UNMANNED AERIAL VEHICLE (UAV) HIGH-THROUGHPUT PHENOTYPING (HTP) OF SELECTED NEGLECTED AND UNDER-UTILISED CROP SPECIES (NUS) FOR IMPROVED WATER USE AND PRODUCTIVITY ON SMALLHOLDER FARMS

Report to the

WATER RESEARCH COMMISSION

by

M Sibanda[1], O Mutanga[2], J O Odindi[2], V G.P Chimonyo[2,3], R Kunz[2], A Clulow[2], J Magidi[4], T Matongera[2], H S Ndlovu[2], S Buthelezi[2], M Abrahams[1], Y Kamteni[1], R Masemola[1], T Mabhaudhi[2,5]

[1]University of the Western Cape, Faculty of Arts and Humanities
[2]University of KwaZulu-Natal, School of Agriculture and Science
[3]International Maize and Wheat Improvement Centre (CIMMYT)-Zimbabwe
[4]Tshwane University of Technology, Faculty of Engineering and the Built Environment
[5]Centre on Climate Change and Planetary Health, London School of Hygiene and Tropical Medicine

WRC report no. 3251/1/26
ISBN 978-0-6392-0793-3

May 2026

This is the final report of WRC project no. C2022/2023-00930.

Obtainable from:
Water Research Commission
Private Bag X03
Gezina, 0031
South Africa

Download from www.wrc.org.za

EXECUTIVE SUMMARY

South Africa's population is projected to reach 87 million by 2050, which will intensify the food and nutrition insecurity and increase the pressure on the already limited land and water resources. A significant proportion of the population is both water- and food-insecure, particularly those living in marginal environments, who rely on rainfed smallholder agriculture for their subsistence. Maize, the dominant staple across southern Africa, performs poorly in these systems, due to water deficits, episodic flooding and climate variability, which render many production environments unsuitable.

Neglected and Under-utilised Crop Species (NUS), including sorghum, cowpeas, amaranth, sweet potatoes and taro, offer an alternative crop because of their inherent tolerance to droughts and flooding. However, these crops remain under-researched and are primarily cultivated by smallholder farmers who are familiar with their adaptive traits.

Smallholder farming systems are highly heterogeneous, yet current production guidelines are overly generalised and are not adequately resilient to climate extremes. Improved, evidence-based insights derived from rapid and accurate field data are needed. Crop phenotyping is central to optimising crop management; however, the conventional methods are manual, laborious, subjective and lack spatial detail. While satellite remote sensing offers non-invasive phenotyping capabilities, its use in smallholder systems is constrained by their spatial and temporal heterogeneity and the high cost of fine-resolution imagery.

Proximal remote sensing, using UAV-based high-throughput phenotyping platforms, has emerged as a cost-effective, rapid and efficient alternative for acquiring detailed crop data. Such platforms have proven themselves to be valuable for characterising NUS phenotypes. This project has sought to apply UAV imagery and crop simulation models to monitor the crop health and to quantify the yields of selected NUS in smallholder systems, more specifically, sweet potatoes, taro and Bambara groundnuts. The near-real-time monitoring of the crop–soil water status will support drought assessment and inform irrigation decision-making. This project extended the objective of WRC Project No. K5/2717//4, which focuses on estimating the water use of crops by integrating UAV-based phenotyping to advance precision agriculture in the smallholder context.

This project aimed to assess the application of Unmanned Aerial Vehicle (UAV)-based high-throughput phenotyping for Neglected and Under-utilised crop Species (NUS), in order to improve the water use and productivity on smallholder farms. To achieve the overarching objective, the project set out the following specific objectives:

1) to review the literature on the utility of earth observation data in characterising the productivity of Neglected and Under-utilised crop Species (NUS) on smallholder farms;
2) to map the spatial distribution and health of neglected and under-utilised crop species on smallholder farms, using UAV data;
3) to assess the relationship between the NUS crop productivity variations and the edaphic factors (i.e. soil moisture, temperature and salinity) on smallholder farms;

4) to evaluate the use of unmanned aerial vehicle platforms in quantifying the spatial and temporal variability of canopy crop water status and yield by NUS crops on smallholder farms; and
5) to assess the potential of unmanned aerial vehicle high-throughput phenotyping data in estimating the yield of NUS crops in smallholder croplands.

New Knowledge and Innovation

This project aimed to address a critical knowledge gap in agricultural research regarding the limited use of drone-based remote sensing for the mapping and monitoring of Neglected and Under-utilised Crop species, with particular focus on sweet potatoes and taro grown in smallholder farming systems. A longstanding challenge in Applied Geospatial Applications in smallholder croplands of southern Africa is the limited availability of accessible spatial data that are suitable for mapping and monitoring small, fragmented and heterogeneous croplands owned by subsistence and smallholder farms. To bridge this gap, the project utilised drone technologies to generate new, practical insights into how these crops can be identified, assessed and monitored at a fine spatial resolution.

Through applied research, the project generated models, techniques and high-resolution spatial maps, with resolutions finer than 10 cm, that clearly depict the spatial distribution of these often-overlooked crops within smallholder croplands. Beyond mapping, the drone imagery was used to derive the key phenotypic crop attributes, including their Leaf Area index, chlorophyll content and foliar moisture levels. These are valuable indicators for assessing a crop's health and performance.

In addition, the project contributed significantly to the scientific literature by publishing its findings in internationally peer-reviewed journals and by combining literature reviews with applied research articles. These articles contribute to the global discussion on the practical use of drones in agriculture and they highlight the potential of NUS crops as a viable alternative to mainstream crops, in the smallholder farming context.

Overall, the project integrates drone technology with agricultural research to develop innovative approaches for the discrimination, mapping and phenotyping of NUS crops, by offering a valuable framework for improving the monitoring, management and promotion of these important, but under-represented, agricultural resources in smallholder systems.

Capacity Building

The project allocated funding for three full-time MSc students over its four-year span. It also employed one Postdoctoral Fellow, whose role was vital for the University of KwaZulu-Natal's capacity development goals for early-career researchers, which focus on nurturing the next generation of scientific talent. In line with its planned commitments, three Master's candidates were recruited. Two graduated cum laude, while the third is submitting her thesis this year. Subsequently, a fourth MA student was recruited and supported for the fieldwork. Recognising the strategic importance of increased postgraduate engagement, the project team included two female PhD students from the previous Flagship WRC Research Project (K5/2971//4). One of

these students has graduated, while the other is expected to submit her thesis before the end of the year. These doctoral candidates received partial funding from the National Research Foundation (NRF). In addition, the first MA student, who completed her degree in 2023, was recruited as a PhD student. Beyond student development, the Project Leader, who is a mid-career researcher, was also enhanced through this project. His NRF rating improved from Y2 to C2, and he received a promotion at UWC from being a Senior Lecturer to an Associate Professor, following the graduation of the students from this project.

Conclusion

The overall goal of this project was to evaluate the use of Unmanned Aerial Vehicle (UAV)-based high-throughput phenotyping for neglected and under-utilised crop species to improve the water use and productivity on smallholder farms. In a systematic review of the literature on the progress, challenges, gaps and opportunities associated with using drone-derived remotely sensed data to map the spatial distribution and health of NUS crops, to specifically address Objective 1, it was found that studies, especially those regarding spatial mapping, remain limited. Most research focuses on assessing crop health and productivity, while barriers such as the high costs, regulatory restrictions and skill shortages further hinder their widespread adoption. Despite these issues, drones have great potential for providing timely and accurate data, which are essential for farm-level decision-making. Successfully integrating them into smallholder systems could empower farmers, preserve the local food cultures, and support broader goals, such as gender equity and biodiversity conservation. Our analysis also showed that machine learning algorithms (GTB and RF) can effectively delineate NUS crop boundaries on smallholder farms by using UAV data. The GTB algorithm proved particularly resilient when combining spectral datasets, and it outperformed other methods for crops like sweet potatoes and taro. Regarding the relationship between NUS crop productivity variations and the edaphic factors on smallholder farms, which specifically address Objective 2, we evaluated the effectiveness of thermal remote sensing and index-based segmentation techniques for predicting the EWT canopy in smallholder taro fields. It was concluded that combining UAV thermal data with multispectral imagery and the ExGR segmentation technique is optimal for estimating the water content of the taro canopy. Moreover, the project demonstrates that integrating UAV thermal and multispectral data offers an efficient way to assess the water status of smallholder taro crops. The most accurate predictive models for key indicators, such as the equivalent water thickness and stomatal conductance, were obtained by processing the raw sensor data alongside the thermal and spectral indices. This multi-modal approach provided highly reliable results, which highlight its practical utility for precision farming. Using drone data and Random Forest regression to predict the spatial fluctuations of the Leaf Chlorophyll Content (LCC) and Canopy Chlorophyll Content (CCC) for taro and sweet potatoes, it was concluded that CCC is best modelled during mid-vegetative growth, by using NIR and red-edge wavelengths. In contrast, the LCC prediction varied by crop and growth stage. The most precise CCC estimates were achieved by combining spectral bands with vegetation indices. Overall, the CCC proved to be a more reliably predicted variable than the LCC for both crops. Concerning the use of UAV multispectral data for estimating the taro Above-Ground Biomass (AGB) and the Below-Ground Biomass (BGB), this study identified

two key findings. Firstly, it established a strong positive relationship between the AGB and BGB over the growing season, as captured by UAV imagery. Most notably, the data could effectively predict the yield-critical BGB before harvest during the vegetative phase, which provided a valuable tool for the pre-harvest yield estimation. Lastly, this project confirms that the early taro yield can be accurately forecast on smallholder farms by using UAV data, which addresses the specific objective. The most dependable method combines segmented imagery from the red, red-edge and NIR bands with the Leaf Area Index (LAI) measurements, which are processed through a Random Forest algorithm. The results show that image segmentation and canopy architecture data are vital for enhancing prediction accuracy.

Overall, the project demonstrates that UAV-based high-throughput phenotyping is a powerful tool for improving the water use and productivity of neglected and under-utilised crops on smallholder farms. By developing and validating the specific methodologies, it demonstrated that drones can accurately map the crop distribution, predict the critical water status indicators, estimate the biomass, and forecast the final yield, for crops such as taro and sweet potatoes. The integration of multispectral and thermal sensors with advanced machine learning and image segmentation techniques was key for generating reliable, farm-level data for better agricultural decision-making. Ultimately, these findings provide a practical framework for empowering smallholder farmers, for optimising resource use and for enhancing the sustainability of local food systems that are centred on these vital crops. By generating precise, actionable data on the crop status and resource use, UAVs may enable farmers to optimise their water management and boost productivity, which are foundational steps toward improved food security and climate adaptation.

Recommendations for Future Studies

Despite the success of the project, several limitations and research gaps remain. The project's limitations span the methodological (analytical techniques, data fusion), scope (crop variety, growth stages, nutrient modelling), and technological (sensor resolution, noise filtering) areas, which highlight multiple avenues for future research. Future studies could consider the following,

- This project was constrained to two crops, namely, taro and sweet potatoes. Future UAV-based high-throughput phenotyping projects could assess the influence of different varieties of taro and sweet potatoes
- In this project, the classification procedure was limited to machine learning algorithms in the Google Earth Engine. However, the potential of big data analytics, including deep machine learning algorithms and artificial intelligence, remains untested in modelling the spatial distribution of these smallholder crops, with a particular focus on the NUS in Africa.
- In order to translate the canopy water status models into practical tools, the research must validate the canopy water status models against the real-world water inputs. A critical next step is to apply these models in order to compare the water-use dynamics of different NUS crop varieties under both controlled irrigation and natural rainfed conditions.

- Future yield estimation models should incorporate a wider range of canopy traits across the entire growing season, in order to account for critical phenological variability. This expansion of input data is expected to improve a model's robustness and accuracy, but it requires systematic evaluation, in order to quantify the specific contribution of each trait.
- In addition, to fully achieve the goal of high-resolution, field-scale yield mapping, the data quality must be improved at the end of the season. Investing in higher-resolution sensors and advanced noise-filtering algorithms is recommended. This will ensure that yield models are built on precise canopy data throughout the year, which will lead to more accurate and actionable forecasts.
- There is a need to bridge the agronomic-nutritional gap. Future studies should prioritise research that models the seasonal relationship between the water use and nutrient content in taro and sweet potatoes. This is necessary to optimise crops for their nutritional density and to build a credible 'farm-to-health' narrative for marketing NUS crops.
- To enable the large-scale phenotyping of NUS crops in smallholder landscapes, future research must actively integrate drone and satellite datasets. Specifically, studies should leverage the synergies between high-resolution drone imagery and freely-available, frequent Sentinel-2 multispectral imager and Landsat data. This fusion is key for developing robust models that can accurately upscale the critical crop traits, from field plots to regional croplands, and for optimising monitoring where it is needed most.

ACKNOWLEDGMENTS

The research presented here was a component of an unsolicited project that was initiated, funded and overseen by the Water Research Commission (WRC). The project team expresses its sincere gratitude to the WRC for its funding and project management. The project team also wishes to sincerely thank the following members of the Reference Group for their valuable contributions and guidance:

Dr L Nhamo	:	Water Research Commission (Project Manager)
Prof S Mpandeli	:	Water Research Commission
Dr S Hlophe-Ginindza	:	Water Research Commission
Prof K Ayisi	:	University of Limpopo
Dr H Araya	:	Agricultural Research Council
Dr I du Plooy		Agricultural Research Council
Dr S Figlan	:	University of South Africa
Dr T Masupha	:	Agricultural Research Council
Mr T Newby		Geo Terra
Dr E Z Mashimbye	:	Stellenbosch University
Dr Z Ntshidi	:	National Research Foundation-South African Environmental Observation Network: Arid Lands Node
Dr BCE le Roux		Food and Water Research
Dr A Nyamugama		Agricultural Research Council- NRM

We would also like to thank the following individuals:

- the South African Research Chair (SARChI) in Land-use management in the School of Agriculture and Science, Discipline of Geography, University of KwaZulu-Natal, for its logistical support and leverage funding;
- the uMgeni Resilience Project, funded by the Adaptation Fund, for partially supporting the fieldwork as part of building the resilience of smallholder farmers to climate change;
- the postgraduate students who participated in the project: Dr Helen S Ndlovu, Siphiwokuhle Buthelezi, Mishkah Abrahams, Yola Kamteni, Esethu Bacela, Reitumetse Masemola and Dimpho Ndala;
- the administrative staff from the Centre for Transformative Agricultural and Food Systems (CTAFS) and the Centre for Water Resources Research at UKZN, the Project administrators at UWC (D Ndala), as well as the WRC (S Fritz), for supporting the project;
- the community members from Swayimane, for allowing the field trials on their land;
- Dr Trylee Matongera, Dr Shaeden Gokool and Mr Vivek Naiken for their assistance with image acquisition; and
- the many stakeholders whose contributions have laid the groundwork for the data used in the report.

TABLE OF CONTENTS

EXECUTIVE SUMMARY iii
ACKNOWLEDGMENTS viii
LIST OF FIGURES xvi
LIST OF TABLES xix
ABBREVIATIONS xx
REPOSITORY OF DATA xxi
1 Introduction and Background 1
1.1 Introduction 1
1.2 Aims and Objectives 2
1.3 Scope and the Overview of the Report 3
1.3.1 Description of the study area 3
1.3.2 Report outline 4
2 A Systematic Review of Remote Sensing Applications for Neglected and Under-utilised Crop Species 6
2.1 Introduction 6
2.2 Methods 8
2.2.1 Phase One: Literature search 8
2.2.2 Phase Two: Data extraction 9
2.2.3 Phase Three: Data analysis 10
2.3 Results 10
2.3.1 Searched literature characteristics 10
2.3.2 Progress in mapping the spatial distribution and health status of neglected and under-utilised crop species 14
2.4 Discussion 23
2.4.1 Evolution of drone technology applications in remote sensing 23
2.4.2 Frequency of publication and its geographic distribution 24
2.4.3 NUS crop attributes that have been remotely sensed using drone-acquired data 24
2.4.4 Sensors and platforms used in remote sensing NUS 27
2.4.5 Performance of vegetation indices, classification and estimation algorithms 29
2.4.6 The challenges and opportunities of mapping the spatial distribution and health of NUS crops 30

2.4.7 Research gaps and opportunities....31
2.4.8 The way forward: Closing the gaps in the utilisation of drone technology in mapping the spatial distribution and health status of NUS crops....31
2.5 Conclusion....32
3 Mapping Taro and Sweet Potatoes on smallholder farms Using UAV Multispectral Imagery and Machine Learning....34
3.1 Introduction....34
3.2 Methods and Materials....36
3.2.1 Data processing....37
3.2.2 Image classification....39
3.2.3 Accuracy assessment....41
3.3 Results....41
3.3.1 Spectral reflectance curve....41
3.3.2 Comparative cropland classification using SVM, RF and GTB, based on bands only 42
3.3.3 Comparative cropland classification using SVM, RF and GTB-based vegetation indices only....44
3.3.4 Comparative performance of spectral variables....44
3.3.5 Comparative classification performance of SVM, RF and GTB....45
3.3.6 Final classification of the cropland using combined data....46
3.3.7 Spatial distribution of land cover types and their areal extents....48
3.4 Discussion....49
3.4.1 The comparative performance of raw spectral bands and vegetation indices in mapping the spatial distribution of NUS in the smallholder cropland....49
3.4.2 The comparative performance of machine learning algorithms in mapping the spatial distribution of NUS in the smallholder cropland....50
3.5 Conclusion....52
4 Using UAV Multispectral-Thermal Data and Index-Based Segmentation to Enhance Equivalent Water Thickness Estimation in Taro Crops....53
4.1 Introduction....53
4.2 Materials and Methods....56
4.2.1 Field sampling and in-situ measurements....56
4.2.2 UAV platform and multispectral-thermal camera....56
4.2.3 Image acquisition and pre-processing....57
4.2.4 Index-based image segmentation of the taro crops' spectral signatures....58

4.2.5 Model development and statistical analysis 59
4.2.6 Accuracy assessment 60
4.3 Results 62
4.3.1 Descriptive statistics of the in-situ EWT-canopy of taro crops 62
4.3.2 Predicting EWT canopy of taro crops using the Deep Neural Network algorithm and image segmentation techniques 63
4.3.3 Assessment of the influence of the thermal band in conjunction with the Deep Neural Network algorithm 65
4.3.4 Spatial distribution of the EWT canopy of smallholder taro crops 66
4.4 Discussion 68
4.4.1 Prediction of EWT canopy using UAV multispectral thermal imagery 68
4.4.2 Performance of the thermal band in predicting the EWT canopy of taro crops 69
4.4.3 Performance of index-based segmentation techniques for the estimation of the taro EWT canopy 69
4.4.4 Implications of the study, its limitations and recommendations for future research 70
4.5 Conclusion 70
5 Assessing the Water Status of Neglected and Under-utilised Taro Crops Using Physiological Indicators and UAV Multispectral-Thermal Data 72
5.1 Introduction 72
5.2 Materials and Methods 74
5.2.1 Field sampling and in-situ measurements 74
5.2.2 Equivalent water thickness and fuel moisture content 74
5.2.3 Stomatal conductance 75
5.2.4 Foliar temperature 75
5.2.5 Chlorophyll content 75
5.2.6 Image acquisition and pre-processing 76
5.2.7 Crop canopy index-based segmentations 78
5.2.8 Model development and statistical analysis 79
5.2.9 Accuracy assessment 80
5.2.10 A methodological overview on assessing the influence of thermal spectral data in characterising crop water status physiological indicators 80
5.3 Results 82
5.3.1 Descriptive analysis of taro crop water status indicators 82

5.3.2 Estimation of the taro crop water status physiological indicators using selected spectral variables....82
5.3.3 Optimised regression models of taro crop water status physiological indicators integrating thermal and multispectral data....84
5.3.4 Mapping the spatial distribution of the physiological indicators of the taro crop water status....86
5.4 Discussion....87
5.4.1 The performance of integrating multispectral and thermal bands with thermal and spectral indices....87
5.4.2 The performance of selected spectral variables in estimating the crop water status, physiological indicators of smallholder taro crops....89
5.4.3 Implications of the study limitations and recommendations for future research 90
5.5 Conclusion....91
6 UAV-Based Estimation of Canopy Chlorophyll Content for Taro and Sweet Potatoes on smallholder farms....92
6.1 Introduction....92
6.2 Materials and Methods....94
6.2.1 Foliar chlorophyll content sampling strategy....94
6.2.2 Measuring leaf chlorophyll content of taro and sweet potatoes....94
6.2.3 Leaf area index and canopy chlorophyll content....95
6.2.4 Acquisition and processing of remotely sensed data using an UAV....95
6.2.5 Estimation of chlorophyll content by using the Random Forest....96
6.2.6 Accuracy assessment....96
6.2.7 Estimation of chlorophyll content of taro and sweet potatoes....97
6.3 Results....98
6.3.1 Descriptive analysis....98
6.3.2 Comparison of the LCC and CCC estimations of spectral variables in taro and sweet potatoes....98
6.3.3 Optimised regression models of taro and sweet potato canopy chlorophyll content at different phenological stages....101
6.3.4 Comparing the leaf and canopy chlorophyll content estimations....106
6.3.5 Spatial distribution of the chlorophyll content of the sweet potato and taro crops at selected growth stages....106
6.4 Discussion....108

6.4.1 Comparative performance of Vegetation Indices, bands and combined datasets 108

6.4.2 Estimation of taro and sweet potato leaf chlorophyll content at different phenological stages 109

6.4.3 Estimation of the taro and sweet potato canopy chlorophyll content at different phenological stages 110

6.4.4 Comparing the leaf and canopy chlorophyll content estimations 112

6.5 Conclusion 113

..... EVALUATING DRONE-BASED REMOTE SENSING FOR MONITORING THE SOIL WATER CONTENT AND LEAF CHLOROPHYLL IN TARO (*COLOCASIA ESCULENTA*) ACROSS PHENOLOGICAL STAGES 114

7 114

7.1 Introduction 114

7.2 Methods and Materials 116

7.2.1 Field data collection 116

7.2.2 Spectral variables used for the prediction of the soil moisture content and chlorophyll content of taro crops 116

7.2.3 Statistical analysis 118

7.2.4 Accuracy assessment 118

7.3 Results 118

7.3.1 Descriptive statistics 118

7.3.2 Comparing the performance of bands, vegetation indices and combined data in estimating the soil moisture and chlorophyll content 120

7.3.3 Estimating soil moisture and chlorophyll content at different phenological stages 120

7.3.4 Spatial variation of soil moisture content and chlorophyll content at different growth stages 125

7.4 Discussion 127

7.4.1 Association between the soil moisture content and chlorophyll 127

7.4.2 Optimal estimation of soil moisture and chlorophyll across the phenological stages 128

7.4.3 Comparative performance of bands, vegetation indices and combined data sets in estimating the soil moisture and chlorophyll across the phenological stages 128

7.4.4 Implications of the study 129

7.5 Conclusion 129

8 Estimating THE Taro Below-Ground Biomass in Smallholder Croplands Using UAV Multispectral Imagery 130
8.1 Introduction 130
8.2 Materials and Methods 131
8.2.1 Study area 131
8.2.2 Field sampling and taro above-ground biomass and below-ground biomass/yield measurements 131
8.2.3 UAV platform, image acquisition and processing 131
8.2.4 Statistical analysis and model development for biomass and yield prediction 133
8.2.5 Selection of Vegetation Indices 133
8.2.6 Taro AGB and BGB prediction and accuracy assessment 134
8.3 Results 135
8.3.1 Descriptive analysis of AGB and BGB 135
8.3.2 Correlation analysis between the AGB and BGB of taro 135
8.3.3 Predicting the AGB and BGB of taro using the Random Forest regression ensemble 136
8.3.4 Modelling the spatial distribution of the AGB and BGB of taro 139
8.4 Discussion 140
8.4.1 Characterising the taro AGB and BGB across the growing season using UAV-derived data 141
8.4.2 The most optimal stage and variables for the prediction of taro's BGB yield 142
8.4.3 Implications of the findings 143
8.4.4 Limitations and recommendations 143
8.5 Conclusion 143
9 Integrating Image Segmentation and Leaf Area Index with UAV Multispectral Data to Improve Taro Yield Prediction on smallholder farms 145
9.1 Introduction 145
9.2 Materials and Methods 146
9.2.1 Image acquisition and pre-processing 146
9.2.2 Field crop assessment 147
9.2.3 Image processing and data extraction 147
9.2.4 Vegetation indices 148
9.2.5 Data analysis 148
9.2.6 Accuracy assessment of taro yield 149

9.2.7 Comparing the performance of different datasets 149
9.3 Results 150
9.3.1 Ground truth data 150
9.3.2 Evaluation of taro yield regression models 150
9.3.3 Mapping the spatial distribution of the taro yield 151
9.4 Discussion 152
9.4.1 Prediction of taro yield in smallholder croplands 152
9.4.2 Assessing the impact of image segmentation on prediction accuracies 153
9.4.3 Implications of the study 153
9.5 Conclusion 154
10 General discussion, conclusions and recommendations 155
10.1 General Discussion 156
10.2 Limitations 158
10.3 General Conclusion 158
10.4 Recommendations 160
Appendices 184
INFORMATION DISSEMINATION: COMPLETED, SUBMITTED AND PUBLISHED MANUSCRIPTS 190
Research Articles for Publication 190
Conference attendance: 190
Published Popular articles 191
Session Chair 192
CAPACITY BUILDING REPORT 193
Postgraduate student recruitment 193
Professional Advancement of the Project Leader 193

LIST OF FIGURES

Figure 1-1 Aerial diagram of Swayimane, KwaZulu-Natal (Source :Sibanda et al. (2023)) 4
Figure 2-1 PRISMA flow diagram for selection of studies considered in the review.....9
Figure 2-2 Changes in topical concepts associated with the remote sensing productivity of NUS crops based on titles.....11
Figure 2-3 Topical concepts identified by using a bibliometric analysis of titles and abstracts of articles that utilised (a) all remote sensing sensors, and (b) only drone-acquired data in mapping the NUS.....13
Figure 2-4 Frequency of published articles on remote sensing applications of NUS, based on (a) all sensors, and (b) UAVs.....14
Figure 2-5 Spatial distribution of studies on remote sensing, the attributes and spatial distribution of NUS.....15
Figure 2-6 Frequency of studies that remotely sensed a specific crop attribute based on (a) all satellite and drone-borne sensors, and drone-borne sensors only.....16
Figure 2-7. Frequency of (a) satellite, and (b) drone-borne sensors that have been used to map the spatial distribution of NUS and their attributes.....17
Figure 2-8 Frequency of the utilisation of a specific section of the electromagnetic spectrum covered by (a) drone and (b) satellite-borne sensors.....18
Figure 2-9 Frequency of (a) summarised drone platforms, (b) and specific drone types that were utilised in the literature to map the spatial distribution of NUS and their health attributes 19
Figure 2-10 Frequency of statistical methods used in remote sensing NUS attributes, based on (a) all sensors and (b) drone-borne sensors only 20
Figure 2-11 Frequency of machine learning and general regression techniques used in remote sensing NUS attributes, based on (a) all sensors and (b) drone-borne sensors only ...21
Figure 2-12 Frequency of generic GIS classification techniques used in remote sensing NUS attributes, based on (a) all sensors and (b) drone-borne sensors only 22
Figure 2-13 Frequency of multivariate techniques used in remote sensing NUS attributes, based on (a) all sensors and (b) drone-borne sensors only 22
Figure 2-14 Evolution of concepts in drone-based remote sensing of NUS attributes.....23
Figure 3-1 Flowchart of main processing steps in this study 36
Figure 3-2 Spectral reflectance curve of all classes.....42
Figure 3-3: Comparative classification performance of bands, vegetation indices and combined data 45
Figure 3-4 Comparative classification performance of Support Vector Machines (SVM), Random Forest (RF) and Gradient Tree Boosting algorithms (GTB) 46
Figure 3-5 User and producer accuracies of (a) RF, (b) GTB and (c) SVM, in conjunction with Dataset 3 47
Figure 3-6 Variable importance scores of (a) RF and (b) GTB with Dataset 3.....47
Figure 3-7 Areal extents per class of (a) RF, (b) GTB and (c) SVM with Dataset 3 48
Figure 3-8 NUS crop distribution maps of (a) RF, (b) GTB and (c) SVM 49

Figure 4-1 Taro crop canopy extraction: a) raw UAV image, b) ExG, c) ExR and d) ExGR canopy extraction 59
Figure 4.4-2 Graphical representation of the DNN structure, composed of input, hidden and output layers 60
Figure 4-3 Workflow diagram for this study 62
Figure 4-4 Relationship between the predicted and observed EWT_{canopy} based on the a) raw UAV image excluding thermal, b) raw UAV image including thermal, c) ExG excluding thermal, d) ExG including thermal, e) ExR excluding thermal, f) ExR including thermal, g) ExGR excluding thermal and h) ExGR including thermal 64
Figure 4-5 SHAP-generated variable importance of predictor variables in estimating EWT_{canopy} based on the a) raw UAV image excluding thermal, b) raw UAV image including thermal, c) ExG excluding thermal, d) ExG including thermal e) ExR excluding thermal, f) ExR including thermal, g) ExGR excluding thermal and h) ExGR including thermal 65
Figure 4-6 Comparative analysis of the mean performance metrics, including a) R^2, b) RMSE and c) rRMSE across the segmentation techniques between the inclusion and exclusion of the thermal band and overall mean d) R^2, e) RMSE and f) rRMSE obtained between including and excluding the thermal band 66
Figure 4-7 Spatial distribution of EWT_{canopy} within the taro field based on the a) raw UAV image, b) ExG, c) ExR and d) ExGR canopy extraction techniques 67
Figure 5-1 Taro crop canopy extraction by using the ExGR index-based segmentation technique 79
Figure 5-2 Graphical representation of the DNN structure, composed of input, hidden and output layers 80
Figure 5-3 Methodological overview of the study 81
Figure 5-4: Relationship between measured and predicted a) equivalent water thickness, b) fuel moisture content, c) stomatal conductance, d) leaf temperature and e) chlorophyll content 85
Figure 5-5 SHAP-generated variable importance values of predictor variables used in developing the optimal estimation models of a) equivalent water thickness, b) fuel moisture content, c) stomatal conductance, d) leaf temperature and e) chlorophyll content 86
Figure 5-6 Spatial distribution of modelled a) equivalent water thickness, b) fuel moisture content, c) stomatal conductance, d) leaf temperature and e) chlorophyll content 87
Figure 6-1 The comparative effectiveness of bands, indices and integrated datasets in estimating the CCC based on (a) average R-squared values, (b) typical RMSEs, and (c) mean rRMSEs 99
Figure 6-2 The comparative effectiveness of bands, indices and integrated datasets in estimating the CCC based on (a) average R-squared values, (b) typical RMSEs, and (c) mean rRMSEs 100
Figure 6-3 Regression models illustrating the connections between the observed and forecasted LCC throughout the taro and sweet potato phenological cycles: (a) 11th February 2021, (b) 12th April 2021, and (c) 14th May 2021 102
Figure 6-4 Variable importance scores of optimal LCC bands and VIs across the phenological cycle: (a) 11th February 2021, (b) 12th April 2021, and (c) 14th May 2021 ..103

Figure 6-5 Regression models illustrating the connections between the observed and forecasted CCC throughout the taro and sweet potato phenological cycles: (a) 11th February 2021, (b) 12th April 2021, and (c) 14th May 2021 104
Figure 6-6 Variable importance scores of optimal CCC bands and VIs across the phenological cycle: (a) 11th February 2021, (b) 12th April 2021, and (c) 14th May 2021 ..105
Figure 6-7 Comparison between the accuracies (a) average R2 and (b) average rRMSE derived in estimating LCC and CCCC 106
Figure 6-8 Spatial distribution of (a) taro and (b) sweet potato chlorophyll content in the experimental fields 107
Figure 7.7-1 Correlation between the SMC and Chlorophyll for (a) Establishment stage Week 5, (b) Establishment stage Week 8, (c) Vegetative Week Pink and (d) Late Vegetative stage 120
Figure 7.7-2 Soil moisture content for the (a) Establishment Stage (i), (b) Establishment Stage (ii), (c) Vegetative stage (i), (d) vegetation Stage (ii) and their associated variable importance 123
Figure 7.7-3 Chlorophyll content accuracies for (a) Establishment stage (i), (b) Establishment stage (ii), (c) Vegetative stage (i), (d) Vegetation stage (ii) and their associated variable importance 125
Figure 7.7-4 Spatial variation of soil moisture content at different growth stages (a) establishment (i), (b) establishment (ii), (c) vegetative (i), (d) vegetative (ii) 126
Figure 7.7-5 Spatial variation of chlorophyll content at different growth stages (a) establishment I, (b) establishment II, (c) vegetative I, (d) vegetative II) 127
Figure 8.8-1 a) Matrice 300 UAV integrated with the Altum sensor serving as the imaging platform used in this study, b) the Altum camera, c) the calibrated reflectance panel and d) the flight path plan for the study image 133
Figure 8.8-2 Correlation between the measured AGB and BGB at the a) establishment stage (DAP =58), b) vegetative stage (DAP =154) and c) maturity stage (DAP = 249) 135
Figure 8.8-3 Relationship between the predicted and observed AGB in the (a) establishment stage, (b) the vegetative stage, and (c) the maturity stage of taro derived from a combination of bands and Vis, using the RF algorithm and the model variable importance scores 137
Figure 8.8-4 Comparison of accuracies derived in estimating the AGB and BGB of taro using UAV-acquired remotely-sensed data 139
Figure 8.8-5 The spatial distribution of the modelled AGB (a, c and e) and BGB (b, d and f) of smallholder taro crops 140
Figure 9.9-1 a) DJI Matrice 300 platform, and b) Micasense Altum multispectral and thermal sensor 147
Figure 9.9-2 Prediction of taro yield using UAV image data from (a) unsegmented image, (b) segmented image, (c) segmented image and LAI, and (i, ii and iii) are the associated variable importance scores 151
Figure 9.9-3 Spatial distribution of modelled taro yield 152

LIST OF TABLES

Table 3-1 UAV-derived vegetation indices....38
Table 3-2 Training and testing data used for pixel-based image classification 39
Table 3-3 Overall accuracies, kappa statistics and F1 scores for RF, GTB and SVM....43
Table 3-4 Band 1 JM distances 43
Table 4-1 UAV derived vegetation indices used in this study....57
Table 4-2 Descriptive statistics of the in-situ measured EWT_{canopy} of taro crops 62
Table 5-1 UAV-derived spectral indices used in this study....77
Table .5-2 Descriptive analysis of the taro crop water status physiological indicators 82
Table 5-3 Estimation accuracies of taro crop water status physiological indicators using selected spectral variables....83
Table 6-1 Stages followed in estimating chlorophyll content....97
Table 6-2 Descriptive statistics of the leaf chlorophyll content of sweet potatoes and taro 98
Table 6-3 Descriptive statistics of the canopy chlorophyll content of sweet potatoes and taro 98
Table 7.7-1 List of vegetation indices used in modelling the soil moisture and chlorophyll content 117
Table 7.7-2 Descriptive statistics of soil moisture content and chlorophyll across different phenological stages 119
Table 8.8-1 DJI M300 UAV specifications....132
Table 8.8-2 List of vegetation indices used in the modelling of AGB and BGB and related source references....134
Table 8.8-3 Descriptive statistics of the AGB and BGB across selected growth stages of taro 135
Table 8.8-4 Relationship between the predicted and observed BGB in the (a) establishment stage, (b) vegetative stage and (c) maturity stage of taro derived from a combination of bands and Vis, using the RF algorithm and the model variable importance scores 138
Table 9.9-1 List of vegetation indices used in the modelling of yield and related source references 148
Table 9.9-2 Descriptive statistics of taro yield....150

ABBREVIATIONS

AAPA	African Association for Precision Agriculture
ANN	Artificial Neural Network
AWS	Automatic Weather Station
CCI	Chlorophyll Carotenoid Index
CWSI	Crop Water Status Index
DEM	Digital Elevation Model
EGI	Excess Green Index
EVI	Enhanced Vegetation Index
EWT	Equivalent Water Thickness
EXG	Excess Green Index
FMC	Fuel Moisture Content
GLM	Generalised Linear Models
GNDVI	Green Normalised Difference Vegetation Index
LAI	Leaf Area Index
LR	Linear Regression
MTVI1	Modified Triangular Vegetation Index 1
NDRE	Normalised Difference Red-Edge Index
NDVI	Normalised Difference Vegetation Index
NGRDI	Normalised Green, Red Difference Index
NUS	Neglected and Under-utilised Crop Species
OSAVI	Optimised Soil-Adjusted Vegetation Index
PCA	Principal Components Analysis
PLS	Partial Least Squares Regression
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RE	Red Edge
RF	Random Forest
RGB	Red Green Blue
SLA.	Specific Leaf Area
SRI	Simple Ratio
SSA	Sub-Saharan Africa
SVM	Support Vector Machine
SVM	Support Vector Machines
UAV-HTPPs	Unmanned Aerial Vehicle - High Throughput Phenotyping system
US	United States
WRC	Water Research Commission

REPOSITORY OF DATA

For details related to the project's data, please contact:

Mbulisi Sibanda
Department of Geography, Environmental Studies and Tourism
Faculty of Arts and Humanities
University of the Western Cape
Landline: +27 (0)21 959 2668
Email: msibanda@uwc.ac.za

1 INTRODUCTION AND BACKGROUND

1.1 Introduction

The observed climate change has been extensively demonstrated as directly impacting the water resources, agriculture and food systems. The rise in global temperatures has led to unpredictable weather patterns, marked by substantial alterations in precipitation and a heightened frequency of extreme events like droughts. These dramatic shifts have been demonstrated to have a notable impact on agricultural mainstream crops, including maize, wheat and sugar beet. The diminishing agricultural productivity has been demonstrated to be contributing to food and nutrition insecurity, especially in the drylands of Africa and the high mountain regions of Asia and South America. The compounding factor of rapid population growth and increased food demand further exacerbates these challenges.

A considerable segment of the growing population resides in marginal areas and depends on smallholder agriculture for their livelihood, by either actively engaging in farming or benefiting indirectly (Masipa 2017). Smallholder farmers play a pivotal role in addressing the complex challenge of ensuring food security for the growing population, all the while minimising the adverse environmental impacts and adapting to the effects of climate change (Dhillon and Moncur 2023, Touch *et al.* 2024). More than 80% of the world's farms (475 million) rely on less than two hectares of land (Fan and Rue 2020). Regardless of the fact that they account for only 12% of the world's agriculture, these farms offer an estimated 80% of the food grown in sub-Saharan Africa (SSA) (Fan and Rue 2020). Despite the significance of smallholder agriculture in the global food supply and poverty alleviation, agricultural productivity in the SSA region has decreased (Mugiyo *et al.* 2021). It has been well established that underlying water shortages, compounded by climate variability and changes in land use, have led to a reduction in the amount of land that is accessible for agricultural development, particularly in resource-poor farming systems, in order to produce the main crops (Mugiyo *et al.* 2021).

The agricultural sector needs a transformative shift from mainstream cropping systems to unconventional options, such as the Neglected and Under-utilised crop Species (NUS), as future crops. Research indicates that the NUS exhibit an adaptability to various agroecological conditions, they are rich in nutrients and they offer promising prospects for regions with low crop yields (Mabhaudhi *et al.* 2017, Mugiyo *et al.* 2021). The NUS complement, and are alternatives to, many commercially-important food crops, and they can aid in addressing the food and nutritional deficiencies in marginalised communities (Mugiyo *et al.* 2021). NUS crops, such as sorghum, cowpea, amaranth, sweet potatoes and taro, thrive in marginal areas that are prone to heavy droughts and flash floods (Mugiyo *et al.* 2021).

Therefore, research and consumer interest in these crops have gained traction due to their potential role in strengthening the local food systems and mitigating the risks and effects of climate change shocks in vulnerable agricultural production systems (Padulosi *et al.* 2013). The absence of precise, consistent, spatially-explicit and high-throughput phenotyping technologies poses a challenge in complementing the rapid advancements in phenotyping and genotyping technologies, which are crucial for high-speed crop breeding. Considering this

challenge, field phenotyping becomes essential. This approach demonstrates the outstanding performance of crop traits, in alignment with productivity, by spatially showcasing phenotypes at specific levels of their statistical significance. Field phenotyping is indispensable for enhancing the knowledge base of agricultural production in smallholder farming systems. This will set up a backbone for effectively monitoring crop health and productivity.

Unmanned Aerial Vehicles (UAV)-based phenotyping holds the potential to provide ultra-high spatial resolution data, and it facilitates precise, consistent and high-throughput phenotyping in smallholder cropping plots. The resolution of UAV-based phenotyping operates at a plot level, which enables immediate data capture for single or multiple plots and makes it well-suited for plant breeding in smallholder croplands. Utilising UAVs with high-resolution sensors is an emerging and relatively cost-effective technique for high-throughput crop phenotyping at a farm scale. However, the capability of UAV-mounted sensors to differentiate crop types, based on their spectral responses as a mechanism for plausible high-throughput field phenotyping, is yet to be determined (Chivasa *et al.* 2020). Remote sensing and machine learning techniques have greatly aided high-throughput phenotyping technologies in the recent past. For example, Zheng *et al.* (2021) used remote sensing and machine learning in crop phenotyping and management and emphasised their application in strawberry farming. Ampatzidis and Partel (2019) examined UAV-based high-throughput phenotyping in citrus farms, by using multispectral imaging and artificial intelligence. These studies illustrate that UAV-derived data are suitable for the local-scale, high-throughput phenotyping of crops.

However, despite the rapid growth of literature on the UAV-based phenotyping of crops and forage, there is a need to extend the research efforts to the NUS in smallholder croplands, since many works are based on experimental plots in controlled sites of maize and wheat. Therefore, this project aims to assess the utility of UAV-derived remotely-sensed data in the mapping and monitoring of NUS (taro and sweet potatoes) crop productivity and their health attributes, such as their yield, biomass, leaf area index, chlorophyll content and crop water stress index, in smallholder croplands.

1.2 Aims and Objectives

This project aimed to assess the application of unmanned aerial vehicle high-throughput phenotyping of NUS for the improved water use and productivity on smallholder farms. To achieve this overarching aim, the project has set out specific objectives, namely:

1) to review the literature on the utility of earth observation data in characterising the productivity of NUS crop species on smallholder farms;
2) to map the spatial distribution and health of NUS, by using UAV data on smallholder farms;
3) to assess the relationship between the NUS crop productivity variations and edaphic factors (i.e. soil moisture, temperature and salinity) on smallholder farms;
4) to evaluate the use of unmanned aerial vehicle platforms for quantifying the spatial and temporal variability of canopy crop water status and yield by NUS crops on smallholder farms; and

5) to assess the potential of unmanned aerial vehicle high-throughput phenotyping data in estimating the yield of NUS crops in smallholder croplands

This project leveraged the data and infrastructure that were established through the WRC Flagship project K5/2971//4 to address some of the objectives. The WRC Flagship project K5/2971//4 was titled "*The use of drones in monitoring crop health, water stress and crop water requirements for improving crop water productivity in the context of precision agriculture in smallholder croplands.*"

1.3 Scope and the Overview of the Report

1.3.1 Description of the study area

The geographic scope of this project encompasses Swayimane, a communal area situated within the uMshwathi Municipality, north-east of the city of Pietermaritzburg in South Africa (Ndlovu *et al.* 2021, Brewer *et al.* 2022) (Figure 1.1). The communal area spans approximately 36 km² and is dominated by smallholder farming that is practised by the local community. The major crops include white and yellow maize, sugarcane, amadumbe (taro) and sweet potatoes. The farming systems are largely traditional, and they rely on rainfed cultivation, livestock manure for fertilisation, as well as manual labour for planting, maintenance and harvesting, with some farmers using back-pack herbicide sprayers for weed control. Agricultural production underpins the food security and livelihoods of the locals, with the produce being sold mainly in the local markets.

Swayimane is located within the moist Midlands mist belt bioresource area and it experiences average temperatures ranging from 11.8°C to 24°C, with hot, wet-to-cool summers and dry winters (Ndlovu *et al.* 2021, Brewer *et al.* 2022). The annual rainfall ranges between 600 and 1100 mm (Ndlovu *et al.* 2021). The area is characterised by arable clay loam soils and is ranked among the top-2% of high-potential agricultural land in South Africa (Ndlovu *et al.* 2021). Considering that all the studies, which are presented as separate chapters, were conducted in the same experimental plots, a description of the study area was removed from all the other chapters.

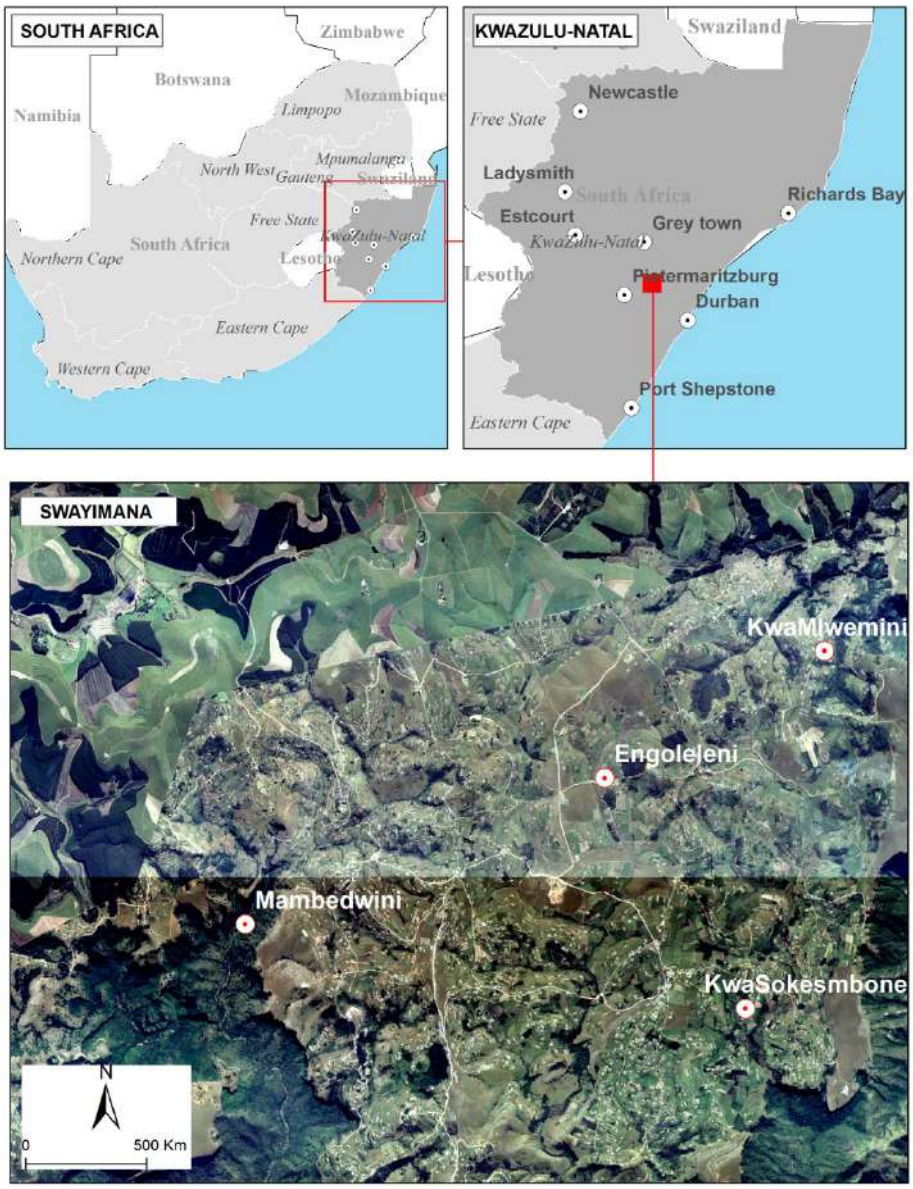


Figure 1-1 Aerial diagram of Swayimane, KwaZulu-Natal (Source :Sibanda *et al.* (2023))

1.3.2 Report outline

This report consolidates the contributions of multiple project authors, with each chapter being designed as a stand-alone unit that addresses specific contractual objectives under the guidance of the WRC project managers and the technical reference group. In line with this 'paper' format, the study area is detailed in Section 1.3.1, and a single, unified methodology chapter is omitted, in favour of distinct methodological descriptions within each chapter, which may lead to some necessary overlap, especially in methods sections, to preserve the integrity of the aim of each chapter. Several chapters are adaptations of peer-reviewed articles that were previously published by the project team. The report is structured in the following sequence, in order to logically fulfil the objectives of the project:

Chapter One: A comprehensive introduction to the study is provided, including its background, conceptual framework and rationale. It clearly outlines the project's overarching goals and specific contractual objectives, thereby establishing the necessary foundation for the detailed research presented in the subsequent chapters.

Chapter Two: A systematic review of the literature is given on the use of earth observation data in characterising the productivity of NUS crop species on smallholder farms. This chapter specifically addresses Objective 1.

Chapter Three: The potential of UAV-acquired multispectral imagery, combined with machine learning techniques, is assessed in mapping the spatial distribution of taro and sweet potatoes in smallholder farms, and it specifically addresses Objective 2.

Chapter Four: Objectives 3 and 4 are addressed by developing a method to enhance the estimation of the Equivalent Water Thickness in taro crops, by utilising UAV-acquired multispectral-thermal imagery and index-based image segmentation.

Chapter Five: Objective 4 is specifically addressed by assessing the neglected and under-utilised water status of the taro crop, using physiological indicators and UAV multi-modal thermal-multispectral data

Chapter Six: The canopy chlorophyll content of taro and sweet potatoes is estimated by using UAV-remotely-sensed data in smallholder croplands, which is specifically in partial fulfilment of Objective 3.

Chapter Seven: The potential of drone remotely sensed data is assessed for detecting the soil water content and taro leaf chlorophyll content across the different phenological stages, and it is also specifically in partial fulfilment of Objective 3.

Chapter Eight: The potential of UAV-derived multispectral imagery for estimating the taro below-ground biomass in smallholder croplands is assessed, and it specifically addresses Objective 5.

Chapter Nine: In partial and specific fulfilment of Objective 5, this chapter also explores the utility of image segmentation techniques and leaf area index for improving the taro yield prediction, based on the UAV-acquired multispectral data in smallholder croplands

Chapter Ten: The general synthesis and discussion of the overall findings of the project are presented. This chapter outlines the general and the overall conclusion, and it addresses the overarching aim of the project. Finally, it outlines the contextual and analytical limitations, along with some recommendations for consideration in future projects.

Supplementary materials that document the skills and knowledge transfer activities, which were conducted as part of the project’s capacity development, are documented at the end of the report.

2 A SYSTEMATIC REVIEW OF REMOTE SENSING APPLICATIONS FOR NEGLECTED AND UNDER-UTILISED CROP SPECIES

2.1 Introduction

Climate change, which is evidenced by the increased occurrence and frequency of extreme climatic conditions, has become unequivocal, as it affects water resources, agriculture and food systems (Adhikari *et al.* 2015). Meanwhile, the world is confronted by the problem of feeding a growing population, while trying to minimise the adverse environmental impacts and to adjust to the changing climate (Mugiyo *et al.* 2021). Food, nutrition and water insecurity are affecting more than two-thirds of the population (Mugiyo *et al.* 2021). For the first time in history, more than one billion people in the world are malnourished, which emphasises the importance of addressing the food and nutrition insecurity in agriculture that is being accelerated by climate change (Fan and Rue 2020). More specifically, Africa is epitomised as the continent that is most vulnerable to climate change (Masipa 2017). The annual population growth rate in Africa is anticipated to be 2.4%, and the expanding population is expected to increase the demand for food (Nyasimi *et al.* 2014). By 2050, the continent's population will have more than doubled from 0.9 billion people, and by 2100, it will have quadrupled to 3.9 billion (Nyasimi *et al.* 2014). The bulk of these people live in marginal areas and rely, either actively or passively, on smallholder agriculture for their livelihoods (Masipa 2017). In the global food security conundrum, smallholder farmers in developing nations play a vital role. More than 80% of the world's farms (475 million) rely on less than two hectares of land (Fan and Rue 2020). Regardless of accounting for only 12% of the world's agriculture, these farms offer an estimated 80% of the food grown in sub-Saharan Africa (SSA) (Fan and Rue 2020). Despite the significance of smallholder agriculture in the global food supply and in poverty alleviation, agricultural productivity in the SSA region has decreased (Mugiyo *et al.* 2021). It has been well established that the underlying water shortages, which are compounded by climate variability and changes in land use, have led to a reduction in land that is accessible for agricultural development and the production of the main crops, particularly in resource-poor farming systems (Mugiyo *et al.* 2021).

There has been a need for a paradigm change in agriculture to investigate traditional avenues, such as the planting of Neglected and Under-utilised crop Species (NUS) as future crops. Literature shows that NUS are adaptable to a variety of agro-ecologies, are highly nutritious, and they offer a better chance in low-yielding regions (Mabhaudhi *et al.* 2017, Mugiyo *et al.* 2021). NUS are a feasible alternative for resolving food and nutritional deficiencies and stresses in marginalised communities (Chivenge *et al.* 2015). NUS crops, such as sorghum, cowpea, amaranth, sweet potatoes and taro are known to thrive in marginal areas that are prone to heavy droughts and flash floods (Mugiyo *et al.* 2021). Despite their agronomic, genomic, socioeconomic and cultural importance, these crops have long been overlooked by conventional agriculture, but they are now gaining traction due to their potential role in mitigating the risk and effects of the climate change shocks in vulnerable agricultural production systems (Mabhaudhi *et al.* 2017).

NUS crops benefit the subsistence market by offering smallholder farmers the possibility of profiting from greater crop yields, despite the climate variability and reduced water resources. Furthermore, cultivating NUS will safeguard the well-being of minority populations by alleviating poverty and malnutrition and by tapping into indigenous knowledge systems, while facilitating their traditional food and cultural heritage. The cultivation of NUS assists in achieving SDGs 1, 2 and 3, thereby enhancing the agro-economic sector. For this reason, smallholder farmers require spatially-explicit information on the biophysical and morphological characteristics of NUS to optimise their productivity. The lack of precise, consistent, spatially-explicit and high-throughput phenotyping technologies is a hurdle hurdle that needs to be overcome in supplementing the rapid advancements in phenotyping NUS to optimise the production in smallholder croplands.

Traditional or in-field observations were first utilised to measure the spatial extent, suitability, crop vegetative growth and morphological attributes. However, this method of analysis is deemed to be too time-intensive and expensive, which make it unsuitable for continuous and precision crop monitoring in areas with numerous crop varieties. Furthermore, in-situ measurements are not always as accurate; their spatial representativeness is inadequate, and the data may be difficult to obtain for users, such as farmers (Sibanda *et al.* 2021). They are often associated with destructive harvesting, which makes it impossible to examine the same plant over time. Over the years, satellite-based earth observation technologies have proven to be effective in monitoring plant growth and health changes (Qader *et al.* 2021).

In this regard, the application of data from earth observation and geo-spatial techniques has become more effective and reliable for monitoring and estimating crop health at a field and farm level. Utilising platforms such as satellites, unmanned aerial vehicles, vehicle-mounted sensors, as well as hand-held cameras or sensors, high-throughput image-based phenotyping can be performed from a landscape level, to a cellular level. On a wider scale, satellite remote sensing technologies provide non-invasive, accurate, fast and cost-effective data for estimating traits, such as the chlorophyll concentration in plants, as a proxy for a crop's productivity and for quantifying its diversity and spatial distribution. However, the spectral and temporal resolution of freely-available satellite sensors limits a plot-level variety analysis and data acquisition in plant phenotyping (Chivasa *et al.* 2020). Very high spatial resolution remote sensing solutions and manned aerial platforms are expensive to use in smallholder croplands, and they are constrained by operational complexities (Chivasa *et al.* 2020). By detecting the interactions between plant components and the light spectrum, sensors can provide crucial information concerning plant features. Sensors and cameras can detect the physiological and phenotypic variations across different crop species by using certain spectral portions of the electromagnetic spectrum, such as the visible (400-780 nm) and near-infrared (781-1000 nm) sections (Homolová *et al.* 2013).

Meanwhile, the advent of Unmanned Aerial Vehicle (UAV)-based phenotyping has the prospect of offering ultra-high spatial resolution data that are suitable for precise, consistent and high-throughput phenotyping in smallholder cropping plots. The resolution of UAV-based phenotyping is at the plot level; hence, it allows for the immediate records of single or several plots and makes it suitable for plant breeding in smallholder croplands. The use of Unmanned

Aerial Vehicles (UAVs) equipped with high-resolution sensors is an emerging, relatively cost-effective technique for high-throughput crop phenotyping, at a farm scale. However, the capacity of UAV-mounted sensors to differentiate crop types, based on their spectral responses as a mechanism for plausible high-throughput field phenotyping, is yet to be determined (Maimaitijiang *et al.* 2020). Remote sensing and machine learning techniques have both aided high-throughput phenotyping technologies in the recent past. For example, (Maimaitijiang *et al.* 2020) examined the potential of combining canopy spectral and structural information for the crop monitoring of a heterogeneous soybean field, using satellite/UAV data fusion, in conjunction with machine learning, for the prediction of a crop's AGB, LAI.

Despite the usefulness of UAVs, their application in agriculture, rural development and, more importantly, crop water resources management, remains limited (Sibanda *et al.* 2021). Although some studies have attempted to assess the literature on the application of drone-acquired data, most did not systematically and quantitatively assess the literature on mapping the spatial distribution and health of NUS crops, with a special interest in the Global South trends. Hence, the objective of this work was to evaluate and present an in-depth systematic evaluation of the publications on the progress, challenges, opportunities and gaps in the application of the UAVs' remotely-sensed data in mapping the health status of NUS on smallholder farms, with a special focus on the Global South.

2.2 Methods

2.2.1 Phase One: Literature search

In the initial phase of the literature search, keywords, terms and phrases for searching the literature were generated from other literature reviews on NUS. The following keywords and variants were used in this study to search for the relevant: "neglected and under-utilised crop species," "orphan crops," "traditional crops," "unmanned aerial vehicle(s)," "drone(s)," "remote sensing," "GIS," "crop health," "stomatal conductance", "leaf area index" and chlorophyll". Furthermore, the following keywords and variants were used to search for literature pertaining to taro and sweet potatoes: "Taro," "sweet-potatoes," "unmanned aerial vehicle(s)," "drone(s)," "remote sensing," "GIS," "crop health," "stomatal conductance", "leaf area index" and chlorophyll".

SCOPUS, Web of Science and Google Scholar were utilised to collect literature by using the established key search terms. The PRISMA statement served as the framework for the literature search procedure. This search was not restricted, in terms of time. Google Scholar, Scopus and the Web of Science literature searches yielded 109, 193 and 90 articles, respectively. In preparation for screening, all the obtained material was organised in EndNote. Furthermore, a bibliographic investigation was conducted in this phase. This consists of first filtering the diverse articles and eliminating duplicates (n = 77), based on comparable key search phrases. In this case, literature that was not written in English was excluded from the analysis in the second phase. The next step was to assess whether the different publications were concentrated on mapping the spatial distribution and assessing the health of neglected and under-utilised crops. Full-length articles of the selected abstracts were then sought and downloaded. After the

screening procedure, 134 articles were retained (Figure 2.1). Furthermore, a Microsoft Excel spreadsheet was constructed to record the details and topics of each research study. The constructed database was used to extract the quantitative data from each article, as indicated in the preceding phase.

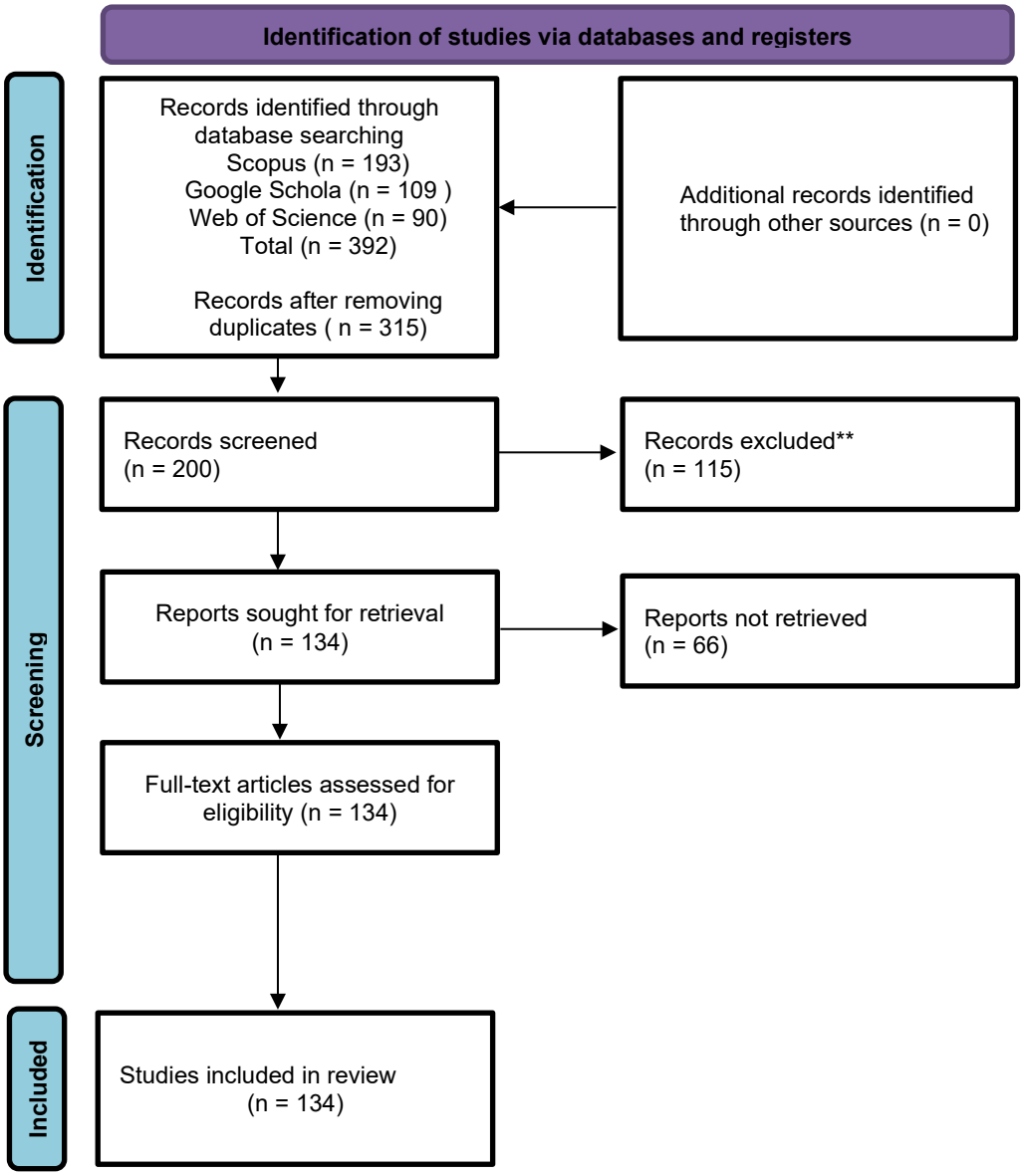


Figure 2-1 PRISMA flow diagram for selection of studies considered in the review

2.2.2 Phase Two: Data extraction

In the preceding phase, the Excel database was utilised to generate and extract information on the progress, gaps, challenges and opportunities in the use of UAV technologies in mapping the spatial distribution and health of NUS crops. The second stage of the research gathered data from the selected articles, in order to fulfil these study objectives. The main information related to the study region, the type of UAV or satellite sensor used to map the distribution, the type

of crop attribute investigated, sensor and platform type, vegetation indices, predictive or classification algorithms, as well as the optimum spectral variables obtained, were all retrieved from the literature and documented in the spreadsheet. In planning for data *analysis*, all the categorical variables were converted into numerical values. During this step, relevant bibliometric information was also collected. These included authors' names, the region, year of publication, article title, journal name and abstract were among the bibliometric information gathered.

2.2.3 Phase Three: Data analysis

The retrieved literature and extracted data were subjected to quantitative and qualitative analyses during this phase. Basic statistical frequencies were calculated for the quantitative analysis (Sibanda *et al.* 2021). In addition, an exploratory trend analysis was carried out to assess the progress made in mapping the spatial distribution and health of NUS crops by utilising satellite and drone-borne sensors. A bibliometric analysis was also conducted to identify trends in the co-occurring key terms of literature that assessed the spatial distribution and health of NUS crops. The trends were identified by quantitatively examining the occurrence and co-occurrence of key terms used in the VOSviewer software. Furthermore, the titles and abstracts of publications in the final database (134 articles), including the database of publications that particularly utilised UAV-derived datasets (with 29 articles), were included in the VOSviewer software to evaluate how concepts and topics evolved. To address the research objectives, the review was divided into two main sections. The first section investigated recent advances in mapping the spatial distribution and health of NUS crops by using remotely-sensed data. This section presents and discusses quantitative literature trends in analysing the spatial distribution and health of NUS. Throughout this phase, the crop health attributes, earth observation sensors (cameras), sensor platforms, algorithms, and optimum spectral variables that had been deployed and are used by the community, in practice, to date, were highlighted. The last phase uncovered and discussed the challenges, gaps and opportunities for knowledge generation in mapping the spatial distribution and health of NUS, by using drone-derived, remotely-sensed data.

2.3 Results

2.3.1 Searched literature characteristics

Figure 2.2 illustrates the shifts of the topical concepts in remote sensing, the spatial distribution and the health attributes of NUS crops based on information derived from the titles only. These specific terms included "drought stress", "climate change", "estimation", "photosynthesis", "stomatal conductance", "productivity" "unmanned aerial vehicle" "vegetation index", "sweet potatoes" and "Bambara" (Figure 2.2). This implies the utility of unmanned aerial vehicles in estimating the productivity elements of NUS, such as sweet potatoes and bambara groundnuts. Based on the titles of the retrieved literature published until 2014, the most co-occurring concepts were "assessment of "drought stress" and "water stress", "stomatal conductance", of "sweet potatoes" in the context of "precision farming". From 2014 to 2019, there was a shift

in the co-occurring terms towards crop "productivity", "estimation", "detection" and "response" of NUS such as "Bambara groundnuts".

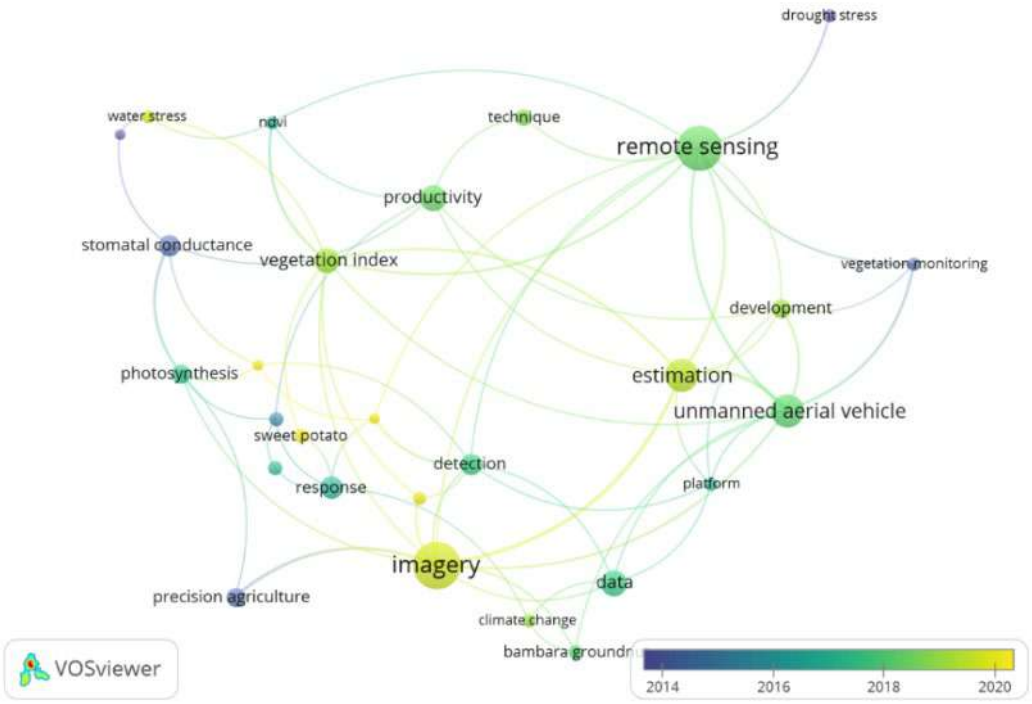


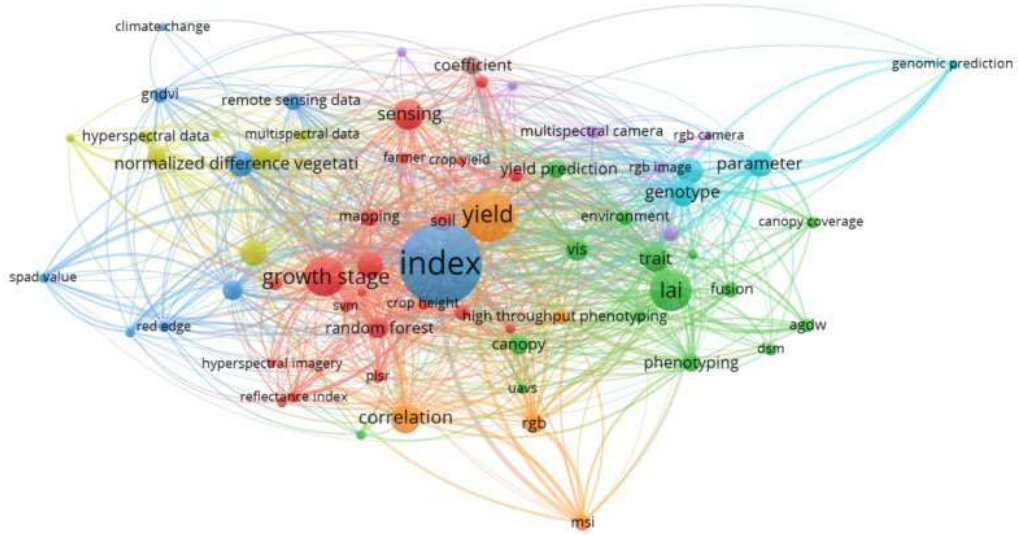
Figure 2-2 Changes in topical concepts associated with the remote sensing productivity of NUS crops based on titles

Meanwhile, Figure 2.3 shows the co-occurrence of topical concepts derived from titles and abstracts. Figure 2.3a illustrates seven topical clusters, dark blue, light blue, red, green, purple, yellow and orange, in mapping the crop spatial extent and health. The key terms from the "red" cluster were "agriculture", "spad value", "low cost", "remote sensing data", "hyperspectral data", "multispectral data", "VIS", "processing", "size", "crop field", which directly imply the utility of "low-cost remote sensing" systems for mapping and monitoring NUS crop productivity (Spad-value) remotely-sensed data in smallholder crop fields (Figure 2.3a). The second-largest cluster linked to UAVs was in yellow and had "growth", "variety", "trait", "detection", "plant-height", "UAVs", "drone", "dsm", "rgb" and "msi". This cluster articulates the use of drone remotely-sensed ("UAVs") data in detecting and mapping the relevant NUS phenotypical attributes ("growth", "variety", "trait", "detection" and "plant-height",). The third cluster in dark blue had 'growth stage', "lai", "crop height", "agdw", "canopy nitrogen", "weight", "fusion", "plsr" and "rmse" as the key terms, in order of importance (Figure 2.4). This cluster relates to the estimation of NUS crop productivity attributes ("lai", "crop height", "agdw", "canopy nitrogen", "weight", "lai", "crop height", "agdw", "canopy nitrogen" and "weight") by using remotely sensed data and regression techniques. The fourth cluster in green had "population", "climate change", "water stress "food security", "African leafy vegetable", "suitable area", "moisture", "SSA" (sub-Saharan Africa) and "validation", amongst others. This links with the issues of food and nutrition security issues in SSA, which are highly impacted and driven by climate variability, such that NUS crops ("African leafy vegetable"),

and only crops that are suitable because of their drought tolerance capabilities. The fifth cluster which is light blue, features co-occurring terms such as “species”, “reflectance”, “density”, “waveband”, “nir” “classification”, “soybean” and “palmer amaranth”. This cluster relates to the impact of species variability and plant, or foliage, density, which makes the spectral signatures of NUS crops different at different wavebands of the electromagnetic spectrum, and which facilitates optimal classification accuracies. The sixth cluster is characterised by the co-occurrence of terms such as “plant”, “amaranth”, “phenological” and “growth stage”, which relate to the assessment of the phenological characteristics of NUS crops.

Meanwhile, Figure 2.3b illustrates seven topical clusters, namely, dark blue, light blue, red, green, purple, yellow and orange, that are rederived by using abstracts and titles by using drone remotely-sensed data in mapping the spatial extent and health of NUS crops. The key terms from red clusters were “mapping”, “crop height”, “growth stage”, “yield” and “crop yield”, “reflectance index”, “hyperspectral imagery”, “Random Forest” and “svm”. This cluster relates to the use of high-resolution, remotely-sensed data in the mapping and monitoring of NUS crop productivity elements, such as the growth stage, using the application of machine learning techniques. The key terms in the green cluster are “phenotyping”, “LAI”, “canopy coverage”, “trait”, “agdw”, “prediction”, “uavs”, “dsm”, “fusion”, “VIS” and “environment”. The cluster relates to the use of UAV remotely-sensed data that could be fused with other data to assess the structural attributes of NUS crops (phenotyping). The third cluster in dark blue has key terms, such as “remotely-sensed data”, “gndvi”, “normalised difference vegetation”, “red-edge”, “spad value” and “climate change”. The clusters relate to the utility of the widely-used spectral variables (“gndvi”, “normalised difference vegetation”, “red-edge”) in monitoring the health of NUS crops. The fourth cluster is orange and has co-occurrence terms, such as the “high throughput phenotyping”, “rgb”, “msi”, “yield” and “correlation”. This cluster can be attributed to the application of high-throughput phenotyping of NUS through RGB and multispectral spectrums in estimating crop yield. The fifth cluster in light blue had “genotype”, “parameter” and “genomic parameters”. This cluster suggests the examination of the genetic composition of NUS crops through the assessment of optimal parameters to assess and infer their crop health. The sixth cluster has “multispectral camera”, “rgb camera” and “rgb image”. This cluster relates to the use of general colour imagery (in red, green and blue spectrums) that are acquired by using drones for the mapping and monitoring of NUS. The last cluster has “hyperspectral data” and “multispectral data”. This cluster relates to the use of multispectral and hyperspectral data in the mapping and monitoring of the spatial distribution and health of NUS crops.

(a)



(b)

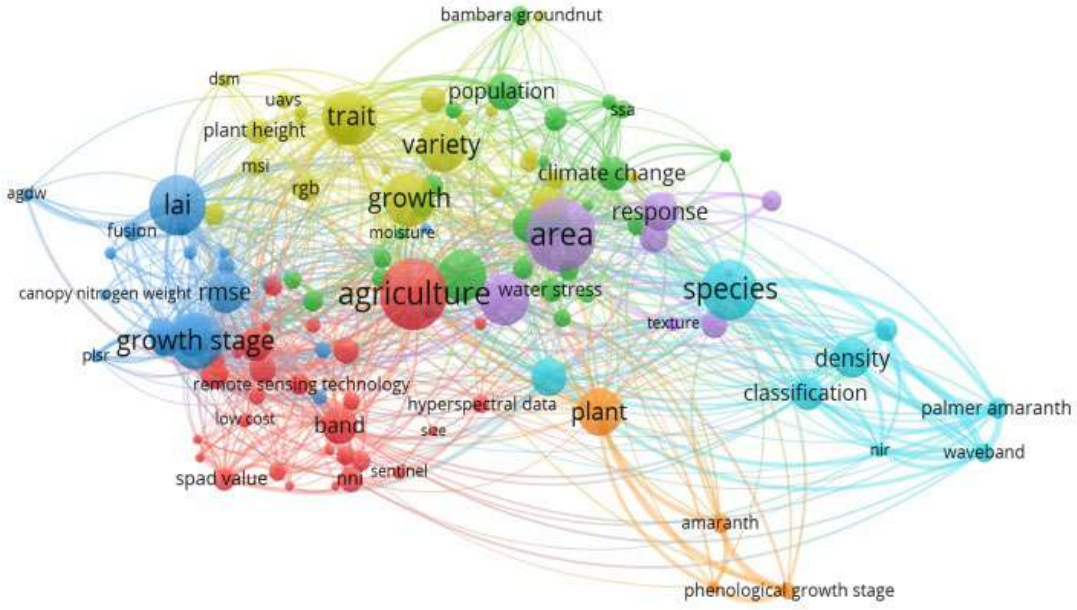


Figure 2-3 Topical concepts identified by using a bibliometric analysis of titles and abstracts of articles that utilised (a) all remote sensing sensors, and (b) only drone-acquired data in mapping the NUS

2.3.2 Progress in mapping the spatial distribution and health status of neglected and under-utilised crop species

Significant progress has been attained in detecting, mapping and monitoring the spatial distribution and health status of NUS crops (Figure 2.4). The period between 2003 and 2013 is marked by a low frequency of published literature, based on all sources of earth observation sensors (Figure 2.4(a)). Similarly, there was a low frequency of published studies that were conducted by using UAVs that acquired remotely-sensed data to characterise the health and spatial distribution of NUS (Figure 2.4(b)). Between 2014 and 2022, there was a rapid increase in the number of articles that mapped the health and spatial distribution of NUS by using all earth observation sensors. Again, from 2014 to 2022, there was a rapid growth in the body of literature that utilised UAV remote-sensed data to characterise the NUS attributes. Moreover, 20 studies utilised satellite-borne, remotely-sensed data, while 29 studies used drones to acquire remotely-sensed data in mapping the attributes of NUS crops and their spatial distribution. From the retrieved literature, very few studies assessed any remotely-sensed data for characterising the spatial distribution of NUS crops or discriminating the varieties.

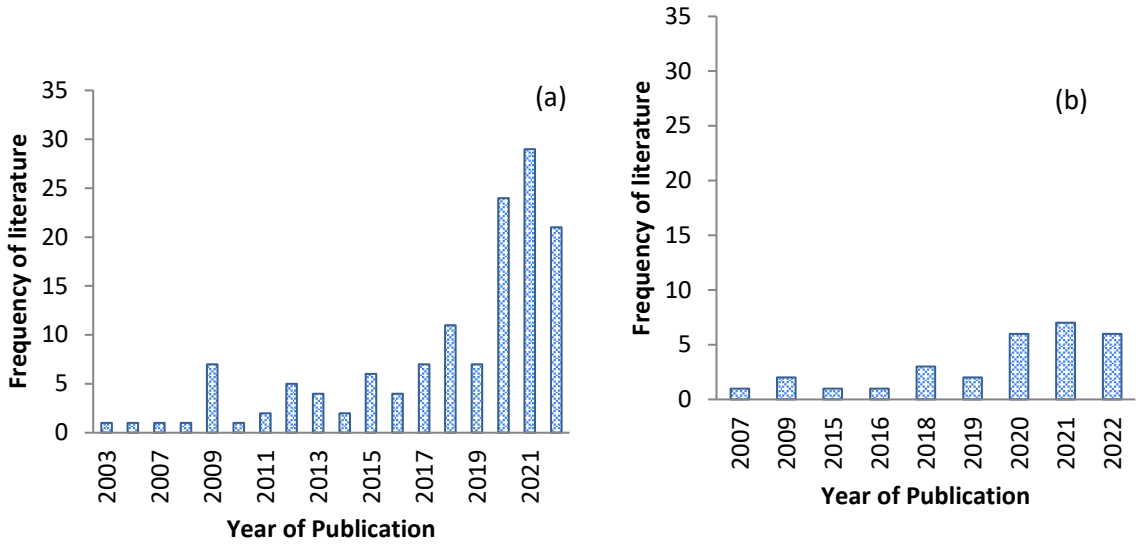


Figure 2-4 Frequency of published articles on remote sensing applications of NUS, based on (a) all sensors, and (b) UAVs

Most studies were conducted in Asia, America and the African continent. Most of the studies were conducted in America (n = 21), followed by China and Africa (Figure 2.5). Interestingly, most of the studies that were conducted in Africa were conducted in southern Africa, where South Africa had the highest number of studies where earth observation data were utilised to map the health attributes and spatial distribution of NUS.

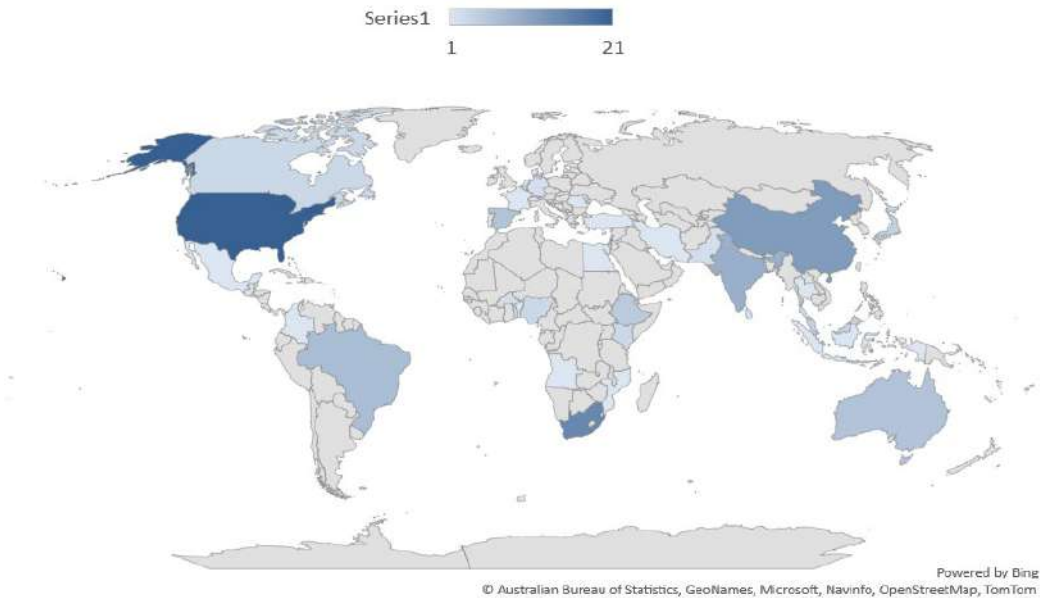


Figure 2-5 Spatial distribution of studies on remote sensing, the attributes and spatial distribution of NUS

Most of the retrieved literature mapped the productivity and water stress-related elements. Specifically, the most widely-researched NUS health attributes, based on all sensors, included the crop yield, leaf water relational status, LAI, canopy height, biomass, chlorophyll content, photosynthesis, canopy cover, leaf water attributes, canopy temperature, root growth attributes, stomatal conductance and leaf nitrogen (Figure 2.6(a)). The most researched NUS attributes in the context of UAV-based remote sensing are the crop yield (n = 18), chlorophyll content (n = 12), biomass (n = 12), crop health (n=9), LAI (n = 8), cover (n = 8), canopy height (n = 7), leaf nitrogen (n =7), leaf water content elements (n =7) and leaf temperature (n = 7). As mentioned above, very few studies classified and characterised the spatial distribution of NUS (Figure 2.6(b)).

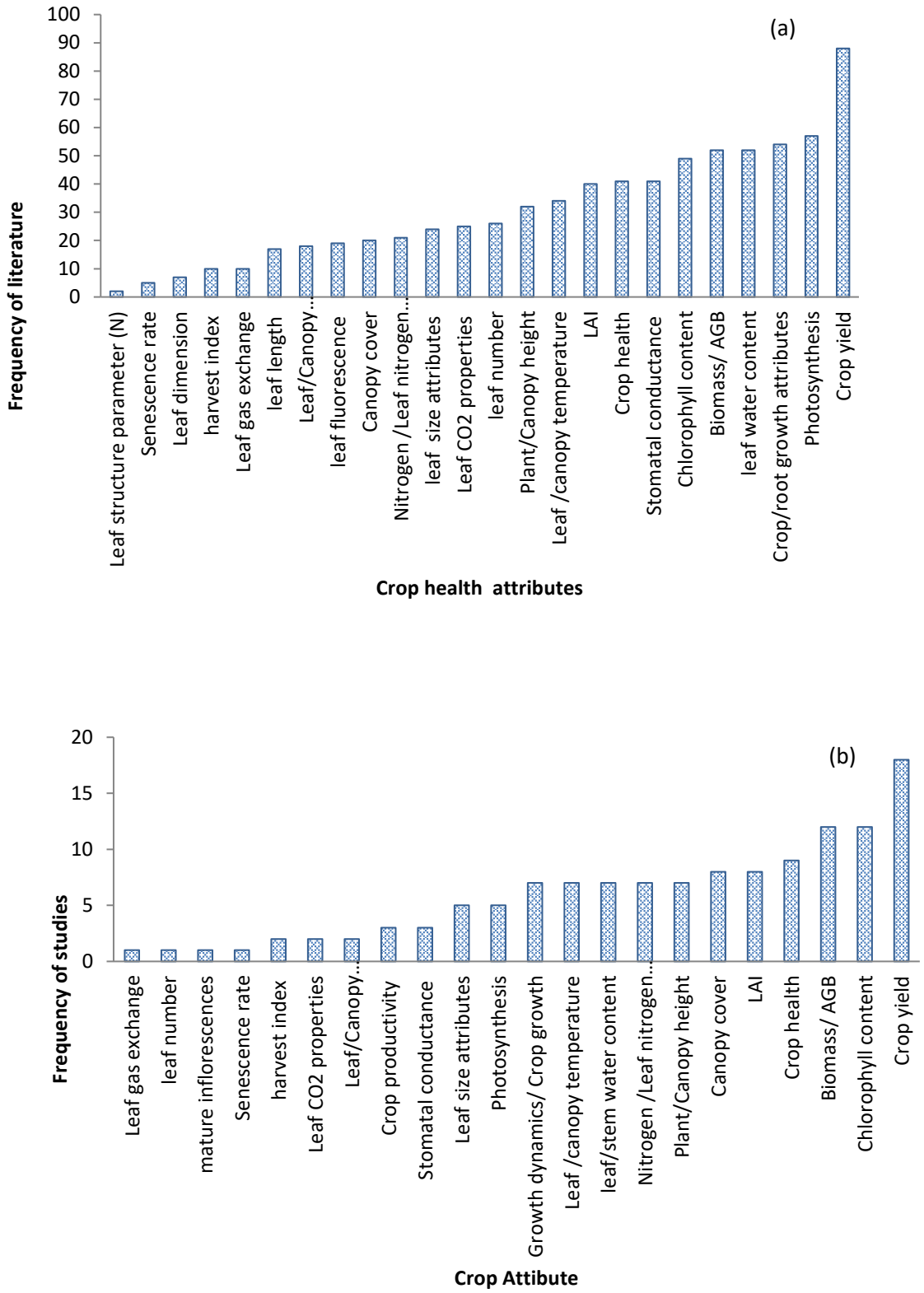


Figure 2-6 Frequency of studies that remotely sensed a specific crop attribute based on (a) all satellite and drone-borne sensors, and drone-borne sensors only

2.3.2.1 Types of sensors and their spectral resolutions

The findings of this review show that UAVs emerged as the most widely-used sensor for characterising the health status of NUS crops, as they have been utilised in 29 studies. Based on the retrieved literature in this study, the most widely-used satellite-borne sensors are Landsat, Sentinel- 2 MSI and MODIS (Figure 2.7(a)). The Landsat fleet and Sentinel 2 MSI were utilised in 11 studies. In terms of the drone-borne sensors, Thermal cameras (in seven studies) and Parrot Sequoia (in five studies) were the most widely used in the retrieved literature. These were followed by the Canon, RedEdge-MX, Micasense Altum, MCA6 and Sentera cameras, in order of frequency in the retrieved literature (Figure 2.7(b)).

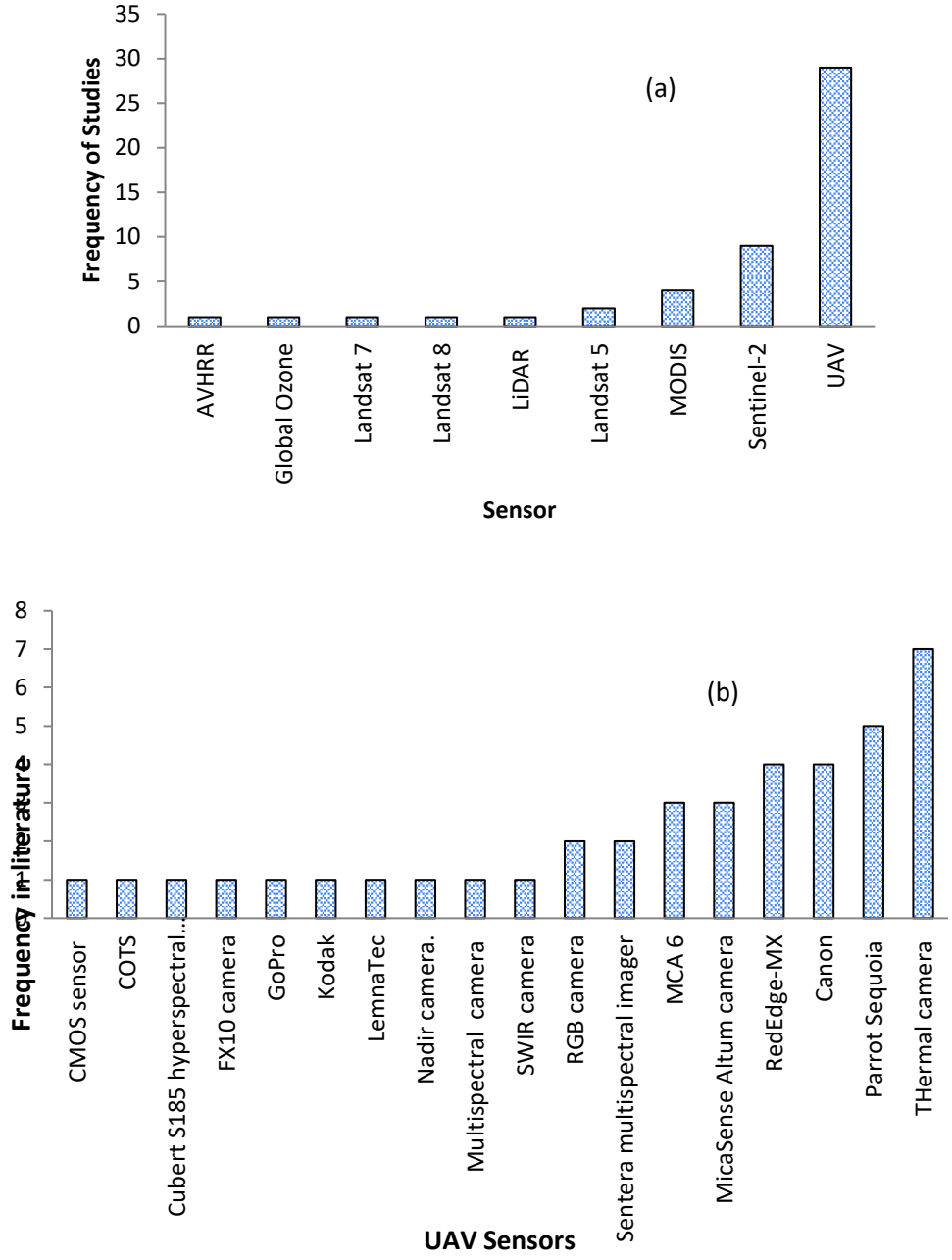


Figure 2-7. Frequency of (a) satellite, and (b) drone-borne sensors that have been used to map the spatial distribution of NUS and their attributes

The visible section of the electromagnetic spectrum, specifically the red, green and blue (RGB) sections, is primarily the most utilised wavelengths in mapping the spatial distribution of NUS crops and their health attributes (Figure 2.8). Specifically, the RGB section of the electromagnetic spectrum (EM) was utilised in 26 and 20 studies, based on drone and satellite-borne sensors, respectively (Figure 2.8). The second most widely-used section of the electromagnetic spectrum in the literature was the NIR section, which were utilised in 24 studies, based on drones, and 20 studies, based on satellite-borne sensors. When considering only the drone-borne sensors, the 12 studies utilised the Red Edge (RE) section of the electromagnetic spectrum, while only nine studies utilised the satellite remotely-sensed RE section of the EM to map the spatial distribution of NUS and their health attributes. Relatively few studies engaged the thermal bands in characterising the spatial distribution of NUS and their health attributes. In this regard, 10 studies used the drone-acquired thermal remotely sensed data and only eight studies utilised the satellite-acquired thermal bands (Figure 2.8). Across all platforms, the multispectral (broadbands) were highly utilised in literature, when compared to the hyperspectral (narrow) bands.

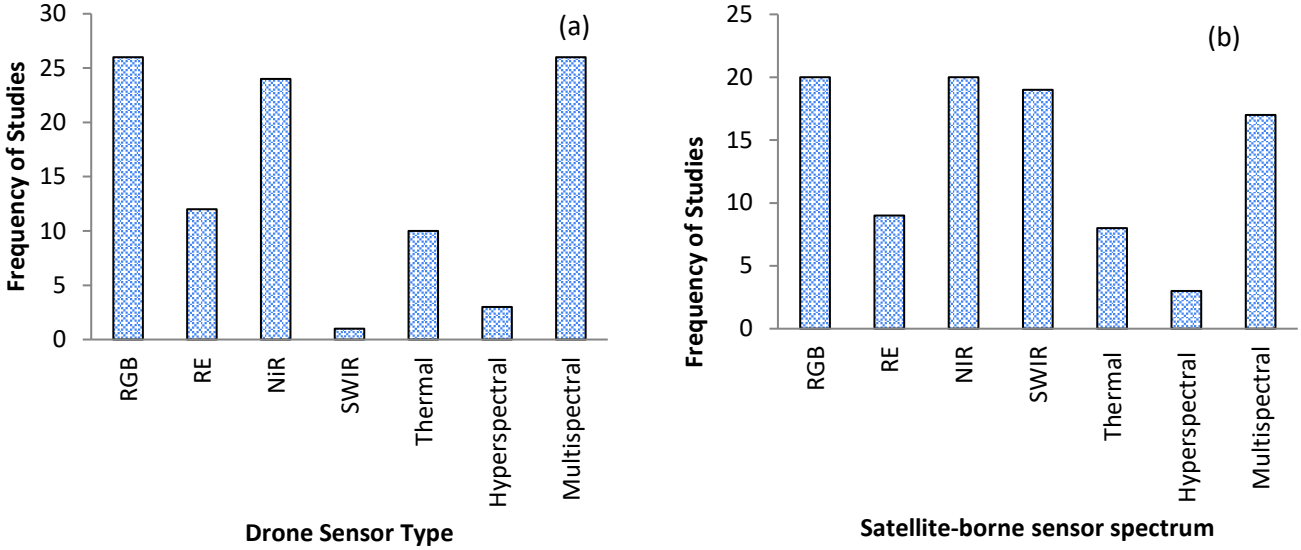


Figure 2-8 Frequency of the utilisation of a specific section of the electromagnetic spectrum covered by (a) drone and (b) satellite-borne sensors

2.3.2.2 UAV platforms utilised in literature

The DJI fleet had a marginally higher number of studies in relation to all the other platforms (n = 19). The other dominant drone platforms that were utilised in the literature to map the spatial distribution of NUS and their productivity and health attributes were the Octocopter in three studies, G-Q45 Quadcopter in two studies, Benzin Acrobatic in two studies, as well as Sensfly eBee in one study (Figure 2.9(a)). When assessing the frequency of the use of each specific sensor in the retrieved literature, it was observed that the DJI Phantom four Pro was the most widely-used platform across the board (appearing in eight studies), followed by the DJI Matrice, which was utilised in four studies (Figure 2.9(a)). Furthermore, the most widely-used drone platform was the quadcopter, followed by fixed-wing drones (Figure 2.9(b)).

Quadcopter drones were utilised in 21 studies, while fixed-wing drones were used in only two studies.

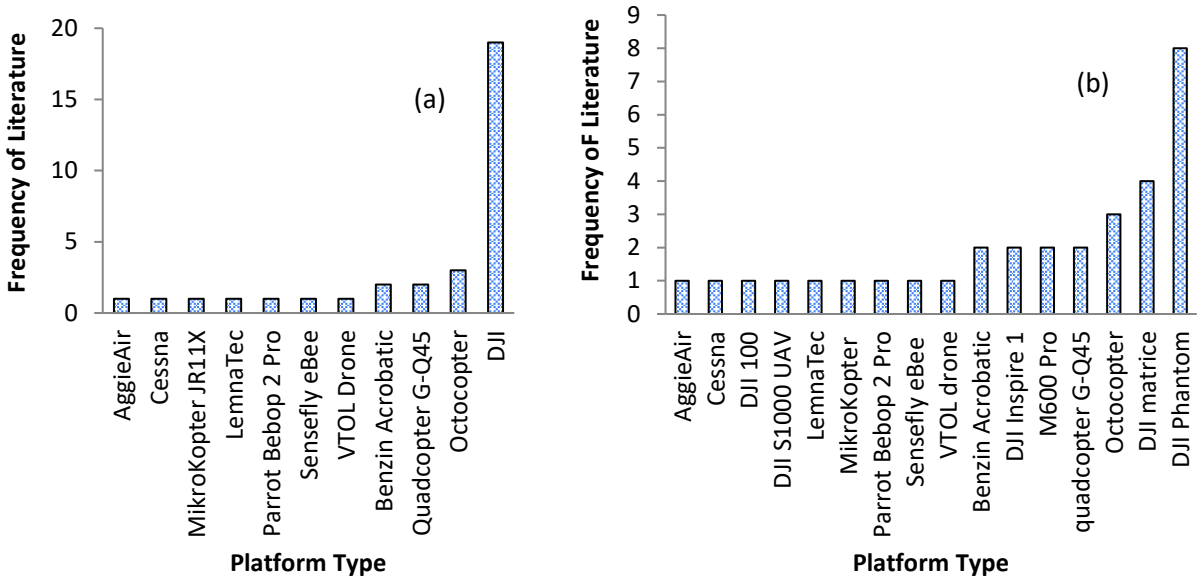


Figure 2-9 Frequency of (a) summarised drone platforms, (b) and specific drone types that were utilised in the literature to map the spatial distribution of NUS and their health attributes

2.3.2.3 Derived vegetation indices in remote the spatial distribution and health of NUS crops

Vegetation Indices (VI's) are algebraic combinations that are derived from multiple spectral bands, which are designed to estimate the vegetation vigour and vegetative characteristics (canopy biomass, absorbed radiation and chlorophyll content, among others) (Candiago *et al.* 2015). Due to the increasing prevalence and advancement of remote sensing technologies in agriculture, vegetation indices are more frequently being used to estimate various crop parameters/variables. The retrieved literature showed that the visible, near-infrared and red-edge sections of the electromagnetic spectrum are most commonly used in assessing crop health, e.g. Green (530 – 570 nm), Red (640 – 680 nm), Red-edge (730 – 740 nm), and NIR (770 – 810 nm) (Adam *et al.* 2014). The reflectance values of these prominent wavelengths are used to calculate the vegetation indices, such as the Normalised Difference Vegetation Index (NDVI), NDVI-red Edge (NDRE), Simple Ratio (SR), green normalised difference vegetation index (GNDVI), green chlorophyll index (CIgreen) and Red-edge Chlorophyll Index (CIred-edge), which were noted to be highly sensitive in predicting crop health and productivity parameters of NUS and they were frequently used in the retrieved literature, as illustrated in Figure 2.6.

2.3.2.4 Statistical and machine algorithms utilised in mapping the spatial distribution and health of NUS crops

The findings of this study show that a number of algorithms have been used in mapping NUS crops, ranging from basic statistical procedures, such as the simple regression techniques, to machine learning procedures (Figure 2.10). Machine learning procedures were the most widely-utilised algorithms in mapping the spatial distribution and health of NUS, as well as crop/plant productivity attributes, based on satellite and drone remotely-sensed data (Figure 2.10). Specifically, 71 and 33 articles utilised machine learning and regression algorithms that were based on satellite and drone-acquired data, respectively. Other than machine learning procedures, generic classification procedures (i.e. multicriteria decision-based mapping, fuzzy logic, maximum likelihood and minimum to mean distance classification) were widely used in mapping the spatial distribution of NUS and their health, as well as productivity elements based on both satellite and drone-borne datasets (Figure 2.10). The Principal Component Analysis (PCA), OBIA, BPNN and Houghton Line transformations were relatively moderate to less frequently utilised in the retrieved literature, in order of frequency.

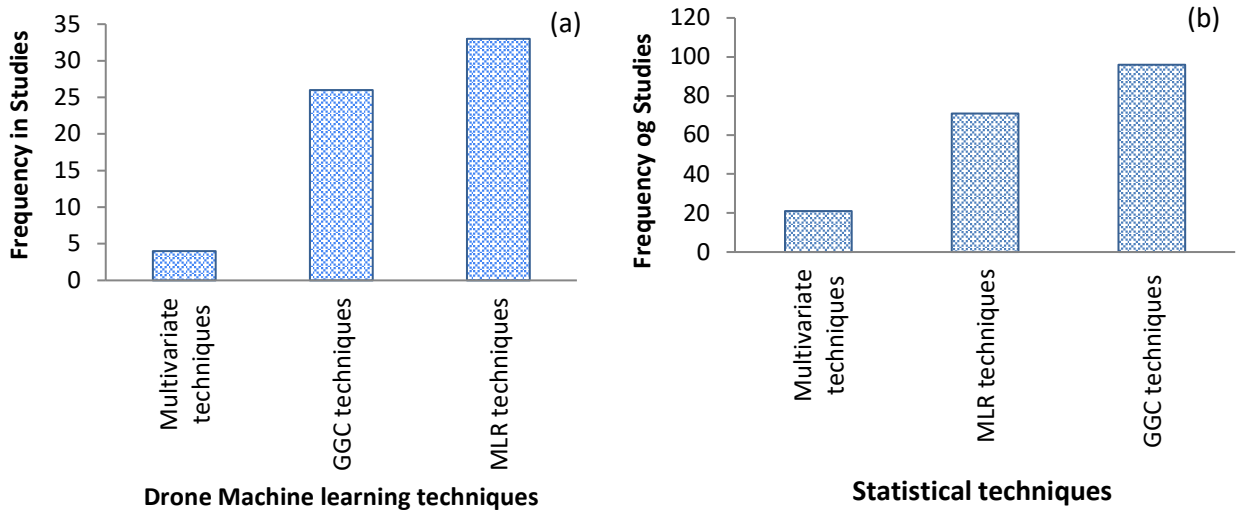


Figure 2-10 Frequency of statistical methods used in remote sensing NUS attributes, based on (a) all sensors and (b) drone-borne sensors only

The main machine learning algorithms utilised in the literature by drone-based systems were Linear Regression (35%), Support Vector Machine (SVM) (21%), Random Forest (17 %), and Partial Least Squares regression (PLS) (14%), in order of frequency (Figure 2.11). Linear regression techniques were the most frequent regression algorithms in the assessment of NUS crop health (Figure 2.11). The frequency of linear regression and machine learning-based algorithms was also detected during the bibliometric analysis, as illustrated in Figure 2.3 (b) (red cluster). The SVM was the second most widely-used machine learning ensemble in the retrieved literature, followed by RF, PLS, GLM and Artificial Neural Network (ANN), in order of frequency.

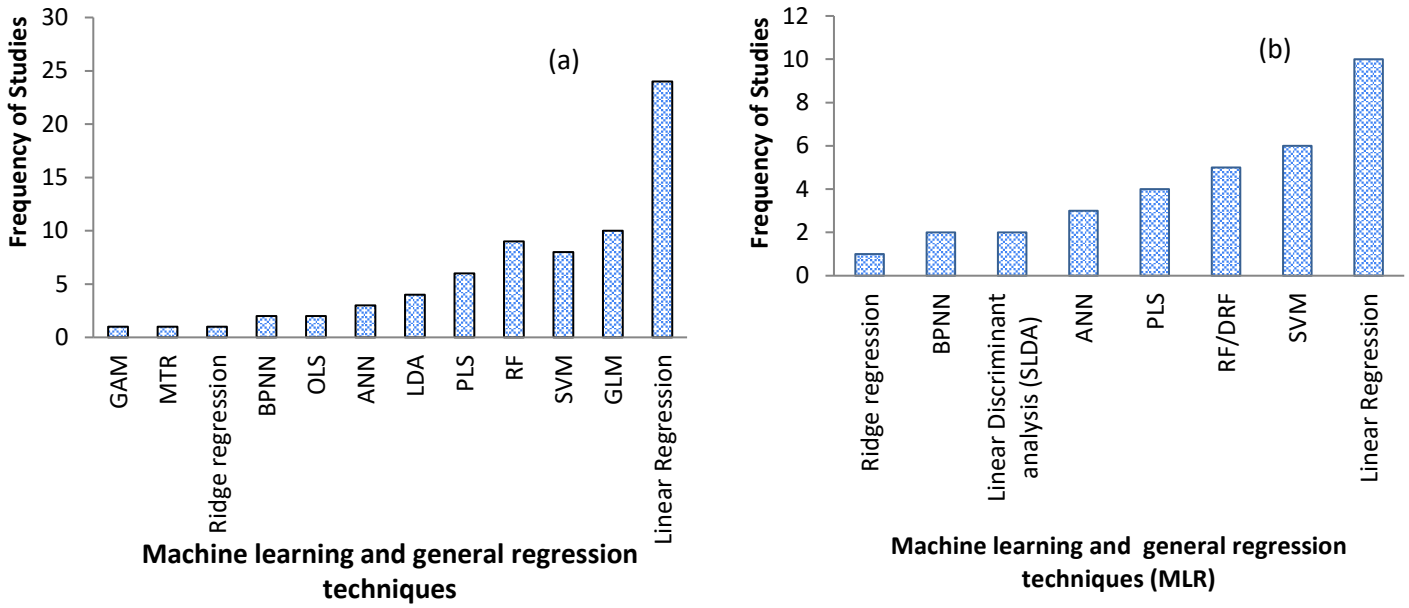


Figure 2-11 Frequency of machine learning and general regression techniques used in remote sensing NUS attributes, based on (a) all sensors and (b) drone-borne sensors only

In terms of the generic classification techniques, Pearson correlation, ANOVA, Maximum Likelihood (ML) and Empirical Line Method (ELM) were the most widely-used algorithms, based on the satellite and drone-borne remotely-sensed data (Figure 2.12). The Mahalanobis distance, parallel-piped, k-means and Canny Edge Filtering were among the popular generic classification algorithms used to map NUS, using both satellite- and drone-acquired remotely-sensed data. Studies based on the generic classification were relatively few for each algorithm (<5), compared to all other algorithms that were utilised in the retrieved literature. Generic classification algorithms, such as the Analytical Hierarchical Process (AHP), cluster analysis, fuzzy logic and multilayer perceptron, have not yet been utilised in the context of drone-based remote-sensing applications in crop mapping.

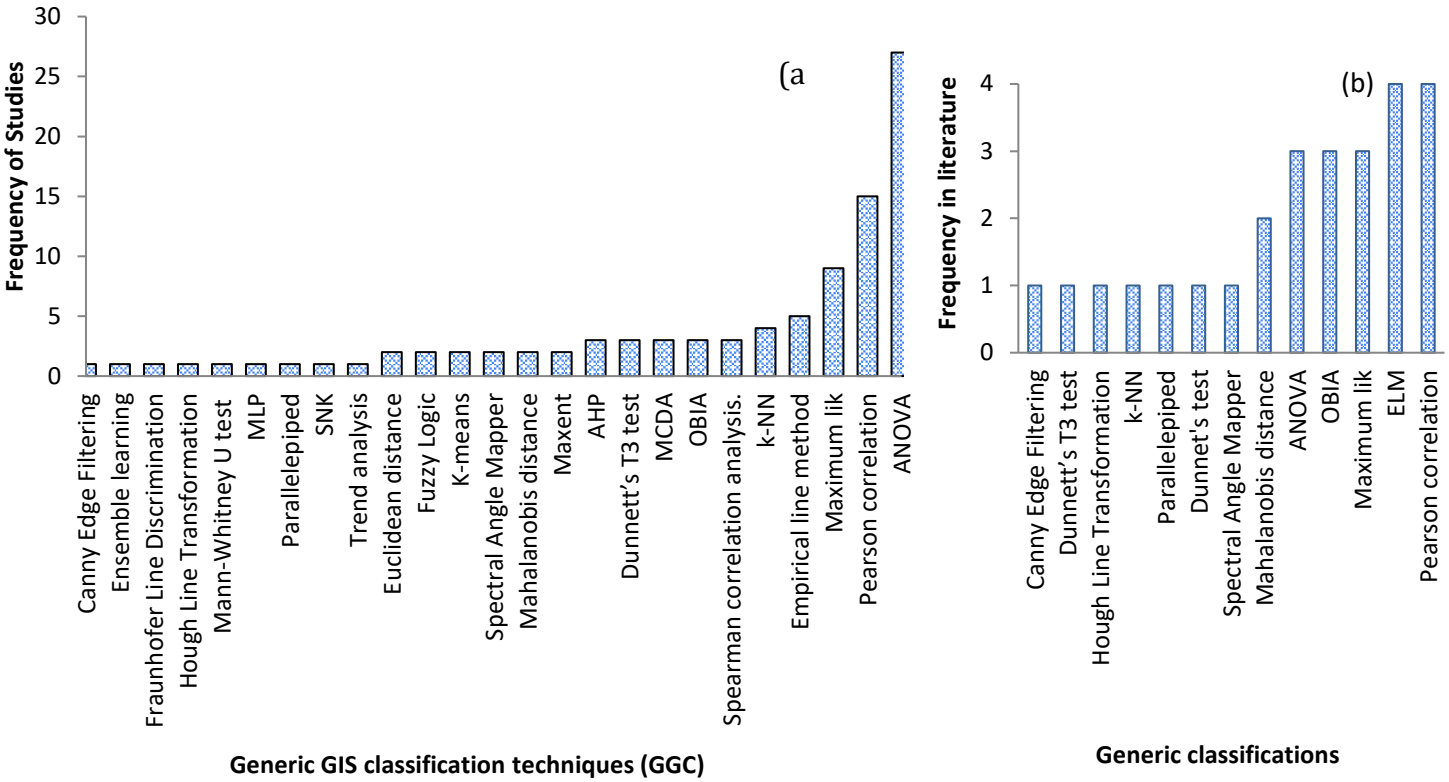


Figure 2-12 Frequency of generic GIS classification techniques used in remote sensing NUS attributes, based on (a) all sensors and (b) drone-borne sensors only

In terms of multivariate techniques, PCA, followed by cluster analysis and multiple regression, were the most widely-used, based on the satellite and drone-based studies (Figure 13.1). Studies based on multivariate classifications were relatively few for each algorithm (<5), compared to all the other algorithms utilised in the retrieved literature.

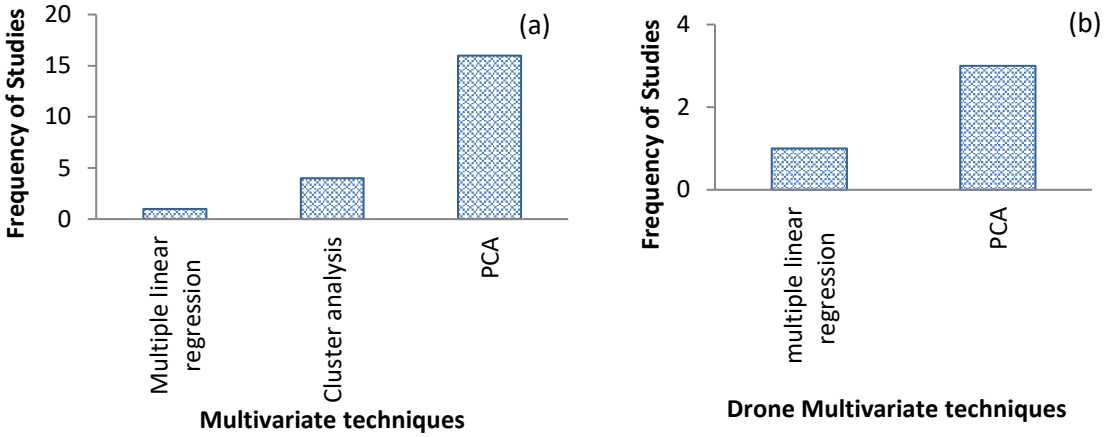


Figure 2-13 Frequency of multivariate techniques used in remote sensing NUS attributes, based on (a) all sensors and (b) drone-borne sensors only

2.4 Discussion

2.4.1 Evolution of drone technology applications in remote sensing

There have been numerous shifts in the utilisation of remotely-sensed data in mapping the spatial distribution of NUS and their health and productivity elements. These transitions can be clearly depicted and summarised by the bibliometric analysis of the co-occurrence of topical concepts derived from the titles and abstracts of studies that utilised drone remotely-sensed data in mapping the spatial distribution of NUS and their health attributes (Figure 2.14). Figure 2.14 shows a shift in the topical terms, from mere correlations based on RGB remotely-sensed data in mapping the NUS attributes from 2018 to 2019, to the utilisation of hyperspectral data to predict and map NUS attributes, such as yield and AGB, in 2020. Currently, the research efforts are focused on the fusion of drone remotely-sensed datasets, in conjunction with robust machine learning algorithms such as PLSR and SVM, for the characterising yield genotypes and canopy coverages, amongst others. The progress is discussed, in detail, in the following sub-sections.

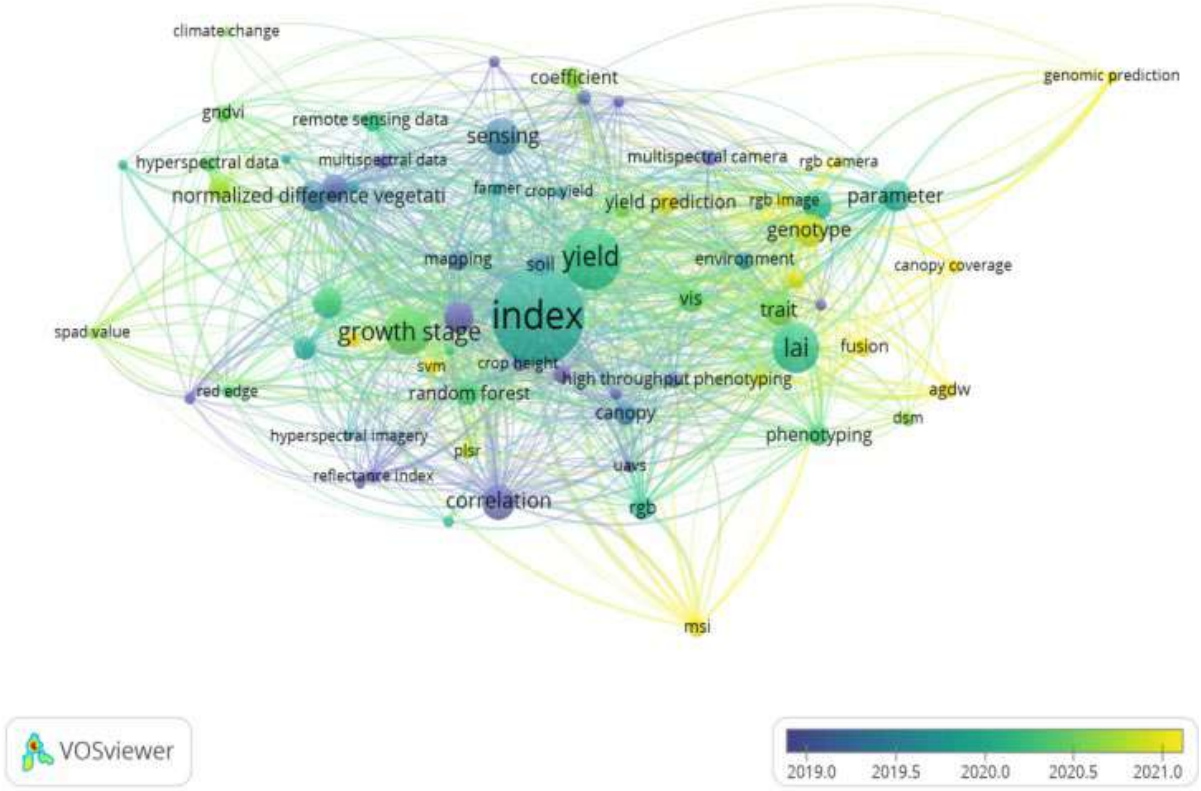


Figure 2-14 Evolution of concepts in drone-based remote sensing of NUS attributes

2.4.2 Frequency of publication and its geographic distribution

The results of this study showed that the studies that utilised drone-based remotely-sensed data to map and model the spatial distribution and health of NUS, as well as their productivity elements, gently increased from 2009 to 2014 (Figure 2.4). This trend was similar to those exhibited by published studies that utilised satellite remotely-sensed data (Figure 2.4 (b)). This could be explained by the rapid changes and improvements in all earth observation sensors and platforms. During that period, there was limited access to high spatial resolution remotely-sensed data, and fewer research activities were noticed in the context of NUS crops. In the context of satellite-borne sensors, this research period was mainly dominated by the use of Landsat and MODIS, which could not capture the fragmentation and heterogeneity associated with NUS. Furthermore, the only available fine spatial resolution images were dominantly commercial sensors, such as the WorldView and QuickBird sensors, which are associated with exorbitant costs that restrict the research activities. Above all, the role of NUS during that period was not recognised well, in relation to mainstream crops. Subsequently, the period between 2015 and 2022 is marked by a rapid increase in the number of drone platforms and the associated sensors, hence the rapid increase in the number of studies that utilised them in mapping the spatial distribution of the health and productivity elements of NUS. It is interesting to note that most of the studies were conducted in China, America, South America, Europe and Australia. This could be attributed to the fact that the earliest drone technologies began in Europe, the USA and China between 1849 and 1916 (Sibanda *et al.* 2021). Since then, the technology has been spreading. However, very few studies have been conducted in South Africa. Hence, there is an increasing requirement to improve the application of UAV technologies for precision agriculture, that area in line with the spatial extent and health assessments of the NUS crops,. Furthermore, it was noted that most of the studies were conducted on experimental plots, and in some instances, with irrigation facilities. No studies were conducted in the retrieved literature on remote sensing NUS in rainfed smallholder croplands.

2.4.3 NUS crop attributes that have been remotely sensed using drone-acquired data

The findings of this study showed that 11 NUS crop productivity elements were the most widely researched. More specifically, the crop yield, chlorophyll content, biomass, LAI, crop health, canopy height, cover, leaf nitrogen, leaf water content elements and stomatal conductance were the highly-researched productivity and health elements of NUS. These are the typical crop productivity and health elements; hence, they are anticipated to be covered extensively in the literature. The principal role of remote sensing applications is to optimise the yield production to improve the secure and balanced supply of nutrient-rich food. In addition, all research efforts have been focused on optimising the agricultural productivity (increasing the yields) on existing agricultural land, primarily to circumvent the perpetual increase in greenhouse emissions caused by the reduced vegetation cover, in order to maintain the functions of the ecosystem, while addressing Sustainable Goals 1 and 2, which address poverty and hunger (Van Wart *et al.* 2013, Mutanga *et al.* 2017, Duku *et al.* 2018, Jewan *et al.* 2022). Therefore, it is indeed expected that the yield will be the most widely-researched NUS

attribute, based on both the drone and satellite-borne remotely-sensed data (Van Wart *et al.* 2013, Jewan *et al.* 2022). It is worth noting that the yield, in most instances, is synonymous with the above-ground biomass. The findings of this study showed that 41% of drone-based remote-sensing studies on NUS estimated the crop's Above-Ground Biomass (AGB). AGB is an important parameter, which can also be utilised to predict the crop growth, yield and productivity. In this regard, it is also not surprising that AGB is amongst the most researched NUS attributes, in the context of remote sensing applications, for improving food and nutrition security in the light of the increasing climate variability and rapid population growth. Optimising the ABG and yield estimations is the principal role of remote sensing applications in precision agriculture, and it is anticipated that they will help to balance the increasing climate variability and rapid population growth with food production.

The findings of this study showed that Leaf Area Index (LAI) is one of the most researched elements of NUS in the retrieved literature, as it featured in 28% of the retrieved articles. This can be attributed to the fact that LAI is another accurate proxy of plant growth that has been widely proven to be linked to the chlorophyll content, biomass and, to an extent, the yield. More specifically, LAI represents the structural attributes of the leaf components estimated by the area of the leaf per unit of the ground surface area (Shuang Wu et al., 2022). The area covered by leaves per unit of the ground surface area changes with each growth stage of a plant. In this regard, this variable can better represent the space available for photon interception across the phenological stages of a crop, which can affect the yield. Subsequently, LAI is a plausible indicator of the canopy health and development that can be accurately remotely sensed, because it regulates and is linked to how light moves through the canopy. In addition, LAI can affect the canopy's surrounding microclimate. Radiation from the sun is intercepted by the leaves, which affects transpiration and, by consequence latent heat. Radiation can also affect leaf surface temperature and, by consequence, the surrounding air. This ultimately influences the photosynthetic nature of the leaf, as well as its stomatal conductance. It is in this regard that LAI was one of the most researched NUS attributes in the retrieved literature.

Chlorophyll content is another predominant NUS attribute, as it constituted 41% of the retrieved literature in this study (Raji *et al.* 2017, Singhal *et al.* 2019). The chlorophyll content in leaf tissue is a good indicator of a plant's physiological structure, nutritional composition and health (Tahir *et al.* 2018, Brewer *et al.* 2022). The antenna pigments in chloroplasts absorb incoming solar radiation during the photosynthetic activities; the resulting radiation is then transferred to the reaction centre pigments, which discharge electrons to activate the photochemical process (Monteoliva *et al.* 2021). More specifically, chlorophyll is a very relevant indicator of crop health, since it absorbs higher in the red (650-700 nm) and blue (400-500 nm) regions of the electromagnetic spectrum, in order to increase photosynthesis. The chlorophyll responsible for the higher absorption within the visible spectrum is chlorophyll a and b (Monteoliva *et al.* 2021). Hence, there is a positive relationship between the leaf total-chlorophyll content (chlorophyll a+b) and the solar radiation absorbed by the leaf tissues and the photosynthetic rate of the crop (Monteoliva *et al.* 2021). Therefore, the leaf's biophysical pigment and biochemical photosynthetic processes are linked to the plant's health and productivity; hence, it is regarded as a proxy for both its health and productivity (Tahir *et al.*

2018). It is in this regard that the chlorophyll content is the third highly researched NUS attribute, based on the remotely-sensed data.

Meanwhile, the chlorophyll content has been widely proven to correlate positively with the nitrogen content, in a wide range of crop species (Bojović and Marković 2009, Musa *et al.* 2016, Liu *et al.* 2019, Blumenthal *et al.* 2020). Indeed, approximately 75% of the total nitrogen is stored within the leaf chloroplasts (Li *et al.* 2013). In this regard, the results of this study showed that 24% of the retrieved studies estimated the nitrogen concentration by using drone remotely-sensed data, which it another sought-after attribute of NUS. The high nitrogen concentration is associated with an increase in chlorophyll concentrations and green pigments, which are all proxies of health and productivity that can be optimally characterised by using remotely sensed data. Furthermore, the literature shows that the application of nitrogen increases the CO^2 assimilation rate, which produces chlorophyll and the green pigment and are accurate proxy of the crop yield (Anas *et al.* 2020). For example, Muhammad *et al.* (2021) conducted a related study to evaluate the impact of various nitrogen and phosphorus levels, along with beneficial microbes, to enhance canola productivity. The results revealed that the nitrogen applied at a rate of 180 kg ha-1 increased the seed pod, the seed filling duration, the seed weight, the biological yield and the seed yield.

Findings of this review also showed that the structural parameters of a crop, such as the crop height, crop growth and canopy cover, were measured in 24%, 24% and 28% of drone-based studies, respectively. These crop structural parameters are related to AGB, LAI and chlorophyll content, which are all measured and used to predict the yield, and which, in turn, is directly linked to food security. The crop height and growth attributes directly interact with the incident electromagnetic energy that is typically measured and used to model productivity. Furthermore, these physiological crop variables are sensitive to variations in the environmental conditions (precipitation, temperature, soil type, etc.) and biochemical conditions (fertility, weeds, pests and diseases). In this regard, these structural attributes are instrumental in monitoring the health of crops in order to optimise their production.

The findings of this study showed that stomatal conductance is another optimal health parameter of NUS, as assessed by 11% of the retrieved studies that utilised drone data. The stomatal conductance is a measure of the degree of stomatal opening and can be used as an indicator of the leaf gas exchange (Iseki and Olaleye 2020). Stomatal conductance is strongly associated with the leaf transpiration rate, photosynthetic efficiency, chlorophyll concentration and nitrogen concentration (Wijewardana *et al.* 2019). More specifically, periods of crop water stress or drought stress will be limited by the CO^2 concentration at carboxylation sites (Cc) inside the chloroplast, which is determined by the CO^2 diffusion components, i.e. stomatal conductance (gs) and mesophyll conductance (gm) (Ouyang *et al.* 2017). Since the stomata control the CO^2 diffusion into the leaf tissue and water diffusion/transpiration out of the plant, it has been therefore proven that under water-deficit conditions, the plant stomata will close to prevent a major water loss, which consequently reduces photosynthesis via the decreased influx of CO_2 (Ouyang *et al.* 2017). In this regard, higher stomatal conductance and high photosynthetic efficiency are associated with limited moisture stress; hence, plants with a high chlorophyll content and green pigment are healthier. Therefore, stomatal conductance was also

researched in the retrieved literature, because it is an accurate proxy of the moisture stress and health status of NUS, and it could help to regulate irrigation schedules and to optimise food production. A study by Chai *et al.* (2016) measured the morpho-physiological traits of Bambara groundnuts exposed to a progressive mild drought in a controlled environment. Drought stress reduced their stomatal conductance significantly (p\0.01). Furthermore, a higher stomatal density and reduced leaf area were observed in drought treatment plants (p\0.01), which suggests that NUS crops are more resistant to biotic and abiotic stresses, such as drought and water stress.

In most studies, there was a strong correlation between stomatal conductance and the water relation status, or leaf water content. Specifically, 24% of studies based on drone-remote sensing researched the NUS leaf water content. The water content is an important indicator of a crop's health; a plant with a higher water potential will produce a greener pigmentation and increase crop productivity. According to (Ouyang *et al.* 2017), under water deficit or conditions of mild water stress, the stomatal conductance, internal CO_2 concentration and/or net assimilation rate within a plant will decline, and the An/gs ratio increases. Various studies have noted that crop water stress allows for a decline in chlorophyll a, chlorophyll b and the total chlorophyll of various crops (Maes and Steppe 2012, Majid and Roza 2012, Pineda *et al.* 2021). However, when the severity and advancement of water stress is associated with the non-stomatal regulation of photosynthesis, a decreased light and CO_2 concentration, a reduction in photochemistry, a decline in the activity of photosynthetic enzymes, as well as a lower mesophyll conductance will occur. According to a study by Chibarabada (2018), the stomatal conductance of three grain legume crops (groundnut, dry bean and Bambara groundnut) was measured under three water treatments. The results showed that, under varying water regimes, NUS crops adapted to limited soil water through stomatal regulation and a reduction in the canopy size and duration. Furthermore, the Bambara groundnut showed a positive attribute for water-limited conditions; it had the lowest stomatal conductance of all the water regimes, compared to the other crops. This example illustrates why stomatal conductance and foliar temperature have also been considered in the context of NUS production.

2.4.4 Sensors and platforms used in remote sensing NUS

The results of this study showed that there were more drone-based remote sensing studies on NUS than those based on satellite remotely-sensed data (Figure 2.7). The dominance of drone-based remote sensing studies on NUS could be explained by the fact that most of these crop species are orphaned, neglected and under-utilised. In this regard, they are generally planted on smaller areal extents, in relation to mainstream crops, such as maize, except sorghum. As a result, very few freely-accessible satellite-borne sensors have finer spatial resolutions that could capture the dynamics of these crops, when compared to UAVs. Our findings o supported this aspect by showing that Sentinel-2 MSI was the most widely-used satellite-borne sensor in mapping the spatial distribution, health and productivity parameters of NUS. Sentinel 2 MSI has a minimum spatial resolution of 10 m and a spectral resolution that covers the red edge section of the electromagnetic spectrum, which is sensitive to crop health. These Sentinel attributes make it the most suitable sensor for mapping NUS, which are often grown in small, fragmented fields.

In terms of the drone sensors, our findings showed that the RGB and multispectral cameras (>3 bands) were the most widely used in the retrieved literature, when compared to hyperspectral sensors (Figure 2.8). This could be explained by the exorbitant expenses associated with the hyperspectral sensors, in relation to multispectral cameras. RGB and Multispectral cameras generally cover the visible, including the red edge, and the NIR sections of the electromagnetic spectrum, usually in less than six broader spectral band resolutions (i.e. the MicaSense Altum) (Ndlovu *et al.* 2021). Meanwhile, most of the hyperspectral sensors, such as the Cubert S185 hyperspectral sensors, cover a wavelength range between 450 nm and 950 nm that are composed of 125 channels at a resolution of 8 nm at 532 nm, and they range from visible to the near-infrared (450 nm–950 nm), with a sampling interval of 4 nm (Zheng *et al.* 2020). The RGB type of sensors are less accurate than multispectral or hyperspectral images, which cover a wider spectral range. Meanwhile, hyperspectral sensors have been widely proven to offer robust, narrow, relatively-contiguous bands that are sensitive to subtle crop variations, when compared to broadband sensors, despite their association with exorbitant prices (Thenkabail *et al.* 2002, Thenkabail *et al.* 2013, Marshall and Thenkabail 2015). With respect to sensors, the findings of this study also showed that very limited studies utilised a camera that captured data in the red edge section of the electromagnetic spectrum. Multispectral sensors that capture the red edge section in retrieved studies are the MicaSense series and Parrot Sequoia, which were utilised to map and monitor NUS health attributes in seven and five studies, respectively. The red edge section has been widely proven to be sensitive to minute variations in the plant attributes associated with health and productivity, such as the LAI, chlorophyll content, stomatal conductance, AGB and nitrogen content. For instance, the increased chlorophyll content, LAI and biomass generally result in increased absorption in the red region and push the red edge to longer wavelengths. Considering this phenomenon, the red edge has become one of the most sought-after sections of the electromagnetic spectrum in crop monitoring. There is a need to engage and increase the robustness of hyperspectral and red edge sensors for assessing and monitoring the health elements of NUS, if their productivity is to be enhanced.

The findings of this study also showed that multi-rotor drones were the most widely used platforms in the retrieved literature, when compared to the fixed-wing drones (Figure 2.9). This could be explained by the fact that, despite the long endurance associated with fixed-wing drones, they often require take-off and landing space, which was not always available in the experimental sites where most studies were conducted (Zaludin and Harituddin 2019). Meanwhile, the multirotor drones are characterised by a Vertical Take-Off and Landing (VTOL) (Zaludin and Harituddin 2019). This makes them more suitable for utilisation in research areas where there is limited take-off space. Furthermore, the multi-rotor drones are relatively cheaper than fixed-wing drones. This could also explain the high frequency in the utilisation of DJI drones in the retrieved literature (Figure 2.9). Most of the DJI drones are quadcopter and multi-rotor platforms that are associated with VTOL. The advantage of DJI drones is that some of them are associated with an automated VTOL, which makes them easy to fly as they require less expertise in drone piloting. According to Sibanda *et al.* (2021), the DJI platforms are generally compatible with many types of sensors, in relation to other platforms, hence they are widely utilised.

2.4.5 Performance of vegetation indices, classification and estimation algorithms

The findings of this review showed that vegetation indices, such as NDVI, NDRE, SR, GNDVI, CIgreen and CIred-edge, were the most frequently used in mapping the spatial distribution of NUS and their health parameters. The engagement of vegetation indices can be explained by the optimal performance that has been demonstrated by a large and growing body of literature (Thenkabail *et al.* 2002, Thenkabail *et al.* 2013, Marshall and Thenkabail 2015). More specifically, vegetation indices are robust because they can suppress the background effects, in comparison to the bands (Thenkabail *et al.* 2013). Furthermore, vegetation indices derive their strength from two or more sections of the electromagnetic spectrum that are sensitive to crop elements, such as the chlorophyll content, LAI, ABG, nitrogen concentration and stomatal conductance (Brewer *et al.* 2022). Furthermore, a large body of literature illustrates that red-edge-based vegetation indices are more robust than conventional vegetation indices, such as NDVI and SR. NDVI tends to be insensitive to increases in the chlorophyll, LAI and biomass. As previously mentioned, the red edge has a high association with plant physiological traits, such as the leaf angle distribution, chlorophyll concentration and leaf area index, which directly influence the vegetation spectral reflectance. Subsequently, red-edge derived vegetation harnesses this robust sensitivity to variations in the LAI biomass and chlorophyll content, among others (Guyot *et al.* 1992, Filella and Penuelas 1994, Mutanga and Skidmore 2004, Mutanga *et al.* 2012). Their optimal performance could explain the frequency of their utilisation in the retrieved literature.

As was suggested in many studies, the combination of sensitive spectral variables with robust and efficient algorithms produces accurate models (Brewer *et al.* 2022). The findings of this study showed that the machine learning algorithms that were utilised in conjunction with drone-acquired remotely-sensed data were primarily linear regressions (35%), followed by SVM (21%), RF (17%) and PLS (14 %). Machine learning regression algorithms present a potential approach for generating adaptive, robust and fast estimates. A growing body of recent studies has demonstrated the efficiency and optimal performance of the machine learning algorithms for estimating crop biophysical parameters (Sapkota *et al.* 2020, Brewer 2021, Wu *et al.* 2022). The high frequency in the use of linear regression models could be attributed to the fact that they are simple to implement, based on various platforms ranging from Microsoft Excel to complex programming programmes, such as R- Statistical software and the Google Earth Engine. Despite their high frequency in the retrieved literature of this current study, LR models are parametric algorithms that are associated with data normality assumptions, which are often challenging to attain. Furthermore, these models tend to be susceptible to overfitting and outliers. Therefore, more efforts have been channelled towards machine learning algorithms, such as SVM and RF, which are non-parametric, robust and less prone to outliers. Specifically, Random Forest has (i) hyper parameters that are easy to adjust, (ii) it is robust to outliers, overfitting and data dimensionality, (iii) it has low bias and moderate-to-minimal variance, due to the averaged trees, (vi) it works well with both continuous and categorical variables, (v) it has a capability of discerning the importance of predictor variables, and (iv) it is relatively resistant to multicollinearity and its tree depths are greater (Mutanga *et al.* 2012, Ehlers *et al.* 2022, Singh *et al.* 2022). However, numerous generic classification algorithms were utilised in

the retrieved literature, such as the maximum likelihood, and minimum distance to mean classification algorithms. Despite the optimal performance in the literature, these algorithms are less robust, when compared to machine learning algorithms.

2.4.6 The challenges and opportunities of mapping the spatial distribution and health of NUS crops

Even though much progress has been made regarding the utility of drones in crop mapping and health quantification, several challenges are impeding the propagation of drones, especially in Africa. Many African countries are battling to address a variety of drone-related issues, such as privacy, public safety and the possession of malicious drones (Nguyen and Nguyen 2021). Moreover, restrictive UAV or drone regulations across many developing regions, including Africa, are holding the industry back. Among several other issues, in order to utilise UAVs, users are required to seek permission from landowners and municipalities, and they can only operate in certain areas (Kemp *et al.* 2021).

For example, the civil aviation authorities in many countries aim to prevent UAVs from entering the flight paths of other aircraft and are attempting to construct an inclusive system that accommodates them into their respective air navigation and surveillance systems (Grote *et al.* 2022). Only a few African countries, including Ghana, Kenya, South Africa, Rwanda and Tanzania, have established complex legal requirements that govern the use of Unmanned Aerial Vehicles (UAVs) in varied airspace practices. For example, some drone restrictions in South Africa include that UAVs weighing no more than 7 kilograms (15.4 pounds) are permitted to operate (Kemp *et al.* 2021). In this regard, the regulation on the mass of a UAV at take-off tends to indirectly restrict the areal extent that can be covered, as well as the size of the camera to be mounted for research purposes, amongst other uses (Sibanda *et al.* 2021). Specifically, due to the weight restrictions, many of the sensor types that are frequently used tend to be lightweight, small-sized and general consumer grades, with limited spectral resolutions (Sibanda *et al.* 2021).

Furthermore, SACAA asserts that drone operators should maintain a continuous Visual Line of Sight (VLOS) with their drone during flight and that RPAs are not permitted to travel Beyond Visual-Line-of-Sight (BVLOS) (Kemp *et al.* 2021). Sibanda *et al.* (2021) stated that supporting the regulation and operationalisation of BVLOS drone technology applications will facilitate the coverage of greater areas on a single mission, which will improve the cost-effectiveness in the acquisition of VHR imagery.

Other restrictions include a requirement for drone operators to obtain an RPAS Operator Certificate (ROC) from the civilian aviation authority before flying in overcrowded regions. UAV operators should also obtain insurance to cover their liability in the case of committing physical or bodily harm to another individual while operating their drone. Aside from these regions, numerous African countries are still attempting to establish the necessary regulations that endorse UAV operations. As a result, these regulations are increasingly complicated and expensive. More notably, the cost of obtaining a drone pilot's license is high, with estimates ranging from USD 1 500 to USD 2 000 (Sibanda *et al.* 2021). Furthermore, the value of drone

platforms, including sensors, is prohibitively expensive for several minority groups or scholars, which makes these technologies unattainable for research purposes in the vast bulk of sub-Saharan African countries (Sibanda *et al.* 2021).

2.4.7 Research gaps and opportunities

The following gaps were noted in assessing the use of drone-based remotely-sensed data in mapping the spatial distribution and health of NUS:

- Few studies have sought to evaluate the utility of drone technology for characterising crop phenomena in the Global South.
- The observation of NUS crop health has garnered minimal attention and interest from the scientific community.
- Although NUS crops are reported to be resistant to abiotic stresses, such as droughts and heat stress, most of this information is anecdotal and inconsistent. Hence, there is a requirement to generate more empirical information on the ecophysiology and morphology of NUS. This should be associated with agro-ecologies, in order to enhance the crop spatial distribution that is used for classification.
- Similarly, with regard to the agronomic characteristics of NUS, comparable information on the nutritional value of NUS is inadequate. Addressing this information gap would help to emphasise the nutritional benefits of NUS crops and to market them as healthy alternatives.
- Only a few research studies have sought to evaluate the effectiveness of robust machine learning algorithms in conjunction with vegetation indices for predicting the spatial distribution and health of NUS crops
- Few studies have attempted to assess and leverage the potential synergies between drone and satellite-borne datasets, particularly since the release of the freely-accessible Sentinel 2 MSI and Landsat series.

It is therefore essential to conduct market research in order to determine the prospects for commercialising or interfering with the existing market and to ensure inclusivity for NUS crops within the agriculture sector.

2.4.8 The way forward: Closing the gaps in the utilisation of drone technology in mapping the spatial distribution and health status of NUS crops

Scholarly attention needs to be bestowed on promoting the application of UAVs for the assessment of the spatial extent and health of NUS crops at a field scale. There is a need to increase and extend research efforts toward rainfed smallholder croplands in remote sensing NUS, in order to optimise their productivity, while strengthening the livelihood and food security of marginalised subsistence farmers.

UAVs are becoming increasingly common in agricultural research as relevant sources of high temporal resolution data for agricultural health monitoring, which has been previously limited on a local scale. UAVs provide immense opportunities by incorporating climate-smart

agriculture and precision agriculture into smallholder farming, since they deliver high-resolution imagery at user-defined temporal resolutions, which are beneficial for crop health monitoring and independent decision-making.

For example, the integration of multi-rotor drones, such as the DJI series, in conjunction with multispectral sensors, such as the MicaSense RedEdgeMX and MicaSense Altum cameras, has been demonstrated to deliver accurate models of crop health characteristics at a field scale. Furthermore, the progressive adoption of advanced narrow-band hyperspectral technology offers the perfect opportunity to evaluate the diverse crop health attributes at a high optical resolution. This assessment would also allow for the creation of a wide variety of optimal VI's, which are directly associated with the morphological and physiological characteristics of a NUS crop, and which allows for the quantification of NUS crop health. Moreover, in order to improve the research, based on the application of UAVs to estimate NUS crop health, it is essential to compare, identify and implement robust and reliable non-parametric machine learning algorithms, such as RF, that enhance the prediction accuracy with fewer assumptions.

The use of multi-fusion techniques, based on the combination of ideal vegetation indices and machine learning algorithms, will ensure the attainment of the best optimal accuracy. As a result, additional research is required to evaluate the use of UAVs with the distinct spectral characteristics of NUS. It would be valuable to research whether UAV sensors that measure spectral reflectance, along the SWIR and NIR sections of the electromagnetic spectrum, enhance the prediction of NUS water stress at a farm scale. As a result, near-real-time spatially-explicit information on the spatial extent and health of NUS crops can certainly assist smallholder farmers in the field to optimise their yields, by detecting variations in their crop morphology. Near-real-time fine-resolution NUS information could be used to draw up early warning systems for smallholder farmers, in order to prevent any damage or reduction to the crop yields.

2.5 Conclusion

The objective of this study was to systematically assess the literature on the progress, challenges, gaps and opportunities of utilising drone-derived remotely-sensed data to map the spatial distribution and health of NUS crops. The review established that a limited number of studies have focused on mapping the spatial distribution of NUS, while most of the literature has focused on mapping their health and productivity elements. Overall, these studies are still relatively few, in relation to those on the main crops, partly due to exorbitant prices of drones, pilot licenses, the shortage of personnel with the necessary skills, and the stringent regulations for securing and operating drones, amongst others. On the other hand, UAVs are cutting-edge technology with a high potential for offering spatially-explicit data, as well as timely, efficient and reliable assessments of the NUS health attributes. Furthermore, there is an urgent need to enhance the application and enrolment of UAV remotely-sensed data within smallholder farming systems, so as to provide farmers with the opportunity to acquire expertise, based on the operation and application of drone platforms and sensors, to improve optimal decision-making and to further enhance crop productivity on a farm scale. Improving the productivity of NUS helps to sustain and develop the indigenous local food systems and cultural identities,

while promoting gender equality in indigenous communities and for the women who are predominantly rearing these crops. Furthermore, this presents a chance to restore a local food culture and food systems, while empowering smallholder farmers, who are frequently the stewards of this agro-biodiverse resource.

3 MAPPING TARO AND SWEET POTATOES ON SMALLHOLDER FARMS USING UAV MULTISPECTRAL IMAGERY AND MACHINE LEARNING

3.1 Introduction

Food security and sustainability are major global concerns, due to population growth, climate change and the decreasing availability of arable land (Vågsholm *et al.* 2020). The global population is predicted to rise to nine billion by 2050, which includes 1.5 billion in SSA. In the year 2022, the African continent witnessed a notable surge in the number of malnourished individuals, which amounted to nearly 282 million and denoted a substantial escalation of 57 million people, since the onset of the COVID-19 pandemic (FAO et al., 2023). Approximately 868 million people experienced moderate or severe food insecurity, which highlights the significant scale of the issue. Moreover, 78% of the African population needed help to afford a nutritionally-adequate diet throughout 2021 (FAO et al., 2023). Since 2023, the progress toward achieving universal food and nutrition targets has been slow, which amplifies the pressure on agricultural lands that are already affected by climate change to meet the rising food demands (Kumar *et al.* 2022). Smallholder farmers who produce most of the local food are especially vulnerable, as 80% of the crops they grow are rainfed (Fan and Rue 2020). Therefore, to address these nutrition issues, it is crucial to identify and grow alternative drought-tolerant crops on smallholder farms (Mabhaudhi *et al.* 2017). Neglected and under-utilised crops (NUS) like sweet potatoes and taro are stress-tolerant, nutrient-rich crops that can grow under adverse environmental conditions. They are renowned for having a natural resistance and tolerance to pests, diseases and droughts (Mabhaudhi *et al.* 2017). These NUS are also suitable for sustainable farming because of their adaptability, lower inputs and diversification (El Bilali *et al.* 2023).

The incorporation of NUS crops into both commercial and subsistence food systems has been difficult, partly due to a lack of knowledge on their spatial distribution and phenotypical attributes (Mabhaudhi *et al.* 2017). Furthermore, there need to be a clear criterion or spatially-explicit techniques for assessing their spatial distribution and health withon smallholder farms in developing countries. Therefore, it is imperative to develop accurate and time-effective quantitative spatial techniques for detecting and mapping their distribution. This information could support better resource management and decision-making and ultimately increase productivity, sustainability, as well as food and nutrition security, through optimised production.

Generally, traditional field surveys have commonly been used to measure the spatial extent, suitability, growth and morphology of NUS crops. However, such methods are time-consuming and expensive, as well as impractical for continuous and effective crop monitoring (Ndlovu *et al.* 2021). Over several decades, satellite-based earth observation technologies have demonstrated their effectiveness in discriminating and mapping the spatial distribution of crops to optimise production. For example, according to Wang *et al.* (2022), Sentinel-2 Multispectral Instrument (MSI) data with machine learning can differentiate alfalfa (*Medicago sativa L.*)

from other crop types with overall accuracies of over 90%. Moreover, Sharifi (2020) effectively combined Sentinel-2 MSI bands and vegetation indices to achieve accurate crop nutrition mapping. Nevertheless, even with the advancements in satellite remote sensing applications, the effectiveness of satellite data for identifying and delineating the scope of NUS on a farm scale is constrained by the relatively coarse spatial and temporal resolution of openly accessible satellite remotely-sensed datasets (Ndlovu *et al.* 2021). Satellite-borne datasets such as Sentinel-2 MSI and Landsat, typically have a coarser spatial resolution, which often masks out the fine-scale variability within farms or individual plants. In recent years, Unmanned Aerial Vehicles (UAVs), also known as drones, have gained significant traction in precision agriculture for remotely sensing crops (Sun *et al.* 2021). UAVs equipped with advanced high-resolution sensors provide precise and timely data, which are well-suited for monitoring crops at a plot level (Ndlovu *et al.* 2021). They allow flexible altitude control, which makes them well-suited for assessing the spatial distribution of crops at a farm scale. By flying at customised altitudes, UAVs can capture high-definition imagery, which enables the detailed analysis of minute crop formations at a farm scale. Several studies have employed UAVs for proximal sensing in environmental applications (Everitt *et al.* 2007, Yang *et al.* 2012, Sankaran *et al.* 2018). For instance, Everitt *et al.* (2007) mapped the spatial distribution of wild taro by using a Kodak colour-infrared camera, in combination with supervised image analysis techniques. This approach yielded kappa accuracies that ranged from 83.3% to 100%. Although significant progress has been achieved in the utility of high-throughput phenotyping technologies for assessing crop distribution, very few studies have attempted to assess their accuracies in mapping and monitoring the spatial distribution of NUS crops, especially in typical smallholder croplands in the Global South. It is hypothesised that high-resolution UAV data, in conjunction with robust classification algorithms, can accurately characterise the spatial distribution of NUS crops within smallholder fields.

A variety of classification techniques have been proposed for mapping the spatial distribution of crops. These techniques leverage various data sources, like remote sensing imagery and ancillary data, to differentiate crop types at different spatial scales. Machine learning and deep learning algorithms, such as Support Vector Machines (SVM), Artificial Neural Networks (ANN), Random Forests (RF) and Naïve Bayes (NB) are widely used, due to their ability to accurately detect and quantify the spatial extent of crops (Sarker 2021). These algorithms can learn complex patterns and relationships from training data, and they can classify pixels or objects into crop classes. For instance, a study by Mazarire *et al.* (2020) explored the efficacy of SVM and RF in classifying crops within heterogenous landscapes in South Africa. SVM demonstrated the highest performance, by achieving an overall accuracy of 95% and a kappa accuracy of 94%. Furthermore, Ayele and Tamiru (2020) designed a chickpea-type classification model using ANN, SVM and Decision Tree (DT) algorithms. The experimental results demonstrated that the best algorithm was DT, which achieved an accuracy of 97.5%.

A large body of research has demonstrated that Vis, like the Normalised Difference Vegetation Index (NDVI), the Soil Adjusted Vegetation Index (SAVI) and the Normalised Difference Red-Edge index (NDRE) often improve the accuracy of the classifications. For instance, Asgari and Hasanlou (2023) created a Vis mapping approach to accurately identify specific crop types,

such as rapeseed, by utilising phenological and spectral metrics that were extracted from Sentinel-2 MSI images. Among the various Vis employed for crop mapping, the Atmospherically Resistant Vegetation Index (ARVI) yielded superior accuracies. The RF, KNN and GB models, utilising the ARVI index, achieved overall accuracies of 95%, 90% and 88%, respectively. VIs are known to overcome the influence of shadows and other background noise (Mutanga *et al.* 2023). Although this has been demonstrated extensively, based on satellite-borne data, there is a need to expand these research efforts by including crops, such as taro and sweet potatoes, on smallholder farms.

Therefore, this study sought to assess the potential of advanced machine learning methods and UAV-derived multispectral data in mapping the spatial distribution of taro and sweet potatoes, amongst other crops, on smallholder farms in South Africa. To address this objective, the performances of Gradient Tree Boosting, Random Forest and Support Vector Machine were compared in mapping the spatial distribution of sweet potatoes and taro in a smallholder cropland. Special attention was given to the classification model's ability to capture field boundaries. In addition, the relative contribution of bands, vegetation indices, as well as both datasets combined was assessed.

3.2 Methods and Materials

Figure 3.1 presents a flowchart outlining the key stages of this study, including the image data collection, image pre-processing, the extraction of spectral traits and the subsequent statistical analysis.

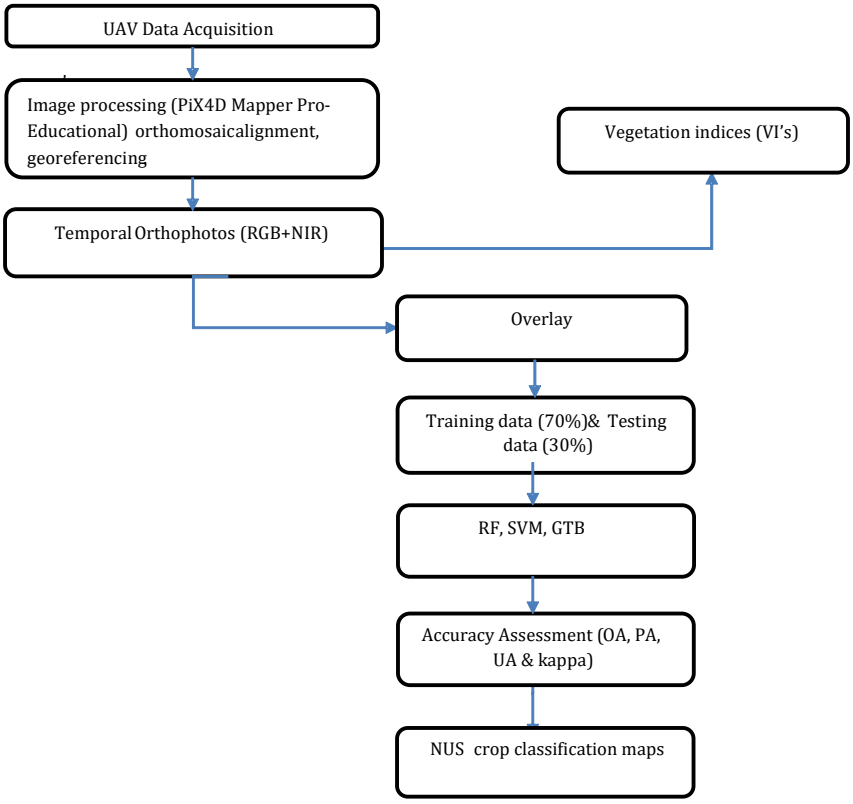


Figure 3-1 Flowchart of main processing steps in this study

Data collection

Before capturing the remotely-sensed data, a field survey was conducted within the study area. A hand-held Trimble Garmin GPS with an accuracy of ±+-1m was utilised to measure the location of different land-use and land-cover classes. Specifically, eight classes were considered in this study. These were taro, sweet potatoes, natural vegetation, sugarcane, bare land, build-up, maize Growth Stage One and maize Growth Stage Two. A total of 310 points were collected in Google Earth Pro. The collected data were converted into a point map that was imported into the Google Earth Engine for model training and validation.

This study utilised a DJI Matrice (M300) quadcopter fitted with a Micasense Altum imaging sensor to capture the images spanning the smallholder fields. The DJI Matrice is an advanced, high-precision quadcopter that is fitted with a high-accuracy Global Navigation Satellite System receiver (GNSS)(KEGA 2021). The DJI has a maximum flight time of 55 minutes and a maximum control range of up to 15 kilometres, which provides a wide operational range. The drone has a 1P45 rating, which protects it against solid particles, such as dust and water (KEGA 2021). The Micasense Altum camera captures data in five spectral bands: red, green, blue, Red-edge (RE) and near-infrared (NIR). It also includes a radiometric thermal sensor that operates within the 8-14 nm wavelength range. The Altum camera incorporates several features to enhance its functionality. Firstly, it is equipped with a solar irradiance sensor called DLS 2, which allows for accurate radiometric calibration by measuring the incident light conditions during image capture (KEGA 2021). The camera also has a built-in GPS, which enables the georeferencing of the captured images (KEGA 2021). The camera's Ground Sample Distance (GSD) is 5.2 cm per pixel, when operated at 120 meters Above Ground Level (AGL) (KEGA 2021). The multispectral bands of the camera have a sensor resolution of 3.2 cm per pixel, while the thermal band has a spatial resolution of 81 cm (KEGA 2021).

A shapefile defining the study area was created by using the Google Earth Pro application to formulate a UAV flight plan. This shapefile was then transferred to the DJI matrice hand-held console device. The formulated flight plan facilitated an autonomous drone mission, which allowed for the seamless aerial coverage of the study field and its surrounding regions. This automated flight mission was conducted at a flight height 120 meters Above Ground Level (AGL), with an image overlap of 80%. Images of the radiometric calibration target (CRP) and a white balance card were captured before and after the flights. The captured imagery was then pre-processed by using Pix4D Fields photogrammetry software.

3.2.1 Data processing

In this stage, 3 576 images were stitched together and adjusted for radiometric accuracy by using Pix4Dfields 1.8.0 software, San Francisco, CA, USA. The radiometric correction process involved utilising the CRP images captured before and after the flights to account for atmospheric conditions and variations in the incent light conditions during image acquisition. This calibration process ensured accurate reflectance values in the image. After processing, an orthomosaic (high-resolution aerial view) and Digital Elevation Model (DEM) GeoTIFF image

were produced. The georeferencing of the orthomosaic was performed in ArcGIS 10.5, by utilising ground reference points that were obtained from Google Earth Pro. This georeferencing stage yielded a Root Mean Square Error (RMSE) that was less than half the size of a pixel. The orthomosaic was aligned with the Universal Transverse Mercator (UTM zone 36S) projection. After that, the drone orthomosaic of the smallholder crop field was imported into the Google Earth Engine (GEE), a cloud-based platform for geospatial data analyses. Within the GEE, the orthomosaic underwent classification by using the Gradient Tree Boost (GTB), Random Forest (RF) and Support Vector Machine (SVM) classifiers. These classifiers employ machine learning algorithms to classify the pixels, based on their spectral properties, which enables the identification and mapping of crops, including sweet potatoes and taro. Three distinct datasets were used for this purpose within the study area.

A total of eleven VIs, which have been previously documented in the literature, were chosen and evaluated to effectively differentiate the NUS species. The VIs were derived by using the reflectance data obtained from the Altum multispectral and thermal bands. These indices typically relied on the red, red-edge and near-infrared (NIR) segments of the electromagnetic (EM) spectrum. Table 3.1 presents a compilation of the VIs that were chosen for this study, considering their direct and indirect correlation with the image classification and crop health attributes, such as the NDVI, the NDRE, OSAVI and NDRGI.

Table 3-1 UAV-derived vegetation indices

Vegetation Index	Abbreviation	Equation	Reference
Normalized difference vegetation index	NDVI	$\frac{NIR - RED}{NIR + RED}$	Xue and Su (2017)
Green normalized difference vegetation index	GNDVI	$\frac{NIR - GREEN}{NIR + GREEN}$	Gitelson *et al.* (1996)
Excess green index	EGI	$2.5\frac{GREEN - RED}{GREEN + RED + 1}$	(Camargo Neto 2004)
Normalized difference red-edge index	NDRE	$\frac{NIR - RED\ EDGE}{NIR + RED\ EDGE}$	(Fitzgerald *et al.* 2006)
Excess green index	EXG	$2 \times GREEN - RED - BLUE$	Woebbecke *et al.* (1995)
Chlorophyll carotenoid index	CCI	$\frac{RED - GREEN}{RED + GREEN}$	(Jäger *et al.* 2022)
Optimized soil adjusted vegetation index	OSAVI	$\frac{NIR - RED}{NIR + RED + 0.16}$	Xue and Su (2017)
Enhanced vegetation index	EVI	$2.5\frac{NIR - RED}{NIR + 6 \times RED - 7.5 \times BLUE + 1}$	(Xing *et al.* 2019)
Normalised green, red difference index	NGRDI	$\frac{GREEN - RED}{GREEN + RED}$	Meyer and Neto (2008)

Simple ratio	SRI	$\frac{NIR}{RED}$	Jordan (1969)
Modified triangular vegetation index 1	MTVI1	$1.2 \times [1.2 \times (NIR - GREEN) - (2.5 \times (RED - GREEN))]$	(Xing *et al.* 2019)

3.2.2 Image classification

In the GEE platform, three classification methods, namely, Gradient tree boost, Support vector machine and Random Forest classifiers, were implemented. These algorithms are renowned for their accuracy as they can effectively select spectral features for discriminating different cover classes (Judson *et al.* 2008). The training dataset utilised for image classification was derived by using a meticulous visual examination of the orthomosaic, which had been acquired on February 12, 2021. Eight major land cover classes were determined, and the training data were acquired by identifying the pixels that corresponded to these classes, as outlined in Table 3.2. The sampled data were randomly split into 70% training (n = 217) and 30% validation (n = 93) for all classifiers. The training dataset was utilised for model development, while the validation data were utilised to assess the classification accuracy. An area computation was performed in the GEE to estimate the number of pixels within each class, thereby determining the total areal extent for each landcover class.

Table 3-2 Training and testing data used for pixel-based image classification

Class name	**Code**	**Training (*n*)**	**Testing (*n*)**
Taro	T	28	12
Sweet potatoes	SP	28	12
Natural Vegetation	NV	21	9
Sugarcane	SU	28	12
Bare land	BL	21	9
Built-up	BU	21	9
Maize growth stage 1	M1	35	15
Maize growth stage 2	M2	35	15

***Random Forest (RF)*:** RF is an ensemble machine learning method that incorporates bootstrap aggregation and binary recursive partitioning to grow multiple decision trees on randomly-selected subsets of the features and data. Groups of decision trees vote on the best class for each sample (Breiman 2001). The algorithm was optimised by using its hyperparameters, namely: the number of features randomly selected at each node (*mtry*) and the number of total trees (*ntree*). Specifically, the *ntree* was set to 300, the maximum depth of each tree (max Nodes), minimum samples required to split nodes (minSamples), and the randomisation seed was set to the default in the GEE. The bag fraction, set at 0.5, facilitates the introduction of randomness during the construction of the tree-building process. According to (Mutanga *et al.* 2023), Random Forest (RF) has several advantageous characteristics:

(i) RF offers easily-adjustable hyperparameters;

(ii) The training process for RF is efficient, resulting in quick model training;

(iii) RF exhibits robustness against outliers, overfitting and high-dimensional data, which makes it suitable for various datasets;

(iv) Due to the averaging of multiple trees, RF demonstrates low bias and a moderate-to-minimal variance;

(v) RF is effective in handling both continuous and categorical variables;

(vi) RF possesses the ability to assess the importance of predictor variables; and

(vii) RF shows a relative resistance to multicollinearity between predictor variables.

These factors make it well-suited for mapping the crop distribution in this study area. This study used all the spectral bands and VIs that were derived from the drone composite as predictor variables.

***Support Vector Machine (SVM)*:** Support Vector Machines (SVMs) were initially developed for binary classification, but they have also proven their efficacy in solving regression problems (Zhao *et al.* 2020). For a multiclass classification, as in this study, SVMs use a one-vs-all or one-vs-one approach with voting, in order to determine the correct class. A Radial Basis Function (RBF) kernel was used, as it performed better than linear or polynomial kernels on this data. The two hyperparameters that were used to optimise SVM-RBF were cost C and gamma γ. A higher C increases the penalty for misclassifications, while a higher γ decreases the variance. SVM is relatively insensitive to noisy inputs, which reduces thee estimation errors and improves its robustness (Singla *et al.* 2020). It performs well on high-dimensional datasets and is robust to outliers, which makes it advantageous for many problems (Hsu *et al.* 2003). Three parameters were tuned for the SVR model, namely, the penalty parameter (C), the precision parameter (ε) and the kernel parameter (γ). Through this process, the SVR model achieved an optimal performance, with a C value of 10, a ε value of 0.5 and the default γ value of 1 was maintained.

***Gradient Tree Boost (GTB)*:** Gradient Boosting Machines (GBMs) are a family of powerful ensemble machine learning techniques that fall under sequential models. Each model in the sequence learns from the mistakes of previous models, in order to incrementally improve the overall performance (Natekin and Knoll 2013). Unlike Random Forests, which average the predictions, GBMs use a gradient descent approach to minimise the error at each stage. GBMs are robust, due to their flexibility in customising the loss functions during model optimisation. This has led to their widespread success in real-world applications, compared to the single models. They rely on three main components - a loss function, weak learner models like decision trees, as well as an additive model that combines the predictions from each weak learner (Natekin and Knoll 2013). Importantly, tree-based GBM algorithms were designed to handle large datasets efficiently. They can run over ten times faster than other popular algorithms on large data, which makes them highly scalable for different scenarios. The GTB classifier has several adjustable parameters that require configuration, including the number of trees, the shrinkage, the sampling rate, the maximum nodes and the seed values. The number

of trees (*nTree*) was adjusted to 300. The maximum nodes, seeds, shrinkage and sampling parameters were set to default in GEE.

3.2.3 Accuracy assessment

An accuracy assessment was conducted on the multispectral UAV image to assess the sensor's capability to effectively discriminate between the NUS crops at the plot level. Confusion matrices were generated to compute the Overall Accuracy (OA), the kappa coefficient, F1 score, as well as the user and producer's accuracies for RF, GTB and SVM classifiers. A confusion matrix is an overly-complex process that compares the classified classes with the reference points and provides a count of the correct and incorrect classifications for each class. It helps to quantify the classification accuracy and identify the patterns of misclassification. The Overall Accuracy (OA) metric was utilised to assess the efficiency of the algorithms employed in this study (Gxokwe 2022). The Producer's Accuracy (PA) was employed to determine the likelihood of correctly classifying the reference data on the map. The User's Accuracy (UA) was employed to assess the probability that a classified pixel accurately correlates to the corresponding category on the ground (Gxokwe 2022). The kappa statistic, which ranges from 0 to 1, measures the agreement between the classified map and the reference data. Values greater than 0.80 indicate a high level of agreement. In contrast, values below 0.40 indicate a weak level of agreement, whereas scores between 0.40 to 0.80 indicate a moderate level of agreement (Dondofema *et al.* 2023). The F1 score harmonises the weighted averages of precision and recall metrics, which makes it a commonly-used score for validating the accuracy of the classification process (Pham *et al.* 2023). As the values of the metrics increase, the model's confidence in accurately assigning the predefined classes in the study also increases. Furthermore, line plots and Jeffries-Matusita (JM) distances were employed in the GEE to assess the spectral separability between the landcover classes. This analysis examined the distances between the class means and the distribution of values derived from those means (Gxokwe 2022). The JM criterion is a parametric measure that ranges from 0 to 2 (Gxokwe 2022). A value approaching 2 signifies a higher degree of spectral distinctiveness between the two classes.

3.3 Results

3.3.1 Spectral reflectance curve

The spectral reflectance curve analysis revealed a distinct separability among the classes, as depicted in Figure 3.2. The responses of most spectral classes exhibited notable distinctions in the NIR and RE regions of the electromagnetic spectrum. The Blue and Red regions of the EM spectrum exhibited limited discriminatory potential in distinguishing between the different landcover classes. Sweet potatoes and taro exhibited a pronounced spectral distinctiveness, compared to the other classes, particularly in the RE and NIR segments of the EM spectrum (717 to 842 nm).

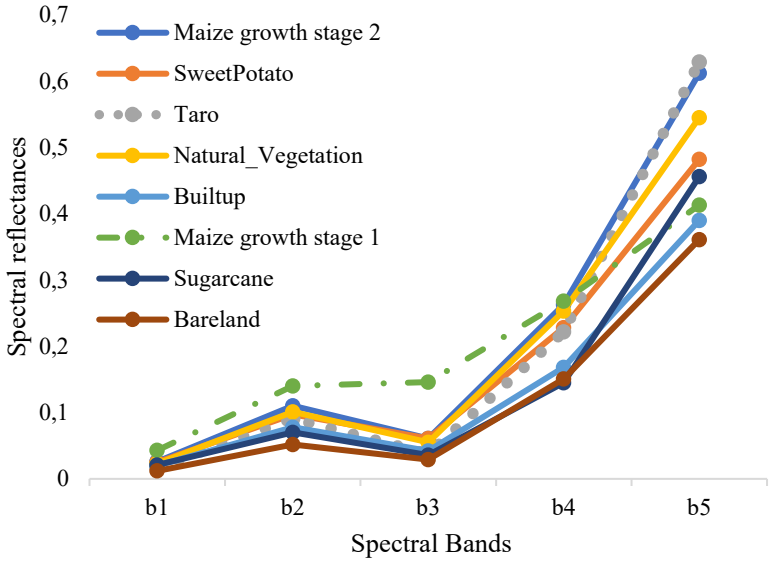


Figure 3-2 Spectral reflectance curve of all classes

3.3.2 Comparative cropland classification using SVM, RF and GTB, based on bands only

The land cover classification results, using spectral bands as an independent dataset, are shown in Table 3.3. The findings revealed that the RF and GTB models exhibited superior accuracies when utilising the raw spectral band dataset. These models achieved exceptional results regarding the OA, kappa and F1 scores, surpassing 80% when using optimal bands, such as B1, B4 and B5. Consequently, these results demonstrate a noteworthy concurrence between the predicted and actual land cover categories, as visually depicted in Table 3.3. Moreover, the RF and GTB models also attained superior producer accuracies for most land cover classes (>75%) . Sweet potatoes and taro exhibited notably higher user accuracies, reaching >56% with the RF and GTB models. These outcomes are further supported by Figure 3.2, as the spectral reflectance of sweet potatoes and taro are more discernible, particularly within the Red-edge and NIR segments of the electromagnetic spectrum. In contrast, the SVM classifier yielded lower accuracies, with an OA of 0.42, a kappa statistic of 0.32 and an F1 score of 0.31. The SVM user and producer accuracies were comparatively lower, with values of 25-60%.

Table 3-3 Overall accuracies, kappa statistics and F1 scores for RF, GTB and SVM

Analysis Stage	Variables	Accuracy unit	SVM	RF	GTB
1	Bands only	OA	0.42	0.86	0.83
		Kappa	0.32	0.84	0.80
		F1	0.31	0.84	0.80
2	VI's only	OA	0.74	0.84	0.83
		Kappa	0.70	0.81	0.80
		F1	0.71	0.82	0.81
3	Bands & VI'S	OA	0.71	0.86	0.88
		Kappa	0.67	0.84	0.85
		F1	0.66	0.80	0.84

The separability between all classes in the single image was investigated by using the JM distances in Table 3.4. Most of the classes exhibited notable disparities among their respective outputs. The findings indicated that the observable bands, such as B1, demonstrated lower JM distances, which suggests a high degree of overlap between the classes. This was particularly evident for taro, sweet potatoes and maize Growth Stage Two, which is often confused with bare land and natural vegetation. In contrast, B4 and B5 exhibited a higher separability (with JM distances ranging from 1.7 to 1.8) for classes, such as taro, sweet potatoes, sugarcane and maize Growth Stage One. This suggests that classification algorithms are likely to effectively distinguish these pairs of classes (Appendix A Tables 5 & 6). The results also showed that combining all auxiliary variables slightly decreased the separability between the sweet potatoes, bare land and the maize growth stages by 1-8%, but they increased the separability between the taro and natural vegetation.

Table 3-4 Band 1 JM distances

B1	TA	SP	NV	SU	BL	BU	M1	M2
TA	0	0.47	0.53	0.46	0.62	1.02	0.81	0.35
SP	0.47	0	0.51	0.44	0.54	1.05	0.8	0.34
NV	0.53	0.51	0	0.71	0.45	0.85	0.64	0.58
SU	0.46	0.44	0.713	0.00	0.72	1.11	0.92	0.24
BL	0.62	0.54	0.45	0.72	0	0.95	0.58	0.61
BU	1.02	1.05	0.85	1.11	0.95	0	0.87	1.06
M1	0.81	0.8	0.64	0.91	0.58	0.87	0	0.84
M2	0.35	0.34	0.58	0.24	0.61	1.06	0.84	0

3.3.3 Comparative cropland classification using SVM, RF and GTB-based vegetation indices only

The utilisation of Vis, in combination with RF, SVM and GTB models, resulted in superior outcomes, as indicated in Table 3.3. This superiority is evident from the variable importance scores, where EGI, EXG and GNDVI were identified as the most optimal performing variables (Appendix A Figure 11). Notably, upon incorporating Vis, the SVM model exhibited a significant increase in the kappa statistic, which rose from 0.32 to 0.74, (Table 3.3). Regarding the performance of RF's, the inclusion of VIs led to a slight decrease of approximately 2% in the kappa statistics and accuracy. The GTB classifier demonstrated a consistently high accuracy, which was equivalent to the dataset of the band, with an Overall Accuracy (OA) of 0.83 and a kappa statistic of 0.80. There was a noticeable improvement in the user accuracy for the RF and GTB algorithms, with values ranging from 66% to 94% for most classes. Similarly, the SVM accuracies demonstrated an average increase of 10% to 26% across most classes. The combination of multiple bands, such as Blue, Red-edge and Near-Infrared, contributed to the improved performance of vegetation indices. However, the highest overall accuracy was achieved, when combining both spectral bands and VIs, which indicated that their joint utilisation maximised the separability for this algorithm.

3.3.4 Comparative performance of spectral variables

The results indicated that the raw spectral bands underperformed, compared to the vegetation indices and the combined dataset across all models, as demonstrated by a mean OA of 70% and a mean kappa statistic of 0.65 (Figure 3.3). This observation is further supported by the variable importance scores, whereby B2 and B4 ranked relatively lower (see Appendix Figure 10). More specifically, VIs outperformed the bands by 10%, but yielded similar accuracies to the combined dataset (Table 3.3). The performance of the optimal bands, such as EGI, NDRE and GNDVI, contributed significantly to the improved performance of machine learning algorithms. Furthermore, the combined dataset exhibited the highest performance across all models, with a mean OA of 0.82 and a kappa statistic of 0.79. This is evident from the close alignment between the OA and the kappa statistic, which indicates a strong agreement between the two measures.

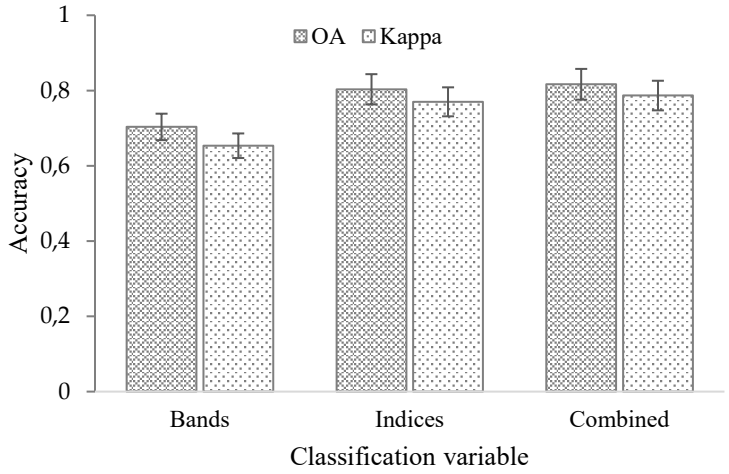


Figure 3-3: Comparative classification performance of bands, vegetation indices and combined data

3.3.5 Comparative classification performance of SVM, RF and GTB

The RF and GTB models achieved the highest overall performance, based on the classification mean OA and kappa statistics (Figure 3.4). RF and GTB produced accuracies >80% and mean kappa statistic values ranging from 0.80 to 0.85, which demonstrated a very good agreement between the predicted and actual land cover classes (Table 3.3). Furthermore, the mean F1 scores ranged between 0.82 and 0.81 for RF and GTB, respectively. In contrast, the SVM classifier performed worse, with a mean OA of 0.62 and kappa statistic of 0.56. Of all the models that were tested, GTB showed the greatest increase (a 4% difference) in performance across the input data, with accuracies sharply rising from 83% for the initial dataset, to 88% for the combined dataset (Table 3.3). In comparison, RF displayed a smaller magnitude change in accuracy across datasets; it decreased slightly from the bands-only to vegetation indices-only, before increasing by 2% with the combined dataset. The findings of the GTB classifier illustrate that it can generate consistent results that are slightly better than those of RF.

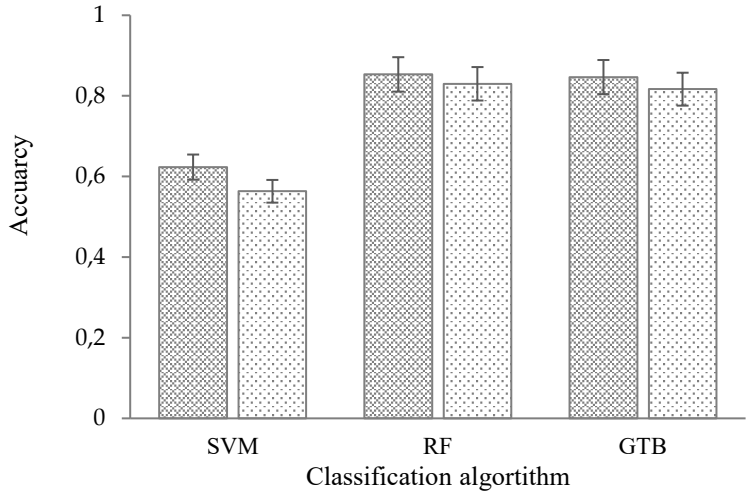


Figure 3-4 Comparative classification performance of Support Vector Machines (SVM), Random Forest (RF) and Gradient Tree Boosting algorithms (GTB)

3.3.6 Final classification of the cropland using combined data

The combination of spectral bands and VIs yielded the highest user and producer accuracies, which ranged from 71.4% to 100% for all classes, and which effectively represented the relationship between the spectral data and land cover (Figure 3.5). The final RF classification, based on the combined data, achieved an OA of 0.86 and a kappa statistic of 0.84 (Table 3.3), utilising B1, NDRE and EGI as the most optimal classification variables (Figure 3.6). Similarly, the GTB model showcased a superior performance, achieving an OA of 0.88 and a kappa statistic of 0.85 (Table 3.3), when utilising the B1, B3 and B5 variables. At the individual class level, RF and GTB most accurately mapped maize Growth Stage One and sugarcane (with a 95-100% accuracy across datasets). Sweet potatoes, natural vegetation and built-up areas were the most frequently misclassified (12.5-50% accuracy), which suggests a weaker spectral representation of these classes. More specifically, maize had the highest classification accuracy when the RF and GTB models and Datasets 2 and 3 were used. SVM exhibited lower accuracies, with an OA of 0.71 and a kappa of 0.67 (Table 3.3). The SVM model experienced more misclassifications overall, and it required extensive hyper-parametrisation to effectively model the complex spectral-class relationships of the real-world environment (Figure 3.5C). Overall, the final classification models suggested that GTB could slightly outperform the RF ensemble.

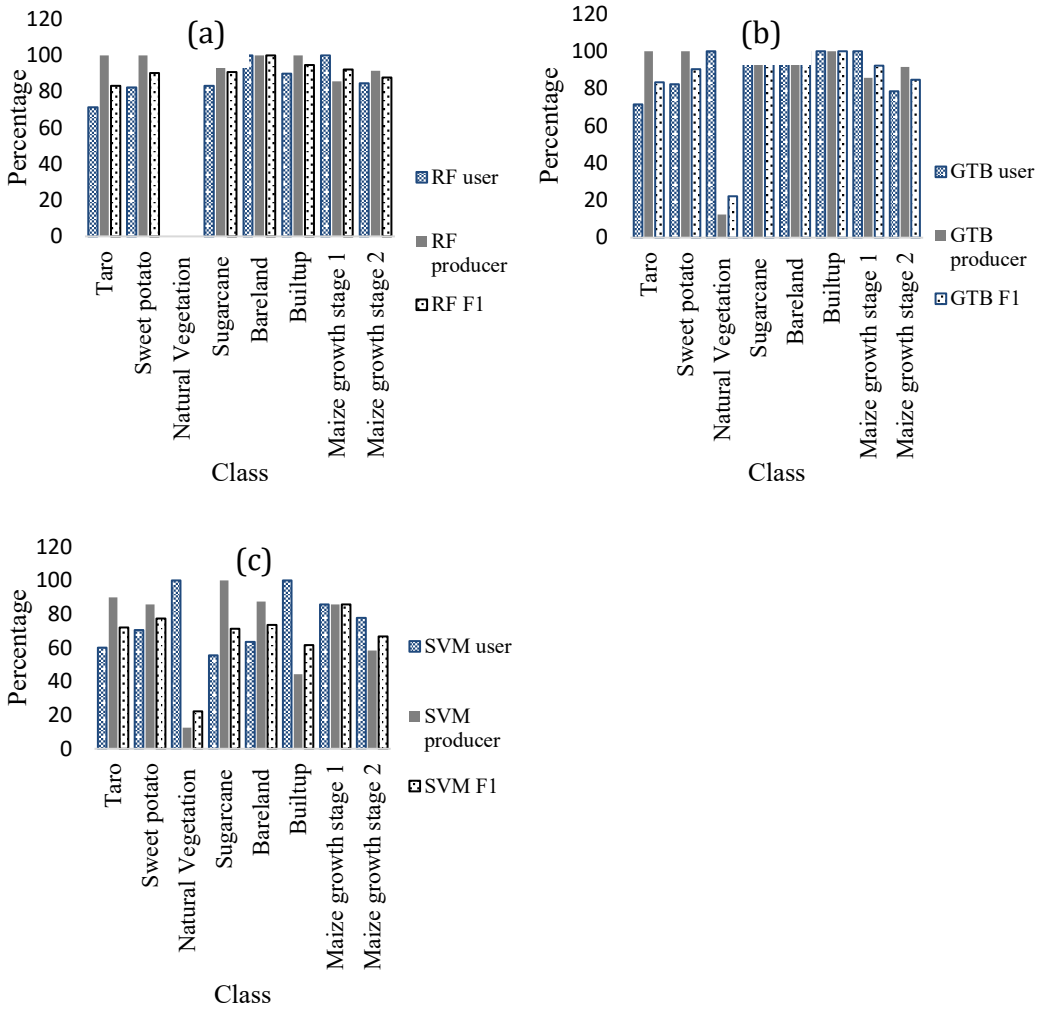


Figure 3-5 User and producer accuracies of (a) RF, (b) GTB and (c) SVM, in conjunction with Dataset 3

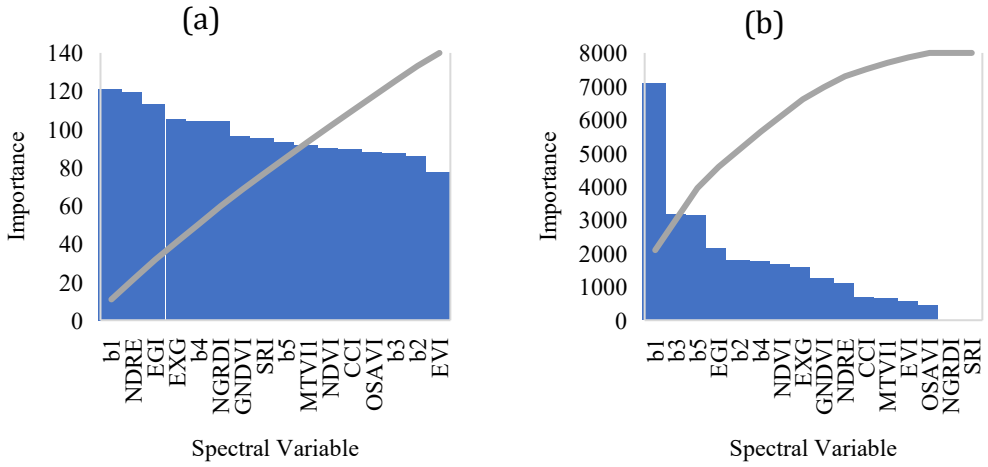


Figure 3-6 Variable importance scores of (a) RF and (b) GTB with Dataset 3

3.3.7 Spatial distribution of land cover types and their areal extents

Figure 3.7 presents the percentage of the total area occupied by each class, as determined by the selected classification approaches. A visual analysis of the classification maps reveals differences in how accurately the algorithms represented the spatial distribution of crops (Figure 3.8). The maps generated by the RF and GTB models exhibited distinct and accurate boundaries between sweet potatoes, maize and taro, effectively capturing the shape of the fields. However, the depiction of crop boundaries was distorted on the map that was classified using SVM. Field shapes were irregular and inaccurately represented. Pixel blocks were misclassified, rather than capturing the fine-scale variability that RF and GTB had achieved. Some crops, like sweet potatoes, appeared to be fragmented into separate small patches, rather than compact field units. The transition zones between different crops appeared less clear, which indicated a lack of precise delineation of the adjacent fields by the SVM model. On the other hand, the RF and GTB models demonstrated a sharper and more distinct demarcation of the fields located next to one another.

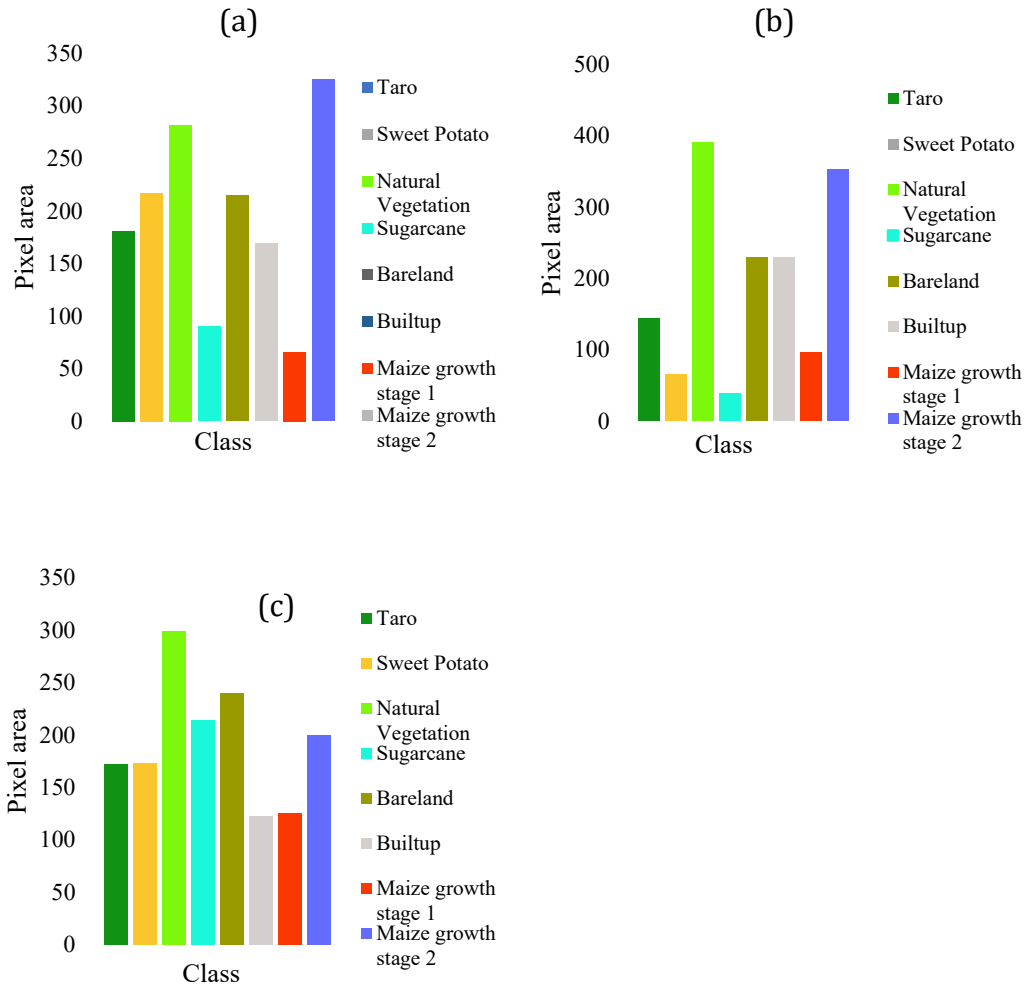


Figure 3-7 Areal extents per class of (a) RF, (b) GTB and (c) SVM with Dataset 3

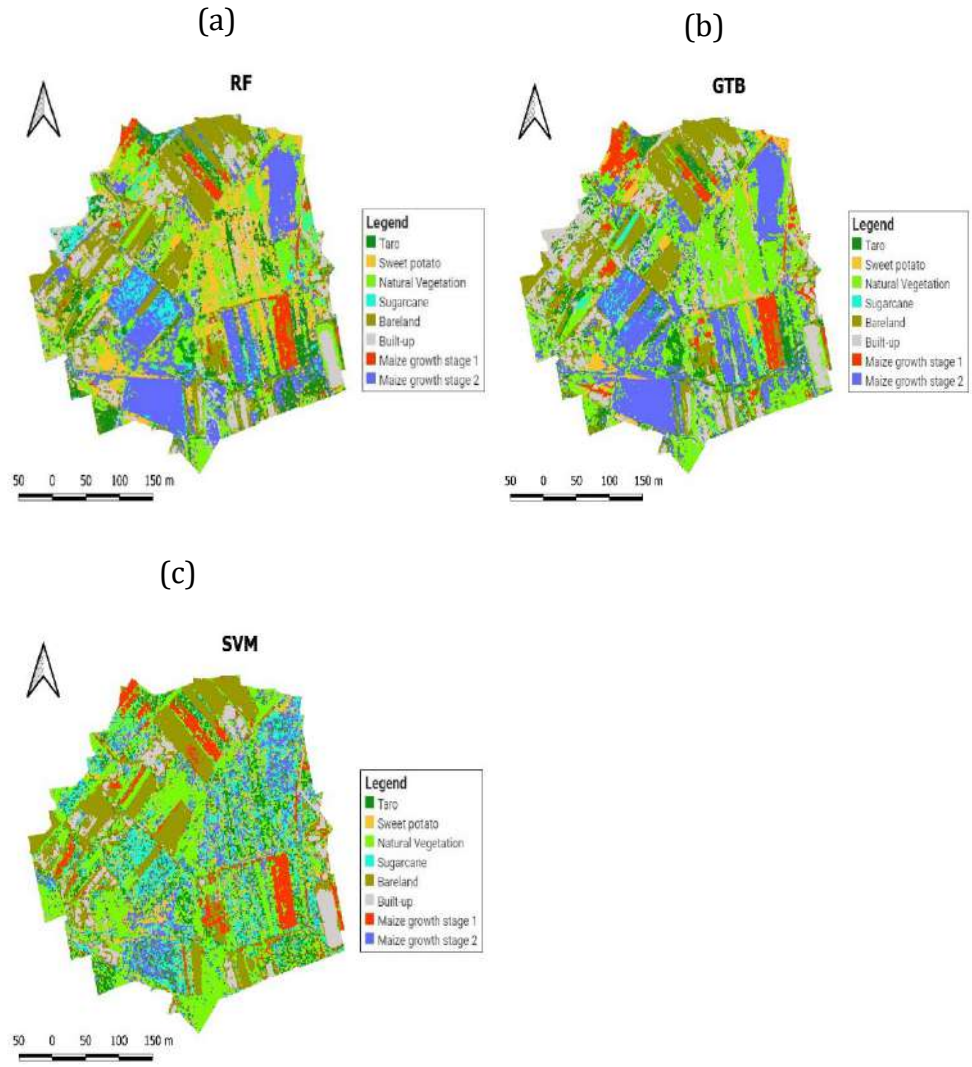


Figure 3-8 NUS crop distribution maps of (a) RF, (b) GTB and (c) SVM

3.4 Discussion

3.4.1 The comparative performance of raw spectral bands and vegetation indices in mapping the spatial distribution of NUS in the smallholder cropland

The results of this study indicated that both the individual bands and vegetation indices can effectively map the spatial distribution of the NUS crop species (Table 3.3). However, there were some notable differences in their predictive performance. In examining the variable importance graphs, it was apparent that certain bands, specifically Band 1 (Blue), Band 4 (Red-edge) and Band 5 (NIR), consistently yielded higher classification accuracies, compared to the other individual raw spectral bands that were depicted (Figure 3.6). This is supported by the spectral separability between the classes, as measured by the JM distances. The bands produced higher JM distances that ranged from 1.5 to 1.7 for some classes, like taro, sweet potatoes and maize Growth Stage Two, which indicated a better spectral separability (Table 3..4). However, the contribution of vegetation indices to the classification algorithms was more effective than that of the bands. This also suggests that the VIs differentiated the crop types effectively, with

similar spectral characteristics within the homogeneous fields. The vegetation indices that exhibited the highest importance scores in both the Random Forest (RF) and Gradient Boosting (GTB) models were the RGB-based indices EGI, EXG and NDGRI. EGI consistently demonstrated the highest predictive capacity among the vegetation indices across all crop types (Figure 3.6).

It has been extensively demonstrated that VIs exhibit robustness, compared to the bands, as expected. Moreover, the utilisation of VIs led to a significant increase in both the PA and UA results, by surpassing the outcomes obtained in the previous stages of analysis (Analysis I). This improvement indicates an enhanced capacity of the dataset variables to distinguish specific classes from others, as depicted in Figure 3.3. The results of this study underscore the significance of integrating VIs in mapping NUS crops within smallholder croplands. The amalgamation of spectral bands and VIs contributed to an increase in the OA, which can be attributed to the VIs' ability to mitigate the influence of the soil background, the sensor zenith angle, the sun angle and other atmospheric impurities (Zeng *et al.* 2022). Moreover, incorporating multiple spectral bands from different sections of the electromagnetic spectrum when computing vegetation indices, enhances their resilience for image classification, compared to the influence of only one band from a single section of the EM spectrum. For example, a study by Niederheiser *et al.* (2021) demonstrated that integrating different spectral and textural indices improved the accuracy of RF models in mapping vegetation cover. Asgari and Hasanlou (2023) also reported similar outcomes in their study, where they utilised Sentinel-2 MSI extracted vegetation indices to delineate various crop types, including rapeseed. The overall accuracy of the RF, GTB and K-Nearest Neighbours (KNN) models increased to 95%, 88% and 90%, respectively, when employing the Atmospherically Resistant Vegetation Index (ARVI).

More specifically, VIs such as EGI and NDRE are a robust proxy for leaf area, green biomass and general photosynthetic activity, as well as all factors that influence crop spectral separability (Meyer and Neto 2008). While other indices, like EVI and NDVI, incorporate additional bands and parameters, they did not substantially outperform NDRE, EXG or GNDVI for this application. This finding suggests that the NUS fields likely exhibited distinct spectral responses in the NIR, RE and Red bands, and applying more complex adjustments provided limited additional predictive insights. Overall, both individual spectral bands and vegetation indices demonstrate their usefulness in mapping neglected crops by using remote sensing techniques.

3.4.2 The comparative performance of machine learning algorithms in mapping the spatial distribution of NUS in the smallholder cropland

When evaluating the classification performance of Support Vector Machine (SVM), Random Forest (RF) and Gradient Tree Boost (GTB) in mapping the spatial distribution of NUS crop species in smallholder croplands, the results showed that RF and GTB achieved the highest classification accuracies, based on all the spectral variables. The optimal performance of RF can be attributed to its utilisation of decision trees and the combination of their predictions, which enable accurate classifications (Sipper and Moore 2021). The ability of the algorithm to

capture complex interactions and handle high-dimensional data contributed to its superior performance (Hornung and Boulesteix 2022).

Furthermore, the GTB algorithm, in conjunction with the combined data, exhibited superior capabilities in detecting and mapping the field boundaries of sweet potatoes and taro, amongst others. More specifically, this combination produced the highest accuracy of 88%, a kappa statistic of 0.85 and an F1 score of 0.84, compared to the other models/datasets (Table 3.3). Moreover, the GTB classifier significantly improved the accuracy from Dataset 1 to Dataset 3, which outperformed the RF and SVM algorithms. GTB can identify subtle variations in the spectral responses that may not be discernible by using linear classifiers, such as SVM (Ghimire *et al.* 2010). In a seminal study that explored land cover mapping with remote sensing data, Abdi (2020) found that the GTB algorithm outperformed the RF algorithm by achieving more accurate results for mapping and detecting changes in the land use and land cover types within a boreal (northern forest) landscape. GTB achieved higher overall and user accuracies, which demonstrated a better predictive performance.

While Ensemble classifiers generally outperformed SVM in their overall performance, there was no substantial difference in the performance of the ensemble methods. The mapped distribution of land cover classes revealed comparable spatial coverage for taro, bare land and maize at different growth stages, as classified by GTB and RF. RF and GTB can capture complex patterns, while controlling overfitting (Thenkabail and Lyon 2016). Therefore, RF and GTB were better able to learn the intricate spectral signatures of different classes, hence accurately discriminating the NUS crops at a plot level from the remote sensing data. In contrast, the SVM results demonstrated divergent representations, with an exacerbated misclassification of some classes. This increased misassignment was prominently visible in the north-eastern and south-western regions of the map, which highlighted a disproportionate error in those areas, relative to the other algorithms (Figure 3.8). SVM is relatively prone to overfitting, and it is sensitive to noisy inputs, which further exacerbates its weakness in crop mapping on smallholder farms (Singla *et al.* 2020). Furthermore, according to Mountrakis *et al.* (2011), SVM models rely on finding a maximal margin hyperplane, which can be heavily influenced by outliers and noisy samples. The comparable performance of GTB and RF suggests that ensemble decision tree methods excel in capturing and delineating the intricate relationships between predictor variables and crop classes (Aguilar *et al.* 2018). By combining many individual decision trees, they can model nonlinear interactions that aid in the differentiation of spectrally-similar crops, such as NUS (Saarela and Jauhiainen 2021). Similarly, Yulianto *et al.* (2023) compared the performance of various algorithms for predicting and mapping land degradation by using remote sensing techniques. Their results demonstrated the superiority and effectiveness of GTB and RF, both of which achieved accuracies that surpassed 85%, in contrast to the less effective performance of SVM. The superior performance metrics imply that these algorithms have the potential to create more reliable land cover maps with limited misclassifications.

Despite its ability to handle high-dimensional data and non-linear relationships, SVM might need help to capture the complex patterns and interactions that are present in the spatial

distribution of NUS. This indicates that the crop classifications may not have been perfectly separable by simple hyperplanes in the feature space (Gove and Faytong 2012). SVM accuracies were slightly lower than those of RF and GTB, with a 42% to 72% OA across all datasets (Table 3.3). The SVM classifier produced the lowest accuracies and kappa coefficients for all datasets, which confirms its strong assumption that independence between the features is inappropriate for this remote sensing task. Furthermore, classes such as taro, natural vegetation and maize at Growth Stage Two were frequently misclassified, hence the lower user accuracies <80%.

3.5 Conclusion

The objective of this study was to comparatively assess the classification performance of the Gradient Tree Boosting, Random Forest and Support Vector Machine in mapping the field boundaries and the spatial distribution of sweet potatoes and taro in smallholder cropland, by using remotely sensed multispectral data acquired via UAVs. The relative contribution of the bands and vegetation indices were also assessed in addressing this objective. GTB and RF could effectively discriminate the field boundaries of NUS, and they could effectively map their spatial distribution in smallholder croplands, based on the bands and indices derived from the Red-edge and the NIR, amongst others. More specifically, the GTB algorithm, in conjunction with the combined data, exhibited superior capabilities in detecting and mapping the field boundaries of sweet potatoes and taro, amongst other crops. Overall, the contribution of bands in classifying smallholder croplands with NUS, among other crops, was found to be less significant than the contribution of vegetation indices.

This study adds to the effectiveness of advanced classification algorithms for mapping NUS in smallholder farming systems. The thematic maps depicting the distribution of NUS crops offer valuable insights for farm-scale management. The precise identification and comprehension of the distribution patterns of NUS crops can play a pivotal role in enabling precision agriculture interventions for their conservation, cultivation and utilisation. This, in turn, could contribute significantly to enhancing the food and nutrition security in vulnerable environments across sub-Saharan Africa. At the same time, national departments can benefit from implementing robust classifiers, in order to develop accurate and dependable agricultural land cover maps.

4 USING UAV MULTISPECTRAL-THERMAL DATA AND INDEX-BASED SEGMENTATION TO ENHANCE EQUIVALENT WATER THICKNESS ESTIMATION IN TARO CROPS

4.1 Introduction

The world is being challenged by the pressing need to sustain the food supply and ensure food security, due to climate change and the increasing global population (Mugiyo *et al.* 2021, Din *et al.* 2022). In southern Africa, climate variability, among other factors, has led to water scarcity, which restricts crop production, and significantly reduces the available arable land (Mugiyo *et al.* 2021, Zulu 2022, Kapari *et al.* 2023). This has particularly affected smallholder farmers, who contribute up to 80% of the food production in sub-Saharan Africa and who play a crucial role in addressing the issue of local food security (Hlophe-Ginindza and Mpandeli 2020). Despite their crucial role, smallholder farms commonly encounter sporadic water stress and droughts, which lead to substantial losses in their crop yields. Addressing this challenge necessitates a shift toward crop diversification and incorporating crop species that demonstrate resilience to abiotic stresses (Chivenge *et al.* 2015, Hilary van Wyk and Oscar Amonsou 2021, Mugiyo *et al.* 2021).

Driven by water scarcity, there has recently been an increasing interest in the potential use of NUS for addressing the food and nutrition challenges (Chivenge *et al.* 2015, Mabhaudhi *et al.* 2017, Mugiyo *et al.* 2021). NUS, which is characterised by historical domestication with limited scientific research and which is predominantly confined to smallholder farming systems, have emerged as key drought-tolerant crops for diversifying the communal cropping systems (Chivenge *et al.* 2015, Mabhaudhi *et al.* 2017). Taro (*Colocasia esculenta* (L)) is one of the oldest and most widely-cultivated NUS in the world's tropical and subtropical regions (Mabhaudhi *et al.* 2011, Mawoyo *et al.* 2017, Van Wyk 2021). In South Africa, taro, which is locally known as *amadumbe,* is known to be heat-tolerant and it is primarily cultivated for subsistence, especially within small and marginalised communities (Mabhaudhi *et al.* 2014, Joshi *et al.* 2020, Oyeyinka and Amonsou 2020, Van Wyk 2021). Taro is identified as a future smart food under the NUS category because of its edible tubers, which are rich in carbohydrates, protein and vitamins (Li and Siddique 2018, Kapoor *et al.* 2022). Despite the value of taro and other NUS, the literature shows that they have largely been ignored.

Whereas advocacy for taro has gained traction as a valuable drought and heat-tolerant crop in diversifying the smallholder farming systems, the evidence is predominantly anecdotal (Joshi *et al.* 2020). Notably, there is a significant research gap in understanding the water status of the crop; hence, knowledge on its ability to adapt to diverse agroecological environments remains limited (Mabhaudhi *et al.* 2017). Consequently, this presents challenges for decision-ma*kers who aim to streamline taro into the existing agricultural systems (Mugiyo *et al.* 2021). Therefore, it is imperative that further research on the canopy water status of taro be conducted, as a proxy of water use, to optimise its productivity and ultimately facilitate its integration into sustainable agricultural practices.

Equivalent Water Thickness (EWT), which is the ratio between the crop leaf area and the quantity of water per unit area, is a critical metric for assessing and understanding the canopy water status of a crop (Traore *et al.* 2021). The indicator quantifies the amount of water that is present in the plant canopy, hence providing a valuable measure of a crop's hydration level (Traore *et al.* 2021). Furthermore, EWT is critical for understanding bio-geochemical processes, such as photosynthesis, primary productivity and the overall crop health (Zhang and Zhou 2019). Traditionally, EWT is measured by using in-situ measurements that are laborious, time-consuming and often require trained experts in field sampling (Chivasa *et al.* 2020, Traore *et al.* 2021). However, over the past few decades, the use of satellite-borne earth observation technologies has provided a valuable alternative to quantifying the agronomic properties and understanding the changes in a crop's water status (Pasqualotto *et al.* 2018). For example, Yilmaz *et al.* (2008) utilised a time series of Landsat 5 Thematic Mapper (TM), Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) and Advanced Wide Field Sensor (AWiFS) imagery to estimate the water status of maize and soybean at optimal accuracies ($R^2 = 0.87$ and $R^2 = 0.47$, respectively). Then, Neinavaz *et al.* (2017) successfully estimated the canopy EWT by using hyperspectral thermal infrared data to an R^2 of 0.81 and a RMSE of 0.003 gcm^{-2}. The success of these studies could be explained by the presence of the thermal band, which facilitates the accurate estimation of the canopy water content (Khanal *et al.* 2017).

Thermal infrared remote sensing has emerged as a valuable tool for crop water assessment and monitoring, and it offers a direct correlation with the biophysical and biochemical elements of crop water (Khanal *et al.* 2017, Messina and Modica 2020). More specifically, the emitted radiance in the thermal infrared region exhibits a sensitivity to variations in canopy temperature, which is influenced by changes in crop water status (Gerhards *et al.* 2019). In a state of water deficit, the rate of leaf transpiration decreases, which leads to a linear reduction in the leaf and canopy temperature, inducing a cooling effect (Maes and Steppe 2012, Hou *et al.* 2018). Therefore, thermal imagery can be used to identify areas of water stress, which is valuable for the implementation of appropriate management strategies to improve the crop water status and overall productivity. The adoption of thermal remote sensing for estimating the crop water status has gained popularity over the decades (Krishna *et al.* 2021). Various satellite-based thermal sensors, including the Moderate Resolution Imaging Spectroradiometer (MODIS), ASTER and the Landsat series (Malbéteau *et al.* 2018, Gerhards *et al.* 2019, Masina *et al.* 2020, Xue *et al.* 2020), have been used to assess and monitor the crop water status. However, these datasets are constrained by their relatively-coarse spatial resolutions, which limits their suitability for monitoring the taro water status on smallholder farms (Malbéteau *et al.* 2018).

The recent advancements in image acquisition, like Unmanned Aerial Vehicles (UAVs) mounted with light-weight multispectral sensors, have provided spatially-explicit near-real-time information on the water status of crops (Hussain *et al.* 2020). UAV proximal sensors with a sub-centimetre resolution deliver rapid, cost-effective and accurate measurements that are suitable for monitoring the water stress and health attributes of crops at a plot level (Chivasa *et al.* 2020). Specifically, UAVs that have been integrated with very-high resolution Thermal

Infrared Sensors (TIRS), provide an optimal approach for retrieving crop attributes for the accurate quantification of crop water status at a field scale (Berni *et al.* 2009).

In addition to the ultra-high spatial resolutions of UAV multispectral thermal imagery, image enhancement techniques and robust algorithms have been demonstrated to improve a model's accuracy. For instance, Index-Based Image Segmentation has been demonstrated to be effective in robustly segmenting plants into colour images, which enables the extraction of vegetation cover and removes the soil background for enhanced crop spectral signatures (Hamuda *et al.* 2016). Meanwhile, colour indices have the potential to enhance the visual contrast of plants against the soil background and to reduce the effects of the light conditions, thereby improving the plant segmentation accuracy, in comparison to original red-green-blue (RGB) imagery (Lu *et al.* 2022). The Excess Green (ExG) and Excess Red (ExR) indices were proposed by Woebbecke *et al.* (1995) and Meyer *et al.* (1999), respectively, to enhance the plant segmentation accuracy by emphasising the greenness of the plant by accounting for the relative proportions of red and physiological green. In addition, Meyer and Neto (2008) leveraged the strength of both the ExG and ExR to develop the Excess Green minus Excess Red (ExGR) index to improve crop water assessment and monitoring, by using thermal remote sensing systems. While colour index-based segmentations have been proven to be effective in improving estimation accuracies, they remain relatively underexplored, which prompts the need for further inquiry, in order to fully leverage their potential in enhancing crop water assessment and monitoring through UAV multispectral thermal remote sensing systems.

In addition to traditional techniques, the use of machine learning methods, particularly deep learning, has recently shown great potential for assessing the crop water status (Zhou *et al.* 2021). Deep learning techniques, such as Deep Neural Networks (DNN), have shown robustness and high accuracy in identifying and quantifying water deficits in various crops (e.g. sugarcane and maize) using thermal remote sensing (de Melo *et al.* 2022, Rajwade *et al.* 2023). Unlike traditional machine learning algorithms, the DNN can exploit feature representations that have been exclusively learnt from data, in order to understand the complicated nonlinear relationships between different environmental attributes (Yuan *et al.* 2020, Odebiri *et al.* 2022). Despite these advantages, the integration of the DNN algorithm with UAV thermal remote sensing for crop water estimations, especially in smallholder farming systems, has been infrequent (Abrahams *et al.* 2023). Consequently, there is a dire need for further research to explore and incorporate the potential of DNN algorithms in precision agriculture, for more accurate and efficient crop water status assessments.

In this regard, by leveraging the capabilities of deep learning, this study sought to assess the performance of thermal remote sensing and index-based segmentation techniques for improving the canopy EWT estimation of smallholder taro crops, using UAV multispectral thermal imagery. Specifically, the study sought: (1) to assess the potential of the UAV thermal band in estimating the EWT of smallholder taro; (2) to compare the performance of crop canopy images that were extracted using the ExG, ExR, and ExGR colour indices to improve the EWT estimations of the taro crop; and (3) to evaluate the potential of UAV multispectral thermal imagery in the EWT estimations of taro crop in smallholder farming systems.

4.2 Materials and Methods

The study area is detailed in Chapter One.

4.2.1 Field sampling and in-situ measurements

A polygon delineating the taro field was created by using Google Earth Pro and it was imported into ArcMap 10.6 to facilitate the generation of 100 stratified random sampling points. This approach was adopted to ensure the variability and accurate representation of all taro crops within the field. These sampling points were subsequently uploaded into a Trimble hand-held Global Positioning System (GPS) with a sub-centimetre accuracy, which enabled the precise location of each sampling point within the taro field. In-situ measurements were obtained at each sampling point, in order to compute the respective EWT values.

A portable LiCOR-2200C Plant Canopy Analyser was used to obtain the Leaf Area Index (LAI) of the crops. The LAI measurements were obtained by using the 38° zenith angle with a 270° view cap and the ABBBB sequence, where A corresponds to a reference reading 'above' the canopy, and B corresponds to a reading 'below' the canopy. Thereafter, the above-ground biomass of each sampled crop was obtained, and the Fresh Weight (FW) was obtained by using a calibrated scale with a 0.5 g measurement error. The sampled biomass was then placed in a labelled brown paper bag and dried in an oven at 60°C, until a constant Dry Weight (DW) was reached (approximately 72 hrs). The LAI, FW and DW were then used as input variables to compute the taro EWT_{canopy} indicators, using the following equation:

$$EWTcanopy = \frac{\mathrm{FW-DW}}{\mathrm{LAI}} \qquad \text{Units: g/m}^2 \qquad (1)$$

Lastly, in-situ measurements were conducted on a sunny day between 10:00 am and 2:00 pm local time, as it aligns with the optimal period for crop photosynthesis, which ensures that the collected data accurately reflects the maximum reflectance and photosynthetic activity of the taro crops.

4.2.2 UAV platform and multispectral-thermal camera

The DJI Matrice 300 (M300) platform, mounted with a MicaSense Altum camera and Downwelling Light Sensor 2 (DLS 2), was used to collect multispectral-thermal imagery. The M300 platform is equipped with four rotary wings and incorporates Vertical Take-Off and Landing (VTOL) technology; hence, it is highly suitable for imaging smallholder farms, especially those near settlements. The MicaSense Altum camera is a multispectral and thermal-imaging sensor that incorporates five high-resolution narrow bands that measure reflectance in the blue (475 nm), green (560 nm), red (668 nm), red-edge (717 nm) and NIR (840 nm) regions, along with a radiometric longwave infrared thermal channel (11 μm). This advanced camera provides the synchronised capture of multispectral and thermal images, by utilising a global shutter that facilitates a 1-second capture rate, and it ensures clear and precisely-aligned imagery (Hutton *et al.* 2020). The multispectral channels have a sensor resolution of 2 064 × 1 544 pixels (equivalent to 3.2 megapixels per multispectral band), while the thermal band has a sensor resolution of 160 × 120 pixels, at a 120 m platform altitude. Lastly, the calibrated

reflectance panel and equipped DLS 2 allow the Altum sensor to measure the ambient light and sun angle, which improves the quality of the captured radiation intensity.

4.2.3 Image acquisition and pre-processing

A shapefile outlining the study area was generated in Google Earth Pro and exported to the M300's hand-held console for the development of a UAV flight plan. The Altum sensor was calibrated pre- and post-flight by using a Calibrated Reflectance Panel (CRP) to compensate for incidental light conditions. This involved the use of known reflectance values across the spectrum to radiometrically calibrate the sensor, by taking an unshaded image of the CRP to account for the illumination and atmospheric conditions prevalent during the flight. The automated flight mission was executed at a 100 m flight height, with an 80% forward and sideward image overlap, thereby obtaining a 10.08 cm per pixel spatial resolution imagery. The flight was conducted under clear sky conditions between 10:00 am and 2:00 pm local time, when the solar zenith angle is minimal, and solar radiation reaches its peak. The UAV flight and in-situ measurements of taro water status were carried out concurrently.

A total of 1 626 raw images of the experimental field were obtained and pre-processed in Pix4D photogrammetry software. The software allows for the stitching of the imagery to create a composite ortho-mosaic of the taro field, while accounting for the atmospheric conditions during image acquisition. The CRP images, together with the raw images, are automatically recognised by the software and used to perform radiometric corrections by calibrating the image reflectance to align with the atmospheric conditions during the time of image acquisition. The ground reference points surveyed before the fieldwork were then used to improve the geometric accuracy of the acquired images in ArcGIS 10.6. Lastly, the EWT_{canopy} in-situ measurements and the location of each sampled taro point were overlaid with a UAV multispectral-thermal image. The multispectral and thermal reflectance data of taro were extracted from the UAV imagery and used to derive Vegetation Indices (VIs) for the development of the EWT_{canopy} prediction model. A list of VIs used in this study is summarised in Table 4.1. These VIs were selected based on their optimal performance in the literature and their relationship with crop water status (Baluja *et al.* 2012, Zhang and Zhou 2019, Ozelkan 2020).

Table 4-1 UAV derived vegetation indices used in this study

Vegetation Index	Abbreviation and Equation	Reference
Normalised difference vegetation index	$NDVI = \frac{R_{Nir} - R_{red}}{R_{Nir} + R_{red}}$	(Rouse *et al.* 1974)
Soil-Adjusted Vegetation Index	$SAVI = \frac{R_{Nir} - R_{red}}{R_{Nir} + R_{red} + L} \times (1 + L)$	(Huete 1988)
Normalised Difference Red Edge Index	$NDRE = \frac{R_{Nir} - R_{Rededge}}{R_{Nir} + R_{Rededge}}$	(Gitelson and Merzlyak 1994)

Transformed Chlorophyll Absorption in Reflectance Index	$TCARI = 3 \times (R_{Rededge} - R_{Red}) - 0.2 \times (R_{Rededge} - R_{Green}) \times \frac{R_{Redegde}}{R_{Red}}$	(Haboudane *et al.* 2002)
Modified Chlorophyll Absorption in Reflectance Index	$MCARI = (R_{Nir} - R_{Rededge}) - 0.2\,x\,(R_{Nir} - R_{Green}) \times \frac{R_{Nir}}{R_{Rededge}}$	(Daughtry *et al.* 2000)
Normalised Green-Red Difference Index	$NGRDI = \frac{R_{Green} - R_{red}}{R_{Green} + R_{red}}$	(Hunt Jr *et al.* 2013)
Normalised difference water index	$NDWI = \frac{R_{Green} - R_{Nir}}{R_{Green} + R_{Nir}}$	(Gao 1996)
Red edge chlorophyll index	$CIrededge = \frac{R_{Nir}}{R_{rededge}} - 1$	(Gitelson *et al.* 2003, Gitelson *et al.* 2006)

4.2.4 Index-based image segmentation of the taro crops' spectral signatures

Image segmentation, to remove the soil background, is crucial for accurately assessing the crop water status, especially in set-ups like smallholder taro fields, where the soil background effect is pronounced, due to the low planting density and high interrow spacing. The expansive growth of taro tubers necessitates substantial space beneath the soil surface, to access sufficient moisture and nutrients and to reduce competition, for optimal growth and development (Tumuhimbise 2015). An index-based segmentation technique was employed to delineate the crop canopy and eliminate the soil background from the multispectral thermal image. Specifically, the Excess Green (ExG), Excess Red (ExR) and Excess Green minus Excess Red (ExGR) colour indices were computed by using the green, red, and blue bands of the UAV multispectral thermal imagery (Woebbecke *et al.* 1995, Meyer *et al.* 1999, Meyer and Neto 2008, Hamuda *et al.* 2016). The ExG, ExR and ExGR were calculated by using the following equations:

$$ExG = 2 \times R_{Green} - R_{Red} - R_{Blue} \quad (2)$$

$$ExR = 1.4 \times R_{Red} - R_{Green} \quad (3)$$

$$ExGR = ExG - ExR \quad (4)$$

Finally, the threshold method was used to generate a binary image from the grey-level histograms that were obtained during the index-based segmentation process (Shu *et al.* 2021). The ExG and ExR binary images of the taro crop canopy cover were generated by using an automatic Otsu threshold, which was determined by the maximum interclass variance, as defined by the Otsu method (Otsu 1979). The Otsu threshold was executed in Mathworks MatLab using the image toolbox gray-thresh function to develop a binary ExG image. Unlike the ExG index, the ExGR index does not require the calculation of a specific threshold, as all plant pixel values are positive, while the background pixels are negative (Meyer and Neto 2008, Riehle *et al.* 2020). Consequently, a fixed zero threshold was implemented to automatically generate a binary image with a consistent threshold of zero (Hamuda *et al.* 2016). Thereafter,

the image pixels were reclassified to create a new layer, where vegetation and soil pixels were assigned the value 1 and 0, respectively. The vegetation mask was then converted into a shapefile and applied to the UAV image to effectively remove the soil background, as shown in Figure 4.1.

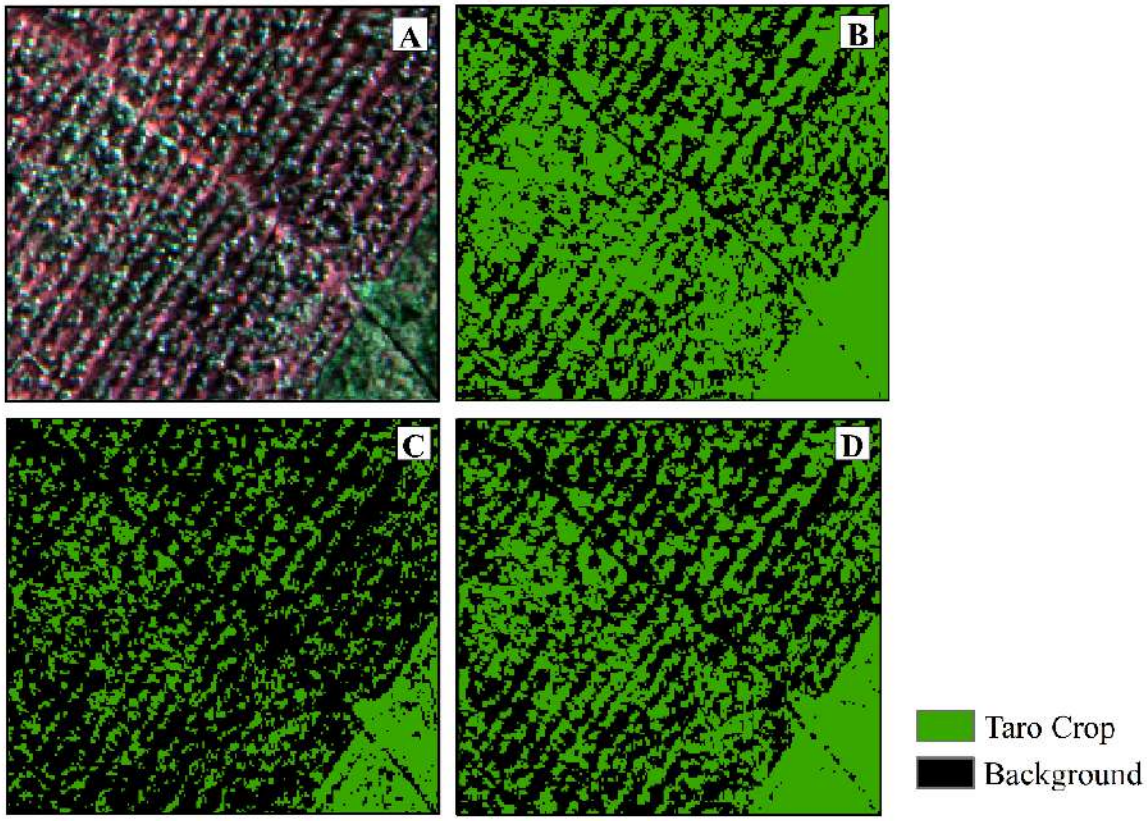


Figure 4-1 Taro crop canopy extraction: a) raw UAV image, b) ExG, c) ExR and d) ExGR canopy extraction

4.2.5 Model development and statistical analysis

In this study, we employed a deep machine learning approach to estimate the EWT_{canopy}, by using UAV-derived multispectral optical and thermal datasets. Specifically, we utilised a Deep Neural Network (DNN), an artificial neural network that is characterised by multiple interconnected hidden layers (Traore *et al.* 2021). These hidden layers work in tandem to progressively learn higher-level features, which facilitates the transformation of input data into meaningful output information (Chew *et al.* 2020, Bouguettaya *et al.* 2022). In contrast to the traditional machine learning techniques that rely heavily on a prior knowledge of the parameters, DNN leverages feature representations that are learned exclusively from the input dataset (Odebiri *et al.* 2021). This unique approach allows the model to understand and capture the complex non-linear relationships among variables, which enhances its ability to uncover intricate patterns and dependencies within the data (Odebiri *et al.* 2021, Traore *et al.* 2021). Therefore, the combination of very-high resolution UAV multispectral-thermal imagery and sophisticated DNN architecture facilitated the precise and accurate estimation of the EWT_{canopy} over the smallholder taro farmland.

The study utilised a three-layer neural network model consisting of an input layer, a hidden layer and an output layer (Figure 4.2). A Rectified Linear Unit (ReLU) was applied to stimulate the EWT_{canopy} prediction model, with the maximum epochs set to 200 iterations, which indicated that the weights in the hidden layers were iteratively adjusted 200 times to reduce error and enhance the EWT_{canopy} prediction accuracy. Thereafter, the SoftMax activation function was used to transform the raw outputs of the neural network into a vector of probabilities, and the Adaptive Moment estimation (Adam) optimiser was used to optimise the results of the output model. Furthermore, the drop-out regularisation technique was applied to avoid overfitting and improve the generalisation of the model (Deepan and Sudha 2020). The hyperparameters of the DNN model were tuned to a learning rate of 0.001, a batch size of 32 and an input and hidden layer dropout of 0.4 and 0.2, respectively.

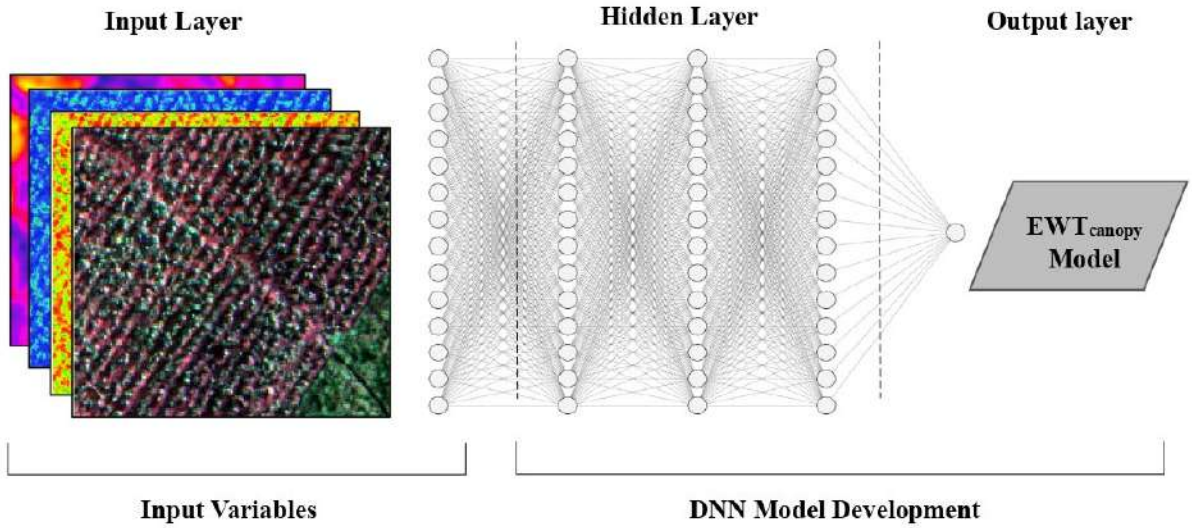


Figure 4.4-2 Graphical representation of the DNN structure, composed of input, hidden and output layers

4.2.6 Accuracy assessment

To evaluate the performance EWT_{canopy} estimation model, the acquired dataset (n = 100) was divided into 70% training data (n = 70) and 30% testing data (n = 30). The training data were employed for the DNN model development, and the testing data were used to assess the accuracy of the predictive model. The evaluation metrics included the coefficient of determination (R^2), the Root Mean Square Error (RMSE) and the Relative Root Mean Square Error (rRMSE). The R^2 measures the variance between the in-situ measured and the estimated taro EWT_{canopy}, while the RMSE assesses the magnitude of error between the field measurements and the modelled EWT_{canopy} field. Lastly, the rRMSE was used to compare the DNN model accuracies across the different image analysis techniques. The statistical indices are presented as follows:

$$R^2 = \frac{\left[\sum_{i=1}^{n}(y_i - \bar{y}_i)\left(\hat{y}_i - \overline{\hat{y}_i}\right)\right]^2}{\left[\sum_{i=1}^{n}(y_i - \bar{y}_i)^2 \sum_{i=1}^{n}\left(\hat{y}_i - \overline{\hat{y}_i}\right)^2\right]} \quad (5)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \quad (6)$$

$$rRMSE = \frac{RMSE}{\bar{y}_i} \times 100 \quad (7)$$

where $\hat{y}_i$ and y_i are the predicted and actual values, $\bar{y}_i$ and $\overline{\hat{y}_i}$ are the means of the observed and predicted values, respectively, and *n* is the total number of data points. The higher the value of R^2, and the lower the values of RMSE and rRMSE, the greater the precision and accuracy of the EWT_{canopy} estimation model. The accuracy metrics were subjected to a statistical analysis by using a two-way Analysis of Variance (ANOVA) in IBM SPSS software Version 29.0.2, to assess the significant differences between the means of the prediction accuracies.

Lastly, the SHapley Additive exPlanations (SHAP) approach was used to assess the impact and contribution of predictor variables within the developed models. SHAP ranks the importance of the model variables by calculating the average marginal contribution value of each predictor variable (Nahiduzzaman *et al.* 2023). By leveraging the SHAP approach, a comprehensive understanding of the relative importance of each predictor variable is gained, which enhances the interpretability of the developed EWT_{canopy} estimation models. An overall workflow of this study is presented in Figure 4.3.

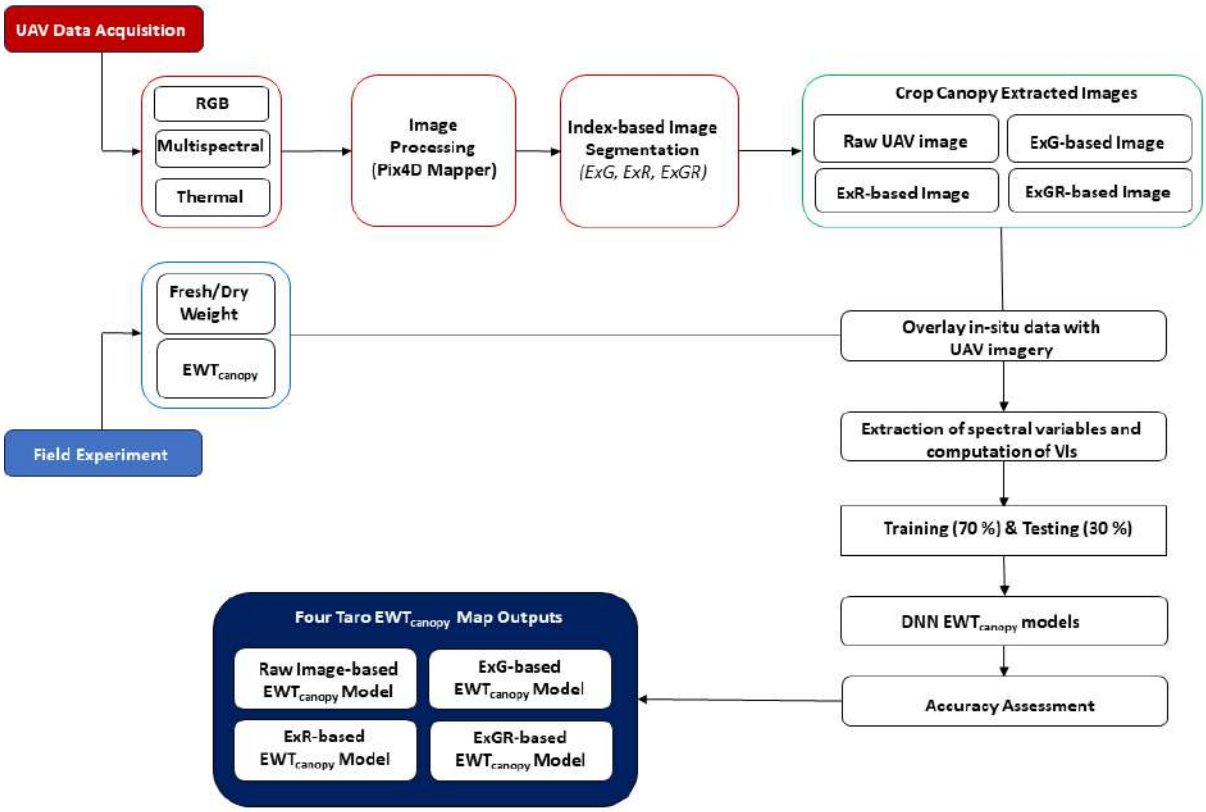


Figure 4-3 Workflow diagram for this study

4.3 Results

4.3.1 Descriptive statistics of the in-situ EWT-canopy of taro crops

Table 4.2 presents the descriptive statistics of the in-situ measurements of EWT_{canopy} across the taro field. A total of 100 samples (N) were collected during the peak vegetative growth stage. The mean and median of EWT_{canopy} were 60.90 g/m^2 and 47.70 g/m^2, respectively, with a Standard Error of Mean (SEM) of 4.45 g/m^2. A high spatial variability with a range of 8.72 g/m^2 – 220.77 g/m^2 was observed, which underscores the variability of the crop water content across the sampled taro field. The standard deviation (Std.) was 44.49 $g/m^{2,}$ and the Coefficient of Variation (CV) was 72.70%.

Table 4-2 Descriptive statistics of the in-situ measured EWT_{canopy} of taro crops

EWTcanopy (gm2)	n	Mean	Median	Minimum	Maximum	SEM	Std	CV%
Entire Dataset	100	60.70	47.70	8.72	220.77	4.45	44.49	72.70
Training Data	70	62.85	47.72	8.79	220.77	5.61	46.94	74.14
Testing Data	30	56.33	41.54	8.72	173.30	7.04	38.54	67.26

4.3.2 Predicting EWT canopy of taro crops using the Deep Neural Network algorithm and image segmentation techniques

Figure 4.4 illustrates the model accuracies achieved in predicting the taro EWT_{canopy} across the various index-based image segmentation techniques, including the ExG, ExR and ExGR crop canopy extraction applied to the raw UAV imagery. Figure 4.4a–d represents the prediction accuracies from the analysis conducted without the thermal band, while Figure 4.4e-h illustrates the estimation accuracies based on the models, and including the thermal band.

Generally, the exclusion of the thermal band resulted in relatively low model accuracies for all index-based segmentation methods. For instance, when estimating the EWT_{canopy} based on the raw UAV image, excluding the thermal band, it yielded a poor R^2 of 0.35, an RMSE of 34.96 g/m^2 and an rRMSE of 60.51 (Figure 4.4a), with blue, green, red, TCARI, NIR, red-edge, NGRDI, CIrededge, NDVI and MCARI emerging as the important predictor variables (Figure 4.5a). Meanwhile, the EWT_{canopy} estimation derived from the raw UAV image, including the thermal band, resulted in a moderately higher prediction accuracy of R^2 of 0.61, an RMSE of 25.35 g/m^2 and an rRMSE of 43.87% (Figure 4.4e). The most influential predictor variables were CIrededge, NDRE, blue, NGRDI, green, TCARI, thermal, NDWI, rededge and NDVI, in descending order of importance (Figure 4.5e).

Overall, an increase in the prediction accuracy was observed when index-based image segmentation techniques and thermal data were applied to the raw UAV imagery. The ExR-based EWT_{canopy} model, including the thermal band, yielded a R^2 of 0.79, an RMSE of 14.53 g/m^2 and an rRMSE of 27.76% (Figure 4.4g), while the ExG-based model, including thermal, achieved an R^2 of 0.90, an RMSE of 10.69 g/m^2 and an rRMSE of 18.82% (Figure 4.4f). The most important predictor variables were CIrededge, thermal, SAVI, red and NIR, NGRDI, TCARI, blue, red-edge and NDWI for the ExR-based EWT_{canopy} model (Figure 4.5g), whereas the ExG-based model had CIrededge, TCARI, red, green, NGRDI, NDWI, thermal, red-edge, SAVI and blue band as the most influential predictor variables (Figure 4.5f). The ExGR-based EWT_{canopy} model, including the thermal, exhibited the highest prediction accuracy, by achieving an optimal R^2 of 0.92, an RMSE of 8.04 g/m^2 and an rRMSE of 15.31%, based on TACRI, CIrededge, MCARI, NGRDI, NDWI, blue red-edge, red, thermal and NDVI variables, in descending order of importance (Figure 4.4h & 4.5h).

Overall, the fitting degree between the EWT_{canopy} ExGR-based model (R^2 = 0.92) was significantly higher than that of the raw UAV image (R^2 of 0.61). By applying the ExGR technique and including the thermal band, the prediction accuracy of the target taro EWT_{canopy} was significantly improved. These outcomes underscore the efficacy of the DNN model, particularly when it is integrated with the thermal data and advanced image segmentation approaches in enhancing the precision of EWT_{canopy} estimations for taro crops.

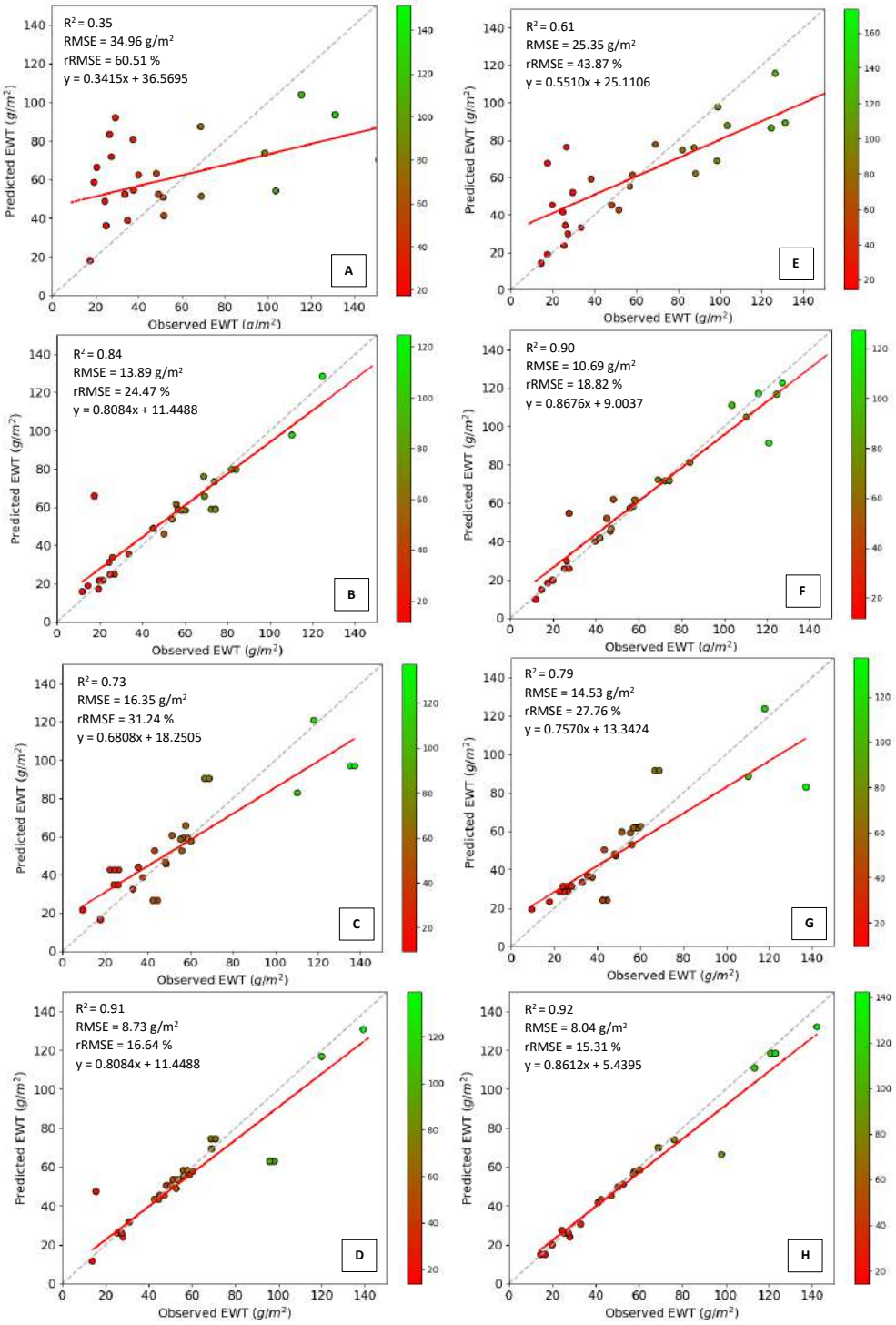


Figure 4-4 Relationship between the predicted and observed EWT_{canopy} based on the a) raw UAV image excluding thermal, b) raw UAV image including thermal, c) ExG excluding thermal, d) ExG including thermal, e) ExR excluding thermal, f) ExR including thermal, g) ExGR excluding thermal and h) ExGR including thermal

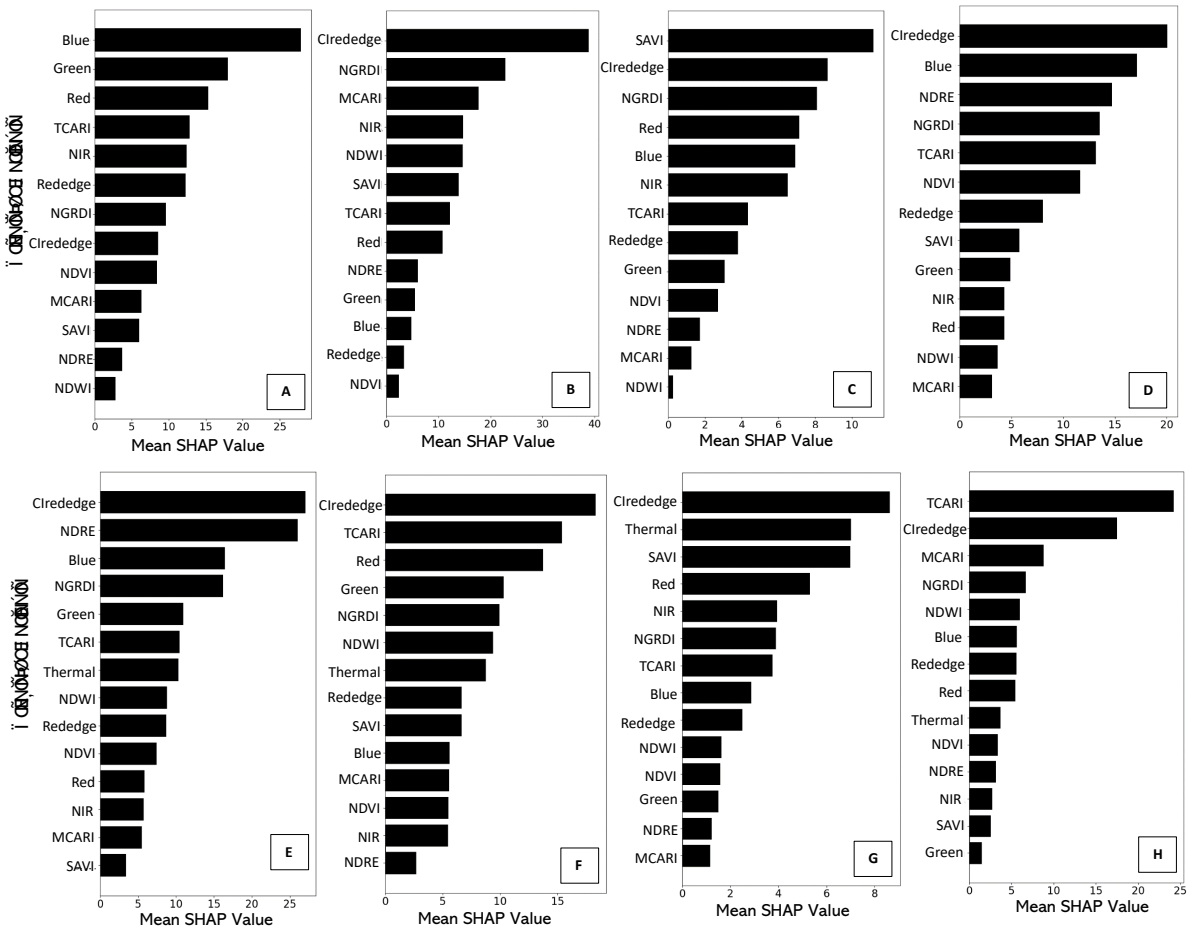


Figure 4-5 SHAP-generated variable importance of predictor variables in estimating EWT_{canopy} based on the a) raw UAV image excluding thermal, b) raw UAV image including thermal, c) ExG excluding thermal, d) ExG including thermal e) ExR excluding thermal, f) ExR including thermal, g) ExGR excluding thermal and h) ExGR including thermal

4.3.3 Assessment of the influence of the thermal band in conjunction with the Deep Neural Network algorithm

Generally, excluding the thermal band results in comparatively lower model accuracies across the index-based segmentation methods. This pattern is consistent across the R^2, RMSE and rRMSE metrics for all segmentation techniques, with generally higher estimation accuracies being achieved when the thermal band is considered in the taro EWT_{canopy} analysis. To investigate the impact of including the thermal band on prediction accuracy, an ANOVA test was conducted to assess the differences in accuracy among the segmentation techniques and the presence, or absence, of the thermal channel. The two-way ANOVA revealed a significant interaction between the prediction accuracies and the segmentation techniques, based on the inclusion and exclusion of the thermal band, which implies that the trends of variation of the parameters change with the inclusion of the thermal band and different segmentation techniques. This suggests that variations in prediction accuracies exhibit different trends, depending on whether the thermal band is incorporated and whether distinct segmentation techniques are employed. A statistical significance was observed in the accuracies of predicting

the EWT_{canopy} across all segmentation techniques ($P < 0.05$) (Figure 4.6a-c). In addition, significant differences in prediction accuracies ($P < 0.001$) were noted between the inclusion and exclusion of the thermal band across the segmentation methods (Figure 4.6d-f). The segmentation factor, the thermal band factor, and their interaction significantly contributed to the differences in the R^2, RMSE and rRMSE (Figure 4.6). However, no significant difference in prediction accuracy was observed for the inclusion and exclusion of thermal characteristics for the ExGR-based technique. The ExGR data were the most optimal model, as they exhibited an R^2 of 0.92, an RMSE of 8.04 g/m^2 and an rRMSE of 15.31% with the thermal band (Figure 4.6h), and they achieved an R^2 of 0.91, an RMSE of 8.73 g/m^2 and an rRMSE of 16.64%, even with the exclusion of the thermal band (Figure 4.4d). Following the optimal performance of the models that were based on the index-based image segmentation methods, in conjunction with the thermal band, these models were adopted in mapping the spatial distribution of EWT_{canopy} in the study area.

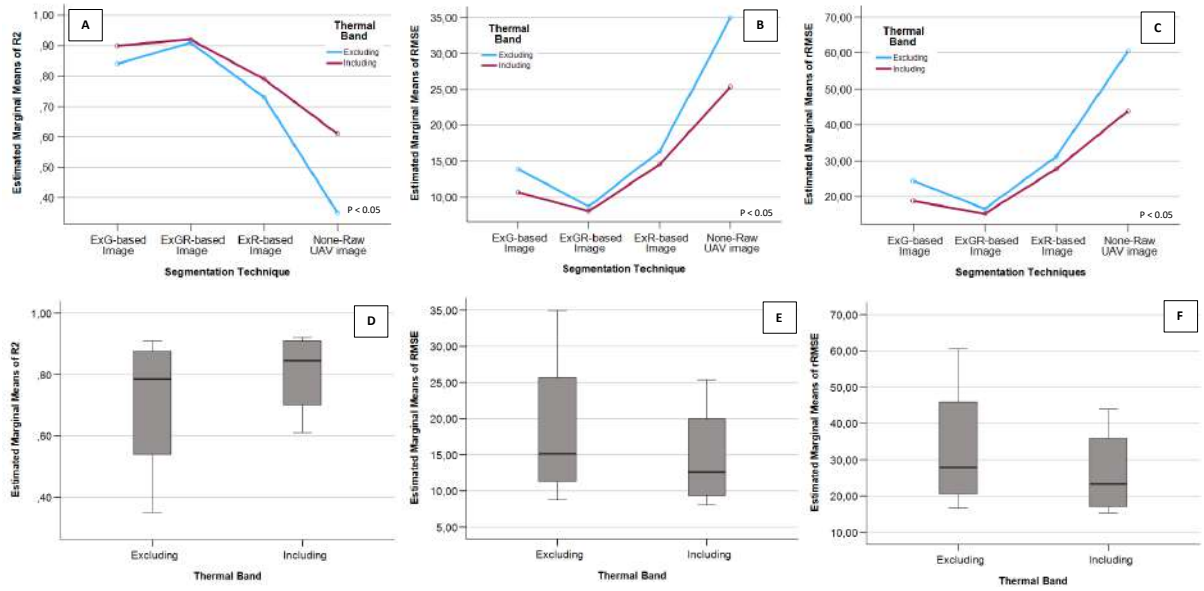


Figure 4-6 Comparative analysis of the mean performance metrics, including a) R^2, b) RMSE and c) rRMSE across the segmentation techniques between the inclusion and exclusion of the thermal band and overall mean d) R^2, e) RMSE and f) rRMSE obtained between including and excluding the thermal band

4.3.4 Spatial distribution of the EWT canopy of smallholder taro crops

Figure 4.7 presents the spatial distribution of EWT_{canopy} that was estimated by using UAV remotely sensed multispectral thermal data, as well as the different image segmentation techniques, including ExG, ExR and ExGR crop canopy extraction that were applied to the raw UAV imagery. While the raw UAV image achieved a moderate prediction accuracy, this approach could not accurately assign EWT_{canopy} because of the overfitting problem (Figure 4.7a). On the contrary, the ExGR-based EWT_{canopy} model produced results close to the observed EWT_{canopy} within the taro field and it achieved the highest accuracy in visual comparison (Figure 4.7d). Similarly, the ExG-based EWT_{canopy} model demonstrated comparable results in capturing the spatial distribution of the EWT_{canopy} of taro crops (Figure

4.7d). A general trend was observed in the south-western area (high elevation) of the field, which exhibited higher EWT_{canopy} estimates in relation to the eastern parts (Figure 4.7).

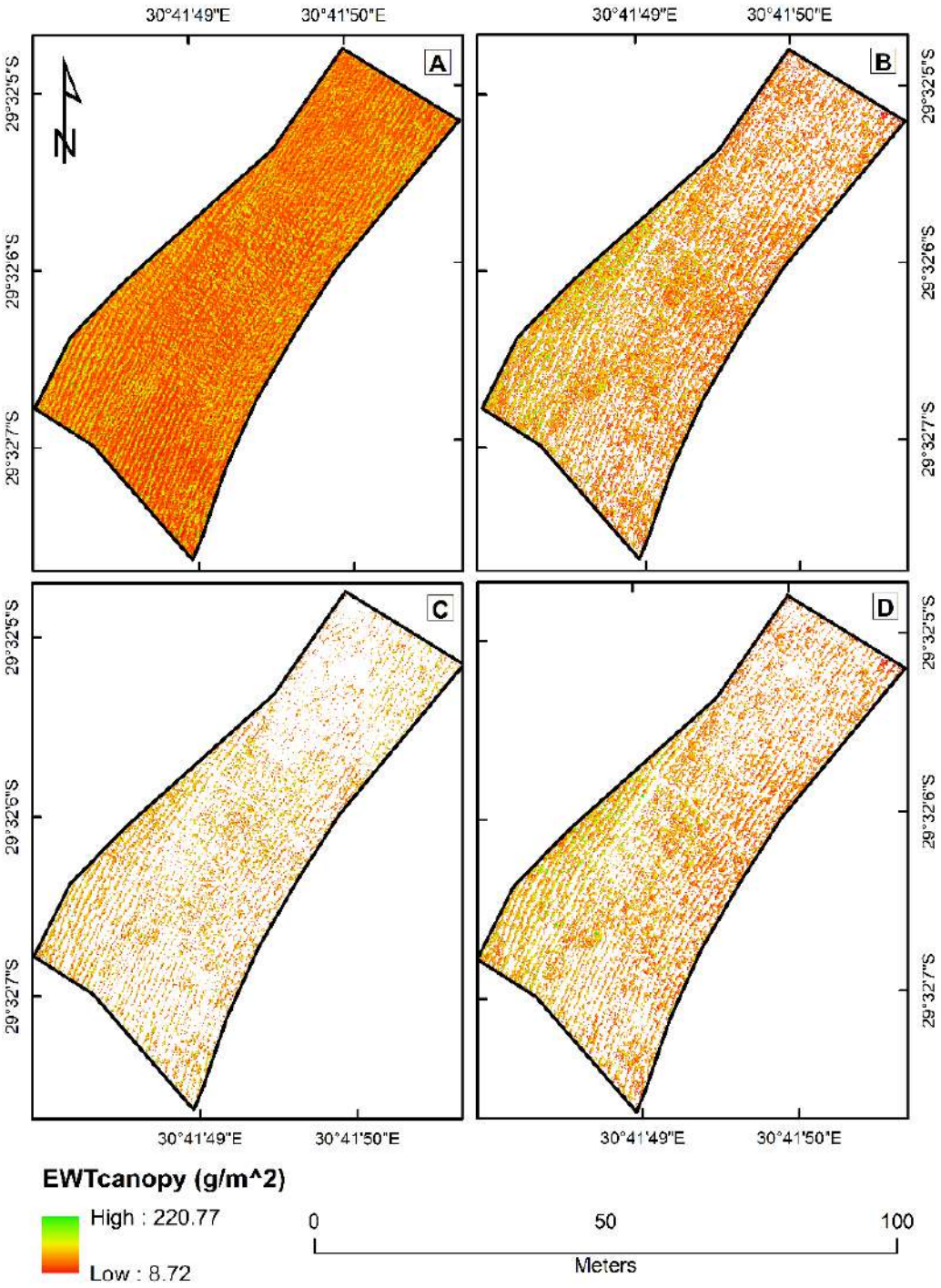


Figure 4-7 Spatial distribution of EWT_{canopy} within the taro field based on the a) raw UAV image, b) ExG, c) ExR and d) ExGR canopy extraction techniques

4.4 Discussion

This study sought to evaluate the use of thermal remote sensing and index-based segmentation techniques for enhancing the estimation of EWT_{canopy} on a smallholder taro farm. Specifically, the study investigated the capability of the UAV thermal band in estimating the EWT_{canopy} of the taro crop. In addition, this study assessed the influence of the soil background removal, by using ExG, ExR and ExGR segmentation indices, in enhancing the prediction accuracy of the taro EWT_{canopy}. Furthermore, the study aimed to explore the capabilities of UAV-derived multispectral thermal data in predicting EWT_{canopy} of taro crops.

4.4.1 Prediction of EWT canopy using UAV multispectral thermal imagery

The findings of this study showed a wide variety of the EWT_{canopy} measured within the taro field, ranging from 8.72 g/m^2 to 220.77 $g/m^{2,}$, with the highest values recorded at the lower elevations of the field. This spatial variability can generally be attributed to water drainage and accumulation, particularly the gravity-driven flow that tends to accumulate at lower elevations. This fosters a wetter environment that enhances soil moisture availability for crops in the area and leads to higher EWT_{canopy} values. A study by Basile *et al.* (2020) also confirmed the contribution of topographic characteristics, such as slope, on the crop water status, with down-slope soils exhibiting a higher water retention capacity, which increases water available to the crop. Therefore, future studies should consider quantifying the relative contribution of the topography to the spatial variability of the crop moisture content within small-scale farming systems.

The results of this study also showed that taro EWT_{canopy} could be optimally estimated by using the ExGR segmentation method, including thermal ($R^2 = 0.92$, RMSE = 8.04 g/m^2 and rRMSE = 15.31%), with the most influential variable being TACRI, CIrededge, MCARI, NGRDI, NDWI, blue, red-edge, red, thermal and NDVI (Figure4.5). These findings demonstrate the sensitivity of EWT_{canopy} to the NIR and red-edge sections of the electromagnetic spectrum. Literature has confirmed that the NIR region is valuable in characterising the crop water status, due to its strong water absorption ability that can detect variations in the crop water through changes in canopy reflectance under different water status (Pasqualotto *et al.* 2018, El-Hendawy *et al.* 2019, Brewer *et al.* 2022). The importance of red-edge-based indices, such as TCARI, CIrededge and MCARI, is attributed to the unique properties of the red-edge region, which is particularly sensitive to subtle changes in the vegetation structure and leaf pigmentation, due to water stress, hence they are invaluable for assessing the crop water status (Ballester *et al.* 2019, Colovic *et al.* 2022). The results of this study also reported indices, such as CIrededge and TCARI, as being among the influential variables in estimating EWT_{canopy} of taro crops (Figure 4.5). These findings are consistent with those of Zhang and Zhou (2019) and Colovic *et al.* (2022), which demonstrated the CIred-edge to be sensitive to variations in the crop water conditions and highly influenced by changes in the irrigation treatments. Similarly, Baluja *et al.* (2012) identified TCARI as a valuable predictor variable of the water status of vineyards, by yielding an optimal correlation of R^2 of 0.80.

4.4.2 Performance of the thermal band in predicting the EWT canopy of taro crops

The performance of the thermal band in predicting the EWT_{canopy} of taro crops revealed a consistent trend across the various index-based segmentation techniques. It was observed that the exclusion of the thermal band in the EWT_{canopy} analysis resulted in lower estimation accuracies ($P < 0.05$), which emphasises the importance of the thermal band in characterising the taro crop water status (Figure 4.6). Surprisingly, our study found no significant difference in the prediction accuracies when thermal data were considered, in comparison to their exclusion in the ExGR-based model. These results underscore the effectiveness of the ExGR-based technique, particularly its robust performance, irrespective of the inclusion and exclusion of the thermal channel.

In addition, it was observed that the thermal band was among the top predictor variables across all the EWT_{canopy} models. The literature confirms the invaluable role of thermal infrared remote sensing in assessing and monitoring the crop water status, as it establishes a direct correlation with crop water biophysical and biochemical elements (Khanal *et al.* 2017, Messina and Modica 2020, Krishna *et al.* 2021). The use of thermal remote sensing is based on the premise that the thermal characteristics of crop leaves are affected by leaf transpiration, which decreases in a state of the water deficit, resulting in a reduction of the leaf and canopy temperatures (Maes and Steppe 2012, Gerhards *et al.* 2019). The findings of this study align with a recent study by Guan and Grote (2023), which achieved an R^2 of 0.74 when incorporating the thermal channel, compared to an R^2 of 0.63 with the thermal band excluded, which highlights the integration of multispectral and thermal data and its combined value in understanding the crop water status. The findings of this study are further corroborated by García-Tejero *et al.* (2018), who concluded that the thermal band is feasible for monitoring almond water stress for irrigation scheduling, and by Cheng *et al.* (2023), who highlighted the applicability of thermal imaging in the assessment of the crop water conditions of the summer maize crop.

4.4.3 Performance of index-based segmentation techniques for the estimation of the taro EWT canopy

This study shows that the inclusion of the soil background reduces the accuracy of the EWT_{canopy} predictions within the taro crop (an R^2 of 0.61, an RMSE of 25.35 g/m^2 and an rRMSE of 43.87%). It was noted that the prediction accuracy of taro EWT_{canopy} improved significantly after the removal of the soil background by using the ExGR-based image segmentation technique, and that it yielded an optimal R^2 of 0.92, an RMSE of 8.04 g/m^2 and an rRMSE of 15.31%. These results align with the broader consensus in the literature. For instance, Xu *et al.* (2021) and Li *et al.* (2022) emphasised the challenge that is posed by the soil background in influencing the crop canopy spectra, particularly in UAV-derived imagery. Notably, while the ExG and ExR techniques demonstrated an acceptable accuracy in quantifying the taro EWT_{canopy} (an R^2 of 0.90 and 0.76, respectively), the ExGR method outperformed both these techniques. This notable enhancement can be attributed to the inherent capabilities of the ExGR technique in effectively mitigating soil background interference (Zhai *et al.* 2023). The comprehensive nature of the ExGR method combines the advantages of both the ExG and ExR by simultaneously leveraging the ExG for extracting the crop canopy and the

ExR for eliminating background noise (Meyer *et al.* 2004, Hamuda *et al.* 2016, Riehle *et al.* 2020, Upendar *et al.* 2021).

Overall, the removal of the soil background has been proven to be imperative for enhancing the accuracy of taro EWT_{canopy} predictions. These findings are further supported by Shu *et al.* (2021), who reported a significant increase in the prediction accuracy, from an R^2 of 0.45 and an RMSE of 7.13 before, to an R^2 of 0.74 and an RMSE of 3.68, after performing soil background removal in estimating the SPAD chlorophyll content of a maize crop. These parallel findings underscore the significance of addressing the soil background interference for accurate and reliable estimations in water-related crop assessments.

4.4.4 Implications of the study, its limitations and recommendations for future research

The results of this study underscore the significant potential of accurately assessing and monitoring the EWT_{canopy} of taro in smallholder rainfed croplands by using UAV-derived multispectral-thermal and the ExGR image segmentation. The value of the UAV thermal band in this study highlights its significance in enhancing the precision of EWT estimations for taro crops. The application of UAV-based multispectral and thermal data offers exceptional accuracy in predicting the water status of taro, and it provides valuable insights into the smallholder taro-farming practices. These findings emphasise the significance of precision agricultural technologies, particularly UAVs, in enhancing smallholder agricultural management. The near-real-time capabilities of UAV-derived data enable rapid and effective decision-making, and they contribute to the improved crop health and productivity of crops, particularly NUS crops, such as taro. Despite the promising outcomes, it is worth noting the limitations. The high spatial resolution provided by the UAV onboard sensor contrasts with the limitations in the spectral resolution, which constrains the available multispectral bands and derived vegetation index options. In addition, while the study employed a high-precision GPS system, with up to a one-centimetre accuracy, careful consideration is essential during the sampling procedure to maintain the precision and accuracy of the sample locations. Furthermore, in order to advance the understanding of the crop water status, future studies should explore indicators beyond the EWT, including the stomatal conductance and more detailed assessments of canopy temperature, by incorporating thermal-related indices. Further research should also consider integrating structural features, such as crop cover, plant height, chlorophyll content and topographic variables, namely the aspect, slope and elevation, to enhance the precision of crop water predictions. In addition, exploring alternative image segmentation techniques can contribute to refining the accuracy and applicability of the models.

4.5 Conclusion

This study sought to evaluate the use of thermal remote sensing and index-based segmentation techniques in predicting the EWT_{canopy} in smallholder taro farmlands. The results showed that the UAV thermal data integrated with the multispectral data, and that the ExGR segmentation

technique is optimal for estimating the EWT_{canopy} of taro crops. Based on the findings of this study, it can therefore be concluded that:

- the exclusion of the thermal band in the EWT_{canopy} estimations reduces the prediction accuracies of the taro water status;
- the ExGR segmentation technique proved most effective in extracting crop canopy spectra and removing soil background, thereby enhancing the prediction accuracies of the taro EWT_{canopy}; and
- the NIR, red-edge and thermal sections of the electromagnetic spectrum were identified as being the most influential in estimating the EWT_{canopy} of taro crops.

These findings underscore the prospects of using UAV proximal remote sensing techniques to provide near-real-time spatial information for assessing the crop water status of smallholder NUS taro crops. The results of this study will assist the development of a robust, spatially-explicit monitoring framework for taro water status by filling the knowledge gap, which is essential for integrating taro into sustainable agricultural practices. The findings serve as a stepping-stone to improved agricultural management practices and resource optimisation in the small-scale cultivation of NUS crops, and they align with the ongoing efforts to bridge the gap between research and practical implementation. Therefore, the insights gained from this study are imperative for decision-makers who are aiming to streamline NUS crops, particularly taro, into the existing agricultural systems, in order to enhance food security and to promote resilient farming practices in diverse agroecological environments.

5 ASSESSING THE WATER STATUS OF NEGLECTED AND UNDER-UTILISED TARO CROPS USING PHYSIOLOGICAL INDICATORS AND UAV MULTISPECTRAL-THERMAL DATA

5.1 Introduction

Neglected and Under-utilised crop Species (NUS) have emerged as vital assets in sustainable agricultural practices, and they offer the opportunity to diversify cropping systems and enhance resilience, in the light of the escalating climate change impacts and the need to ensure food security (Chivenge *et al.* 2015, Popoola *et al.* 2019, Mugiyo *et al.* 2021). In developing regions, such as southern Africa, where pre-existing challenges of hunger and nutrition insecurity persist, the cultivation of NUS, which is commonly confined to smallholder farms, plays a critical role in sustaining the local food security (Chivenge *et al.* 2015, Kapari *et al.* 2023). Despite contributing to approximately 80% of food production, smallholder systems are disproportionately susceptible to the effects of climate change and frequently experience periods of water stress and drought, which lead to substantial crop yield losses (Gomez y Paloma *et al.* 2020, Hlophe-Ginindza and Mpandeli 2020). Taro *(Colocasia esculenta (L))*, which is renowned for its drought and heat tolerance, holds great potential as a future smart crop that is capable of thriving in diverse agroecological environments (Mabhaudhi and Modi 2015, Oyeyinka and Amonsou 2020, Talucder *et al.* 2024). Locally (in South Africa) taro is referred to as *amadumbe*, and it has garnered significant attention as a priority crop, due to its resilience to drought and heat stress, and its edible tubers are of a high nutritional value (Mabhaudhi *et al.* 2017, Munialo *et al.* 2024). Despite its promising attributes, there is a dearth in literature on its crop water status variations, which has hindered a comprehensive understanding of its adaptive capabilities. Therefore, research on the water status of taro becomes imperative for optimising its productivity, and it is also a step towards its integration into the mainstream agricultural systems.

Having an understanding of the variations of the physiological indicators of the taro crops, in response to water deficit, is essential for designing and implementing effective crop water stress management strategies (Huang *et al.* 2020, Parkash and Singh 2020). During periods of water deficit, plants undergo a variety of physiological changes that are aimed at conserving water and mitigating the stress (Shafiq *et al.* 2019, Wang *et al.* 2022). For example, stomatal conductance, which is the measure of gas exchange through the leaf stomata, typically decreases under water stress conditions, the as plants close their stomata to reduce water loss through transpiration (Gerhards *et al.* 2019, Wang *et al.* 2022). This decrease limits the uptake of carbon dioxide for photosynthesis, which leads to reduced plant productivity and growth (Ahmad *et al.* 2021). In addition, the chlorophyll content, which is essential for absorbing light energy during photosynthesis, declines due to the detrimental effects of water stress on the chloroplast membrane and photo-oxidation (Shafiq *et al.* 2019, Parkash and Singh 2020). Meanwhile, the leaf temperature, which is influenced by evaporative cooling processes, often increases under water stress, as plants close their stomata to conserve water, which results in reduced transpiration cooling (Gerhards *et al.* 2019, Parkash and Singh 2020). Then the

equivalent water thickness and fuel moisture content, which reflect the moisture levels within plant tissues, typically decrease during water stress, as the plants experience dehydration (Ahmad *et al.* 2021, Ndlovu *et al.* 2024). Therefore, by monitoring the variations in these physiological indicators, we can effectively quantify the crop water status and implement timely interventions to mitigate the impacts of a water deficit on the productivity of the taro crop. Although traditional methods of retrieving the physiological indicators of the crop water status are accurate, these techniques are labour-intensive, time-consuming, and they generally rely on an exhaustive chemical analysis (Yin *et al.* 2023). Moreover, these measurements are prone to inaccuracies and do not sufficiently indicate the spatial variability of the crop water status across the entire field (Mobasheri and Fatemi 2013, Traore *et al.* 2021).

Thermal remote sensing has proven to be a robust tool for assessing crop water status and the related physiological indicators, due to its inherent ability to detect subtle variations in the surface temperature, which are indicative of a crop"s water status (Khanal *et al.* 2017, Awais *et al.* 2022). By capturing thermal emissions from the earth's surface, thermal sensors provide valuable insights into the crop water status, as well as the related heat dissipation processes that are influenced by transpiration, evaporation and the plant water uptake dynamics (Gerhards *et al.* 2019, Zhai *et al.* 2023). Satellite-based thermal sensors, such as the Landsat Thermal Infrared Sensor (TIRS), Advanced Spaceborne Thermal Emission and Reflection Radiometer and the Ecosystem Spaceborne Thermal Radiometer Experiment on Space Station have been instrumental in obtaining regional-scale crop water status assessments (Gerhards *et al.* 2019, Awais *et al.* 2022). Despite their utility, satellite-based thermal sensors are restricted by their coarse spatial resolution, which limits the characterisation of the crop water status at a field scale (Khanal *et al.* 2017). Consequently, alternative approaches are necessary to capture finer-scale variations in the crop water status and to facilitate precision agricultural practices.

In recent years, advancements in earth observation technologies, specifically the use of Unmanned Aerial Vehicles (UAVs) that are equipped with multi-modal thermal multispectral sensors, have revolutionised the characterisation of the crop water status and associated physiological indicators, particularly on a local scale (Hussain *et al.* 2020). UAVs, equipped with very-high-spatial resolution sensors of up to sub-centimetre precision, enable the precise assessment of the crop water status at a field scale (Chivasa *et al.* 2020). In comparison to satellites, UAVs offer flexibility in flight planning, and the ability to capture data on-demand, which makes them ideal for monitoring a crop's physiological dynamics, such as taro's crop water stress, in fragmented heterogeneous smallholder agricultural systems (Zhu *et al.* 2021). This is demonstrated by studies by Marques *et al.* (2020) who utilised UAV-acquired thermal data to characterise the stomatal conductance and relative water content of a commercial olive orchard, and studies by Han *et al.* (2021) that achieved an optimal R^2 value of 0.96 in estimating the canopy temperature of apple trees, by using UAV thermal imagery.

The integration of UAVs with machine learning methods, particularly deep learning approaches, such as Deep Neural Networks (DNNs), has further enhanced the analysis of remotely-sensed agricultural data (de Melo *et al.* 2022). The DNN algorithm is robust in characterising complex relationships, which allows for the more accurate prediction of crop attributes, including the water status (Omosalewa *et al.* 2022, Rajwade *et al.* 2023). However,

despite these advantages, the integration of DNN algorithms with UAV thermal remote sensing for crop water estimations on smallholder farms remains elusive. As a result, there is need for further research to fully exploit and incorporate the capabilities of DNN algorithms into precision agricultural applications. Therefore, this study aimed to leverage the capabilities of UAV multi-modal thermal multispectral remote sensing and deep neural network techniques to assess the crop water status of taro. This study specifically aimed to evaluate the performance of selected thermal and multispectral variables, along with their derived indices, in estimating the status of crop water physiological indicators (i.e. the equivalent water thickness, fuel moisture content, stomatal conductance, canopy temperature and chlorophyll content), and to optimise the prediction models by utilising optimal multi-modal variables derived from UAV-acquired thermal multispectral data.

5.2 Materials and Methods

5.2.1 Field sampling and in-situ measurements

The study area is detailed in Chapter One. To establish sampling points within the taro field, a polygon delineating the field boundary was constructed by using Google Earth Pro. After importing this polygon into ArcGIS Pro, 100 stratified random sampling points were created to ensure comprehensive coverage and accuracy in representing the diversity of taro crops across the field. Subsequently, these sampling points were transferred to a Trimble hand-held Global Positioning System (GPS) device (Trimble Inc., Sunnyvale, California, USA) with sub-centimetre accuracy, which allowed for the precise location determination of each sampling point within the taro field. At each sampling point, in-situ measurements were taken to calculate the respective values of the equivalent water thickness, fuel moisture content, stomatal conductance, canopy temperature and chlorophyll content. The sampling procedure of each physiological indicator is explained further below.

5.2.2 Equivalent water thickness and fuel moisture content

A hand-held LiCOR-2200C Plant Canopy Analyser (LI-COR Biosciences, Lincoln, Nebraska, USA) was used to measure the Leaf Area Index (LAI) of the taro crop. LAI measurements were taken at a 45° zenith angle with a 180° view cap, and a total of five readings were collected above the (A) canopy and below (B) the canopy, following the ABBBB sequence. Subsequently, the above-ground biomass of each sampled crop was recorded, with the Fresh Weight (FW) being measured on an analytical balance with a measurement error of 0.5 g. The collected biomass was then transferred to a brown paper bag, which was labelled to identify it, then placed to dry in an oven at 60°C for approximately 72 hours until a consistent Dry Weight (DW) was reached. LAI, FW and DW served as the input variables for computing the taro Equivalent Water Thickness (EWT) and Fuel Moisture Content (FMC) by using the following equations:

$$Equivalent\ Water\ Thickness = \frac{FW-DW}{LAI} \qquad \text{Units: g/m}^2 \qquad (1)$$

$$Fuel\ Moisture\ Content = \frac{FW-DW}{DW}\ x\ 100 \qquad \text{Units: \%} \qquad (2)$$

5.2.3 Stomatal conductance

An SC-1 leaf porometer (Decagon Devices Inc., Pullman, Washington, USA) was used to measure the stomatal conductance. This device was selected based on its capability to record leaf conductance rates, in relation to the external environmental factors (Byambadorj *et al.* 2023). Prior to conducting measurements in the field, the calibration of the SC-1 leaf porometer was conducted to ensure a thermal equilibrium between the leaf clip and the environment. During the calibration process, filter paper that was saturated with distilled water from the sensor kit was placed over the aperture in the calibration plate. Subsequently, the sensor head was fixed to the calibration plate, and an initial measurement period of 30 seconds was established. The sensor was equilibrated after each measurement and the head of the sensor was reattached for the subsequent calibrations (Brewer *et al.* 2022). The calibration process was iterated up to ten times, to ensure stable measurements, thus standardising the readings of the instrument. Subsequently, stomatal conductance measurements were conducted at each sampling point by positioning the SC-1 leaf porometer at the centre of the leaf blade, perpendicular to the midrib of a sunlit taro leaf (Brewer *et al.* 2022).

The leaf blades receiving full exposure to sunlight were selected to ensure accurate stomatal conductance measurements and to avoid potential shading effects that could alter the photosynthetic characteristics of the crop (Brewer *et al.* 2022). The SC-1 leaf porometer measured the air temperature and humidity simultaneously and automatically recorded the leaf stomatal conductance (in mmol m^{-2} s^{-1}) for 30 seconds. Stomatal conductance measurements close to 0 mmol m^{-2} s^{-1} suggest the closure of the stomata to minimise water loss, which indicates an extreme water deficit, while values close to 500 mmol m^{-2} s^{-1} suggest open leaf stomata, which indicate minimal crop water stress and optimal physiological activity (Byambadorj *et al.* 2023).

5.2.4 Foliar temperature

The taro foliar temperature was assessed by using a digital laser infrared GM320 hand-held thermometer (Shenzhen Jumaoyuan Science and Technology Co., Ltd, Nanshan District, Shenzhen, China). The GM320 has a 1.5°C error margin and is characterised by a wide measurement range spanning from -50°C to 380°C and a spectral response of 8-14 µm. At each sampling point, five foliar temperature measurements were taken and averaged, to ensure accuracy. These measurements were acquired at specific locations on the taro leaf, namely: (1) the centre of each leaf blade, near the primary leaf vein, (2) roughly one-third along from the tip of the leaf, (3) about two-thirds of the way down from the leaf tip, and (4) at the midpoint between the leaf vein and the leaf margin on both sides. This comprehensive approach facilitated a thorough assessment of the leaf temperature distribution and variability across the taro foliage, which provided valuable insights into its thermal characteristics.

5.2.5 Chlorophyll content

The amount of chlorophyll in taro leaves was determined by using a Soil Plant Analysis Development (SPAD) 502 chlorophyll meter (Minolta Corporation Ltd., Osaka, Tokyo, Japan).

The SPAD meter facilitates the rapid non-destructive measurement of the chlorophyll content by emitting light through the crop leaves and quantifying the light transmitted in the red (650 nm) and near-infrared (940 nm) regions (Delegido *et al.* 2011). The device instantaneously generates a unitless value, which correlates with the chlorophyll content that is present in the sampled leaf (Uddling *et al.* 2007, Bayat *et al.* 2016). Five chlorophyll measurements were obtained per sunlit leaf sampled at the top of the taro canopy at specific locations: (a) the middle point of every leaf blade, adjacent to the primary leaf vein, (b) about one-third down from the leaf tip, (c) about two-thirds from the leaf tip, and (d) at the middle point between the leaf vein and leaf margin on either side. These measurements were then averaged for each leaf and noted accordingly. The SPAD meter was protected from direct sunlight, in order to ensure accuracy when taking the chlorophyll readings. Finally, using the equation provided by Markwell *et al.* (1995) that demonstrated an R^2 of 0.94, SPAD measurements were converted to a value for chlorophyll content:

$$Chlorophyll\ Content = 10^{M^{0.265}} \qquad \text{units: } \mu\text{mol m}^{-2} \quad (3)$$

where the chlorophyll content, measured in μmol m^{-2}, is represented by the SPAD value (M), which indicates the total chlorophyll per unit leaf area (Ling *et al.* 2011, Bayat *et al.* 2016). All field measurements were performed on a sunny day from 10:00 am to 2:00 pm local time, which coincided with the ideal time-frame for crop photosynthesis. This timing was selected to ensure that the gathered data captures the peak reflectance and photosynthetic function of the taro plants precisely.

5.2.6 Image acquisition and pre-processing

The DJI Matrice 300 (M300) (DJI-Innovations Inc., Shenzhen, China) multi-rotor platform and MicaSense Altum multispectral-thermal camera (MicaSense, Seattle, WA, USA) were flown over the taro field. The MicaSense Altum camera combines five narrow high-resolution bands to assess the reflectance across various spectral ranges, including the blue (475 nm), green (560 nm), red (668 nm), red-edge (717 nm) and near-infrared (NIR) (840 nm), and it features a radiometric longwave infrared thermal channel (11 μm). To achieve optimal field coverage, a flight path was generated by transferring a shapefile of the experimental field directly to the smart controller of the M300. Prior to, and post, each flight, the MicaSense Altum camera was calibrated by using a Calibrated Reflectance Panel (CRP), based on known reflectance values across the spectrum, to perform the radiometric calibration on the sensor. Subsequently, an image of the CRP was captured under consistent lighting conditions, and it was utilised to compensate for the atmospheric and environmental conditions that were encountered during the flight. The autonomous aerial survey was carried out at an altitude of 100 m, by maintaining an 80% forward and sideward image overlap, which achieved 10.08 cm per pixel spatial resolution imagery. The drone flight was conducted during periods of peak solar radiation, typically from 10:00 am to 2:00 pm local time. This flight mission was executed concurrently with the field measurements, in order to ensure accurate synchronisation and comprehensive data collection.

The automated flight mission produced 1 626 raw images, which were then imported into Pix4D photogrammetry software (Pix4D, Lausanne, Switzerland) for pre-processing. Within Pix4D, the software automatically identifies the CRP images, leveraging them to execute radiometric corrections and generate an ortho-mosaic of the smallholder taro field. Subsequently, the ground reference points that were surveyed prior to the fieldwork were utilised to refine the geometric accuracy of the acquired images within the ArcGIS Pro software. The in-situ field measurements were superimposed on top of the obtained multispectral-thermal imagery, and both were georeferenced to the Universal Transverse Mercator zone 36S projection. Crop reflectance characteristics from both the multispectral and thermal wavelengths were extracted to compute the thermal and spectral indices, as outlined in Table 5.1. These specific indices were chosen due to their documented strong correlation with the plant water conditions, which made them reliable indicators for assessing and monitoring the crop water status (Baluja *et al.* 2012, Gago *et al.* 2017, Zhang and Zhou 2019, Ozelkan 2020). In calculating the thermal indices, the Normalised Relative Canopy Temperature (NRCT) and Second Formulation of the Stomatal Conductance Index (I3), T_{canopy} were determined by the canopy temperature of each sampled taro crop. Meanwhile, the T_{wet} was determined by using the lower boundary temperature, assuming that these crops were experiencing maximum transpiration (open stomata), and the T_{dry} was determined by using the higher boundary temperature, assuming that these crops were experiencing minimum transpiration conditions (closed stomata) (Baluja *et al.* 2012, Crusiol *et al.* 2020).

Table 5-1 UAV-derived spectral indices used in this study

Vegetation Index	Abbreviation and Equation	Reference
Thermal Indices		
Normalised Relative Canopy Temperature	$NRCT = \frac{T_{canopy} - T_{wet}}{T_{dry} - T_{wet}}$	(Elsayed *et al.* 2015)
Second Formulation of the Stomatal Conductance Index	$I3 = \frac{T_{canopy} - T_{wet}}{T_{dry} - T_{canopy}}$	(Baluja *et al.* 2012)
Spectral Indices		
Normalised difference vegetation index	$NDVI = \frac{R_{Nir} - R_{red}}{R_{Nir} + R_{red}}$	(Rouse *et al.* 1974)
Soil-Adjusted Vegetation Index	$SAVI = \frac{R_{Nir} - R_{red}}{R_{Nir} + R_{red} + L} \times (1 + L)$	(Huete 1988)
Normalised Difference Red Edge Index	$NDRE = \frac{R_{Nir} - R_{Rededge}}{R_{Nir} + R_{Rededge}}$	(Gitelson and Merzlyak 1994)
Normalised Green-Red Difference Index	$NGRDI = \frac{R_{Green} - R_{red}}{R_{Green} + R_{red}}$	(Hunt Jr *et al.* 2013)
Photochemical Reflectance Index	$PRI = \frac{R_{Rededge} - R_{Green}}{R_{Rededge} + R_{Green}}$	(Gamon *et al.* 1992)

Transformed Chlorophyll Absorption in Reflectance Index	$TCARI = 3 \times (R_{Rededge} - R_{Red}) - 0.2 \times (R_{Rededge} - R_{Green}) \times \frac{R_{Redegde}}{R_{Red}}$	(Haboudane *et al.* 2002)
Modified Chlorophyll Absorption in Reflectance Index	$MCARI = (R_{Nir} - R_{Rededge}) - 0.2\,x\,(R_{Nir} - R_{Green}) \times \frac{R_{Nir}}{R_{Rededge}}$	(Daughtry *et al.* 2000)
Red edge chlorophyll index	$CIrededge = \frac{R_{Nir}}{R_{rededge}} - 1$	(Gitelson *et al.* 2003, Gitelson *et al.* 2006)

5.2.7 Crop canopy index-based segmentations

Crop canopy index-based segmentation was executed on the UAV multispectral-thermal image by using the Excess Green minus Excess Red (ExGR) index technique (Hamuda *et al.* 2016). This method was chosen to effectively delineate the crop canopy, while eliminating the soil background from the multispectral thermal image, to ensure the accuracy of crop water predictions across the taro field. Unlike other segmentation techniques, such as the Excess Green index, the ExGR index does not rely on a specific threshold for calculation, as positive pixel values are classified as the crop canopy, while negative values represent the background (Riehle *et al.* 2020). Following the calculation detailed below, a fixed zero threshold was used to automatically generate a binary image from the grey-level histogram in Mathworks MatLab (Hamuda *et al.* 2016, Shu *et al.* 2021). The resulting image was then reclassified into vegetation pixels (assigned Value 1) and soil pixels (assigned Value 0). The crop canopy mask was then converted into a shapefile and overlaid onto the original UAV image in ArcGIS Pro, effectively removing the soil background (Figure 5.1).

$$ExGR = (2 \times R_{Green} - R_{Red} - R_{Blue}) - (1.4 \times R_{Red} - R_{Green}) \quad (4)$$

where *R* represents the reflectance values obtained from the respective spectral band.

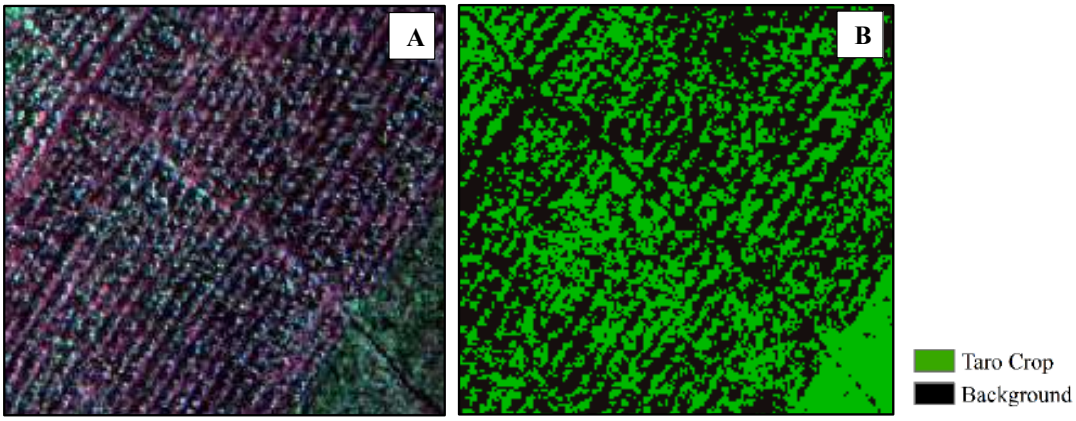


Figure 5-1 Taro crop canopy extraction by using the ExGR index-based segmentation technique

5.2.8 Model development and statistical analysis

This study utilised a deep machine learning approach to develop models for estimating the physiological indicators of the taro crop water status, which included th equivalent water thickness, the fuel moisture content, the stomatal conductance, the temperature and the chlorophyll content. Specifically, we leveraged the abilities of the Deep Neural Network (DNN), which is an artificial neural network that is distinguishable by its multiple interconnected hidden layers, to estimate the crop water-related indicators (Traore *et al.* 2021). These hidden layers work collaboratively to learn the complex features, which enables the transformation of input data into valuable output information (Chew *et al.* 2020, Bouguettaya *et al.* 2022). Unlike conventional machine learning methods that require a prior knowledge of the parameters, DNN relies on the feature representations acquired directly from the input dataset (Omosalewa *et al.* 2021). This distinct strategy enables the model to comprehend and detect complex non-linear relationships among the variables, thereby improving its capacity to reveal intricate patterns and dependencies within the data (Omosalewa *et al.* 2021, Traore *et al.* 2021).

A three-layer neural network model consisting of a comprehensive set of 14 input variables, including five spectral channels, one thermal infrared band and eight vegetation indices, was assessed individually and used collectively to develop the optimal crop water content models (Figure 5.2). The models were stimulated and refined by using a Rectified Linear Unit (ReLU) with a maximum of 200 epochs to enhance the prediction accuracy. To address the challenge of overfitting, a drop-out regularisation technique was implemented, with the input and hidden layer drop-out rates set at 0.4 and 0.2, respectively. Furthermore, the SoftMax activation function was used for the transformation of raw neural network outputs into a vector of probabilities, while the Adaptive moment estimation (Adam) optimiser was chosen to achieve optimal results. Finally, the DNN models for crop-water related parameters were fine-tuned with a learning rate of 0.001 and a batch size of 32.

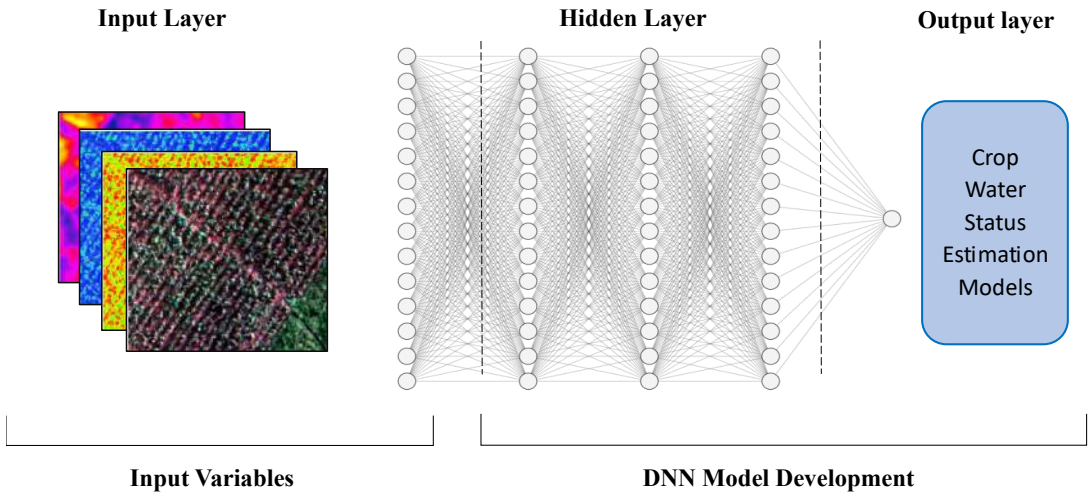


Figure 5-2 Graphical representation of the DNN structure, composed of input, hidden and output layers

5.2.9 Accuracy assessment

An accuracy assessment was performed by using the training dataset (n = 30) to evaluate the effectiveness of the DNN model in estimating the taro crop water indicators. Three key accuracy metrics were utilised for this evaluation: the coefficient of determination (R^2), the Root Mean Square Error (RMSE), and the Relative Root Mean Square Error (rRMSE). These metrics are detailed in Ndlovu *et al.* (2024). To understand the contribution of predictor variables in model development, the SHapley Additive exPlanations (SHAP) approach was employed. This approach calculates the marginal contribution value of each predictor variable, thereby ranking its importance in the model (Nahiduzzaman *et al.* 2023). Utilising SHAP provides a comprehensive understanding of the significance of individual predictor variables in influencing the accuracy of the crop water status physiological indicators estimation models.

5.2.10 A methodological overview on assessing the influence of thermal spectral data in characterising crop water status physiological indicators

Figure 5.3 illustrates the methodology that was used to evaluate the influence of integrating thermal spectral data on the prediction accuracy of the physiological indicators of the crop water status. The analysis followed a comparative approach across the different stages (a-d) to determine the significance of thermal data in enhancing the model's performance. Firstly, a DNN model was developed for each crop water status physiological indicator by using multispectral bands (blue, green, red, NIR and red-edge) to establish a baseline prediction accuracy (a). Thereafter, the thermal band was integrated into the model in the second stage, to assess the impact of the raw thermal data on model accuracy (b). In the third stage (c), thermal indices were incorporated to assess the influence of including the raw thermal band and derived thermal indices (NRCT and I3) in the prediction accuracies of the crop water indicators. Lastly (d), for comparative purposes against thermal indices, a final analysis was conducted for each crop water status indicator, using only the spectral indices derived from multispectral wavelengths (i.e. spectral indices (NDVI, SAVI, NDRE, NGRDI, PRI, TCARI,

MCARI, CIrededge)), to compare the efficiency of spectral indices in characterising the water status variations of taro crops. The influence of thermal remote sensing in enhancing the prediction accuracy of the crop water status indicators is determined by the model with the highest R^2 and the lowest RMSE and rRMSE values.

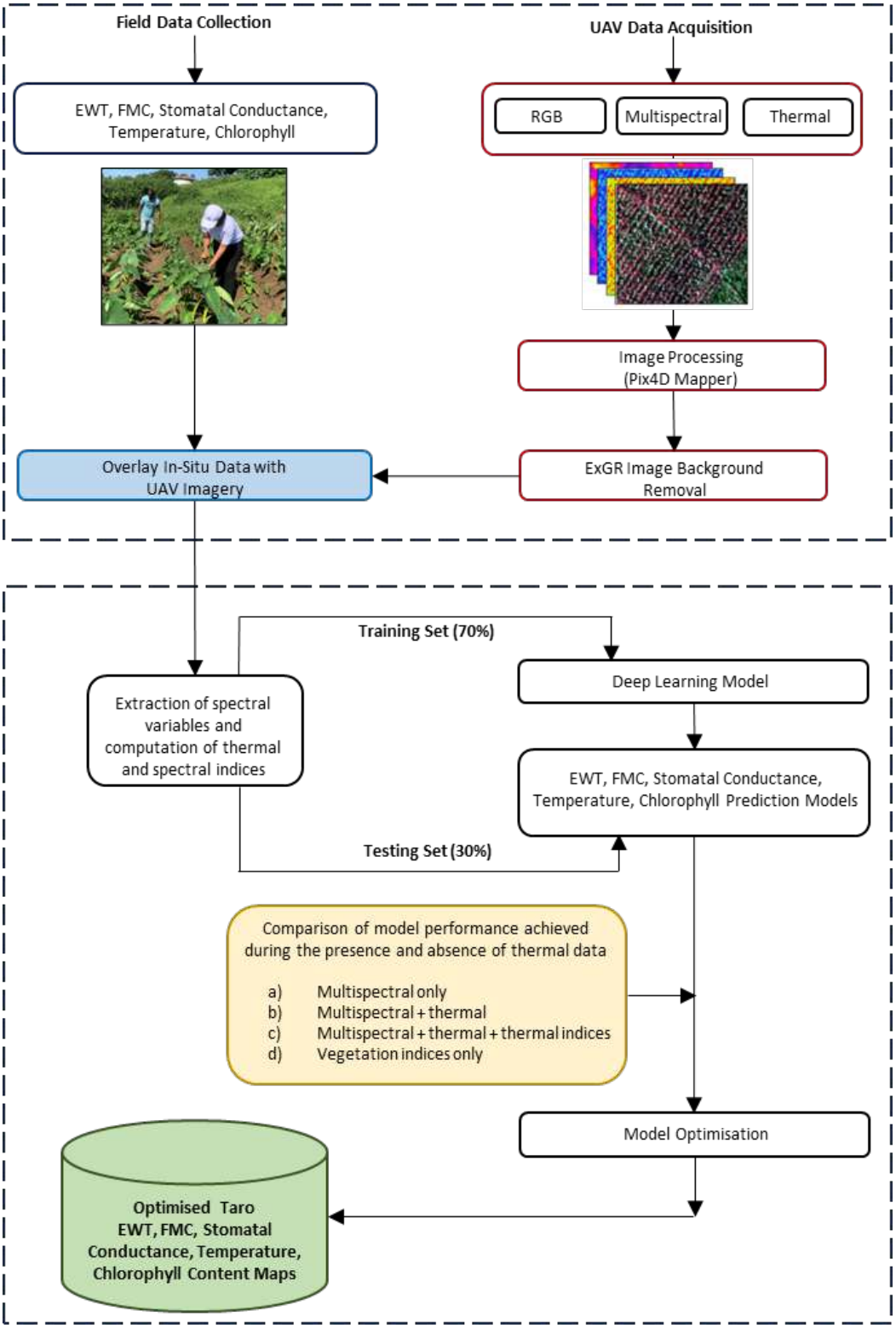


Figure 5-3 Methodological overview of the study

5.3 Results

5.3.1 Descriptive analysis of taro crop water status indicators

Table 5.2 presents the descriptive analysis results and variability of the taro crop water status indicators across the smallholder field. For equivalent water thickness, the range spanned from 8.72 g/m² to 220.77 g/m², with a mean value of 59.77 ± 42.17 g/m². The fuel moisture content ranged from 28.68% to 84.43%, with a mean of 68.96 ± 10.19%. Similarly, the stomatal conductance varied from 43.10 mmol m^{-2} s^{-1} to 575.90 mmol m^{-2} s^{-1}, with a mean value of 234.09 ± 149.77 mmol m^{-2} s^{-1}. The leaf temperature varied from 26.10°C to 33.40°C, with a mean of 29.90 ± 1.65°C, while the chlorophyll ranged from 587.89 µmol/m² to 1643.95 µmol/m², with a mean value of 1118.07 ± 229.59 µmol/m².

Table .5-2 Descriptive analysis of the taro crop water status physiological indicators

Indicator	Min	Max	Mean	Median	SEM	Std	CV%
Equivalent Water Thickness (g/m^2)	8.72	220.77	59.77	48.20	4.77	42.17	70.54
Fuel Moisture Content (%)	28.68	84.43	68.96	70.69	1.15	10.19	14.78
Stomatal Conductance (mmol m^{-2} s^{-1})	43.10	575.90	234.09	174.85	16.96	149.77	63.98
Leaf Temperature (°C)	26.10	33.40	29.90	29.80	0.19	1.65	5.51
Chlorophyll Content (µmol/m^{-2})	587.89	1643.95	1118.07	1135.20	26.00	229.59	20.53

The standard error of mean is denoted by SEM, the standard deviation by Std., and the coefficient of variation by CV

5.3.2 Estimation of the taro crop water status physiological indicators using selected spectral variables

The estimation accuracies of the physiological indicators of the taro crop water status, using selected spectral variables, are presented in Table 5.3. When considering the prediction accuracies based on multispectral bands alone, a poor-to-moderate model performance was observed across all physiological indicators. Specifically, an equivalent water thickness, fuel moisture content and chlorophyll content reported R^2 values between 0.30 and 0.36, while the stomatal conductance and leaf temperature recorded R^2 values of 0.46 and 0.47, respectively. Furthermore, both the RMSE and rRMSE values were relatively high across the indicators, which revealed significant variations between the predicted and observed measurements that were achieved, based on multispectral bands. Meanwhile, incorporating the thermal band into the analysis improved the prediction accuracies, particularly for the stomatal conductance and leaf temperature, and it yielded an impressive R^2 of 0.80 and 0.65, respectively. The corresponding RMSE values were 66.54 mmol m^{-2} s^{-1} and 0.89°C, with rRMSE values of 29.16% and 2.96%, respectively (Table 5.3).

Similarly, the integration of multispectral and thermal bands with thermal indices improved the prediction accuracies further, across all crop water status physiological indicators, with higher R^2 and lower RMSE and rRMSE values. Notably, the leaf temperature achieved the highest model accuracy, based on this integration, yielding an R^2 of 0.74, an RMSE of 0.76°C and an rRMSE of 2.54%. Meanwhile, the incorporation of thermal indices did not significantly

enhance the predictive performance of fuel moisture content and chlorophyll content, with both indicators achieving moderate accuracies (R^2 = 0.55, RMSE = 6.48%, rRMSE = 9.22% and R^2 = 0.53, RMSE = 148.99 μmol/m^{-2}, rRMSE = 13.6 %, accordingly) (Table 5.3).

The analysis revealed that spectral indices yielded the highest prediction accuracies across several physiological indicators. For instance, the fuel moisture content and chlorophyll content achieved their peak prediction accuracies when the spectral indices were selected as input variables, yielding R^2 values of 0.77 and 0.76, respectively. The corresponding RMSE and rRMSE values were 4.60% and 6.55% for the fuel moisture content, and 115.77 μmol/m^{-2} and 10.56%, for the chlorophyll content. While the equivalent water thickness obtained an R^2 of 0.73, the model produced a relatively high rRMSE of 33.82%, which indicated a notable variability between the predicted and actual values. Interestingly, while most physiological indicators performed best with general spectral indices, the stomatal conductance and leaf temperature achieved their highest prediction accuracies when integrating the multispectral and thermal bands with the thermal indices. Stomatal conductance, for example, achieved an impressive R^2 of 0.84, with an RMSE of 59.23 mmol m^{-2} s^{-1} and rRMSE of 25.96%, while the leaf temperature exhibited an R^2 of 0.74, with an RMSE of 0.76℃ and an rRMSE of 2.54% (Table 5.3).

Table 5-3 Estimation accuracies of taro crop water status physiological indicators using selected spectral variables

Predictor Variables	**Accuracy Metric**	**Equivalent Water Thickness (g/m^2)**	**Fuel Moisture Content (%)**	**Stomatal Conductance (mmol m^{-2} s^{-1})**	**Leaf Temperature (°C)**	**Chlorophyll Content (μmol/m^{-2})**
Multispectral Bands	R^2	0.36	0.34	0.46	0.47	0.30
	RMSE	36.28	7.88	108.83	1.10	196.09
	rRMSE	52.62	11.21	47.69	3.64	17.88
Multispectral Bands Thermal Band	R^2	0.47	0.43	0.80	0.65	0.43
	RMSE	32.95	7.31	66.54	0.89	177.22
	rRMSE	47.80	10.41	29.16	2.96	16.16
Multispectral Bands Thermal Band Thermal Indices	R^2	0.64	0.55	**0.84**	**0.74**	0.53
	RMSE	27.21	6.48	**59.23**	**0.76**	148.99
	rRMSE	39.47	9.22	**25.96**	**2.54**	13.16
Spectral Indices	R^2	**0.73**	**0.77**	0.77	0.67	**0.76**
	RMSE	**23.32**	**4.60**	71.30	0.86	**115.77**
	rRMSE	**33.82**	**6.55**	31.25	2.84	**10.56**

The highest accuracies achieved for each physiological indicator using optimal spectral variables selected by the DNN algorithm are highlighted in bold text.

5.3.3 Optimised regression models of taro crop water status physiological indicators integrating thermal and multispectral data

Optimised regression models were developed for all physiological indicators of crop water status by leveraging combinations of multispectral bands, the thermal band, thermal indices, and spectral indices. For instance, in estimating the equivalent water thickness of taro crops, the optimised regression model achieved an impressive R^2 of 0.95, a corresponding RMSE of 9.76 g/m^2 and an rRMSE of 14.15 (Figure 5.4a). The top-most influential predictor variables in the model development were MCARI, Thermal, CIrededge, TCARI, NIR, PRI and NDRE, in descending order of importance (Figure 5.5a). Similarly, the optimised model for the fuel moisture content yielded an exceptional R^2 of 0.94, with an RMSE of 2.33% and an rRMSE of 3.32% (Figure 5.4b). Variables such as MCARI, CIrededge, and PRI emerged as the most important predictors, indicating the significance of chlorophyll-based indices in estimating fuel moisture content (Figure 5.5b).

Furthermore, the optimised regression model for stomatal conductance exhibited an optimal R^2 of 0.96, with an RMSE of 29.34 mmol m^{-2} s^{-1} and an rRMSE of 12.86% (Figure 5.4c). Among the top predictors, variables such as CIrededge, thermal band, NRCT, PRI and MCARI (Figure 5.5c), emerged as significant contributors, which highlighted the importance of thermal data and certain spectral indices in predicting stomatal conductance. Similarly, the optimised regression model for leaf temperature produced an impressive R^2 of 0.95, an RMSE of 0.33°C and an rRMSE of 1.11% (Figure 5.4d). Again, thermal data, including the thermal band and NRCT, along with spectral indices, such MCARI, CIrededge, PRI, NDWI and NDRE, were among the most important predictor variables (Figure 5.5d). Lastly, the chlorophyll content estimation model achieved a strong R^2 of 0.91, with an RMSE of 71.51 µmol/m^{-2} and an rRMSE of 6.52% (Figure 5.4e). Despite being slightly lower than the other indicators, these results still demonstrate the model's effectiveness in estimating the chlorophyll content of smallholder taro crops. The top-most influential predictor variables were CIrededge, thermal, MCARI, PRI, NDRE, TCARI and NDVI, which indicates the significance of red-edge-derived spectral indices and the thermal band in predicting the chlorophyll content (Figure 5.5e).

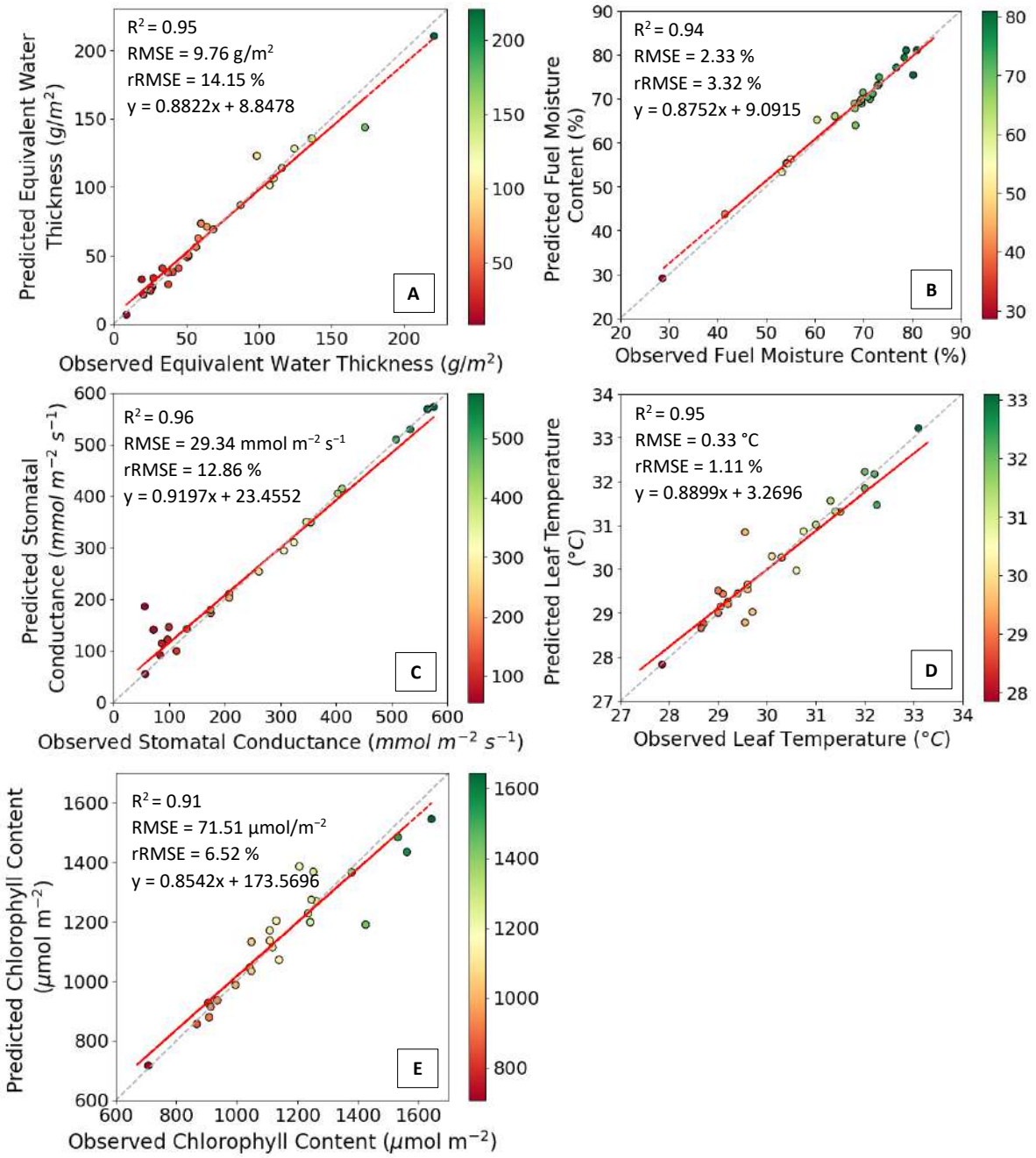


Figure 5-4: Relationship between measured and predicted a) equivalent water thickness, b) fuel moisture content, c) stomatal conductance, d) leaf temperature and e) chlorophyll content

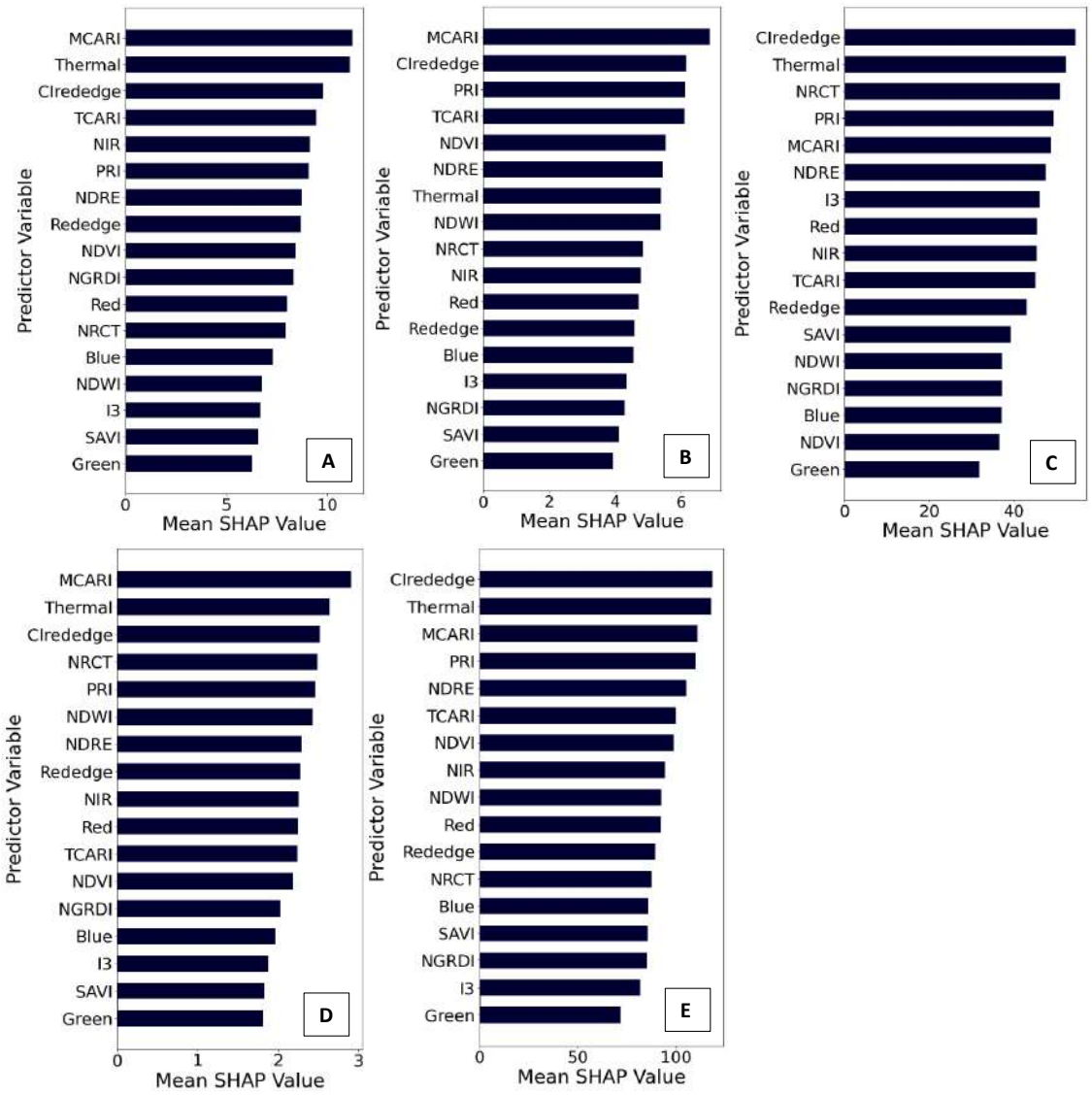


Figure 5-5 SHAP-generated variable importance values of predictor variables used in developing the optimal estimation models of a) equivalent water thickness, b) fuel moisture content, c) stomatal conductance, d) leaf temperature and e) chlorophyll content

5.3.4 Mapping the spatial distribution of the physiological indicators of the taro crop water status

The integration of multispectral and thermal bands with thermal and spectral indices presented the highest prediction accuracies. Figure 5.6 depicts the spatial distribution of the crop water status physiological indicators of smallholder taro crops. Notably, the southwestern area of the field consistently exhibited higher estimated values for the equivalent water thickness, fuel moisture content, stomatal conductance and chlorophyll content, which contrasts with the relatively lower values observed in the eastern up-slope regions. In addition, the leaf temperature patterns revealed distinct spatial variations, with lower temperatures being

prevalent in the southwestern parts and higher temperatures in the northern and southern extremities of the field (Figure 5.6).

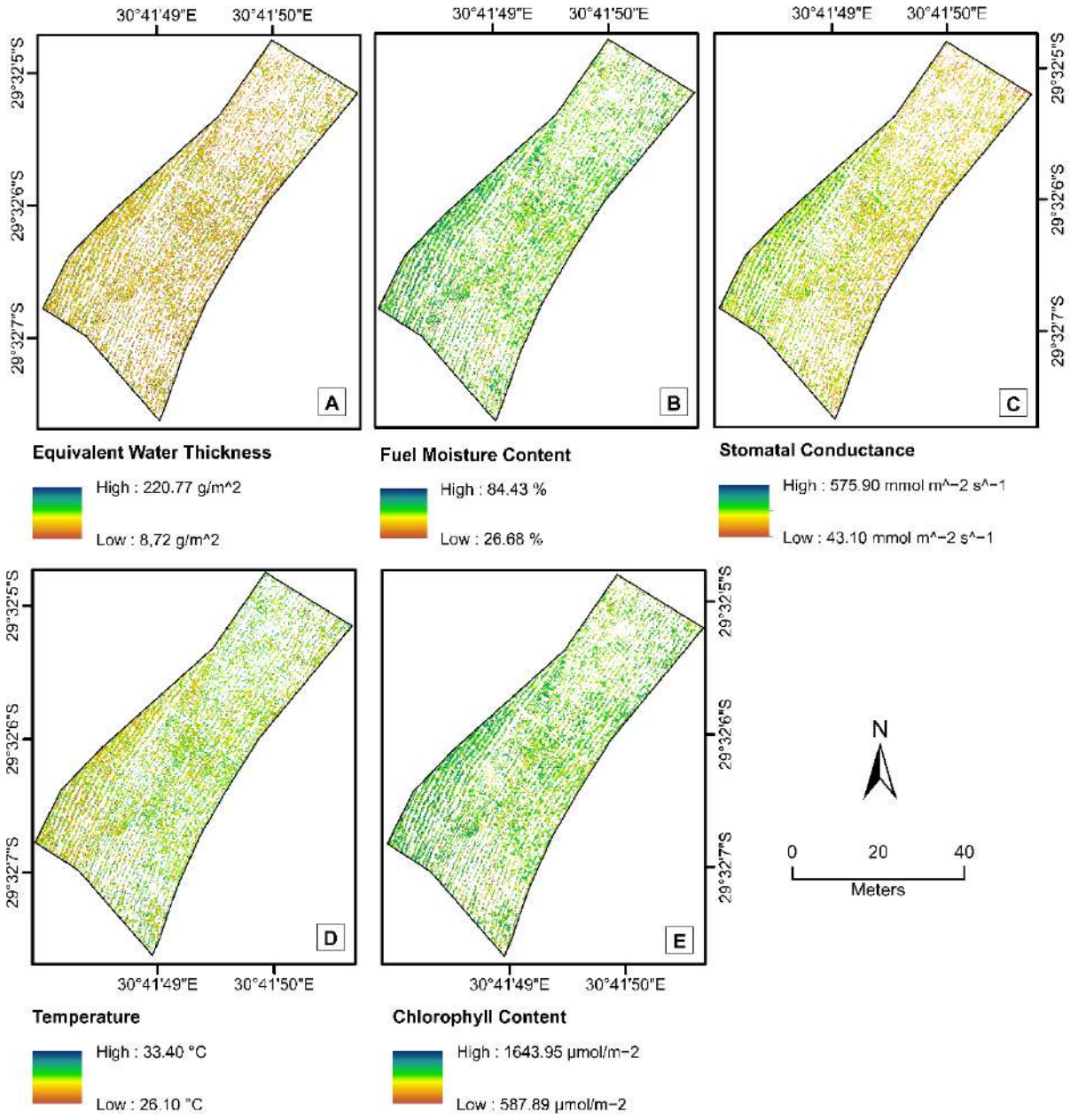


Figure 5-6 Spatial distribution of modelled a) equivalent water thickness, b) fuel moisture content, c) stomatal conductance, d) leaf temperature and e) chlorophyll content

5.4 Discussion

5.4.1 The performance of integrating multispectral and thermal bands with thermal and spectral indices

The results of this study illustrate that integrating multi-modal features, such as thermal and multispectral data, along with their derived spectral indices, leads to higher estimation accuracies, compared to using the single-modal variables alone. For instance, the combination

of thermal and multispectral data resulted in the most accurate predictions of the crop water status physiological indicators, as evidenced by the findings presented in Table 5.3 and Figure 5.5. This improvement can be attributed to the complementary nature of the thermal and multispectral data, which captured different aspects of the crop physiology and water status, thereby enhancing the overall predictive capability of the model (Sobejano-Paz *et al.* 2020). While thermal imaging provides valuable temperature information that is associated with the leaf moisture, pigmentation and canopy structure features, it is important to note that a wide range of environmental factors influence the crop water status beyond what thermal imaging alone can capture (Zhu *et al.* 2021). For example, during their evaluation of the water status variability using UAV data, Baluja *et al.* (2012) found that certain vegetation indices derived from multispectral data, as opposed to thermal images, exhibited a stronger correlation with the stomatal conductance and leaf water potential within a commercial rainfed vineyard. In addition, physiological indicators such as the equivalent water thickness, the fuel moisture content and the chlorophyll content were not extensively associated with the variations in the thermal section of the electromagnetic spectrum (Gerhards *et al.* 2019). This could be attributed to the fact that these physiological elements of the crop are more aligned with biomass accumulation (Quan *et al.* 2015, Liu *et al.* 2019, Sibanda *et al.* 2021). In contrast, spectral indices derived from multispectral data that focus on the crop greenness and health, complement thermal imaging by providing additional information on the plant's physiological processes that are related to its water status (Pineda *et al.* 2020, Kang *et al.* 2021). Therefore, the integration of various types of variables from multiple modalities could have effectively reduced the influence of noise, thereby enhancing the predictive capabilities of the estimation models (Sobejano-Paz *et al.* 2020, Zhai *et al.* 2023). The findings of this study are supported by those of Zhai *et al.* (2023), who integrated the spectral capabilities of thermal and multispectral data to optimise the estimation of the chlorophyll content of maize crops, which achieved an R^2 of 0.75 and an RMSE of 8.36%.

This study demonstrated that the equivalent water thickness of taro was optimally predicted by using a combination of spectral variables, including MCARI, Thermal, CIrededge, TCARI, NIR, PRI and NDRE, which resulted in an R^2 of 0.95 and an rRMSE of 14.15% (Figures 5.4a and 5.5a). Similarly, for fuel moisture content, the optimised model revealed MCARI, CIrededge and PRI as the most influential predictors, by achieving a high R^2 of 0.94 and an rRMSE of 3.32% (Figures 5.4b and 5.5b). Even though the thermal band and derived thermal indices did not emerge as the top-most predictor variables, it was not surprising, as the equivalent water thickness and fuel moisture content are closely linked to the leaf matter and accumulated above-ground biomass, which decreases under soil water stress (Zhang and Zhou 2019, Qi *et al.* 2020). These findings are supported by those of Gunawardena *et al.* (2015) who observed a poor negative correlation between the above-ground biomass and land surface temperature, with an R^2 value of 0.35, hence signifying the poor relationship between the thermal characteristics and the biomass-derived indices.

Moreover, the result highlighted the influential role of the thermal band and derived indices in predicting the stomatal conductance and leaf temperature in smallholder taro crops. Specifically, influential predictor variables, such as CIrededge, the thermal band, NRCT, PRI

and MCARI, yielded optimal predictions for stomatal conductance, by achieving an impressive R^2 of 0.96 and an rRMSE of 12.86% (Figures 5.4c and 5.5c). Similarly, for the leaf temperature, the thermal band, NRCT, MCARI, CIrededge, PRI, NDWI and NDRE emerged as crucial predictor variables, resulting in an optimal R^2 of 0.95 and an rRMSE of 1.11% (Figures 5.4d and 5.5d). The significance of the thermal band can be attributed to its ability to capture variations in the canopy temperature, which directly affect the stomatal conductance and water use (Ahmad *et al.* 2021, Wang *et al.* 2022). As the stomata regulate the release of water vapour from crops, changes in the stomatal conductance are reflected in the leaf temperature, which makes the thermal band a valuable predictor for this physiological indicator (Yu *et al.* 2015, Gerhards *et al.* 2019). In a similar study, Brewer *et al.* (2022) explored the use of integrating optical and thermal infrared UAV imagery to estimate the stomatal conductance and leaf temperature, by achieving R^2 values of 0.85 and 0.81, respectively, with the thermal band emerging amongst the top most influential variables. Furthermore, PRI has been shown to exhibit a strong correlation with the stomatal conductance, as it captures changes in photosynthetic activity that are associated with stomatal regulation (Magney *et al.* 2016, Yang *et al.* 2020). Therefore, the inclusion of thermal data and spectral indices, such as PRI, in estimation models enhances the ability to accurately estimate stomatal conductance, which offers insightful information about the water status of taro crops.

Finally, the findings revealed that the estimation of the chlorophyll content can be optimally achieved to an R^2 of 0.91 and an rRMSE of 6.52%, with thermal data and red-edge-based indices, including CIrededge, MCARI, PRI, NDRE and TCARI emerging as the most influential predictor variables (Figures 5.4e and 5.5e). Similarly, Brewer *et al.* (2022) found that smallholder maize farms attained a comparable correlation between the chlorophyll content and the red-edge-based indices, yielding optimal chlorophyll content prediction models, with R^2 values exceeding 0.75 throughout the growing season. Meanwhile, in a related study, Clevers and Gitelson (2013) identified red-edge-based indices, such as CIrededge, as optimal variables for precisely quantifying the chlorophyll content of maize and soybean, by achieving R^2 values of 0.92 and 0.94, respectively.

5.4.2 The performance of selected spectral variables in estimating the crop water status and physiological indicators of smallholder taro crops

The findings of this study revealed a suboptimal performance of UAV-derived multispectral bands in predicting the physiological indicators of crop water status. This limitation can be attributed to the limited spectral resolution of UAV-acquired multispectral bands, which constrain their ability to capture the intricate variations in the physiological indicators (Lu and He 2017, Barbedo 2019, Huang *et al.* 2019). These findings are consistent with those of Zhai *et al.* (2023), who investigated the efficacy of gradient boosting machine, Random Forest regression and ridge regression in estimating the chlorophyll content of maize crops, by solely using multispectral bands and achieving R^2 values ranging from 0.54 to 0.61. Consequently, achieving optimal predictions of the crop water status of taro necessitates the integration of additional data sources, such as thermal bands and spectral indices, in order to enhance the accuracy and reliability of the estimations.

Meanwhile, the findings demonstrated a notable increase in the prediction accuracies when the thermal band and thermal indices were integrated with multispectral bands, especially for stomatal conductance and leaf temperature of taro. The literature has detailed that thermal data and derived thermal indices offer valuable insights into crop water stress by detecting variations in the leaf temperature that are linked to water availability and the transpiration rate (Baluja *et al.* 2012, García-Tejero *et al.* 2016, Gago *et al.* 2017). For example, stomatal conductance, which regulates the leaf temperature dynamics, influences the closure of stomata under water stress, which reduces the water vapour loss and increases the leaf temperature (Yu *et al.* 2015, Wang *et al.* 2022). Therefore, by correlating the crop temperature and stomatal activity, thermal infrared data can provide rapid and robust data that are essential for quantifying the stomatal conductance of smallholder taro crops.

Moreover, spectral vegetation indices produced the highest model accuracies for estimating the equivalent water thickness, the fuel moisture content and the chlorophyll of smallholder taro crops. These findings are consistent with a substantial and expanding body of literature that demonstrates the efficacy of spectral indices that have been developed from water-sensitive portions of the spectrum for estimating the physiological indicators of a crop's water status (Zhang and Zhou 2015, Pasqualotto *et al.* 2018, Zhang *et al.* 2019, Sibanda *et al.* 2021). For instance, water stress-induced changes in the moisture content and chlorophyll levels are intricately linked to their leaf pigmentation and biochemical composition, which influence the spectral reflectance of the crop canopy, especially in the near-infrared and red edge regions of the electromagnetic spectrum (Pasqualotto *et al.* 2018, Parkash and Singh 2020). Therefore, by leveraging the unique spectral properties captured by these bands, vegetation indices provide enhanced sensitivity to changes in the chlorophyll content and canopy moisture content (Brewer *et al.* 2022, Ndlovu *et al.* 2024).

5.4.3 Implications of the study limitations and recommendations for future research

The study highlights the prospects for high-resolution UAV-acquired thermal remote-sensing techniques in assessing crop water status physiological indicators, particularly of NUS smallholder crops like taro. Demonstrating the effectiveness of thermal data when they are used in conjunction with multispectral bands highlights the importance of integrating various spectral variables for comprehensive agricultural monitoring. Understanding the dynamics of crop water status physiological indicators, including the equivalent water thickness, fuel moisture content, stomatal conductance, canopy temperature and chlorophyll content, enables the development of targeted adaptation strategies for smallholder farmers who are facing climate-related risks. Furthermore, the in-depth analysis of the crop water status of taro could contribute to agricultural frameworks that are aimed at redefining existing agricultural landscapes and diversifying cropping systems to include future smart NUS crops, such as taro, thus ensuring sustainable food production and food security.

However, this study acknowledges several limitations. Firstly, the occlusion effects, as a result of the top-down perspective of the UAV sensor, may obscure certain parts of the crop canopy due to overlapping leaves, and therefore, they may potentially influence the accuracy of the crop water status measurements. Therefore, future studies should explore the integration of

Light Detection and Ranging technology to account for the effects of occlusion, by providing a three-dimensional view of the canopy structure and its influence on crop water monitoring. Furthermore, while the interlacing of leaves was not a significant factor in this study, due to the discontinuous canopy structure of taro, further research is required to understand the effect of the intra-plant overlapping of leaves, particularly for leaf-level estimations of crop water status indicators. In addition, environmental and atmospheric attributes, such as solar radiation, can introduce interferences in the thermal reflectance properties, thereby affecting the quality of the data and reducing the accuracy of the estimated crop water physiological indicators. Therefore, further research on the refinement and incorporation of these aspects is required, in order to enhance the reliability and accuracy of the estimated crop water status indicators. Furthermore, in order to gain a comprehensive understanding of crop water status, further research is required to investigate the phenology and water status variations of smallholder taro crops across the growing season. In addition, there is a need to explore the integration of additional crop water-related indicators, such as the soil moisture content, to provide a more holistic understanding of the water dynamics in smallholder taro farming systems.

5.5 Conclusion

This study aimed to assess the applicability of UAV multi-modal thermal multispectral data and deep neural network techniques to assess the crop water status of smallholder taro crops. Considering the findings presented in this study, the following conclusions can be drawn:

- The physiological indicators of taro crop water content could be optimally estimated by using multi-modal UAV-acquired thermal and multispectral data, achieving impressive R^2 values greater than 0.91 and rRSME values less than 14.15% of the equivalent water thickness, fuel moisture content, stomatal conductance, canopy temperature and chlorophyll content.
- Integrating the thermal and multispectral bands with thermal and spectral indices resulted in the most optimal estimation models for an equivalent water thickness, fuel moisture content, stomatal conductance and canopy temperature.

This study serves as a stepping stone toward advancing precision agriculture practices for taro. Demonstrating the effectiveness of UAV multi-modal thermal multispectral remote sensing techniques paves the way for the more efficient and sustainable management of taro cultivation practices. Furthermore, the integration of advanced remote sensing technologies and deep learning algorithms allows for new avenues for enhancing agricultural monitoring and decision-making processes, that ultimately contribute to global efforts towards achieving food security and sustainable agricultural systems.

6 UAV-BASED ESTIMATION OF CANOPY CHLOROPHYLL CONTENT FOR TARO AND SWEET POTATOES ON SMALLHOLDER FARMS

6.1 Introduction

Overcoming the well-known contradiction between the rapid population increase and the limited food supply is a challenge that motivates an increasing number of studies on smart agriculture (Ren *et al.* 2024). Neglected and under-utilised crop species, also known as traditional crops, since they are grown in rural areas by marginalised communities, are some of the smart crops that are considered as alternative crops that have the potential to optimise food production and supply. The neglected taro and sweet potatoes are renowned for their adaptability to dryland agro-ecologies, they are highly nutritious and they offer greater opportunities in low-yielding regions (Mabhaudhi *et al.* 2017, Mugiyo *et al.* 2021). For several agronomic, genomic, socioeconomic and cultural beliefs, these crops have long been overlooked by conventional farming, but they are now gaining traction, due to their potential role in mitigating the risks associated with climate change in agricultural production systems (Padulosi *et al.* 2013). To optimise their production, there is a need for robust and effective phenotyping and monitoring techniques. One of the elements that is often associated with crop growth and health is the chlorophyll content concentration.

The chlorophyll content serves as a proxy for the general photosynthetic productivity of a crop population, and it is crucial for assessing the growth and nutritional health of individual crops (Daughtry *et al.* 2000, Yang *et al.* 2022). This is attributed to the biophysical pigment that is found in leaves and the biochemical processes that are involved in photosynthesis, also indicates the crop yield (Tahir *et al.* 2018, Falcioni *et al.* 2020, Petibon and Wiesenberg 2022). Insufficient chlorophyll levels often indicate that plants may be under stress, due to a water scarcity, a lack of nutrients, pests or diseases (Parkash and Singh 2020, Bhattacharya and Bhattacharya 2021). Therefore, tracking its concentration and variability in plants can assist in the early detection of stress, which allows for timely and targeted interventions to enhance the productivity and yield of a crop. The Leaf Chlorophyll Content (LCC) is the concentration of chlorophyll within a single leaf at a particular time, while the Canopy Chlorophyll Content (CCC) is the chlorophyll concentration over the entire canopy of plants over a given unit of area. This includes the chlorophyll content of each leaf, together with the spatial and structural arrangement of the plant's foliage (Sun *et al.* 2021, Narmilan *et al.* 2022). CCC provides a thorough assessment of the chlorophyll levels across the canopy scale; the measurement of CCC is influenced by the vertical arrangement (leaf area index) and concentration of the LCC (Ciganda *et al.* 2009, Li *et al.* 2020, Yang *et al.* 2022). While LCC is a fundamental component of CCC, it is not yet understood whether they could yield similar accuracies. Nevertheless, both the LCC and CCC are traditionally assessed by following visual inspections or invasive methods, such as a chemical analysis. These methods are not spatially explicit, are point sample-based, expensive and laborious. There is therefore a need for precise, consistent, spatially-explicit and high-throughput phenotyping technologies for assessing crop health by using its proxy chlorophyll content.

Non-destructive, precise, quick and consistent measurements of crop traits, such as the chlorophyll content, have become possible through remote sensing techniques (Miao *et al.* 2008, Duveiller *et al.* 2013, Nhamo *et al.* 2020). Remote sensing-based chlorophyll content phenotyping offers repeated measurements which are non-invasive, as well as time- and cost-effective, while it also covers both the local to regional scales. However, conventional remote sensing-based chlorophyll content phenotyping techniques often rely on freely-accessible satellite data (Nhamo *et al.* 2018, Berra and Peppa 2020). A significant limitation of publicly-accessible satellite data is its failure to meet the growing demand for the high spatial and temporal resolution information that is required in near-real-time chlorophyll content phenotyping in small and fragmented croplands, where sweet potatoes and taro are often grown (Psirofonia *et al.* 2017). In addition, UAVs can address the challenges related to spatio-temporal resolution; however, the expenses involved pose a significant barrier for smallholder farms.

In the recent past, Unmanned Aerial Vehicles (UAVs) have gained worldwide recognition as ground-breaking, economical and efficient precision tools for crop phenotyping. UAVs provide advanced analytics for image data transmission at Very High Resolutions (VHR) and have been found to effectively address the shortcomings of satellite imagery (Maes and Steppe 2019, Nhamo *et al.* 2020). In addition, UAVs possess the capability for versatile movement and a wide field of vision that is suitable for capturing the chlorophyll content variability of taro and sweet potatoes throughout their phenological stages in fragmented smallholder fields. Thus, accurately mapping and monitoring the chlorophyll content of sweet potatoes and taro crops by using UAV-acquired remotely-sensed data presents significant opportunities for providing farmers with an ultra-high spatial resolution and near real-time insights into a crop's health throughout the phenological cycle, which enables timely and targeted interventions, to optimise the productivity and yield.

Modelling the chlorophyll content of crops is often effectively conducted with Vegetation Indices (VIs) alongside strong machine learning techniques like Random Forest. Notably, chlorophyll-specific VIs have demonstrated greater significance than normalised Vis, in certain cases, as they encompass a range of combinations from bands that reflect strongly in the vegetation (Brewer *et al.* 2022). Examples of these indices include the Canopy Chlorophyll Content Index (CCCI) and the Modified Chlorophyll Absorption Ratio Index (MCARI), both of which demonstrate a significant correlation with the chlorophyll content in crops (Haboudane *et al.* 2002, Raper and Varco 2014). Meanwhile, Random Forest regression ensembles have been proven to be effective in assessing the crop chlorophyll content and health status (Abdel-Rahman *et al.* 2012, Han *et al.* 2019, Guo *et al.* 2020, Hassanijalilian *et al.* 2020). RF regression has consistently outperformed algorithms, such as the support vector machines, artificial neural networks and multiple regressions, in estimating crop attributes, including the chlorophyll content (Yao *et al.* 2013, Ramos *et al.* 2020). As a result, using data obtained from UAVs (such as vegetation indices and spectral bands), in combination and conjunction with a strong RF regression model, is anticipated to produce precise outcomes in measuring the chlorophyll content of taro and sweet potatoes, which serves as a health indicator on smallholder farms.

As taro and sweet potatoes play a crucial role in agrobiodiversity, as they are noted for their resilience against biotic pressures and their capacity to yield reliable harvests under challenging climatic conditions, it is essential to evaluate the utility of UAVs remotely-sensed data and its derivatives, in concert with the robust Random Forest in estimating their chlorophyll content as a proxy for their health on smallholder farms. Thus, this study assessed the capabilities of multispectral remote sensing data obtained from drones, in conjunction with Random Forest regression, in estimating the chlorophyll content of taro and sweet potato crops across selected phenological stages, to enhance the understanding of their health and efficiency in smallholder farming systems. Therefore, the specific objectives of the study were: i) to comparatively assess the relative contribution of bands, as well as the traditional and chlorophyll content-based vegetation indices, in estimating chlorophyll content variations of taro and sweet potatoes at different phenological stages, by using UAV-derived data; and ii) to compare the estimation accuracies of the LCC and CCC that were derived using UAV-derived multispectral remotely-sensed data, in conjunction with Random Forest regression.

6.2 Materials and Methods

6.2.1 Foliar chlorophyll content sampling strategy

The study area is detailed in Chapter One. To enhance the sampling process, a polygon encompassing all the experimental fields was created in Google Earth Pro and imported into a GIS. Subsequently, 50-point locations were established within each polygon of the experimental field plots by using a stratified random sampling method to identify the sampling points. The points were subsequently uploaded into a Trimble hand-held Global Positioning System (GPS) that offers sub-meter precision. The GPS was utilised to find and navigate to the sampling locations in the field. At every site, a plot measuring approximately one square meter was defined. Sweet potato and taro plants were selectselected,heir chlorophyll content was measured.

6.2.2 Measuring leaf chlorophyll content of taro and sweet potatoes

The Soil–Plant Analysis Development (SPAD) 502 DL Plus chlorophyll meter (SPAD chlorophyll meter, hereafter) was used to non-destructively measure the foliar chlorophyll content in the sweet potato and taro leaves. A SPAD chlorophyll meter was selected for its portability, affordability, non-destructive nature, quicker process, as well as its ease of use, relative to destructive techniques, such as the Kjeldahl digestion and Dumas combustion (Sibanda et al., 2020). The SPAD chlorophyll meter provides immediate measurements in ‘unitless SPAD values’ that range from 1 to 100. Ten chlorophyll meter readings were measured per plot. The SPAD readings were obtained roughly at the midpoint of the leaves, ensuring that the midribs were avoided (Lang et al., 2019). During the chlorophyll measurements, the chlorophyll meter was protected from direct sunlight. The chlorophyll readings were taken at the same sites as where the remotely-sensed data was acquired. As a result, the SPAD measurements were converted to chlorophyll content by using the well-known universal equation established by Markwell et al. (1995):

$$Chlorophyll\ Content = 10^{M^{0.265}} \qquad \text{units: } \mu mol\ m^{-2} \quad (1)$$

where Chl represents the total quantity of chlorophyll in μmol m² and S denotes the dimensionless value of the SPAD measurements. Given the molecular weights of Chl *a* and Chl *b*, conversions were performed from μmol m^{-2} (Chl *a*) to μg cm^{-2} (Chl *b*). The conversion of SPAD values was conducted after capturing the data in Microsoft Excel as a spreadsheet, along with each coordinate of the sampling point. After deriving the chlorophyll content, the spreadsheet was imported into QGIS and overlaid with drone imagery to extract spectral signatures.

6.2.3 Leaf area index and canopy chlorophyll content

The Leaf Area Index (LAI) reflects the nutritional status of the crop population and is directly linked to the crop health and yield. The LAI of the sample plants were measured by using the LI COR LAI-2200 C Plant Canopy Analyser (LI-COR Biosciences, Lincoln, Nebraska, USA). The LAI was measured in the crops within the identified plot. When measuring the LAI of sweet potato and taro crops within the selected plots, a single unobstructed measurement of light above the canopy, as well as four readings measured below the canopy at evenly spaced intervals within the plot, were measured by using the instrument's fish-eye optical sensor at multiple zenith angles, to account for variability. It calculated the LAI by comparatively assessing the four below-canopy light measurements with an unobstructed reference light reading measured above the canopies. Subsequently, the Canopy Chlorophyll Content (CCC) of both the sweet potatoes and taro was calculated by multiplying the LAI and LCC values, based on the following formulae:

$$\text{CCC} = \text{Leaf Chlorophyll Content (LCC)} \times \text{LAI} \qquad (2)$$

where CCC is the Canopy Chlorophyll Content, and LAI is Leaf Area Index estimates measured using the LI COR LAI-2200 C Plant canopy analyser, as in Kganyago *et al.* (2022). The CCC estimates were also added to the table with coordinates, LCC and LAI estimates. These were then transformed into a point map, which facilitated the extraction of the spectral signatures that were used in predicting both the CCC and LCC, using the reflectance data acquired by using the UAVs, based on the approach detailed in the preceding section.

6.2.4 Acquisition and processing of remotely sensed data using an UAV

The DJI Matrice 300 series (M300), along with the MicaSense Altum imaging sensors, were utilised to obtain spectral imagery of the taro and sweet potato fields reviewed in this study. The Altum camera features a thermally-calibrated sensor that is equipped with five spectral channels, which allow for the capture of reflectance from both the visible and non-visible light ranges, including the blue (475 nm), red-edge (717 nm), red (668 nm), green (560 nm), thermal (8 000-14 000 nm) and NIR (840 nm) wavelengths. The primary benefit of this imaging system is its capability to simultaneously collect synchronised multispectral and thermal data. The study shapefile was generated by using Google Earth Pro and it was subsequently loaded into

the remote control of the M300, to assist in creating a light plan for the drone. To compensate for the varying light conditions, an automatically-calibrated reflectance panel was utilised before and after the flight, which relied on the known reflectance values to perform radiometric calibration of the Altum sensor. A flight mission was carried out automatically at an altitude of 100 m, which achieved a ground sampling distance of 9.6 cm per pixel, with an image overlap of 80%. The weather conditions during the fights were conducive for UAV operations, as detailed in (Buthelezi *et al.* 2023) and (Ndlovu *et al.* 2024). The images were uploaded, ortho-mosaiced and pre-processed by using Pix4D Fields photogrammetry software, based on the calibration images. The pre-processed images were then orthorectified in QGIS to less than a pixel. Subsequently, all images were overlaid with the point map containing the LCC and CCC estimates. The extracted spectral signatures were used to conduct the RF regression algorithm, as detailed in the preceding section.

6.2.5 Estimation of chlorophyll content by using the Random Forest

The extracted sample data were randomly divided into a training set (70%) and a validation set (30%). The training set was utilised to develop the Random Forest models, whereas the testing datasets were utilised to assess the accuracy of the generated models. The RF regression was utilised to forecast the chlorophyll content. RF was selected and utilised in this study because it is a non-parametric method that is reliable and effective, and it can be easily parameterised and implemented with ease (Liakos *et al.* 2018, Yue *et al.* 2018). Random Forest derives its strength from bootstrapping in selecting training data points to construct decision trees, in order to predict the chlorophyll content (Sibanda *et al.* 2021). The algorithm used supervised learning to construct decision trees by recursively splitting the data, based on the most informative spectral features (bands, vegetation indices). Each tree was trained on a random subset of the data, which assisted in circumventing overfitting issues, while optimising generalisations. The results were then aggregated through averaging, which enabled it to effectively handle the high dimensionality that is associated with bands and their derived vegetation indices. The RF regression is well-known for its capacity to deliver high accuracy in predictions, and it is easy to apply (Sibanda *et al.* 2021). The effectiveness of the RF model was optimised by the careful adjustment of its hyperparameters, namely, *Ntree*, which indicates the number of decision trees to create, and *Mtry*, which is the count of prediction elements that are evaluated for the optimal split during tree construction (Ndlovu 2021). Instead of using default values, the *Ntree* and *Mtry* parameters were established after several iterations, in order to identify the most effective settings for predicting the chlorophyll content of taro and sweet potatoes (Singhal *et al.* 2019).

6.2.6 Accuracy assessment

The accuracy assessment of the chlorophyll prediction models utilised 30% of the designated test datasets. The evaluations in this study were based on the Root Mean Squared Error (RMSE), the Relative Root Mean Squared Error (RMSE%) and the coefficient of determination (R^2). The RMSE and R^2 were computed to quantify the predictive accuracy. The relative RMSE (rRMSE) was derived by dividing the RMSE by the average measured value, as depicted by Despotovic et al. (2016) and as represented in the following formula:

$$rRMSE = (sqrt (sum ((y\_predicted - y\_actual)^2) / n)) / (mean(y\_actual)) \quad (3)$$

RMSE represents a commonly-adopted error metric that is suitable for the general evaluation of numerical predictions. The RMSE error metric is widely used and is regarded as a good general-purpose error metric for numerical predictions. In a regression model, the coefficient of determination (R) determines the goodness of fit. Mathematically, it is the square of the multiple correlation coefficient relating to the predicted and actual values. Importantly, the R^2 can only reliably characterise the fit when the measurements that are obtained are devoid of any measurement error.

6.2.7 Estimation of chlorophyll content of taro and sweet potatoes

The literature suggests that vegetation indices augment the strength of models in predicting crop and plant attributes (Haboudane 2004, Panda *et al.* 2010, Farella *et al.* 2022). However, there is no consensus. In estimating the chlorophyll content for both the taro and sweet potato crops, the initial phase involved utilising spectral reflectance derived from UAV bands (blue, green, red, red-edge and NIR) (Table 6.1). Then, the traditional vegetation indices, which include the Normalised Difference Vegetation Index (NDVI), the Normalised Difference Green/Red Index (NGRDI), the Normalised Difference Red-edge Index (NDRE) and the Optimised Soil Adjusted Vegetation Index (OSAVI), were calculated. These traditional indices were chosen because of their ability to assess the canopy structure and overall vegetation health. Consequently, the chlorophyll-based vegetation indices were also chosen and calculated, based on their sensitivity to the chlorophyll variations. The final modelling stage involved combining all the data (bands and vegetation indices) in predicting the chlorophyll content of both taro and sweet potatoes (Table 6.1). The final analysis stage compared whether the modelling LCC differed from the modelling CCC of the taro and sweet potato crops. This evaluation focused on comparing the performance, accuracy and parameter sensitivity of predictive models for LCC, based on the SPAD measurements and UAV spectral data, and for CCC, which was derived from the integration of LCC and LAI estimates. More specifically, the analysis tested whether the final prediction model metrics, including R^2, RMSE and RRMSE, showed significant differences ($p < 0.05$) between the LCC and CCC predictions of the taro and sweet potato crops.

Table 6-1 Stages followed in estimating chlorophyll content

Analysis stage	**Spectral Variables utilised**
1	Bands
2	Traditional Vegetation Indices
3	Chlorophyll vegetation indices
4	Combined data (bands and vegetation indices)
5	Comparison of model accuracies for LCC and CCC

6.3 Results

6.3.1 Descriptive analysis

The minimum chlorophyll concentration that was observed was during the late vegetative growth phase, and it measured 30.3 µmol m^{-2} for both sweet potatoes and taro (Table 6.2). The peak chlorophyll content was noted in the early development phase, with values of 57.1 µmol m^{-2} for sweet potatoes and 58.4 µmol m^{-2} for taro. The overall median and mean chlorophyll concentrations for sweet potatoes and taro across the different growth stages were 54.96 µmol m^{-2} and 55.23 µmol m^{-2}, respectively. The mean standard deviation of the chlorophyll measurements was 7.1, which signified a considerable deviation of the values from the mean, which stands at 54.96.

Table 6-2 Descriptive statistics of the leaf chlorophyll content of sweet potatoes and taro

Date in 2021	LCC (µmol m-2)	Minimum	Maximum	Mean	Median	St deviation
11th February	Sweet potatoes	38.1	57.1	47.44	44.4	6.09
	Taro	33.5	58.4	46.36	46	7.37
12thApril	Sweet potatoes	32.4	53.2	41.6	39.9	5.72
	Taro	24.4	57.4	44.43	43.5	7.64
14th May	Sweet potatoes	34	56.1	46.07	45.6	6.11
	Taro	35.4	61.8	47.07	45.6	6.89

Table 6-3 Descriptive statistics of the canopy chlorophyll content of sweet potatoes and taro

Date in 2021	CCC (µmol m-2)	Minimum	Maximum	Mean	Median	St deviation
11th February	Sweet potatoes	44.4	75.1	64.2	64.45	8.38
	Taro	37.2	64	50.81	51.05	6.01
12thApril	Sweet potatoes	30.03	74.7	56.62	58.3	8.41
	Taro	49.2	71.6	58.17	56.85	6.07
14th May	Sweet potatoes	33.8	58.5	43.31	42.4	5.33
	Taro	30.3	74.7	56.62	58.3	8.41

6.3.2 Comparison of the LCC and CCC estimations of spectral variables in taro and sweet potatoes

For taro, the traditional vegetation indices consistently exhibited a significantly higher mean RMSE, compared to the bands, chlorophyll indices and the combined datasets. This suggests that the traditional VIs did not enhance the prediction of LCC. While the chlorophyll vegetation indices produced a relatively lower mean RMSE than the traditional vegetation indices, their mean RMSE remained significantly greater than what was obtained from the combined datasets and the bands-only dataset (Figure 6.1(b)). These findings indicate that the bands-only and combined datasets significantly enhanced the accuracy of LCC estimation, and that they outperformed the traditional and chlorophyll vegetation indices when evaluated with the RF algorithm (Figure 6.1).

Similarly, for sweet potatoes, the bands-only dataset exhibited a significantly higher mean RMSE, compared to the traditional vegetation indices, the chlorophyll vegetation indices and the combined datasets. This implies that relying solely on bands did not sufficiently improve the LCC estimation, when using the RF algorithm. However, the combined dataset yielded a relatively lower mean RMSE than the bands-only dataset, although its performance was relatively lower than that of the traditional and chlorophyll vegetation indices (Figure 6.1). This further suggests that chlorophyll vegetation indices played a more substantial role in improving the model's accuracy for the estimation of the LCC, by surpassing the combined datasets, bands and traditional vegetation indices (Figure 6.1).

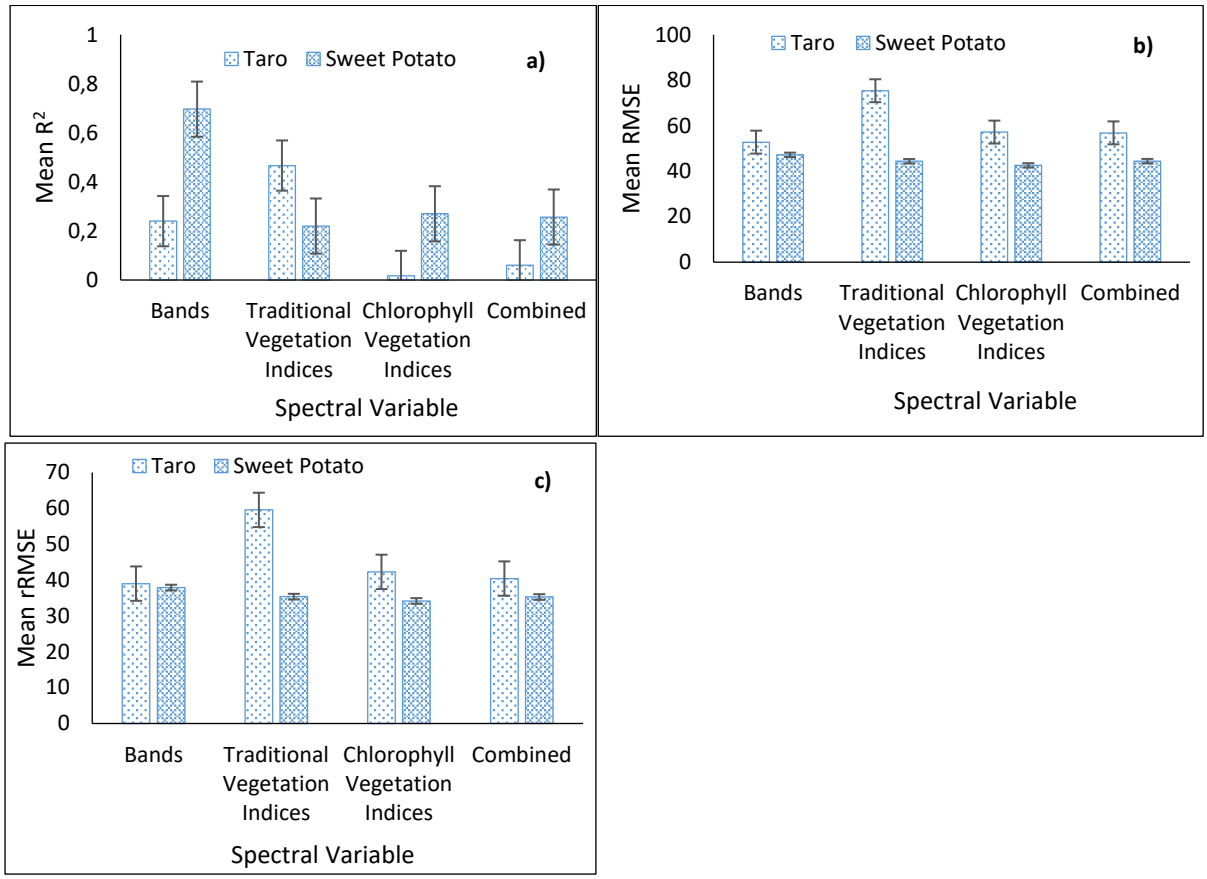


Figure 6-1 The comparative effectiveness of bands, indices and integrated datasets in estimating the CCC based on (a) average R-squared values, (b) typical RMSEs, and (c) mean rRMSEs

Regarding the estimation of the CCC of taro, the traditional vegetation indices resulted in a significantly higher mean RMSE, compared to the bands-only, chlorophyll vegetation indices and combined datasets. This reinforces the initial finding that the traditional vegetation indices were ineffective in improving the estimation of CCC by using the RF algorithm. While the bands-only yielded a lower mean RMSE than the traditional vegetation indices, it still performed less than the chlorophyll vegetation indices and combined datasets in accuracy (Figure 6.2(b)). The combined datasets demonstrated the greatest improvement, achieving a

significantly higher model accuracy than the bands-only dataset, as well as the traditional and chlorophyll vegetation indices (Figure 6.2).
For sweet potatoes, the traditional vegetation indices exhibited similarly high mean RMSE values, compared to the bands-only, chlorophyll indices and the combined datasets. This underscores their inferiority in enhancing the CCC estimation through the RF algorithm. The chlorophyll vegetation indices, while producing a lower mean RMSE than the traditional VIs, were surpassed by the bands-only and combined datasets. Among all the datasets, the combined dataset achieved the most substantial improvement in the model accuracy for CCC estimation, and it outperformed the bands-only dataset, traditional vegetation indices and chlorophyll indices (Figure 6.2).

These findings highlight the varying effectiveness in the performance and sensitivity of various spectral variables in estimating both the LCC and CCC across the selected growth stages of taro and sweet potatoes. The combined datasets consistently demonstrated a superior performance, particularly for CCC estimation, while the chlorophyll vegetation indices showed a notable use in the LCC estimation for sweet potatoes. Traditional vegetation indices were consistently outperformed across both crop types and target variables.

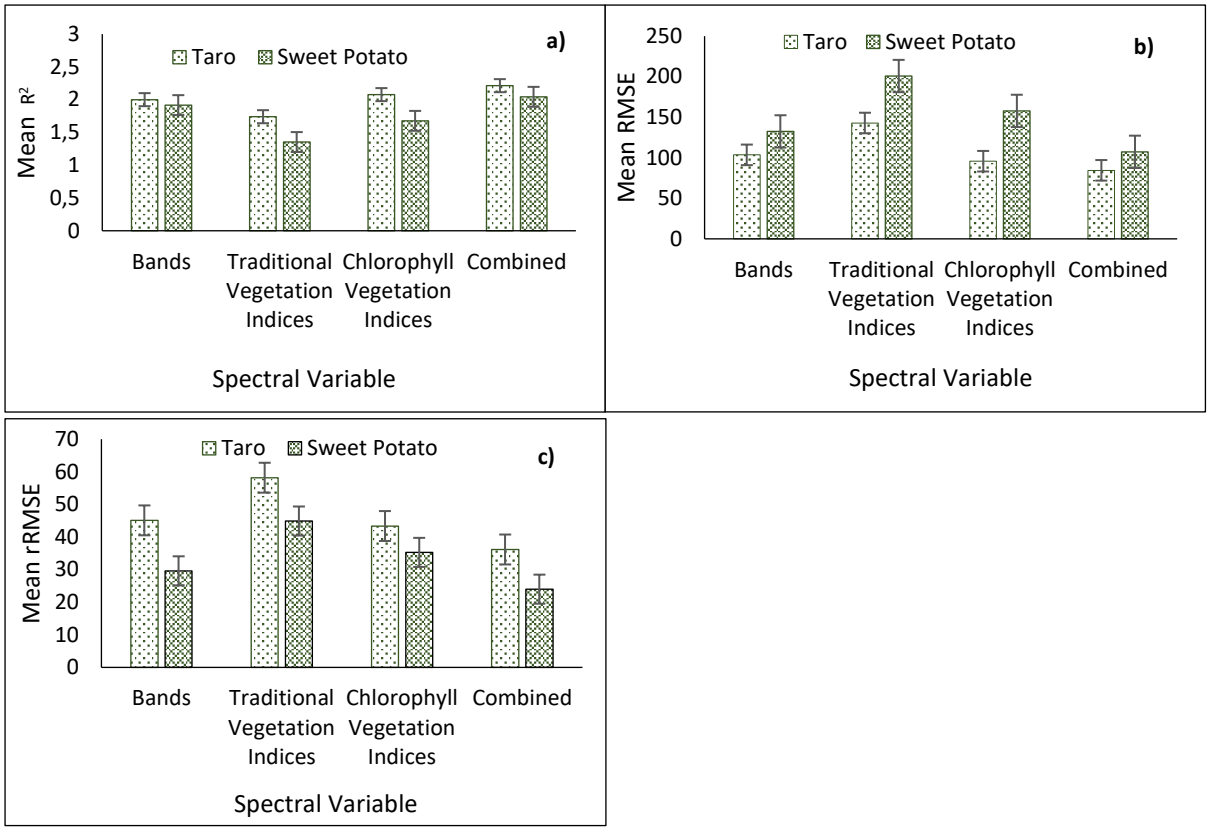
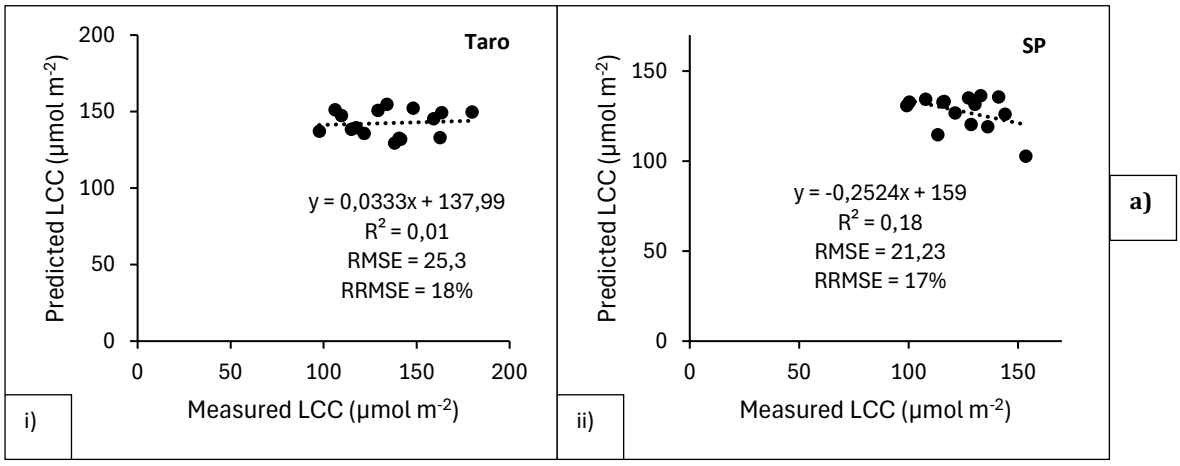


Figure 6-2 The comparative effectiveness of bands, indices and integrated datasets in estimating the CCC based on (a) average R-squared values, (b) typical RMSEs, and (c) mean rRMSEs

6.3.3 Optimised regression models of taro and sweet potato canopy chlorophyll content at different phenological stages

For sweet potatoes, LCC prediction was the most accurate during the late vegetative development phase, achieving an R^2 of 0.23, an RMSE of 13.6 µmol m^{-2} and an rRMSE of 11% (Figure 6.3c(ii)), based on SR1, CIRE, Ci_{green}, NDVI, RRI1, SAVI, BLUE, GCI and EXG as optimal prediction variables, according to their significance (Figure 6.4c(ii)). In contrast, during the mid-vegetative growth stage, the model's accuracy decreased to an RMSE of 18.7 µmol m^{-2}, an R^2 of 0.0 and an rRMSE of 15% (Figure 6.3b(ii)) based on SRI, NGRDI, $NDVI_{rededge}$ and CIRE as the key variables (Figure 6.4b(ii)). The early vegetative stage produced the lowest RMSE of 21.23 µmol m^{-2} with an R^2 of 0.18, but it maintained a moderate rRMSE of 17% (Figure 6.3a(ii)) for sweet potatoes. For taro, the best LCC estimation was attained during the mid-vegetative development phase, with an R^2 of 0.05, an RMSE of 23.3 µmol m^{-2} and an rRMSE of 17% (Figure 6.3b(i)), based on ExG, $NDVI_{rededge}$, NIR, CCI, GCI, OSAVI, EVI, RRI1 and SRI, also ranked in order of importance (Figure 6.4b(i)). During the early vegetative stage, the model exhibited an RMSE of 25.3 µmol m^{-2}, an R^2 of 0.01 and an rRMSE of 18% (Figure 6.3a(i)) based on MCARI, CCCI, RED, SRI and Red-edge as the key variables (Figure 6.4a(i))., Comparatively, the late vegetative growth stage showed a similar performance with an RMSE of 23.27 µmol m^{-2}, an R^2 of 0.05 and an rRMSE of 17% (Figure 6.3b(i)) based on ExG, $NDVI_{rededge}$, NIR, CCCI and GCI as optimal variables, in order of importance (Figure 6.4b(i)). Most of the LCC models exhibited a very low R^2, which may be indicative of the fact that the models were not sufficiently capturing the underlying patterns of taro and sweet potatoes at a leaf level.



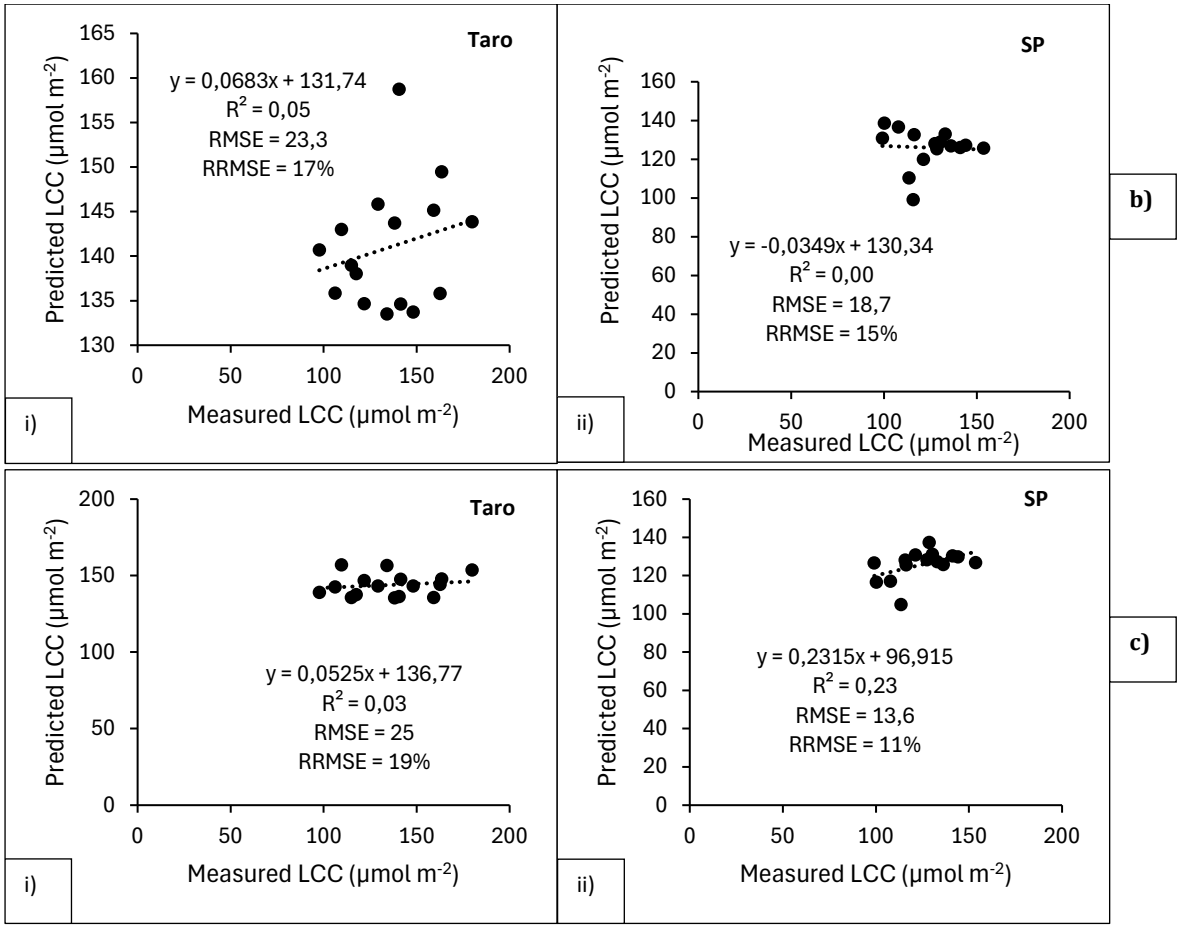
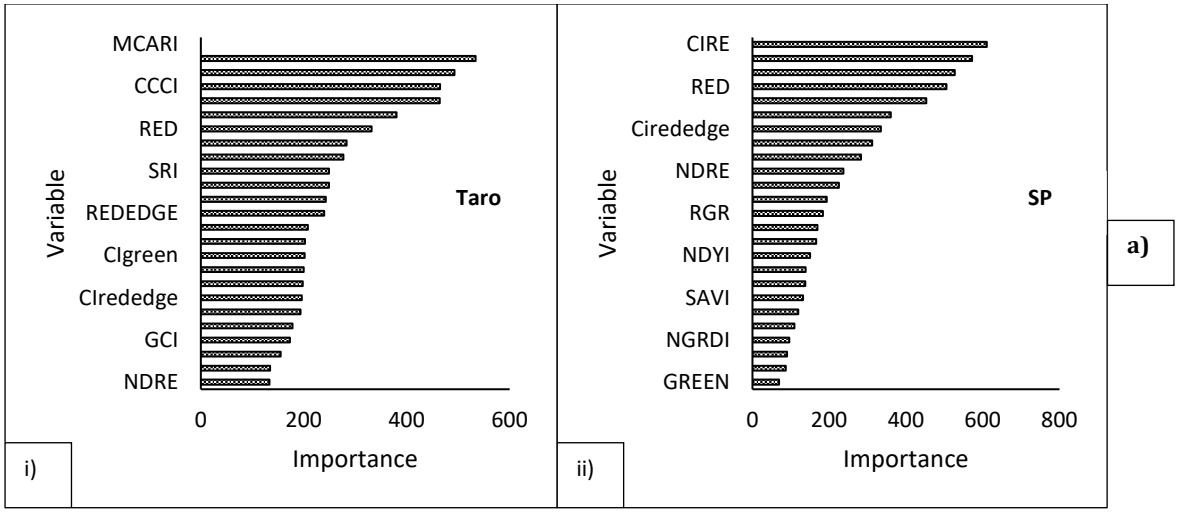


Figure 6-3 Regression models illustrating the connections between the observed and forecasted LCC throughout the taro and sweet potato phenological cycles: (a) 11th February 2021, (b) 12th April 2021, and (c) 14th May 2021



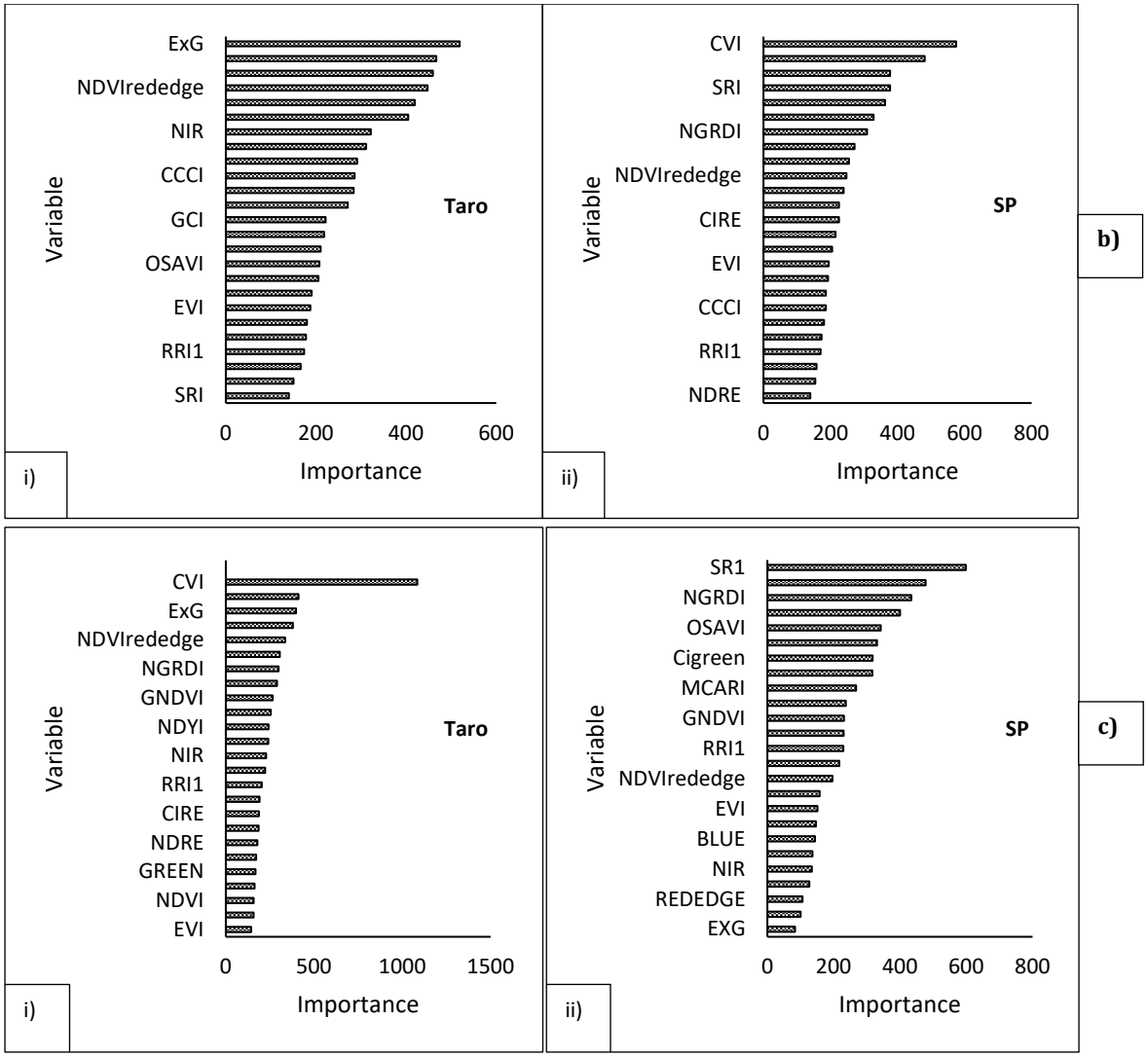


Figure 6-4 Variable importance scores of optimal LCC bands and VIs across the phenological cycle: (a) 11th February 2021, (b) 12th April 2021, and (c) 14th May 2021

When evaluating the optimised CCC estimations, both sweet potatoes and taro demonstrated optimal prediction accuracies during the mid-vegetative growth stage. For sweet potatoes, an R^2 of 0.91, an RMSE of 42.3 µmol m^{-2} and an rRMSE of 9.2% (Figure 6.5b(ii)) based on $Ci_{rededge}$, RED, $NDVI_{rededge}$, Rededge, CIRE, and Ci_{green} as the most important variables (Figure 6.6b(ii)). Similarly, for taro, an R^2 of 0.96, an RMSE of 15.8 µmol m^{-2} and an rRMSE of 14.5% were recorded (Figure 6.5b(i)) by using $CI_{rededge}$, GREEN, NIR, CIRE, and RGR as optimal variables, ranked by their importance (Figure 6.6b(i)). During the early vegetative stage, CCC prediction accuracy decreased slightly, with sweet potatoes achieving an R^2 of 0.85, an RMSE of 48.2 µmol m^{-2} and an RRMSE of 10.8%, based on REDEDGE, $Ci_{rededge}$, NDRE, NIR and Ci_{green} as the optimal prediction features, also in order of importance. Meanwhile, taro exhibited an R^2 of 0.95, an RMSE of 49.1 µmol m^{-2} and an rRMSE of 15% (Figure 6.5a) using ExG, $NDVI_{rededge}$, CCCI, RED, CVI, NGRDI and RGR, in order of importance (Figure 6.6b(ii)). The late vegetative growth stages for both crops had the lowest CCC prediction accuracies. For sweet potatoes, this stage had an R^2 of 0.85, an RMSE of 53.6 µmol m^{-2}, and an rRMSE of

12% using NIR, OSAVI, NDRE, RED and CIRE. On the other hand, taro exhibited an R^2 of 0.91, an RMSE of 58.7 µmol m^{-2} and an rRMSE of 18.5% (Figure 6.6c) based on $CI_{rededge}$, OSAVI, CIRE, GNDVI and Ci_{green} (Figure 6.6a). These findings indicate that, while both the LCC and CCC predictions show variability across the growth stages, the mid-vegetative stage generally provides the most reliable estimations, particularly for CCC. This underscores the importance of growth-stage-specific modelling in enhancing the precision of the chlorophyll content assessments for sweet potatoes and taro. Overall, all CCC prediction models exhibited optimal accuracy metrics.

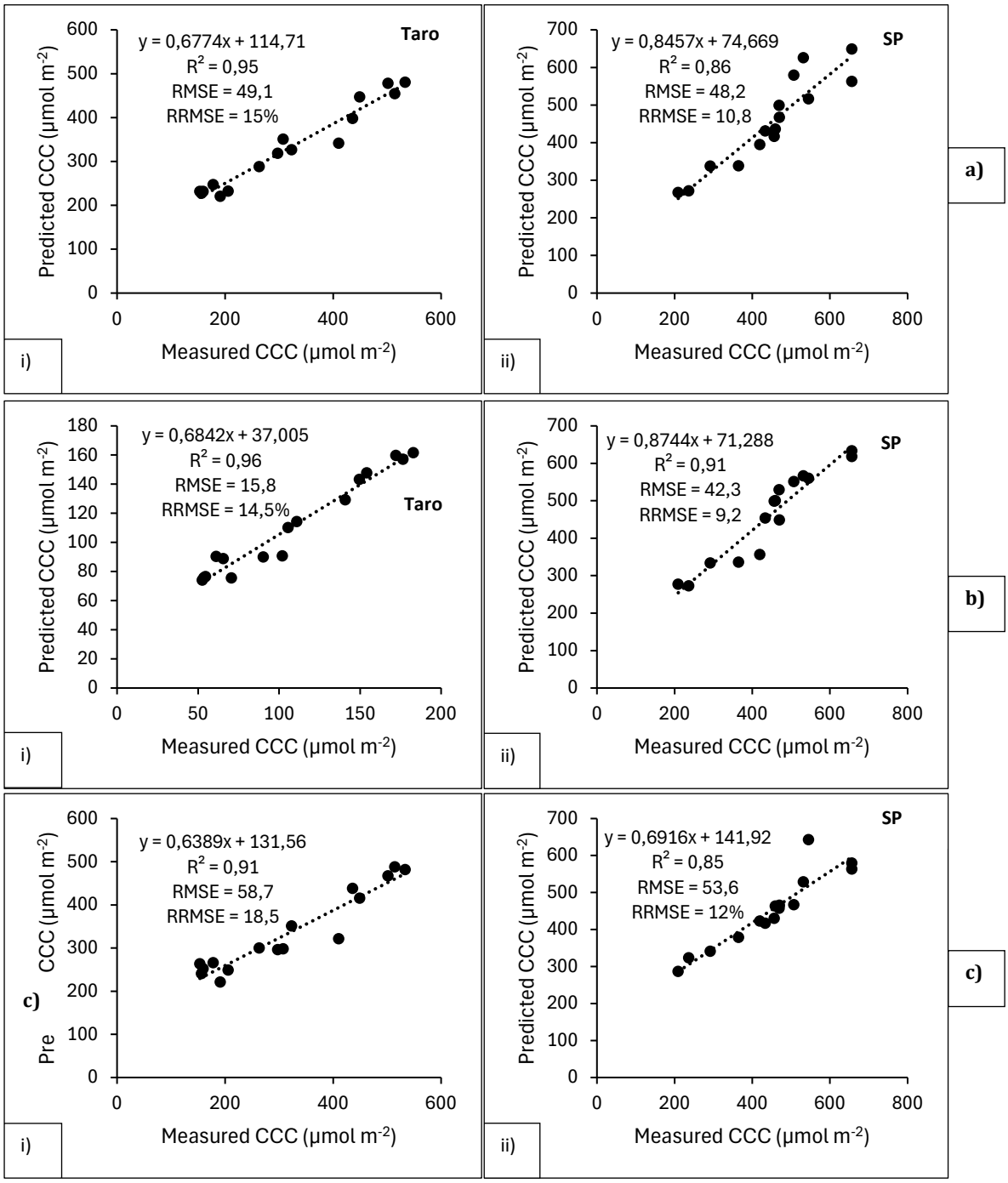


Figure 6-5 Regression models illustrating the connections between the observed and forecasted CCC throughout the taro and sweet potato phenological cycles: (a) 11th February 2021, (b) 12th April 2021, and (c) 14th May 2021

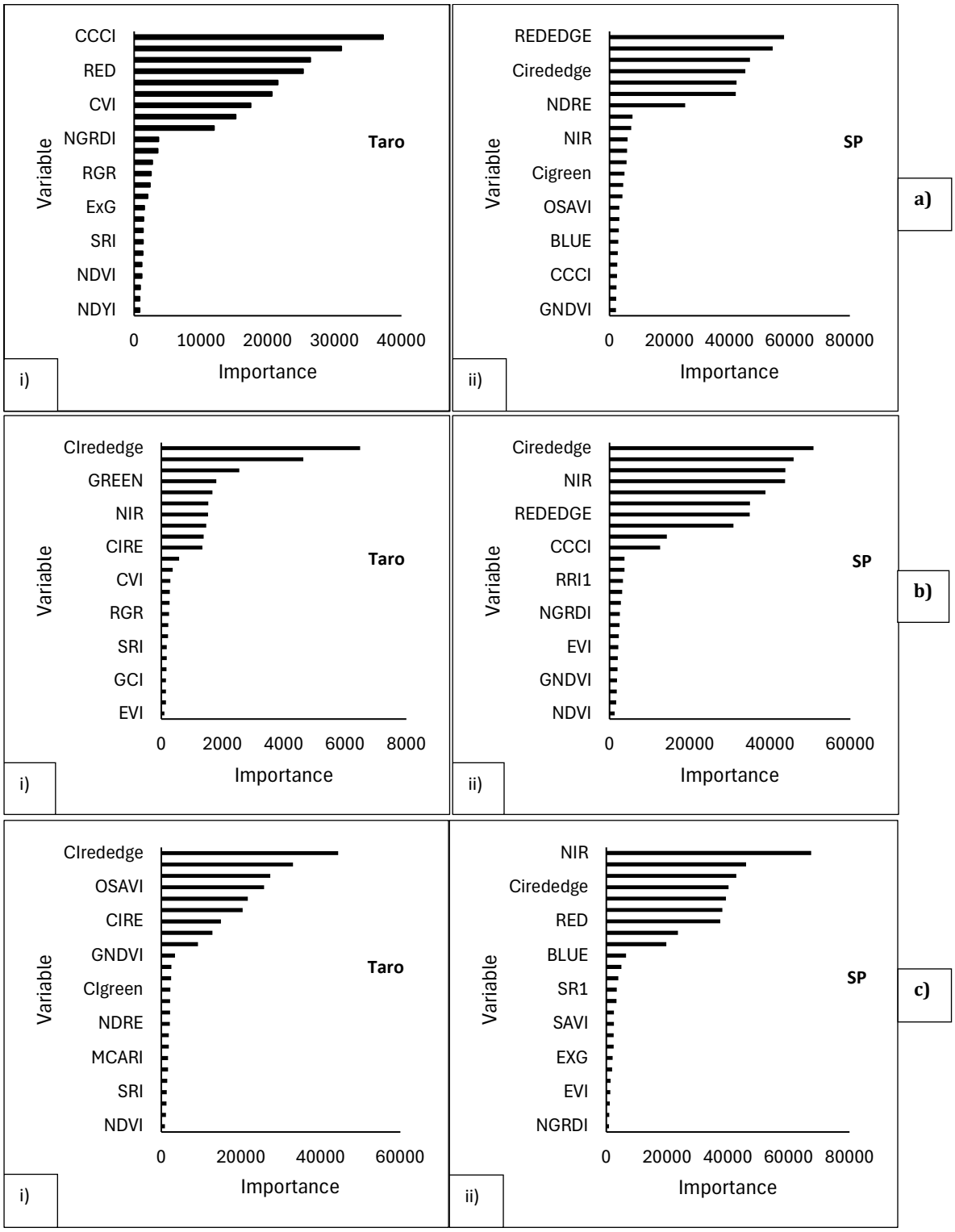


Figure 6-6 Variable importance scores of optimal CCC bands and VIs across the phenological cycle: (a) 11th February 2021, (b) 12th April 2021, and (c) 14th May 2021

6.3.4 Comparing the leaf and canopy chlorophyll content estimations

In the estimation of the chlorophyll content at both the leaf (LCC) and canopy (CCC) levels, CCC estimations were notably more precise. They recorded a significantly higher R^2 value, nearing 0.9, in contrast to the significantly lower R^2 of approximately 0.1 for the LCC estimations. The error bars associated with the CCC R^2 demonstrated a relatively minor variability, while those for the LCC R^2, although minimal, consistently indicated low values. Moreover, the average rRMSE for CCC estimation was approximately 12%, whereas the LCC estimation had a higher rRMSE of about 16% (Figure 6.7).

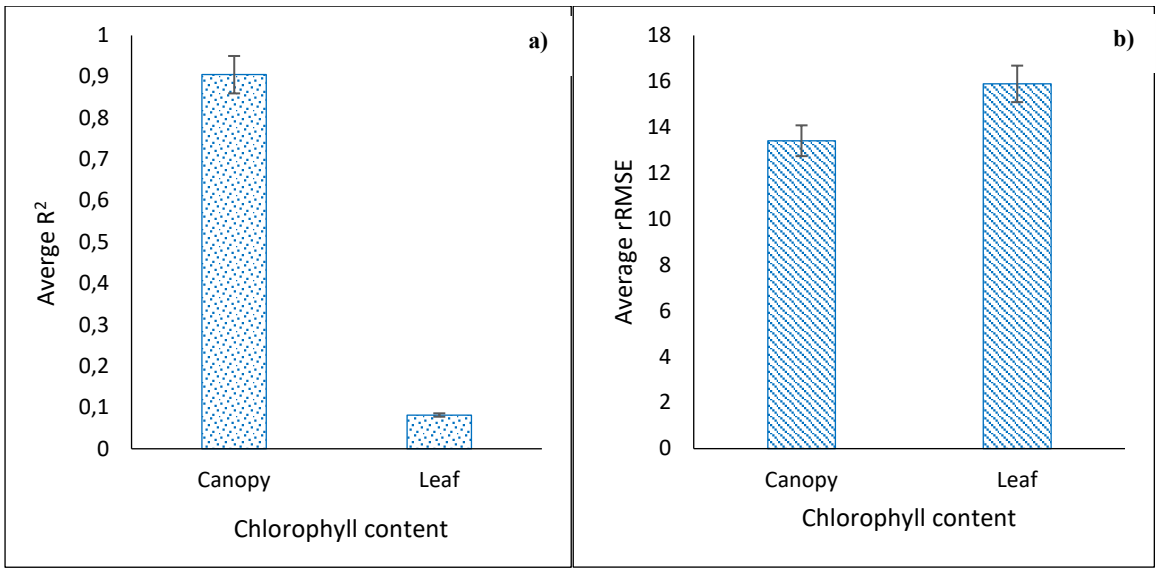


Figure 6-7 Comparison between the accuracies (a) average R2 and (b) average rRMSE derived in estimating LCC and CCCC

6.3.5 Spatial distribution of the chlorophyll content of the sweet potato and taro crops at selected growth stages

The estimated chlorophyll levels varied between 37 µmol m-2 and 133 µmol m-2 for taro (Figure 6.8a), and 71 µmol m-2 to 143 µmol m-2 for sweet potatoes (Figure 6.8b). The chlorophyll content of taro and sweet potatoes reached their peak during the mid-vegetative growth phases, followed by a gradual decline throughout the different stages of growth. The highest concentrations of chlorophyll were observed during the mid-reproductive, early- and late-reproductive growth periods. Afterwards, the chlorophyll content decreased during the late reproductive phase.

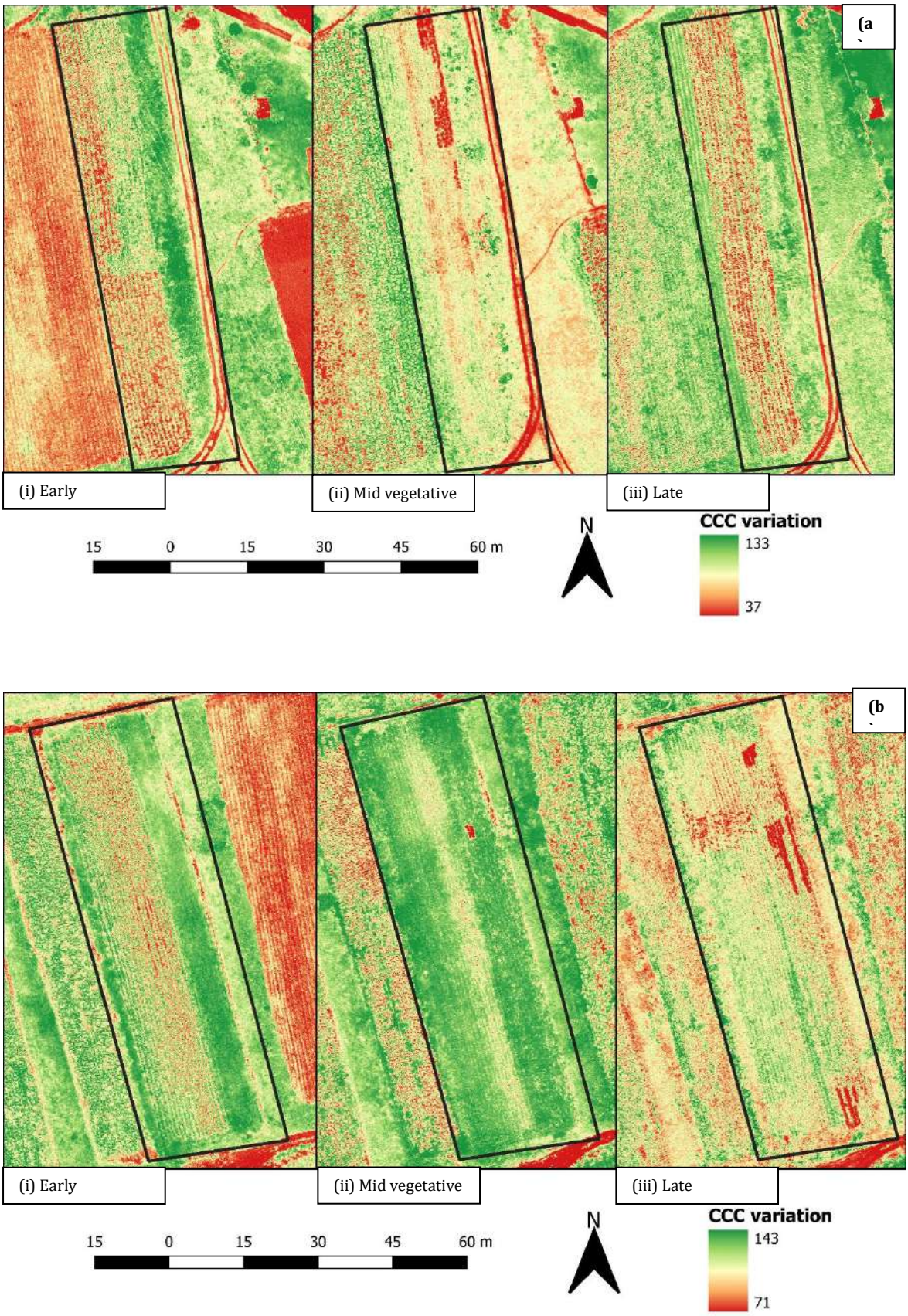


Figure 6-8 Spatial distribution of (a) taro and (b) sweet potato chlorophyll content in the experimental fields

6.4 Discussion

The objective of the study was to monitor the canopy chlorophyll content of taro and sweet potato crops using UAV remote sensing data in smallholder farming areas that are commonly found in southern Africa. To achieve this, the significance of individual vegetation indices, the bands and the combination of both datasets was analysed.

6.4.1 Comparative performance of Vegetation Indices, bands and combined datasets

When predicting the CCC of sweet potatoes, the use of combined data modelling achieved the highest prediction accuracy, as determined by CVI, SRI, NGRDI, NDVIred-edge and CIRE, which were identified as the key predictor variables ranked by their significance. Similarly, for taro, the combined dataset provided a superior performance in estimating the chlorophyll content, with a mean RMSE of 84.51 μmol m^{-2}, compared to 95.75 μmol m^{-2} for chlorophyll vegetation indices, 103.60 μmol m^{-2} for bands, and 142.87 μmol m^{-2} for traditional vegetation indices. The key predictor variables for the combined model included ExG, $NDVI_{rededge}$, NIR, CCCI and GCI, which were also ranked by their significance. This denotes that employing an integration of spectral bands and indices enhanced the effectiveness of elements that may have been limited, due to dense canopy cover. This improvement is due to the integration of bands and vegetation indices, which offers a wider spectrum coverage. Each band collects spectral information from a designated section of the electromagnetic spectrum, by reflecting a particular feature of crop growth or attribute, whereas VIs are formulated from the combination of bands spanning various electromagnetic spectrum segments. As a result, vegetation indices gain their effectiveness from various electromagnetic spectrum regions, which allows them to respond to a wider range of features (Kapari *et al.* 2024).

In addition, VIs generally diminish the effects of noise on the spectral signatures of crops. As a result, the integration of different bands and VIs improves their sensitivity to multiple facets of crop physiological and health status, including the canopy chlorophyll content. The combination of these datasets creates a synergistic impact that incorporates both the physiological and structural traits of plants to enhance the precision and sensitivity of the prediction procedure overall. In addition, health-related VIs are noted to be the more effective estimation features for evaluating the chlorophyll content in crops, rather than conventional indices, such as NDVI. The superior performance of Vis, such as CCCI, can be attributed to their capability to avoid spectral saturation when the crop canopy completely covers the ground. For example, at peak canopy densities, the NDVI reaches a point of spectral saturation. Therefore, the findings indicate that integrating the bands with VIs like CCCI enhanced the accuracy of the estimations. When only VIs were utilised, the prediction accuracies dropped notably for taro (RMSE = 142.86) and for sweet potatoes (RMSE = 200.81). This implies that traditional indices, when integrated with chlorophyll indices, were unable to effectively evaluate the chlorophyll content in the crop canopy. This remains the case, despite the effectiveness of Vis, like TCARI, in minimising the interference from backgrounds that do not involve photosynthesis.

6.4.2 Estimation of taro and sweet potato leaf chlorophyll content at different phenological stages

The modelling results demonstrated a strong performance across the vegetative growth phases, which indicated that the late vegetative growth phase for sweet potatoes achieved the greatest RMSE accuracy of 13.6 and R^2 = 0.23 μmol m^{-2} (rRMSE of 11%), based on the SR1, CIRE, Ci_{green}, NDVI, RRI1, SAVI, BLUE, GCI and EXG as the most influential elements (according to their significance). The enhanced performance is due to the greater uniformity and elevated concentration of chlorophyll in the foliage, which improves the spectral signals that are acquired by the model. This aligns with the denser canopy formation of the crop, which minimises the soil background reflectance and generates a more uniform spectral signal. The vegetation indices accurately reflected variations in leaf reflectance, which was associated with chlorophyll content, canopy structure and pigment concentration (Gitelson *et al.* 2003). During the mid-vegetative growth stage, the model achieved an RMSE = 18,7, an rRMSE of 15% and an R^2 = 0.0 by using CVI, SRI, NGRDI, $NDVI_{rededge}$, and CIRE as the key variables. These key variables are known for their sensitivity to vegetation indices and chlorophyll estimation. While these indices most likely captured some aspects of the leaf reflectance and chlorophyll characteristics, structural and environmental factors or inherent noise may have limited their effectiveness. This performance underscores the necessity for the additional refinement of variable combinations and adjustments to models, in order to improve their precision and dependability during the mid-vegetative period (Thenkabail Prasad S *et al.* 2000). At the early vegetative growth stage, the model recorded the lowest accuracies, with an RMSE of 21,2, an rRMSE of 17% and an R^2 = 0.18, with the key variables including CIRE, RED, $Cir_{rededge}$, NDRE and RDR. The reduced performance can be attributed to the developmental characteristics of the plants at this phase, such as the reduced leaf size, the decreased chlorophyll levels and a less consistent distribution, which diminish the spectral signals employed for estimation (Zhang *et al.* 2023). In addition, factors that are related to the environment, like soil moisture and light settings, may exert a stronger influence on the reflectance measurements during this stage, which may contribute to the noise in the key predictors. Although these indices work well for mature plants, they may not be as responsive to the nuanced chlorophyll changes in young leaves (Kima *et al.* 2022).

In contrast, the LCC model performed best for taro over the mid-vegetative development phase, with the highest RMSE accuracy of 23,3 and an R^2 = 0.05 μmol m^{-2}. (rRMSE of 17%), based on the ExG, $NDVI_{rededge}$, NIR, CCI, GCI, OSAVI, EVI, RRI1 and SRI as the most influential variables (in order of importance). In particular, the NIR and red-edge wavelengths were identified as vital for the model's estimation of the chlorophyll content, as these wavelengths relate to the health of the plants. The red-edge wavelength is particularly sensitive to plants that exhibit a high chlorophyll content, nitrogen levels and biomass, which makes predictions more accurate in this spectrum (Brewer *et al.* 2022). Moreover, the NIR region is crucial for estimating the chlorophyll content, due to its sensitivity to elevated foliar reflectance, which is driven by the pigment concentrations in the plant canopy structures (Brewer *et al.* 2022). In addition, vegetation indices derived from the red-edge and NIR wavelengths tend to mitigate the influence of atmospheric disturbances, the varying visible light, different background

effects and the geometric arrangement of scenes, compared to the traditional bands (Brewer *et al.* 2022). As a result, the red-edge and NIR wavelengths enabled the accurate prediction of the taro chlorophyll content. The reduced foliage density in the initial stages of vegetative growth did not create saturation and caused spectral confusion for the sensor, when capturing the images. In addition, the connection between NIR, Red-edge, the chlorophyll concentration and the vegetation indices acquired from them has been recognised as essential for evaluating crop health (Brewer *et al.* 2022). In the initial phase of vegetative growth, the model achieved an RMSE= 25.3, an rRMSE of 17,7% and an R^2 = 0.01 by using MCARI, CCI. RED, SRI and Red-edge as the key variables. Similarly, in the mid-vegetative growth phases, the red-edge band emerged as a crucial factor in the prediction of the chlorophyll levels in taro, due to its effective prediction of peak chlorophyll reflectance in these areas. The biochemical characteristics of dense leaf canopy, including the thick waxy cuticle, air spaces, chloroplasts and the thickness of mesophyll cells, all contribute to elevating the reflectance of red-edge, which is directly linked to the chlorophyll content (Mutanga *et al.* 2012, Sibanda *et al.* 2021, Brewer *et al.* 2022). Consequently, the elevated levels of chlorophyll found during these phases of the study correlated with a more robust crop. In this context, the high chlorophyll concentrations were advantageous for model prediction, which made these phases the most precise for estimation. During the late vegetative development phase, the model recorded an RMSE of 25, an rRMSE of 18,5% and an R^2 = 0.03, with the key variables including CVI, CI_{green}, NGRDI, SRI and NIR. The low accuracy of model predictions can be explained by the negative impact of a hailstorm that compromised the structure of the maize canopy, which led to a decrease in the crop's chlorophyll levels. The impairment in the canopy of maize uncovered the underlying soil surface, which consisted of damaged and decaying maize leaves. This situation created spectral confusion, due to the combined reflectance of the soil and rotting leaves captured by the MicaSense Altum camera, which made it difficult for the model to accurately identify the existing variations in chlorophyll. This challenge arose because the prevailing brown hue from the ageing leaves and soil led to several bands and vegetation indices being obscured, due to the reduced chlorophyll levels, while wavelengths like the red band reflected at much higher levels (Brewer *et al.* 2022).

6.4.3 Estimation of the taro and sweet potato canopy chlorophyll content at different phenological stages

The estimation of the CCC of sweet potatoes and taro was optimal during the mid-vegetative stages, with an RMSE of 42.3 and 15.8, and an rRMSEs of 9.2% and 14.5%, respectively. For sweet potatoes, the most vital predictor elements, according to their significance, were the spectral reflectance of CVI, SRI, NGRDI and $NDVI_{rededge}$ (Figure 6.3b). Similarly, for taro, the key variables were ExG, $NDVI_{rededge}$, NIR, CCCI and GCI, in order of significance. The NIR and red-edge regions were highlighted as the key wavelengths for predicting the chlorophyll content in the model, as they correlated with those of healthier plants (Chang-Hua *et al.* 2010, Ansar and Muhammad 2020, Brewer *et al.* 2022). The red-edge wavelength is particularly sensitive to vegetation that has an elevated chlorophyll level, nitrogen content and biomass, which make it more effective for predictions in these cases (Brewer *et al.* 2022). In addition, the optimal influence of the Near-Infrared (NIR) region may be attributed to the fact that plants

with high chlorophyll concentrations tend to reflect highly in such regions (Broge and Leblanc 2001, Sankaran *et al.* 2013). Consequently, the vegetation indices acquired from NIR and red-edge are key areas that diminish the impacts of atmospheric interference, the variations in visible light, the diverse background influences and the geometric layout of a scene, when contrasted with the traditional bands (Curran *et al.* 1990, Goodbody *et al.* 2020). Therefore, the NIR and red-edge wavelengths enhanced the accuracy of predicting the sweet potato and taro chlorophyll, since the sparse foliage density during the early growth stage did not overwhelm the sensor and lead to spectral confusion during the image capture.

Moreover, the superior accuracies that are seen in the models for sweet potatoes and taro during the mid-vegetative stage can be linked to the specific physiological and structural traits of these crops at this growth period. The mid-vegetative stage represents an optimal phase for canopy development, which is characterised by dense, uniform foliage and increased chlorophyll concentrations. These factors synergistically enhance the signals that are captured by remote sensing tools, while reducing the effects of noise from the non-vegetative components, such as the soil or sparse vegetation (Prabawardani 2007, Lencha *et al.* 2016, Kunz *et al.* 2024). The increased chlorophyll concentrations at this phase enhance the near-infrared reflectance, which reduces the reflectivity in the red region of the electromagnetic spectrum. These spectral traits have a strong association with the health of the vegetation and canopy attributes, which facilitates more precise estimates of essential biophysical parameters, such as the CCC (Pinty *et al.* 2009, Brewer *et al.* 2022). An important element that leads to the improved performance of the model at this stage is the development of robust canopies in both crops. By the mid-vegetative stage, the leaves have grown to form a thick and consistent layer that effectively conceals the soil beneath. This coverage of the canopy diminishes the impact of soil reflectance, which can interfere with the spectral signals received by remote sensing devices. In addition, the uniformity of the canopy decreases the variability in reflectance that might occur due to uneven or incomplete coverage, thus improving the reliability and consistency of the data that are collected (Prabawardani 2007, Din *et al.* 2017, Yang *et al.* 2022). Furthermore, the combination of a dense canopy, a consistent leaf coverage and a heightened chlorophyll concentration leads to a better signal-to-noise ratio. This enhancement ensures that the spectral information captures the vegetation signal more accurately, while reducing interference from unrelated factors, like the bare soil or environmental artefacts. By decreasing such interference, the models gain clearer and more accurate input data, which is crucial for reliable predictions (Brewer *et al.* 2022, Kapari *et al.* 2024).

During the early and mid-vegetative stages, taro and sweet potatoes exhibited their highest levels of chlorophyll. During these stages, the plants are engaged in active photosynthesis, which results in a maximum leaf chlorophyll content, which strengthens and simplifies the detection of the reflectance signal. This is particularly important because spectral indices like NDVI, CVI and GCI are responsive to variations in the chlorophyll level. At the early to mid-vegetative stages, chlorophyll absorption in the red spectrum and reflectance in the near-infrared (NIR) range are most effective for capturing the Canopy Chlorophyll Content (CCC). As the plants progress to the later stages, there is a decline in chlorophyll levels, due to senescence, while the levels of other pigments, like carotenoids, may rise. These alterations

can shift the spectral reflectance, which complicates the accurate assessment of the chlorophyll content, thus diminishing the effectiveness of the regression models (Song and Wang 2022).

The low prediction accuracies of the model during the late vegetative phase may be linked to the detrimental impact of a hailstorm that compromised the taro and sweet potato canopy structure, which led to a decrease in the crop's chlorophyll concentration. The canopy suffered physical deterioration, by revealing the underlying soil that was covered in deteriorated and decayed crop leaves. The exposure resulted in spectral confusion because of the blending of the reflectance from soil and decaying leaves captured by the MicaSense Altum camera, which impeded the model's capacity to identify the differences in the chlorophyll levels. The dominant brown hues from the senescing leaves and soil led to the absorption of several VIs and bands because of the low chlorophyll levels, while wavelengths like the red band had significantly higher reflections. Nevertheless, a noticeable drop in the NDVI values before the hailstorm was apparent and it correlated with the reduction in chlorophyll. This could be attributed to the crop redirecting its nutritional resources and energy towards fruit development (Walker *et al.* 2018, Brewer *et al.* 2022), which corresponded with the observed decrease in chlorophyll concentration.

6.4.4 Comparing the leaf and canopy chlorophyll content estimations

In the estimation of chlorophyll content, the CCC estimations were notably more accurate than the LCC estimations, with R^2 values of 0.9 and 0.1, respectively. This is because CCC offers a more comprehensive and precise depiction of the chlorophyll content at a canopy level, compared to estimates based on individual leaves (Halme *et al.* 2019, Le Saint *et al.* 2024). CCC estimations utilise remote sensing technologies, such as hyperspectral and multispectral imaging, to capture the chlorophyll signals from entire canopies. This approach integrates the spatial variability, which provides a more comprehensive representation of chlorophyll content over large areas (Narmilan *et al.* 2022). The use of vegetation indices, including the Normalised Difference Vegetation Index (NDVI) and the Chlorophyll/Carotenoid Index (CCI), improves the accuracy by reducing the impact of environmental factors, such as soil reflectance and atmospheric conditions (Gao *et al.* 2023). A wealth of research has provided substantial evidence that VIs are highly correlated with various vegetation parameters, including structural, phenological and biophysical attributes like LAI (Potithep *et al.* 2013, Cao *et al.* 2017, Xing *et al.* 2019). Moreover, CCC estimations consider the structural characteristics of the canopy, such as the LAI, the distribution of leaf angles and the density of the canopy. These characteristics affect how light is absorbed and reflected within the canopy, which enables remote sensing models to incorporate three-dimensional complexity (Yang *et al.* 2022). This capacity to model the canopy structure and merge the spatial data results in the enhanced accuracy of CCC estimations, when compared to the LCC methods. Furthermore, the accuracy of CCC estimations is also affected by the observation scale. At a canopy level, the measurements combine the chlorophyll signals from many leaves, which lessens the influence of localised anomalies and outliers. On the other hand, LCC estimations concentrate on single leaves, which makes them more vulnerable to variability and less indicative of the total chlorophyll content in a plant or ecosystem.

6.5 Conclusion

This research sought to predict and monitor the spatial fluctuations of the LCC and CCC of the neglected and under-utilised taro and sweet potato crops, by utilising data from remote sensing derived from drones in smallholder croplands, using a Random Forest regression. Therefore, based on the results of the research, it was determined that:

- accurate CCC prediction accuracies were generated over the mid-vegetative development phase stages for both taro and sweet potato crops, based on UAV-derived NIR and red-edge wavelengths;
- accurate LCC prediction accuracies were generated over the late vegetative development stage for sweet potatoes, and the mid-vegetative growth stage for taro, based on UAV-derived SR1, NDVI and ExG wavelengths;
- combining bands, traditional vegetation indices and chlorophyll vegetation indices results in optimal prediction accuracies for the CCC of taro and sweet potatoes, compared to using these variables separately.
- CCC exhibits better model estimations, when compared to LCC, for both taro and sweet potato crops

Given that the canopy chlorophyll content has been extensively shown to serve as an indicator of crop health, the results of the research suggest that data derived from drones could be accurately employed to assess the overall health condition of taro and sweet potatoes in smallholder croplands, with notably enhanced spatial precision. The advancements in precision technology offer cost-effective and precise methods for smallholder farmers to enhance their decision-making and manage their agricultural practices. Multispectral UAV systems provide spatially-detailed and almost instantaneous information to assess the condition of crops by measuring the chlorophyll, a key biochemical indicator. Moreover, the Random Forest model showed a relatively-high level of accuracy in predicting the chlorophyll levels on smallholder farms. Thus, multispectral UAV technology serves as a valuable tool for smallholder agriculture, and it equips farmers with an insight into crop growth patterns that are tailored to specific temporal and spatial requirements, which leads to better management and increased productivity.

7 EVALUATING DRONE-BASED REMOTE SENSING FOR MONITORING THE SOIL WATER CONTENT AND LEAF CHLOROPHYLL IN TARO (*COLOCASIA ESCULENTA*) ACROSS PHENOLOGICAL STAGES

7.1 Introduction

Food insecurity has been intensifying and remains a global concern. Agriculture is the backbone of food security; however, as the population continues to grow, so does the demand for food. Climate change poses a threat to the existing agricultural systems. Erratic weather events, such as droughts or flooding, have led to a reduction in the crop yield, which further exacerbates the food insecurity issues (Ngcamu and Chari 2020). Erratic weather conditions affect the soil moisture content, which plays an important role in influencing the health and overall productivity of crops. The literature shows that the soil moisture content is one of the most important parameters affecting the crop yield and overall productivity (Zhao *et al.* 2021, Furtak and Wolińska 2023, Kumar *et al.* 2024). This poses a challenge to regions such as southern Africa, which is impacted by high climate variability and which, in turn, induces extensive variations in the amount and pattern of precipitation and soil moisture (Fahad *et al.* 2017). This affects 70% of the smallholder croplands in Africa, where rainfed crop farming is conducted, which makes them prone to the impacts of climate change (Nyoni *et al.* 2024). Maintaining optimal moisture levels is important for the crop yield and for ensuring food security, especially in smallholder croplands where water and resources are limited. However, there are very limited spatially-explicit methods of accurately assessing the linkages between soil moisture and crop productivity in near-real-time.

Soil moisture stress occurs when there is not enough water available in the soil to support the photosynthetic process, which affects the optimal growth of plants (Fu *et al.* 2022). The morphology and physiological functions of plants change under soil moisture stress (Vennam *et al.* 2023). This includes stunted growth, a reduction in photosynthetic activities, wilting, lower crop yields and an increase in their susceptibility to pests (Ahluwalia *et al.* 2021, Vennam *et al.* 2023, Muhammad *et al.* 2024). According to Vennam *et al.* (2023), plants can adapt to soil moisture stress, depending on the duration, growth stage and severity of the water stress. For instance, during the vegetative and tasselling stages, moisture stress results in a reduced plant height and it affects leaf development and reproduction (Cakir 2004). To combat the impact of climate variability and soil moisture stress, smallholder farmers are adopting NUS, such as taro, which has the capability to thrive under drought conditions that are associated with high soil moisture stress. One of the most effective, yet indirect, ways of assessing soil moisture is through the chlorophyll content in crops, such as taro (Li *et al.* 2018, Li *et al.* 2023). Chlorophyll is the green pigment of plants that is responsible for photosynthesis, and it is quite sensitive to the availability of soil water (Kancheva *et al.* 2014). A low soil moisture results in plants experiencing water stress, which leads to a lower chlorophyll content (Kancheva *et al.* 2014, Parkash and Singh 2020). Therefore, monitoring the chlorophyll levels in 'climate smart' crops, such as taro, may reveal more about the field's moisture conditions. For example, crops

that reflect lower chlorophyll levels in a field could indicate that there is less soil moisture; therefore, a decision can be made on whether to irrigate the field and enhance its overall growth.

Soil moisture stress and its interaction with the productivity elements of such crops as taro, particularly the chlorophyll content, is traditionally and non-invasively measured by using instruments that include the Konica Minolta SPAD 502 plus and the soil moisture probe. Although these instruments are highly accurate, they are point-based, time-consuming and monotonous (Nduwamungu *et al.* 2009). Furthermore, these methods are often unable to capture the spatial variability of soil moisture across the fields, as they are limited in scope (Rasheed 2022). There is therefore a need for the spatially-explicit and accurate assessment of linkages between the soil moisture and crop productivity in near-real-time.

The literature demonstrates that remote sensing technologies can be effectively used to accurately assess the soil moisture content and its linkage to the chlorophyll content of a crop (Ma *et al.* 2020). For instance, Ma *et al.* (2020) utilised a combination of Sentinel-1 SAR and Sentinel-2 multispectral imagery to estimate the soil moisture, with an Root Mean Square Errors (RMSEs) ranging from 0.039 to 0.078 m^3/m^3 and R^2 ranging from 0.47 to 0.67. However, these sensors are limited, in terms of their relatively coarse spatial and temporal resolution (Bukowiecki *et al.* 2021). They cannot capture the finest details and have a much longer revisit time, which makes them less suitable for monitoring crops in smallholder croplands, which are small, heterogenous and highly fragmented (Feng *et al.* 2024). Therefore, there is a need for remote sensing technologies that can provide high spatial and temporal resolution data that are suitable for accurately monitoring variations in the soil moisture content and chlorophyll content of crops, such as taro, in smallholder croplands across their different phenological stages.

The recent advances in remote sensing technology, particularly using UAVs, which are widely known as drones, have demonstrated promising potential for spatially quantifying aspects of precision agriculture, such as soil moisture stress, and its linkage to crop productivity elements, such as the chlorophyll content (Ndlovu *et al.* 2021, Brewer *et al.* 2022, Brewer *et al.* 2022). These technologies are suitable for high-throughput remote sensing-based phenotyping in smallholder croplands because they fly at lower altitudes, they capturing ultra-high spatial resolution data that are suitable for soil and crop productivity assessments (Zhang *et al.* 2021). Furthermore, UAVs can capture near-real-time remotely-sensed data at user-defined spatial and temporal resolutions, on a field scale. They are relatively more affordable and time-efficient, which makes them a suitable technology for settings, such as smallholder croplands, where resources are limited and the fields are small and fragmented. For example, Ge *et al.* (2021) demonstrated that UAV-acquired hyperspectral data could estimate the agricultural soil moisture content at optimal accuracies (with an R^2 of 0.921 and an RMSE of 1.943). Despite their capabilities, UAV-borne sensors have not been utilised to their fullest potential in mapping the soil moisture content, in relation to productivity elements, such as the chlorophyll content of neglected and under-utilised crops, such as taro. Therefore, there is a need to further explore their potential in assessing the soil moisture content and chlorophyll content of taro in smallholder croplands. The remote sensing the of soil moisture content and the chlorophyll content of taro contributes towards optimising its agricultural productivity in smallholder

croplands. Indirectly, this work contributes towards addressing the grand global challenges of poverty, food and nutrition insecurity, particularly in sub-Saharan Africa, where food shortages are a looming concern. This aligns with the objectives of Sustainable Development Goals 1 and 2.

7.2 Methods and Materials

7.2.1 Field data collection

The study area was divided into multiple sampling plots to ensure the accurate collection of data and for soil moisture variation purposes. The soil moisture content and chlorophyll were monitored throughout the different phenological stages. The field data were collected on 16th December 2022, 08th January, 11th April and 02nd June 2023. Points were sampled at random within the experimental field, and a Trimble hand-held GPS was used to record the location of these points. The soil moisture measurements were taken by using the HH2 hand-held Moisture Meter with the ThetaProbe ML3. The HH2 Moisture Meter is equipped with Frequency Domain Reflectometry technology (FDR) to accurately and quickly measure the soil moisture content (Tapia *et al.* 2019). The HH2 probe was inserted into the soil across the different sampling plots to ensure representative sampling. The probe transmits electromagnetic pulses into the soil, and the time it takes for the pulses to return is proportional to the amount of moisture present (Tapia *et al.* 2019, Loconsole *et al.* 2025). The chlorophyll content was measured by using the Soil and Plant Analyser Development (SPAD) 502 plus chlorophyll meter. The meter is put directly onto the leaf, and the light that is reflected is used to determine the chlorophyll, as unitless values (Masemola *et al.* 2025). The unitless values measured by the meter are then converted to the chlorophyll content, as in (Brewer *et al.* 2022), based on the model by Markwell *et al.* (1995)The readings were conducted on a single leaf of each plant, to avoid any external light (Melash et al. 2023). The sampled data for both the soil moisture and chlorophyll content were captured in a Microsoft Excel spreadsheet against the coordinates of each sampling point and converted into point maps. The point maps were overlaid with the image and its derivatives.

7.2.2 Spectral variables used for the prediction of the soil moisture content and chlorophyll content of taro crops

In order to accurately predict the soil moisture content and the chlorophyll content, a combination of UAV-derived spectral bands and VIs were used. The spectral signatures of the soil moisture content and the chlorophyll content were extracted from the reflectance values. Vegetation Indices (VIs) were calculated by using the spectral bands (Table 7.1). The spectral data were overlaid with the field-collected point data on the soil moisture and chlorophyll content to develop the predictive models.

Table 7-1List of vegetation indices used in modelling the soil moisture and chlorophyll content

Vegetation Index	Abbreviation	Formula	Reference
Normalised Difference Vegetation Index	NDVI	(NIR - Red) / (NIR + Red)	(Giovos, Tassopoulos et al. 2021, Ji, Pan et al. 2021)
Soil Adjusted Vegetation Index	SAVI	(1 + 0.5) (NIR – RED)/ (NIR + RED + 0.5)	(Zhen, Chen et al. 2021)
Normalised Difference Water Index	NDWI	(GREEN – NIR) / (GREEN + NIR)	(Pizarro Carcausto, Pricope et al. 2023, Patil, Jagtap et al. 2024)
GNDVI	GNDVI	(NIR + GREEN) / (NIR – GREEN)	(Pizarro Carcausto, Pricope et al. 2023)
Normalised Difference Red-Edge	NDRE	(NIR - Red-edge) / (NIR + Red-edge)	(Voitik, Kravchenko et al. 2023)
Chlorophyll red-edge	Clred-edge	((NIR / Red-edge) – 1)	(Giovos, Tassopoulos et al. 2021)
Chlorophyll green	Clgreen	((NIR / GREEN) – 1)	(Kurbanov and Zakharova 2020)
Chlorophyll Vegetation Index	CVI	(NIR*RED) / ($GREEN^2$)	(Gao, Yan et al. 2024)
Modified Chlorophyll Absorption in Reflectance Index	MCARI	((Red-edge – RED) – 0.2(Red-edge - GREEN))(Red-edge / RED)	(Giovos, Tassopoulos et al. 2021)
Enhanced Vegetation Index	EVI	2.5 × (NIR - Red) / (NIR + 6 × Red - 7.5 × Blue)	(Zeng, Hao et al. 2022)
Transformed Vegetation Index	TVI	0.5(120(NIR − GREEN) − 200(RED − GREEN))	(Marques, Pádua et al. 2023)
Photochemical Reflectance Index	PRI	(GREEN – BLUE) / (GREEN + BLUE)	(Marques, Pádua et al. 2023)
Simple Ratio	SR	NIR / RED	(Kurbanov and Zakharova 2020)
Canopy Chlorophyll Content Index	CCCI	(NIR – Red-edge) / (NIR + Red-edge) / (NIR – RED) / (NIR + RED)	(Aslan, Sabanci et al. 2024)

7.2.3 Statistical analysis

Exploratory assessments were conducted to assess the magnitude of association between the soil moisture and chlorophyll content by using the Pearson correlation test. The Random Forest (RF) algorithm was then used to predict the soil moisture and chlorophyll content of crops across different phenological stages, by using remotely-sensed data in RStudio Version R i386 4.1.3. RF is a machine learning algorithm that uses multiple decision trees to accurately predict and obtain results (Belgiu and Drăguţ 2016, Sheykhmousa 2020). RF was chosen and used in this study because it is a very robust and versatile algorithm, which uses the bootstrap aggregating approach and splits variables randomly from the training data (Sheykhmousa 2020). RF is useful for predicting continuous variables, such as the soil moisture and chlorophyll content. The algorithm constructed decision trees, each using subsets of the spectral features (bands and vegetation indices) that were chosen randomly, and then their results were averaged to generate a prediction (Manafifard 2024). More specifically, the variables used in generating models included UAV-derived bands, particularly blue, green, red, red-edge, NIR and thermal bands. The VIs were chosen, based on their performance in the literature (Brewer *et al.* 2022, Abrahams *et al.* 2023, Buthelezi *et al.* 2023). Before predicting the soil moisture and chlorophyll content, the data were split into two subsets: 70% for training the models and 30% as testing data for evaluating the model's performance.

7.2.4 Accuracy assessment

The accuracy assessment of the soil moisture and chlorophyll content prediction models was assessed, based on the Root Mean Square Error (RMSE), the Relative Root Mean Square Error (rRMSE) and the coefficient of determination (R^2), based on the training and testing data. The RSME is commonly used as a quantitative measure of the differences between the observed and predicted values. R^2 is commonly used to assess the goodness of fit in a model (Wu *et al.* 2023). The rRMSE is commonly used to assess the performance of predictive models. Interpreting the statistical metrics involved evaluating how well the model predictions match the observed data. Lower RMSE and rRMSE values indicate a better model performance, and an R^2 value closer to 1 indicates a strong relationship between the predicted and observed values (Chicco *et al.* 2021).

7.3 Results

7.3.1 Descriptive statistics

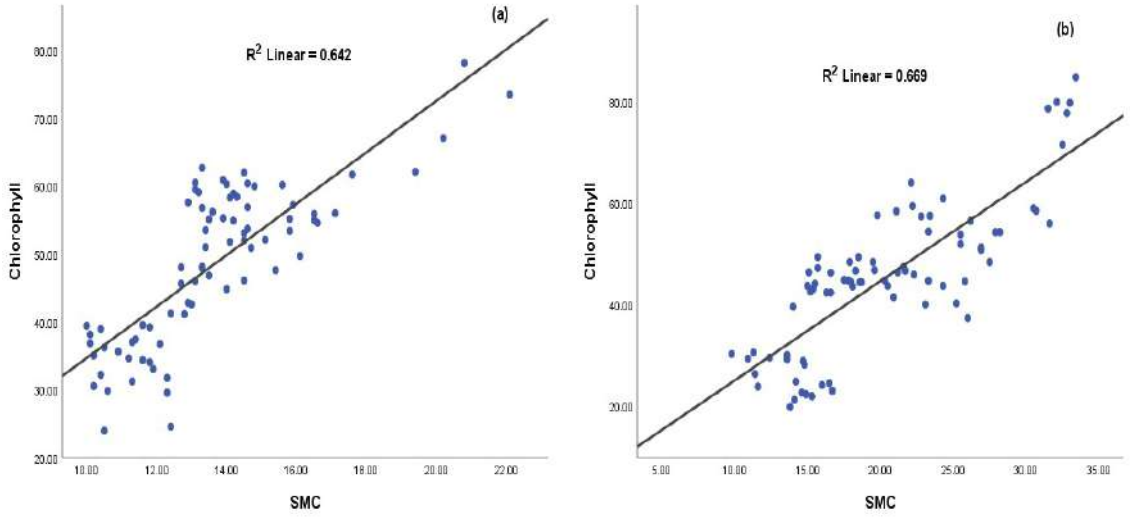
Table 7.2 shows the variation of soil moisture and chlorophyll content across the four phenological stages. At the early establishment stage (16 Dec 2022), the soil moisture content had a moderate mean of 13,64% and a standard deviation of 2.46, while the chlorophyll had a mean of 48,48 and a standard deviation of 11.65. The mean of the soil moisture then increased to 20,40% and a standard deviation of 6.16, while the chlorophyll content exhibited a mean of 45.43 and standard deviation of 14.72 during the establishment stage on the 08th January 2023.

Then the mean of the soil moisture content decreased during the vegetative stage to 12.04% and a consistent standard deviation of 2.96 on the 11th of Apr 2023. Meanwhile, chlorophyll had the highest variability in this stage, with a standard deviation of 17.56 and a mean of 47.03%. The mean of the soil moisture content then increased to 18.8%, with a standard deviation of 5.47, while the chlorophyll content exhibited a mean of 62.28%, with a standard deviation of 12.05, during the late vegetative stage on the 2nd of June 2023.

Table 7-2 Descriptive statistics of soil moisture content and chlorophyll across different phenological stages

Phenological stage	variable	Minimum	Maximum	Mean	Standard deviation
Early Establishment	SMC	10	21.1	13.64	2.46
	Chlorophyll	23.9	78.2	48.48	11.65
Late Establishment	SMC	9.8	33.4	20.40	6.16
	Chlorophyll	19.9	84.8	45.43	14.72
Early Vegetative	SMC	6.1	24.9	12.04	2.96
	Chlorophyll	10.4	94.9	47.03	17.56
Late Vegetative	SMC	9.6	33.4	8.8	5.74
	Chlorophyll	29.3	89.8	62.28	12.05

In evaluating the magnitude of the relationship between the soil moisture content and the chlorophyll content, the Pearson correlation test showed that there were statistically significant correlations ($p < 0.05$) between the soil moisture content and the chlorophyll content across all the phenological stages. During the early establishment stage (Week 5), the association between the soil moisture content and chlorophyll content was significant ($R^2 = 0.64$, $r = 0.80$, $p < 0.05$) (Figure 7.1a). The establishment stage (Week 8) also showed a significant association between soil moisture content and chlorophyll content of taro crops ($R^2 = 0.67$, $r = 0.82$, $p < 0.05$). The vegetative exhibited an R^2 of 0.64, an r of 0.81, and a $p < 0.05$ (Figure 7.1(c). Similarly, a significant association between the soil moisture content and chlorophyll content was noted during the vegetative stage ($R^2 = 0.61$, $r = 0.78$, $p < 0.05$).



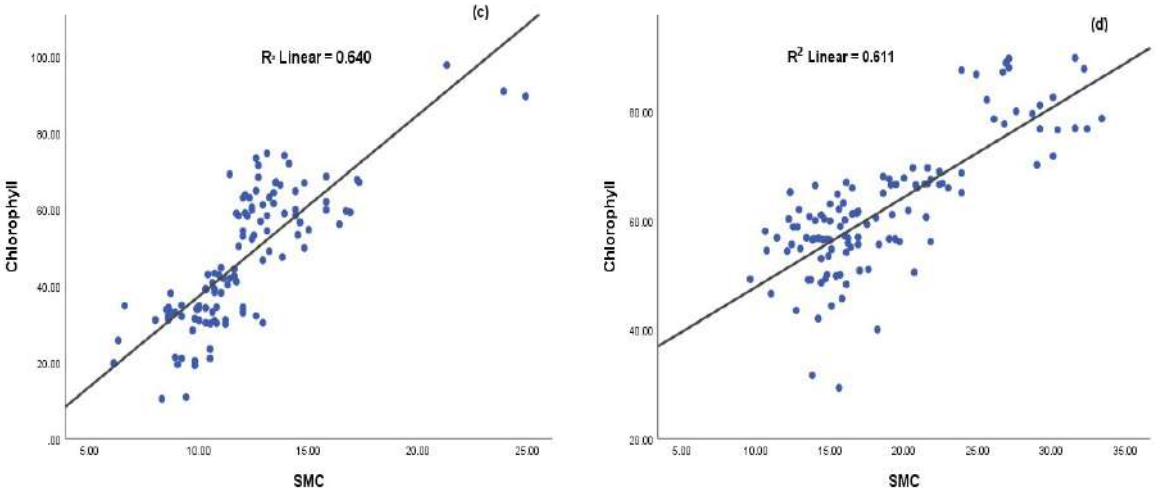


Figure 7-1 Correlation between the SMC and chlorophyll for (a) establishment stage in week 5, (b) establishment stage in week 8, (c) vegetative week pink and (d) late vegetative stage

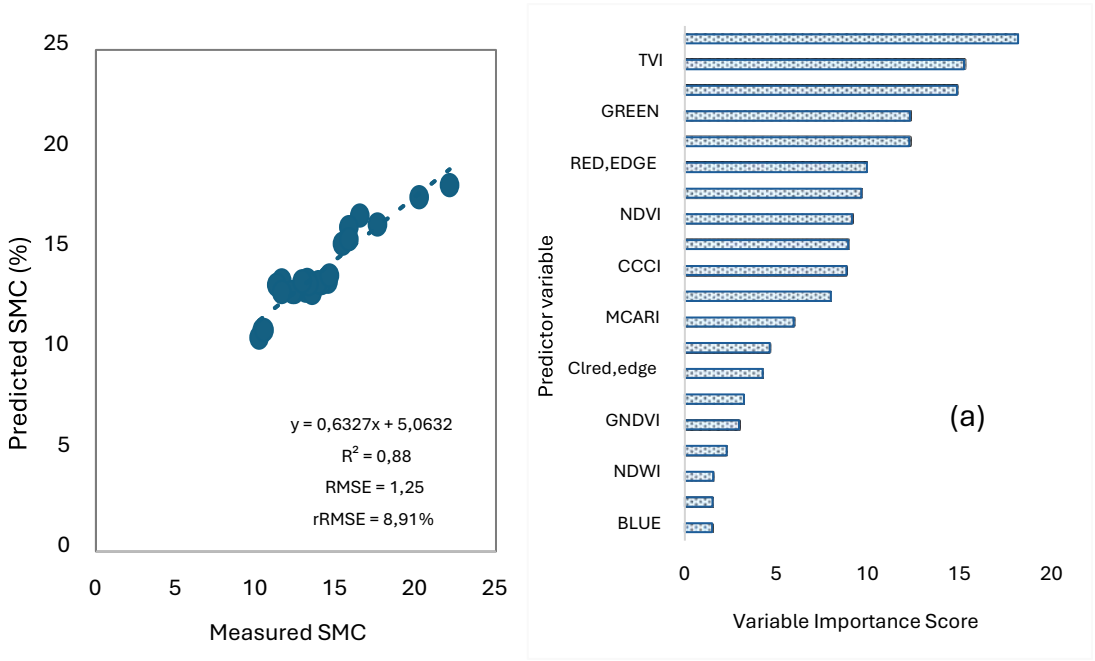
7.3.2 Comparing the performance of bands, vegetation indices and combined data in estimating the soil moisture and chlorophyll content

The UAV bands outperformed the VIs in predicting the chlorophyll content of taro crops across the growing season. For example, the UAV bands showed an R^2 value of 0.67 and an RMSE of 12.09 for 08 Jan 2023, whereas the VIs exhibited a higher rRMSE of 18.35%. The UAV bands provided better results for chlorophyll predictions, especially in the late vegetative state (02 June 2023), where the R^2 was 0.90, and the RMSE was 4.24. It can be noted that the combination of both the UAV bands and VIs did improve the model accuracies. However, the combination did not outperform UAV bands alone, but often matched the performance of the UAV bands. It can be said that the VIs exhibited better accuracies during the early growth stages, particularly in predicting SMC, but they underperformed for chlorophyll. On the other hand, the UAV bands were consistent throughout the phenological stages in the prediction of of both SMC and chlorophyll. The results suggest that UAV bands are more reliable for monitoring crop health attributes.

7.3.3 Estimating soil moisture and chlorophyll content at different phenological stages

Figures 7.2 and 7.3 show the accuracies obtained in estimating both the soil moisture content and the chlorophyll, based on UAV bands, Vegetation Indices (VIs), as well as a combination of both. UAV bands provided better prediction accuracies for the chlorophyll content. The first establishment stage had an R^2 of 0.88, an RMSE of 1.67 and an rRMSE of 11.94%, based on the Thermal, Red, Green, NIR, Red-Edge and Blue spectral variables, in order of importance. The chlorophyll content was optimally estimated, based on the combined data only during the first vegetative stage, with an R^2 value of 0.8, an RMSE of 13.69 and an rRMSE of 26.57, based on SR, CCCI, NDVI, TVI, SAVI, EVI, PRI, Red, NIR, NDWI, Blue, MCARI, NDRE, Red-Edge, Thermal, Green, CVI, GNDVI, Clred-edge and Clgreen, as the optimal spectral variables, in order of importance (Figure 7.3c). Meanwhile, the SMC was optimally estimated

by using the combined data across all phenological stages, except for the second vegetative stage. The stage which exhibited the optimal soil moisture content estimation was during the establishment stage (an R^2 = 0.88, an RMSE =1.25% and an rRMSE = 8.91) based on Red-Edge, Thermal, Red, TVI, MCARI, NDVI, SAVI, SR, Clred-edge, EVI, NDRE, NIR, GNDVI, NDWI, GREEN, BLUE, CVI, PRI and Clgreen, in order of importance (Figure 7.2c). Meanwhile, chlorophyll content was optimally estimated during the vegetative growth stage of taro to an R^2 of 0.9, an RMSE of 4.24 and an rRMSE of 6.72%, based on Clred-edge, Red-Edge, SR, TVI, MCARI, NIR, PRI, SAVI, Red, EVI, NDVI, NDRE, CCCI, Blue, Thermal, GNDVI, NDWI, CVI, Clgreen and Green as the optimal spectral variables, in order of importance.


Predicted SMC (%)
Measured SMC
y = 0,6327x + 5,0632
R² = 0,88
RMSE = 1,25
rRMSE = 8,91%
Predictor variable
TVI
GREEN
RED,EDGE
NDVI
CCCI
MCARI
Clred,edge
GNDVI
NDWI
BLUE
(a)
Variable Importance Score

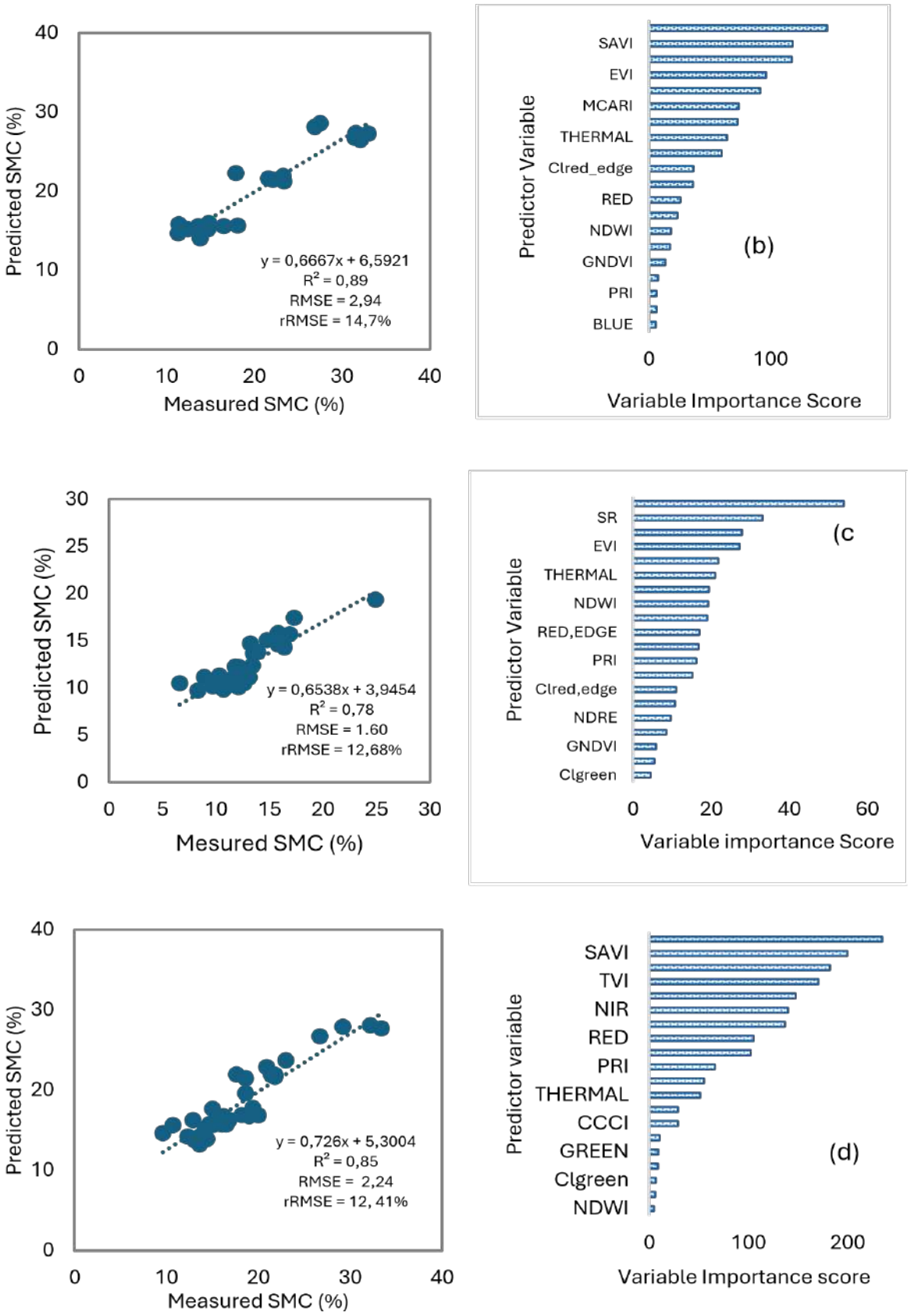
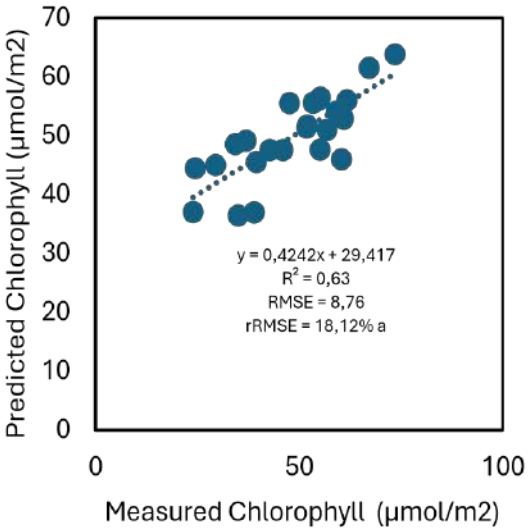
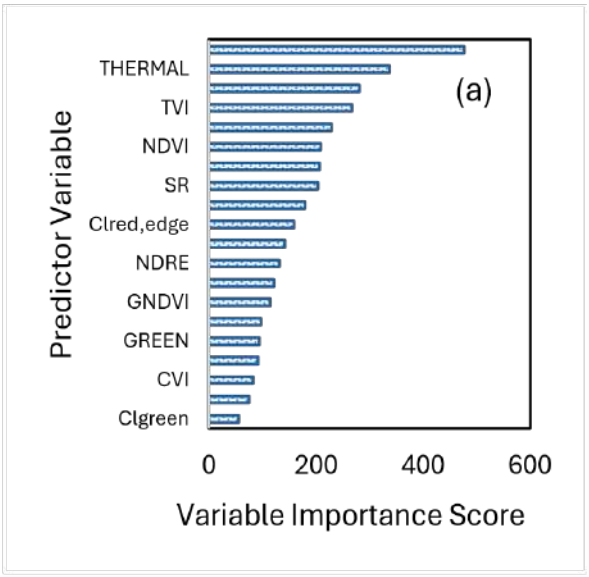
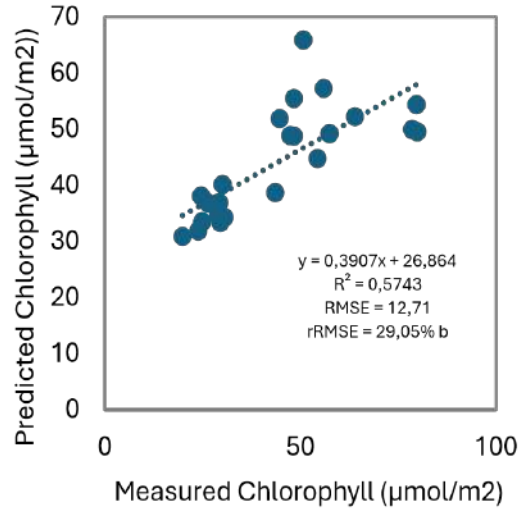
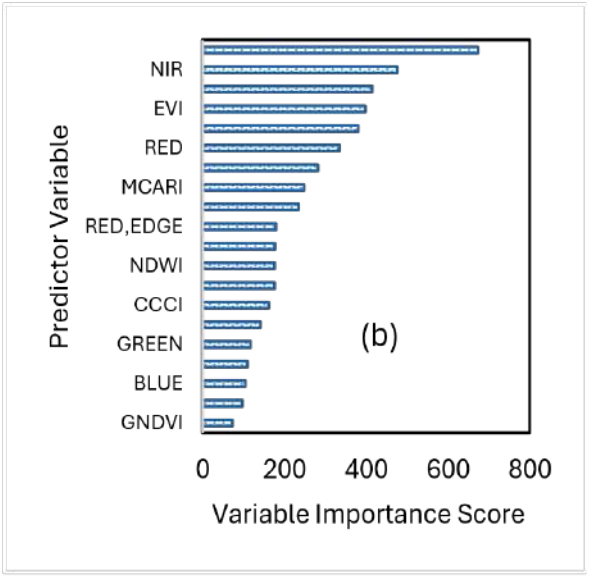


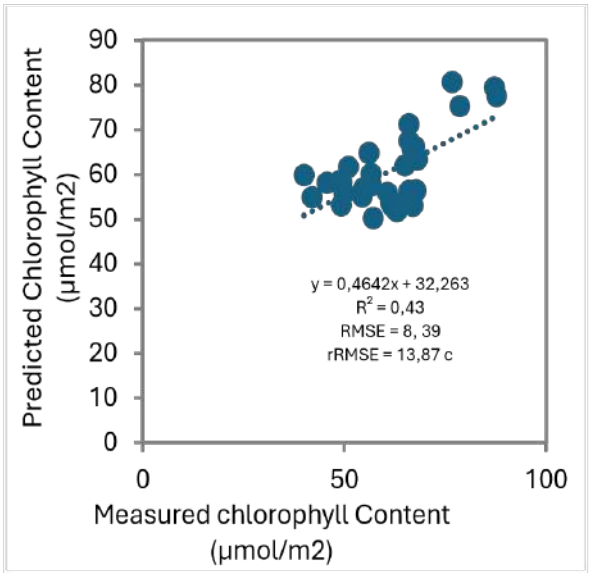
Figure 7-2 Soil moisture content for the (a) Establishment Stage (i), (b) Establishment Stage (ii), (c) Vegetative stage (i), (d) Vegetation Stage (ii) and their associated variable importance

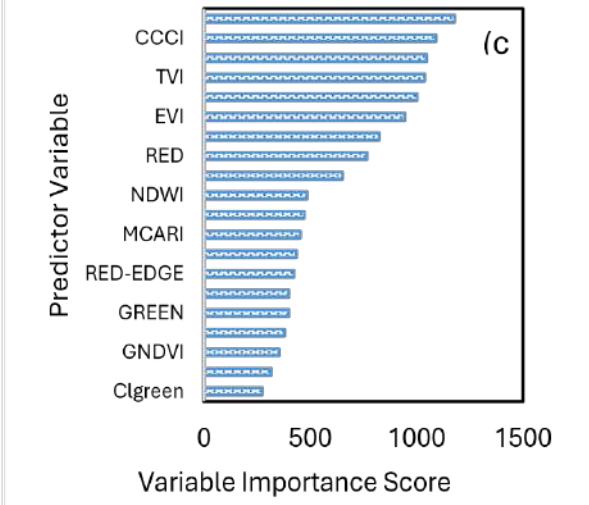

Predicted Chlorophyll (µmol/m2)
70
60
50
40
30
20
10
0
0
50
100
Measured Chlorophyll (µmol/m2)
y = 0,4242x + 29,417
R² = 0,63
RMSE = 8,76
rRMSE = 18,12% a


Predictor Variable
THERMAL
TVI
NDVI
SR
Clred,edge
NDRE
GNDVI
GREEN
CVI
Clgreen
(a)
0
200
400
600
Variable Importance Score


Predicted Chlorophyll (µmol/m2))
70
60
50
40
30
20
10
0
0
50
100
Measured Chlorophyll (µmol/m2)
y = 0,3907x + 26,864
R² = 0,5743
RMSE = 12,71
rRMSE = 29,05% b


Predictor Variable
NIR
EVI
RED
MCARI
RED,EDGE
NDWI
CCCI
GREEN
BLUE
GNDVI
(b)
0
200
400
600
800
Variable Importance Score


Predicted Chlorophyll Content (µmol/m2)
90
80
70
60
50
40
30
20
10
0
0
50
100
Measured chlorophyll Content (µmol/m2)
y = 0,4642x + 32,263
R² = 0,43
RMSE = 8, 39
rRMSE = 13,87 c


Predictor Variable
CCCI
TVI
EVI
RED
NDWI
MCARI
RED-EDGE
GREEN
GNDVI
Clgreen
(c
0
500
1000
1500
Variable Importance Score

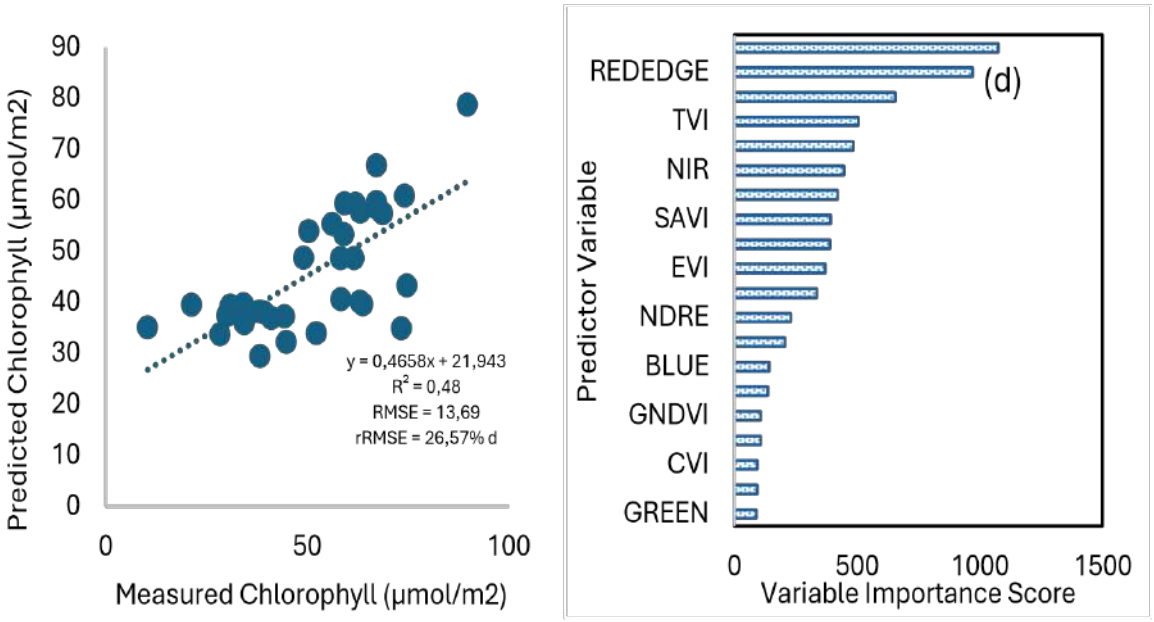


Figure 7-3 Chlorophyll content accuracies for (a) Establishment stage (i), (b) Establishment stage (ii), (c) Vegetative stage (i), (d) Vegetation stage (ii) and their associated variable importance

7.3.4 Spatial variation of soil moisture content and chlorophyll content at different growth stages

Figures 7.4 and 7.5 illustrate the spatial distribution of the soil moisture content and chlorophyll content across the four phenological stages of taro crops. It can be observed that SMC was higher in both the establishment stages and lower in the vegetative stages, whereas the chlorophyll was higher in the late vegetative stage and lower in the establishment stages.

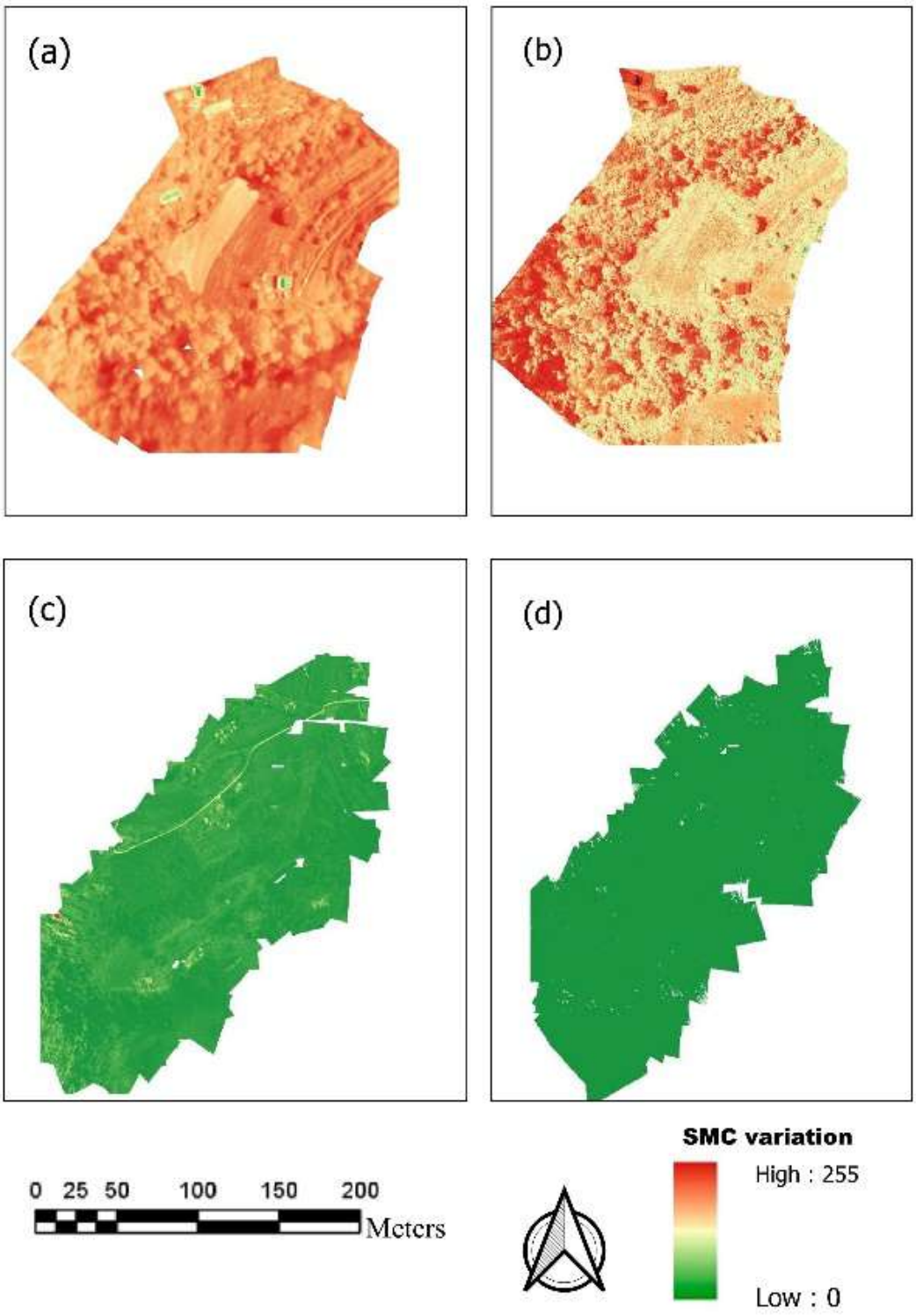


Figure 7-4 Spatial variation of soil moisture content at different growth stages (a) establishment (i), (b) establishment (ii), (c) vegetative (i), (d) vegetative (ii)

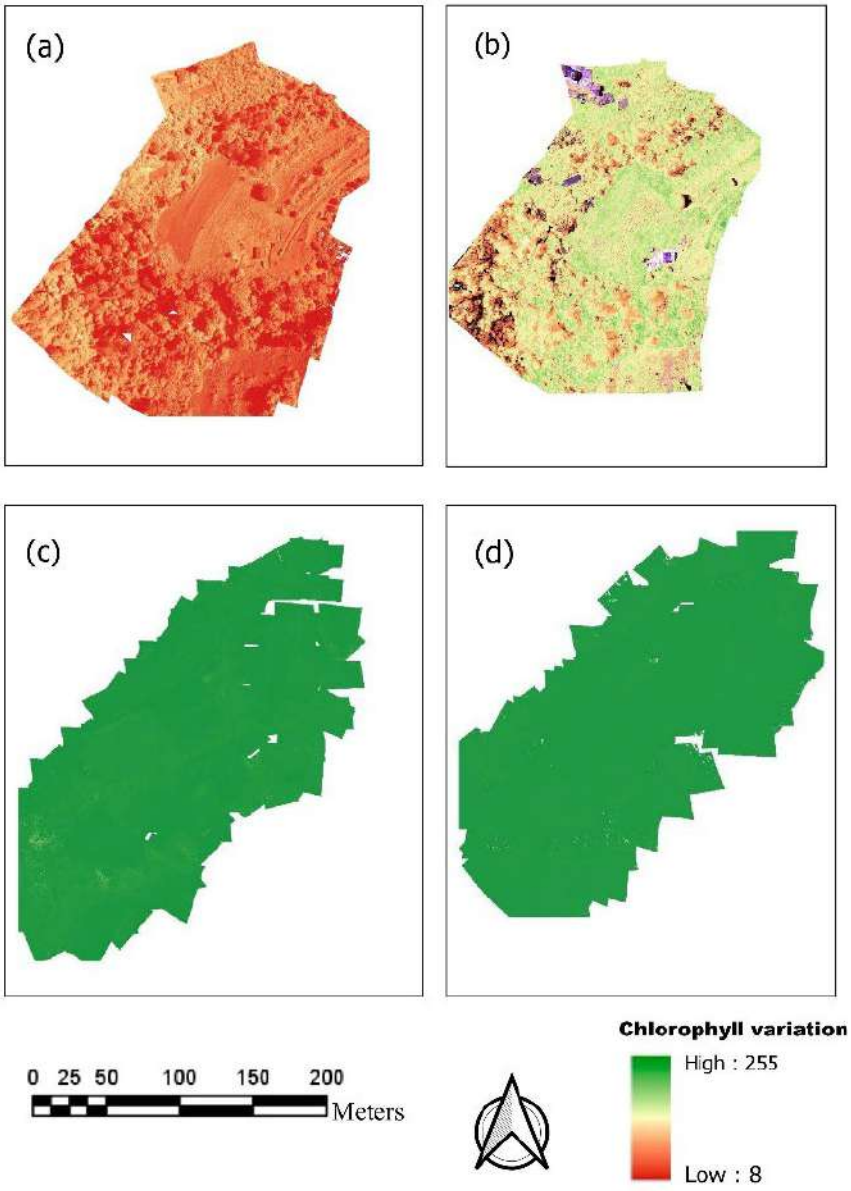


Figure 7-5 Spatial variation of chlorophyll content at different growth stages (a) establishment (i), (b) establishment (iii), (c) vegetative (i), (d) vegetative (ii)

7.4 Discussion

The objective of this work was to test the utility of UAV-acquired data in predicting the SMC and chlorophyll content across the smallholder taro croplands at different phenological stages.

7.4.1 Association between the soil moisture content and chlorophyll

A significant association between the SMC and chlorophyll across the different phenological stages of sweet potato and taro crops was observed in this study (Figure 7.1). This suggests that a low soil moisture is associated with less chlorophyll content. Soil moisture is critical for photosynthetic processes and the health of a plant (Gavrilescu 2021). When there is limited moisture in the soil, plants fail to draw up the water and nutrients required for photosynthesis.

Generally, this is often associated with chlorosis and reduced photosynthesis, as the leaves turn yellow (Gholamin and Khayatnezhad 2020). This implies that the chlorophyll content could be used as a surrogate measure to infer the soil moisture content. A few studies have demonstrated the linkage between the SMC and chlorophyll content. For example, the study by Gholamin and Khayatnezhad (2020) demonstrated the effects of the SMC on chlorophyll, and the results showed a decline in the maize yield in instances where there was low soil moisture. Sun *et al.* (2023) focused on how spinach responds to different moisture contents in soil under irrigation; the results indicated high chlorophyll levels where there was adequate soil moisture. The least optimal estimation of SMC was observed in both the vegetative stages with SR, SAVI, EVI, TVI, Clred-edge, NIR, NDVI, Red, Red-edge, PI, RED, PRI, MCARI, Thermal, NDRE, CCCI, CVI, Green, GNDVI, Blue and NDWI as optimal spectral variables, in order of their importance (Figure 7.3(c)(d)). For chlorophyll, the least optimal estimations were observed in the early establishment, with Red-Edge, Thermal, Red, TVI, MCARI, NDVI, SAVI, SR, CCCI, Clrededge, EVI, NDRE, NIR, GNDVI, NDWI, Green, Blue, CVI, PRI and Clgreen as optimal spectral variables, in order of their importance (Figure 7.2).

7.4.2 Optimal estimation of soil moisture and chlorophyll across the phenological stages

The results of this study showed that the optimal estimation of SMC was achieved during the early establishment stage (16 December 2022) based on the Red-Edge, Thermal, Red, TVI, MCARI, NDVI, SAVI, SR, Clred-edge, EVI, NDRE, NIR, GNDVI, NDWI, GREEN, BLUE, CVI, PRI and Clgreen as optimal spectral variables, in order of importance (Figure 7.2a). This could be explained by the fact that crops are still young and highly responsive to soil moisture content variations at the establishment stage (Ge *et al.* 2021, Zhao *et al.* 2021). In addition, variables such as Red-edge, Red, Thermal, NDVI and SAVI are very sensitive to moisture and may reflect highly (Liu *et al.* 2021). Red-edge and NIR bands can detect variations in the chlorophyll levels and even changes in the plant structure, which are indicators of the moisture content (Giovos *et al.* 2021). The chlorophyll content was optimally estimated during the late vegetative stage of taro, based on Clred-edge, Red-Edge, SR, TVI, MCARI, NIR, PRI, SAVI, Red, EVI, NDVI, NDRE, CCCI, Blue, Thermal, GNDVI, NDWI, CVI, Clgreen and Green as the optimal prediction variables, in order of importance (Figure 7.3d). This could be attributed to the fact that taro crops are well-developed in this stage and therefore exhibit high chlorophyll levels (Melash *et al.* 2023). The variables have the capability of detecting chlorophyll and differentiating chlorophyll levels, based on the plant density. The chlorophyll content in the vegetative stage is consistent, which makes it easy for these spectral bands and indices to detect any variations (Kancheva *et al.* 2014, Li *et al.* 2023, Wu *et al.* 2023, Lao *et al.* 2024).

7.4.3 Comparative performance of bands, vegetation indices and combined data sets in estimating the soil moisture and chlorophyll across the phenological stages

The findings of this study showed that there were no significant differences between the performance of all variables combined and bands, when estimating the soil moisture and chlorophyll content. This could be explained by the fact that incident energy from the sun interacts directly with the soil and crop foliage before it is captured by the sensor as reflectance.

This explains why bands are more robust, when compared to vegetation indices. Despite the renowned performance of vegetation indices in mapping other fields and crop attributes, in this study, they did not optimally estimate the soil moisture content and chlorophyll content. This could be explained by the fact that the vegetation indices could be generating redundant information, resulting in lower accuracies, due to their failure to estimate the soil moisture content.

7.4.4 Implications of the study

UAV-borne sensors are becoming important tools for precision agriculture. They provide opportunities for smallholder farmers to improve their crop health and to develop water management strategies. By mapping the soil moisture and chlorophyll content variations across different phenological stages, farmers can make timely and data-driven decisions to enhance their crop productivity and food and nutrition security. The results of this study can be used by farmers to develop and promote climate adaptation.

7.5 Conclusion

The objective of this study was to predict the soil moisture content and chlorophyll content across the different phenological stages of taro crops in smallholder croplands. Grounded on the findings of this study, it can be concluded that:

- the soil moisture content could be optimally estimated during the early establishment stage, based on Red-Edge, Thermal, Red, TVI, MCARI, NDVI, SAVI, SR, Clred-edge, EVI, NDRE, NIR, GNDVI, NDWI, GREEN, BLUE, CVI, PRI and Clgreen as the optimal spectral variables, in order of importance;
- the chlorophyll content may be optimally estimated during the late vegetative stage, based on Clred-edge, Red-Edge, SR, TVI, MCARI, NIR, PRI, SAVI, Red, EVI, NDVI, NDRE, CCCI, Blue, Thermal, GNDVI, NDWI, CVI, Clgreen and Green as the optimal spectral variables, in order of importance; and
- there was a significant positive correlation between the soil moisture and chlorophyll content across all the growth stages of taro. This suggests that the chlorophyll content may be utilised as a proxy for assessing the soil moisture content.

The findings of this study imply that UAV-acquired data can optimally estimate the SMC and chlorophyll content across different phenological stages by using specific spectral bands and vegetation indices. These findings pave the way towards the operationalisation of the UAV remote-sensing-based high-throughput phenotyping of crops in smallholder croplands. In the long run, this technology could be effectively used to optimise agricultural production. Increasing agricultural production will assist in achieving the objectives of the Sustainable Development Goals Numbers 1 and 2, by advocating a reduction in hunger, food and nutrition insecurity and ending poverty, especially in the majority of the world, including sub-Saharan Africa.

8 ESTIMATING THE TARO BELOW-GROUND BIOMASS IN SMALLHOLDER CROPLANDS USING UAV MULTISPECTRAL IMAGERY

8.1 Introduction

Agricultural production plays a crucial role in feeding the rapidly-growing world population (Awad 2019). By 2050, the crop yields in sub-Saharan African (SSA) countries will need to double, in order to meet the national food demand, and the irrigation water will need to increase by 40% to 50% (Abbas *et al.* 2020). Maize and wheat are among the most widely-cultivated food crops in SSA and they are all high-water-demand crops. As a result, their capacity to sustain a continued increase in yield is proving to be limited and more difficult, as a result of the increasing climate variability and frequent drought episodes (Kavhiza *et al.* 2022). However, tuber crops such as taro, are some of the most widely-grown traditional food crops in SSA, as they are rich in carbohydrate nutrients and are highly adaptable to extreme weather conditions, such as droughts (Vitor *et al.* 2019, Chimonyo *et al.* 2020). More specifically, taro has the potential to produce high yields, which makes it a promising smart future crop for addressing the looming food and nutrition insecurity and it will contribute immensely towards improving the resilience of the food systems in SSA (Kunene *et al.* 2022). There is a need to understand the biomass dynamics of taro across the growing season, in order to optimise the productivity and yields. More specifically, understanding the spatial and temporal variations of taro's AGB and its relationship to the yield is important for developing and implementing site-specific crop management measures to improve crop production. The existing methods of measurement, such as cutting and weighing, are invasive and traditional. Furthermore, these traditional approaches tend to be costly, labour-intensive and destructive, while yield prediction depends largely on on-site visits, which can be subjective, inaccurate and labour-intensive. In this regard, there is a need to establish comprehensive, spatially-explicit techniques for the monitoring and early detection of taro's productivity dynamics, in order to initiate remedial actions for optimising its yield production.

Remote sensing has been demonstrated to be an efficient alternative approach for measuring the growing season crop canopies and to provide information on the spatial variability of the crop AGB and yield. The Above-Ground Biomass (AGB) is closely related to a crop's health and its nutritional status; however, it is often hypothesised and unclear as to whether there is a tangible relationship between AGB and the Below-Ground Biomass (BGB)/yield. As demonstrated in the literature, AGB can be used as an indicator of a crop's growth status (Li *et al.* 2020, Liu *et al.* 2021). However, the remote sensing systems capture data on the crop's canopy cover and less often on the BGB (i.e. the tuber), which is the critical component of taro's productivity and yield. Since the tuber grows in the soil, it would require ground-penetrating sensors to acquire data, which makes it complex to establish accurate models, especially for fragmented and heterogeneous smallholder croplands. UAV-borne multi-spectral sensors are capable of acquiring data on the AGB, which is a significant agronomic parameter in characterising the growth, development and nutritional status of a crop. Such data is generally used to evaluate the crop growth and estimate the yield, especially in smallholder

croplands (Buthelezi *et al.* 2023). In this regard, the fast, non-destructive and accurate monitoring of the temporal and spatial dynamic variations in the AGB and BGB of taro is essential for facilitating the use of multispectral data for their estimation.

Furthermore, machine learning algorithms, such as Random Forest, support vector machines and multiple linear regressions have been proven to be instrumental in characterising the crop biomass and yield (Li *et al.* 2020, Liu *et al.* 2021). The Random Forest ensemble has been widely proven to outperform the other two aforementioned algorithms (Liu *et al.* 2021, Liu *et al.* 2022). Hence, it is anticipated that the combination of UAV-derived VIs and the robust Random Forest regression could produce accurate results to quantitatively assess the relationship between the AGB and BGB taro. Therefore, this study aimed to test the use of UAV multi-spectral remotely-sensed data in estimating the AGB and BGB of taro. The specific objectives of this study were: (1) to explore and assess the relationship between the AGB and BGB of taro, and (2) to identify the optimum phenological stage for the prediction of BGB before harvest.

8.2 Materials and Methods

8.2.1 Study area

This experiment was conducted in a smallholder farm located in Swayimane, a communal area within the uMshwathi Municipality, north-east of Pietermaritzburg, KwaZulu-Natal, South Africa. The details of Swayimane and the experimental fields are outlined in Chapter One.

8.2.2 Field sampling and taro above-ground biomass and below-ground biomass/yield measurements

Field visits were conducted on 8 January 2023, 11 April 2023 and 14 July 2023 at the study site during the establishment stage (58 days after planting, or DAP), at the vegetative stage (DAP = 154) and at the mature stage (DAP = 249), respectively. A stratified random sampling method was used to generate a total of 80 random sampling points within the taro field plot. This method was selected because it provides a representative sample of the study area. A Trimble hand-held Global Position System (GPS) with a sub-metre accuracy was used to navigate to these randomly generated sample points within the plot. The biomass and yield data of taro were measured by harvesting the crops at the 80 sampling points. Sampling was conducted between 12:00 and 14:00, in order to coincide with the image acquisition. The crop samples were then oven-dried at 70°C until a constant Dry Weight (DW) was reached (approximately 48 h for AGB and 72 h for BGB) (Mhango *et al.* 2021). The DW was obtained by using a calibrated scale with a 0.5 g measurement error. The DW of the ABG and BGB taro was then integrated with the GPS location and converted into a point map that was overlaid with the UAV multispectral images of the study area.

8.2.3 UAV platform, image acquisition and processing

The DJI Matrice 300 series (M300) and the MicaSense Altum imaging sensor were used to acquire multispectral images covering the taro field (Figure 8.1a). The specifications of the

M300 are detailed in Table 8.1. The Altum camera consists of a thermal and five visible and near infrared spectral channels that specifically acquire data in the blue (475 nm), green (560 nm), red (668 nm), red-edge (717 nm), NIR (840 nm) and thermal (8-14 nm) at a ground-sampling distance of 9.6 cm per pixel at a flight height of 100 m (Figure 8.1b). The key advantage of this imaging platform is its ability to capture synchronised thermal and multispectral data simultaneously in an automated manner. A shapefile of the study area that was digitised in Google Earth Pro was imported into the M300's hand-held console to automatically determine a flight path plan, based on the area and the provided flight height (Figure 8.1c). To calibrate for any variations in the incidental atmospheric conditions, a calibration reflectance panel was used (Figure 8.1d). This was done by capturing the image of the reflectance panel before and after capturing the data, by using the UAV across the growing stages. The automated flight missions were conducted on cloudless days and at low wind speeds, between 10:00 and 12:00 local time. This was done to minimise the variation in illumination and solar zenith angle (Raeva *et al.* 2019). The flight path was set up to have an 80% forward overlap and an 80% side overlap, which was deemed to be sufficient to generate high-quality mosaics. The captured images were recorded in a 16-bit local digital memory card in raw Geographic Tagged Image File Format (GeoTIFF) for further post-georeferencing and mosaicking. An ortho-mosaic of the captured imagery was generated and pre-processed by using the Pix4D Fields photogrammetry software.

Table 8-1 DJI M300 UAV specifications

Parameter	**Specification**
UAV type	Rotary wing
Weight	~ 4.53 kg
Size	887 (width) x 880 (length) x 378 (height)
Flight Duration	55 min
Maximum speed	27 m/s
Maximum altitude	7 000 m
Company	Da-Jiang Innovations
Location	Shenzen, Guangdong, China

Figure 8-1 a) Matrice 300 UAV integrated with the Altum sensor serving as the imaging platform used in this study, b) the Altum camera, c) the calibrated reflectance panel and d) the flight path plan for the study image

8.2.4 Statistical analysis and model development for biomass and yield prediction

8.2.4.1 Establishing the relationship between AGB and BGB

Correlation analyses were conducted to test and establish whether there were significant relationships between the AGB and BGB at the selected growth stages within the growing season. A Pearson product-moment correlation test was conducted by following a Kolmogorov-Smirnov (K-S) normality test, which indicated that the data did not significantly deviate from the normal distribution ($P > 0.05$). Pearson's correlation coefficient (r) was used to assess the magnitude of association between the two biomass variables. The correlation coefficient is a scale that measures the strength of a linear association between the variables. It ranges between -1 and +1, with -1 to -0.49 indicating a weak negative correlation, while -0.5 to -0.99 indicates a strong negative correlation, 0 to +0.49 a weak positive correlation and +0.5 to +1.0 a strong positive correlation.

8.2.5 Selection of Vegetation Indices

The pre-processed multispectral bands were used to compute the vegetation indices (Aliche *et al.*). Vegetation Indices (VIs) were used to augment the spectral bands in predicting the AGB and BGB of taro. Table 8.2 illustrates the VIs in this study and their respective formulae. These VIs were selected for this study based on their direct and indirect correlation with the plant biomass, as detailed in the literature (Elmetwalli *et al.* 2014, Abou Ali *et al.* 2020). The prepared multispectral imagery data were then overlaid with the point data with the measured

AGB and BGB, in order to extract the spectral signatures from each sampling point. The spectral signatures were exported as points and used in the statistical prediction procedure.

Table 8-2 List of vegetation indices used in the modelling of AGB and BGB and related source references

Index	Full Name	Formula	Reference
NDVI	Normalised Difference Vegetation Index	NIR – Red/NIR+ Red	(Tan *et al.* 2020)
ENDVI	Enhanced Normalised Difference Vegetation Index	((NIR + Green) – (2*Blue)) / ((NIR + Green) + (2*Blue))	(Dhau *et al.* 2018)
NDRE	Normalised Difference Red-Edge Index	NIR – Red-edge / NIR + Red-edge	(Sibanda *et al.* 2017)
NDWI	Normalised Difference Water Index	Green – NIR / Green + NIR	(Ndlovu *et al.* 2021)
SR	Simple Ratio	NIR/R	(Brewer *et al.* 2022)
SAVI	Soil Adjusted Vegetation Index	((NIR – Red) / (NIR + Red + 0.5)) * (1.5)	(Al-Gaadi *et al.* 2016)
OSAVI	Optimised Soil Adjusted Vegetation Index	(NIR – Red) / (NIR +Red + 0.16)	(Al-Gaadi *et al.* 2018)

8.2.6 Taro AGB and BGB prediction and accuracy assessment

To explore the use of UAV-acquired, remotely-sensed data in predicting the AGB and BGB of taro, the extracted data, characterised by the 80 samples of the crop's spectral signature (a combination of bands and VIs data), were randomly split into the training (70%) and validation (30%) datasets, respectively, for each growth stage. The Random Forest (RF) algorithm was adopted and used in this study because it is a robust, nonparametric statistical technique that uses a bagging-based approach to build an ensemble of regression trees, while ranking the important variables that produce an independent measure of prediction error (Adam *et al.* 2014). RF was conducted in R statistical software. In R, the number of decision trees (*ntree*) and the number of predictor variables tested for the best split when growing the trees (*mtry*) parameters, were optimised by using the *doBest* function. The function selected the *ntree* and *mtry* parameters with the lowest RMSE, to determine the most influential parameters. These parameters were tuned to 500 for *ntree* and 3 for *mtry*. Test data were used to assess the accuracy of the predictive models. Performance indicators, which include the coefficient of determination (R^2), the Root Mean Square Error (RMSE) and the Relative Root Mean Square (rRMSE) were determined and used to assess the accuracy of each model. More specifically, R^2 was used to measure the magthe nitude of variation between the measured and predicted AGB and BGB, while the RMSE was used to assess the magnitude of error between the field samples and the modelled AGB and BGB. The rRMSE was used to compare the performance

of the models across the different predictor variables. To compute rRMSE, the RMSEs from each model were normalised by using the mean of each variable,variable andressed as a percentage. RF Gini impurity index was employed to select optimal spectral features for the estimation of the AGB and BGB.

8.3 Results

8.3.1 Descriptive analysis of AGB and BGB

Significant variations were observed in the ABG and BGB of taro across the growing season. Table 8.3 presents the descriptive statistics of AGB and BGB. The averages for AGB and BGB were 22.91 g and 30.38 g during the establishment stage, 54.21 g and 106.22 g during the vegetative stage and 2.54 g and 129.12 g during the maturity stage, respectively.

Table 8-3 Descriptive statistics of the AGB and BGB across selected growth stages of taro

Growth Stage	Parameter	Min	Max	Mean	Median	Std
Establishment	AGB	6	146.2	22.91	16.6	21.72
	BGB	0	133.8	30.38	21.95	33.38
Vegetative	AGB	27.1	413.8	54.21	44	44.31
	BGB	12.9	374.5	106.22	113.3	73.25
Maturity	AGB	2	6	2.54	7.5	27.26
	BGB	10.4	445.5	129.12	83.2	79.95

8.3.2 Correlation analysis between the AGB and BGB of taro

Strong positive correlations were observed between the AGB and BGB during the establishment, vegetative and maturity stages of the growing season (Figure 8.2). During the establishment stage, there was a significant ($p < 0.05$) positive correlation coefficient of 0.81, while 0.55 and 0.81 were observed between AGB and BGB for the vegetative and the mature growth stages, respectively. This suggests that there is a positive relationship between the AGB and BGB. However, during the vegetative stage, the relationship between the AGB and BGB was relatively weaker in relation to the establishment and maturity stages.

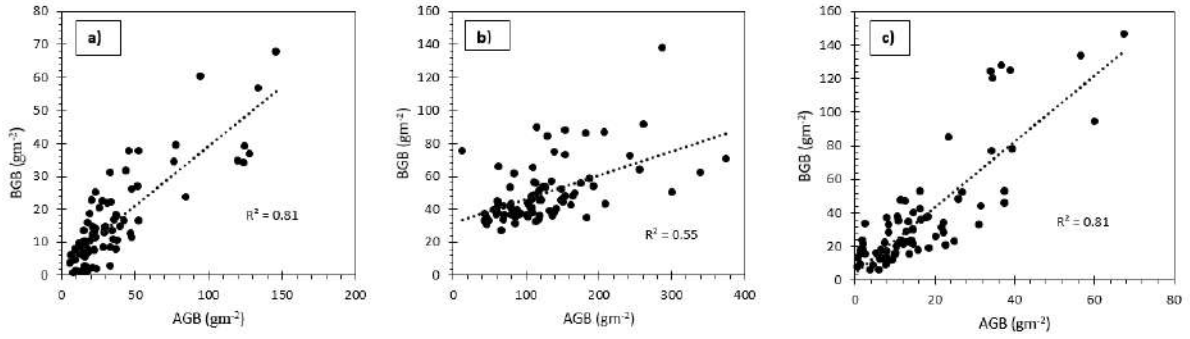


Figure 8-2 Correlation between the measured AGB and BGB at the a) establishment stage (DAP =58), b) vegetative stage (DAP =154) and c) maturity stage (DAP = 249)

8.3.3 Predicting the AGB and BGB of taro using the Random Forest regression ensemble

When estimating the AGB, the maturity model demonstrated the highest prediction accuracy (R^2 = 0.94 and RMSE 0.14 g/m^2), with SAVI and the thermal band being the most optimal estimation spectral variables. This was followed by the vegetative model. with an R^2 of 0.72 and an RMSE of 10.85 g/m^2, with SAVI and OSAVI being the most influential predictor variables. The prediction accuracy decreased slightly for the establishment growth stage model (an R^2 of 0.90, an RMSE of 13.91 g/m^2), thus making the maturity model the most optimal for the prediction of AGB (Figure 8.3). At the maturity stage, SAVI and Thermal proved to be the most suitable AGB predictor variables (Figure 8.3).

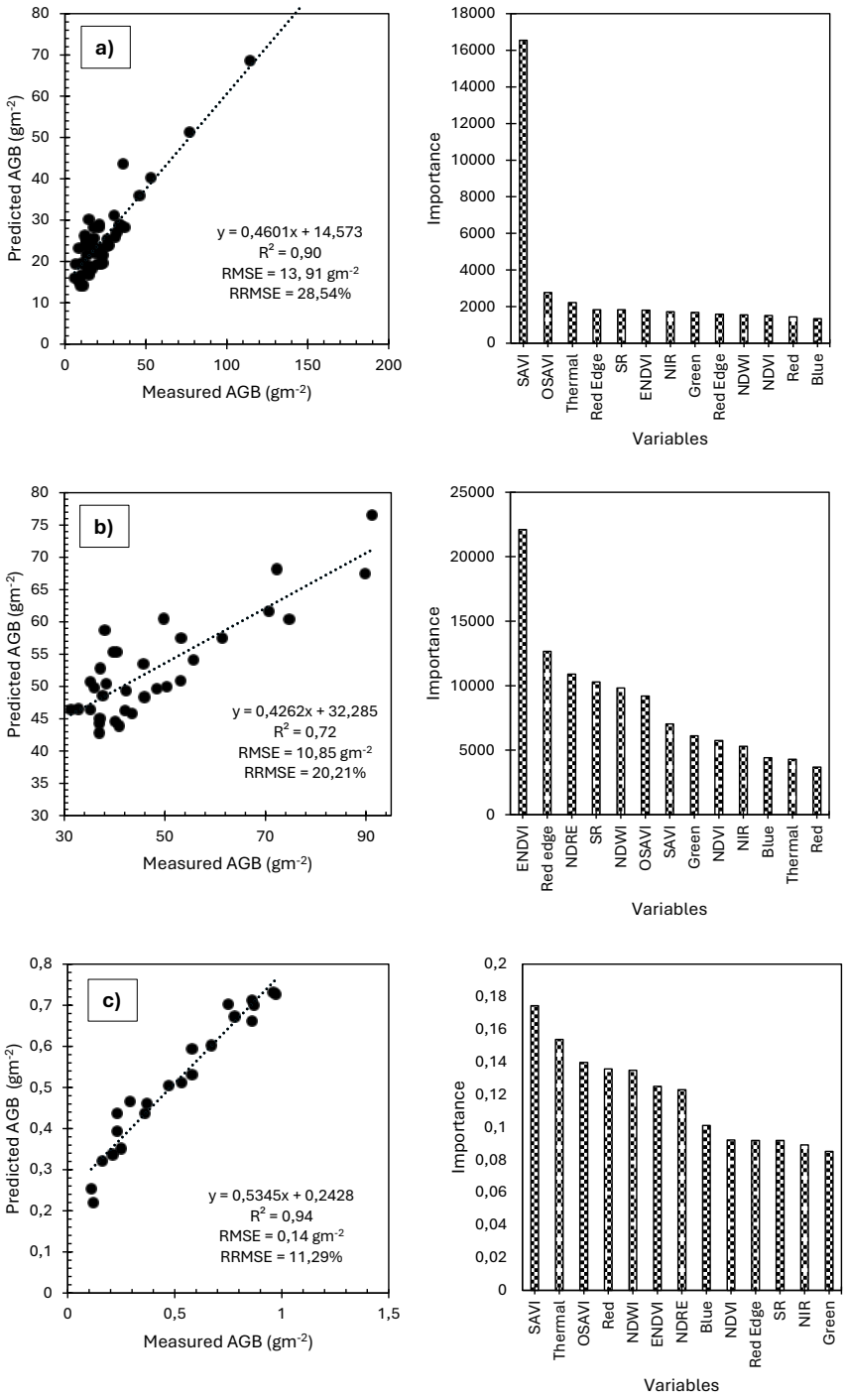


Figure 8-3 Relationship between the predicted and observed AGB in the (a) establishment stage, (b) the vegetative stage, and (c) the maturity stage of taro derived from a combination of bands and Vis, using the RF algorithm and the model variable importance scores

When predicting the BGB, the establishment stage model produced a model with a prediction accuracy with an R^2 of 0.92 and an RMSE of 16.06 g/m^2. The optimal model for estimating the BGB was established during the vegetative growth stage, which exhibited an R^2 of 0.96 and the lowest RMSE of 7.26 g/m^2 (Figure 8.4). The most suitable predictor variables during the

establishment and maturity stages included SAVI and OSAVI (Figure 8.4). At the vegetative stage, the red-edge, ENDVI and NDRE were the most influential predictor variables, in order of importance. The BGB prediction accuracies drastically depreciated for the maturity growth stage, which exhibited an R^2 of 0.93 and an RMSE of 21.35 g/m^2, based on OSAVI, thermal, and SAVI, in order of importance.

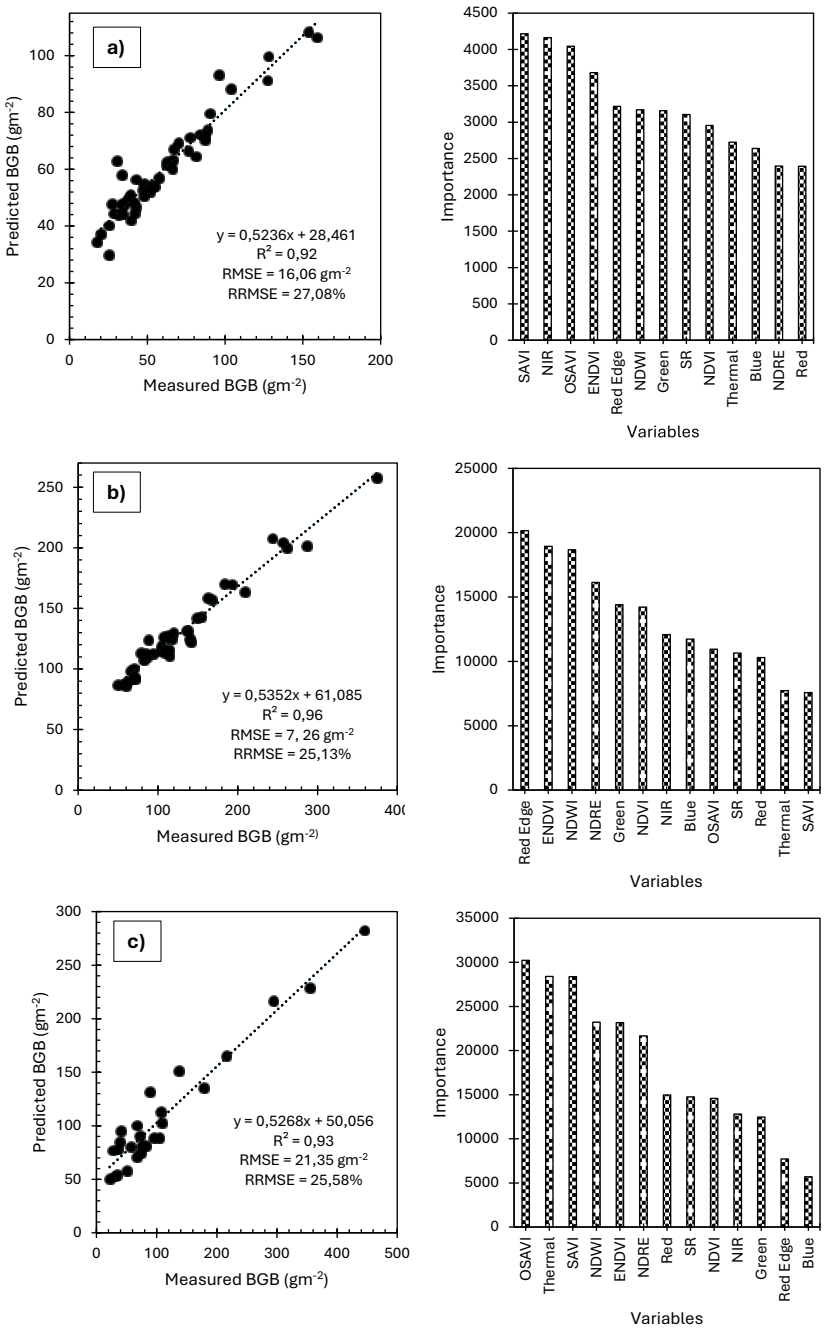


Figure 8-4 Relationship between the predicted and observed BGB in the (a) establishment stage, (b) vegetative stage and (c) maturity stage of taro derived from a combination of bands and Vis, using the RF algorithm and the model variable importance scores

When comparing the performance of the AGB and BGB models across all the growth stages, the results did not show significant variations (Figure 8.5). For instance, when estimating the biomass at the establishment growth stage, the AGB model exhibited the poorest prediction accuracy with an rRMSE of 28.54%, followed by an improvement in the BGB model with an rRMSE of 27.08% (Figure 8.5). Similarly, the most optimal model for estimating the biomass during the vegetative stage was the AGB model with an rRMSE of 20.21%, followed by the BGB model, with a rRMSE of 25.13% (Figure 8.5). During the maturity stage, the AGB model yielded the most optimal biomass prediction accuracy, with an rRMSE of 11.29%, followed by a drop in the AGB model (25.58%) (Figure 8.5).

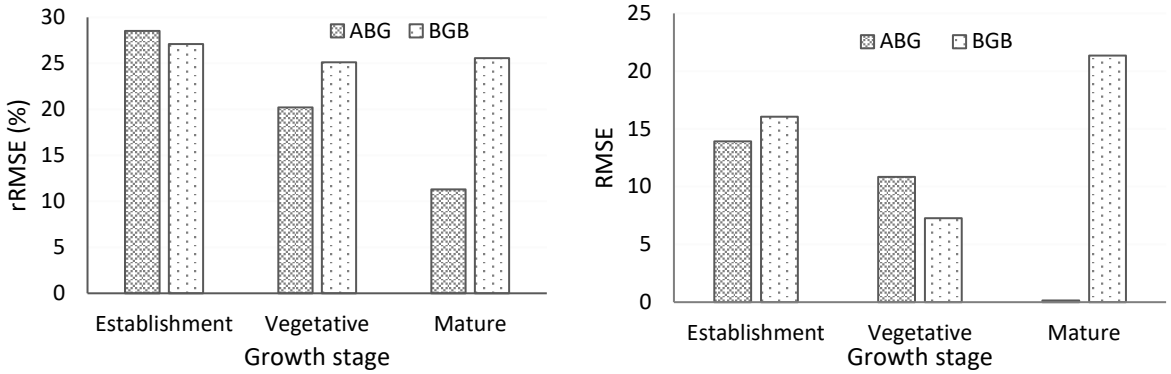


Figure 8-5 Comparison of accuracies ((a) rRMSE & (b) RMSE) derived in estimating the AGB and BGB of taro using UAV-acquired remotely-sensed data

8.3.4 Modelling the spatial distribution of the AGB and BGB of taro

The spatial distribution of AGB and BGB was estimated based on the optimal models. Figure 8.6 illustrates the spatial distribution of taro's AGB and BGB. It can be observed that the AGB is relatively low during the establishment stage, followed by an increase in the vegetative stage and a decrease during the maturity stage. Compared to the AGB, the BGB is also low at the establishment stage, lower than the AGB, followed by an increase during the vegetative and maturity stages, and resulting in a high BGB at the last two stages.

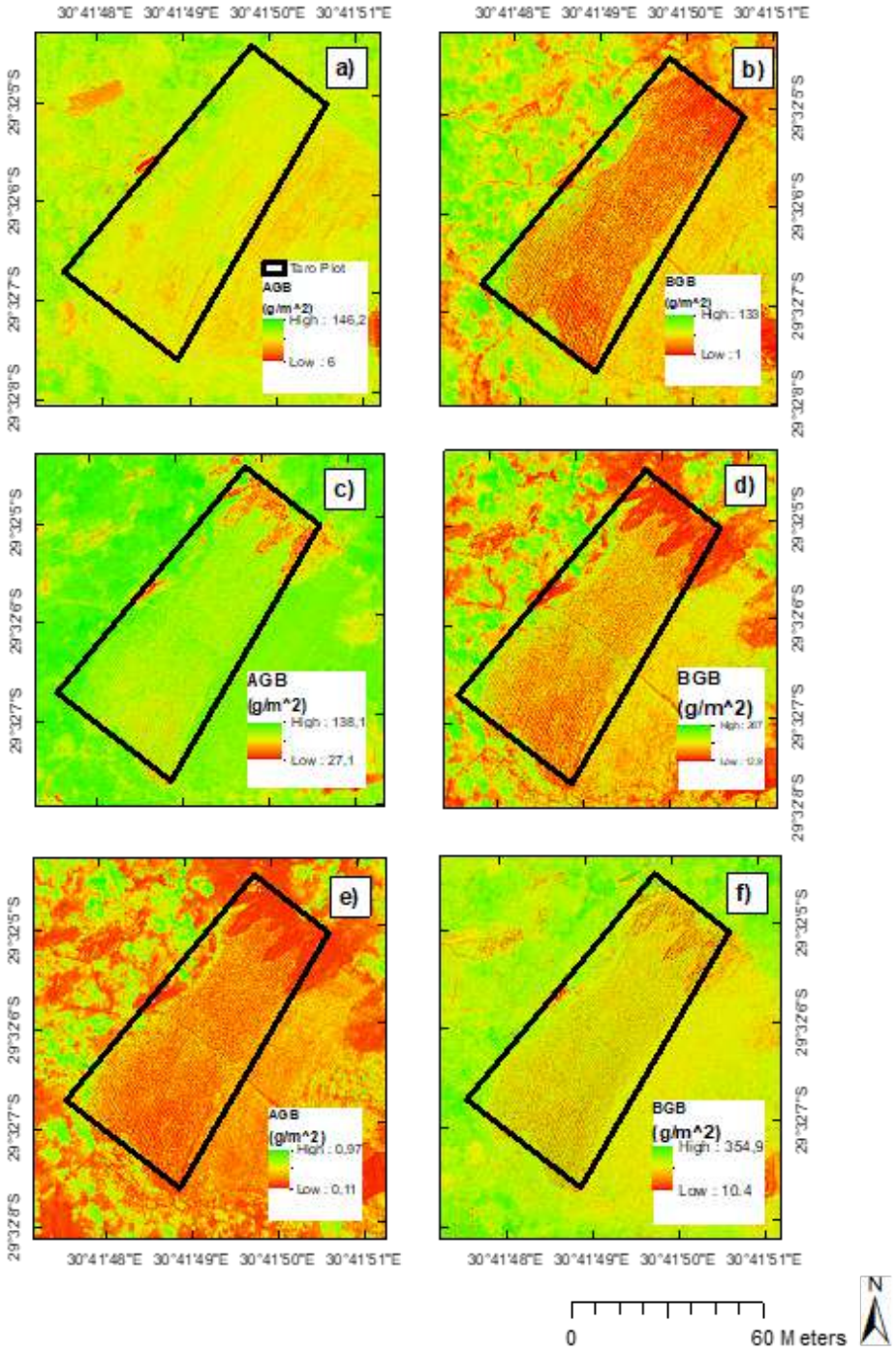


Figure 8-6 The spatial distribution of the modelled AGB (a, c and e) and BGB (b, d and f) of smallholder taro crops

8.4 Discussion

The study aimed to test the utility of UAV multi-spectral remotely-sensed data in estimating the AGB and BGB of taro. In doing so, we assessed the relationship between taro's AGB and BGB and identified the optimum phenological stage for the prediction of BGB before harvest. It is evident that the AGB and BGB of taro varied over the phenological stages, and Random Forest could discern the optimal growth stage for estimating the BGB of taro before the harvest. The biomass variations of taro over the growing season are useful for estimating the

productivity status of the smallholder crop field. In this regard, our findings can assist in management strategies that are aimed at improving taro production.

8.4.1 Characterising the taro AGB and BGB across the growing season using UAV-derived data

The findings of the models in this study performed well throughout the growth stages. The establishment stage yielded an RMSE of 13.91 g/m^2, an R^2 of 0.90 and an rRMSE of 28.54% for AGB, and an RMSE of 16.06 g/m^2, an R^2 of 0.92 and an rRMSE of 27.08% for BGB. These results were based on SAVI, OSAVI, thermal and NIR, as the most influential variables. Specifically, these variables were identified as crucial in the model's prediction of AGB and BGB at this stage, because of their ability to surpass the variable background effects. The establishment stage (58 DAP) is normally characterised by a small leaf area and adventitious root production (Prikaziuk *et al.* 2022). This causes some limitations when monitoring the early stage by using remote sensing, where areas with a low canopy cover can be misinterpreted due to the soil background effect, which varies with the soil texture, organic matter and moisture (Jasim *et al.* 2020, Elsayed *et al.* 2021). Thus, NIR, SAVI and OSAVI facilitated the optimal taro AGB and BGB prediction at this stage. Low biomass concentrations were associated with the crop establishment stage in this study. The leaf-area index was low, resulting in higher photosynthetic rates in the crop and facilitating high reflectance in the NIR region. The low foliage density of the establishment growth stage did not saturate and cause spectral confusion of the sensor during image acquisition. Furthermore, the relationship between the biomass, NIR and its derived vegetation indices has been noted to be critical in crop production assessments (Atanasov *et al.* 2025).During the vegetative growth stage, the major variables of importance were derived from the red, red-edge and NIR regions; however, the green waveband was also identified as a significant contributor to the model estimation. This is attributed to the high biomass accumulation and an increase in the leaf-foliage density (Moussa Kourouma *et al.* 2021). The AGB rapidly increased at this stage, from a mean of 22.91 g/m^2 during the establishment stage, to 54.21 g/m^2 during the vegetative stage. Similarly, the BGB also rapidly increased at this stage from a mean of 30.38 g/m^2 during the establishment stage, to 129.12 g/m^2 during the vegetative stage. Moreover, the red, red-edge, and NIR wavelengths are sensitive to plants with high biomass and chlorophyll content concentrations, and they are thus better predicted by using these wavebands (Li *et al.* 2021). A high photosynthetic activity rate and chlorophyll content are usually observed during this growth period, thereby enhancing the response of the red-edge and NIR bands, which are most sensitive to chlorophyll. Consequently, the predictor variables sensitive to changes in chlorophyll content may be useful to define management strategies and to explore the productive potential of taro, as this crop responds well to intensive management practices. In addition, the NIR regions strongly influence the prediction of biomass, as it is sensitive to the high foliar reflectance induced by the pigment concentrations of the plant canopy structure (Ramírez *et al.* 2014). Furthermore, the literature has stated that a moderate-to-high increase in biomass during tuber growth results in sensitivity in the red and green bands (Vitor *et al.* 2019). This also explains why the green band was identified as a variable that was significantly important only during the vegetative growth stage.

Meanwhile, low AGB concentrations and high BGB concentrations were associated with the mature growth stage. This is where tuberisation increases the root mass until the end of the shoot growth. Thereafter, the nutrients in the leaves are transported to the tuberous roots until leaf senescence (Wang *et al.* 2021). At the end of this growth stage, the rainfed field reached senescence and its leaf area was reduced after yellowing and leaf fall, thereby decreasing the canopy reflectance. This explains the importance of SAVI, OSAVI and the thermal band in predicting taro's AGB and BGB at this stage. Although SAVI and OSAVI are composed of the same spectral bands as NDVI, they have a factor to correct the soil background effect. During this growth stage, there were regions in the production field with a higher proportion of exposed soils, which explains the higher importance of these predictor variables, compared to the others. Furthermore, tuberisation is strongly associated with canopy growth and, in turn, it can be influenced by environmental factors, such as the temperature and the availability of water and nutrients, which explains the importance of the thermal band, which is sensitive to surface heat and moisture content (Al-Gaadi *et al.* 2018).

8.4.2 The most optimal stage and variables for the prediction of taro's BGB yield

To identify the most optimal stage for the prediction of taro's BGB, based on taro's AGB, the correlation between AGB and BGB and the model accuracies and their variations were considered. At the establishment stage, the best-fit model was the AGB model, followed closely by the BGB model, with SAVI, OSAVI, NIR and the thermal band being the most important variables for biomass prediction at this stage. The significance of these variables in the prediction model of taro's AGB at this stage can be attributed to the fact that this is the initial stage of the crop, where the crop canopy is significantly low, which results in significant soil exposure (Ma *et al.* 2021). This then results in SAVI and OSAVI being crucial for suppressing the soil background effects and a high reflectance of the soil in the NIR and thermal wavelengths.

At the vegetative stage, the best-fit model was again the AGB model, followed closely by the BGB model. The most influential predictor variables for biomass prediction at this stage were the red-edge band and ENDVI. The influence of these predictor variables in biomass prediction at this stage could be explained by the good relationship between ENDVI, the biomass and yield at mid-density canopies before saturation (Mhango *et al.* 2021). Furthermore, the literature highlights that the red-edge band relates to chlorophyll, which directly relates to the biomass and yield (Sibanda *et al.* 2015, Nyawade *et al.* 2021). Mid-density canopies are generally characterised by high biomass concentrations, which equate to high chlorophyll concentrations and which are sensitive to the red-edge waveband (Sibanda *et al.* 2019). Lastly, at the mature stage, the best-fit model was the AGB model, followed at a distance by the BGB model, with SAVI, OSAVI and the thermal band being critical in the prediction of biomass at this stage. The influence of these variables could be attributed to the fact that, at this stage of the crop, the leaves have reached senescence, and most have fallen off, which results in significantly low leaf foliage and thus soil exposure. In this regard, SAVI and OSAVI were critical for eliminating this soil background effect.

Overall, the vegetative state (154 DAP) proved to have higher prediction accuracies for AGB and BGB, and it had less variation between the AGB and BGB model accuracies, when compared to the establishment and mature stages of the crop. Based on our findings, this is the most optimal stage at which the taro yield could be predicted before harvesting. Estimating the yield at this stage may seem premature, since taro harvesting occurs towards the end of the maturity stage. However, it is promising since, as stated above, the state of the crop at the establishment stage determines its subsequent growth and yield (Li *et al.* 2021). The most significant variables for the optimal yield model were the red edge band and ENDVI. This reveals that the key wavelengths for yield prediction are located across the visible and near infrared range. The findings of this study are in line with those of Li *et al.* (2020) and Tedesco *et al.* (2021), who demonstrated the use of variables derived from the visible and near infrared regions of the electromagnetic spectrum in estimating the AGB and yield of potato crops.

8.4.3 Implications of the findings

Taro is a promising smart future crop for addressing the looming food and nutrition insecurity, as it has the potential to produce high yields and can contribute immensely to improving the resilience of food systems in SSA. However, the crop is a NUS in most of SSA. Thus, such findings demonstrate the use of precision agricultural technologies (i.e. UAVs), which can facilitate improved taro production. Specifically, using UAV-derived data to monitor the biomass dynamics of taro across the growing season will help to remotely identify areas with a higher incidence of roots that will be economically important, which consequently implies higher yields, and those requiring intervention. The growth dynamics could also assist in determining the most appropriate time to start harvesting, when considering the workforce that is available. Therefore, near-real-time UAV technology is beneficial in taro production as it allows for rapid and informed decision-making to limit further crop deficit issues.

8.4.4 Limitations and recommendations

Despite the high spatial resolution offered by the UAV onboard sensor, its spectral resolution limited the choice of spectral derivatives that could be generated to optimise the model's accuracy. A higher spectral resolution allows for more precise spectral extraction, especially during the establishment and maturity growth stages, when the crop canopy cover is low and there is a significant presence of soil, and moreover, when there is a presence of weeds. There is a need for future studies to consider other approaches when generating yield prediction models, such as masking or image segmentation. This may result in improved model performance and more detailed map outputs of biomass across the field. Nonetheless, the application of UAV-derived data, in conjunction with the RF algorithm, proved to be instrumental in assisting taro farmers in near-real-time, to make the necessary management decisions.

8.5 Conclusion

This study aimed to test the use of UAV multi-spectral remotely-sensed data in estimating the AGB and BGB of taro, with the following specific objectives: (1) to explore the utility of using

UAV image data in assessing and characterising the relationship between taro's AGB and BGB, and (2) to identify the optimum phenological stage for the prediction of BGB before harvest. The following conclusions were drawn:

- UAV image data can be used to assess and characterise the relationship between the AGB and BGB of taro;
- There is a positive relationship between taro's AGB and BGB across the growing season; and
- UAV-derived data optimally predicted taro's BGB before the harvest, at the vegetative stage.

The characterised variations in AGB and BGB can assist farmers and decision makers in identifying deficits in crop production within the field, in order to adjust their farming practices and achieve maximum farm productivity. These findings highlight the applicability of UAV multi-spectral image data in the prediction of the BGB, as well as the use of UAV systems to optimise crop production through precision farming on smallholder farms, which is crucial for poverty alleviation and improving food and nutritional security.

9 INTEGRATING IMAGE SEGMENTATION AND LEAF AREA INDEX WITH UAV MULTISPECTRAL DATA TO IMPROVE TARO YIELD PREDICTION ON SMALLHOLDER FARMS

9.1 Introduction

Taro is one of the most important traditional staple food crops grown on smallholder farms in Asia, the Pacific Islands and the Caribbean. It is extensively cultivated on a subsistence scale to ensure household income generation and food security, particularly in disadvantaged communities (Moussa Kourouma *et al.* 2021). However, smallholder farmers generally lack adequate the resources and information to optimise their agricultural productivity. Thus, it is imperative to provide them with innovative, effective and low-cost solutions to optimise their productivity and to produce increased and healthy taro yields. The early prediction of the taro yield could assist smallholder farmers to identify yield deficits at an early stage, which would allow them to implement the necessary remedial solutions to optimise their production. Yield prediction is commonly dependent on subjective, inaccurate and labour-intensive ground-based visits (Li *et al.* 2021). Remote sensing, on the other hand, has been proven to be a useful alternative for monitoring the growing season crop canopies and for providing information on the spatial variability of a crop yield. More specifically, UAV-based imaging is a low-cost solution that can be used for yield prediction. For instance, Li *et al.* (2021) improved the prediction of the potato yield by using UAV-derived data in conjunction with cultivar information. Moreover, Sun *et al.* (2020) predicted the potato yield by using UAV-based hyperspectral imagery and machine learning.

In comparison to the ground-based and satellite-based remote sensing techniques, UAV-based imaging offers satisfactory temporal, spatial and spectral resolution data (Yin *et al.* 2023). Thus, using UAV-derived data could assist in achieving timely and accurate taro yield predictions. UAV imaging methods for crop yield prediction are commonly based on broad-band Vegetation Indices (VIs), such as the Normalised Difference Vegetation Index (NDVI) (Tedesco *et al.* 2021). Although these VIs are related to the yield, some of them can be affected by other factors, including the soil background, light conditions and the presence of weeds, which results in low prediction accuracies, especially during the early stages of the growing season, where the crop is characterised by a low canopy cover (Luo *et al.* 2020, Liu *et al.* 2021). Consequently, improving the accuracy of these prediction models remains a challenge to their operational use. An alternative solution that has been demonstrated to be effective in eliminating the soil background and weed effects, while improving the prediction accuracy, has been the incorporation of VIs. These include the Soil Adjusted Vegetation Index (SAVI), the Optimised Soil Adjusted Vegetation Index (OSAVI), the Excess Green Index (ExGI) and the Excess Red index (ExRI), which are designed to address the influence of the soil brightness in areas where the vegetation cover is low. Another alternative approach includes incorporating image enhancement techniques (Kastens *et al.* 2005). Image segmentation techniques, or colour indices, have the great potential of circumventing the soil background noise, while minimising the impact of the light conditions, which significantly improves the crop signals and the associated prediction accuracies. For instance, Woebbecke *et al.* (1995) proposed the

Excess Green (ExGI) and Excess Red (ExRI) indices for enhancing the detection of plant pixels in digital images. These indices leverage the distinct reflectance properties of green plants in the visible (red and green) spectrum, which optimises the discrimination of live plants from non-plants. Zeng *et al.* (2021) demonstrated that the wheat yield could be optimally estimated by using colour indices during the flowering stage (R^2 = 0.93, RMSE = 32.18 g/m^2) by using UAV-acquired remotely-sensed data. Although these colour indices have been demonstrated to be effective in detecting and mapping various crop elements, their application in the yield mapping of NUS, such as taro, remain underexplored.

Moreover, yield mapping is generally associated with the crop canopy cover parameters, which include the Leaf Area Index (LAI), for optimising the prediction accuracies (Sun *et al.* 2020). The relative contribution of canopy cover traits, such as LAI, has not been extensively explored in taro crop production. Recent studies on incorporating multi-source data in the yield prediction of other tuber crops have achieved better results than studies that are based on remotely sensed data alone (Li *et al.* 2021, Mhango *et al.* 2021, Tedesco *et al.* 2021).

Other than the ultra-spatial resolution of UAV-acquired remotely-sensed data, in conjunction with image enhancement techniques and multisource data, machine learning algorithms have been demonstrated to be effective in optimising the prediction accuracies of crop yields in smallholder croplands (Sibanda *et al.* 2023). In particular, Random Forest has been extensively demonstrated to be robust and effective in optimising yield estimation accuracies (Ramos *et al.* 2020, Li *et al.* 2021, Cheng *et al.* 2022). For instance, Li *et al.* (2021), demonstrated that combining high spatial-resolution UAV images and cultivar information, using machine learning algorithms, can significantly improve potato yield prediction. However, the capabilities of Random Forest in estimating the taro yield in smallholder croplands still needs to be explored.

In this regard, using multisource data in conjunction LAI data could assist in optimising the yield forecasting techniques for fragmented and heterogeneous smallholder farmers, by building models of area-specific inputs that are relevant to a crop of interest that has influential crop traits. This is critical and urgently required, in order to achieve resilient and timely crop early warning systems. Therefore, this study aimed to test the potential of UAV-derived data, in concert with image segmentation and LAI, in predicting the taro yield in smallholder croplands. The specific objectives of this study were: (1) to predict and characterise the taro yield by using UAV-acquired data and image segmentation; (2) to evaluate the potential of improving taro yield mapping by incorporating LAI into the model.

9.2 Materials and Methods

9.2.1 Image acquisition and pre-processing

The study area and the experimental fields are detailed in Chapter One. Multi-spectral and thermal imaging data were obtained under clear sky conditions on the 11th April 2023, approximately 154 Days After Planting (DAP) during the vegetative stage of the crop. The vegetative stage was chosen because of the strong correlation between the above-ground

biomass of tuberous crops during this period and their final yield, as was proven in the literature (Wang and Rosen 2003, Tanabe *et al.* 2019, Vitor *et al.* 2019). The image was taken by a lightweight UAV (DJI Matrice 300 series) (Figure 9.1a) equipped with the MicaSense Altum imaging sensor (Figure 9.1b) at a flight altitude of 100 m, which is equivalent to a spatial resolution of 9.6 cm/pixel. The flight survey was configured with an 80% side and an 80% forward overlap. Radiometric and geometric corrections were applied to the raw image strips by using corresponding onboard navigation information and in-situ grey-white reflectance calibration panels for each flight, in order to produce georeferenced reflectance images. The imagery and corresponding Position and Orientation System (POS) data were used to generate an ortho-mosaic image of the site by using the Pix4D software.

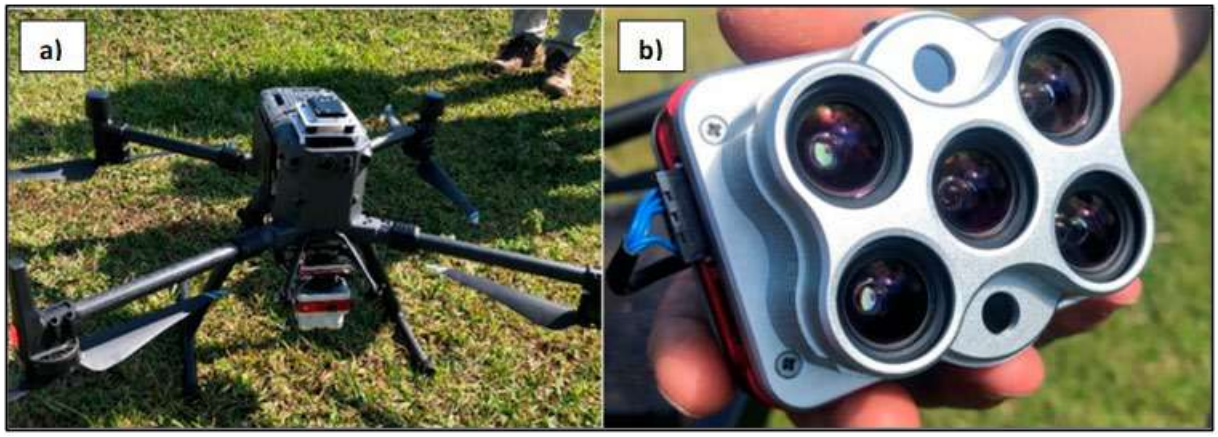


Figure 9-1 a) DJI Matrice 300 platform, and b) Micasense Altum multispectral and thermal sensor

9.2.2 Field crop assessment

Field measurements were conducted on the 11th of April 2023 and 14th of July 2023 during the vegetative stage and harvesting period, respectively, to provide ground truth data. Eighty plants were randomly selected across the experimental plot. The LAI was measured by using the LiCOR plant analyser, and the yield was measured by harvesting the crop. The corresponding dry weight was obtained after the fresh samples were oven-dried at 70°C until a constant dry weight was reached (approximately 72 h). The yield data were measured by weighing the total weight of each of the 80 collected taro tubers.

9.2.3 Image processing and data extraction

To extract the spectra that corresponded to the taro canopy, and to segment the image between the soil, weeds and taro, it was necessary to create a taro crop mask. In order to achieve this, a binary mask image was generated by segmenting the taro canopy from the soil background and weeds. The Excess Green (ExG), Excess Red (ExR)and Excess Green minus Excess Red (ExGR) colour indices were computed and compared. The ExR index was ultimately used as a robust index that facilitated a contrast enhancement between the taro canopy and the soil background, as well as the weeds:

$$ExR = 1.4 \times (R - G) / (G + R + B) \quad (1)$$

where R, G and B are the reflectance intensities in the red, green and blue regions, respectively. The Otsu thresholding method (Otsu 1979) was applied to convert the ExG greyscale image to a binary image, with zero value assigned to soil background and weeds, and the spectra were extracted from the non-zero pixels (taro mask) as a region of interest.

9.2.4 Vegetation indices

VIs are mathematical transformations of the spectra at predefined wavelengths. With the use of multi-spectral sensors, many broadband VIs have been developed in recent years for estimating the biophysical parameters of crops. Several vegetation indices have been applied to tuber crops for estimating their leaf area index, chlorophyll content, biomass and yield (Liu *et al.* 2021, Nyawade *et al.* 2021, Yang *et al.* 2022, Yu *et al.* 2023). Based on these studies, 11 vegetation indices (Table 9.1) that showed a good correlation with the biophysical parameters, crop yield and the biomass of tuber crops, were selected for use in this study.

Table 9-1 List of vegetation indices used in the modelling of yield and related source references

Index	Full Name	Formula	Reference
NDVI	Normalised Difference Vegetation Index	NIR – Red/NIR+ Red	(Tan *et al.* 2020)
ENDVI	Enhanced Normalised Difference Vegetation Index	((NIR + Green) – (2*Blue)) / ((NIR + Green) + (2*Blue))	(Dhau *et al.* 2018)
NDRE	Normalised Difference Red-Edge Index	NIR – Red-edge / NIR + Red-edge	(Sibanda *et al.* 2017)
NDWI	Normalised Difference Water Index	Green – NIR / Green + NIR	(Ndlovu *et al.* 2021)
SR	Simple Ratio	NIR/R	(Brewer *et al.* 2022)
SAVI	Soil Adjusted Vegetation Index	((NIR – Red) / (NIR + Red + 0.5)) * (1.5)	(Al-Gaadi *et al.* 2016)
OSAVI	Optimised Soil Adjusted Vegetation Index	(NIR – Red) / (NIR +Red + 0.16)	(Al-Gaadi *et al.* 2018)

9.2.5 Data analysis

The RF algorithm was used to predict taro yield because of its simplicity and robustness, which enables it to perform well, despite the sample size (Gómez *et al.* 2019, Sun *et al.* 2020). The RF ensemble is a machine learning algorithm that builds multiple trees on a subset of samples from the training data, by using bootstrap aggregation (Li *et al.* 2020). Decision trees are grown

to their threshold through a randomised subset of predictors (UAV image data), and each node is split by using random subsets of input variables. Moreover, the RF algorithm can rank the predictor variables, from the most to the least influential, in the prediction model, based on the sum of the reduction in Gini impurity across the nodes of the feature (Liu *et al.* 2022). Specifically, the R statistical software was used to develop three RF regression models through numerical inputs. Before running the RF algorithm, the data (n = 80) were split into training and test groups, with a split ratio of 70:30. The training data were used to train the models, while the testing data were used to validate the derived models. In tuning, the *mtry*, which is the number of variables that are randomly sampled as candidates at each split, and the *ntree*, which is the number of trees to grow, were used in this study to train each model. All the models were user-defined and fine-tuned to an optimal 150 trees and three variables. These hyperparameters were attained for numerous iterations. Although a similar number of *ntree* and *mtry* were used for all the models, the model calibration and validation were stage-specific.

9.2.6 Accuracy assessment of taro yield

Accuracy assessments were conducted to evaluate the performance of the regression models of the predicted yield. The accuracy metrics that were used were the coefficient of determination (R^2), the Root Mean Square Error (RMSE) and the Relative Root Mean Square Error (rRMSE). The R^2 measured the variation between the measured and predicted yield. The RMSE assessed the error magnitude between the field measurements and the modelled yield outputs, while the rRMSE evaluated the accuracy of the model, and it was used to compare the performance of the two regression models. The rRMSE is calculated by normalising the mean of each variable RMSE value, and it is expressed as a percentage, where lower percentages are considered to be more accurate (Yadav *et al.* 2021).

9.2.7 Comparing the performance of different datasets

In order to predict the taro yield, a comparative analysis of the performance of three datasets was conducted. The first model was developed by using predictor variables that were based on the multispectral and thermal bands and vegetation indices derived from the unsegmented UAV image. This was done to test the potential of the vegetation indices that were targeted at correcting the influence of soil brightness in areas where the vegetation cover was low or non-existent. The second model was developed by using predictor variables, based on the multispectral and thermal bands and vegetation indices derived from the segmented AV image. This model was sought to assess the potential of image segmentation for improving the yield forecasts. The final model was computed from UAV-acquired multispectral bands and vegetation indices that were derived after image segmentation and combined with LAI. This model sought to evaluate the strength and potential of ultra-high spatial resolution UAV-acquired data, in concert with the image segmentation and Leaf Area Index (LAI), for improving the yield prediction of tuber crops.

9.3 Results

9.3.1 Ground truth data

The minimum, maximum, mean and standard deviation of the yield data are shown in Table 9.2. Significant variations were observed in the taro yield across the plot. The average for the taro yield was 166.67g. The minimum and maximum yields were 71.47g and 523.1g, respectively.

Table 9-2 Descriptive statistics of taro yield

Parameter	Min	Max	Mean	Median	Std
Yield	71.47g	523.1g	166.67g	139g	81.34

9.3.2 *Evaluation of taro yield regression models*

When predicting the taro yield, RF Model 3 demonstrated the highest prediction accuracy (R^2 = 0.90 and RMSE 10.29 g/m^2) with the LAI, red-edge and NIR bands being the most optimal predictor variables. RF Models 1 and 2 also produced satisfactory results with an R^2 of 0.90 and 0.92, and an RMSE of 39.7 g/m^2 and 23.63 g/m^2, respectively with red, red-edge, NIR, ExR, ExGR and LAI being the most influential predictor variables (Figure 9.2).

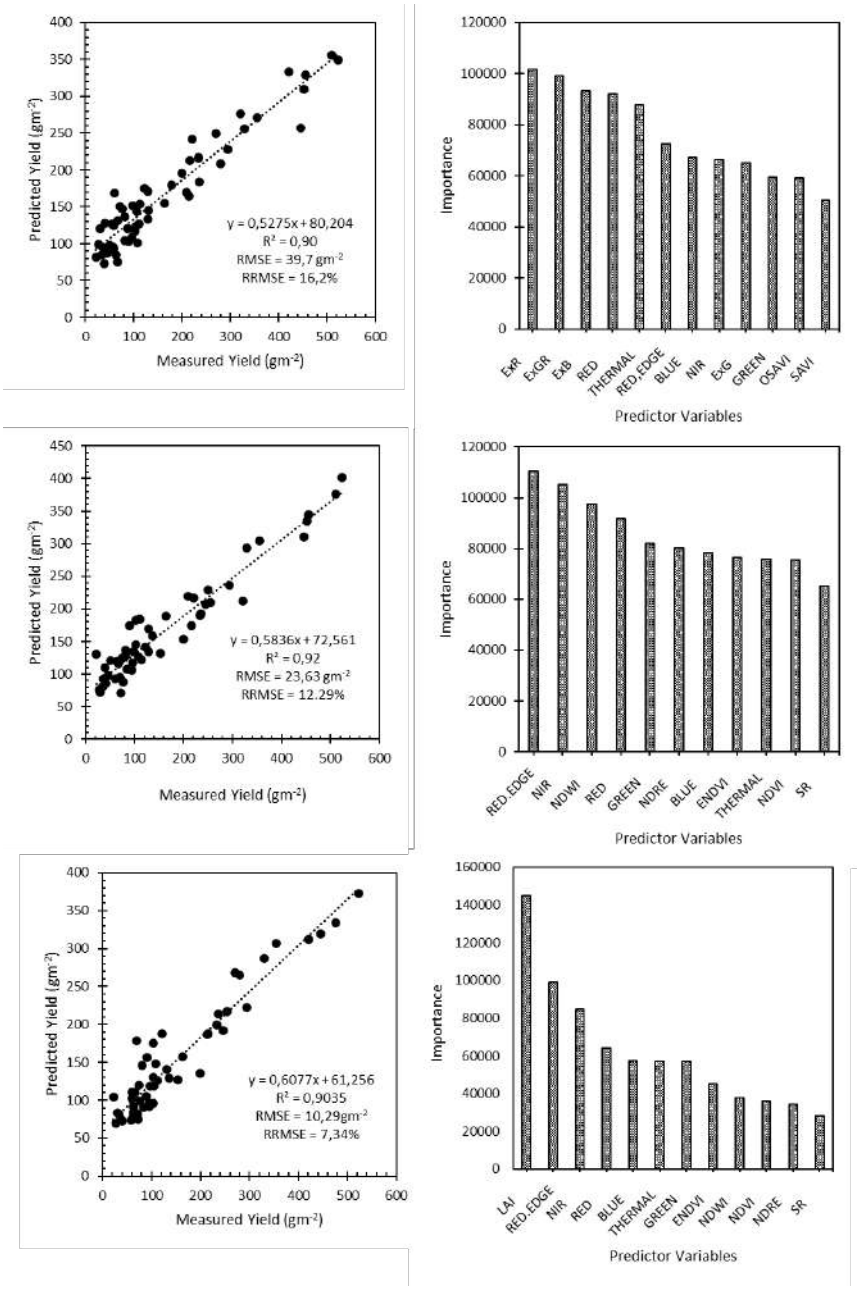


Figure 9-2 Prediction of taro yield using UAV image data from (a) unsegmented image, (b) segmented image, (c) segmented image and LAI, and (i, ii and iii) are the associated variable importance scores

9.3.3 Mapping the spatial distribution of the taro yield

The spatial distribution of the taro yield was predicted, based on the most optimal model, i.e. the RF model. The modelled yield ranged from 71 g/m^2 to 401 g/m^2. Figure 9.3 illustrates the spatial distribution of taro yield. It can be observed that the yield is relatively lower on the upper portions of the plot (the south portion of the plot map), followed by an increase in the mid- and lower portions (the mid- and north portions of the plot map) of the plot.

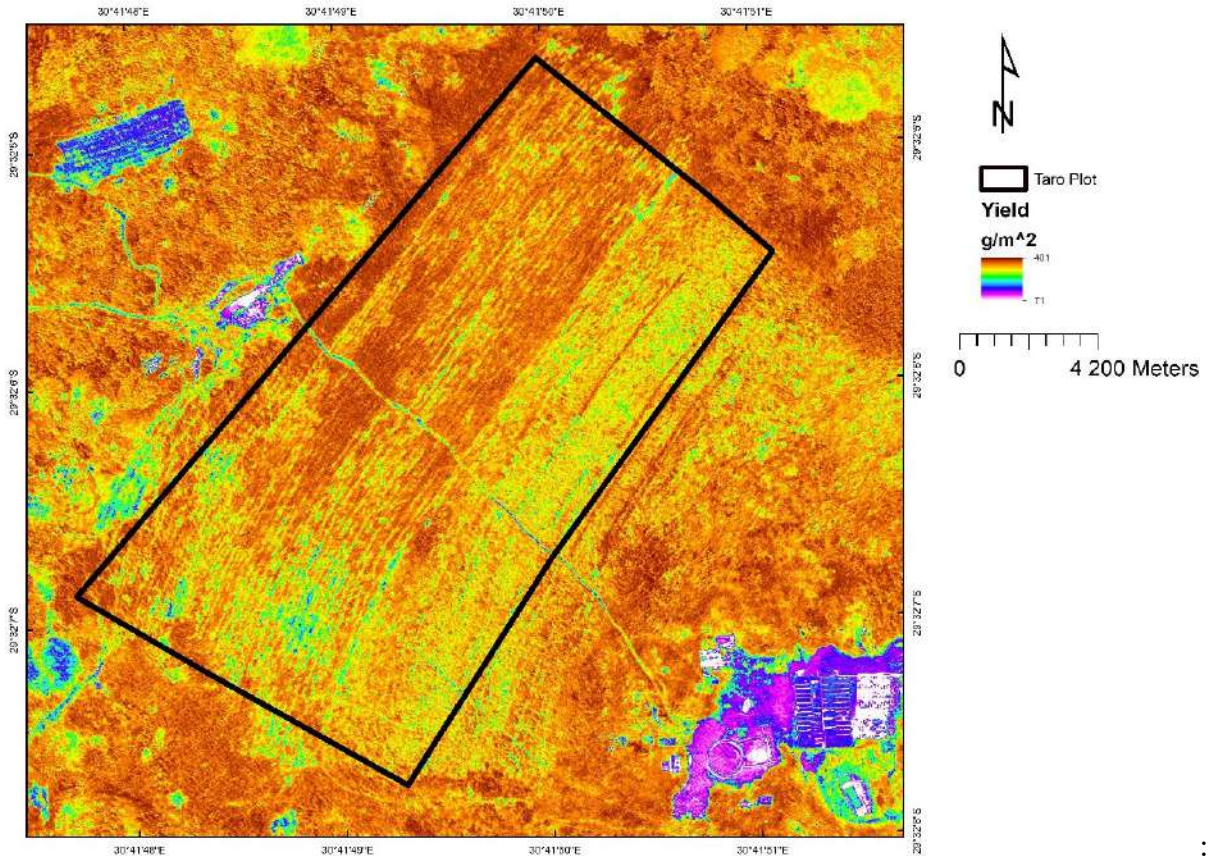


Figure 9-3 Spatial distribution of modelled taro yield

9.4 Discussion

The objective of this study was to determine the potential of UAV-derived data, in concert with image segmentation and LAI, for predicting the taro yield in smallholder croplands. To address this overarching aim, the prediction performance of segmented data, in concert with LAI, was compared with that of unsegmented data.

9.4.1 Prediction of taro yield in smallholder croplands

In this study, the taro yield was optimally estimated to an RMSE of 1029 g/m^2 based on the LAI, combined with the red-edge, NIR and red bands derived from a segment UAV-acquired multispectral image. The optimal performance of LAI could be explained by the fact that it is one of the principal components of a plant that is central to the process of photosynthesis, yet it is directly correlated with biomass accumulation. In particular, the LAI regulates the capability of a crop to intercept sunlight, which drives photosynthesis. With regard to tuber crops, photosynthetic efficiency directly regulates the accumulation of biomass in the tubers up to the mature phenological stage. Therefore, a higher LAI often positively correlates with a higher biomass and tuber yield. Generally, aa high foliage coverage and development (LAI) suggest efficient carbon assimilation, which directly contributes to tuber growth (Li *et al.* 2023). In a related study, Luo *et al.* (2020) demonstrated that the optimal combination weighting method, combined with LAI, estimated the potato yield better (with an adjusted R^2 value of 0.8333, and the estimation error of about 8%), when compared to other models. They

indicated that the inclusion of LAI in their model significantly improved the accuracy in predicting the yield of potato crops, because the saturation phenomena of the spectra were reduced.

Meanwhile, the optimal performance of the red-edge NIR and red sections of the electromagnetic spectrum could be attributed to the foliar attributes of the crops across the different stages. The edge section of the electromagnetic spectrum is renowned for being highly sensitive to subtle changes in the canopy architecture of crops and plants. For chlorophyll content, the absorption in the red section of the electromagnetic spectrum increases and the red-edge inflexion point gravitates towards the longer wavelengths, in relation to biomass accumulation. Meanwhile, as the foliage density increases (LAI), the reflectance in the near infrared and red-edge energy is also incrementally scattered by the leaves, which optimises the spectral signal of crops in the red-edge region and increases its sensitivity to changes in the LAI. In another related study, Li *et al.* (2020) showed that the combination of the robust Random Forest regression ensemble with the UAV-acquired red-edge spectral derivatives improved the estimation of both above- and below-ground biomass of potatoes with a coefficient of determination (R^2) > 0.90.

9.4.2 Assessing the impact of image segmentation on prediction accuracies

When comparing the effect of segmentation on the estimation of the taro yield, a prediction accuracy improvement to a magnitude of 3.59 g/m^2 was observed, when the image was segmented. This indicates that segmentation significantly improves the estimation of the taro yield. This could be explained by the fact that segmentation optimises the spectral signature of crops by excluding the non-crop regions on the image, thereby exclusively offering the data from the crops, which enhances the model's performance. Comparatively, an unsegmented image tends to be impacted by noise from other background features, including noise, which taints the spectral signature of crops and reduces the model's performance in predicting the crop yield. The results of this study are consistent with those of Li *et al.* (2020), who demonstrated that image segmentation significantly improved the estimation accuracy of potato yields.

The incorporation of LAI as a predictor variable among the spectral variables derived from a segmented ultra-high spatial resolution UAV-acquired image, significantly improved the yield prediction accuracies of taro with a magnitude of 13.34 g/m^2. This could be attributed to the combination of the pure crop signatures from segmentation with the foliage density data, which is strongly correlated with a crop's biomass accumulation. Subsequently, the magnitude of agreement between the taro yield and its predictor variables is extensively magnified, which reduces errors, while it significantly improves the prediction accuracy (Li *et al.* 2020, Luo *et al.* 2020).

9.4.3 Implications of the study

The results of this study suggest that the segmentation of UAV-acquired remotely-sensed data, in concert with machine learning algorithms and the biophysical elements of crops, such as

LAI, has great potential for predicting the yield of tuber crops during the growing season in smallholder crops at near-real-time. This implies that the use of these techniques could facilitate improved and accurate yield forecasting, which could indirectly optimise crop production, in the light of the current climate change and food insecurities. However, in this study, a limited number of phenological stages were considered for estimating the yield. This may not adequately represent the variability and contribution of the changes in the canopy traits across the growing season. Thus, it is recommended that a further analysis of the model's performance should be undertaken in the future, by using more crop canopy traits as input variables.

9.5 Conclusion

The study tested the potential of UAV-derived data, in concert with image segmentation and LAI, for improving the early prediction of the taro yield in smallholder croplands, by using the Random Forest regression algorithm. The findings of this study concluded that:

- the taro yield can be optimally predicted by using Random Forest in concert with the red, red-edge and NIR bands derived from a segmented high spatial resolution UAV-acquired image, and combined with LAI in smallholder croplands; and
- image enhancement techniques, such as segmentation, combined with data on the crop canopy architecture, can significantly improve the accuracy in predicting the tuber yield of taro crops in smallholder croplands.

The findings of the study demonstrate that precise UAV crop phenotyping and advanced image enhancement techniques, which are time- and cost-effective, are suitable for smallholder farmers and could improve informed decision-making and agricultural productivity in climate-vulnerable regions. Specifically, UAV image data, LAI and the image segmentation technique are spatially explicit and provide near-real-time data for understanding crop productivity. This approach potentially overcomes some of the limitations that are associated with the prediction of crop yield, by using satellite-borne remotely-sensed data.

10 GENERAL DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

The human population in South Africa is anticipated to rapidly increase to between 79 and 87 million by 2050, which will intensify such challenges as food and nutrition security, water scarcity and land resources. This implies that South Africa will need to produce at least one-third more food, in order to feed the rapidly-growing population. The growing population, combined with climate change and land degradation, will place increased pressure on the already-limited water resources. Furthermore, climate variability and frequent droughts associated with El Niño-Southern Oscillation (ENSO), are already impacting mainstream crops, such as maize (Pretorius and Geyser 2025). However, some of our most resilient crops are also the most neglected. Sorghum, cowpea, amaranth, sweet potatoes and taro are naturally tolerant to both droughts and floods. Their champions are smallholder farmers, who rely on this inherited wisdom. The challenge, and the opportunity, is for modern science to fully catch up to their potential.

Despite being subjected to droughts, climate variability and change, which drastically impact their productivity, smallholder farms play a critical role in global food production, by producing about 70-80% of the world's food (Ricciardi *et al.* 2018). There is a critical gap between the diverse reality of smallholder farms and the generic guidelines that are meant to support them. To build real climate resilience, precise and timely field data are necessary. Traditional crop monitoring is slow and subjective, while satellite imagery is often blurry or too costly to capture the fine details of these heterogeneous and fragmented landscapes. Therefore, there is an urgent need to develop time-efficient and spatially explicit approaches for high-throughput phenotyping. This would enable the precise monitoring of the crop health, water use and yield forecasts, in order to support on-farm decision-making and optimise production in the light of climate variability and change.

Proximal remote sensing by means of UAV-based high-throughput phenotyping platforms has become a cost-effective, rapid and efficient approach for acquiring high-resolution crop data. These platforms have demonstrated considerable potential for the characterisation of neglected and under-utilised species (NUS)(Ndlovu *et al.* 2021, Ndlovu *et al.* 2025). The proposed project leveraged UAV-derived imagery, in combination with crop simulation models, to assess the crop health and estimate yields of selected NUS, namely sweet potatoes, taro and Bambara groundnuts, within smallholder farming systems. The near-real-time monitoring of the crop–soil water status will facilitate drought assessments and it will support informed irrigation management. In addition, this study will build on WRC Project No. K5/2717//4, which focused on estimating the green water use of maize by integrating UAV-based phenotyping to enhance precision agriculture applications in smallholder environments. This project aimed to assess the application of unmanned aerial vehicle high-throughput phenotyping of Neglected and Under-utilised crop Species (NUS) (taro and sweet potatoes) for their improved water use and productivity on smallholder farms. In order to achieve the overarching objective, this project set out the specific objectives below:

1) to review the literature on the utility of earth observation data in characterising the productivity of Neglected and Under-utilised crop Species on smallholder farms;
2) to map the spatial distribution and health of neglected and under-utilised crop species using UAV data on smallholder farms;
3) to assess the relationship between NUS crop productivity variations and edaphic factors (i.e. soil moisture temperature and salinity) on smallholder farms;
4) to evaluate the utility of unmanned aerial vehicle platforms in quantifying the spatial and temporal variability of the canopy crop water status and yield by NUS crops on smallholder farms; and
5) to assess the potential of unmanned aerial vehicle high-throughput phenotyping data in estimating the yield of NUS crops on smallholder croplands.

10.1 General Discussion

A systematic review of the literature was undertaken on the progress, challenges, gaps and opportunities associated with using drone-derived remotely-sensed data to map the spatial distribution and health of NUS crops, in order to specifically address Objective 1; however, a notable imbalance was found in the literature. While several studies have focused on assessing crop health and productivity, very few have mapped where these crops are spatially located. The findings implied that overall, research on the NUS remains limited, compared to the research on staple crops. This gap is partly explained by the high costs of drones and pilot licensing, a shortage of trained personnel and the restrictive regulations. Despite these barriers, the review highlights the strong potential of UAVs to deliver detailed, timely and reliable spatial data for the monitoring of crops. The findings stress an urgent need to integrate drone technology into smallholder farming systems. Doing so, would equip farmers with the skills to operate drones and to interpret the data, which would lead to better on-farm decisions and increased productivity. Beyond the farm-scale benefits, the wider adoption of drones for NUS could help to sustain indigenous food systems and cultural heritage, to support gender equality, since women often manage these crops, and to empower smallholder farmers as the custodians of agricultural biodiversity. Ultimately, this study frames drone technology as not just a technical tool, but as a catalyst for ecological, social and cultural resilience in farming communities.

In order to specifically address Objective 2, the study compared the machine learning algorithms (Gradient Tree Boosting, Random Forest and Support Vector Machine) for mapping sweet potatoes and taro in smallholder fields by using UAV data. It was found that Gradient Tree Boosting performed best, particularly when using vegetation indices derived from the Red-edge and NIR bands, rather than the raw spectral bands. The resulting maps provide accurate, field-scale data that support targeted farm management and the creation of reliable agricultural landcover maps for improving food security in sub-Saharan Africa.

In order to specifically address Objective 3, the project explored the utility of drone data and Random Forest regression to predict the chlorophyll content of taro and sweet potato crops in smallholder systems. The findings indicate that combining multispectral bands and vegetation indices, particularly from the NIR and red-edge regions, enables the accurate prediction of the

canopy chlorophyll content, which is a key health indicator, during the key vegetative stages. These results confirmed the viability of drone-derived data as a precise, high-resolution tool for monitoring the health and condition of these under-utilised crops. In addition, this project demonstrated that UAV data can effectively predict the soil moisture and chlorophyll content in taro crops at specific growth stages, with the optimal estimation of soil moisture in the early establishment phase and the estimation of chlorophyll in the late vegetative stage, using key spectral bands and indices. A significant positive correlation was found between these two parameters, which indicates that the chlorophyll content can act as a reliable proxy for soil moisture. These findings highlight the practical utility of drone-based remote sensing for precise, stage-specific crop- and soil moisture-monitoring in smallholder farming systems.

In order to specifically address Objective 4, this project aimed to assess the use of thermal remote sensing and index-based segmentation techniques for predicting the canopy Equivalent Water Thickness (EWTcanopy) of taro crops on smallholder farms. The findings revealed that an integrated approach, using UAV-based thermal and multispectral data, combined with the ExGR segmentation method, provided the most accurate estimates of the taro water status. Notably, the exclusion of thermal data reduced the prediction accuracy, while segmentation effectively isolated the crop canopies from soil interference. The most influential spectral regions for modelling the EWTcanopy were the near-infrared, red-edge and thermal bands. These results highlight the significant potential of UAV remote sensing to deliver timely, high-resolution insights into the crop water status, which offers a practical tool for water management in smallholder systems that are cultivating neglected and under-utilised species, such as taro. In addition, the project explored the use of UAV thermal-multispectral data and deep learning to assess the water status of taro crops on smallholder farms. The results were highly promising, as they demonstrated that combining thermal and multispectral data can optimally estimate the key crop water indicators, including the equivalent water thickness, fuel moisture content, stomatal conductance and canopy temperature, with a very high accuracy ($R^2 > 0.91$). Integrating raw spectral bands with thermal and spectral indices proved to be the most effective modelling approach, as it showcases a practical and precise method for monitoring taro health.

In order to specifically address Objective 5, when assessing the potential of unmanned aerial vehicle high-throughput phenotyping data for estimating the yield of NUS crops, the project demonstrated that UAV multispectral data can reliably estimate both the above- and below-ground biomass in taro crops. It revealed a positive correlation between the two across the growing season and it identified the vegetative stage as being optimal for predicting the below-ground yield before harvest. These findings highlight the practical value of UAVs for precision farming, which will enable smallholders to monitor their crop performance, adjust their practices, and ultimately enhance their productivity, food security and livelihoods. Finally, the project also confirmed that UAV-derived data, when combined with image segmentation and the Leaf Area Index (LAI), can effectively provide early and accurate predictions of the taro yield. Using the red, red-edge and NIR spectral bands with a Random Forest model, this approach overcomes the limitations of satellite data by delivering timely, high-resolution insights. This cost-effective method equips smallholder farmers with precise, field-scale

information to improve their decision-making and boost their productivity in climate-vulnerable regions.

10.2 Limitations

While this project established significant insights on drone remotely-sensed data high-throughput phenotyping of the neglected and under-utilised taro and sweet potato crops, its findings were shaped by numerous methodological and contextual constraints, namely:

The limited scope of the crop assessment: The project was constrained to only two crop species (taro and sweet potatoes) and it did not differentiate between the varieties within these species. This limits the understanding of varietal-specific traits, resilience and performance.

The limited analytical methodology: The classification and modelling procedures were restricted to standard machine learning algorithms within the Google Earth Engine platform. The project did not explore, or leverage, more advanced computational techniques, such as deep learning or artificial intelligence, which could improve the accuracy and scalability.

The lack of an integrated data synergy: There was a missed opportunity to synergise high-resolution drone (UAV) data with freely-available, moderate-resolution satellite data (e.g. Sentinel-2, Landsat). This limits the ability to effectively upscale the findings from a field plot to a regional landscape.

The incomplete seasonal and phenological data: The analysis did not fully capture the crop traits across all key phenological stages, particularly during early establishment and late maturity. This was mainly due to bad weather conditions, a shortage of pilots, and technical repairs to the drone. This gap reduces the robustness and seasonal accuracy of the yield estimation models.

The sensor and data quality constraints: The Altum MicaSense sensor is limited in its spectral resolution. During the critical growth stages with sparse canopy cover, the project was limited by the spectral resolution of the sensors that were used and the lack of advanced data processing (e.g. vegetation-soil masking) to filter out weeds. This challenge mostly affected the discrimination of crops and yield modelling.

10.3 General Conclusion

Based on the systematic review, this study concludes that drone technology holds significant potential for advancing the research on Neglected and Under-utilised Species (NUS) crops, particularly in mapping their spatial distribution - an area that is currently lacking in the literature. However, its widespread adoption is hindered by the high costs, regulatory barriers and a shortage of technical skills, especially within the smallholder systems. Overcoming these challenges is essential, as integrating UAVs into farming practices would not only improve crop monitoring and productivity, but it would also empower communities, preserve the indigenous food systems and promote greater gender equity in agriculture. In mapping the spatial distribution and health of Neglected and Under-utilised crop Species by using UAV

data on smallholder farms, this study demonstrated that Gradient Tree Boosting is an effective machine learning algorithm for accurately mapping sweet potatoes and taro. These precise, field-scale distribution maps provide actionable data to support on-farm decision-making and the development of reliable agricultural landcover maps. Ultimately, this approach is a valuable tool for targeting interventions that can enhance food security and the sustainable management of under-utilised crops. The project also successfully demonstrates that integrating UAV thermal data with multispectral imagery and ExGR segmentation provides the most accurate method for estimating the taro water status on smallholder farms. Excluding thermal information reduces the accuracy, while the red-edge, NIR and thermal bands are critical for prediction. This approach offers a practical, high-resolution tool for real-time water management in under-utilised crop systems. The findings of the project also demonstrated that integrating UAV thermal data with multispectral imagery and ExGR segmentation provides the most accurate method for estimating the water status of taro on smallholder farms. Excluding thermal information reduces the accuracy, while the red-edge, NIR and thermal bands are critical for prediction. This approach offers a practical, high-resolution tool for real-time water management in under-utilised crop systems. In mapping the crop productivity elements, such as chlorophyll, the findings of the project confirmed that drone-based remote sensing, combined with machine learning, provides a viable and precise method for monitoring the chlorophyll content, as a proxy for the health and productivity of under-utilised crops like taro and sweet potatoes in smallholder systems. Accurate predictions are achieved by integrating multispectral bands with vegetation indices, especially from the NIR and red-edge regions. This project also demonstrates that UAV-based remote sensing can effectively monitor both the soil moisture and chlorophyll content in taro crops across their growth cycle. The chlorophyll content serves as a reliable indicator of the soil moisture, which allows for the non-invasive, real-time assessment of the crop and soil health. By identifying key spectral variables for each growth stage, this approach provides smallholder farmers with a practical tool for precision agriculture and informed water management. Based on the findings, this project conclusively demonstrates that multispectral imagery can reliably estimate taro biomass, which reveals a direct link between above- and below-ground growth and identifies the vegetative stage as being the best time to predict yield. Finally, the project illustrated that UAV-based monitoring, and integrating specific spectral data and image analysis, provides an early, accurate and practical yield prediction tool for taro, which will enable smallholder farmers to make better decisions and increase their resilience.

Overall, the studies significantly addressed the overarching aim of the project by validating the precise UAV-based methods to map these crops, to monitor their water status and chlorophyll content, and to predict their biomass and yield. Collectively, this evidence establishes drone technology as a practical and transformative tool for smallholder agriculture. By generating precise, actionable data on the crop status and resource use, UAVs may enable farmers to optimise their water management and boost their productivity, which are foundational steps for improving their food security and climate adaptation.

10.4 Recommendations

- Future projects must move beyond species-level mapping to conduct the varietal-level phenotyping of taro and sweet potatoes. Deploying UAVs to assess the distinct varieties will uncover critical differences in their canopy architecture, stress response and yield potential. These data are essential for developing targeted breeding programmes and providing farmers with evidence-based recommendations for the most resilient and productive varieties that are suited to their local conditions.
- In order to accurately map the complex spatial patterns of NUS crops in Africa, the research must advance beyond basic machine learning. We recommend integrating big data analytics, including deep learning and AI, to process high-resolution UAV and satellite imagery. These advanced techniques are needed for automatically detecting the subtle spectral-spatial features of fragmented and highly-heterogeneous smallholder plots, which will significantly improve the classification accuracy and enable scalable, dynamic crop monitoring.
- It is recommended that future projects should expand the geospatial modelling of the canopy water status as a proxy for crop water use in NUS crops (taro and sweet potatoes), to conduct systematic variety and water management trials. Current research lacks a direct comparison between irrigated and rainfed production across genetically-diverse material. This gap hinders the identification of varieties that are inherently water-efficient, or best-suited to specific water regimes. Implementing targeted field studies that correlate canopy-level spectral data with actual water use and yield metrics will enable the development of variety-specific recommendations for optimising water use efficiency.
- Future projects could enhance yield estimation models by integrating multi-temporal canopy data that span all the major phenological stages. The current models often rely on limited snapshots, and they miss the critical seasonal variability in canopy development. Incorporating a broader suite of traits across the growing season is essential for robust and accurate predictions. Therefore, future projects should systematically evaluate the predictive power of these new input variables, to determine their individual and synergistic value for forecasting the yield.
- The yield model accuracy declines during the early and late growth stages, due to the sparse canopy cover and significant background interference from the soil and weeds. This could be addressed by integrating (1) sensors with a higher spectral resolution for precise spectral feature extraction, and (2) advanced data processing pipelines that feature vegetation-soil masking and image segmentation. These strategies will generate cleaner canopy data, enhance the predictive accuracy of a model, and support the creation of more reliable, high-resolution biomass and yield maps.
- In order to fully realise the value of taro and sweet potatoes, future studies must link agronomic practice to nutritional outcome. We recommend pioneering research that models how nutrients in these crops accumulate, or vary, across the growing season, in direct response to their water use. This goes beyond traditional phenotyping to answer a vital question: how can we manage water to enhance the nutritional quality?

Closing this knowledge gap is essential for scientifically validating and marketing these NUS crops as truly sustainable, nutrient-optimised foods.

- A significant opportunity remains to fuse drone-level detail with satellite-level coverage for the phenotyping of NUS crops. We recommend that future studies prioritise methodologies that integrate high-resolution drone data with the free, moderate-resolution time series from Sentinel-2 and Landsat. By building these synergies, researchers can create scalable models that translate precise, field-level observations into reliable landscape-level insights, which will dramatically improve crop monitoring and management for smallholder farmers.

REFERENCES

Abbas, F., H. Afzaal, A. A. Farooque and S. Tang (2020). "Crop yield prediction through proximal sensing and machine learning algorithms." Agronomy **10**(7): 1046.

Abdel-Rahman, E. M., F. B. Ahmed and R. Ismail (2012). "Random forest regression and spectral band selection for estimating sugarcane leaf nitrogen concentration using EO-1 Hyperion hyperspectral data." International Journal of Remote Sensing **34**(2): 712–728.

Abdi, A. M. (2020). "Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data." GIScience & Remote Sensing **57**(1): 1–20.

Abou Ali, H., D. Delparte and L. M. Griffel (2020). FROM PIXEL to YIELD: FORECASTING POTATO PRODUCTIVITY in Lebanon and IDAHO. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, International Society for Photogrammetry and Remote Sensing.

Abrahams, M., M. Sibanda, T. Dube, V. G. Chimonyo and T. Mabhaudhi (2023). "A systematic review of UAV applications for mapping neglected and underutilised crop species' spatial distribution and health." Remote Sensing **15**(19): 4672.

Adam, E., O. Mutanga, E. M. Abdel-Rahman and R. Ismail (2014). "Estimating standing biomass in papyrus (Cyperus papyrus L.) swamp: exploratory of in situ hyperspectral indices and random forest regression." International Journal of Remote Sensing **35**(2): 693–714.

Adam, E., O. Mutanga, J. Odindi and E. M. Abdel-Rahman (2014). "Land-use/cover classification in a heterogeneous coastal landscape using RapidEye imagery: evaluating the performance of random forest and support vector machines classifiers." International Journal of Remote Sensing **35**(10): 3440–3458.

Adhikari, U., A. P. Nejadhashemi and S. A. Woznicki (2015). "Climate change and eastern Africa: a review of impact on major crops." Food and Energy Security **4**(2): 110–132.

Aguilar, R., R. Zurita-Milla, E. Izquierdo-Verdiguier and R. A. de By (2018). "A cloud-based multi-temporal ensemble classifier to map smallholder farming systems." Remote sensing **10**(5): 729.

Ahluwalia, O., P. C. Singh and R. Bhatia (2021). "A review on drought stress in plants: Implications, mitigation and the role of plant growth promoting rhizobacteria." Resources, Environment and Sustainability **5**: 100032.

Ahmad, U., A. Alvino and S. Marino (2021). "A review of crop water stress assessment using remote sensing." Remote Sensing **13**(20): 4155.

Al-Gaadi, K. A., A. A. Hassaballa, E. Tola, A. G. Kayad, R. Madugundu, B. Alblewi and F. Assiri (2016). "Prediction of potato crop yield using precision agriculture techniques." PloS one **11**(9).

Al-Gaadi, K. A., A. A. Hassaballa, E. Tola, A. G. Kayad, R. Madugundu, F. Assiri and B. Alblewi (2018). "Characterization of the spatial variability of surface topography and moisture content and its influence on potato crop yield." International journal of remote sensing **39**(23): 8572–8590.

Aliche, E. B., M. Oortwijn, T. P. Theeuwen, C. W. Bachem, R. G. Visser and C. G. van der Linden (2018). "Drought response in field grown potatoes and the interactions between canopy growth and yield." Agricultural Water Management **206**: 20–30.

Ampatzidis, Y. and V. Partel (2019). "UAV-based high throughput phenotyping in citrus utilizing multispectral imaging and artificial intelligence." Remote Sensing **11**(4): 410.

Anas, M., F. Liao, K. K. Verma, M. A. Sarwar, A. Mahmood, Z.-L. Chen, Q. Li, X.-P. Zeng, Y. Liu and Y.-R. Li (2020). "Fate of nitrogen in agriculture and environment: agronomic, eco-physiological and molecular approaches to improve nitrogen use efficiency." Biological Research **53**(1): 1–20.

Ansar, A. and I. Muhammad (2020). "Evaluating the potential of red edge position (REP) of hyperspectral remote sensing data for real time estimation of LAI & chlorophyll content of kinnow mandarin (Citrus reticulata) fruit orchards." Scientia Horticulturae **267**.

Asgari, S. and M. Hasanlou (2023). "A Comparative Study of Machine Learning Classifiers for Crop Type Mapping Using Vegetation Indices." ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences **10**: 79–85.

Atanasov, A. I., A. Z. Atanasov and B. I. Evstatiev (2025). "A remote sensing approach for biomass assessment in winter wheat using the NDVI second derivative in terms of NIR." Sustainability **17**(16): 7299.
Awad, M. M. (2019). "An innovative intelligent system based on remote sensing and mathematical models for improving crop yield estimation." Information Processing in Agriculture **6**(3): 316–325.
Awais, M., W. Li, M. Cheema, Q. Zaman, A. Shaheen, B. Aslam, W. Zhu, M. Ajmal, M. Faheem and S. Hussain (2022). "UAV-based remote sensing in plant stress imagine using high-resolution thermal sensor for digital agriculture practices: a meta-review." International Journal of Environmental Science and Technology: 1–18.
Ayele, N. A. and H. K. Tamiru (2020). "Developing classification model for chickpea types using machine learning algorithms." International Journal of Innovative Technology and Exploring Engineering **10**(1): 5–11.
Ballester, C., J. Brinkhoff, W. C. Quayle and J. Hornbuckle (2019). "Monitoring the effects of water stress in cotton using the green red vegetation index and red edge ratio." Remote Sensing **11**(7): 873.
Baluja, J., M. P. Diago, P. Balda, R. Zorer, F. Meggio, F. Morales and J. Tardaguila (2012). "Assessment of vineyard water status variability by thermal and multispectral imagery using an unmanned aerial vehicle (UAV)." Irrigation Science **30**: 511–522.
Barbedo, J. G. A. (2019). "A review on the use of unmanned aerial vehicles and imaging sensors for monitoring and assessing plant stresses." Drones **3**(2): 40.
Basile, A., R. Albrizio, D. Autovino, A. Bonfante, R. De Mascellis, F. Terribile and P. Giorio (2020). "A modelling approach to discriminate contributions of soil hydrological properties and slope gradient to water stress in Mediterranean vineyards." Agricultural Water Management **241**: 106338.
Bayat, B., C. van der Tol and W. Verhoef (2016). "Remote sensing of grass response to drought stress using spectroscopic techniques and canopy reflectance model inversion." Remote Sensing **8**(7): 557.
Belgiu, M. and L. Drăguţ (2016). "Random forest in remote sensing: A review of applications and future directions." ISPRS journal of photogrammetry and remote sensing **114**: 24–31.
Berni, J. A. J., P. J. Zarco-Tejada, G. Sepulcre-Cantó, E. Fereres and F. Villalobos (2009). "Mapping canopy conductance and CWSI in olive orchards using high resolution thermal remote sensing imagery."
Berra, E. F. and M. V. Peppa (2020). "Advances and Challenges of Uav Sfm Mvs Photogrammetry and Remote Sensing: Short Review." The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **XLII-3/W12-2020**: 267–272.
Bhattacharya, A. and A. Bhattacharya (2021). "Mineral nutrition of plants under soil water deficit condition: A Review." Soil water deficit and physiological issues in plants: 287–391.
Blumenthal, J., D. B. Megherbi and R. Lussier (2020). "Unsupervised machine learning via Hidden Markov Models for accurate clustering of plant stress levels based on imaged chlorophyll fluorescence profiles & their rate of change in time." Computers and Electronics in Agriculture **174**: 105064.
Bojović, B. and A. Marković (2009). "Correlation between nitrogen and chlorophyll content in wheat (Triticum aestivum L.)." Kragujevac Journal of Science **31**(5827): 69–74.
Bouguettaya, A., H. Zarzour, A. Kechida and A. M. Taberkit (2022). "Deep learning techniques to classify agricultural crops through UAV imagery: A review." Neural Computing and Applications **34**(12): 9511–9536.
Breiman, L. (2001). "Random forests." Machine learning **45**: 5–32.
Brewer, K., A. Clulow, M. Sibanda, S. Gokool, V. Naiken and T. Mabhaudhi (2022). "Predicting the Chlorophyll Content of Maize over Phenotyping as a Proxy for Crop Health in Smallholder Farming Systems." Remote Sensing.
Brewer, K., A. Clulow, M. Sibanda, S. Gokool, V. Naiken and T. Mabhaudhi (2022). "Predicting the Chlorophyll Content of Maize over Phenotyping as a Proxy for Crop Health in Smallholder Farming Systems." Remote Sensing **14**(3).

Brewer, K., A. Clulow, M. Sibanda, S. Gokool, V. Naiken and T. Mabhaudhi (2022). "Predicting the Chlorophyll Content of Maize over Phenotyping as a Proxy for Crop Health in Smallholder Farming Systems." Remote Sensing **14**(3): 518.
Brewer, K., A. Clulow, M. Sibanda, S. Gokool, J. Odindi, O. Mutanga, V. Naiken, V. G. Chimonyo and T. Mabhaudhi (2022). "Estimation of maize foliar temperature and stomatal conductance as indicators of water stress based on optical and thermal imagery acquired using an unmanned aerial vehicle (UAV) platform." Drones **6**(7): 169.
Brewer, K., A. Clulow, M. Sibanda, S. Gokool, J. Odindi, O. Mutanga, V. Naiken, V. G. P. Chimonyo and T. Mabhaudhi (2022). "Estimation of Maize Foliar Temperature and Stomatal Conductance as Indicators of Water Stress Based on Optical and Thermal Imagery Acquired Using an Unmanned Aerial Vehicle (UAV) Platform." Drones **6**(7): 169.
Brewer, K. R. (2021). Assessment of maize crop health and water stress based on multispectral and thermal infrared unmanned aerial vehicle phenotyping in smallholder farms.
Broge, N. H. and E. Leblanc (2001). "Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density." Remote Sensing of Environment **76**(2): 156–172.
Bukowiecki, J., T. Rose and H. Kage (2021). "Sentinel-2 data for precision agriculture?—A UAV-based assessment." Sensors **21**(8): 2861.
Buthelezi, S., O. Mutanga, M. Sibanda, J. Odindi, A. D. Clulow, V. G. P. Chimonyo and T. Mabhaudhi (2023). "Assessing the Prospects of Remote Sensing Maize Leaf Area Index Using UAV-Derived Multi-Spectral Data in Smallholder Farms across the Growing Season." Remote Sensing **15**(6): 1597.
Byambadorj, S.-O., J. O. Hernandez, S. Lkhagvasuren, G. Erma, K. Sharavdorj, B. B. Park and B. Nyam-Osor (2023). "Leaf morpho-physiological traits of Populus sibirica and Ulmus pumila in different irrigation regimes and fertilizer types." PeerJ **11**: e16107.
Cakir, R. (2004). "Effect of water stress at different development stages on vegetative and reproductive growth of corn." Field Crops Research **89**(1): 1–16.
Camargo Neto, J. (2004). "A combined statistical-soft computing approach for classification and mapping weed species in minimum-tillage systems."
Candiago, S., F. Remondino, M. De Giglio, M. Dubbini and M. Gattelli (2015). "Evaluating multispectral images and vegetation indices for precision farming applications from UAV images." Remote sensing **7**(4): 4026–4047.
Cao, Z., T. Cheng, X. Ma, Y. Tian, Y. Zhu, X. Yao, Q. Chen, S. Liu, Z. Guo, Q. Zhen and X. Li (2017). "A new three-band spectral index for mitigating the saturation in the estimation of leaf area index in wheat." International Journal of Remote Sensing **38**(13): 3865–3885.
Chai, H. H., F. Massawe and S. Mayes (2016). "Effects of mild drought stress on the morpho-physiological characteristics of a bambara groundnut segregating population." Euphytica **208**(2): 225–236.
Chang-Hua, J. U., T. I. A. N. Yong-Chao, Y. A. O. Xia, C. A. O. Wei-Xing, Z. H. U. Yan and D. Hannaway (2010). "Estimating leaf chlorophyll content using red edge parameters." Pedosphere **20**(5): 633–644.
Cheng, M., X. Jiao, L. Shi, J. Penuelas, L. Kumar, C. Nie, T. Wu, K. Liu, W. Wu and X. Jin (2022). "High-resolution crop yield and water productivity dataset generated using random forest and remote sensing." Scientific data **9**(1): 641.
Cheng, M., C. Sun, C. Nie, S. Liu, X. Yu, Y. Bai, Y. Liu, L. Meng, X. Jia, Y. Liu, L. Zhou, F. Nan, T. Cui and X. Jin (2023). "Evaluation of UAV-based drought indices for crop water conditions monitoring: A case study of summer maize." Agricultural Water Management **287**.
Chew, R., J. Rineer, R. Beach, M. O'Neil, N. Ujeneza, D. Lapidus, T. Miano, M. Hegarty-Craver, J. Polly and D. S. Temple (2020). "Deep neural networks and transfer learning for food crop identification in UAV images." Drones **4**(1): 7.
Chibarabada, T. P. (2018). Water use and nutritional water productivity of selected major and underutilised grain legumes.

Chicco, D., M. J. Warrens and G. Jurman (2021). "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation." PeerJ Computer Science **7**: e623.
Chimonyo, V. G., E. M. Wimalasiri, R. Kunz, A. T. Modi and T. Mabhaudhi (2020). "Optimizing traditional cropping systems under climate change: a case of maize landraces and Bambara groundnut." Frontiers in Sustainable Food Systems **4**: 562568.
Chivasa, W., O. Mutanga and C. Biradar (2020). "UAV-based multispectral phenotyping for disease resistance to accelerate crop improvement under changing climate conditions." Remote Sensing **12**(15): 2445.
Chivenge, P., T. Mabhaudhi, A. T. Modi and P. Mafongoya (2015). "The potential role of neglected and underutilised crop species as future crops under water scarce conditions in Sub-Saharan Africa." International journal of environmental research and public health **12**(6): 5685–5711.
Ciganda, V., A. Gitelson and J. Schepers (2009). "Non-destructive determination of maize leaf and canopy chlorophyll content." J Plant Physiol **166**(2): 157–167.
Clevers, J. G. and A. A. Gitelson (2013). "Remote estimation of crop and grass chlorophyll and nitrogen content using red-edge bands on Sentinel-2 and-3." International Journal of Applied Earth Observation and Geoinformation **23**: 344–351.
Colovic, M., K. Yu, M. Todorovic, V. Cantore, M. Hamze, R. Albrizio and A. M. Stellacci (2022). "Hyperspectral vegetation indices to assess water and nitrogen status of sweet maize crop." Agronomy **12**(9): 2181.
Crusiol, L. G. T., M. R. Nanni, R. H. Furlanetto, R. N. R. Sibaldelli, E. Cezar, L. M. Mertz-Henning, A. L. Nepomuceno, N. Neumaier and J. R. B. Farias (2020). "UAV-based thermal imaging in the assessment of water status of soybean plants." International Journal of Remote Sensing **41**(9): 3243–3265.
Curran, P. J., J. L. Dungan and H. L. Gholz (1990). "Exploring the relationship between reflectance red edge and chlorophyll content in slash pine " Tree Physiology(7): 33–48.
Daughtry, C. S., C. Walthall, M. Kim, E. B. De Colstoun and J. McMurtrey Iii (2000). "Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance." Remote sensing of Environment **74**(2): 229–239.
Daughtry, C. S., C. L. Walthall, M. S. Kim, E. B. De Colstoun and J. E. McMurtrey Iii (2000). "Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance." Remote sensing of Environment **74**(2): 229–239.
de Melo, L. L., V. G. M. L. de Melo, P. A. A. Marques, J. A. Frizzone, R. D. Coelho, R. A. F. Romero and T. H. da Silva Barros (2022). "Deep learning for identification of water deficits in sugarcane based on thermal images." Agricultural Water Management **272**: 107820.
Deepan, P. and L. Sudha (2020). Object classification of remote sensing image using deep convolutional neural network. The cognitive approach in cloud computing and internet of things technologies for surveillance tracking systems, Elsevier**:** 107–120.
Delegido, J., C. Vergara, J. Verrelst, S. Gandía and J. Moreno (2011). "Remote Estimation of Crop Chlorophyll Content by Means of High-Spectral-Resolution Reflectance Techniques." Agronomy journal **103**(6): 1834–1842.
Dhau, I., E. Adam, O. Mutanga, K. Ayisi, E. M. Abdel-Rahman, J. Odindi and M. J. G. I. Masocha (2018). "Testing the capability of spectral resolution of the new multispectral sensors on detecting the severity of grey leaf spot disease in maize crop." **33**(11): 1223–1236.
Dhillon, R. and Q. Moncur (2023). "Small-scale farming: A review of challenges and potential opportunities offered by technological advancements." Sustainability **15**(21): 15478.
Din, M., W. Zheng, M. Rashid, S. Wang and Z. Shi (2017). "Evaluating Hyperspectral Vegetation Indices for Leaf Area Index Estimation of Oryza sativa L. at Diverse Phenological Stages." Front Plant Sci **8**: 820.
Din, M. S. U., M. Mubeen, S. Hussain, A. Ahmad, N. Hussain, M. A. Ali, A. E. Sabagh, M. Elsabagh, G. M. Shah and S. A. Qaisrani (2022). World nations priorities on climate change and food security. Building Climate Resilience in Agriculture, Springer**:** 365–384.

Dondofema, F., N. Nethengwe, P. Taylor and A. Ramoelo (2023). "Comparison of Satellite Platform for Mapping the Distribution of Mauritius Thorn (Caesalpinia decapetala) and River Red Gum (Eucalyptus camaldulensis) in the Vhembe Biosphere Reserve." Remote Sensing **15**(11): 2753.
Duku, C., S. J. Zwart, L. G. van Bussel and L. Hein (2018). "Quantifying trade-offs between future yield levels, food availability and forest and woodland conservation in Benin." Science of the total environment **610**: 1581–1589.
Duveiller, G., R. López-Lozano and B. Baruth (2013). "Enhanced Processing of 1-km Spatial Resolution fAPAR Time Series for Sugarcane Yield Forecasting and Monitoring." Remote Sensing **5**(3): 1091–1116.
Ehlers, D., C. Wang, J. Coulston, Y. Zhang, T. Pavelsky, E. Frankenberg, C. Woodcock and C. Song (2022). "Mapping Forest Aboveground Biomass Using Multisource Remotely Sensed Data." Remote Sensing **14**(5): 1115.
El-Hendawy, S. E., N. A. Al-Suhaibani, S. Elsayed, W. M. Hassan, Y. H. Dewir, Y. Refay and K. A. Abdella (2019). "Potential of the existing and novel spectral reflectance indices for estimating the leaf water status and grain yield of spring wheat exposed to different irrigation rates." Agricultural Water Management **217**: 356–373.
El Bilali, H., G. Cardone, E. De Falcis, A. K. Naino Jika, S. Rokka, A. B. Diawara, B. Nouhou and A. Ghione (2023). "Neglected and underutilised species (NUS): an analysis of strengths, weaknesses, opportunities and threats (SWOT)."
Elmetwalli, A. H., T. Z. Fouda and E. Y. Ali (2014). High resolution satellite imagery to detect stress in potato. Acta Horticulturae, International Society for Horticultural Science.
Elsayed, S., S. El-Hendawy, M. Khadr, O. Elsherbiny, N. Al-Suhaibani, M. Alotaibi, M. U. Tahir and W. Darwish (2021). "Combining thermal and rgb imaging indices with multivariate and data-driven modeling to estimate the growth, water status, and yield of potato under different drip irrigation regimes." Remote Sensing **13**(9).
Elsayed, S., P. Rischbeck and U. Schmidhalter (2015). "Comparing the performance of active and passive reflectance sensors to assess the normalized relative canopy temperature and grain yield of drought-stressed barley cultivars." Field Crops Research **177**: 148–160.
Everitt, J., C. Yang, M. Davis, J. Everitt and M. Davis (2007). "Mapping wild taro with color-infrared aerial photography and image processing." J. Aquat. Plant Manage **45**: 106–110.
Fahad, S., A. A. Bajwa, U. Nazir, S. A. Anjum, A. Farooq, A. Zohaib, S. Sadia, W. Nasim, S. Adkins and S. Saud (2017). "Crop production under drought and heat stress: plant responses and management options." Frontiers in plant science **8**: 1147.
Falcioni, R., T. Moriwaki, M. Pattaro, R. Herrig Furlanetto, M. R. Nanni and W. Camargos Antunes (2020). "High resolution leaf spectral signature as a tool for foliar pigment estimation displaying potential for species differentiation." J Plant Physiol **249**: 153161.
Fan, S. and C. Rue (2020). "The role of smallholder farms in a changing world." The role of smallholder farms in food and nutrition security: 13–28.
Fan, S. and C. Rue (2020). The role of smallholder farms in a changing world. The role of smallholder farms in food and nutrition security, Springer, Cham**:** 13–28.
Farella, M. M., M. L. Barnes, D. D. Breshears, J. Mitchell, W. J. D. van Leeuwen and R. E. Gallery (2022). "Evaluation of vegetation indices and imaging spectroscopy to estimate foliar nitrogen across disparate biomes." Ecosphere **13**(3).
Feng, Z., Z. Cheng, L. Ren, B. Liu, C. Zhang, D. Zhao, H. Sun, H. Feng, H. Long and B. Xu (2024). "Real-time monitoring of maize phenology with the VI-RGS composite index using time-series UAV remote sensing images and meteorological data." Computers and Electronics in Agriculture **224**: 109212.
Filella, I. and J. Penuelas (1994). "The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status." International Journal of Remote Sensing **15**(7): 1459–1470.
Fitzgerald, G., D. Rodriguez, L. Christensen, R. Belford, V. Sadras and T. Clarke (2006). "Spectral and thermal sensing for nitrogen and water status in rainfed and irrigated wheat environments." Precision agriculture **7**: 233–248.

Fu, Z., P. Ciais, A. F. Feldman, P. Gentine, D. Makowski, I. C. Prentice, P. C. Stoy, A. Bastos and J.-P. Wigneron (2022). "Critical soil moisture thresholds of plant water stress in terrestrial ecosystems." Science Advances **8**(44): eabq7827.
Furtak, K. and A. Wolińska (2023). "The impact of extreme weather events as a consequence of climate change on the soil moisture and on the quality of the soil environment and agriculture–A review." Catena **231**: 107378.
Gago, J., A. R. Fernie, Z. Nikoloski, T. Tohge, S. Martorell, J. M. Escalona, M. Ribas-Carbo, J. Flexas and H. Medrano (2017). "Integrative field scale phenotyping for investigating metabolic components of water stress within a vineyard." Plant Methods **13**: 1–14.
Gamon, J., J. Penuelas and C. Field (1992). "A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency." Remote Sensing of environment **41**(1): 35–44.
Gao, B.-C. (1996). "NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space." Remote sensing of environment **58**(3): 257–266.
Gao, S., R. Zhong, K. Yan, X. Ma, X. Chen, J. Pu, S. Gao, J. Qi, G. Yin and R. B. Myneni (2023). "Evaluating the saturation effect of vegetation indices in forests using 3D radiative transfer simulations and satellite observations." Remote Sensing of Environment **295**.
García-Tejero, I., J. Costa, R. Egipto, V. Durán-Zuazo, R. Lima, C. Lopes and M. Chaves (2016). "Thermal data to monitor crop-water status in irrigated Mediterranean viticulture." Agricultural Water Management **176**: 80–90.
García-Tejero, I., A. Rubio, I. Viñuela, A. Hernández, S. Gutiérrez-Gordillo, C. Rodríguez-Pleguezuelo and V. Durán-Zuazo (2018). "Thermal imaging at plant level to assess the crop-water status in almond trees (cv. Guara) under deficit irrigation strategies." Agricultural water management **208**: 176–186.
Gavrilescu, M. (2021). "Water, soil, and plants interactions in a threatened environment." Water **13**(19): 2746.
Ge, X., J. Ding, X. Jin, J. Wang, X. Chen, X. Li, J. Liu and B. Xie (2021). "Estimating agricultural soil moisture content through UAV-based hyperspectral images in the arid region." Remote Sensing **13**(8): 1562.
Gerhards, M., M. Schlerf, K. Mallick and T. Udelhoven (2019). "Challenges and future perspectives of multi-/Hyperspectral thermal infrared remote sensing for crop water-stress detection: A review." Remote Sensing **11**(10): 1240.
Ghimire, B., J. Rogan and J. Miller (2010). "Contextual land-cover classification: incorporating spatial dependence in land-cover classification models using random forests and the Getis statistic." Remote Sensing Letters **1**(1): 45–54.
Gholamin, R. and M. Khayatnezhad (2020). "Assessment of the correlation between chlorophyll content and drought resistance in corn cultivars (Zea Mays)." Helix **10**(05): 93–97.
Giovos, R., D. Tassopoulos, D. Kalivas, N. Lougkos and A. Priovolou (2021). "Remote sensing vegetation indices in viticulture: A critical review." Agriculture **11**(5): 457.
Gitelson, A. and M. N. Merzlyak (1994). "Spectral reflectance changes associated with autumn senescence of Aesculus hippocastanum L. and Acer platanoides L. leaves. Spectral features and relation to chlorophyll estimation." Journal of plant physiology **143**(3): 286–292.
Gitelson, A. A., Y. Gritz and M. N. Merzlyak (2003). "Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves." J Plant Physiol **160**(3): 271–282.
Gitelson, A. A., Y. Gritz and M. N. Merzlyak (2003). "Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves." Journal of plant physiology **160**(3): 271–282.
Gitelson, A. A., Y. J. Kaufman and M. N. Merzlyak (1996). "Use of a green channel in remote sensing of global vegetation from EOS-MODIS." Remote sensing of Environment **58**(3): 289–298.

Gitelson, A. A., G. P. Keydan and M. N. Merzlyak (2006). "Three-band model for noninvasive estimation of chlorophyll, carotenoids, and anthocyanin contents in higher plant leaves." Geophysical research letters **33**(11).
Gómez, D., P. Salvador, J. Sanz and J. L. Casanova (2019). "Potato yield prediction using machine learning techniques and sentinel 2 data." Remote Sensing **11**(15): 1745.
Gomez y Paloma, S., L. Riesgo and K. Louhichi (2020). The Role of Smallholder Farms in Food and Nutrition Security, Springer Nature.
Goodbody, T. R. H., P. Tompalski, N. C. Coops, C. Hopkinson, P. Treitz and K. van Ewijk (2020). "Forest Inventory and Diversity Attribute Modelling Using Structural and Intensity Metrics from Multi-Spectral Airborne Laser Scanning Data." Remote Sensing **12**(13).
Gove, R. and J. Faytong (2012). Machine learning and event-based software testing: classifiers for identifying infeasible GUI event sequences. Advances in computers, Elsevier. **86:** 109–135.
Grote, M., A. Pilko, J. Scanlan, T. Cherrett, J. Dickinson, A. Smith, A. Oakey and G. Marsden (2022). "Sharing airspace with Uncrewed Aerial Vehicles (UAVs): Views of the General Aviation (GA) community." Journal of Air Transport Management **102**: 102218.
Guan, Y. and K. Grote (2023). "Assessing the Potential of UAV-Based Multispectral and Thermal Data to Estimate Soil Water Content Using Geophysical Methods." Remote Sensing **16**(1): 61.
Gunawardena, A., S. Nissanka, N. Dayawansa and T. Fernando (2015). "Estimation of above ground biomass in Horton Plains National Park, Sri Lanka using Optical, thermal and RADAR remote sensing data."
Guo, Y., G. Yin, H. Sun, H. Wang, S. Chen, J. Senthilnath, J. Wang and Y. Fu (2020). "Scaling Effects on Chlorophyll Content Estimations with RGB Camera Mounted on a UAV Platform Using Machine-Learning Methods." Sensors (Basel) **20**(18).
Guyot, G., F. Baret and S. Jacquemoud (1992). Imaging spectroscopy for vegetation studies, Kluwer Academic Publishers: Norwell, MA, USA.
Gxokwe, S. (2022). "Developing an integrated remotely sensed framework for the detection and monitoring of seasonally-flooded wetlands in semi-arid environments of southern Africa."
Haboudane, D. (2004). "Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture." Remote Sensing of Environment **90**(3): 337–352.
Haboudane, D., J. R. Miller, N. Tremblay, P. J. Zarco-Tejada and L. Dextraze (2002). "Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture." Remote sensing of environment **81**(2-3): 416–426.
Halme, E., P. Pellikka and M. Mõttus (2019). "Utility of hyperspectral compared to multispectral remote sensing data in estimating forest biomass and structure variables in Finnish boreal forest." International Journal of Applied Earth Observation and Geoinformation **83**.
Hamuda, E., M. Glavin and E. Jones (2016). "A survey of image processing techniques for plant extraction and segmentation in the field." Computers and electronics in agriculture **125**: 184–199.
Han, L., G. Yang, H. Dai, B. Xu, H. Yang, H. Feng, Z. Li and X. Yang (2019). "Modeling maize above-ground biomass based on machine learning approaches using UAV remote-sensing data." Plant Methods **15**: 10.
Han, Y., B. A. Tarakey, S. J. Hong, S. Y. Kim, E. Kim, C. H. Lee and G. Kim (2021). "Calibration and Image Processing of Aerial Thermal Image for UAV Application in Crop Water Stress Estimation."
Hassanijalilian, O., C. Igathinathane, C. Doetkott, S. Bajwa, J. Nowatzki and S. A. H. Esmaeili (2020). "Chlorophyll Estimation in Soybean Leaves Infield with Smartphone Digital Imaging and Machine Learning." Computers and Electronics in Agriculture **174**: 105433.
Hilary van Wyk, R. and E. Oscar Amonsou (2021). "Physiochemical and functional properties of albumin and globulin from amadumbe (Colocasia esculenta) corms." Food Science and Technology.
Hlophe-Ginindza, S. N. and N. Mpandeli (2020). "The role of small-scale farmers in ensuring food security in Africa." Food Security in Africa: 1–12.

Homolová, L., Z. Malenovský, J. G. Clevers, G. García-Santos and M. E. Schaepman (2013). "Review of optical-based remote sensing for plant trait mapping." Ecological Complexity **15**: 1–16.
Hornung, R. and A.-L. Boulesteix (2022). "Interaction forests: Identifying and exploiting interpretable quantitative and qualitative interaction effects." Computational Statistics & Data Analysis **171**: 107460.
Hou, M., F. Tian, L. Zhang, S. Li, T. Du, M. Huang and Y. Yuan (2018). "Estimating crop transpiration of soybean under different irrigation treatments using thermal infrared remote sensing imagery." Agronomy **9**(1): 8.
Hsu, C.-W., C.-C. Chang and C.-J. Lin (2003). A practical guide to support vector classification, Taipei, Taiwan.
Huang, C.-y., H.-L. Wei, J.-Y. Rau and J.-P. Jhan (2019). "Use of principal components of UAV-acquired narrow-band multispectral imagery to map the diverse low stature vegetation fAPAR." GIScience & remote sensing **56**(4): 605–623.
Huang, Z., Y. Liu, F.-P. Tian and G.-L. Wu (2020). "Soil water availability threshold indicator was determined by using plant physiological responses under drought conditions." Ecological Indicators **118**: 106740.
Huete, A. R. (1988). "A soil-adjusted vegetation index (SAVI)." Remote sensing of environment **25**(3): 295–309.
Hunt Jr, E. R., P. C. Doraiswamy, J. E. McMurtrey, C. S. Daughtry, E. M. Perry and B. Akhmedov (2013). "A visible band index for remote sensing leaf chlorophyll content at the canopy scale." International journal of applied earth observation and Geoinformation **21**: 103–112.
Hussain, S., K. Gao, M. Din, Y. Gao, Z. Shi and S. Wang (2020). "Assessment of UAV-Onboard Multispectral Sensor for non-destructive site-specific rapeseed crop phenotype variable at different phenological stages and resolutions." Remote Sensing **12**(3): 397.
Hutton, J., G. Lipa, D. Baustian, J. Sulik and R. Bruce (2020). "High accuracy direct georeferencing of the Altum multi-spectral UAV camera and its application to high throughput plant phenotyping." The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **43**: 451–456.
Iseki, K. and O. Olaleye (2020). "A new indicator of leaf stomatal conductance based on thermal imaging for field grown cowpea." Plant Production Science **23**(1): 136–147.
Jäger, T., A. Mokos, N. I. Prasianakis and S. Leyer (2022). "first_page settings Order Article Reprints Open AccessArticle Pore-Level Multiphase Simulations of Realistic Distillation Membranes for Water Desalination." Membranes.
Jasim, A., A. Zaeen, L. K. Sharma, S. K. Bali, C. Wang, A. Buzza and A. Alyokhin (2020). "Predicting phosphorus and potato yield using active and passive sensors." Agriculture **10**(11): 564.
Jewan, S. Y. Y., V. Pagay, L. Billa, S. D. Tyerman, D. Gautam, D. Sparkes, H. H. Chai and A. Singh (2022). "The feasibility of using a low-cost near-infrared, sensitive, consumer-grade digital camera mounted on a commercial UAV to assess Bambara groundnut yield." International Journal of Remote Sensing **43**(2): 393–423.
Jordan, C. F. (1969). "Derivation of leaf-area index from quality of light on the forest floor." Ecology **50**(4): 663–666.
Joshi, B. K., R. Shrestha, D. Gauchan and A. Shrestha (2020). "Neglected, underutilized, and future smart crop species in Nepal." Journal of Crop Improvement **34**(3): 291–313.
Judson, R., F. Elloumi, R. W. Setzer, Z. Li and I. Shah (2008). "A comparison of machine learning algorithms for chemical toxicity classification using a simulated multi-scale data model." BMC bioinformatics **9**(1): 1–16.
Kancheva, R., D. Borisova and G. Georgiev (2014). "Chlorophyll assessment and stress detection from vegetation optical properties." Ecological engineering and environment protection **1**: 34–43.
Kang, J., X. Hao, H. Zhou and R. Ding (2021). "An integrated strategy for improving water use efficiency by understanding physiological mechanisms of crops responding to water deficit: Present and prospect." Agricultural Water Management **255**: 107008.

Kapari, M., S. Hlophe-Ginindza, L. Nhamo and S. Mpandeli (2023). "Contribution of smallholder farmers to food security and opportunities for resilient farming systems." Frontiers in Sustainable Food Systems **7**: 1149854.
Kapari, M., M. Sibanda, J. Magidi, T. Mabhaudhi, L. Nhamo and S. Mpandeli (2024). "Comparing Machine Learning Algorithms for Estimating the Maize Crop Water Stress Index (CWSI) Using UAV-Acquired Remotely Sensed Data in Smallholder Croplands." Drones **8**(61).
Kapari, M., M. Sibanda, J. Magidi, T. Mabhaudhi, L. Nhamo and S. Mpandeli (2024). "Comparing Machine Learning Algorithms for Estimating the Maize Crop Water Stress Index (CWSI) Using UAV-Acquired Remotely Sensed Data in Smallholder Croplands." Drones **8**(2).
Kapoor, B., S. Singh and P. Kumar (2022). "Taro (Colocasia esculenta): Zero wastage orphan food crop for food and nutritional security." South African Journal of Botany **145**: 157–169.
Kastens, J. H., T. L. Kastens, D. L. Kastens, K. P. Price, E. A. Martinko and R.-Y. Lee (2005). "Image masking for crop yield forecasting using AVHRR NDVI time series imagery." Remote Sensing of Environment **99**(3): 341–356.
Kavhiza, N. J., M. Zargar, S. I. Prikhodko, E. N. Pakina, K. M.-S. Murtazova and M. R. Nakhaev (2022). "Improving crop productivity and ensuring food security through the adoption of genetically modified crops in sub-Saharan Africa." Agronomy **12**(2): 439.
KEGA, S. (2021). EVALUATING THE POTENTIAL FOR INCREASED FORAGE PRODUCTIVITY AND SOIL CARBON SEQUESTRATION IN STRIP-THINNED SILVOPASTURES, Department of Natural Resource Sciences, Thompson Rivers University.
Kemp, L., M. Roux, M. Kemp and R. Kock (2021). "Application of drones and image processing for bridge inspections in South Africa." Civil Engineering= Siviele Ingenieurswese **29**(8): 25–30.
Kganyago, M., C. Adjorlolo, M. Sibanda, P. Mhangara, G. Laneve and T. Alexandridis (2022). "Testing Sentinel-2 spectral configurations for estimating relevant crop biophysical and biochemical parameters for precision agriculture using tree-based and kernel-based algorithms." Geocarto International: 1–25.
Khanal, S., J. Fulton and S. Shearer (2017). "An overview of current and potential applications of thermal remote sensing in precision agriculture." Computers and Electronics in Agriculture **139**: 22–32.
Kima, D.-H., S. Son, J.-Y. Jung, J.-C. Lee and P.-G. Kim (2022). "Photosynthetic characteristics and chlorophyll content of Cypripedium japonicum in its natural habitat." Forest Science and Technology **18**(4): 160–171.
Krishna, G., R. N. Sahoo, P. Singh, H. Patra, V. Bajpai, B. Das, S. Kumar, R. Dhandapani, C. Vishwakarma and M. Pal (2021). "Application of thermal imaging and hyperspectral remote sensing for crop water deficit stress monitoring." Geocarto International **36**(5): 481–498.
Kumar, L., N. Chhogyel, T. Gopalakrishnan, M. K. Hasan, S. L. Jayasinghe, C. S. Kariyawasam, B. K. Kogo and S. Ratnayake (2022). Climate change and future of agri-food production. Future Foods, Elsevier**:** 49–79.
Kumar, N., S. Sow, L. Rana, V. Kumar, J. Kumar, B. Pramanick, A. Singh, L. A. Alkeridis, S. Sayed and A. Gaber (2024). "Productivity, water use efficiency and soil properties of sugarcane as influenced by trash mulching and irrigation regimes under different planting systems in sandy loam soils." Frontiers in Sustainable Food Systems **8**: 1340551.
Kunene, T., S. Hlophe-Ginindza, V. G. Chimonyo, A. T. Modi, S. Mpandeli, L. Nhamo and T. Mabhaudhi (2022). Contribution of Underutilised Indigenous Crops to Enhanced Food and Nutrition Security in the Advent of Climate Change. Food Security for African Smallholder Farmers, Springer**:** 295–310.
Kunz, R., K. Reddy, T. Mthembu, S. Lake, T. Mabhaudhi, V. Chimonyo and V. Naiken (2024). Crop and nutritional water productivity of sweet potato and taro. Pretoria, Water Research Commission.
Lao, Z., B. Fu, Y. Wei, T. Deng, W. He, Y. Yang, H. He and E. Gao (2024). "Retrieval of chlorophyll content for vegetation communities under different inundation frequencies using UAV images and field measurements." Ecological Indicators **158**: 111329.

Le Saint, T., J. Nabucet, L. Hubert-Moy and K. Adeline (2024). "Estimation of Urban Tree Chlorophyll Content and Leaf Area Index Using Sentinel-2 Images and 3D Radiative Transfer Model Inversion." Remote Sensing **16**(20).
Lencha, B., A. Birksew and G. Dikale (2016). "The Evaluation of growth performance of sweet potato (Ipomoea Batatas L.) Awassa var. by using different type of vine cuttings." Food Science and Quality Management **54**(4): 887–891.
Li, B., X. Xu, L. Zhang, J. Han, C. Bian, G. Li, J. Liu and L. Jin (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**: 161–172.
Li, C., C. Ma, H. Pei, H. Feng, J. Shi, Y. Wang, W. Chen, Y. Li, X. Feng and Y. Shi (2020). "Estimation of Potato Biomass and Yield Based on Machine Learning from Hyperspectral Remote Sensing Data." J. Agric. Sci. Technol. B **10**: 195–213.
Li, D., Y. Miao, S. K. Gupta, C. J. Rosen, F. Yuan, C. Wang, L. Wang and Y. Huang (2021). "Improving potato yield prediction by combining cultivar information and UAV remote sensing data using machine learning." Remote Sensing **13**(16): 3322.
Li, D., C. Quan, Z. Song, X. Li, G. Yu, C. Li and A. Muhammad (2020). "High-Throughput Plant Phenotyping Platform (HT3P) as a Novel Tool for Estimating Agronomic Traits From the Lab to the Field." Front Bioeng Biotechnol **8**: 623705.
Li, J., H. Jiang, W. Luo, X. Ma and Y. Zhang (2023). "Potato LAI estimation by fusing UAV multi-spectral and texture features." Journal of South China Agricultural University **44**(1): 93–101.
Li, J., N. K. Wijewardane, Y. Ge and Y. Shi (2023). "Improved chlorophyll and water content estimations at leaf level with a hybrid radiative transfer and machine learning model." Computers and Electronics in Agriculture **206**: 107669.
Li, W., C. Liu, Y. Yang, M. Awais, W. Li, P. Ying, W. Ru and M. Cheema (2022). "A UAV-aided prediction system of soil moisture content relying on thermal infrared remote sensing." International Journal of Environmental Science and Technology **19**(10): 9587–9600.
Li, X. and K. H. Siddique (2018). "Future smart food." Rediscovering hidden treasures of neglected and underutilized species for Zero Hunger in Asia, Bangkok.
Li, Y., N. He, J. Hou, L. Xu, C. Liu, J. Zhang, Q. Wang, X. Zhang and X. Wu (2018). "Factors influencing leaf chlorophyll content in natural forests at the biome scale." Frontiers in Ecology and Evolution **6**: 64.
Li, Y., B. Ren, L. Ding, Q. Shen, S. Peng and S. Guo (2013). "Does chloroplast size influence photosynthetic nitrogen use efficiency?" PloS one **8**(4): e62036.
Liakos, K. G., P. Busato, D. Moshou, S. Pearson and D. Bochtis (2018). "Machine Learning in Agriculture: A Review." Sensors (Basel) **18**(8).
Ling, Q., W. Huang and P. Jarvis (2011). "Use of a SPAD-502 meter to measure leaf chlorophyll concentration in Arabidopsis thaliana." Photosynthesis research **107**(2): 209–214.
Liu, C., Y. Liu, Y. Lu, Y. Liao, J. Nie, X. Yuan and F. Chen (2019). "Use of a leaf chlorophyll content index to improve the prediction of above-ground biomass and productivity." PeerJ **6**: e6240.
Liu, N., P. A. Townsend, M. R. Naber, P. C. Bethke, W. B. Hills and Y. Wang (2021). "Hyperspectral imagery to monitor crop nutrient status within and across growing seasons." Remote Sensing of Environment **255**.
Liu, Y., H. Feng, J. Yue, Y. Fan, X. Jin, X. Song, H. Yang and G. Yang (2022). "Estimation of Potato Above-Ground Biomass Based on Vegetation Indices and Green-Edge Parameters Obtained from UAVs." Remote Sensing **14**(21).
Liu, Y., H. Feng, J. Yue, Y. Fan, X. Jin, Y. Zhao, X. Song, H. Long and G. Yang (2022). "Estimation of Potato Above-Ground Biomass Using UAV-Based Hyperspectral images and Machine-Learning Regression." Remote Sensing **14**(21).
Liu, Y., J. Huang, Q. Sun, H. Feng, G. Yang and F. Yang (2021). "Estimation of plant height and above ground biomass of potato based on UAV digital image." National Remote Sensing Bulletin **25**(9): 2004–2014.

Liu, Y., J. Qian and H. Yue (2021). "Comprehensive evaluation of Sentinel-2 red edge and shortwave-infrared bands to estimate soil moisture." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **14**: 7448–7465.
Liu, Y., Q. Sun, J. Huang, H. K. Feng, J. J. Wang and G. J. Yang (2021). "Estimation of Potato Above Ground Biomass Based on UAV Multispectral Images." Guang Pu Xue Yu Guang Pu Fen Xi/Spectroscopy and Spectral Analysis **41**(8): 2549–2555.
Liu, Y., H. Zhang, H. K. Feng, Q. Sun, J. Huang, J. J. Wang and G. J. Yang (2021). "Estimation of Potato Above Ground Biomass Based on Hyperspectral Images of UAV." Guang Pu Xue Yu Guang Pu Fen Xi/Spectroscopy and Spectral Analysis **41**(9): 2657–2664.
Loconsole, D., M. Elia, G. Conversa, B. De Lucia, G. Cristiano and A. Elia (2025). "Soil Moisture Sensing Technologies: Principles, Applications, and Challenges in Agriculture." Agronomy **15**(12): 2788.
Lu, B. and Y. He (2017). "Species classification using Unmanned Aerial Vehicle (UAV)-acquired high spatial resolution imagery in a heterogeneous grassland." ISPRS Journal of Photogrammetry and Remote Sensing **128**: 73–85.
Lu, Y., S. Young, H. Wang and N. Wijewardane (2022). "Robust plant segmentation of color images based on image contrast optimization." Computers and Electronics in Agriculture **193**: 106711.
Luo, S., Y. He, Q. Li, W. Jiao, Y. Zhu and X. Zhao (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**: 1–14.
Luo, S., Y. He, Q. Li, W. Jiao, Y. Zhu and X. Zhao (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).
Ma, C., X. Li and M. F. McCabe (2020). "Retrieval of high-resolution soil moisture through combination of Sentinel-1 and Sentinel-2 data." Remote Sensing **12**(14): 2303.
Ma, Y., Z. Zhang, Y. Kang and M. Özdoğan (2021). "Corn yield prediction and uncertainty analysis based on remotely sensed variables using a Bayesian neural network approach." Remote Sensing of Environment **259**: 112408.
Mabhaudhi, T., V. Chimonyo and A. Modi (2017). "Status of Underutilised Crops in South Africa: Opportunities for Developing Research Capacity." Sustainability **9**(9).
Mabhaudhi, T., V. G. Chimonyo, T. P. Chibarabada and A. T. Modi (2017). "Developing a roadmap for improving neglected and underutilized crops: A case study of South Africa." Frontiers in plant science **8**: 2143.
Mabhaudhi, T., V. G. Chimonyo and A. T. Modi (2017). "Status of underutilised crops in South Africa: opportunities for developing research capacity." Sustainability **9**(9): 1569.
Mabhaudhi, T., V. G. P. Chimonyo, T. P. Chibarabada and A. T. Modi (2017). "Developing a roadmap for improving neglected and underutilized crops: A case study of South Africa." Frontiers in plant science **8**: 2143.
Mabhaudhi, T. and A. Modi (2015). "Drought tolerance of selected South African taro (Colocasia esculenta L. Schott) landraces." Experimental Agriculture **51**(3): 451–466.
Mabhaudhi, T., A. Modi and Y. Beletse (2011). Growth response of selected taro [Colocasia esculenta (L.) schott] landraces to water stress. II International Symposium on Underutilized Plant Species: Crops for the Future-Beyond Food Security 979.
Mabhaudhi, T., A. T. Modi and Y. G. Beletse (2014). "Parameterisation and evaluation of the FAO-AquaCrop model for a South African taro (Colocasia esculenta L. Schott) landrace." Agricultural and Forest Meteorology **192**: 132–139.
Maes, W. and K. Steppe (2012). "Estimating evapotranspiration and drought stress with ground-based thermal remote sensing in agriculture: a review." Journal of experimental botany **63**(13): 4671–4712.
Maes, W. H. and K. Steppe (2012). "Estimating evapotranspiration and drought stress with ground-based thermal remote sensing in agriculture: a review." JOURNAL OF EXPERIMENTAL BOTANY **63**(13): 4671–4712.

Maes, W. H. and K. Steppe (2019). "Perspectives for Remote Sensing with Unmanned Aerial Vehicles in Precision Agriculture." Trends in Plant Science **24**(2): 152–164.
Magney, T. S., L. A. Vierling, J. U. Eitel, D. R. Huggins and S. R. Garrity (2016). "Response of high frequency Photochemical Reflectance Index (PRI) measurements to environmental conditions in wheat." Remote Sensing of Environment **173**: 84–97.
Maimaitijiang, M., V. Sagan, P. Sidike, A. M. Daloye, H. Erkbol and F. B. Fritschi (2020). "Crop monitoring using satellite/UAV data fusion and machine learning." Remote Sensing **12**(9): 1357.
Majid, K. and G. Roza (2012). "The effect of drought stress on leaf chlorophyll content and stress resistance in maize cultivars (Zea mays)." African Journal of Microbiology Research **6**(12): 2844–2848.
Malbéteau, Y., S. Parkes, B. Aragon, J. Rosas and M. F. McCabe (2018). "Capturing the diurnal cycle of land surface temperature using an unmanned aerial vehicle."
Manafifard, M. (2024). "A new hyperparameter to random forest: application of remote sensing in yield prediction." Earth Science Informatics **17**(1): 63–73.
Markwell, J., J. C. Osterman and J. L. Mitchell (1995). "Calibration of the Minolta SPAD-502 leaf chlorophyll meter." Photosynthesis research **46**(3): 467–472.
Marques, P., L. Padua, T. Brito, J. J. Sousa and A. Fernandes-Silva (2020). Monitoring of Olive Trees Temperatures under Different Irrigation Strategies by UAV Thermal Infrared Imagery.
Marshall, M. and P. Thenkabail (2015). "Advantage of hyperspectral EO-1 Hyperion over multispectral IKONOS, GeoEye-1, WorldView-2, Landsat ETM+, and MODIS vegetation indices in crop biomass estimation." ISPRS Journal of Photogrammetry and Remote Sensing **108**: 205–218.
Masemola, R., M. Sibanda, O. Mutanga, R. Kunz, V. G. Chimonyo and T. Mabhaudhi (2025). "Assessing the Potential of Drone Remotely Sensed Data in Detecting the Soil Moisture Content and Taro Leaf Chlorophyll Content Across Different Phenological Stages." Water **17**(19): 2796.
Masina, M., A. Lambertini, I. Daprà, E. Mandanici and A. Lamberti (2020). "Remote sensing analysis of surface temperature from heterogeneous data in a maize field and related water stress."
Masipa, T. (2017). "The impact of climate change on food security in South Africa: Current realities and challenges ahead." Jàmbá: Journal of Disaster Risk Studies **9**(1): 1–7.
Mawoyo, B., P. Adebola, A. S. Gerrano and E. Amonsou (2017). "Effect of genotypes and growth locations on composition and functional properties of amadumbe flours." Journal of food science and technology **54**(11): 3577–3586.
Mazarire, T. T., P. E. Ratshiedana, A. Nyamugama, E. Adam and G. Chirima (2020). "Exploring machine learning algorithms for mapping crop types in a heterogeneous agriculture landscape using Sentinel-2 data. A case study of Free State Province, South Africa." South African Journal of Geomatics **9**(2): 333–347.
Melash, A. A., B. Bytyqi, M. S. Nyandi, A. M. Vad and É. B. Ábrahám (2023). "Chlorophyll meter: A precision agricultural decision-making tool for nutrient supply in durum wheat (Triticum turgidum L.) cultivation under drought conditions." Life **13**(3): 824.
Messina, G. and G. Modica (2020). "Applications of UAV thermal imagery in precision agriculture: State of the art and future research outlook." Remote Sensing **12**(9): 1491.
Meyer, G. E., T. W. Hindman and K. Laksmi (1999). Machine vision detection parameters for plant species identification. Precision agriculture and biological quality, SPIE.
Meyer, G. E. and J. C. Neto (2008). "Verification of color vegetation indices for automated crop imaging applications." Computers and electronics in agriculture **63**(2): 282–293.
Meyer, G. E., J. C. Neto, D. D. Jones and T. W. Hindman (2004). "Intensified fuzzy clusters for classifying plant, soil, and residue regions of interest from color images." Computers and electronics in agriculture **42**(3): 161–180.
Mhango, J. K., E. W. Harris, R. Green and J. M. Monaghan (2021). "Mapping potato plant density variation using aerial imagery and deep learning techniques for precision agriculture." Remote Sensing **13**(14).

Mhango, J. K., W. E. Harris and J. M. Monaghan (2021). "Relationships between the spatio-temporal variation in reflectance data from the Sentinel-2 satellite and potato (Solanum tuberosum L.) yield and stem density." Remote Sensing **13**(21): 4371.
Miao, Y., D. J. Mulla, G. W. Randall, J. A. Vetsch and R. Vintila (2008). "Combining chlorophyll meter readings and high spatial resolution remote sensing images for in-season site-specific nitrogen management of corn." Precision Agriculture **10**(1): 45–62.
Mobasheri, M. R. and S. B. Fatemi (2013). "Leaf Equivalent Water Thickness assessment using reflectance at optimum wavelengths." Theoretical and Experimental Plant Physiology **25**: 196–202.
Monteoliva, M. I., M. C. Guzzo and G. A. Posada (2021). "Breeding for drought tolerance by monitoring chlorophyll content."
Mountrakis, G., J. Im and C. Ogole (2011). "Support vector machines in remote sensing: A review." ISPRS journal of photogrammetry and remote sensing **66**(3): 247–259.
Moussa Kourouma, J., E. Eze, E. Negash, D. Phiri, R. Vinya, A. Girma and A. Zenebe (2021). "Assessing the spatio-temporal variability of NDVI and VCI as indices of crops productivity in Ethiopia: a remote sensing approach." Geomatics, Natural Hazards and Risk **12**(1): 2880–2903.
Mugiyo, H., V. Chimonyo, M. Sibanda, R. Kunz, C. Masemola, A. Modi and T. Mabhaudhi (2021). "Evaluation of land suitability methods with reference to neglected and underutilised crop species: A scoping review." Land **10**(2): 125.
Mugiyo, H., V. G. Chimonyo, M. Sibanda, R. Kunz, L. Nhamo, C. R. Masemola, C. Dalin, A. T. Modi and T. Mabhaudhi (2021). "Multi-criteria suitability analysis for neglected and underutilised crop species in South Africa." PloS one **16**(1): e0244734.
Mugiyo, H., V. G. P. Chimonyo, M. Sibanda, R. Kunz, L. Nhamo, C. R. Masemola, C. Dalin, A. T. Modi and T. Mabhaudhi (2021). "Multi-criteria suitability analysis for neglected and underutilised crop species in South Africa." PLOS ONE **16**(1): e0244734.
Muhammad, A., M. Alam, I. Ahmad and A. Jalal (2021). "Role of beneficial microbes with nitrogen and phosphorous levels on canola productivity." Brazilian Journal of Biology **82**.
Muhammad, M., A. Waheed, A. Wahab, M. Majeed, M. Nazim, Y.-H. Liu, L. Li and W.-J. Li (2024). "Soil salinity and drought tolerance: An evaluation of plant growth, productivity, microbial diversity, and amelioration strategies." Plant Stress **11**: 100319.
Munialo, S., K. H. Siddique, N. P. Barker, C. M. Onyango, J. N. Amissah, L. N. Wamalwa, Q. Qwabe, A. J. Dougill and L. M. Sibanda (2024). "Reorienting research investments toward under-researched crops for sustainable food systems." Food and Energy Security **13**(2): e538.
Musa, M., F. Massawe, S. Mayes, I. Alshareef and A. Singh (2016). "Nitrogen fixation and N-balance studies on Bambara groundnut (Vigna subterranea L. Verdc) landraces grown on tropical acidic soils of Malaysia." Communications in Soil Science and Plant Analysis **47**(4): 533–542.
Mutanga, O., E. Adam and M. A. Cho (2012). "High density biomass estimation for wetland vegetation using WorldView-2 imagery and random forest regression algorithm." International Journal of Applied Earth Observation and Geoinformation **18**: 399–406.
Mutanga, O., E. Adam and M. A. Cho (2012). "High density biomass estimation for wetland vegetation using WorldView-2 imagery and random forest regression algorithm." International Journal of Applied Earth Observation and Geoinformation **18**(0): 399–406.
Mutanga, O., T. Dube and O. Galal (2017). "Remote sensing of crop health for food security in Africa: Potentials and constraints." Remote Sensing Applications: Society and Environment **8**: 231–239.
Mutanga, O., A. Masenyama and M. Sibanda (2023). "Spectral saturation in the remote sensing of high-density vegetation traits: A systematic review of progress, challenges, and prospects." ISPRS Journal of Photogrammetry and Remote Sensing **198**: 297–309.
Mutanga, O. and A. K. Skidmore (2004). "Narrow band vegetation indices overcome the saturation problem in biomass estimation." International Journal of Remote Sensing **25**(19): 3999–4014.
Nahiduzzaman, M., M. E. Chowdhury, A. Salam, E. Nahid, F. Ahmed, N. Al-Emadi, M. A. Ayari, A. Khandakar and J. Haider (2023). "Explainable deep learning model for automatic mulberry leaf disease classification." Frontiers in Plant Science **14**.

Narmilan, A., F. Gonzalez, A. S. A. Salgadoe, U. W. L. M. Kumarasiri, H. A. S. Weerasinghe and B. R. Kulasekara (2022). "Predicting canopy chlorophyll content in sugarcane crops using machine learning algorithms and spectral vegetation indices derived from UAV multispectral imagery." Remote Sensing **14**(5): 1140.
Narmilan, A., F. Gonzalez, A. S. A. Salgadoe, U. W. L. M. Kumarasiri, H. A. S. Weerasinghe and B. R. Kulasekara (2022). "Predicting Canopy Chlorophyll Content in Sugarcane Crops Using Machine Learning Algorithms and Spectral Vegetation Indices Derived from UAV Multispectral Imagery." Remote Sensing **14**(5).
Natekin, A. and A. Knoll (2013). "Gradient boosting machines, a tutorial." Frontiers in neurorobotics **7**: 21.
Ndlovu, H. S. (2021). "The utility of very-high resolution unmanned aerial vehicles (UAV) imagery in monitoring the spatial and temporal variations in leaf moisture content of smallholder maize farming systems." Doctoral dissertation.
Ndlovu, H. S., J. Odindi, M. Sibanda and O. Mutanga (2024). "A systematic review on the application of UAV-based thermal remote sensing for assessing and monitoring crop water status in crop farming systems." International Journal of Remote Sensing **45**(15): 4923–4960.
Ndlovu, H. S., J. Odindi, M. Sibanda and O. Mutanga (2025). "Multi-Temporal Analysis of Taro Crop Water Stress Using High-Resolution Thermal and Multispectral Proximal Sensing for Improved Resilience of Smallholder Farming Systems." Smart Agricultural Technology: 101337.
Ndlovu, H. S., J. Odindi, M. Sibanda, O. Mutanga, A. Clulow, V. G. Chimonyo and T. Mabhaudhi (2021). "A comparative estimation of maize leaf water content using machine learning techniques and unmanned aerial vehicle (UAV)-based proximal and remotely sensed data." Remote Sensing **13**(20): 4091.
Ndlovu, H. S., J. Odindi, M. Sibanda, O. Mutanga, A. Clulow, V. G. Chimonyo and T. Mabhaudhi (2024). "Use of unmanned aerial vehicle-derived multi-spectral data for the early detection of multi-temporal maize leaf equivalent water thickness and fuel moisture content for the improved resilience of smallholder maize farming." Journal of Applied Remote Sensing **18**(1): 014520–014520.
Ndlovu, H. S., J. Odindi, M. Sibanda, O. Mutanga, A. Clulow, V. G. Chimonyo and T. J. R. S. Mabhaudhi (2021). "A Comparative Estimation of Maize Leaf Water Content Using Machine Learning Techniques and Unmanned Aerial Vehicle (UAV)-Based Proximal and Remotely Sensed Data." **13**(20): 4091.
Ndlovu, H. S., J. Odindi, M. Sibanda, O. Mutanga, A. Clulow, V. G. P. Chimonyo and T. Mabhaudhi (2021). "A Comparative Estimation of Maize Leaf Water Content Using Machine Learning Techniques and Unmanned Aerial Vehicle (UAV)-Based Proximal and Remotely Sensed Data." Remote Sensing **13**(20): 4091.
Nduwamungu, C., N. Ziadi, L.-É. Parent, G. F. Tremblay and L. Thuriès (2009). "Opportunities for, and limitations of, near infrared reflectance spectroscopy applications in soil analysis: A review." Canadian Journal of soil science **89**(5): 531–541.
Neinavaz, E., A. K. Skidmore, R. Darvishzadeh and T. A. Groen (2017). "Retrieving vegetation canopy water content from hyperspectral thermal measurements." Agricultural and forest meteorology **247**: 365–375.
Ngcamu, B. S. and F. Chari (2020). "Drought influences on food insecurity in Africa: a systematic literature review." International Journal of Environmental Research and Public Health **17**(16): 5897.
Nguyen, H. P. D. and D. D. Nguyen (2021). "Drone application in smart cities: The general overview of security vulnerabilities and countermeasures for data communication." Development and Future of Internet of Drones (IoD): Insights, Trends and Road Ahead: 185–210.
Nhamo, L., J. Magidi, A. Nyamugama, A. D. Clulow, M. Sibanda, V. G. P. Chimonyo and T. Mabhaudhi (2020). "Prospects of Improving Agricultural and Water Productivity through Unmanned Aerial Vehicles." Agriculture **10**(7).
Nhamo, L., R. Van Dijk, J. Magidi, D. Wiberg and K. Tshikolomo (2018). "Improving the Accuracy of Remotely Sensed Irrigated Areas Using Post-Classification Enhancement Through UAV Capability." Remote Sensing **10**(5).

Niederheiser, R., M. Winkler, V. Di Cecco, B. Erschbamer, R. Fernández, C. Geitner, H. Hofbauer, C. Kalaitzidis, B. Klingraber and A. Lamprecht (2021). "Using automated vegetation cover estimation from close-range photogrammetric point clouds to compare vegetation location properties in mountain terrain." GIScience & remote sensing **58**(1): 120–137.
Nyasimi, M., D. Amwata, L. Hove, J. Kinyangi and G. Wamukoya (2014). "Evidence of impact: climate-smart agriculture in Africa." CCAFS Working Paper.
Nyawade, S. O., H. I. Gitari, N. N. Karanja, C. K. Gachene, E. Schulte-Geldermann and M. L. Parker (2021). "Yield and evapotranspiration characteristics of potato-legume intercropping simulated using a dual coefficient approach in a tropical highland." Field Crops Research **274**: 108327.
Nyoni, R. S., G. Bruelle, R. Chikowo and N. Andrieu (2024). "Targeting smallholder farmers for climate information services adoption in Africa: A systematic literature review." Climate Services **34**: 100450.
Odebiri, O., O. Mutanga, J. Odindi and R. Naicker (2022). "Modelling soil organic carbon stock distribution across different land-uses in South Africa: A remote sensing and deep learning approach." ISPRS Journal of Photogrammetry and Remote Sensing **188**: 351–362.
Odebiri, O., J. Odindi and O. Mutanga (2021). "Basic and deep learning models in remote sensing of soil organic carbon estimation: A brief review." International Journal of Applied Earth Observation and Geoinformation **102**: 102389.
Omosalewa, O., O. Mutanga, J. Odindi and R. Naicker (2022). "Modelling soil organic carbon stock distribution across different land-uses in South Africa: A remote sensing and deep learning approach." ISPRS Journal of Photogrammetry and Remote Sensing **188**: 351–362.
Omosalewa, O., J. Odindi and O. Mutanga (2021). "Basic and deep learning models in remote sensing of soil organic carbon estimation: A brief review." International Journal of Applied Earth Observation and Geoinformation **102**: 102389.
Otsu, N. (1979). "A threshold selection method from gray-level histograms." IEEE transactions on systems, man, and cybernetics **9**(1): 62–66.
Ouyang, W., P. C. Struik, X. Yin and J. Yang (2017). "Stomatal conductance, mesophyll conductance, and transpiration efficiency in relation to leaf anatomy in rice and wheat genotypes under drought." Journal of Experimental Botany **68**(18): 5191–5205.
Oyeyinka, S. A. and E. O. Amonsou (2020). "Composition, pasting and thermal properties of flour and starch derived from amadumbe with different corm sizes." Journal of Food Science and Technology **57**(10): 3688–3695.
Ozelkan, E. (2020). "Water body detection analysis using NDWI indices derived from landsat-8 OLI." Polish Journal of Environmental Studies **29**(2): 1759–1769.
Padulosi, S., J. Thompson and P. Rudebjer (2013). "Fighting poverty, hunger and malnutrition with neglected and underutilized species: needs, challenges and the way forward."
Padulosi, S., J. Thompson and P. G. Rudebjer (2013) "Fighting poverty, hunger and malnutrition with neglected and underutilized species: needs, challenges and the way forward."
Panda, S. S., D. P. Ames and S. Panigrahi (2010). "Application of Vegetation Indices for Agricultural Crop Yield Prediction Using Neural Network Techniques." Remote Sensing **2**(3): 673–696.
Parkash, V. and S. Singh (2020). "A review on potential plant-based water stress indicators for vegetable crops." Sustainability **12**(10): 3945.
Pasqualotto, N., J. Delegido, S. Van Wittenberghe, J. Verrelst, J. P. Rivera and J. Moreno (2018). "Retrieval of canopy water content of different crop types with two new hyperspectral indices: Water Absorption Area Index and Depth Water Index." International journal of applied earth observation and geoinformation **67**: 69–78.
Petibon, F. and G. L. B. Wiesenberg (2022). "Characterization of complex photosynthetic pigment profiles in European deciduous tree leaves by sequential extraction and reversed-phase high-performance liquid chromatography." Front Plant Sci **13**: 957606.
Pham, H.-T., H.-Q. Nguyen, K.-P. Le, T.-P. Tran and N.-T. Ha (2023). "Automated Mapping of Wetland Ecosystems: A Study Using Google Earth Engine and Machine Learning for Lotus Mapping in Central Vietnam." Water **15**(5): 854.

Pineda, M., M. Barón and M.-L. Pérez-Bueno (2020). "Thermal imaging for plant stress detection and phenotyping." Remote Sensing **13**(1): 68.
Pineda, M., M. Baron and M. L. Perez-Bueno (2021). "Thermal Imaging for Plant Stress Detection and Phenotyping." REMOTE SENSING **13**(1).
Pinty, B., T. Lavergne, J. L. Widlowski, N. Gobron and M. M. Verstraete (2009). "On the need to observe vegetation canopies in the near-infrared to estimate visible light absorption." Remote Sensing of Environment **113**(1): 10–23.
Popoola, J., O. Ojuederie, C. Omonhinmin and A. Adegbite (2019). Neglected and underutilized legume crops: Improvement and future prospects. Recent advances in grain crops research, IntechOpen.
Potithep, S., S. Nagai, K. N. Nasahara, H. Muraoka and R. Suzuki (2013). "Two separate periods of the LAI–VIs relationships using in situ measurements in a deciduous broadleaf forest." Agricultural and Forest Meteorology **169**: 148–155.
Prabawardani, S. (2007). Physiological and growth responses of selected sweet potato (Ipomoea batatas (L.) Lam.) cultivators to water stress. PhD, James Cook University.
Pretorius, A. and M. Geyser (2025). "The Impact of El Niño-Southern Oscillation Events on Price Volatility: The Case of South African Maize." Agriculture **15**(22): 2361.
Prikaziuk, E., G. Ntakos, T. ten Den, P. Reidsma, T. van der Wal and C. van der Tol (2022). "Using the SCOPE model for potato growth, productivity and yield monitoring under different levels of nitrogen fertilization." International Journal of Applied Earth Observation and Geoinformation **114**: 102997.
Psirofonia, P., V. Samaritakis, P. Eliopoulos and I. Potamitis (2017). "Use of Unmanned Aerial Vehicles for Agricultural Applications with Emphasis on Crop Protection: Three Novel Case - studies." International Journal of Agricultural Science and Technology **5**(1): 30–39.
Qader, S. H., J. Dash, V. A. Alegana, N. R. Khwarahm, A. J. Tatem and P. M. Atkinson (2021). "The role of earth observation in achieving sustainable agricultural production in arid and semi-arid regions of the world." Remote Sensing **13**(17): 3382.
Qi, D., T. Hu and T. Liu (2020). "Biomass accumulation and distribution, yield formation and water use efficiency responses of maize (Zea mays L.) to nitrogen supply methods under partial root-zone irrigation." Agricultural Water Management **230**: 105981.
Quan, X., B. He, X. Li and Z. Tang (2015). "Estimation of grassland live fuel moisture content from ratio of canopy water content and foliage dry biomass." IEEE Geoscience and Remote Sensing Letters **12**(9): 1903–1907.
Raeva, P. L., J. Šedina and A. Dlesk (2019). "Monitoring of crop fields using multispectral and thermal imagery from UAV." European Journal of Remote Sensing **52**(sup1): 192–201.
Raji, S. N., G. N. Aparna, C. N. Mohanan and N. Subhash (2017). "Proximal remote sensing of herbicide and drought stress in field grown colocasia and sweet potato plants by sunlight-induced chlorophyll fluorescence Imaging." Journal of the Indian Society of Remote Sensing **45**(3): 463–475.
Rajwade, Y. A., N. S. Chandel, K. Dubey, S. Anakkallan, K. Upender and D. Jat (2023). "Assessment of water stress in rainfed maize using RGB and thermal imagery." Arabian Journal of Geosciences **16**(2): 119.
Ramírez, D., W. Yactayo, R. Gutiérrez, V. Mares, F. De Mendiburu, A. Posadas and R. Quiroz (2014). "Chlorophyll concentration in leaves is an indicator of potato tuber yield in water-shortage conditions." Scientia Horticulturae **168**: 202–209.
Ramos, A. P. M., L. P. Osco, D. E. G. Furuya, W. N. Gonçalves, D. C. Santana, L. P. R. Teodoro, C. A. da Silva Junior, G. F. Capristo-Silva, J. Li and F. H. R. Baio (2020). "A random forest ranking approach to predict yield in maize with uav-based vegetation spectral indices." Computers and Electronics in Agriculture **178**: 105791.
Ramos, M. A. P., L. P. Osco, E. G. D. Furuya, N. W. Gonçalves, C. D. Santana, P. R. L. Theodoro, A. d. S. C. Junior, F. G. Capristo-Silva, J. Li, H. R. F. Baio, M. J. Juniour, E. P. Teodoro and H. Pistori (2020). "A random forest ranking approach to predict yield in maize with uav-based vegetation spectral indices." Computers and Electronics in Agriculture **178**.

Raper, T. B. and J. J. Varco (2014). "Canopy-scale wavelength and vegetative index sensitivities to cotton growth parameters and nitrogen status." Precision Agriculture **16**(1): 62–76.
Rasheed, M. W., Tang, J., Sarwar, A., Shah, S., Saddique, N., Khan, M. U., Imran Khan, M., Nawaz, S., Shamshiri, R. R., Aziz, M., & Sultan, M. (2022). "Soil Moisture Measuring Techniques and Factors Affecting the Moisture Dynamics: A Comprehensive Review." Sustainability **14**(18).
Ren, Z., H. Zheng, J. Chen, T. Chen, P. Xie, Y. Xu, J. Deng, H. Wang, M. Sun and W. Jiao (2024). "Integrating UAV, UGV and UAV-UGV collaboration in future industrialized agriculture: Analysis, opportunities and challenges." Computers and Electronics in Agriculture(227): 109631.
Ricciardi, V., N. Ramankutty, Z. Mehrabi, L. Jarvis and B. Chookolingo (2018). "How much of the world's food do smallholders produce?" Global food security **17**: 64–72.
Riehle, D., D. Reiser and H. W. Griepentrog (2020). "Robust index-based semantic plant/background segmentation for RGB-images." Computers and Electronics in Agriculture **169**: 105201.
Rouse, J., J. Haas, J. Schell and D. Deering (1974). Monitoring vegetation systemsin the Great Plains witherts. Proceedings of the 3rd ERTS Symposium, Washington, DC, USA.
Saarela, M. and S. Jauhiainen (2021). "Comparison of feature importance measures as explanations for classification models." SN Applied Sciences **3**: 1–12.
Sankaran, S., J. M. Maja, S. Buchanon and R. Ehsani (2013). "Huanglongbing (citrus greening) detection using visible, near infrared and thermal imaging techniques." Sensors (Basel) **13**(2): 2117–2130.
Sankaran, S., J. Zhou, L. R. Khot, J. J. Trapp, E. Mndolwa and P. N. Miklas (2018). "High-throughput field phenotyping in dry bean using small unmanned aerial vehicle based multispectral imagery." Computers and Electronics in Agriculture **151**: 84–92.
Sapkota, B., V. Singh, D. Cope, J. Valasek and M. Bagavathiannan (2020). "Mapping and Estimating Weeds in Cotton Using Unmanned Aerial Systems-Borne Imagery." AGRIENGINEERING **2**(2).
Sarker, I. H. (2021). "Machine learning: Algorithms, real-world applications and research directions." SN computer science **2**(3): 160.
Shafiq, S., N. A. Akram and M. Ashraf (2019). "Assessment of physio-biochemical indicators for drought tolerance in different cultivars of maize (Zea mays L.)." Pakistan Journal of Botany **51**(4): 1241–1247.
Sharifi, A. (2020). "Remotely sensed vegetation indices for crop nutrition mapping." Journal of the Science of Food and Agriculture **100**(14): 5191–5196.
Sheykhmousa, M., Mahdianpari, M., Ghanbari, H., Mohammadimanesh, F., Ghamisi, P. and Homayouni, S., (2020). "Support vector machine versus random forest for remote sensing image classification: A meta-analysis and systematic review." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **13**: 6308–6325.
Shu, M., J. Zuo, M. Shen, P. Yin, M. Wang, X. Yang, J. Tang, B. Li and Y. Ma (2021). "Improving the estimation accuracy of SPAD values for maize leaves by removing UAV hyperspectral image backgrounds." International Journal of Remote Sensing **42**(15): 5862–5881.
Sibanda, M., S. Buthelezi, O. Mutanga, J. Odindi, A. D. Clulow, V. Chimonyo, S. Gokool, V. Naiken, J. Magidi and T. Mabhaudhi (2023). "Exploring the prospects of UAV-Remotely sensed data in estimating productivity of Maize crops in typical smallholder farms of Southern Africa." ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences **10**: 1143–1150.
Sibanda, M., O. Mutanga, V. G. Chimonyo, A. D. Clulow, C. Shoko, D. Mazvimavi, T. Dube and T. Mabhaudhi (2021). "Application of drone technologies in surface water resources monitoring and assessment: a systematic review of progress, challenges, and opportunities in the global south." Drones **5**(3): 84.
Sibanda, M., O. Mutanga, T. Dube, T. S Vundla and P. L Mafongoya (2019). "Estimating LAI and mapping canopy storage capacity for hydrological applications in wattle infested ecosystems using Sentinel-2 MSI derived red edge bands." GIScience & remote sensing **56**(1): 68–86.

Sibanda, M., O. Mutanga and M. Rouget (2015). "Examining the potential of Sentinel-2 MSI spectral resolution in quantifying above ground biomass across different fertilizer treatments." ISPRS Journal of Photogrammetry and Remote Sensing **110**: 55–65.
Sibanda, M., O. Mutanga, M. Rouget and L. Kumar (2017). "Estimating biomass of native grass grown under complex management treatments using worldview-3 spectral derivatives." Remote Sensing **9**(1): 55.
Sibanda, M., H. S. Ndlovu, K. Brewer, S. Buthelezi, T. N. Matongera, O. Mutanga, J. Odidndi, A. D. Clulow, V. G. Chimonyo and T. Mabhaudhi (2023). "Remote sensing hail damage on maize crops in smallholder farms using data acquired by remotely piloted aircraft system." Smart Agricultural Technology **6**: 100325.
Sibanda, M., M. Onisimo, T. Dube and T. Mabhaudhi (2021). "Quantitative assessment of grassland foliar moisture parameters as an inference on rangeland condition in the mesic rangelands of southern Africa." International Journal of Remote Sensing(4): 1474–1491.
Sibanda, M., M. Onisimo, T. Dube and T. Mabhaudhi (2021). "Quantitative assessment of grassland foliar moisture parameters as an inference on rangeland condition in the mesic rangelands of southern Africa." International Journal of Remote Sensing **42**(4): 1474–1491.
Singh, C., S. K. Karan, P. Sardar and S. R. Samadder (2022). "Remote sensing-based biomass estimation of dry deciduous tropical forest using machine learning and ensemble analysis." Journal of Environmental Management **308**: 114639.
Singhal, G., B. Bansod, L. Mathew, J. Goswami, B. Choudhury and P. Raju (2019). "Estimation of leaf chlorophyll concentration in turmeric (Curcuma longa) using high-resolution unmanned aerial vehicle imagery based on kernel ridge regression." Journal of the Indian Society of Remote Sensing **47**(7): 1111–1122.
Singhal, G., B. Bansod, L. Mathew, J. Goswami, B. U. Choudhury and P. L. N. Raju (2019). "Chlorophyll estimation using multi-spectral unmanned aerial system based on machine learning techniques." Remote Sensing Applications: Society and Environment **15**(100235).
Singla, M., D. Ghosh and K. K. Shukla (2020). "Improved sparsity of support vector machine with robustness towards label noise based on rescaled α-hinge loss with non-smooth regularizer." Neural Processing Letters **52**(3): 2211–2239.
Sipper, M. and J. H. Moore (2021). "Conservation machine learning: a case study of random forests." Scientific Reports **11**(1): 3629.
Sobejano-Paz, V., T. N. Mikkelsen, A. Baum, X. Mo, S. Liu, C. J. Köppl, M. S. Johnson, L. Gulyas and M. García (2020). "Hyperspectral and thermal sensing of stomatal conductance, transpiration, and photosynthesis for soybean and maize under drought." Remote Sensing **12**(19): 3182.
Song, G. and Q. Wang (2022). "Developing Hyperspectral Indices for Assessing Seasonal Variations in the Ratio of Chlorophyll to Carotenoid in Deciduous Forests." Remote Sensing **14**(6): 1324.
Sun, C., L. Feng, Z. Zhang, Y. Ma, T. Crosby, M. Naber and Y. Wang (2020). "Prediction of end-of-season tuber yield and tuber set in potatoes using in-season uav-based hyperspectral imagery and machine learning." Sensors (Switzerland) **20**(18): 1–13.
Sun, Q., Q. Jiao, X. Qian, L. Liu, X. Liu and H. Dai (2021). "Improving the retrieval of crop canopy chlorophyll content using vegetation index combinations." Remote Sensing **13**(3): 470.
Sun, Y., J. Wang, Q. Wang and C. Wang (2023). "Responses of the growth characteristics of spinach to different moisture contents in soil under irrigation with magnetoelectric water." Agronomy **13**(3): 657.
Sun, Z., X. Wang, Z. Wang, L. Yang, Y. Xie and Y. Huang (2021). "UAVs as remote sensing platforms in plant ecology: review of applications and challenges." Journal of Plant Ecology **14**(6): 1003–1023.
Tahir, M. N., S. Z. A. Naqvi, Y. Lan, Y. Zhang, Y. Wang, M. Afzal, M. J. M. Cheema and S. Amir (2018). "Real time estimation of chlorophyll content based on vegetation indices derived from multispectral UAV in the kinnow orchard." International Journal of Precision Agricultural Aviation **1**(1).
Tahir, N., Z. A. N. Muhammad, S. Lan, Y. Zhang, Y. Wang, Y. Afzal, M. J. M. M. Cheema and M. S. Amir (2018). "Real time estimation of chlorophyll content based on vegetation indices derived from

multispectral UAV in the kinnow orchard." International Journal of Precision Agricultural Aviation **1**(1): 24–31.
Talucder, M. S. A., U. B. Ruba and M. A. S. Robi (2024). "Potentiality of Neglected and Underutilized Species (NUS) as a future resilient food: A systematic review." Journal of Agriculture and Food Research: 101116.
Tan, C.-W., P.-P. Zhang, X.-X. Zhou, Z.-X. Wang, Z.-Q. Xu, W. Mao, W.-X. Li, Z.-Y. Huo, W.-S. Guo and F. J. S. r. Yun (2020). "Quantitative monitoring of leaf area index in wheat of different plant types by integrating NDVI and Beer-Lambert law." **10**(1): 1–10.
Tanabe, D., S. Ichiura, A. Nakatsubo, T. Kobayashi and M. Katahira (2019). Yield prediction of potato by unmanned aerial vehicle. TAE 2019-Proceeding of 7th International Conference on Trends in Agricultural Engineering 2019, TAE Prague.
Tapia, F. G., M. J. Pavek and Z. Holden (2019). Modern soil moisture monitoring methods. Proceedings of the Washington, Oregon Potato Conf.
Tedesco, D., B. R. D. Almeida Moreira, M. R. Barbosa Júnior, J. P. Papa and R. P. D. Silva (2021). "Predicting on multi-target regression for the yield of sweet potato by the market class of its roots upon vegetation indices." Computers and Electronics in Agriculture **191**.
Tedesco, D., M. F. de Oliveira, A. F. dos Santos, E. H. C. Silva, G. de Souza Rolim and R. P. da Silva (2021). "Use of remote sensing to characterize the phenological development and to predict sweet potato yield in two growing seasons." European Journal of Agronomy **129**: 126337.
Thenkabail Prasad S, Smith Ronald B and P. E. De (2000). "Hyperspectral Vegetation Indices and Their Relationships with Agricultural Crop Characteristics." Remote sensing of Environment **71**(2): 158–182.
Thenkabail, P. S. and J. G. Lyon (2016). Hyperspectral remote sensing of vegetation, CRC Press.
Thenkabail, P. S., I. Mariotto, M. K. Gumma, E. M. Middleton, D. R. Landis and K. F. Huemmrich (2013). "Selection of hyperspectral narrowbands (HNBs) and composition of hyperspectral twoband vegetation indices (HVIs) for biophysical characterization and discrimination of crop types using field reflectance and Hyperion/EO-1 data." Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of **6**(2): 427–439.
Thenkabail, P. S., R. B. Smith and E. De Pauw (2002). "Evaluation of narrowband and broadband vegetation indices for determining optimal hyperspectral wavebands for agricultural crop characterization." Photogrammetric Engineering and Remote Sensing **68**(6): 607–622.
Touch, V., D. K. Tan, B. R. Cook, D. Li Liu, R. Cross, T. A. Tran, A. Utomo, S. Yous, C. Grunbuhel and A. Cowie (2024). "Smallholder farmers' challenges and opportunities: Implications for agricultural production, environment and food security." Journal of Environmental Management **370**: 122536.
Traore, A., S. T. Ata-Ul-Karim, A. Duan, M. K. Soothar, S. Traore and B. Zhao (2021). "Predicting Equivalent Water Thickness in Wheat Using UAV Mounted Multispectral Sensor through Deep Learning Techniques." Remote Sensing **13**(21): 4476.
Tumuhimbise, R. (2015). "Plant spacing and planting depth effects on corm yield of taro (Colocasia esculenta (L.) Schott)." Journal of Crop Improvement **29**(6): 747–757.
Uddling, J., J. Gelang-Alfredsson, K. Piikki and H. Pleijel (2007). "Evaluating the relationship between leaf chlorophyll concentration and SPAD-502 chlorophyll meter readings." Photosynthesis research **91**(1): 37–46.
Upendar, K., K. Agrawal, N. Chandel and K. Singh (2021). "Greenness identification using visible spectral colour indices for site specific weed management." Plant Physiology Reports **26**: 179–187.
Vågsholm, I., N. S. Arzoomand and S. Boqvist (2020). "Food security, safety, and sustainability—getting the trade-offs right." Frontiers in Sustainable Food Systems: 16.
Van Wart, J., K. C. Kersebaum, S. Peng, M. Milner and K. G. Cassman (2013). "Estimating crop yield potential at regional to national scales." Field Crops Research **143**: 34–43.
Van Wyk, H. R. A., O E (2021). "Physiochemical and functional properties of albumin and globulin from amadumbe (Colocasia esculenta) corms." Food Science and Technology.

Vennam, R. R., P. Ramamoorthy, S. Poudel, K. R. Reddy, W. B. Henry and R. Bheemanahalli (2023). "Developing functional relationships between soil moisture content and corn early-season physiology, growth, and development." Plants **12**(13): 2471.
Vitor, A. B., R. P. Diniz, C. V. Morgante, R. P. Antônio and E. J. de Oliveira (2019). "Early prediction models for cassava root yield in different water regimes." Field Crops Research **239**: 149–158.
Walker, B. J., D. T. Drewry, R. A. Slattery, A. VanLoocke, Y. B. Cho and D. R. Ort (2018). "Chlorophyll Can Be Reduced in Crop Canopies with Little Penalty to Photosynthesis." Plant Physiol **176**(2): 1215–1232.
Wang, D. and C. J. Rosen (2003). Determining growth and yield limiting factors in potato from canopy spectral reflectance. Proceedings of SPIE - The International Society for Optical Engineering.
Wang, J., X. Zhang, Z. Han, H. Feng, Y. Wang, J. Kang, X. Han, L. Wang, C. Wang and H. Li (2022). "Analysis of physiological indicators associated with drought tolerance in wheat under drought and re-watering conditions." Antioxidants **11**(11): 2266.
Wang, Q., Y. Che, H. Chai, K. Shao, C. Yu, B. Li and Y. Ma (2021). "Monitoring of sugar beet growth using canopy spectrum and structural characteristics with UAV images." Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering **37**(20): 90–98.
Wang, R., F. Shi and D. Xu (2022). "The Extraction Method of Alfalfa (Medicago sativa L.) Mapping Using Different Remote Sensing Data Sources Based on Vegetation Growth Properties." Land **11**(11): 1996.
Wijewardana, C., F. A. Alsajri, J. T. Irby, L. J. Krutz, B. Golden, W. B. Henry, W. Gao and K. R. Reddy (2019). "Physiological assessment of water deficit in soybean using midday leaf water potential and spectral features." Journal of Plant Interactions **14**(1): 533–543.
Woebbecke, D. M., G. E. Meyer, K. Von Bargen and D. A. Mortensen (1995). "Color indices for weed identification under various soil, residue, and lighting conditions." Transactions of the ASAE **38**(1): 259–269.
Wu, Q., Y. Zhang, Z. Zhao, M. Xie and D. Hou (2023). "Estimation of relative chlorophyll content in spring wheat based on multi-temporal UAV remote sensing." Agronomy **13**(1): 211.
Wu, S., L. Deng, L. J. Guo and Y. J. Wu (2022). "Wheat leaf area index prediction using data fusion based on high-resolution unmanned aerial vehicle imagery." Plant Methods **18**(1).
Xing, N., W. Huang, Q. Xie, Y. Shi, H. Ye, Y. Dong, M. Wu, G. Sun and Q. Jiao (2019). "A Transformed Triangular Vegetation Index for Estimating Winter Wheat Leaf Area Index." Remote Sensing **12**(1).
Xing, N., W. Huang, Q. Xie, Y. Shi, H. Ye, Y. Dong, M. Wu, G. Sun and Q. Jiao (2019). "A transformed triangular vegetation index for estimating winter wheat leaf area index." Remote Sensing **12**(1): 16.
Xu, X., L. Fan, Z. Li, Y. Meng, H. Feng, H. Yang and B. Xu (2021). "Estimating leaf nitrogen content in corn based on information fusion of multiple-sensor imagery from UAV." Remote Sensing **13**(3): 340.
Xue, J., M. C. Anderson, F. Gao, C. Hain, L. Sun, Y. Yang, K. R. Knipper, W. P. Kustas, A. Torres-Rua and M. Schull (2020). "Sharpening ECOSTRESS and VIIRS land surface temperature using harmonized Landsat-Sentinel surface reflectances."
Xue, J. and B. Su (2017). "Significant remote sensing vegetation indices: A review of developments and applications." Journal of Sensors **2017**.
Yadav, S. S. P., R. Reddy, D. Dharun, A. Niveditha and N. Hema (2021). "Crop Yield Prediction Using Image Processing." NVEO-NATURAL VOLATILES & ESSENTIAL OILS Journal| NVEO: 11981–11992.
Yang, C., J. H. Everitt, Q. Du, B. Luo and J. Chanussot (2012). "Using high-resolution airborne and satellite imagery to assess crop growth and yield variability for precision agriculture." Proceedings of the IEEE **101**(3): 582–592.
Yang, H., Y. Hu, Z. Zheng, Y. Qiao, K. Zhang, T. Guo and J. Chen (2022). "Estimation of Potato Chlorophyll Content from UAV Multispectral Images with Stacking Ensemble Algorithm." Agronomy **12**(10): 2318.
Yang, H., B. Ming, C. Nie, B. Xue, J. Xin, X. Lu, J. Xue, P. Hou, R. Xie, K. Wang and S. Li (2022). "Maize Canopy and Leaf Chlorophyll Content Assessment from Leaf Spectral Reflectance: Estimation and Uncertainty Analysis across Growth Stages and Vertical Distribution." Remote Sensing **14**(9).

Yang, J. C., T. S. Magney, D. Yan, J. F. Knowles, W. K. Smith, R. L. Scott and G. A. Barron-Gafford (2020). "The photochemical reflectance index (PRI) captures the ecohydrologic sensitivity of a semiarid mixed conifer forest." Journal of Geophysical Research: Biogeosciences **125**(11): e2019JG005624.
Yao, D., J. Yang and X. Zhan (2013). "An improved random forest algorithm for class-imbalanced data classification and its application in PAD risk factors analysis." The Open Electrical & Electronic Engineering Journal **7**(1).
Yilmaz, M. T., E. R. Hunt Jr and T. J. Jackson (2008). "Remote sensing of vegetation water content from equivalent water thickness using satellite imagery." Remote Sensing of Environment **112**(5): 2514–2522.
Yin, H., W. Huang, F. Li, H. Yang, Y. Li, Y. Hu and K. Yu (2023). "Multi-temporal UAV Imaging-Based Mapping of Chlorophyll Content in Potato Crop." PFG–Journal of Photogrammetry, Remote Sensing and Geoinformation Science **91**(2): 91–106.
Yin, Q., Y. Zhang, W. Li, J. Wang, W. Wang, I. Ahmad, G. Zhou and Z. Huo (2023). "Estimation of winter wheat SPAD values based on UAV multispectral remote sensing." Remote Sensing **15**(14): 3595.
Yu, M.-H., G.-D. Ding, G.-L. Gao, Y.-Y. Zhao, L. Yan and K. Sai (2015). "Using plant temperature to evaluate the response of stomatal conductance to soil moisture deficit." Forests **6**(10): 3748–3762.
Yu, T., J. Zhou, J. Fan, Y. Wang and Z. Zhang (2023). "Potato Leaf Area Index Estimation Using Multi-Sensor Unmanned Aerial Vehicle (UAV) Imagery and Machine Learning." Remote Sensing **15**(16): 4108.
Yuan, Q., H. Shen, T. Li, Z. Li, S. Li, Y. Jiang, H. Xu, W. Tan, Q. Yang and J. Wang (2020). "Deep learning in environmental remote sensing: Achievements and challenges." Remote Sensing of Environment **241**: 111716.
Yue, T., Y. Wang, L. Zhang, C. Gu, H. Xue, W. Wang, Q. Lyu and Y. Dun (2018). Deep learning for genomics: A concise overview.
Yulianto, F., P. D. Raharjo, I. B. Pramono, M. A. Setiawan, G. A. Chulafak, G. Nugroho, A. D. Sakti, S. Nugroho and S. Budhiman (2023). "Prediction and mapping of land degradation in the Batanghari watershed, Sumatra, Indonesia: utilizing multi-source geospatial data and machine learning modeling techniques." Modeling Earth Systems and Environment: 1–22.
Zaludin, Z. and A. S. M. Harituddin (2019). Challenges and Trends of Changing from Hover to Forward Flight for a Converted Hybrid Fixed Wing VTOL UAS from Automatic Flight Control System Perspective. 2019 IEEE 9th International Conference on System Engineering and Technology (ICSET), IEEE.
Zeng, L., G. Peng, R. Meng, J. Man, W. Li, B. Xu, Z. Lv and R. Sun (2021). "Wheat yield prediction based on unmanned aerial vehicles-collected red–green–blue imagery." Remote Sensing **13**(15): 2937.
Zeng, Y., D. Hao, A. Huete, B. Dechant, J. Berry, J. M. Chen, J. Joiner, C. Frankenberg, B. Bond-Lamberty and Y. Ryu (2022). "Optical vegetation indices for monitoring terrestrial ecosystems globally." Nature Reviews Earth & Environment **3**(7): 477–493.
Zhai, W., C. Li, Q. Cheng, F. Ding and Z. Chen (2023). "Exploring Multisource Feature Fusion and Stacking Ensemble Learning for Accurate Estimation of Maize Chlorophyll Content Using Unmanned Aerial Vehicle Remote Sensing." Remote Sensing **15**(13): 3454.
Zhang, F. and G. Zhou (2015). "Estimation of canopy water content by means of hyperspectral indices based on drought stress gradient experiments of maize in the north plain China." Remote Sensing **7**(11): 15203–15223.
Zhang, F. and G. Zhou (2019). "Estimation of vegetation water content using hyperspectral vegetation indices: A comparison of crop water indicators in response to water stress treatments for summer maize." BMC ecology **19**(1): 1–12.

Zhang, F. and G. Zhou (2019). "Estimation of vegetation water content using hyperspectral vegetation indices: A comparison of crop water indicators in response to water stress treatments for summer maize." BMC ecology **19**: 1–12.
Zhang, H., L. Wang, X. Jin, L. Bian and Y. Ge (2023). "High-throughput phenotyping of plant leaf morphological, physiological, and biochemical traits on multiple scales using optical sensing." The Crop Journal **11**(5): 1303–1318.
Zhang, H., L. Wang, T. Tian and J. Yin (2021). "A review of unmanned aerial vehicle low-altitude remote sensing (UAV-LARS) use in agricultural monitoring in China." Remote Sensing **13**(6): 1221.
Zhang, L., H. Zhang, Y. Niu and W. Han (2019). "Mapping Maize Water Stress Based on UAV Multispectral Remote Sensing." Remote Sensing **11**(6): 605.
Zhao, H., L. Di, Z. Sun, P. Hao, E. Yu, C. Zhang and L. Lin (2021). Impacts of soil moisture on crop health: A remote sensing perspective. 2021 9th International Conference on Agro-Geoinformatics (Agro-Geoinformatics), IEEE.
Zhao, Y.-P., J.-J. Wang, X.-Y. Li, G.-J. Peng and Z. Yang (2020). "Extended least squares support vector machine with applications to fault diagnosis of aircraft engine." ISA transactions **97**: 189–201.
Zheng, C., A. Abd-Elrahman and V. Whitaker (2021). "Remote sensing and machine learning in crop phenotyping and management, with an emphasis on applications in strawberry farming." Remote Sensing **13**(3): 531.
Zheng, Q., W. Huang, H. Ye, Y. Dong, Y. Shi and S. Chen (2020). "Using continous wavelet analysis for monitoring wheat yellow rust in different infestation stages based on unmanned aerial vehicle hyperspectral images." Applied Optics **59**(26): 8003–8013.
Zhou, Z., Y. Majeed, G. D. Naranjo and E. M. Gambacorta (2021). "Assessment for crop water stress with infrared thermal imagery in precision agriculture: A review and future prospects for deep learning applications." Computers and Electronics in Agriculture **182**: 106019.
Zhu, W., Z. Sun, Y. Huang, T. Yang, J. Li, K. Zhu, J. Zhang, B. Yang, C. Shao and J. Peng (2021). "Optimization of multi-source UAV RS agro-monitoring schemes designed for field-scale crop phenotyping." Precision Agriculture **22**(6): 1768–1802.
Zhu, W., Z. Sun, Y. Huang, T. Yang, J. Li, K. Zhu, J. Zhang, B. Yang, C. Shao, J. Peng, S. Li, H. Hu and X. Liao (2021). "Optimization of multi-source UAV RS agro-monitoring schemes designed for field-scale crop phenotyping."
Zulu, N. N. (2022). Water Scarcity and Household Food Security: A Case of Ulundi Local Municipality in KwaZulu-Natal, South Africa. Handbook of Research on Resource Management and the Struggle for Water Sustainability in Africa, IGI Global**:** 127–148.

APPENDICES

B4

	Taro	SP	NV	SU	BL	BU	M1	M2
Taro	0	0.963	0.699	0.783	0.836	0.849	0.694	0.770
SP	0.963	0	0.606	0.841	0.605	0.472	0.736	0.549
NV	0.699	0.606	0	0.698	0.574	0.531	0.401	0.470
SU	0.783	0.841	0.698	0	0.517	0.688	0.802	0.607
BL	0.836	0.605	0.574	0.517	0	0.361	0.749	0.318
BU	0.849	0.472	0.531	0.688	0.361	0	0.734	0.22
M1	0.69	0.736	0.401	0.802	0.745	0.734	0	0.676
M2	0.77	0.548	0.470	0.607	0.318	0.22	0.676	0

Chapter 3 Appendix

Table S2.2 Band 5 JM distances

B5

	Taro	SP	NV	SU	BL	BU	M1	M2
Taro	0	0.508	0.313	0.868	0.636	0.617	0.602	0.358
SP	0.508	0	0.365	1.034	0.687	0.653	0.813	0.538
NV	0.313	0.365	0	0.966	0.622	0.591	0.711	0.355
SU	0.868	1.034	0.966	1.490	1.050	1.053	0.806	0.859
BL	0.636	0.687	0.622	1.050	0	0.187	0.712	0.747
BU	0.617	0.653	0.591	1.053	0.187	0	0.717	0.736
M1	0.602	0.813	0.711	0.806	0.712	0.717	0	0.658
M2	0.358	0.538	0.355	0.859	0.747	0.736	0.658	0

Table S2.3 EXG index JM distances

EXG

	Taro	SP	NV	SU	BL	BU	M1	M2
Taro	0	0.686	0.334	0.523	0.706	0.705	0.613	0.254
SP	0.686	0	0.504	0.894	0.739	0.592	0.943	0.667
NV	0.334	0.504	1.490	0.681	0.629	0.553	0.755	0.332
SU	0.523	0.894	0.681	0	0.799	0.845	0.257	0.583
BL	0.706	0.739	0.629	0.799	1.490	0.425	0.836	0.734

BU	0.705	0.592	0.553	0.845	0.425	1.490	0.886	0.726	
M1	0.613	0.943	0.755	0.257	0.836	0.886	0	0.663	
M2	0.254	0.667	0.332	0.583	0.734	0.726	0.663		0

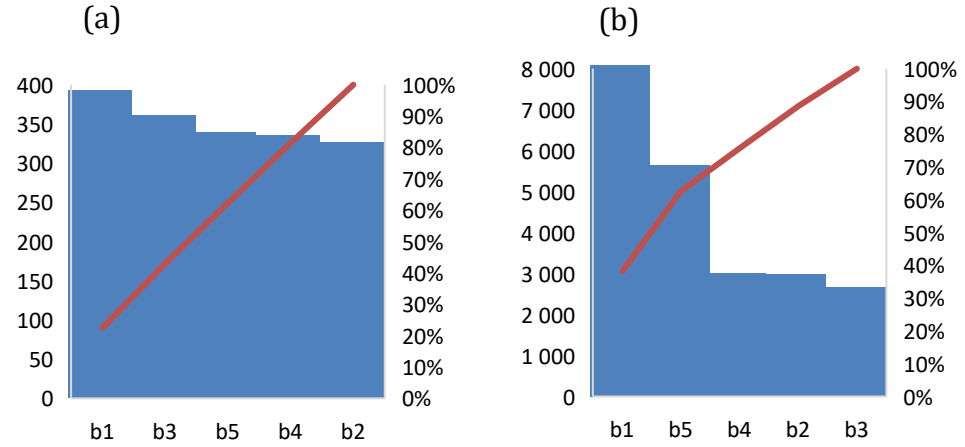


Figure S2.1 Variable importance scores of (a) RF and (b) GTB with Dataset 1

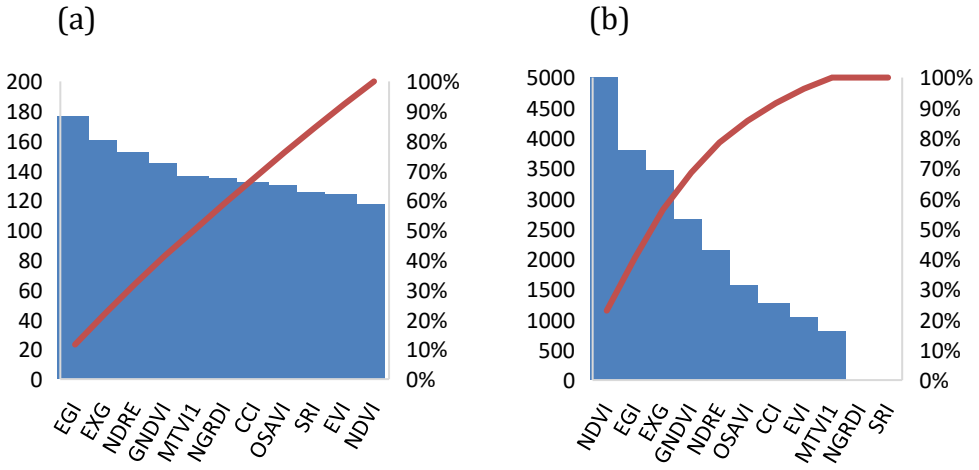


Figure S2.2 Variable importance scores of (a) RF and (b) GTB with Dataset 2

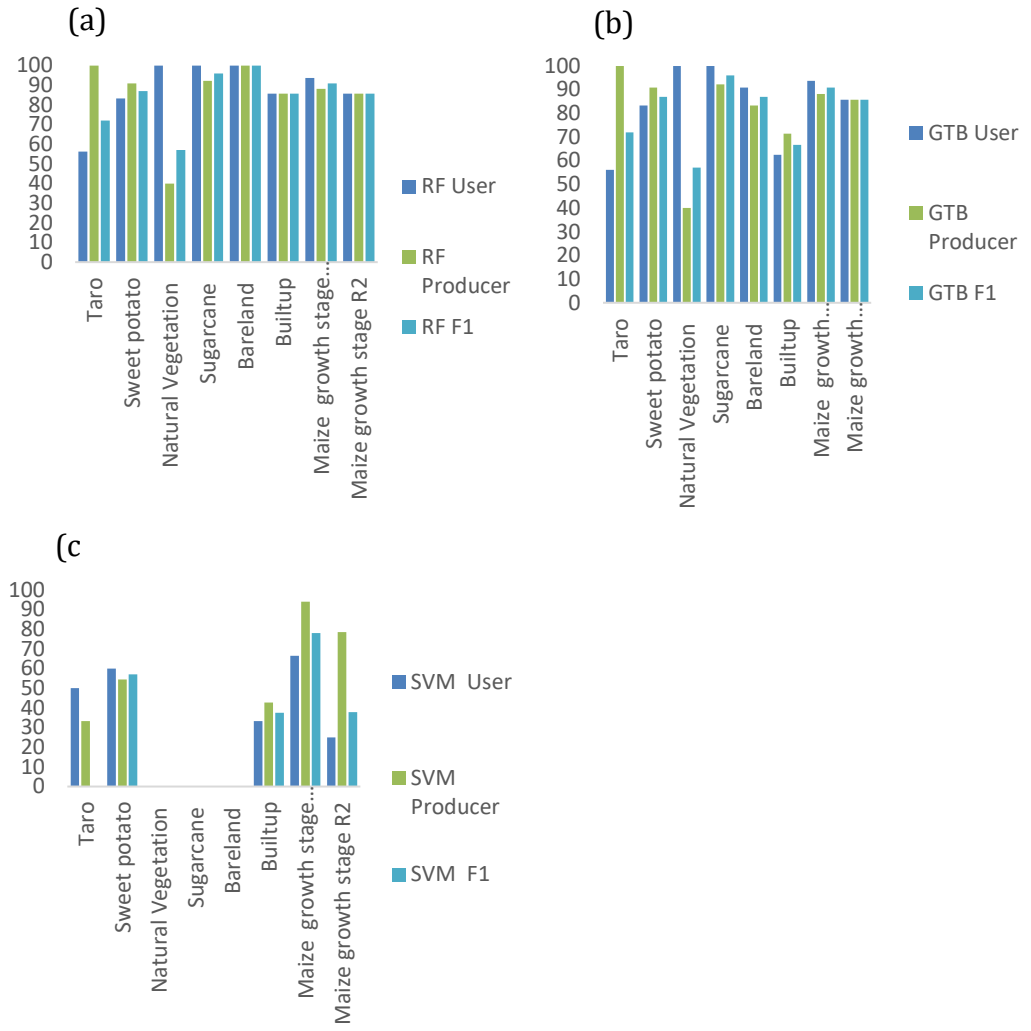
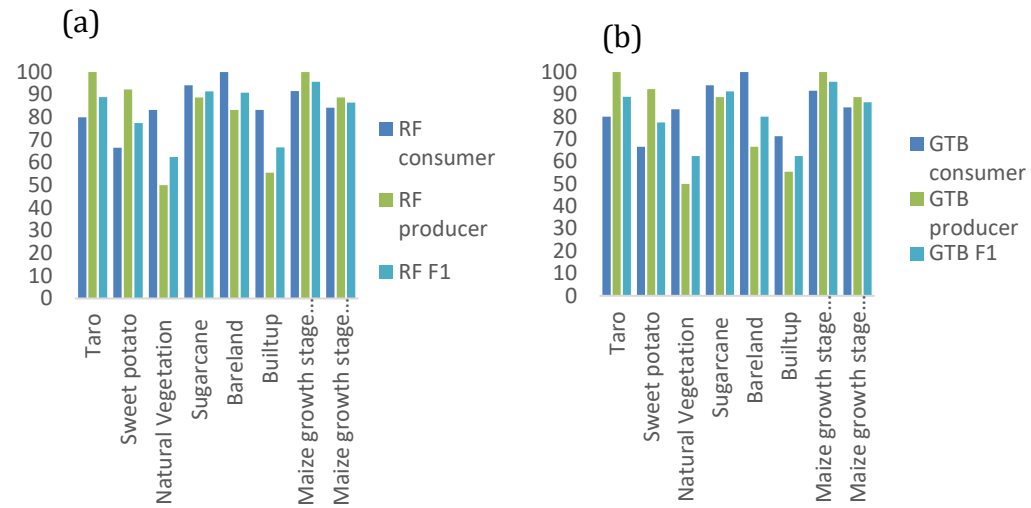


Figure S2.3 User and producer accuracies of (a) RF, (b) GTB and (c) SVM in conjunction with Dataset 1



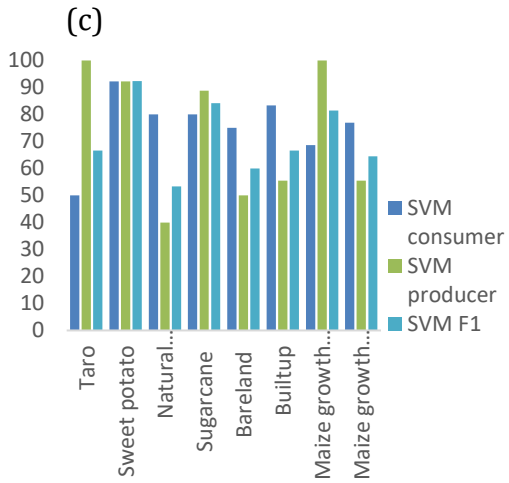


Figure S2.4 User and producer accuracies of (a) RF, (b) GTB and (c) SVM in conjunction with Dataset 2

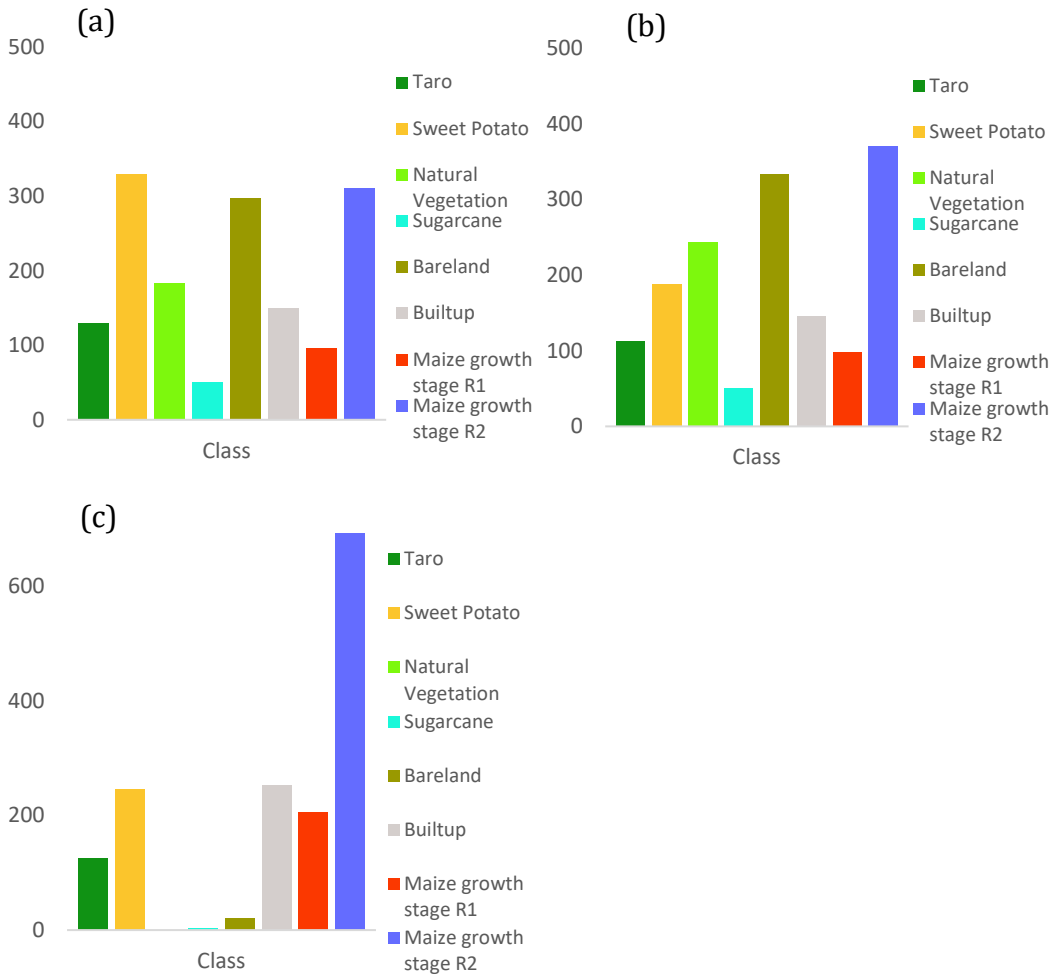


Figure S2.5 Areal extents per class of (a) RF, (b) GTB and (c) SVM with Dataset 1

(a) (b)

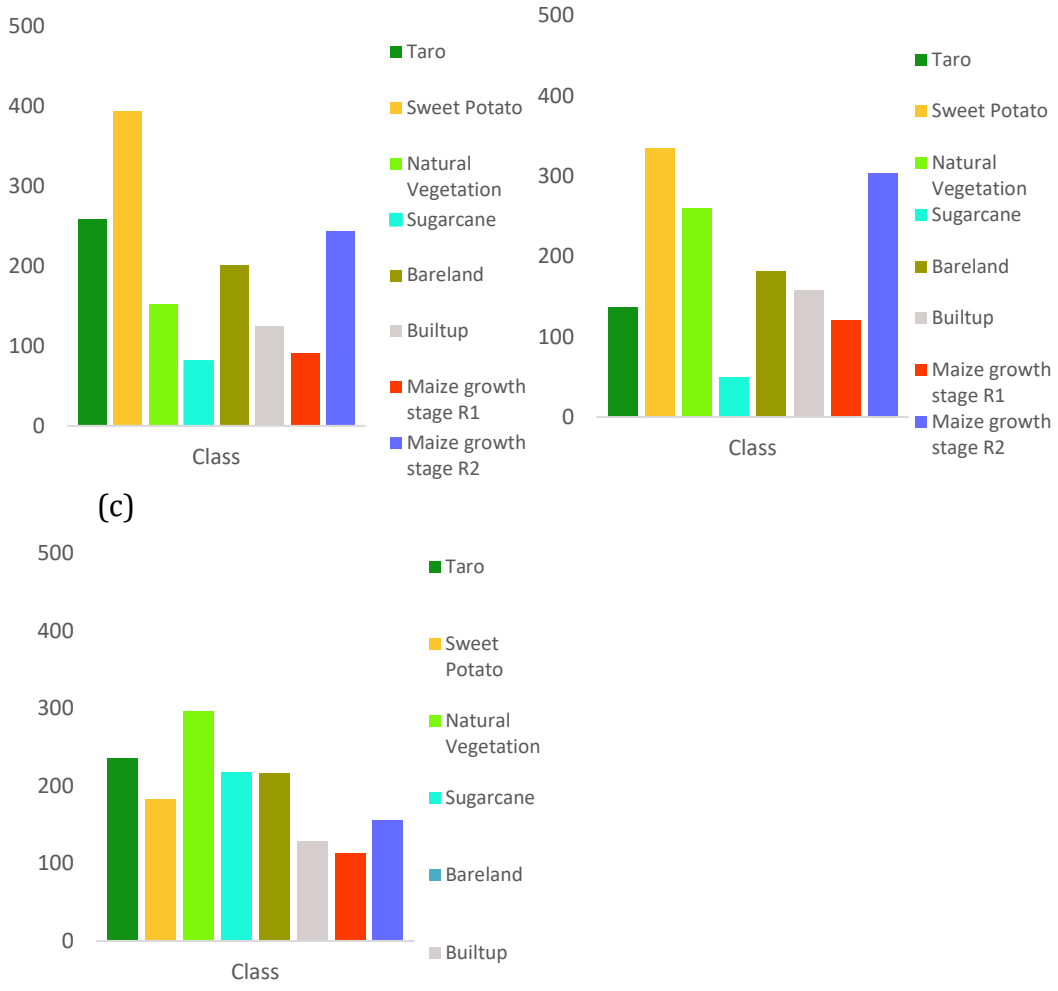
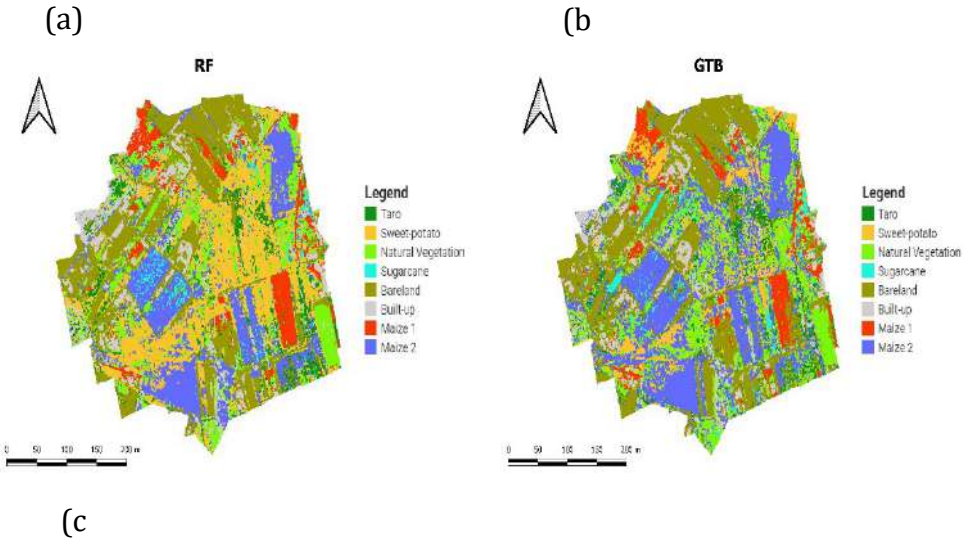


Figure S2.6 Areal extents per class of (a) RF, (b) GTB and (c) SVM with Dataset 2



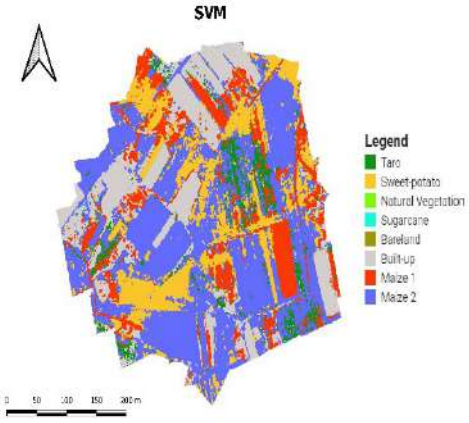


Figure S2.7 NUS crop distribution maps of (a) RF (b) GTB and (c) SVM with Dataset 1

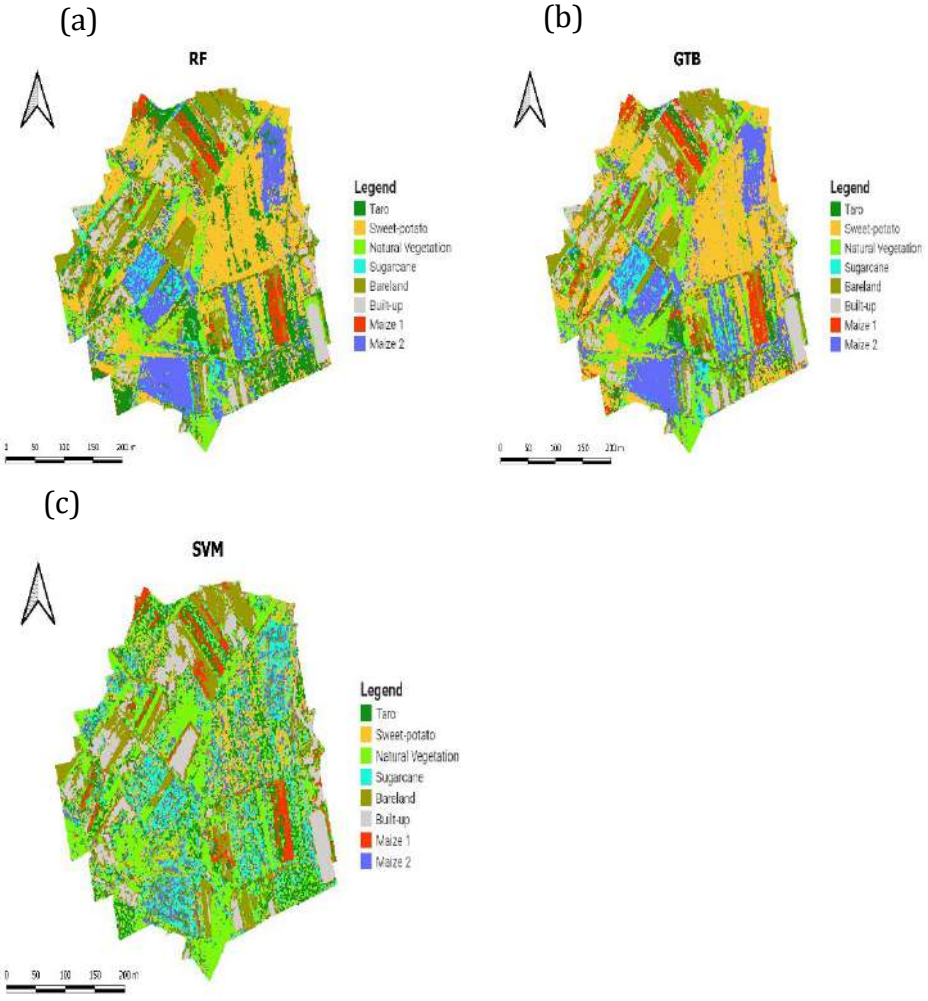


Figure S2.8 NUS crop distribution maps of (a) RF (b) GTB and (c) SVM with Dataset 2

INFORMATION DISSEMINATION: COMPLETED, SUBMITTED AND PUBLISHED MANUSCRIPTS

Research Articles for Publication

To date, the following manuscripts have been generated and some published as a vehicle for information dissemination:

- Ndlovu, H.S., Odindi, J., Sibanda, M. and Mutanga, O., 2026. Assessing neglected and underutilised taro crop water status using physiological indicators and UAV multi-modal thermal-multispectral data. Precision Agriculture, 27(1), p.15.
- Ndlovu, H.S., Odindi, J., Sibanda, M. and Mutanga, O., 2025. Enhancing the Estimation of Equivalent Water Thickness in Neglected and Underutilized Taro Crops using UAV acquired Multispectral Thermal Image data and Index-Based Image Segmentation. Remote Sensing Applications: Society and Environment, p.101758.
- Ndlovu HS, Odindi J, Sibanda M and Mutanga O 2024. A systematic review on the application of UAV-based thermal remote sensing for assessing and monitoring crop water status in crop farming systems. *International Journal of Remote Sensing* **45** (15) 4923-4960.https://www.tandfonline.com/doi/full/10.1080/01431161.2024.2368933
- Ndlovu, H.S., Odindi, J., Sibanda, M. and Mutanga, O., 2025. Multi-temporal analysis of taro crop water stress using high-resolution thermal and multispectral proximal sensing for improved resilience of smallholder farming systems. Smart Agricultural Technology, p.101337.
- Abrahams M, Sibanda M, Dube T, Chimonyo VGP and Mabhaudhi T 2023. A systematic review of UAV applications for mapping neglected and under-utilised crop species' spatial distribution and health. *Remote Sensing* **15** 4672. https://doi.org/10.3390/rs15194672

Conference attendance:

- Yola Kamteni (2025) *Assessing the spatial variability* of *Neglected and Underutilized Crop Species (NUS) leaf and canopy chlorophyll content in Kwazulu-Natal smallholder farms using unmanned aerial vehicle (UAV)-based high-throughput phenotyping* 26th WaterNet/WARFSA/GWPSA Symposium held online and at the Ciela Resort, Lusaka, Zambia, 29-30 October 2025
- Mishkah Abrahams: *Assessing the Potential of UAV Acquired Multispectral Imagery Combined with Machine Learning Techniques in Mapping the Spatial Distribution of Taro and Sweet Potatoes in Smaller Holder Farms.* Institute for Water Studies (IWS) Open Day, 21 May 2025, University of the Western Cape, South Africa
- **Ndlovu HS**, Odindi J, Sibanda M and Mutanga O (2024). *Enhancing Taro Equivalent Water Thickness Estimation through Index-Based Image Segmentation using Unmanned Aerial Vehicle Multispectral Thermal Imagery*. Proceedings of the 3rd

African Conference on Precision Agriculture, 229-235. African Plant Nutrition Institute, Benguérir, Morocco.

- **Ndlovu HS**, Odindi J, Sibanda M and Mutanga O (2024). Enhancing Taro Equivalent Water Thickness Estimation through Index-Based Image Segmentation using Unmanned Aerial Vehicle Multispectral Thermal Imagery. *RCMRD International Conference*. Regional Centre for Mapping of Resources for Development, Nairobi, Kenya.
- Sibanda M, Bacela E, Mutanga O, Odindi J, Chimonyo VGP, Magidi J, Clulow A, Kunz R and Mabhaudhi T. *Leveraging UAV Remotely Sensed Data to Estimate Crop Water Stress of Underutilized Taro Crops in Smallholder Croplands,* 3rd Biennial Africa Climate Smart Agriculture Stakeholders Conference, 29-31 July 2024, Kigali, Rwanda.
- Mishkah Abrahams 2024 *Assessing the Potential of UAV Acquired Multispectral Imagery Combined with Machine Learning Techniques in Mapping the Spatial Distribution of Taro and Sweet Potatoes in Smaller Holder Farms.* 3rd African Conference on Precision Agriculture (AfCPA) 3 December 2024 University of KwaZulu-Natal Pietermaritzburg, South Africa
- Helen S Ndlovu 2024 *Enhancing the Estimation of Equivalent Water Thickness in Neglected and Underutilized Taro Crops using UAV-acquired Multispectral Thermal Image data and Index-Based Image Segmentation* 3rd African Conference on Precision Agriculture (AfCPA) 3 December 2024 University of KwaZulu-Natal Pietermaritzburg, South Africa
- Siphiwokuhle Buthelezi 2024 *Assessing the Utility of Image Segmentation Techniques and Leaf Area Index in Improving Taro Yield Prediction Based on UAV-Acquired Multispectral Data in Smallholder Croplands,* 3rd African Conference on Precision Agriculture (AfCPA) 3 December 2024 University of KwaZulu-Natal Pietermaritzburg, South Africa
- Reitumetse Masemola 2024 *Assessing the potential of drone remotely sensed data in detecting the soil water content, and taro leaf chlorophyll content across different phenological stages,* 3 December 2024, University of KwaZulu-Natal, Pietermaritzburg, South Africa
- Sibanda M, Bacela E, Mutanga O, Odindi J, Chimonyo VGP, Magidi J, Clulow A, Kunz R and Mabhaudhi T 2024. *Leveraging UAV Remotely Sensed Data to Estimate Crop Water Stress of Underutilised Taro Crops in Smallholder Croplands, 3rd Biennial Africa Climate Smart Agriculture Stakeholders Conference, 29-31 July 2024, Kigali Rwanda.*

Published Popular articles

- Unlocking Precision Farming: How Drones Revolutionise the Mapping of Neglected and Under-utilised Crops like Taro and Sweet Potatoes in Smallholder Agriculture; Newsletter of the African Association for Precision Agriculture (AAPA) Newsletters, May 2025

https://paafrica.org/AAPA#:~:text=AAPA%20::%20AfCPA,Publishes%20a%20quarterly%20newsletter

Session Chair

- Mbulisi Sibanda Session Chair: Precision Agriculture: U.S. National Academy of Sciences U.S.-Africa Frontiers Symposium, February 18-20, 2025, Kigali, Rwanda

CAPACITY BUILDING REPORT

This project's four-year budget included provisions to fund three full-time master's students. The following summarises capacity-building progress to date.

Postgraduate student recruitment

The project engaged one postdoctoral fellow. A portion of the fellow's time was dedicated to the project as follows:

- Dr Trylee Matrongera – GIS and Remote Sensing – Agroecological and Land Use Modelling

The fellow was part of the project's capacity development initiative for Early Career Researchers (ECRs) at the University of KwaZulu-Natal, which aims to foster the next generation of researchers. Their defined role included organising field trips, acquiring UAV remote sensing data, co-supervising postgraduate students, and contributing to deliverable reports, all integral to their professional development.

In line with the project's capacity-building commitments, three master's candidates were recruited. Two graduated cum laude, while the third is submitting her thesis this year. Subsequently, a fourth MA student was recruited and supported for fieldwork. Recognising the strategic importance of increased postgraduate engagement, the project team included two female PhD students from the previous Flagship WRC Research Project (K5/2971//4). One of these students has graduated, while the other is expected to submit her thesis before the year's end. These doctoral candidates received partial funding from the National Research Foundation (NRF). Additionally, the first MA student, who completed her degree in 2023, was recruited as a PhD student. Overall, the postgrad students who were recruited under this project are:

- Yola Kamteni (MA 1) (Completed in 2024 *Cum laude*)
- Mishkah Abrahams (MA 2) (Completed 2024 *Cum laude*) (PhD 3) (Year 3)
- Reitumetse Masemola (MA 3) (Completed Honours in 2024) (Year 2) (Submitting)
- Dimpho Ndala (MA 4) (Year 2) (Submitting)
- Snethemba Ndlovu (PhD 1) (Completed)
- Siphiwokuhle Buthelezi (PhD 2) (Completing) (Year 4)

Professional Advancement of the Project Leader

Beyond student development, the Project Leader, a mid-career researcher, was also enhanced through this project. Their NRF rating improved from Y2 to C2, and they received a promotion at UWC from Senior Lecturer to Associate Professor following the graduation of the students from this project.