

---

**ASSESSING THE UTILITY OF UNMANNED AERIAL VEHICLE  
REMOTELY SENSED DATA FOR ESTIMATING MAIZE LEAF AREA  
INDEX (LAI) AND YIELD ACROSS THE GROWING SEASON**

**By**

**Siphiwokuhle Buthelezi**

**216002673**

**Supervisor: Professor O Mutanga**

**Co-supervisor: Dr M Sibanda**

Submitted in fulfilment of the academic requirements for the degree of Master of Science in  
the Discipline of Applied Environmental Science

School of Agricultural, Earth Environmental Sciences

University of KwaZulu-Natal

Pietermaritzburg

December 2021



**UNIVERSITY OF  
KWAZULU-NATAL** <sup>TM</sup>  
**INYUVESI  
YAKWAZULU-NATALI**

---

## Abstract

Monitoring maize leaf area index (LAI) and yield over smallholder farms is important for understanding crop productivity and for developing early warning systems to improve crop production. The advances in spatio-temporal remote sensing have made it feasible to monitor the development and productivity of crops LAI and predict their yields at a farm to field scales in smallholder croplands. It is, in this regard, that this study employed Unmanned Aerial vehicles (UAV) remotely sensed data in predicting maize LAI and yield of maize across the growing season in smallholder farms located in the KwaZulu-Natal province of South Africa. Specifically, five images were acquired using the DJI Matrice 300 and the Mica sense Altum across the maize growing season. Maize LAI samples were measured during the image acquisition times while the yield samples were harvested at the end of the growing season. The acquired multispectral images were used to generate 57 vegetation indices (VIs) that were used in this study to estimate maize LAI and yield based on the Random Forest regression (RF) ensemble to address two specific objectives. These specific objectives were i) to estimate LAI of maize crops using UAV derived VIs and RF regression across the growing season in smallholder croplands, and ii) to estimate maize yield across the phenological cycle based on UAV derived data in conjunction with RF regression in smallholder croplands. The results of this study showed that the optimal stage for estimating maize LAI using UAV derived VIs in concert with the RF ensemble was during the vegetative stage (V8- V10) with an RMSE of 0.15,  $R^2$  of 0.91 (RRMSE = 8 %) based on the blue, green and thermal variables. Across the growing season LAI was estimated to RMSE of 0.15  $m^2/m^2$ , 0.17  $m^2/m^2$ , 0.65  $m^2/m^2$ , 0.19  $m^2/m^2$  and 0.32  $m^2/m^2$ ,  $R^2$  of 0.91, 0.93, 0.91, 0.89 and 0.91; and RRMSE = 8.13%, 8.97%, 19.61%, 10.78% and 15, 22% for the V8-V10, V12-V14, VT-R1, R2-R3 and R3-R4 growth stages, respectively. Meanwhile, the combination of UAV derived VIs and bands facilitated an optimal estimation of yield to an  $R^2$  of 0.80-0.95, RMSE of 0.03-0.94  $kg/m^2$  and RRMSE of 2.21%-39.91% across the phenological cycle. The predictor variables derived from the blue, red, red edge and NIR sections of the electromagnetic spectrum (EMS) proved to be the most optimal variables for maize yield predictions. These results demonstrate the prospects of utilizing UAV derived data in predicting maize LAI and yield at field scale – a previously challenging task with freely available spatial resolution satellite sensors. This offers detailed spatially explicit information needed for optimizing agricultural production in smallholder farms especially in data-scarce regions such as sub-Saharan Africa.

**Keywords:** Maize, Yield, Leaf Area Index, Smallholder Farms, UAV, Random Forest

## **Dedication**

To my grandparents, my mother, my uncle, my family and friends.

“He who began a good work in you will carry it on to completion.”

- Philippians 1: 6

## **Acknowledgements**

First, I would like to thank God for his enduring love and faithfulness towards me.

To my supervisor Prof Onesimo Mutanga, thank you for your assistance, guidance and supervision throughout this study and most importantly thank you for funding my studies, I will forever be grateful.

To my co-supervisor Dr Mbulisi Sibanda, thank you for everything, from your assistance and guidance to your unwavering support, commitment and faith in me. Thank you for being my number one cheerleader, it is highly appreciated. I would also like to extend a huge appreciation to Prof John Odindi, thank you Prof for your immense contribution to the success of this project.

I would like to acknowledge Mr Vivek Naiken, Trylee Matongera, Serge Kiala, Kiara Brewer, Amanda Nyawose, Israel Odebiri, Lwando Royimani and Mr Welcome Ngcobo for their assistance with fieldwork and laboratory logistics.

Also, I would like to acknowledge the NRF Research Chair in Land Use Planning and Management for funding this study.

To my family and friends, thank you for your continued support and prayers. Your love and care kept me going. To Lethu, thank you for loving, supporting and for holding my hand throughout this journey. To Snethemba Ndlovu and Mthokozisi Buthelezi, thank you so much for your support and care during this project, you are appreciated.

Last but not least, to my grandpa, thank you for your undying love and support Mvelase!! Words fail me.

## Preface

The work described in this paper was completed at the University of KwaZulu-Natal, Pietermaritzburg, from March 2020 to December 2021, under the supervision of Professor O Mutanga and Dr M Sibanda.

The research represented in this document is original work by the author and has not otherwise been submitted in any form for any degree or diploma to any tertiary institution. Where use has been made of the work of others it is duly acknowledged in the text.

Signature (



Date: 02/12/2021

Siphwokuhle Buthelezi



Signature (Supervisor)

Date: 04/12/2021

Professor O Mutanga



Signature (Co-supervisor)

Date: 03/12/2021

Dr M Sibanda

## Declaration

I Siphwokuhle Buthelezi declare that the research presented in this dissertation, except where otherwise indicated is my original work. Wherever contributions of others were included, every effort was made to indicate this through in-text referencing and an additional reference list.

.....

Siphwokuhle Buthelezi

02/12/2021

.....

Date

As the candidate's supervisor, I certify the aforementioned statement and have approved this thesis for submission.

.....

Professor O Mutanga

04/12/2021

.....

Date

# Table of Contents

Abstract.....	i
Dedication.....	iii
Acknowledgements.....	iv
Preface.....	v
Declaration.....	vi
List of Figures.....	x
List of Tables.....	x
CHAPTER ONE: GENERAL INTRODUCTION-MAPPING LEAF AREA INDEX AND YIELD OF MAIZE USING UNMANNED AERIAL VEHICLE DERIVED DATA.....	1
1.1 Introduction.....	1
1.2 Aims and objectives.....	5
1.2.1 Specific objectives.....	5
1.3 Significance of the study.....	5
1.4 Structure of the thesis.....	6
1.5 References.....	7
CHAPTER TWO: ESTIMATING MAIZE LEAF AREA INDEX USING UAV-DERIVED MULTI-SPECTRAL REMOTELY SENSED DATA IN SMALLHOLDER FARMS.....	11
Abstract.....	11
2.1 Introduction.....	11
2.2 Materials and methods.....	15
2.2.1 Study site.....	15
2.2.2 Measurements of LAI.....	17
2.2.3 Image acquisition and pre-processing.....	18
2.2.4 Data analysis.....	19
2.3 Results.....	21
2.3.1 Descriptive statistics.....	21

2.3.2 Derived Maize LAI prediction models and their accuracies .....	21
2.4 Discussion .....	25
2.4.1 Predicting maize LAI.....	25
2.4.2 The performance of combining UAV derived traditional, red edge-based and new VIs in predicting maize LAI.....	27
2.5 Conclusion.....	27
2.6 References .....	29
 CHAPTER THREE: ESTIMATING MAIZE YIELD USING UAV-DERIVED MULTI- TEMPORAL DATA IN SMALLHOLDER FARMS OF KWAZULU-NATAL, SOUTH AFRICA.....	
Abstract .....	34
3.1 Introduction .....	35
3.2 Materials and methods .....	37
3.2.1 Study area .....	37
3.2.2 Agricultural practices .....	38
3.2.3 Sampling strategy for yield measurements.....	38
3.2.4 UAV system and imaging sensor .....	39
3.2.5 Image processing and field data collection.....	40
3.2.6 Data Analysis.....	41
3.3 Results .....	42
3.3.1 Descriptive statistics .....	42
3.3.2 Derived maize yield prediction models and their accuracies .....	43
3.4. Discussion .....	47
3.4.1 Maize yield prediction models .....	47
3.4.2 Determining the most optimal growth stages and variables for yield prediction ....	49
3.5 Conclusion.....	51
3.6 References .....	52

CHAPTER FOUR: ASSESSING THE UTILITY OF UNMANNED AERIAL VEHICLE REMOTELY SENSED DATA FOR ESTIMATING MAIZE LEAF AREA INDEX (LAI) AND YIELD ACROSS THE GROWING SEASON: A SYNTHESIS .....	56
4.1 Introduction .....	56
4.2 Estimating Maize Leaf Area Index using UAV-derived multi-spectral remotely sensed data in smallholder farms. ....	57
4.3 Estimating Maize yield using UAV-derived multi-temporal data in smallholder farms of KwaZulu-Natal, South Africa .....	57
4.4 Conclusion.....	58
4.5 Recommendations for future studies.....	58

## List of Figures

<b>Figure 2.1:</b> a) Location of study site in Swayimane, KwaZulu-Natal, South Africa and b) The maize plot.....	16
<b>Figure 2.2:</b> a) DJI Matrice 300 UAV b) Mica Sense Altum sensor utilized in this study c) Flight path plan for the study area image acquisition and d) Altum sensor calibrated reflectance panel.....	19
<b>Figure 2.3:</b> Relationship between observed and predicted LAI based on the combination of traditional, red edge-based and new VIs using the RF Model for the a) V8-V10 b) V12-V14 c) VT-R1 d) R2-R3 and e) R3-R4 maize growth stages. ....	23
<b>Figure 2.4:</b> Variable importance scores of selected variables that exhibited the highest scores in predicting maize LAI for the a) V8-V10 b) V12-V14 c) VT-R1 d) R2-R3 and e) R3-R4 maize growth stages.....	24
<b>Figure 3.2:</b> a) DJI Matrice 300 UAV and b) Mica Sense Altum sensor utilized in this study. ....	39
<b>Figure 3.3:</b> Correlation between the grain yield and biomass at the R3-R4 stage using all samples.....	42
<b>Figure 3.5:</b> Variable importance of the i) biomass ii) grain yield and iii) proportional yield models for a) V8-V10 b) V12-V14 c) VT-R1 d) R2-R3 and e) R3-R4 maize growth stages. 46	
<b>Figure 3.6:</b> Spatial distribution of modelled maize a) biomass b) grain yield and c) proportional yield based on the most optimal RF models. ....	47

## List of Tables

<b>Table 2.1:</b> Maize growth stages. ....	17
<b>Table 2.2:</b> UAV derived VIs used in this study. ....	19
<b>Table 2.3:</b> Descriptive statistics of the actual maize LAI .....	21
<b>Table 3.1:</b> VIs used in this study.....	40

---

# CHAPTER ONE: GENERAL INTRODUCTION-MAPPING LEAF AREA INDEX AND YIELD OF MAIZE USING UNMANNED AERIAL VEHICLE DERIVED DATA

---

## 1.1 Introduction

Despite the great progress that has been made in meeting the basic needs of the poor and most vulnerable, food insecurity is still affecting a large number of people globally (Conceição *et al.*, 2016). Approximately, one in every four people in sub-Saharan Africa is food insecure (Mabhaudhi *et al.*, 2016). This is expected to increase over the years mainly because of population growth which will increase the demand for agricultural and horticultural production by 60% (Van Ittersum *et al.*, 2016). While it will be very difficult to eliminate this challenge, it can be minimized by access to timely information that would allow farmers to make informed food crop management decisions, especially for main staple crops such as maize to optimize crop production.

Maize (*Zea mays* L.), is one of the major crops grown in sub-Saharan Africa and is a staple food for approximately 50% of the population. Out of the 53 countries found in sub-Saharan Africa, 46 grow maize including, South Africa. Maize covers 60% of South Africa's cropping area and constitutes 70% of its grain production (Ndlovu *et al.*, 2021). Currently, South Africa is the leading maize producer on the African continent and its production is concentrated in the smallholder farms of the Free State, North West, Mpumalanga and KwaZulu-Natal provinces (Baloyi, 2011). Half the production of maize produced in South Africa is for human food consumption (Mditshwa, 2017). It is predicted that the South African growing population will result in an increased demand for food. The provision of timely accurate maize production estimates such as yield is critical for intervention measures to cover for possible deficits and leakages (Mkhabela *et al.*, 2005).

Crop yield is the production of a crop per unit area and is determined by the complex interaction between management, meteorological, chemical and physical conditions of a specific crop (Awad, 2019). Meanwhile, another key for assessing yield before harvest is measuring leaf area index (LAI) as a major biophysical parameter for determining crop development. LAI provides a measure of the density of foliage and has a close link to the evapotranspiration and

photosynthetic capacities of plants (Aboelghar *et al.*, 2011). It is a critical crop growth parameter that is closely associated with not only biomass accumulation, but also the yield at different phenological stages.

Before the development of new, advanced technologies, traditional methods such as physical field observations were used to determine plant productivity and yield. However, these are typically time-consuming, resource, and labour intensive and not ideal for continuous smallholder farm crop monitoring (Stein, 2018). Other traditional methods for estimating maize yield throughout its phenological cycle included models that integrate soils, climate and other environmental factors. These variables act as response functions to describe photosynthesis, development and evapotranspiration and yield for a specific crop as well as crop simulation models (Labus *et al.*, 2002). Even though these models are based on strong physical and physiological concepts, they lack spatial representativeness.

Geospatial technologies such as remote sensing have since provided an economical and efficient alternative for estimating and characterizing maize productivity based on detailed quantitative information available at different spatio-temporal scales (Labus *et al.*, 2002). Remotely sensed data can characterize photosynthetic active radiation (PAR), crop development, biomass accumulation, and LAI as proxies of productivity (Adam *et al.*, 2014; Pantazi *et al.*, 2016). Furthermore, geospatial technologies can be used to characterize actual crop yields in a spatially explicit manner. LAI can be considered as the main morphological parameter of the vegetation canopy that links remotely sensed data and plant photosynthesis, growth, productivity and yield (Son *et al.*, 2013; Peng *et al.*, 2019). For instance, Ahmad *et al.* (2020) predicted the annual variability of maize yield using Landsat 8 imagery and reported an  $R^2$  of 0.78 while Sun *et al.* (2019) used Sentinel-2 MSI derived red edge VIs to estimate maize LAI to an  $R^2$  of 0.85. However, despite the success of these applications, satellite-borne earth observation sensors are less suitable for monitoring crops at a farm scale. This is because satellite-borne earth observation sensors do not offer adequate spatial and temporal resolution suitable for characterizing maize crop attributes at a farm scale. For example, Landsat is restricted by the 16-day overpass frequency, which results in insufficient images being available for phenological changes, especially in areas that are consistently covered by clouds. Besides, more problems arise when faced with challenging weather conditions (Veroustraete, 2015). In addition, Sentinel-2 MSI's and Landsat's spatial resolutions of 10 m and 30 m, respectively, are also far too coarse for crop management applications at a farm scale. Taking

into consideration these challenges, there is a need for the farming industry to explore and embrace innovative precision farming technologies to improve agricultural productivity, in this case, in smallholder maize farms.

Recently, UAVs have offered a new solution for improving maize productivity i.e. maize yield and LAI estimates (Noureldin *et al.*, 2013). UAVs are complementary to high-altitude systems and can be an alternative source of information on maize crops at a farm scale. Their application helps bridge the gap between satellites and manned aircraft as well as the time-consuming, labour and resource-intensive conventional field surveys (Khaliq *et al.*, 2019). Moreover, UAVs are more advantageous in relation to satellite-borne earth observation systems, especially for the monitoring of plants, because they are less affected by weather conditions as they can be operated during overcast days (Psirofonina *et al.*, 2017). They offer flexibility in terms of flight planning and image acquisition scheduling which can be easily changed in near-real-time according to field conditions.

UAVs are relatively cheaper making them more suitable for farm-scale remote sensing applications (Khaliq *et al.*, 2019). Khaliq *et al.* (2019) compared satellite and UAV derived multispectral imagery for assessing the relative strength of each platform in representing vineyard variability. They concluded that satellite imagery could not be effectively used to describe vineyard variability, because of its lower spatial resolution and that UAVs are more advantageous for relatively smaller areas. While the groundbreaking study of Stroppiana *et al.* (2015) demonstrated the possibilities of remotely estimating rice yield through the use of aerial photography, the increasing application of UAVs is promising for in-field decision making. Using an S1000, SZ DJI UAV mounted with a Mini MCA camera with 6 channels to acquire imagery, Liu *et al.* (2019) computed VIs and texture metrics to estimate biomass of winter oilseed rape. They concluded that UAVs have a great potential in estimating the plot level above ground biomass through the combination of VIs and texture metrics. Using UAV imagery, Yao *et al.* (2017) sought to estimate the LAI of maize and concluded that this approach can be useful in estimating LAI in a shorter period of time and at lower costs. So, the different dynamics of UAV remotely sensed data together with their close relation to plant characteristics could play a role in establishing an effective method for estimating LAI and yield before harvest.

Most remote sensing systems including UAVs acquire spectral data of crops from the visible to the mid-infrared sections of the electro-magnetic spectrum which is often influenced by

strong chlorophyll absorption in the red (0.45-0.67 $\mu\text{m}$ ) and structural characteristics in the near-infrared region (0.7-0.9  $\mu\text{m}$ ). As a result, the red and near-infrared region is very convenient for vegetation monitoring and mapping because of such high reflectance values (Aboelghar *et al.*, 2011).

Several studies have illustrated that plant development, stress, LAI and yield can be characterized using conventional spectral reflectance data, but more accurately with VIs (Labus *et al.*, 2002; Mahajan and Raj, 2016; Raeva *et al.*, 2019). Vegetation Indices are sensitive to biochemical and biophysical variations in vegetation because they are calculated using the reflectance of two or more spectral wavelengths particularly from the visible and near-infrared regions of the EMS. VIs have been proven to be more robust than raw bands as they can overcome atmospheric impurities, soil background effects, effects of the zenith, and viewing angle while improving the signature of vegetation. Numerous studies have illustrated that VIs significantly improve crop yield estimations (Haboudane *et al.*, 2002; Aboelghar *et al.*, 2011; Noureldin *et al.*, 2013; Wahab *et al.*, 2018; Raeva *et al.*, 2019). Several studies have established the relationship between VIs and green biomass and concluded that VIs can be used to estimate yield before harvest (Satir and Berberoglu, 2016; Sun *et al.*, 2019; Ahmad *et al.*, 2020). However, the relationship between VIs and plant development can be negatively affected by canopy shadows and rapid vegetation growth rates as VIs perform differently (Awad, 2019).

It is therefore perceived that vegetation transforms such as VIs derived from drone remotely sensed data could be effective in estimating crop LAI and yields (Mahajan and Raj, 2016; Ahirwar *et al.*, 2019). VIs and LAI correlate strongly with crop productivity and plant physiological conditions under various dimensions and growth stages with remote sensing data from multi-sources (Jégo *et al.*, 2012). In this regard, identifying suitable variables for the estimation of crop yields (biomass) is critical because certain variables are weakly correlated with crop productivity attributes (LAI and yield ) or they are extremely auto-correlated (Lu, 2006). Given this challenge, a robust algorithm for distinguishing the most optimal variables to increase the estimation of crop yields is essential.

Advanced machine learning algorithms such as the RF regression are currently being used to resolve the overfitting problems and to select a subset of variables that best explain crop attributes such as LAI and yield (Mutanga and Adam, 2011; Mutanga *et al.*, 2012; Mditshwa, 2017; Ngie and Ahmed, 2018). The RF algorithm is widely used because it is a non-parametric statistical technique that uses a bagging-based approach to build an ensemble of regression

trees while being able to rank important variables that produce an independent measure of prediction error (Prasad *et al.*, 2006). To the best of our knowledge, very few studies have used the combination of bands and VIs derived from UAV imagery to estimate maize LAI and yield across its growing season in concert with the RF machine learning ensemble in smallholder farms. Therefore, this study aims to estimate maize LAI and yield using a combination of bands and VIs derived from UAV imagery in smallholder farms of KwaZulu-Natal, South Africa.

## **1.2 Aim and objectives**

The study aims to estimate maize LAI and yield across the growing season in smallholder farms of KwaZulu-Natal province using the RF algorithm on UAV derived data.

### **1.2.1 Specific objectives**

- To estimate LAI of maize crops using UAV derived VIs and RF regression across the growing season in smallholder croplands
- To estimate maize yield across the phenological cycle based on UAV derived data in conjunction with RF regression in smallholder croplands

## **1.3 Significance of the study**

Assessing crop health and productivity of smallholder farms using very high-resolution remotely sensed images will immensely contribute towards identifying and addressing the leakages and gaps between the actual and attainable yield in smallholder farms of southern Africa. Meanwhile, the monitoring of crop production through satellite remote sensing techniques has been affected by cloud cover, limited flight times associated with a coarse resolution which are inadequate for capturing the complex and heterogeneous nature of smallholder farming systems. Therefore, the application of a UAV in this study will bridge the gap between satellites and remotely sensed data products in characterizing crop production attributes in smallholder farms while reducing the high costs, intensive labour and lengthy times associated with conventional field surveys. In this regard, this research will monitor crop productivity throughout the growing season, which is important for immediate informed tactical decision making on the farm at any point in time during the cropping season. Lastly, this study will test and provide remotely sensed data-based yield estimation models which could help in informing the farmers on the expected yield before harvesting. This will be a step

towards designing an early warning technique enhancing preparedness with regards to either crop yield deficit or surplus, facilitating a timely decision-making process on possible intervention mechanisms.

#### **1.4 Structure of the thesis**

This thesis is comprised of four chapters. The first and last chapters represent the general introduction and the synthesis and concluding remarks, respectively. The second and third chapters present two research papers that seek to meet the above-mentioned objectives. Chapters two and three are presented as stand-alone chapters.

Chapter Two estimates maize LAI across the growing season using UAV derived VIS and field-collected LAI in concert with the RF algorithm at a smallholder farm. The study also gives insight into the most suitable variables in predicting maize LAI for each growth stage. Chapter Three focuses on predicting collected maize yields and finding the most suitable time and variables to do this. The RF algorithm was employed on the combination of bands and VIS derived from UAV imagery to achieve the results. Since the two chapters are presented as standalone manuscripts, overlaps between these two chapters were inevitable as they are addressing a single overarching objective of the study.

## 1.5 References

- Aboelghar, M, Arafat, S, Yousef, MA, El-Shirbeny, M, Naeem, S, Massoud, A, Saleh, NJTEJoRS and Science, S. 2011. Using SPOT data and leaf area index for rice yield estimation in Egyptian Nile delta. 14 (2): 81-89.
- Adam, E, Mutanga, O, Abdel-Rahman, EM and Ismail, R. 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.
- Ahirwar, S, Swarnkar, R, Bhukya, S and Namwade, G. 2019. Application of drone in agriculture. *International Journal Curr. Microbiological Applied Science* 8 2500-2505.
- Ahmad, I, Singh, A, Fahad, M, Waqas, MMJC and Agriculture, Ei. 2020. Remote sensing-based framework to predict and assess the interannual variability of maize yields in Pakistan using Landsat imagery. 178 105732.
- Awad, MM. 2019. Toward precision in crop yield estimation using remote sensing and optimization techniques. *Agriculture* 9 (3): 54.
- Baloyi, RT. 2011. Technical efficiency in maize production by small-scale farmers in Gathimba, Limpopo Province, South Africa. Unpublished thesis, University of Limpopo (Turfloop Campus),
- Conceição, P, Levine, S, Lipton, M and Warren-Rodríguez, AJFP. 2016. Toward a food secure future: Ensuring food security for sustainable human development in Sub-Saharan Africa. 60 1-9.
- Haboudane, D, Miller, JR, Tremblay, N, Zarco-Tejada, PJ and Dextraze, L. 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.
- Jégo, G, Pattey, E and Liu, JJFCR. 2012. Using Leaf Area Index, retrieved from optical imagery, in the STICS crop model for predicting yield and biomass of field crops. 131 63-74.
- Khaliq, A, Comba, L, Biglia, A, Ricauda Aimonino, D, Chiaberge, M and Gay, PJRS. 2019. Comparison of satellite and UAV-based multispectral imagery for vineyard variability assessment. 11 (4): 436.
- Labus, M, Nielsen, G, Lawrence, R, Engel, R and Long, D. 2002. Wheat yield estimates using multi-temporal NDVI satellite imagery. *International Journal of Remote Sensing* 23 (20): 4169-4180.

- Liu, Y, Liu, S, Li, J, Guo, X, Wang, S and Lu, J. 2019. Estimating biomass of winter oilseed rape using vegetation indices and texture metrics derived from UAV multispectral images. *Computers and Electronics in Agriculture* 166 105026.
- Lu, D. 2006. The potential and challenge of remote sensing-based biomass estimation. *International journal of remote sensing* 27 (7): 1297-1328.
- Mabhaudhi, T, Chibarabada, T, Modi, AJIjoer and health, p. 2016. Water-food-nutrition-health nexus: Linking water to improving food, nutrition and health in Sub-Saharan Africa. 13 (1): 107.
- Mahajan, U and Raj, B.2016. Drones for normalized difference vegetation index (NDVI), to estimate crop health for precision agriculture: A cheaper alternative for spatial satellite sensors. *International Conference on Innovative Research in Agriculture, Food Science, Forestry, Horticulture, Aquaculture, Animal Sciences, Biodiversity, Ecological Sciences and Climate Change (AFHABEC-2016), At Jawaharlal Nehru University*, 38-41.
- Mditshwa, S. 2017. Estimating Maize Grain Yield from Crop Growth Stages Using Remote Sensing and GIS in the Free State Province, South Africa. Unpublished thesis, University of Fort Hare,
- Mkhabela, MS, Mkhabela, MS, Mashinini, NNJA and Meteorology, F. 2005. Early maize yield forecasting in the four agro-ecological regions of Swaziland using NDVI data derived from NOAA's-AVHRR. 129 (1-2): 1-9.
- Mutanga, O and Adam, E.2011. High density biomass estimation: Testing the utility of Vegetation Indices and the Random Forest Regression algorithm. *34th International Symposium for Remote Sensing of the Environment (ISRSE), Sydney, Australia*,
- Mutanga, O, Adam, E and Cho, MA. 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.
- Ndlovu, HS, Odindi, J, Sibanda, M, Mutanga, O, Clulow, A, Chimonyo, VG and Mabhaudhi, TJRS. 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.
- Ngie, A and Ahmed, FJSAJoG. 2018. Estimation of Maize grain yield using multispectral satellite data sets (SPOT 5) and the random forest algorithm. 7 (1): 11-30.
- Noureldin, N, Aboelghar, M, Saady, H, Ali, AJTEJoRS and Science, S. 2013. Rice yield forecasting models using satellite imagery in Egypt. 16 (1): 125-131.

- Pantazi, XE, Moshou, D, Alexandridis, T, Whetton, RL and Mouazen, AM. 2016. Wheat yield prediction using machine learning and advanced sensing techniques. *Computers and Electronics in Agriculture* 121 57-65.
- Peng, Y, Li, Y, Dai, C, Fang, S, Gong, Y, Wu, X, Zhu, R, Liu, KJA and Meteorology, F. 2019. Remote prediction of yield based on LAI estimation in oilseed rape under different planting methods and nitrogen fertilizer applications. 271 116-125.
- Prasad, AM, Iverson, LR and Liaw, A. 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems* 9 (2): 181-199.
- Psirofonía, P, Samaritakis, V, Eliopoulos, P and Potamitis, I. 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.
- Raeva, PL, Šedina, J and Dlesk, A. 2019. Monitoring of crop fields using multispectral and thermal imagery from UAV. *European Journal of Remote Sensing* 52 (sup1): 192-201.
- Satir, O and Berberoglu, S. 2016. Crop yield prediction under soil salinity using satellite derived vegetation indices. *Field crops research* 192 134-143.
- Son, N, Chen, C, Chen, C, Chang, L, Duc, H and Nguyen, LJIJoRS. 2013. Prediction of rice crop yield using MODIS EVI– LAI data in the Mekong Delta, Vietnam. 34 (20): 7275-7292.
- Stein, N. 2018. The future of drones in the modern farming industry. *GEOMedia* 21 (5):
- Stroppiana, D, Migliazzi, M, Chiarabini, V, Crema, A, Musanti, M, Franchino, C and Villa, P. 2015. Rice yield estimation using multispectral data from UAV: A preliminary experiment in northern Italy. *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 4664-4667. IEEE,
- Sun, Y, Qin, Q, Ren, H, Zhang, T, Chen, SJIToG and Sensing, R. 2019. Red-edge band vegetation indices for leaf area index estimation from sentinel-2/msi imagery. 58 (2): 826-840.
- Van Ittersum, MK, Van Bussel, LG, Wolf, J, Grassini, P, Van Wart, J, Guilpart, N, Claessens, L, de Groot, H, Wiebe, K and Mason-D’Croz, DJPotNAoS. 2016. Can sub-Saharan Africa feed itself? 113 (52): 14964-14969.
- Veroustraete, F. 2015. The rise of the drones in agriculture. *EC agriculture* 2 (2): 325-327.

- Wahab, I, Hall, O and Jirström, MJD. 2018. Remote sensing of yields: Application of uav imagery-derived ndvi for estimating maize vigor and yields in complex farming systems in sub-saharan africa. 2 (3): 28.
- Yao, X, Wang, N, Liu, Y, Cheng, T, Tian, Y, Chen, Q and Zhu, YJRS. 2017. Estimation of wheat LAI at middle to high levels using unmanned aerial vehicle narrowband multispectral imagery. 9 (12): 1304.

---

## CHAPTER TWO: ESTIMATING MAIZE LEAF AREA INDEX USING UAV-DERIVED MULTI-SPECTRAL REMOTELY SENSED DATA IN SMALLHOLDER FARMS.

---

### Abstract

Understanding maize LAI is critical in assessing maize crop productivity. Spatially explicit information on smallholder farm maize production, specifically in sub-Saharan Africa remains scarce due to lack of incentives and appropriate technologies. These are required for efficient and timely monitoring and assessment of maize LAI to combat the challenges of food insecurity and end poverty. UAV imagery in concert with VIs obtained at the high spatial resolution, provide appropriate technologies for determining maize LAI at a farm scale. This study, therefore, evaluated the robustness of using UAV derived VIs in concert with the RF algorithm in estimating maize LAI across the growing season in a smallholder farm in KwaZulu-Natal, South Africa. The results showed that the optimal stage for estimating maize LAI using UAV derived VIs in concert with the RF ensemble in this study was during the vegetative stage (V8- V10) with a RMSE of 0.15, R<sup>2</sup> of 0.91 (RRMSE = 8 %). Generally, the findings also showed that UAV derived traditional, red edge-based and new VIs were able to accurately predict maize LAI across the growing season with R<sup>2</sup> of 0.89 - 0.93, RMSE of 0.15 – 0.65 m<sup>2</sup>/m<sup>2</sup> and RRMSE of 8.13 – 19.61%. The blue, red edge and NIR sections of the EMS were critical in predicting maize LAI. Furthermore, combining traditional, red edge-based and new VIs was useful in attaining high LAI estimation accuracies. These results are a step towards achieving robust, efficient and spatially explicit monitoring frameworks for sub-Saharan African smallholder farm productivity.

**Keywords:** Smallholder farming, Maize, Leaf Area Index, Remote Sensing, UAV, Vegetation Indices, Random Forest algorithm.

### 2.1 Introduction

Smallholder agriculture is a very important sector in sub-Saharan African economies. This is because most households in this region depend on it for their livelihoods (Gollin, 2014). According to Mango *et al.* (2017) smallholder croplands support about 70% of the households.

Above all, smallholder agriculture contributes about 15% of the Gross Domestic Product in Africa and it is estimated that it contributes 2.5% to the GDP of South Africa (Kamara *et al.*, 2019). Maize (*Zea Mays* L) is the most important and most widely grown grain crop in smallholder farms of sub-Saharan Africa. In addition, the maize industry plays a significant role in the region's economy because of its contribution to the domestic markets for local consumption and its importance in foreign exchange (Ndlovu *et al.*, 2021). It is therefore imperative to increase and optimize the production of maize for the country's food and nutritional security as well as for economic benefit. Considering that the population is projected to increase, demand for food will increase rapidly. Therefore, there is a dire need to generate monitoring frameworks for optimizing the agricultural productivity of staple crops such as maize in a spatially explicit manner by looking at the crops elements such as LAI, which is half the area of all leaves per unit of surface area. Furthermore, the actual contribution of smallholder farmers is generally unknown because they are widely distributed, spatially small, fragmented and highly diverse in terms of crop types.

Generally, crop productivity is evaluated based on its elements such as LAI, chlorophyll content concentration and yield. Amongst these elements, LAI is one important parameter that can be monitored to assess crop health status, canopy physiology and nutritional supply (Luo *et al.*, 2020b). As previously mentioned LAI is defined as half the area of all leaves per unit of surface area and its estimation has long been a research focus in the domain of remote sensing (Dong *et al.*, 2019). This is because LAI has a huge effect on the physiological process of the plant canopy, which is closely related to crop productivity. In addition, the total accumulation of LAI is strongly related to biomass accumulation and crop yield (Gitelson *et al.*, 2014). Therefore, monitoring LAI of maize at a farm-scale can assist in assessing crop condition variation across space and time for the detection of crop phenology and to model biomass and yield to optimize production in smallholder farms. The monitoring of LAI in crops is a technique used to diagnose and assess crop deficiencies and necessities such as fertilization (Tunca *et al.*, 2018). Therefore, monitoring and estimating maize LAI is of vital importance as it can assist in improving grain production, which is very critical in combating food insecurity while addressing the sustainable development goals of reducing hunger and poverty (Jin *et al.*, 2019; Peng *et al.*, 2019).

Crop LAI can be monitored and estimated through traditional methods and/or remote sensing techniques which are often associated with field surveys and point sample measurements (Tunca *et al.*, 2018). Despite the high accuracy associated with the traditional methods of

measuring LAI, they tend to be time-consuming and labour-intensive while lacking spatial representativeness (Martínez-Guanter *et al.*, 2019). In contrast, remote sensing technologies have increasingly become popular in agricultural research, because they offer fast and non-destructive ways of monitoring and estimating crop productivity parameters such as maize LAI (Yao *et al.*, 2017). Remote sensing provides both spatial and temporal information on crop responses to dynamic environmental conditions or information that relates directly to LAI (Peng *et al.*, 2019). Remote sensing data has been successfully used to derive important crop parameters, including among others; LAI, water use efficiency, chlorophyll, biomass fraction of photosynthetically active radiation (Tumlisan, 2017).

There are numerous ways of using remotely sensed information to estimate LAI. The simplest way is to establish an empirical relationship between the remotely sensed data such as spectral bands and VIs and measured LAI (Gao *et al.*, 2016). Several earth observation sensors have been used in estimating the LAI of maize in various continents resulting in optimal accuracies. These range from Landsat (González-Sanpedro *et al.*, 2008; Su *et al.*, 2019), moderate resolution imaging spectral radiometer (MODIS) (Kira *et al.*, 2017; Yu *et al.*, 2021) and recently Sentinel-2 multispectral instrument (MSI) (Luo *et al.*, 2020a; Amin *et al.*, 2021). Despite the optimal accuracies associated with the data from these satellite-borne sensors in the estimation of LAI, the trade-off between its spatial and temporal resolution limits its use in capturing crop LAI heterogeneity and dynamics at a farm-scale level (Martínez-Guanter *et al.*, 2019). Yang *et al.* (2021) states that medium spatial resolution products e.g., Landsat and Sentinel-2 have the potential to miss observations at critical growth stages because of their long revisit time (16 and 10 days respectively) as well as their coarser spatial resolution which is inadequate for small holder fields of less 5 Ha. In this regard, there is still a need to assess other sources of spatial data which could be cheaper, more flexible while offering very high spatial resolution data suitable for capturing crop LAI at farm to field scales.

The introduction of UAV remote sensing technology has offered an optimal source of remotely sensed data that is suitable for estimating crop productivity elements such as LAI (Martínez-Guanter *et al.*, 2019). UAV remote sensing technologies offer maximum flexibility in terms of temporal resolution since the flying times are user-determined. Their ability to fly at low altitude and portability makes them more suitable for farm-scale research when compared to satellite remote sensing, as they provide very high spatial resolution data of up to 5cm. (Gao *et al.*, 2016). It is anticipated that the very high resolution (VHR) spatial resolution in concert with a multispectral resolution which covers the red edge section of the EMS renowned for

mapping LAI of plants, could optimize the estimation of maize productivity in smallholder croplands. These UAVs have been used in estimating other crops yielding optimal accuracies. For instance, Kanning *et al.* (2018) successfully estimated wheat LAI with an  $R^2$  of 0.79 and an RMSE of 0.18. The proven compatibility of UAVs enables LAI estimation daily and at high resolution. However, most of these studies have been conducted based on single images in experimental plots outside the third world smallholder croplands. For an accurate estimation and outlook on a specific crop's productivity and yield, multitemporal images are required to understand the growth trajectory of the crop for informed decision making before the harvesting period. There is, therefore, a need to assess the utility of UAV derived multispectral data in assessing the productivity of staple crops such as maize in smallholder croplands of regions such as southern Africa where much hunger and poverty are rife and the need for optimizing crop production is imperative.

Literature also illustrates that the combination of VIs with robust machine learning algorithms improves the accuracies of crop productivity models. VIs depict biophysical parameters of the plant canopy such as biomass, greenness and LAI and are calculated using the reflectance of two or more spectral bands (Zhang *et al.*, 2009). VIs enhance the sensitivity to a specific crop parameter while suppressing the influence of other factors such as leaf and canopy structure (Sun *et al.*, 2019). Also, VIs counteract the impacts of soil background, atmospheric conditions, leaf pigment and inclination among others (Ngie and Ahmed, 2018). Several VIs have been proven to be strongly correlated with the LAI of maize (Sun *et al.*, 2019). These include the soil-adjusted VIs (SAVI and OSAVI), that were developed to reduce the impact of soil reflectance when LAI is low. In recent years, the advancement in sensor technologies has allowed application of red edge based VIs such as NDVI based on the red edge (NDVI<sub>RE</sub>), the normalized difference red edge (NDRE), the modified simple ratio red edge (MSR<sub>RE</sub>) and the red edge-based chlorophyll index (CI<sub>RE</sub>) were developed (Dong *et al.*, 2019). The new VIs and red edge-based VIs are very effective in estimating LAI especially from moderate to high LAI and are less influenced by canopy structures (Martínez-Guanter *et al.*, 2019). To the best of our knowledge, few studies have tested the effectiveness and performance of UAV derived VIs in assessing and monitoring crop growth.

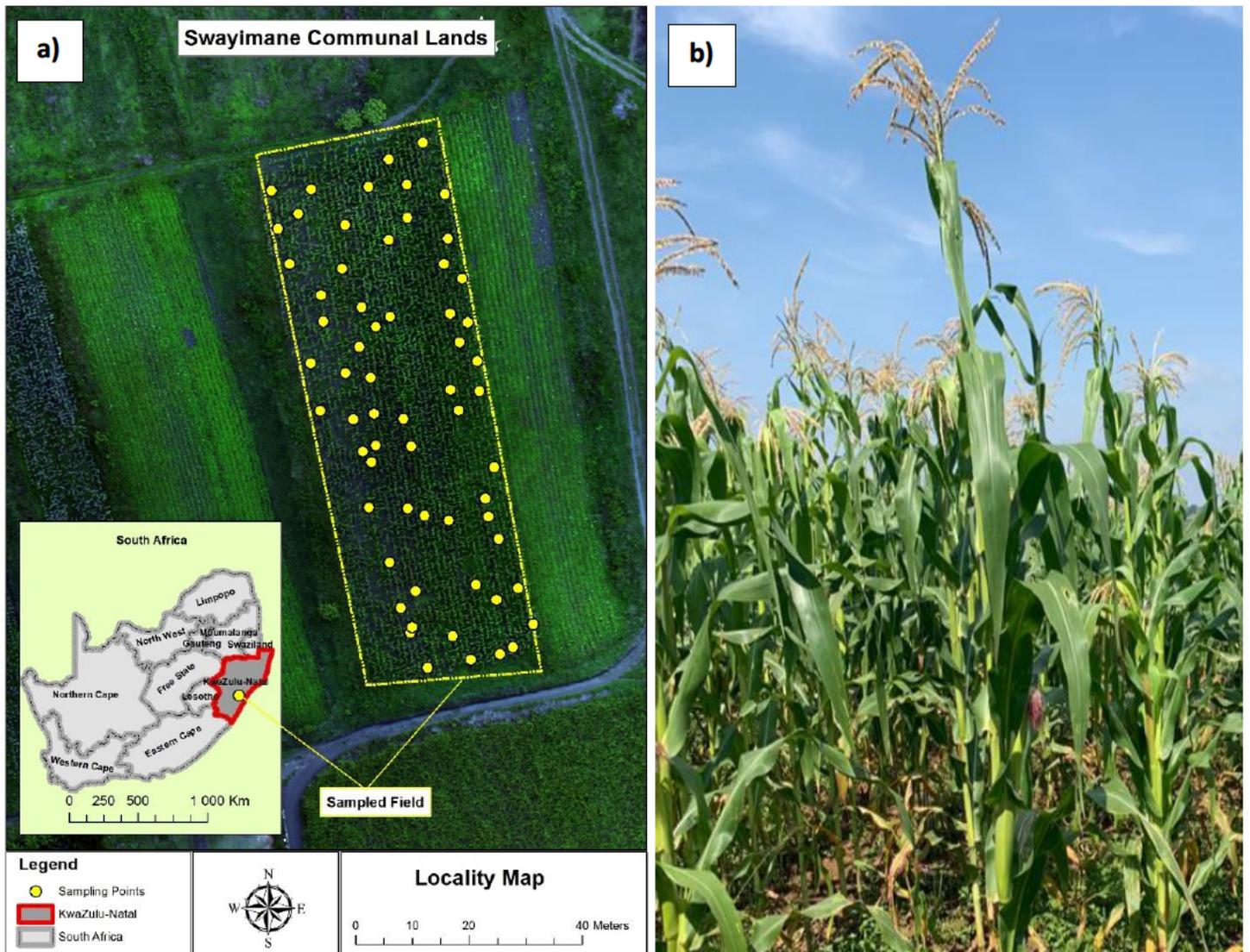
It is in this regard that this study sought to test the robustness of using UAV derived VIs in estimating maize LAI across the growing season. To achieve this, a robust algorithm, RF regression was used. This algorithm was chosen and used in this study based on its performance in previous studies and also considering that it resolves the overfitting problems,

select a subset of variables that best explain crop attributes such as LAI while being insensitive to small sample sizes. The specific objectives of this study were (i) to accurately estimate LAI using a combination of traditional, new and red edge-based VIs in conjunction with the RF algorithm and (ii) to develop a model for the estimation of maize LAI at each growth stage based on UAV data and field-collected LAI measurements.

## **2.2 Materials and methods**

### **2.2.1 Study site**

This study was conducted in a maize crop field in a smallholder farm located in Swayimane within the province of KwaZulu-Natal in South Africa (29°31'24''S and 30°41'37'' E) covering an area of 2699.005 m<sup>2</sup> (Figure 2.1). The area of Swayimane is characterized by a sub-humid climate with hot and humid summers, warm and dry winters. According to Miya *et al.* (2018), the area is characterized by a uni-modal rainfall pattern from November to March with an average precipitation of 900-1200 mm and an average temperature of 20°C. The people of Swayimane practice small scale farming involving sugarcane and maize production.



**Figure 2.1:** a) Location of study site in Swayimane, KwaZulu-Natal, South Africa and b) The maize plot.

The maize growth stages were divided into vegetative and reproductive stages (Table 2.1). The vegetative stages were further divided into VE, V(n) and VT where VE was the emergence stage (i.e. 0 days after emergence), V(n), where n represents the number of leaves as the plant grows from 7 to 55 days after emergence and VT the tassling stage. The reproductive stages were further divided into R1 to R6 (i.e. the silking to maturity, 56 to 160 days after emergence).

**Table 2.1:** Maize growth stages.

	<b>Growth stage</b>	<b>Name of growth stage</b>	<b>Days after emergence</b>	<b>Brief description</b>
Vegetative	(VE)	Emergence	0	Germination and Emergence
	V1	First leaf collar		
	V2	Second leaf collar	7	
	V3	Third leaf collar		
	V(n)	Nth leaf collar	21 - 55	Plant population established, cob development, active growth: cob size determined
	VT	Tassling	56	Pollination
Reproductive	R1	Silking	63	
	R2	Blister	70	Kernel development
	R3	Milk	91	Grain filling: nutrients transported to cob
	R4	Dough	105	
	R5	Dent	112	Physiological maturity and ready for harvest
	R6	Maturity	160	

### 2.2.2 Measurements of LAI

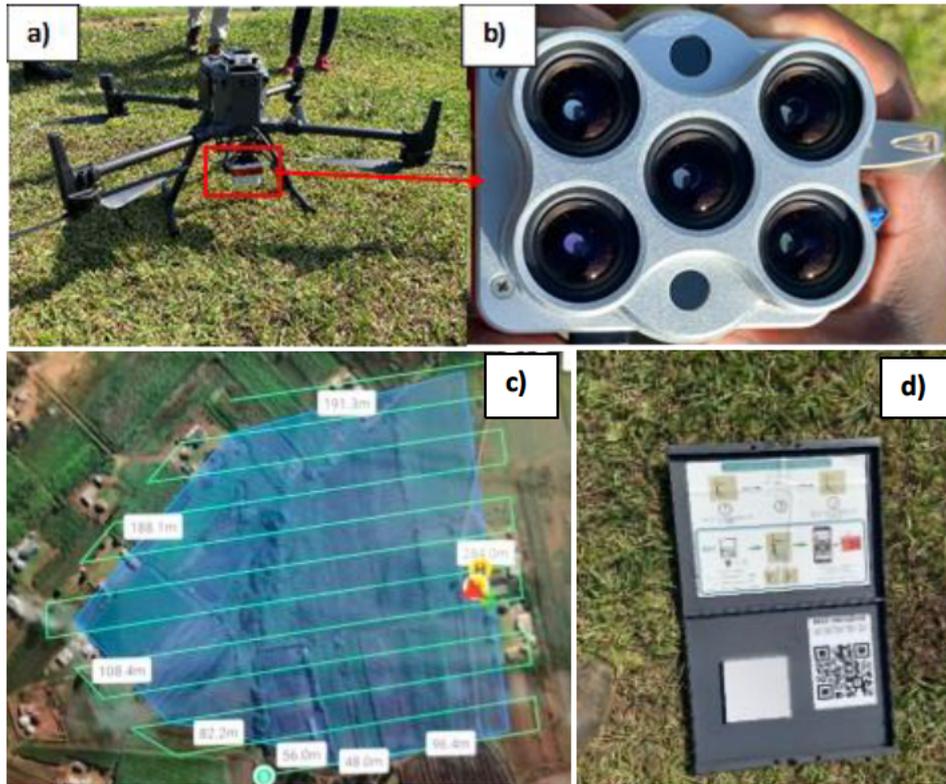
To estimate the LAI of maize in this study, a polygon map was generated in Google Earth Pro covering the maize fields. The polygon was imported into ArcMap 10.6 as a keyhole markup language (KML) file. This polygon was used to conduct the stratified random sampling in generating sampling points as well as in determining the flight path (section 2.2.3). A total of 63 points were generated and used for this analysis. These sampling points were loaded into a Trimble handheld Global Positioning System (GPS) with a sub-meter accuracy of 30 cm and used to locate the sampling points in the plot. At each of the selected and located sample points, a maize plant that was close to the location was marked for ease of identification and used for further sampling. Specifically, five field surveys were conducted during the growing season of maize at vegetative (V) and reproductive (R) growth stages. These were V8-V10 (18 March 2021), V10-V12 (31 March 2021), VT-R1 (12 April 2021), R2-R3 (28 April 2021) and R3-R4 (14 May 2021) stages. During the field surveys, images were acquired simultaneously with the measurement of maize LAI.

LAI was determined by using the LiCOR 2200C Plant Canopy Analyzer. The LiCOR 2200C has a fisheye optical sensor with five concentric rings centered at zenith angles 7°, 22°, 38°, 52° and 68° measuring radiation above and below the canopy to estimate canopy light interception and transmittance at five angles, by which LAI can be calculated from inversion of the Beer-Lambert law. This study utilized the 38° zenith angle.

### **2.2.3 Image acquisition and pre-processing**

The Mica Sense multi-spectral camera (Altum) (Figure 2.2 b) was mounted on a UAV (DJI Matrice 300) (Figure 2.2 a) in this study to acquire multi-spectral images of the study area. The Altum consists of five spectral bands (Blue, Green, Red, Red Edge and NIR) with a radiometric thermal camera for the thermal region of the EMS, which enables it to take multispectral and thermal imagery in a single flight. Before the flights, a flight plan (Figure 2.2 c) was established using the polygon of the study area that was created in Google Earth Pro and the polygon of the study area was imported as a kml file into the drone controller to generate the flight path (Figure 2.2 c). Calibration was also conducted just before flying the drone by acquiring images of the radiometric calibration target provided (Figure 2.2 d), which was set to be horizontal and not covered by any shadows. This was done to account for the illumination and atmospheric conditions prevalent during the flight. The flights were carried out on clear days between 10:00 AM and 1:00 PM local time. This was the most optimum time of the day when the solar zenith angle is minimal and radiation from the sun is at maximum. The flight altitude was kept at 100 m above the ground obtaining images with a spatial resolution of 5 cm.

The Pix4D software was then used to pre-process the UAV images. This was done to account for radiometric and geometric errors. The images were imported into the Pix4D software and thereafter relative calibration and radiometric correction followed by the stitching of the images to create ortho-images of the study area.



**Figure 2.2:** a) DJI Matrice 300 UAV b) Mica Sense Altum sensor utilized in this study c) Flight path plan for the study area image acquisition and d) Altum sensor calibrated reflectance panel.

#### 2.2.4 Data analysis

UAV derived VIs were computed and used to estimate LAI of maize across the growing season. The UAV bands were used to calculate 57 VIs (Table 2.2). Specifically, traditional, red edge-based and new VIs (nDVI) based on all possible combination of the 6 spectral bands were calculated in geographic information systems (GIS). The UAV data used in this study is summarized in Table 2.2.

**Table 2.2:** UAV derived VIs used in this study.

	Vegetation Index	Formula	Reference
Traditional	NDVI	$(\text{NIR}-\text{R})/(\text{NIR}+\text{R})$	(Aboelghar <i>et al.</i> , 2011)
	PNDVI	$(\text{NIR}-(\text{G}+\text{R}+\text{B}))/(\text{NIR}+(\text{G}+\text{R}+\text{B}))$	(Liu <i>et al.</i> , 2019)
	RBNDVI	$(\text{NIR}-(\text{R}+\text{B}))/(\text{NIR}+(\text{R}+\text{B}))$	(Li <i>et al.</i> , 2020)
	ENDVI	$((\text{NIR}+\text{G})-(2*\text{B}))/((\text{NIR}+\text{G})+(2*\text{B}))$	(Mditshwa, 2017)
	GBNDVI	$(\text{NIR}-(\text{G}+\text{B}))/(\text{NIR}+(\text{G}+\text{B}))$	(Peng <i>et al.</i> , 2019)

	Vegetation Index	Formula	Reference
	GRNDVI	$(\text{NIR}-(\text{G}+\text{R})) / (\text{NIR}+(\text{G}+\text{R}))$	(Peng <i>et al.</i> , 2019)
	GDVI	$\text{NIR}-\text{G}$	(Ramos <i>et al.</i> , 2020)
	CIgreen	$(\text{NIR}/\text{G})-1$	(Stroppiana <i>et al.</i> , 2015)
	CVI	$\text{NIR} * (\text{R} / (\text{G} * \text{G}))$	(Stroppiana <i>et al.</i> , 2015)
	GLI	$((2 * \text{G}) - \text{R} - \text{B}) / ((2 * \text{G}) + \text{R} + \text{B})$	(Tumlisan, 2017)
	EVI	$2,5 * ((\text{NIR}-\text{R}) / (\text{NIR}+(6 * \text{B}) - (7,5 * \text{B}))) + 1$	(Potgieter <i>et al.</i> , 2007)
	EVI2	$2,4 * ((\text{NIR}-\text{R}) / (\text{NIR}+\text{R}+1))$	(Zheng <i>et al.</i> , 2019)
	EVI3	$2,5 * ((\text{NIR}-\text{R}) / (\text{NIR}+(2,4 * \text{R})+1))$	(Sibanda <i>et al.</i> , 2017a)
	CI	$(\text{R}-\text{B})/\text{B}$	(Yao <i>et al.</i> , 2017)
	IPVI	$(\text{NIR}/\text{NIR}+\text{R})/2 * (\text{NDVI}+1)$	(Zhang <i>et al.</i> , 2019c)
	SAVI	$((\text{NIR}-\text{R}) / (\text{NIR}+\text{R}+0,5)) * (1+0,5)$	(Mditshwa, 2017)
	OSAVI	$(\text{NIR}-\text{R}) / (\text{NIR}+\text{R}+0,16)$	(Peng <i>et al.</i> , 2019)
	SR	$(\text{NIR}/\text{R})$	(Peng <i>et al.</i> , 2019)
Red edge based	NDRE	$(\text{NIR}-\text{RE}) / (\text{NIR}+\text{RE})$	(Sun <i>et al.</i> , 2019)
	CI <sub>RE</sub>	$(\text{NIR}/\text{RE})-1$	(Sun <i>et al.</i> , 2019)
	CCCI	$((\text{NIR}-\text{RE}) / (\text{NIR}+\text{RE})) / ((\text{NIR}-\text{R}) / (\text{NIR}+\text{R}))$	(Al-Gaadi <i>et al.</i> , 2016)
	NDVI <sub>RE</sub>	$(\text{RE}-\text{R}) / (\text{RE}+\text{R})$	(Dong <i>et al.</i> , 2019)
New	nDVI	$(\text{R}_{\text{Yi}}) - (\text{R}_{\text{Yj}}) / (\text{R}_{\text{Yi}}) + (\text{R}_{\text{Yj}})$	This study

\*Where  $\text{R}_{\text{Yi}}$  and  $\text{R}_{\text{Yj}}$  are different Altum spectral bands.

### 2.2.5 Maize LAI prediction

RF algorithm was used to estimate maize LAI across the growing season. RF is amongst the group of machine learning techniques developed to advance the classification and regression trees algorithm through the compilation of a huge set of decision trees. It is advantageous in that it can optimize the regression trees (*n-tree*) method by the combination of a large set of decision trees. The machine learning technique was implemented using the R interface. In R, the *doBest* function was used to optimize the *n-tree* and *mtry* parameters to 200 and 5, respectively, which was the best combination of parameters after testing the *n-tree* values in increments of 100 to 2500 and the *mtry* values in increments of 1 to 5. The resulting models of each growth stage were then compared to assess the best performing model.

### 2.2.6 Accuracy assessment

To assess the performance of the models, the dataset ( $n = 63$ ) was split into 70% training ( $n = 44$ ) and 30% test ( $n = 19$ ) dataset. The training data was used to train the model and the test

data was used to evaluate the estimation models. The performance of each model in estimating LAI was evaluated using the coefficient of determination, the root mean square error (RMSE) and relative root mean square error (RRMSE). The model that yielded a high  $R^2$  and low RMSE was then used to create a LAI map for the study site in ArcMap 10.6.

$$RMSE = \sqrt{\frac{SSE^2}{n}} \quad (1)$$

$$RRMSE \% = \frac{RMSE}{MEAN} \times 100 \quad (2)$$

## 2.3 Results

### 2.3.1 Descriptive statistics

Descriptive statistics of LAI measured in the field for all the growth stages (i.e. V8-V10, V10-V12, VT-R1, R2-R3, R3-R4) are shown in Table 2.3. The highest average maize LAI of 3.44 was obtained from the R3-R4 growth stage and the lowest was observed for the V8-V10 growth stage which was 1.78. Furthermore, the R3-R4 growth stage had the highest maximum LAI of 6.29 compared to the rest. The V8-V10 stage had the lowest LAI of 0.47 compared to the rest. The mean of LAI increased along with an increase in maize crop productivity.

**Table 2.3:** Descriptive statistics of the actual maize LAI

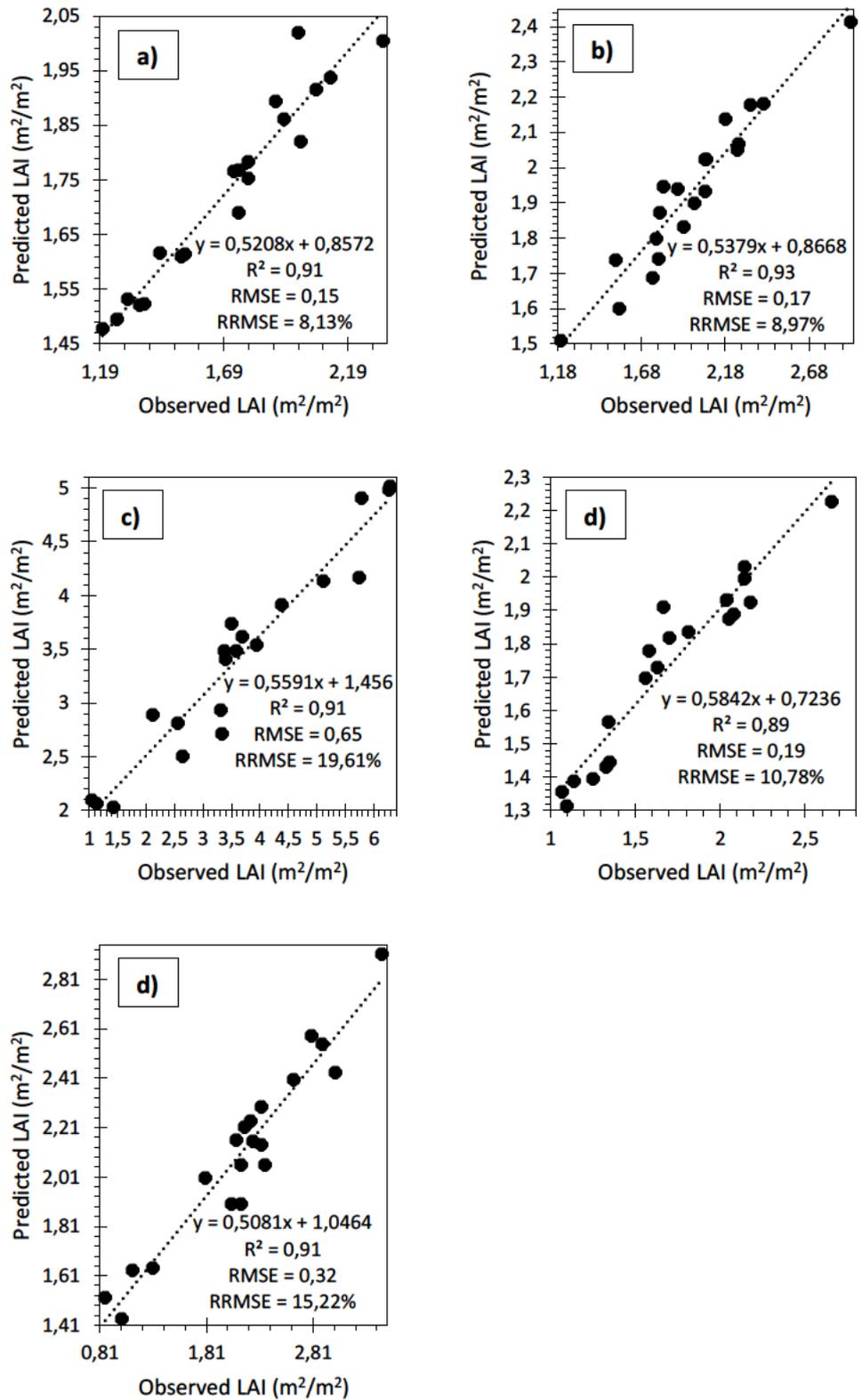
Growth Stage	N	Mean	Std. dev	Min	Max
V8-V10	63	1.78	0.35	0.47	1.37
V12-V14	63	1.82	1.37	1.01	2.93
VT-R1	63	2.07	1.14	2.24	3.46
R2-R3	63	3.29	1.1	2.66	5.15
R3-R4	63	3.44	0.63	3.53	6.29

### 2.3.2 Derived Maize LAI prediction models and their accuracies

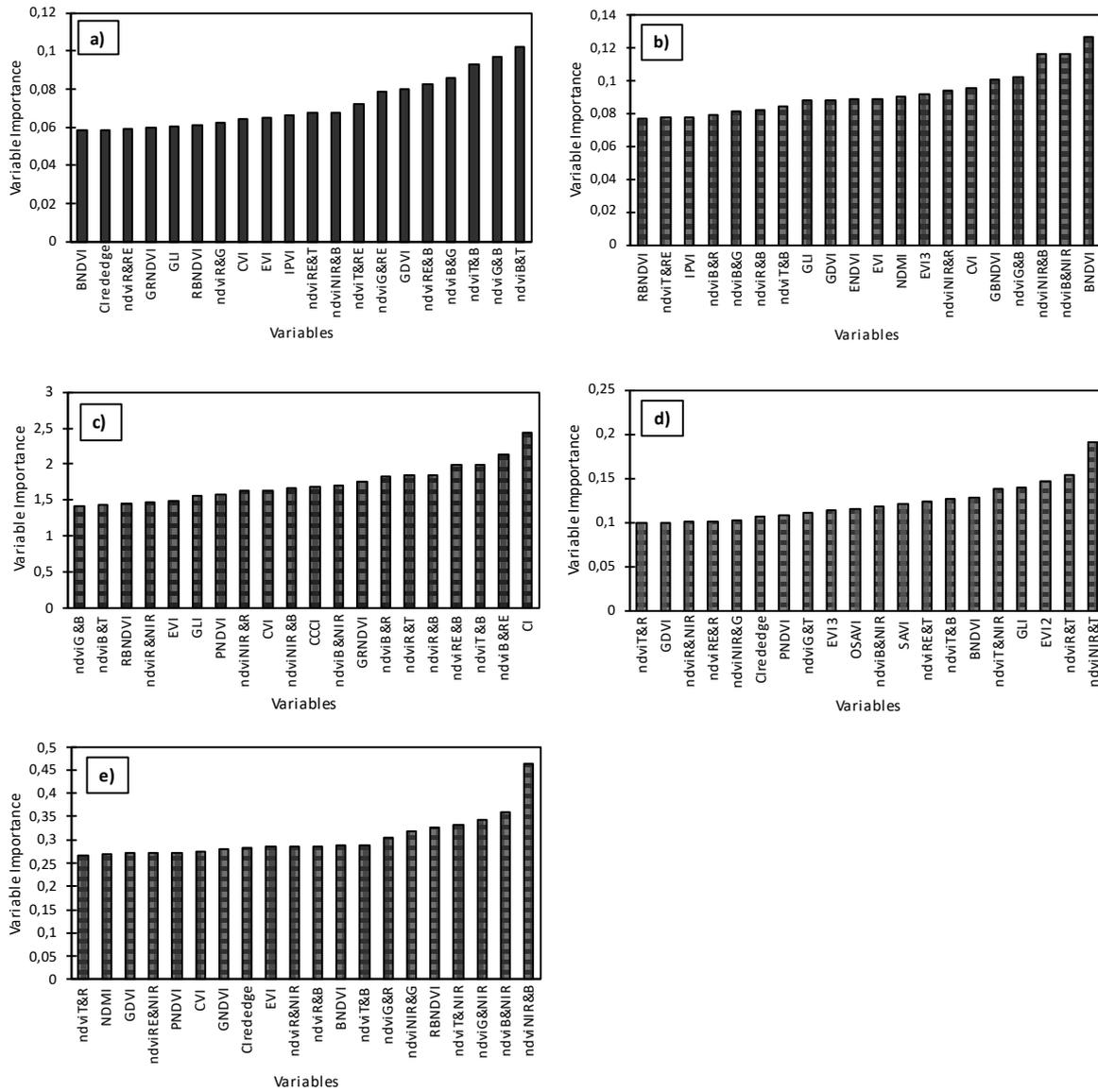
Figure 2.3 demonstrates the model accuracies obtained in estimating maize LAI based on the RF algorithm. The accuracies of the prediction models were moderate to high across the different growth stages of maize. For instance, when predicting LAI at the V8-V10 growth stage, a good model with an  $R^2$  of 0.91, RMSE of 0.15  $m^2/m^2$  and RRMSE of 8.13% was obtained (Figure 2.3 a). The most optimal variables in estimating maize LAI at this stage were *ndviB&T* and *ndviG&B* (Figure 2.4 a). Similarly, the V12-V14 growth stage predicted maize

LAI at an  $R^2$  of 0.93, RMSE of  $0.17 \text{ m}^2/\text{m}^2$  and RRMSE of 8.97% (Figure 2.3 b) with BNDVI and ndviB&NIR being more influential for this model (Figure 2.4 b).

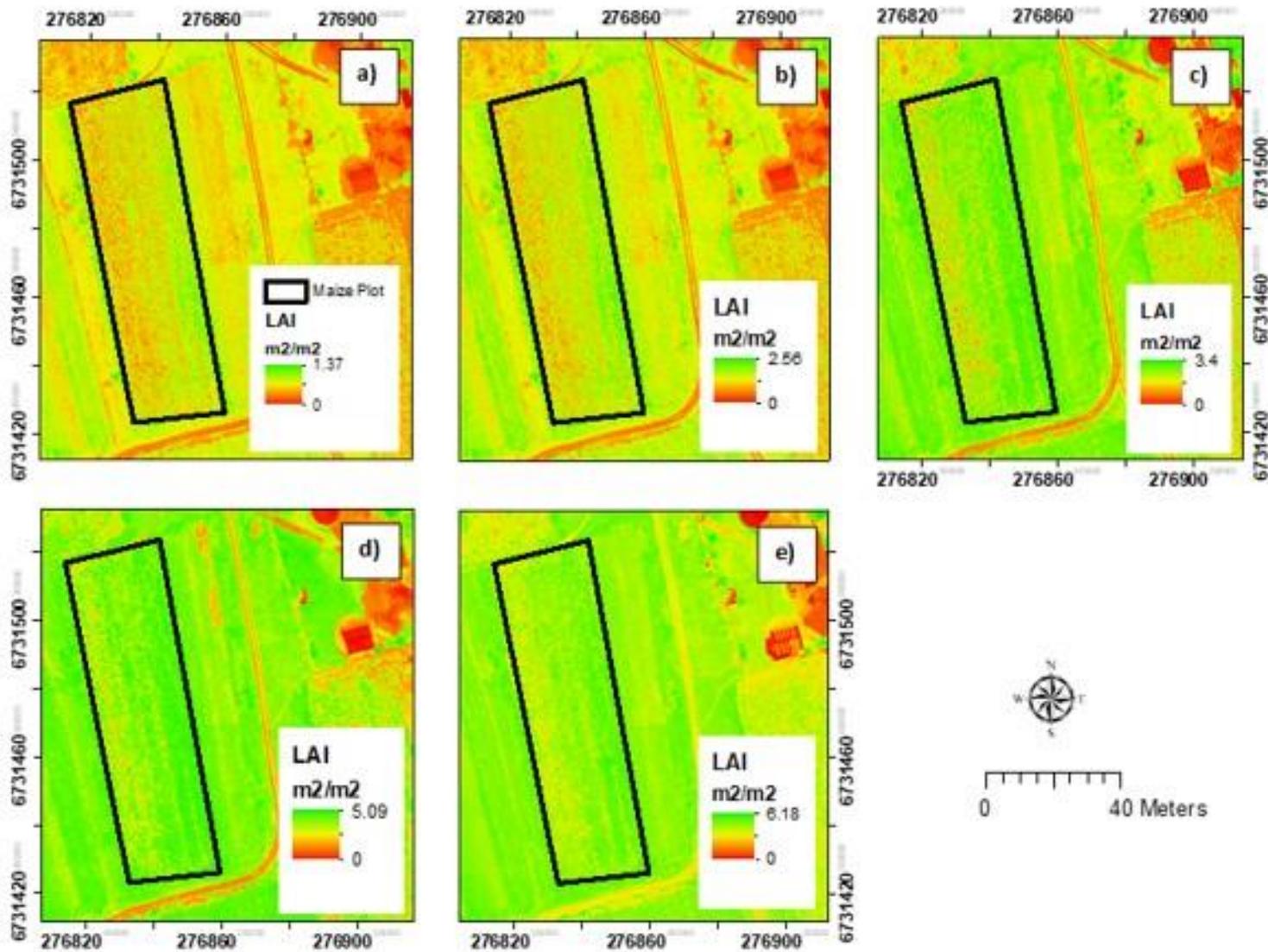
The VT-R1 growth stage demonstrated a moderate prediction accuracy in estimating maize LAI ( $R^2 = 0.91$ , RMSE =  $0.65 \text{ m}^2/\text{m}^2$  and RRMSE = 19.61%) (Figure 2.3 c). The most suitable predictor variables for this stage included ndviNIR&T and ndviR&T (Figure 2.4 c). This was followed by a drastic improvement in the R2-R3 growth stage with an  $R^2$  of 0.89, RMSE of  $0.19 \text{ m}^2/\text{m}^2$  and RRMSE of 10.78% (Figure 2.3 d). The most influential variables for this prediction were CI and ndviB&RE (Figure 2.4 d). The R3-R4 growth stage also yielded a moderate model ( $R^2 = 0.91$ , RMSE =  $0.32 \text{ m}^2/\text{m}^2$  and RRMSE = 15.22%) (Figure 2.3 e). The most optimal variables for the prediction of maize LAI at the R3-R4 growth stage were ndviNIR&B and ndviB&NIR (Figure 2.4 e).



**Figure 2.3:** Relationship between observed and predicted LAI based on the combination of traditional, red edge-based and new VIs using the RF Model for the a) V8-V10 b) V12-V14 c) VT-R1 d) R2-R3 and e) R3-R4 maize growth stages.



**Figure 2.4:** Variable importance scores of selected variables that exhibited the highest scores in predicting maize LAI for the a) V8-V10 b) V12-V14 c) VT-R1 d) R2-R3 and e) R3-R4 maize growth stages



**Figure 2.5:** Spatial distribution of modelled maize LAI for the a) V8-V10 b) V12-V14 c) VT-R1 d) R2-R3 and e) R3-R4 growth stages based on the RF models.

## 2.4 Discussion

This study sought to test the utility of UAV derived VIs in estimating maize LAI across the growing season based on the Altum sensor mounted on the DJI Matrice 300 UAV data. This study specifically sought to estimate LAI using the combination of UAV derived traditional, new and red edge-based and the RF algorithm across the growing season in the context of smallholder farms.

### 2.4.1 Predicting maize LAI

The results of this study showed that maize LAI was optimally estimated at the V8-V10 growth stage to an  $R^2$  of 0.91, RMSE of  $0.15 \text{ m}^2/\text{m}^2$  and RRMSE of 8.13% with the most influential variables being the ndviG&B and ndviB&T derived using the green, blue and thermal spectral variables. This demonstrates the sensitivity of maize LAI to the blue, green and thermal regions of the EMS in the early growth stages. Literature states that the blue section of the EMS is sensitive to green vegetation as it is used by plants during the process of photosynthesis, which results in its absorption by vegetation, hence its influence in predicting LAI (Dou *et al.*, 2019; Grajek *et al.*, 2020). Previous studies also states that the presence of bright green vegetation on the ground during the early stages of plant growth results in a high reflectance in the green region of the EMS which explains the sensitivity of maize LAI to the green section of the EMS at the V8-V10 for this study (Ren and Zhou, 2019; Sharifi and Agriculture, 2020). These findings are in agreement with a study by Motohka *et al.* (2010) who noticed a decrease in green reflection when leaves changed from a bright green in the early stages of the season to a dark green colour towards the later stages of the season. This was caused firstly by the end of the formation of new leaves which was also detected using spectral variables derived from the green section of the EMS.

In addition, the thermal band was also amongst the most influential spectral predictor variables. This could be explained by the fact that during the V8-V10 growth stage, there is low foliage density and previous studies states that when there is low foliage density, the soil tends to absorb more heat resulting in a high reflectance of the thermal region from the ground, explaining the sensitivity to the thermal band during this stage for this study (Filgueiras *et al.*, 2019).

In estimating maize LAI during the V12-V14 growth stage, UAV derived VIs yielded an  $R^2$  of 0.93, a RMSE of  $0.17 \text{ m}^2/\text{m}^2$  and an RRMSE of 8.97% based on spectral variables derived from the blue and NIR regions of the EMS (BNDVI). The results of this growth stage signify

the sensitivity of maize LAI to the blue and NIR sections of the EMS to maize LAI during the V12-V14 growth stage. As mentioned earlier, the blue region of the EMS plays an important role in plant photosynthesis which is a daily process that a plant undergoes, hence the importance of the blue waveband in this growth stage as well (Grajek *et al.*, 2020). In explaining the sensitivity of maize LAI to the NIR section of the EMS, literature states that this section is very important in vegetation monitoring as healthy vegetation tends to reflect highly in this section, hence its influence in estimating LAI (Fu *et al.*, 2014; Liu *et al.*, 2019; Martínez-Guanter *et al.*, 2019). Specifically, the presence and increased foliage density of maize plants result in leaves strongly reflecting in the NIR section of the EMS.

In predicting maize LAI at the VT-R1 growth stage, UAV derived VIs produced a prediction model with an  $R^2$  of 0.91, a RMSE of 0.65  $m^2/m^2$  and an RRMSE of 19.61% based on the combination of spectral variables derived from the red and NIR regions of the EMS (ndviR&T and ndviNIR&T). The red and NIR sections of the EMS are of significance in vegetation monitoring. Specifically, vegetation tends to strongly absorb in the red section and as mentioned earlier reflect highly in the NIR section explaining the sensitivity of maize LAI to these sections of the EMS. These sections of the EMS are of great value in explaining LAI, because the level of absorption in the red section and reflection in the NIR section is based on the amount of vegetation present on the ground. Therefore the higher the absorption and reflection in the red and NIR sections respectively, the higher the amount of vegetation on the ground and vice versa (Ramos *et al.*, 2020).

When predicting maize LAI in the R2-R3 growth stage using UAV derived VIs, a model with an  $R^2$  of 0.89, a RMSE of 0.19  $m^2/m^2$  and an RRMSE of 10.78% was obtained again, based on the indices derived using the blue and red wavebands together with the red edge wavebands (ndviB&RE and CI). This indicates a sensitivity of maize LAI to the blue, red and red edge sections of the EMS in the R2-R3 growth stage. The contribution of the red edge could be attributed to the fact that chlorophyll and biomass are sensitive to the red edge (Tumlisan, 2017). Specifically, LAI is correlated to chlorophyll and biomass, hence the influence of the red edge in predicting LAI (Sun *et al.*, 2019). Finally, in the R3-R4 growth stage, maize LAI was sensitive to the blue and NIR sections of the EMS. These produced an optimal model with  $R^2$  of 0.91, a RMSE of 0.32  $m^2/m^2$  and an RRMSE of 15.22%. The influence of the blue and NIR bands in predicting maize LAI, could be explained by the role of the blue band in the process of photosynthesis and the strong reflection of vegetation in the NIR section of the EMS as mentioned previously.

#### **2.4.2 The performance of combining UAV derived traditional, red edge-based and new VIs in predicting maize LAI**

Results in this study show that combining traditional, red edge-based and new VIs produced good yield prediction models for all the growth stages. This could be caused by the sensitivity of the red edge region of the EMS together with the ability of VIs to enhance vegetation features to the variation in LAI changes (Sun *et al.*, 2019). Across the growing season, LAI changes as was shown in Table 2.3. During the early stages (V8-V10 and V12-V14) of the growing season, leaves are small and as maize grows so does the leaves. This results in the alteration of LAI across the phenological cycle. Therefore, the red edge section of the EMS better detects the spectral reflectance of these growth stages which shifts with vegetation growth, expanding on the performance of VIs (Sibanda *et al.*, 2021). Additionally, the red edge region of the EMS is also sensitive to chlorophyll content variability, which increases as maize grows. This also contributes to the high accuracies of the estimation of maize LAI when VIs are combined with the red edge.

Meanwhile, VIs are sensitive to distinctive spectral properties of green vegetation in the image caused by the reflectance of maize at various growth stages on particular spectral bands such as the red, red edge and NIR (Kanning *et al.*, 2018). Furthermore, VIs are highly correlated with LAI. This then boosts the robustness of VIs in estimating LAI. VIs are also sensitive to the LAI variability caused by the different stages of the phenological cycle as well as the accumulating chlorophyll content throughout the crops growing season (Leroux *et al.*, 2019). It is in this regard that high estimation accuracies of LAI are realized when the traditional, red edge-based and new VIs are combined. In addition, VIs optimize the characterization of spatial information on vegetation while increasing the range of LAI to optimal levels (He *et al.*, 2019). The results of this study are consistent with those of Fu *et al.* (2014) who reported that models derived from the combination of VIs and band parameters could effectively increase the accuracy of winter wheat biomass estimation when compared with the performance of bands or VIs as stand-alone data. Another study by He *et al.* (2019) estimated the LAI of rice-based on a new vegetation index and concluded that the combination of the NIR and red edge bands was the best in predicting rice LAI ( $R^2 = 0.6$ ,  $RMSE = 1.41 \text{ m}^2/\text{m}^2$ ).

#### **2.5 Conclusion**

This study sought to test the utility of UAV derived VIs in estimating maize LAI across the growing season based on the Altum sensor mounted on the DJI Matrice 300 UAV data in a smallholder farm. Based on the findings of this study it can be concluded that:

- Maize LAI can be optimally estimated using UAV derived VIs across the growing season;
- The blue, green, red edge and NIR sections of the EMS were influential in estimating Maize LAI;
- Combining traditional, red edge-based and new VIs was useful in attaining high LAI estimation accuracies.

Quantitative assessments of maize LAI attained in this study are a step towards developing non-destructive and cost-effective methods for routine and timely monitoring of maize LAI in smallholder farms for precision farming and increasing crop productivity. The findings indirectly contribute towards poverty alleviation by ensuring food security.

## 2.6 References

- Aboelghar, M, Arafat, S, Yousef, MA, El-Shirbeny, M, Naeem, S, Massoud, A, Saleh, NJTEJoRS and Science, S. 2011. Using SPOT data and leaf area index for rice yield estimation in Egyptian Nile delta. 14 (2): 81-89.
- Al-Gaadi, KA, Hassaballa, AA, Tola, E, Kayad, AG, Madugundu, R, Alblewi, B and Assiri, F. 2016. Prediction of potato crop yield using precision agriculture techniques. *PloS one* 11 (9):
- Amin, E, Verrelst, J, Rivera-Caicedo, JP, Pipia, L, Ruiz-Verdú, A and Moreno, J. 2021. Prototyping Sentinel-2 green LAI and brown LAI products for cropland monitoring. *Remote Sensing of Environment* 255 112168.
- Dong, T, Liu, J, Shang, J, Qian, B, Ma, B, Kovacs, JM, Walters, D, Jiao, X, Geng, X and Shi, YJRSoE. 2019. Assessment of red-edge vegetation indices for crop leaf area index estimation. 222 133-143.
- Dou, H, Niu, G and Gu, MJH. 2019. Photosynthesis, morphology, yield, and phytochemical accumulation in basil plants influenced by substituting green light for partial red and/or blue light. 54 (10): 1769-1776.
- Filgueiras, R, Mantovani, EC, Dias, SHB, Fernandes Filho, EI, da Cunha, FF and Neale, CMUJEJoA. 2019. New approach to determining the surface temperature without thermal band of satellites. 106 12-22.
- Fu, Y, Yang, G, Wang, J, Song, X, Feng, HJC and Agriculture, Ei. 2014. Winter wheat biomass estimation based on spectral indices, band depth analysis and partial least squares regression using hyperspectral measurements. 100 51-59.
- Gao, L, Yang, G, Li, H, Li, Z, Feng, H, Wang, L, Dong, J and He, PJZSNXCJoE-A. 2016. Winter wheat LAI estimation using unmanned aerial vehicle RGB-imaging. 24 (9): 1254-1264.
- Gitelson, AA, Peng, Y, Arkebauer, TJ and Schepers, J. 2014. Relationships between gross primary production, green LAI, and canopy chlorophyll content in maize: Implications for remote sensing of primary production. *Remote Sensing of Environment* 144 65-72.
- Gollin, DJIWPI, London. 2014. Smallholder agriculture in Africa.
- González-Sanpedro, M, Le Toan, T, Moreno, J, Kergoat, L and Rubio, E. 2008. Seasonal variations of leaf area index of agricultural fields retrieved from Landsat data. *Remote Sensing of Environment* 112 (3): 810-824.
- Grajek, H, Rydzyński, D, Piotrowicz-Cieślak, A, Herman, A, Maciejczyk, M and Wieczorek, ZJC. 2020. Cadmium ion-chlorophyll interaction–Examination of spectral properties

- and structure of the cadmium-chlorophyll complex and their relevance to photosynthesis inhibition. 261 127434.
- He, J, Zhang, N, Su, X, Lu, J, Yao, X, Cheng, T, Zhu, Y, Cao, W and Tian, YJRS. 2019. Estimating leaf area index with a new vegetation index considering the influence of rice panicles. 11 (15): 1809.
- Jin, Z, Azzari, G, You, C, Di Tommaso, S, Aston, S, Burke, M and Lobell, DBJRSOE. 2019. Smallholder maize area and yield mapping at national scales with Google Earth Engine. 228 115-128.
- Kamara, A, Conteh, A, Rhodes, ER and Cooke, RA. 2019. The relevance of smallholder farming to African agricultural growth and development. *African Journal of Food, Agriculture, Nutrition and Development* 19 (1): 14043-14065.
- Kanning, M, Kühling, I, Trautz, D and Jarmer, TJRS. 2018. High-resolution UAV-based hyperspectral imagery for LAI and chlorophyll estimations from wheat for yield prediction. 10 (12): 2000.
- Kira, O, Nguy-Robertson, AL, Arkebauer, TJ, Linker, R and Gitelson, AA. 2017. Toward generic models for green LAI estimation in maize and soybean: Satellite observations. *Remote Sensing* 9 (4): 318.
- Leroux, L, Castets, M, Baron, C, Escorihuela, M-J, Bégué, A and Seen, DLJEJoA. 2019. Maize yield estimation in West Africa from crop process-induced combinations of multi-domain remote sensing indices. 108 11-26.
- Li, B, Xu, X, Zhang, L, Han, J, Bian, C, Li, G, Liu, J, Jin, LJJoP and Sensing, R. 2020. Above-ground biomass estimation and yield prediction in potato by using UAV-based RGB and hyperspectral imaging. 162 161-172.
- Liu, Y, Liu, S, Li, J, Guo, X, Wang, S and Lu, J. 2019. Estimating biomass of winter oilseed rape using vegetation indices and texture metrics derived from UAV multispectral images. *Computers and Electronics in Agriculture* 166 105026.
- Luo, P, Liao, J and Shen, G. 2020a. Combining spectral and texture features for estimating leaf area index and biomass of maize using Sentinel-1/2, and Landsat-8 data. *IEEE Access* 8 53614-53626.
- Luo, P, Liao, J and Shen, GJIA. 2020b. Combining spectral and texture features for estimating leaf area index and biomass of maize using Sentinel-1/2, and Landsat-8 data. 8 53614-53626.

- Mango, N, Siziba, S and Makate, C. 2017. The impact of adoption of conservation agriculture on smallholder farmers' food security in semi-arid zones of southern Africa. *Agriculture & Food Security* 6 (1): 1-8.
- Martínez-Guanter, J, Egea, G, Pérez-Ruiz, M and Apolo-Apolo, O. 2019. Estimation of the leaf area index in maize based on UAV imagery using deep learning techniques. In: *Precision agriculture '19*. Wageningen Academic Publishers.
- Mditshwa, S. 2017. Estimating Maize Grain Yield from Crop Growth Stages Using Remote Sensing and GIS in the Free State Province, South Africa. Unpublished thesis, University of Fort Hare,
- Miya, SP, Modi, AT, Tesfay, SZ, Mabhaudhi, TJSaJoP and Soil. 2018. Maize grain soluble sugar and protein contents in response to simulated hail damage. 35 (5): 377-383.
- Motohka, T, Nasahara, KN, Oguma, H and Tsuchida, S. 2010. Applicability of Green-Red Vegetation Index for Remote Sensing of Vegetation Phenology. 2 (10): 2369-2387.
- Ndlovu, HS, Odindi, J, Sibanda, M, Mutanga, O, Clulow, A, Chimonyo, VG and Mabhaudhi, TJRS. 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.
- Ngie, A and Ahmed, FJSaJoG. 2018. Estimation of Maize grain yield using multispectral satellite data sets (SPOT 5) and the random forest algorithm. 7 (1): 11-30.
- Peng, Y, Li, Y, Dai, C, Fang, S, Gong, Y, Wu, X, Zhu, R, Liu, KJA and Meteorology, F. 2019. Remote prediction of yield based on LAI estimation in oilseed rape under different planting methods and nitrogen fertilizer applications. 271 116-125.
- Potgieter, AB, Apan, A, Dunn, P and Hammer, G. 2007. Estimating crop area using seasonal time series of Enhanced Vegetation Index from MODIS satellite imagery. *Australian Journal of Agricultural Research* 58 (4): 316-325.
- Ramos, APM, Osco, LP, Furuya, DEG, Gonçalves, WN, Santana, DC, Teodoro, LPR, da Silva Junior, CA, Capristo-Silva, GF, Li, J, Baio, FHRJC and Agriculture, Ei. 2020. A random forest ranking approach to predict yield in maize with uav-based vegetation spectral indices. 178 105791.
- Ren, H and Zhou, GJEI. 2019. Estimating green biomass ratio with remote sensing in arid grasslands. 98 568-574.
- Sharifi, AJJotSoF and Agriculture. 2020. Remotely sensed vegetation indices for crop nutrition mapping. 100 (14): 5191-5196.

- Sibanda, M, Mutanga, O, Rouget, M and Kumar, L. 2017. Estimating biomass of native grass grown under complex management treatments using worldview-3 spectral derivatives. *Remote Sensing* 9 (1): 55.
- Sibanda, M, Onesimo, M, Dube, T and Mabhaudhi, TJJJoRS. 2021. Quantitative assessment of grassland foliar moisture parameters as an inference on rangeland condition in the mesic rangelands of southern Africa. 42 (4): 1474-1491.
- Stroppiana, D, Migliazzi, M, Chiarabini, V, Crema, A, Musanti, M, Franchino, C and Villa, P.2015. Rice yield estimation using multispectral data from UAV: A preliminary experiment in northern Italy. *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 4664-4667. IEEE,
- Su, W, Huang, J, Liu, D and Zhang, M. 2019. Retrieving corn canopy leaf area index from multitemporal Landsat imagery and terrestrial LiDAR data. *Remote Sensing* 11 (5): 572.
- Sun, Y, Qin, Q, Ren, H, Zhang, T, Chen, SJIToG and Sensing, R. 2019. Red-edge band vegetation indices for leaf area index estimation from sentinel-2/msi imagery. 58 (2): 826-840.
- Tumlisan, GY. 2017. Monitoring growth development and yield estimation of maize using very high-resolution UAV-images in Gronau, Germany. Unpublished thesis, University of Twente,
- Tunca, E, Köksal, ES, Çetin, S, Ekiz, NM, Balde, HJEm and assessment. 2018. Yield and leaf area index estimations for sunflower plants using unmanned aerial vehicle images. 190 (11): 1-12.
- Yang, K, Gong, Y, Fang, S, Duan, B, Yuan, N, Peng, Y, Wu, X and Zhu, RJRS. 2021. Combining Spectral and Texture Features of UAV Images for the Remote Estimation of Rice LAI throughout the Entire Growing Season. 13 (15): 3001.
- Yao, X, Wang, N, Liu, Y, Cheng, T, Tian, Y, Chen, Q and Zhu, YJRS. 2017. Estimation of wheat LAI at middle to high levels using unmanned aerial vehicle narrowband multispectral imagery. 9 (12): 1304.
- Yu, H, Yin, G, Liu, G, Ye, Y, Qu, Y, Xu, B and Verger, A. 2021. Validation of Sentinel-2, MODIS, CGLS, SAF, GLASS and C3S Leaf Area Index Products in Maize Crops. *Remote Sensing* 13 (22): 4529.
- Zhang, H, Lan, Y, Lacey, R, Hoffmann, W and Huang, Y. 2009. Analysis of vegetation indices derived from aerial multispectral and ground hyperspectral data. *International Journal of Agricultural and Biological Engineering* 2 (3): 33-40.

- Zhang, X, Zhao, J, Yang, G, Liu, J, Cao, J, Li, C, Zhao, X and Gai, JJRS. 2019. Establishment of Plot-Yield Prediction Models in Soybean Breeding Programs Using UAV-Based Hyperspectral Remote Sensing. 11 (23): 2752.
- Zheng, H, Cheng, T, Zhou, M, Li, D, Yao, X, Tian, Y, Cao, W and Zhu, YJPA. 2019. Improved estimation of rice aboveground biomass combining textural and spectral analysis of UAV imagery. 20 (3): 611-629.

---

## CHAPTER THREE: ESTIMATING MAIZE YIELD USING UAV-DERIVED MULTI-TEMPORAL DATA IN SMALLHOLDER FARMS OF KWAZULU-NATAL, SOUTH AFRICA

---

### Abstract

Designing relevant food and nutrition security measures and frameworks on smallholding staple crops require robust, efficient and spatially explicit yield monitoring and estimation techniques. This is particularly necessary in developing countries where hunger, malnutrition and food and nutrition insecurity are rife. Smallholder maize croplands are a key source of livelihood in rural communities, especially in regions such as southern Africa where rapid population increases and climate change have compounded the challenge of food and nutrition insecurity, hunger and poverty. Advances in remote sensing techniques have made it feasible to accurately monitor and predict crop yields throughout the growing season. Specifically, UAVs equipped with multispectral high spatial resolution sensors offer spatially explicit near real-time data suitable for maize monitoring and yield estimation at a smallholder farm scale. Therefore, this study sought to predict maize yield in smallholder croplands of southern Africa using UAVs derived multi-temporal remotely sensed datasets in concert with the RF regression ensemble. The value of using the grain biomass, absolute plant biomass, grain biomass as a proportion of the absolute maize plant biomass, VIs and combined spectral data were evaluated. Results showed that UAV derived data could accurately predict yield with  $R^2$  ranging from 0.80 - 0.95, RMSE ranging from 0.03 - 0.94 kg/m<sup>2</sup> and RRMSE ranging from 2.21% - 39.91% based on the spectral datasets combined. Results of this study further revealed that the VT-R1 (56-63 days after emergence) vegetative growth stage was the most optimal stage for the early prediction of maize biomass ( $R^2 = 0.89$ , RMSE = 0.77, RRMSE = 14.47%), grain yield ( $R^2 = 0.85$ , RMSE = 0.1, RRMSE = 5.08%) and proportional yield ( $R^2 = 0.92$ , RMSE = 0.06, RRMSE = 17.56%), with the Normalized Difference Vegetation Index (NDVI), Enhanced Normalized Difference Vegetation Index (ENDVI), Soil Adjusted Vegetation Index (SAVI) and the red edge band being the most optimal prediction variables. The grain yield models produced more accurate results in estimating maize yield when compared to the biomass and proportional yield models. The results demonstrate the value of UAV derived data in predicting maize yield on smallholder farms – a previously challenging task with coarse spatial resolution satellite sensors.

**Keywords:** Smallholder farms, Maize, Remote Sensing, UAV, Vegetation indices, Random Forest algorithm, Biomass, Grain yield, Proportional yield.

### 3.1 Introduction

Agriculture continues to be the mainstay of the economies of most southern African countries, providing over 35% of their gross domestic product, 70-80% of the available employment opportunities and about 30% foreign exchange. Furthermore, the agricultural sector provides livelihoods to over 70% of the country's population in the form of smallholder farming who constitute the majority of food producers (Jin *et al.*, 2019). However, despite the sector's fundamental role in the region's economies and food security, like many other regions of the African continent, abject poverty and deepening hunger continue to stall the development prospects. This is demonstrated by among others the increasing number of people living below the poverty line and malnourished children, trends that could reduce the region's objective of ending poverty and hunger by 2030 as stipulated in the current sustainable development goals. Whereas there are a plethora of challenges accelerating food insecurity, the principal cause is the decline in the production of staple crops. Specifically, decreasing yields of critical food crops such as maize are attributed to, among others, the utility of rudimentary farming practices, the low inputs that characterize conventional farming systems, lack of incentives and appropriate technologies to optimize production, especially on smallholder farms (Tan *et al.*, 2020). Despite their potential, the characteristic nature of smallholder farming systems has so far presented a low predisposition to invest in improved agricultural technologies that can optimize agricultural productivity, hence mitigating food insecurity and poverty.

As aforementioned, smallholder farming is the most prevalent form of maize production in southern African countries. In a recent sub-national census, Jin *et al.* (2019) showed that 50% of food calories in the region were produced on farms of less than 5 ha in size. The annual demand for maize in this region is expected to increase at a rate of 2.4% per annum up to 2025 (Dhau *et al.*, 2018). Hence, it is necessary to explore approaches to maximize maize production on smallholder farms to mitigate poverty and address food and nutritional insecurities. To achieve this frequent monitoring across the growing season is critical in assessing the value of the adopted techniques and approaches to improve smallholder farm productivity (Tunca *et al.*, 2018).

Traditionally, several approaches that include ground observations, surveys and measurements have been adopted in crop monitoring (Mditshwa, 2017). However, these approaches are

limited by their high labour and financial costs and therefore not ideal for continuous and time-efficient crop monitoring (Jégo *et al.*, 2012). Meanwhile, satellite remote sensing has emerged as a better alternative for crop monitoring and yield estimation (Fernandez-Ordoñez and Soria-Ruiz, 2017; Leroux *et al.*, 2019). For instance, Aghighi *et al.* (2018) demonstrated that Landsat 8 multispectral remotely sensed data could predict silage maize yield with an optimal  $R^2$  of 0.87, while Kayad *et al.* (2019) utilized Sentinel-2 multispectral instrument derived VIs to estimate corn grain yield spatial variability with an  $R^2$  of 0.6. Despite the successes associated with these studies, the utilization of such multispectral satellite datasets in crop monitoring and yield estimation in smallholder farms is limited by their relatively coarse spatial and temporal resolutions (Stratoulas *et al.*, 2017). Whereas there are numerous satellite images with high spatial resolutions (e.g. SPOT, Worldview and QuickBird and PlanetScope), they are not cost-effective for monitoring smallholder crops. Moreover, they are often associated with processing complexities which makes them unsuitable for monitoring and estimating maize yield at a farm-scale (Jin *et al.*, 2019; Chivasa *et al.*, 2020).

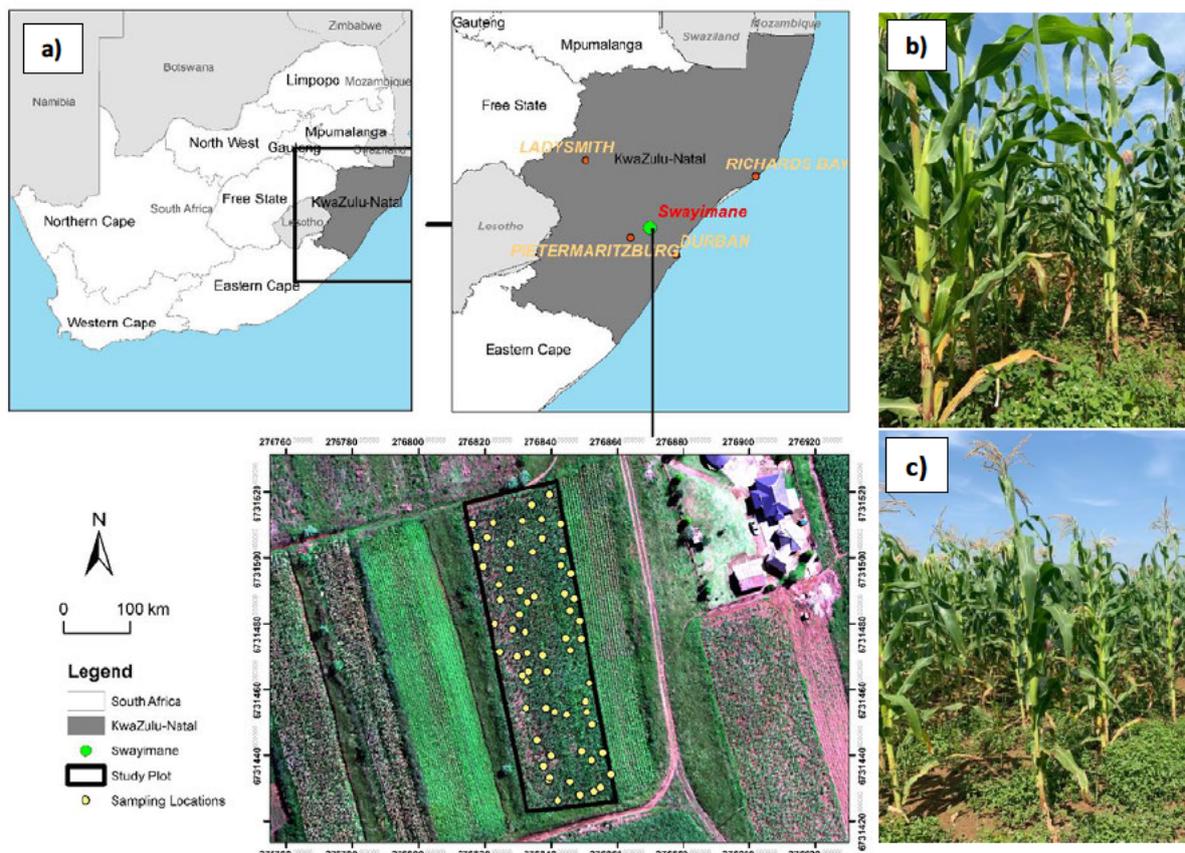
On the other hand, UAVs, also known as drones have emerged as a prospective alternative source of remotely sensed data suitable for mapping and monitoring crop productivity at a farm to field scale (Maes *et al.*, 2018). With advancements in technology, the weight and size of multispectral cameras have been drastically reduced to ease mounting on UAVs for use in precision agriculture (Candiago *et al.*, 2015). UAV systems provide high spatial resolution remotely sensed data at user-defined revisit frequencies and areas of interest, hence time-efficient and cost-effective agricultural applications such as yield modelling (Schut *et al.*, 2018; Ziliani *et al.*, 2018). Furthermore, in estimating maize crop yield using temporal remotely sensed datasets, it is not very clear whether the ultimate plant biomass (inclusive of the grains), the actual grain biomass (excluding the foliage and stem) or the biomass of grain yield as a proportion of ultimate plant biomass exhibits more accurate yield estimates. This has further compounded the challenge in using remotely sensed data to estimate the yield of crops such as maize when compared with crops such as cabbages and spinach (Abdel-Rahman *et al.*, 2014), where biomass is derived from the foliage which in turn directly interacts with the spectral signatures used in yield estimation. In this regard, very few studies have utilized UAV derived data in estimating maize yield at smallholder farms in sub-Saharan Africa. Hence, there is a need to test the utility of multispectral and thermal drone derived remotely sensed datasets to not only estimate maize yield in smallholder farms of the southern African region, but also identify the specific yield variables that optimally facilitate the accurate estimation of yield.

Testing drone derived remotely sensed data in estimating maize yield is important for optimizing agricultural production, a challenge using coarse spatial resolution image data. Therefore, this study aimed to test the utility of UAV derived data in estimating maize yield across the growing season in a smallholder farm. To address this overarching objective, the study sought to; i) predict maize yield using UAV remotely sensed data in conjunction with the RF algorithm and determine the most optimal growth stage for yield prediction, and ii) compare the performance of using the actual grain biomass (excluding the foliage), the ultimate plant biomass (inclusive of the grains and foliage) and the biomass of grain yield as a proportion of ultimate plant biomass in estimating maize yield. To achieve this, the combination of bands and VIs and the RF algorithm regression ensemble were used.

## **3.2 Materials and methods**

### **3.2.1 Study area**

This study was conducted on a smallholder farm located in Swayimane, KwaZulu-Natal, South Africa. The farm is located between 29°31'24''S and 30°41'37'' E (Figure 3.1). The area has a sub-humid climate with an average temperature of 20 °C and average precipitation of 900-1200 mm per annum (Miya *et al.*, 2018). The study was conducted on a 2699.005 m<sup>2</sup> maize field where the maize was sown in November with approximately 160 days of the growing season.



**Figure 3.1:** a) Location of the experimental field plot in Swayimane, KwaZulu-Natal, South Africa, and b) & c) The maize field.

The maize growth stages were divided into two sub-groups, the vegetative growth stages which are the early growth stages covering the V8-V10, V12-V14 and VT-R1 growth stages and the reproductive growth stages covering the R2-R3 and R3-R4 growth stages

\*Refer to Table 2.1 in Chapter Two for more details on the growth stages.

### 3.2.2 Agricultural practices

Maize seeds were sown by hand in February 2021 and weeds were constantly hand removed throughout the growing season. Cow manure, instead of chemical fertilizers were used to optimize soil fertility. The maize crops in the study area were rain-fed.

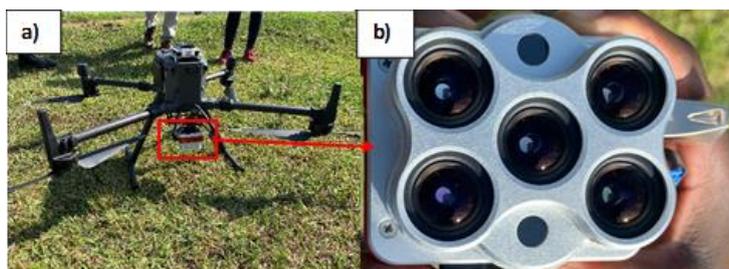
### 3.2.3 Sampling strategy for yield measurements

To optimize the sampling procedure, a polygon of the entire experimental field was generated in Google Earth Pro and imported into ArcGIS 10.6. Subsequently, 63-point locations were generated inside the experimental field plot polygon based on stratified random sampling to determine the sampling points for yield data collection. These points were then uploaded into

a Trimble handheld GPS with a sub-meter accuracy of 30 cm. The GPS was then used to locate and navigate to the sampling points in the field. At each location, a square meter plot was established and maize plants in proximity to each sample point selected for yield estimation. To determine the absolute maize plant biomass, the sample plants (the entire stalk and the cob with the grains) were harvested manually during the reproductive stage R3-R4, which marked the end of the growing season of maize. These were lightly shredded to fit in the brown bags and appropriately labelled. The entire plant biomass was first oven-dried at 60 °C for 48 hours then weighed to determine the entire plant biomass before separating the cobs from the plant. After the separation, the grains were shelled to determine the grain yield biomass. The dry grains were weighed and grain yield was calculated as the weight in kg/m<sup>2</sup>. The dry grains were then divided by the absolute plant biomass to determine the proportional yield. These weights were then recorded on an excel spreadsheet together with coordinates of each sampling point.

### 3.2.4 UAV system and imaging sensor

A DJI Matrice 300 UAV was used in this study for acquiring remotely sensed data (Figure 3.2 a). The device weighs 6.3 kg and has a 30 min flight duration. The DJI M300 flight controller was used for autonomous flights and a DJI Data Link was used to transmit flight parameters to the controller and to remotely control the UAV. A MicaSense Altum multi-spectral camera in conjunction with a DSL 2 was used for UAV spectral imaging of the study site (Figure 3.2 b). MicaSense Altum sensor is equipped with a DSL2 GPS to determine image coordinates during the acquisition period. The device acquires images simultaneous at a 5.2 cm spatial resolution in the blue (475 – 559 nm), green (560 – 667 nm), red (668 – 716 nm), red edge (717 - 839 nm), NIR (840 nm) and thermal (8-14 um) regions of the EMS. Each multispectral image had 2064 x 1544 (rows and columns) pixels and 160 x 120 for the thermal band. Horizontal and vertical field of view angles were 48° and 37°, respectively, for the multispectral bands and 57° and 44° for the thermal band at a flight altitude of 100 m. The sensor can store the images both in its internal memory and on a USB.



**Figure 3.2:** a) DJI Matrice 300 UAV and b) Mica Sense Altum sensor utilized in this study.

### 3.2.5 Image processing and field data collection

#### 3.2.5.1 Image acquisition and pre-processing

A polygon was digitized on Google Earth and exported as a kml file. The polygon was then imported into the controller and used to establish the flight plan, flight altitude and speed parameters for image acquisition. Prior to image acquisition, the sensor was calibrated by acquiring images of the radiometric calibration panel before and after the reconnaissance flight. Five images were acquired at different times across the growing season between February and May of 2021 (days after emergence, 35, 49, 62, 78 and 94). These images, covering the V8 to R3-R4 growth stages were acquired under clear sky conditions between 10:00 AM to 1:00 PM local time, which is the period of the day when changes in solar zenith angle are minimal and the radiation from the sun at maximum. The images from calibration targets were used in calibrating and correcting the reflectance of images. The calibrated images were then exported alongside all the other images into Pix4 D for stitching and radiometric correction. To accurately retrieve georeferenced ortho-mosaicked images of the study plot for the different growth stages, the Altum camera was set to 80% overlap mode using the sensors Wi-Fi. This facilitated the stitching of the images using Pix4D. After transferring the images into Pix4D fields, they were calibrated, radiometrically corrected and stitched to create ortho-images for the entire study site. Geometric correction was done in QGIS 3.12.3 using field collected ground control points.

#### 3.2.5.2 Calculation of VIs

The UAV derived image bands were used to compute VIs and both spectral bands and indices used to predict maize yield. The VIs selected in and utilized in this study are summarized in Table 3.1. These VIs were chosen based on their performance in previous studies (Bala and Islam, 2009; Mditshwa, 2017; Tumlisan, 2017). The indices directly relate to the chlorophyll content of the plant, which indirectly relates to yield.

**Table 3.1:** VIs used in this study.

Vegetation Index	Abbreviation	Equation	Reference
Normalized Vegetation Index	NDVI	$\frac{NIR-RED}{NIR+RED}$	(Wahab <i>et al.</i> , 2018)
Enhanced Vegetation Index	ENDVI	$\frac{(NIR + GREEN) - (2 \times BLUE)}{(NIR + GREEN) + (2 \times BLUE)}$	(Zhang <i>et al.</i> , 2009)

Vegetation Index	Abbreviation	Equation	Reference
Soil Adjusted Vegetation Index	SAVI	$\frac{(NIR - RED)}{(NIR + RED + L)} \times (1 + L)$	(Ngie and Ahmed, 2018)
Optimized Adjusted Vegetation Index	OSAVI	$\frac{(NIR - RED)}{(NIR + RED + 0.16)} \times (1 + 0.16)$	(Liu <i>et al.</i> , 2019)
Simple Ratio	SR	$\frac{RED}{NIR}$	(Kanning <i>et al.</i> , 2018)

### 3.2.6 Data Analysis

#### 3.2.6.1 Correlation between grain yield and the entire plant biomass

A correlation between the grain and the biomass data was determined to evaluate whether there was a link between the accumulated biomass and the actual yield at the R3-R4 growth stage. A Pearson product-moment correlation test was conducted in this regard following a data normality test, which indicated that the data did not significantly deviate from the normal distribution.

#### 3.2.6.2 Maize yield prediction

To test the relationship between biomass, grain yield and proportional yield determined at the R3-R4 stage, the collected 63 yield samples and UAV data (i.e. combination of bands and VIs data) were divided into training (70%) and test (30%) datasets to derive models using the RF algorithm in R statistical package. The RF algorithm was adopted in this study as it is a non-parametric statistical technique that uses a bagging-based approach to build an ensemble of regression trees while ranking important variables that produce an independent measure of prediction error Prasad *et al.* (2006). In R, the *ntree* and *mtry* parameters were optimized 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 600 for *ntree* and five for *mtry*. In addition, the most optimal growth stage at which the combination of bands and VIs were highly correlated to the yield was assessed to determine the most suitable period to predict maize yield before harvest.

#### 3.2.6.3 Model validation and accuracy assessment

Test data (30%) was used to evaluate model performance of the derived models. Performance indicators such as  $R^2$ , RMSE) and RRMSE were determined and used to assess the accuracy of each model.

$$RMSE = \sqrt{\frac{SSE^2}{n}} \quad (1)$$

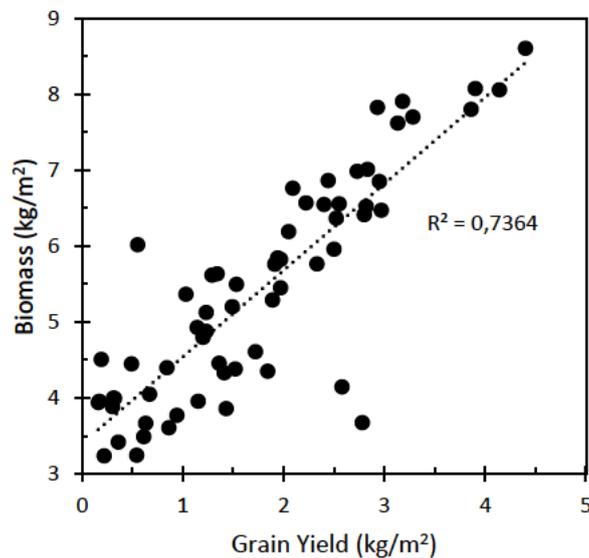
$$RRMSE \% = \frac{RMSE}{MEAN} \times 100 \quad (2)$$

Where the SSE notation symbolizes the sum of errors of (measure yield - predicted yield) and n symbolizes the number of predictor values in the model construction. The model that yielded a low RMSE and high  $R^2$  at a stage with adequate time before harvest for intervention was set aside and used to predict yield. The most optimal model was then used to create a yield map for the study site in ArcMap 10.6.

### 3.3 Results

#### 3.3.1 Descriptive statistics

The highest maize biomass, grain yield and proportional yield was 8.61 kg/m<sup>2</sup>, 4.4 kg m<sup>2</sup> and 0.76 kg/m<sup>2</sup> and the lowest were 3.24 kg/m<sup>2</sup>, 0.16kg/m<sup>2</sup>, 0.04 kg/m<sup>2</sup>, respectively. There was considerable variation in maize yields samples in the study. The standard deviation was 1.45, 1.08 and 0.15 for biomass, grain yield and proportional yield, respectively. Furthermore, a strong ( $R^2$  of 0.74) positive correlation between the grain yield samples and the overall biomass of the maize plants was attained. Figure 3.3 shows the relationship between yield and biomass.



**Figure 3.3:** Correlation between the grain yield and biomass at the R3-R4 stage using all samples.

### 3.3.2 Derived maize yield prediction models and their accuracies

Figure 3.4 illustrates the model accuracies obtained in predicting the biomass, grain yield and proportional yield based on the RF algorithm. The accuracies of the prediction models varied greatly across the maize growing season. For example, when estimating the absolute plant biomass, the V8-V10 growth stage yielded the poorest model, with an  $R^2$  of 0.80 and RMSE of 0.94 kg/m<sup>2</sup>. The prediction of biomass improved in the V12-V14 growth stage model ( $R^2 = 0.85$  and RMSE = 0.72 kg/m<sup>2</sup>). Similarly, the VT-R1 and R2-R3 models predicted biomass at an improved  $R^2 = 0.89$ , RMSE = 0.77 kg/m<sup>2</sup> and  $R^2 = 0.89$ , RMSE = 0.88 kg/m<sup>2</sup>, respectively. The optimal model in estimating biomass was derived from the R3-R4 model, with an  $R^2$  of 0.91 and RMSE of 0.61 kg/m<sup>2</sup> (Figure 3.4 e). The most optimal variables in estimating biomass were ENDVI, the red edge band, NIR and NDVI, in order of importance (Figure 3.5 e).

Similarly, the V8-V10 model demonstrated the lowest prediction accuracy in estimating the grain yield ( $R^2 = 0.85$  and RMSE = 0.6 kg/m<sup>2</sup>). This was followed by V12-V14 and VT-R1 with an  $R^2$  of 0.89, RMSE of 0.12 kg/m<sup>2</sup> and  $R^2$  of 0.85, RMSE of 0.1 kg/m<sup>2</sup>, respectively. The prediction accuracy increased significantly with the R2-R3 model ( $R^2 = 0.95$  and RMSE = 0.09 kg/m<sup>2</sup>). The R3-R4 model optimally predicted the grain yield with the lowest RMSE = 0.03 kg/m<sup>2</sup> and  $R^2 = 0.92$  (Figure 3.4 e). The variables that had the highest influence in the grain yield model were ENDVI, NIR, NDVI and the red edge band in ascending order of importance (Figure 3.5 e).

When predicting the proportional yield, the V12-V14 model produced the lowest prediction accuracy with an  $R^2$  of 0.92 and RMSE of 0.11 kg/m<sup>2</sup>. The prediction of proportional yield improved in the V8-V10, VT-R1 and R2-R3 models with an  $R^2$  of 0.91, RMSE of 0.09 kg/m<sup>2</sup>;  $R^2$  of 0.92, RMSE of 0.06 kg/m<sup>2</sup> and  $R^2 = 0.92$ , RMSE = 0.07 kg/m<sup>2</sup>. The optimal model for estimating proportional yield produced an  $R^2$  of 0.95 and RMSE = 0.07 kg/m<sup>2</sup> (Figure 3.4 e). The most suitable predictor variables included NDVI, the green, NIR and red edge bands (Figure 3.5 e).

In comparing the performance of the maize biomass, grain yield and proportional yield variables in predicting yield across all growth stages, the results varied greatly (Figure 3.4). For example, when estimating yield at the V8-V10 growth stage, the proportional yield model exhibited the poorest prediction accuracy with an RRMSE of 30.43% followed by the grain yield model with an RRMSE of 27.99%. Comparatively, the most optimal model in estimating yield during the V8-V10 growth stage was the biomass model with an RRMSE of 15.42%

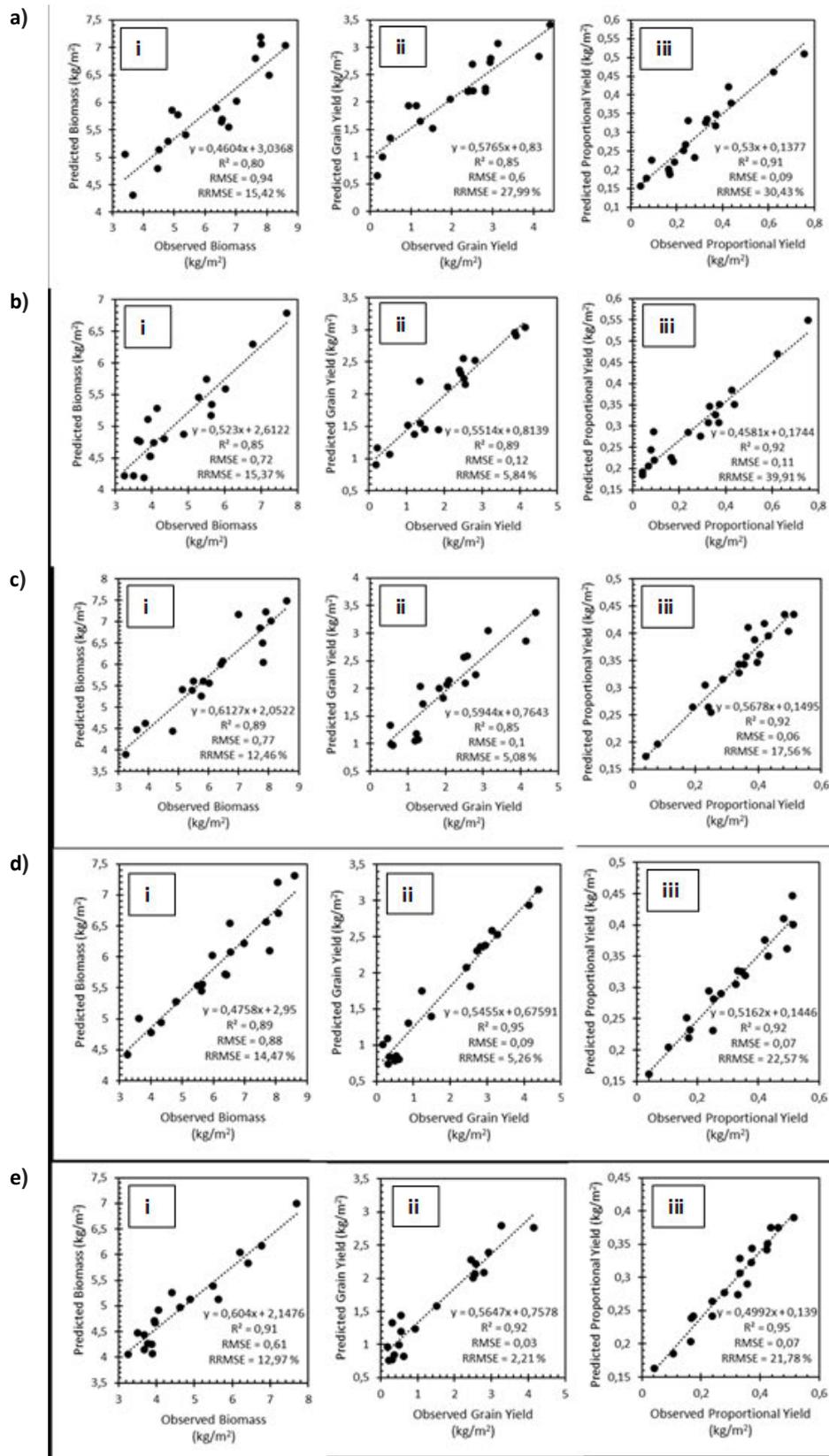
(Figure 3.4 a). The most important variables include the red and blue bands, SAVI and OSAVI (Figure 3.5 a).

Similarly, the proportional yield model yielded the poorest model with an RRMSE of 39.91% followed by the biomass model with an RRMSE of 15.37% at the V12-V14 growth stage. The grain yield model optimally predicted maize yield with the lowest RRMSE = 5.44% at the V12-14 (Figure 3.4 b). The most optimal variables for this prediction were the green, red edge, red and blue bands (Figure 3.5 b).

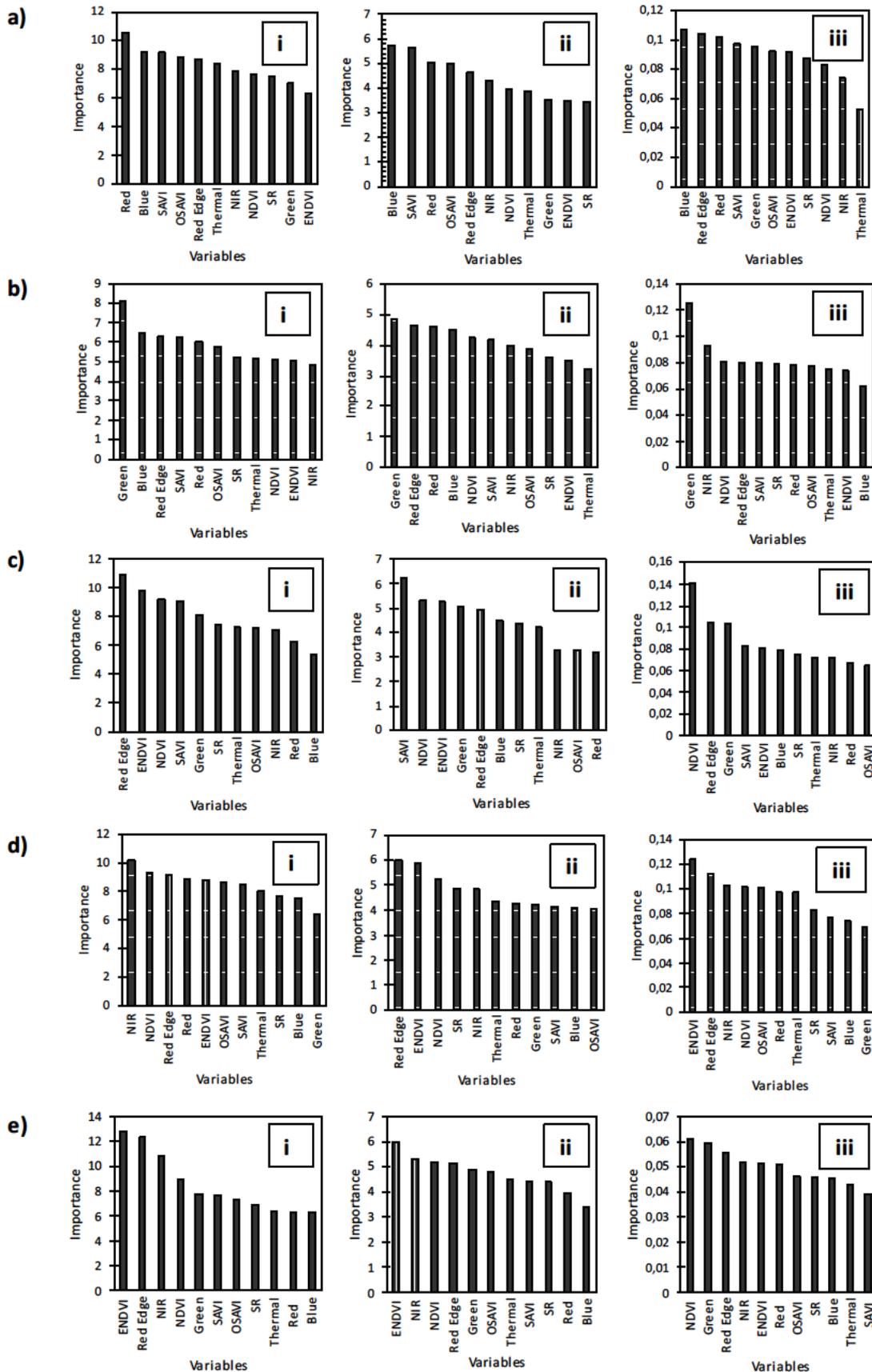
In predicting yield at the VT-R1 growth stage, the proportional yield model produced the highest RRMSE of 17.56%. The prediction accuracy improved with the biomass and grain yield models (RRMSE = 12.56% and 5.08%, correspondingly) (Figure 3.4 c). The variables that had the highest influence in the grain yield model were SAVI, NDVI, ENDVI and the green band, in order of importance (Figure 3.5 c).

When predicting yield in the R2-R3 growth stage, the highest RRMSE of 22.57% was obtained by the proportional yield model. The biomass model improved the prediction by a magnitude of 8.1%, i.e., RRMSE = 14.47%. Similarly, the grain yield model was the optimal model in estimating yield at the R2-R3 growth stage (Figure 3.4 d). The red-edge band, NDVI, ENDVI and SR were the most influential variables for this model (Figure 3.5 d).

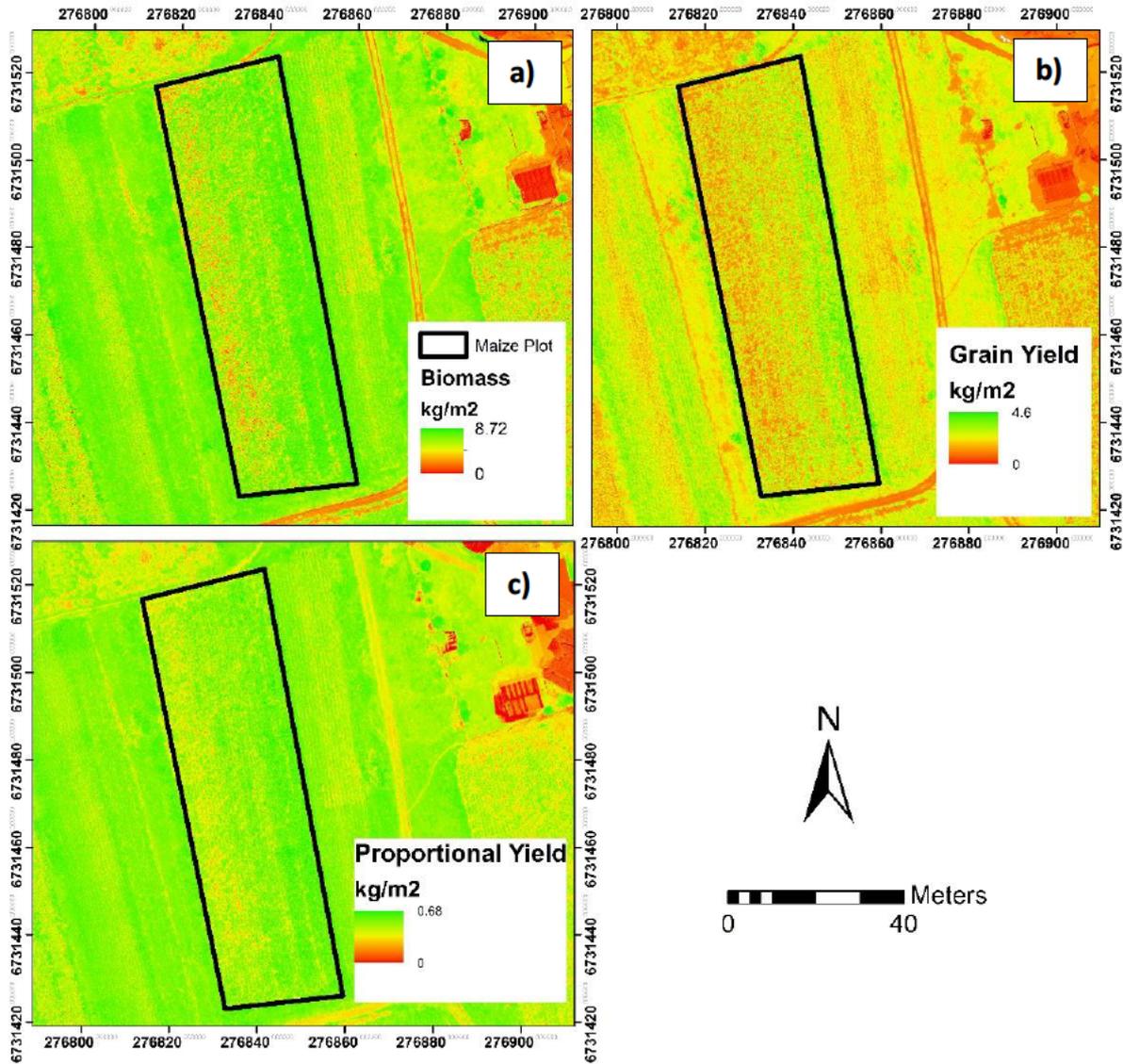
For the R3-R4 growth stage, the proportional yield exhibited the lowest prediction accuracy with an RRMSE of 21.78%. The prediction of yield improved significantly with the biomass model (RRMSE = 12.97%) and even greater with the grain yield model (RRMSE = 2.21%) (Figure 3.4 e). The most influential variables for this prediction were NDVI, NIR, ENDVI and the red edge band (Figure 3.5 e).



**Figure 3.4:** Relationship between observed and predicted i) biomass, ii) grain yield and iii) proportional yield based on the combination of bands and VIs using the RF Model for a) V8-V10 b) V12-V14 c) VT-R1 d) R2-R3 and e) R3-R4 maize growth stages.



**Figure 3.5:** Variable importance of the i) biomass ii) grain yield and iii) proportional yield models for a) V8-V10 b) V12-V14 c) VT-R1 d) R2-R3 and e) R3-R4 maize growth stages.



**Figure 3.6:** Spatial distribution of modelled maize **a)** biomass **b)** grain yield and **c)** proportional yield based on the most optimal RF models.

### 3.4. Discussion

This study sought to test the capability of UAV derived data in estimating maize yield across the growing season based on the Altum sensor mounted on the DJI Matrice 300 UAV. Specifically, this study sought to predict maize yield using UAV images and the RF algorithm in smallholder farms.

#### 3.4.1 Maize yield prediction models

The results of this study show that the early growth stages of the crop yielded lower overall accuracies for the biomass, grain yield and proportional yield followed by some improvements in the later stages of growth (Figure 3.4). Specifically, the V8-V10, V12-V14 and VT-R1

growth stages had lower overall accuracies when compared to the R2-R3 and R3-R4 growth stages. Several studies (Guindin-Garcia, 2010; Son *et al.*, 2013; Al-Gaadi *et al.*, 2016; Chivasa *et al.*, 2017) have noted that in the early stages of crop development, vegetation reflectance is affected by the soil background, which explains the low performance of UAV data in predicting maize biomass, grain yield and proportional yield at the early (vegetative growth) stages of this study. At this stage, the maize leaves are not fully grown, exposing the surrounding soil, which then interferes with the plant's reflectance as the sensor also picks up the soil reflectance (Zhang *et al.*, 2019a).

In contrast, the later growth stages of the crop yielded higher overall accuracies. Specifically, the R2-R3 and R3-R4 growth stages had higher accuracies when compared to the V8-V10, V12-V14 and VT-R1 stages. The high performance of the UAV data in predicting maize yield at the R2-R3 and R3-R4 stages of the growth cycle can be explained by existing literature which has reported significantly high accuracies in the prediction of maize yield at the late (reproductive) stages of the crop (Guindin-Garcia, 2010; Mditshwa, 2017). Literature notes that at this stage, the maize leaves have grown to mid-density covering the surrounding soil and therefore crop reflectance is not impacted by the soil background (Mkhabela *et al.*, 2005; Tumliyan, 2017; Tunca *et al.*, 2018). When plants have grown to mid-density, there is canopy coverage, meaning the biomass production has reached its most mature stage, making it possible to remotely sense vegetation without any interferences from the ground such as soil (Kayad *et al.*, 2019). At this stage, when biomass production has reached its peak, it is most closely related to yield, which explains why the model accuracies for these stages were higher than of the earlier stages of growth (Ngie and Ahmed, 2018; Li *et al.*, 2020).

Regarding model variable importance, SAVI, OSAVI, the blue and red bands were more important in the prediction at the early stages than in the late stages of the crops phenological cycle. The value of SAVI and OSAVI can be attributed to their ability to suppress soil background, hence better prediction at minimal leaf coverage resulting and soil exposure (Ren and Zhou, 2019; Zhang *et al.*, 2019b). The importance of the blue and red bands for these models can be explained by soil being more dominant than vegetation in the early stages of the crop resulting in high reflectance in the blue and red region of the EMS (Ngie and Ahmed, 2018).

Comparatively, NDVI, ENDVI, the green, red, red edge and NIR bands were of significant importance in the prediction models at the R2-R3 and R3-R4 crop growth stages. The

importance of NDVI and ENDVI in these models could be as a result of the fact that when the reflectance measurements for the R2-R3 and R3-R4 growth stages were taken, a saturation of the plant canopy had not occurred, the plant canopy had only accumulated to mid-density and there is a good relationship between NDVI and ENDVI and biomass and yield at mid-density canopies, which characterize the R2-R3 and R3-R4 maize growth stages (Awad, 2019). The importance of the green, red, red edge and NIR bands in the models of the R2-R3 and R3-R4 growth stages for this study can be attributed to the fact that there was a dominance of vegetation which reflects strongly in the green and NIR regions of the EMS and highly absorbs in the red and red edge regions of the EMS (Khaliq *et al.*, 2019; Marcial-Pablo *et al.*, 2019).

### **3.4.2 Determining the most optimal growth stages and variables for yield prediction**

The best-fit model for predicting maize biomass and grain yield was obtained at the R3-R4 growth stage, with ENDVI and the red edge band being the most important variables for the prediction of maize biomass and ENDVI and NDVI being the most important for the prediction of grain yield. The influence of the ENDVI, NDVI and the red edge in the prediction at this stage could be explained by the good relationship between the two indices and biomass and yield at mid-density canopies before saturation (Mutanga *et al.*, 2012; Tan *et al.*, 2020). On the other hand, previous studies note that the red edge section of the EMS is related to chlorophyll and biomass, which directly relates to yield (Dube *et al.*, 2017; Sibanda *et al.*, 2017b). Generally, mid-density canopies are characterized by a high amount of biomass, which is associated with high chlorophyll content and carbon assimilation which are sensitive to the red edge section of the EMS (Sibanda *et al.*, 2021). Furthermore, a mid-density canopy like in the R3-R4 growth stage results in high ENDVI and NDVI values as well as a strong absorption in the red edge region of the EMS, hence their strong influence in the prediction of biomass and grain yield when compared to the other variables (Raeva *et al.*, 2019). In addition, the best-fit model for predicting maize proportional yield was obtained in the VT-R1 growth stage with NDVI and SAVI being the most important variables for the prediction of proportional yield. The significance of NDVI and SAVI in the prediction model of maize of proportional yield at this stage can be attributed to the fact that this is the middle stage where the canopy has not grown to mid-density resulting in significant soil exposure (Mditshwa, 2017). This then results in SAVI being important in suppressing the soil background effect and allows NDVI to perform well as it has a good relationship with the biomass and yield at this stages' canopy level because the canopy has not yet reached saturation, as canopy saturation hinders the performance of NDVI (Mutanga *et al.*, 2012).

Regarding the best-fit model for maize biomass and grain yield which was obtained at the R3-R4 reproductive development stage and proportional yield at the VT-R1 vegetative development stage (78 and 62 days after emergence) of the growth cycle. Using the R3-R4 growth stage for biomass and grain yield prediction could be late for the adoption of any effective measure before harvest. A significant relationship was found at the VT-R1 (62 days after emergence) growth stage for biomass as well as grain yield. Based on our findings, this is the optimal stage at which maize yield could be predicted before harvesting. The most significant variables for the optimal biomass, grain yield and proportional yield prediction models were the red edge band and ENDVI, SAVI and NDVI, ENDVI and the red edge band respectively. Furthermore, the grain yield produced higher prediction accuracies in estimating maize yield for most of the crop's growth stages (V12-V14, VT-R1, R2-R3 and R3-R4) when compared to the absolute plant biomass and the biomass of grain yield as a proportion of ultimate plant biomass. The absolute plant biomass was only optimal in the V8-V10 growth stage and the proportional yield produced the poorest yield prediction accuracies in all of the growth stages. Therefore, the grain yield proved to be the most optimal in estimating maize yield.

The obtained validation accuracies and their variable importance match those from previous findings, where the prediction was conducted on maize using different space-borne sensors. For example, Battude *et al.* (2016) estimated the biomass and maize yield over a large area using Sentinel-2 data and concluded that remotely sensed data can accurately be used to predict the biomass and yield throughout the phenological cycle, with prediction accuracies ranging from 0.8-0.9. Ngie and Ahmed (2018) successfully estimated maize grain yield using SPOT 5 data in the Free State province of South Africa, where prediction models with accuracies of 0.92 and 0.9 were achieved using SAVI and NDVI. Mditshwa (2017) used GIS and remote sensing to estimate maize grain yield from the different growth stages and concluded that NDVI and SAVI are good yield predictors. Unlike the current study which was conducted on a small area using high spatial and temporal resolutions datasets, the above-mentioned studies were conducted at large spatial extents using Sentinel-2 and Landsat-8. In this study, the adoption of a sensor mounted on a UAV has demonstrated its value in predicting maize yield in a smallholder farm.

### **3.5 Conclusion**

This study aimed to predict maize yield (biomass, grain yield and proportional yield) across the growing season in a smallholder farm based on UAV remotely sensed data. The following conclusions were drawn:

- UAV derived data optimally predicted maize yield during the R3-R4 growth stage using ENDVI, NDVI and the red edge band;
- The VT-R1 stage was the most optimal stage for the early prediction of maize yields using SAVI, NDVI, ENDVI and the red edge band;
- The grain yield models produced higher accuracies in estimating maize yield when compared to the absolute plant biomass and the biomass of grain yield as a proportion of absolute plant biomass models.

The characterised variations in field productivity can assist farmers and decision-makers in identifying low yield areas within the field so as to adjust their management practices to maximize farm productivity. These findings highlight the utility of UAV systems in optimizing agricultural production through precision farming on smallholder farms, necessary for poverty alleviation and food and nutritional security.

### 3.6 References

- Abdel-Rahman, EM, Mutanga, O, Odindi, J, Adam, E, Odindo, A and Ismail, R. 2014. A comparison of partial least squares (PLS) and sparse PLS regressions for predicting yield of Swiss chard grown under different irrigation water sources using hyperspectral data. *Computers and Electronics in Agriculture* 106 11-19.
- Aghighi, H, Azadbakht, M, Ashourloo, D, Shahrabi, HS, Radiom, SJIJoSTiAEO and Sensing, R. 2018. Machine learning regression techniques for the silage maize yield prediction using time-series images of landsat 8 OLI. 11 (12): 4563-4577.
- Al-Gaadi, KA, Hassaballa, AA, Tola, E, Kayad, AG, Madugundu, R, Alblewi, B and Assiri, F. 2016. Prediction of potato crop yield using precision agriculture techniques. *PloS one* 11 (9):
- Awad, MM. 2019. Toward precision in crop yield estimation using remote sensing and optimization techniques. *Agriculture* 9 (3): 54.
- Bala, S and Islam, AJIJoRS. 2009. Correlation between potato yield and MODIS-derived vegetation indices. 30 (10): 2491-2507.
- Battude, M, Al Bitar, A, Morin, D, Cros, J, Huc, M, Sicre, CM, Le Dantec, V and Demarez, VJRSoE. 2016. Estimating maize biomass and yield over large areas using high spatial and temporal resolution Sentinel-2 like remote sensing data. 184 668-681.
- Candiago, S, Remondino, F, De Giglio, M, Dubbini, M and Gattelli, M. 2015. Evaluating multispectral images and vegetation indices for precision farming applications from UAV images. *Remote sensing* 7 (4): 4026-4047.
- Chivasa, W, Mutanga, O and Biradar, CJIJoRS. 2017. Application of remote sensing in estimating maize grain yield in heterogeneous African agricultural landscapes: a review. 38 (23): 6816-6845.
- Chivasa, W, Mutanga, O and Biradar, CJRS. 2020. UAV-based multispectral phenotyping for disease resistance to accelerate crop improvement under changing climate conditions. 12 (15): 2445.
- Dhau, I, Adam, E, Mutanga, O, Ayisi, K, Abdel-Rahman, EM, Odindi, J and Masocha, MJGI. 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.
- Dube, T, Mutanga, O, Sibanda, M, Shoko, C, Chemura, AJP and Chemistry of the Earth, PABC. 2017. Evaluating the influence of the Red Edge band from RapidEye sensor in

- quantifying leaf area index for hydrological applications specifically focussing on plant canopy interception. 100 73-80.
- Fernandez-Ordoñez, YM and Soria-Ruiz, J. 2017. Maize crop yield estimation with remote sensing and empirical models. *2017 IEEE international geoscience and remote sensing symposium (IGARSS)*, 3035-3038. IEEE,
- Guindin-Garcia, N. 2010. Estimating maize grain yield from crop biophysical parameters using remote sensing.
- Jégo, G, Pattey, E and Liu, JJFCR. 2012. Using Leaf Area Index, retrieved from optical imagery, in the STICS crop model for predicting yield and biomass of field crops. 131 63-74.
- Jin, Z, Azzari, G, You, C, Di Tommaso, S, Aston, S, Burke, M and Lobell, DBJRSoE. 2019. Smallholder maize area and yield mapping at national scales with Google Earth Engine. 228 115-128.
- Kanning, M, Kühling, I, Trautz, D and Jarmer, TJRS. 2018. High-resolution UAV-based hyperspectral imagery for LAI and chlorophyll estimations from wheat for yield prediction. 10 (12): 2000.
- Kayad, A, Sozzi, M, Gatto, S, Marinello, F and Pirotti, FJRS. 2019. Monitoring within-field variability of corn yield using Sentinel-2 and machine learning techniques. 11 (23): 2873.
- Khaliq, A, Comba, L, Biglia, A, Ricauda Aimonino, D, Chiaberge, M and Gay, PJRS. 2019. Comparison of satellite and UAV-based multispectral imagery for vineyard variability assessment. 11 (4): 436.
- Leroux, L, Castets, M, Baron, C, Escorihuela, M-J, Bégué, A and Seen, DLJEJoA. 2019. Maize yield estimation in West Africa from crop process-induced combinations of multi-domain remote sensing indices. 108 11-26.
- Li, B, Xu, X, Zhang, L, Han, J, Bian, C, Li, G, Liu, J, Jin, LJJoP and Sensing, R. 2020. Above-ground biomass estimation and yield prediction in potato by using UAV-based RGB and hyperspectral imaging. 162 161-172.
- Liu, Y, Liu, S, Li, J, Guo, X, Wang, S and Lu, J. 2019. Estimating biomass of winter oilseed rape using vegetation indices and texture metrics derived from UAV multispectral images. *Computers and Electronics in Agriculture* 166 105026.
- Maes, WH, Huete, AR, Avino, M, Boer, MM, Dehaan, R, Pendall, E, Griebel, A and Steppe, KJRS. 2018. Can UAV-based infrared thermography be used to study plant-parasite interactions between mistletoe and eucalypt trees? 10 (12): 2062.

- Marcial-Pablo, MdJ, Gonzalez-Sanchez, A, Jimenez-Jimenez, SI, Ontiveros-Capurata, RE and Ojeda-Bustamante, WJ. 2019. Estimation of vegetation fraction using RGB and multispectral images from UAV. 40 (2): 420-438.
- Mditshwa, S. 2017. Estimating Maize Grain Yield from Crop Growth Stages Using Remote Sensing and GIS in the Free State Province, South Africa. Unpublished thesis, University of Fort Hare,
- Miya, SP, Modi, AT, Tesfay, SZ, Mabhaudhi, TJSaJoP and Soil. 2018. Maize grain soluble sugar and protein contents in response to simulated hail damage. 35 (5): 377-383.
- Mkhabela, MS, Mkhabela, MS, Mashinini, NNJA and Meteorology, F. 2005. Early maize yield forecasting in the four agro-ecological regions of Swaziland using NDVI data derived from NOAA's-AVHRR. 129 (1-2): 1-9.
- Mutanga, O, Adam, E and Cho, MA. 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.
- Ngie, A and Ahmed, FJSaJoG. 2018. Estimation of Maize grain yield using multispectral satellite data sets (SPOT 5) and the random forest algorithm. 7 (1): 11-30.
- Prasad, AM, Iverson, LR and Liaw, A. 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems* 9 (2): 181-199.
- Raeva, PL, Šedina, J and Dlesk, A. 2019. Monitoring of crop fields using multispectral and thermal imagery from UAV. *European Journal of Remote Sensing* 52 (sup1): 192-201.
- Ren, H and Zhou, GJEI. 2019. Estimating green biomass ratio with remote sensing in arid grasslands. 98 568-574.
- Schut, AG, Traore, PCS, Blaes, X and Rolf, AJFCR. 2018. Assessing yield and fertilizer response in heterogeneous smallholder fields with UAVs and satellites. 221 98-107.
- Sibanda, M, Mutanga, O, Rouget, M and Kumar, LJRS. 2017. Estimating biomass of native grass grown under complex management treatments using worldview-3 spectral derivatives. 9 (1): 55.
- Sibanda, M, Onesimo, M, Dube, T and Mabhaudhi, TJIJoRS. 2021. Quantitative assessment of grassland foliar moisture parameters as an inference on rangeland condition in the mesic rangelands of southern Africa. 42 (4): 1474-1491.

- Son, N, Chen, C, Chen, C, Chang, L, Duc, H and Nguyen, LJIJoRS. 2013. Prediction of rice crop yield using MODIS EVI– LAI data in the Mekong Delta, Vietnam. 34 (20): 7275-7292.
- Stratoulas, D, Tolpekin, V, De By, RA, Zurita-Milla, R, Retsios, V, Bijker, W, Hasan, MA and Vermote, EJRS. 2017. A workflow for automated satellite image processing: From raw VHSR data to object-based spectral information for smallholder agriculture. 9 (10): 1048.
- Tan, C-W, Zhang, P-P, Zhou, X-X, Wang, Z-X, Xu, Z-Q, Mao, W, Li, W-X, Huo, Z-Y, Guo, W-S and Yun, FJSr. 2020. Quantitative monitoring of leaf area index in wheat of different plant types by integrating NDVI and Beer-Lambert law. 10 (1): 1-10.
- Tumlisan, GY. 2017. Monitoring growth development and yield estimation of maize using very high-resolution UAV-images in Gronau, Germany. Unpublished thesis, University of Twente,
- Tunca, E, Köksal, ES, Çetin, S, Ekiz, NM, Balde, HJEm and assessment. 2018. Yield and leaf area index estimations for sunflower plants using unmanned aerial vehicle images. 190 (11): 1-12.
- Wahab, I, Hall, O and Jirstrom, MJD. 2018. Remote sensing of yields: Application of uav imagery-derived ndvi for estimating maize vigor and yields in complex farming systems in sub-saharan africa. 2 (3): 28.
- Zhang, H, Lan, Y, Lacey, R, Hoffmann, W and Huang, Y. 2009. Analysis of vegetation indices derived from aerial multispectral and ground hyperspectral data. *International Journal of Agricultural and Biological Engineering* 2 (3): 33-40.
- Zhang, L, Niu, Y, Zhang, H, Han, W, Li, G, Tang, J and Peng, XJFips. 2019a. Maize canopy temperature extracted from UAV thermal and RGB imagery and its application in water stress monitoring. 10 1270.
- Zhang, L, Zhang, H, Niu, Y and Han, WJRS. 2019b. Mapping maize water stress based on UAV multispectral remote sensing. 11 (6): 605.
- Ziliani, MG, Parkes, SD, Hoteit, I and McCabe, MFJRS. 2018. Intra-season crop height variability at commercial farm scales using a fixed-wing UAV. 10 (12): 2007.

---

## CHAPTER FOUR: ASSESSING THE UTILITY OF UNMANNED AERIAL VEHICLE REMOTELY SENSED DATA FOR ESTIMATING MAIZE LEAF AREA INDEX (LAI) AND YIELD ACROSS THE GROWING SEASON: A SYNTHESIS

---

### 4.1 Introduction

In southern Africa, the number of people living below the poverty datum line, as well as the number of malnourished children, are exponentially increasing which could potentially reduce the prospects of the region to meet the sustainable development goals of ending poverty and hunger by 2030. The principal cause that is accelerating food insecurities is the decline in the production of staple crops such as maize (*Zea mays* L), especially in smallholder agricultural lands. Smallholder farms have so far presented a low predisposition to invest in advanced agricultural technologies that will optimize agricultural productivity while addressing food insecurity and poverty challenges. To close the gap of poverty while addressing food and nutritional insecurities there is an urgent need to maximize the production of maize in smallholder farms. This can be achieved through the reduction of input costs and the increment of maize yields as well as optimizing agricultural productivity. For this optimization to be successful, crops should be monitored across their growing season and crop growth management techniques implemented on time and at required quantities. However, spatial data sources suitable for estimating crop yield and productivity elements in smallholder agricultural lands have been scanty due to their relatively small size (< 5 ha), fragmentation and crop diversity. The advent of UAVs has presented better prospects for monitoring crop productivity at a field scale. It is in this regard that this study sought to assess the utility of UAV derived data in estimating maize LAI and yield across the growing season in smallholder farms of KwaZulu-Natal based on the RF algorithm. The specific objectives of this study were:

- To estimate LAI of maize crops using UAV derived VIs and RF regression across the growing season in smallholder croplands;
- To estimate maize yield across the phenological cycle based on UAV derived data in conjunction with RF regression in smallholder croplands.

#### **4.2 Estimating Maize Leaf Area Index using UAV-derived multi-spectral remotely sensed data in smallholder farms.**

In this objective, this study sought to test the utility of UAV derived VIs in estimating maize LAI across the growing season in a smallholder farm. Based on the RF regression algorithm, the results of this study showed that UAV derived VIs optimally estimated maize LAI ( $R^2 = 0.91, 0.93, 0.91, 0.89$  and  $0.91$ ;  $RMSE = 0.15, 0.17, 0.65, 0.19$  and  $0.32 \text{ m}^2/\text{m}^2$  and  $RRMSE = 8.13\%, 8.97\%, 19.61\%, 10.78\%$  and  $15, 22\%$  for the five growth stages, respectively). Specifically, the V8-V10 stage exhibited the most accurate model in estimating LAI to an  $RMSE$  of  $0.15$ ,  $R^2$  of  $0.91$  and  $RRMSE$  of  $8.13\%$  based on the  $ndviG\&B$  and  $ndviB\&T$  spectral variables. These results illustrate the utility of UAV derived VIs in estimating maize LAI in a smallholder farm with spectral variables derived from the blue, green and thermal regions of the EMS were the most optimal variables for these predictions. This offers the explicit information needed for optimizing agricultural production in smallholder farms in data-scarce regions such as sub-Saharan Africa.

#### **4.3 Estimating Maize yield using UAV-derived multi-temporal data in smallholder farms of KwaZulu-Natal, South Africa**

For the second objective, UAV remotely sensed data in concert with RF regression ensemble were employed in predicting maize yield in a smallholder farm located in the KwaZulu-Natal province of South Africa. To answer this objective, the biomass of maize grain, the biomass of grain as a proportion of the absolute biomass, as well as absolute biomass of the maize (plant and cob) were compared based on the RF regression. UAV images were acquired across the maize growing season and VIs were computed, combined with the bands and used in this element of the study.

The findings of this study showed that yield could be optimally estimated during the R3-R4 growth stage to an  $RMSE$  of  $0.61$  and  $0.03 \text{ kg}/\text{m}^2$  and  $R^2$  of  $0.91$  and  $0.92$  for the biomass and grain yield variables, respectively, based on the  $NDVI$ ,  $ENDVI$ ,  $SAVI$  and the red edge band spectral variables. However, using the R3-R4 growth stage for biomass and grain yield prediction could be late for the adoption of any effective measure before harvest and recommended the use of the VT-R1 growth stage for the early prediction of yield as a significant relationship was found at this stage. The proportional yield was optimally predicted during the VT-R1 growth stage to an  $RMSE$  of  $0.06 \text{ kg}/\text{m}^2$  and  $R^2$  of  $0.92$  based on  $NDVI$  and the green band. The findings of this study also indicated that grain yield was optimally

predicted in relation to biomass and proportional yield as a variable for characterizing the yield of maize in smallholder croplands.

#### **4.4 Conclusion**

This study sought to assess the utility of UAV derived data in concert with the RF algorithm in predicting maize LAI and yield in smallholder farms of KwaZulu-Natal, South Africa. Based on the findings deduced from each chapter and or objective, this study concluded that:

- i. Maize LAI can be optimally estimated at the V8-V10 growth stage using UAV derived VIs based on the blue, green and thermal sections of the EMS;
- ii. UAV derived data can optimally predict maize biomass and grain yield during the R3-R4 growth stage using ENDVI, NDVI and the red edge band and proportional yield during the VT-R1 growth stage using NDVI, SAVI, green and the red edge band;
- iii. The grain yield models produced higher accuracies in estimating maize yield when compared to the absolute plant biomass and the biomass of grain yield as a proportion of absolute plant biomass models.

These findings are a fundamental step towards the establishment of timely accurate maize production estimates critical for intervention measures to cover for possible deficits and leakages. Optimizing food production will facilitate the attainment of the sustainable development goals of drastically reducing hunger and poverty while improving food and nutrition security.

#### **4.5 Recommendations for future studies**

Despite the high accuracies derived in this study, there is still a gap in research that requires further inquiry, particularly on smallholder maize farms. We recommend that future studies should;

- i. Seek to evaluate the utility of UAV derived data in predicting maize yield using plant crop height as a proxy of yield;
- ii. Assess the performance of UAV remotely sensed data in discriminating and mapping yield of maize considering the multiple cropping patterns associated with smallholder croplands within a single plot;
- iii. Test the utility of combining UAV derived texture models and red edge waveband derivatives in estimating maize LAI and yield.