

MAPPING CROPLAND COVER IN MOZAMBIQUE WITH EXPERT KNOWLEDGE AND REMOTE SENSING

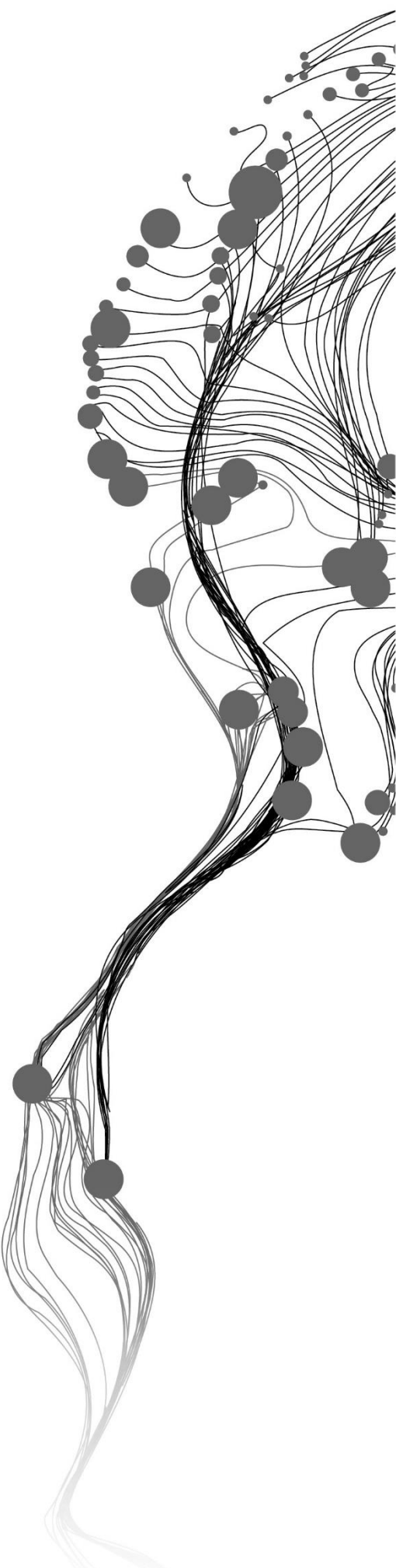
LAWRENCE MAWUTOR LEMGO
July 2024

SUPERVISORS:

Dr. RER. NAT. Florian.J. Ellsäßer
Dr. Panagiotis. NYKTAS

Advisor:

Dr. Claudia Paris



MAPPING CROPLAND COVER IN MOZAMBIQUE WITH EXPERT KNOWLEDGE AND REMOTE SENSING

LAWRENCE MAWUTOR LEMGO

Enschede, The Netherlands, July 2024

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.

Specialization: Natural Resources Management

SUPERVISORS:

Dr. RER. NAT. Florian J. Ellsäßer

Dr. Panagiotis NYKTAS

Advisor:

Dr. Claudia Paris

THESIS ASSESSMENT BOARD:

Prof. Andy D. Nelson (Chair)

Dr. Raian Vargas Maretto (External Examiner, University of Twente)

DISCLAIMER

This document describes work undertaken as part of a programme of study at the Faculty of Geo-Information Science and Earth Observation of the University of Twente. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the Faculty.

ABSTRACT

Creating accurate and up-to-date land cover maps are essential for managing natural resources, strategic land use planning, and monitoring changes that occur in our environment. This study employed remote sensing to evaluate the effectiveness of Sentinel-2 and PlanetScope satellite imagery in distinguishing different maize cropping systems in Gaza province, Mozambique. Using a Random Forest (RF) classifier, the research investigated the relative importance of spectral indices, textural, and topographic features derived from high spectral resolution (Sentinel-2) and high spatial resolution (PlanetScope) data.

Key research questions addressed included identifying the most effective features for distinguishing maize cropping systems and determining whether the two datasets had a statistically significant difference in classification accuracy. The study area encompassed diverse landscapes characterized by smallholder farms and mixed cropping systems. Sentinel-2 and PlanetScope imagery from 2023 were processed using the Google Earth Engine (GEE) platform followed by the extraction of spectral, textural, and topographic features.

The results indicated that elevation, red edge bands (Sentinel-2) and visible bands (PlanetScope) were among the most significant features for classification. Feature importance analysis using Mean Decrease Gini (MDG), SHapley Additive exPlanations (SHAP), and Permutation Importance methods consistently highlighted these features. The RF model showed high accuracy for classes like trees, shrubs, and water bodies but recorded low performance with grass and mixed fields primarily due to spectral overlaps and imbalanced training samples.

Comparative analysis using McNemar's test revealed no statistically significant difference in classification accuracy between Sentinel-2 and PlanetScope for discriminating maize cropping systems. Despite the high spatial resolution of PlanetScope its classification performance was similar to Sentinel-2. This emphasizes the importance of spectral resolution.

This research contributes to the understanding of the strengths and limitations of high spectral and spatial resolution datasets in agricultural monitoring. It provides insights into cost-effective methods for accurate land cover mapping, essential for informed agricultural policymaking and resource management in heterogeneous landscapes like Gaza Province. The study also emphasizes the need for balanced datasets and high-precision data collection to improve model performance and reliability.

Keywords: Land cover mapping, Sentinel-2, PlanetScope, Random Forest classifier, spectral indices, textural metrics, topographic features, maize cropping systems, agricultural monitoring, SHapley Additive exPlanations (SHAP), Mean Decrease Gini (MDG), Google Earth Engine (GEE).

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to God for His blessings and the good health granted to me throughout my study. My heartfelt thanks go to my supervisors, Dr. RER. NAT. Florian J. Ellsäßer and Dr. Panagiotis NYKTAS for their unwavering guidance and patience.

Special thanks to my advisor, Dr. Claudia Paris for her invaluable support, particularly with Google Earth Engine (GEE) and her expertise in supervised classification. I am also grateful to Prof. Andy D. Nelson for his critical feedback during my proposal defense which taught me the importance of using the right methods and understanding their limitations.

I owe a significant debt of gratitude to my family, especially my brother Godwin Lemgo and his wife Nina Konate for their immense support and encouragement. My sincere thanks extend to the ITC faculty and the NRM department for their impact on my personal and professional life and to the ITC Geospatial Computing Platform for providing the necessary computational resources when my local environment fell short.

Lastly, I am deeply thankful to my colleagues: Charlie Shawa, Umar Abdul, Prince Boateng, Ofori-Karikari (Prof), Chimbi Anesu, Joel Ketu, Emmanuel Salami (Omo), Jane-Marie, Zhu Jingwen, Swai Calvin, Sally Sampson, and all the esteemed members of HQ (Room 605 ITC Hotel) where countless brainstorming sessions were held. Our sleepless nights ultimately paid off and good luck to all of you in your future endeavors.

TABLE OF CONTENTS

Table of Contents

1.	INTRODUCTION.....	1
1.1.	Background.....	1
1.2.	Problem Statement.....	6
1.3.	Research Objectives, and Questions.....	7
2.	METHODS.....	8
2.1.	Study area.....	8
2.2.	Data.....	10
2.2.1.	Remote Sensing data.....	11
2.2.2.	Field Data.....	12
2.2.3.	Existing Landcover Products.....	13
2.3.	Data pre-processing.....	15
2.3.1.	Sentinel-2 Image pre-processing.....	15
2.3.2.	Image composites calculations.....	15
2.4.	Preparation and refinement of field observations.....	17
2.5.	Feature Extraction.....	19
2.5.1.	Preparation of input features for RF model.....	23
2.5.2.	Correlation analysis on input features.....	24
2.5.3.	Input features experimental setup.....	25
2.5.4.	Implementation of Random Forest.....	27
2.5.5.	Variable Importance Analysis.....	28
2.5.6.	Evaluation of Classification Model.....	29
2.6.	McNemar's Test for Comparing Classification Models.....	30
3.	RESULTS.....	31
3.1.	Correlation results for Sentinel-2 input data.....	31
3.2.	Results of Input Features Experiments.....	34
3.3.	Variable Importance Results.....	36
3.3.1.	Variable importance for Sentinel-2 (Rain season composite-best experiment).....	36
3.3.2.	Variable importance for PlanetScope (Annual composite-best experiment).....	38
3.4.	Classification Results for Sentinel 2.....	41
3.4.1.	Classification Results for PlanetScope data.....	43
3.4.2.	Consistency in Land Cover Classification between Sentinel-2 and PlanetScope data.....	45
3.4.3.	Inconsistencies in land cover classification between Sentinel-2 and PlanetScope data.....	47
3.5.	Evaluation of Classification Results for Sentinel-2.....	48
3.5.1.	Evaluation of Classification Results for PlanetScope.....	50
3.5.2.	Comparison of the best classification outputs for Sentinel-2 and PlanetScope.....	52
3.6.	McNemar's Test.....	54
4.	DISCUSSION.....	55
4.1.	Correlation Analysis and Experiments for input variables.....	55
4.2.	Feature Importance for the discrimination of different maize cropping systems.....	56
4.3.	Model performance.....	57
4.4.	Comparison of Classification Accuracy between Sentinel-2 and PlanetScope.....	58

4.5.	Limitations and Recommendations.....	59
4.6.	Implications	60
5.	CONCLUSIONS	61
6.	ETHICAL CONSIDERATIONS	62

LIST OF FIGURES

Figure 1: Map showing the study area (Gaza Province).....	9
Figure 2: Photos showing Maize, Mixed crop fields and Rice fields.....	13
Figure 3: Overall flowchart of the study methods. This flowchart outlines the process of data acquisition, preprocessing, and analysis.	14
Figure 4: Workflow for Pre-Processing Sentinel-2 Imagery.....	16
Figure 5: Comparison of Raw and Pre-Processed Sentinel-2 Imagery after cloud removal and median composite calculation.....	16
Figure 6: Refinement of Crop Field Polygon Data Using K-means Clustering Algorithm.....	18
Figure 7: Workflow for refining crop field data to be used as input into the RF model.....	19
Figure 8 : workflow for preparing input features in GEE.....	24
Figure 9: Distribution of training and validation points across different land cover classes.	28
Figure 10: Pearson correlation matrix for Sentinel-2 spectral bands.	32
Figure 11: Pearson correlation matrix of vegetation indices computed from Sentinel-2 image.....	32
Figure 12: Pearson correlation matrix for PlanetScope spectral bands.....	33
Figure 13: Pearson correlation matrix of vegetation indices computed from PlanetScope image.....	33
Figure 14: Bar chart showing feature importance using Mean Decrease Gini (MDG) Method on Sentinel-2 rain season composite data.	36
Figure 15: Bar chart showing feature importance using permutation importance Method on Sentinel-2 rain season composite data.	37
Figure 16: Bar chart showing feature importance using SHAP Method on Sentinel-2 rain season composite data.	37
Figure 17: Bar chart showing feature importance using Mean Decrease Gini (MDG) Method on PlanetScope annual composite data.....	38
Figure 18: Bar chart showing feature importance using permutation importance method on PlanetScope annual composite data.	39
Figure 19: Bar chart showing feature importance SHAP Method on PlanetScope annual composite data.....	39
Figure 20: Land Cover Map Produced Using Sentinel-2 Data with the Best Identified Features.....	41
Figure 21: Land Cover Map Produced Using Sentinel-2 Multi-stack Composite Data with the Best Identified Features.....	42
Figure 22: Land Cover Map Produced Using PlanetScope Data with the Best Identified Features.	43
Figure 23: Land cover map produced using PlanetScope data with the best identified features.	44
Figure 24: Comparison of land cover classification using PlanetScope data for areas A, B, C, and D. The top row shows satellite imagery snapshots, the middle row displays the best feature maps, and the bottom row presents the multi-stack maps.	45
Figure 25: Comparison of land cover classification using Sentinel-2 data for areas A, B, C, and D. The top row shows satellite imagery snapshots, the middle row displays the best feature maps, and the bottom row presents the multi-stack maps.....	46
Figure 26: Comparison of land cover classification inconsistencies using Sentinel-2 and PlanetScope data.	47
Figure 27: Precision, Recall, and F1-Score by Class for the Sentinel-2 Best Features Map.	48
Figure 28: Precision, Recall, and F1-Score by Class for the Sentinel-2 Multi-Stack Composite Map.	49
Figure 29: Precision, Recall, and F1-Score by Class for the PlanetScope Best Features Map.....	50
Figure 30: Precision, Recall, and F1-Score by Class for the PlanetScope Multi-Stack Composite Map.	51

Figure 31: Agreement and disagreement in land cover classification between PlanetScope and Sentinel-2 data for various classes.53

LIST OF TABLES

Table 1: Summary of Datasets used in this study.....	10
Table 2: Sentinel-2 spectral and spatial resolution	11
Table 3:PlanetScope spectral and spatial resolution.....	12
Table 4: GLCM Textural Features Computed for Sentinel-2 Bands	21
Table 5: GLCM Textural Features Computed for PlanetScope Bands	21
Table 6: Summary Table of Vegetation indices Calculated for Sentinel-2 and PlanetScope Images.....	22
Table 7: Summary of Experimental Setup for Sentinel-2 Datasets	26
Table 8: Summary of Experimental Setup for PlanetScope Datasets	27
Table 9: Summary of performance metrics for the input features experiments using Sentinel-2 data across different composite types (rain season, dry season, and annual composite).....	34
Table 10: Summary of performance metrics for the input features experiments using PlanetScope data across different composite types (rain season, dry season, and annual composite)	35
Table 11: Summary of McNemar's Test for Maize and Mixed Fields	54

LIST OF ABBREVIATIONS

RF	Random Forest
GEE	Google Earth Engine
MDG	Mean Decrease Gini
SHAP	SHapley Additive exPlanations
MODIS	Moderate Resolution Imaging Spectroradiometer
ESA	European Space Agency
NIR	Near-Infrared
SWIR	Short-Wave Infrared
GLCM	Gray-Level Co-occurrence Matrix
NDVI	Normalized Difference Vegetation Index
EVI	Enhanced Vegetation Index
NDWI	Normalized Difference Water Index
GNDVI	Green Normalized Difference Vegetation Index
SAVI	Soil-Adjusted Vegetation Index
NDMI	Normalized Difference Moisture Index
DT	Decision Trees
BSI	Bare Soil Index
SRTM	Shuttle Radar Topography Mission
SVM	Support Vector Machine
GIS	Geographic Information Systems
ML	Machine learning

1. INTRODUCTION

1.1. Background

The precise and timely compilation of land cover maps plays a pivotal role in understanding the evolving dynamics of human societies and their impact on the environment (Somvanshi et al., 2020). Researchers have increasingly employed remote sensing and Geographic Information Systems (GIS) tools to enhance accuracy and accelerate the making of maps showing land use land cover and various features of the Earth surface across the world (Schulz et al., 2021). Land cover mapping has grown exponentially in importance in contemporary society, playing a pivotal role across various fields (Giri et al., 2013; Schulz et al., 2021). For example, within the field of agriculture, accurate cropland mapping enables proper cropland management and provides insight into cropland and land cover dynamics (Hao et al., 2018). Land cover maps are particularly useful in providing information regarding suitable areas for the growing and cultivation of different categories of crops. This aids in effective farm management, policymaking, and resource allocation (Singh & Sharma, 2022). Using satellite images for land cover mapping can offer significant advantages over traditional field data collection by providing large-scale coverage, consistent, and frequent observations (Villegas Rugel et al., 2023). Various sensors like MODIS, Landsat, and Sentinel-2 among others provide valuable datasets for agriculture mapping (F. Gao & Gao, 2021). Various useful information can then be extracted from these datasets for the monitoring of crop health, soil moisture, and crop yield among others (Pratik et al., 2023).

Within the field of remote sensing, the selection of an appropriate sensor and algorithm is important for accurate image classification due to the complexities associated with remotely sensed data (Belgiu & Drăgu, 2016). Different types of sensors in relation to remote sensing image analysis relate to different types of roles (Melesse et al., 2007). There are many different types of remote sensing satellites that are used to gather information about the Earth's surface. For example, The Sentinel-2 sensor, launched by the European Space Agency (ESA) has become popular with respect to cropland mapping applications due to its unique capabilities that surpass those of other sensors (Pazúr et al., 2022). The extended red-edge bands of Sentinel-2 offer a distinct advantage for detailed vegetation monitoring and cropland discrimination. This has been demonstrated by (Otunga et al., 2019). The red-edge bands positioned near the chlorophyll absorption edge in combination with the NIR band provide enhanced sensitivity to variations in vegetation health (Delegido et al., 2011). This enables more accurate identification of croplands. Additionally, the 13 spectral bands of the Sentinel-2 sensor encompass key wavelengths that are crucial for cropland identification and mapping (Masimula, 2020). The presence of narrow-wavebands in the blue, green, red, and near-infrared regions allow for effective discrimination between different vegetation types, including croplands (Phiri et al., 2020). This spectral depth contributes to improved classification accuracy compared to sensors with fewer bands (Zhang et al., 2019). In comparison to other sensors, the Sentinel-2 sensor outperforms other sensors like Landsat and MODIS in terms of spatial resolution (Meyer et al., 2019). With a spatial resolution ranging from 10 to 20 meters, Sentinel-2 provides a detailed view of the landscape surpassing these sensors in clarity (Mandanici & Bitelli, 2016). This enhanced spatial resolution enables the delineation of

smaller and more fragmented cropland areas. This can be particularly important in regions with complex land cover (Song et al., 2021). Another important feature of the Sentinel-2 sensor has to do with its revisit period (Spoto et al., 2012). Revisit time in remote sensing refers to the time it takes for a satellite to complete one orbit cycle and capture images of the same location on Earth again. The revisit time of every 5 days provides a significant advantage over the longer revisit times of other sensors like Landsat and MODIS (J. Li et al., 2017). The superior capabilities of Sentinel-2 for cropland mapping have been recognized in some previous studies. For instance, Belgiu & Csillik, (2018) successfully employed Sentinel-2 imagery to map croplands within three different study areas in Europe and North America. Another study by Rawat & Saini, (2022) used a single date Sentinel-2 image to map and detect changes in cropland sizes in Uttar Pradesh, India.

On the other hand, another satellite which has gained popularity with regards to cropland mapping is the PlanetScope satellite. The PlanetScope constellation, which has been launched and operated by Planet Labs has gained important recognition in cropland mapping (Rao et al., 2021). This is due to the satellite's high spatial resolution and frequent revisit times which offers an advantage over other satellite sensors (Houborg & McCabe, 2016). PlanetScope satellites provide imagery with a spatial resolution which ranges from 3 to 5 meters. The high spatial resolution capability of PlanetScope allows for a detailed monitoring and classification of cropland areas (Menefee et al., 2022). Additionally, with daily revisit capability, PlanetScope surpasses many other satellites including the Sentinel-2 satellite in terms of temporal resolution. Daily images captured by the PlanetScope satellite can help minimize data gaps caused by cloud cover providing consistent datasets for cropland analysis (Pickering et al., 2021). PlanetScope's multispectral capabilities includes four spectral bands (blue, green, red, and near infrared). These spectral bands are essential for vegetation monitoring and health assessment. These bands enable the calculation of vegetation indices like the Normalized Difference Vegetation Index (NDVI) which helps assess crop vigor and biomass (Houborg & McCabe, 2016). Additionally, other vegetation indices like the Enhanced Vegetation Index (EVI), the Soil Adjusted Vegetation Index (SAVI), and the Green Normalized Difference Vegetation Index (GNDVI) among others can also be calculated. These indices are relevant for agriculture and cropland mapping as they offer more accurate assessments of plant health, biomass, and soil conditions (Vidican et al., 2023). Although PlanetScope satellite has fewer spectral bands compared to other satellites like the Sentinel-2 sensor, Planet scope's higher spatial resolution offers more precise delineation of crop boundaries and intra-field variability (Rao et al., 2021). Some previous studies have successfully used PlanetScope for cropland and land cover mapping. For instance, a study in Northern Mozambique used PlanetScope data to map active cropland and short-term fallows in smallholder landscapes. This research combined PlanetScope mosaics and an iterative learning method to achieve high mapping accuracy. This method utilized PlanetScope's high detail and frequent updates to provide precise cropland and fallow period estimates (Rufin et al., 2022). Another study by (Trivedi et al., 2023) focused on mapping cropland in tropical smallholder systems. In this study, the high resolution and frequent revisit times of PlanetScope helped overcome challenges like cloud cover and field variability. This study employed a Random Forest classifier with recursive feature elimination to effectively map arable fields in complex landscapes. Finally, a study by Liu et al., (2022) demonstrated PlanetScope's effectiveness in mapping and monitoring smallholder farms when combined with other satellite sensors. This study highlighted the superior performance of PlanetScope in detecting and classifying diverse crop types in smallholder farms. These studies show that PlanetScope's high spatial resolution and frequent revisit capabilities can enhance the accuracy and reliability of cropland and land cover mapping.

There are two main methods commonly used in remote sensing regarding image classification. These are unsupervised and supervised classification (Wang & Cheng, 2010). Unsupervised classification is a way to identify and group pixels in an image based on their spectral characteristics, without the need for prior knowledge or labeled data (Wang & Cheng, 2010). Two of the most used algorithms for unsupervised classification include ISODATA and K-means clustering. The k-means clustering algorithm groups pixels based on their spectral characteristics. This enables the algorithm to identify different land cover types. ISODATA clustering is another clustering algorithm used in remote sensing for classifying images. The term ISODATA stands for Iterative Self-Organizing Data Analysis Technique (de Bie et al., 2011). This is an unsupervised clustering algorithm commonly used in remote sensing and image analysis. The algorithm iteratively splits a dataset into clusters based on data point similarity. The algorithm calculates the mean and standard deviation of each cluster, reassigning data points to the nearest cluster based on a user-defined distance threshold. ISODATA is useful for land use and land cover classification in remote sensing. Compared to K-means, ISODATA is more computationally complex, requiring multiple iterations, but it can be more accurate as it adapts to changes and adjusts the number of clusters accordingly (de Bie et al., 2008; lowast et al., 2020; Reda et al., 2019; Richards, 2022).

Supervised classification involves selecting training samples and classifying the image based on the chosen samples. The training samples are key because they determine which class each pixel inherits in the overall image (Samaniego et al., 2008). Some of the most used supervised classification algorithms include maximum likelihood, support vector machines (SVM), decision trees (DT), and random forests (RF). Each algorithm has its own strengths and weaknesses. The choice of algorithm depends on the specific application and the type of data being analysed. For example, traditional classifiers like the maximum likelihood classifier is often used for image classification because it is simple and easy to use (Sunar Erbek et al., 2004). However, one of the main drawbacks is that the maximum likelihood classifier assumes that the data is normally distributed which may not always be the case in practice. Additionally, maximum likelihood can be sensitive to outliers in the data, which can lead to inaccurate classification results (Peng et al., 2019). SVM is a supervised learning algorithm primarily used for both classification and regression tasks (Belavagi & Muniyal, 2016). SVM works by identifying the optimal hyperplane that maximizes the margin between different classes and can handle non-linear datasets using a kernel trick (Nie et al., 2020). This approach transforms data into a higher-dimensional space to achieve linear separation, making it effective for complex datasets. SVM have proven useful in remote sensing for land cover mapping, achieving high classification accuracies (Ding et al., 2020; ElMannai et al., 2019; Shi & Yang, 2015). Despite its strengths, SVM can be computationally demanding and requires careful parameter selection (Dabija et al., 2021; Sheykhmousa et al., 2020).

Another popular machine learning algorithm used in remote sensing for image classification is a decision tree (DT). A decision tree is a machine learning algorithm which operates by breaking down a dataset into smaller subsets while at the same time developing an associated decision tree (Bansal et al., 2022). DT models work like a tree with branches (Costa & Pedreira, 2022). They divide the data step by step using simple rules which can be associated with the characteristic or features with the data. These steps lead to the end points of the tree called leaves where the final categories or decision is made. DT stand out for their flexibility and robustness in managing diverse datasets (Pal & Mather, 2003). One major advantages of DT algorithms are their capability to process data across various scales which makes them adaptable to different types of data (Brodley & Friedl, 1997). Notably, DT operate without making any prior assumptions about how data is distributed within each category (Fletcher & Islam, 2019). This allows for a

more thorough analysis of the data at hand. However, the simplicity of DT can also be a limitation as DT are prone to overfitting especially with complex datasets (Holloway et al., 2019). This means they can create overly complicated trees that do not generalize well to new unseen data.

Addressing the main limitation of DTs, RF algorithms leverages the collective power of ensembles (Louppe, 2014). RF algorithms construct an ensemble of decision trees which collectively help predict outcomes with high accuracy (Talekar, 2020). Each individual decision tree in the ensemble adds a vote to assign the input data to the most frequent class. The RF algorithm initially constructs numerous binary classification trees (ntree) using diverse bootstrap samples drawn with replacements from the original dataset. This majority vote strategy ensures that the final classification output is based on a collective decision (Dogan & Birant, 2019). This makes the RF algorithm resilient to noise and outliers in the data. Within the field of remote sensing, the Random Forest algorithm has emerged as a good foundation for image classification especially land cover mapping (Belgiu & Drăgu, 2016). RF algorithm distinguishes itself by its ability to integrate multiple decision trees thereby mitigating overfitting and reducing variance (Briem et al., 2002). The incorporation of the bagging method within the random forest algorithm ensures robustness by training each tree on a random subset of data (X. Zhou et al., 2019). This results in the development of a stable and accurate model. Moreover, RF is capable of handling multi-modal input datasets without assuming a normal distribution making it ideal for capturing complex spatial patterns (DeFries & Chan, 2000). For instance, studies by Belgiu & Csillik, (2018); Ghimire et al., (2012); Gislason et al., (2006) successfully employed the random forest algorithm for land cover classification achieving high accuracy. The ability of the random forest algorithm to adapt to diverse spectral data, its proficiency in modeling non-linear relationships, and its ability to reduce the risk of overfitting makes it an ideal choice for extracting meaningful information from satellite imagery. While Random Forest is a powerful machine learning algorithm that has improved image classification, it also has some limitations. One of the main limitations is that it can be computationally expensive and requires a lot of memory to run, especially when dealing with large datasets. Additionally, random forest can be prone to overfitting if the number of trees in the ensemble is too high. This can lead to poor generalization performance (Louppe, 2014). The incorporation of machine learning algorithms has improved image classification within remote sensing, which has led to more detailed analysis and better interpretation of remotely sensed data.

Previous studies have demonstrated the potential of high-resolution satellite data such as those from PlanetScope and Sentinel-2 for general cropland mapping and vegetation monitoring (Liu et al., 2022; Pickering et al., 2021; Rao et al., 2021; Zagajewski et al., 2024). However, there is still limited research specifically focused on extracting and comparing useful features from these datasets to accurately distinguish between different maize cropping systems over a large area. Features such as vegetation indices, spectral information, and temporal patterns have been used in various studies but a systematic comparison of these features in the context of maize crop systems is lacking. While high spectral resolution data can provide detailed information on crop health and stress through various vegetation indices, high spatial resolution data offers the advantage of capturing fine-scale spatial patterns within fields. Studies have shown the effectiveness of both types of data in general land cover mapping but there is a gap in understanding what type of data is more effective for distinguishing different maize cropping systems especially in regions where the landscape is heterogeneous. The specific spectral bands and spatial details necessary to differentiate maize crop fields from mixed crop systems have not been thoroughly explored. Therefore, the current gap lies in the comparative analysis of features extracted from high

spectral resolution (Sentinel-2) and high spatial resolution (PlanetScope) optical satellite data specifically for distinguishing maize crop systems.

This research aims to fill this gap by systematically comparing the outputs from Sentinel-2 and PlanetScope in effectively mapping different maize cropping systems. The focus will be on their efficacy in mapping and differentiating the various features that can help produce the best outputs. This comparison will help to determine the strengths and limitations of each dataset which can help contribute to improved methodologies for cropland monitoring and mapping.

1.2. Problem Statement

Heterogeneous landscapes are recognized as significant challenges for accurate land cover map production crucial for Natural Resource Management (NRM). Environmental changes over time are tracked by these maps making them essential for sustainable resource management and development. This is because the natural environment is relied upon by society for sustenance and economic stability. For instance, land cover maps are used by policy makers in making informed decisions regarding agriculture, forest monitoring and management among others which in tend affect many lives.

In Gaza Mozambique, agriculture particularly farming is characterized by smallholder farms, shifting cultivation, and mixed cropping systems. Local food production is important for supporting livelihoods. This makes accurate crop field monitoring essential. The high variability in farming practices combined with the heterogeneous landscape makes mapping land cover, such as agricultural fields particularly challenging in Gaza province, Mozambique.

Field based monitoring and measurement is recognized as expensive and time-consuming often requiring extensive manpower and resources to cover large areas effectively. It also poses logistical challenges and risks especially in remote or conflict-prone regions. In contrast, remote sensing is seen as a potentially cheaper and more objective alternative.

The research problem focuses on finding a cost-effective and efficient method to map land cover in the complex landscape of Gaza province, Mozambique. Field survey in this region is often expensive and time-consuming. This proves the need to find alternative means for mapping land cover in this region. As a result, this study relies on both a freely available dataset (Sentinel-2) and a commercial dataset (PlanetScope) which is also freely available upon request for specific purposes to assess cost effective and accurate means of mapping land cover in this region. By comparing the effectiveness of these two datasets using RF, this study aims to determine the most suitable approach for accurately differentiating maize cropping systems. This research is important as it seeks to provide reliable Insights into cost effective means of mapping which also guarantees accuracy and reliability.

1.3. Research Objectives, and Questions

The overall aim of this research is to assess the effectiveness of high spectral resolution (Sentinel-2) and high spatial resolution (PlanetScope) optical satellite to distinguish different maize cropping systems and compare their classification accuracies.

Research Objective 1

To assess the usefulness and relative importance of spectral indices, textural and topographic features from optical satellite data for the discrimination of different maize cropping systems using RF algorithm.

Research Question 1.1: Which spectral indices, textural metrics, topographic features, and spectral bands derived from optical satellite data are most effective in distinguishing different maize cropping systems.

Hypothesis1.1:

H0: All the spectral indices, textural metrics, topographic features, and spectral bands derived from optical satellite data will be effective in distinguishing different maize cropping systems.

H1: Some of the spectral indices, textural metrics, topographic features, and spectral bands derived from optical satellite data will be more effective than others in distinguishing different maize cropping systems.

Research Objective 2

To compare the classification accuracy of high spectral resolution (Sentinel-2) and high spatial resolution (PlanetScope) satellite imagery for the discrimination of different maize cropping systems.

Research Question 2.1: Is there a statistically significant difference in the classification accuracy between high spectral resolution (Sentinel-2) and high spatial resolution (PlanetScope) optical satellite imagery for discriminating different maize cropping systems?

Hypothesis2.1:

H0: There is no significant difference in the classification accuracy between high spectral resolution (Sentinel-2) and high spatial resolution (PlanetScope) optical satellite imagery for discriminating different maize cropping systems.

H1: There is a significant difference in the classification accuracy between high spectral resolution (Sentinel-2) and high spatial resolution (PlanetScope) optical satellite imagery for discriminating different maize cropping systems.

2. METHODS

This chapter contains a description of the study area, the data, and the methods used to achieve the objectives of this research.

2.1. Study area

This study is conducted in Gaza Province situated in southern Mozambique which can be found in the southern part of the African continent. Gaza Province spans an extensive area of 75,709 square kilometers and features diverse landscapes including wide plains ideal for agriculture (Ponguane et al., 2023). The landscape includes unique ecological areas such as peat bogs, raised moors, and significant waterways like the Limpopo and Changane rivers (Saveca et al., 2022).

The province experiences a tropical savanna climate characterized by distinct wet and dry seasons, with average annual temperatures ranging from 23°C along the coast to 25°C inland and annual rainfall of approximately 600mm (Muhala et al., 2021). The geology of Gaza Province varies from sandy soils in the southeastern regions to clay sediments and traces of swamps in the northern parts (Lächelt, 2004). According to the 2017 population census, the population of this region is approximately 1.4 million (Azar et al., 2023).

Agriculture is the primary economic activity in Gaza Province, with maize, rice, and cassava being among the main crops cultivated (Salite, 2019). The region predominantly relies on subsistence farming characterized by smallholder farms, shifting cultivation, and mixed cropping methods. Agriculture is not just an economic activity but a vital part of life for the inhabitants with many families depending directly on their farming outputs for food, income, and overall livelihood. This dependence on agriculture means that the success of crop yields directly impacts the community's well-being, influencing food security and economic stability.

Over the last 30 years, Gaza Province has seen rapid changes, particularly in response to climate variability, making it a key area for agricultural research. (World Bank, 2017). This research is fundamental to explore cost-effective methods for providing reliable and accurate spatial information which is essential for informed agricultural policy making.

2.2. Data

To achieve the research objectives, this study used satellite imagery from Sentinel-2 and PlanetScope to map maize cropping systems within the study area. These datasets include both vector and raster data which served different purposes. The data utilized is detailed below in **Table 1**.

Table 1: Summary of Datasets used in this study.

Dataset	Type	Source	Purpose
Country Administrative Data	Vector data	Mozambique National Cartography and Remote Sensing Centre (CENACARTA)	Administrative boundaries and geographical context
Sentinel-2 Satellite Image	Raster data	Copernicus Open Access Hub (https://scihub.copernicus.eu/dhus)	High-resolution spectral data for land cover analysis.
Google Earth Imagery	Raster data	Google Earth	High-resolution imagery for additional training and validation samples collection
PlanetScope Satellite Image	Raster data	https://developers.planet.com/docs/data/PlanetScope	High spatial resolution data for land cover analysis
Dynamic World Land Cover Map	Raster data	https://dynamicworld.app/explore	Used as guide for collection of additional training samples
Cropland Boundary	Vector data	Stats from Space project (ITC)	Delineating the extent of croplands within the study area
Field Observations	Vector data	Stats from Space project (ITC)	Observations of crop types
Esa Land Cover	Raster data	European Space Agency (ESA) (https://www.esa-landcover-cci.org)	Used as guide for collection of additional training samples

2.2.1. Remote Sensing data

Harmonized atmospherically corrected Sentinel-2 images acquired throughout the year 2023 were considered for the study. Before further analysis, a cloud masking function was applied to reduce noise caused by atmospheric conditions. This process was conducted using the Google Earth Engine platform which facilitates efficient processing and analysis of large satellite datasets. Table 2 below details the sensor properties of the Sentinel-2 satellite.

Table 2: Sentinel-2 spectral bands and spatial resolution

Band	Description	Wavelength (nm)	Spatial Resolution (m)
B2	Blue	490	10
B3	Green	560	10
B4	Red	665	10
B5	Red Edge 1	705	20
B6	Red Edge 2	740	20
B7	Red Edge 3	783	20
B8	Near-Infrared (NIR)	842	10
B8A	Narrow NIR	865	20
B11	Short-Wave Infrared (SWIR) 1	1610	20
B12	Short-Wave Infrared (SWIR) 2	2190	20

This study also made use of high-resolution imagery from PlanetScope acquired throughout the year 2023. Unlike Sentinel-2, PlanetScope imagery did not require a cloud masking process as the daily revisit capability allowed for the selection of cloud-free images. This high temporal resolution ensures that enough clear images are available throughout the duration of this study. This reduced the need for cloud filtering and allowed for consistent data acquisition. PlanetScope data was also processed using the Google Earth Engine platform.

Table 3: PlanetScope spectral bands and spatial resolution

Band	Description	Wavelength (nm)	Spatial Resolution (m)
B	Blue	455-515	3-5
G	Green	500-590	3-5
R	Red	590-670	3-5
NIR	Near-Infrared (NIR)	780-860	3-5

2.2.2. Field Data.

This study used data collected from the field within the study area to complement existing datasets. The field data consisted of geotagged photos from various agricultural fields within the study area. This data was collected using mobile phones which contained inbuilt GPS. These geotagged photos provided the precise geolocations of different cropland types within the study area. Among the numerous farms in the study area, three primary cropping patterns stand out: maize fields, mixed fields, and rice fields. This pattern largely influenced the categorization of the croplands that can be mapped.

Given the predominance of maize fields, mixed fields, and rice fields within the field data collected, the classification scheme for cropland mapping had to be adapted to accurately reflect these patterns. The intercropping practices particularly in the mixed fields where crops such as cassava and cowpeas, groundnuts among others are often grown together with maize posed a significant challenge for distinct crop mapping. Therefore, the classification performed in this study focused on these major categories ensuring that the primary agricultural landscapes were accurately represented. See (Figure 2) for an example of how some of the crop fields appear within the study area.

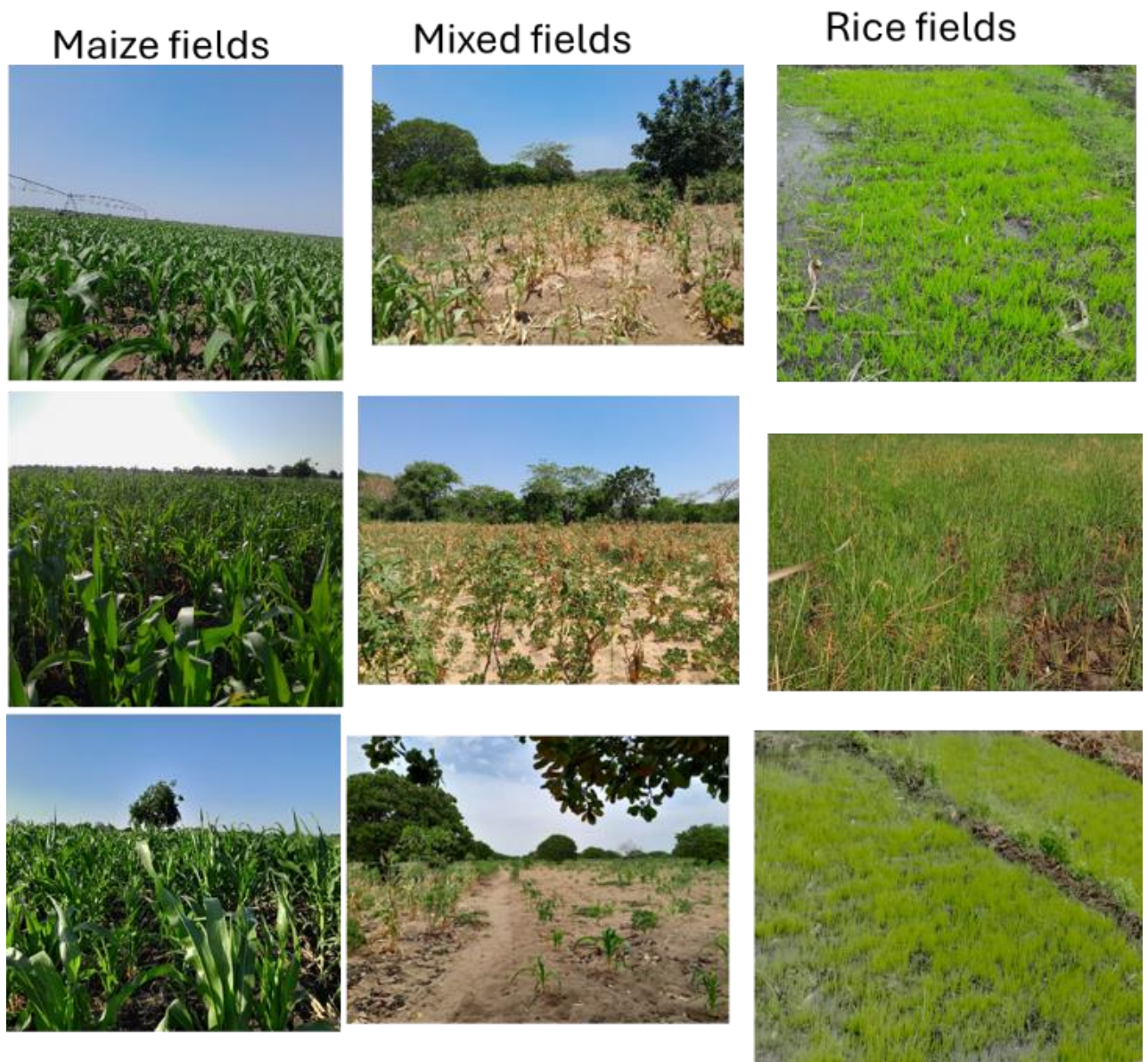


Figure 2: Photos showing Maize, Mixed crop fields and Rice fields.

2.2.3. Existing Landcover Products.

This study used two existing land cover products to guide the collection of training samples for other landcover classes present within the study area for supervised image classification. This action was computed in the Google Earth Engine (GEE) platform. These datasets included the Dynamic World Land Cover Map and the ESA Land Cover dataset.

The Dynamic World Land Cover Map which has been developed jointly by Google and the World Resources Institute provided up-to-date classifications of various land cover types. This dataset is a 10-meter resolution near real-time global land cover dataset. The Dynamic World Land Cover Map was produced using a deep learning method. It offers per-pixel probabilities across nine land cover classes and updates globally every 2-5 days depending on location (Dynamic World, 2023). The ESA Land Cover dataset (Raster data) which has been developed by the European Space Agency (ESA) is an essential resource for land cover classification and analysis.

The ESA World Cover product provides global land cover maps at a 10-meter resolution. This dataset fuses data from Sentinel-1 and Sentinel-2 satellites to produce a global landcover products which is made up of 11 land cover classes. These landcover classes are in alignment with the UN FAO's Land Cover Classification System to offer a broad view of the Earth's surface. Both products were loaded into GEE and intersected with high-resolution imagery of the study area for the year 2023 which was provided within the GEE platform. This intersection aimed to verify the land cover within the study area to facilitate the collection of accurate additional training samples. A summary table for the land cover class distribution and description for both Dynamic World and ESA land cover products can be found in the Appendix.

The flow chart below (**Figure 3**) represents the overall methodology used to analyze the data for distinguishing maize cropping systems in the Gaza Province of Mozambique. The methodology involves several key steps, each contributing to the comprehensive analysis and classification process.

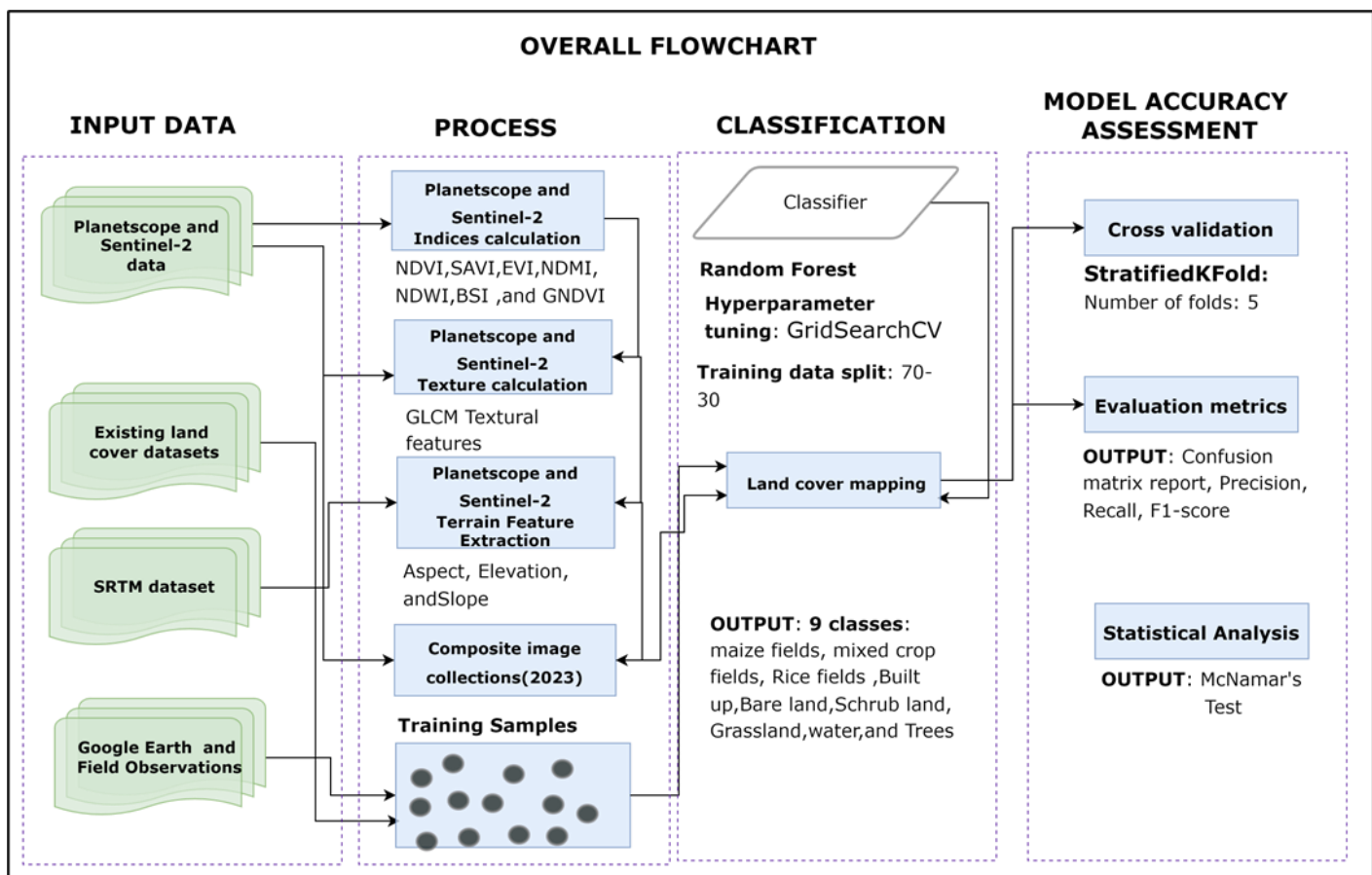


Figure 3: Overall flowchart of the study methods. This flowchart outlines the process of data acquisition, preprocessing, and analysis.

2.3. Data pre-processing

Data pre-processing is a critical step in ensuring the quality and reliability of remote sensing data for land cover mapping. This process involves a series of techniques to enhance image clarity, minimize atmospheric noise, and prepare the data for analysis (Macarringue et al., 2022). In this study pre-processing was applied to the Sentinel-2 satellite imagery to ensure high-quality inputs for further analysis. The next section details the specific pre-processing steps undertaken for the Sentinel-2 data.

2.3.1. Sentinel-2 Image pre-processing.

In this study a thorough pre-processing was conducted on the Sentinel-2 satellite imagery. This was to ensure high-quality data for mapping maize cropping systems in the Gaza province of Mozambique. The Sentinel-2 image collection from the COPERNICUS/S2_SR_HARMONIZED dataset was used encompassing images acquired throughout the year 2023. To enhance image clarity and minimize atmospheric noise, a few additional key steps were undertaken. Firstly, a cloud masking function “maskS2clouds” was applied using the QA60 band to identify and mask pixels affected by clouds and cirrus. This method is effective for reducing atmospheric interference. Such cloud masking techniques are widely used in remote sensing to improve data quality (Mateo-García et al., 2018). Secondly, pre-filtering was performed to select images with less than 10% cloud cover. This was done to ensure the inclusion of only the clearest images. This step is crucial as it significantly reduces the impact of cloud cover which can conceal surface features and affect the accuracy of land cover classification. Furthermore, the Cloud Score+ collection (GOOGLE/CLOUD_SCORE_PLUS/V1/S2_HARMONIZED) was integrated with the Sentinel-2 images to enhance cloud detection and masking. This integration was achieved using the linkCollection function in GEE. This cloud identification process further improved the quality of the dataset.

2.3.2. Image composites calculations.

After performing masking on the Sentinel-2 imagery, the next step involved calculating a median composite image. A median composite is created by determining the median value for each pixel across a series of images. This approach is aimed at reducing noise and eliminating anomalies like clouds thereby providing a clearer and more consistent representation of the land surface. This technique is fundamental for enhancing data quality and ensuring accurate land cover classification, as demonstrated in studies such as Phan et al. (2020). See **Figure 5** for a snapshot of the before and after the application of cloud masking and the computation of a median composite on the Sentinel-2 imagery. For this study, four distinct median composites: annual, dry season, wet season, and a multi-stack composite that combines the annual, dry, and wet season composites to capture temporal variations were created.

The annual median composite was created by collecting multiple images of the same area taken throughout the year 2023. For each pixel location the pixel values are sorted over time and the median value is calculated and assigned to the corresponding pixel in the composite image. This procedure is repeated for both the dry season and wet season composites for the year 2023. Finally, the three composite images were stacked on top of each other to create a multi-stack image composite. The same procedure except for cloud masking, was performed on the PlanetScope satellite imagery as well. The entire workflow was carried out in GEE. **Figure 4** below presents an illustration of the workflow used for the generation of an annual median composite image for the Sentinel-2 satellite imagery.

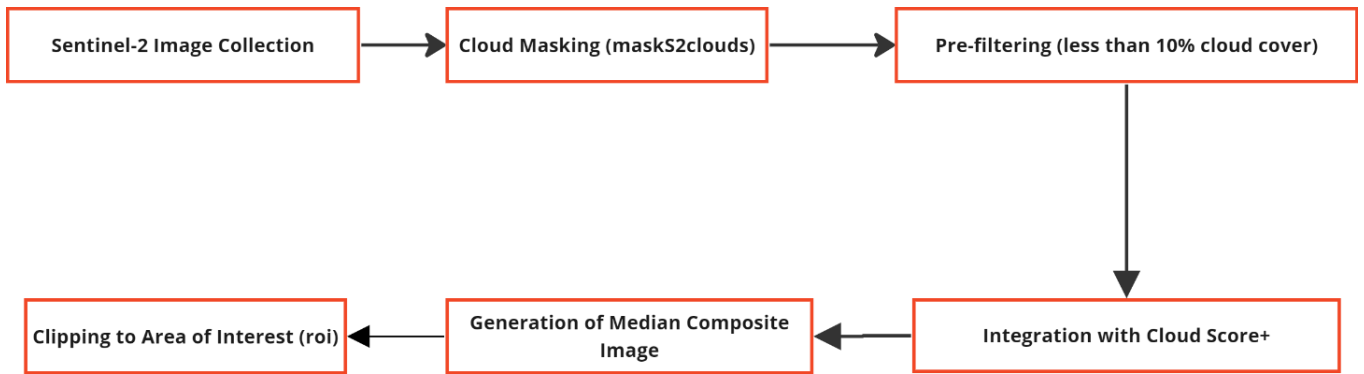


Figure 4: Workflow for Pre-Processing Sentinel-2 Imagery

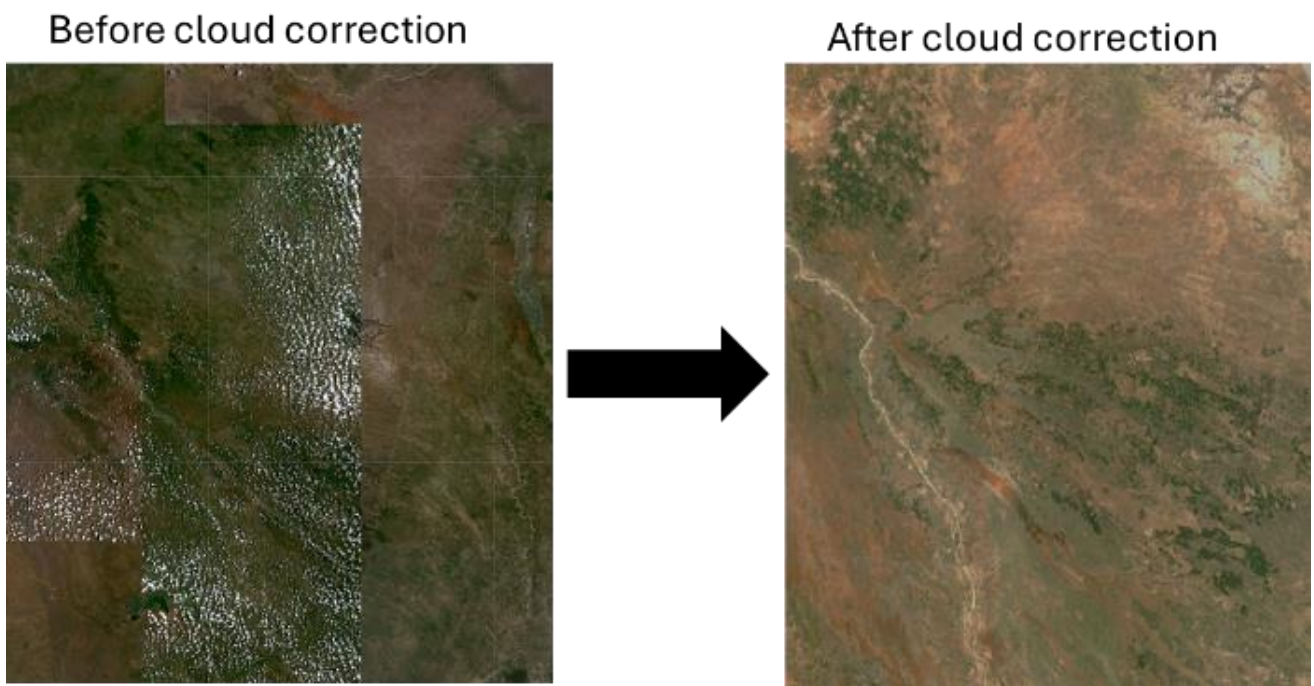


Figure 5: Comparison of Raw and Pre-Processed Sentinel-2 Imagery after cloud removal and median composite calculation.

2.4. Preparation and refinement of field observations

To prepare the field data to be used as input into the RF model which as later implemented, a few systematic steps were followed to ensure data accuracy. Firstly, already existing plot segments delineated as polygons for the study area provided the necessary information regarding the spatial location of different land cover classes or features within the study area. A spatial selection was done to select all cultivated crop fields identified for the study area.

Geotagged field photos were collected from various identified farms spread across the study area using a mobile phone GPS. The geotagged field photos were then converted to vector data (points) using the “GeoTagged Photos To Points” tool within the ArcGIS software (Esri, 2020). The conversion of these geotagged photos to point data was a crucial step because it allowed for the accurate association of crop types identified from the field observation with the crop field segmented plots. Furthermore, the converted point data was intersected with the crop segment polygon data and only polygons which fully contained point observations were selected for further analysis. The points that did not fall within any crop segment (polygon) were discarded indicating positional error on the part of the GPS devices used in collecting the field observations. This action was to ensure data quality to generate reliable outputs. The extracted crop field polygons were subsequently overlaid on high-resolution PlanetScope satellite imagery in the Google Earth Engine platform and inspected visually for consistency. This visual inspection step was necessary to validate the field data against the satellite imagery.

The aim was to ensure that the polygons accurately represented the crop fields. To further enhance data quality, the K-means clustering method was performed on the satellite imagery clipped with the crop field polygons. See **(Figure 6)**. The K-means parameter like number of clusters used varied depending on the land cover composition of each plot. Typically, clustering was performed with (n) clusters ranging from 2 to 4. This means that the data points were divided into 2 to 4 distinct groups which was sufficient to capture the spectral variations within each plot.

For a description of the K-means algorithm, see (Background) in this study's introduction. This step was vital as it enabled the refinement of the initial polygons to ensure that the training data was representative of actual crop fields not other land cover classes that were not relevant to the study. The refined polygons were merged to form one wholistic polygon containing all the individual crop fields boundaries. After that, the next step was to convert the polygon data to point data. Converting the training polygons to points has several benefits for many tasks performed in GEE. Firstly, this action reduces the computational load and memory usage making data processing more efficient. This simplification enhances spatial analysis by decreasing data volume hence leading to faster and more manageable workflows. Additionally, points ensure consistent data sampling across regions improving the reliability and accuracy of geospatial analysis. A flow diagram illustrating the entire workflow for preparing and refining field data is shown below in **(Figure 7)**

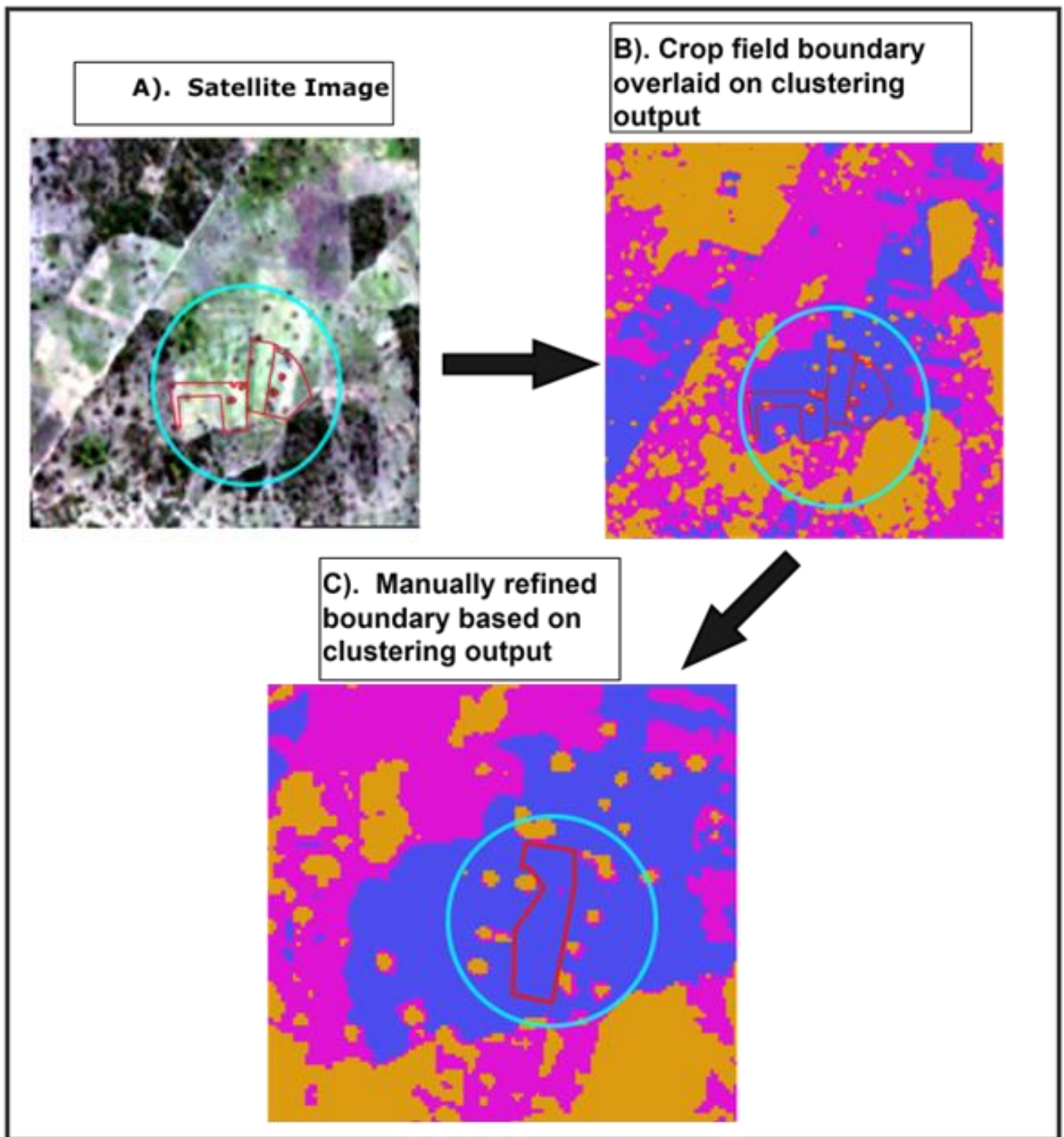


Figure 6: Refinement of Crop Field Polygon Data Using K-means Clustering Algorithm

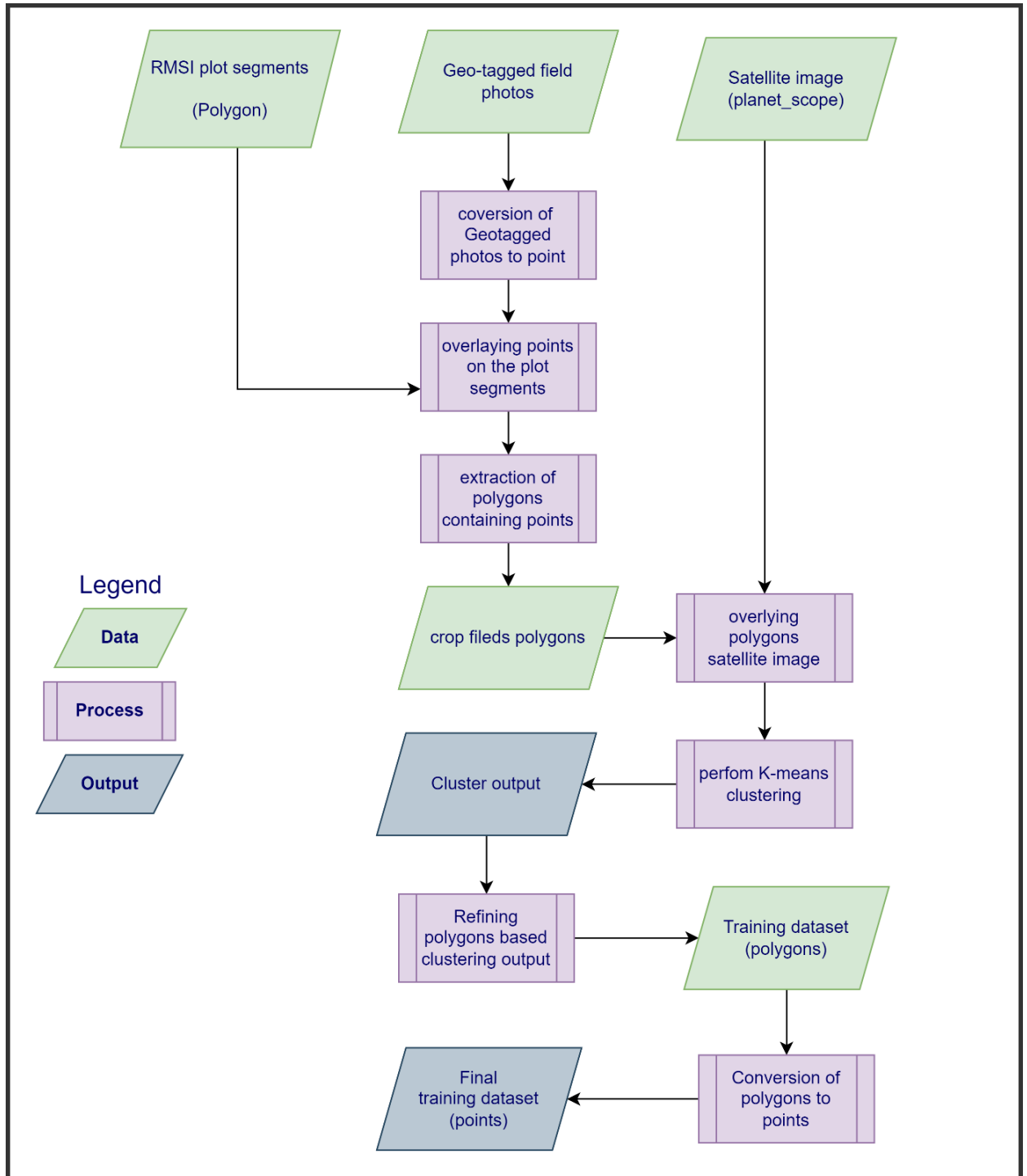


Figure 7: Workflow for refining crop field data to be used as input into the RF model.

2.5. Feature Extraction

Feature extraction is an important step in any image-based classification. This process involves transforming raw data into a set of measurable characteristics or features that can be used to categorize or analyze the image. The focus of image classification research has always been on feature extraction as it forms the basis of the classification process by isolating specific geometric, textural, and spectral information that distinguishes one land cover type from another. (Lu et al., 2023; Puls et al., 2023). This is

important for achieving precise and accurate classification outputs in remote sensing applications (Lu et al., 2023). For this study, four main different types of features were considered namely **spectral, textural, spectral indices and topographic features**.

A) Spectral features from Sentinel-2 and PlanetScope image

In this study, all bands from the Sentinel-2 satellite were considered for further analysis. This approach included the visible, near-infrared (NIR), and shortwave infrared (SWIR) bands as well as the red edge bands. This selection was based on common practices observed in previous studies on similar topics. For example, study by Inglada et al., (2022) has demonstrated the value of using the full spectral range of Sentinel-2 for diverse land cover classification tasks. This justifies the inclusion of all available bands in this research. Just like the case for the Sentinel-2 image, all the 4 spectral bands of PlanetScope were also considered for further analysis. See **(Table 2) and (Table 3)** for information regarding the band specifications of the Sentinel-2 and PlanetScope satellite respectively.

B) Textural Features

Texture is an important aspect of all satellite images, describing how pixel values are spread out across an image and revealing patterns that help us understand and categorize the content (Pratt et al., 1978). Extracting these textural features helps identify the spatial arrangement of pixels which is especially useful in remote sensing for classifying different types of land cover or crops. In image analysis various methods are employed to derive textural information that can aid in classification tasks.

In this study, the Gray-Level Co-occurrence Matrix (GLCM) method was implemented to extract textural features for the classification task. The Gray Level Co-occurrence Matrix (GLCM) is a robust method for extracting textural features from images particularly in the context of remote sensing. This method was developed by (Haralick et al., 1973). GLCM analyzes the spatial distribution of pixel intensities in an image by creating a matrix that captures how often pairs of pixel values occur in specific spatial relationships. This method transforms raw image data into a set of statistical measures that describe texture, making it useful for classifying different land cover types or crops. A total of 18 GLCM textural features can be derived from the spectral bands of interest for both the Sentinel-2 and PlanetScope satellite images. These are Angular Second Moment (asm), Contrast (contrast), Correlation (corr), Sum of Squares, Variance (svar), Inverse Difference Moment (idm), Sum Average (savg), Sum Variance (var), Sum Entropy (sent), Entropy (ent), Difference Variance (dvar), Difference Entropy (dent), Information Measures of Correlation1 (imcorr1), Information Measures of Correlation2 (imcorr2), Maximal Correlation Coefficient (maxcorr), Dissimilarity (diss), Inertia (inertia), Prominence (prom), and Shade (shade). However, for this study a total of seven features were used. See **(Table 4 & 5)**. This selection was based on their proven effectiveness in similar studies. The aim of this action was to increase computational efficiency and reduce redundancy and irrelevant features (Clausi, 2002).

Table 4: GLCM Textural Features Computed for Sentinel-2 Bands

GLCM Feature	Sentinel-2 Bands	Computed Textures
Angular Second Moment (Asm)	B4 (Red), B8 (NIR), B3 (Green), B7 (Red Edge 3)	asm_B4, asm_B8, asm_B3, asm_B7
Contrast	B4 (Red), B8 (NIR), B3 (Green), B7 (Red Edge 3)	contrast_B4, contrast_B8, contrast_B3, contrast_B7
Correlation (Corr)	B4 (Red), B8 (NIR), B3 (Green), B7 (Red Edge 3)	corr_B4, corr_B8, corr_B3, corr_B7
Inverse Difference Moment (Idm)	B4 (Red), B8 (NIR), B3 (Green), B7 (Red Edge 3)	idm_B4, idm_B8, idm_B3, idm_B7
Entropy	B4 (Red), B8 (NIR), B3 (Green), B7 (Red Edge 3)	ent_B4, ent_B8, ent_B3, ent_B7
Dissimilarity (Diss)	B4 (Red), B8 (NIR), B3 (Green), B7 (Red Edge 3)	diss_B4, diss_B8, diss_B3, diss_B7
Variance	B4 (Red), B8 (NIR), B3 (Green), B7 (Red Edge 3)	var_B4, var_B8, var_B3, var_B7

Table 5: GLCM Textural Features Computed for PlanetScope Bands

GLCM Feature	PlanetScope Bands	Computed Textures
Angular Second Moment (Asm)	R (Red), N (NIR), G (Green)	R_asm, N_asm, G_asm
Contrast	R (Red), N (NIR), G (Green)	R_contrast, N_contrast, G_contrast
Correlation (Corr)	R (Red), N (NIR), G (Green)	R_corr, N_corr, G_corr
Inverse Difference Moment (Idm)	R (Red), N (NIR), G (Green)	R_idm, N_idm, G_idm
Entropy	R (Red), N (NIR), G (Green)	R_ent, N_ent, G_ent
Dissimilarity (Diss)	R (Red), N (NIR), G (Green)	R_diss, N_diss, G_diss
Variance	R (Red), N (NIR), G (Green)	R_var, N_var, G_var

C) Vegetation Indices

Vegetation indices are mathematical combinations of spectral bands from remote sensing images designed to highlight specific characteristics of the Earth's surface. Vegetation indices are important in various fields such as agriculture, forestry, and environmental monitoring because they transform raw spectral data into interpretable metrics. This enhances our ability to monitor, analyze, and manage natural resources effectively (Prasad et al., 2022; Tran et al., 2022). In this study, seven vegetation indices were calculated to be used as part of the input features to the classification model. Five of these indices were consistently computed across both the Sentinel-2 and PlanetScope image. These indices included the Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), Normalized Difference Water Index (NDWI), Green Normalized Difference Vegetation Index (GNDVI), and Soil-Adjusted Vegetation Index (SAVI). The other two indices the Normalized Difference Moisture Index (NDMI) and Bare Soil Index (BSI) were specifically computed using the Sentinel-2 imagery due to the availability of additional spectral bands which are not present in the PlanetScope data.

Table 6: Summary Table of Vegetation indices Calculated for Sentinel-2 and PlanetScope Images

Index	Image	Formula	Description & Explanation
NDVI	Sentinel-2, PlanetScope	$(\text{NIR} - \text{Red}) / (\text{NIR} + \text{Red})$	Indicates vegetation health and density. High values suggest healthy, dense vegetation.
EVI	Sentinel-2, PlanetScope	$2.5 * ((\text{NIR} - \text{Red}) / (\text{NIR} + 6 * \text{Red} - 7.5 * \text{Blue} + 1))$	Enhances vegetation signal with improved sensitivity in high biomass regions and better noise reduction.
NDWI	Sentinel-2, PlanetScope	$(\text{NIR} - \text{Green}) / (\text{NIR} + \text{Green})$	Reflects water content in vegetation and soil. Higher values indicate higher water content.
GNDVI	Sentinel-2, PlanetScope	$(\text{NIR} - \text{Green}) / (\text{NIR} + \text{Green})$	Sensitive to chlorophyll content in vegetation, useful for monitoring plant health.
SAVI	Sentinel-2, PlanetScope	$((\text{NIR} - \text{Red}) / (\text{NIR} + \text{Red} + 0.5)) * 1.5$	Similar to NDVI but adjusts for soil brightness, improving vegetation detection in areas with less dense canopy.
NDMI	Sentinel-2	$(\text{NIR} - \text{SWIR}) / (\text{NIR} + \text{SWIR})$	Indicates moisture content in vegetation and soil. Higher values signify more moisture.
BSI	Sentinel-2	$((\text{SWIR} + \text{Red}) - (\text{NIR} + \text{Blue})) / ((\text{SWIR} + \text{Red}) + (\text{NIR} + \text{Blue}))$	Differentiates bare soil from vegetation and other land covers. Higher values indicate bare soil presence.

D) Topographic Features

Topographic features provide important information about the physical characteristics of the Earth's surface. The use of terrain features is essential for remote sensing applications, including image classification (Marzi et al., 2023). These features are derived from Digital Elevation Models (DEMs) and include metric such as elevation, slope, aspect, curvature, and roughness. For this study, the Shuttle Radar Topography Mission (SRTM) dataset was used to extract aspect, elevation and slope data which were then used as input features for the RF model. SRTM data provides detailed elevation models with a resolution of 30 meters at a global scale. This makes it a useful resource for various applications including remote sensing image analysis.

2.5.1. Preparation of input features for RF model

For this study, the input features were prepared in GEE and then exported to a local Python environment to facilitate various experimental setups. This step was essential because it allowed for the manipulation and testing of different configurations of the input features. This was done to ensure that a comprehensive analysis was performed on the input features and ensure the RF is well optimized. The workflow for creating input features for the RF model in GEE involved several key steps. Firstly, the study area boundary and training points were imported to ensure the entire process was restricted to the region of interest. The various image composites as described in section 2.4.2 served as the basis for the extraction of key spectral indices and textural features such as NDVI, EVI, NDWI, NDMI, SAVI, BSI, and GNDVI as well as the Grey Level Co-occurrence Matrix (GLCM) textural features. Topographic data such as elevation, slope, and aspect, were incorporated to account for terrain influences. The next step involved the combination of all processed into a single collection and sampled using the training points. The geometry parameter was set to true to retain the spatial information of each sample point. Finally, the data is exported as a CSV file for further analysis. This process was performed on both satellite imagery thus Sentinel-2 and PlanetScope. This workflow is illustrated in (figure 8) below. The exported CSV file containing all the input features was then joined with the training points shapefile using a unique ID common to both datasets. This step was to ensure that each point data is properly linked to the corresponding feature data.

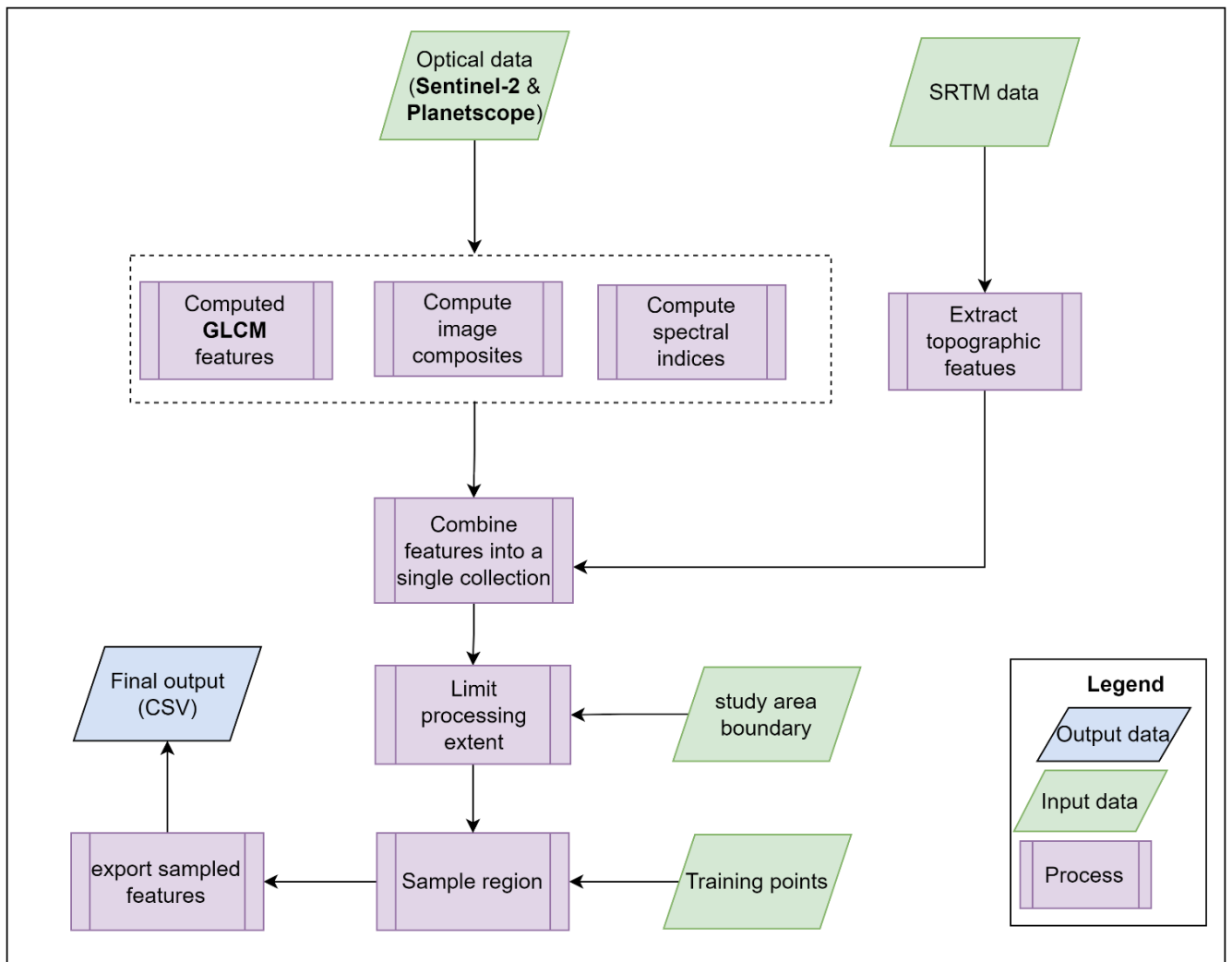


Figure 8 : workflow for preparing input features in GEE.

2.5.2. Correlation analysis on input features.

To ensure relevance and non-redundancy of the input features to the RF model, this study implemented a Pearson correlation analysis on the spectral bands, topographic features, spectral indices, and GLCM texture features. Pearson correlation is a commonly used statistical method for assessing linear relationships between pairs of variables. This helps to identify variables that are highly correlated and might contribute redundant information to the model. By reducing redundancy, the RF model can become more efficient and effective. This is because machine learning models rely on the most informative and independent features for decision-making. This method has been extensively documented in scientific literature for its effectiveness in such analyses (Schober & Schwarte, 2018).

The results of the correlation analysis for both datasets (Sentinel-2 and PlanetScope) were used to organize the input features into different experimental groups. Each group represented a separate combination of features, allowing for a comparative analysis to determine which set produced the best classification outcome. This approach not only helps in improving model performance but also contributes to a better understanding of the significance of each feature set in the classification process.

2.5.3. Input features experimental setup.

In total, four experiments were setup on the various image composites in line with the correlation coefficients of the input features. This approach was applied to both datasets thus Sentinel-2 and PlanetScope. The experiments were designed with the following goals: **Experiment 1** aimed to include moderately correlated features for a balanced dataset. **Experiment 2** focused on highly correlated texture features and essential indices for vegetation and moisture. **Experiment 3** combined various vegetation indices with detailed texture features for comprehensive analysis. **Experiment 4** used features with varied correlation levels for extensive coverage.

Topographic features were constant in all experiments for both datasets because they were not highly correlated with any spectral bands, texture features, or vegetation indices. This constancy ensured that topographic information was always considered in the classification process without contributing to redundancy.

Sentinel-2 Experiments:

- **Experiment 1:** This experiment included moderately correlated features for a balanced dataset. The texture features used were B4_contrast, B8_contrast, B3_diss, and B7_var. The indices included EVI, NDWI, and BSI. The spectral bands were B2, B5, B8, and B11, along with the topographic features slope, elevation, and aspect.
- **Experiment 2:** This experiment focused on highly correlated texture features and essential indices for vegetation and moisture. The texture features were B4_var, B8_diss, B3_var, and B7_contrast. The indices included NDVI, SAVI, and NDMI. The spectral bands used were B3, B6, B8A, and B12, with the topographic features slope, elevation, and aspect.
- **Experiment 3:** This experiment combined various vegetation indices with detailed texture features for comprehensive analysis. The texture features included B4_diss, B8_var, B3_contrast, and B7_diss. The indices were EVI, NDVI, SAVI, GNDVI, NDMI, and BSI. The spectral bands were B3, B4, B5, B7, and B8A, along with slope, elevation, and aspect.

Experiment 4: This experiment used features with varied correlation levels for extensive coverage. The texture features were B4_contrast, B8_contrast, B3_var, and B7_var. The indices included EVI, NDWI, and GNDVI. The spectral bands were B11, B2, B5, B6, and B8, along with the topographic features slope, elevation, and aspect.

PlanetScope Experiments:

- **Experiment 1:** This experiment selected low to moderately correlated features to cover a broad spectrum. The texture features used were R_contrast, G_contrast, N_diss, and G_var. The indices included EVI, NDWI, and SAVI. The spectral bands were B, G, and N, along with the topographic features slope, elevation, and aspect.
- **Experiment 2:** This experiment chose features with low to moderate correlations for a diverse dataset. The texture features were R_diss, N_var, G_diss, and G_var. The indices included NDVI, GNDVI, and SAVI. The spectral bands were R, G, and N, with the topographic features slope, elevation, and aspect.
- **Experiment 3:** This experiment selected a range of correlations to cover more variance. The texture features used were R_var, N_contrast, R_diss, and G_diss. The indices included EVI, NDVI, GNDVI, and NDWI. The spectral bands were B, R, and G, along with slope, elevation, and aspect.
- **Experiment 4:** This experiment included a wide range of data with varied correlations for comprehensive analysis. The texture features used were R_contrast, N_contrast, G_diss, and G_var. The indices included NDWI, GNDVI, and SAVI. The spectral bands were B, R, G, and N, along with the topographic features slope, elevation, and aspect.

By structuring the input features into these experimental groups, the study aimed to identify the most effective combination for classifying maize cropping systems and to understand the impact of different feature sets on model performance. A summary of the experimental setup can be found below in **(table 7 & 8)** for Sentinel-2 and PlanetScope datasets respectively.

Table 7: Summary of Experimental Setup for Sentinel-2 Datasets

Experiment	Texture Features	Indices Features	Bands Features	Terrain Features	Reason
1	B4_contrast, B8_contrast, B3_diss, B7_var	EVI, NDWI, BSI	B2, B5, B8, B11	slope, elevation, aspect	To include moderately correlated features for a balanced dataset.
2	B4_var, B8_diss, B3_var, B7_contrast	NDVI, SAVI, NDMI	B3, B6, B8A, B12	slope, elevation, aspect	To focus on highly correlated texture features and essential indices for vegetation and moisture.
3	B4_diss, B8_var, B3_contrast, B7_diss	EVI, NDVI, SAVI, GNDVI, NDMI, BSI	B3, B4, B5, B7, B8A	slope, elevation, aspect	To combine various vegetation indices with detailed texture features for comprehensive analysis.
4	B4_contrast, B8_contrast, B3_var, B7_var	EVI, NDWI, GNDVI	B11, B2, B5, B6, B8	slope, elevation, aspect	To use features with varied correlation levels for extensive coverage

Table 8: Summary of Experimental Setup for PlanetScope Datasets

Experiment	Texture Features	Indices	Bands	Topographic Features	Reason
1	R_contrast, G_contrast, N_diss, G_var	EVI, NDWI, SAVI	B, G, N	slope, elevation, aspect	Selected low to moderately correlated features to cover a broad spectrum.
2	R_diss, N_var, G_diss, G_var	NDVI, GNDVI, SAVI	R, G, N	slope, elevation, aspect	Chose features with low to moderate correlations for a diverse dataset.
3	R_var, N_contrast, R_diss, G_diss	EVI, NDVI, GNDVI, NDWI	B, R, G	slope, elevation, aspect	Selected a range of correlations to cover more variance.
4	R_contrast, N_contrast, G_diss, G_var	NDWI, GNDVI, SAVI	B, R, G, N	slope, elevation, aspect	Included a wide range of data with varied correlations for comprehensive analysis.

2.5.4. Implementation of Random Forest

In this study, the RF classifier was used to perform supervised classification on the various combinations of input features derived from Sentinel-2 and PlanetScope datasets. To ensure an unbiased assessment, the input data was split into training (70%) and testing (30%) sets using a stratified approach to ensure a balance class distribution. This is a typical split also used in other studies for training random forests on land cover classification (Eisavi et al., 2015; C. Zhou et al., 2022). See Figure 9 for a summary of the distribution of training and validation points across the various land cover classes respectively.

The ‘random_state’ was controlled for reproducibility and set to 100 ensuring that the same random processes such as data shuffling and model initialization produce consistent splits and results across different runs (Pedregosa et al., 2011). This action enhances the reliability and comparability of the outputs generated. A description of the RF model and its characteristics can be found within the introduction section of this study. Hyperparameter tuning was conducted using ‘GridSearchCV’ with a 5-fold ‘StratifiedKfold’ cross-validation. This meant that the 70% training data was divided into 5 folds using the ‘StratifiedKfold’ method to ensure that each fold preserved the class distribution.

GridSearchCV is a technique used in machine learning to perform hyperparameter tuning (Pedregosa et al., 2011). It systematically works through multiple combinations of hyperparameter values cross-validating each combination to determine the best-performing parameters (Yang & Shami, 2020). Hyperparameter tuning in general is the process of optimizing the parameters that control the learning process of a machine learning model (Yang & Shami, 2020). Through the ‘GridSearchCV’, parameters such as the number of trees (n_estimators), tree depth (max_depth), minimum samples required to split a node (min_samples_split), minimum samples required at each leaf node (min_samples_leaf), and the use of bootstrap samples (bootstrap) were explored. The choice of cross-validation in this study was motivated

by the need to ensure a robust and unbiased evaluation of the model's performance. This method divides the training data into five equal parts or folds. After which the model iteratively trains on four folds while validating it on the remaining one-fold. This process is repeated five times with each fold serving as the validation set once. By doing so ensures that every data point is used for both training and validation. This provides a thorough assessment of the model's performance and reduce the risk of overfitting.

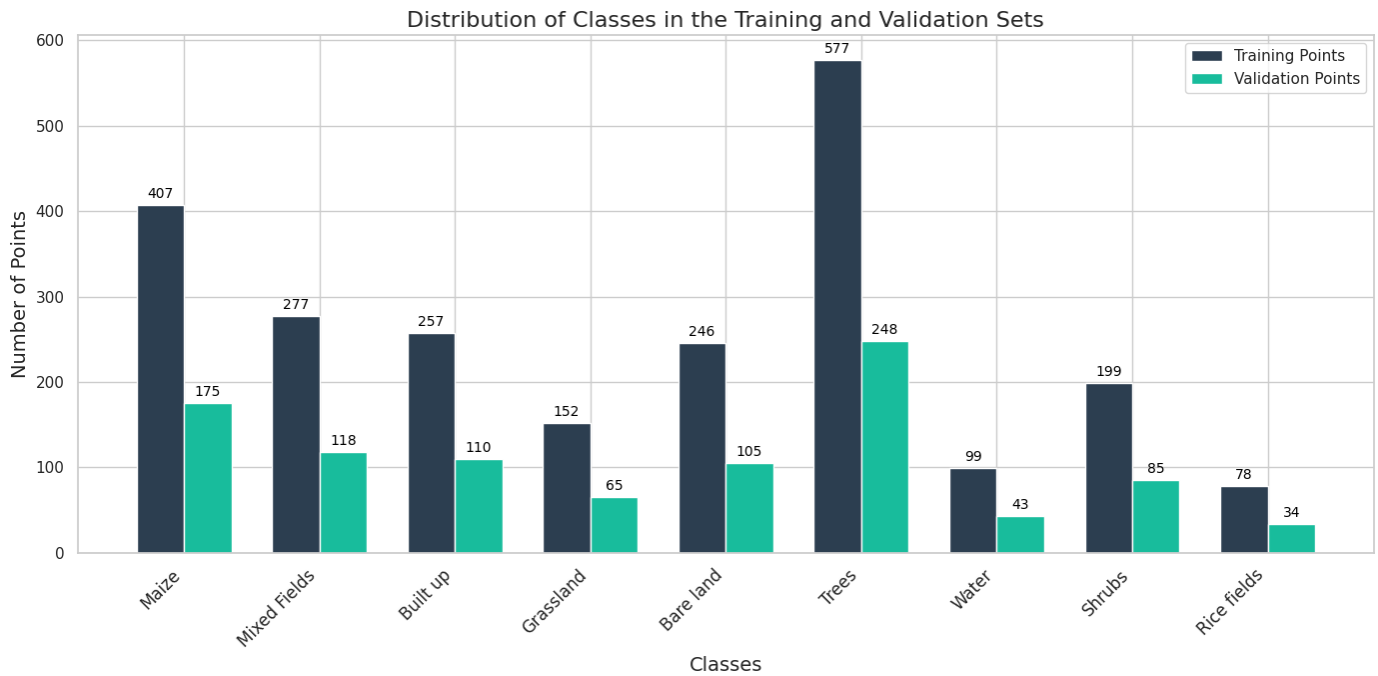


Figure 9: Distribution of training and validation points across different land cover classes.

2.5.5. Variable Importance Analysis

This study used three methods to assess variable importance: mean decrease Gini, SHAP (SHapley Additive exPlanations), and permutation importance.

Mean decrease Gini measures feature importance by calculating the average reduction in Gini impurity when a feature is used to split a node in tree-based models like Random Forests (Sandri & Zuccolotto, 2010).

SHAP values explain each feature's impact on model predictions by distributing the prediction among features, considering all possible subsets (Lundberg et al., 2017).

Permutation importance measures the change in model accuracy when feature values are shuffled (Pereira et al., 2021). Comparing these methods was aimed at gathering a complete understanding of feature importance from different perspectives. Each method offers unique advantages. For instance, mean decrease Gini is efficient for tree-based models, SHAP offers consistent and detailed explanations. Finally, permutation importance is flexible and applicable to any model.

2.5.6. Evaluation of Classification Model

In this study, the performance of the RF model in predicting the land cover classes of interest was evaluated using key metrics such as precision, recall, F1-score, and the confusion matrix.

Precision measures the proportion of true positive predictions among all positive predictions. This indicates the model's accuracy in identifying the correct land cover classes (Navnath et al., 2022).

Precision is calculated as follows:

$$\mathbf{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

On the other hand, recall evaluates the ratio of true positive predictions to the total actual positives. It measures the model's ability to capture all relevant instances of each land cover class (Navnath et al., 2022). The formula for calculating recall is as follows:

$$\mathbf{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

F1-score represents the harmonic mean of precision and recall. This provides a balanced assessment by considering both false positives and false negatives (Maung et al., 2023). The F1-score is calculated as follows:

$$\mathbf{F1 - score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Additionally, the confusion matrix offers an inclusive view of the model's performance by displaying the counts of true positive, true negative, false positive, and false negative predictions. It enables the calculation of various metrics and helps to identify specific areas where the model may be misclassifying (Cheng et al., 2021).

2.6. McNemar's Test for Comparing Classification Models

To compare the performance of the two classification models (**Sentinel-2 and PlanetScope**) in correctly distinguishing between maize fields and mixed fields, the McNemar's test was used. McNemar's test is useful for evaluating differences in performance between paired data (Leon, 1998). This test is conducted based on Chi-square (χ^2) test for goodness of fit (Kavzoglu, 2017). The aim is to compare an observed count distribution to the expected distribution under a null hypothesis. The test creates a 2x2 contingency table. This table includes the counts of samples correctly and incorrectly classified by two different models. The McNemar test statistic is calculated using the following formula:

$$\chi^2 = \frac{(|n_{ij} - n_{ji}| - 1)^2}{n_{ij} + n_{ji}}$$

Where:

n_{ij} represents the count of samples that were misclassified in model i but correctly classified in model j .

n_{ji} represents the count of samples that were misclassified in model j but not in model i .

This formula is used to determine whether there is a statistically significant difference between the two classification models. If the calculated Chi-square statistic exceeds the critical value of 3.84 at a 95% confidence level, it indicates that there is a significant difference in the performance of the two models. For this study, the 2x2 contingency table was created based on the confusion matrices derived from both models (Sentinel-2 and PlanetScope) focusing only on the land cover class of interest thus (maize fields and mixed fields).

3. RESULTS

This chapter presents the findings of the study, organized according to the research objectives. The outcomes of various experiments and analyses are reported to address each objective systematically.

3.1. Correlation results for Sentinel-2 input data.

To assess the usefulness and relative importance of spectral indices, textural, and topographic features from optical satellite data for the discrimination of different maize cropping systems using the Random Forest (RF) algorithm, for this study, a series of experiments with different feature combinations was setup. The correlation analysis revealed key findings regarding the relationships between various input features.

For Sentinel-2 bands, high correlations were observed among adjacent spectral bands such as B12, B11, B8, B8A, B7, B6, B5, and B4 (**Figure 10**) with correlation coefficients often exceeding 0.90. This high correlation indicates similar information capture due to their spectral proximity. Similarly, texture features derived from these bands also exhibited high correlations. For instance, features such as B4_contrast and B4_diss with a correlation of 0.87, and B8_contrast and B8_diss with a correlation of 0.90 suggest redundancy. Spectral indices like EVI, NDVI, and SAVI were also highly correlated with coefficients often above 0.90 (**Figure 11**), reflecting their similar derivation and purpose.

In the PlanetScope dataset, high correlations were noted among visible bands (B, R, G), with coefficients ranging from 0.93 to 0.98 (**Figure 12**). Texture features such as R_contrast and R_diss had a correlation of 0.89 indicating some redundancy. Indices such as EVI, NDVI, and NDWI for PlanetScope also showed high correlations, often exceeding 0.90 (**Figure 13**). In contrast, topographic features (slope, elevation, aspect) displayed low correlations with each other and with the spectral bands, texture features, and vegetation indices. This suggests their unique contribution to the model.

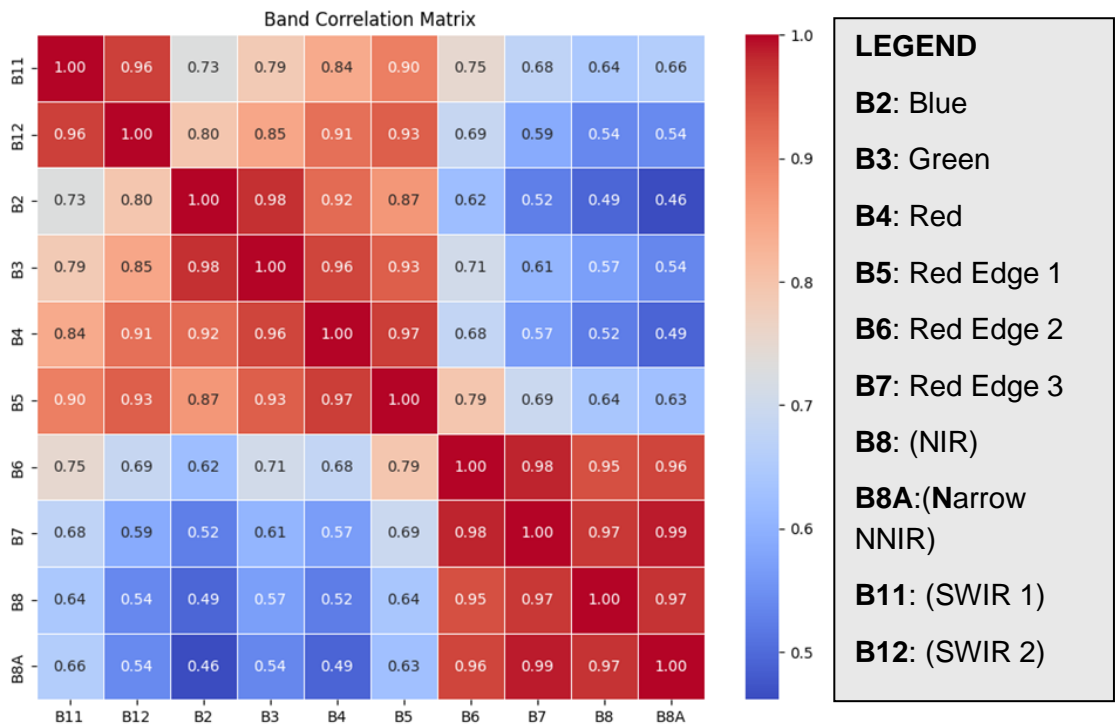


Figure 10: Pearson correlation matrix for Sentinel-2 spectral bands.

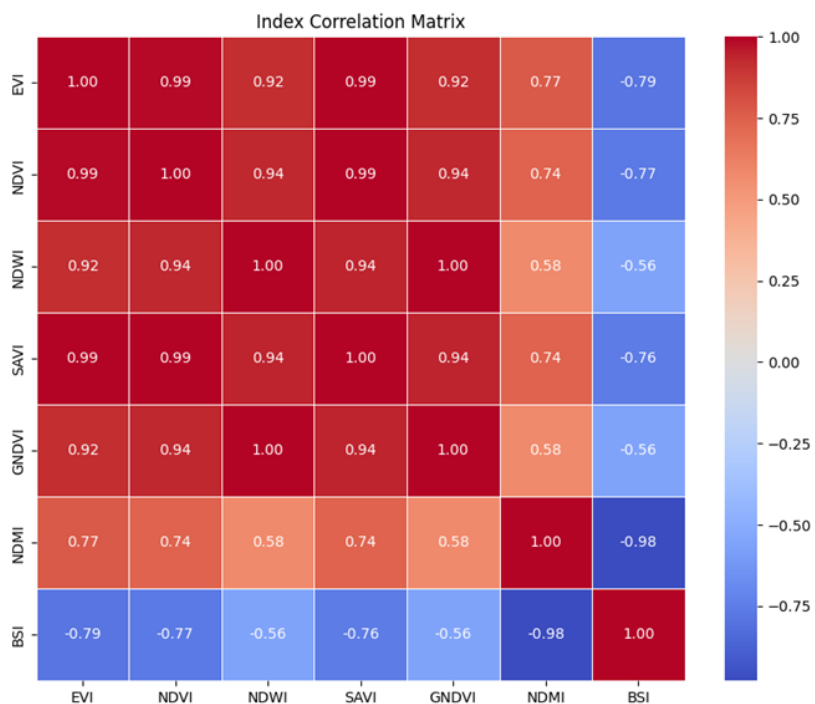


Figure 11: Pearson correlation matrix of vegetation indices computed from Sentinel-2 image.

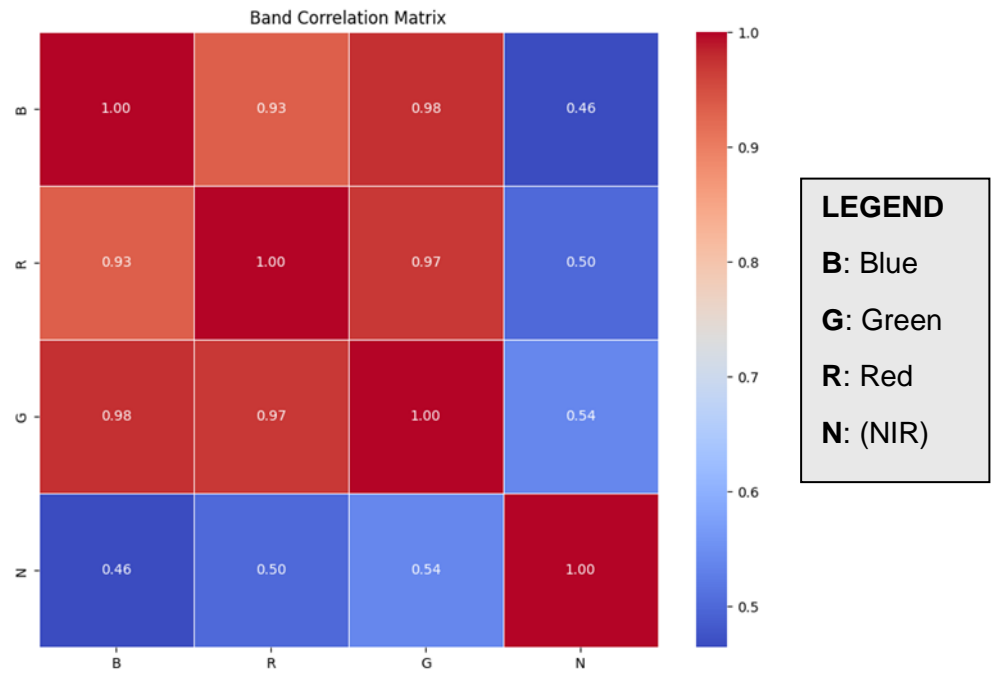


Figure 12: Pearson correlation matrix for PlanetScope spectral bands

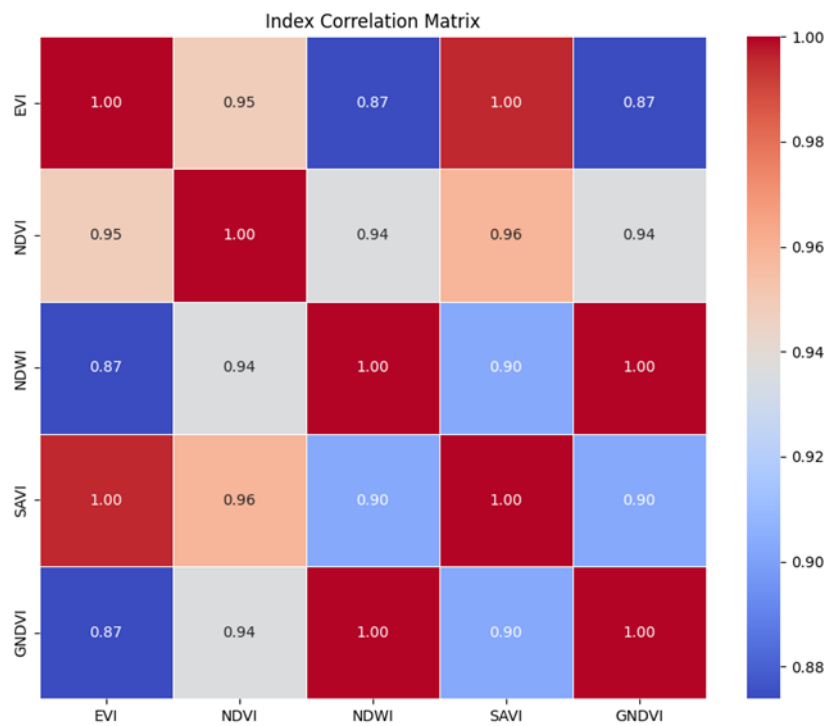


Figure 13: Pearson correlation matrix of vegetation indices computed from PlanetScope image.

3.2. Results of Input Features Experiments

In this study, four experiments were setup on the various image composites in line with the correlation coefficients of the input features. The aim was to identify the best performing experiments for both datasets. In line with this, the tables displayed below show a comparison of the cross-validation accuracy, test accuracy, and OOB Score for the best combination of input features across the different experiments for Sentinel-2 and PlanetScope composite data.

The performance metrics for the input features experiments using both Sentinel-2 and PlanetScope datasets revealed key results. For Sentinel-2 data, across the rain season, dry season, and annual composites, the cross-validation accuracy ranged from 0.82 to 0.84, the test accuracy ranged from 0.83 to 0.85, and the OOB scores were consistent between 0.83 to 0.84 (**Table 9**). Similarly, for PlanetScope data, the cross-validation accuracy varied from 0.80 to 0.82. The test accuracy ranged from 0.81 to 0.83 and the OOB scores were around 0.82 (**Table 10**). These consistent results demonstrate the robust performance of the RF model across the different feature sets.

Table 9: Summary of performance metrics for the input features experiments using Sentinel-2 data across different composite types (rain season, dry season, and annual composite)

Composite Type	Experiment	Cross-validation Accuracy	Test Accuracy	OOB Score
Rain Season	1	0.84	0.85	0.83
	2	0.82	0.83	0.83
	3	0.83	0.83	0.83
	4	0.82	0.83	0.83
Dry season	1	0.84	0.84	0.84
	2	0.83	0.84	0.84
	3	0.83	0.83	0.84
	4	0.83	0.84	0.84
Annual Composite	1	0.84	0.83	0.84
	2	0.83	0.83	0.84
	3	0.83	0.84	0.84
	4	0.83	0.84	0.84

Table 10: Summary of performance metrics for the input features experiments using PlanetScope data across different composite types (rain season, dry season, and annual composite)

Composite Type	Experiment	Cross-validation Accuracy	Test Accuracy	OOB Score
Rain Season	1	0.82	0.81	0.83
	2	0.82	0.81	0.83
	3	0.82	0.82	0.83
	4	0.83	0.81	0.83
Dry season	1	0.81	0.80	0.82
	2	0.80	0.81	0.82
	3	0.81	0.81	0.82
	4	0.81	0.81	0.82
Annual Composite	1	0.82	0.82	0.82
	2	0.81	0.82	0.82
	3	0.82	0.82	0.82
	4	<u>0.82</u>	<u>0.83</u>	<u>0.82</u>

3.2.1 Comparative Performance of Experiments Across Datasets (Sentinel-2 vs PlanetScope)

Comparing the performance of experiments across both datasets, it is evident that the Sentinel-2 data consistently outperformed PlanetScope in terms of accuracy and robustness. For Sentinel-2, **Experiment 1** which used the rain season composite displayed the highest test accuracy (0.85) and cross-validation accuracy (0.84) indicating that this combination of features and image composite was particularly effective (**Table 9**). The dry season and annual composites for Sentinel-2 also showed high and consistent performance across experiments with only minor variations in accuracy metrics.

In contrast, for PlanetScope dataset, the annual composite in **Experiment 4** exhibited the highest test accuracy (0.83) and consistent cross-validation accuracy (0.82) suggesting this combination was the most effective for this dataset (**Table 10**). The dry season and rain season composites for PlanetScope demonstrated slightly lower but uniform performance with Experiment 4 achieving marginally better test accuracy.

Overall, the Sentinel-2 rain season composite (**Experiment 1**) and the PlanetScope annual composite (**Experiment 4**) emerged as the top performers.

3.3. Variable Importance Results

To assess the usefulness and relative importance of spectral indices, textural, and topographic features from (Sentinel-2 and PlanetScope) data for discriminating different maize cropping systems, this study computed feature importance using three methods: Mean Decrease Gini (MDG), SHapley Additive exPlanations (SHAP), and Permutation Importance.

3.3.1. Variable importance for Sentinel-2 (Rain season composite-best experiment)

The feature importance analysis using MDG, SHAP, and Permutation Importance methods highlights key features which enabled the predictive capabilities of the RF model is shown in **Figures 14, 15, and 16**. **Figure 14** reveals that Band 5, Elevation, Band 2, and Band 11 are among the most significant features according to the MDG method. **Figure 15** shows Permutation Importance scores and identifies Elevation, Band 7 Variance, and Band 3 Dissimilarity as the most critical features. **Figure 16** with SHAP values shows Band 8, Band 3, and Band 7 as having the highest impact on the model output across various land cover classes.

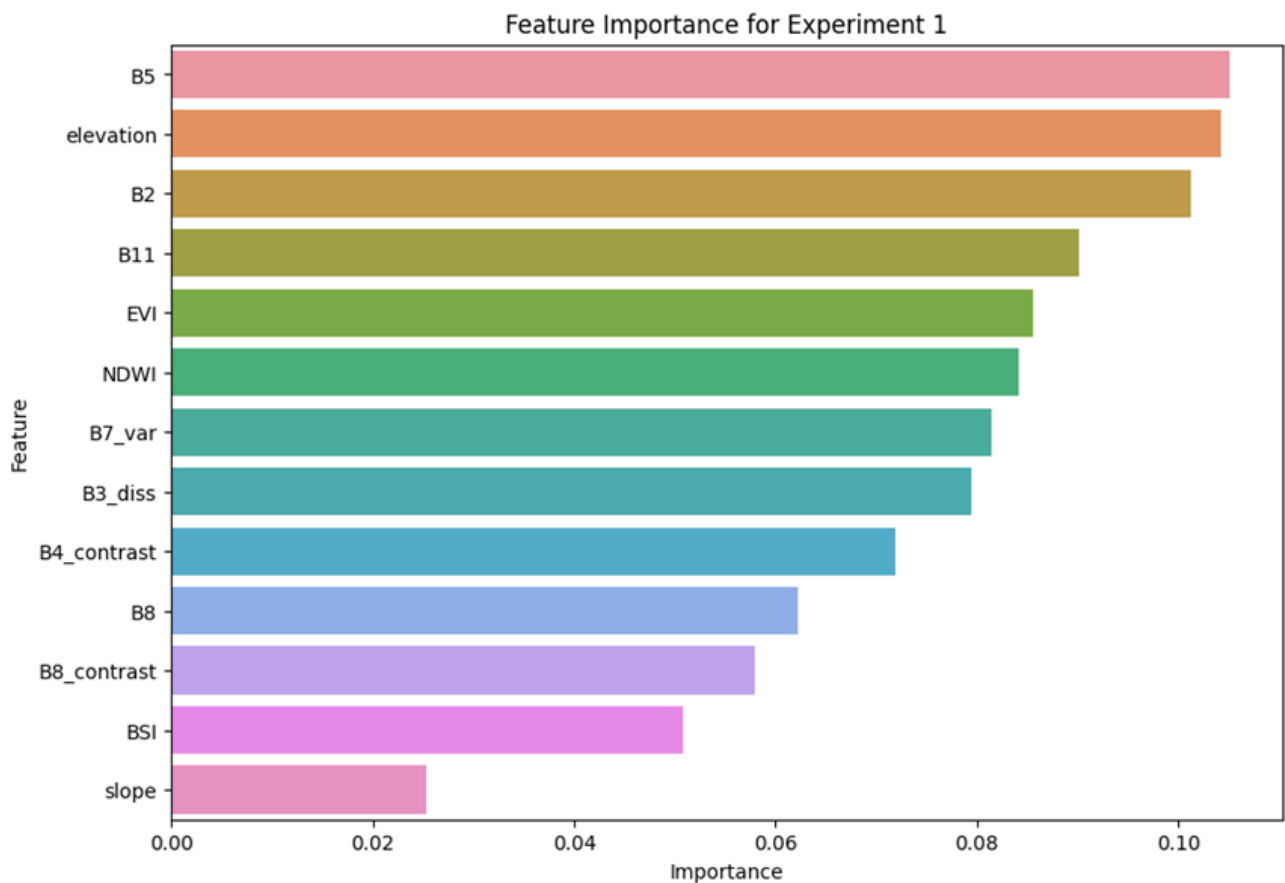


Figure 14: Bar chart showing feature importance using Mean Decrease Gini (MDG) Method on Sentinel-2 rain season composite data.

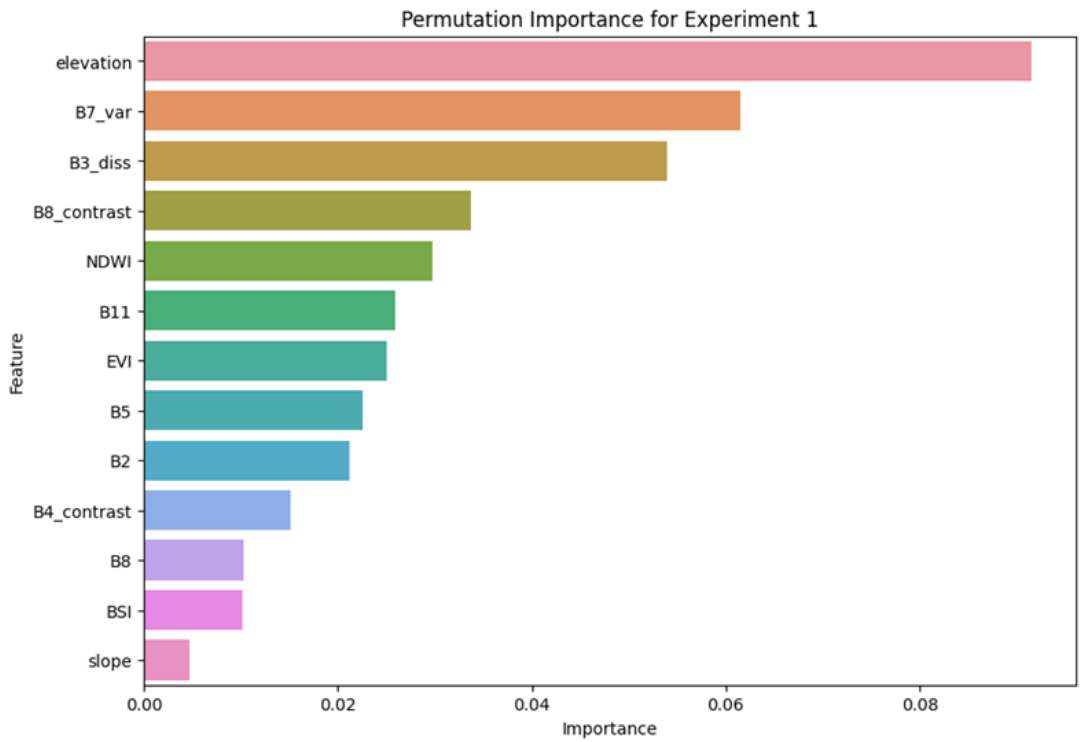


Figure 15: Bar chart showing feature importance using permutation importance Method on Sentinel-2 rain season composite data.

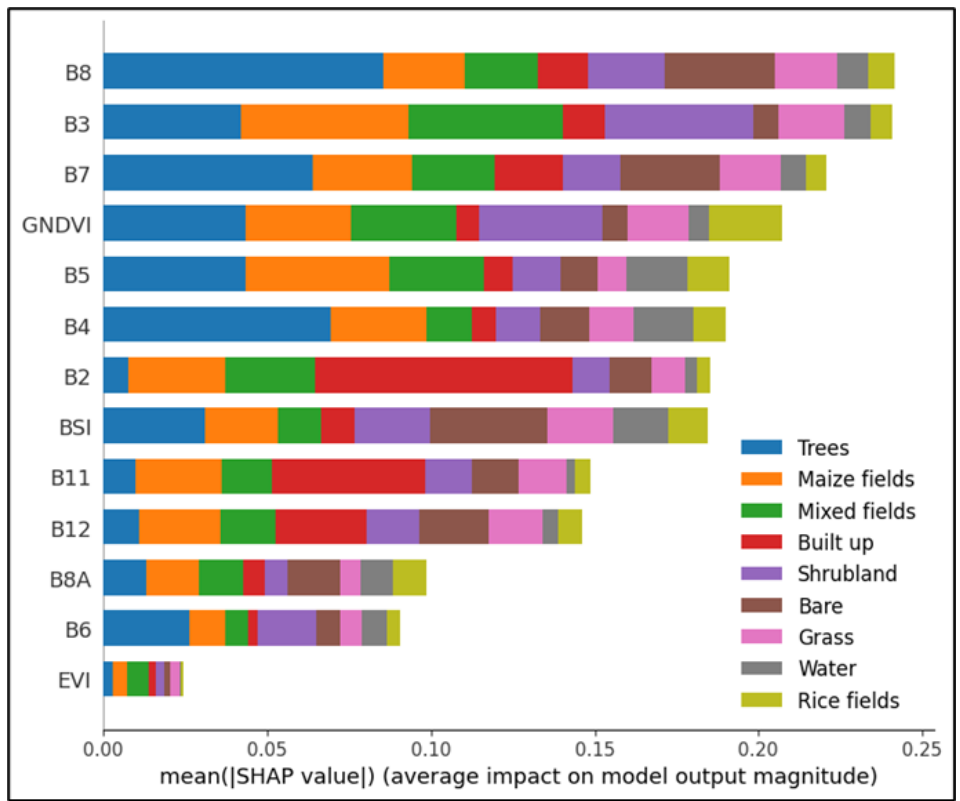


Figure 16: Bar chart showing feature importance using SHAP Method on Sentinel-2 rain season composite data.

3.3.2. Variable importance for PlanetScope (Annual composite-best experiment)

Similarly, feature importance analysis using MDG, SHAP, and Permutation Importance methods for PlanetScope annual composite data is shown in **Figures 17, 18, and 19**. **Figure 17** shows that the Green Band, Elevation, Red Band, and Blue Band are among the most important features based on the MDG method. **Figure 18** shows that Elevation, Green Dissimilarity, and the NIR Band are the most crucial features according to Permutation Importance scores. **Figure 19** using SHAP values, reveals that Red Dissimilarity, Red ASM, and SAVI have the greatest impact on the model output across various land cover classes. Similar analysis has been conducted on the multi-stack composite data for both Sentinel-2 and PlanetScope datasets and the results can be found in the appendix section of this study.

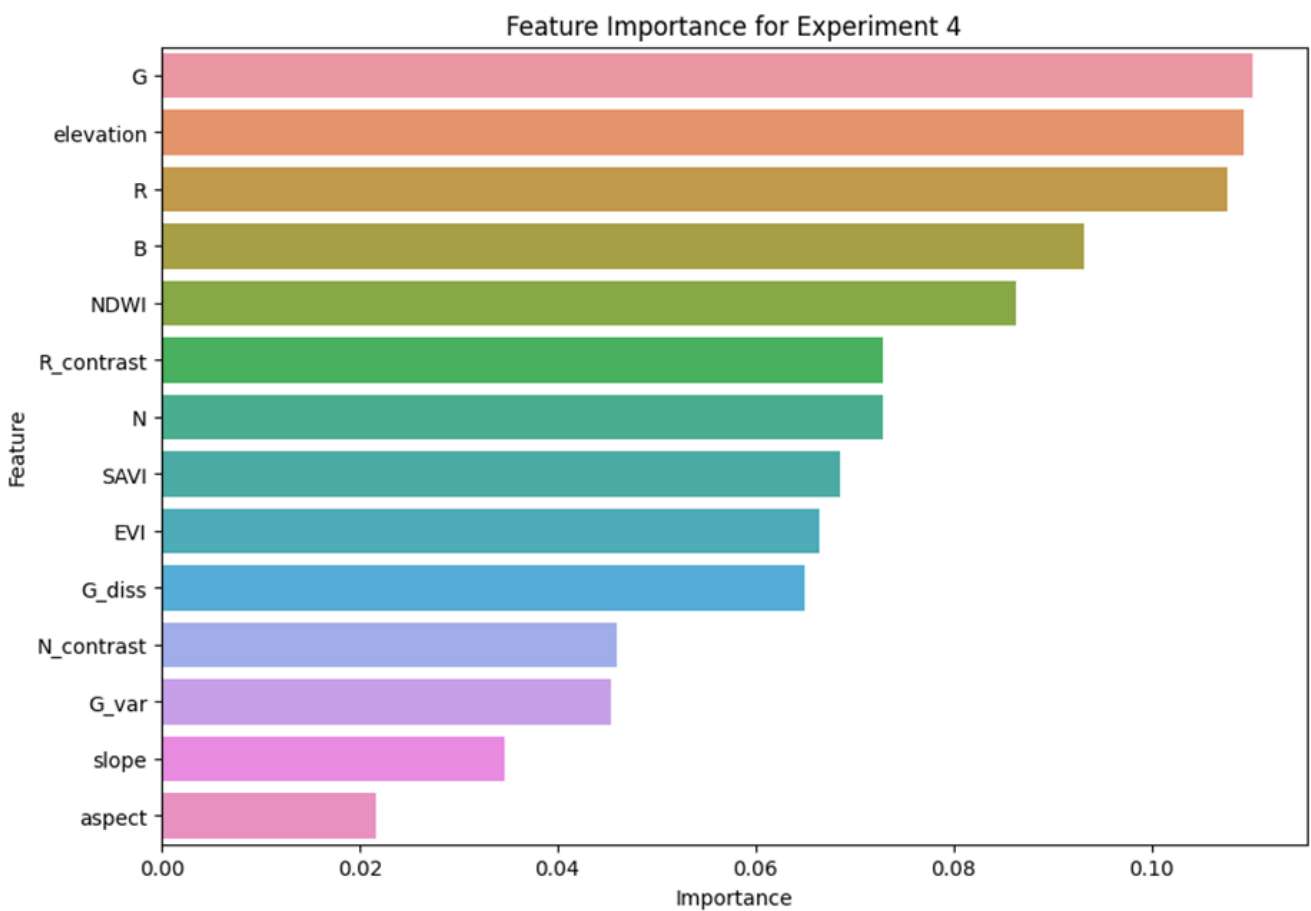


Figure 17: Bar chart showing feature importance using Mean Decrease Gini (MDG) Method on PlanetScope annual composite data.

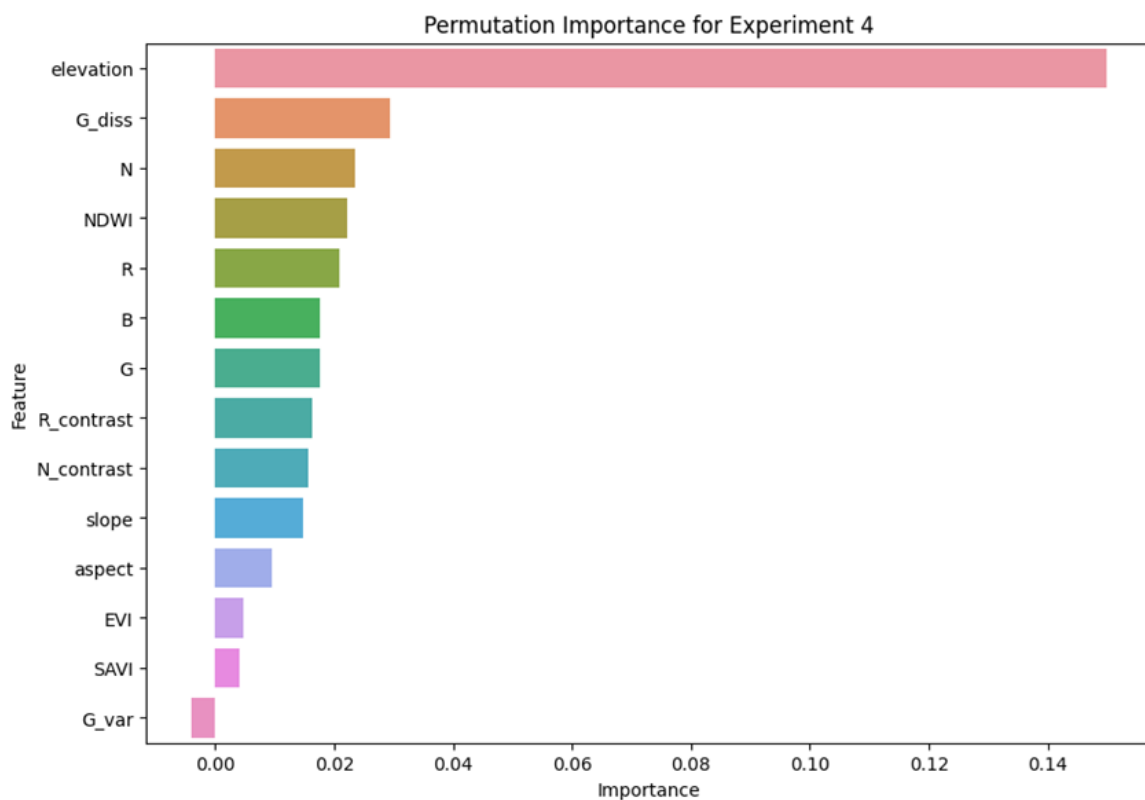


Figure 18: Bar chart showing feature importance using permutation importance method on PlanetScope annual composite data.

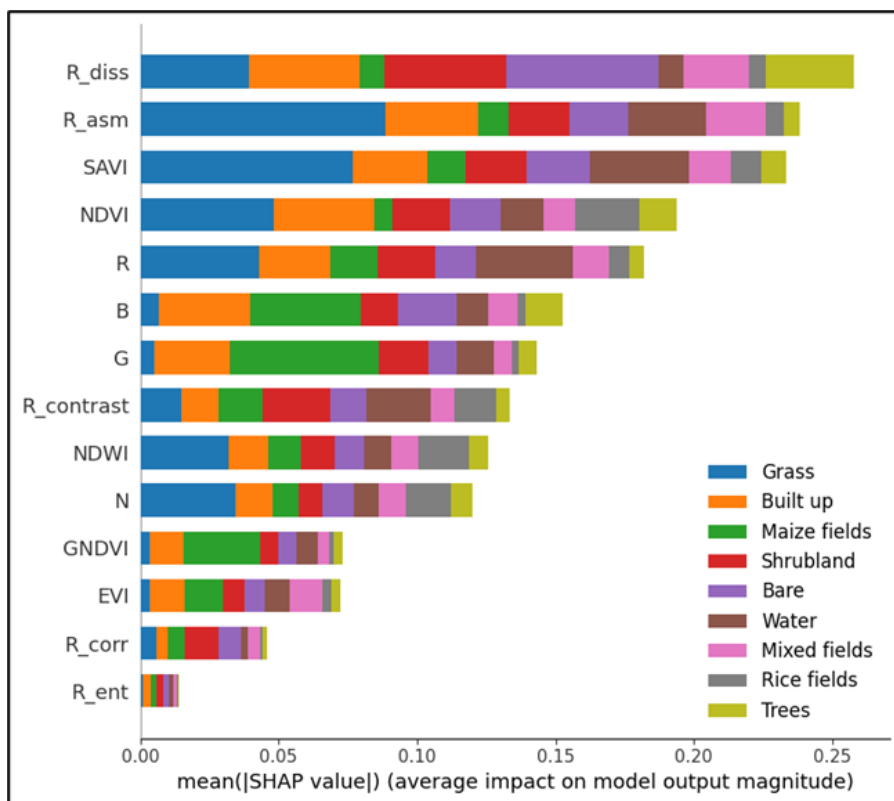


Figure 19: Bar chart showing feature importance SHAP Method on PlanetScope annual composite data.

The comparison between the feature importance of Sentinel-2 and PlanetScope data shows both shared and unique features across the datasets. Common features identified as important in both datasets include Elevation, NDWI, EVI, NIR, and Blue highlighting their universal relevance in identifying different maize cropping systems. Sentinel-2 emphasized specific spectral bands such as B5, B2, B11, and B8, as well as textural features like B3_diss (Dissimilarity) and B4_contrast (Contrast). In contrast, PlanetScope highlights spectral bands like G (Green), R (Red), and B (Blue) and focuses on vegetation indices like SAVI, NDWI, EVI, and textural measures such as G_diss (Green Dissimilarity) and R_asm (Red Angular Second Moment).

3.4. Classification Results for Sentinel 2

In line with the second objective (section 1.3) of this study, the figures below illustrate the land cover classification outputs for the same region generated using Sentinel-2 data. The land cover maps were produced using RF classifier. In **Figure 20**, the classification uses the best features achieving an overall accuracy of 85%. This map distinguishes various land cover classes such as Maize Fields, Mixed Crop fields, Built-up areas, Shrubs, Trees, Bareland, Grass, and Rice fields. In **Figure 21**, the classification uses a multi-stack composite resulting in a slightly lower OA of 83%. This map similarly differentiates land cover classes as demonstrated in **Figure 30**.

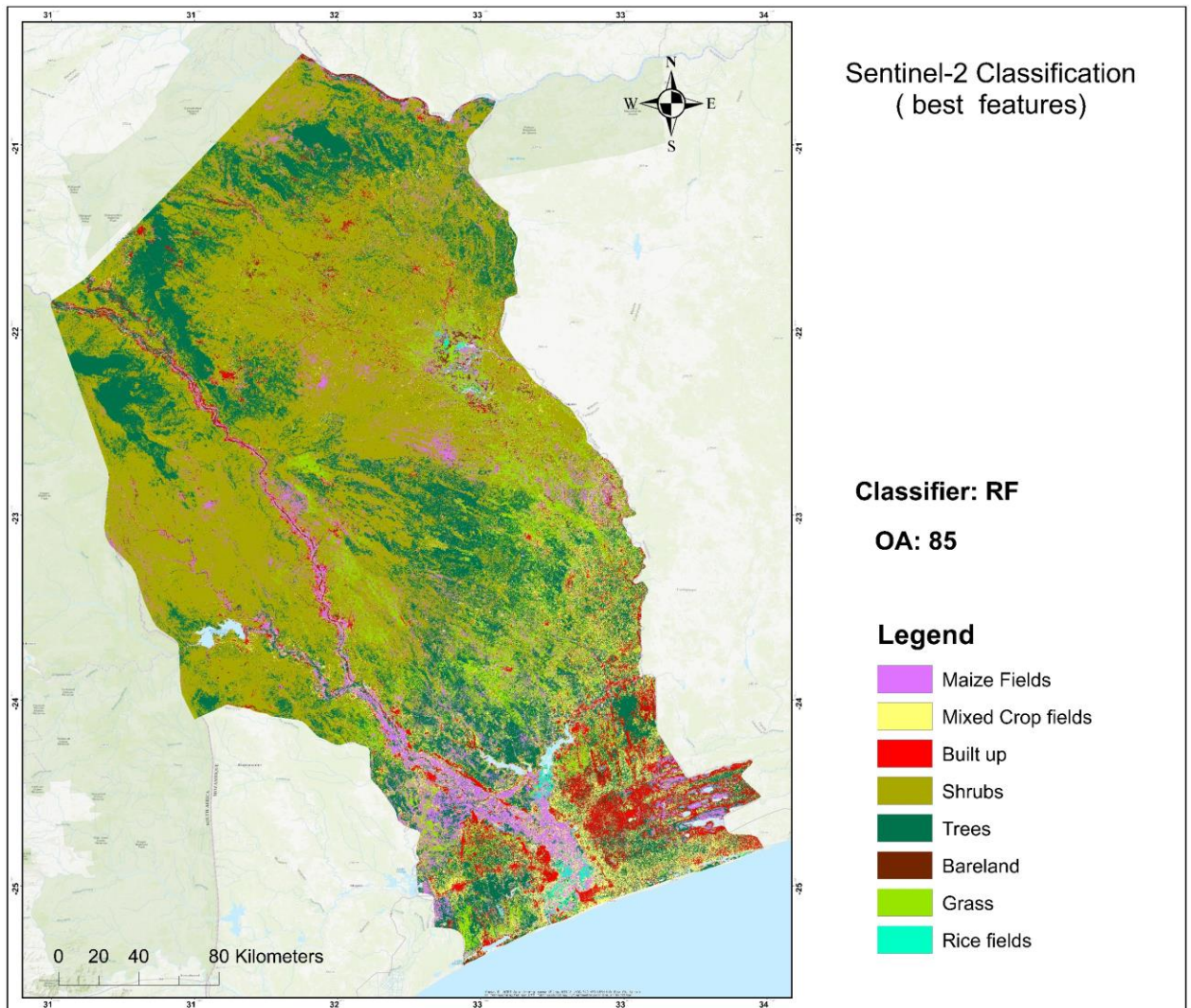


Figure 20: Land Cover Map Produced Using Sentinel-2 Data with the Best Identified Features.

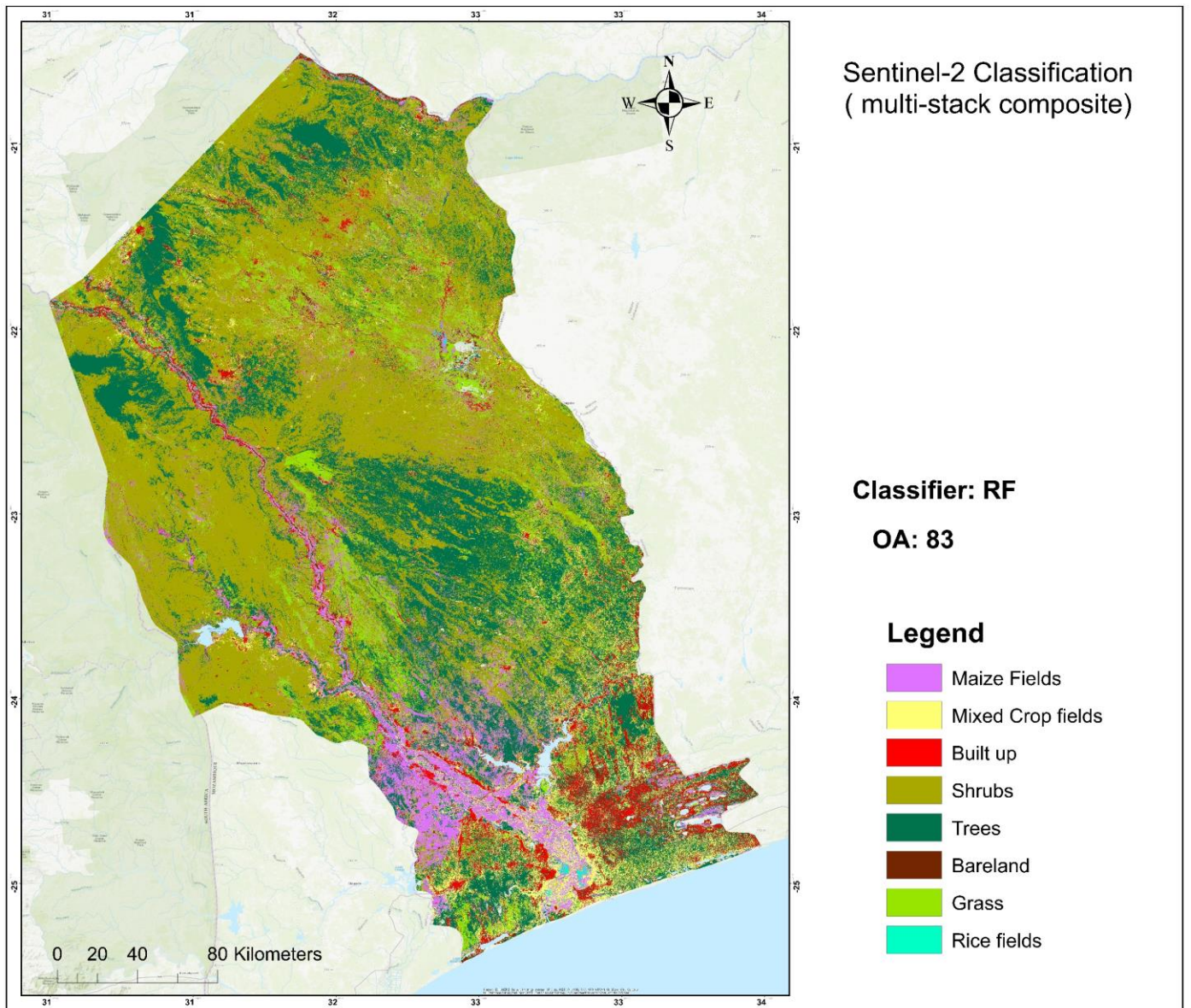


Figure 21: Land Cover Map Produced Using Sentinel-2 Multi-stack Composite Data with the Best Identified Features

3.4.1. Classification Results for PlanetScope data

Similarly, figures below show the classification results for the study area using PlanetScope data produced using RF classifier. In **Figure 22**, the classification used the best features resulting in an overall accuracy of 83%. The map identifies various land cover classes such as Maize Fields, Mixed Crop fields, Built-up areas, Shrubs, Trees, Bareland, Grass, and Rice fields. In **Figure 23**, the classification used a multi-stack composite achieving a higher overall accuracy of 86%. This map similarly distinguishes the land cover classes with notable differences in classification accuracy and distribution compared to **Figure 22**.

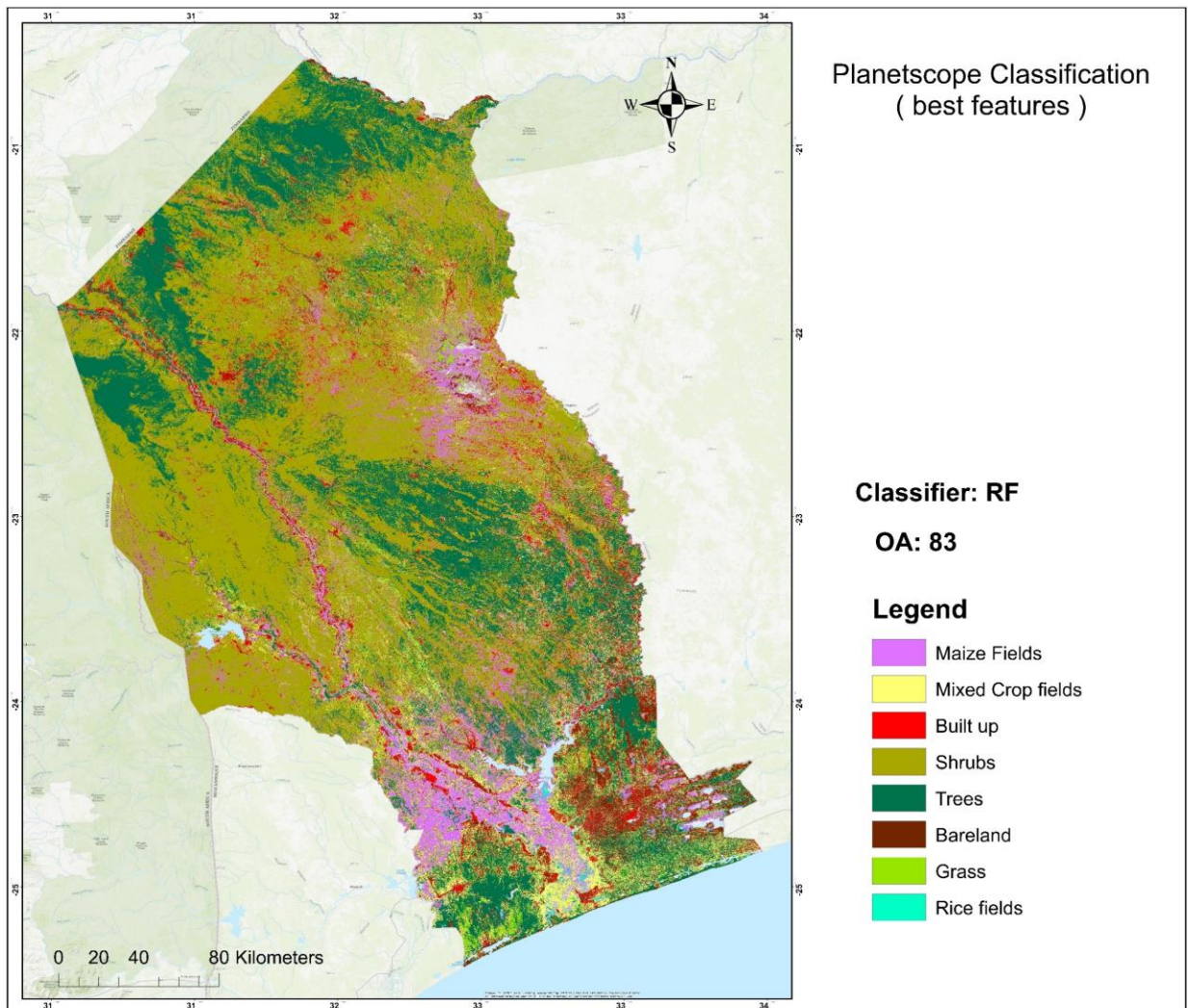


Figure 22: Land Cover Map Produced Using PlanetScope Data with the Best Identified Features.

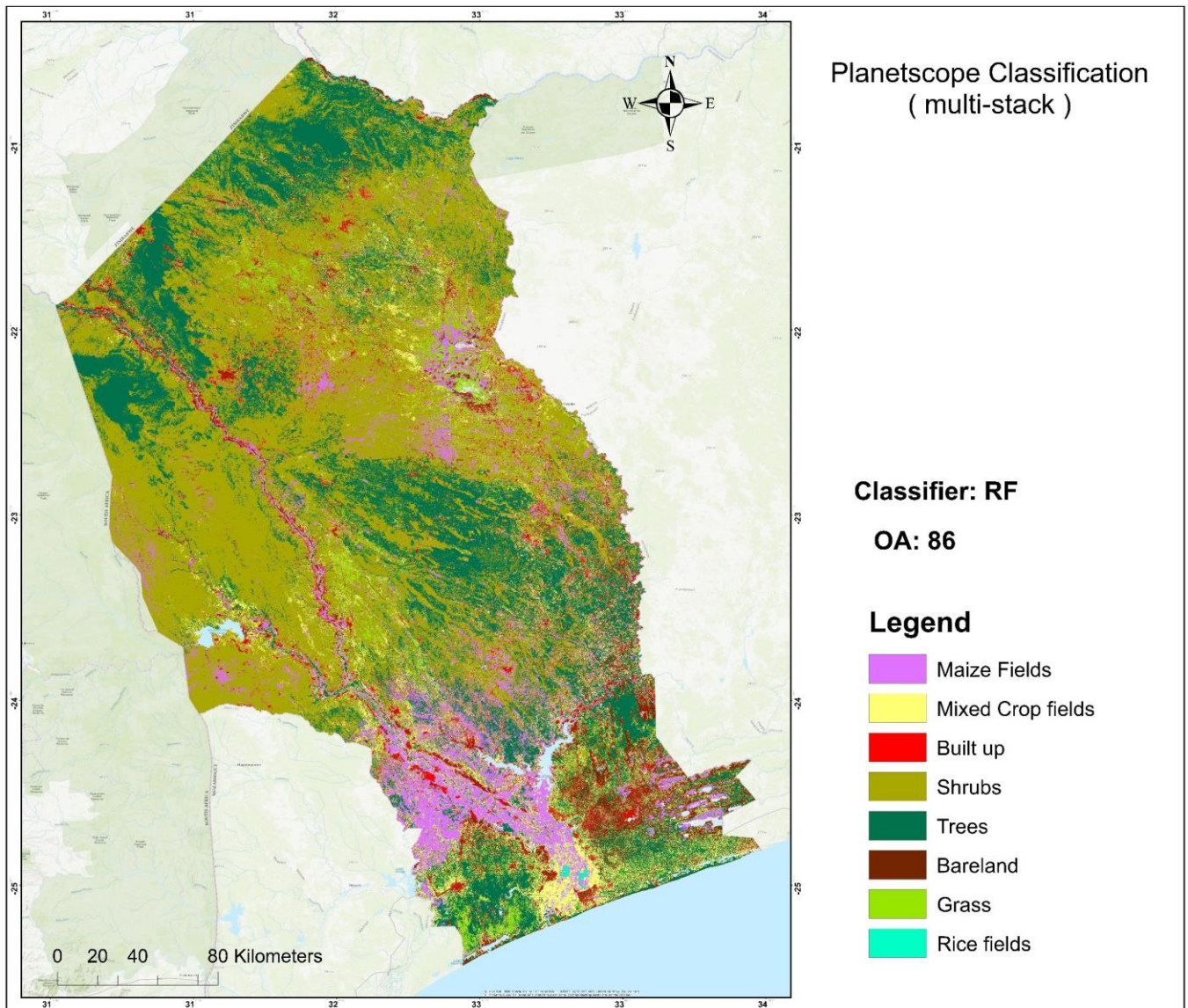


Figure 23: Land cover map produced using PlanetScope data with the best identified features.

3.4.2. Consistency in Land Cover Classification between Sentinel-2 and PlanetScope data

Figure 24 shows a comparative analysis of land cover classification using PlanetScope data for areas A, B, C, and D. The top row presents satellite imagery snapshots of the study area while the subsequent rows display the classification outputs from the PlanetScope best feature maps and multi-stack maps. In area A, both datasets demonstrate consistent classification for rice fields and maize fields. Area B shows consistent classification for maize fields and built-up areas (red). In area C, both maps consistently classify trees and shrubs. Lastly, area D shows consistent classification for trees and built-up areas.

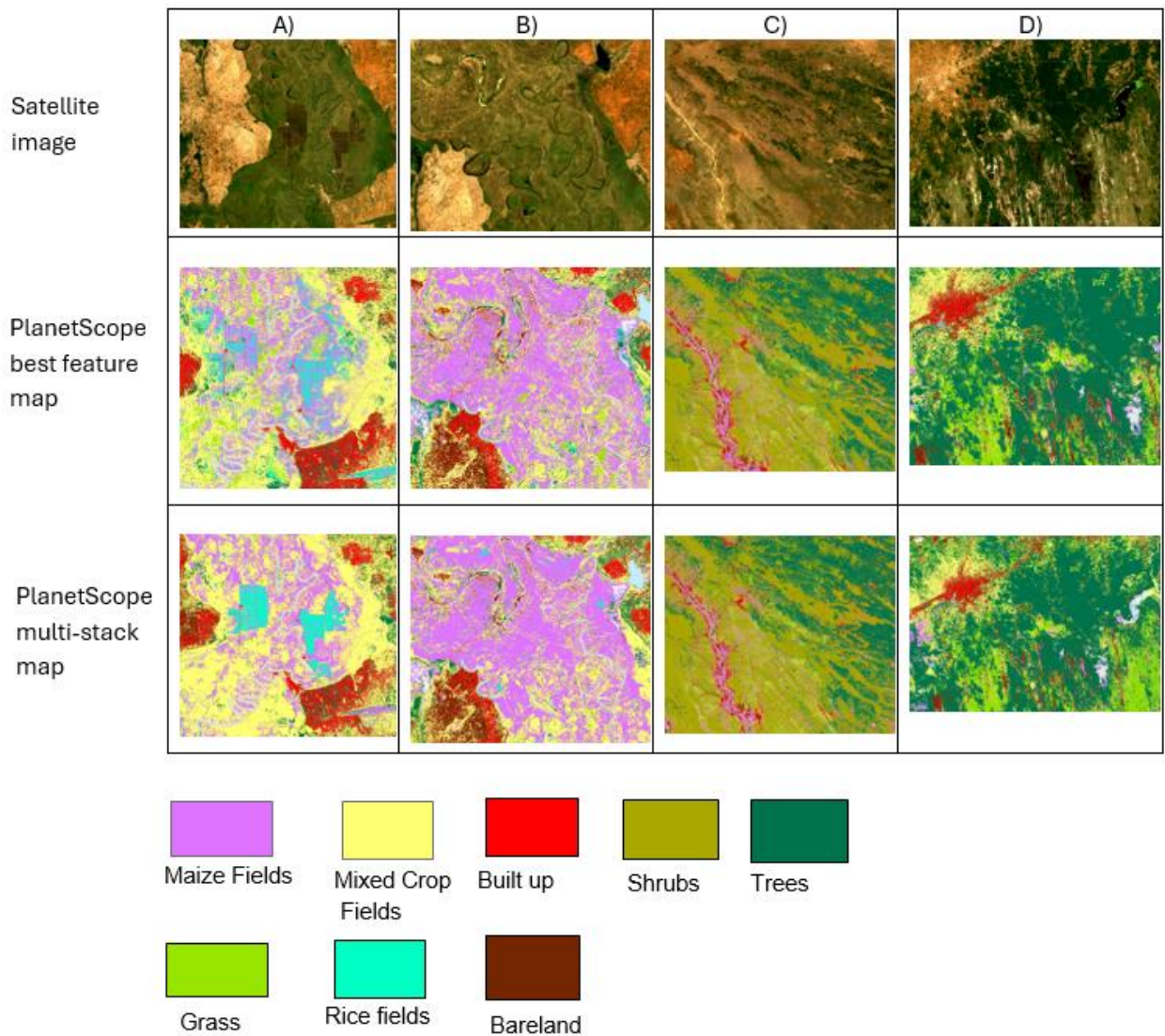


Figure 24: Comparison of land cover classification using PlanetScope data for areas A, B, C, and D. The top row shows satellite imagery snapshots, the middle row displays the best feature maps, and the bottom row presents the multi-stack maps.

Just like the PlanetScope data, a similar trend can be observed in consistency of classifications using Sentinel-2 data as shown in **figure 25** below. In area A, both datasets consistently classify rice fields and maize fields. Area B shows consistent classification for maize fields and built-up areas. In area C, both maps consistently classify trees and shrubs. Lastly, area D shows consistent classification for trees and built-up areas.

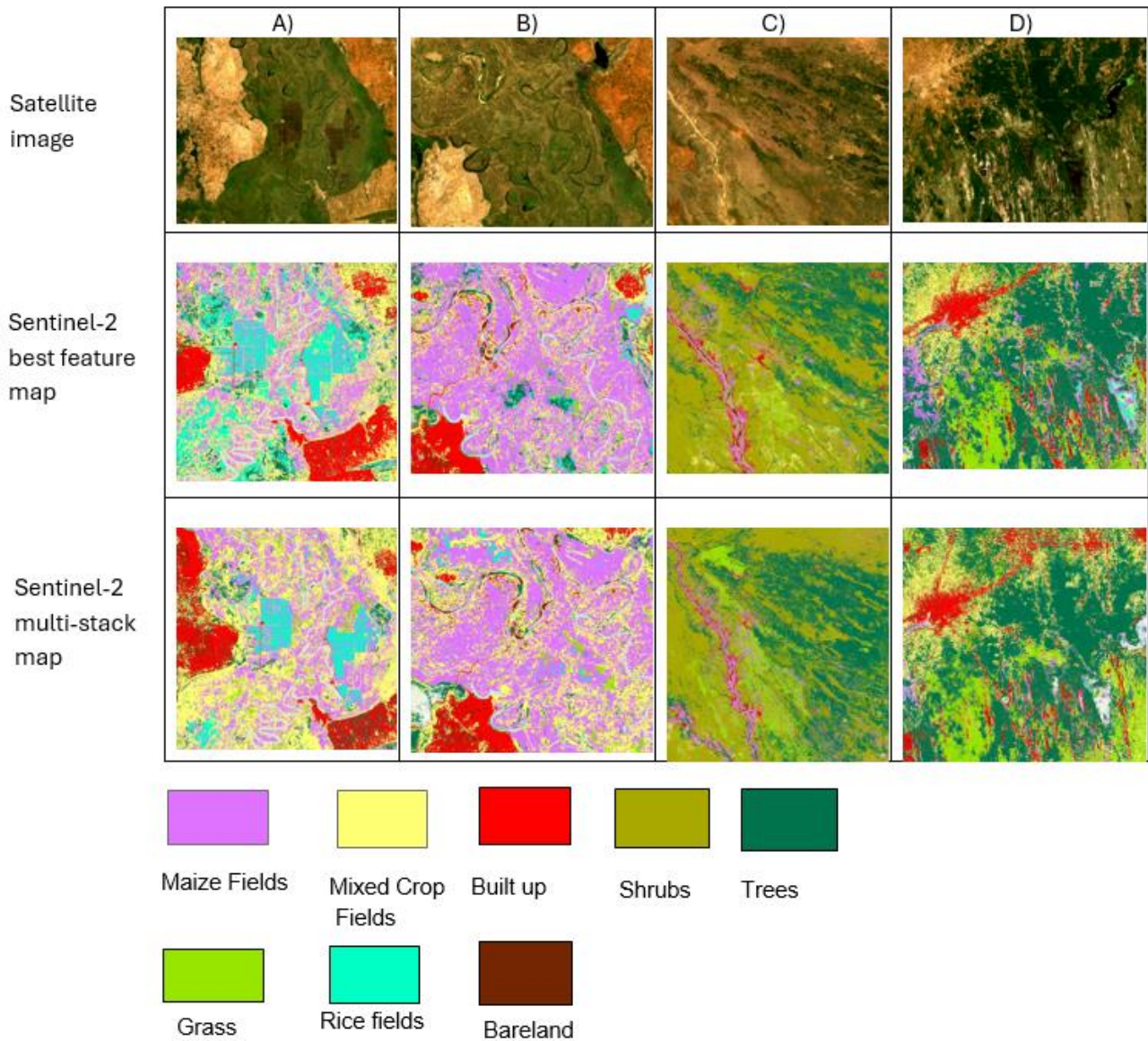


Figure 25: Comparison of land cover classification using Sentinel-2 data for areas A, B, C, and D. The top row shows satellite imagery snapshots, the middle row displays the best feature maps, and the bottom row presents the multi-stack maps.

3.4.3. Inconsistencies in land cover classification between Sentinel-2 and PlanetScope data

Figure 26 shows a comparison of the land cover classification inconsistencies for Sentinel-2 and PlanetScope data for areas A, B, C, and D. In this context, inconsistencies refer to instances where the land cover classes identified by the PlanetScope and Sentinel-2 maps do not match.

Inconsistencies can be observed across the datasets in different areas. In area A, there is a discrepancy in the classification of maize fields and grass. Area B shows inconsistencies in identifying mixed crop fields and rice fields. Area C demonstrates variability in classifying shrubs and grass. Lastly, area D displays differences in the classification of shrubs, maize fields, and grass.

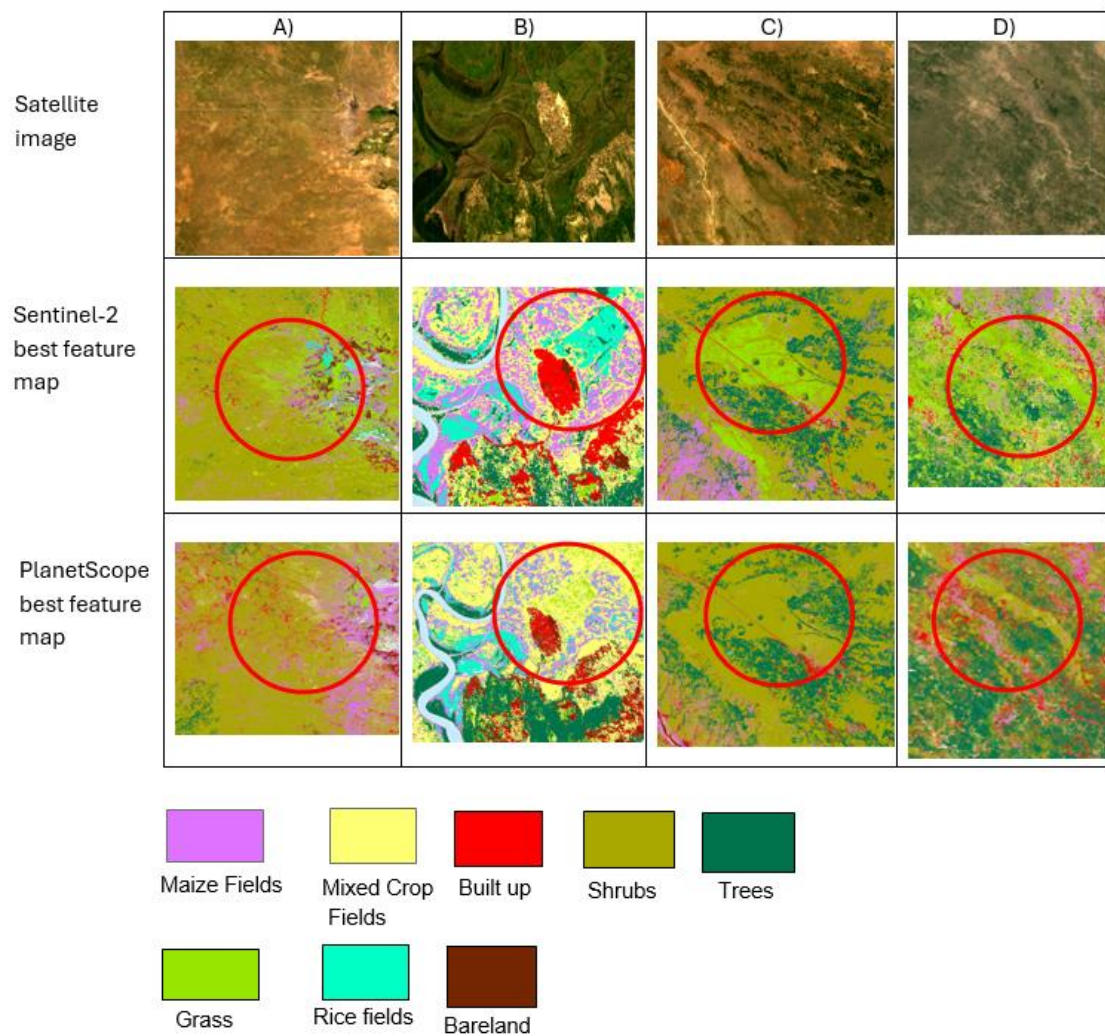


Figure 26: Comparison of land cover classification inconsistencies using Sentinel-2 and PlanetScope data.

3.5. Evaluation of Classification Results for Sentinel-2

In this study, the classification model was evaluated to assess its performance in predicting the landcover classes of interest using numerous metrics. These metrics included precision, recall, and F1-score. The performance of different land cover classes for the Sentinel-2 classification maps including the ‘best features map’ (Figure 27) and the ‘multi-stack composite map’ (Figure 28) reveals both high and low classification results.

High classification performance is observed for water, trees, and shrubs, with precision, recall, and F1-Score values consistently above 0.90 across both maps. This indicates the model’s strong ability to accurately classify these classes. Built-up areas also demonstrate strong performance with precision, recall, and F1-Score values around 0.87, 0.90, and 0.88 respectively in the best features map and slightly lower in the multi-stack composite map (Figure 27 & 28).

On the contrary, grass and mixed crop fields show lower performance metrics with precision, recall, and F1-Score values significantly lower highlighting challenges in accurately detecting these classes. Specifically, mixed crop fields have the lowest scores observed in both composites. Notably, maize fields show a minor improvement in recall in the multi-stack composite map. However, the performance in terms of precision and F1-Score remains close in both composites (Figure 24 & 25). Overall, while the model performs well for water, trees, shrubs, and built-up areas, it performs low with accurately classifying grass and mixed crop fields.

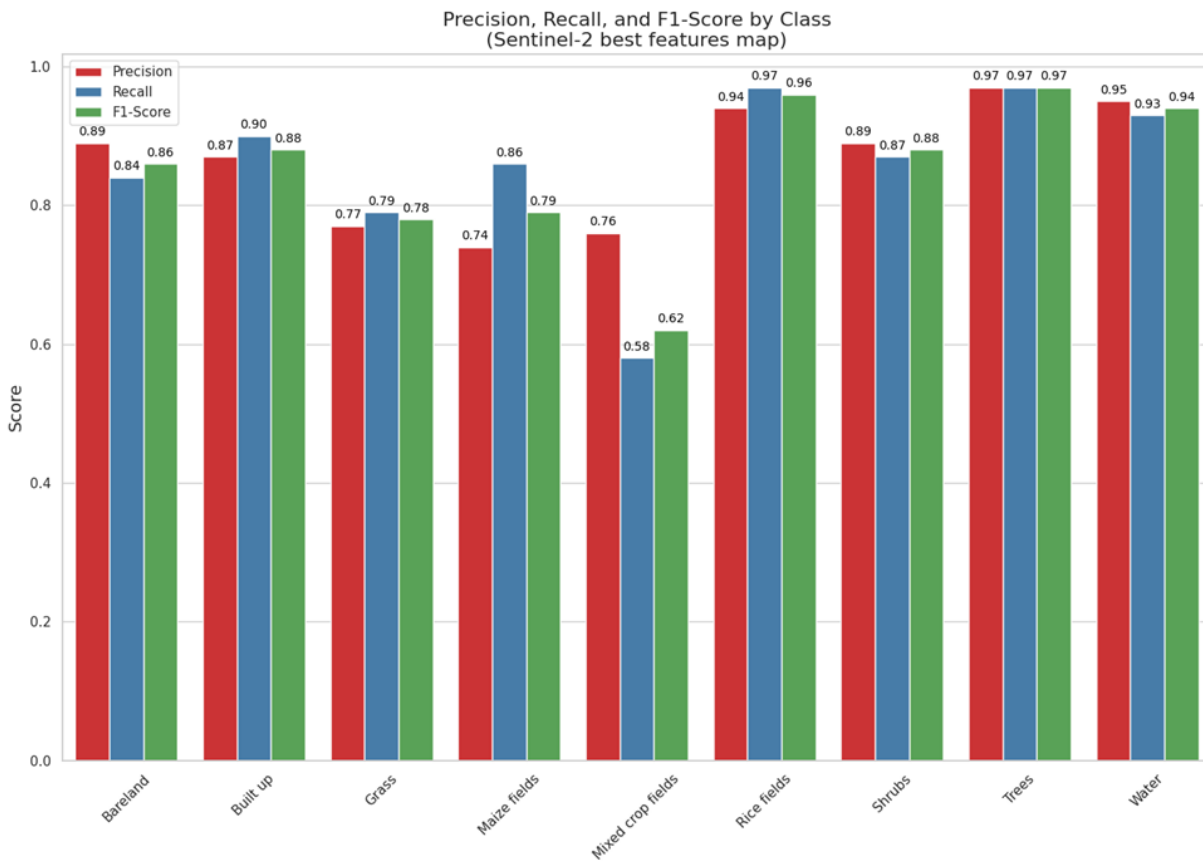


Figure 27: Precision, Recall, and F1-Score by Class for the Sentinel-2 Best Features Map.

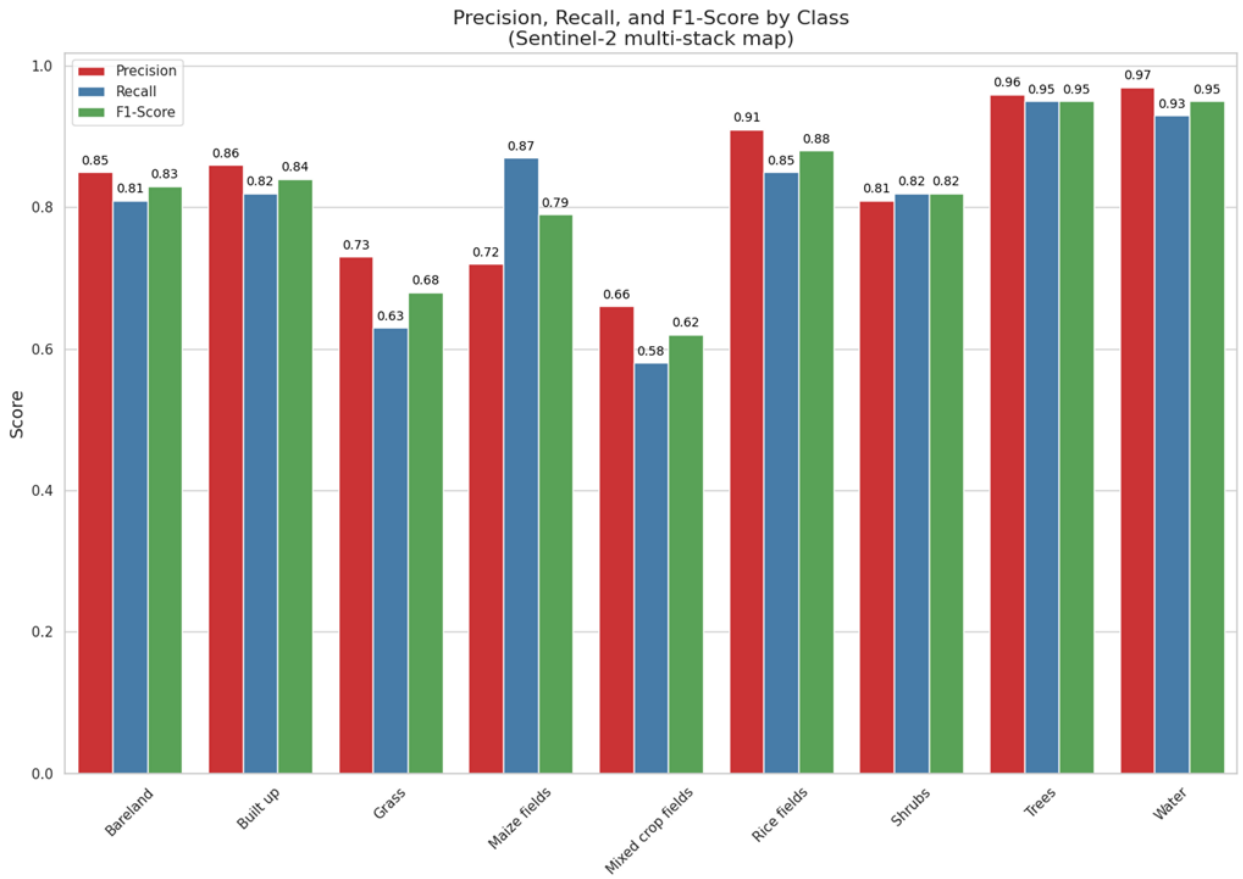


Figure 28: Precision, Recall, and F1-Score by Class for the Sentinel-2 Multi-Stack Composite Map.

3.5.1. Evaluation of Classification Results for PlanetScope

Similarly, **Figure 29 and Figure 30** below shows the performance of different land cover classes for the PlanetScope classification maps thus ‘best features map’ and the ‘multi-stack composite map’ respectively. In Figure 29, high precision, recall, and F1-Score values are observed across most land cover classes indicating effective classification. Specifically, built-up areas, shrubs, trees, and water show notably high scores. On the contrary, classes such as grass and mixed crop fields display lower performance metrics. In Figure 30, a similar trend can be observed with slightly varying scores. Built-up areas and water maintain high classification metrics while trees continue to show high scores across all metrics. The precision for maize fields improved compared to the best features map whereas grass and mixed crop fields still show lower scores.

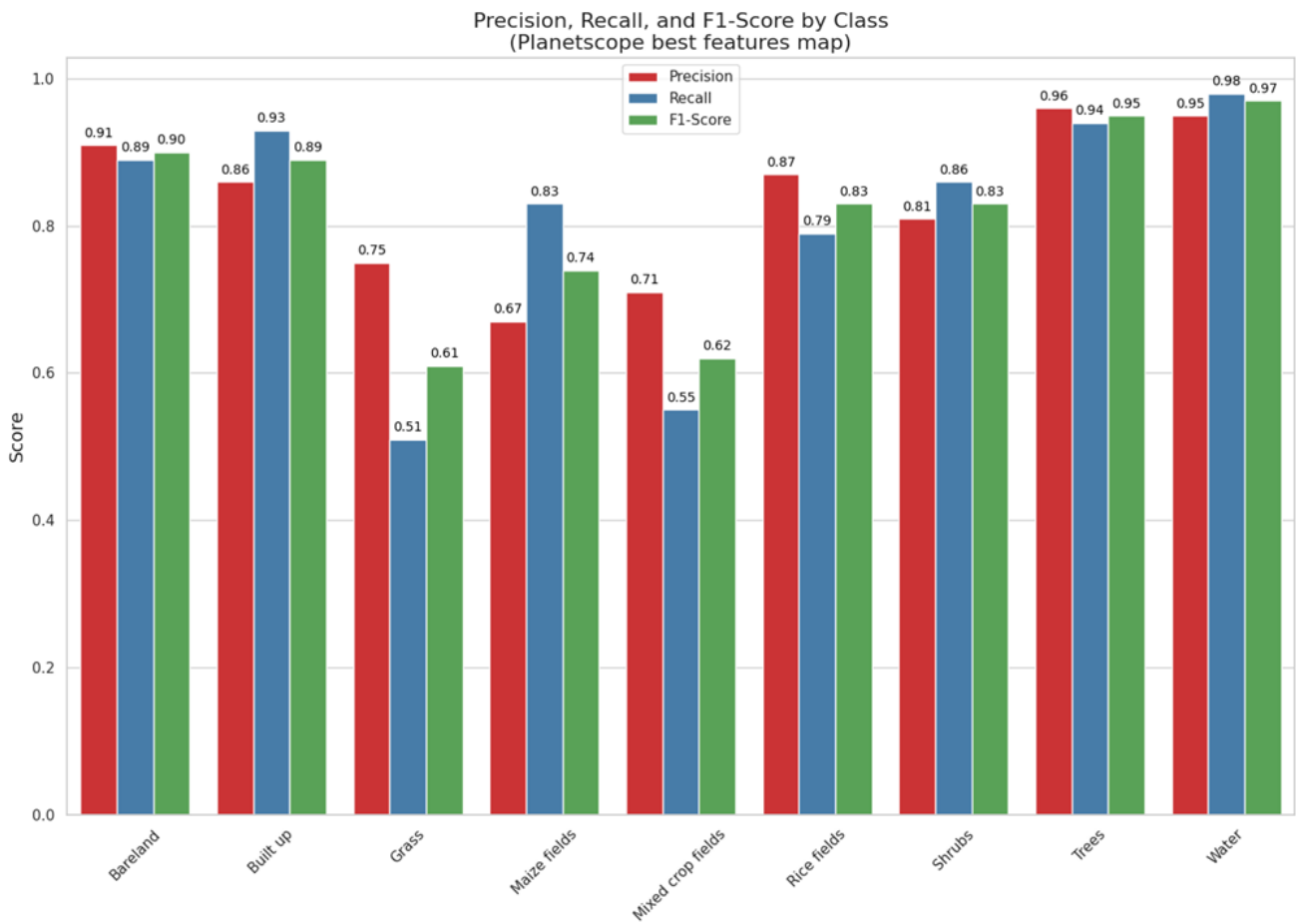


Figure 29: Precision, Recall, and F1-Score by Class for the PlanetScope Best Features Map

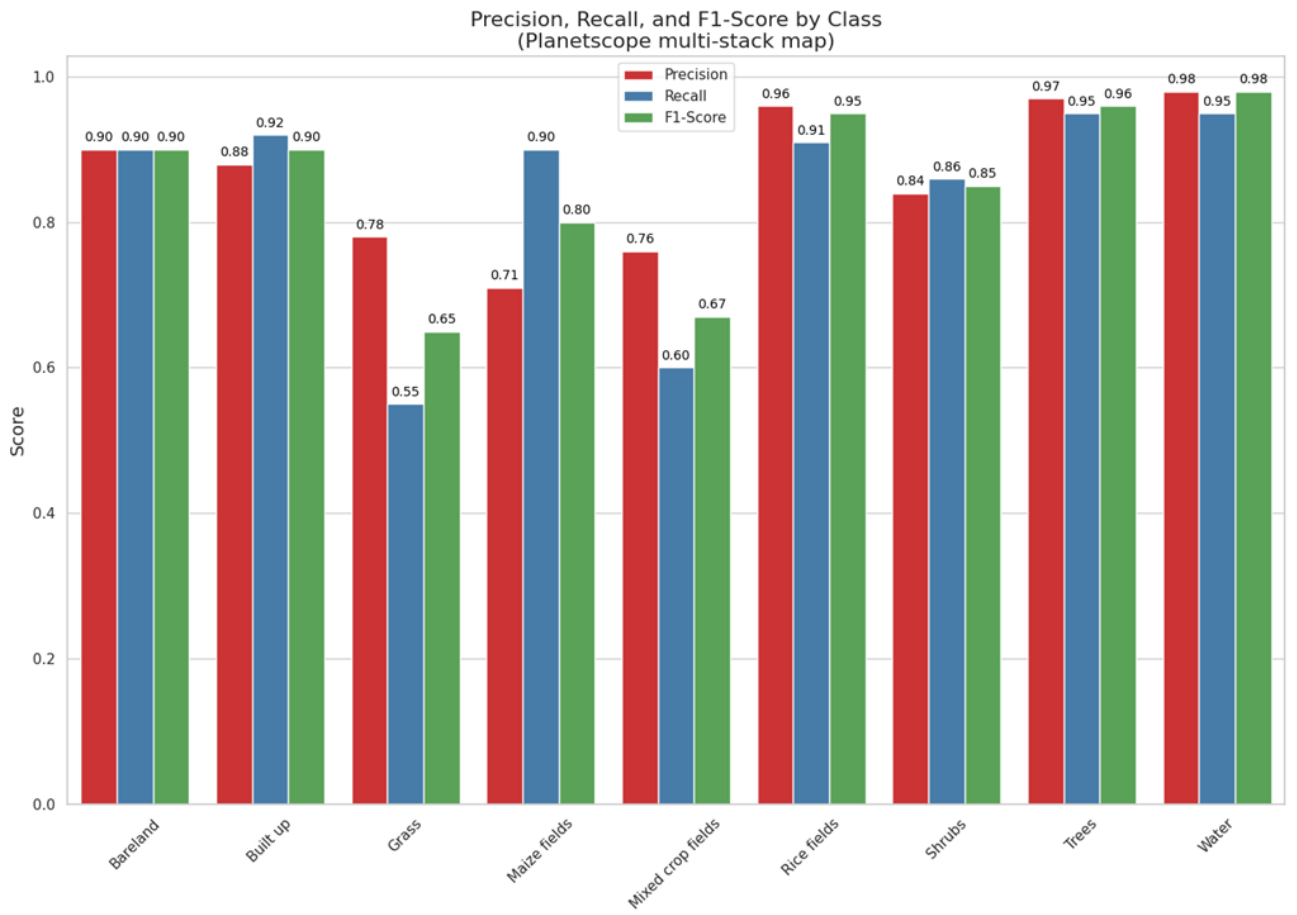


Figure 30: Precision, Recall, and F1-Score by Class for the PlanetScope Multi-Stack Composite Map.

3.5.2. Comparison of the best classification outputs for Sentinel-2 and PlanetScope.

To compare the best classification maps, agreement and disagreement metrics were used to evaluate the consistency between the classification outputs for both PlanetScope and Sentinel-2. This was computed using the confusion matrix of both classification outputs. Agreement indicates the proportion of land cover classifications where both datasets match reflecting consistency and reliability. Disagreement on the other hand, highlights areas where the classifications differ indicating potential differences in how each dataset interprets the land cover. Figure 31 below shows the agreement and disagreement in land cover classification between PlanetScope and Sentinel-2 data for various classes.

For Bare Land, PlanetScope shows 90.48% agreement and 9.52% disagreement while Sentinel-2 shows 86.17% agreement and 13.83% disagreement. Built-up areas have high agreement in both datasets with 93.64% for PlanetScope and 90.09% for Sentinel-2. Grassland shows notable differences with 64.62% agreement for PlanetScope and 52.38% for Sentinel-2 coupled with higher disagreement (35.38% and 47.62% respectively). Maize fields show variation with PlanetScope showing 91.43% agreement compared to 82.29% for Sentinel-2.

Mixed fields show similar agreement levels with PlanetScope at 58.47% and Sentinel-2 at 59.32% though the disagreement percentages are close (41.53% and 40.68% respectively). Rice fields show a significant difference with PlanetScope at 91.18% agreement and 8.82% disagreement, compared to Sentinel-2's 76.47% agreement and 23.53% disagreement reflecting the challenges in accurately classifying this category. Shrubs, Trees, and Water categories generally exhibit high agreement levels in both datasets with PlanetScope consistently showing slightly higher agreement percentages than Sentinel-2.

Agreement and Disagreement by Land Cover Class for PlanetScope and Sentinel-2

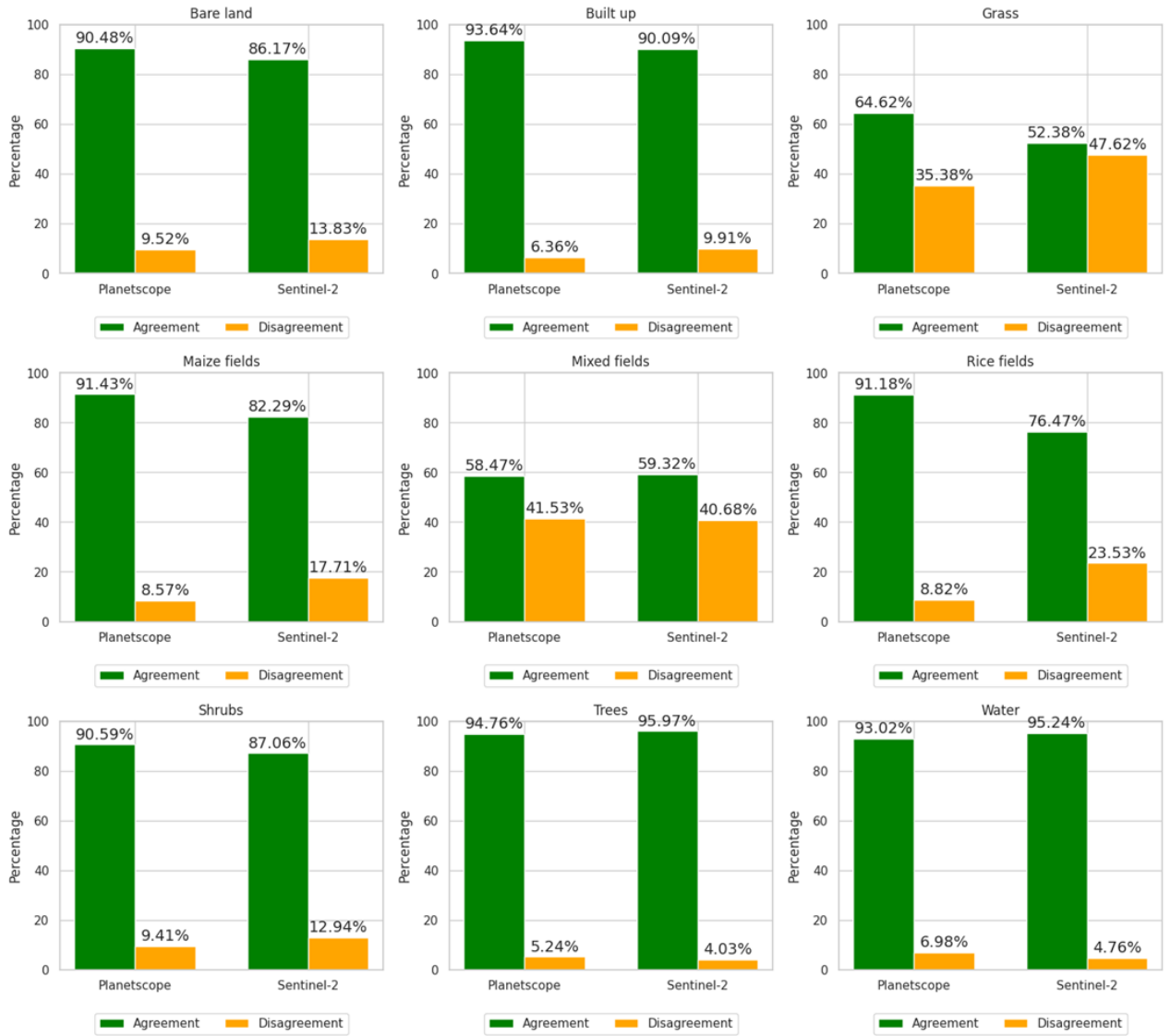


Figure 31: Agreement and disagreement in land cover classification between PlanetScope and Sentinel-2 data for various classes.

3.6. McNemar's Test.

McNemar's test was conducted to assess whether the classification of maize and mixed fields with Sentinel-2 and PlanetScope were statistically significant. **Table 9** summarizes this analysis presenting counts for various scenarios: fields where both models agreed (correct or incorrect), and where they disagreed (one dataset correct, the other incorrect). Based on this p-value (0.8283) there is no a statistically significant difference between the two datasets on classes of interest. A summary of the contingency table is displayed in **Figure 32**. The table presents a comparison between Sentinel-2 and PlanetScope classification results for maize fields and mixed fields. Out of the total data points, 229 were correctly classified by both models, while 65 cases were incorrectly classified by both. Additionally, there were 44 cases where Sentinel-2 correctly classified the data points, but PlanetScope did not and 41 cases where PlanetScope was correct, but Sentinel-2 was not. The McNemar test statistic for this comparison is 0.0471 with a p-value of 0.8283 indicating no significant difference in the classification performance between the two models. This analysis suggests that both models perform similarly in classifying maize fields and mixed fields.

Table 11: Summary of McNemar's Test for maize and mixed fields

Summary of McNemar's Test for Maize Fields and Mixed Fields

Statistic	Value
Both Correct (n11)	229
Both Incorrect (n00)	65
Sentinel-2 Correct, PlanetScope Incorrect (n10)	44
PlanetScope Correct, Sentinel-2 incorrect (n01)	41
McNemar Test Statistic	0.0471
P-value	0.8283

4. DISCUSSION

4.1. Correlation Analysis and Experiments for input variables

In general, machine learning (ML) models rely on useful features to make accurate predictions (Theng & Bhojar, 2024). It is good practice to supply these ML models with as much useful information as possible both in terms of correct sample data and explanatory variables including spatial information. The challenge lies in determining how much information is too much. While it might seem important to include all available variables to capture all possible characteristics, this approach can fail (Theng & Bhojar, 2024). Too many variables, especially those that are redundant or irrelevant, lead to increased computational processing, overfitting, and eventually, poor model performance (Danasingh et al., 2020; Sarker, 2021). This phenomenon is known as the curse of dimensionality. This occurs when adding too many features makes the model less effective. With too many variables, the space in which the model operates becomes so vast that the data points become sparse and spread out (Crespo Márquez, 2022). This makes it harder for the model to learn and make accurate predictions. In this study, the issue of curse of dimensionality and variable redundancy was handled by conducting a Pearson correlation analysis to identify variables that might potentially be providing redundant information to the model. Then the correlation coefficients of the various variables were used to set up different experiments as seen in (section 2.5.3, Table 7).

By conducting a Pearson correlation analysis, this study aimed at inspecting the overall model performance across different experiments with different variable combinations. This was to ensure an optimal balance between input variables and model efficiency. While popular methods like principal component analysis (PCA) and feature selection techniques (forward selection, backward elimination, and recursive feature elimination) are commonly used, they have limitations (Nick et al., 2015; Srinivas et al., 2023). PCA, for instance, reduces dimensionality but can make transformed features difficult to interpret. Feature selection methods simplify models by selecting important features but may not always capture the best feature set. However, the approach used in this study is straightforward to implement and allows for direct tracking of variable importance, reduces computational complexities, and increases efficiency.

Unlike PCA, which can make features hard to interpret, this correlation-based approach keeps the original features understandable while ensuring the model is manageable and efficient. Similar ideas are demonstrated by (Chen et al., 2019; Hall, 2024). The findings from the correlation analysis in this study showed expected high correlations among adjacent spectral bands and texture features derived from these bands. This was particularly evident in the band correlation matrix (Section 3) where adjacent bands exhibited very high correlations. Similarly, texture features derived from these bands also showed high correlations. This is expected according to existing knowledge in remote sensing. This is because adjacent bands tend to often capture similar information because of their close wavelengths (Đideliija et al., 2023). By understanding and managing these correlations, this study was able to identify high-performing experiments as shown in (Table 10 and 11). This helped to fast-track the variable selection process for both datasets ensuring optimal performances and robust model predictions.

4.2. Feature Importance for the discrimination of different maize cropping systems

An important aspect of this study was determining which variables from the Sentinel-2 and PlanetScope datasets were most effective in distinguishing different maize cropping systems. This directly addresses the first research objective of this study. To achieve this, this study employed three methods to assess feature importance: Mean Decrease Gini (MDG), SHapley Additive exPlanations (SHAP), and Permutation Importance. Each method offers a unique perspective on the role of each feature in the model's predictive capabilities, as explained in **Section 2.5.4**. In this study, MDG and Permutation Importance helped assess overall the variables that were considered as important by the RF model. Additionally, SHAP provided detailed insights into the predictive power of each feature for predicting the specific classes under study.

The overall variable importance analysis for Sentinel-2 and PlanetScope datasets generated using MDG and permutation methods revealed both expected and unexpected findings. For Sentinel-2, the high ranking of B5, B11, B2, EVI, and NDWI was expected due to their known effectiveness in vegetation monitoring and distinguishing different land cover types (Al-Ali et al., 2024; Misra et al., 2020; Phiri et al., 2020).

Elevation emerged as the top feature for both datasets indicating that topographic features played a significant role in the classification model. This might be because elevation is often uncorrelated with other remote sensing data providing unique information. Additionally, elevation influences climate conditions which affect crop suitability. In the study area, crops typically grow in valleys or where water is accessible while other land cover classes are distributed in higher and lower areas.

For instance, Noi Phan et al., (2020) also demonstrated that elevation was a highly influential factor in their land cover classification study using the RF classifier on five different datasets. For PlanetScope dataset, the high importance of the green (G) and red (R) bands were anticipated due to their spectral sensitivity to vegetation health and biomass (Houborg & McCabe, 2016; Vidican et al., 2023). The presence of texture features like G_diss and R_contrast also aligns with expectations as they capture spatial patterns within the imagery. However, for PlanetScope, vegetation indices were less influential compared to Sentinel-2. This could likely be attributed to PlanetScope's lower spectral resolution which limits the range of spectral information available for calculating indices.

The SHAP analysis across different composites for the Sentinel-2 dataset revealed some key observations specifically for maize and mixed crops. For the rain season composite, Bands B8, B3, B7, and the GNDVI index were most influential in predicting maize and mixed crop fields. In the annual composite, Bands B4, B12, B3, and elevation emerged as significant with textural features like B3_diss, B4_contrast, and B7_var also playing important roles. For the multi-stack composite, Bands B11, B3, B2, were the most prominent in increasing the model's predictive power to map maize and mixed fields.

These findings are expected and align with existing knowledge in vegetation monitoring and remote sensing. This is because Bands B8 (near-infrared), B3 (green), and B7 (red edge) are well-known for their sensitivity to vegetation characteristics making them valuable for distinguishing crop types (Al-Ali et al., 2024; Misra et al., 2020; Phiri et al., 2020). Indices like GNDVI are effective in assessing crop health and biomass. (Huang et al., 2021). For instance, study by Hung et al., (2019) also found that by incorporating spectral bands and GLCM features their model achieved the highest accuracy highlighting the importance of both spectral bands and textural features in attaining an optimal classification performance.

In contrast, topographic features like elevation were crucial but their influence on increasing the model's predictive power for the two agricultural field types (only maize and mixed) was less significant. However, the influence of elevation in boosting the predictive power of the model for land cover classes like trees, shrubs, bare land, and grassland was high. This was expected as agricultural fields within the study area are

mostly in low-lying areas, whereas other classes are widely distributed across both low- and high-lying areas. Similar patterns were observed for the PlanetScope datasets. Given their limited spectral depth, the red and green bands contributed positively to the model in terms of discriminating maize from mixed fields with texture features generated from the visible bands being largely influential and elevation showing similar trends.

To conclude, this study largely achieved its first research objective by identifying the most important features for distinguishing different maize cropping systems from the Sentinel-2 and PlanetScope datasets through variable importance analysis. This analysis helped to answer the research question associated with research objective 1 thoroughly by revealing variables which were useful for the discrimination of different maize cropping systems across the two datasets used in this study (Sentinel-2 and planetScope composites).

4.3. Model performance

Evaluating the overall performance of any ML model is necessary to ensure its effectiveness in performing designated tasks. This study used key metrics such as cross-validation, precision, recall, and F1-score, to assess the performance of the RF model (**Section 3.4**). These metrics are widely used in machine learning tasks such as classification to provide a robust assessment of model performance (Naidu et al., 2023). Cross-validation was particularly chosen to address sample imbalances. As demonstrated in (**Figure 9**), the distribution of training and validation points used in this study were not entirely balanced therefore there was a need to implement K-fold cross validation to confirm the model's performance is consistent and generalizable across different subsets of the data. The K-fold method has proven to be useful in situations of sample imbalances (Fontanari et al., 2022a).

The performance of the classification model across different land cover classes evaluated using precision, recall, and F1-score metrics revealed both strengths and areas for improvement. While the model demonstrated robust performance for classes such as **trees**, **shrubs**, and **water bodies** across both datasets (**Sentinel-2 and PlanetScope**) due to their distinct spectral signatures, it recorded low performances with mixed crop fields, maize fields, and grasslands. Low performance for some of these classes has also been reported by other studies. For example, Rujoiu-Mare et al., (2017) revealed that the most significant omission errors observed in their classification result were that of grassland and pine trees attributed to spectral similarities. Similar conclusions can be drawn in this case for grasslands which seem to be confused up with maize, mixed fields, and shrubs by the model. Although, in this study, this confusion could partly be attributed to spectral overlaps between these classes as well as limited number of samples for training and validation.

Furthermore, mixed crop fields and maize fields exhibited significant confusion as could be seen in their confusion matrix reports (**Appendix, figures 13&14**). This confusion was highly anticipated. The main reason could be attributed to the nature of the cropping system largely practiced within the study area. The mixed cropping system practiced within the study area makes it difficult to separate distinct crop types using optical data see (**Section 2.3.3**). In this case mixed fields also contain maize crops making it hard for the classifier to differentiate due to the spectral and textural similarity between these crop fields. This is further demonstrated through the input class variability check performed in this study (see **Appendix**). These two classes do not show enough variability in terms of spectral information. In other words, the two classes (maize and mixed fields) have very similar spectral information which makes it difficult for the model to properly separate them.

Rice fields on the other hand, showed varied high performance which could be attributed to their distinct spectral signatures. Classification result for rice fields confirms that rice fields were best captured using the multi-stack composites for both Sentinel-2 and PlanetScope datasets. This is clearly shown by the high

precision, recall, and F1-scores. The multi-stack composites combine multiple temporal snapshots. This could have enhanced the model's ability to detect rice fields more accurately across different growth stages. This high performance is to be expected due to several reasons. Rice fields have unique spectral characteristics that change distinctly with different growth stages making them easier to identify when multiple temporal images are used. The multi-stack composite approach captures these temporal variations providing a more comprehensive dataset that improves the model's ability to distinguish rice fields from other land cover types as confirmed in a study by Kustiyo et al., (2024). Their study indicated that the integration of data from different seasonal periods, specifically the rainy and dry seasons enhanced the accuracy of classification results.

In summary, the RF model performed well overall excelling with distinct classes like trees and water but had lower performance with similar classes like mixed crops and maize. Multi-stack composites effectively identified rice fields.

4.4. Comparison of Classification Accuracy between Sentinel-2 and PlanetScope

Although PlanetScope and Sentinel-2 models both performed moderately well, and the classification accuracies obtained for both models were close (**Section 3.4**), **research objective 2** of this study remained unattained. For a definitive answer to **research objective 2**, a statistical test was needed to determine whether the accuracies obtained from the Sentinel-2 and PlanetScope models were significantly different in discriminating different maize cropping systems. This was done by focusing only on the maize and mixed fields class and required statistical testing to confirm the formulated hypothesis.

Therefore, a McNemar's test was conducted as it is designed for paired nominal data which fits our scenario where the same set of fields is classified by both models (Pembury Smith & Ruxton, 2020). The test specifically evaluates the differences in paired proportions, making it suitable for assessing whether the two models significantly differ in their classification performance (Kavzoglu, 2017). McNemar's test has been used in closely similar application like comparing classification models (Abdi, 2019). This is because of its simplicity and clarity offering a straightforward result in terms of a test statistic and p-value.

McNemar's test reaches its conclusion by calculating a test statistic based on the differences in the paired classifications of the two models. Specifically, it examines the cases where the models disagree and where one model correctly classifies an instance that the other model misclassifies, and vice versa. The test statistic is then compared to a critical value from the chi-squared distribution to determine statistical significance. If the test statistic exceeds the critical value, the null hypothesis (that there is no difference between the models) is rejected see (**Section 2.5.6**).

The results of the McNemar's test concluded that there was not a statistically significant differences between Sentinel-2 and PlanetScope models for discriminating maize and mixed fields (different maize cropping systems) as summarized in (**Table 9 & Figure 29**). This conclusion is reasonable considering the p-value obtained (0.8283). This was higher than the typical significance threshold (0.05) indicating that any observed difference in performance could be due to random chance rather than a true difference in model capability.

For instance, a study by Rösch et al., (2022) comparing the same datasets (PlanetScope and Sentinel-2) in mapping pines found that Sentinel-2 achieved comparable results to PlanetScope with accuracies of 90.65% and 90.96% respectively, despite its lower spatial resolution. Again, another study by (Zagajewski et al., 2024) found that Sentinel-2 outperformed PlanetScope in some cases when classifying goldenrod. Specifically, the use of multitemporal Sentinel-2 images classified with the RF classifier achieved higher F1-scores (0.92–0.95 for goldenrod-dominated areas and 0.85–0.89 for heterogeneous areas) compared to

PlanetScope data. These studies highlight not only the significance of high spatial resolution but also the crucial role of high spectral resolution in classification tasks.

This result was contrary to the study's initial expectation that PlanetScope would perform exceptionally well. The expectation was based on the assumption that finer spatial details would aid in better discrimination of crop types. However, the results demonstrated that both datasets have their strengths and spatial resolution alone may not be the determining factor in classification accuracy.

4.5. Limitations and Recommendations

Despite some promising results, this study has a few limitations that need to be acknowledged. A primary limitation was the accuracy of geotagged photos used to capture the location of various crop fields within the study area. These photos were taken with a variety of mobile phones which could have led to inconsistencies in GPS accuracy due to the varying quality of devices used. This variability was evident during the data preparation stage where a notable number of photos could not be accurately associated with any existing plot due to positional inaccuracies. (Zandbergen & Barbeau, 2011) highlighted the need for caution when using mobile phones as GPS devices for data collection. To overcome this limitation, future study could consider standardizing the data collection process by using high-precision GPS devices across all data collection efforts. Using high-precision GPS devices to standardize data collection enhances geographical data accuracy and consistency which is crucial for mapping and classification tasks. This uniformity will help to reduce positional errors, improve data integration, and promote data quality leading to the generation of more reliable and accurate results (Dauwalter et al., 2006; Gao, 2002.; Lunetta et al., 1991).

Secondly, this study also encountered difficulties in ensuring a representative distribution of samples across all land cover classes. The challenge of achieving a representative distribution of samples across all land cover classes is a common issue acknowledged in land cover classification studies (Fontanari et al., 2022b; C. Li et al., 2021) particularly, when dealing with naturally imbalanced datasets. In this study, a stratified sampling approach was employed to split data into training and validation sets to maintain balance (section 2.5.4). However, because the original dataset itself was imbalanced, the stratified approach might still not fully mitigate the effects of class imbalance on model performance.

This bias potentially could have also affected the model's ability to generalize effectively particularly for underrepresented land cover classes as evidently shown to be case for land cover classes such grass and mixed fields (Figure 9). To improve this in future studies, increasing the number of field observations to cover various land cover classes more adequately could strengthen the stratification method. Such enhancements would help ensure a more balanced dataset potentially leading to better model performance and a more reliable generalization across diverse land cover types. For instance, studies by Mellor et al., (2015); Z. Zhou et al., (2023) emphasized the importance of having balanced training data for achieving high classification accuracy. Mellor et al., (2015) found that balanced datasets result in the lowest overall error rates for both binary and multiclass classifications. Similarly, Zhou et al., (2023) highlighted that unbalanced data can negatively impact classification accuracy in remote sensing image segmentation and proposed a dynamic weighting method to improve the accuracy of underrepresented classes while maintaining overall segmentation performance. These findings collectively emphasize the key role of balanced datasets in ensuring robust and accurate classification outcomes across various ML models not just RF.

Finally, another potential limitation of this study stems from the use of standalone satellite images from Sentinel-2 and PlanetScope datasets without exploring the benefits of data fusion. While both datasets performed adequately on their own, merging these datasets could potentially enhance the classification accuracy and robustness. Data fusion could capitalize on the unique strengths of each dataset which could

allow them to complement each other where one falls short (Zhu et al., 2018). For instance, Sentinel-2 offers high spectral resolution (**Table 2**) which is beneficial for distinguishing between different types of vegetation and other land covers based on their spectral signatures. On the other hand, PlanetScope provides high spatial resolution (**Table 3**) which is also important for identifying smaller features and finer details on the ground. By fusing these datasets, the combined data could offer a broad view integrating the high spatial resolution information from PlanetScope with the spectral diversity of Sentinel-2. The approach of fusing datasets has been proven effective in other studies (Albanwan et al., 2024; He & Wong, 2024; Zhu et al., 2018) where data fusion significantly improved the performance of the applied models.

4.6. Implications

Considering that the latest land cover maps for this study area produced by the Mozambique National Cartography and Remote Sensing Centre (CENACARTA) per the knowledge of this study dates to 2013. The updated land cover maps produced by this study offer vital insights into the evolving landscape dynamics of the Gaza province. These new maps reveal the current distribution of land cover in the region providing stakeholders with accurate and up-to-date information. Policymakers, for instance, can compare these maps to previous versions to identify areas of significant change. For example, in instances where tree cover loss has increased, this could prompt immediate conservation efforts and actions. Given the moderate overall accuracy (OA) of the maps produced in this study, these maps hold significant potential as inputs for future models especially for studies involving suitability analysis that require accurate land cover data. These reliable land cover maps can improve the precision of analyses in areas such as agricultural planning and environmental conservation. For instance, these maps can be used to identify suitable sites for new agricultural projects or conservation efforts by providing current information on land cover.

Furthermore, the maps produced in this study can assist in improving the methodological approaches of future studies. By analyzing the techniques and variables that contributed to the high accuracy of the maps in this study, future researchers can refine their own methods to achieve similar or better results. This iterative improvement process can lead to more robust and reliable land cover results. In practical terms, these maps can also facilitate better resource allocation for field validation and sampling efforts. Knowing which areas have been accurately classified can help future researchers focus their ground-truthing efforts on regions where the classification might be less certain thereby optimizing the use of time and resources.

5. CONCLUSIONS

This research aimed to evaluate the effectiveness of high spectral resolution (Sentinel-2) compared to high spatial resolution (PlanetScope) optical satellite data in distinguishing different maize cropping systems in Gaza province, Mozambique. The study focused on two main objectives: first, assessing the usefulness and importance of spectral indices, textural, and topographic features from these datasets; and second, comparing the classification accuracy of Sentinel-2 and PlanetScope imagery to determine any statistically significant differences. To ensure the accuracy and reliability of the training data, a careful sample cleaning process was employed. This included aligning point data with crop segments, removing misaligned points, and visually validating the data with high-resolution satellite imagery. Finally, K-means clustering was used to refine the polygons ensuring only homogeneous areas were considered and merging them into a comprehensive polygon of crop fields. This thorough process substantially enhanced the robustness of the data and is a key methodological strength of this study.

5.1. Effectiveness of spectral indices, textural metrics, topographic features, and spectral bands derived from optical satellite data in distinguishing different maize cropping systems

This study uncovered that Sentinel-2 red edge bands ranked among the top 10 features across all three methods used and were collectively the most useful for discriminating both maize and mixed crops. Elevation emerged as the most important topographic feature overall, but it was not relevant for distinguishing between different maize cropping systems. Textural features derived from the green and red edge bands were important overall but not particularly influential for maize and mixed fields. Specifically, texture features from the red band were important for mixed fields but not for maize fields, where blue and green bands were more influential. For PlanetScope data, vegetation indices contributed less to discriminating different maize cropping systems. Instead, the red and green bands were more effective in distinguishing maize from mixed fields which was expected as previous research indicates that spectral bands sensitive to vegetation characteristics like the green band play important role in crop discrimination. Again, it was the study's expectation that the indices calculated from PlanetScope's high spatial resolution dataset would be highly influential in discriminating the different maize cropping systems, but this turned out not to be the case in this study.

5.2. Comparison of classification accuracy between high spectral resolution (sentinel 2) and high spatial resolution (PlanetScope) satellite imagery for discriminating different maize cropping systems

The findings revealed that while the RF model demonstrated robust performance for distinct classes like trees, shrubs, and water bodies across both datasets (Sentinel-2 and PlanetScope). Classification performance in discriminating different maize cropping systems with Sentinel-2 and PlanetScope was not statistically different. This outcome supports the null hypothesis (H0) that there is no significant difference in the classification accuracy between high spectral resolution (Sentinel-2) and high spatial resolution (PlanetScope) optical satellite imagery for discriminating different maize cropping systems. Despite the higher spatial and temporal resolution of PlanetScope, it did not result in significant performance improvements. This emphasizes the critical role of other factors such as spectral resolution. Additionally, the study also revealed that employing a multi-stack composite for PlanetScope improved the overall model accuracy in predicting the maize class. In contrast, the use of a multi-stack composite for Sentinel-2 resulted in slightly lower accuracy compared to using the rain season composite data. The results demonstrated that both datasets have their strengths, and that spatial resolution alone may not be the determining factor in classification accuracy.

6. ETHICAL CONSIDERATIONS

This research utilized secondary data, therefore, direct interactions with human subjects were not required. Therefore, the primary ethical concerns focused on safeguarding data privacy. The study adhered strictly to the ethical principles and guidelines outlined by the University of Twente's Research Ethics Policy, emphasizing the responsible use and handling of data. This included obtaining informed consent when necessary and implementing stringent measures to preserve data privacy. Specifically, all identifiable information was anonymized and stored securely. Access to sensitive data was restricted to authorized personnel only. These practices highlighted the commitment to maintaining the highest ethical standards in data handling throughout the research process.

LIST OF REFERENCES

- Abdi, A. M. (2019). *GIScience & Remote Sensing Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data*.
<https://doi.org/10.1080/15481603.2019.1650447>
- Al-Ali, Z., Abulibdeh, A., Al-Awadhi, T., Mohan, M., Al Nasiri, N., Al-Barwani, M., Al Nabbi, S., & Abdullah, M. (2024). Examining the potential and effectiveness of water indices using multispectral sentinel-2 data to detect soil moisture as an indicator of mudflow occurrence in arid regions. *International Journal of Applied Earth Observation and Geoinformation*, 130, 103887.
<https://doi.org/10.1016/J.JAG.2024.103887>
- Albanwan, H., Qin, R., & Tang, Y. (2024). *Image Fusion in Remote Sensing: An Overview and Meta Analysis*.
<https://arxiv.org/abs/2401.08837v1>
- Azar, A. T., Barretta, R., & Cambaza, E. (2023). Mozambique: Country Profile. *Encyclopedia 2023, Vol. 3, Pages 143-167*, 3(1), 143–167. <https://doi.org/10.3390/ENCYCLOPEDIA3010011>
- Bansal, M., Goyal, A., & Choudhary, A. (2022). A comparative analysis of K-Nearest Neighbor, Genetic, Support Vector Machine, Decision Tree, and Long Short Term Memory algorithms in machine learning. *Decision Analytics Journal*, 3, 100071. <https://doi.org/10.1016/J.DAJOUR.2022.100071>
- Belavagi, M. C., & Muniyal, B. (2016). Performance Evaluation of Supervised Machine Learning Algorithms for Intrusion Detection. *Procedia Computer Science*, 89, 117–123.
<https://doi.org/10.1016/J.PROCS.2016.06.016>
- Belgiu, M., & Csillik, O. (2018). Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. *Remote Sensing of Environment*, 204, 509–523.
<https://doi.org/10.1016/J.RSE.2017.10.005>
- Belgiu, M., & Drăgu, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24–31.
<https://doi.org/10.1016/J.ISPRSJPRS.2016.01.011>
- Briem, G. J., Benediktsson, J. A., & Sveinsson, J. R. (2002). Multiple classifiers applied to multisource remote sensing data. *IEEE Transactions on Geoscience and Remote Sensing*, 40(10), 2291–2299.
<https://doi.org/10.1109/TGRS.2002.802476>
- Brodley, C. E., & Friedl, M. A. (1997). Decision tree classification of land cover from remotely sensed data. *Remote Sensing of Environment*, 61(3), 399–409. [https://doi.org/10.1016/S0034-4257\(97\)00049-7](https://doi.org/10.1016/S0034-4257(97)00049-7)
- Chen, S. B., Ding, C. H. Q., Zhou, Z. L., & Luo, B. (2019). Feature selection based on correlation deflation. *Neural Computing and Applications*, 31(10), 6383–6392. <https://doi.org/10.1007/S00521-018-3467-4/TABLES/4>
- Cheng, K. S., Ling, J. Y., Lin, T. W., Liu, Y. T., Shen, Y. C., & Kono, Y. (2021). Quantifying Uncertainty in Land-Use/Land-Cover Classification Accuracy: A Stochastic Simulation Approach. *Frontiers in Environmental Science*, 9, 628214. <https://doi.org/10.3389/FENVS.2021.628214/BIBTEX>
- Clausi, D. A. (2002). *An analysis of co-occurrence texture statistics as a function of grey level quantization*.
- Costa, V. G., & Pedreira, C. E. (2022). Recent advances in decision trees: an updated survey. *Artificial Intelligence Review* 2022 56:5, 56(5), 4765–4800. <https://doi.org/10.1007/S10462-022-10275-5>
- Crespo Márquez, A. (n.d.). *The Curse of Dimensionality*. https://doi.org/10.1007/978-3-030-97660-6_7
- Dabija, A., Kluczek, M., Zagajewski, B., Raczko, E., Kycko, M., Al-Sulttani, A. H., Tardà, A., Pineda, L., & Corbera, J. (2021). Comparison of Support Vector Machines and Random Forests for Corine Land

- Cover Mapping. *Remote Sensing* 2021, Vol. 13, Page 777, 13(4), 777.
<https://doi.org/10.3390/RS13040777>
- Danasingh, A. A. G. S., Subramanian, A. alias B., & Epiphany, J. L. (2020). Identifying redundant features using unsupervised learning for high-dimensional data. *SN Applied Sciences*, 2(8), 1–10.
<https://doi.org/10.1007/S42452-020-3157-6/FIGURES/2>
- Dauwalter, D. C., Fisher, W. L., & Belt, K. C. (2006). Mapping stream habitats with a global positioning system: Accuracy, precision, and comparison with traditional methods. *Environmental Management*, 37(2), 271–280. <https://doi.org/10.1007/S00267-004-0270-Z/FIGURES/5>
- de Bie, C. A. J. M., Khan, M. R., Smakhtin, V. U., Venus, V., Weir, M. J. C., & Smaling, E. M. A. (2011). Analysis of multi-temporal SPOT NDVI images for small-scale land-use mapping. *International Journal of Remote Sensing*, 32(21), 6673–6693. <https://doi.org/10.1080/01431161.2010.512939>
- De Bie, C. A., Khan, M. R., Toxopeus, A. G., Venus, V., & Skidmore, A. K. (2008). *HYPERTEMPORAL IMAGE ANALYSIS FOR CROP MAPPING AND CHANGE DETECTION*. www.VGT.vito.be.
- DeFries, R. S., & Chan, J. C. W. (2000). Multiple criteria for evaluating machine learning algorithms for land cover classification from satellite data. *Remote Sensing of Environment*, 74(3), 503–515.
[https://doi.org/10.1016/S0034-4257\(00\)00142-5](https://doi.org/10.1016/S0034-4257(00)00142-5)
- Delegido, J., Verrelst, J., Alonso, L., & Moreno, J. (2011). Evaluation of Sentinel-2 Red-Edge Bands for Empirical Estimation of Green LAI and Chlorophyll Content. *Sensors* 2011, Vol. 11, Pages 7063-7081, 11(7), 7063–7081. <https://doi.org/10.3390/S110707063>
- Didelija, M., Kulo, N., Mulahusić, A., Tuno, N., & Topoljak, J. (2023). Correlation analysis of different optical remote sensing indices for drought monitoring: a case study of Canton Sarajevo, Bosnia and Herzegovina. *Environmental Monitoring and Assessment*, 195(11), 1–19.
<https://doi.org/10.1007/S10661-023-11930-2/FIGURES/8>
- Ding, K., Wang, C., Tao, M., Xiao, H., Yang, C., & Huang, P. (2020). A Classification Method of Land Cover Based on Support Vector Machines. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 12488 LNCS, 48–54.
https://doi.org/10.1007/978-3-030-62463-7_5
- Dogan, A., & Birant, D. (2019). A Weighted Majority Voting Ensemble Approach for Classification. *2019 4th International Conference on Computer Science and Engineering (UBMK)*, 366–371.
<https://doi.org/10.1109/UBMK.2019.8907028>
- Eisavi, V., Homayouni, S., Yazdi, A. M., & Alimohammadi, A. (2015). Land cover mapping based on random forest classification of multitemporal spectral and thermal images. *Environmental Monitoring and Assessment*, 187(5), 1–14. <https://doi.org/10.1007/S10661-015-4489-3/FIGURES/8>
- ElMannai, H., Hamdi, M., & AlGarni, A. (2019). Enhanced Support Vector Machine Applied to Land-Use Classification. *Communications in Computer and Information Science*, 1097 CCIS, 236–244.
https://doi.org/10.1007/978-3-030-36365-9_20
- Fletcher, S., & Islam, M. Z. (2019). 83 Decision Tree Classification with Differential Privacy: A Survey. *ACM Comput. Surv.*, 52. <https://doi.org/10.1145/3337064>
- Fontanari, T., Fróes, T. C., & Recamonde-Mendoza, M. (2022a). Cross-validation Strategies for Balanced and Imbalanced Datasets. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 13653 LNAI, 626–640. https://doi.org/10.1007/978-3-031-21686-2_43/TABLES/4
- Fontanari, T., Fróes, T. C., & Recamonde-Mendoza, M. (2022b). Cross-validation Strategies for Balanced and Imbalanced Datasets. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial*

- Intelligence and Lecture Notes in Bioinformatics*), 13653 LNAI, 626–640. https://doi.org/10.1007/978-3-031-21686-2_43/TABLES/4
- Gao, F., & Gao, F. (2021). *Remote Sensing for Agriculture*. 7–24. https://doi.org/10.1007/978-3-030-66387-2_2
- Gao, J. (n.d.). *Integration of GPS with Remote Sensing and GIS: Reality and Prospect*.
- Ghimire, B., Rogan, J., Galiano, V., Panday, P., & Neeti, N. (2012). An evaluation of bagging, boosting, and random forests for land-cover classification in Cape Cod, Massachusetts, USA. *GIScience and Remote Sensing*, 49(5), 623–643. <https://doi.org/10.2747/1548-1603.49.5.623>
- Giri, C., Pengra, B., Long, J., & Loveland, T. R. (2013). Next generation of global land cover characterization, mapping, and monitoring. *International Journal of Applied Earth Observation and Geoinformation*, 25(1), 30–37. <https://doi.org/10.1016/J.JAG.2013.03.005>
- Gislason, P. O., Benediktsson, J. A., & Sveinsson, J. R. (2006). Random forests for land cover classification. *Pattern Recognition Letters*, 27(4), 294–300. <https://doi.org/10.1016/J.PATREC.2005.08.011>
- Hall, M. A. (n.d.). *Correlation-based Feature Selection for Discrete and Numeric Class Machine Learning*.
- Hao, P., Löw, F., & Biradar, C. (2018). Annual Cropland Mapping Using Reference Landsat Time Series—A Case Study in Central Asia. *Remote Sensing 2018, Vol. 10, Page 2057, 10(12)*, 2057. <https://doi.org/10.3390/RS10122057>
- Haralick, R. M., Dinstein, I., & Shanmugam, K. (1973). Textural Features for Image Classification. *IEEE Transactions on Systems, Man and Cybernetics, SMC-3(6)*, 610–621. <https://doi.org/10.1109/TSMC.1973.4309314>
- He, S., & Wong, S. W. K. (2024). *Spatio-temporal data fusion for the analysis of in situ and remote sensing data using the INLA-SPDE approach*.
- Holloway, J., Helmstedt, K. J., Mengersen, K., & Schmidt, M. (2019). A Decision Tree Approach for Spatially Interpolating Missing Land Cover Data and Classifying Satellite Images. *Remote Sensing 2019, Vol. 11, Page 1796, 11(15)*, 1796. <https://doi.org/10.3390/RS11151796>
- Houborg, R., & McCabe, M. F. (2016). High-Resolution NDVI from Planet’s Constellation of Earth Observing Nano-Satellites: A New Data Source for Precision Agriculture. *Remote Sensing 2016, Vol. 8, Page 768, 8(9)*, 768. <https://doi.org/10.3390/RS8090768>
- Huang, S., Tang, L., Hupy, J. P., Wang, Y., & Shao, G. (2021). A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing. *Journal of Forestry Research*, 32(1), 1–6. <https://doi.org/10.1007/S11676-020-01155-1/FIGURES/2>
- Hung, C.-C., Song, E., & Lan, Y. (2019). Image Texture, Texture Features, and Image Texture Classification and Segmentation. *Image Texture Analysis*, 3–14. https://doi.org/10.1007/978-3-030-13773-1_1
- Inglada, J., Michel, J., & Hagolle, O. (2022). Assessment of the Usefulness of Spectral Bands for the Next Generation of Sentinel-2 Satellites by Reconstruction of Missing Bands. *Remote Sensing 2022, Vol. 14, Page 2503, 14(10)*, 2503. <https://doi.org/10.3390/RS14102503>
- Kavzoglu, T. (2017). Object-Oriented Random Forest for High Resolution Land Cover Mapping Using Quickbird-2 Imagery. *Handbook of Neural Computation*, 607–619. <https://doi.org/10.1016/B978-0-12-811318-9.00033-8>
- Kustiyo, K., Rokhmatuloh, R., Saputro, A. H., & Kushardono, D. (2024). Rice fields classification through spectral-temporal data fusion during the rainy and dry seasons using Sentinel-2 optical images in Subang Regency, West Java, Indonesia. *Paddy and Water Environment*, 1–11. <https://doi.org/10.1007/S10333-024-00972-Y/FIGURES/7>

- Lächelt, Siegfried. (2004). *The geology and mineral resources of Mozambique*. 515.
https://www.researchgate.net/publication/237383440_Sedimentary_rocks_of_the_Mapai_formation_in_the_Massingir-Mapai_region_Gaza_province_Mozambique
- Leon, A. C. (1998). Descriptive and Inferential Statistics. *Comprehensive Clinical Psychology*, 243–285.
[https://doi.org/10.1016/B0080-4270\(73\)00264-9](https://doi.org/10.1016/B0080-4270(73)00264-9)
- Li, C., Ma, Z., Wang, L., Yu, W., Tan, D., Gao, B., Feng, Q., Guo, H., & Zhao, Y. (2021). Improving the Accuracy of Land Cover Mapping by Distributing Training Samples. *Remote Sensing 2021, Vol. 13, Page 4594, 13(22)*, 4594. <https://doi.org/10.3390/RS13224594>
- Li, J., Roy, D. P., Atzberger, C., & Zhou, G. (2017). A Global Analysis of Sentinel-2A, Sentinel-2B and Landsat-8 Data Revisit Intervals and Implications for Terrestrial Monitoring. *Remote Sensing 2017, Vol. 9, Page 902, 9(9)*, 902. <https://doi.org/10.3390/RS9090902>
- Liu, Y., Rao, P., Zhou, W., Singh, B., Srivastava, A. K., Poonia, S. P., Van Berkel, D., & Jain, M. (2022). Using Sentinel-1, Sentinel-2, and Planet satellite data to map field-level tillage practices in smallholder systems. *PLOS ONE, 17(11)*, e0277425.
<https://doi.org/10.1371/JOURNAL.PONE.0277425>
- Louppe, G. (2014). *Understanding Random Forests: From Theory to Practice*. <https://arxiv.org/abs/1407.7502v3>
- lowast, V. R., A, M., S, V., & R, P. (2020). Unsupervised ISODATA algorithm classification used in the landsat image for predicting the expansion of Salem urban, Tamil Nadu. *Indian Journal of Science and Technology, 13(16)*, 1619–1629. <https://doi.org/10.17485/IJST/V13I16.271>
- Lu, W., Wang, X., Sun, L., & Zheng, Y. (2023). Spectral–Spatial Feature Extraction for Hyperspectral Image Classification Using Enhanced Transformer with Large-Kernel Attention. *Remote Sensing 2024, Vol. 16, Page 67, 16(1)*, 67. <https://doi.org/10.3390/RS16010067>
- Lundberg, S. M., Allen, P. G., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems, 30*. <https://github.com/slundberg/shap>
- Lunetta, R. S., Fenstermaker, L. K., Lensen, J. R., Mcgwire, K. C., & Tinny, L. R. (n.d.). *Remote Sensing and Geographic Information System Data Integration: Error Sources and Research Issues*.
- Macarringue, L. S., Bolfe, É. L., & Pereira, P. R. M. (2022). Developments in Land Use and Land Cover Classification Techniques in Remote Sensing: A Review. *Journal of Geographic Information System, 14(01)*, 1–28. <https://doi.org/10.4236/JGIS.2022.141001>
- Mandanici, E., & Bitelli, G. (2016). Preliminary comparison of sentinel-2 and landsat 8 imagery for a combined use. *Remote Sensing, 8(12)*. <https://doi.org/10.3390/RS8121014>
- Marzi, D., Sorriso, A., & Gamba, P. (2023). Automatic wide area land cover mapping using Sentinel-1 multitemporal data. *Frontiers in Remote Sensing, 4*, 1148328.
<https://doi.org/10.3389/FRSEN.2023.1148328/BIBTEX>
- Masimula, N. (2020). *Cropland and tree cover mapping using Sentinel-2 data in an agroforestry landscape, Burkina Faso*. <https://gupea.ub.gu.se/handle/2077/65237>
- Mateo-García, G., Gómez-Chova, L., Amorós-López, J., Muñoz-Marí, J., & Camps-Valls, G. (2018). Multitemporal Cloud Masking in the Google Earth Engine. *Remote Sensing 2018, Vol. 10, Page 1079, 10(7)*, 1079. <https://doi.org/10.3390/RS10071079>
- Maung, W. S., Tsuyuki, S., & Guo, Z. (2023). Improving Land Use and Land Cover Information of Wunbaik Mangrove Area in Myanmar Using U-Net Model with Multisource Remote Sensing Datasets. *Remote Sensing 2024, Vol. 16, Page 76, 16(1)*, 76. <https://doi.org/10.3390/RS16010076>
- Melesse, A. M., Weng, Q., Thenkabail, P. S., & Senay, G. B. (2007). Remote Sensing Sensors and Applications in Environmental Resources Mapping and Modelling. *Sensors 2007, Vol. 7, Pages 3209-3241, 7(12)*, 3209–3241. <https://doi.org/10.3390/S7123209>

- Mellor, A., Boukir, S., Haywood, A., & Jones, S. (2015). Exploring issues of training data imbalance and mislabelling on random forest performance for large area land cover classification using the ensemble margin. *ISPRS Journal of Photogrammetry and Remote Sensing*, *105*, 155–168. <https://doi.org/10.1016/J.ISPRSJPRS.2015.03.014>
- Menefee, D., Rajan, N., Shafian, S., & Cui, S. (2022). Modeling Carbon Uptake of Dryland Maize Using High Resolution Satellite Imagery. *Frontiers in Remote Sensing*, *3*, 810030. <https://doi.org/10.3389/FRSEN.2022.810030/BIBTEX>
- Meyer, L. H., Heurich, M., Beudert, B., Premier, J., & Pflugmacher, D. (2019). Comparison of Landsat-8 and Sentinel-2 Data for Estimation of Leaf Area Index in Temperate Forests. *Remote Sensing* *2019*, *Vol. 11*, Page 1160, *11*(10), 1160. <https://doi.org/10.3390/RS11101160>
- Misra, G., Cawkwell, F., & Wingler, A. (2020). Status of Phenological Research Using Sentinel-2 Data: A Review. *Remote Sensing* *2020*, *Vol. 12*, Page 2760, *12*(17), 2760. <https://doi.org/10.3390/RS12172760>
- Muhala, V., Rumieque, A., & Hasimuna, O. J. (2021). Aquaculture production in Mozambique: Approaches and practices by farmers in Gaza province. *The Egyptian Journal of Aquatic Research*, *47*(1), 87–92. <https://doi.org/10.1016/J.EJAR.2020.11.004>
- Naidu, G., Zuva, T., & Sibanda, E. M. (2023). A Review of Evaluation Metrics in Machine Learning Algorithms. *Lecture Notes in Networks and Systems*, *724 LNNS*, 15–25. https://doi.org/10.1007/978-3-031-35314-7_2/FIGURES/3
- Navnath, N. N., Chandrasekaran, K., Stateczny, A., Sundaram, V. M., & Panneer, P. (2022). Spatiotemporal Assessment of Satellite Image Time Series for Land Cover Classification Using Deep Learning Techniques: A Case Study of Reunion Island, France. *Remote Sensing* *2022*, *Vol. 14*, Page 5232, *14*(20), 5232. <https://doi.org/10.3390/RS14205232>
- Nick, W., Shelton, J., Bullock, G., Esterline, A., & Asamene, K. (2015). Comparing dimensionality reduction techniques. *Conference Proceedings - IEEE SOUTHEASTCON, 2015-June*(June). <https://doi.org/10.1109/SECON.2015.7132997>
- Nie, F., Zhu, W., & Li, X. (2020). Decision Tree SVM: An extension of linear SVM for non-linear classification. *Neurocomputing*, *401*, 153–159. <https://doi.org/10.1016/J.NEUCOM.2019.10.051>
- Noi Phan, T., Kuch, V., & Lehnert, L. W. (2020). Land cover classification using google earth engine and random forest classifier-the role of image composition. *Remote Sensing*, *12*(15). <https://doi.org/10.3390/RS12152411>
- Otunga, C., Odindi, J., Mutanga, O., & Adjorlolo, C. (2019). Evaluating the potential of the red edge channel for C3 (*Festuca* spp.) grass discrimination using Sentinel-2 and Rapid Eye satellite image data. *Geocarto International*, *34*(10), 1123–1143. <https://doi.org/10.1080/10106049.2018.1474274>
- Pal, M., & Mather, P. M. (2003). An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sensing of Environment*, *86*(4), 554–565. [https://doi.org/10.1016/S0034-4257\(03\)00132-9](https://doi.org/10.1016/S0034-4257(03)00132-9)
- Pazúr, R., Huber, N., Weber, D., Ginzler, C., & Price, B. (2022). A national extent map of cropland and grassland for Switzerland based on Sentinel-2 data. *Earth System Science Data*, *14*(1), 295–305. <https://doi.org/10.5194/ESSD-14-295-2022>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Louppe, G., Prettenhofer, P., Weiss, R., Weiss, R. J., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*. <https://doi.org/10.5555/1953048.2078195>
- Pembury Smith, M. Q. R., & Ruxton, G. D. (2020). Effective use of the McNemar test. *Behavioral Ecology and Sociobiology*, *74*(11), 1–9. <https://doi.org/10.1007/S00265-020-02916-Y/TABLES/5>

- Peng, J., Li, L., & Tang, Y. Y. (2019). Maximum Likelihood Estimation-Based Joint Sparse Representation for the Classification of Hyperspectral Remote Sensing Images. *IEEE Transactions on Neural Networks and Learning Systems*, 30(6), 1790–1802. <https://doi.org/10.1109/TNNLS.2018.2874432>
- Pereira, J., Stroes, E. S. G., Zwinderman, A. H., & Levin, E. (n.d.). *Covered Information Disentanglement: Model Transparency via Unbiased Permutation Importance*. Retrieved June 25, 2024, from <https://github.com/JBPereira/CID>.
- Phiri, D., Simwanda, M., Salekin, S., Nyirenda, V. R., Murayama, Y., & Ranagalage, M. (2020). Sentinel-2 Data for Land Cover/Use Mapping: A Review. *Remote Sensing 2020, Vol. 12, Page 2291, 12*(14), 2291. <https://doi.org/10.3390/RS12142291>
- Pickering, J., Tyukavina, A., Lima, A., Khan, A., Potapov, P., Adusei, B., & Hansen, M. C. (2021). Using multi-resolution satellite data to quantify land dynamics: Applications of planetscope imagery for cropland and tree-cover loss area estimation. *Remote Sensing*, 13(11). <https://doi.org/10.3390/RS13112191>
- Ponguane, S., Mucavele, N., & Mussumbuluco, B. (2023). LAND GRABBING OR RICE SECTOR DEVELOPMENT OPPORTUNITY? *African Journal of Land Policy and Geospatial Sciences*, 3(4), 343–353. <https://doi.org/10.22004/AG.ECON.334443>
- Prasad, A. D., Ganasala, P., Hernández-Guzmán, R., & Fathian, F. (2022). Remote sensing satellite data and spectral indices: an initial evaluation for the sustainable development of an urban area. *Sustainable Water Resources Management*, 8(1), 1–16. <https://doi.org/10.1007/S40899-022-00607-2/METRICS>
- Pratik, J., Vishwambhar, K., Aakash, M., & Kavita, J. (2023). Crop detection using satellite image processing. *I-Manager's Journal on Image Processing*, 10(2), 50. <https://doi.org/10.26634/JIP.10.2.19800>
- Puls, E. da S., Todescato, M. V., & Carbonera, J. L. (2023). *An evaluation of pre-trained models for feature extraction in image classification*. <http://arxiv.org/abs/2310.02037>
- Rao, P., Zhou, W., Bhattarai, N., Srivastava, A. K., Singh, B., Poonia, S., Lobell, D. B., & Jain, M. (2021). Using sentinel-1, sentinel-2, and planet imagery to map crop type of smallholder farms. *Remote Sensing*, 13(10), 1870. <https://doi.org/10.3390/RS13101870/S1>
- Rawat, S., & Saini, R. (2022). Cropland Mapping Using Single Date Sentinel-2 Imagery Using Machine Learning Classifiers. *2022 International Conference on Advances in Computing, Communication and Materials, ICACCM 2022*. <https://doi.org/10.1109/ICACCM56405.2022.10009406>
- Reda, W., Elazhary, H., & Hassanein, E. (2019). Comparing K-means, K-medoids and ISODATA clustering algorithms for a cloud service search engine. *International Journal of Recent Technology and Engineering*, 8(3), 4978–4982. <https://doi.org/10.35940/IJRTE.C5632.098319>
- Richards, J. A. (2022). Clustering and Unsupervised Classification. *Remote Sensing Digital Image Analysis*, 369–401. https://doi.org/10.1007/978-3-030-82327-6_9
- Rösch, M., Sonnenschein, R., Buchelt, S., & Ullmann, T. (2022). Comparing PlanetScope and Sentinel-2 Imagery for Mapping Mountain Pines in the Sarntal Alps, Italy. *Remote Sensing 2022, Vol. 14, Page 3190, 14*(13), 3190. <https://doi.org/10.3390/RS14133190>
- Rujoiu-Mare, M. R., Olariu, B., Mihai, B. A., Nistor, C., & Săvulescu, I. (2017). Land cover classification in Romanian Carpathians and Subcarpathians using multi-date Sentinel-2 remote sensing imagery. *European Journal of Remote Sensing*, 50(1), 496–508. <https://doi.org/10.1080/22797254.2017.1365570>
- Salite, D. (2019). Traditional prediction of drought under weather and climate uncertainty: analyzing the challenges and opportunities for small-scale farmers in Gaza province, southern region of Mozambique. *Natural Hazards*, 96(3), 1289–1309. <https://doi.org/10.1007/S11069-019-03613-4/FIGURES/1>

- Samaniego, L., Bárdossy, A., & Schulz, K. (2008). Supervised classification of remotely sensed imagery using a modified k-NN technique. *IEEE Transactions on Geoscience and Remote Sensing*, 46(7), 2112–2125. <https://doi.org/10.1109/TGRS.2008.916629>
- Sandri, M., & Zuccolotto, P. (2010). Analysis and correction of bias in Total Decrease in Node Impurity measures for tree-based algorithms. *Statistics and Computing*, 20(4), 393–407. <https://doi.org/10.1007/S11222-009-9132-0/METRICS>
- Sarker, I. H. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*, 2(3), 1–21. <https://doi.org/10.1007/S42979-021-00592-X/FIGURES/11>
- Saveca, P. S. L., Abi, A., Stigter, T. Y., Lukas, E., & Fourie, F. (2022). Assessing Groundwater Dynamics and Hydrological Processes in the Sand River Deposits of the Limpopo River, Mozambique. *Frontiers in Water*, 3, 731642. <https://doi.org/10.3389/FRWA.2021.731642/BIBTEX>
- Schober, P., & Schwarte, L. A. (2018). Correlation coefficients: Appropriate use and interpretation. *Anesthesia and Analgesia*, 126(5), 1763–1768. <https://doi.org/10.1213/ANE.0000000000002864>
- Schulz, D., Yin, H., Tischbein, B., Verleysdonk, S., Adamou, R., & Kumar, N. (2021). Land use mapping using Sentinel-1 and Sentinel-2 time series in a heterogeneous landscape in Niger, Sahel. *ISPRS Journal of Photogrammetry and Remote Sensing*, 178, 97–111. <https://doi.org/10.1016/j.isprsjprs.2021.06.005>
- Sheykhmousa, M., Mahdianpari, M., Ghanbari, H., Mohammadimanesh, F., Ghamisi, P., & Homayouni, S. (2020). Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 6308–6325. <https://doi.org/10.1109/JSTARS.2020.3026724>
- Shi, D., & Yang, X. (2015). *Support Vector Machines for Land Cover Mapping from Remote Sensor Imagery*. 265–279. https://doi.org/10.1007/978-94-017-9813-6_13
- Singh, P. K., & Sharma, A. (2022). An intelligent WSN-UAV-based IoT framework for precision agriculture application. *Computers and Electrical Engineering*, 100. <https://doi.org/10.1016/j.compeleceng.2022.107912>
- Somvanshi, S. S., Bhalla, O., Kunwar, P., Singh, M., & Singh, P. (2020). Monitoring spatial LULC changes and its growth prediction based on statistical models and earth observation datasets of Gautam Budh Nagar, Uttar Pradesh, India. *Environment, Development and Sustainability*, 22(2), 1073–1091. <https://doi.org/10.1007/s10668-018-0234-8>
- Song, X. P., Huang, W., Hansen, M. C., & Potapov, P. (2021). An evaluation of Landsat, Sentinel-2, Sentinel-1 and MODIS data for crop type mapping. *Science of Remote Sensing*, 3, 100018. <https://doi.org/10.1016/J.SRS.2021.100018>
- Spoto, F., Sy, O., Laberinti, P., Martimort, P., Fernandez, V., Colin, O., Hoersch, B., & Meygret, A. (2012). Overview of sentinel-2. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 1707–1710. <https://doi.org/10.1109/IGARSS.2012.6351195>
- Srinivas, P., Guggari, S., Darapaneni, N., Paduri, A. R., & Sudha, B. G. (2023). Feature Selection Algorithms: A Comparative Study. *Lecture Notes in Networks and Systems*, 648 LNNS, 402–412. https://doi.org/10.1007/978-3-031-27524-1_38/FIGURES/5
- Sunar Erbek, F., Özkan, C., & Taberner, M. (2004). Comparison of maximum likelihood classification method with supervised artificial neural network algorithms for land use activities. *International Journal of Remote Sensing*, 25(9), 1733–1748. <https://doi.org/10.1080/0143116031000150077>
- Talekar, B. (2020). A Detailed Review on Decision Tree and Random Forest. *Bioscience Biotechnology Research Communications*, 13(14), 245–248. <https://doi.org/10.21786/BBRC/13.14/57>

- Theng, D., & Bhoyar, K. K. (2024). Feature selection techniques for machine learning: a survey of more than two decades of research. *Knowledge and Information Systems*, 66(3), 1575–1637. <https://doi.org/10.1007/S10115-023-02010-5/TABLES/6>
- Tran, T. V., Reef, R., & Zhu, X. (2022). A Review of Spectral Indices for Mangrove Remote Sensing. *Remote Sensing 2022, Vol. 14, Page 4868, 14(19)*, 4868. <https://doi.org/10.3390/RS14194868>
- Trivedi, M. B., Marshall, M., Estes, L., de Bie, C. A. J. M., Chang, L., & Nelson, A. (2023). Cropland Mapping in Tropical Smallholder Systems with Seasonally Stratified Sentinel-1 and Sentinel-2 Spectral and Textural Features. *Remote Sensing 2023, Vol. 15, Page 3014, 15(12)*, 3014. <https://doi.org/10.3390/RS15123014>
- Vidican, R., Mălinaș, A., Ranta, O., Moldovan, C., Marian, O., Ghețe, A., Ghișe, C. R., Popovici, F., & Cătunescu, G. M. (2023). Using Remote Sensing Vegetation Indices for the Discrimination and Monitoring of Agricultural Crops: A Critical Review. *Agronomy*, 13(12), 3040. <https://doi.org/10.3390/AGRONOMY13123040/S1>
- Villegas Rugel, G. M., Ochoa, D., Menendez, J. M., & Van Coillie, F. (2023). Evaluating the Applicability of Global LULC Products and an Author-Generated Phenology-Based Map for Regional Analysis: A Case Study in Ecuador's Ecoregions. *Land*, 12(5). <https://doi.org/10.3390/LAND12051112>
- Wang, D., & Cheng, B. (2010). An Unsupervised Classification Method of Remote Sensing Images Based on Ant Colony Optimization Algorithm. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6440 LNAI(PART 1), 294–301. https://doi.org/10.1007/978-3-642-17316-5_29
- World Bank. (2017). *Climate-smart agriculture (CSA) considerations*.
- Zagajewski, B., Kluczek, M., Zdunek, K. B., & Holland, D. (2024). Sentinel-2 versus PlanetScope Images for Goldenrod Invasive Plant Species Mapping. *Remote Sensing 2024, Vol. 16, Page 636, 16(4)*, 636. <https://doi.org/10.3390/RS16040636>
- Zandbergen, P. A., & Barbeau, S. J. (2011). Positional Accuracy of Assisted GPS Data from High-Sensitivity GPS-enabled Mobile Phones. *The Journal of Navigation*, 64(3), 381–399. <https://doi.org/10.1017/S0373463311000051>
- Zhang, T. X., Su, J. Y., Liu, C. J., & Chen, W. H. (2019). Potential Bands of Sentinel-2A Satellite for Classification Problems in Precision Agriculture. *International Journal of Automation and Computing*, 16(1), 16–26. <https://doi.org/10.1007/S11633-018-1143-X>
- Zhou, C., Feng, H., Yang, Y., Zhang, N., & Ibrahim, ad. (2022). Improving Land Use/Cover Classification Accuracy from Random Forest Feature Importance Selection Based on Synergistic Use of Sentinel Data and Digital Elevation Model in Agriculturally Dominated Landscape. *Agriculture 2023, Vol. 13, Page 98, 13(1)*, 98. <https://doi.org/10.3390/AGRICULTURE13010098>
- Zhou, X., Ding, P. L. K., & Li, B. (2019). Improving Robustness of Random Forest Under Label Noise. *IEEE Workshop/Winter Conference on Applications of Computer Vision*, 950–958. <https://doi.org/10.1109/WACV.2019.00106>
- Zhou, Z., Zheng, C., Liu, X., Tian, Y., Chen, X., Chen, X., & Dong, Z. (2023). A Dynamic Effective Class Balanced Approach for Remote Sensing Imagery Semantic Segmentation of Imbalanced Data. *Remote Sensing 2023, Vol. 15, Page 1768, 15(7)*, 1768. <https://doi.org/10.3390/RS15071768>
- Zhu, X., Cai, F., Tian, J., & Williams, T. K. A. (2018). Spatiotemporal Fusion of Multisource Remote Sensing Data: Literature Survey, Taxonomy, Principles, Applications, and Future Directions. *Remote Sensing 2018, Vol. 10, Page 527, 10(4)*, 527. <https://doi.org/10.3390/RS10040527>

APPENDIXES

Classification Scheme for Dynamic world and ESA land cover maps

Table: 2 Classification and description of land cover types in the ESA 2021 map including colour representation and class values.

Class Name	Description	Color
Tree Cover	Areas predominantly covered by trees	Dark Green
Shrubland	Areas predominantly covered by shrubs	Light Orange
Grassland	Areas predominantly covered by grass	Yellow
Cropland	Areas used for agricultural purposes	Light Purple
Built-up	Urban and industrial areas	Red
Bare / Sparse Vegetation	Areas with little to no vegetation	Light Gray
Snow and Ice	Areas covered with snow or ice	White
Permanent Water Bodies	Rivers, lakes, reservoirs, and other permanent water bodies	Blue
Herbaceous Wetland	Areas with herbaceous plants that are permanently or seasonally inundated with water	Cyan
Mangroves	Coastal wetlands with mangrove vegetation	Light Green
Moss and Lichen	Areas predominantly covered by mosses and lichens	Light Yellow

Table: 2 Classification and description of land cover types in the Dynamic world 2023 map including colour representation

Name	Description	Color
Tree Cover	Areas covered with trees	Dark Green
Shrubland	Areas dominated by shrubs	Light Green
Grassland	Areas covered with grass	Yellow
Cropland	Areas used for growing crops	Brown
Built-up	Urban areas and other man-made structures	Red
Bare/Sparse Vegetation	Areas with minimal vegetation	Light Brown
Snow and Ice	Regions covered with snow or ice	White
Permanent Water Bodies	Bodies of water such as lakes and rivers	Blue
Herbaceous Wetland	Wetlands dominated by herbaceous plants	Light Blue
Mangrove	Coastal wetlands with mangrove trees	Dark Blue
Moss and Lichen	Areas dominated by moss	Grey

Variability plots

This study used box plots to examine the variability of key for both Sentinel-2 and PlanetScope satellite imagery along with terrain features such as elevation, slope, and aspect. These plots help to evaluate the dispersion of data points within each land cover class offering insights into the distinctiveness and separability of these classes based on their properties. The plots provide a visual representation of this variability and can be seen below (1 to 12).

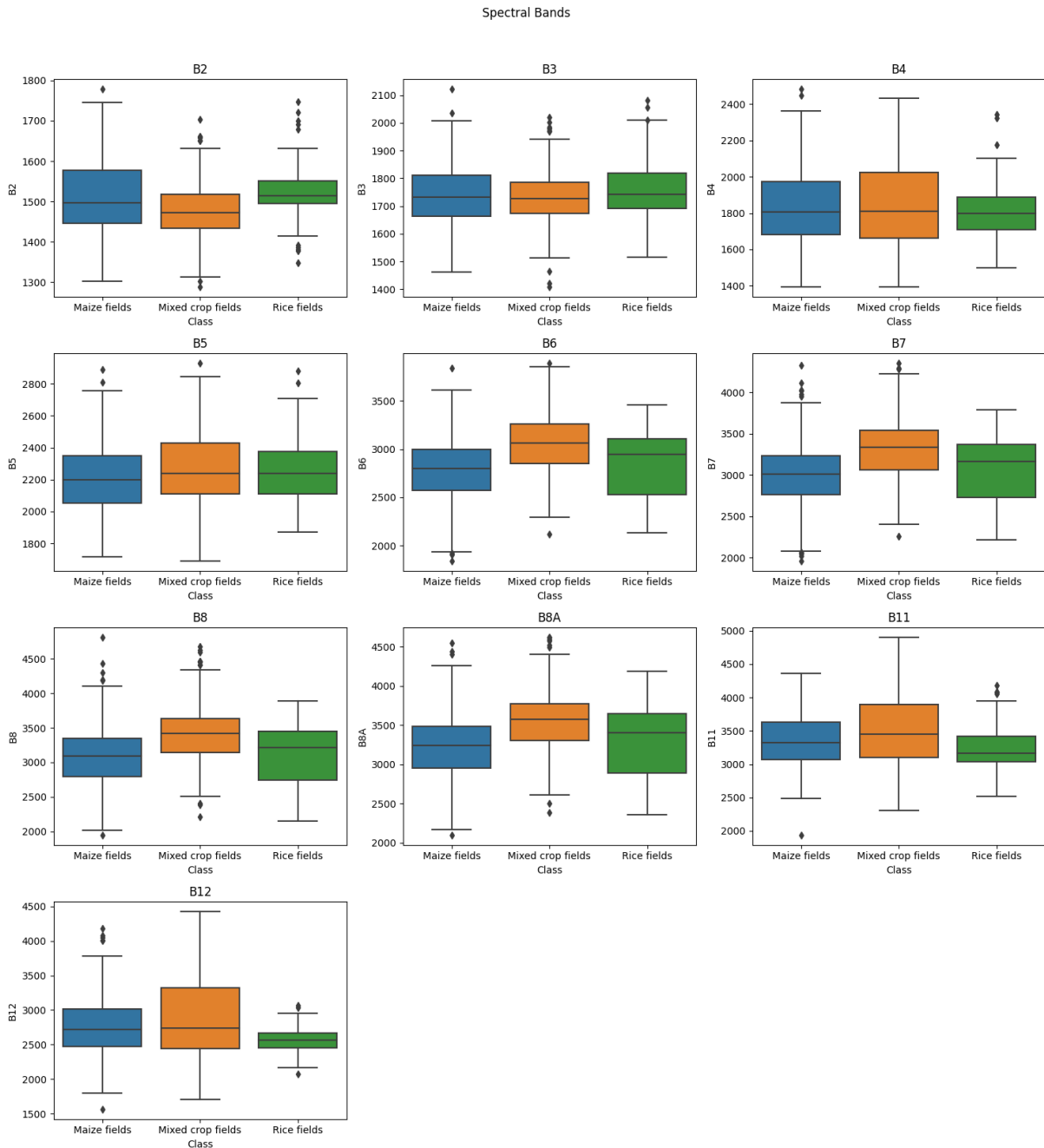


Figure 1: Box plots showing the distribution of Sentinel-2 spectral band values (B2, B3, B4, B5, B6, B7, B8, B8A, B11, and B12) across maize fields, mixed crop fields, and rice fields. Each plot displays the interquartile range, median, and outliers for the respective spectral band showing the variability between the land cover classes.

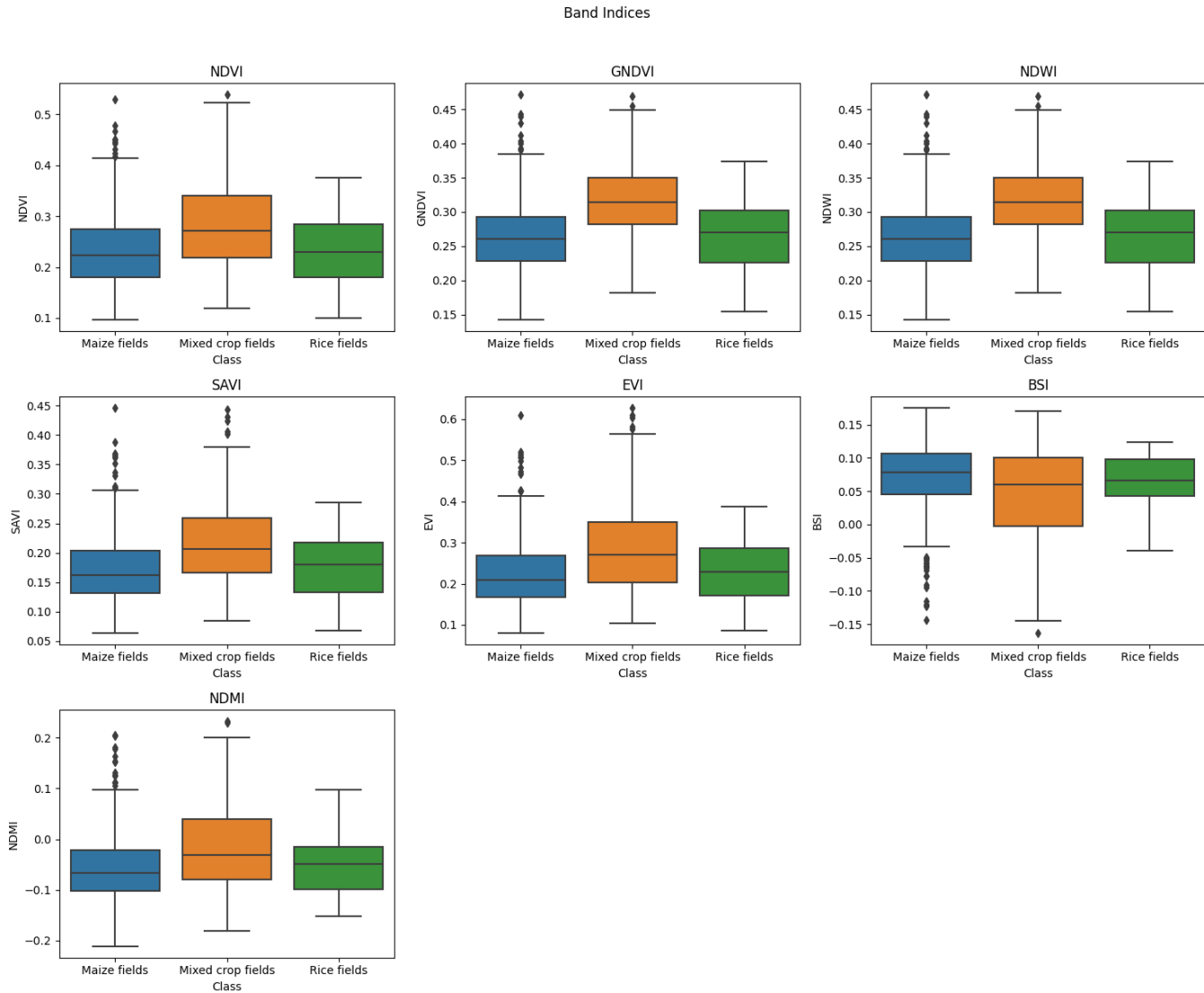


Figure 2: Box plots showing the distribution of spectral index values (NDVI, GNDVI, NDWI, SAVI, EVI, BSI, and NDMI) from Sentinel-2 data across maize fields, mixed crop fields, and rice fields.

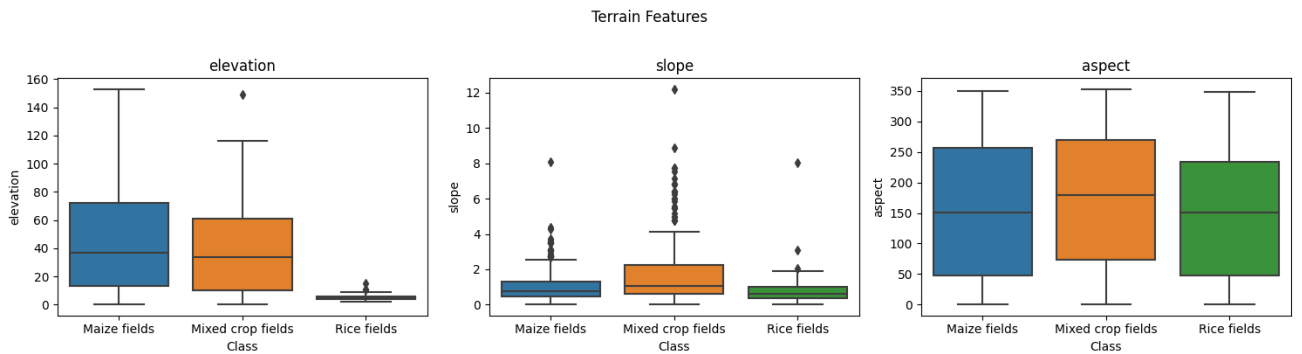


Figure 3: Box plots showing the distribution of terrain feature values (elevation, slope, and aspect) across maize fields, mixed crop fields, and rice fields.

Spectral Bands

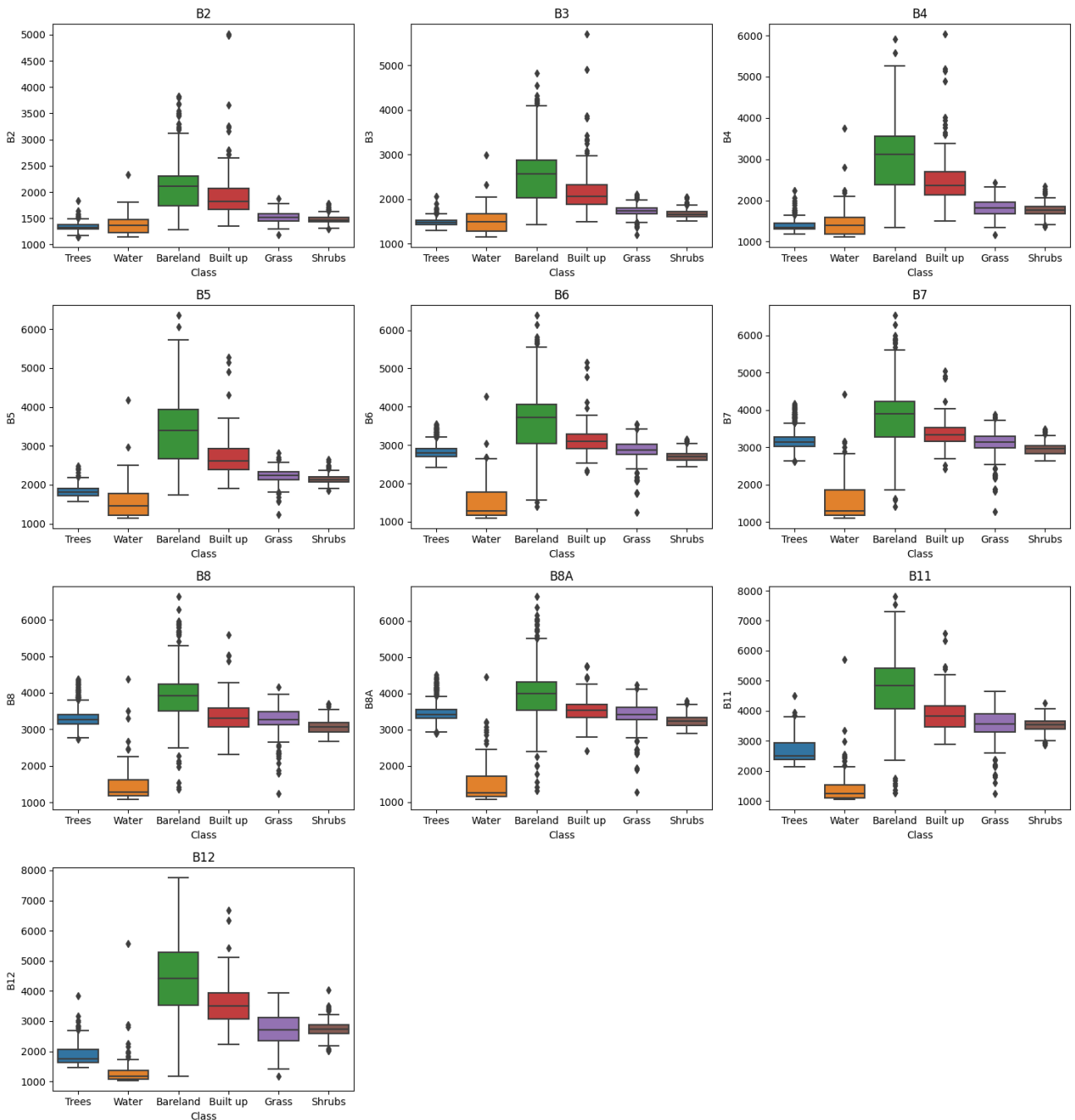


Figure 4: Box plots showing the distribution of Sentinel-2 spectral band values (B2, B3, B4, B5, B6, B7, B8, B8A, B11, and B12) across different land cover classes: trees, water, bareland, built-up areas, grass, and shrubs.

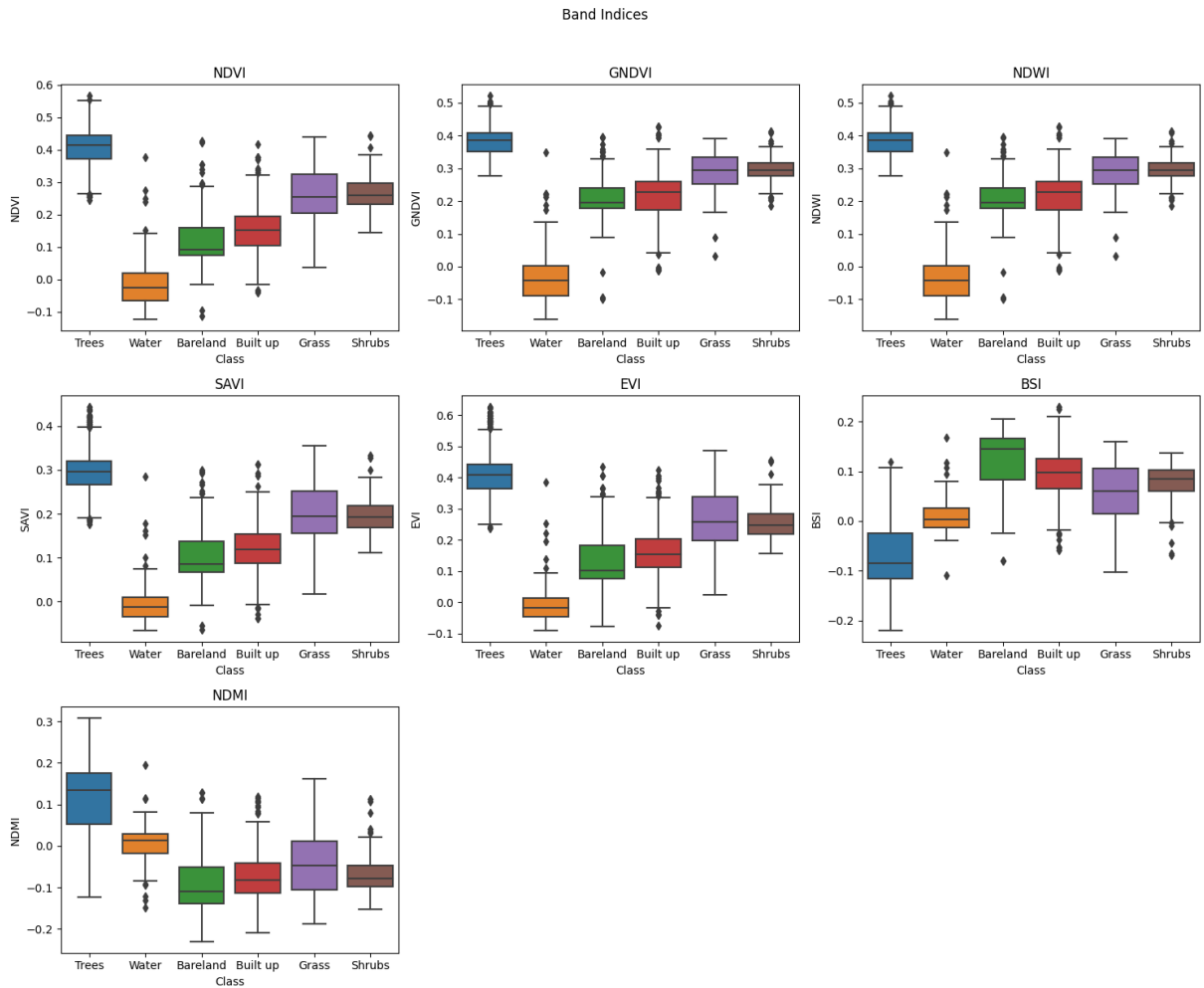


Figure 5: Box plots showing the distribution of spectral index values (NDVI, GNDVI, NDWI, SAVI, EVI, BSI, and NDMI) from Sentinel-2 data across different land cover classes: trees, water, bareland, built-up areas, grass, and shrubs.

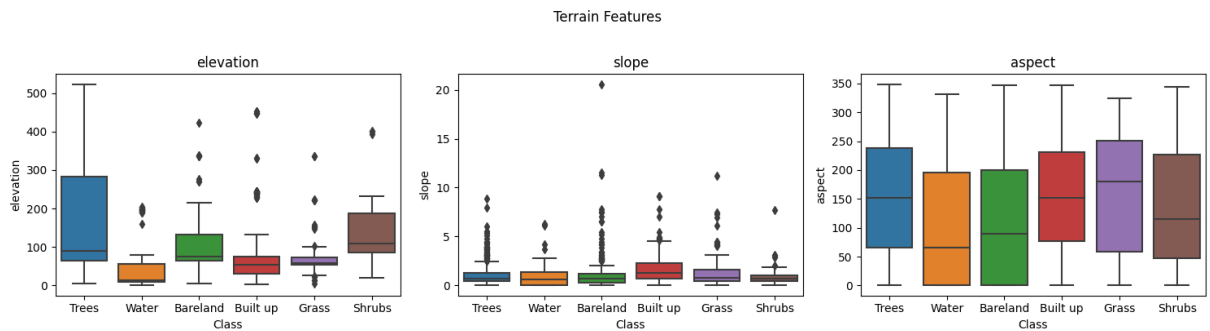


Figure 6: Box plots showing the distribution of terrain feature values (elevation, slope, and aspect) across different land cover classes: trees, water, bareland, built-up areas, grass, and shrubs.

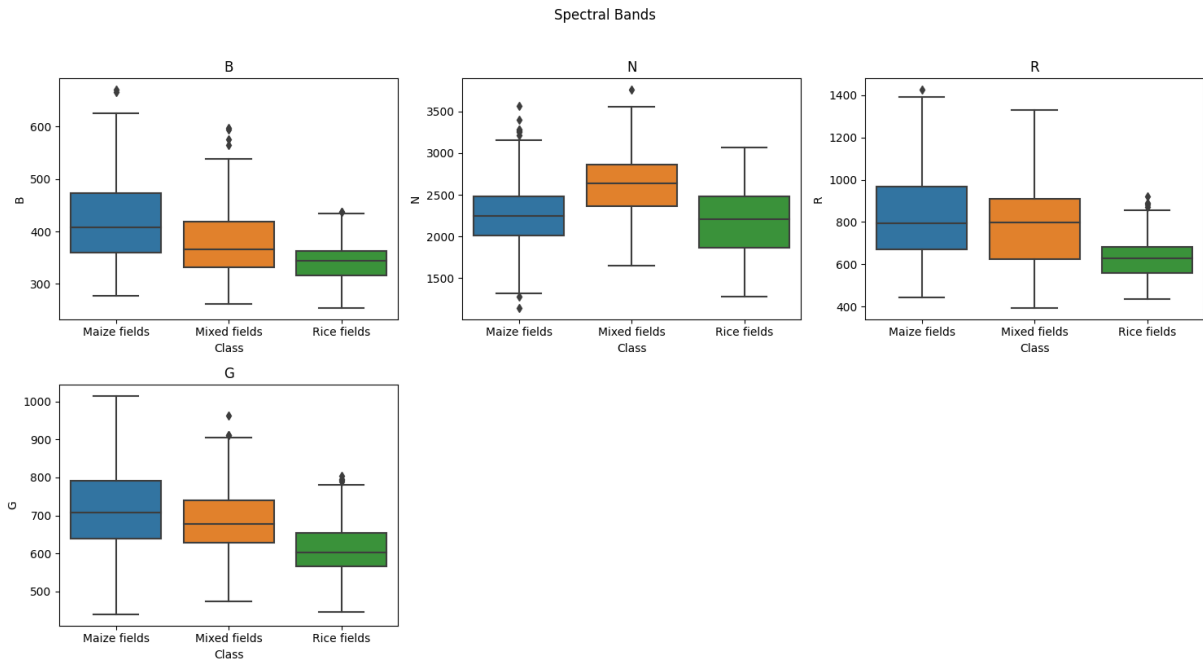


Figure 7: Box plots showing the distribution of PlanetScope spectral band values (Blue, Green, Red, and Near-Infrared (NIR)) across maize fields, mixed crop fields, and rice fields.

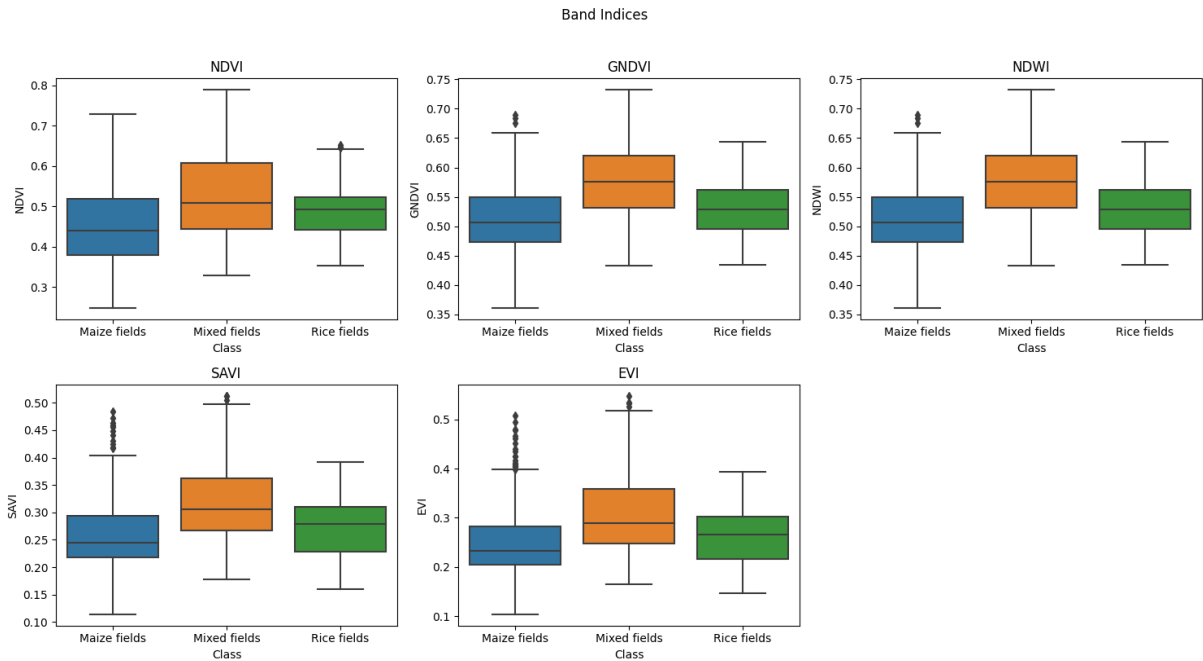


Figure 8: Box plots showing the distribution of spectral index values (NDVI, GNDVI, NDWI, SAVI, and EVI) from PlanetScope data across maize fields, mixed crop fields, and rice fields.

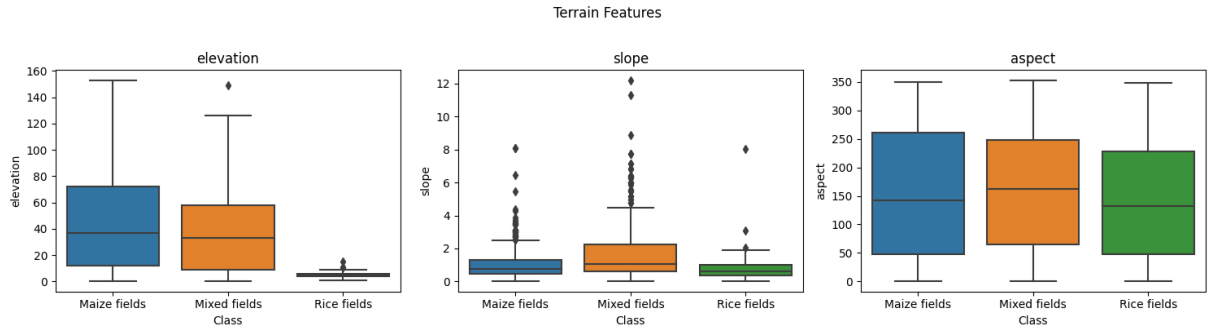


Figure 9: Box plots showing the distribution of terrain feature values (elevation, slope, and aspect) across maize fields, mixed crop fields, and rice fields using PlanetScope data.

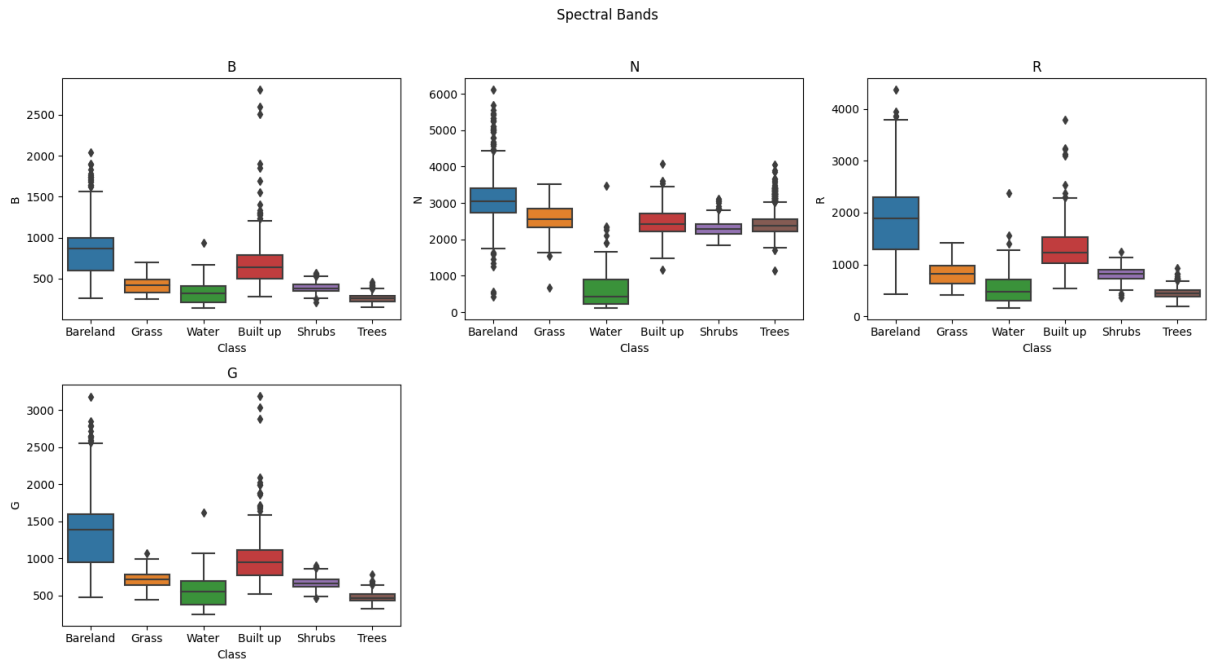


Figure 10: Box plots showing the distribution of PlanetScope spectral band values (Blue, Green, Red, and Near-Infrared (NIR)) across different land cover classes: bareland, grass, water, built-up areas, shrubs, and trees.

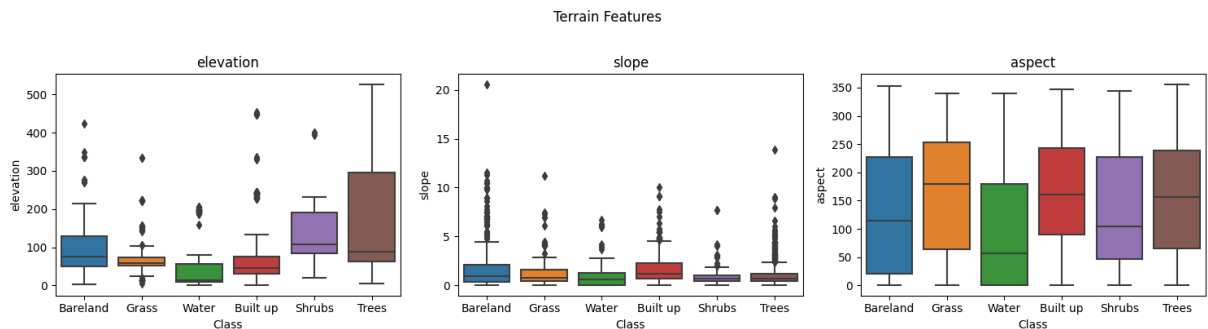


Figure 11: Box plots showing the distribution of terrain feature values (elevation, slope, and aspect) across different land cover classes: bareland, grass, water, built-up areas, shrubs, and trees.

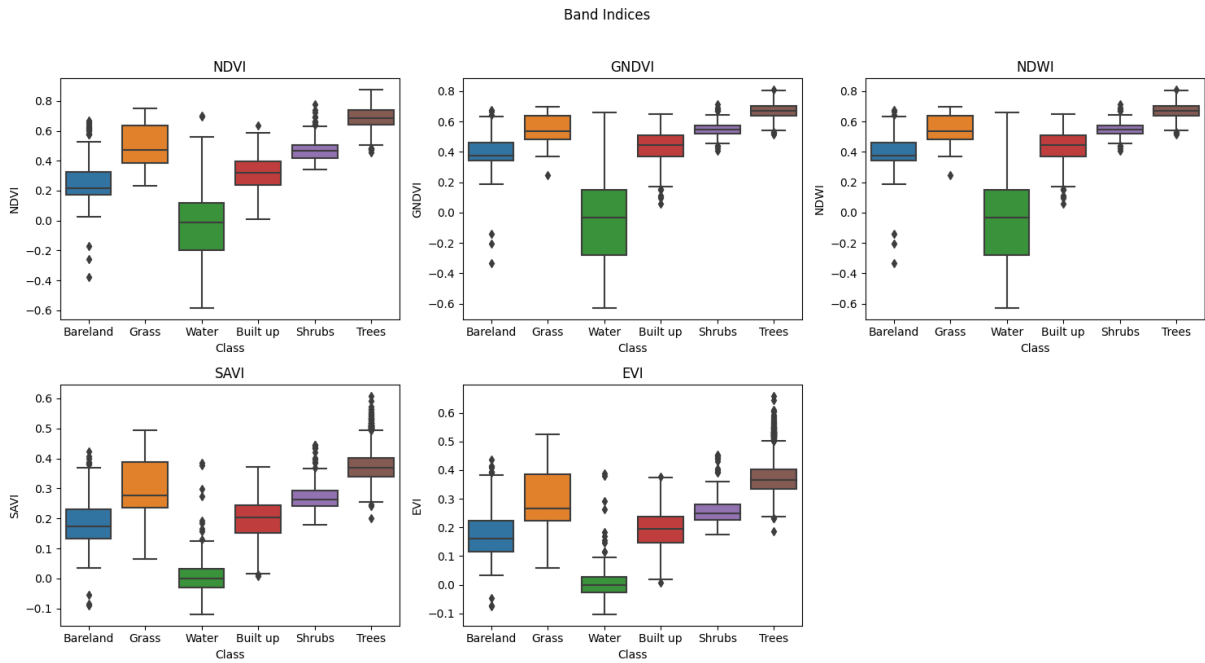


Figure 12: Box plots showing the distribution of spectral index values (NDVI, GNDVI, NDWI, SAVI, and EVI) from PlanetScope data across different land cover classes: bareland, grass, water, built-up areas, shrubs, and trees.

Confusion matrix

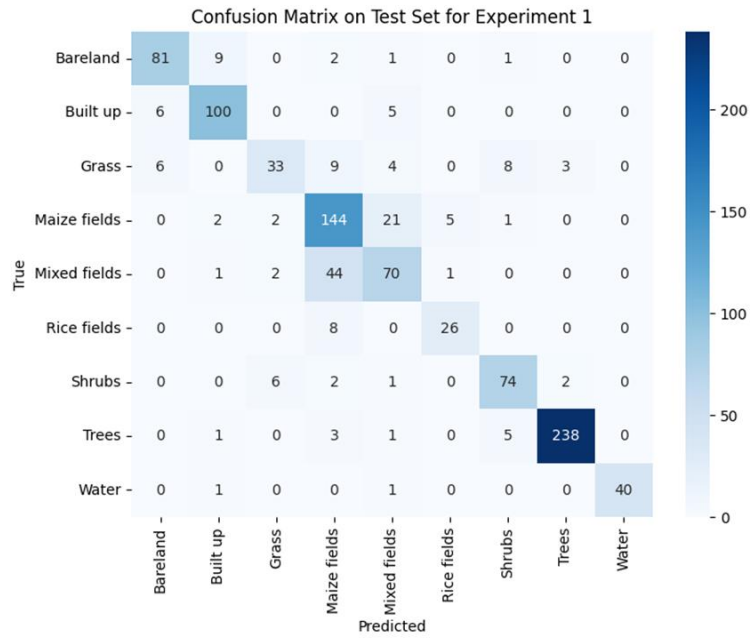
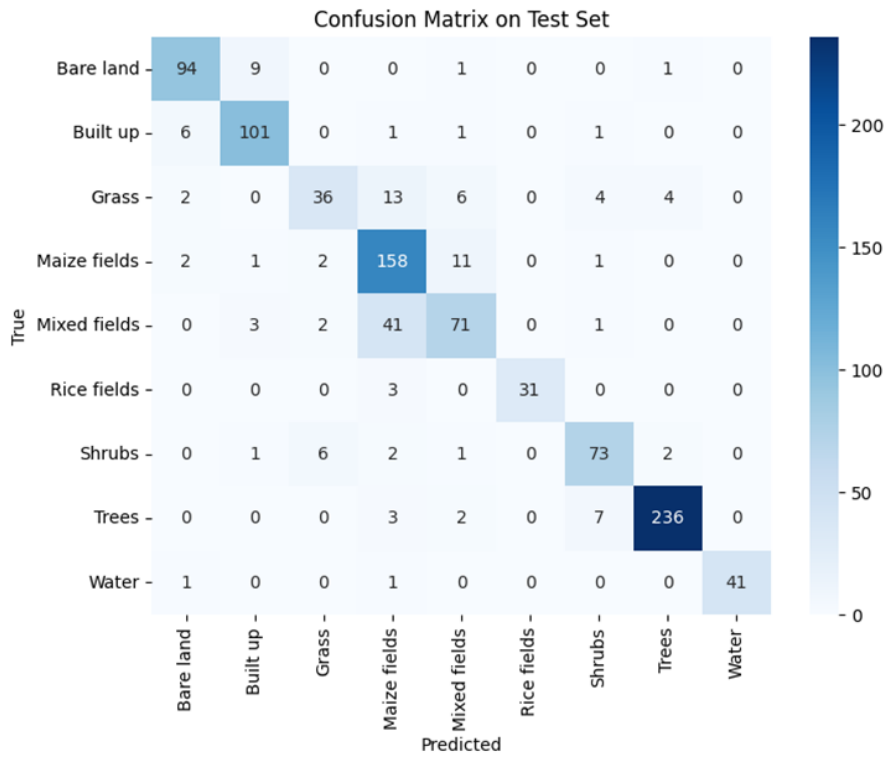


Figure 13: Confusion Matrix for Sentinel-2 best features map



14: Confusion Matrix for PlanetScope multi-stack map

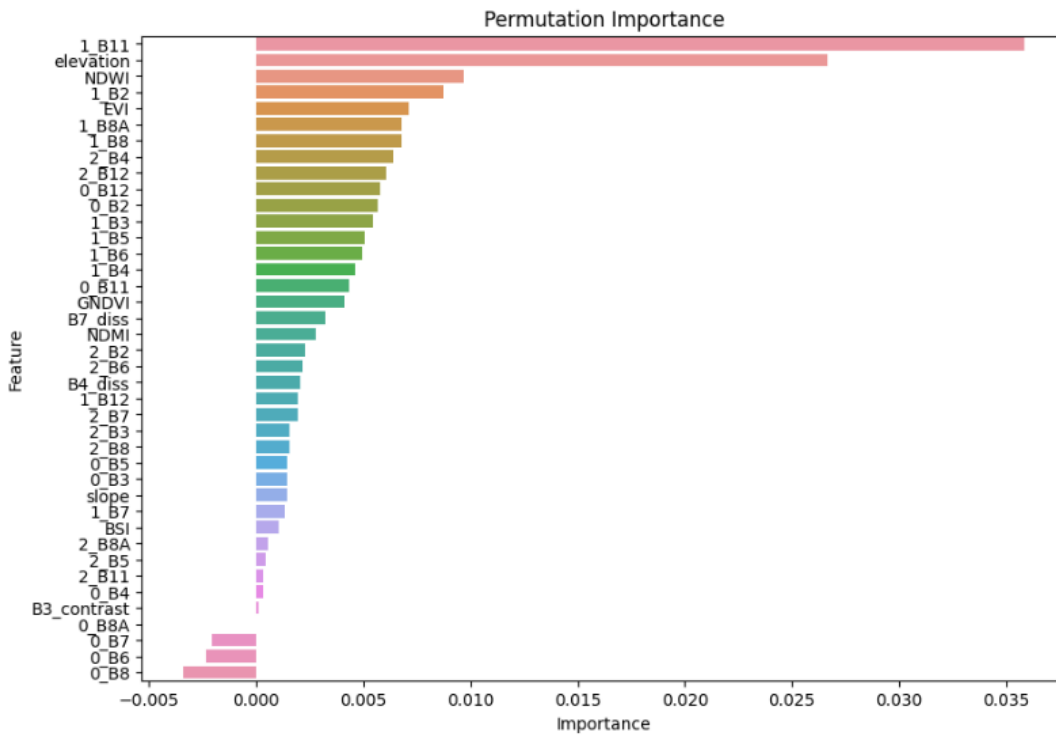


Figure15: Bar chart showing feature importance using permutation importance Method on Sentinel-2 multi-stack composite data.

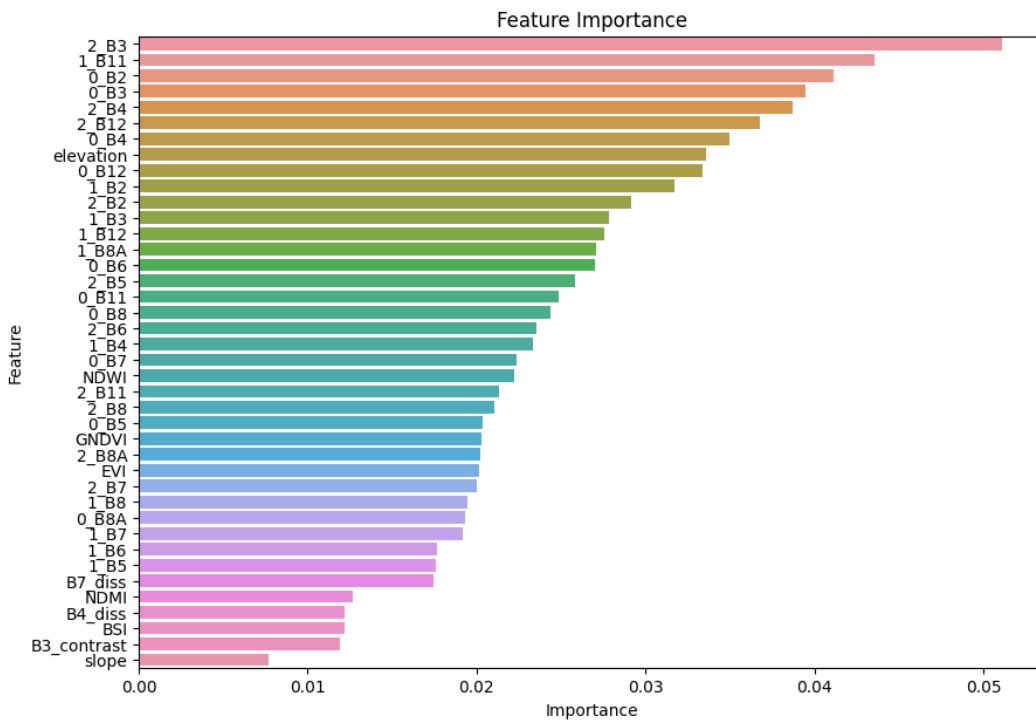


Figure16: Bar chart showing feature importance using MDG Method on Sentinel-2 multi-stack composite data.

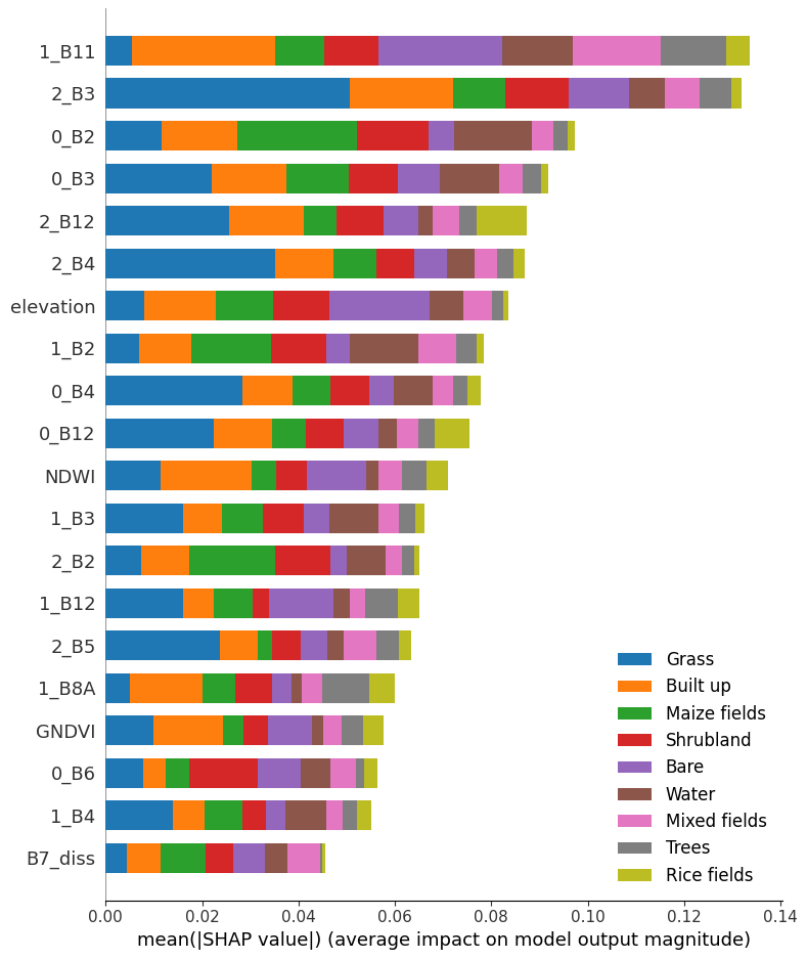


Figure 16: Bar chart showing feature importance using SHAP Method on Sentinel-2 multi-stack composite data.

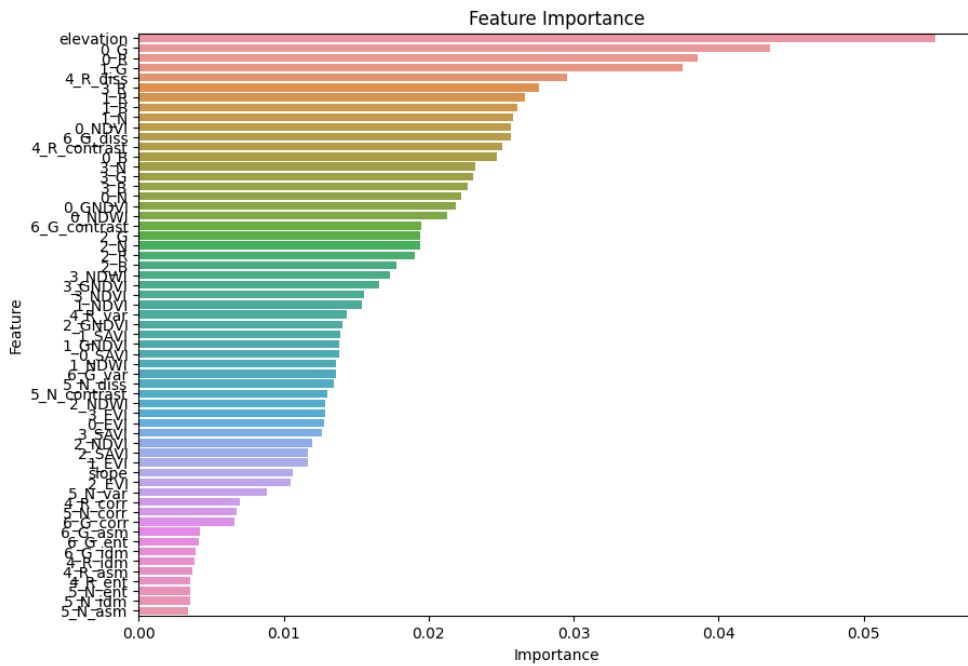


Figure 17: Bar chart showing feature importance using MDG Method on PlanetScope multi-stack composite data.

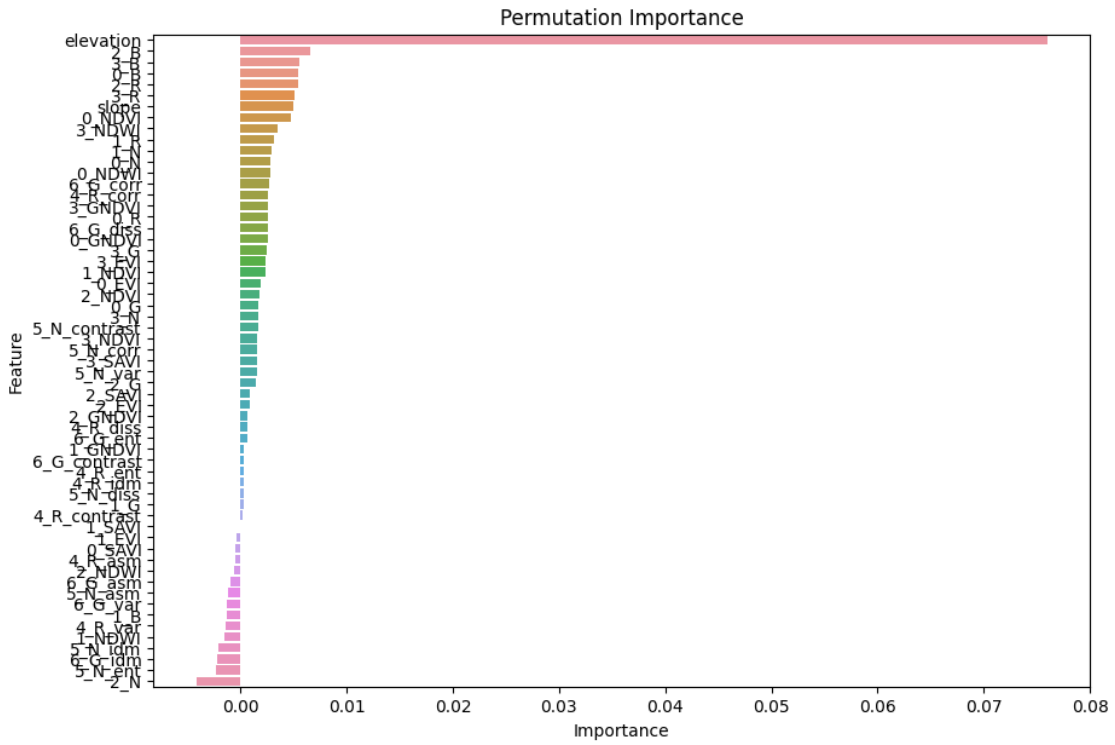


Figure 18: Bar chart showing feature importance using Permutation Method on PlanetScope multi-stack composite data.

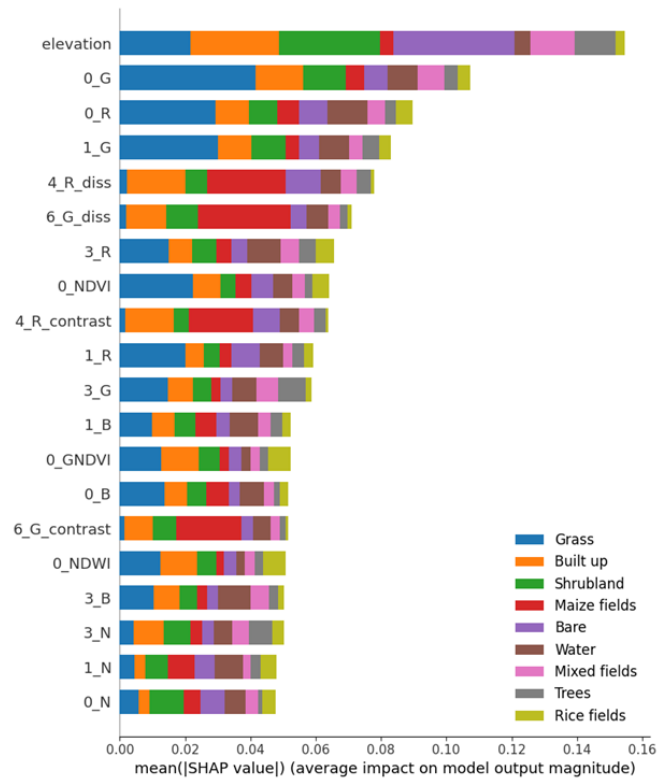


Figure 19: Bar chart showing feature importance using SHAP Method on PlanetScope multi-stack composite data.

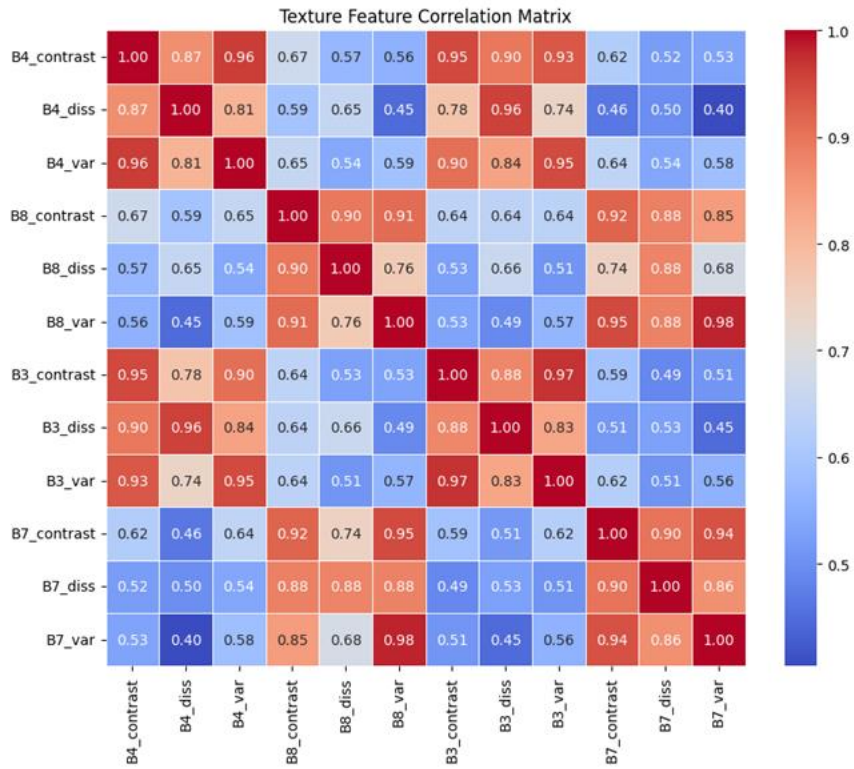


Figure 20: Pearson correlation matrix of GLCM features computed from Sentinel-2 image.

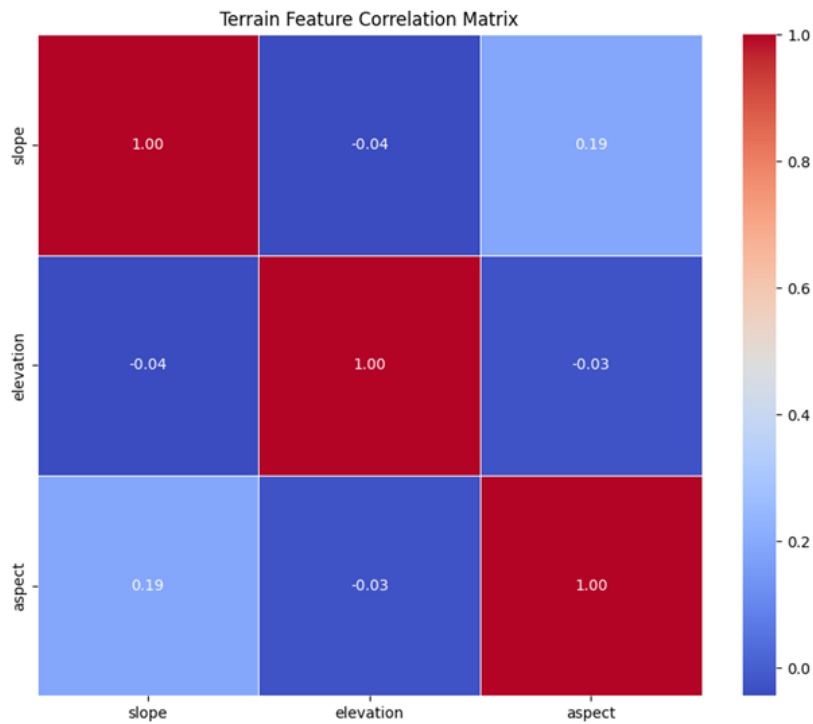


Figure 21: Pearson correlation matrix for topographic features for Sentinel-2

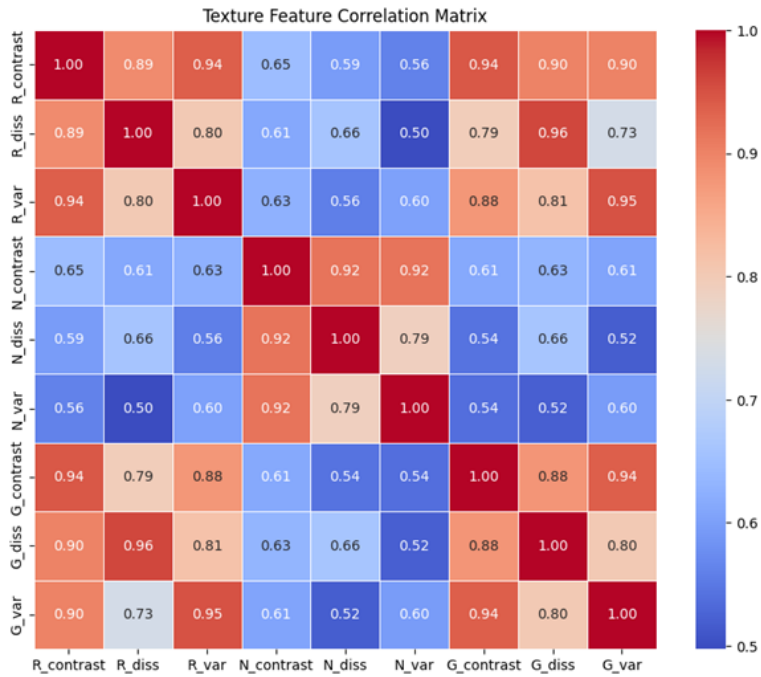


Figure 22: Pearson correlation matrix of GLCM features computed from PlanetScope image.

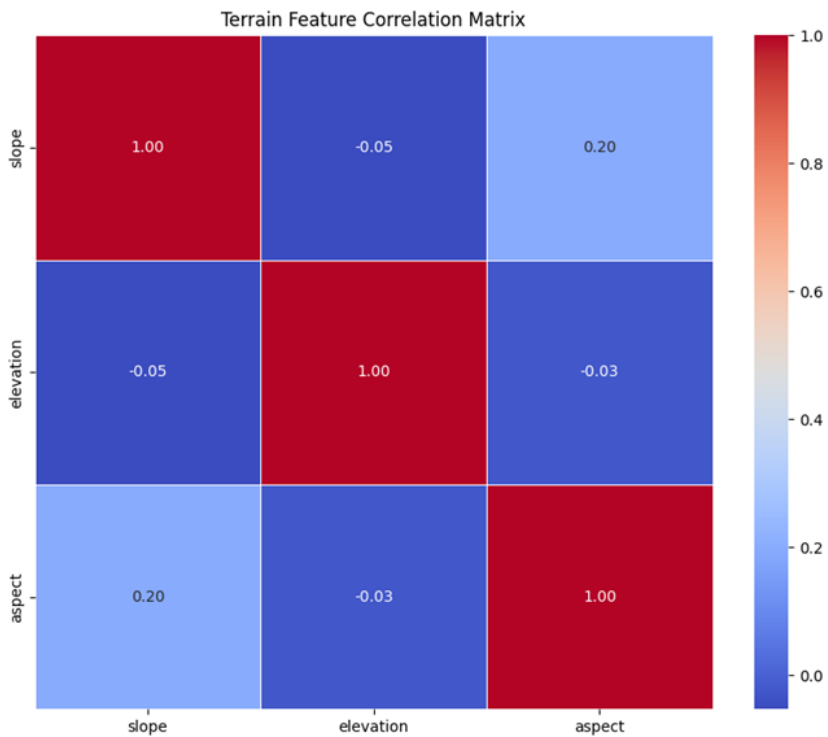


Figure 23: Pearson correlation matrix for topographic features for PlanetScope image

