extraction method across different models and software and evaluating visual observation against these methodologies in terms of accuracy and precision. Additionally, identifying the method that provides consistent results over multiple years and fields involved calculating metrics that reflect stability over time including the standard deviation of accuracy and precision metrics across years and fields.

This structured approach allows for a thorough comparison at different data extraction methods and field from models or software including visual observation. By focusing on accuracy, precision, and consistency across multiple datasets and years. Thus, we can identify the most reliable method for the potato study. Although, the best method may vary depending on the context, so it is important to consider the specific requirements of each application.

5.8 RESULTS AND DISCUSSION

5.8.1 NUMBER OF PLANT PREDICTION

A semi-automated recognition of plant illustrated in **Figure 52**; the red dot mask the distinctive potato canopies, effectively distinguish individual plants from one another.

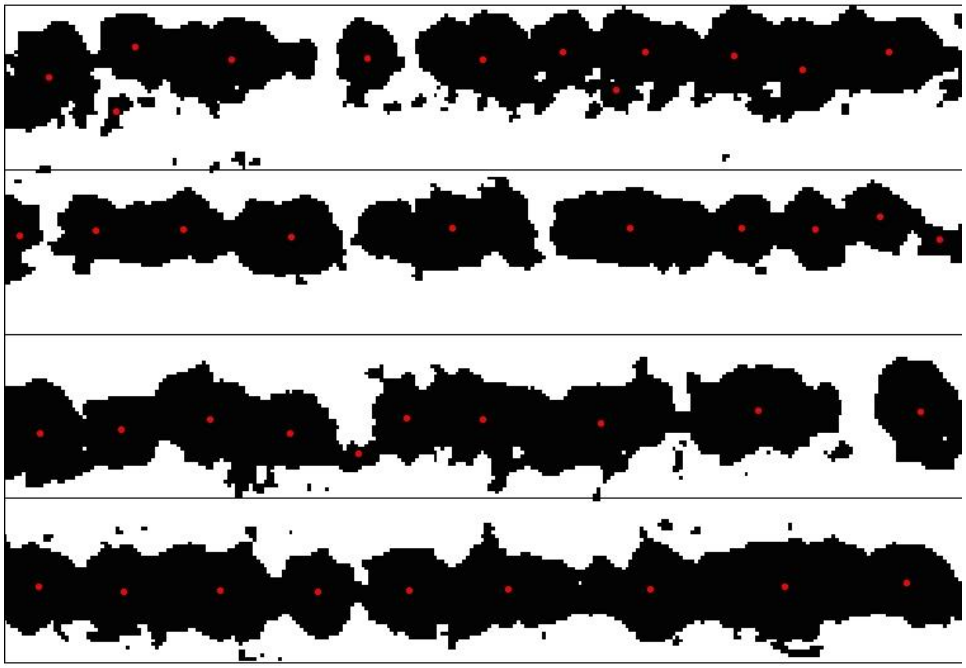

Figure 52 Automated plant count process from the UAV-based imagery.

The overall accuracy of the automated plant counting process utilising UAV-based imagery at 82.88% with a Mean Absolute Error (MAE) of 0.216, for the UAV flight capturing the potato plant at the initial or establishment growth stage. However, this accuracy declined in UAV flight 2 and 3, the accuracy is dropped to 61.26% (MAE = 0.523) and 57.37% (MAE = 0.558), respectively. The precision of the automated plant count tends to be highest when plants are still small. This is because as plant grow bigger, the software struggles and sometimes confusing weeds for potato plants, when there are weed present which impacts its effectiveness.

$$Accuracy = \frac{Number\ of\ correct\ predictions}{Total\ number\ of\ predictions}$$

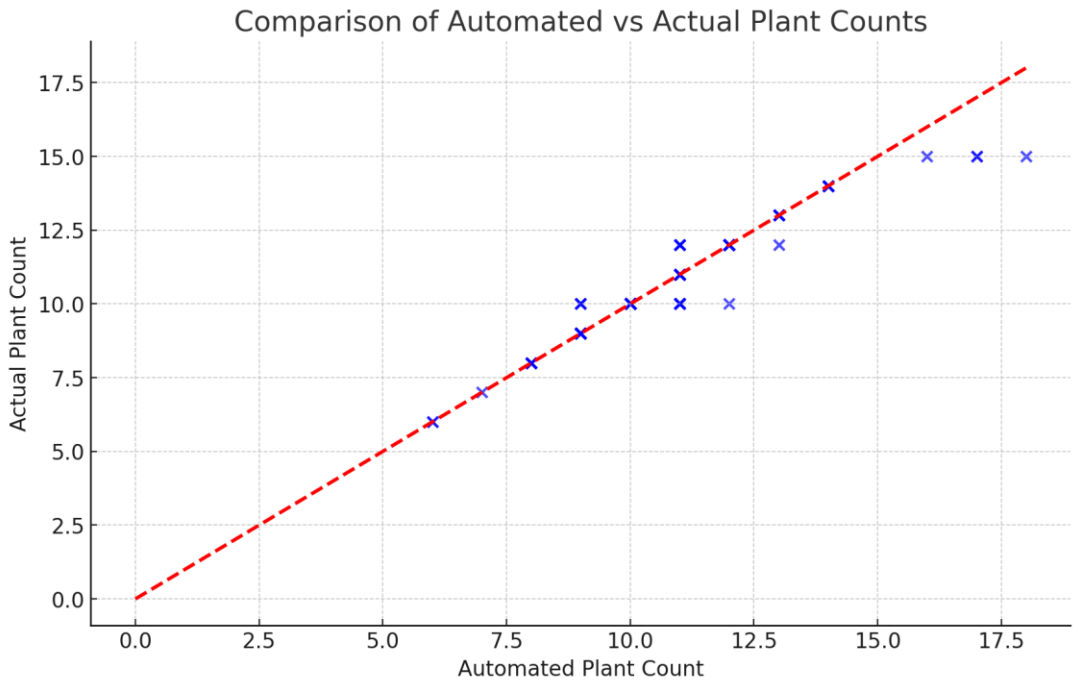

Figure 53 A comparison between automated plant counted actual plant counts from flight 1 (initial stage) y = 0.8658x+1.3659, R^2 = 0.9275.

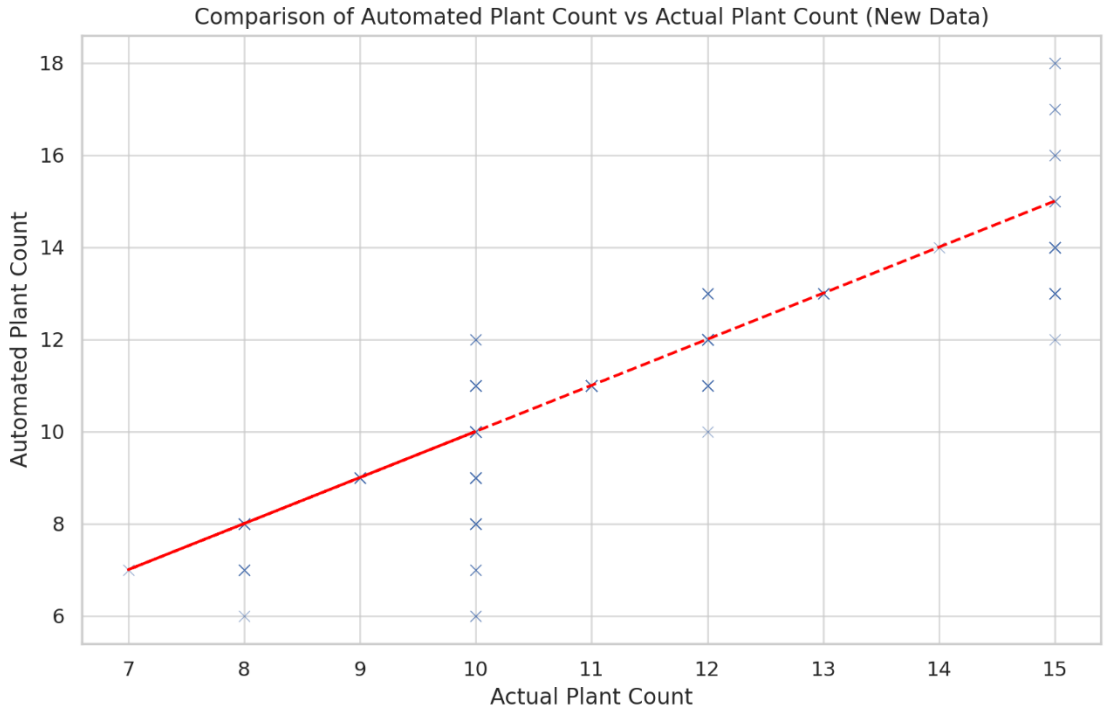

Figure 54 A comparison between automated plant counted actual plant counts from flight 2 (tuber initiation stage) y = 0.8027x+2.402, R^2 = 0.7821.

Within the context of precision agriculture, this component of the study presents a comparison analysis between automated plant counting approaches and traditional field inspector counts conducted during the 2021 trial. The investigation revealed a significant correlation between automated counts and observed plant number per plot, as evidenced by associated r-squared values of 0.9275 and 0.782 were observed (**Figure 53**and **Figure 54**). This suggest that while automated prediction method is reasonably effective, there is room for improvement.

The variation detected in the automated counts in comparison to manual inspections can be related to various factors such as the plant growth stage and the presence of weeds, which may introduce complexities to the automated detection step. For practical application, the prediction model can be used as an initial screening tool. Nonetheless, it is crucial emphasise the importance of calibration and potential integration of manual inspections, especially in organic production where weed can be present with high density (Gallandt, 2014). This would improve precision and reliability, which is particularly crucial in precision agriculture and research setting where exact plant counts are necessary.

A significant result of this research refers to the optimal timing for using unmanned aerial vehicle (UAV)-based plant counting. The study also reveals that the most accurate model performance can be achieved during the initial flight, where plants are relatively small. This observation supports the conclusion established in **Chapter 3 (**Development and Validation for High-throughput Field Phenotyping and Disease Detection of Potatoes Using Multispectral Imagery and Plant 3D-model**)**, which confirm that the initial UAV flight is the most suitable timeframe for precise plant stand count. Following conducting a more comprehensive examination, incorporating data from three years trials, it was found that the model's accuracy decreased from 92.75% to 82.33%. This pattern is consistent with the findings in Chapter 3, where the accuracy decreased from 93.69% to 75.45%. Despite the observed decrease in accuracy when combined three years data, these findings still reflect a significant advancement in the area of automated plant counting, providing accuracies that exceed those documented in previous studies which includes several crops (Jin *et al.*, 2017; Oh *et al.*, 2020; Pathak *et al.*, 2022).

5.8.2 YIELD PERFORMANCE

The bar charts below (**Figure 55**, **Figure 56**, and **Figure 57**) display the average yield of potato varieties, colour coded by yield performance levels: high (green), medium (yellow),

and low (red). Yield performance groups were defined based on quantiles of average yield, dividing yields into “Hight”, “Medium” and “Low” categorises.

- “High” ~ average yield ≥ quantile (average yield, 0.75)
- “Medium” ~ average yield < quantile (average yield, 0.75) & average yield ≥ quantile (average yield, 0.25)
- “Low” ~ average yield ≤ quantile (average yield, 0.25)

This visual distribution highlights the significant differences in yield potential among varieties, providing a clear and comparative view on their performance over three years trials.

Table 21 Annual yield comparison of potato varieties over three years (2020, 2021, and 2022).

Year	Top 10		Least yield	
2020	Variety name	Yield (kg)	Variety name	Yield (kg)
	LEVANTE	75.9	KIS VIPAVA	26.5
	CARA	71.0	MAYAN GOLD	29.4
	TAYFUN	68.5	BELANA	29.8
	GARDENA	66.0	OTOLIA	31.9
	AMBO	63.1	CAPUCINE	32.8
	NOFY	62.7	KIS SALAVNIK	32.9
	TWISTER	62.6	CAPRS	33.0
	OMEGA	60.9	INCA BELLA	35.6
	OSPREY	60.8	DITTA	36.3
	BELMONDA	60.4	KARLINA	36.3
2021	Variety name	Yield (kg)	Variety name	Yield (kg)
	LEVANTE	92.6	12-LHI-8	25.1
	FIDELIA	83.6	MAYAN GOLD	26.4
	VALOR	83.6	12-LHI-6	30.2
	TAYFUN	81.2	12-LHI-7	30.6
	WHITE LADY	80.4	INCA BELLA	38.4
	WEGA	79	BELANA	44.3
	BOTOND	77.1	DITTA	46.4
	DESIREE	76.9	KIS VIPAVA	46.8
	YONA	76.8	SALOME	47.2
	COLLEEN	76.3	TINCA	49.3
2022	Variety name	Yield (kg)	Variety name	Yield (kg)
	LEVANTE	64.3	MAYAN GOLD	10.7
	NOFY	59.2	INCA BELLA	13.0
	TAYFUN	57.3	BASA	23.7
	YONA	56.4	TINCA	27.0
	AGRIA	53.5	KIS VIPAVA	27.4
	VALOR	53.1	DITTA	28.2

	DESIREE	52.4	BELANA	28.5
	CARA	51.1	CAPUCINE	28.8
	LILLY	50.9	SALOME	29.3
	12-LHI-6	49.0	SALAVNIK	29.3

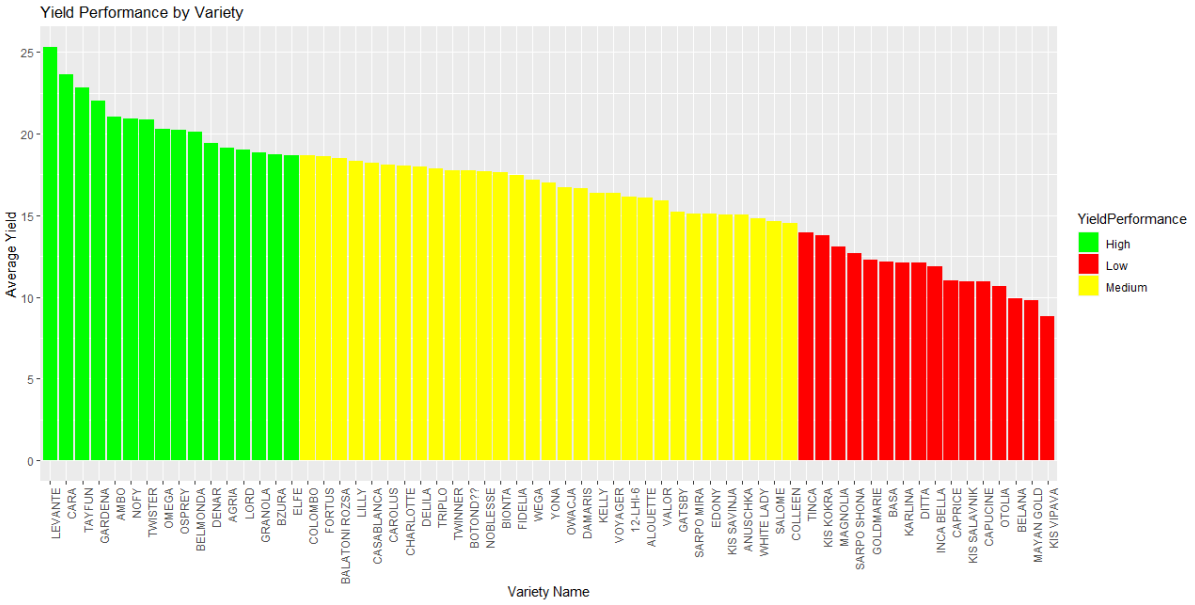

Figure 55 Yield performance in 2020 by variety, with the x-axis representing the variety names and the y-axis indicating the average yield. The bars are colour-coded to categorise the yield performance as high (green), medium (yellow), and low (red).

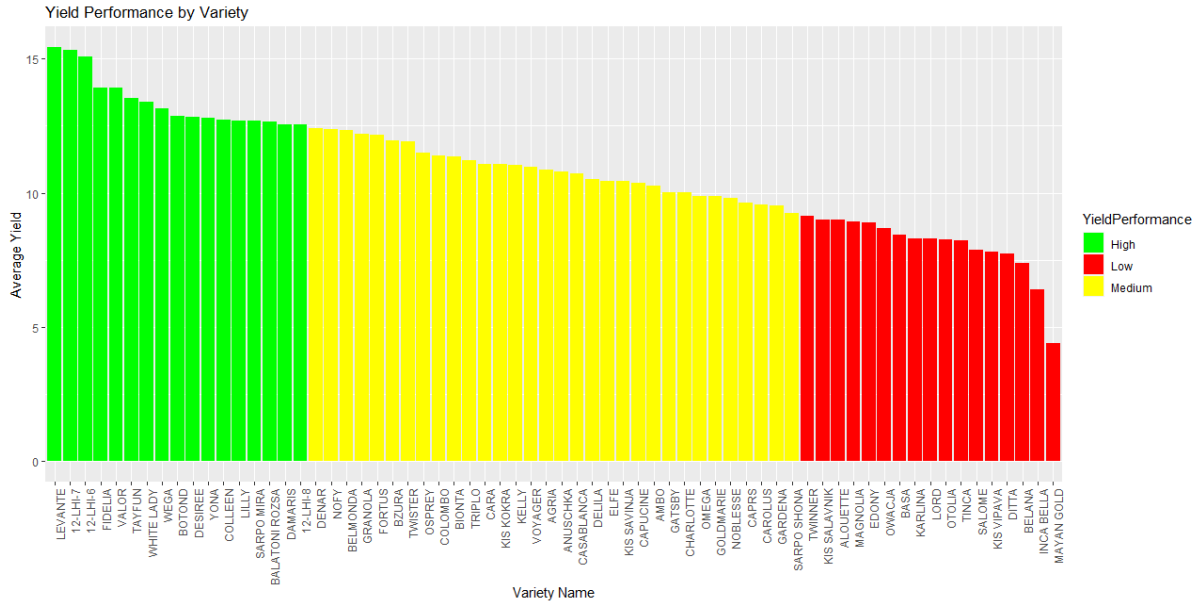

Figure 56 Yield performance in 2021 by variety, with the x-axis representing the variety names and the y-axis indicating the average yield. The bars are colour-coded to categorise the yield performance as high (green), medium (yellow), and low (red).

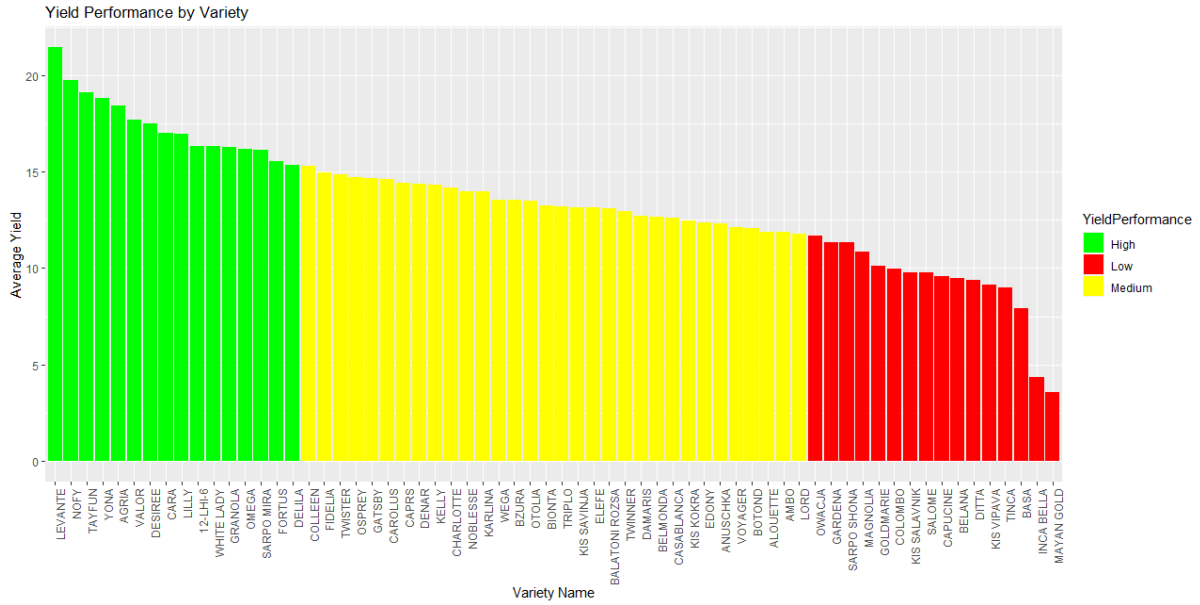

Figure 57 Yield performance in 2022 by variety, with the x-axis representing the variety names and the y-axis indicating the average yield. The bars are colour-coded to categorise the yield performance as high (green), medium (yellow), and low (red).

5.8.3 YIELD PREDICTION MODEL

A Random Forest (RF) regression model was computed with the aim of forecasting potato yield from crop growth parameters. To train this model, 80% of the data set was set as training features with the remaining 20% as testing features. The selection of crop growth parameters as prediction was selected based on the correlation analysis of actual yield data and other canopy parameters. The model was tested 63 varieties of potatoes (Error! Reference s ource not found.). The effectiveness of this model illustrates the potential for using machine leaning approaches in yield prediction in order to enhance yield prediction and inform crop management decision.

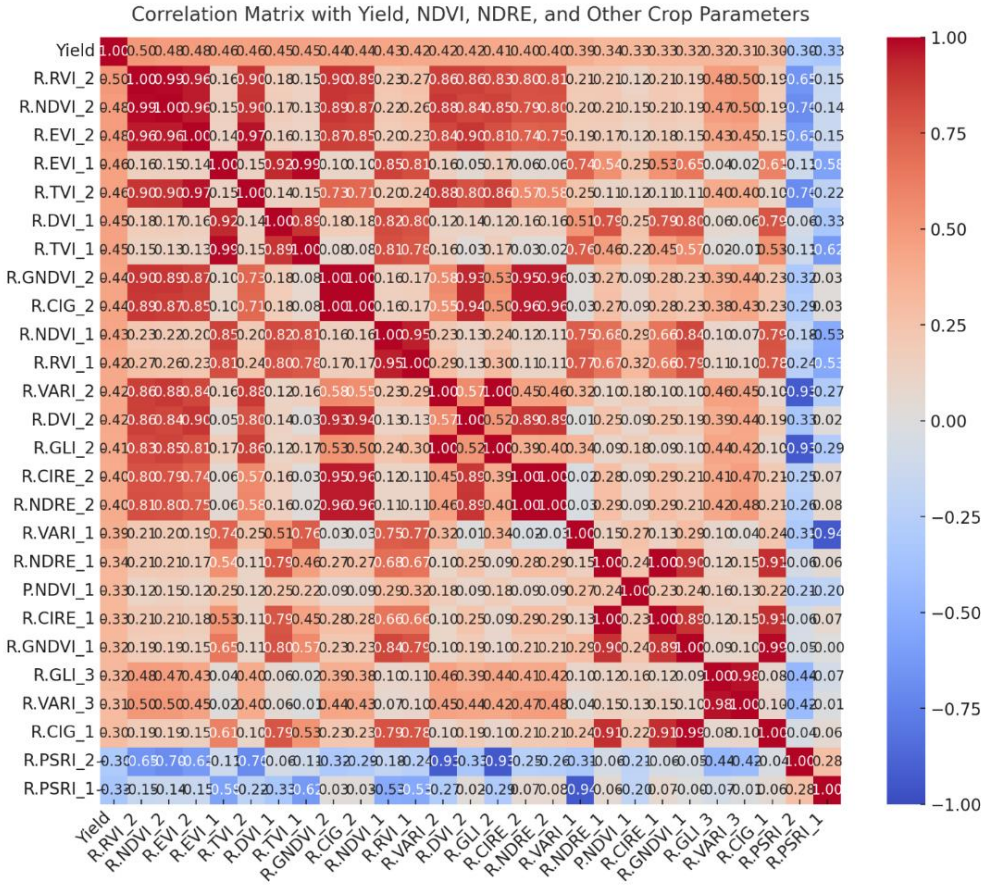

Figure 58 Heat map correlation between crop parameter for Ecobreed field 2020

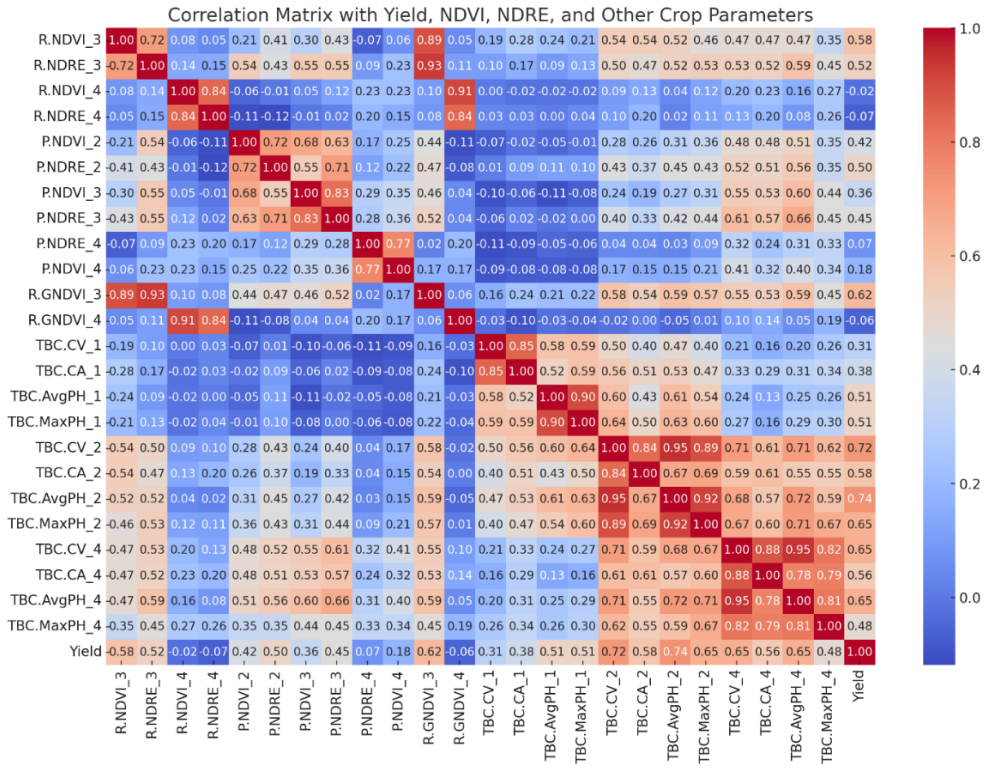

Figure 59 Heat map correlation between crop parameter for Ecobreed field 2021

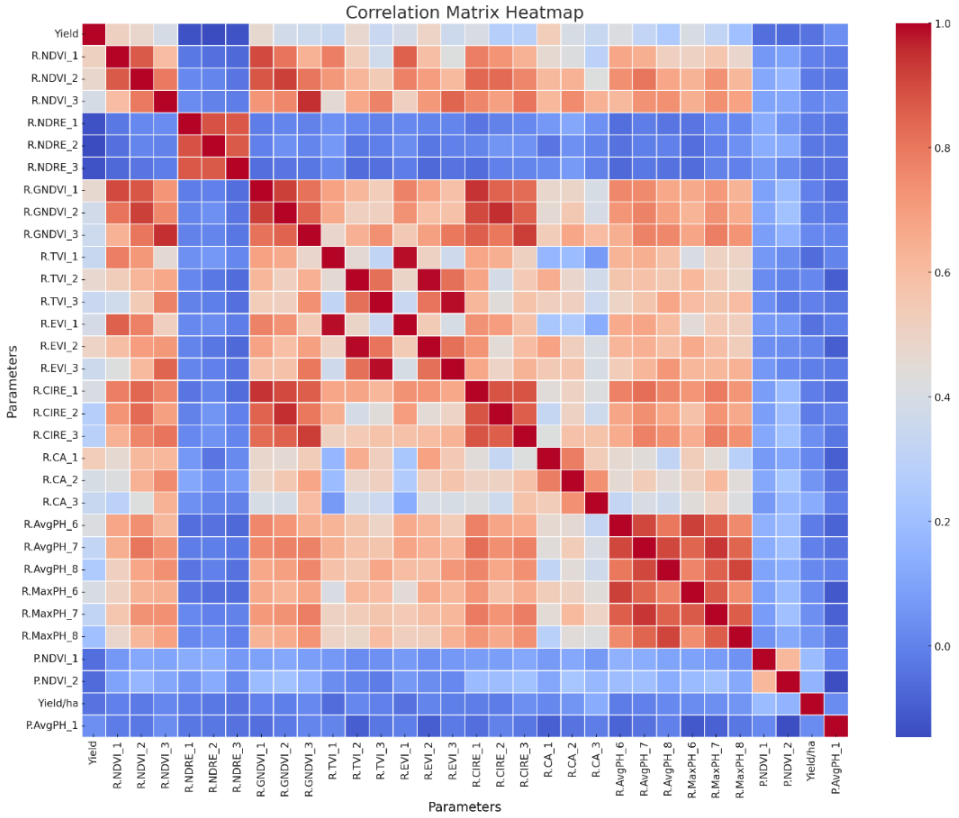

Figure 60 Heat map correlation between crop parameter for Ecobreed field 2022

2020

The correlation analysis conducted with fresh yield confirmed that these three variables, extracting from UAV imagery captured during tuber filling stage, were the most closely correlated with yield. Further analysis identified canopy area as the most importance variable in the model follows by NDVI and average canopy height as illustrated in the heatmap in **Figure 58**, **Figure 59**, and **Figure 60** Error! Reference source not found.**.**

Table 22 Random Forest model performance summary in predicting the 2020 potato yield.

Formular	Crop parameter	% Var explained	MSE	RMSE	MAE	R-squared
Yield = train parameters, y = train target	All crop parameters	29.08	0.951	0.975	MAE	0.412
Formular = Yield ~ Crop parameters	NDVI + Canopy area + Average height + Maximum height	30.43	0.951	0.975	0.794	0.412
Formular = Yield ~ Crop parameters	NDVI + Canopy area + Average height	30.73	0.312	0.558	0.423	0.807
Formular = Yield ~ Crop parameters	NDVI + Canopy area	32.02	0.333	0.577	0.444	0.794

The RF model was developed using all the parameters as the train features (80% of the data) and tested on the test features to predict the yield (**Table 22**). The model explained 29.08 % with a Mean Squared Error (MSE) of 0.951, a Root Mean Squared Error (RMSE) of 0.975,

and a Mean Absolute Error (MAE) of 0.794, R^2 value of 0.412. This suggests that while the model has predictive potential, a substantial portion of the yield variation remains unexplained. After selecting the top three variables that had the strongest correlation with yield, the model accuracy remained at 30.43% - 32.02%. However, this approach significantly improves the R^2 value to 0.807, suggesting a closer fit to the actual yield data when key variables are considered. This improvement highlights the importance of these variables in predicting model.

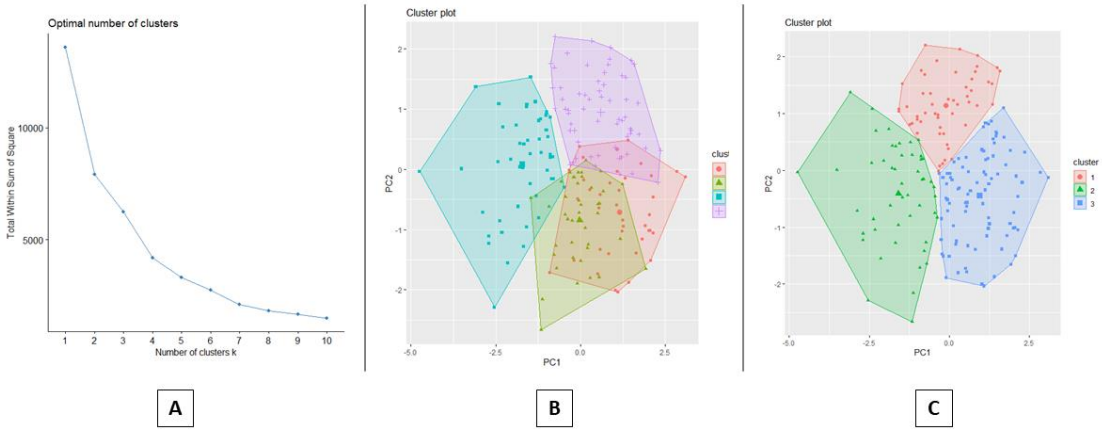

Figure 61 Clustering analysis for potato variety grouping (2020 trial). (A) shows the elbow plot determine the optimal number of clusters for grouping based on crop parameter. (B) and (C) illustrate cluster plots created using Principal Component Analysis (PCA), where each cluster represents group of potato varieties with similar characteristics.

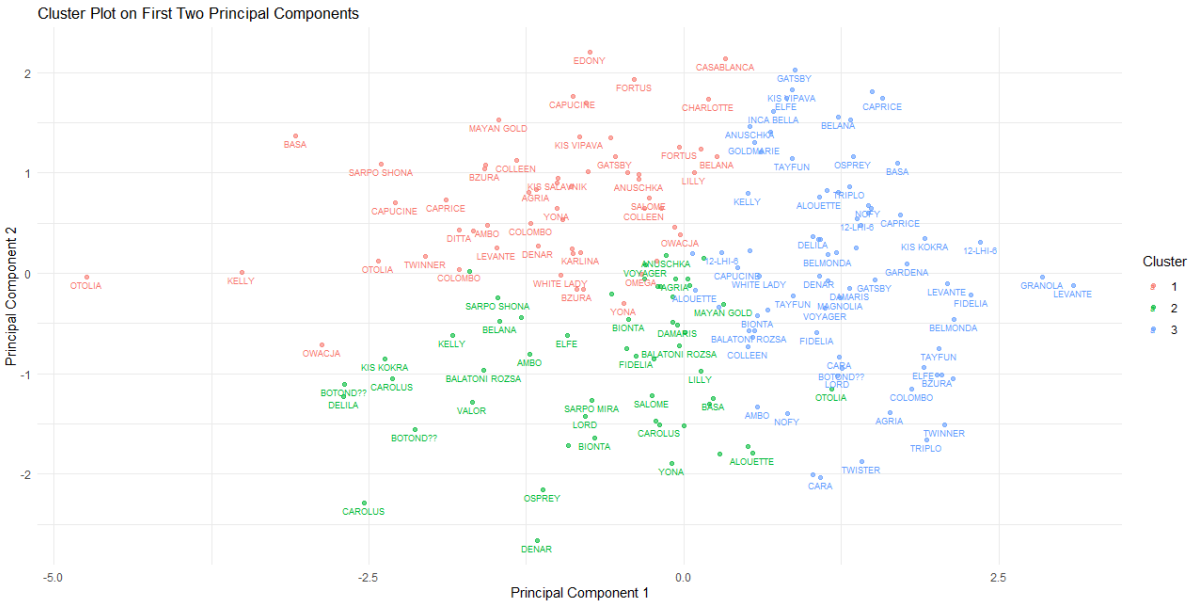

Figure 62 Clusters of potato varieties created using Principal Component Analysis (PCA), categorised into three distinct clusters with each dot labelled with variety name.

2021

Based on the correlation metrics, the data extracted from UAV-based imagery captured during tuber initiation stage, provided the strongest correlation with yield. With Ratio Vegetation Index (RVI) was the most correlated, follow by Normalised Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI) and Green Normalised Difference Vegetation Index (GNDVI) that captured during the tuber initiation stage (**Figure 59**).

Table 23 Random Forest model performance summary in predicting the 2021 potato yield.

Formular	Crop parameter	% Var explained	MSE	RMSE	MAE	R-squared
Yield = train parameters, y = train target	All crop parameters	41.64	3.501	1.871	1.465	0.541
Formular = Yield ~ Crop parameters	RVI + NDVI + EVI + GNDVI	34.78	0.816	0.904	0.679	0.893

The Random Forest model which utilised all available parameters as the train features (80% of the dataset) and tested on the test features (20%). The model accounted for 41.64 % explained MSE = 3.501, RMSE = 1.871, MAE = 1.465 and R-squared = 0.541, indicating a moderate fit to the data (**Table 23**). When considering only vegetation indices (NDVI, EVI, GNDVI, and RVI), we found that RF outperformed the previous model, the variable explained was slightly decreased to 34.78%, but the fit of the model improved with a higher R-squared at 0.893 and lower means od MAE, MSE, and RMSE. This suggests that while using a more targeted set of variables reduced the accuracy, the model's prediction became more accurate and reliable. Recent studies have classified that the highest performance found here using RF possibly occurred due to the internal structure of the algorithm, which is based on multiple decision tree sets (Pantazi *et al.*, 2016; Teodoro *et al.*, 2021).

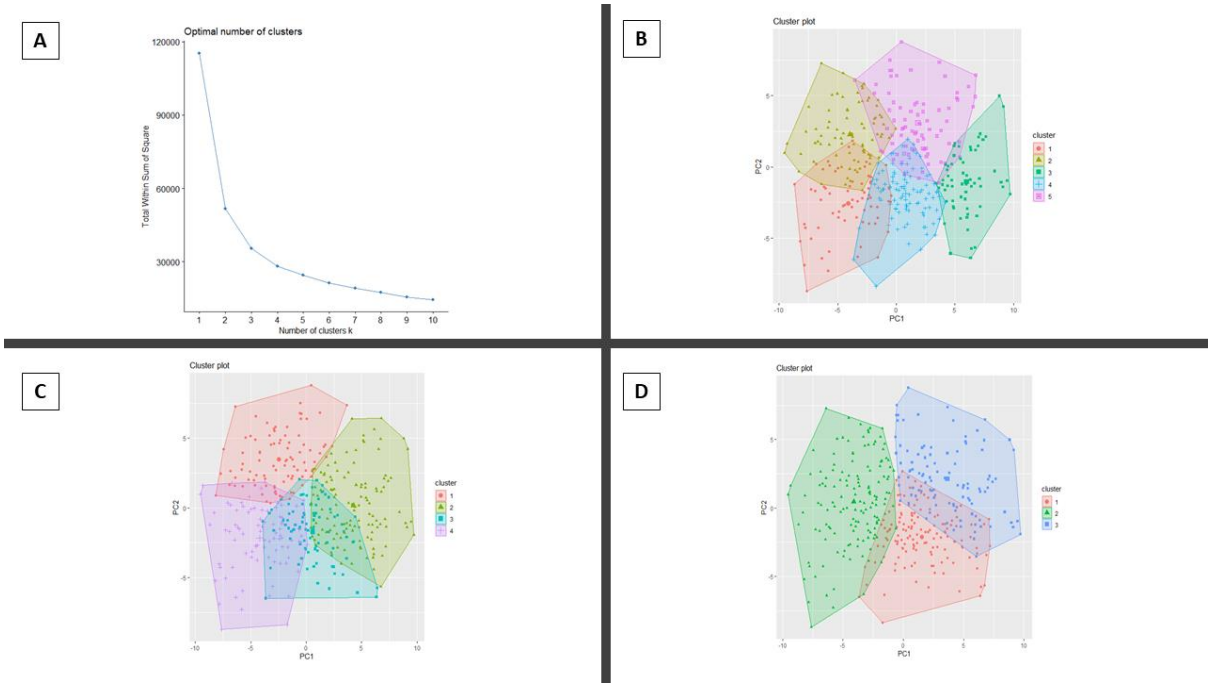

Figure 63 Clustering analysis for potato variety grouping (2020 trial). (A) shows the elbow plot determine the optimal number of clusters for grouping based on crop parameter. (B), (C), and (D) illustrate cluster plots created using Principal Component Analysis (PCA), where each cluster represents group of potato varieties with similar characteristics.

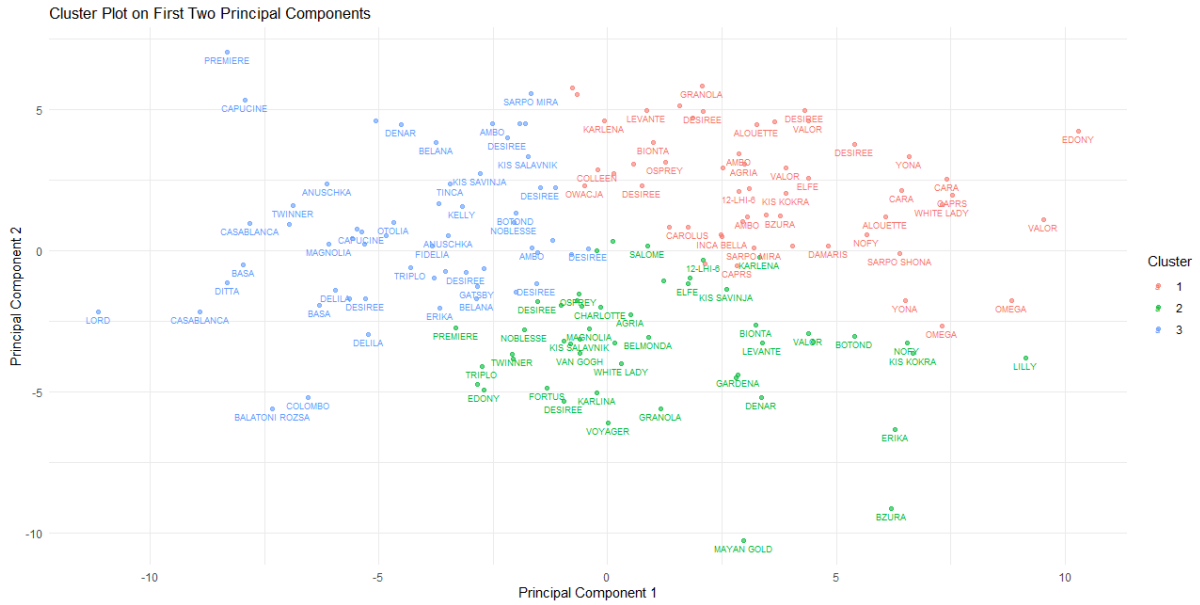

Figure 64 Clusters of potato varieties created using Principal Component Analysis (PCA), categorised into three distinct clusters with each dot labelled with variety name.

2022

The correlation analysis with yield highlighted that average plant height, extracted from UAV-based imagery captured during tuber initiation stage, showed the highest correlation follow by canopy area and maximum plant hight (0.739, 0.724, and 0.651, respectively) (**Figure 59**).

Table 24 Random Forest model performance summary in predicting the 2022 potato yield.

Formular	Crop parameter	% Var explained	MSE	RMSE	MAE	R-squared
Yield = train parameters, y = train target	All crop parameters	56.85	0.295	0.543	1.465	0.645
Formular = Yield ~ Crop parameters	Canopy area + Canopy volume + Average height +	56.73	0.089	0.298	0.217	0.925

	Maximum Height					
Formular = Yield ~ Crop parameters	Average height + Canopy area + Maximum plant height	58.06	0.048	0.219	0.179	0.942

The RF model trained with all the variable explained 56.85% of the yield variation (MSE = 0.295, RMSE = 0.543, MAE = 0.431 and R-squared = 0.645) (**Table 24**). The enhanced RF model, which included only the crop parameters with the highest correlation to yield, showed a significant enhancement in its ability to predict yield. The model explained 58.06% of the yield variation and significantly enhanced the model's accuracy as evidenced by a lower MSE of 0.048, RMSE of 0.219, an improved MAE of 0.179, and R^2 value to 0.942. These improvements indicate a stronger predictive ability and a closer fit between the predicted yields and the actual yield data.

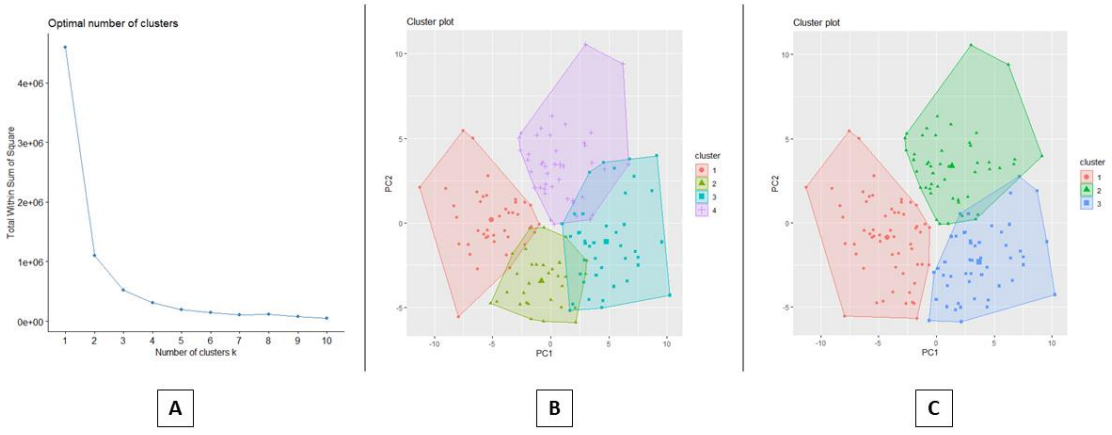

Figure 65 Clustering analysis for potato variety grouping (2020 trial). (A) shows the elbow plot determine the optimal number of clusters for grouping based on crop parameter. (B) and

(C) illustrate cluster plots created using Principal Component Analysis (PCA), where each cluster represents group of potato varieties with similar characteristics.

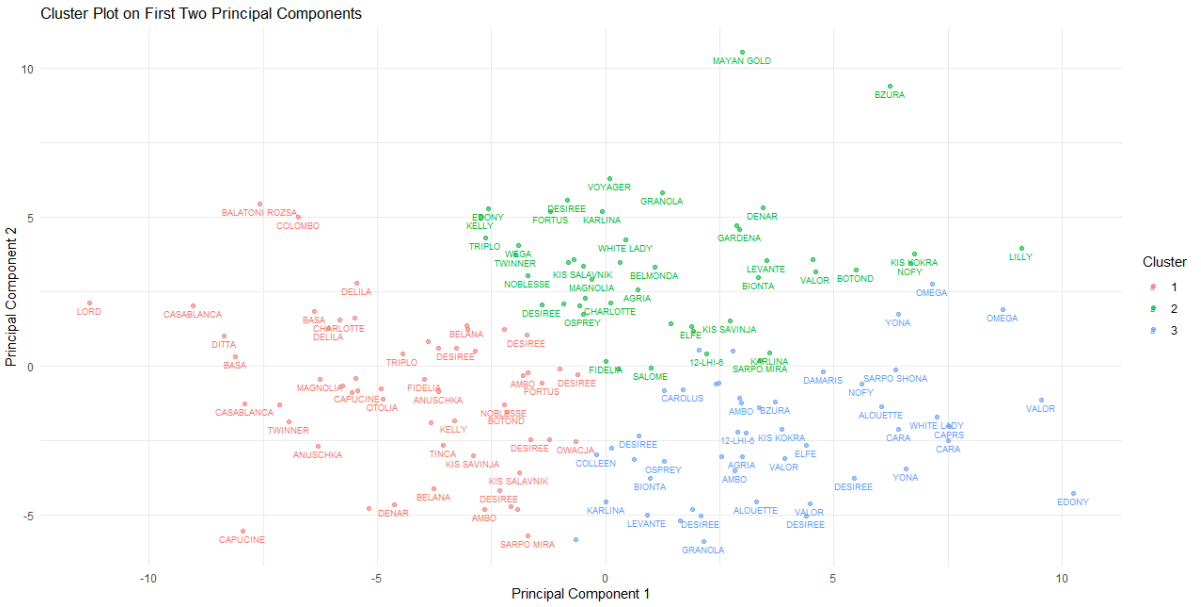

Figure 66 Clusters of potato varieties created using Principal Component Analysis (PCA), categorised into three distinct clusters with each dot labelled with variety name.

Overall, the Random Forest (RF) regression models have demonstrated significant accuracy in potato yield prediction over three years trials. However, when all crop parameters were used to train the model the R^2 value of 0.412 in 2020, 0.541 in 2021, and 0.645 in 2022. Whereas, when the model was further adjusted to choose the variable with the highest correlation to yield. The R^2 values were improved to 0.807, 0.893, and 0.942, respectively. This suggests that the model was highly predictive of actual yield with a significantly reduced error rate when using variables that are highly correlated with the actual yield performance. The high performance of RF regression model is evidently achieved when predictor or explanatory variable are highly correlated (Pantazi *et al.*, 2016).

The outcome from the correlation analysis also reveals that crop parameters are more correlated with the yield at tuber initiation stage, where plant reach the maximum vegetative growth. Whereas the result from previous studies found that in order to achieved a high accuracy the most suitable UAV flight is close to the plant maturity stage such as Anderson II *et al.* (2019) , Gómez *et al.* (2019) and Al-Gaadi *et al.* (2016). In this study we found that as closer to maturity stage results in less vegetative growth and shorter plant.

In conclusion, the RF regression model performed well in predicting potato yield. This results also demonstrated the suitability of our model to predict potato yields in the organic varietal studied and based on these prediction model we can categorised potato varieties based on the yield performance into three main groups (**Figure 62, Figure 64, and Figure 66**) as suggest by the elbow plot determine the optimal number of clusters for grouping based on crop parameter that used in the model (**Figure 61, Figure 63, and Figure 65**).

5.8.4 DISEASE PREDICTION

In this part, we utilised the machine leaning approach to classify the disease score compared to the field visual assessment in different potato varieties in order to predict the disease incidence. This approach aimed to enhance the accuracy and efficiency of disease incidence prediction. The K-means clustering, and Random Forest (RF) classification model were computed based on various crop parameters such as canopy area, canopy height, and vegetation indices (VIs) include NDVI, NDRE, GNDVI, TVI, EVI, and CIRE extracted from the multispectral UAV-based imagery.

In potatoes, "disease scoring" refers to a scoring system that use to determine how serious a disease is to the crop (**Table 3**). Typically, this procedure comprises physically inspecting the plants in the field for signs of infection, such as discolouration, lesions, or other symptoms, and assessing the disease's severity on a scale of one to ten. This score system helps researchers and farmers keep an eye on the emergence of diseases, assess the disease resistance, and make well-informed decisions on disease management strategies. The "k-means clustering by PCA" technique evaluates the spectral properties and disease severity of potato crops by combining data from the VI and disease score. Crop health is reflected in disease score data, whereas biophysical parameters are indicated by VI data from multispectral UAV images. Through the integration of these datasets, the technique finds patterns and correlations, using PCA to reduce dimensionality and k-means to cluster data according to spectral and disease severity features. With this method, potato crops can be divided into various categories according to their spectral characteristics and overall health.

The results from K-means clustering illustrates in **Figure 67**. Where **Figure 67 A**, shows that the k-means clusters exhibit overlap, with no clear distinction among disease score clusters when utilising the NDVI and NDRE index. However, in **Figure 67 B**, upon removing disease score 9 from the dataset, cluster classification appears to enhance, revealing two distinct clusters representing disease score 10, indicative of healthy potato crops. In response to the

challenges faced by the k-means clustering approach, Random Forest (RF) classification method was chosen. As RF algorithm has shown outstanding performance for disease detection in various crop, such as ground nut (Chaudhary *et al.*, 2016), tomato (Govardhan and V, 2019) and rice (Rajpoot *et al.*, 2023). Therefore, we have selected RF approach for potato diseases classification.

The RF model was employed to assess the disease score of potatoes using VI and additional canopy parameters, such as canopy area and plant height, during the tuber filling stage and maturation growth stages in Ecobreed potatoes across three different years: 2020, 2021, and 2022. The objective was to determine the most effective stage for monitoring potato diseases. The RF model achieved an accuracy of 73% during the tuber filling stage and 50% during the maturity stage in classifying disease scores based on VIs. This indicates that the tuber filling stage is more suitable for accurately classifying disease scores. When we focused only on the tuber filling stage, the results show an accuracy of 63% for 2020, 78% for 2021, and 65% for 2022. The higher accuracy achieved in 2021 and 2022 can be attributed to the inclusion of more vegetation indices compared to 2020.

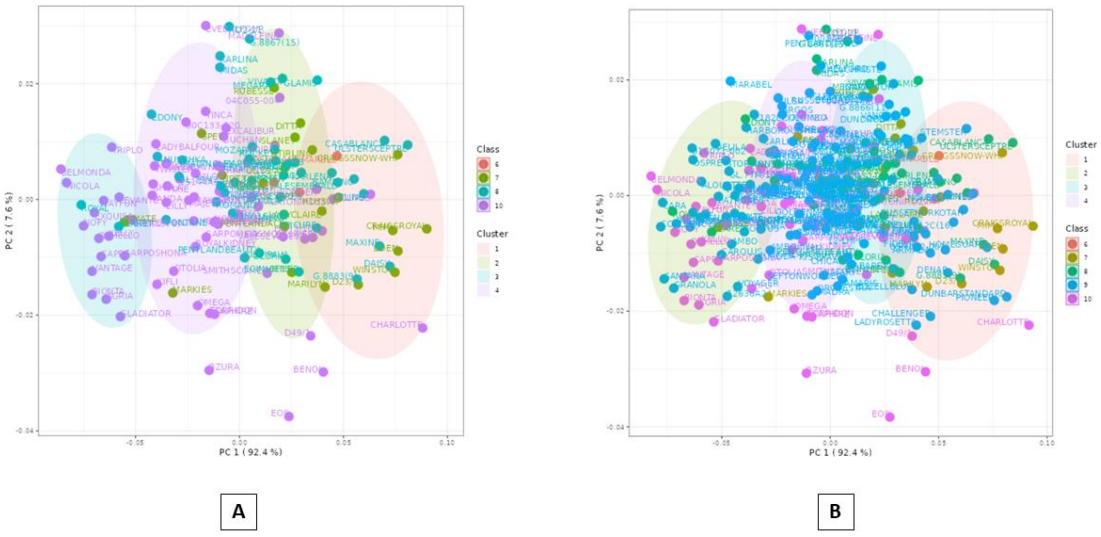

Figure 67 (A) K-means clustering using PCA1 and PCA2 to classify the disease score using NDVI and NDRE. (B) K-means clustering using PC1 and PC2 to classify the disease score potato excluding score 9 by NDVI and NDRE values.

K-means clustering and RF classification model were performed best when using crop parameters and VIs extracted from the multispectral UAV-based imagery and field visual assessments based on flight and date that captured potato plants during tuber filling stage in all three years data. This is due to the early field inspection was not applicable to the model as most of the potato crop has a healthy with the disease score of 10 and some with score 9 (healthy plant) (**Table 3**). The best accuracy to classify disease score were at tuber filling stage in 2021 with an accuracy of 78 %, 65% in 2022, and 63% in 2020. The less accuracy occurred in 2020 due to the limited vegetation indices that has been observed compared to 2021 and 2022.

The model that included plant height and other crop growth parameters was outperformed the model that trained only on VIs, although the model accuracy was slightly dropped to 73%. However, we cannot excluded VIs data as stated by Khan *et al.* (2018) that VIs observed from aerial images were effective in capturing both intra-variety and inter-variety differenced among crops. Moreover, the variable importance plot (**Figure 68**) indicates that plant height is the most significant crop parameter contributing to the classification of disease scores using the random forest model. This confirmed by the study from Kwambai *et al.* (2023) and found that traits that influence genotype and environment interactions is plant height (49.3%). This finding is significant as it enables the implementation of early management strategies to mitigate potato diseases effectively (Access code in Appendix).

To further adjust and improve model accuracy, advancements in image processing techniques are necessary. As computerised image processing techniques are critical for detecting and classifying plant diseases and prevent their spreading (Demilie, 2024). A promising direction is the development of end-to-end deep learning model (CropdocNet) proposed by Shi *et al.* (2022), which demonstrated a high accuracy in extracting hierarchical vector features from UAV-based hyperspectral data, achieved average accuracy of 98.09%.

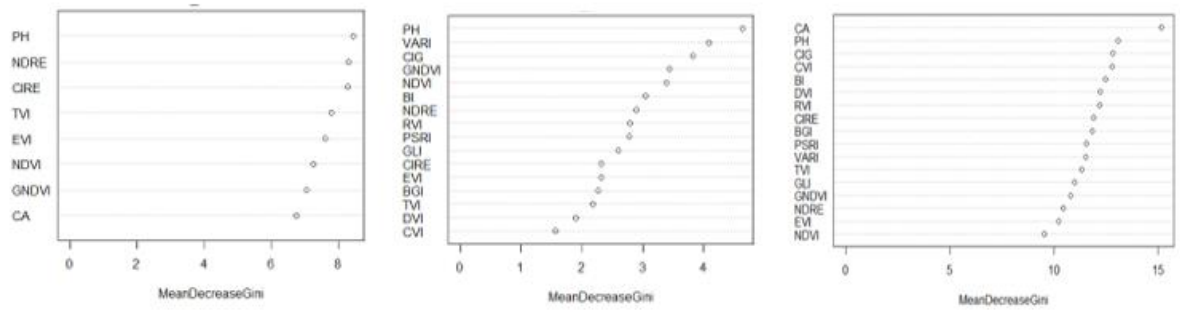

Figure 68 Variable importance plot by Random Forest (A) 2020 (B) 2021 (C) 2022.

In summary, the RF classification model has proven to be an effective tool and more suitable approach for classifying disease scores compared to k-means clustering, especially during the tuber filling stage. Despite the successes achieved with the RF model, there is a room for improvement in the model in order to achieved higher accuracy rate. The ongoing development and refinement of machine learning applications in agriculture hold the promise of transforming how crop diseases are detected, understood, and managed.

5.8.5 GROWING PATTERNS

Initially, Random Forest classification model was carried out in R language as an approach to classify the different variety of potato using different vegetation indices (VIs), canopy area, and canopy height. Since there were only two replicates per UAV flight, the accuracy was relatively low. Hence, here the study examined how the temperature affect the yield for each potato variety across the years. It focused on the average temperature during the relevant months of the growing season for each year, hypothesising that these factors might affect yield differently. The study then examined the correlation between yield and weather parameters for each variety across the years to identify any noticeable pattern.

In 2020, the average temperature during the growing season was 12.60 degree Celsius. 12.34 degree Celsius in 2021, and 13.47 degree Celsius in 2022. Where the average relative humidity (RH) is approximately 93.12%, 88.68%, and 90.28%, respectively.

Given many potato varieties, this study only focuses on the top 5 and bottom 5 based on the yield performance analysis (**Table 21**). For example, ‘Lavante’ showed a significant increase in yield from 8.33 kg in 2020 to 15.65 kg in 2021 and reaching 19.09 kg in 2022. Similarly, ‘Cara’ showed a significant increase in yield from 3.02 kg in 2020 to 11.36 kg in 2021 and reached 17.03 kg in 2022. ‘Gardena’ also showed an increasing trend in yield from 4.37 kg in 2020, to 9.43 in 2021, and 11.35 kg in 2022. The varieties ‘Tayfun’ and ‘Twister’ are also demonstrated similar pattern. This might due at the higher temperature, the photoperiod effect is stronger (Struik, 2007).

For the varieties with lower yield, there was a consistent patterns of yield pattern peaking in 2021, which was a slightly cooler year, followed by a sharp decline in 2022, the warmest of the three years. For instance, ‘Inca Bella’ started with 2.69 kg in 2020, increased to 6.53 kg in 2021, and then decreased to 4.33 kg in 2022. ‘Mayan Gold’ showed an improvement from 1.43 kg in 2020 to 4.53 kg in 2021, with a slight reduction to 3.55 kg in

2022. Meanwhile ‘Ditta’ showed a significant yield increase from 3.2 kg in 2020 to 7.93kg in 2021 and rose further to 9.39 kg in 2022. These trends suggest that while temperature may play a significant role in affecting yields, the exact relationship varies by variety and likely influenced by a complex interaction of weather parameters. This also confirm in the study from Den *et al.* (2022) where the effect of daytime and night-time temperature on photosynthesis was explored during the field trials. There is no literature suggesting other values than the defaults ones and low model sensitivity to the parameters.

Figure 69 Yield trends for ‘Levente’ and ‘Inca Bella’ against the average temperature and relative humidity for the years 2020, 2021, and 2022.

For a general overview of RH, it appears that highest yielding years (2022) has intermediate RH levels (**Figure 69** Yield trends for ‘Levente’ and ‘Inca Bella’ against the average temperature and relative humidity for the years 2020, 2021, and 2022.). This indicates the complex relationship between yield performance and weather data. This preliminary analysis suggests that while temperature might influence yield, the relationship is complex and

likely influenced by multiple factors, including possibly the variety's specific tolerance to temperature changes and other stresses that might occur in the field such as disease incidence.

The complexity suggests that changing in environmental condition and plant genotype variation has an influence on tuber yield. The result from this analysis aligns with the findings from Tessema *et al.* (2020), which concluded the integration between varietal characteristics and environmental variations has considerable influence on the tuber yield and the potato phenotypes. Den *et al.* (2022) further contribute to this understanding by highlighting that average yields were highly influenced by potato genotypes and their reaction with the growing location.

Additionally, when we look at the growth pattern of each variety across flights, some variety show a wide range of heights, while others are more uniform. By comparing the same variety growth parameter across different flights, we can see how plant height and other crop parameters change over time. Compared to the study in chapter 3 and 4, the potato plot in this chapter is longer and have more plant (15 seeds). This reference of the work from Liebisch *et al.* (2015) suggested the significance of experimental design in effectively evaluating phenotypic features and yield potential. It indicates a preference for utilising two-row plots considered the best selection for aerial phenotyping.

In conclusion, the research highlights the complex correlation between potato productivity and environmental variables, emphasising the need for a broader set of varieties studied and combing this data with other field observations. Adopting a comprehensive viewpoint is essential for creating more efficient agricultural methods and will provide deeper insights into how weather conditions may influence yield variations across different years.

5.9 CONCLUSION

In the context of the current climate change impact on potato production in the UK, where could precipitating shifts in the agricultural calendar, affecting potato production through earlier or late planting and harvest times, the call for a change to better adapted varieties is urgent. The use of high-throughput phenotyping and remote sensing techniques, including UAV-based imagery coupled with machine learning algorithms, represents a significant advancement in this field. This study evaluated the efficacy of Random Forest (RF) regression for yield prediction and classification for disease detection in organic potato. The result from this study shows strong potential for the implementation of RF algorithm as an

alternative statistical modelling method for yield prediction and disease detection. Despite the promising results achieved with RF models trained on multispectral UAV imagery, challenges remain in terms of computational demands and the effectiveness of multispectral data for disease detection and yield prediction. The results also reveal a non-linear correlation between weather parameter and potato productivity, where changes in temperature have a significant impact on the growth patterns of high-performing potato varieties. In contrast, the lower-yielding varieties recorded peak yields during the slightly cooler year. This indicates the possibility of a temperature threshold above genotypes. The integration of weather data into yield models has showed the need to consider yearly climate fluctuations in order to enhance the precision of these predictive models. This perspective highlights the complex relationship between environmental variables and agricultural productivity, emphasising the need for a deeper understanding of the effects of climate change on crop productivity. Therefore, future research is needed to explore the most effective classification approaches, that involves identifying the most suitable growth stages for disease monitoring and yield prediction and assessing the reliability of these methods in various environmental settings. In conclusion, this study contributes valuable insights into the dynamic interaction between environmental factors and agricultural productivity, offering guidance for future research and practical applications in the field of precision agriculture.

CHAPTER 6. GREENING IN POTATO TUBERS AND THE RELATIONSHIP WITH GLYCOALKALOID CONTENTS

In the fresh market, greening potatoes pose a perceived health risk due to their association with the accumulation of glycoalkaloids. Besides the perception of toxicity, greening in potato tubers is one of the major issues affecting tuber quality and leading to potato waste. The synthesis of chlorophyll and potato glycoalkaloids (PGAs) in potato tubers occurs in direct response to light exposure. In this study, total PGAs and chlorophyll accumulation were determined in nine potato varieties across different greening scales (using an arbitrary scale of 0-5) and calibrating the scale for PGA content using high performance liquid chromatography (HPLC) with UV/Visible spectrometer detection. Potato variety 'Craigs Royal' exhibited the highest total PGA concentration (2,635 ± 683 mg/kg fresh weight, mg kg^{-1} FW) at a greening scale of 5, followed by the varieties Majestic, Anna, Agria, and Yukon Gold. Significant positive correlations were observed for both α-chaconine and α-solanine ($p \leq 0.001$) and for chlorophyll and carotenoid concentrations ($p \leq 0.001$). However, the PGA concentration did not increase proportionally with the greening level and chlorophyll content, particularly in very green tubers. This finding revealed the complex relationship between PGA concentration and greening level. In addition, and of particular concern, the PGA concentration exceeded the recommended food safety limit (200 mg kg^{-1} FW) even at relatively low levels of greening in some tubers from varieties, such as, Dundrod (525.9 mg kg^{-1} FW, level 1), Anna (729.5 mg kg^{-1} FW, level 1), and Craigs Royal (1453 mg kg^{-1} FW, level 1). In contrast, all non-green potato tubers had a total PGA content within the safety limit for human consumption, ranging from 21.8 to 189.5 mg kg^{-1} FW. In addition, the chlorophyll content in non-green potatoes ranged from 0.7 to 5.5 mg kg^{-1} FW. These finding emphasise the importance of avoiding consumption of green potatoes and provide valuable insights for both consumer and breeders, while highlighting the importance of quality control measurements to ensure safe potato consumption and reduce potential health risks.

6.1 INTRODUCTION

The *Solanaceae* family (or deadly nightshade)to which potato belongs , is one of the most important plant families as an agricultural food commodity, yet it is also known for containing toxic compounds, such as alkaloids and glycoalkaloids (GAs) (Friedman, 2006; Barceloux, 2009). GAs are not required for plant growth and function; however, they have been linked with plant resistance to pests and pathogens and have been shown to be toxic to

organisms ranging from fungi to humans (Milner *et al.*, 2011). Solanine, chaconine, capsaicin and nicotine are example of the chemicals that exist in this family. Solanine poisoning is the most common when the amount of solanine increases substantially from green or sprouted in commercial potatoes (Barceloux, 2009).

Potato glycoalkaloids (PGAs) are natural secondary plant metabolic compounds that contain nitrogen in a steroidal structure and carbohydrate side chain (aglycons) (**Figure 70**) (Barceloux, 2009). There are other glycoalkaloids found in potatoes, 95 percent of the total potato glycoalkaloids are primarily comprised of α-chaconine and α-solanine (Friedman *et al.*, 1997; Duke Gekonge *et al.*, 2016). α-chaconine and α-solanine are trisaccharide steroidal glycoalkaloids but differ at the composition of the sugar side chain, α-chaconine consists of two rhamnose molecules and one glucose, whereas α-solanine is formed up of galactose, glucose and rhamnose (Friedman *et al.*, 2003).

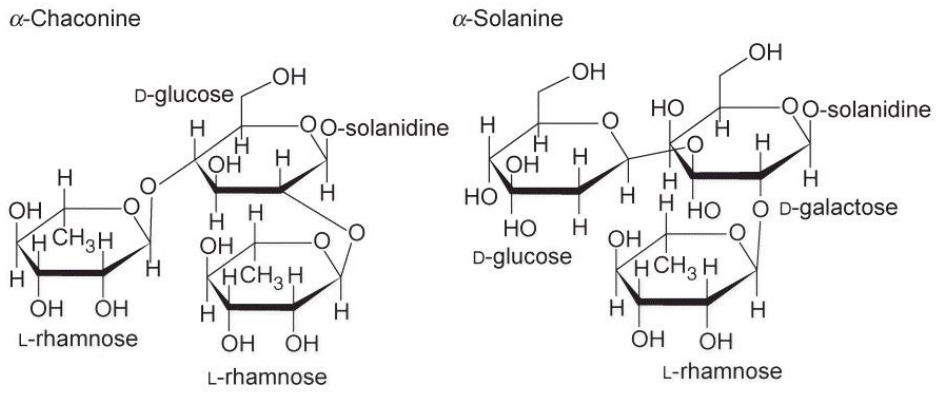

Figure 70 Chemical structures of potato glycoalkaloids α-chaconine and α-solanine (Jensen *et al.*, 2007).

6.1.1 GREENING IN POTATOES

Tuber quality such as flavour, taste, texture, cooking quality and skin appearance all influence consumer acceptance of potatoes (Nema *et al.*, 2008). Potatoes are prone to quality degradation by water loss, diseases, colour changing or build-up of toxin (Tilahun *et al.*, 2020a). Greening in potato tubers is the single biggest cause of tuber quality loss. It is considered as waste due to its association with the accumulation of toxin. Green potatoes indicate an increase in chlorophyll and glycoalkaloids levels. Potato plant contains potato glycoalkaloids (PGAs), primarily α-solanine and α-chaconine (Phillips *et al.*, 1996). Even though the concentrations found in potato tubers are generally low and mostly concentrated to

the tuber's periderm (Tanios *et al.*, 2020), their toxicity and bitterness are detrimental to the product's quality (Tilahun *et al.*, 2020a).

Potato tubers are non-photosynthetic tissues; however, the exposure of tuber to light result in greening potatoes, as the tuber can accumulate chlorophyll in the peripheral layers (Zhu *et al.*, 1984; Tanios *et al.*, 2018). These pigments reflect predominantly in green light, which gives potato tubers a green colour. Greening can occur in the field, storage cellars, on the store shelf, and at home on the counter and is influenced by varietal genetic, tuber physiology and the environmental interaction (Tanios *et al.*, 2020). Light also induces accumulation of glycoalkaloids, which causes a bitter taste in potato tubers, occurs concurrently with tuber greening, despite being under independent genetic control (Nahar *et al.*, 2017; Tanios *et al.*, 2020). The formation of chlorophyll may be look unappealing, but it does not affect the taste or cause toxicity concern, as increased glycoalkaloid levels do. Green tuber still poses a major rejection among consumers. Therefore, this led to economic loss (Tanios *et al.*, 2020; Dhalsamant *et al.*, 2022).

Important factors influencing the formation of chlorophyll and glycoalkaloids in potato are cultivar, growing conditions, injuries and diseases (Nema *et al.*, 2008). Numerous studies have shown that susceptibility to greening and total glycoalkaloid contents vary within cultivars and are influenced by environmental interaction (Bejarano *et al.*, 2000; Friedman *et al.*, 2003; Grunenfelder *et al.*, 2006a). As genetics/genotype is a significant factor, strategies to prevent greening should begin with understanding varietal variation in the field which includes selection of greening resistance cultivars and implementation of appropriate agronomic practices in the field (Pavek and Thornton, 2009; Tanios *et al.*, 2018; Tanios *et al.*, 2020). These practices including planting tubers at the appropriate depth, sufficient soil hilling during the growing season and providing adequate irrigation to prevent soil erosion and cracking (Pavlista, 2001). However, after harvest PGAs content can increase during storage, transportation, and on the shelf, when tubers are under influence of light, heat, damaging, and sprouting (Friedman, 2006). To prevent greening and PGAs development after harvesting, numerous studies have been carried out on the ideal postharvest condition. Post-harvest measurement for controlling greening includes managing light condition and using dehumidifiers along with controlling the temperature and pre-treatment (Dhalsamant *et al.*, 2022).

PGAs are found in all parts of the potato plant with the highest concentrations in flowers and the lowest in tubers, these levels also vary between different plants and varieties (Grunenfelder *et al.*, 2006a). The PGAs function as a defence molecule to protect plants against pathogens and other pests; however, they are relatively toxic compared to other common plant poisons if consumed in high doses (Abreu *et al.*, 2007). The potential human toxicity of PGAs has led to the establishment of guidelines as the glycoalkaloid composition is unaffected by food preparation steps (boiling, baking, or frying) (Barceloux, 2009; Schrenk *et al.*, 2020). Acute toxic effects of potato GAs in humans include gastrointestinal symptoms of various severity, such as vomiting, diarrhoea, and abdominal pain, which can develop after consuming 1 mg kg^{-1} body weight or more of total PGAs (European Food Safety, 2020; Schrenk *et al.*, 2020). A safe level of total PGAs intake is 200 mg kg^{-1} fresh weight, but at levels above 280 mg kg^{-1} fresh weight can cause serious illness and even death (Nema *et al.*, 2008). It is important to note that potato tubers naturally contain a small quantity of potato glycoalkaloids prior to any light exposure (Petermann and Morris, 1985). For tuber protected from light, the concentration of PGAs has been found range to on average from 12-20 mg kg^{-1} fresh weight (Grunenfelder *et al.*, 2006b).

Table 25 Physical properties of α-chaconine and α-solanine (adapted from Nema *et al.* (2008)).

Information	**α-chaconine**	**α-solanine**
Chemical identification	$C_{45}H_{73}NO_{14}$	$C_{45}H_{73}NO_{15}$
Physical stage	Solid	Solid or Slender needles from 85% alcohol
Melting point (˚C)	243	190-285[a]
Solubility	Insoluble in water	Soluble in water
Organic solvent	Soluble in organic solvent e.g., ethanol and acetone	Readily soluble in hot alcohol; practically insoluble in ether and chloroform
Molecular weight	852.1	868.1
Biological half-life (h)	19.1	10.7

Plasma level (ng mL^{-1})	6-11	3-111

[a] Decomposes at this temperature

6.1.2 POTATO GLYCOALKALOIDS AND CHLOROPHYLL DETERMINATION/QUANTIFICATION TECHNIQUES

For analysing glycoalkaloid contents in potato tubers are usually measured using destructive analysis approaches. The methodological approaches for determination/quantification of potato material includes calorimeter, gas chromatography and mass spectrometry (GC - MS), high performance liquid chromatography (HPLC), thin-layer chromatography (TLC), enzyme-linked immunosorbent assay (ELISA), biosensor-based method, hydrolysis, and isotachyphoresis (Nema *et al.*, 2008). As PGAs are only partially soluble in aqueous solution at pH ≥ 7 (**Table 25**), the extraction solvent needs to be organic, acidic or both (Friedman *et al.*, 1997). Some methods have low sensitivity chemical or require derivatisation, while other require expensive compounds or take an excessively long preparation time. HPLC separation with UV/VIS detection appears to be the most commonly preferred approach of analysis because its rapid, reliable, repeatable and ability to distinguish between two glycoalkaloids in different plant species without required derivatisation (Matsuda *et al.*, 2004; Friedman, 2006; Nema *et al.*, 2008).

Multiple methods have been proposed to extract and quantify chlorophyll pigments in plant tissues; however, there is no specific protocol, which is applicable to all the varieties of known potatoes due to the lack of applicability of common protocol to all know varieties. Thus, protocol detailed here is a modification from Petermann and Morris (1985) and Tanios *et al.* (2020). This protocol is designed to extract and analyse the chlorophyll and carotenoid contents in potato tuber periderm using N, N – dimethylformamide (DMF). The solvent DMF efficiently extracted chlorophyll and carotenoid from intact tuber periderm of potatoes (Bergweiler and Lütz, 1986). Complete extraction occurred within 8-48 hours of immersion. The most important reason to use DMF is storage time, with DMF can store up to 10 days at 4 degree Celsius with no pigment degradation (Bergweiler and Lütz, 1986).

Despite extensive efforts to control greening in potato tubers, the complex relationship between tuber greening and PGA accumulation has yet to be fully understood and requires accurate method to measure the content of individual glycoalkaloids (Friedman *et al.*, 2003). Some potato varieties have many preferable industrial quality traits such as diseases resistance,

high yield, and high dry matter content but they produce bitter tubers and must be removed from the market due to high levels of glycoalkaloids (Peng *et al.*, 2019). HPLC separation with UV/VIS detection method are widely used to determine the concentration of individual glycoalkaloids of different part of potatoes (Nema *et al.*, 2008). However, peak separation of α-chaconine and α-solanine on HPLC chromatograms is not always present. The aims of this study are therefore: 1. to separate α-chaconine and α-solanine in potato extracts by HPLC with UA/Visible spectrometer detection to allow quantitative determination, 2. to understand the relationship between greening and PGA concentration in different potato varieties, and 3. to further investigate the relationship between PGA concentration and chlorophyll and carotenoid content. These findings will provide valuable insights for both consumers and the potato industry by highlighting the importance of quality control in producing potatoes safe for consumption.

6.2 MATERIAL AND METHODS

6.2.1 POTATO SAMPLES

This study examined the greening response within nine different potato varieties (listed in **Table 26**). Potato tuber samples used in this study were part of a large study in 2020 that performed at Nafferton Farm, Newcastle University, United Kingdom. The germplasm panel were part of a tetraploid variety association panel available at the James Hutton Institute (Sharma *et al.*, 2018). Individual variety was planted in 2020 and grown under field condition. During the growing season, ridges were regularly topped up to protect growing tubers from light exposure, approximately once every four weeks. At the time of harvest, each tuber was harvest once it reached natural senescence. After harvesting, tubers were then stored, in the dark, at room temperature (4-10 °C) for at least 40 days to allow postharvest maturation. This process ensured conditions for the subsequent analyses and examination of greening and glycoalkaloid content in mature tubers.

Table 26 List of potato varieties that use in this study.

Variety	*Skin Colour*	*Type*
Agria	Yellow	Intermediate
Alouette	Pink	Early
Anna	Yellow	Intermediate
Craigs Royal	Yellow	Early
Dundrod	Yellow	Early
Majestic	Yellow	Main Crop
Maxine	Pink	Early
Mimi	Pink	Early
Yukon Gold	Yellow	Early

6.2.2 DATA COLLECTION

Visual assessment

Sixteen tubers of each variety, which were undamaged and similar in size, were selected and arranged in a row in a metal tray. The tray was left exposed to natural light at room temperature (4-10 °C, February 2021, UK). To avoid any bias in light intensity, the orientation of the tubers was kept constant. After the light exposure, the colour change was visually evaluated based on the greening percentage area on the tuber surface, using an arbitrary scoring scale of 0 - 5 where a score of 0 was applied when < 1% was green and a score of 5 applied when > 50 % of the surface was green (**Table 27**, **Figure 71**). For visual assessment, on the initial day (day 0), all tubers (each tuber considered as a replicate) were examined for appearance and scored for greening. In addition to the assessment, one tuber from each verity (non-green tuber), was represented as a control for the experiment. This allowed establishing a baseline for a comparison and using as a reference point. On day 14, tubers were scored for their degree of greening, and then periderm disks were collected for further chlorophyll, carotenoid, and PGAs analyses.

Table 27 Greening percentage coverage on the tuber surface scoring 0-5 scale.

SCORE	Percentage coverage on tuber surface
0	0-1%
1	>1%-5%
2	>5%-10%
3	>10%-25%
4	>25%-50%
5	>50%

Figure 71 The percentage coverage on the tuber surface look like (AHDB)

6.2.3 SAMPLE PREPARATION

After a visual assessment, three periderm disks were cut from each tuber' skin tissue as replicates (n = 3). The disks were cut from the stem, the middle and the bud end of each tuber periderm using a cork-borer with a thickness of 3 mm and a 1 cm diameter. The disks were then ground to powder in liquid nitrogen using a mortar and pestle. Between each sample, the mortar and pestle were rinsed with ethanol and wiped with clean tissue until all visible material was removed. The grounded sample were then stored in centrifuge tubes at -80 °C for further analysis including the assessment of chlorophyll, carotenoid, and PGAs content.

6.2.4 CHLOROPHYLL AND CAROTENOID ANALYSIS: SPECTROPHOTOMETER

For chlorophyll analysis, a sub-sample was extracted with 9 mL of N, N – dimethylformamide (DMF) (Bergweiler and Lütz, 1986). The extract was vortex mixed and

stored in the dark at 4 °C for 24 hours. The extract was then centrifuged for 15 minutes at 2500 g, prior to measurement of the absorbance at 480, 647, and 664 nm using a Biochrom Libra S12 UV/Vis spectrophotometer (Biochrom, UK). The total chlorophyll and carotenoid components were calculated as described below (adapted from Porra *et al.* (1989) and Wellburn (1994)).

The equations for determination of chlorophyll a and b, and carotenoid concentrations in N, N – dimethylformamide are derived from the difference extinction coefficients at the indicated wavelengths. The spectrophotometer was zeroed at 750 nm, and the extinction coefficients were determined as 3.11 and 12.00 at 647 nm and 664 nm, respectively for chlorophyll a. For chlorophyll b, the extinction coefficients were 20.78 and 4.88 at 647 nm and 664 nm, respectively. The total chlorophyll content was determined as the sum of chlorophyll a and chlorophyll b, where the extinction coefficients used were 17.67 and 7.12, respectively. The carotenoid content was determined using the extinction coefficients of 1000, 1.12 and 34.07 nm.

Corrected equations for determination of chlorophyll a and b, and carotenoid concentrations in N, N – dimethylformamide.

$$\text{Chlorophyll a} = 12.00(\text{A664}) - 3.11(\text{A647})$$

$$\text{Chlorophyll b} = 20.78(\text{A647}) - 4.88(\text{A664})$$

$$\text{Total chlorophyll} = 17.67(\text{A647}) + 7.12(\text{A664})$$

$$\text{Total carotenoid} = \frac{1000(\text{A480}) - 1.12(Chl\ a) - 34.07(Chl\ b)}{245}$$

Total pigments show the sum of chlorophyll and carotenoid concentration and expressed in mg kg^{-1} fresh weight (mg kg^{-1}FW).

[Note: The A term represents the determined absorbance at the stated wavelength.]

6.2.5 GLYCOALKALOID CONCENTRATIONS ANALYSIS: HPLC

SAMPLE EXTRACTION

Extracts were prepared for HPLC analysis from all the samples (**Table 26**). Tuber periderm samples were accurately weighed to four decimal places and extracted with 99.5%

HPLC grade methanol. After sonication in a sonic bath for 15 minutes, the supernatant was filtered through a nylon 0.22-µm nylon membrane. Then, the extract was evaporated to dryness under a gentle stream of nitrogen, prior to reconstitution and repeat filtration. Samples were then transferred to an auto sampler HPLC vial (0.1 ml).

The analytical standard for glycoalkaloids (α-chaconine and α-solanine) were obtained from Sigma Aldrich (Dorset, UK) with a purity of ≥ 95%. Stock solution of α-chaconine and α-solanine were prepared in methanol at a concentration of 0.5 mg ml^{-1} and stored in the dark at 4 °C prior to analysis. Calibration standards were prepared at various concentrations (0, 0.025, 0.05, 0.1, 0.15, 0.2, 0.25 and 0.5 mg ml^{-1}) for both α-chaconine and α-solanine. At the start and end of each chromatographic cycle, a complete set of calibration standards was run to create a standard calibration curve.

SAMPLE DETECTION

High performance liquid chromatography (HPLC) separations were performed using Agilent 1260 Infinity II HPLC-MWD with UV/VIS detector incorporating with chromatography data system (CDS) software from Agilent (USA). A reversed phase pentafluoro phenyl column (Supelco 2.1 m F5, 100 x 2.1 mm) from Sigma Aldrich was used to achieve chromatographic separation (Dorset, UK). The mobile phases for the analytes were water + 0.2 percent formic acid (A) and methanol + 0.2 percent formic acid (B). UV detection was measured at 210 nm using a 2.5 Hz bandwidth, 16 nm slit, and 1000 mA attenuation.

6.2.6 DATA ANALYSIS

For the Agilent LC-UV system data analysis was performed using OpenLab™ version C.01.07 supplied by Agilent Technologies (Cheadle, UK). Each sample was extracted as two technical replicates. All data was reported as the mean, analyzed using analysis of variance (ANOVA) with Tukey's post hoc test and significant differences between means determined at $p < 0.05$ (95 percent), principal component analysis (PCA), and measured with a Pearson's correlation test using R studio version 3.6.0 in R software (Team, 2021).

6.3 RESULTS AND DISCUSSION

6.3.1 GREENING IN POTATO TUBERS

Table 28 The level of greening on potato tubers visual assessment by varieties.

Variety	*Skin Colour*	*Type*	*Level of greening*
Agria	Yellow	Intermediate	0, 1, 2, 3, 4, 5
Alouette	Pink	Early	0, 1, 2, 3
Anna	Yellow	Intermediate	0, 1, 2, 3, 4, 5
Craigs Royal	Yellow	Early	0, 1, 2, 3, 4, 5
Dundrod	Yellow	Early	0, 1, 2
Majestic	Yellow	Main Crop	0, 1, 2, 3, 4, 5
Maxine	Pink	Early	0, 1, 2, 3
Mimi	Pink	Early	0, 1
Yukon Gold	Yellow	Early	0, 1, 2, 3, 4

All nine potato varieties studied exhibited a greening response after 14 days of exposure to light. The extent of greening was visually evaluated using a qualitative appearance scale ranging from zero to five (**Table 27**). In this study, potato varieties were categorised into two groups based on their skin colour: yellow and pink. Among the yellow-skinned varieties (Agria, Anna, Craigs Royal, Dundrod, Majestic and Yukon Gold), greening was clearly visible on the surface of some tubers and the full range of greening levels, from 0 to 5 was displayed (**Table 28**). Among pink-skinned varieties (Alouette, Maxine and Mimi), greening was less clear and difficult to distinguish due to the presence of pigments that mask the green colouration (**Figure 72**). Consequently, the recorded greening level only ranged from 0 to 3, with the variety Mimi only showing level 1 greening and the varieties Alouette and Maxine mainly showing greening levels at 1, 2 and 3. This suggested that the pigments present in pink-skinned potatoes may to some extent slow down the greening process and chlorophyll accumulation or a masking effect of the pigment is being evidenced.

Figure 72 Nine potato varieties before periderm disk were collected for chlorophyll and glycoalkaloid analysis.

6.3.2 THE RELATIONSHIP BETWEEN CHLOROPHYLL CONTENTS AND GREENING LEVELS

Changes in chlorophyll and carotenoid concentration in tubers was influenced by light. While all nine varieties showed increase in chlorophyll and carotenoid as the level of greening increase. The increase in and carotenoid significantly varied between variety ($P<0.001$) and greening level ($P<0.001$). While the increase in chlorophyll significantly different between greening level ($P<0.001$). At high greening level (4 and 5) they are not significantly different from each other (**Figure 73**).

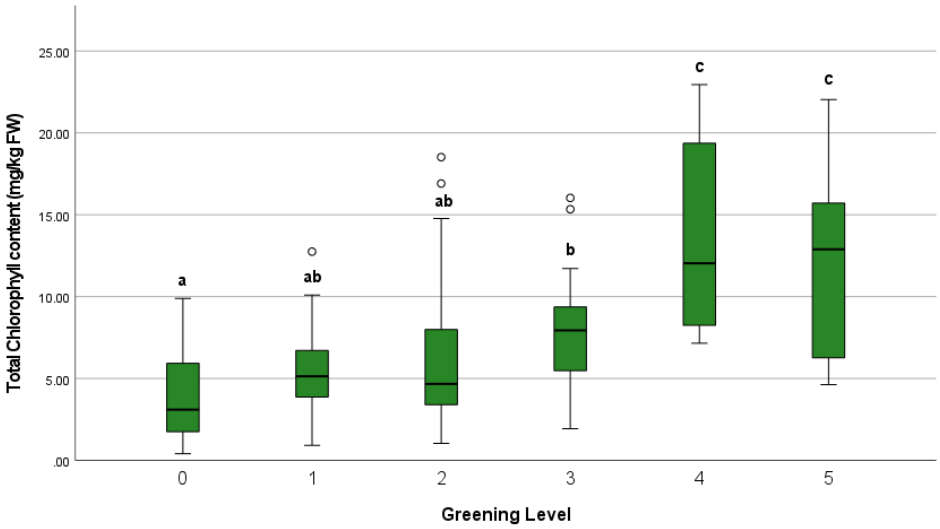

Figure 73 The distribution of the total chlorophyll content between different greening level.

A Pearson's correlation analysis showed that increasing in chlorophyll and carotenoid were highly correlated (R = 0.852, P<0.001). Whereas greening level and chlorophyll content were at some point correlated (r = 0.535, P < 0.001). Testing of tuber of the same variety revealed most clustered together showing a similar greening response with moderate significant difference between varieties (P < 0.05). Nevertheless, changes in chlorophyll contents may also influenced by the tuber maturity as the highest was observed in variety classified as early crop, Craigs Royal (2635 mg kg^{-1} FW).

In pink potato varieties, the results obtained from variety Alouette, Maxine, and Mimi suggested that the total glycoalkaloid varied and was not correlated with chlorophyll concentration as higher greening level. For example, variety Mimi, total PGA content was around 677.7 mg kg^{-1} FW although the level of greenness was at 0 and 630.8 mg kg^{-1} FW at green-1 (**Table 29** in the appendix).

In contrast, the results showed the similar trend among yellow/white varieties where the changes in total glycoalkaloid concentration increased linearly as greening level increased, except (**Figure 75**) variety Dundrod. For example, variety Agria total chlorophyll concentration increased relatively slow from level 0 to 2. Total glycoalkaloid concentrations increased significantly with level of greening and the total PGA of a very green tuber (score 5) was about 1900 mg kg^{-1} FW that nine times higher than the upper limit food safety.

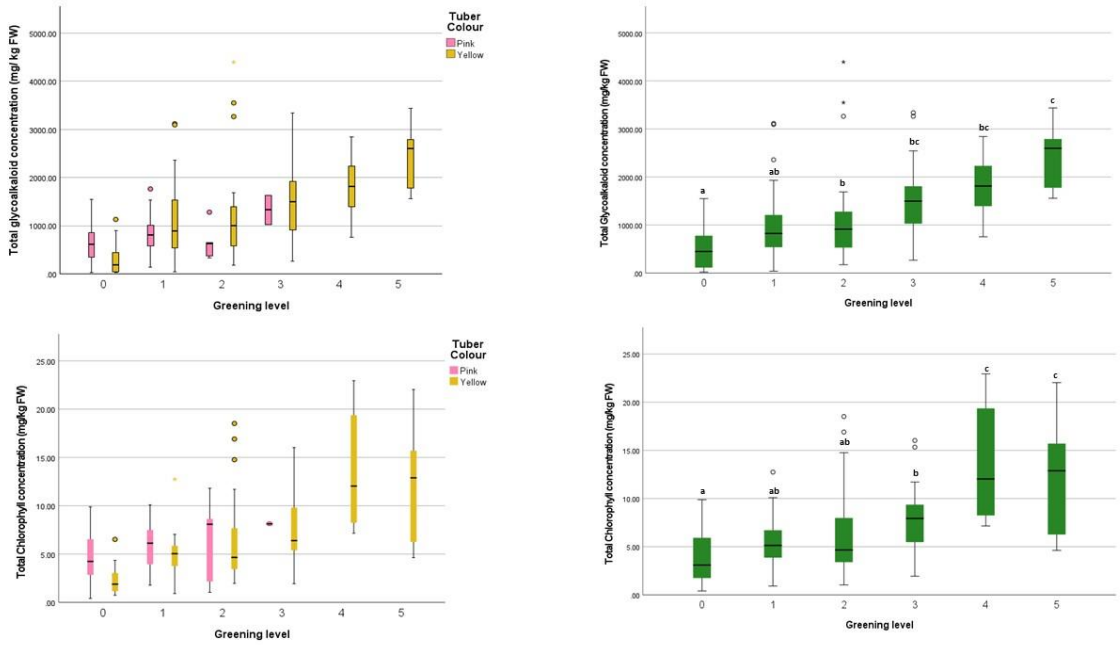

Figure 74 Mean of total glycoalkaloid content (mg/kg FW) at different greening levels by tuber colour.

For carotenoid concentration in greening potato tubers. When varieties were classified into two groups based on tuber skin colour. For pink varieties (Alouette, Maxine and Mimi), the correlation between chlorophyll and carotenoid increase were highly correlated (r=0.834, P < 0.001). Carotenoid contents are significantly difference between pink varieties (P<0.001), where Alouette appears to have highest carotenoid than the rest.

For yellow varieties (Agria, Anna, Craigs Royal, Dundrod, and Yukon Gold), the correlation between chlorophyll and carotenoid increase were highly correlated (r=0.629, P < 0.001). However, there were not statistically significant different between yellow varieties (P>0.05) (**Figure 74**).

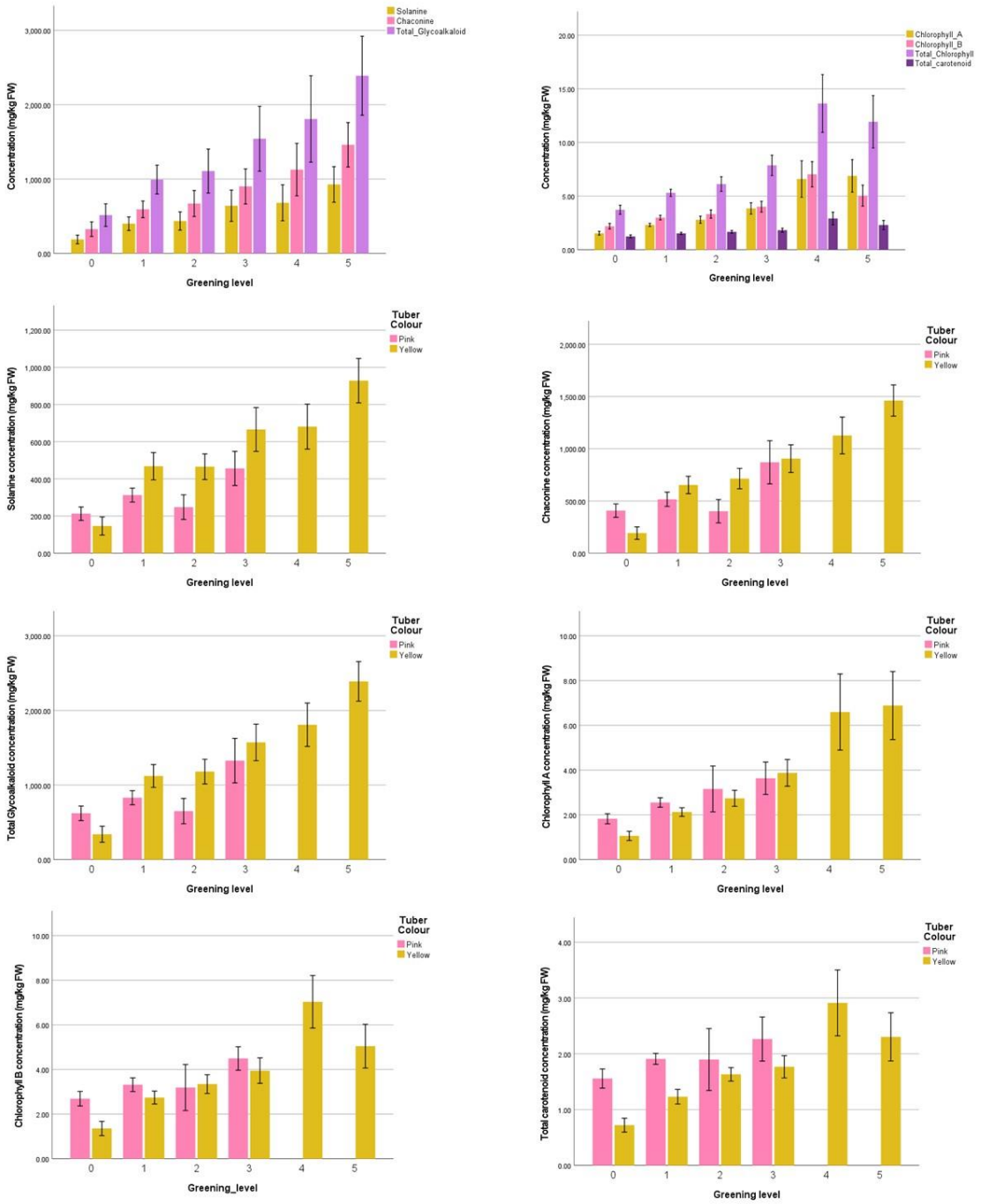

Figure 75 The distribution of chlorophyll A, chlorophyll B, total chlorophyll and total carotenoid concentration (mg/kg FW) between nine varieties by greening levels. The bar represents the means and the error bar shows ± 1 SE.

6.3.3 DETERMINATION OF CHACONINE AND SOLANINE

In addition to the experimental sample set, reference, quality control and blank samples were analyzed to provide high-level precision. The developed LC-UV/VIS method is both

sensitive and selective for α-solanine and α-chaconine with limit of quantitation (LOQ) values of 6 µg/mL and 4 µg/mL and limit of detection (LOD) of 2 and 1 µg/mL, respectively. These values were calculated based on the standard curve method; LOD = (3.3σ)/S and LOQ = (10σ)/S where σ is the standard deviation and S is the slope of the curve. The concentration value of the glycoalkaloid compound was calculated from the compound peak area which identified by plotting the unknown peaks against the standard calibration graph of the reference standards of α-chaconine and α-solanine (**Figure 76**). **Figure 76** shows the separation of standard α-solanine and α-chaconine peak on HPLC chromatograms where the retention time were approximately 5.42 and 6.17 min, respectively. However, each of tuber samples showed several unidentified peaks in addition to those for α-chaconine and α-solanine. There were minor differences in the pattern of peaks observed with different, but the chromatogram shown is typical of all the varieties consistently showed the first peak at about 9.7-9.8 min.

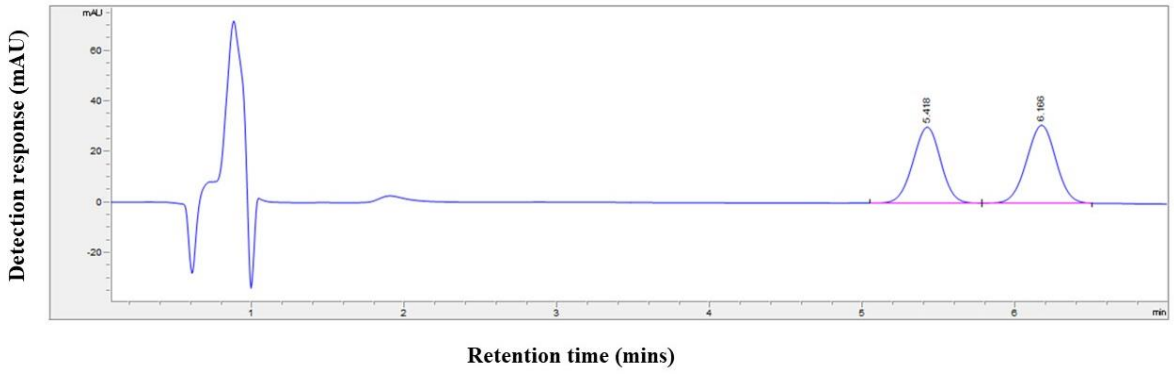

Figure 76 HPLC chromatograms of standard α-solanine and α-chaconine on a pentafluoro phenyl column (Supelco 2.1 m F5, 100 x 2.1 mm). Peaks; α-solanine (1) retention time 5.42 min, and α-chaconine (2) retention time 6.17 min.

The concentration of α-chaconine and α-solanine and total glycoalkaloid are shown in **Table 29**. A Pearson's correlation test showed that a high correlation exists between the increase in α-solanine and α-chaconine concentration for all varieties ($r = 0.96$, $p < 0.001$). In all tuber samples analysed, the concentration of α-chaconine was approximately one to two times greater than α-solanine. The total PGA concentrations in some varieties were significantly higher than the recommended food safety, even at low greening levels, such as, Craigs Royal (1453 mg kg^{-1} FW, level 2), Majestic (972 mg kg^{-1} FW, level 2), Dundrod (1153 mg kg^{-1} FW, level 1), Anna (1142 mg kg^{-1} FW, level 2), and Maxine (668.7 mg kg^{-1} FW, level

0) (**Table 29**). **Figure 74** shows the statistically significant differences between greening levels and total PGA content, while **Figure 77** illustrates the mean distribution and variation of α-chaconine, α-solanine and total PGAs concentration between nine potato varieties at different greening levels. Solanine and chaconine are highly correlated (correlation = 0.986) suggesting they tend to vary together in the potato samples. This is confirmed as the total glycoalkaloid is also highly correlated with both solanine and chaconine (correlation value close to 1).

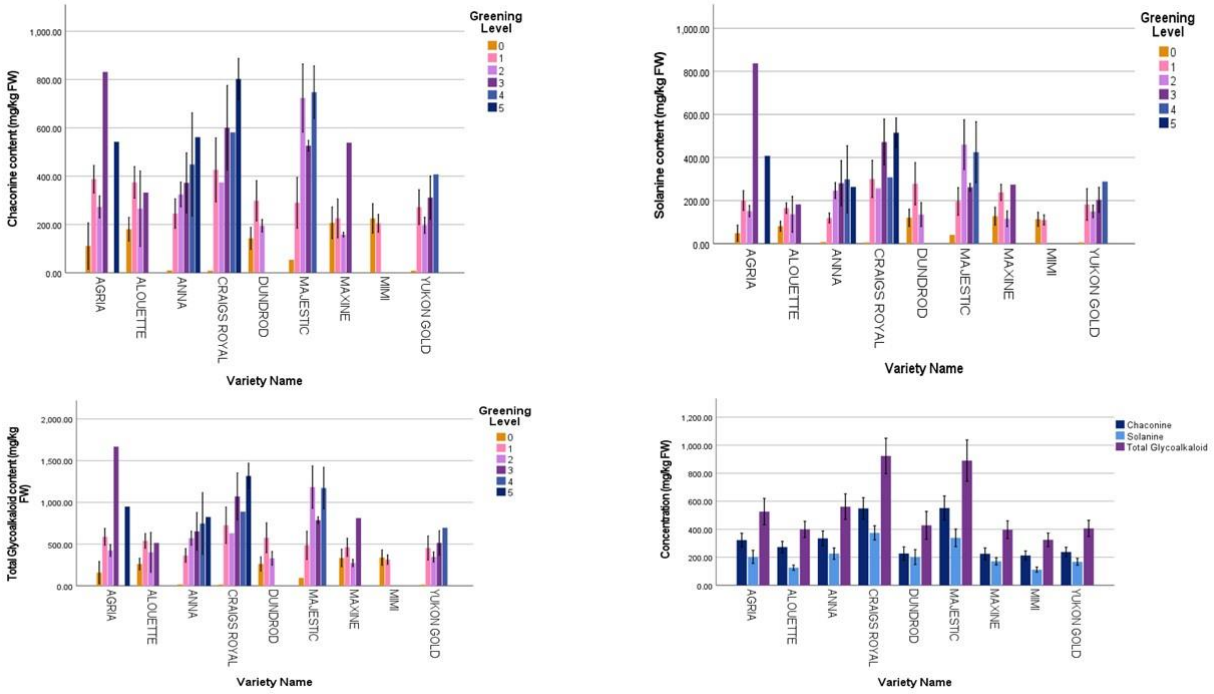

Figure 77 The distribution of chaconine, solanine and total glycoalkaloid concentration (mg/kg FW) between nine varieties by greening levels. The bar represents the means ± 1 SE.

6.3.4 RELATIOHSIP BETWEEN CHLOROPHYLL AND TOTAL GLYCOALKALOID CONTENTS

In general, the total potato glycoalkaloid (PGA) content was found to be much higher in the tubers with high levels of greening compared to less green tubers ($F_{(5, 105)} = 4.9049$, $p < 0.001$), and the content varied significantly among different varieties ($F_{(8, 105)} = 3.6531$, $p < 0.001$). However, there was considerable variation between tuber samples ($F_{(27, 105)} = 1.5902$, $p > 0.05$). A large variation between the same varieties masked differences between varieties (same variety different plot number / one variety two-plot numbers). Regardless of the variety, the increase in PGA content did not show a direct proportion relationship with greening level and chlorophyll content. A Pearson's correlation analysis showed no significant correlation

between chlorophyll and PGA content ($r = 0.052$, $p > 0.05$). These results indicate that there is a complex relationship between greening, chlorophyll concentration and PGA content in green potato tubers, and suggesting other factors may also be involved such as tuber skin colours (**Figure 78** and **Figure 79**).

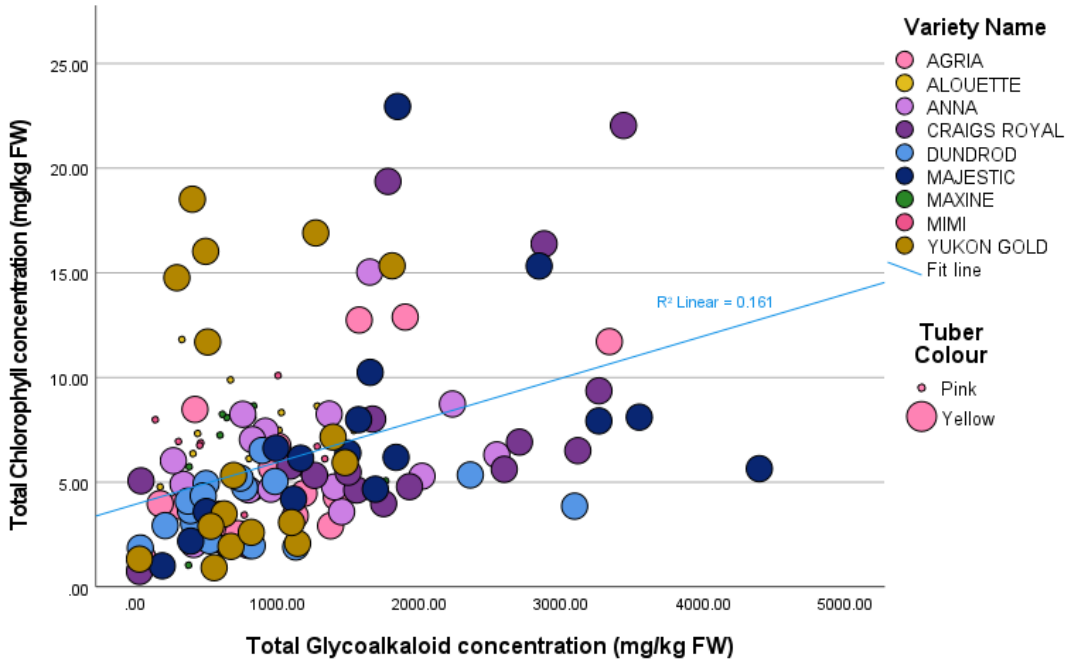

Figure 78 The relationship between glycoalkaloid and chlorophyll content by varieties and tuber skin colours (pink and yellow).

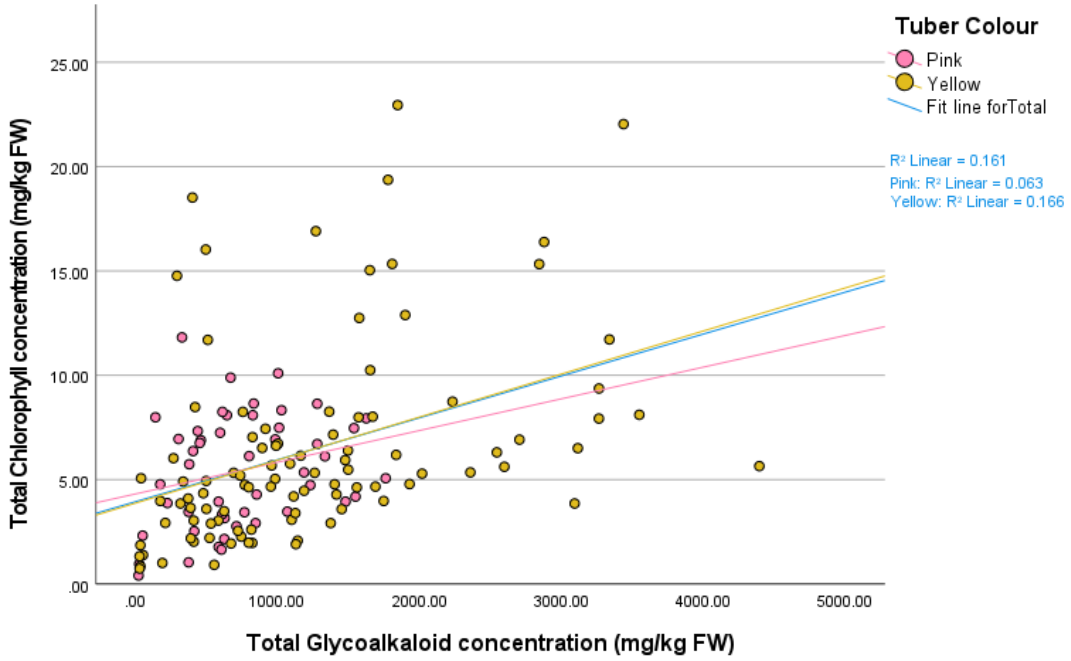

Figure 79 The relationship between glycoalkaloid and chlorophyll content by the colour of tuber.

6.4 DISCUSSION

It has long been recognised that exposure to light causes potato tuber to turn green and potato glycoalkaloid build up. Exposure to light in the field can occur when potato tubers protrude out from the ground. This is sometimes associated with heat necrosis, a hollowing of the centre of the tuber, caused by exposure to high temperatures. In consideration of field greening, it can be prevented by avoiding light and selecting resistance varieties along with good management practices. Potato tubers for the fresh market should be graded for external disorders and diseases, especially green tubers, the skin should be thin, smooth, bright and blemish free. Tubers with such skin are required in most markets selling washed and prepacked potatoes because consumers make a preliminary selection based on tuber appearance. While the changes in tuber colour during greening decrease tuber appeal to consumers and industry, discrimination against greened tubers due to the perception of potato glycoalkaloid toxicity is subject to debate. Even though the increase in glycoalkaloid concentration is not directly linked to potato greening or toxicity, it can be varied depending on environmental and genetic conditions. yellow/white varieties. Therefore, for potato fresh markets, European Food Safety (2020) suggested that because of greening, glycoalkaloids should be utilised as a toxicity indicator, and the relationship between greening and glycoalkaloid concentration should be investigated to see if incidental potato greening poses a health risk.

In addition to the above suggestions, this study investigated the relationship between greening in potato tubers and glycoalkaloid concentration in different varieties. Light exposure increases the total glycoalkaloid contents of the tuber. An increase in both α-chaconine and α-solanin has been reported by numerous studies (Friedman and McDonald, 1999; Grunenfelder *et al.*, 2006a; Okamoto *et al.*, 2020). HPLC analysis has been developed as a standard method for separation and quantification of glycoalkaloids in many plant species (Matsuda *et al.*, 2004). In this study, we improved the chromatographic separation of α-chaconine and α-solanine in extracts of potatoes. Several lines of evidence are presented to support the core finding of this study: potato glycoalkaloids are naturally available depending on variety/cultivar, and potato skin colour is an important determinant of greening sensitivity.

Potato glycoalkaloids (PGAs) are naturally available depending on variety/cultivar. The total glycoalkaloid content of non-green potatoes was found in all varieties. Among the common steroidal alkaloids in potatoes, α-chaconine was the most abundance in the potato skin followed by α-solanine. There are, however, other glycoalkaloids found in potatoes. In this study, HPLC chromatogram of each tuber samples showed several unidentified peaks in addition to those for α-chaconine and α-solanine (). High temperature, which was used during sample re-concentration step, might have increase the solubility of many compounds resulting in unknown compounds being extracted. According to Hossain *et al.* (2015) study indicates that extraction yield of total steroidal alkaloids was affected by temperature. Thus, it is not surprising, given that rising temperatures increase the solubility of many compounds. Although without a definitive identification of the substances detected, it is difficult to make inferences about unusual/uncommon glycoalkaloids in certain potato cultivars.

The value of the total PGA contents found in this study was extremely high in very green potato tubers. However, consider the variation that was found between samples based on tuber skin colours. In pink varieties, the total glycoalkaloid varied and was not correlated with chlorophyll concentration as higher greening level. It also accumulated less chlorophyll that yellow varieties as less high level of greening was found. This could be explained by the present of anthocyanin pigments in the periderm of pink varieties that competing with chlorophyll and can limit the accumulation of PGAs (Tanios *et al.*, 2020). Another factor that may be responsible for the spike in variation is the damage/disease or poor postharvest handling. In this study found that in yellow varieties at higher greening levels, chlorophyll content varied. This is similar to the result from Grunenfelder *et al.* (2006a) found for cultivars Yukon Gold and White Rose at a particular green scale; the amount of chlorophyll content varies somewhat

and it is not the sole determinant of how green the tuber looks. Whereas Tilahun *et al.* (2020b) found that total glycoalkaloids were accumulated at the late phases of greening and were associated with the highest chlorophyll concentration. Abreu *et al.* (2007) also found that the differences between cultivars are greater than between plants of the same cultivar under different conditions. Nonetheless, further research is required to fully understand and characterise the relationship between total glycoalkaloids and chlorophyll concentration at the high level of green. And whether genetic conditions (variety) influence total potato glycoalkaloid and level of greening and consumption of these varieties would cause glycoalkaloid intoxication in consumers.

6.5 CONCLUSION

This study investigated the relationship between greening and glycoalkaloid content in nine different potato varieties by measuring chlorophyll and glycoalkaloid accumulation using high performance liquid chromatogram (HPLC) and a UV/Visible spectrometer. The association between greening and potato glycoalkaloid (PGA) accumulation was explored in nine distinct potato varieties by measuring PGA and chlorophyll concentrations using high performance liquid chromatography (HPLC) and a spectrometer, respectively. The results indicated that while PGA and chlorophyll accumulation occur simultaneously, the increase in PGA concentration is not proportional to the greening level and chlorophyll content. Interestingly, certain potato varieties, including Dundrod, Anna, and Craigs Royal, have exceptionally high PGA contents even at low levels of greening. On the other hand, non-green potatoes were found to have PGA content within the safe limits. The observed variation in PGA contents among tuber samples and even within the same variety indicated that PGA could be influenced by multiple factors. The presence of the anthocyanin pigment in pink-skinned varieties was identified as an influential factor in modulating the greening response and affecting PGA accumulation. These findings highlight the complexity of the relationship between greening, PGAs concentration, and pigments in potato tubers. Further research to fully understand these interactions is needed. However, measuring the PGA content requires a destructive approach that is time-consuming and requires expensive chemicals and equipment. Therefore, there is a need for a non-destructive, rapid, cost-effective, and reliable approach to estimate the PGA contents in potato tubers.

CHAPTER 7. GENERAL DISCUSSION

This chapter presents a summary of the research findings from this thesis on the use of unmanned aerial vehicles (UAVs)-based imagery for breeding and potato production. It highlights the importance of combining UAV-based imagery with plant features evaluated from ground-measured techniques. Overall, the application of UAVs and remote sensing technologies has significantly enhanced the field of plant phenotyping and crop production, providing exceptional spectral, spatial, and temporal resolution. These technological improvements enable the comprehensive capture of crop traits and offer a multi-directional perspective, which is crucial for predicting crop production, crop monitoring, and disease detection.

The main objective of this thesis was to investigate the feasibility of UAV-based remote sensing technology in the context of crop monitoring, plant health evaluation, and yield estimation. This study provides a comprehensive analysis of the accuracy of yield prediction and disease detection by utilising imaging sensors, including RGB and multispectral, in conjunction with advanced data extraction and analysis techniques. Chapter 2 provided a full discussion of the methodologies employed, which include statistical analysis and machine learning techniques such as Partial Least Squares Regression (PLSR), Random Forest (RF), and K-means clustering. The experiment demonstrates that multispectral sensors on UAVs have significant efficacy in potato varietal study trials and large-scale production. Significant discoveries presented in Chapters 3, 4, and 5, coupled with the identification of the toxin present in greening potato tubers, are discussed in Chapter 6. These findings highlight the important effect of this study on the field of potato monitoring and phenotyping. The Random Forest algorithm demonstrates a significant advantage in improving the timely identification of plant diseases, improving yield prediction, and supporting crop monitoring. Additionally, the study found that UAV flights recorded during the tuber filling stage provide the most suitable plant information for both yield estimation and disease detection models. This significant advancement emphasises the importance of UAV-based remote sensing in improving potato phenotyping. It provides accurate disease diagnosis, yield predictions, and continuous crop monitoring throughout the crop growth cycle.

Furthermore, this study has established a connection between phenotyping and genotyping studies, particularly through the varietal studies of potatoes. By integrating all the chapters in this study, the research not only opens new avenues for collaborative research, as highlighted by the findings related to the European potato study (Chapter 5), it also demonstrates that in-field phenotyping may accurately detect important crop characteristics in known genotypes when challenged to environmental stresses. Nevertheless, it is essential to recognise the constraints and challenges associated with handheld crop sensors, particularly in terms of data accuracy and analysis reliability, in order to facilitate future progress.

Overall, the present thesis provides a significant contribution to the field of potato phenotyping study and breeding programmes by presenting novel methodology and interdisciplinary approaches. The acquired insights provide a foundation for future progress in potato varietal study trials, with a special focus on the application of UAV remote sensing in phenotyping studies. In conclusion, this study offers a strong basis for future investigations on potato phenotyping and genotyping, emphasising its significance and capacity to improve crop production techniques and the integration of machine learning tools for further automation.

7.1 WHAT DO WE KNOW

This section of the thesis explores the utilisation of handheld crop sensors and the shift toward more advanced unmanned aerial vehicles (UAVs) with remote sensing technology for the purpose of this study. Handheld crop sensors, which are primarily used by field inspectors to get a general idea of how crops behave in the field. For more than five decades, handheld sensors such as RapidSCAN, GreenSeeker, and ceptometers have played a crucial role in field inspection by facilitating the collection of first data on crop behaviour (Singh *et al.*, 2006; Kipp *et al.*, 2014). Nevertheless, these technologies have difficulties in gathering and retaining data. For example, the GreenSeeker, although small and easily transportable, does not have the ability to save data, necessitating the manual recording of NDVI readings following each measurement. On the other hand, RapidSCAN demonstrates advancements by incorporating the capacity to store a maximum of 25,000 readings and do statistical calculations for NDVI and NDRE indices. This highlights the progression and variety of capabilities in sensor technology. The study from Kipp *et al.* (2014) compared three different active canopy sensors (GreenSeeker, CropCircle and Active Flash Sensor) and found that varying devices have variable results depending on sensors and spectral indices. As we mention many times throughout this thesis, that the conventional method in plant phenotype study is labour

intensive and time consuming, therefore, we need a more robust and accurate method that can accelerate field data collection.

The limitations of portable sensors highlight the need for more resilient and precise techniques for accelerating the data collection in the field. The challenges can be addressed by combining UAV technology with multispectral photography, a technique that has been demonstrated in multiple studies to improve agricultural research by quickly providing high-resolution data. The emergence of UAVs and remote sensing technology offers an innovative and non-destructive data collection approach, providing the ability to gather images with outstanding spatial and temporal accuracy (Zhou *et al.*, 2017).

In addressing the labour-intensive and time-consuming nature of conventional plant phenotyping methods, this study explores the feasibility of using low-altitude UAVs equipped with multispectral sensors and RGB cameras. This innovative approach not only aims to accelerate data collection but also to validate data extracted from UAV-based imagery against traditional manual approaches. The study's comprehensive experimental design, encompassing over 290 potato varieties under both organic and conventional management practices, presents a unique challenge in manual data collection, further emphasising the potential of UAV technology in agricultural research. Multispectral imaging has been utilised in the field of agriculture, as demonstrated in various studies on many crops and for various purposes (Zaman-Allah *et al.*, 2015; Tripodi *et al.*, 2018; Watt *et al.*, 2020; Zhou *et al.*, 2023).

A well-defined methodology, including sensor calibration and the use of user-friendly software for data processing, is crucial for ensuring the reliability of UAV-captured data. This study employed R language software in R Studio and TBC software for data extraction and analysis, highlighting R's accessibility and support from a vast online community (Team, 2021). Using R to extract drone-based imagery compared to other studies has been proved in this study to be viable (Matias *et al.*, 2020). Additionally, machine learning algorithms such as Random Forest, K-means clustering, Partial Least Square Regression, and Decision Tree are utilized for yield prediction and disease detection, demonstrating the versatility and effectiveness of these methods in potatoes and previously shown across various crops such as soybean (Bolton and Friedl, 2013), wheat (Pantazi *et al.*, 2016), olive tree (Calderón *et al.*, 2015), and tomato (Zhu *et al.*, 2018).

7.2 A RESULT FROM THIS RESEARCH

The results of this study highlight the effectiveness of utilising data obtained from integrated systems to estimate plant characteristics and improve agricultural productivity. A comprehensive regression analysis was conducted to compare plant attributes measured in the field with those inferred from data obtained from unmanned aerial vehicles (UAVs), using both R and TBC methodologies. The precision of estimating plant height using multispectral photography was 68%, however 3D plant models enhanced this percentage to 78%. Similarly, the accuracy in measuring vegetation indices varied and ranged from 69% to 72%, illustrating the efficiency of UAV methodologies.

Crucially, this study highlighted the importance of feature selection in the process of model development. Yield prediction models that focused on plant traits closely correlated with actual yields showed significantly better fit, with R^2 values ranging between 0.807 and 0.942. This contrasted with models that incorporated all available parameters, where R^2 values range from 0.412 to 0.645. Moreover, the research demonstrated the relevance of selecting the most appropriate data acquisition times that correspond with different plant growth stages. Data from the tuber filling stage proved most effective for training models and aimed at yield prediction and disease detection, with accuracy improving from 50% at the maturity stage to 78% at the tuber filling stage for disease detection models and from 43% to 78% for yield prediction models.

In comparison with other models such as regression and decision trees, the Random Forest (RF) model achieved a remarkable accuracy rate of 94.2%, outperforming other models in the investigation of potato yield prediction. The RF demonstrated higher computational efficiency and predictive power in forecasting potato crop yields and disease classification. Although other studies employing different methodologies have reported varied results, this research indicates potential areas for accuracy improvement. For instance, integrating vegetation indices with a stacking algorithm approach or employing deep learning architectures like Faster R-CNN for feature extraction could significantly enhance model precision (Wishart *et al.*, 2014; Yang *et al.*, 2022; Rajpoot *et al.*, 2023; Demilie, 2024).

Nevertheless, the selection of specific traits for accurate crop or yield prediction models requires careful consideration, especially given the diversity in canopy characteristics among varieties. This work confirms that conducting phenotyping experiments in the field is crucial for discovering important characteristics in established genotypes when they are subjected to

environmental stress. This highlights the significant impact that UAV-based remote sensing can have on agricultural research and practical applications.

7.3 OPPORTUNITIES AND CHALLENGES

The implementation of unmanned aerial vehicles and remote sensing technologies in potato phenotyping indicates a significant progression in agricultural research and operation. However, several opportunities and challenges emerge from this research due to the diversity of potato varieties in the field. On one hand, it provides potential for comprehensive phenotypic examination, allowing researchers to study a wide range of plant features and responses across various management practices. On the other hand, this diversity increases complexity resulting from the limitations associated with managing a large quantity of variations. The methodology employed in this study, which involved planting only two replicates for each variety under both organic and conventional management fields, highlighted the difficulties in conducting a comprehensive statistical analysis due to the small number of data points available. A key requirement was identified for a more complete dataset, which would improve the accuracy of crop prediction models for both yields and diseases. This highlights the necessity for more extensive data collection with more repetitions to enhance the precision of crop prediction models, especially for yield estimation and disease identification.

CHALLENGES IN DATA COLLECTION AND FUTURE DIRECTION

Although the methodology used in this study works well with potatoes across varieties, key challenges remain in applying these methods across different crops. One of the primary challenges is the variability in canopy structure, growth patterns, and growing conditions, such as greenhouse and field settings. Unlike uniform and vertically structured growth crops like cereals, potatoes have complex stems in the form of a shrub, which leads to difficulty extracting accurate canopy traits using photogrammetry-based methods (de Jesus Colwell *et al.*, 2021). Weed, overlapping canopy, wind, and shadow in the field also created false positives in crop traits extraction, such as plant count, canopy area, heights, and vegetation indices (Xie and Yang, 2020). Additionally, the UAV-based image approach has to take soil and ridge effects into account, as they can interfere with reflectance. These differences in crop parameters can affect measurement accuracy and model performance. To address these challenges, there is a need to develop a standardised analysis approach that can be adapted to a wide range of crops and growing conditions.

A key point from this research highlights the vital role of both data quantity and quality for improving precision agriculture. During the study of the Ecobreed project that featured fewer potato varieties but was planted with larger plots and more replicates. This facilitated a more precise analysis, particularly for supervised classification in machine learning, which requires a significant amount of data for model improvement. The complex nature of agricultural research, especially in disease detection and management, requires an augmentation in the quantity of replicates and the duration of data collection in order to develop more precise models (Liebisch *et al.*, 2015). This demonstrates that improving data collection, including both the number of replicates and the duration of monitoring, is crucial for developing precise and practical models. Nevertheless, there is also a need for imaging techniques and data processing algorithms to account for differences in crop architectures and cultivation conditions. To handle the complexity and variability in phenotyping studies, 'big data' technology in agriculture is a trend among researchers due to its potential to handle large amounts of data and powerful processing capacity (Andrivon, 2017; Sourav and Emanuel, 2021).

Moreover, key trends in agricultural research include the increasing use of machine learning and deep learning techniques, as well as the application of the Internet of Things (IoT) and hyperspectral imaging. Researchers are focusing on developing cost-effective and robust technologies with sophisticated algorithms to improve crop yield estimation and phenotyping studies. There is a trend toward developing rapid real-time and field image processing using image devices that incorporate artificial intelligence such as machine learning (ML), deep learning, and machine vision technology (Condran *et al.*, 2022; Song *et al.*, 2022; Shin *et al.*, 2023). Effective use of data required interdisciplinary work that bridged the gap between computer science and agricultural expertise (Condran *et al.*, 2022). However, the accuracy and uncertainty of ML prediction are dependent on several factors, including the quality of the data (Yang *et al.*, 2024). Therefore, future research should focus on not only choosing the right model but also data collection, data processing, and handling using applications with less complicated and user-friendly interfaces for all users. This is crucial for the implementation and utilisation of this technology in the agricultural industry.

PRACTICAL BENEFITS FOR RESEARCHERS AND FARMERS

One of the most significant challenges in applying remote sensing technology to the real field conditions and precision agriculture is associated with the cost of sensor technology.

Although this technology provides significant potential to improve efficiency and productivity, its high cost is still considered unaffordable for small-scale farmers (Lajoie-O'Malley *et al.*, 2020; Nduku *et al.*, 2023). It may be more affordable for many farmers to hire the field inspection for an overview of the farm during the critical growth stage of potatoes. The ability to early detect disease occurrence and estimate the best harvesting time is crucial to maximising potato production (Zhang and Kovacs, 2012). Although experienced farmers can identify the differences in the potato leaves visually (Chakraborty *et al.*, 2022), technological solutions offer more accurate and fast observations. However, as the technology develops, the price is expected to drop. Nowadays, the rapid development of low-cost UAVs is making remote sensing more accessible for small-scale crop production. Although these systems may lack the extensive functionalities of advanced research-grade equipment, they offer a practical and cost-effective alternative for farmers to assess field variability and can be served as a decision support system. With technological advancements and reduced prices, wider adoption of UAVs and remote sensing becomes increasingly feasible, creating a possibility to integrate advanced equipment with traditional farming practices into a complete operational production system.

INTEGRATION OF TECHNOLOGY AND SUSTAINABILITY

The integration of advanced technology with traditional farming knowledge is crucial for optimising the advantages of precision agriculture. UAVs, along with remote sensing technology, improve marketable yield by reducing crop loss through early detection of diseases and stresses. It also enhances efficiency in input applications, such as fertilisers and pesticides, leading to both economic and environmental benefits. For agronomics, UAV-based phenotyping provides detailed insights into varietal performance and crop resilience under varying conditions, enabling better breeding and management strategies. Remote sensing can detect disease symptoms prior to their visibility to the naked eye, which allows for early decision-making and action. Data obtained by UAVs can also assist farmers in identifying the optimal harvest time, hence enhancing crop quality and productivity. This technique promotes more sustainable agricultural practices by minimising synthetic chemical application and enhancing resource usage. Early crop detection and decreased total chemical usage result in cost savings for farmers and promote environmental sustainability. Agronomists and farmers will require training to operate the UAV, analyse data, and interpret data effectively. Nevertheless, advancements in user-friendly software are diminishing this obstacle, making the technology more accessible for non-specialists.

Overall, the existing gaps include the need for cost-effective and flexible applications that integrate technological advancements with traditional farming expertise. This combination promises to enhance productivity and disease management, leading precision agriculture towards sustainable practices capable of meeting the demands of food production. This approach outlines a path forward that could significantly impact global food security and farming efficiency. In summary, addressing these challenges while leveraging potential would accelerate the implementation of UAV-based precision agriculture techniques, thereby significantly enhancing global food security and sustainable farming practices.

7.4 CONCLUSION

This thesis conducted a comprehensive investigation into the feasibility of low-altitude unmanned aerial vehicle (UAV)-based remote sensing technologies for potato breeding and production. The field of potato phenotyping has been significantly advanced through a comprehensive investigation that incorporates RGB and multispectral imaging sensors, as well as advanced data extraction and analytical techniques. These techniques include statistical analysis and machine learning approaches such as Partial Least Squares Regression, Random Forest, and K-means clustering. This thesis also provides a comprehensive explanation of the procedures and data analysis techniques used to make important findings on predicting potato yield and detecting diseases. Additionally, it covers the identification of toxins in greening potato tubers. The results highlight the efficacy of UAVs equipped with multispectral sensors in the execution of varietal research and the improvement of large-scale production. The Random Forest algorithm has particularly demonstrated its efficacy as a prominent approach for disease detection, yield estimation, and crop monitoring throughout their growth cycle. Furthermore, this study emphasised the most advantageous timing for unmanned aerial vehicle (UAV) flights in order to acquire crucial crop characteristics during the tuber filling phase, hence augmenting the precision of the model.

In summary, this thesis showcases the practicality of using unmanned aerial vehicles (UAVs) for remote sensing in plant studies. Additionally, it makes a substantial contribution to the disciplines of potato phenotyping and genotyping and the benefit of integrated sensing approaches. By forging a strong correlation between phenotyping and genotyping investigations, it creates new opportunities for future research, offering the potential for additional progress in crop production methods and disease control strategies.

APPENDIX A

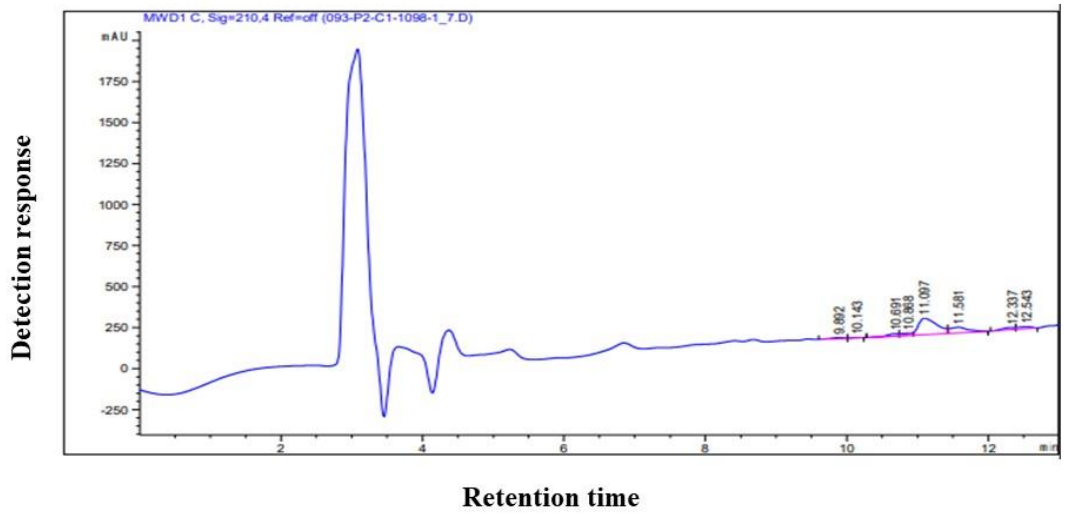

Figure 80 HPLC chromatogram of green potato peel sample, the peaks correspond to the presence of glycoalkaloids with retention times indicating the separation of α-chaconine and α-solanine.

Table 29 Average concentrations of solanine, chaconine, total glycoalkaloid, chlorophyll a, chlorophyll b, total chlorophyll, and total carotenoid contents (mg/kg fresh weight) of nine potato varieties at different level of greening (0-5 scale).

VARIETY NAME	N	LEVEL OF GREEN	SOLANINE	CHACONINE	TOTAL GLYCOALKALOID	CHLOROPHYLL-A	CHLOROPHYLL-B	TOTAL CHLOROPHYLL	TOTAL CAROTENOID
AGRIA									
	1	0	23.4	29.9	53.3	0.7	0.7	1.4	0.6
	4	1	343.7 ± 172.0	684.5 ± 241.8	1028 ± 409.6	2.7 ± 1.6	4.3 ± 2.6	7.0 ± 4.1	2.1 ± 0.9
	10	2	303.0 ± 159.5	545.6 ± 287.1	848.7 ± 442.4	1.7 ± 0.7	2.6 ± 1.3	4.3 ± 1.9	1.7 ± 0.3
	1	3	1675	1663	338.1	4.4	7.3	11.7	3.0
	1	5	815.5	1084	1900	6.9	6	12.9	3.6
ALOUETTE									
	7	0	162.6 ± 120.6	360.6 ± 256.2	523.3 ± 364.1	2.2 ± 1.1	3.3 ± 1.5	5.5 ± 2.6	2.1 ± 0.7
	6	1	330.7 ± 117.9	784.8 ± 319.4	1079 ± 428.8	2.5 ± 1.0	2.9 ± 1.1	5.4 ± 2.0	2.1 ± 0.5
	2	2	273.2 ± 232.9	530.8 ± 440.6	804.0 ± 673.5	5.0 ± 1.2	5.3 ± 1.0	10.2 ± 2.2	3.2 ± 0.1
	1	3	364.5	663.9	1028	4.4	4.0	8.3	2.7
ANNA									
	1	0	15.4	19.2	34.6	0.4	0.4	0.9	0.4
	3	1	238.7 ± 77.4	490.8 ± 209.2	729.5 ± 283.7	2.4 ± 1.4	2.1 ± 1.4	4.6 ± 2.5	1.2 ± 0.4
	4	2	493.4 ± 145.2	648.7 ± 203.3	1142 ± 333.8	2.6 ± 0.6	2.6 ± 1.0	5.1 ± 1.6	1.1 ± 0.2
	5	3	562.0 ± 465.9	744.5 ± 553.9	1306 ± 1008	2.8 ± 0.8	3.3 ± 0.7	6.2 ± 1.3	1.5 ± 0.4
	2	4	597.4 ± 440.6	897.3 ± 603.6	1494 ± 1044	3.6 ± 2.9	4.9 ± 2.6	8.5 ± 0.3	2.2 ± 0.1
	1	5	527.6	1123	1650	7.7	7.3	15.0	2.9
CRAIGS ROYAL									
	1	0	11.5	17.8	29.3	0.3	0.5	0.7	0.3
	6	1	601.1 ± 419.3	852.3 ± 649.5	1453 ± 1062	2.0 ± 1.0	2.7 ± 1.6	4.7 ± 1.6	1.2 ± 0.8
	1	2	514.2	747.8	1262	3.5	1.8	5.3	1.1
	3	3	943.6 ± 368.3	1201 ± 605.9	2145 ± 973.7	4.6 ± 1.8	3.1 ± 0.7	7.6 ± 2.0	1.7 ± 0.4

	1	4	616.0	1162	1779	10.0	9.4	19.4	2.9
	5	5	1030 ± 305.3	1605 ± 379.9	2635 ± 683.3	6.7 ± 4.9	4.4 ± 2.8	11.1 ± 7.7	1.9 ± 1.1
DUNDROD									
	6	0	241.0 ± 192.4	284.9 ± 222.3	525.9 ± 413.7	1.6 ± 0.6	1.8 ± 1.3	3.4 ± 1.8	0.9 ± 0.4
	8	1	556.6 ± 551.5	596.5 ± 468.4	1153 ± 1011	1.8 ± 0.7	2.5 ± 0.5	4.3 ± 1.0	0.9 ± 0.5
	2	2	270.8 ± 156.9	389.7 ± 379.9	660.5 ± 229.8	1.4 ± 0.6	2.0 ± 1.5	3.5 ± 2.1	1.1 ± 0.4
MAJESTIC									
	1	0	82.0	107.5	189.5	0.2	0.8	1.0	0.1
	4	1	392.6 ± 253.6	579.6 ± 209.3	972.2 ± 334.8	2.2 ± 0.9	2.4 ± 1.1	4.5 ± 2.0	1.2 ± 0.4
	7	2	920.3 ± 609.9	1447 ± 371.8	2368 ± 672.9	3.1 ± 0.9	3.4 ± 0.8	6.5 ± 1.6	1.6 ± 0.4
	2	3	523.3 ± 49.2	1052 ± 30.8	1575 ± 55.4	3.8 ± 0.8	4.5 ± 1.9	8.3 ± 2.7	1.8 ± 0.8
	2	4	849.6 ± 398.2	1495 ± 153.4	2345 ± 352.5	9.8 ± 3.7	9.3 ± 1.7	19.1 ± 5.4	4.4 ± 1.5
MAXINE									
	6	0	255.2 ± 201.9	413.5 ± 318.6	668.7 ± 513.3	1.3 ± 0.9	1.7 ± 1.0	3.0 ± 1.9	1.0 ± 0.4
	5	1	475.5 ± 164.0	450.7 ± 358.3	926.2 ± 480.9	3.3 ± 0.5	4.2 ± 1.2	7.5 ± 1.4	1.8 ± 0.4
	3	2	231.4 ± 124.8	316.6 ± 31.3	548.0 ± 150.6	2.1 ± 2.0	1.8 ± 1.7	3.8 ± 3.8	1.0 ± 0.5
	1	3	548.3	1077	1625	2.9	5.0	7.9	1.9
MIMI									
	7	0	227.5 ± 169.5	450.2 ± 320.9	677.7 ± 484.7	1.9 ± 1.0	2.9 ± 1.5	4.8 ± 2.4	1.6 ± 0.7
	10	1	220.7 ± 144.2	410.1 ± 229.0	630.8 ± 349.6	2.2 ± 1.0	3.1 ± 1.6	5.4 ± 2.5	1.9 ± 0.5
YUKON GOLD									
	1	0	11.9	17.2	29.1	0.6	0.8	1.3	0.4
	3	1	364.1 ± 251.8	543.7 ± 248.9	907.8 ± 497.9	1.8 ± 1.2	2.3 ± 1.6	4.2 ± 2.7	1.0 ± 0.6
	8	2	300.3 ± 160.3	393.0 ± 188.4	693.3 ± 343.1	4.0 ± 3.5	5.2 ± 3.9	9.2 ± 6.9	2.0 ± 1.2
	4	3	407.3 ± 230.5	622.4 ± 357.1	1030 ± 586.4	4.6 ± 4.3	4.3 ± 3.9	8.8 ± 7.9	1.8 ± 1.3
	1	4	577.4	815.0	1392	2.7	4.4	7.2	1.3

APPENDIX B

PLS-DA score plot with K-means clustering (**Table 30**) shows four groups of potato varieties based on vegetation indices where we have standardised and transformed into the new PLS-DA scores to each sample based on its position. Based on the senescence rating score system (**Table 5**) and the results from PLS-DA clustering, we can interpret the health conditions of each cluster in relation to plant senescence stages.

Table 30 Mapping PLS-DA clustering to senescence rating score system.

PLS-DA cluster	Average senescence score	Interpretation
1	2.00	Some browning and yellowing of leaves
2	2.52	Around 75 percent of the greening material remains
3	1.67	Overall excellent health with some browning and yellowing of leaves
4	2.32	Some browning and yellowing of leaves

APPENDIX C

Conference attended

1. Newcastle university PGR conference 2020
2. European Association for Potato Research (EAPR), PGR conference 2020, Krakow, Poland: Abstract accepted but did not attend due to the visa issue. Supervisor presented on my behalf.
3. Ecobreed organic breeding conference, Ljubljana, Slovenia: Abstract present title "Exploring High Throughput Tools for Decision Making in Potato"
4. 19th INWEPF Steering Meeting and Symposium (Smart farming innovations and efficient water management practices), 12-13 December 2024, Thailand.

Presentation

1. Presented the used of phenotyping tools for potato varietal trail at Advanced phenotyping training event: August 4-5, 2022, at Nafferton Farm, Northumberland, NE43 7XD
2. Presented greening in potatoes research at Spud Fest: December 15, 2022, at The Boiler House, Newcastle University.

Upcoming conference

1. Abstracted accepted for oral presentation at the 4th world irrigation forum (WIF), 7-13 September 2025 in the theme of "technology and modernisation in agricultural sector towards food security".

Publications

1. Book chapter title "Optical Imaging Resources for Crop Phenotyping and Stress Detection"
   Waiphara P, Bourgenot C, Compton LJ, Prashar A. Optical Imaging Resources for Crop Phenotyping and Stress Detection. Methods Mol Biol. 2022; 2494:255-265. doi: 10.1007/978-1-0716-2297-1_18. PMID: 35467213.
2. Manuscripts potential submit to journal of Food and Function
   "Total potato glycoalkaloid and chlorophyll synthesis in green potato tubers".
3. Conference Proceeding under review
   Title "Advancing Food Security Through the Development of IoT based Water Allocation Model", 4th World Irrigation Forum (WIF).

REFERENCES

1) Aakash, C., Joost van, H., Hanna, B., Oscar, B., Erik, A. and Rodomiro, O. (2019) 'High-throughput field-phenotyping tools for plant breeding and precision agriculture', *Agronomy*, 9(5), p. 258.

2) Abreu, P., Relva, A., Matthew, S., Gomes, Z. and Morais, Z. (2007) 'High-performance liquid chromatographic determination of glycoalkaloids in potatoes from conventional, integrated, and organic crop systems', *Food Control*, 18, pp. 40-44.

3) Acs, S., Berentsen, P.B.M. and Huirne, R.B.M. (2005) 'Modelling conventional and organic farming: A literature review', *NJAS - Wageningen Journal of Life Sciences*, 53(1), pp. 1-18.

4) Adam, E., Deng, H., Odindi, J., Abdel-Rahman, E.M. and Mutanga, O. (2017) 'Detecting the early stage of phaeosphaeria leaf spot infestations in maize crop using in situ hyperspectral data and guided regularized random forest algorithm', *Journal of Spectroscopy*, 2017(1).

5) Adamchuk, V.I., Ferguson, R.B. and Hergert, G.W. (2010) *Soil heterogeneity and crop growth*. Dordrecht: Dordrecht: Springer Netherlands.

6) Adao, T., Hruska, J., Padua, L., Bessa, J., Peres, E., Morais, R. and Sousa, J. (2017) 'Hyperspectral imaging: A review on uav-based sensors, data processing and applications for agriculture and forestry', *Remote Sensing*, 9(11).

7) Adekanmbi, T., Wang, X., Basheer, S., Nawaz, R.A., Pang, T., Hu, Y. and Liu, S. (2023) 'Assessing future climate change impacts on potato yields — a case study for prince edward island, canada', *Foods*, 12(6), p. 1176.

8) AHDB (2024a) *Climate change impacts on uk potato production: Industry options and responses*. Available at: https://potatoes.ahdb.org.uk/climate-change-impacts-on-uk-potato-production-industry-options-and-responses (Accessed: 24 Febuary).

9) AHDB (2024b) 'Identify the symptoms caused by fungi, fungal-like organisms and bacteria that are capable of causing disease in potato tubers.', *Potato diseases and defects*. Available at: https://potatoes.ahdb.org.uk/disease-defects (Accessed: 24 Febuary).

10) Akca, D. and Grün, A. (2007) 'Generalized least squares multiple 3d surface matching', *IAPRS*. ISPRS, pp. 1-7.

11) Akino, S., Takemoto, D. and Hosaka, K. (2014) 'Phytophthora infestans: A review of past and current studies on potato late blight', *Journal of General Plant Pathology*, 80(1), pp. 24-37.

12) Al-Gaadi, K.A., Hassaballa, A.A., Tola, E., Kayad, A.G., Madugundu, R., Alblewi, B. and Assiri, F. (2016) 'Prediction of potato crop yield using precision agriculture techniques', *PLOS ONE*, 11(9), p. e0162219.

13) Al-jabery, K.K., Obafemi-Ajayi, T., Olbricht, G.R. and Wunsch Ii, D.C. (2020) '2 - data preprocessing', in Al-jabery, K.K., Obafemi-Ajayi, T., Olbricht, G.R. and Wunsch Ii, D.C. (eds.) *Computational learning approaches to data analytics in biomedical applications*. Academic Press, pp. 7-27.

14) Alamar, M.C., Tosetti, R., Landahl, S., Bermejo, A. and Terry, L.A. (2017) 'Assuring potato tuber quality during storage: A future perspective', *Frontiers in Plant Science*, 8(2034).

15) Aliche, E.B., Oortwijn, M., Theeuwen, T.P.J.M., Bachem, C.W.B., van Eck, H.J., Visser, R.G.F. and van der Linden, C.G. (2019) 'Genetic mapping of tuber size distribution and marketable tuber yield under drought stress in potatoes', *Euphytica*, 215, pp. 1-19.

16) Almeida, M.R., Fidelis, C.H., Barata, L.E. and Poppi, R.J. (2013) 'Classification of amazonian rosewood essential oil by raman spectroscopy and pls-da with reliability estimation', *Talanta*, 117, pp. 305-311.

17) Anderson II, S.L., Murray, S.C., Malambo, L., Ratcliff, C., Popescu, S., Cope, D., Chang, A., Jung, J. and Thomasson, J.A. (2019) 'Prediction of maize grain yield before maturity using improved temporal height estimates of unmanned aerial systems', *Plant Phenome Journal*, 2(1), pp. 1-15.

18) Andrivon, D. (2017) 'Potato facing global challenges: How, how much, how well?', *Potato Research*, 60(3), pp. 389-400.

19) Appeltans, S., Guerrero, A., Nawar, S., Pieters, J. and Mouazen, A.M. (2020) 'Practical recommendations for hyperspectral and thermal proximal disease sensing in potato and leek fields', *Remote Sensing*, 12(12), pp. 19-39.

20) Araus, J.L. and Cairns, J.E. (2014) 'Field high-throughput phenotyping: The new crop breeding frontier', *Trends in Plant Science*, 19(1), pp. 52-61.

21) Araus, J.L., Kefauver, S.C., Zaman-Allah, M., Olsen, M.S. and Cairns, J.E. (2018) 'Translating high-throughput phenotyping into genetic gain', *Trends Plant Sci*, 23(5), pp. 451-466.

22) Ávila-Valdés, A., Quinet, M., Lutts, S., Martínez, J.P. and Lizana, X.C. (2020) 'Tuber yield and quality responses of potato to moderate temperature increase during tuber bulking under two water availability scenarios', *Field Crops Research*, 251, p. 107786.

23) Bamberg, J., Moehninsi, Navarre, R. and Suriano, J. (2015) 'Variation for tuber greening in the diploid wild potato solanum microdontum', *American Journal of Potato Research*, 92(3), pp. 435-443.

24) Bannari, A., Morin, D., Bonn, F. and Huete, A. (1995) 'A review of vegetation indices', *Remote sensing reviews*, 13(1-2), pp. 95-120.

25) Barber-Rowe, H. (2021) 'Soil association standards farming and growing version 18.6: Updated on 12 february 2021'.

26) Barceloux, D.G.M.D. (2009) 'Potatoes, tomatoes, and solanine toxicity (solanum tuberosum l., solanum lycopersicum l.)', *Disease-a-month*, 55(6), pp. 391-402.

27) Bejarano, L., Mignolet, E., Devaux, A., Espinola, N., Carrasco, E. and Larondelle, Y. (2000) 'Glycoalkaloids in potato tubers: The effect of variety and drought stress on the α-solanine and α-chaconine contents of potatoes', *Journal of the science of food and agriculture*, 80(14), pp. 2096-2100.

28) Bendig, J., Bolten, A. and Bareth, G. (2013a) 'Uav-based imaging for multi-temporal, very high resolution crop surface models to monitor crop growth variability', *Photogrammetrie - Fernerkundung - Geoinformation*, 6, pp. 551-562.

29) Bendig, J., Bolten, A., Bennertz, S., Broscheit, J., Eichfuss, S. and Bareth, G. (2014) 'Estimating biomass of barley using crop surface models (csms) derived from uav-based rgb imaging', *Remote sensing*, 6(11), pp. 10395-10412.

30) Bendig, J., Willkomm, M., Tilly, N., Gnyp, M., Bennertz, S., Qiang, C., Miao, Y., Lenz-Wiedemann, V. and Bareth, G. (2013b) 'Very high resolution crop surface models (csms) from uav-based stereo images for rice growth monitoring in northeast china', *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40, pp. 45-50.

31) Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., Gnyp, M.L. and 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, pp. 79-87.

32) Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelder, I. and Heynen, M. (2004) 'Multi-resolution, object-oriented fuzzy analysis of remote sensing data for gis-ready information', *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(3), pp. 239-258.

33) Bergweiler, P. and Lütz, C. (1986) 'Determination of leaf pigments by hplc after extraction with n,n-dimethylformamide: Ecophysiological applications', *Environmental and Experimental Botany*, 26(3), pp. 207-210.

34) Biau, G. (2012) 'Analysis of a random forests model', *The Journal of Machine Learning Research*, 13(1), pp. 1063-1095.

35) Birenboim, M., Kenigsbuch, D. and Shimshoni, J.A. (2023) 'Novel fluorescence spectroscopy method coupled with n-pls-r and pls-da models for the quantification of cannabinoids and the classification of cannabis cultivars', *Phytochemical Analysis*, 34(3), pp. 280-288.

36) Bishop, C., Rees, D., Cheema, M.U.A., Harper, G. and Stroud, G. (2012) 'Potatoes', in *Crop post-harvest: Science and technology*. pp. 334-359.

37) Bivand, R.S., Pebesma, E.J., Gómez-Rubio, V. and Pebesma, E.J. (2008) *Applied spatial data analysis with r*. Springer.

38) Bolton, D.K. and Friedl, M.A. (2013) 'Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics', *Agricultural and Forest Meteorology*, 173, pp. 74-84.

39) Bonierbale, M.W., Amoros, W.R., Salas, E. and de Jong, W. (2020) 'Potato breeding', in Campos, H. and Ortiz, O. (eds.) *The potato crop: Its agricultural, nutritional and social contribution to humankind*. Cham: Springer International Publishing, pp. 163-217.

40) Boschiero, M., De Laurentiis, V., Caldeira, C. and Sala, S. (2023) 'Comparison of organic and conventional cropping systems: A systematic review of life cycle assessment studies', *Environmental Impact Assessment Review*, 102, pp. 107-187.

41) Bradshaw, J.E. (2007) 'Potato-breeding strategy', in *Potato biology and biotechnology*. Elsevier, pp. 157-177.

42) Bradshaw, J.E. (2021) 'Improving potato quality: A problem of definition and measurement', in *Potato breeding: Theory and practice*. Cham: Springer International Publishing, pp. 195-246.

43) Bradshaw, J.E. and Bonierbale, M. (2010) 'Potatoes', in Bradshaw, J.E. (ed.) *Root and tuber crops*. New York, NY: Springer New York, pp. 1-52.

44) Brazinskiene, V., Asakaviciute, R., Miezeliene, A., Alencikiene, G., Ivanauskas, L., Jakstas, V., Viskelis, P. and Razukas, A. (2014) 'Effect of farming systems on the yield, quality parameters and sensory properties of conventionally and organically grown potato (solanum tuberosum l.) tubers', *Food Chemistry*, 145, pp. 903-909.

45) Bréda, N.J. (2003) 'Ground-based measurements of leaf area index: A review of methods, instruments and current controversies', *Journal of experimental botany*, 54(392), pp. 2403-2417.

46) Breiman, L. (2001) 'Random forests', *Machine learning*, 45, pp. 5-32.

47) Britannica, T. (2022) *Potato*. Available at: https://www.britannica.com/plant/potato.

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

49) Bucksch, A., Burridge, J., York, L.M., Das, A., Nord, E., Weitz, J.S. and Lynch, J.P. (2014) 'Image-based high-throughput field phenotyping of crop roots', *Plant Physiology*, 166(2), pp. 470-486.

50) Burger, J. and Gowen, A. (2011) 'Data handling in hyperspectral image analysis', *Chemometrics and Intelligent Laboratory Systems*, 108(1), pp. 13-22.

51) Bussan, A.J. (2009) *Tuber maturation and potato storability: Optimizing skin set, sugars, and solids*. Division of Cooperative Extension of the University of Wisconsin--Extension.

52) Bussan, A.J., Mitchell, P.D., Copas, M.E. and Drilias, M.J. (2007) 'Evaluation of the effect of density on potato yield and tuber size distribution', *Crop Science*, 47(6), pp. 2462-2472.

53) Calderón, R., Navas-Cortés, J.A. and Zarco-Tejada, P.J. (2015) 'Early detection and quantification of verticillium wilt in olive using hyperspectral and thermal imagery over large areas', *Remote Sensing*, 7(5), pp. 5584-5610.

54) Caldiz, D.O., Fernandez, L.V. and Struik, P.C. (2001) 'Physiological age index: A new, simple and reliable index to assess the physiological age of seed potato tubers based on haulm killing date and length of the incubation period', *Field Crops Research*, 69(1), pp. 69-79.

55) Campos, H. (2020) *The potato crop : Its agricultural, nutritional and social contribution to humankind*. 1st 2020.. edn. Cham: Cham Springer Nature.

56) Campos, J., Llop, J., Gallart, M., García-Ruiz, F., Gras, A., Salcedo, R. and Gil, E. (2019) 'Development of canopy vigour maps using uav for site-specific management during vineyard spraying process', *Precision Agriculture*.

57) Canonico, P., De Nito, E., Esposito, V., Fattoruso, G., Iacono, M.P. and Mangia, G. (2021) 'Visualizing knowledge for decision-making in lean production development settings. Insights from the automotive industry', *Management Decision*, 60(4), pp. 1076-1094.

58) Carputo, D., Aversano, R. and Frusciante, L. (2004) *Meeting of the Physiology Section of the European Association for Potato Research 684*.

59) Celis-Gamboa, B.C. (2002) *The life cycle of the potato (solanum tuberosum l.): From crop physiology to genetics*. Wageningen University and Research.

60) Celis-Gamboa, C., Struik, P.C., Jacobsen, E. and Visser, R.G.F. (2003) 'Sprouting of seed tubers during cold storage and its influence on tuber formation, flowering and the duration of the life cycle in a diploid population of potato', *Potato Research*, 46(1), pp. 9-25.

61) Cendrero-Mateo, M.P., Muller, O., Albrecht, H., Burkart, A., Gatzke, S., Janssen, B., Keller, B., Körber, N., Kraska, T., Matsubara, S., Li, J., Müller-Linow, M., Pieruschka, R., Pinto, F., Rischbeck, P., Schickling, A., Steier, A., Watt, M. and Schurr, U. (2017) 'Field phenotyping: Challenges and opportunities', in, pp. 53-80.

62) Chakraborty, K.K., Mukherjee, R., Chakroborty, C. and Bora, K. (2022) 'Automated recognition of optical image based potato leaf blight diseases using deep learning', *Physiological and Molecular Plant Pathology*, 117, p. 101781.

63) Chapman, S.C., Merz, T., Chan, A., Jackway, P., Hrabar, S., Dreccer, M.F., Holland, E., Zheng, B., Ling, T.J. and Jimenez-Berni, J. (2014) 'Pheno-copter: A low-altitude, autonomous remote-sensing robotic helicopter for high-throughput field-based phenotyping', *Agronomy*, 4(2), p. 279.

64) Chatzivassiliou, E.K., Moschos, E., Gazi, S., Koutretsis, P. and Tsoukaki, M. (2008) 'Infection of potato crops and seeds with potato virus y and potato leafroll virus in greece', *Journal of Plant Pathology*, 90(2), pp. 253-261.

65) Chaudhary, A., Kolhe, S. and Kamal, R. (2016) 'An improved random forest classifier for multi-class classification', *Information Processing in Agriculture*, 3(4), pp. 215-222.

66) Chawade, A., van Ham, J., Blomquist, H., Bagge, O., Alexandersson, E. and Ortiz, R. (2019) 'High-throughput field-phenotyping tools for plant breeding and precision agriculture', *Agronomy*, 9(5), p. 258.

67) Comba, L., Biglia, A., Ricauda Aimonino, D. and Gay, P. (2018) 'Unsupervised detection of vineyards by 3d point-cloud uav photogrammetry for precision agriculture', *Computers and Electronics in Agriculture*, 155, pp. 84-95.

68) Condran, S., Bewong, M., Islam, M.Z., Maphosa, L. and Zheng, L. (2022) 'Machine learning in precision agriculture: A survey on trends, applications and evaluations over two decades', *IEEE Access*, 10, pp. 73786-73803.

69) Coxon, D. (1984) 'Methodology for glycoalkaloid analysis', *American potato journal*, 61, pp. 169-183.

70) D'hoop, B.B., Paulo, M.J., Mank, R.A., van Eck, H.J. and van Eeuwijk, F.A. (2008) 'Association mapping of quality traits in potato (solanum tuberosum l.)', *Euphytica*, 161(1), pp. 47-60.

71) Daccache, A., Keay, C., Jones, R.J., Weatherhead, E., Stalham, M. and Knox, J.W. (2012) 'Climate change and land suitability for potato production in england and wales: Impacts and adaptation', *The Journal of Agricultural Science*, 150(2), pp. 161-177.

72) Dammer, K.-H., Dworak, V. and Selbeck, J. (2016) 'On-the-go phenotyping in field potatoes using camera vision', *Potato Research*, 59(2), pp. 113-127.

73) Das, B., Mahajan, G. and Singh, R. (2019) 'Hyperspectral remote sensing: Use in detecting abiotic stresses in agriculture', in, pp. 317-335.

74) Datiles, M.J. and Acevedo-Rodríguez, P. (2022) *Solanum tuberosum (potato)*. CABI International.

75) David, D., Jose, J.-B., Hamlyn, J., Xavier, S. and Robert, F. (2014) 'Proximal remote sensing buggies and potential applications for field-based phenotyping', *Agronomy (Basel)*, 4(3), pp. 349-379.

76) Daviet, B., Fernandez, R., Cabrera-Bosquet, L., Pradal, C. and Fournier, C. (2022) 'Phenotrack3d: An automatic high-throughput phenotyping pipeline to track maize organs over time', *Plant Methods*, 18(1), p. 130.

77) De Castro, A., Jiménez-Brenes, F., Torres-Sánchez, J., Peña-Barragán, J.M., Borra-Serrano, I. and López-Granados, F. (2018) '3-d characterization of vineyards using a novel uav imagery-based obia procedure for precision viticulture applications', *Remote Sensing*, 10, p. 584.

78) de Jesus Colwell, F., Souter, J., Bryan, G.J., Compton, L.J., Boonham, N. and Prashar, A. (2021) 'Development and validation of methodology for estimating potato canopy structure for field crop phenotyping and improved breeding', *Frontiers in Plant Science*, 12(139).

79) Dean, B.B. (2018) *Managing the potato production system: 0734*. CRC Press.

80) Deery, D.M., Greg, J.R., Jose Antonio, J.-B., Richard Alexander, J., Anthony, G.C., William, D.B., Paul, H., Jamie, S., Robert, D. and Robert, T.F. (2016) 'Methodology for high-throughput field phenotyping of canopy temperature using airborne thermography', *Frontiers in plant science*, 7(2016).

81) Demilie, W.B. (2024) 'Plant disease detection and classification techniques: A comparative study of the performances', *Journal of Big Data*, 11(1), p. 5.

82) Den, T.t., van de Wiel, I., de Wit, A., van Evert, F.K., van Ittersum, M.K. and Reidsma, P. (2022) 'Modelling potential potato yields: Accounting for experimental differences in modern cultivars', *European Journal of Agronomy*, 137, p. 126510.

83) Devadas, R., Devadas, D.W., Lamb, D., Lamb, S., Backhouse, S. and Simpfendorfer, S. (2015) 'Sequential application of hyperspectral indices for delineation of stripe rust infection and nitrogen deficiency in wheat', *Precision Agriculture*, 16(5).

84) Devaux, A., Goffart, J.-P., Kromann, P., Andrade-Piedra, J., Polar, V. and Hareau, G. (2021) 'The potato of the future: Opportunities and challenges in sustainable agri-food systems', *Potato Research*, 64(4), pp. 681-720.

85) Dhalsamant, K., Singh, C.B. and Lankapalli, R. (2022) 'A review on greening and glycoalkaloids in potato tubers: Potential solutions', *Journal of Agricultural and Food Chemistry*.

86) Dordas, C. (2008) 'Role of nutrients in controlling plant diseases in sustainable agriculture. A review', *Agronomy for sustainable development*, 28, pp. 33-46.

87) Dourado, C., Pinto, C., Barba, F.J., Lorenzo, J.M., Delgadillo, I. and Saraiva, J.A. (2019) 'Innovative non-thermal technologies affecting potato tuber and fried potato quality', *Trends in food science & technology*, 88, pp. 274-289.

88) Duarte-Carvajalino, J., Alzate, D., Ramirez, A., Santa-Sepulveda, J., Fajardo-Rojas, A. and Soto-Suárez, M. (2018) 'Evaluating late blight severity in potato crops using unmanned aerial vehicles and machine learning algorithms. Remote sens. 10, 1513 (2018)'.

89) Duarte, H., Zambolim, L., Capucho, A., Nogueira Júnior, A.F., Rosado, A., Carine, C.R.C., Paul, P. and Mizubuti, E. (2013) 'Development and validation of a set of standard area diagrams to estimate severity of potato early blight', *European Journal of Plant Pathology*, 137.

90) Duke Gekonge, O., George Ooko, A. and Michael Wandayi, O. (2016) 'A review of occurrence of glycoalkaloids in potato and potato products', *Current research in nutrition and food science*, 4(3), p. 195.

91) Ecobreed (2020) *Ecobreed: Increasing the efficiency and competitiveness of organic crop breeding*. Available at: https://ecobreed.eu/.

92) Economou, F., Papamichael, I., Voukkali, I., Loizia, P., Klontza, E., Lekkas, D.F., Vincenzo, N., Demetriou, G., Navarro-Pedreño, J. and Zorpas, A.A. (2023) 'Life cycle assessment of potato production in insular communities under subtropical climatic conditions', *Case Studies in Chemical and Environmental Engineering*, 8, p. 100419.

93) El-Amin, S.M., Valkonen, J.P.T., Bremer, K. and Pehu, E. (1994) 'Elimination of viruses and hypersensitivity to potato virus y (pvyo) in an important sudanese potato stock (zalinge)', *American Potato Journal*, 71(4), pp. 267-272.

94) Elavarasan, D. and Vincent, P.D. (2020) 'Crop yield prediction using deep reinforcement learning model for sustainable agrarian applications', *IEEE access*, 8, pp. 86886-86901.

95) Eriksson, L., Byrne, T., Johansson, E., Trygg, J. and Vikström, C. (2013) *Multi- and megavariate data analysis basic principles and applications*. Umetrics Academy.

96) European Food Safety, A. (2020) 'Outcome of a public consultation on the draft risk assessment of glycoalkaloids in feed and food, in particular in potatoes and potato-derived products', *EFSA Supporting Publications*, 17(8), p. n/a.

97) Fiers, M., Chatot, C., Edel-Hermann, V., Le Hingrat, Y., Konate, A.Y., Gautheron, N., Guillery, E., Alabouvette, C. and Steinberg, C. (2010) 'Diversity of microorganisms associated with atypical superficial blemishes of potato tubers and pathogenicity assessment', *European Journal of Plant Pathology*, 128(3), pp. 353-371.

98) Fiorani, F. and Schurr, U. (2013) 'Future scenarios for plant phenotyping', *Annu Rev Plant Biol*, 64, pp. 267-91.

99) Franceschini, M., Bartholomeus, H., Van Apeldoorn, D., Suomalainen, J. and Kooistra, L. (2017) 'Assessing changes in potato canopy caused by late blight in organic production systems through uav-based pushbroom imaging spectrometer', *ISPRS* -

International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-2/W6, pp. 109-112.

100) Franceschini, M.H.D., Bartholomeus, H., van Apeldoorn, D.F., Suomalainen, J. and Kooistra, L. (2019) 'Feasibility of unmanned aerial vehicle optical imagery for early detection and severity assessment of late blight in potato', *Remote Sensing*, 11(3).

101) Friedli, M., Kirchgessner, N., Grieder, C., Liebisch, F., Mannale, M. and Walter, A. (2016) 'Terrestrial 3d laser scanning to track the increase in canopy height of both monocot and dicot crop species under field conditions', *Plant Methods*, 12(1), p. 9.

102) Friedman, M. (2006) 'Potato glycoalkaloids and metabolites: Roles in the plant and in the diet', *Journal of Agricultural and Food Chemistry*, 54(23), pp. 8655-8681.

103) Friedman, M. and McDonald, G.M. (1999) 'Postharvest changes in glycoalkaloid content of potatoes', in Jackson, L.S., Knize, M.G. and Morgan, J.N. (eds.) *Impact of processing on food safety*. Boston, MA: Springer US, pp. 121-143.

104) Friedman, M., McDonald, G.M. and Filadelfi-Keszi, M. (1997) 'Potato glycoalkaloids: Chemistry, analysis, safety, and plant physiology', *Critical Reviews in Plant Sciences*, 16(1), pp. 55-132.

105) Friedman, M., Roitman, J.N. and Kozukue, N. (2003) 'Glycoalkaloid and calystegine contents of eight potato cultivars', *Journal of Agricultural and Food Chemistry*, 51(10), pp. 2964-2973.

106) Fry, W.E. (2020) 'Phytophthora infestans: The itinerant invader; “late blight”: The persistent disease', *Phytoparasitica*, 48(1), pp. 87-94.

107) Gago, J., Douthe, C., Coopman, R.E., Gallego, P.P., Ribas-Carbo, M., Flexas, J., Escalona, J. and Medrano, H. (2015) 'Uavs challenge to assess water stress for sustainable agriculture', *Agricultural Water Management*, 153(C), pp. 9-19.

108) Gallandt, E. (2014) 'Weed management in organic farming', *Recent advances in weed management*, pp. 63-85.

109) Gao, J., Westergaard, J.C., Sundmark, E.H.R., Bagge, M., Liljeroth, E. and Alexandersson, E. (2021) 'Automatic late blight lesion recognition and severity quantification based on field imagery of diverse potato genotypes by deep learning', *Knowledge-Based Systems*, 214, p. 106723.

110) Garcia-Gonzalez, J., Mehl, H.L., Langston, D.B. and Rideout, S.L. (2022) 'Planting date and cultivar selection to manage southern blight in potatoes in the mid-atlantic united states', *Crop Protection*, 162, p. 106077.

111) Gattinger, A., Muller, A., Haeni, M., Skinner, C., Fliessbach, A., Buchmann, N., Mäder, P., Stolze, M., Smith, P., Scialabba, N.E.-H. and Niggli, U. (2012) 'Enhanced top soil carbon stocks under organic farming', *Proceedings of the National Academy of Sciences*, 109(44), pp. 18226-18231.

112) Gerard, J. (1995) 'The herball or generall historie of plantes (1597)', *Bonham & John Norton, London pp [xviii]*, 1392, p. 72.

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

114) Gitelson, A.A., Kaufman, Y.J. and Merzlyak, M.N. (1996) 'Use of a green channel in remote sensing of global vegetation from eos-modis', *Remote Sensing of Environment*, 58(3), pp. 289-298.

115) Gitelson, A.A., Kaufman, Y.J., Stark, R. and Rundquist, D. (2002) 'Novel algorithms for remote estimation of vegetation fraction', *Remote sensing of Environment*, 80(1), pp. 76-87.

116) Goffart, J.-P., Haverkort, A., Storey, M., Haase, N., Martin, M., Lebrun, P., Ryckmans, D., Florins, D. and Demeulemeester, K. (2022) 'Potato production in northwestern europe (germany, france, the netherlands, united kingdom, belgium): Characteristics, issues, challenges and opportunities', *Potato Research*, 65(3), pp. 503-547.

117) Gold, K.M., Townsend, P.A., Herrmann, I. and Gevens, A.J. (2020) 'Investigating potato late blight physiological differences across potato cultivars with spectroscopy and machine learning', *Plant Science*, 295, p. 110316.

118) Golzarian, M.R., Frick, R.A., Rajendran, K., Berger, B., Roy, S., Tester, M. and Lun, D.S. (2011) 'Accurate inference of shoot biomass from high-throughput images of cereal plants', *Plant methods*, 7, pp. 1-11.

119) Gómez, D., Salvador, P., Sanz, J. and Casanova, J.L. (2019) 'Potato yield prediction using machine learning techniques and sentinel 2 data', *Remote Sensing*, 11(15), p. 1745.

120) Gondwe, R.L., Kinoshita, R., Suminoe, T., Aiuchi, D., Palta, J.P. and Tani, M. (2020) 'Yield and quality characteristics of popular processing potato (solanum tuberosum l.) cultivars in two contrasting soil types under grower management in hokkaido, japan', *Potato Research*, 63(3), pp. 385-402.

121) Gong, H.-L., Dusengemungu, L., Igiraneza, C. and Rukundo, P. (2021a) 'Molecular regulation of potato tuber dormancy and sprouting: A mini-review', *Plant Biotechnology Reports*, 15(4), pp. 417-434.

122) Gong, Y., Yang, K., Lin, Z., Fang, S., Wu, X., Zhu, R. and Peng, Y. (2021b) 'Remote estimation of leaf area index (lai) with unmanned aerial vehicle (uav) imaging for different rice cultivars throughout the entire growing season', *Plant Methods*, 17(1), p. 88.

123) Gopal, J. and Khurana, S. (2006) *Handbook of potato production, improvement, and postharvest management*. CRC Press.

124) Gottschalk, K. and Ezekiel, R. (2006) 'Storage', in *Handbook of potato production, improvement, and postharvest management*. CRC Press, pp. 489-522.

125) Govardhan, M. and V, M.B. (2019) *2019 Global Conference for Advancement in Technology (GCAT)*. 18-20 Oct. 2019.

126) Grasel, F.S. and Ferrão, M.F. (2016) 'A rapid and non-invasive method for the classification of natural tannin extracts by near-infrared spectroscopy and pls-da', *Analytical Methods*, 8(3), pp. 644-649.

127) Griffin, T.S., Johnson, B.S. and Ritchie, J.T. (1993) *A simulation model for potato growth and development: Substor-potato version 2.0*. Michigan State University, Department of Crop and Soil Sciences East Lansing ….

128) Großkinsky, D.K., Svensgaard, J., Christensen, S. and Roitsch, T. (2015) 'Plant phenomics and the need for physiological phenotyping across scales to narrow the genotype-to-phenotype knowledge gap', *J Exp Bot*, 66(18), pp. 5429-40.

129) Grunenfelder, L., Hiller, L.K. and Knowles, N.R. (2006a) 'Color indices for the assessment of chlorophyll development and greening of fresh market potatoes', *Postharvest biology and technology*, 40(1), pp. 73-81.

130) Grunenfelder, L.A., Knowles, L.O., Hiller, L.K. and Knowles, N.R. (2006b) 'Glycoalkaloid development during greening of fresh market potatoes (solanum tuberosum l.)', *Journal of Agricultural and Food Chemistry*, 54(16), pp. 5847-5854.

131) Guo, C., Zhang, L., Zhou, X., Zhu, Y., Cao, W., Qiu, X., Cheng, T. and Tian, Y. (2018) 'Integrating remote sensing information with crop model to monitor wheat growth and yield based on simulation zone partitioning', *Precision Agriculture*, 19(1), pp. 55-78.

132) Guo, W., Fukano, Y., Noshita, K. and Ninomiya, S. (2020) 'Field-based individual plant phenotyping of herbaceous species by unmanned aerial vehicle', *Ecology and Evolution*, 10(21), pp. 12318-12326.

133) Guo, Y., Ren, G., Zhang, K., Li, Z., Miao, Y. and Guo, H. (2021) 'Leaf senescence: Progression, regulation, and application', *Molecular Horticulture*, 1, pp. 1-25.

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

135) Hagman, J.E., Mårtensson, A. and Grandin, U. (2009) 'Cultivation practices and potato cultivars suitable for organic potato production', *Potato Research*, 52(4), pp. 319-330.

136) Han, L., Yang, G., Yang, H., Xu, B., Li, Z. and Yang, X. (2018) 'Clustering field-based maize phenotyping of plant-height growth and canopy spectral dynamics using a uav remote-sensing approach', *Frontiers in plant science*, 9, pp. 1638-1638.

137) Hansen, J., Koppel, M., Valskyte, A., Turka, I. and Kapsa, J. (2005) 'Evaluation of foliar resistance in potato to phytophthora infestans based on an international field trial network', *Plant Pathology*, 54(2), pp. 169-179.

138) Hao, Y., Sun, X., Gao, R., Pan, Y. and Liu, Y. (2010) 'Application of visible and near infrared spectroscopy to identification of navel orange varieties using simca and pls-da methods', *Transactions of the Chinese Society of Agricultural Engineering*, 26(12), pp. 373-377.

139) Haverkort, A. (1990) 'Ecology of potato cropping systems in relation to latitude and altitude', *Agricultural systems*, 32(3), pp. 251-272.

140) Haverkort, A. and Struik, P. (2015) 'Yield levels of potato crops: Recent achievements and future prospects', *Field Crops Research*, 182, pp. 76-85.

141) Haverkort, A.J., Linnemann, A.R., Struik, P.C. and Wiskerke, J.S.C. (2023) 'On processing potato. 4. Survey of the nutritional and sensory value of products and dishes', *Potato Research*, 66(2), pp. 429-468.

142) Haverkort, A.J., Struik, P.C., Visser, R.G.F. and Jacobsen, E. (2009) 'Applied biotechnology to combat late blight in potato caused by phytophthora infestans', *Potato Research*, 52(3), pp. 249-264.

143) Heltoft, P., Wold, A.-B. and Molteberg, E.L. (2017) 'Maturity indicators for prediction of potato (solanum tuberosum l.) quality during storage', *Postharvest Biology and Technology*, 129, pp. 97-106.

144) Hendricks, R.L., Olsen, N., Thornton, M.K. and Hatzenbuehler, P. (2022) 'Susceptibility of potato cultivars to blackspot and shatter bruise at three impact heights', *American Journal of Potato Research*, 99(5), pp. 358-368.

145) Herrero-Huerta, M., González-Aguilera, D., Rodriguez-Gonzalvez, P. and Hernández-López, D. (2015) 'Vineyard yield estimation by automatic 3d bunch modelling in field conditions', *Computers and Electronics in Agriculture*, 110, pp. 17-26.

146) Hirut, B.G., Shimelis, H.A., Melis, R., Fentahun, M. and De Jong, W. (2017) 'Yield, yield-related traits and response of potato clones to late blight disease, in north-western highlands of ethiopia', *Journal of Phytopathology*, 165(1), pp. 1-14.

147) Holman, F.H., Riche, A.B., Michalski, A., Castle, M., Wooster, M.J. and Hawkesford, M.J. (2016) 'High throughput field phenotyping of wheat plant height and growth rate in field plot trials using uav based remote sensing', *Remote Sensing*, 8(12), p. 1031.

148) Horsfield, A., Wicks, T., Davies, K., Wilson, D. and Paton, S. (2010) 'Effect of fungicide use strategies on the control of early blight (alternaria solani) and potato yield', *Australasian Plant Pathology*, 39, pp. 368-375.

149) Hossain, M.B., Rawson, A., Aguiló-Aguayo, I., Brunton, N.P. and Rai, D.K. (2015) 'Recovery of steroidal alkaloids from potato peels using pressurized liquid extraction', *Molecules (Basel, Switzerland)*, 20(5), pp. 8560-8573.

150) Hu, P., Chapman, S.C., Wang, X., Potgieter, A., Duan, T., Jordan, D., Guo, Y. and Zheng, B. (2018) 'Estimation of plant height using a high throughput phenotyping platform based on unmanned aerial vehicle and self-calibration: Example for sorghum breeding', *European Journal of Agronomy*, 95, pp. 24-32.

151) Hu, Q., Sulla-Menashe, D., Xu, B., Yin, H., Tang, H., Yang, P. and Wu, W. (2019) 'A phenology-based spectral and temporal feature selection method for crop mapping from satellite time series', *International Journal of Applied Earth Observation and Geoinformation*, 80, pp. 218-229.

152) Hu, Q., Tang, C., Zhou, X., Yang, X., Luo, Z., Wang, L., Yang, M., Li, D. and Li, L. (2023) 'Potatoes dormancy release and sprouting commencement: A review on current and future prospects', *Food Frontiers*, 4(3), pp. 1001-1018.

153) Huang, Z. (1998) 'Extensions to the k-means algorithm for clustering large data sets with categorical values', *Data mining and knowledge discovery*, 2(3), pp. 283-304.
154) Huber, D., Römheld, V. and Weinmann, M. (2012) 'Chapter 10 - relationship between nutrition, plant diseases and pests', in Marschner, P. (ed.) *Marschner's mineral nutrition of higher plants (third edition)*. San Diego: Academic Press, pp. 283-298.
155) Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X. and Ferreira, L.G. (2002) 'Overview of the radiometric and biophysical performance of the modis vegetation indices', *Remote sensing of environment*, 83(1-2), pp. 195-213.
156) Huete, A.R. (1988) 'A soil-adjusted vegetation index (savi)', *Remote Sensing of Environment*, 25(3), pp. 295-309.
157) Huete, A.R. (2004) '11 - remote sensing for environmental monitoring', in Artiola, J.F., Pepper, I.L. and Brusseau, M.L. (eds.) *Environmental monitoring and characterization*. Burlington: Academic Press, pp. 183-206.
158) Hurtado, P.X., Schnabel, S.K., Zaban, A., Veteläinen, M., Virtanen, E., Eilers, P.H.C., van Eeuwijk, F.A., Visser, R.G.F. and Maliepaard, C. (2012) 'Dynamics of senescence-related qtls in potato', *Euphytica*, 183(3), pp. 289-302.
159) Hussein, Z., Fawole, O.A. and Opara, U.L. (2020) 'Harvest and postharvest factors affecting bruise damage of fresh fruits', *Horticultural Plant Journal*, 6(1), pp. 1-13.
160) Ingram, K.T. and McCloud, D.E. (1984) 'Simulation of potato crop growth and development1', *Crop Science*, 24(1), p. cropsci1984.0011183X002400010006x.
161) Jackson, K., Kerr, J., Kilpatrick, J., Henderson, C. and Lovatt, J. (1997) 'Potato information kit. Agrilink, your growing guide to better farming guide'. Brisbane, Queensland: Queensland Horticulture Institute.
162) Jansky, S. (2009) 'Chapter 2 - breeding, genetics, and cultivar development', in Singh, J. and Kaur, L. (eds.) *Advances in potato chemistry and technology*. San Diego: Academic Press, pp. 27-62.
163) Jemison Jr, J.M., Sexton, P. and Camire, M.E. (2008) 'Factors influencing consumer preference of fresh potato varieties in maine', *American Journal of Potato Research*, 85(5), pp. 388-389.
164) Jensen, P.H., Harder, B.J., Strobel, B.W., Svensmark, B. and Hansen, H.C.B. (2007) 'Extraction and determination of the potato glycoalkaloid α -solanine in soil', *International Journal of Environmental Analytical Chemistry*, 87(12), pp. 813-824.

165) Ji, Y., Chen, Z., Cheng, Q., Liu, R., Li, M., Yan, X., Li, G., Wang, D., Fu, L., Ma, Y., Jin, X., Zong, X. and Yang, T. (2022) 'Estimation of plant height and yield based on uav imagery in faba bean (vicia faba l.)', *Plant Methods*, 18(1), p. 26.

166) Jimenez-Berni, J.A., Deery, D.M., Rozas-Larraondo, P., Condon, A.G., Rebetzke, G.J., James, R.A., Bovill, W.D., Furbank, R.T. and Sirault, X.R.R. (2018) 'High throughput determination of plant height, ground cover, and above-ground biomass in wheat with lidar', *Frontiers in Plant Science*, 9(237).

167) Jin, X., Liu, S., Baret, F., Hemerlé, M. and Comar, A. (2017) 'Estimates of plant density of wheat crops at emergence from very low altitude uav imagery', *Remote Sensing of Environment*, 198, pp. 105-114.

168) Johansen, T.J. and Mølmann, J.A.B. (2018) 'Seed potato performance after storage in light at elevated temperatures', *Potato Research*, 61(2), pp. 133-145.

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

170) Jung, J., Maeda, M., Chang, A., Bhandari, M., Ashapure, A. and Landivar-Bowles, J. (2021) 'The potential of remote sensing and artificial intelligence as tools to improve the resilience of agriculture production systems', *Current Opinion in Biotechnology*, 70, pp. 15-22.

171) Jurado, J.M., López, A., Pádua, L. and Sousa, J.J. (2022) 'Remote sensing image fusion on 3d scenarios: A review of applications for agriculture and forestry', *International journal of applied earth observation and geoinformation*, 112, p. 102856.

172) Kazimierczak, R., Średnicka-Tober, D., Hallmann, E., Kopczyńska, K. and Zarzyńska, K. (2019) 'The impact of organic vs. Conventional agricultural practices on selected quality features of eight potato cultivars', *Agronomy*, 9(12), p. 799.

173) Kempenaar, C. and Struik, P.C. (2007) 'The canon of potato science: 33. Haulm killing', *Potato Research*, 50(3), pp. 341-345.

174) Khan, Z., Rahimi-Eichi, V., Haefele, S., Garnett, T. and Miklavcic, S.J. (2018) 'Estimation of vegetation indices for high-throughput phenotyping of wheat using aerial imaging', *Plant methods*, 14, pp. 1-11.

175) Kipp, S., Mistele, B. and Schmidhalter, U. (2014) 'The performance of active spectral reflectance sensors as influenced by measuring distance, device temperature and light intensity', *Computers and Electronics in Agriculture*, 100, pp. 24-33.

176) Kirkman, M.A. (2007) 'Chapter 2 - global markets for processed potato products', in Vreugdenhil, D., Bradshaw, J., Gebhardt, C., Govers, F., Mackerron,

D.K.L., Taylor, M.A. and Ross, H.A. (eds.) *Potato biology and biotechnology*. Amsterdam: Elsevier Science B.V., pp. 27-44.

177) Koch, M., Naumann, M. and Pawelzik, E. (2019) 'Cracking and fracture properties of potato (solanum tuberosum l.) tubers and their relation to dry matter, starch, and mineral distribution', *Journal of the science of food and agriculture*, 99(6), pp. 3149-3156.

178) Koch, M., Naumann, M., Pawelzik, E., Gransee, A. and Thiel, H. (2020) 'The importance of nutrient management for potato production part i: Plant nutrition and yield', *Potato Research*, 63(1), pp. 97-119.

179) Kooman, P., Fahem, M., Tegera, P. and Haverkort, A. (1996) 'Effects of climate on different potato genotypes 2. Dry matter allocation and duration of the growth cycle', *European Journal of Agronomy*, 5(3-4), pp. 207-217.

180) Kreuze, J.F., Souza-Dias, J.A.C., Jeevalatha, A., Figueira, A.R., Valkonen, J.P.T. and Jones, R.A.C. (2020) 'Viral diseases in potato', in Campos, H. and Ortiz, O. (eds.) *The potato crop: Its agricultural, nutritional and social contribution to humankind*. Cham: Springer International Publishing, pp. 389-430.

181) Kumar, R., Bhardwaj, A., Singh, L.P. and Singh, G. (2023) 'Quantifying ecological impacts: A comparative life cycle assessment of conventional and organic potato cultivation', *Ecological Modelling*, 486, p. 110510.

182) Kwambai, T.K., Struik, P.C., Gorman, M., Nyongesa, M., Rop, W., Kemboi, E. and Griffin, D. (2023) 'Understanding genotype × environment interactions in potato production to guide variety adoption and future breeding strategies', *Potato Research*.

183) Lacomme, C. and Jacquot, E. (2017) 'General characteristics of potato virus y (pvy) and its impact on potato production: An overview', in Lacomme, C., Glais, L., Bellstedt, D.U., Dupuis, B., Karasev, A.V. and Jacquot, E. (eds.) *Potato virus y: Biodiversity, pathogenicity, epidemiology and management*. Cham: Springer International Publishing, pp. 1-19.

184) Lajoie-O'Malley, A., Bronson, K., van der Burg, S. and Klerkx, L. (2020) 'The future(s) of digital agriculture and sustainable food systems: An analysis of high-level policy documents', *Ecosystem Services*, 45, p. 101183.

185) Laurent, T., Ruiz-Gazen, A. and Thomas-Agnan, C. (2008) 'Geoxp: An r package for exploratory spatial data analysis', *Journal of Statistical Software*, 47.

186) Lee, L.C., Liong, C.-Y. and Jemain, A.A. (2018) 'Partial least squares-discriminant analysis (pls-da) for classification of high-dimensional (hd) data: A review

of contemporary practice strategies and knowledge gaps', *Analyst*, 143(15), pp. 3526-3539.

187) Li, B., Xu, X., Han, J., Zhang, L., Bian, C., Jin, L. and Liu, J. (2019) 'The estimation of crop emergence in potatoes by uav rgb imagery', *Plant Methods*, 15(1), pp. 1-13.

188) Li, B., Xu, X., Zhang, L., Han, J., Bian, C., Li, G., Liu, J. and Jin, L. (2020) 'Above-ground biomass estimation and yield prediction in potato by using uav-based rgb and hyperspectral imaging', *ISPRS Journal of Photogrammetry and Remote Sensing*, 162, pp. 161-172.

189) Li, D., Miao, Y., Gupta, S.K., Rosen, C.J., Yuan, F., Wang, C., Wang, L. and Huang, Y. (2021) 'Improving potato yield prediction by combining cultivar information and uav remote sensing data using machine learning', *Remote Sensing*, 13(16), p. 3322.

190) Li, L., Zhang, Q. and Huang, D. (2014) 'A review of imaging techniques for plant phenotyping', *Sensors (Basel)*, 14(11), pp. 20078-111.

191) Li, M., Shamshiri, R.R., Schirrmann, M., Weltzien, C., Shafian, S. and Laursen, M.S. (2022) 'Uav oblique imagery with an adaptive micro-terrain model for estimation of leaf area index and height of maize canopy from 3d point clouds', *Remote sensing (Basel, Switzerland)*, 14(3), p. 585.

192) Liang, L., Di, L., Zhang, L., Deng, M., Qin, Z., Zhao, S. and Lin, H. (2015) 'Estimation of crop lai using hyperspectral vegetation indices and a hybrid inversion method', *Remote Sensing of Environment*, 165, pp. 123-134.

193) Liebisch, F., Kirchgessner, N., Schneider, D., Walter, A. and Hund, A. (2015) 'Remote, aerial phenotyping of maize traits with a mobile multi-sensor approach', *Plant Methods*, 11(1).

194) Lillesand, T., Kiefer, R.W. and Chipman, J. (2015) *Remote sensing and image interpretation*. John Wiley & Sons.

195) Lin, H., Chen, Z., Qiang, Z., Tang, S.-K., Liu, L. and Pau, G. (2023) 'Automated counting of tobacco plants using multispectral uav data', *Agronomy*, 13(12), p. 2861.

196) Lin, Y.-H., Johnson, D.A. and Pappu, H.R. (2014) 'Effect of potato virus s infection on late blight resistance in potato', *American Journal of Potato Research*, 91(6), pp. 642-648.

197) Liu, N., Zhao, R., Qiao, L., Zhang, Y., Li, M., Sun, H., Xing, Z. and Wang, X. (2020) 'Growth stages classification of potato crop based on analysis of spectral response and variables optimization', *Sensors*, 20(14), p. 3995.

198) Liu, Y., Feng, H., Yue, J., Fan, Y., Jin, X., Song, X., Yang, H. and Yang, G. (2022) 'Estimation of potato above-ground biomass based on vegetation indices and green-edge parameters obtained from uavs', *Remote Sensing*, 14(21), p. 5323.

199) Liu, Y., Yuan, H., Zhao, X., Fan, C. and Cheng, M. (2023) 'Fast reconstruction method of three-dimension model based on dual rgb-d cameras for peanut plant', *Plant Methods*, 19(1), p. 17.

200) Louhaichi, M., Borman, M.M. and Johnson, D.E. (2001) 'Spatially located platform and aerial photography for documentation of grazing impacts on wheat', *Geocarto International*, 16(1), pp. 65-70.

201) Lulai, E.C. (2007) 'Chapter 22 - skin-set, wound healing, and related defects', in Vreugdenhil, D., Bradshaw, J., Gebhardt, C., Govers, F., Mackerron, D.K.L., Taylor, M.A. and Ross, H.A. (eds.) *Potato biology and biotechnology*. Amsterdam: Elsevier Science B.V., pp. 471-500.

202) Lumme, J., Karjalainen, M., Kaartinen, H., Kukko, A., Hyyppä, J., Hyyppä, H., Jaakkola, A. and Kleemola, J. (2008) 'Terrestrial laser scanning of agricultural crops', *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci*, 37, pp. 563-566.

203) Luo, S., He, Y., Li, Q., Jiao, W., Zhu, Y. and Zhao, X. (2020) 'Nondestructive estimation of potato yield using relative variables derived from multi-period lai and hyperspectral data based on weighted growth stage', *Plant Methods*, 16(1), p. 150.

204) Lutaladio, N. and Castaldi, L. (2009) 'Potato: The hidden treasure', *Journal of Food Composition and Analysis*, 22(6), pp. 491-493.

205) Machida-Hirano, R. (2015) 'Diversity of potato genetic resources', *Breeding science*, 65(1), pp. 26-40.

206) Mack, J., Lenz, C., Teutrine, J. and Steinhage, V. (2017) 'High-precision 3d detection and reconstruction of grapes from laser range data for efficient phenotyping based on supervised learning', *Computers and Electronics in Agriculture*, 135, pp. 300-311.

207) Mackerron, D.K.L. and Davies, H.V. (1986) 'Markers for maturity and senescence in the potato crop', *Potato Research*, 29(4), pp. 427-436.

208) Maes, W.H. and Steppe, K. (2019) 'Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture', *Trends in Plant Science*, 24(2), pp. 152-164.

209) Maggio, A., Carillo, P., Bulmetti, G.S., Fuggi, A., Barbieri, G. and De Pascale, S. (2008) 'Potato yield and metabolic profiling under conventional and organic farming', *European Journal of Agronomy*, 28(3), pp. 343-350.

210) Mahlein, A.-K. (2016a) 'Plant disease detection by imaging sensors - parallels and specific demands for precision agriculture and plant phenotyping', *Plant disease*, 100(2), pp. 241-251.

211) Mahlein, A.-K. (2016b) 'Plant disease detection by imaging sensors – parallels and specific demands for precision agriculture and plant phenotyping', *Plant Disease*, 100(2), pp. 241-251.

212) Malambo, L., Popescu, S.C., Murray, S.C., Putman, E., Pugh, N.A., Horne, D.W., Richardson, G., Sheridan, R., Rooney, W.L. and Avant, R. (2018) 'Multitemporal field-based plant height estimation using 3d point clouds generated from small unmanned aerial systems high-resolution imagery', *International Journal of Applied Earth Observation and Geoinformation*, 64, pp. 31-42.

213) Maoka, T. (2020) 'Carotenoids as natural functional pigments', *Journal of Natural Medicines*, 74(1), pp. 1-16.

214) Martinelli, F., Scalenghe, R., Davino, S., Panno, S., Scuderi, G., Ruisi, P., Villa, P., Stroppiana, D., Boschetti, M., Goulart, L., Davis, C. and Dandekar, A. (2015) 'Advanced methods of plant disease detection. A review', *Agronomy for Sustainable Development*, 35(1), pp. 1-25.

215) Matias, F.I., Caraza-Harter, M.V. and Endelman, J.B. (2020) 'Fieldimager: An r package to analyze orthomosaic images from agricultural field trials', *Plant phenome journal*, 3(1), p. n/a.

216) Matsuda, F., Morino, K., Miyazawa, H., Miyashita, M. and Miyagawa, H. (2004) 'Determination of potato glycoalkaloids using high-pressure liquid chromatography-electrospray ionisation/mass spectrometry', *Phytochemical analysis*, 15(2), pp. 121-124.

217) McBratney, A., Whelan, B., Ancev, T. and Bouma, J. (2005) 'Future directions of precision agriculture', *Precision Agriculture*, 6(1), pp. 7-23.

218) McGregor, I. (2007) 'Chapter 1 - the fresh potato market', in Vreugdenhil, D., Bradshaw, J., Gebhardt, C., Govers, F., Mackerron, D.K.L., Taylor, M.A. and Ross, H.A. (eds.) *Potato biology and biotechnology*. Amsterdam: Elsevier Science B.V., pp. 3-26.

219) Meglic, V., Hauptvogel, P., Bilsborrow, P., Janovska, D., Grausgruber, H., Dolnicar, P., Pagnotta, M.A., Petrovic, K., Kuhar, A. and Vogt-kaute, W. (2020) 'Ecobreed: Increasing the efficiency and competitiveness of organic crop breeding'.

220) Meyer, G.E. and Neto, J.C. (2008) 'Verification of color vegetation indices for automated crop imaging applications', *Computers and electronics in agriculture*, 63(2), pp. 282-293.

221) Michael, S., Antje, G., Franziska, G., Michael, P., Jan, L. and Karl-Heinz, D. (2016) 'Monitoring agronomic parameters of winter wheat crops with low-cost uav imagery', *Remote Sensing*, 8(9), p. 706.

222) Mikitzel, L. (2014) 'Tuber physiological disorders', *CABI*, pp. 237–254.

223) Milner, S.E., Brunton, N.P., Jones, P.W., O' Brien, N.M., Collins, S.G. and Maguire, A.R. (2011) 'Bioactivities of glycoalkaloids and their aglycones from solanum species', *Journal of Agricultural and Food Chemistry*, 59(8), pp. 3454-3484.

224) Moghimi, A., Yang, C. and Marchetto, P.M. (2018) 'Ensemble feature selection for plant phenotyping: A journey from hyperspectral to multispectral imaging', *IEEE Access*, 6, pp. 56870-56884.

225) Mølmann, J.A. and Johansen, T.J. (2020) 'Sprout growth inhibition and photomorphogenic development of potato seed tubers (solanum tuberosum l.) under different led light colours', *Potato Research*, 63(2), pp. 199-215.

226) Moran, S., Fitzgerald, G., Rango, A., Walthall, C., Barnes, E., Bausch, W., Clarke, T., Daughtry, C., Everitt, J., Escobar, D., Hatfield, J., Havstad, K., Jackson, T., Kitchen, N., Kustas, W., McGuire, M., Pinter, P., Sudduth, K., Schepers, J.s. and Upchurch, D. (2003) 'Sensor development and radiometric correction for agricultural applications', *Photogrammetric Engineering & Remote Sensing*, 69, pp. 705-718.

227) Morgan, M.R.A., Coxon, D.T., Bramham, S., Chan, H.W.-S., Van Gelder, W.M.J. and Allison, M.J. (1985) 'Determination of the glycoalkaloid content of potato tubers by three methods including enzyme-linked immunosorbent assay', *Journal of the Science of Food and Agriculture*, 36(4), pp. 282-288.

228) Mori, K., Asano, K., Tamiya, S., Nakao, T. and Mori, M. (2015) 'Challenges of breeding potato cultivars to grow in various environments and to meet different demands', *Breeding Science*, 65(1), pp. 3-16.

229) Moysiadis, V., Sarigiannidis, P., Vitsas, V. and Khelifi, A. (2021) 'Smart farming in europe', *Computer Science Review*, 39, p. 100345.

230) Nahar, N., Westerberg, E., Arif, U., Huchelmann, A., Olarte Guasca, A., Beste, L., Dalman, K., Dutta, P.C., Jonsson, L. and Sitbon, F. (2017) 'Transcript profiling of two potato cultivars during glycoalkaloid-inducing treatments shows differential expression of genes in sterol and glycoalkaloid metabolism', *Scientific Reports*, 7(1), p. 43268.

231) Namugga, P., Sibiya, J., Melis, R. and Barekye, A. (2018) 'Yield response of potato (solanum tuberosum l.) genotypes to late blight caused by phytophthora infestans in uganda', *American Journal of Potato Research*, 95, pp. 423-434.

232) Naumann, M., Koch, M., Thiel, H., Gransee, A. and Pawelzik, E. (2020) 'The importance of nutrient management for potato production part ii: Plant nutrition and tuber quality', *Potato Research*, 63(1), pp. 121-137.

233) Nduku, L., Munghemezulu, C., Mashaba-Munghemezulu, Z., Kalumba, A.M., Chirima, G.J., Masiza, W. and De Villiers, C. (2023) 'Global research trends for unmanned aerial vehicle remote sensing application in wheat crop monitoring', *Geomatics*, 3(1), pp. 115-136.

234) Neilson, J.A., Smith, A.M., Mesina, L., Vivian, R., Smienk, S. and De Koyer, D. (2021) 'Potato tuber shape phenotyping using rgb imaging', *Agronomy*, 11(9), p. 1781.

235) Nema, P.K., Ramayya, N., Duncan, E. and Niranjan, K. (2008) 'Potato glycoalkaloids: Formation and strategies for mitigation', *Journal of the Science of Food and Agriculture*, 88(11), pp. 1869-1881.

236) Njane, S.N., Tsuda, S., van Marrewijk, B.M., Polder, G., Katayama, K. and Tsuji, H. (2023) 'Effect of varying uav height on the precise estimation of potato crop growth', *Frontiers in Plant Science*, 14.

237) Nolte, P., Miller, J., Duellman, K.M., Gevens, A.J. and Banks, E. (2020) 'Disease management', in *Potato production systems*. Cham: Springer International Publishing, pp. 203-257.

238) Norman, J.M. and Campbell, G.S. (1989) 'Canopy structure', in Pearcy, R.W., Ehleringer, J.R., Mooney, H.A. and Rundel, P.W. (eds.) *Plant physiological ecology: Field methods and instrumentation*. Dordrecht: Springer Netherlands, pp. 301-325.

239) Nuijten, R.J.G., Kooistra, L. and De Deyn, G.B. (2019) 'Using unmanned aerial systems (uas) and object-based image analysis (obia) for measuring plant-soil feedback effects on crop productivity', *Drones*, 3(3), p. 54.

240) Obidiegwu, J., Bryan, G., Jones, H. and Prashar, A. (2015) 'Coping with drought: Stress and adaptive responses in potato and perspectives for improvement', *Frontiers in Plant Science*, 6(542).

241) Odilbekov, F., Armoniené, R., Henriksson, T. and Chawade, A. (2018) 'Proximal phenotyping and machine learning methods to identify septoria tritici blotch disease symptoms in wheat', *Frontiers in Plant Science*, 9.

242) Oh, S., Chang, A., Ashapure, A., Jung, J., Dube, N., Maeda, M., Gonzalez, D. and Landivar, J. (2020) 'Plant counting of cotton from uas imagery using deep learning-based object detection framework', *Remote Sensing*, 12(18), p. 2981.

243) Okamoto, H., Ducreux, L.J.M., Allwood, J.W., Hedley, P.E., Wright, A., Gururajan, V., Terry, M.J. and Taylor, M.A. (2020) 'Light regulation of chlorophyll and glycoalkaloid biosynthesis during tuber greening of potato s. Tuberosum', *Frontiers in Plant Science*, 11.

244) Olsen, N. and Kleinkopf, G. (2020) 'Storage management', in Stark, J.C., Thornton, M. and Nolte, P. (eds.) *Potato production systems*. Cham: Springer International Publishing, pp. 523-545.

245) Pandey, J., Scheuring, D.C., Koym, J.W., Coombs, J., Novy, R.G., Thompson, A.L., Holm, D.G., Douches, D.S., Miller, J.C. and Vales, M.I. (2021) 'Genetic diversity and population structure of advanced clones selected over forty years by a potato breeding program in the USA', *Scientific Reports*, 11(1), p. 8344.

246) Pantazi, X.E., Moshou, D., Alexandridis, T., Whetton, R.L. and Mouazen, A.M. (2016) 'Wheat yield prediction using machine learning and advanced sensing techniques', *Computers and Electronics in Agriculture*, 121, pp. 57-65.

247) Pathak, H., Igathinathane, C., Zhang, Z., Archer, D. and Hendrickson, J. (2022) 'A review of unmanned aerial vehicle-based methods for plant stand count evaluation in row crops', *Computers and Electronics in Agriculture*, 198, p. 107064.

248) Patil, P., Yaligar, N. and Meena, S. (2017) *2017 IEEE international conference on computational intelligence and computing research (ICCIC)*. IEEE.

249) Pavek, M.J. and Thornton, R.E. (2009) 'Planting depth influences potato plant morphology and economic value', *American Journal of Potato Research*, 86(1), p. 56.

250) Pavlista, A. (2001) 'G1437 green potatoes: The problems and the solution'.

251) Pawelzik, E. and Möller, K. (2014) 'Sustainable potato production worldwide: The challenge to assess conventional and organic production systems', *Potato Research*, 57(3), pp. 273-290.

252) Peddle, D.R., Peter White, H., Soffer, R.J., Miller, J.R. and LeDrew, E.F. (2001) 'Reflectance processing of remote sensing spectroradiometer data', *Computers & Geosciences*, 27(2), pp. 203-213.

253) Peng, Z., Wang, P., Tang, D., Shang, Y., Li, C.-h., Huang, S.-w. and Zhang, C.-z. (2019) 'Inheritance of steroidal glycoalkaloids in potato tuber flesh', *Journal of Integrative Agriculture*, 18(10), pp. 2255-2263.

254) Pentangelo, A., Raimo, F., Parisi, B., Mandolino, G. and Pane, C. (2021) 'Effects of highly concentrated kcl foliar spray for managing the occurrence of the internal brown spot, a physiological disorder of potato tubers', *The Journal of Horticultural Science and Biotechnology*, 96(4), pp. 527-537.

255) Perez-Harguindeguy, N., Diaz, S., Garnier, E., Lavorel, S., Poorter, H., Jaureguiberry, P., Bret-Harte, M.S., Cornwell, W.K., Craine, J.M. and Gurvich, D.E. (2016) 'Corrigendum to: New handbook for standardised measurement of plant functional traits worldwide', *Australian Journal of botany*, 64(8), pp. 715-716.

256) Petermann, J.B. and Morris, S.C. (1985) 'The spectral responses of chlorophyll and glycoalkaloid synthesis in potato tubers (solanum tuberosum)', *Plant Science*, 39(2), pp. 105-110.

257) Phillips, B.J., Hughes, J.A., Phillips, J.C., Walters, D.G., Anderson, D. and Tahourdin, C.S.M. (1996) 'A study of the toxic hazard that might be associated with the consumption of green potato tops', *Food and Chemical Toxicology*, 34(5), pp. 439-448.

258) Pieruschka, R. and Schurr, U. (2019) 'Plant phenotyping: Past, present, and future', *Plant Phenomics*, 2019, p. 7507131.

259) Porra, R.J., Thompson, W.A. and Kriedemann, P.E. (1989) 'Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: Verification of the concentration of chlorophyll standards by atomic absorption spectroscopy', *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 975(3), pp. 384-394.

260) Pound, M.P., Atkinson, J.A., Townsend, A.J., Wilson, M.H., Griffiths, M., Jackson, A.S., Bulat, A., Tzimiropoulos, G., Wells, D.M. and Murchie, E.H. (2017) 'Deep machine learning provides state-of-the-art performance in image-based plant phenotyping', *Gigascience*, 6(10), p. gix083.

261) Prashar, A. and Jones, H. (2014) 'Infra-red thermography as a high-throughput tool for field phenotyping', *Agronomy (Basel)*, 4(3), pp. 397-417.

262) Prashar, A., Yildiz, J., McNicol, J., Bryan, G. and Jones, H. (2013) 'Infra-red thermography for high throughput field phenotyping in solanum tuberosum', *PLoS One*, 8(6), p. e65816.

263) Pugh, N.A., Horne, D.W., Murray, S.C., Carvalho Jr, G., Malambo, L., Jung, J., Chang, A., Maeda, M., Popescu, S. and Chu, T. (2018) 'Temporal estimates of crop growth in sorghum and maize breeding enabled by unmanned aerial systems', *The Plant Phenome Journal*, 1(1), pp. 1-10.

264) Radoglou-Grammatikis, P., Sarigiannidis, P., Lagkas, T. and Moscholios, I. (2020) 'A compilation of uav applications for precision agriculture', *Computer Networks*, 172, p. 107148.

265) Raji, S., Subhash, N., Ravi, V., Saravanan, R., Mohanan, C., MakeshKumar, T. and Nita, S. (2016) 'Detection and classification of mosaic virus disease in cassava plants by proximal sensing of photochemical reflectance index', *Journal of the Indian Society of Remote Sensing*, 44(6), pp. 875-883.

266) Rajpoot, V., Tiwari, A. and Jalal, A.S. (2023) 'Automatic early detection of rice leaf diseases using hybrid deep learning and machine learning methods', *Multimedia Tools and Applications*, 82(23), pp. 36091-36117.

267) Ray, S.S., Jain, N., Arora, R.K., Chavan, S. and Panigrahy, S. (2011) 'Utility of hyperspectral data for potato late blight disease detection', *Journal of the Indian Society of Remote Sensing*, 39(2), p. 161.

268) Rembold, F., Atzberger, C., Savin, I. and Rojas, O. (2013) 'Using low resolution satellite imagery for yield prediction and yield anomaly detection', *Remote Sensing*, 5(4), p. 1704.

269) Rens, L.R., Zotarelli, L., Cantliffe, D.J., Stoffella, P.J., Gergela, D. and Burhans, D. (2015) 'Rate and timing of nitrogen fertilizer application on potato ‘fl1867’ part ii: Marketable yield and tuber quality', *Field Crops Research*, 183, pp. 267-275.

270) Richard, M., Mainassara, Z.-A., Jill, E.C., Cosmos, M., Amsal, T., Mike, O. and Boddupalli, M.P. (2018) 'High-throughput phenotyping of canopy cover and senescence in maize field trials using aerial digital canopy imaging', *Remote Sensing*, 10(2), p. 330.

271) Richardson, A.J. and Wiegand, C. (1977) 'Distinguishing vegetation from soil background information', *Photogrammetric engineering and remote sensing*, 43(12), pp. 1541-1552.

272) Robert, P. (1993) 'Characterization of soil conditions at the field level for soil specific management', *Geoderma*, 60(1), pp. 57-72.

273) Rodríguez, J., Lizarazo, I., Prieto, F. and Angulo-Morales, V. (2021) 'Assessment of potato late blight from uav-based multispectral imagery', *Computers and Electronics in Agriculture*, 184, p. 106061.

274) Rosen, C.J. and Bierman, P.M. (2008) 'Potato yield and tuber set as affected by phosphorus fertilization', *American Journal of Potato Research*, 85, pp. 110-120.

275) Rossi, R., Costafreda-Aumedes, S., Leolini, L., Leolini, C., Bindi, M. and Moriondo, M. (2022) 'Implementation of an algorithm for automated phenotyping through plant 3d-modeling: A practical application on the early detection of water stress', *Computers and Electronics in Agriculture*, 197, p. 106937.

276) Rutkoski, J., Poland, J., Mondal, S., Autrique, E., Pérez, L.G., Crossa, J., Reynolds, M. and Singh, R. (2016) 'Canopy temperature and vegetation indices from high-throughput phenotyping improve accuracy of pedigree and genomic selection for grain yield in wheat', *G3 (Bethesda, Md.)*, 6(9), pp. 2799-2808.

277) Salvador, P., Gómez, D., Sanz, J. and Casanova, J.L. (2020) 'Estimation of potato yield using satellite data at a municipal level: A machine learning approach', *ISPRS International Journal of Geo-Information*, 9(6), p. 343.

278) Sampaio, P.S., Castanho, A., Almeida, A.S., Oliveira, J. and Brites, C. (2020) 'Identification of rice flour types with near-infrared spectroscopy associated with pls-da and svm methods', *European food research and technology*, 246, pp. 527-537.

279) Sarić, R., Nguyen, V.D., Burge, T., Berkowitz, O., Trtílek, M., Whelan, J., Lewsey, M.G. and Čustović, E. (2022) 'Applications of hyperspectral imaging in plant phenotyping', *Trends in Plant Science*, 27(3), pp. 301-315.

280) Scheufele, S. (2022) *Potato, identifying diseases*. Available at: https://ag.umass.edu/vegetable/fact-sheets/potato-identifying-diseases (Accessed: 21 Feruary 2024).

281) Schnabel, S.K., Eilers, P., López, P.H., de Visser, R. and van Eeuwijk, F. (2010) *Proceedings of the 25th International Workshop on Statistical Modelling (IWSM2010), Glasgow*.

282) Schrenk, D., Bignami, M., Bodin, L., Chipman, J.K., Mazo, d.J., Hogstrand, C., Hoogenboom, L., Leblanc, J.C., Nebbia, C.S., Nielsen, E., Ntzani, E., Petersen, A., Sand, S., Schwerdtle, T., Vleminckx, C., Wallace, H., Brimer, L., Cottrill, B., Dusemund, B., Mulder, P., Vollmer, G., Binaglia, M., Ramos Bordajandi, L., Riolo, F.,

Roldán-Torres, R. and Grasl-Kraupp, B. (2020) 'Risk assessment of glycoalkaloids in feed and food, in particular in potatoes and potato-derived products', *EFSA journal*, 18(8).

283) Seufert, V., Ramankutty, N. and Foley, J.A. (2012) 'Comparing the yields of organic and conventional agriculture', *Nature (London)*, 485(7397), pp. 229-232.

284) Shakoor, N., Lee, S. and Mockler, T.C. (2017) 'High throughput phenotyping to accelerate crop breeding and monitoring of diseases in the field', *Current Opinion in Plant Biology*, 38, pp. 184-192.

285) Sharma, S.K., MacKenzie, K., McLean, K., Dale, F., Daniels, S. and Bryan, G.J. (2018) 'Linkage disequilibrium and evaluation of genome-wide association mapping models in tetraploid potato', *G3 Genes|Genomes|Genetics*, 8(10), pp. 3185-3202.

286) Shi, Y., Han, L., Kleerekoper, A., Chang, S. and Hu, T. (2022) 'Novel cropdocnet model for automated potato late blight disease detection from unmanned aerial vehicle-based hyperspectral imagery', *Remote Sensing*, 14(02), p. 396.

287) Shi, Y., Thomasson, J.A., Murray, S.C., Pugh, N.A., Rooney, W.L., Shafian, S., Rajan, N., Rouze, G., Morgan, C.L.S., Neely, H.L., Rana, A., Bagavathiannan, M.V., Henrickson, J., Bowden, E., Valasek, J., Olsenholler, J., Bishop, M.P., Sheridan, R., Putman, E.B., Popescu, S., Burks, T., Cope, D., Ibrahim, A., McCutchen, B.F., Baltensperger, D.D., Avant, R.V., Jr., Vidrine, M. and Yang, C. (2016) 'Unmanned aerial vehicles for high-throughput phenotyping and agronomic research', *PLOS ONE*, 11(7).

288) Shin, J., Mahmud, M.S., Rehman, T.U., Ravichandran, P., Heung, B. and Chang, Y.K. (2023) 'Trends and prospect of machine vision technology for stresses and diseases detection in precision agriculture', *AgriEngineering*, 5(1), pp. 20-39.

289) Sieczka, M. (2001) 'Evaluation of resistance to phytophthora infestans under natural infection pressure', *The methods of evaluation and selection applied in potato research and breeding. Ed. HJ Czembor. Monografie i Rozprawy Naukowe IHAR Radzików, 10a/2001*, pp. 72-74.

290) Singh, I., Srivastava, A.K., Chandna, P. and Gupta, R.K. (2006) 'Crop sensors for efficient nitrogen management in sugarcane: Potential and constraints', *Sugar Tech*, 8(4), pp. 299-302.

291) Sishodia, R.P., Ray, R.L. and Singh, S.K. (2020) 'Applications of remote sensing in precision agriculture: A review', *Remote Sensing*, 12(19), p. 3136.

292) Slater, A.T., Cogan, N.O.I., Rodoni, B.C., Daetwyler, H.D., Hayes, B.J., Caruana, B., Badenhorst, P.E., Spangenberg, G.C. and Forster, J.W. (2017) 'Breeding differently—the digital revolution: High-throughput phenotyping and genotyping', *Potato Research*, 60(3), pp. 337-352.

293) Son, N.T., Chen, C., Chen, C.R., Minh, V. and Trung, N. (2014) 'A comparative analysis of multitemporal modis evi and ndvi data for large-scale rice yield estimation', *Agricultural And Forest Meteorology*, 197, pp. 52-64.

294) Song, C., Ma, W., Li, J., Qi, B. and Liu, B. (2022) 'Development trends in precision agriculture and its management in china based on data visualization', *Agronomy*, 12(11), p. 2905.

295) Sonnewald, S. and Sonnewald, U. (2014) 'Regulation of potato tuber sprouting', *Planta*, 239(1), pp. 27-38.

296) Sourav, A.I. and Emanuel, A.W.R. (2021) 'Recent trends of big data in precision agriculture: A review', *IOP Conference Series: Materials Science and Engineering*, 1096(1), p. 012081.

297) Sousa, C. (2022) 'Anthocyanins, carotenoids and chlorophylls in edible plant leaves unveiled by tandem mass spectrometry', *Foods*, 11(13), p. 1924.

298) Sowokinos, J.R. (2007) 'Chapter 23 - internal physiological disorders and nutritional and compositional factors that affect market quality', in Vreugdenhil, D., Bradshaw, J., Gebhardt, C., Govers, F., Mackerron, D.K.L., Taylor, M.A. and Ross, H.A. (eds.) *Potato biology and biotechnology*. Amsterdam: Elsevier Science B.V., pp. 501-523.

299) Speiser, B., Tamm, L., Amsler, T., Lambion, J., Bertrand, C., Hermansen, A., Ruissen, M.A., Haaland, P., Zarb, J., Santos, J., Shotton, P., Wilcockson, S., Juntharathep, P., Ghorbani, R. and Leifert, C. (2006) 'Improvement of late blight management in organic potato production systems in europe: Field tests with more resistant potato varieties and copper based fungicides', *Biological Agriculture & Horticulture*, 23(4), pp. 393-412.

300) Spoladore, S.F., Brígida dos Santos Scholz, M. and Bona, E. (2021) 'Genotypic classification of wheat using near-infrared spectroscopy and pls-da', *Applied Food Research*, 1(2), p. 100019.

301) Spooner, D.M., McLean, K., Ramsay, G., Waugh, R. and Bryan, G.J. (2005) 'A single domestication for potato based on multilocus amplified fragment length

polymorphism genotyping', *Proceedings of the National Academy of Sciences*, 102(41), pp. 14694-14699.

302) Stark, J.C., Love, S.L. and Knowles, N.R. (2020a) 'Tuber quality', in Stark, J.C., Thornton, M. and Nolte, P. (eds.) *Potato production systems*. Cham: Springer International Publishing, pp. 479-497.

303) Stark, J.C., Thornton, M. and Nolte, P. (2020b) *Potato production systems*. Switzerland: Springer Nature.

304) Storey, M. (2007) 'Chapter 21 - the harvested crop', in Vreugdenhil, D., Bradshaw, J., Gebhardt, C., Govers, F., Mackerron, D.K.L., Taylor, M.A. and Ross, H.A. (eds.) *Potato biology and biotechnology*. Amsterdam: Elsevier Science B.V., pp. 441-470.

305) Struik, P.C. (2007) 'Chapter 18 - responses of the potato plant to temperature', in Vreugdenhil, D., Bradshaw, J., Gebhardt, C., Govers, F., Mackerron, D.K.L., Taylor, M.A. and Ross, H.A. (eds.) *Potato biology and biotechnology*. Amsterdam: Elsevier Science B.V., pp. 367-393.

306) Sugiura, R., Tsuda, S., Tamiya, S., Itoh, A., Nishiwaki, K., Murakami, N., Shibuya, Y., Hirafuji, M. and Nuske, S. (2016) 'Field phenotyping system for the assessment of potato late blight resistance using rgb imagery from an unmanned aerial vehicle', *Biosystems Engineering*, 148, pp. 1-10.

307) Sun, C., Feng, L., Zhang, Z., Ma, Y., Crosby, T., Naber, M. and Wang, Y. (2020) 'Prediction of end-of-season tuber yield and tuber set in potatoes using in-season uav-based hyperspectral imagery and machine learning', *Sensors*, 20(18), p. 5293.

308) Sun, C., Zhou, J., Ma, Y., Xu, Y., Pan, B. and Zhang, Z. (2022) 'A review of remote sensing for potato traits characterization in precision agriculture', *Frontiers in Plant Science*, 13.

309) Suttle, J.C. (2007) 'Chapter 14 - dormancy and sprouting', in Vreugdenhil, D., Bradshaw, J., Gebhardt, C., Govers, F., Mackerron, D.K.L., Taylor, M.A. and Ross, H.A. (eds.) *Potato biology and biotechnology*. Amsterdam: Elsevier Science B.V., pp. 287-309.

310) Tanabe, D., Ichiura, S., Nakatsubo, A., Kobayashi, T. and Katahira, M. (2019) *TAE 2019-Proceeding of 7th International Conference on Trends in Agricultural Engineering 2019*. TAE Prague.

311) Tanios, S., Eyles, A., Tegg, R. and Wilson, C. (2018) 'Potato tuber greening: A review of predisposing factors, management and future challenges', *American Journal of Potato Research*, 95(3), pp. 248-257.

312) Tanios, S., Thangavel, T., Eyles, A., Tegg, R.S., Nichols, D.S., Corkrey, R. and Wilson, C.R. (2020) 'Suberin deposition in potato periderm: A novel resistance mechanism against tuber greening', *The New phytologist*, 225(3), pp. 1273-1284.

313) Tattaris, M., Reynolds, M.P. and Chapman, S.C. (2016) 'A direct comparison of remote sensing approaches for high-throughput phenotyping in plant breeding', *Frontiers in Plant Science*, 7(1131).

314) Taylor, J.A., Chen, H., Smallwood, M. and Marshall, B. (2018) 'Investigations into the opportunity for spatial management of the quality and quantity of production in uk potato systems', *Field Crops Research*, 229, pp. 95-102.

315) Team, R.C. (2021) 'R: A language and environment for statistical computing; r core team: Vienna, austria, 2022', *Available at: www. r-project. org (Accessed February 17, 2022)*.

316) Tein, B., Kauer, K., Runno-Paurson, E., Eremeev, V., Luik, A., Selge, A. and Loit, E. (2015) 'The potato tuber disease occurrence as affected by conventional and organic farming systems', *American Journal of Potato Research*, 92(6), pp. 662-672.

317) ten Harkel, J., Bartholomeus, H. and Kooistra, L. (2020) 'Biomass and crop height estimation of different crops using uav-based lidar', *Remote Sensing*, 12(1), p. 17.

318) Teodoro, P.E., Teodoro, L.P.R., Baio, F.H.R., da Silva Junior, C.A., dos Santos, R.G., Ramos, A.P.M., Pinheiro, M.M.F., Osco, L.P., Gonçalves, W.N., Carneiro, A.M., Junior, J.M., Pistori, H. and Shiratsuchi, L.S. (2021) 'Predicting days to maturity, plant height, and grain yield in soybean: A machine and deep learning approach using multispectral data', *Remote Sensing*, 13(22), p. 4632.

319) Tessema, L., Mohammed, W. and Abebe, T. (2020) 'Evaluation of potato (solanum tuberosum l.) varieties for yield and some agronomic traits', *Open Agriculture*, 5(1), pp. 63-74.

320) Thornton, M. (2020) 'Potato growth and development', in Stark, J.C., Thornton, M. and Nolte, P. (eds.) *Potato production systems*. Cham: Springer International Publishing, pp. 19-33.

321) Thornton, M., Olsen, N. and Liang, X. (2020) 'Physiological disorders', in Stark, J.C., Thornton, M. and Nolte, P. (eds.) *Potato production systems*. Cham: Springer International Publishing, pp. 447-478.

322) Tilahun, S., An, H., Hwang, I., Choi, J., Baek, M., Choi, H., Park, D. and Jeong, C. (2020a) 'Prediction of α-solanine and α-chaconine in potato tubers from hunter color values and vis/nir spectra', *Journal of Food Quality*, 2020.

323) Tilahun, S., An, H., Solomon, T., Baek, M., Choi, H., Lee, H. and Jeong, C. (2020b) 'Indices for the assessment of glycoalkaloids in potato tubers based on surface color and chlorophyll content', *Horticulturae*, 6.

324) Tripodi, P., Massa, D., Venezia, A. and Cardi, T. (2018) 'Sensing technologies for precision phenotyping in vegetable crops: Current status and future challenges', *Agronomy-Basel*, 8(4).

325) Tsedaley, B. (2014) 'Review on early blight (alternaria spp.) of potato disease and its management options', *J Biol Agric Healthc*, 4(27), pp. 191-199.

326) Tucker, C., Holben, B.N., Elgin Jr, J.H. and McMurtrey Iii, J.E. (1980) *Relationship of spectral data to grain-yield variation*.

327) Tunca, E., Köksal, E.S., Taner, S.Ç. and Akay, H. (2024) 'Crop height estimation of sorghum from high resolution multispectral images using the structure from motion (sfm) algorithm', *International Journal of Environmental Science and Technology*, 21(2), pp. 1981-1992.

328) Uarrota, V.G., Moresco, R., Coelho, B., da Costa Nunes, E., Peruch, L.A.M., de Oliveira Neubert, E., Rocha, M. and Maraschin, M. (2014) 'Metabolomics combined with chemometric tools (pca, hca, pls-da and svm) for screening cassava (manihot esculenta crantz) roots during postharvest physiological deterioration', *Food Chemistry*, 161, pp. 67-78.

329) Valderrama, L., Valderrama, P. and Carasek, E. (2022) 'A semi-quantitative model through pls-da in the evaluation of carbendazim in grape juices', *Food Chemistry*, 368, p. 130742.

330) Valente, J., Sari, B., Kooistra, L., Kramer, H. and Mücher, S. (2020) 'Automated crop plant counting from very high-resolution aerial imagery', *Precision Agriculture*, 21, pp. 1366-1384.

331) Vales, M.I., Scheuring, D.C., Koym, J.W., Holm, D.G., Essah, S.Y.C., Wilson, R.G., Sidhu, J.K., Novy, R.G., Whitworth, J.L., Stark, J.C., Spear, R.R., Sathuvalli, V., Shock, C.C., Charlton, B.A., Yilma, S., Knowles, N.R., Pavek, M.J., Brown, C.R.,

Navarre, D.A., Feldman, M., Long, C.M. and Miller, J.C. (2022) 'Vanguard russet: A fresh market potato cultivar with medium-early maturity and long dormancy', *American Journal of Potato Research*, 99(3), pp. 258-267.

332) Van De Vijver, R., Mertens, K., Heungens, K., Nuyttens, D., Wieme, J., Maes, W.H., Van Beek, J., Somers, B. and Saeys, W. (2022) 'Ultra-high-resolution uav-based detection of alternaria solani infections in potato fields', *Remote Sensing*, 14(24), p. 6232.

333) Van De Vijver, R., Mertens, K., Heungens, K., Somers, B., Nuyttens, D., Borra-Serrano, I., Lootens, P., Roldán-Ruiz, I., Vangeyte, J. and Saeys, W. (2020) 'In-field detection of alternaria solani in potato crops using hyperspectral imaging', *Computers and Electronics in Agriculture*, 168, p. 105106.

334) Vilvert, E., Stridh, L., Andersson, B., Olson, Å., Aldén, L. and Berlin, A. (2022) 'Evidence based disease control methods in potato production: A systematic map protocol', *Environmental Evidence*, 11(1), p. 6.

335) Vincini, M., Frazzi, E. and D'Alessio, P. (2008) 'A broad-band leaf chlorophyll vegetation index at the canopy scale', *Precision Agriculture*, 9, pp. 303-319.

336) Virlet, N., Sabermanesh, K., Sadeghi-Tehran, P. and Hawkesford, M.J. (2016) 'Field scanalyzer: An automated robotic field phenotyping platform for detailed crop monitoring', *Functional Plant Biology*, 44(1), pp. 143-153.

337) Visse-Mansiaux, M., Soyeurt, H., Herrera, J.M., Torche, J.-M., Vanderschuren, H. and Dupuis, B. (2022) 'Prediction of potato sprouting during storage', *Field Crops Research*, 278, p. 108396.

338) Walter, A., Liebisch, F. and Hund, A. (2015) 'Plant phenotyping: From bean weighing to image analysis', *Plant Methods*, 11(1), p. 14.

339) Wang, R., Qiu, Y., Zhou, Y., Liang, Z. and Schnable, J.C. (2020) 'A high-throughput phenotyping pipeline for image processing and functional growth curve analysis', *Plant Phenomics*, 2020, p. 7481687.

340) Watanabe, K., Guo, W., Arai, K., Takanashi, H., Kajiya-Kanegae, H., Kobayashi, M., Yano, K., Tokunaga, T., Fujiwara, T., Tsutsumi, N. and Iwata, H. (2017) 'High-throughput phenotyping of sorghum plant height using an unmanned aerial vehicle and its application to genomic prediction modeling', *Front Plant Sci*, 8, p. 421.

341) Watson, C.A., Atkinson, D., Gosling, P., Jackson, L.R. and Rayns, F.W. (2002) 'Managing soil fertility in organic farming systems', *Soil Use and Management*, 18, pp. 239-247.

342) Watt, M., Fiorani, F., Usadel, B., Rascher, U., Muller, O. and Schurr, U. (2020) 'Phenotyping: New windows into the plant for breeders', *Annu Rev Plant Biol*, 71, pp. 689-712.

343) Wellburn, A.R. (1994) 'The spectral determination of chlorophylls a and b, as well as total carotenoids, using various solvents with spectrophotometers of different resolution', *Journal of Plant Physiology*, 144(3), pp. 307-313.

344) White, J.W., Andrade-Sanchez, P., Gore, M.A., Bronson, K.F., Coffelt, T.A., Conley, M.M., Feldmann, K.A., French, A.N., Heun, J.T., Hunsaker, D.J., Jenks, M.A., Kimball, B.A., Roth, R.L., Strand, R.J., Thorp, K.R., Wall, G.W. and Wang, G. (2012) 'Field-based phenomics for plant genetics research', *Field Crops Research*, 133, pp. 101-112.

345) Wishart, J., George, T.S., Brown, L.K., White, P.J., Ramsay, G., Jones, H. and Gregory, P.J. (2014) 'Field phenotyping of potato to assess root and shoot characteristics associated with drought tolerance', *Plant and Soil*, 378(1), pp. 351-363.

346) Wu, J., Wen, S., Lan, Y., Yin, X., Zhang, J. and Ge, Y. (2022) 'Estimation of cotton canopy parameters based on unmanned aerial vehicle (uav) oblique photography', *Plant Methods*, 18(1), p. 129.

347) Xie, C. and Yang, C. (2020) 'A review on plant high-throughput phenotyping traits using uav-based sensors', *Computers and Electronics in Agriculture*, 178, p. 105731.

348) Yang, G., Liu, J., Zhao, C., Li, Z., Huang, Y., Yu, H., Xu, B., Yang, X., Zhu, D., Zhang, X., Zhang, R., Feng, H., Zhao, X., Li, Z., Li, H. and Yang, H. (2017) 'Unmanned aerial vehicle remote sensing for field-based crop phenotyping: Current status and perspectives', *Frontiers in Plant Science*, 8(1111).

349) Yang, H., Hu, Y., Zheng, Z., Qiao, Y., Zhang, K., Guo, T. and Chen, J. (2022) 'Estimation of potato chlorophyll content from uav multispectral images with stacking ensemble algorithm', *Agronomy*, 12(10), p. 2318.

350) Yang, S., Li, L., Fei, S., Yang, M., Tao, Z., Meng, Y. and Xiao, Y. (2024) 'Wheat yield prediction using machine learning method based on uav remote sensing data', *Drones*, 8(7), p. 284.

351) Yu, Q., Gong, P., Clinton, N., Biging, G., Kelly, M. and Schirokauer, D. (2006) 'Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery', *Photogrammetric Engineering & Remote Sensing*, 72(7), pp. 799-811.

352) Yu, T., Zhou, J., Fan, J., Wang, Y. and Zhang, Z. (2023) 'Potato leaf area index estimation using multi-sensor unmanned aerial vehicle (uav) imagery and machine learning', *Remote Sensing*, 15(16), p. 4108.

353) Yuan, L., Pu, R., Zhang, J., Wang, J. and Yang, H. (2016) 'Using high spatial resolution satellite imagery for mapping powdery mildew at a regional scale', *Precision Agriculture*, 17(3), pp. 332-348.

354) Zaman-Allah, M., Vergara, O., Araus, J.L., Tarekegne, A., Magorokosho, C., Zarco-Tejada, P.J., Hornero, A., Albà, A.H., Das, B., Craufurd, P., Olsen, M., Prasanna, B.M. and Cairns, J. (2015) 'Unmanned aerial platform-based multi-spectral imaging for field phenotyping of maize', *Plant Methods*, 11(1), p. 35.

355) Zarzyńska, K. and Boguszewska-Mańkowska, D. (2024) 'Commercial quality of potato tubers of different varieties from organic and conventional production system', *Agronomy*, 14(4), p. 778.

356) Zebarth, B.J., Moreau, G., Dixon, T., Fillmore, S., Smith, A., Hann, S. and Comeau, L.-P. (2022) 'Soil properties and topographic features influence within-field variation in potato tuber yield in new brunswick, canada', *Soil Science Society of America Journal*, 86(1), pp. 134-145.

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

358) Zhang, J., Naik, H.S., Assefa, T., Sarkar, S., Reddy, R.C., Singh, A., Ganapathysubramanian, B. and Singh, A.K. (2017) 'Computer vision and machine learning for robust phenotyping in genome-wide studies', *Scientific Reports*, 7(1), p. 44048.

359) Zhou, X., Xing, M., He, B., Wang, J., Song, Y., Shang, J., Liao, C., Xu, M. and Ni, X. (2023) 'A ground point fitting method for winter wheat height estimation using uav-based sfm point cloud data', *Drones*, 7(7), p. 406.

360) Zhou, X., Zheng, H.B., Xu, X.Q., He, J.Y., Ge, X.K., Yao, X., Cheng, T., Zhu, Y., Cao, W.X. and 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, pp. 246-255.

361) Zhu, W., Chen, H., Ciechanowska, I. and Spaner, D. (2018) 'Application of infrared thermal imaging for the rapid diagnosis of crop disease', *IFAC-PapersOnLine*, 51(17), pp. 424-430.

362) Zhu, Y.S., Merkle-Lehman, D.L. and Kung, S.D. (1984) 'Light-induced transformation of amyloplasts into chloroplasts in potato tubers 1', *Plant Physiology*, 75(1), pp. 142-145.

363) Ziamtsov, I. and Navlakha, S. (2019) 'Machine learning approaches to improve three basic plant phenotyping tasks using three-dimensional point clouds1[open]', *Plant Physiology*, 181(4), pp. 1425-1440.

364) Žibrat, U., Susič, N., Knapič, M., Širca, S., Strajnar, P., Razinger, J., Vončina, A., Urek, G. and Gerič Stare, B. (2019) 'Pipeline for imaging, extraction, pre-processing, and processing of time-series hyperspectral data for discriminating drought stress origin in tomatoes', *MethodsX*, 6, pp. 399-408.