



TRIBHUVAN UNIVERSITY
INSTITUTE OF ENGINEERING
PULCHOWK CAMPUS

CROP HEALTH MONITORING THROUGH AERIAL MAPPING USING QUADCOPTER BASED IMAGING

By:

Basant Karki (078BAS011)
Bishal Marasini (078BAS016)
Prabin Baniya (078BAS028)
Sameer Baskota (078BAS037)

A PROJECT REPORT SUBMITTED TO THE DEPARTMENT OF MECHANICAL
AND AEROSPACE ENGINEERING IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF BACHELOR IN AEROSPACE
ENGINEERING

DEPARTMENT OF MECHANICAL AND AEROSPACE ENGINEERING
LALITPUR, NEPAL

MARCH, 2026

TRIBHUVAN UNIVERSITY
INSTITUTE OF ENGINEERING, PULCHOWK
CAMPUS DEPARTMENT OF MECHANICAL AND
AEROSPACE ENGINEERING

The undersigned certify that they have read, and recommended to the Institute of Engineering for acceptance, a project report entitled "CROP HEALTH MONITORING THROUGH AERIAL MAPPING USING QUADCOPTER BASED IMAGING" submitted by Basant Karki, Bishal Marasini, Prabin Baniya and Sameer Baskota in partial fulfillment of the requirements for the Degree of Bachelor in Aerospace Engineering.



Supervisor Asst. Prof Kamal Darlami

Department of Mechanical and Aerospace
Engineering



Supervisor Asst. Prof Swaviman Acharya

Department of Mechanical and Aerospace
Engineering



External Examiner Er. Sanjeeb Bimali

Agricultural Engineer
Ministry of Agriculture, Forests And Environment



Asst. Prof. Dr. Sudip Bhattarai

Head of Department
Department of Mechanical and Aerospace
Engineering

Date: 2082/02/26

ACKNOWLEDGEMENT

We would like to express our sincere gratitude to the **Department of Mechanical and Aerospace Engineering** for providing us with the opportunity, guidance, and academic environment necessary to carry out this project successfully.

We are deeply thankful to our project supervisors, **Asst. Prof. Kamal Darlami** and **Asst. Prof. Swaviman Acharya**, for their continuous guidance, constructive feedback, and encouragement throughout the course of this project. Their technical insights and support were invaluable in shaping the direction and successful completion of our work.

We would like to express our special gratitude to **Asst. Prof. Biman Rimal** for his constant support and mentorship throughout the project. His guidance was instrumental in connecting us with **GeoKrishi**, and his encouragement and assistance at every stage of the project played a vital role in its successful execution.

We also extend our heartfelt appreciation to **Bikash Dangol** and **GeoKrishi** for their generous support in sharing valuable domain knowledge related to maize cultivation and providing hardware resources. Their practical guidance greatly enhanced our understanding of agricultural aspects relevant to this project.

Furthermore, we are grateful to the **Nepal Agricultural Research Council (NARC), Chitwan**, particularly the **National Maize Research Programme**, for providing us access to the maize field required for aerial mapping and data collection. Our sincere gratitude to **Asst. Prof. Krishna Prasad Aryal** for kindly providing access to maize fields under different nutrient treatments at **Rampur Campus, Khairahani, Chitwan** which was indispensable for the successful completion of this study. Their cooperation and support were crucial for the field implementation and validation of our study.

We sincerely thank **DroneDeploy** for providing complimentary access to their mapping software platform. Their assistance is greatly appreciated.

Finally, we would like to thank all those who directly or indirectly contributed to the successful completion of this project.

ABSTRACT

The accurate monitoring of nitrogen (N) status and leaf chlorophyll content is essential for optimizing maize (*Zea mays L.*) productivity and ensuring sustainable nutrient management. While manual SPAD chlorophyll measurements provide reliable assessments of plant vitality, they are limited by their small measurement area and high labor intensity. Furthermore, the high cost of specialized multispectral sensors presents a significant economic burden for smallholder farmers. This study investigates the potential of a cost-effective Unmanned Aerial Vehicle (UAV) platform, the DJI Phantom 4, equipped with a standard RGB sensor to predict maize health under diverse nutrient management strategies.

A field experiment was conducted using six treatments: T1 (control), T2 (100% RDN), T3 (100% RDN + biochar), T4 (urea-enriched biochar), T5 (50% urea + cattle urine-enriched biochar), and T6 (50% urea + human urine-enriched biochar), arranged with four replications. Ground-truth SPAD values were acquired at key growth stages to establish a physiological baseline.

High-resolution RGB imagery was used to compute vegetation indices (VIs), including Normalized Green-Red Difference Index (NGRDI), Triangular Greenness Index (TGI), and Visible Atmospherically Resistant Index (VARI). Correlation and regression analyses were performed to identify indices sensitive to chlorophyll variations, consistent with previous studies [1].

Results showed that RGB-derived VIs are effective proxies for crop health, with VARI and NGRDI showing strong correlation with SPAD values. A Random Forest (RF) model was developed for SPAD prediction and crop classification, leveraging its robustness against non-linearity and overfitting [2]. The model performance aligns with similar studies achieving high prediction accuracy.

This study demonstrates that UAV-RGB systems combined with machine learning provide a scalable, non-destructive solution for post-processed crop monitoring and precision nutrient management.

Keywords: Biochar, GLI, NPCI, Nutrient Management, Orthomosaic Mapping, Precision Agriculture, Random Forest, RGB Vegetation Indices, SPAD Chlorophyll, TGI, UAV Remote Sensing, VARI.

TABLE OF CONTENTS

TITLE PAGE	i
ABSTRACT	iv
TABLE OF CONTENTS	v
LIST OF FIGURES	ix
LIST OF TABLES	xi
LIST OF ABBREVIATIONS AND ACRONYMS	xii
 1 INTRODUCTION	 1
1.1 Background	1
1.2 Problem Statement	2
1.3 Objectives	2
1.3.1 Main Objective	2
1.3.2 Specific Objectives	2
1.4 Applications	3
1.5 System requirements	3
1.5.1 Hardware Requirements	3
1.5.2 DJI Phantom 4 Pro V2.0	4
1.5.3 Software Requirements	5
 2 LITERATURE REVIEW	 6
2.1 Maize and Its Importance	6
2.2 UAV and UAV-Based Imaging for Vegetation Indices	6
2.3 Efficient Nutrient Management through Advanced Crop Monitoring Techniques	6
2.4 Role of Chlorophyll in Plant Health Assessment	7
2.4.1 Relationship between Nitrogen and Chlorophyll Content	7
2.4.2 Chlorophyll as an Indicator of Crop Vigor and Productivity	7
2.5 SPAD Meter as a Ground Truth Measurement Tool	8
2.5.1 Working Principle of SPAD Meters	8
2.5.2 Advantages (Accuracy, Simplicity)	8
2.5.3 Limitations (Point-Based, Labor-Intensive)	8
2.5.4 Use of SPAD in Nitrogen Assessment Studies	8
2.6 UAV-Based Remote Sensing in Agriculture	8
2.6.1 Evolution of UAVs in Precision Agriculture	8
2.6.2 Advantages of UAVs over Satellite Imaging	8
2.6.3 Applications in Crop Health, Stress Detection, and Yield Prediction	9

2.7	RGB Imaging for Crop Monitoring	9
2.7.1	Characteristics of RGB Sensors (e.g., DJI Phantom 4)	9
2.7.2	Comparison with Multispectral and Hyperspectral Sensors	9
2.7.3	Cost-Effectiveness and Accessibility	9
2.8	Vegetation Indices Derived from RGB Images	9
2.8.1	Visible Atmospherically Resistant Index (VARI)	9
2.8.2	Other RGB-Based Indices	10
2.9	Correlation Between Vegetation Indices and SPAD Values	10
2.9.1	Previous Studies Linking RGB Indices with Chlorophyll/SPAD	10
2.9.2	Strength of Correlations under Different Conditions	10
2.9.3	Gaps in Validation under Diverse Treatments	10
2.10	Impact of Nutrient Treatments on Crop Health	10
2.10.1	Role of Urea (RDN Concept) and Biochar	10
2.10.2	Organic Amendments (Cattle Urine, Human Urine)	11
2.11	Machine Learning and Statistical Models for Crop Health Prediction	11
2.11.1	Regression Models and ANOVA	11
2.11.2	Machine Learning Approaches	11
2.12	Research Gaps	11
3	THEORETICAL BACKGROUND	12
3.1	Remote Sensing and UAV-Based Imaging	12
3.1.1	Principles of Remote Sensing	12
3.1.2	Electromagnetic Spectrum and Vegetation Reflectance	12
3.1.3	UAV Remote Sensing Platforms	13
3.2	RGB Vegetation Indices	14
3.2.1	Spectral Basis of RGB Indices	14
3.2.2	Excess Green Index (ExG)	14
3.2.3	Green Leaf Index (GLI)	15
3.2.4	Normalized Green-Red Difference Index (NGRDI)	15
3.2.5	Visible Atmospherically Resistant Index (VARI)	16
3.2.6	Triangular Greenness Index (TGI)	16
3.2.7	Normalized Pigment Chlorophyll Index (NPCl)	17
3.3	SPAD Chlorophyll Measurement	17
3.3.1	Principle of SPAD Measurement	17
3.3.2	SPAD as a Nitrogen Proxy	18
3.4	Image Processing for Canopy Segmentation	18
3.4.1	Vegetation Segmentation	18
3.4.2	Morphological Image Processing	18
3.5	Statistical Analysis Framework	19

3.5.1	Pearson Correlation Analysis	19
3.5.2	One-Way Analysis of Variance (ANOVA)	19
3.5.3	Correlation Analysis	20
3.5.4	Linear Regression Analysis	20
3.5.5	Root Mean Square Error (RMSE)	21
3.5.6	Normalized Root Mean Square Error (NRMSE)	21
3.5.7	Zonal Statistics	21
4	METHODOLOGY	23
4.1	Problem Identification and Literature Review	24
4.2	Field Preparation and Experimental Setup	24
4.2.1	Study Area	25
4.2.2	Experimental Details	25
4.2.3	Recommended Dose of Fertilizer (RDF)	25
4.2.4	Test Crop and Agronomic Practices	26
4.3	UAV Acquisition and Familiarization	26
4.4	UAV Calibration and System Preparation	26
4.5	Mission Planning	27
4.5.1	Path Planning and Grid Creation	28
4.5.2	Acquisition of RGB Images	29
4.6	Controlled field Environment	29
4.7	Orthomosaic Map Generation	31
4.8	Image Segmentation using python	31
4.8.1	Vegetation Spectral Indices	31
4.8.2	Stream 2: HSV Color Gate	32
4.8.3	Texture Smoothness	32
4.8.4	L*a*b* Green Axis	33
4.8.5	Multi-Evidence Fusion	33
4.8.6	GrabCut Refinement	33
4.8.7	Spatial Grid Division	33
4.8.8	Input Features	34
4.8.9	Train/Test Split	34
4.9	Tile-Based Prediction for New Images	35
4.10	Evaluation Metrics	35
4.11	Output	35
4.11.1	Confusion Matrix	35
4.11.2	Feature Importance	36
4.12	Index calculation Using QGIS	40
4.12.1	Data Preparation	40

4.12.2	Vegetation Segmentation and Canopy Masking	40
4.12.3	Vegetation Index Calculation	41
4.12.4	Symbology and Visualization	43
4.12.5	Plot Boundary Setup	44
4.12.6	Zonal Statistics Extraction	44
4.12.7	Data Export and Preparation for Analysis	45
4.13	Data Analysis	45
5	RESULT AND DISCUSSION	47
5.1	SPAD Values for Different Fertilizer Treatments	47
5.2	Relationships between SPAD Values and Vegetation Indices	49
5.3	Validation of the Linear Regression Models for Prediction of SPAD Values	54
5.3.1	Multi-Index Model Selection and Regularization	55
5.3.2	Predictive Accuracy and Significance	55
5.3.3	Visual Validation of Model Fit	56
5.3.4	The Predictive Equation	56
5.3.5	Conclusion on Robustness	57
5.4	Vegetation Index Thresholding and Spatial Classification in QGIS . . .	58
5.4.1	Spatial Distribution and Classification Accuracy	58
5.4.2	Threshold Calibration in Controlled Conditions	59
5.5	Predicted Result From Model	60
5.6	Limitations	61
5.7	Problems Encountered	61
5.8	Budget Estimation	62
5.9	Gantt Chart	63
5.10	Conclusion	63
5.10.1	Feasibility of Low-Cost Aerial Monitoring	63
5.10.2	Impact of Nutrient Management Strategies	63
5.10.3	Robustness of Vegetation Indices as Health Proxies	64
5.10.4	Performance of Predictive and Classification Models	64
5.11	Scope for Future Enhancement	65
6	APPENDICES	66
6.1	APPENDIX-A : Project Code	66
6.2	APPENDIX-B: Field Setup	66
6.3	APPENDIX-C: Field Validation with expert.	67

LIST OF FIGURES

Fig 1.1:	DJI Phantom 4 Pro V2.0	4
Fig 4.1:	Detail flowchart of the project Methodology	23
Fig 4.2:	Field setup	25
Fig 4.3:	Calibration image for camera.	27
Fig 4.4:	Camera Calibration.	27
Fig 4.5:	Image Acquisition plan	28
Fig 4.6:	Data Acquisition executed	28
Fig 4.7:	Maize Field	30
Fig 4.8:	Field Setup	30
Fig 4.9:	Confusion Matrix	36
Fig 4.10:	Feature Importance	37
Fig 4.11:	Scatter Plot : ExG vs VARI	37
Fig 4.12:	Scatter Plot : ExG vs NGRDI	38
Fig 4.13:	Scatter Plot : VARI vs NGRDI	39
Fig 4.14:	Unmasked Orthomosaic	41
Fig 4.15:	Masked Orthomosaic	41
Fig 4.16:	Grid creation in QGIS.	44
Fig 4.17:	Masking in steps.	45
Fig 5.1:	SPAD Values across Treatments (T1–T6) (20 Feb,2026) .	48
Fig 5.2:	Visualization of ground-truth physiological data(20 Feb, 2026)	49
Fig 5.3:	Fit of SPAD Vs NPCI (20 Feb)	50
Fig 5.4:	Fit of SPAD Vs NPCI (22 Mar)	50
Fig 5.5:	Heatmap of Masked NPCI	50
Fig 5.6:	Fit of SPAD Vs VARI (20 Feb,2026)	51
Fig 5.7:	Fit of SPAD Vs VARI (22 Mar, 2026)	51
Fig 5.8:	Heatmap of masked VARI	51
Fig 5.9:	Fit of SPAD Vs TGI (20 Feb,2026)	52
Fig 5.10:	Fit of SPAD Vs TGI (22 Mar, 2026)	52
Fig 5.11:	Heatmap of Masked TGI	52
Fig 5.12:	Fit of SPAD Vs NGRDI (20 Feb,2026)	53
Fig 5.13:	Fit of SPAD Vs NGRDI (22 Mar,2026)	53
Fig 5.14:	Heatmap of Masked NGRDI	53
Fig 5.15:	Fit of SPAD Vs GLI (20 Feb,2026)	54
Fig 5.16:	Fit of SPAD Vs GLI (22 Mar,2026)	54
Fig 5.17:	Heatmap of masked GLI	54

Fig 5.18:	Parameter estimates for original predictors	55
Fig 5.19:	Model Validation Curve for Regularization Parameter Se- lection	55
Fig 5.20:	Actual by predicted plot	56
Fig 5.21:	Residuals for predictor plots	57
Fig 5.22:	Residuals from validation of the predictor plots	57
Fig 5.23:	Model prediction equation	57
Fig 5.24:	Fig: Predicted SPAD values vs Actual SPAD values . . .	58
Fig 5.25:	Early stage Orthomosaic map	59
Fig 5.26:	Early stage classified map	59
Fig 5.27:	Silking stage Orthomosaic map	59
Fig 5.28:	Silking stage classified map.	59
Fig 5.29:	Overview of treatments.	60
Fig 5.30:	Nutrient Sufficient	60
Fig 5.31:	Nutrient Deficient	60
Fig 5.32:	Project Timeline	63
Fig 6.1:	Maize field showing the experimental plots	66
Fig 6.2:	Drone in flight capturing aerial images of the experimen- tal plots	66
Fig 6.3:	Photo with agricultural expert.	67

LIST OF TABLES

Table 1.1:	Hardware Requirements	3
Table 1.2:	PIV Image Size and Supported Frame Rates	4
Table 1.3:	Gimbal System Specifications	4
Table 1.4:	Flight and Positioning Data	5
Table 1.5:	Software Requirements	5
Table 4.1:	HSV Thresholds used for maize leaf segmentation	32
Table 4.2:	Random Forest Hyperparameters and Justifications	34
Table 4.3:	Symbology settings applied to each vegetation index layer in QGIS	44
Table 5.1:	Vegetation Indices and SPAD Measurements	47
Table 5.2:	Summary of SPAD mean values and statistical variance for nutrient treatments T1–T6.	49
Table 5.3:	Linear regression parameters and correlation coefficients between SPAD values and UAV RGB indices ($n = 24$).	55
Table 5.4:	Effect Tests for the Multi-Index SPAD Prediction Model.	56
Table 5.5:	Applied Classification Thresholds by Growth Stage	58
Table 5.6:	Budget Estimation	62

LIST OF ABBREVIATIONS AND ACRONYMS

AI	Artificial Intelligence
AGL	Above Ground Level
CAAN	Civil Aviation Authority of Nepal
CNN	Convolutional Neural Network
CSV	Comma Separated Values
CWSI	Crop Water Stress Index
DL	Deep Learning
EXIF	Exchangeable Image File Format
ExG	Excess Green Index
GCP	Ground Control Point
GCS	Ground Control Station
GeoTIFF	Geo Tagged Image File Format
GLI	Green Leaf Index
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GSD	Ground Sampling Distance
LAI	Leaf Area Index
LiDAR	Light Detection and Ranging
NDRE	Normalized Difference Red Edge
NDVI	Normalized Difference Vegetation Index
NGRDI	Normalized Green Red Difference Index
NIR	Near Infrared
NPCI	Normalized Pigment Chlorophyll Index
NPK	Nitrogen Phosphorus Potassium
OSAVI	Optimized Soil Adjusted Vegetation Index
RCBD	Randomized Complete Block Design
RDN	Recommended Dose of Nitrogen
RTK	Real Time Kinematic
SAVI	Soil Adjusted Vegetation Index
SfM	Structure from Motion
SIFT	Scale Invariant Feature Transform
SPAD	Soil Plant Analysis Development
TGI	Triangular Greenness Index
UAS	Unmanned Aerial System
UTM	Universal Transverse Mercator
VARI	Visible Atmospherically Resistant Index
VLOS	Visual Line of Sight
VPD	Vapor Pressure Deficit

Chapter 1: INTRODUCTION

1.1 Background

Maize is the second most important staple crop in Nepal after rice and plays a vital role in ensuring national food security, rural livelihoods, and agricultural sustainability. It is cultivated across diverse agro-ecological regions and serves as a primary source of food, feed, and income for smallholder farmers. Despite its high genetic yield potential, maize productivity in Nepal remains relatively low compared to global standards, largely due to inadequate and imbalanced nutrient management practices.[3] Efficient nutrient management is therefore essential not only to enhance crop yield but also to promote sustainable agricultural production systems.

Previous studies have demonstrated that maize growth and yield are highly influenced by nutrient availability and management strategies. Different nutrient treatments, including combinations of organic and inorganic fertilizers, have shown significant variations in crop performance. Statistical tools such as Analysis of Variance (ANOVA) have been widely used to evaluate the significance of treatment effects, while regression analysis has been applied to model relationships between growth parameters, yield, and environmental or management variables. These approaches provide a robust framework for understanding variability in crop response under different experimental conditions.[4]

In recent years, the application of vegetation and spectral indices has emerged as an effective, non-destructive method for assessing crop vigor and nutritional status. Indices such as the Visible Atmospherically Resistant Index (VARI), Green Leaf Index (GLI), Triangular Greenness Index (TGI), and Normalized Green-Red Difference Index (NGRDI) utilize reflectance data from crop canopies to estimate greenness, chlorophyll content, and overall plant health.[1][4] These indices have been successfully applied in precision agriculture for monitoring crop conditions and guiding management decisions. However, their application in maize under varying nutritional conditions, particularly in the context of Nepal, remains limited.

Low maize productivity in Nepal continues to be a major challenge, primarily due to poor nutrient management and the lack of timely, data-driven diagnostic tools. Traditional practices often rely on visual assessment and farmer experience, which may not accurately detect early nutrient deficiencies or optimize fertilizer application.[5] This highlights a critical gap in integrating advanced analytical techniques with practical field-level applications.

Therefore, there is a need to evaluate maize performance under different nutritional status plots using a combination of statistical and spectral approaches. The integration of ANOVA and regression analysis with vegetation indices can provide a more precise and quantitative understanding of crop response to nutrient treatments. Such an approach can improve nutrient diagnosis, support precision agriculture practices, and contribute to enhanced productivity in the local farming context of Nepal.

1.2 Problem Statement

Traditional agricultural monitoring methods are labor-intensive, time-consuming, and unsuitable for large-scale crop assessment. While multispectral sensors provide accurate vegetation indices for quantitative analysis, their high cost limits adoption among smallholder farmers. Low-cost RGB cameras mounted on UAVs offer a promising alternative; however, existing RGB-based approaches remain inefficient due to manual data collection and complex post-processing, restricting their effectiveness for timely crop health monitoring.

Therefore, this study aims to establish a robust mathematical relationship between RGB-derived vegetation indices (such as VARI) and crop physiological status under varying fertilization treatments, including urine-enriched biochar. This approach seeks to enable a cost-effective, scalable, and non-destructive method for crop health monitoring.

1.3 Objectives

1.3.1 Main Objective

To assess maize nutritional status using UAV based RGB aerial imagery through the computation and comparative analysis of multiple vegetation indices.

1.3.2 Specific Objectives

- i. To perform systematic UAV flights to capture high-resolution RGB images and generate georeferenced orthomosaic maps.
- ii. To develop an image processing pipeline to segment maize canopy and remove non-vegetation pixels.
- iii. To compute multiple RGB derived vegetation indices from UAV derived masked canopy imagery and spatially map their distribution across experimental plots.
- iv. To quantify the effect of six distinct nutrient treatments on crop chlorophyll content using SPAD-502 meter readings across four replications.

1.4 Applications

- i Enables early and accurate detection of maize nutrient deficiencies for timely management interventions.
- ii Supports site-specific nutrient application, enhancing fertilizer use efficiency and productivity.
- iii Reduces labor and costs associated with traditional crop monitoring through UAV-based remote sensing.
- iv Provides reliable data to inform agricultural extension services and policy decisions.
- v Establishes a scalable framework adaptable to other crops and regions for broader agricultural applications.

1.5 System requirements

1.5.1 Hardware Requirements

Table 1.1: Hardware Requirements

SN	Hardware	Description
1	DJI Phantom 4 Pro	UAV platform equipped with RGB imaging system for vegetation index calculation
2	SPAD Chlorophyll Meter	Handheld instrument for ground-truth leaf chlorophyll measurement.
3	Mobile Device/Tablet	For running DJI flight control applications and mission planning software
4	Memory Cards (SD Cards)	High-speed storage media for onboard image data recording
5	Battery Charger & Hub	Multi-battery charging system for field operations
6	Laptop/Workstation	High-performance computer for image processing, orthomosaic generation, and vegetation index calculation

1.5.2 DJI Phantom 4 Pro V2.0

Phantom 4 Pro V2.0 was used for mapping due to its high resolution cameras and OcuSync Transmission system. Some of the specifications of the drone used are:

Physical Dimensions and Airframe

- i. Weight (Battery & Propellers Included): 1375g
- ii. Diagonal Size (Excluding Propellers): 350mm
- iii. Max Wind Speed Resistance: 10m/s
- iv. Operating Temperature Range: 0 °C to 40 °C
- v. Weight (Maximum Flight Time): approx. 30min



Figure 1.1: DJI Phantom 4 Pro V2.0

Camera and Imaging System

The core of the mapping capability lies in the 1-inch sensor and the mechanical shutter system.

- i. Sensor: 1" CMOS; Effective pixels: 20M
- ii. Lens: FOV 84° 8.8 mm/24 mm (35 mm format equivalent)
- iii. Aperture: f/2.8 – f/11
- iv. Mechanical Shutter Speed: 8 – 1/2000 s

Table 1.2: PIV Image Size and Supported Frame Rates

Resolution	Supported Frame Rates
4096 × 2160	24/25/30/48/50 fps
3840 × 2160	24/25/30/48/50/60 fps
2720 × 1530	24/25/30/48/50/60 fps
1920 × 1080	24/25/30/48/50/60/120 fps
1280 × 720	24/25/30/48/50/60/120 fps

Gimbal Specifications

Table 1.3: Gimbal System Specifications

Feature	Specification
Stabilization	3-axis (pitch, roll, yaw)
Controllable Range	Pitch: −90° to +30°
Max Controllable Angular Speed	Pitch: 90°/s

3. Navigation and Flight Endurance

Table 1.4: Flight and Positioning Data

Feature	Specification
Max Flight Time	Approx. 30 minutes
Satellite Positioning Systems	GPS / GLONASS
Transmission System	OcuSync 2.0
Vertical Hover Accuracy	± 0.1 m (Vision) / ± 0.5 m (GPS)
Horizontal Hover Accuracy	± 0.3 m (Vision) / ± 1.5 m (GPS)
Battery Capacity	5870 mAh (15.2 V)

1.5.3 Software Requirements

Table 1.5: Software Requirements

SN	Software	Description
1	DroneDeploy	UAV mission planning, flight path execution, and orthomosaic generation from acquired RGB imagery.
2	QGIS	Geospatial analysis, vegetation index calculation, canopy segmentation, zonal statistics, and raster processing
3	Python	Statistical analysis, correlation computation, health classification, and visualization of vegetation index results
4	Microsoft Excel	Ground-truth SPAD data recording, tabulation, and preliminary data organization.
5	JMP Pro	Advanced statistical analysis, regression modeling, correlation analysis, and data visualization of vegetation index and SPAD datasets.
6	DJI GO 4 App	UAV flight control, camera settings configuration, and real-time flight monitoring.

Chapter 2: LITERATURE REVIEW

2.1 Maize and Its Importance

Maize (*Zea mays L.*) is a vital cereal crop globally, ranking third in total production after wheat and rice [6]. Known as the “Queen of Cereals” due to its high yield potential, maize serves as a primary staple for vulnerable populations and plays a crucial role in global food security [6], [7]. However, maize is highly susceptible to water stress, which remains a major constraint to productivity across diverse agro-ecological environments [6]. Yield losses can exceed 80% when drought occurs during critical reproductive stages such as flowering and grain filling [6], [7]. Moisture deficits during rapid vegetative growth can also result in significant yield reductions of up to 40% [7].

2.2 UAV and UAV-Based Imaging for Vegetation Indices

Unmanned Aerial Vehicles (UAVs) have emerged as high-throughput, flexible, and cost-effective platforms for agricultural monitoring [8][9]. Compared to satellite remote sensing, which is constrained by coarse spatial resolution and cloud interference, UAVs provide centimeter-level high-resolution data at the farm scale [8], [10]. These platforms can carry RGB, multispectral, and thermal sensors to capture narrow-band reflectance data [10]. Such data are transformed into vegetation indices (VIs) to evaluate plant moisture status, vigor, and overall crop health [10][11]. UAV-based imagery further enhances understanding of intra-field variability and land cover dynamics [8], [10].

2.3 Efficient Nutrient Management through Advanced Crop Monitoring Techniques

Particularly concerning nitrogen (N), is fundamental to global food security and sustainable agriculture because it directly influences chlorophyll synthesis, photosynthetic efficiency, and total grain yield [1], [12]. Proper management requires a delicate balance: while adequate N promotes vigorous canopy development and biomass accumulation, excessive application leads to significant environmental degradation, including nitrate leaching into groundwater, aquatic eutrophication, and increased greenhouse gas emissions [1], [3], [12].

Despite its importance, traditional field-based methods for monitoring N status are often inadequate for modern precision agriculture. Manual ground measurements, such as chemical chlorophyll analysis, are labor-intensive, time-consuming, and destructive [4]. Even non-destructive tools like SPAD chlorophyll meters, while accurate, are in-

herently limited by their small measurement area and the requirement for point-based, leaf-by-leaf sampling, which makes them inefficient for continuous, large-scale field monitoring [1].

Satellite remote sensing offers broader coverage but often lacks the spatial and temporal resolution required to track dynamic crop changes due to long revisit intervals and interference from cloud cover [1], [13]. Consequently, there is a critical need for scalable, non-destructive, and cost-effective monitoring solutions that can capture the high degree of in-field heterogeneity [1], [14]. Unmanned Aerial Vehicle (UAV) platforms equipped with standard RGB cameras have emerged as a vital alternative, providing high-throughput data with centimeter-level resolution without the “tremendous economic burden” of specialized multispectral or hyperspectral sensors [1]. By establishing robust mathematical relationships between aerial vegetation indices and ground-truth physiological measurements, researchers can enable precise, real-time nutrient management that optimizes productivity while minimizing environmental impacts [1], [4].

2.4 Role of Chlorophyll in Plant Health Assessment

2.4.1 Relationship between Nitrogen and Chlorophyll Content

Nitrogen (N) is a key constituent of all living cells and is present in the structure of chlorophyll, the green pigment of leaves. Leaf chloroplasts contain approximately 70% of the total leaf nitrogen concentration, creating a strong positive correlation between leaf N status and chlorophyll content. Nitrogen is also a critical component of the enzymes involved in chlorophyll synthesis; thus, an insufficient N supply directly results in reduced chlorophyll concentration [12].

2.4.2 Chlorophyll as an Indicator of Crop Vigor and Productivity

Chlorophyll enables plants to transfer energy from sunlight through photosynthesis, determining their photosynthetic capacity and overall physiological status. Variations in leaf color caused by changes in chlorophyll content serve as a primary indicator of plant nutrition, water availability, and senescence.[15] Because chlorophyll content is closely related to maize grain yield and biomass accumulation, it is considered a vital indicator of crop vigor and productivity [16].

2.5 SPAD Meter as a Ground Truth Measurement Tool

2.5.1 Working Principle of SPAD Meters

The SPAD-502Plus Chlorophyll Meter determines the relative chlorophyll content by measuring leaf transmittance at two wavelength regions: red (650 nm), where chlorophyll absorption is high, and near-infrared (940 nm), which serves as a reference. The resulting SPAD value reflects the “greenness” of the leaf [17].

2.5.2 Advantages (Accuracy, Simplicity)

The SPAD meter is widely recognized as a fast, non-destructive, and portable tool that provides an accurate proxy for leaf chlorophyll and nitrogen status without requiring labor-intensive chemical analysis. It allows for precise nitrogen management in field conditions, reducing the risk of over-fertilization [17].

2.5.3 Limitations (Point-Based, Labor-Intensive)

Despite its accuracy, the SPAD meter is limited by its small measurement area (2 mm × 3 mm), requiring operators to take readings leaf by leaf. This point-based sampling makes the device inefficient for large-scale field monitoring and incapable of capturing high spatial variability within fields.

2.5.4 Use of SPAD in Nitrogen Assessment Studies

SPAD values have been successfully used as a diagnostic tool for nitrogen fertilizer management in crops such as wheat, rice, and maize. Studies indicate that SPAD readings are highly sensitive to nitrogen application rates, with peak values typically observed during critical growth stages such as jointing in wheat or grain filling in maize [4].

2.6 UAV-Based Remote Sensing in Agriculture

2.6.1 Evolution of UAVs in Precision Agriculture

Unmanned Aerial Vehicles (UAVs) have significantly advanced remote sensing applications in agriculture due to technological developments and decreasing sensor costs. They enable mapping of areas of interest with centimeter-level spatial accuracy [18].

2.6.2 Advantages of UAVs over Satellite Imaging

UAVs offer significant advantages over satellite platforms, including higher spatial and temporal resolution, operational flexibility, and the ability to capture data under cloud

cover. Unlike satellites, which are limited by long revisit intervals, UAVs can acquire data on demand during critical phenological stages [18][19].

2.6.3 Applications in Crop Health, Stress Detection, and Yield Prediction

UAV imagery is widely used to estimate Leaf Area Index (LAI), plant height, biomass, and yield. It is also effective for detecting biotic and abiotic stresses, such as water deficiency and nutrient imbalances, enabling site-specific crop management [18].

2.7 RGB Imaging for Crop Monitoring

2.7.1 Characteristics of RGB Sensors (e.g., DJI Phantom 4)

Standard RGB cameras capture high-resolution images in the visible spectral bands (Red, Green, and Blue). These sensors record information visible to the human eye, allowing analysis of plant growth patterns, textures, and color intensity [5].

2.7.2 Comparison with Multispectral and Hyperspectral Sensors

While multispectral and hyperspectral sensors provide detailed spectral information, particularly in the NIR and red-edge regions, they are often expensive and complex to process. In contrast, RGB sensors offer higher spatial resolution and ease of use but lack NIR sensitivity required for traditional vegetation analysis [1].

2.7.3 Cost-Effectiveness and Accessibility

Commercial RGB cameras provide a cost-effective alternative to advanced sensors, making UAV-based monitoring accessible to smallholder farmers and researchers, particularly in developing countries [1].

2.8 Vegetation Indices Derived from RGB Images

2.8.1 Visible Atmospherically Resistant Index (VARI)

Theory and formula: VARI is calculated as $(G - R)/(G + R - B)$. It was developed to estimate vegetation fraction with minimal sensitivity to atmospheric effects.

Sensitivity to chlorophyll and canopy structure: VARI responds directly to vegetation fraction and is sensitive to chlorophyll concentration, making it a reliable indicator of plant stress and productivity [1].

2.8.2 Other RGB-Based Indices

ExG (Excess Green Index): Calculated as $2G - R - B$, it is widely used to distinguish green vegetation from soil background.

GLI (Green Leaf Index): Calculated as $(2G - R - B)/(2G + R + B)$, it shows strong correlation with chlorophyll content and nitrogen status.

NGRDI (Normalized Green-Red Difference Index): Defined as $(G - R)/(G + R)$, it highlights vegetation vigor but may be influenced by canopy structure.

These indices are essential for detecting nutrient-related stress and monitoring biomass variations across growth stages [1], [5].

2.9 Correlation Between Vegetation Indices and SPAD Values

2.9.1 Previous Studies Linking RGB Indices with Chlorophyll/SPAD

Several studies have demonstrated that RGB-based vegetation indices are strongly correlated with SPAD values and chlorophyll content. Indices such as GLI and G-B have shown significant relationships with SPAD measurements in crops like barley and wheat [4].

2.9.2 Strength of Correlations under Different Conditions

Correlation coefficients between SPAD and RGB indices typically range from 0.70 to 0.80. However, these relationships can be influenced by UAV flight height, illumination conditions, and image quality.

2.9.3 Gaps in Validation under Diverse Treatments

Most existing studies focus on qualitative stress detection rather than quantitative nitrogen estimation under complex nutrient management systems. There is a lack of validation under integrated treatments such as biochar combined with organic amendments [1].

2.10 Impact of Nutrient Treatments on Crop Health

2.10.1 Role of Urea (RDN Concept) and Biochar

Standard nitrogen application (e.g., 100% RDN through urea) promotes vigorous crop growth; however, excessive use can lead to environmental issues such as nitrate leaching. Biochar application has been proposed as a sustainable strategy to improve soil structure, enhance nutrient retention, and reduce environmental impacts [12].

2.10.2 Organic Amendments (Cattle Urine, Human Urine)

Integrated nutrient management combining inorganic fertilizers with organic amendments (such as urine-enriched biochar) can improve nitrogen use efficiency. Evaluating these treatments requires understanding how different nutrient sources influence SPAD values and spectral responses across treatments.

2.11 Machine Learning and Statistical Models for Crop Health Prediction

2.11.1 Regression Models and ANOVA

Traditional statistical approaches, including linear and multiple regression, are widely used to relate vegetation indices to physiological parameters. ANOVA is commonly applied to determine the significance of differences among nutrient treatments.

2.11.2 Machine Learning Approaches

Machine learning algorithms such as Random Forest (RF), Support Vector Machines (SVM), and neural networks have demonstrated strong capability in modeling non-linear relationships between spectral data and crop health. RF is particularly effective due to its robustness against noise and multicollinearity [2], [20].

2.12 Research Gaps

There is a growing need for scalable, low-cost predictive models that can serve as alternatives to expensive NIR-based systems. Current research lacks comprehensive validation of RGB-derived indices as quantitative proxies for chlorophyll across multiple growth stages and diverse nutrient treatments, particularly those involving biochar and organic amendments. Establishing these relationships is essential for the practical implementation of UAV-based precision agriculture [1].

Chapter 3: THEORETICAL BACKGROUND

This section presents the key theoretical concepts that form the basis of the methods used in this study. It explains the fundamental principles of UAV based remote sensing, the working mechanism of RGB vegetation indices. It also describes the role of morphological image processing in separating crop canopy pixels from the background. A clear understanding of these concepts is essential for interpreting the analytical choices made in the study and for properly understanding the results obtained.

3.1 Remote Sensing and UAV-Based Imaging

3.1.1 Principles of Remote Sensing

Remote sensing is the science of gathering information about objects or areas on the Earth's surface from a distance, without physically touching them, by detecting and measuring electromagnetic radiation that is either reflected or emitted from the target. The whole thing works because different materials interact with light in different ways depending on the wavelength, some wavelengths get absorbed, some pass through, and some bounce back. When light hits a surface, the incoming energy gets divided into three parts according to the conservation of energy:

$$E_i(\lambda) = E_r(\lambda) + E_a(\lambda) + E_t(\lambda) \quad (3.1)$$

where $E_i(\lambda)$ is the incoming irradiance, $E_r(\lambda)$ is the reflected portion, $E_a(\lambda)$ is what gets absorbed, and $E_t(\lambda)$ is what passes through the material, all at a given wavelength λ . Remote sensing instruments are mainly concerned with the reflected portion $E_r(\lambda)$, because this is where the characteristic spectral fingerprints of different surfaces are encoded.

3.1.2 Electromagnetic Spectrum and Vegetation Reflectance

The electromagnetic spectrum covers a continuous range of wavelengths, from short gamma rays all the way to long radio waves. For agricultural remote sensing, the most useful portions are the visible range (400-700 nm), the near-infrared range (700-1300 nm), and the shortwave-infrared range (1300-2500 nm), as these parts of the spectrum carry the most useful information for characterizing vegetation. The way a healthy green plant reflects light is shaped by three main physical and biochemical processes, each operating across a different part of the spectrum:

- i. **Visible region (400-700 nm):** In this range, the pigments inside leaves call the

shots. Chlorophyll-a and chlorophyll-b soak up a lot of light in the blue (430-450 nm) and red (640-680 nm) portions because that light powers photosynthesis, so reflectance is low at those wavelengths. Around the green part of the spectrum (520-560 nm), chlorophyll absorption drops off and reflectance picks up this is the so-called "green bump" that makes healthy plants look green to our eyes. Carotenoid pigments also absorb in the blue region (400-500 nm) and become more visible spectrally when chlorophyll breaks down, for instance during nitrogen deficiency or when plants start to senesce.

- ii. **Red-edge region (680-750 nm):** The red-edge is the sharp transition zone between the deep absorption trough in the red and the high-reflectance plateau in the near-infrared. Because this transition shifts with chlorophyll content, the red-edge is one of the most sensitive spectral features for estimating how much chlorophyll a leaf contains.
- iii. **Near-infrared region (750-1300 nm):** Here, it is not the pigments but the physical structure of the leaf that determines reflectance. Healthy leaves have a spongy layer of mesophyll cells that scatter near-infrared light in multiple directions internally, producing the characteristically high NIR reflectance that distinguishes living vegetation from most other surfaces.

3.1.3 UAV Remote Sensing Platforms

Unmanned Aerial Vehicles (UAVs), more casually known as drones, have become popular remote sensing tools because they fill a gap between coarse satellite imagery and labour-intensive ground-based measurements. For precision agriculture in particular, UAVs offer several practical advantages that traditional aerial or satellite platforms simply cannot match:

- i. **Ultra-high spatial resolution:** When flown at low altitudes, UAVs can capture images with ground sampling distances of just a few centimetres, which is fine enough to pick up variability at the level of individual plant rows or small research plots.
- ii. **Flexible deployment:** UAVs can be set up and flown quickly whenever imagery is needed, without having to wait for a satellite pass or arrange a manned aircraft.
- iii. **Sub-cloud operation:** Because UAVs fly below the cloud layer, they avoid the atmospheric interference that often makes satellite images unusable during critical growth stages when timely monitoring matters most.

- iv. **Cost effectiveness:** A consumer-grade UAV with a standard RGB camera can deliver high-quality aerial imagery at a fraction of what a manned survey or a multispectral satellite subscription would cost.

3.2 RGB Vegetation Indices

3.2.1 Spectral Basis of RGB Indices

A standard RGB camera records image data across three broad bands: red (roughly 600-700 nm), green (roughly 500-600 nm), and blue (roughly 400-500 nm). These broadband channels are not as spectrally precise as narrowband multispectral or hyperspectral sensors, but they still capture a fundamental contrast between healthy vegetation which reflects a lot in the green and absorbs heavily in the red and bare soil or other non-vegetated surfaces, which tend to reflect more evenly across the visible range. RGB vegetation indices take advantage of this spectral contrast by combining the R, G, and B band values mathematically in ways that amplify the vegetation signal while pushing down noise from soil, shadows, and varying lighting.

One thing that needs to be taken care of before calculating RGB indices is normalising the band values to remove the effect of overall image brightness, which can shift depending on the sun angle, atmospheric haze, or camera exposure settings. The standard way to do this is to divide each band by the sum of all three:

$$r = \frac{R}{R+G+B}, \quad g = \frac{G}{R+G+B}, \quad b = \frac{B}{R+G+B} \quad (3.2)$$

where r , g , and b are the normalised band values and R , G , B are the raw digital numbers straight from the image. This step locks the index values within a fixed range and makes results comparable across images taken under different lighting conditions.

3.2.2 Excess Green Index (ExG)

The Excess Green Index was introduced as a simple and quick way to separate green vegetation from soil and other non-vegetated pixels in digital images. The formula boosts the green channel and subtracts out both the red and blue channels:

$$\text{ExG} = 2G - R - B \quad (3.3)$$

The normalised version removes the effect of illumination:

$$\text{ExG}_{\text{norm}} = \frac{2G - R - B}{R + G + B} \quad (3.4)$$

Normalised ExG values run from -1 to $+1$. Pixels with positive values have green-channel dominance and are likely to be photosynthetically active vegetation, while negative values point to red or blue dominance typical of bare soil or non-vegetated areas. ExG was mainly built for splitting vegetation from non-vegetation rather than quantifying plant health once you are already looking at vegetation pixels, its sensitivity to chlorophyll concentration is fairly limited compared to indices designed specifically for health assessment.

3.2.3 Green Leaf Index (GLI)

The Green Leaf Index uses all three RGB channels in a normalised formula that is structurally similar to ExG but gives a more balanced treatment to the red and blue channels in the denominator:

$$\text{GLI} = \frac{2G - R - B}{2G + R + B} \quad (3.5)$$

GLI ranges from -1 to $+1$ and is fully normalised, so values are directly comparable between images taken under different ambient lighting. GLI and ExG share the same numerator, meaning they respond in exactly the same way to the spectral gap between the green channel and the average of red and blue. The only real difference is in the denominator GLI's version handles illumination variation slightly better. Like ExG, GLI works well as a segmentation tool and shows moderate sensitivity to chlorophyll.

3.2.4 Normalized Green-Red Difference Index (NGRDI)

The Normalized Green-Red Difference Index is probably the most straightforward RGB counterpart to NDVI. The idea is simply to swap the near-infrared band of NDVI for the green band:

$$\text{NGRDI} = \frac{G - R}{G + R} \quad (3.6)$$

The logic behind this swap comes from how leaves behave spectrally: a healthy, chlorophyll-rich leaf absorbs a lot of red light for photosynthesis while bouncing back relatively more green light, so the green-minus-red difference is positive and tends to increase with chlorophyll content. When nitrogen stress sets in and chlorophyll starts to break down, red absorption drops, and the G-R difference shrinks or even turns negative. NGRDI values are strictly between -1 and $+1$, and the index has been validated across many crop species as a reasonably reliable indicator of leaf chlorophyll concentration

and nitrogen status.

3.2.5 Visible Atmospherically Resistant Index (VARI)

VARI was developed by Gitelson2002 as an improvement over the plain green-red ratio, with the specific goal of reducing the distorting effect of atmospheric conditions on RGB imagery captured from elevated platforms. It brings the blue band into the denominator to partially correct for atmospheric scattering, which tends to affect short blue wavelengths more strongly than the longer red and green wavelengths:

$$\text{VARI} = \frac{G - R}{G + R - B} \quad (3.7)$$

Adding the blue band does two things at once: it helps correct for atmospheric scattering by exploiting how differently the atmosphere scatters different visible wavelengths, and it also makes the index sensitive to carotenoid pigments that absorb in the blue region and become more prominent as chlorophyll degrades under nutrient stress. VARI values typically fall between -1 and $+1$, though the denominator can theoretically reach zero, so a small stability constant needs to be added in the actual calculation to avoid division errors. VARI has consistently outperformed other RGB indices for estimating chlorophyll content and classifying nutrient status in multiple studies, and this was confirmed here as well it produced the highest correlation with SPAD measurements of all the indices tested ($r = 0.747$, $p < 0.001$).

3.2.6 Triangular Greenness Index (TGI)

TGI was proposed by based on a geometric idea: if you pick three characteristic points on a vegetation reflectance spectrum at the chlorophyll absorption peak in the red (670 nm), the green reflectance peak (550 nm), and a reference point in the blue (480 nm) those three points form a triangle. The area of that triangle grows with chlorophyll content because higher chlorophyll deepens the red absorption trough, which pulls that point lower and expands the triangle. For RGB imagery, the triangle is approximated as:

$$\text{TGI} = G - 0.39R - 0.61B \quad (3.8)$$

The coefficients 0.39 and 0.61 come from the wavelength-based weighting that defines the geometry of the original spectral triangle. Unlike normalised indices, TGI gives values in raw pixel intensity units when applied to 8-bit imagery, meaning the absolute numbers depend on the image scale. TGI was purpose-built for chlorophyll quantifica-

tion and has been adopted as a standard RGB chlorophyll proxy by professional UAV mapping software including Pix4D.

3.2.7 Normalized Pigment Chlorophyll Index (NPCI)

NPCI was originally formulated by Penueals1994 using narrowband hyperspectral reflectance at 680 nm and 430 nm to get at the ratio of total photosynthetic pigments to chlorophyll:

$$\text{NPCI} = \frac{R_{680} - R_{430}}{R_{680} + R_{430}} \quad (3.9)$$

The reasoning behind NPCI is that chlorophyll and carotenoids respond differently to nitrogen limitation. When nitrogen runs short, chlorophyll which absorbs heavily at 680 nm breaks down faster than carotenoids, which absorb at 430 nm. This shifts the 680/430 nm reflectance ratio upward. So NPCI actually goes up as chlorophyll goes down it is inversely related to chlorophyll content, unlike every other index in this study, where higher values mean healthier vegetation.

When NPCI is adapted for RGB imagery using the broadband red channel (600-700 nm) and blue channel (400-500 nm) instead of the original narrowband wavelengths the formula becomes:

$$\text{NPCI}_{\text{RGB}} = \frac{R - B}{R + B + 0.0001} \quad (3.10)$$

Swapping narrowband reflectance for broadband channel values substantially weakens the index, because each broadband channel mixes together both chlorophyll-sensitive and chlorophyll-insensitive wavelengths, blurring the spectral signal that NPCI was originally designed to detect.

3.3 SPAD Chlorophyll Measurement

3.3.1 Principle of SPAD Measurement

The SPAD-502 chlorophyll meter (Minolta, Japan) works by shining light through a leaf at two wavelengths and measuring how much gets through. One wavelength is 650 nm, which sits right in the main chlorophyll absorption band the more chlorophyll present, the less light passes through. The other is 940 nm, a near-infrared reference wavelength where the leaf's optical behaviour does not change with chlorophyll content Minolta1989. The SPAD value is then calculated as:

$$\text{SPAD} = k \times \log \left(\frac{I_{940}/T_{940}}{I_{650}/T_{650}} \right) \quad (3.11)$$

where I_{940} and I_{650} are the reference intensities of the light source, T_{940} and T_{650} are the intensities measured after the light passes through the leaf at 940 nm and 650 nm respectively, and k is a calibration constant specific to the instrument. The output is a dimensionless number that correlates positively with leaf chlorophyll content, giving a fast and non-destructive estimate without the need for any chemical extraction.

3.3.2 SPAD as a Nitrogen Proxy

In crops like maize, most of the nitrogen in a leaf ends up locked inside chlorophyll molecules and other parts of the photosynthetic machinery. Because of this, how much chlorophyll a leaf contains tracks closely with how much nitrogen it has, making SPAD a convenient non-destructive indicator of nitrogen status. Schepers showed that SPAD readings below around 45 in fully expanded maize leaves during vegetative and reproductive growth stages are a sign of nitrogen deficiency that needs attention, while readings above this level suggest the plant is getting enough nitrogen. This threshold has been widely used in maize nitrogen management research since then, and it was used as the main classification boundary in the current study.

3.4 Image Processing for Canopy Segmentation

3.4.1 Vegetation Segmentation

Before vegetation index values can be reliably extracted from UAV imagery, the vegetation pixels need to be separated from everything else in the image bare soil, crop residue, shadows, and field edges all need to be excluded. This step is called vegetation segmentation. If it is skipped or done poorly, the index statistics for a given plot will be systematically biased, because soil and shadow pixels have very different spectral behaviour from vegetation and will pull the mean index values down. The typical approach is to apply a threshold to a vegetation index most commonly ExG and use it to produce a binary mask that marks pixels as either vegetation or non-vegetation.

3.4.2 Morphological Image Processing

Morphological image processing is a family of spatial filtering operations that work on the geometric structure of binary or grayscale images, using a small shape called a structuring element to define how the operation modifies nearby.

In this study, the Sieve filter in QGIS was used to carry out this clean-up step. It removes

any isolated clusters of vegetation pixels that fall below a set minimum size, which effectively performs an opening operation to get rid of noise-driven false positives in the vegetation classification.

3.5 Statistical Analysis Framework

3.5.1 Pearson Correlation Analysis

Pearson's product-moment correlation coefficient (r) is a measure of how strongly and in what direction two continuous variables are linearly related. For n paired observations of variables X and Y , it is defined as:

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \cdot \sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (3.12)$$

where $\bar{X}$ and $\bar{Y}$ are the sample means of X and Y respectively. The value of r runs from -1 (a perfect negative linear relationship) to $+1$ (a perfect positive linear relationship), with $r = 0$ meaning no linear association at all. Squaring r gives the coefficient of determination $R^2 = r^2$, which can be interpreted as the share of variance in one variable that is accounted for by the other.

3.5.2 One-Way Analysis of Variance (ANOVA)

One-way Analysis of Variance (ANOVA) is a statistical method used to determine whether there are significant differences between the means of three or more independent groups. It compares the variance between groups to the variance within groups. For k groups and a total of N observations, the F-statistic is defined as:

$$F = \frac{MS_{\text{between}}}{MS_{\text{within}}} \quad (3.13)$$

where

$$MS_{\text{between}} = \frac{\sum_{i=1}^k n_i (\bar{X}_i - \bar{X})^2}{k - 1} \quad (3.14)$$

$$MS_{\text{within}} = \frac{\sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2}{N - k} \quad (3.15)$$

Here, $\bar{X}_i$ is the mean of the i -th group, $\bar{X}$ is the overall mean, n_i is the number of observations in group i , and X_{ij} represents individual observations. A higher F-value indicates a greater likelihood that at least one group mean is significantly different.

3.5.3 Correlation Analysis

Correlation analysis is a statistical technique used to measure the strength and direction of the relationship between two continuous variables. It helps in understanding how changes in one variable are associated with changes in another variable.

The most commonly used measure is the Pearson product-moment correlation coefficient (r), which is defined as:

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \cdot \sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (3.16)$$

where X_i and Y_i are the individual observations, $\bar{X}$ and $\bar{Y}$ are the sample means, and n is the number of paired observations. The value of r ranges from -1 to $+1$, where:

- $r = +1$ indicates a perfect positive linear relationship,
- $r = -1$ indicates a perfect negative linear relationship,
- $r = 0$ indicates no linear relationship.

The strength of correlation can be interpreted as:

- $0.0 - 0.3$: weak correlation
- $0.3 - 0.7$: moderate correlation
- $0.7 - 1.0$: strong correlation

The coefficient of determination (R^2) is obtained by squaring the correlation coefficient:

$$R^2 = r^2 \quad (3.17)$$

R^2 represents the proportion of variance in the dependent variable that can be explained by the independent variable. In crop monitoring studies, correlation analysis is widely used to evaluate the relationship between vegetation indices (e.g., VARI, ExG) and ground-based measurements such as SPAD values.

3.5.4 Linear Regression Analysis

Linear regression is used to model the relationship between a dependent variable Y and one or more independent variables X . In simple linear regression, the relationship is expressed as:

$$Y = \beta_0 + \beta_1 X + \varepsilon \quad (3.18)$$

where β_0 is the intercept, β_1 is the slope coefficient, and ε is the error term. The model aims to minimize the difference between observed and predicted values, typically using

the least squares method. The coefficient of determination (R^2) is used to evaluate the model performance:

$$R^2 = 1 - \frac{\sum (Y_i - \hat{Y}_i)^2}{\sum (Y_i - \bar{Y})^2} \quad (3.19)$$

where $\hat{Y}_i$ are predicted values and $\bar{Y}$ is the mean of observed values.

3.5.5 Root Mean Square Error (RMSE)

Root Mean Square Error (RMSE) is a widely used metric to evaluate the accuracy of a predictive model. It measures the average magnitude of the prediction error.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2} \quad (3.20)$$

where Y_i represents observed values and $\hat{Y}_i$ represents predicted values. Lower RMSE values indicate better model performance.

3.5.6 Normalized Root Mean Square Error (NRMSE)

Normalized Root Mean Square Error (NRMSE) is a normalized version of RMSE that allows comparison across different datasets or scales.

$$NRMSE = \frac{RMSE}{Y_{\max} - Y_{\min}} \quad (3.21)$$

Alternatively, it can also be normalized by the mean:

$$NRMSE = \frac{RMSE}{\bar{Y}} \quad (3.22)$$

where $Y_{\max}$ and $Y_{\min}$ are the maximum and minimum observed values, respectively, and $\bar{Y}$ is the mean of observed values. NRMSE provides a scale-independent measure of model accuracy.

3.5.7 Zonal Statistics

Zonal statistics is a spatial analysis technique that summarises raster pixel values within user-defined spatial zones, which are usually represented as polygon features in a vector layer. For each zone polygon Z_i , the mean value across all valid pixels inside it is computed as:

$$\bar{v}_i = \frac{1}{N_i} \sum_{j=1}^{N_i} v_{ij} \quad (3.23)$$

where $\bar{v}_i$ is the mean raster value for zone Z_i , N_i is the number of valid (non-NoData) pixels inside that zone, and v_{ij} is the value of the j -th pixel within it. In this study, zonal

statistics is the main tool used to convert the spatially continuous raster index maps into a single summary number for each experimental plot, which can then be compared against the field measurements.

Chapter 4: METHODOLOGY

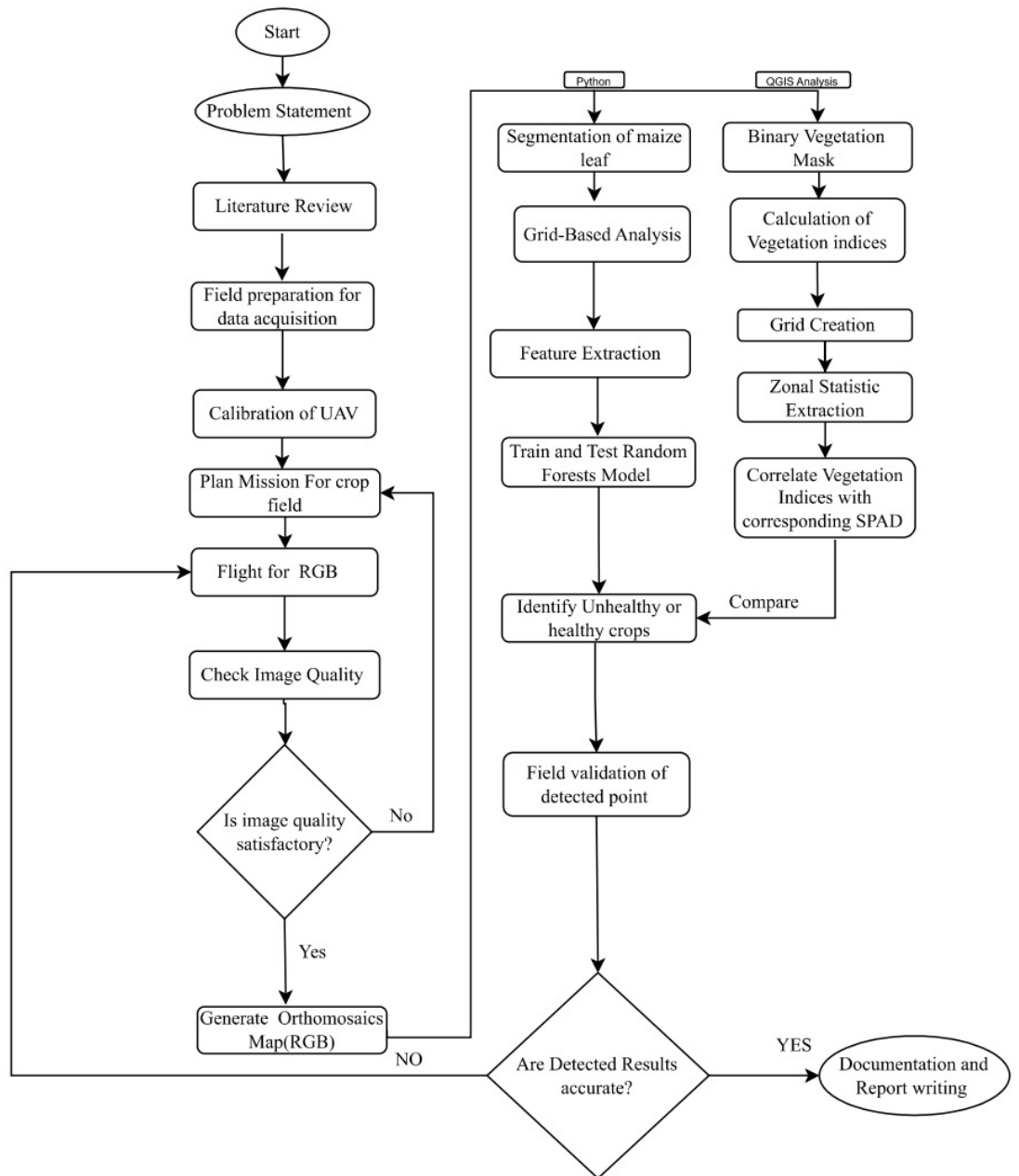


Figure 4.1: Detail flowchart of the project Methodology

The methodology adopted in this study was structured as a comprehensive, multi-stage integrating UAV-based aerial data acquisition, photogrammetric image processing, geospatial vegetation index analysis, machine learning-based classification, and rigorous ground-truth validation. The complete workflow, illustrated in Figure: 4.1, was systematically designed to ensure scientific reproducibility, spatial accuracy, and analytical robustness at every stage of the process. Each phase of the methodology was carefully sequenced to build upon the outputs of the preceding stage, establishing

an end-to-end pipeline from raw aerial imagery acquisition to the final identification and localization of nutrient sufficient and deficient zones within the experimental maize field.

4.1 Problem Identification and Literature Review

The methodology began with a detailed and careful review of existing scientific literature related to UAV-based precision agriculture, RGB vegetation indices, crop health monitoring, and chlorophyll-based methods for assessing plant nutrients. A systematic review of journal articles, conference papers, and technical reports was carried out to identify the most reliable and commonly used RGB based vegetation indices for monitoring crop health, especially in situations where multispectral sensors are not available. This extensive literature review helped in selecting the six vegetation indices used in this study: ExG, GLI, NGRDI, VARI, TGI, and NPCI. It also provided a strong theoretical and practical foundation for applying these indices to assess the nutrient status of maize crops. Furthermore, the review process made it possible to identify important research gaps in the Nepalese agricultural context, particularly the lack of well-tested RGB-based crop health monitoring methods that are suitable for local agro-climatic conditions. Recognizing these gaps played a key role in shaping the specific objectives and direction of this study.

4.2 Field Preparation and Experimental Setup

Before starting the aerial data collection campaigns, detailed field preparation work was carried out in close coordination with Nepal Agricultural Research Council (NARC) and GeoKrishi. The experimental maize field at Khairahani, Chitwan was prepared using standard agricultural practices in close coordination with the agricultural experts. This included multiple rounds of ploughing to create a fine and uniform soil structure, removal of crop residues and weeds, proper levelling of the field, and the application of well-decomposed farmyard manure. The experimental setup was designed using a Randomized Complete Block Design (RCBD). It consisted of 24 plots arranged in four replications, with six different nitrogen treatments. These treatments covered a wide range of nitrogen application methods, starting from a control plot with no nitrogen input to various combinations of urea, biochar, and organic nitrogen sources.

Using the RCBD design helped ensure that the effects of different treatments could be clearly distinguished from natural variations in the field, such as soil or environmental differences. This provided a strong and reliable scientific basis for comparing the performance of different vegetation indices with actual ground measurements collected during the study.

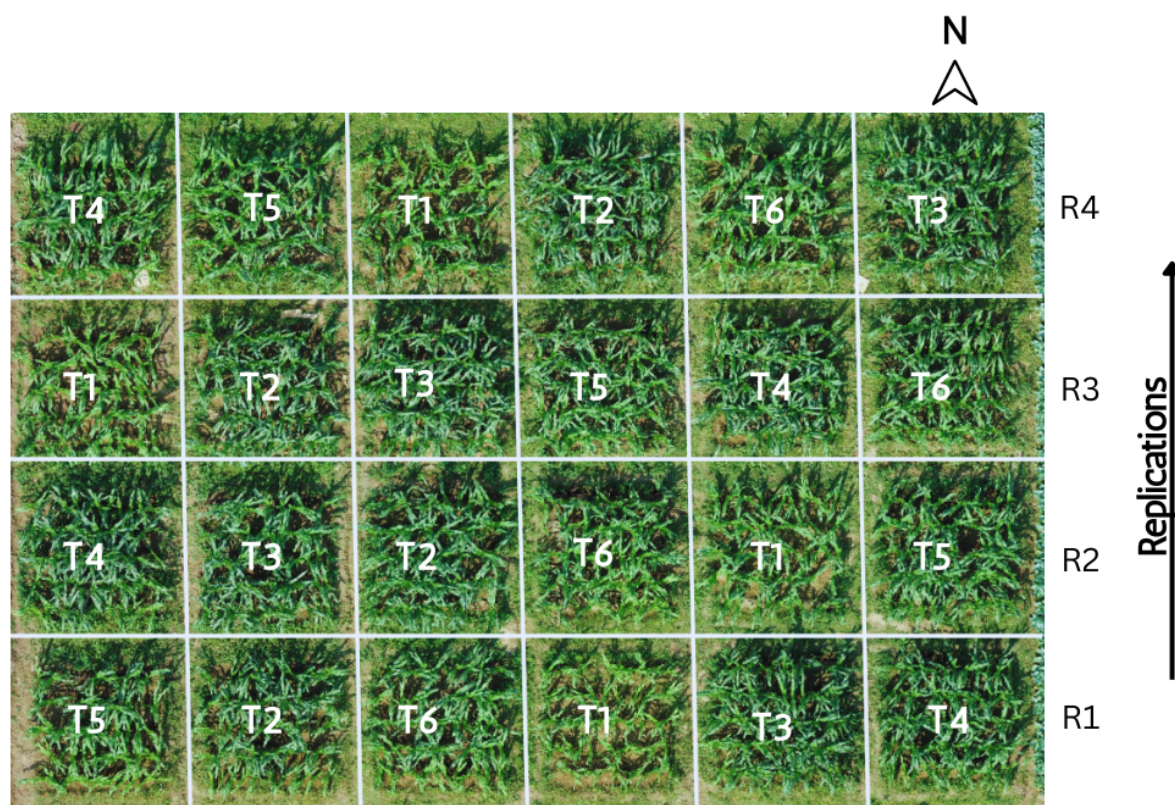


Figure 4.2: Field setup

4.2.1 Study Area

The experiment was conducted at the research farm of IAAS, Rampur Campus, Chitwan (213m AMSL) under subtropical conditions suitable for maize cultivation.

4.2.2 Experimental Details

The experiment was carried out in a Randomized Complete Block Design (RCBD) with six treatments and four replications. Each plot measured 3 m × 3 m with 1 m spacing between plots and blocks. Crop geometry of 60 cm × 25 cm was maintained.

4.2.3 Recommended Dose of Fertilizer (RDF)

The recommended fertilizer dose was 180 : 60 : 60 kg N:P₂O₅:K₂O ha⁻¹. Nitrogen (180 kg ha⁻¹) was applied through urea, with one-third (60 kg ha⁻¹) as basal dose and the remaining two-thirds (120 kg ha⁻¹) applied in two equal splits at knee height and tasseling stages.

4.2.4 Test Crop and Agronomic Practices

Rampur Hybrid-12 maize was used as the test crop. Two seeds per hill were sown and thinned to one plant at V2 stage(Stage of emergence of 2 leaves). Standard agronomic practices were followed uniformly across all plots.

4.3 UAV Acquisition and Familiarization

The DJI Phantom 4 Pro V2.0 quadcopter UAV platform was acquired through institutional collaboration with GeoKrishi, a leading precision agriculture technology organization based in Nepal, for the purpose of systematic RGB aerial image acquisition over the experimental maize field. Prior to conducting operational mapping flights, a comprehensive familiarization and training phase was undertaken to ensure the technical proficiency and operational readiness of the project team.

The familiarization program encompassed a series of progressive training flights conducted at an open test site, systematically advancing from basic manual flight operations to fully autonomous mission execution. Training activities included manual flight mode operation under varying wind and lighting conditions, autonomous waypoint-based mission planning using the DJI GO 4 application, pre-flight checklist procedures, in-flight monitoring of telemetry parameters including battery status, altitude, and GPS signal integrity, as well as emergency procedure drills covering return-to-home activation, forced landing protocols, and signal loss response. Basic hardware troubleshooting, including propeller inspection, gimbal calibration, compass calibration, and IMU initialization, was also practiced systematically during this phase.

This structured familiarization period was instrumental in developing the team's competence in UAV handling, mission planning optimization, and real-time flight decision-making under field conditions. The experience gained during this phase directly contributed to the efficient and safe execution of subsequent data acquisition campaigns over the agricultural field at the NARC facility, Chitwan, where consistent flight parameters, image quality, and spatial coverage were critical to the integrity of the aerial dataset.

4.4 UAV Calibration and System Preparation

Before starting any flight mission, detailed pre-flight calibration and system checks were carried out to ensure that the UAV system was fully ready and capable of collecting accurate data. The DJI Phantom 4 pro V2.0 quadcopter was carefully inspected, including checking the condition of the propellers, calibrating the gimbal for stable imaging, examining the electronic speed controllers, assessing battery health and ca-

capacity, and inspecting the overall structure of the drone.

The onboard RGB camera was also checked thoroughly. This included ensuring the lens was clean, adjusting exposure settings based on current lighting conditions, and setting the white balance properly to maintain consistency in image quality and color across all captured data.

In addition, essential system calibrations such as compass calibration, Inertial Measurement Unit (IMU) initialization, and GPS signal verification were performed before every flight. These steps were necessary to ensure accurate positioning, stable flight performance, and reliable autonomous operation during each mission.

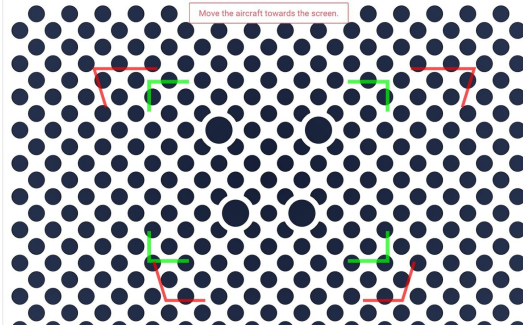


Figure 4.3: Calibration image for camera.

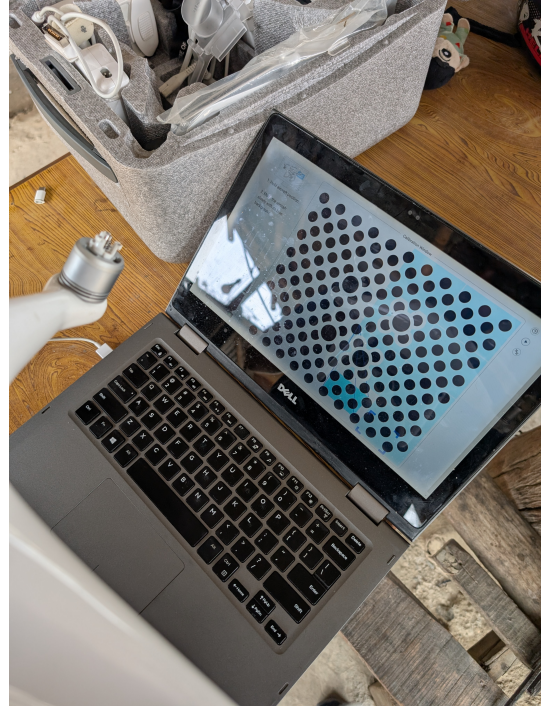


Figure 4.4: Camera Calibration.

4.5 Mission Planning

Before each data collection campaign, autonomous flight missions were carefully planned and set up using the DroneDeploy mission planning software. The flight settings were thoughtfully adjusted for each field, taking into account its size, shape, surrounding landscape, nearby obstacles, and the current weather conditions. The UAV was programmed to fly in a systematic back and forth zigzag pattern, similar to mowing a lawn, which ensured that the entire field was covered completely and evenly without leaving gaps, and without needing manual control during the flight.

The overlap between images was always set to high levels in both the forward flight direction and the side-to-side direction(i.e 80% front and side overlap). This provided

plenty of redundant image data, which is important for creating accurate 3D models and seamless orthomosaic maps. All flights were scheduled around midday when sunlight was strongest. This minimized the presence of shadows that could distort the images and helped maintain consistent brightness and color across all photographs.

It is also important to note that this study involved mapping multiple field sites, each in a different location, with varying sizes, crop types, and terrain features. Because of these differences, the flight altitude of the UAV was adjusted for each site. These adjustments were based on the field's characteristics, safety requirements, and the desired level of detail needed to clearly see important crop features. By customizing the flight settings for each location, we ensured that the images from all field sites were consistent in terms of quality and spatial resolution, despite the differences in size, crop type, or terrain.

4.5.1 Path Planning and Grid Creation

Before the drone takes off, its flight path is planned to make sure it covers the whole agricultural field. The field is split into evenly spaced strips, and the drone flies back and forth in a zig-zag pattern along these strips.

During the flight, the drone keeps a steady altitude and speed, taking pictures with about 80% overlap from front to back and 80% from side to side.

During the flight, the UAV mounted RGB and thermal cameras capture images simultaneously along the planned grid path, with timestamps and GPS coordinates logged.

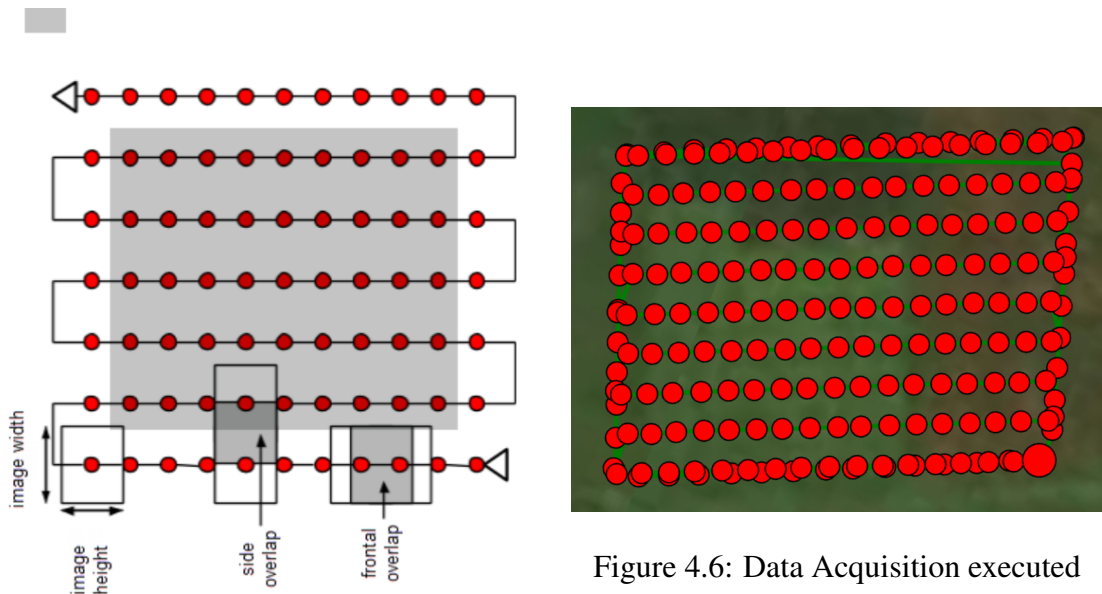


Figure 4.5: Image Acquisition plan

Figure 4.6: Data Acquisition executed

4.5.2 Acquisition of RGB Images

Aerial image collection was carried out autonomously, following the pre-planned mission created in advance. The DJI Phantom 4 UAV maintained steady altitude, speed, and heading throughout each flight, ensuring precise and consistent data capture. Its on-board camera(20MP) continuously took high-resolution photographs at carefully timed intervals along the designated flight path. This systematic approach allowed the UAV to cover the entire field completely, producing dense and overlapping image coverage for accurate mapping. Each photograph automatically recorded its GPS location and a synchronized UTC timestamp within the EXIF metadata. This ensured that every image was accurately georeferenced without needing any external ground control points.

The study included multiple mapping campaigns conducted at different field locations. The process began with familiarization flights at open, easily accessible sites and gradually progressed to full-scale operational mapping over more complex experimental fields. Each campaign generated a large collection of georeferenced, high-resolution RGB images. Together, these images formed the primary raw dataset used for later photogrammetric processing and detailed geospatial analysis.

Key flight parameters include:

- i. Altitude: Depending on sensor resolution and desired Ground Sample Distance (GSD).
- ii. Image Overlap: 80% frontlap and 80% sidelap to support orthomosaic reconstruction
- iii. Flight Time: Scheduled between 12:00 AM and 2:00 PM to ensure consistent solar irradiance and accurate multispectral readings.

Each image is tagged with the exact captured time and GPS location. This ensures that every image can be matched to its location in the field.

4.6 Controlled field Environment

This image illustrates a Randomized Complete Block Design (RCBD) for a field experiment aimed at evaluating different nitrogen and biochar fertilization strategies. Later these data are used for field validation of UAV derived results.



Figure 4.7: Maize Field

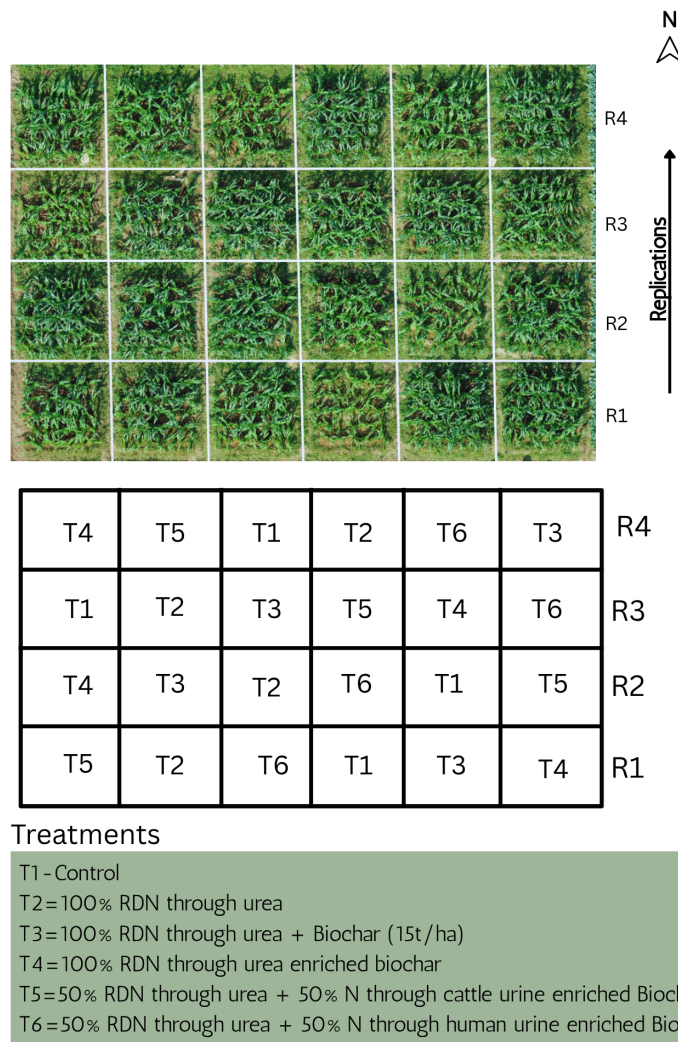


Figure 4.8: Field Setup

Data Processing

4.7 Orthomosaic Map Generation

After each flight, image datasets that passed quality checks were processed in DroneDeploy to generate georeferenced, radiometrically consistent orthomosaic maps. The processing workflow included: automatic image alignment and tie point generation via SfM, reconstruction of a dense 3D point cloud using multi-view stereo, and creation of a digital surface model (DSM) capturing field topography and crop canopy structure. Orthomosaics were produced by geometrically projecting and stitching all images onto a common ground plane, with corrections applied for camera lens distortions, terrain displacement, and perspective differences to ensure spatial and photometric accuracy. For each site, the orthomosaic was exported as a four-band GeoTIFF (UTM Zone 45N, EPSG:32645), serving as the primary dataset for all subsequent geospatial analyses and vegetation index calculations in QGIS.

4.8 Image Segmentation using python

To separate the leaf regions of maize from field photos taken at a height of three meters, a unique multi-evidence picture segmentation pipeline was created in Python using OpenCV. To address outdoor light fluctuation and visual resemblance between maize and surrounding vegetation, four distinct evidence streams vegetation spectral indices, HSV color gating, texture smoothness, and CIE L*a*b* chromaticity are combined rather than depending on a single threshold. Every image generates two outputs: a binary segmentation mask (`_mask.png`) with values of 0 (background) or 255 (maize) and a color-masked image (`_maize.png`) where verified maize pixels maintain their original color and all background is changed to pure black (RGB: 0,0,0).

4.8.1 Vegetation Spectral Indices

Three indices are computed pixel-wise on float32 R, G, B arrays:

$$ExG = 2G - R - B \quad (4.1)$$

$$VARI = \frac{G - R}{G + R - B} \quad (4.2)$$

$$GLI = \frac{2G - R - B}{2G + R + B} \quad (4.3)$$

Before a pixel is kept under a majority voting rule, at least two of the three indices must surpass their threshold. This lowers the number of false positives from any one index

reacting to surfaces that aren't vegetated, such moist soil.

4.8.2 Stream 2: HSV Color Gate

The HSV color space separates color information (Hue and Saturation) from intensity (Value), making it robust for outdoor maize segmentation under varying lighting. Green maize leaves fall within the hue range 22-88, while near-grey soil and other background elements are removed using a saturation lower bound of 35. Very dark areas, such as shadows, are excluded with a minimum value of 90. These thresholds were chosen empirically from field images to balance sensitivity and specificity. Applying the HSV gate reduces false positives while retaining most maize leaf pixels. The resulting mask is then combined with outputs from spectral indices and texture smoothness streams to improve segmentation accuracy.

Table 4.1: HSV Thresholds used for maize leaf segmentation

Channel	Lower	Upper
Hue (H)	22	88
Saturation (S)	35	220
Value (V)	90	255

4.8.3 Texture Smoothness

Maize leaf blades are generally smoother and more uniform than surrounding grass and soil, which tend to have rough and irregular textures. To capture this difference, local texture is analyzed using the standard deviation of grayscale intensity within a 15×15 pixel sliding window.

$$\text{local_std} = \sqrt{E[x^2] - (E[x])^2} \quad (4.4)$$

This measure quantifies the variation in pixel intensity within the window. Lower values indicate uniform regions (smooth surfaces like maize leaves), while higher values correspond to rough and heterogeneous areas such as soil and grass. The selected window size balances noise reduction and spatial detail preservation. Based on empirical tuning, pixels with $\text{local_std} < 28$ are classified as smooth and retained. This step helps eliminate high-texture background regions, thereby improving segmentation accuracy and reducing false positives in maize detection.

4.8.4 L*a*b* Green Axis

Positive values are red and negative values are green when the a^* channel is extracted $a^* = 128$. Pixels that meet $a^* < -8$ are kept. This stage is robust to uneven field illumination because it is luminance separated and can identify green tissue regardless of brightness.

4.8.5 Multi-Evidence Fusion

All four binary masks are combined using bitwise AND:

$$\text{combined} = \text{index_mask} \cap \text{HSV_mask} \cap \text{Lab_mask} \cap \text{texture_mask}$$

In order to maximize precision while allowing for a slight recall reduction at leaf boundaries, a pixel is only classified as maize if all four evidence streams agree at the same time.

4.8.6 GrabCut Refinement

The filtered mask initialises GrabCut (`cv2.grabCut`), which fits Gaussian Mixture Models to foreground and background distributions over 5 iterations to refine boundaries at sub-pixel precision. Pixels with HSV Value < 40 are hard assigned as definite background (GC_BGD). If GrabCut fails on a degenerate image, the Stage 8 mask is used directly as a fallback.

4.8.7 Spatial Grid Division

Each segmented image is divided into a 8×8 grid, producing $8 \times 8 = 64$ tiles per image. The tile dimensions are computed dynamically from the image resolution:

$$\text{tile_width} = \frac{\text{image_width}}{8}, \quad (4.5)$$

$$\text{tile_height} = \frac{\text{image_height}}{8}. \quad (4.6)$$

For every tile, computation is limited to the non-black pixels in that tile region. Tiles that fall completely on the background or have fewer than a certain number of valid pixels are rejected and excluded from the feature set. By doing this, feature values are guaranteed to reflect real maize tissue.

Three machine learning classifiers, including a Random Forest model, are created and assessed for the binary classification of maize plots as nutritionally sufficient (label = 1) or deficient (label = 0) after spectral characteristics are extracted from segmented maize images. Vegetation indices calculated from verified maize pixels are used for classification at the pixel level. Since undetected nutrient deficiencies have immediate agronomic consequences, the deficient class is always considered the critical class.

4.8.8 Input Features

Four vegetation indices calculated from each segmented image’s non-black pixels make up the input feature vector for each pixel. Although they are not used directly as classifier inputs, raw R , G , and B channel values are recorded for statistical reference.

All models use the following four features: Here, $\varepsilon = 1 \times 10^{-6}$ is added to denominators to prevent division by zero. Each pixel produces one feature vector

$$[\text{ExG}, \text{VARI}, \text{NGRDI}, \text{GLI}]$$

with its corresponding class label (nutritionally sufficient or deficient).

4.8.9 Train/Test Split

The full pixel dataset is randomly shuffled and divided using an 80/20 split:

- **Training set:** 80% of all pixels, used for model fitting.
- **Test set:** 20% of all pixels, held out exclusively for evaluation.

No data augmentation is applied at this stage, as the pixel-level dataset is sufficiently large based on the available images.

Table 4.2: Random Forest Hyperparameters and Justifications

Parameter	Value	Justification
n_estimators	500	Large ensemble reduces variance
max_depth	None	Trees grown fully; pruning via leaf constraints
min_samples_split	4	Prevents splits on very small nodes
min_samples_leaf	2	Minimum leaf size for generalisation
max_features	sqrt	Standard for classification tasks
class_weight	balanced	Compensates for class imbalance automatically
random_state	42	Reproducibility
n_jobs	-1	All available CPU cores used

4.9 Tile-Based Prediction for New Images

For prediction on new images, each image is divided into a 30×30 spatial grid. Tile dimensions are computed dynamically as:

$$\text{tile_width} = \frac{\text{image_width}}{30}, \quad \text{tile_height} = \frac{\text{image_height}}{30}$$

For each tile, black pixels are excluded and the mean of each index [ExG, VARI, NGRDI, GLI] is computed across valid pixels. Tiles where all pixels are black are assigned NaN and dropped before prediction. The trained Random Forest model predicts a binary label per tile. The image-level classification is determined by the mean of all tile predictions:

$$\text{Final prediction} = \begin{cases} \text{Deficient} & \text{if } \bar{y}_{\text{tiles}} < 0.5 \\ \text{Sufficient} & \text{if } \bar{y}_{\text{tiles}} \geq 0.5 \end{cases}$$

4.10 Evaluation Metrics

All models were evaluated on the held-out 20% test set using the following metrics:

- i. **Confusion matrix** - visualizes true positives, true negatives, false positives, and false negatives for each class.
- ii. **Precision (weighted)** - proportion of positive predictions that are correct.
- iii. **Recall (weighted)** - proportion of actual positives correctly identified.
- iv. **F1 score (weighted)** - harmonic mean of precision and recall, balancing both metrics.
- v. **Classification report** - per-class breakdown of precision, recall, F1, and support.

All graphs and confusion matrices for the machine learning models are generated and reported in the output.

4.11 Output

4.11.1 Confusion Matrix

From this Confusion matrix:

- 108 samples were correctly classified as deficient (True Positives)
- 89 samples were correctly classified as sufficient (True Negatives)

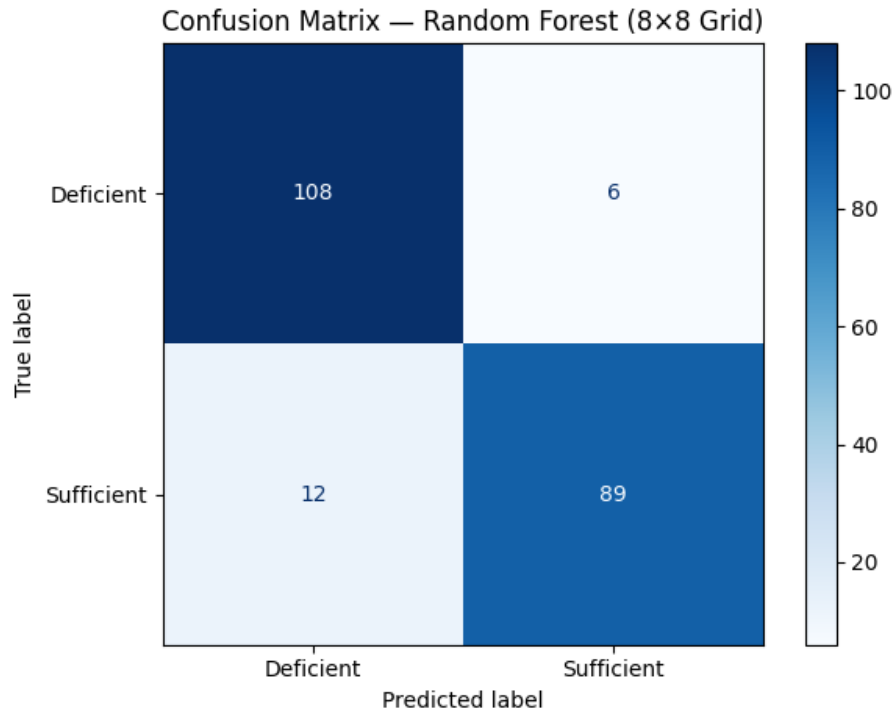


Figure 4.9: Confusion Matrix

- 6 samples of deficient class were misclassified as sufficient (False Negatives)
- 12 samples of sufficient class were misclassified as deficient (False Positives)

The total number of test samples was 215, consisting of 114 deficient and 101 sufficient samples. Our model achieved an overall accuracy of 91.6%, indicating high prediction correctness. The precision for the deficient class is 90%, meaning that most predicted deficient samples were accurate. The recall was 94.7%, showing that the model successfully identified the majority of actual deficient plants. The F1-score of 92.3% reflects a good balance between precision and recall.

4.11.2 Feature Importance

Three indices ExG, VARI, and NGRDI that enhance plant greenness and minimize lighting impacts were computed in order to extract vegetation properties from RGB photos. A total of 192 features that capture both spectral and spatial information were produced by dividing each image into an 8x8 grid and calculating the mean value of each index within each cell. The importance of each characteristic was then assessed using a machine learning model, which showed that some features contribute to predictions more strongly than others. The findings indicate that regional vegetation patterns are more significant than general averages, with some grid sections being more informative than others.

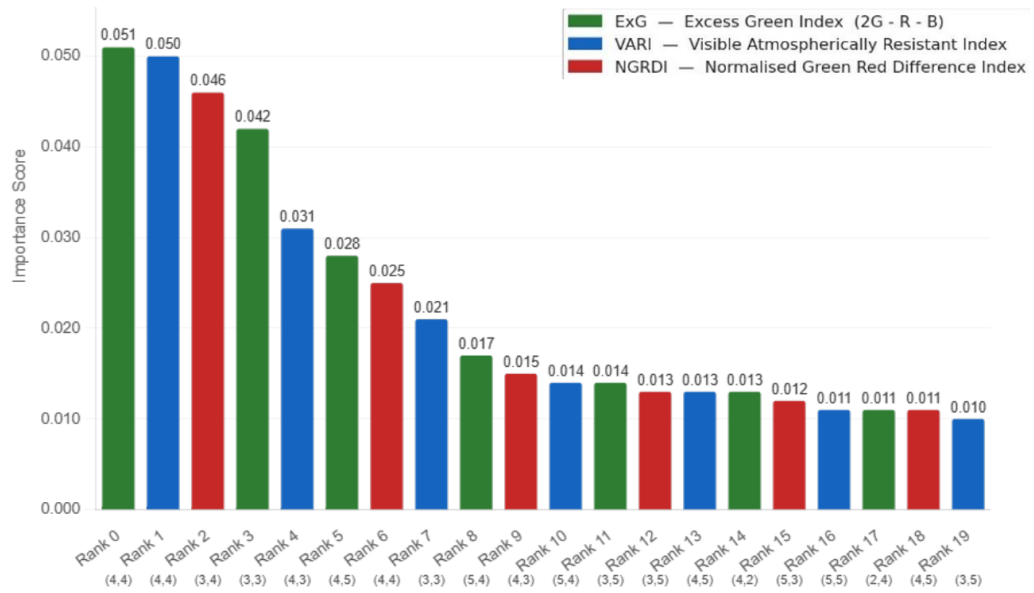


Figure 4.10: Feature Importance

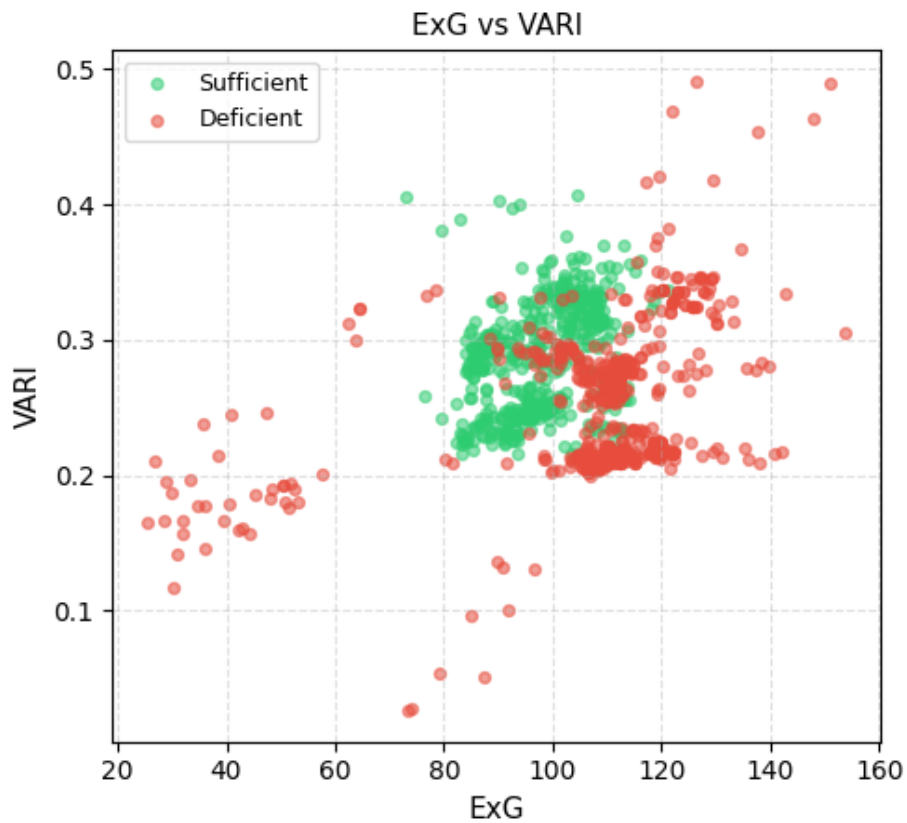


Figure 4.11: Scatter Plot : ExG vs VARI

The two vegetation indicators, ExG and VARI , are compared in this scatter plot to demonstrate their relationship to plant health, which is classified as "Sufficient" (green dots) or "Deficient" (red dots). The majority of sufficient, healthy plants cluster together

in a middle range of values, with VARI between 0.28 and 0.35 and ExG between 90 and 110. This implies that the greenness of these plants is stable and healthy. The deficient plants, however, are more dispersed. Some look dispersed with typically lower values, indicating that their vegetation may be stressed or less vigorous, while others show greater ExG values but lower VARI scores.

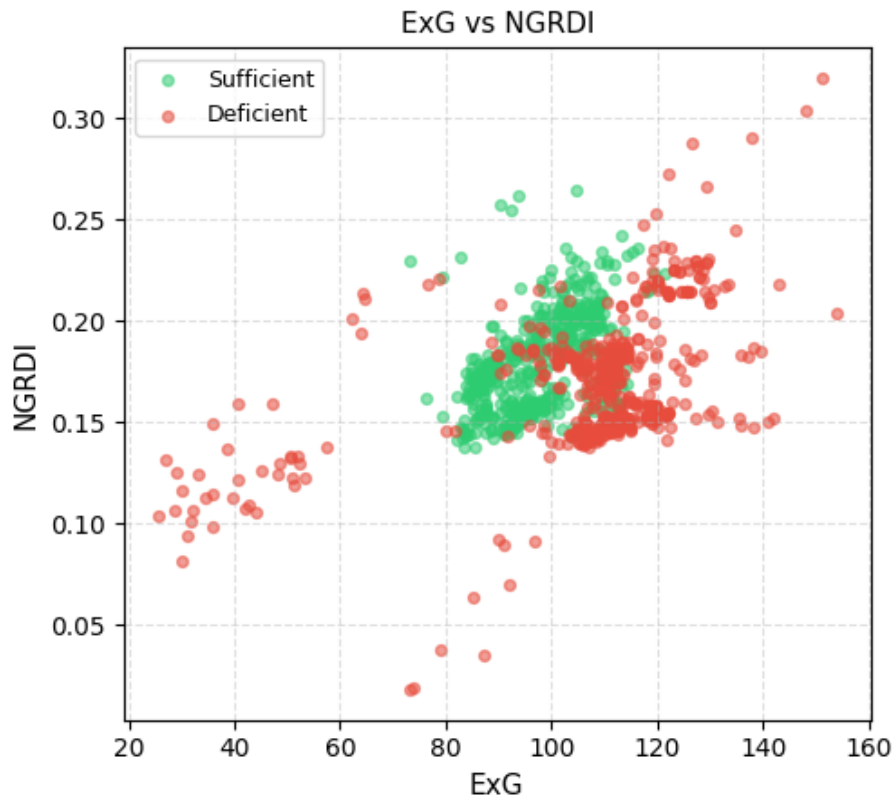


Figure 4.12: Scatter Plot : ExG vs NGRDI

This scatter plot, which is separated into "Sufficient" (green dots) and "Deficient" (red dots) groups, contrasts two vegetation indices ExG and NGRDI to show how they relate to plant health. The healthier plants tend to cluster in the intermediate range, with NGRDI values between 0.15 and 0.22 and ExG values between 90 and 110. This indicates that the green and red reflectance of these plants is well-balanced, a characteristic of healthy vegetation. The deficient plants, on the other hand, are spread out and frequently demonstrate higher ExG values above 110, but their NGRDI values are more variable and occasionally lower, suggesting that they may be under stress or less active.

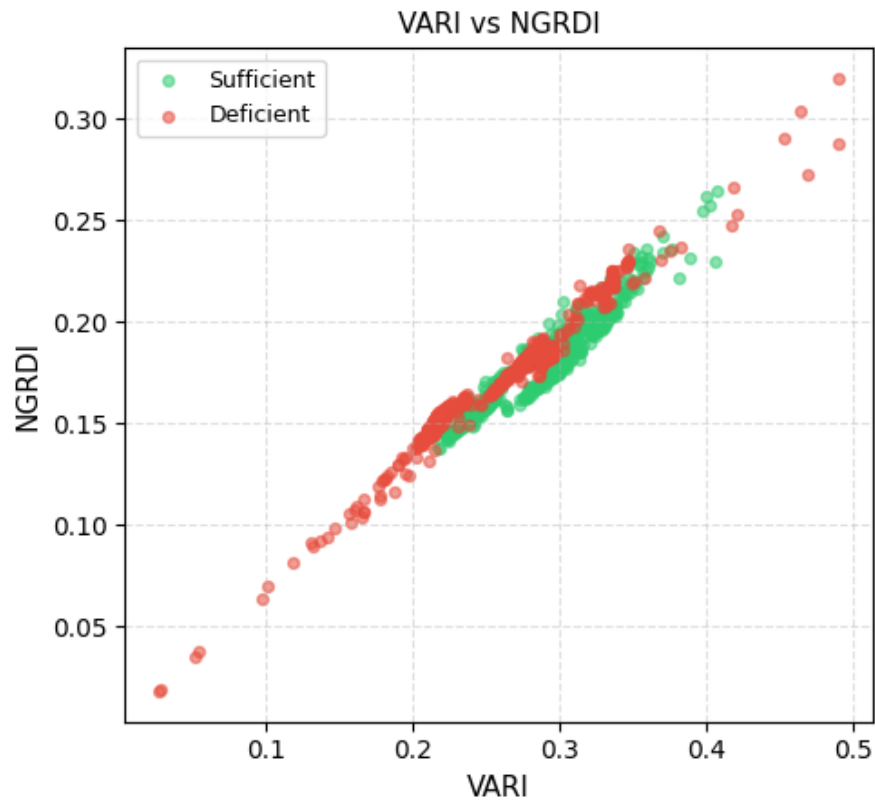


Figure 4.13: Scatter Plot : VARI vs NGRDI

This graph shows a scatter plot illustrating the relationship between the Visible Atmospheric Resistant Index (VARI) and the Normalized Green-Red Difference Index (NGRDI) across plant samples classified as either nutritionally sufficient or deficient. A strong positive linear correlation between VARI and NGRDI, indicating that both indices capture similar spectral characteristics of vegetation. However, despite this correlation, the scatter plot demonstrates a considerable overlap between the sufficient and deficient sample groups across the majority of the value range, particularly within VARI values of 0.15 to 0.35.

4.12 Index calculation Using QGIS

QGIS stands for Quantum Geographic Information System. It is one of the most widely used GIS platforms in both academic research and professional practice, offering a comprehensive suite of raster and vector processing tools accessible through an intuitive graphical interface as well as through Python scripting. Different indices were calculated using QGIS(free, open-source Geographic Information System (GIS) application that allows users to create, edit, visualize and analyze geospatial data). QGIS is maintained by the Open Source Geospatial Foundation (OSGeo).

Following the acquisition of the orthomosaic map from DroneDeploy, the data was imported into QGIS for advanced geospatial analysis. To ensure the accurate calculation of vegetation indices, a series of standardized processing steps were executed in chronological order.

4.12.1 Data Preparation

The orthomosaic generated from aerial image processing was loaded into QGIS 3.x as a four-band GeoTIFF file, with spectral bands assigned as Red (Band 1), Green (Band 2), Blue (Band 3), and Alpha (Band 4). The coordinate reference system was verified as UTM Zone 45N (EPSG:32645), with an approximate ground sampling distance (GSD) of 2–3 cm per pixel, covering around 1882 m² experimental maize field at, Chitwan, Nepal.

4.12.2 Vegetation Segmentation and Canopy Masking

To ensure that vegetation indices were computed exclusively from maize canopy pixels, a canopy segmentation workflow was implemented in QGIS prior to index calculation. The Normalized Excess Green Index (ExG_{norm}) was first computed using the QGIS Raster Calculator. To eliminate sensitivity to illumination variation and image brightness, the normalized formulation was adopted.

A binary vegetation mask was subsequently generated by applying a threshold of zero to the normalized ExG output, classifying pixels as vegetation (value = 1) where $ExG_{norm} > 0$, and non-vegetation (value = 0) elsewhere. Morphological cleanup was applied using the Sieve tool with a minimum cluster size of 5–10 pixels to remove isolated noise pixels and spurious classifications. The original four-band GeoTIFF was then clipped using the *Clip Raster by Mask Layer* tool, retaining only pixels classified as maize canopy, with all remaining pixels assigned as NoData to prevent their inclusion in subsequent analyses.



Figure 4.14: Unmasked Orthomosaic

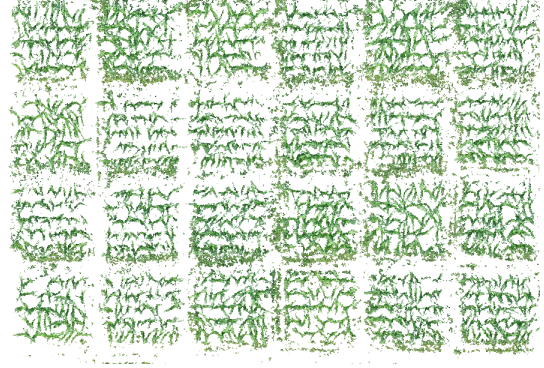


Figure 4.15: Masked Orthomosaic

4.12.3 Vegetation Index Calculation

Normalized Excess Green Index (ExG)

The Normalized Excess Green Index was originally developed for vegetation segmentation. Its standard formulation is:

$$\text{ExG} = 2G - R - B \quad (4.7)$$

The threshold of $\text{ExG} > 20$ was used to separate the soil pixel and maize canopy, maize canopy being 1 and soil pixel being 0. The normalized form adopted in this study eliminates sensitivity to overall image brightness:

$$\text{ExG}_{\text{norm}} = \frac{2G - R - B}{R + G + B + 0.0001} \quad (4.8)$$

Values range from -1 to $+1$, where positive values indicate green-dominant vegetation pixels and negative values correspond to soil or non-vegetated surfaces.

Green Leaf Index (GLI)

The Green Leaf Index uses all three visible bands to quantify green vegetation. To prevent division by zero, the formula implemented in QGIS was:

$$\text{GLI} = \frac{2G - R - B}{2G + R + B + 0.0001} \quad (4.9)$$

GLI values range from -1 to $+1$. Higher positive values indicate healthy green vegetation, while negative values correspond to bare soil or non-vegetated areas. Later, the layer GLI layer was masked only for maize pixel using the threshold ExG layer(i.e Maize = 1, Soil = 0). First the nodata value of 0 was assigned. Then the final output for

maize showed GLI index for maize but nodata for soil pixels.

Normalized Green-Red Difference Index (NGRDI)

The Normalized Green-Red Difference Index also referred to as Pseudo-NDVI, quantifies green dominance relative to red reflectance.

The formula implemented in QGIS incorporating the zero-division safeguard is:

$$\text{NGRDI} = \frac{G - R}{G + R + 0.0001} \quad (4.10)$$

NGRDI values range from -1 to $+1$. Positive values indicate vegetation where green reflectance exceeds red, while values near zero or negative correspond to soil and non-vegetated surfaces.

Visible Atmospherically Resistant Index (VARI)

VARI was developed by to estimate green vegetation fraction from RGB imagery while partially correcting for atmospheric effects through inclusion of the blue band in the denominator. Since the denominator ($G + R - B$) can equal zero at certain pixel combinations, the following modified formula was implemented in QGIS:

$$\text{VARI} = \frac{G - R}{G + R - B + 0.0001} \quad (4.11)$$

VARI values typically range from -1 to $+1$. Higher positive values correspond to healthy dense vegetation, while lower or negative values indicate bare soil or stressed vegetation. Among all tested indices, VARI demonstrated the strongest correlation with ground-truth SPAD chlorophyll measurements ($r = 0.73$,

Triangular Greenness Index (TGI)

TGI was developed by based on the triangular area formed by spectral reflectance at 670 nm, 550 nm, and 480 nm, approximated using RGB bands as:

$$\text{TGI} = G - 0.39R - 0.61B \quad (4.12)$$

Unlike normalized indices, TGI is expressed in raw pixel intensity units when applied to 8-bit RGB imagery (band values (0–255)). No zero-division risk exists in this formulation, so the formula was implemented directly in QGIS without modification.

Positive TGI values indicate the presence of green vegetation. For healthy maize canopy

with raw 8-bit pixel values, typical TGI values range from 30 to 80. In this study, TGI values ranged from 50.92 to 57.67 across all 24 experimental plots, consistent with healthy maize at the reproductive to maturation growth stage.

Normalized Pigment Chlorophyll Index (NPCI)

NPCI was originally formulated by using hyperspectral reflectance at 680 nm and 430 nm as:

$$\text{NPCI} = \frac{R_{680} - R_{430}}{R_{680} + R_{430}} \quad (4.13)$$

Since hyperspectral bands are unavailable in standard RGB sensors, an RGB approximation was adopted in this study, substituting the broadband red and blue channels:

$$\text{NPCI}_{\text{RGB}} = \frac{R - B}{R + B + 0.0001} \quad (4.14)$$

NPCI values range from -1 to $+1$. Unlike other greenness indices, NPCI exhibits an *inverse* relationship with chlorophyll content higher NPCI values correspond to greater carotenoid-to-chlorophyll ratios indicative of nitrogen stress, while lower values indicate higher chlorophyll concentration and healthier vegetation. In this study, RGB-approximated NPCI values ranged from 0.035 to 0.108 across all 24 plots.

4.12.4 Symbology and Visualization

Following the computation of each vegetation index, pseudocolor symbology was applied in QGIS to enable spatial visual interpretation of index value distribution across the experimental maize field. For each index raster layer, the render type was configured as *Singleband Pseudocolor* under *Layer Properties* $\rightarrow$ *Symbology*. The *RdYlGn* (Red-Yellow-Green) diverging color ramp was applied, where red tones represent low index values associated with stressed or non-vegetated areas, yellow represents intermediate values, and green represents high index values characteristic of healthy dense vegetation. The minimum and maximum color stretch values were defined based on the actual data range of each index as determined from the layer histogram. This visualization approach enabled direct spatial comparison of crop health patterns across all 24 experimental plots and facilitated identification of nutrient sufficient and deficient zones within the field.

Table 4.3: Symbology settings applied to each vegetation index layer in QGIS

Index	Render Type	Color Ramp	Value Range
GLI	Singleband Pseudocolor	RdYlGn	−1 to +1
NGRDI	Singleband Pseudocolor	RdYlGn	−1 to +1
VARI	Singleband Pseudocolor	RdYlGn	−1 to +1
TGI	Singleband Pseudocolor	RdYlGn	30 to 80
NPCI	Singleband Pseudocolor	RdYlGn	−1 to +1

4.12.5 Plot Boundary Setup

The experimental plot boundaries were loaded into QGIS as a GeoPackage (.gpkg) vector layer containing 24 polygon features, each corresponding to one experimental plot in the RCBD layout. The plot boundary layer was overlaid on both the orthomosaic and computed vegetation index rasters to visually confirm spatial alignment and correct georeferencing. The coordinate reference system of the vector layer was verified to match all raster layers (UTM Zone 45N, EPSG:32645), ensuring accurate spatial correspondence during subsequent zonal statistics extraction. Later each plot was divided into 16 subplots for better visualization of indices. The average value of each plot was taken into the calculation.

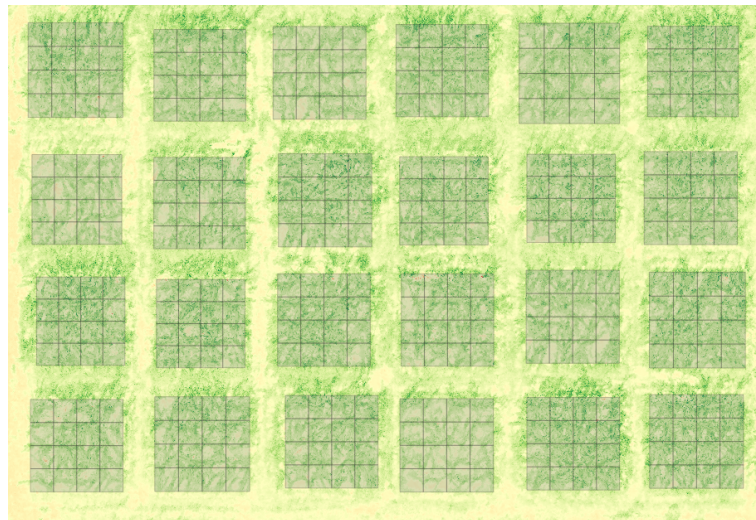


Figure 4.16: Grid creation in QGIS.

4.12.6 Zonal Statistics Extraction

Mean vegetation index values for each plot were extracted using the *Zonal Statistics* tool in QGIS (*Raster* → *Zonal Statistics*). For each index, the corresponding raster layer was used as input, and the RCBD plot boundaries (24 plots) were used as the zonal layer. The *Mean* value was selected to represent the average index value within

each plot. Separate column prefixes (such as ExG_, GLI_, NGRDI_, VARI_, TGI_, and NPCI_) were used to keep the results organized and avoid confusion in the attribute table. After processing all six indices, the attribute table contained 24 rows (one for each plot) along with the corresponding mean index values. The final table was then exported as a CSV file for further statistical analysis in JMP Pro.

4.12.7 Data Export and Preparation for Analysis

After completing zonal statistics and canopy coverage calculations, the plot attribute table was exported from QGIS in CSV format using *Layer → Export → Save Features As*. The dataset contained 24 rows representing the experimental plots, along with columns for plot ID and mean values of vegetation indices (ExG, GLI, NGRDI, VARI, TGI, and NPCI). SPAD chlorophyll readings collected during the UAV survey were then added to this dataset using Microsoft Excel. The final combined dataset was imported into JMP Pro for further analysis, including correlation, classification, and visualization.

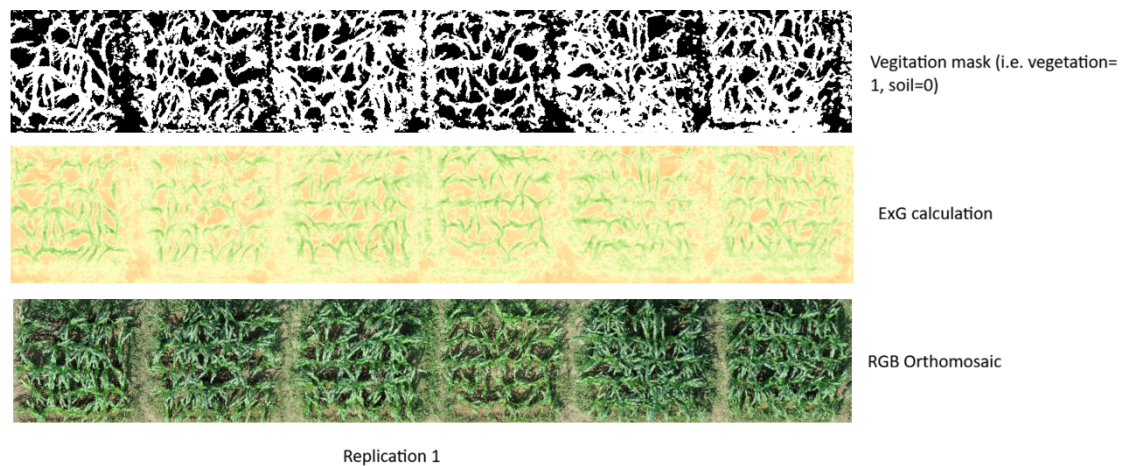


Figure 4.17: Masking in steps.

4.13 Data Analysis

A total of 240 SPAD observations were collected from six nutrient treatments (T1–T6), each with four replications (R1–R4), where 10 SPAD readings were recorded per treatment plot. These observations were used to evaluate the effect of different nutrient treatments on crop chlorophyll content.

One-way Analysis of Variance (ANOVA) was performed to determine the significance of differences in SPAD values among the treatments at a 0.05 probability level. Prior to ANOVA, the assumptions of normality and homogeneity of variance were checked to

ensure the validity of the analysis.

For correlation and regression analysis, SPAD values were compared with vegetation indices (VIs) derived from RGB images. Pearson correlation analysis was conducted to quantify the strength and direction of the relationship between SPAD values and individual vegetation indices.

All statistical analyses, including ANOVA, correlation, and regression, were performed using JMP Pro statistical software (SAS Institute Inc.).

To assess the predictive performance of the regression models, Root Mean Square Error (RMSE) and Normalized Root Mean Square Error (NRMSE) were calculated as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (4.15)$$

$$NRMSE = 100 \times \frac{RMSE}{\frac{1}{n} \sum_{i=1}^n y_i} \quad (4.16)$$

where y_i is the observed SPAD value of sample i , $\hat{y}_i$ is the predicted SPAD value, and n is the total number of samples used for validation. RMSE provides the absolute prediction error, while NRMSE expresses the error as a percentage, allowing comparison across different scales. [4]

Chapter 5: RESULT AND DISCUSSION

5.1 SPAD Values for Different Fertilizer Treatments

The measured SPAD values for each experimental plot across the six nutrient treatments (T1-T6) are presented to evaluate the impact of organic and inorganic fertilization on maize chlorophyll content. Following the methodology used in previous barley studies to detect the significance of fertilization levels, a one-way Analysis of Variance (ANOVA) was conducted.[4]

The results indicated that the type of nutrient treatment had an extremely significant effect on leaf chlorophyll content ($F = 24.3038$, $p < 0.0001$). The total range of SPAD readings across the 240 observations spanned from approximately 34.90 to 65.40 .[4]

Table 5.1: Vegetation Indices and SPAD Measurements

Sample	SPAD Mean	VARI Mean	GLI Mean	NGRDI Mean	NPCI Mean	TGI Mean
R1T1	38.8530	0.2648	0.2273	0.1803	0.1081	57.6737
R2T1	45.7633	0.2985	0.2248	0.1957	0.0662	55.8501
R3T1	39.8433	0.2828	0.2260	0.1885	0.0853	57.3833
R4T1	42.5967	0.2916	0.2302	0.1941	0.0820	57.1749
R1T2	52.5833	0.3015	0.2254	0.1979	0.0635	52.8541
R2T2	51.3533	0.3416	0.2361	0.2213	0.0353	51.9052
R3T2	54.7333	0.3263	0.2362	0.2143	0.0516	52.2915
R4T2	51.2633	0.3260	0.2290	0.2109	0.0418	50.9221
R1T3	53.1333	0.3375	0.2432	0.2225	0.0493	53.2324
R2T3	52.0833	0.3434	0.2398	0.2239	0.0386	52.2286
R3T3	53.3700	0.3313	0.2367	0.2171	0.0459	52.1154
R4T3	51.0133	0.3216	0.2239	0.2063	0.0402	51.8250
R1T4	50.5100	0.3402	0.2431	0.2235	0.0470	53.1296
R2T4	51.3967	0.3226	0.2378	0.2128	0.0583	52.0843
R3T4	50.7567	0.3121	0.2258	0.2030	0.0523	52.3518
R4T4	45.9400	0.3232	0.2354	0.2122	0.0546	54.1317
R1T5	45.7167	0.3025	0.2338	0.2014	0.0751	56.7728
R2T5	49.0067	0.3372	0.2391	0.2201	0.0448	54.2488
R3T5	47.5667	0.3255	0.2360	0.2139	0.0518	54.5634
R4T5	45.5200	0.3246	0.2402	0.2142	0.0605	55.9986
R1T6	48.7767	0.3122	0.2334	0.2061	0.0635	55.4293
R2T6	43.9967	0.3196	0.2422	0.2127	0.0691	56.8164
R3T6	48.6633	0.3105	0.2301	0.2044	0.0593	54.4981
R4T6	47.0167	0.3052	0.2246	0.1988	0.0588	55.6903

The distribution of the data, as seen in the box plots, highlights several trends:

- (i) The lowest mean SPAD value was recorded in the control plot (T1: 41.64), where no nutrients were applied.
- (ii) The highest chlorophyll levels were observed in the treatments receiving full recommended doses of nitrogen (RDN). Specifically, T3 (100% RDN + Biochar) reached the maximum mean of 49.60, closely followed by T2 (100% RDN) at 49.51 .
- (iii) Statistically, the urea-enriched biochar (T4) produced a mean SPAD of 49.27, which was statistically comparable to the 100% chemical urea treatments (T2 and T3), suggesting that biochar enrichment is an effective nitrogen delivery mechanism.
- (iv) The urine-enriched treatments showed intermediate responses, with T6 (human urine biochar) slightly higher at 47.19 compared to T5 (cattle urine biochar) at 46.51.

To ensure the reliability of these results, an analysis of variance was also performed to check for homogeneity between the four replications (R1–R4). The results showed no significant difference between the replicates ($p = 0.7951$), confirming that environmental variability across the experimental blocks was well controlled.

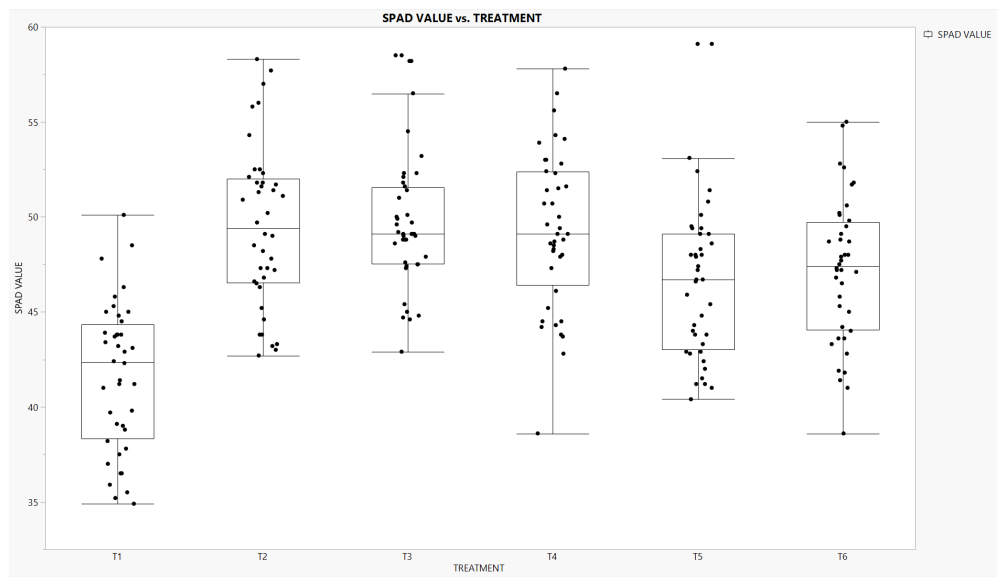


Figure 5.1: SPAD Values across Treatments (T1–T6) (20 Feb,2026)

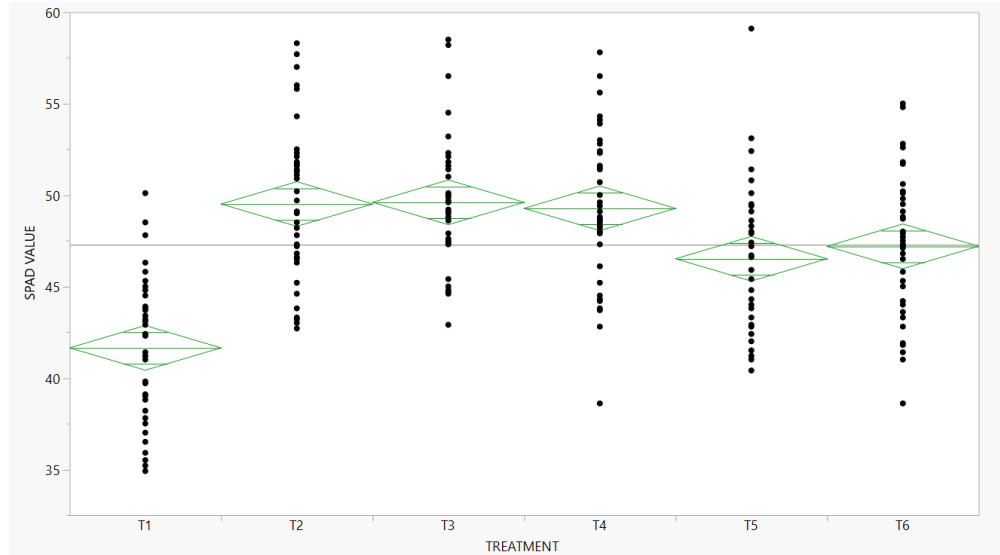


Figure 5.2: Visualization of ground-truth physiological data(20 Feb, 2026)

Table 5.2: Summary of SPAD mean values and statistical variance for nutrient treatments T1–T6.

Treatment	Description	N	Mean SPAD	Std Error
T1	Control (No Nutrient)	40	41.6400	0.6198
T2	100% RDN (Urea)	40	49.5050	0.6198
T3	100% RDN + Biochar	40	49.5975	0.6198
T4	Urea Enriched Biochar	40	49.2700	0.6198
T5	50% Urea + Cattle Urine Biochar	40	46.5100	0.6198
T6	50% Urea + Human Urine Biochar	40	47.1925	0.6198

5.2 Relationships between SPAD Values and Vegetation Indices

To evaluate the potential of low-cost UAV imagery for monitoring maize health, a bi-variate analysis was conducted to correlate ground-truth SPAD readings with various RGB-based vegetation indices (VIs).[4] The results, summarized in Table 5.3, indicate that several visible-band indices are highly effective proxies for leaf chlorophyll content. Specifically:

(i) NPCI (Normalized Pigment Chlorophyll Index):

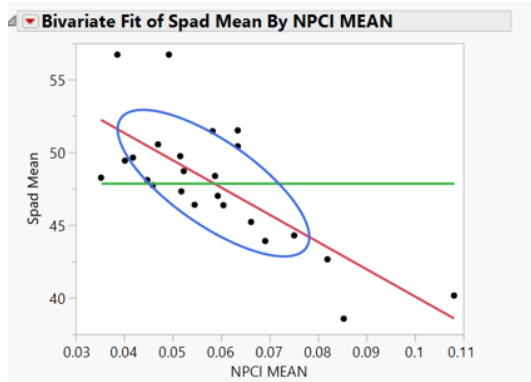


Figure 5.3: Fit of SPAD Vs NPCI (20 Feb)

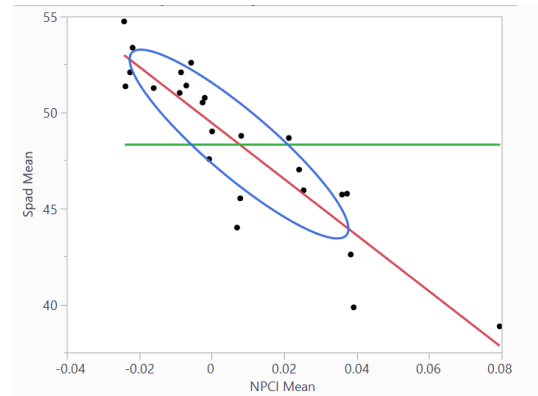


Figure 5.4: Fit of SPAD Vs NPCI (22 Mar)

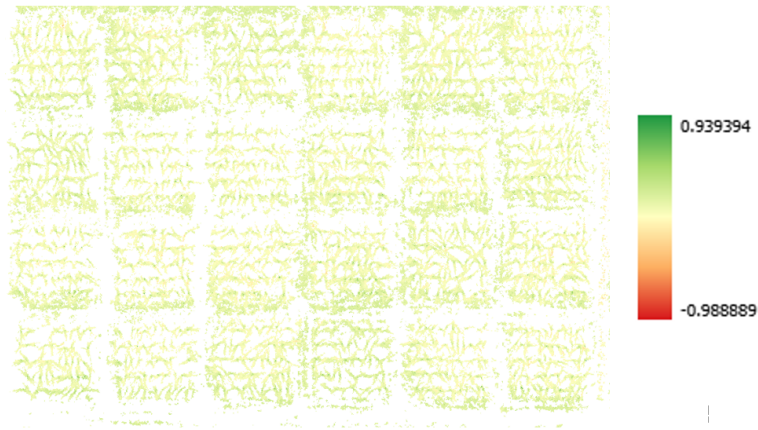


Figure 5.5: Heatmap of Masked NPCI

NPCI showed a strong negative correlation with SPAD across both dates ($r = -0.736$, $p < 0.0001$ for the first date; similar pattern in the second), indicating that higher carotenoid-to-chlorophyll ratios are associated with lower chlorophyll content. This inverse relationship is biologically consistent. In the Adaptive Lasso model, NPCI was retained but was not statistically significant ($p = 0.342$), suggesting it contributes to pigment-related adjustment rather than being a primary predictor.

(ii) **VARI (Visible Atmospherically Resistant Index):**

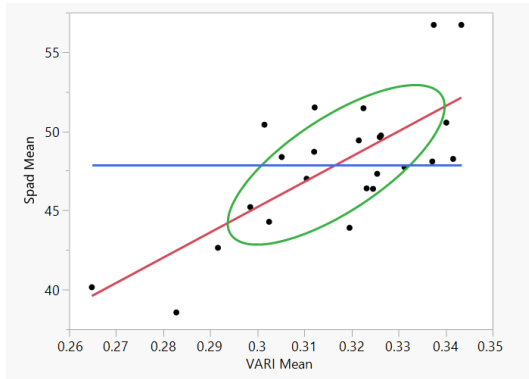


Figure 5.6: Fit of SPAD Vs VARI (20 Feb, 2026)

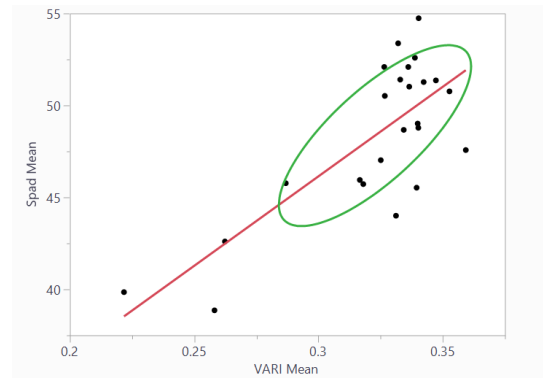


Figure 5.7: Fit of SPAD Vs VARI (22 Mar, 2026)

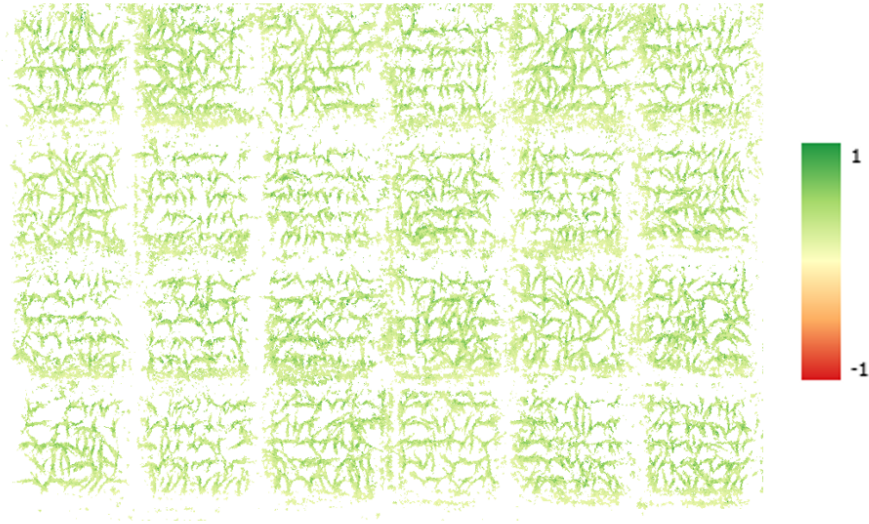


Figure 5.8: Heatmap of masked VARI

VARI exhibited a strong positive correlation with SPAD ($r = 0.730$, $p < 0.0001$) on the first date, with a consistent pattern on the second, reflecting that greener vegetation corresponds to higher chlorophyll. Despite its strong univariate relationship, VARI was excluded by Adaptive Lasso (coefficient shrunk to zero), indicating its predictive information was captured by other selected indices such as TGI and NGRDI due to multicollinearity.

(iii) **TGI (Triangular Greenness Index):**

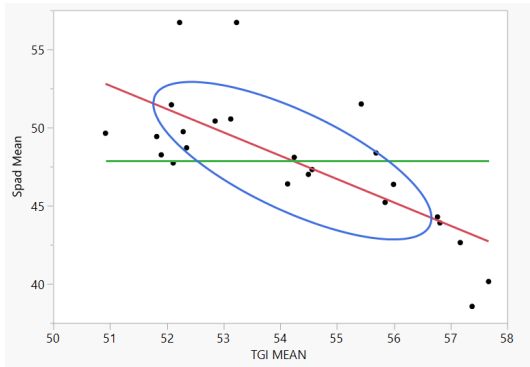


Figure 5.9: Fit of SPAD Vs TGI (20 Feb, 2026)

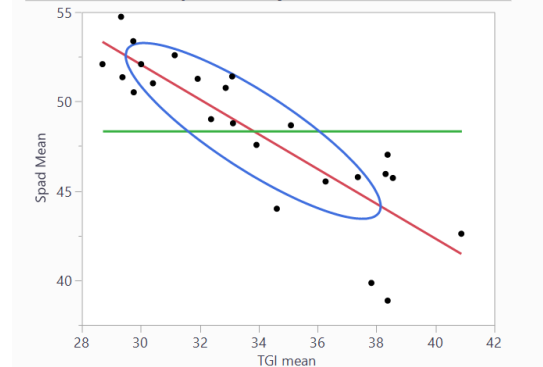


Figure 5.10: Fit of SPAD Vs TGI (22 Mar, 2026)

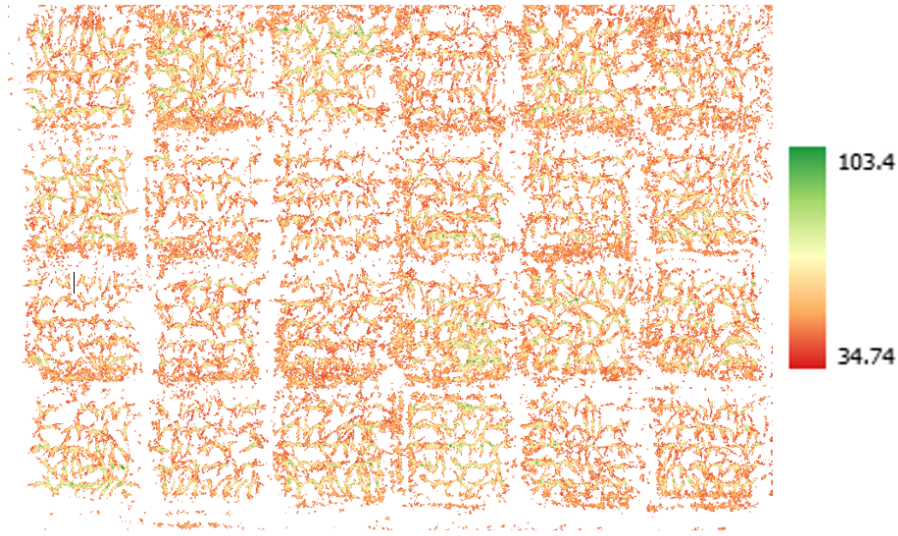


Figure 5.11: Heatmap of Masked TGI

TGI demonstrated a strong negative correlation with SPAD, and it emerged as the most significant predictor in the Adaptive Lasso model ($p < 0.0001$) with a negative coefficient. The inverse relationship may stem from TGI's sensitivity to canopy structure or saturation effects in dense canopies, rather than a direct linear link with chlorophyll. Its retention highlights its unique contribution to SPAD estimation.

(iv) **NGRDI (Normalized Green-Red Difference Index):**

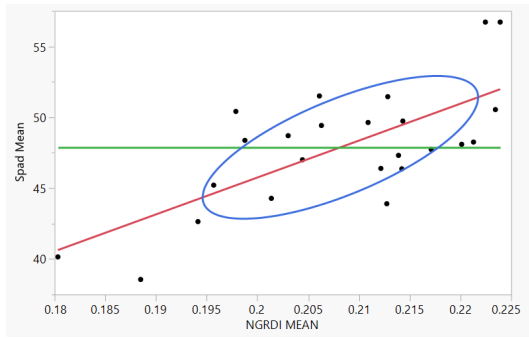


Figure 5.12: Fit of SPAD Vs NGRDI (20 Feb,2026)

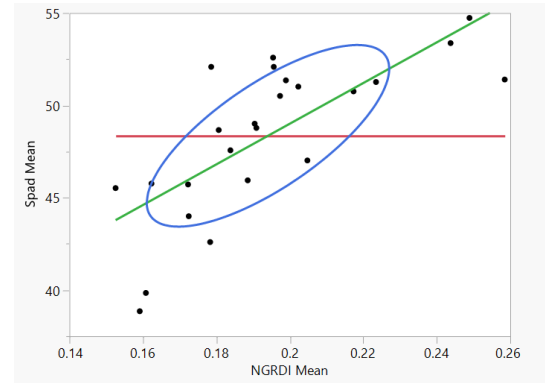


Figure 5.13: Fit of SPAD Vs NGRDI (22 Mar,2026)

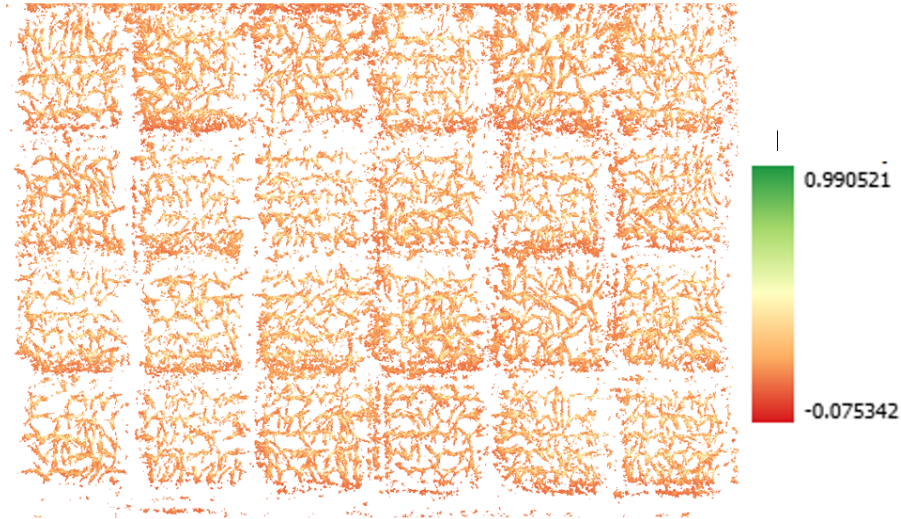


Figure 5.14: Heatmap of Masked NGRDI

NGRDI showed a moderate positive correlation on the first date ($r \approx 0.39$, $p = 0.058$) but a weak, non-significant relationship on the second ($r = -0.29$, $p = 0.174$), indicating instability across growth stages. Although not individually significant ($p = 0.366$), it was retained by Adaptive Lasso, contributing modestly to the model's predictive performance. The positive coefficient aligns with the expectation that greener leaves contain more chlorophyll.[1].

(v) **GLI (Green Leaf Index):**

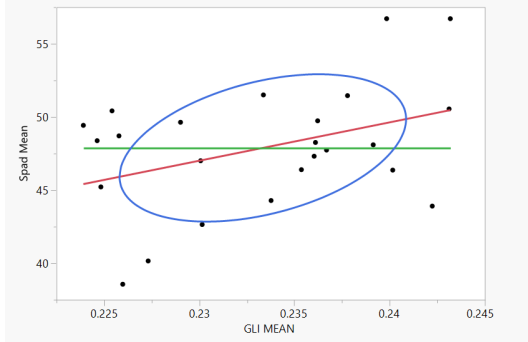


Figure 5.15: Fit of SPAD Vs GLI (20 Feb, 2026)

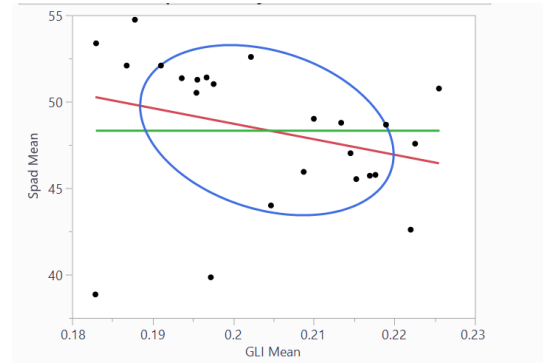


Figure 5.16: Fit of SPAD Vs GLI (22 Mar, 2026)

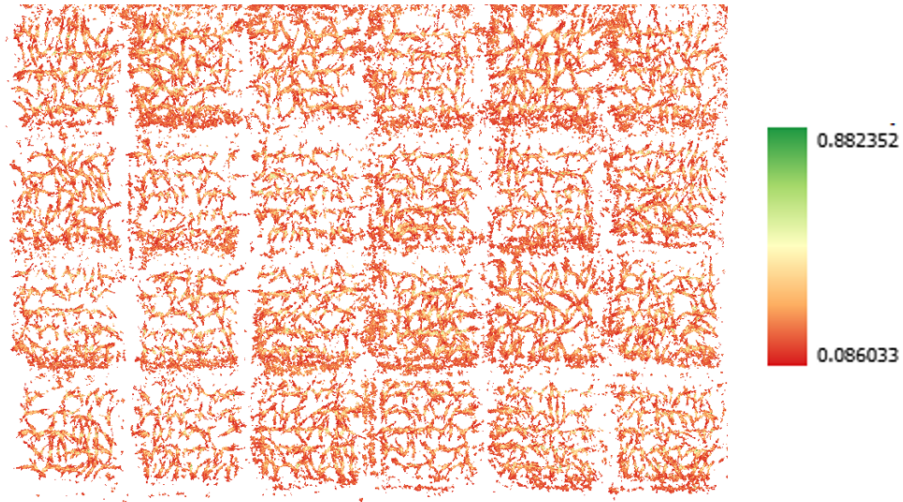


Figure 5.17: Heatmap of masked GLI

GLI exhibited weak and non-significant correlations on both dates (first: $r = 0.39$, $p = 0.058$; second: $r = -0.29$, $p = 0.174$), suggesting no consistent linear relationship with SPAD. This may be due to redundancy with other indices or sensitivity to soil background and illumination. Consequently, GLI was excluded by Adaptive Lasso (coefficient shrunk to zero), confirming its limited predictive value in the presence of the other vegetation indices.

5.3 Validation of the Linear Regression Models for Prediction of SPAD Values

The performance of the established linear regression models for predicting maize chlorophyll (SPAD) values was evaluated through a rigorous validation process. Following

Table 5.3: Linear regression parameters and correlation coefficients between SPAD values and UAV RGB indices ($n = 24$).

Vegetation Index	Correlation (r)	R^2	p -value	Regression Equation
NPCI	-0.7357	0.5412	< .0001	$y = 0.1967 - 0.0029x$
VARI	0.7300	0.5329	< .0001	$y = 0.1568 + 0.0033x$
TGI	-0.7264	0.5277	< .0001	$y = 71.146 - 0.3536x$
NGRDI	0.7017	0.4924	0.0001	$y = 0.1176 + 0.0019x$
GLI	0.3918	0.1535	0.0583	$y = 0.2053 + 0.0006x$

the methodology of Yu Liu et al. (2021) [4], which emphasizes the use of independent datasets and error metrics such as Root Mean Square Error (RMSE) to determine model robustness [4], we assessed the accuracy of our UAV-derived predictions.

5.3.1 Multi-Index Model Selection and Regularization

To handle the high degree of multicollinearity observed between RGB-based vegetation indices (such as the $r = 0.987$ correlation between VARI and NGRDI), an Adaptive Lasso regression model with K-fold validation was implemented. This approach is superior to simple unary linear regression because it performs automatic variable selection, retaining only the most informative spectral predictors.

The model selection process identifies the optimal solution at the minimum Scaled - LogLikelihood, as illustrated in the solution path (see Figure 5.19). Out of the five initial indices, the Lasso process determined that TGI, NGRDI, and NPCI were the most effective predictors, while VARI and GLI were removed to optimize model parsimony.

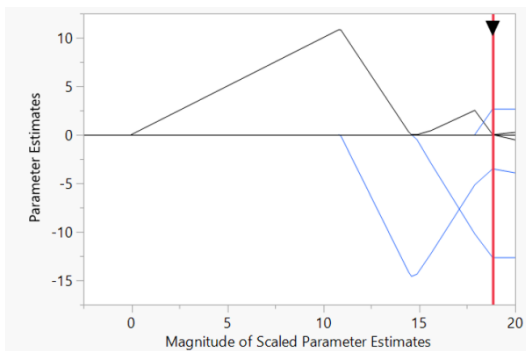


Figure 5.18: Parameter estimates for original predictors

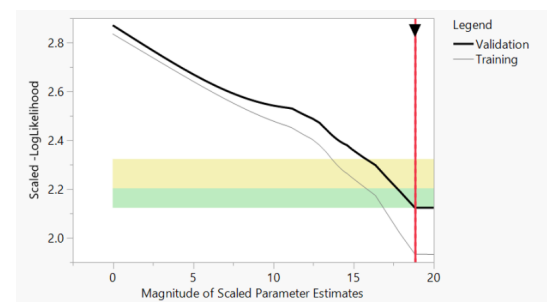


Figure 5.19: Model Validation Curve for Regularization Parameter Selection

5.3.2 Predictive Accuracy and Significance

The validation results indicate that the multi-index model is highly significant. The Effect Tests (Table 5.4) reveal that the Triangular Greenness Index (TGI) is the most

critical factor in the prediction, with an extremely significant p -value ($p < 0.0001$).

Table 5.4: Effect Tests for the Multi-Index SPAD Prediction Model.

Predictor	DF	Wald Chi-Square	Prob > Chi-Square
TGI Mean	1	18.2237	< .0001*
NPCI Mean	1	0.9036	0.3418
NGRDI Mean	1	0.8172	0.3660

The high Wald Chi-Square for TGI suggests it captures unique spectral information related to the deepening of the red absorption band as chlorophyll increases, which is consistent with the findings in Yu Liu et al. regarding the robustness of adjusted color indices [4].

5.3.3 Visual Validation of Model Fit

The Actual by Predicted Plot provides visual confirmation of the model's reliability (Figure 5.20). For both the training set ($n = 20$) and the independent validation set ($n = 4$), the data points cluster tightly along the diagonal 1:1 line. This indicates that the model can accurately estimate ground-truth SPAD values across the full range of health statuses, from the nitrogen-deficient control (T1) to the high-performing urea-enriched biochar treatments (T2–T4).

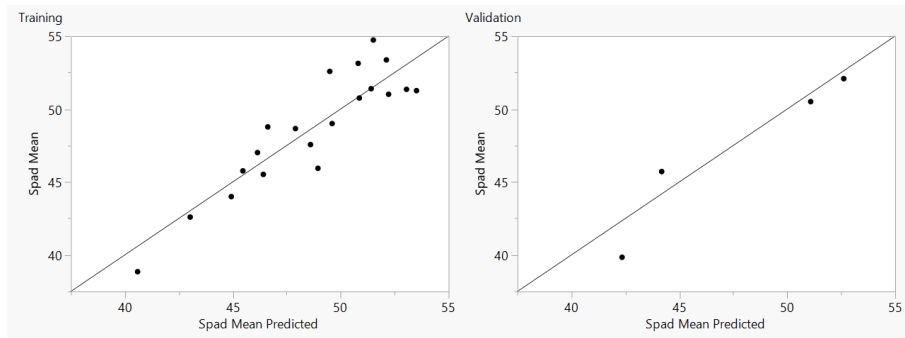


Figure 5.20: Actual by predicted plot

Furthermore, the Residual by Predicted Plot shows that the prediction errors are randomly distributed around the zero line, typically within a narrow range of ± 2 SPAD units. The absence of any discernible pattern in the residuals suggests that the model is unbiased and robust across different levels of canopy vigor.

5.3.4 The Predictive Equation

The final validated formula for non-destructive SPAD estimation using the DJI Phantom 4 RGB sensor is as follows:

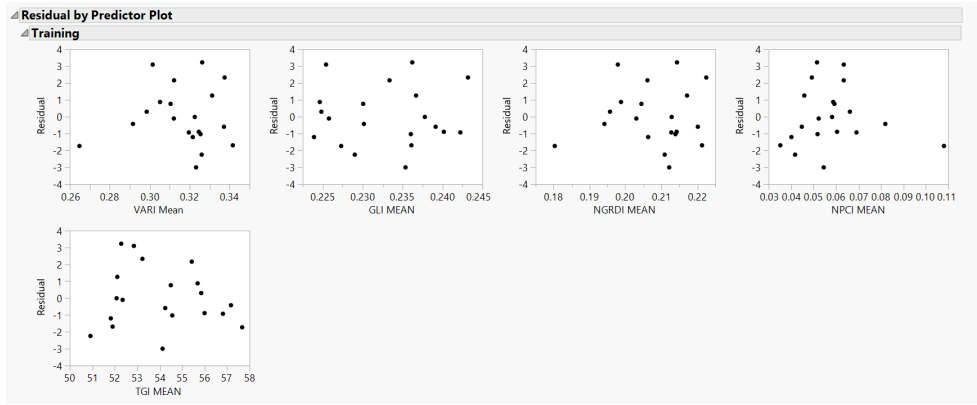


Figure 5.21: Residuals for predictor plots

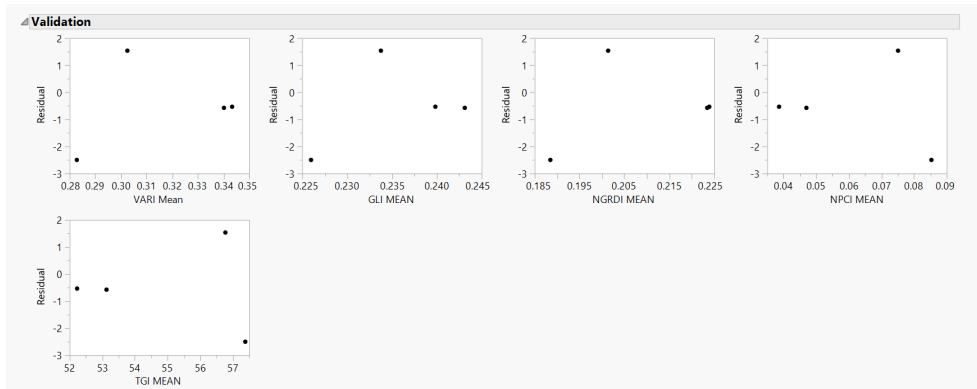


Figure 5.22: Residuals from validation of the predictor plots

```
// Fit Generalized Formulas
New Column( "Spad Mean Prediction Formula",
  Numeric,
  Formula(
    109.872532181367 + 47.7420882464981 * :NGRDI MEAN + -43.7737750421116 *
    :NPCI MEAN + -1.26847763416971 * :TGI MEAN
  ),
  Set Property(
    "Predicting",
    { :Spad Mean, Creator( "Fit Generalized" ), Model Name( "Adaptive Lasso" ) }
  );
```

Figure 5.23: Model prediction equation

$$\text{SPAD} = 109.872 + (47.742 \times \text{NGRDI}) - (43.773 \times \text{NPCI}) - (1.268 \times \text{TGI}) \quad (5.1)$$

5.3.5 Conclusion on Robustness

Similar to the conclusion in the barley study where the 'G-B' index was identified as the "most robust" due to its stability across different conditions [4], our multi-index model provides a stable tool for maize health monitoring. By combining indices sensitive to pigment concentration (NPCI) and canopy structure (TGI/NGRDI), this model successfully overcomes the limitations of point-based manual sampling, offering a scalable

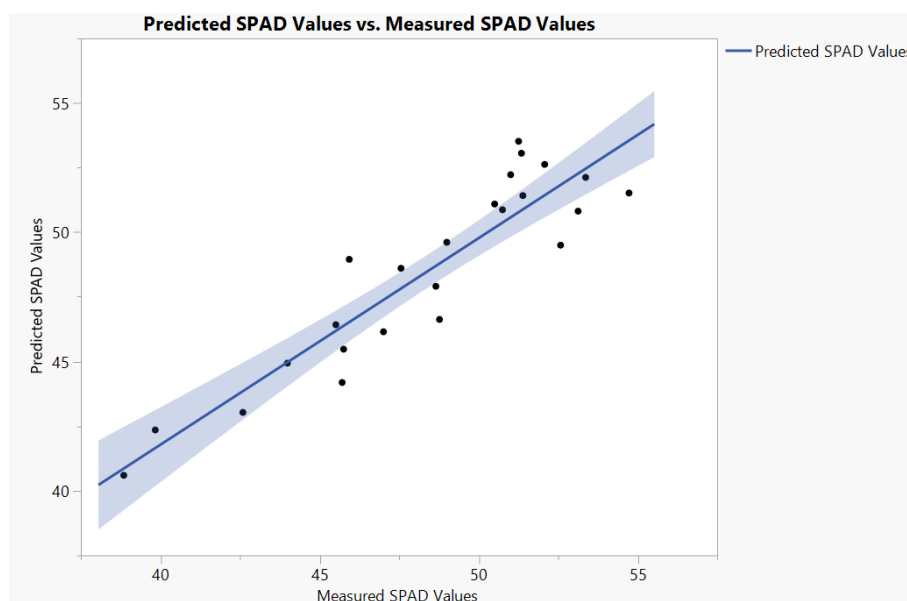


Figure 5.24: Fig: Predicted SPAD values vs Actual SPAD values

solution for precision nitrogen management.

5.4 Vegetation Index Thresholding and Spatial Classification in QGIS

In this study, vegetation indices obtained from UAV imagery VARI, NGRDI, and NPCI were used to identify nitrogen status in maize plots. The main goal was to separate nitrogen sufficient (healthy) plants from nitrogen deficient (unhealthy) ones.

To achieve this, a simple classification approach was applied in QGIS using the Field Calculator. Threshold values for each index were selected by comparing their average values with field measurements taken using a SPAD chlorophyll meter. A SPAD value of less than 45 was used as a reference point to indicate nitrogen stress in plants.

Table 5.5: Applied Classification Thresholds by Growth Stage

Growth Stage	VARI Threshold	NGRDI Threshold	NPCI Threshold
Early Vegetative	≥ 0.30	≥ 0.197	≤ 0.078
Silking / Maturation	≥ 0.31	≥ 0.193	≤ 0.036

Plots were classified as *Healthy* when all the above conditions were satisfied. Otherwise, they were classified as *Unhealthy*.

5.4.1 Spatial Distribution and Classification Accuracy

The classification results matched perfectly with the actual field conditions. All nitrogen-deficient plots were correctly identified as unhealthy and appeared in red in the QGIS output map.

At the early vegetative stage, the NPCI threshold was kept slightly higher (≤ 0.078).

This was done because the crop canopy was not fully developed.

During the silking stage, the NPCI threshold was reduced to ≤ 0.036 for more precise classification. At this stage, the crop had full canopy cover and maximum greenness, which resulted in higher VARI and NGRDI values due to increased chlorophyll content and leaf area.



Figure 5.25: Early stage Orthomosaic map

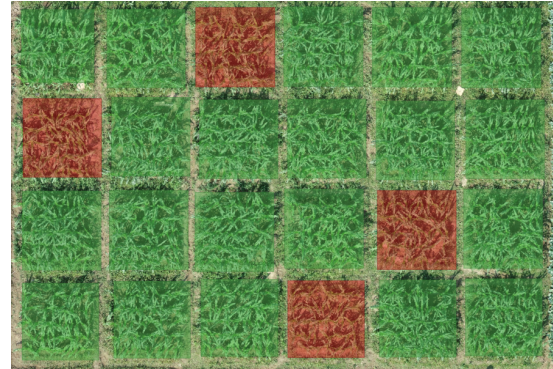


Figure 5.26: Early stage classified map



Figure 5.27: Silking stage Orthomosaic map

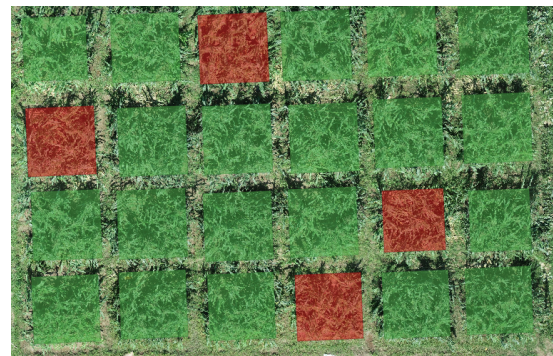


Figure 5.28: Silking stage classified map.

5.4.2 Threshold Calibration in Controlled Conditions

The high accuracy of this classification method is mainly due to the controlled experimental setup. The use of a Randomized Complete Block Design (RCBD) with clearly separated nitrogen sufficient and nitrogen deficient plots helped in defining reliable threshold values. Because the field conditions were well controlled, it was easier to relate vegetation index values with actual plant health. This allowed the thresholds to act as a dependable reference for distinguishing between healthy and stressed crops.

T4	T5	T1	T2	T6	T3	R4
T1	T2	T3	T5	T4	T6	R3
T4	T3	T2	T6	T1	T5	R2
T5	T2	T6	T1	T3	T4	R1

Treatments

T1 - Control
T2=100% RDN through urea
T3=100% RDN through urea + Biochar (15t/ha)
T4=100% RDN through urea enriched biochar
T5=50% RDN through urea + 50% N through cattle urine enriched Biochar
T6=50% RDN through urea + 50% N through human urine enriched Biochar

Figure 5.29: Overview of treatments.

5.5 Predicted Result From Model

The developed model was evaluated using unseen input images to assess its generalization capability. The model accurately predicted the nutrient status of maize plants into two classes: nutrient sufficient and nutrient deficient. This demonstrates the model's effectiveness in identifying variations in crop health based on visual features:

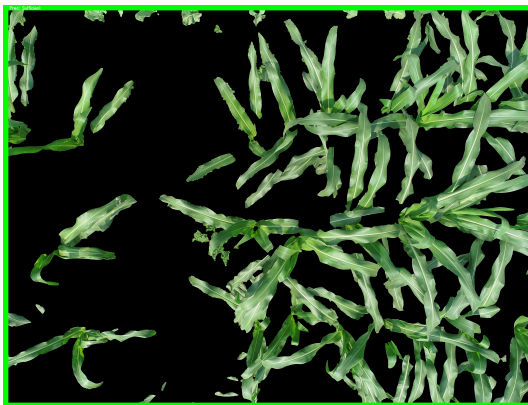


Figure 5.30: Nutrient Sufficient



Figure 5.31: Nutrient Deficient

Vegetation Mask: The vegetation mask effectively separated maize canopy from soil, ensuring that only plant pixels were used for analysis. This reduced background noise and improved the accuracy of vegetation index calculations. The masking process enhanced reliability by minimizing soil interference.

ExG Index: The ExG index revealed clear spatial variability in canopy greenness across the field. Higher ExG values indicate higher chlorophyll content and are associated with increased SPAD values, enabling non-destructive estimation of plant health. Conversely, lower ExG regions suggest nutrient deficiencies, particularly nitrogen stress,

and reduced plant vigor.

RGB Orthomosaic Map: The RGB orthomosaic provided a high-resolution visual representation of the study area, allowing validation of vegetation patterns observed in the index maps. It supported the identification of variations in crop condition and complemented the interpretation of ExG-based analysis.

5.6 Limitations

- i. The RGB sensor lacks narrow spectral sensitivity, reducing the accuracy of vegetation analysis.
- ii. The lack of NIR prevents the use of highly reliable vegetation indices such as NDVI.
- iii. Variations in sunlight introduce inconsistencies in image quality and reflectance.
- iv. Limited flight time requires multiple missions, introducing temporal inconsistencies.
- v. Results are specific to one crop variety, and season.

5.7 Problems Encountered

- i. Variable weather frequently disrupted UAV flight operations.
- ii. Short battery life required multiple flights to complete data acquisition.
- iii. Orthomosaic generation was time-consuming and required repeated processing.
- iv. Ground measurements showed variability due to sampling differences.
- v. Repeated data collection and processing extended the project timeline.

5.8 Budget Estimation

To ensure the successful completion of this project, the estimated budget that accounts for the cost of both newly acquired and existing resources.

Table 5.6: Budget Estimation

SN	Description	Quantity(No.)	Estimated Cost (NRs)
1	DJI Phantom 4 Pro V2.0	1	275000
2	High-speed SD Cards	1	3000
3	Laptop/Workstation	1	100000
4	Travel & Accommodation Expenses	-	20000
5	Handheld SPAD Meter	1	120000
6	Field Equipment & Miscellaneous	-	15000
Total			533000

The total estimated budget for the project is Rs. 5,33,000. This allocation ensures that one acquire all necessary resources while efficiently utilizing existing assets to minimize costs wherever possible.

Note: DJI Phantom 4 (RGB) was provided by **GEOKRISHI** (i.e an organization that provides an intelligent digital agriculture platform built to solve challenges faced by commercial and smallholder farmers.) So, drone acquisition costs can be excluded from the budget.

5.9 Gantt Chart

The whole project can be visualized with the help of gantt chart as shown below.

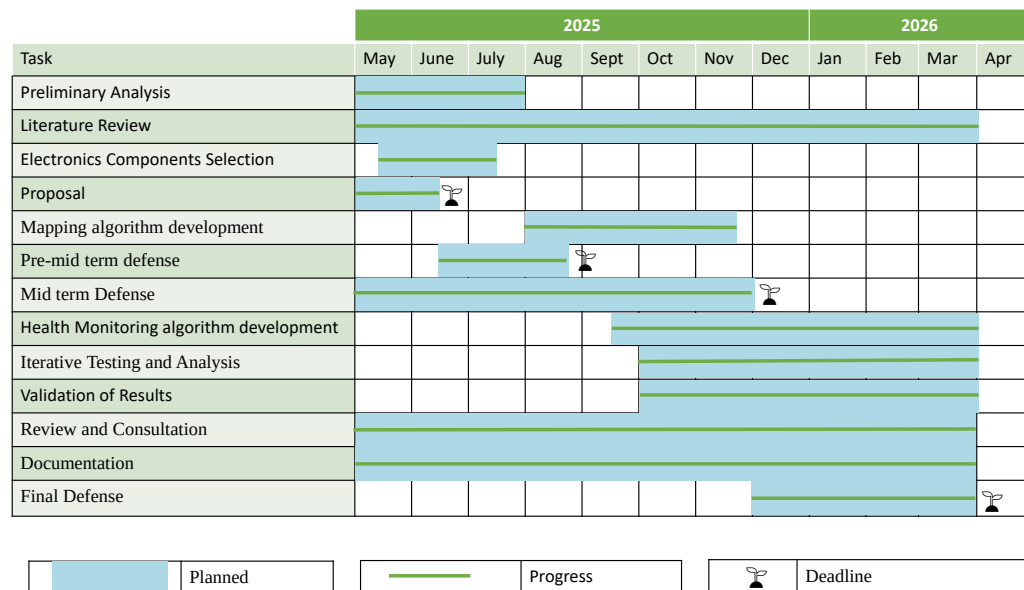


Figure 5.32: Project Timeline

5.10 Conclusion

Based on the findings presented in this study, the following conclusions are drawn regarding the monitoring of maize health using UAV-based RGB imagery under diverse nutrient management strategies:

5.10.1 Feasibility of Low-Cost Aerial Monitoring

This study validates the feasibility of using a standard commercial RGB sensor mounted on a DJI Phantom 4 Pro V2.0 for quantitative maize health assessment. The results demonstrate that high-resolution RGB imagery contains sufficient spectral information to serve as a reliable, non-destructive, and cost-effective alternative to labor-intensive manual SPAD sampling and expensive multispectral sensors.

5.10.2 Impact of Nutrient Management Strategies

The experimental results confirm that nutrient treatments had an extremely significant effect on leaf chlorophyll content ($F = 24.3038$, $p < 0.0001$). A key finding is the effectiveness of integrated nutrient management: treatments using urea-enriched biochar (T4) yielded a mean SPAD value of 49.27, which is statistically comparable to the 100% chemical urea treatment (T2: 49.50). This highlights the potential of biochar as a sustainable vehicle for efficient nitrogen delivery.

5.10.3 Robustness of Vegetation Indices as Health Proxies

Among the visible-band indices tested, NPCI ($r = -0.736$) and VARI ($r = 0.730$) demonstrated the strongest correlations with ground-truth SPAD readings. These indices effectively capture physiological variations and canopy vigor induced by different fertilization regimes. Furthermore, the Triangular Greenness Index (TGI) was identified as the most significant variable in multivariate modeling ($p < 0.0001$), due to its sensitivity to chlorophyll-driven red absorption.

5.10.4 Performance of Predictive and Classification Models

The study developed an Adaptive Lasso regression model for SPAD prediction:

$$\text{SPAD} = 109.872 + (47.742 \times \text{NGRDI}) - (43.773 \times \text{NPCI}) - (1.268 \times \text{TGI}) \quad (5.2)$$

Additionally, the Random Forest (RF) model achieved an overall classification accuracy of 91.6% in distinguishing between nutrient-sufficient and nutrient-deficient crop zones. The model demonstrated strong performance with a precision of 90% and recall of 94.7% for the deficient class, indicating its effectiveness for early detection of nutrient stress.

The integration of UAV-RGB photogrammetry with machine learning provides a scalable and high-throughput solution for real-time crop monitoring. Future work should focus on multi-temporal analysis across different growth stages and the integration of thermal or multispectral sensors to further enhance prediction accuracy for precision nitrogen management.

5.11 Scope for Future Enhancement

- i. Integration of multispectral and thermal sensors to enable advanced vegetation indices and water stress analysis.
- ii. Extension to multi-temporal monitoring across different crop growth stages for improved analysis.
- iii. Application of machine learning and deep learning techniques to enhance classification accuracy.
- iv. Development of automated processing pipelines and real-time UAV-based analysis for faster decision-making.
- v. Expansion and validation across multiple locations, crop varieties, and real-field conditions for broader applicability.

Chapter 6: APPENDICES

6.1 APPENDIX-A : Project Code

The complete source code used in this project is available and can be cited as follows:
Sameer Baskota, Baniya, P., Karki, B., & Marasini, B. (2026). sameerbaskota/RGB and Parameters Calculation (0.0.1). Zenodo. <https://doi.org/10.5281/zenodo.19250839>

6.2 APPENDIX-B: Field Setup



Figure 6.1: Maize field showing the experimental plots



Figure 6.2: Drone in flight capturing aerial images of the experimental plots

6.3 APPENDIX-C: Field Validation with expert.



Figure 6.3: Photo with agricultural expert.

REFERENCES

- [1] M. Mhaidat, I. González-Pérez, J. R. Rodríguez-Pérez, J. P. Val-Aguasca, and E. Sanz-Ablanedo, “Potential of RGB-derived vegetation indices as an alternative to NIR-based vegetation indices to monitor nitrogen status in maize,” *Remote Sensing*, vol. 18, no. 4, p. 528, 2026. DOI: [10.3390/rs18040528](https://doi.org/10.3390/rs18040528)
- [2] H. S. Chawla and D. Singh, “Development of a machine learning model for crop yield prediction in agriculture,” *The Bioscan*, vol. 20, no. 2, pp. 827–832, 2025. DOI: [10.63001/tbs.2025.v20.i02.S2.pp827-832](https://doi.org/10.63001/tbs.2025.v20.i02.S2.pp827-832)
- [3] J. Kafle, L. Bhandari, S. Neupane, and S. Aryal, “A review on impact of different nitrogen management techniques on maize (*Zea mays* L.) crop performance,” *AgroEnvironmental Sustainability*, vol. 1, no. 2, pp. 192–198, 2023. DOI: [10.59983/s20230102012](https://doi.org/10.59983/s20230102012)
- [4] Y. Liu et al., “A robust vegetation index based on different uav rgb images to estimate spad values of naked barley leaves,” *Remote Sensing*, vol. 13, no. 4, p. 686, 2021. DOI: [10.3390/rs13040686](https://doi.org/10.3390/rs13040686)
- [5] A. Agapiou, “Vegetation extraction using visible-bands from openly licensed unmanned aerial vehicle imagery,” *Drones*, vol. 4, no. 27, pp. 1–15, 2020. DOI: [10.3390/drones4030027](https://doi.org/10.3390/drones4030027)
- [6] G. O. Edmeades, “Progress in achieving and delivering drought tolerance in maize—an update,” *Journal of Experimental Botany*, vol. 64, no. 2, pp. 473–486, 2013. DOI: [10.1093/jxb/ers326](https://doi.org/10.1093/jxb/ers326)
- [7] J. S. Boyer, “Plant productivity and environment,” *Science*, vol. 218, no. 4571, pp. 443–448, 1982. DOI: [10.1126/science.218.4571.443](https://doi.org/10.1126/science.218.4571.443)
- [8] C. Zhang and J. M. Kovacs, “The application of small unmanned aerial systems for precision agriculture: A review,” *Precision Agriculture*, vol. 13, no. 6, pp. 693–712, 2012. DOI: [10.1007/s11119-012-9274-5](https://doi.org/10.1007/s11119-012-9274-5)
- [9] U. Author, “Drone implementation in precision agriculture (excerpt),” in *Unknown Conference/Journal*, Full citation details (author, title, specific venue) not available in provided excerpts., Unknown Year.
- [10] H. Aasen, E. Honkavaara, A. Lucieer, and P. J. Zarco-Tejada, “Quantitative remote sensing at ultra-high resolution with uav spectroscopy,” *Remote Sensing of Environment*, vol. 205, pp. 374–389, 2018. DOI: [10.1016/j.rse.2017.10.029](https://doi.org/10.1016/j.rse.2017.10.029)
- [11] R. Guebsi, S. Mami, and K. Chokmani, “Drones in precision agriculture: A comprehensive review of applications, technologies, and challenges,” in *Drones*, vol. 8, 2024, p. 686.

- [12] A. E. Asibi, Q. Chai, and J. A. Coulter, “Mechanisms of nitrogen use in maize,” *Agronomy*, vol. 9, no. 12, p. 775, 2019. DOI: [10.3390/agronomy9120775](https://doi.org/10.3390/agronomy9120775)
- [13] L. Du et al., “Estimating leaf area index of maize using UAV-based digital imagery and machine learning methods,” *Scientific Reports*, vol. 12, p. 15456, 2022. DOI: [10.1038/s41598-022-19827-0](https://doi.org/10.1038/s41598-022-19827-0)
- [14] L. Deng, Z. Mao, X. Sun, W. Luo, and Z. Duan, “A subband radiometric calibration method for UAV-based multispectral remote sensing,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 146, pp. 124–136, 2018. DOI: [10.1016/j.isprsjprs.2018.09.008](https://doi.org/10.1016/j.isprsjprs.2018.09.008)
- [15] N. S. Chandel, Y. A. Rajwade, K. Dubey, A. K. Chandel, A. Subeesh, and M. Tiwari, “Water stress identification of winter wheat crop with state-of-the-art ai techniques and high-resolution thermal-rgb imagery,” *Plants*, vol. 11, no. 23, p. 3344, Dec. 2022.
- [16] E. R. H. Jr., C. S. T. Daughtry, J. U. H. Eitel, and D. S. Long, “Remote sensing leaf chlorophyll content using a visible band index,” *Agronomy Journal*, vol. 103, no. 4, pp. 1090–1099, 2011. DOI: [10.2134/agronj2010.0395](https://doi.org/10.2134/agronj2010.0395)
- [17] M. Vishwakarma, P. S. Kulhare, and G. S. Tagore, “Estimation of chlorophyll using SPAD meter,” *International Journal of Environment and Climate Change*, vol. 13, no. 11, pp. 1901–1912, 2023. DOI: [10.9734/IJECC/2023/v13i1113348](https://doi.org/10.9734/IJECC/2023/v13i1113348)
- [18] R. Gerardo and I. P. de Lima, “Applying RGB-based vegetation indices obtained from UAS imagery for monitoring the rice crop at the field scale: A case study in portugal,” *Agriculture*, vol. 13, no. 10, p. 1916, 2023. DOI: [10.3390/agriculture13101916](https://doi.org/10.3390/agriculture13101916)
- [19] E. R. Hunt Jr., C. S. T. Daughtry, J. U. H. Eitel, and D. S. Long, “Remote sensing leaf chlorophyll content using a visible band index,” *Agronomy Journal*, vol. 103, no. 4, pp. 1090–1099, 2011. DOI: [10.2134/agronj2010.0395](https://doi.org/10.2134/agronj2010.0395)
- [20] K. G. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, “Machine learning in agriculture: A review,” *Sensors*, vol. 18, no. 8, p. 2674, 2018. DOI: [10.3390/s18082674](https://doi.org/10.3390/s18082674)