UCLouvain

AGRO LOUVAIN
BIOINGÉNIEURS

Faculté des bioingénieurs

Exploring Sentinel-2 satellites for crop-type mapping: use case in Kolda, Senegal, using spectral, temporal augmented and non-augmented data

Auteur : Guillaume Jadot
Promoteurs : Pierre Defourny
Sophie Bontemps
Lecteurs : Pierre Houdmont
Sébastien Lambot

Année Académique : 2023 - 2024

Mémoire de fin d'études présenté en vue de l'obtention du diplôme de Bioingénieur : Sciences et technologies de l'environnement

Acknowledgments

I would like to express my sincere gratitude to my supervisor, Professor Pierre Defourny, and my co-supervisor, Sophie Bontemps, for their invaluable guidance and support throughout the development of this master's thesis.

I extend my heartfelt thanks to Boris Norgaard for his collaboration and the numerous productive discussions we shared, which significantly contributed to the progress of this work.

I am particularly grateful to the geomatic team of the ELIE institute for organizing the Friday meetings. These sessions provided a platform for knowledge exchange and fostered a collaborative research environment that greatly enriched my experience.

My appreciation also goes to my family for their unwavering support and encouragement throughout this academic journey.

Finally, I would like to thank the readers, Mr. Pierre Houdmont and Mr. Sebastien Lambot, for taking the time to review this work. I hope you find reading this master's thesis as engaging and insightful as I found the process of creating it.

Contents

Acknowledgments i

List of Figures iv

List of Tables v

List of Acronyms vii

Chapter 1: Introduction 1

Chapter 2 Literature review 2
2.1 Agriculture in land cover typology . . . 2
2.2 Remote sensing data for crop-type mapping . . . 3
2.2.1 Copernicus program . . . 4
2.2.4 Complementary use of Sentinel-1 and Sentinel-2 data . . . 8
2.2.5 Handling missing values . . . 9
2.2.6 Spectral indices time series . . . 10
2.3 Crops classification . . . 12
2.3.1 Supervised classification . . . 13
2.3.2 Unsupervised classification . . . 16
2.3.3 Hughes phenomenon . . . 17
2.3.4 Data distribution and resampling methods . . . 18
2.3.5 Assessing accuracy - Performance metrics . . . 20
2.4 Crop classification in smallholder West African agriculture: challenges and existing approaches . . . 21
2.5 Literature review conclusion . . . 23

Chapter 3: objectives 24
3.1 General Context . . . 24
3.2 Objectives . . . 24

Chapter 4: materials and method 25
4.1 Region of interest . . . 25
4.2 In-situ data preparation . . . 26
4.2.1 Field campaign for in-situ data collection . . . 26
4.2.2 Exploratory analysis . . . 28
4.2.3 Non-crop classes . . . 29
4.2.4 Calibration and validation datasets split . . . 30
4.2.5 Dealing with unbalanced classes . . . 32
4.3 Sentinel-2 data pre-processing . . . 33

4.3.1 Clouds masking, time-series, temporal resampling and gap-filling . 34
4.3.2 Mosaic 35
4.4 Spectral indices 35
4.5 Data transformation and feature extraction 36
4.6 Algorithm training, model and SMOTE algorithms comparison 37
4.6.1 Model and scenarios 37
4.6.2 Training with Sen4Sats cropland mask 38
4.6.3 Resampling with SMOTE algorithms 39
4.7 Spectral domain 43
4.7.1 Spectral signature of each crop class 43
4.7.2 Classification without time-series 44
4.8 Results evaluation 45
4.9 Scale-up and strategy for prediction on the entire ROI 45

Chapter 5 : results **46**
5.1 Time-series classification results 46
5.1.1 Classification with SMOTE algorithm 46
5.1.2 Classification with ADASYN algorithm 48
5.1.3 Classification with SMOTEENN algorithm 49
5.1.4 Prediction over the entire ROI 51
5.1.5 Classification with Sen4Stats cropland mask 54
5.2 Qualitative assessment of class separability in the spectral domain 56
5.2.1 Spectral signature of each crop classes 56
5.2.2 Classification without time-series 63
5.3 Qualitative comparison between the classification results 67

Chapter 6 : Discussion and conclusion **69**

Bibliography **72**

Appendix A: Confusion matrices for all scenarios tested **77**

List of Figures

1 Sentinel 2 MSI spectral bands vs spatial resolution (Gascon et al. (2017)) 5
2 Illustration of ascending and descending satellite orbits ((Sentinel, 2023)) 6
3 Schematic sketch of the three main scattering types considered for SAR data (Flores-Anderson et al., 2019) 8
4 Actual experimental classification results showing the Hughes phenomenon—the loss of classification accuracy as the dimensionality of the spectral domain increases (Richards, 2013) 18
5 Overview of different SMOTE algorithms results (SATPATHY, 2020) . 20
6 ROI location: Kolda region in Senegal under 3 Sentinel 2 tiles 26
7 Area visited during the field campaign in October 28
8 Initial distribution of crop polygons for each class in the Kolda region . 29
9 Count of crop polygons and number of pixels for each class after pre-processing 30
10 Overview of the preparation and splitting of the in-situ data 31
11 Distribution of data samples after SMOTE application 33
12 Overview of the pre-processing steps applied to the Sentinel-2 images . 35
13 NDVI mean curves for crop classes 37
14 Effect of different SMOTE algorithms on the spectral dataset for crops classes 41
15 Effect of different SMOTE algorithms on the metrics dataset for crop classes 42
16 PCA on the spectral dataset for non-crop classes 43
17 PCA on the metrics dataset for non-crop classes 43
18 S2 tile PEV cloud-free images for different dates across the growing season 45
19 Prediction over the entire ROI with the model trained with scenario 9 with non-crop classes masked 53
20 Prediction over the S2 PEV tile with cropland mask, spectral bands and metrics model 55
21 Spectral signatures curves with 95percentile values and standard deviation for crops classes 56
22 Spectral signatures probability density function for crop classes 58
23 Spectral values histogram for each spectral band and each crop class . . 59
24 Spectral values histogram for each spectral band and each non-crop class 60
25 Crop plots visited during the field campaign in October 2023 62
26 Prediction over the PEV tile with the model trained with 3 cloud-free images with spectral bands and metrics (non-crop classes masked) . . . 66
27 Comparison between the four classification models 67
28 Comparison between classification with and without cropland mask . . 68

List of Tables

1 Factors influencing vegetation backscatter (Ulaby, 1975) 8
2 Performance metrics : confusion matric 20
3 Confusion matrix for Random Forest classifier on the Burkina Faso test site (Inglada et al., 2015) 23
4 XGBoost performance for different feature combinations and sampling techniques 46
5 Performance comparison of different feature combinations with SMOTE 46
6 Confusion matrix for SC3_spectral_metrics_smote_strategy 47
7 Performance comparison of different feature combinations with SMOTE across various classes 48
8 Accuracy comparison for crop and non-crop classes across different SMOTE models 48
9 Performance comparison of different feature combinations with ADASYN 48
10 Performance comparison of different feature combinations with ADASYN across various classes 49
11 Accuracy comparison for crop and non-crop classes across different ADASYN models 49
12 Performance comparison of different feature combinations with SMOTEENN 50
13 Confusion matrix for SC9_spectral_metrics_smoteENN_strategy 50
14 Performance comparison of different feature combinations with SMOTEENN across various classes 51
15 Accuracy comparison for crop and non-crop classes across different SMOTEENN models 51
16 Performance comparison for crop classes with cropland mask 54
17 Classification performance metrics by class for classification with cropland mask 55
18 Performance for the classification on cloud-free images 63
19 Confusion matrix for the classification on few clouds free images 64
20 Performance metrics by class for the classification on cloud-free images 64
21 Accuracy comparison for crop and non-crop classes for the classification on cloud-free images 65
22 Confusion matrix for SC1_spectral_smote_strategy (Accuracy: 68.12%) 77
23 Confusion matrix for SC2_metrics_smote_strategy (Accuracy: 69.15%) 77
24 Confusion matrix for SC3_spectral_metrics_smote_strategy (Accuracy: 71.18%) 77
25 Confusion matrix for SC4_spectral_ADASYN_strategy (Accuracy: 65.96%) 78
26 Confusion matrix for SC5_metrics_ADASYN_strategy (Accuracy: 68.85%) 78

27 Confusion matrix for SC6_spectral_metrics_ADASYN_strategy (Accuracy: 70.00%) 78
28 Confusion matrix for SC7_spectral_smoteENN_strategy (Accuracy: 67.41%) 78
29 Confusion matrix for SC8_metrics_smoteENN_strategy (Accuracy: 69.00%) 79
30 Confusion matrix for SC9_spectral_metrics_smoteENN_strategy (Accuracy: 73.05%) 79
31 Confusion matrix for the classification on few clouds free images 79

List of Acronyms

UCL Université catholique de Louvain

LCML Land Cover Meta Language

JECAM Joint Experiment for Crop Assessment and Monitoring

MMU Minimum Mapping Unit

EU European Union

S2 Sentinel-2

SWIR Short-Wave Infrared

S1 Sentinel-1

SAR Synthetic Aperture Radar

H Horizontal polarization

V Vertical polarization

TOA Top of Atmosphere

BOA Bottom of Atmosphere

SC Scene Classification

LULC Land Use and Land Cover

NDVI Normalized Difference Vegetation Index

NIR Near-Infrared

NDWI Normalized Difference Water Index

EVI Enhanced Vegetation Index

BSI Bare Soil Index

REP Red Edge Position

NDVIre2 Normalized Difference Vegetation Index Red Edge 2

GNDVI Green Normalized Difference Vegetation Index

LAI Leaf Area Index

Fcover Fraction of Vegetation Cover

VIs Vegetation Indices

RF Random Forest

SVM Support Vector Machine

MLC Maximum Likelihood Classifier

RBF Radial Basis Function

GBT Gradient Boosting Trees

XGBoost Extreme Gradient Boosting

CNNs Convolutional Neural Networks

RNNs Recurrent Neural Networks

LSTM Long Short-Term Memory

FKM Fuzzy K-Means

SMOTE Synthetic Minority Over-sampling Technique

ADASYN Adaptive Synthetic

ENN Edited Nearest Neighbors

OA Overall Accuracy

DAPSA Direction de l'Analyse, de la Prévision et des Statistiques Agricoles

ROI Region of Interest

Chapter 1: Introduction

Smallholder agriculture is the foundation of food production in Africa. About 90% of the food produced in developing countries comes from this agricultural system. In 2016, there were 51 million farms in Africa, 80% of which were less than 2 hectares in size (Lowder et al., 2016). Yet, the number of farms is still increasing while combined with a drop in farm size (Eastwood et al., 2010).

Meanwhile, Africa has become the center of the world's demographic growth, and its population could reach 26% of the world's population by 2050 (Abramova, 2022). Sub-Saharan Africa is considered to be an area of great concern, as it is particularly vulnerable to food shortages.

The agricultural sector stands as a cornerstone of Senegal's economy, contributing approximately 15 percent to the GDP in 2018. It also underpins the livelihoods of 60 percent of the rural populace, who rely on it directly (FAO AgriSurvey, 2023).

Despite agriculture's importance, Senegal grapples with challenges inherent to the Sahel region, characterized by sporadic rainfall and sub-optimal soil conditions. As a result, Senegal relies on imports to meet approximately 70 percent of its food necessities.

Given the strategic nature of the agricultural sector for the socio-economic development of the country and for the food security of the population, it is essential that policymakers can rely on quality data to make evidence-based decisions.

Given these circumstances, there is a pressing need for systems that monitor and map crops. Such systems serve dual purposes: they act as crucial decision-making aids for small-scale farmers and agricultural workers, while simultaneously providing comprehensive assessments of global land-use changes. These tools could play a role in efforts to enhance food security across developing nations. Agricultural landscape mapping is needed for conducting yield forecasts and facilitating strategic long-term planning, which are essential steps towards bolstering agricultural resilience.

Creating a map based on agricultural land plots can unlock multiple applications, such as enhancing agricultural statistics, refining sampling designs for yield predictions, cadastral functions, and supporting crop insurance programs, to name a few. Nevertheless, procuring such a map poses substantial challenges in regions dominated by smallholder farms due to a number of factors: small field size, the prevalence of mixed cropping, diverse management techniques, a fragmented landscape, and the common occurrence of trees interspersed among the crops.

In addressing these challenges, remote sensing data has proven to be an invaluable resource. The field employs a wide range of sensors, each offering distinct capabilities in terms of image detail and data capture rates. This variety has facilitated the broad

adoption of remote sensing technologies. Moreover, the rapid progress in computational power, coupled with the development of advanced machine learning techniques, has significantly boosted the efficiency and speed at which remote sensing algorithms can process and analyze this wealth of data. However, it has been noted that the accuracy of the crop type maps created with those classifications techniques could fluctuate significantly across different regions due to several factors such as different cloud cover, frequency of image acquisition, etc.

Therefore, this master's thesis aims to develop and compare two classification approaches for accurately distinguishing various crop types in Senegal's smallholder agricultural context using optical remote sensing imagery. The study addresses the challenges of similar crop cycles, frequent cloud cover, high intra-class variability, presence of adventitious plants and class imbalance. It compares a data-intensive time series approach, where the classifier is fed with extensive temporal data, against a selective approach using only a few cloud-free images. This comparison seeks to determine the most effective method for crop-type mapping in these challenging conditions.

Chapter 2 Literature review

This chapter examines the current landscape of scientific research in the field of crop mapping using remote sensing technologies, with a particular emphasis on the classification of different crop types. It aims to provide a comprehensive overview of the latest advancements and methodologies in this domain.

This review is conducted with several objectives. Firstly, current limitations and unresolved challenges in remote sensing-based crop-type mapping are identified and assessed. Through a thorough examination of existing research, areas where knowledge is insufficient or problems remain unaddressed are highlighted. Secondly, the development of the methodology for this study is informed by the analysis of prior work, providing justification for the selected methods and hypotheses. Lastly, this literature review establishes a framework against which the results obtained in this thesis can be compared with findings previously reported in the scientific literature.

2.1 Agriculture in land cover typology

Agricultural mapping, while seemingly straightforward as a category of land cover, is frequently a source of confusion and discrepancy among existing land cover maps. This is particularly true when delineating between cropland and non-cropland areas. The challenge is compounded by the considerable variation in agricultural land types around the globe.

The World Program for the Census of Agriculture 2020 (FAO World Census of Agriculture, 2023) proposes the following definitions, obtained by aggregating the Land Cover Meta Language (LCML) classes [1]:

The term "arable land" refers to plots typically utilized for cultivating temporary crops. "Cropland" encompasses both arable land and areas dedicated to permanent crops. Extending this categorization, "agricultural land" combines cropland with permanent meadows and pastures. Lastly, the broader category of "land used for agriculture" includes all agricultural land plus areas occupied by farm buildings and yards.

The chosen land cover classification system must align with the data source. For satellite remote sensing, the Joint Experiment for Crop Assessment and Monitoring network (JECAM) endorsed a definition for annual cropland due to the annual nature of the Earth Observation time series: “the annual cropland from a remote sensing perspective is a piece of land of minimum 0.25 ha (min. width of 30 m) that is sowed/planted and harvestable at least once within the 12 months after the sowing/planting date. The annual cropland produces an herbaceous cover and is sometimes combined with some tree or woody vegetation."

The definition endorsed by JECAM incorporates the Minimum Mapping Unit (MMU), delineating the smallest area of land to be identified in mapping exercises. As an illustration, the mapping protocol for the European Union’s CORINE Land Cover Database initially established the MMU threshold at 25 hectares, thereby excluding landscape elements below this size. This criterion could result in the omission of smaller agricultural plots, potentially introducing considerable bias into the final agricultural land map.

2.2 Remote sensing data for crop-type mapping

For decades, satellite remote sensing was promoted as a key data source for operational agriculture monitoring. Landsat satellite series permitted developing operational crop type mapping on some specific agricultural landscape such as in the United States and Canada, characterized by large fields. However its 16 days revisit cycle was a problem for near real time crop type mapping in areas with persistent clouds cover due to the limited number of valid observations. Furthemore, most cropping system around the world contains fields that are much smaller than those in the US. The 30-m spatial resolution limits resolving the individuals fields in those agricultural landscapes, preventing any parcel-based application. To resolve those issues, commercial satellite constellations such as RapidEye have emerged. It allows monitoring agricultural landscape with a spatial resolution higher than 5-m. However, their small scene footprint and rather poor

[1] the LCML framework proposed by FAO has been adopted as an international standard by the Technical Committee of the International Organization for Standardization (ISO)

radiometric and atmospheric correction were identified as significant issues preventing consistent coverage over large areas in addition to the cost constraint.

2.2.1 Copernicus program

In order to deal with those challenges, Europe lunched the Copernicus program. It is developed by the European Space Agency and funded by the European Union (EU). It provides information services based both on Earth observation through satellites and on in-situ (non-space) data collection. These services are available for free and assist service providers, public authorities, and other international organizations in enhancing the quality of life for European citizens.

2.2.1.1 Sentinel 2

Building on the Landsat and SPOT missions' legacy, the Sentinel-2 (S2) satellite constellation and its multi spectral instrument were designed in the framework of the European Copernicus program for land surface and agriculture monitoring to measure the reflected solar spectral reflectance in 13 bands ranging from the visible to the Short-Wave Infrared (SWIR) bands as detailed in figure 1. The spectral bands include three narrow bands for cloud screening and atmospheric correction at 60-m, three red-edge bands and two SWIR bands at 20-m providing key information about vegetation, as well as the blue, green, red and near infrared bands at 10-m. Since march 2017, it ensures a 5-days revisit time above almost all landmasses (Gascon et al., 2017).

Optical remote sensing, such as the one offered by S2, is largely and successfully used in agricultural mapping and monitoring. Indeed, the red and near-infrared (NIR) reflectance is strongly related to green biomass. In this process, satellite images are acquired over time, with each pixel representing a specific area on the ground. Temporal profiles are then extracted from these image time series, tracking changes in spectral reflectance values throughout the growing season. These profiles, which capture distinct patterns for different crop types due to variations in growth cycles and canopy structures, form the basis for crop type classification. Various machine learning algorithms, such as Random Forests or Neural Networks, can be applied to these temporal profiles to distinguish between different crop types, resulting in spatial maps of crop distribution.

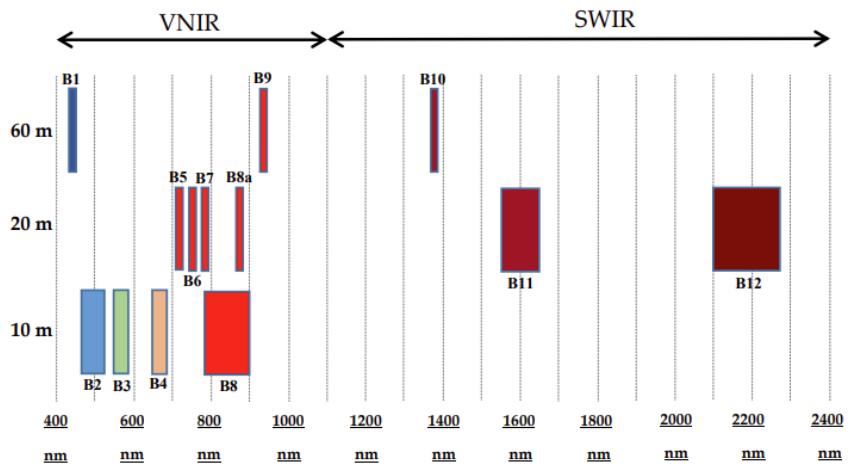


Figure 1: Sentinel 2 MSI spectral bands vs spatial resolution (Gascon et al. (2017))

Pre-processing levels

Image pre-processing is a common way to improve classification accuracy. Radiometric correction is a very commonly used pre-processing technique. Atmospheric correction, as detailed in studies by Chavez et al. (1996), Fukushima et al. (1998), Huang et al. (2008), Song et al. (2001), is a form of radiometric correction aimed at minimizing atmospheric influences on observed radiance. Furthermore, the use of spatial low-pass or smoothing filters serves as an additional strategy to reduce intra-class variability and random noise.

The Level-1C product delivers orthorectified Top-Of-Atmosphere (TOA) reflectance data in multiple spectral bands. Additionally, it includes cloud and land or water masks for detailed environmental analysis.

In addition, the Level-2A product offers orthorectified Bottom-Of-Atmosphere (BOA) reflectance, also with sub-pixel multispectral precision. A feature of this product is its scene classification map, categorizing various geographical and atmospheric features such as clouds, cloud shadows, vegetation, soils/deserts, water bodies, and snow, enriching the data's contextual understanding.

The transformation from Level-1C to Level-2A products is facilitated by the operational processor, Sen2Cor. This processing is split into two parts: the scene classification (SC) algorithm that aims at providing a pixel classification map and atmospheric correction (S2AC) that focuses on converting TOA reflectance into BOA reflectance, ensuring the data reflects more accurately the conditions at the Earth's surface, devoid of atmospheric distortions.

2.2.1.2 Sentinel 1

Sentinel-1 (S1) satellites are part of Europe's Copernicus program. The S1 mission is a constellation of two satellites (A and B units). Sentinel-1A launched in 2014 and Sentinel-1B in 2016, fly in a near polar, sun-synchronized orbit at around 697 km altitude (Geudtner et al., 2014). Their orbit passes close to the poles and takes 12 days to return to the equator, with this period shortening at higher latitudes due to orbit convergence. The satellites operate in both ascending and descending modes (figure 2), which determine their rotation direction around Earth, affecting the observation angles.

The S1 mission features C-band imaging that operates in four unique modes, each offering varying resolutions (as fine as 5 meters) and coverage areas (up to 400 kilometers). It provides dual polarisation capability, very short revisit times and rapid product delivery (Flores-Anderson et al., 2019).

Synthetic Aperture Radar (SAR) operates at wavelengths unaffected by clouds or the absence of light, enabling data collection over any location regardless of the time of day or prevailing weather conditions when there is no heavy rains.

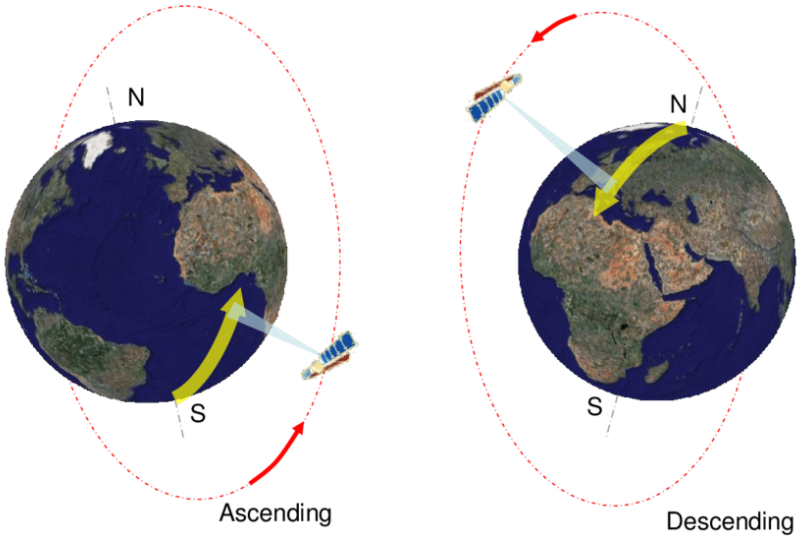


Figure 2: Illustration of ascending and descending satellite orbits ((Sentinel, 2023))

Synthetic aperture radar

In the realm of earth observation, SAR represents a paradigm shift in data collection. Unlike passive optical sensors that rely on existing light sources, SAR actively generates and measures its own energy reflections after they interact with the Earth's surface. This necessitates a unique analytical approach, as SAR data are sensitive to the dielectric

properties of the vegetation and the soil, to the plant structure (i.e., the size, shape, and orientation distributions of the scatterers), to the surface roughness, and to the canopy structure (e.g., row direction and spacing, and cover fraction) (Skriver et al., 1999) in contrast to the more straightforward interpretation akin to analyzing a photograph in optical imagery.

The spatial resolution of radar data hinges on an interplay: it is proportionate to the ratio of the sensor's wavelength to its antenna length. To put this into perspective, consider a satellite employing C-band radar with a wavelength of about 5 cm. Achieving a spatial resolution of 10 meters would require an antenna, approximately 4,250 meters in length, a size that's unfeasible for satellite deployment.

The concept of the synthetic aperture circumvents this limitation. By integrating a sequence of measurements from a shorter antenna, they effectively simulate the effects of a much larger antenna. This innovative approach yields higher resolution data without the need for unwieldy physical structures.

Radar sensors operate with wavelengths in the centimeter to meter range. This attribute bestows upon them unique capabilities. However, S1 is only equiped with a band-C instrument, which operates within the wavelength range of 4 to 8 centimeters. Specifically, S1 functions at a wavelength of 5.5 centimeters, corresponding to a frequency of 5.4 GHz. According to Gu et al. (2011), radars with shorter wavelengths, like the band-C instrument on S1, may undergo considerable attenuation during heavy precipitation. They remain nonetheless impervious to clouds, smoke, and haze, offering the capability for nearly uninterrupted day-and-night imaging in almost all weather conditions, as noted by (Mandal et al., 2019).

The concept of polarization plays a pivotal role. Polarization refers to the plane orientation in which an electromagnetic wave vibrates during transmission. SAR sensors usually employ linear polarization, denoted in two primary forms: horizontal (H) and vertical (V).

Radar sensors have the ability to control signal polarization during both transmission and reception. For instance, a signal transmitted in vertical (V) polarization and received in horizontal (H) is represented as VH. Conversely, a signal transmitted and received in horizontal polarization is denoted as HH. This precise control over polarization opens a window into understanding the surface structure being imaged, based on the signal strength of different polarization combinations (Flores-Anderson et al., 2019).

The backscattering coefficient σ^0 is governed by two basic target parameters: geometry (or roughness) and complex dielectric properties (humidity plays thus a role in the backscattering) (Ulaby, 1975).

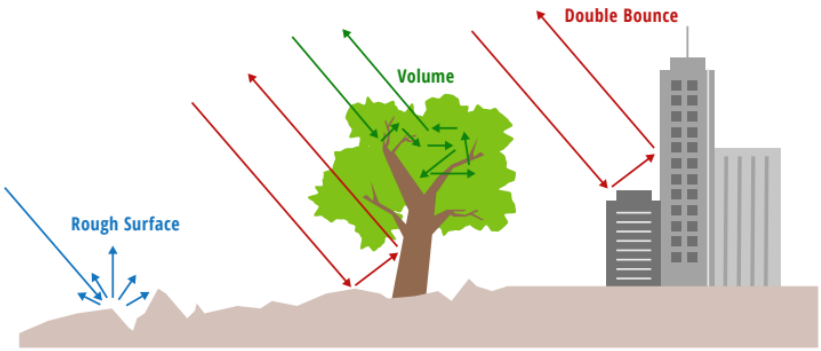


Figure 3: Schematic sketch of the three main scattering types considered for SAR data (Flores-Anderson et al., 2019)

Table 1: Factors influencing vegetation backscatter (Ulaby, 1975)

RADAR PARAMETERS	TARGET PARAMETERS
1.1 Frequency (or wavelength) 1.2 Polarization 1.3 Incidence Angle	2.1 Vegetation Cover 2.1.1 Geometry - Roughness (relative to the wavelength) - Density - Pattern (row direction) - Height 2.1.2 Dielectric Properties - Plant Moisture Content Distribution · Vertically (over the plant height) · Over Various Parts of the Plant (stem, leaves, fruit, etc.) 2.2 Underlying Soil 2.2.1 Surface Roughness (and subsurface if penetration is possible) 2.2.2 Dielectric Properties - Soil Type and Bulk Density Profile with Depth - Soil Moisture Content Profile with Depth

2.2.4 Complementary use of Sentinel-1 and Sentinel-2 data

The integration of S1 radar data with S2 optical imagery presents an opportunity to enhance land cover mapping and monitoring capabilities. The high revisit frequency of both S1 and S2 satellites allows for the creation of satellite image time series, which can effectively track Earth surface changes over time.

Optical sensors, such as those on S2, provide information about surface material properties, while radar sensors like S1 offer insights into the structural characteristics of landscape elements. The combination of these complementary data sources has been

shown to improve performance in various scenarios, including crop monitoring and LULC mapping (Steinhausen et al., 2018).

Several approaches have been developed to integrate radar and optical data for LULC mapping. Many of these methods employ standard machine learning techniques, such as Random Forest and Support Vector Machines, applied to a concatenation of radar and optical input data. While these approaches demonstrate improvements over single-sensor methods, they often treat the two data sources independently, potentially overlooking the interplay between them.

Furthermore, dealing with the different spatial and temporal resolutions of radar and optical data can be challenging. The addition of radar data increases the cost of computation and storage. If the addition of radar data does not significantly improve the classification accuracy, the extra computational burden may not be justified. Therefore, the inclusion of S1 data can help fill gaps in S2 time series, providing more continuous and reliable information for land cover analysis but the additional cost of computation, storage and for pre-processing should be taken into account.

2.2.5 Handling missing values

Optical satellite imagery in remote sensing faces challenges related to data continuity and quality. Atmospheric conditions, particularly cloud cover, can impact the availability and usability of these images. In regions with frequent cloud cover, constructing consistent time series becomes problematic due to data gaps. Additionally, sensor malfunctions or processing errors can contribute to missing or unreliable data points. Addressing these data inconsistencies requires careful consideration to avoid misinterpretation of the results. While linear interpolation between values can be a simple approach to fill gaps, it should be applied judiciously. This method is most appropriate when the time interval between missing data points is relatively short, or when there's minimal risk of significantly altering the time series characteristics.

To tackle the issue of missing data in satellite imagery time series, various algorithms and tools have been developed. One such tool is the ImageTimeSeriesGapFilling algorithm from the Orfeo ToolBox library. This algorithm employs techniques to fill gaps in image time series, taking into account the temporal context of the missing data. The ImageTimeSeriesGapFilling algorithm offers several methods for gap-filling, including local linear interpolation, spline interpolation, linear regression, and adaptive linear regression. Local linear interpolation estimates missing values based on the nearest valid pixels in the time series, while spline interpolation uses a smooth curve to estimate values between known data points. Linear regression fits a linear model to existing data to predict missing values, and adaptive linear regression extends this approach by adapting to local trends in the data.

When applying these gap-filling techniques, the nature of the data and the specific requirements of the analysis should be considered. In crop monitoring applications, for instance, knowledge of typical phenological cycles can inform the selection of an appropriate gap-filling method to ensure that the filled data accurately represents the expected vegetation dynamics. It's important to note that while these methods can improve the completeness of time series data, they introduce a level of uncertainty. Sensitivity analyses can be used to assess the impact of filled data on the results, helping to ensure the reliability of subsequent analyses and interpretations.

2.2.6 Spectral indices time series

According to Siachalou et al. (2017), accurately mapping crops presents several challenges. These include: 1) Limited spectral differentiation between different crops and significant high intraclass spectral variability, arising from analogous growth stages and growth calendars across different crops; 2) Spectral variability in the same crop type due to distinct agricultural practices, local weather conditions, and topography; and 3) A high prevalence of mixed pixels, particularly along extended field boundaries. The complexity of crop mapping is further compounded by the diverse nature of agricultural landscapes, which often feature a wide variety of crops and small-scale field plots. Vegetation indices-based methodologies are thus commonly employed in studies monitoring crop progress, as they yield valuable insights into crop coverage and growth status (Shen et al., 2013).

S2 satellites constellation is equipped with sensors that capture data across 13 distinct spectral bands, as outlined in Table 1. These spectral bands serve as the foundation for computing a range of biophysical variables and spectral indices. By utilizing the reflectance values from these bands, various indices are calculated, which then function as input features for classification algorithms. The following sections will examine several key indices that are commonly applied in crop classification tasks.

The Normalized Difference Vegetation Index (NDVI) is widely utilized to assess green vegetation by leveraging the reflectance in the near-infrared (NIR) and red spectral bands. NDVI serves as an effective indicator of plant vigor and physiological condition. The mathematical formulation of NDVI is presented in equation 1. The resulting index values are constrained within the range of -1 to 1.

$$NDVI = \frac{\rho_{NIR} - \rho_{Red}}{\rho_{NIR} + \rho_{Red}} \tag{1}$$

The Normalized Difference Water Index (NDWI) is designed to detect water content and the presence of water bodies within the observed region. It is calculated using the

NIR and short-wave infrared (SWIR) bands, producing values that range between -1 and 1.

$$NDWI = \frac{\rho_{NIR} - \rho_{SWIR}}{\rho_{NIR} + \rho_{SWIR}} \tag{2}$$

The Enhanced Vegetation Index (EVI) quantifies vegetation by measuring the difference between near-infrared (which vegetation strongly reflects) and red light (which vegetation absorbs). EVI is usually used for areas with high LAI where NDVI is expected to saturate. EVI is used to quantify vegetation greenness and is useful in understanding vegetation density and assessing changes in plant health.

$$EVI = 2.5 * \frac{\rho_{NIR} + \rho_{Red}}{\rho_{NIR} + 6 * \rho_{Red} - 7.5 * \rho_{Blue}} + 1 \tag{3}$$

The BareSoil Index (BSI) combines blue, red, near infrared and short wave infrared spectral bands to capture soil variations. These spectral bands are used in a normalized manner. The short wave infrared and the red spectral bands are used to quantify the soil mineral composition, while the blue and the near infrared spectral bands are used to enhance the presence of vegetation.

$$BSI = \frac{(\rho_{Red} + \rho_{SWIR}) - (\rho_{NIR} + \rho_{Blue})}{(\rho_{Red} + \rho_{SWIR}) + (\rho_{NIR} + \rho_{Blue})} \tag{4}$$

The Red-Edge Position (REP) is a sensitive indicator of chlorophyll content in vegetation. It represents the position of the maximum slope in the reflectance spectrum between the red and near-infrared regions. Guyot et al. (1988) proposed a linear four-point interpolation technique to calculate REP:

$$REP = 705 + 35 * \frac{0.5 * (\rho_{Red} + \rho_{Red-edgeBand7}) - \rho_{Red-egdebande5}}{\rho_{Red-edgebande6} - \rho_{Red-edgebande5}} \tag{5}$$

REP is particularly useful for assessing vegetation stress and nutrient status.

The Red-edge NDVI (NDVIre2) utilizes the red-edge and near-infrared bands to provide a more sensitive measure of chlorophyll content, especially in areas of dense vegetation where traditional NDVI might saturate. Proposed by Gitelson (1997), it is calculated as:

$$NDVIre2 = \frac{\rho_{NIR} - \rho_{Red-edge}}{\rho_{NIR} + \rho_{Red-edge}} \tag{6}$$

This index is useful for detecting subtle changes in vegetation health and stress.

The Green Normalized Difference Vegetation Index (GNDVI), also introduced by Gitelson (1997), is similar to NDVI but uses the green band instead of the red band. It is more sensitive to chlorophyll concentration than NDVI and is calculated as:

$$GNDVI = \frac{\rho_{NIR} - \rho_{Green}}{\rho_{NIR} + \rho_{Green}} \tag{7}$$

where ρ_{Green} is the reflectance in the green band. GNDVI is effective in assessing plant vigor and can be more useful than NDVI when assessing crops at their peak growth stage.

There are also bio-physical indicators such as the Leaf Area Index (LAI), the Fraction of Absorbed Photosynthetically Active Radiation (FAPAR), and the Fraction of Vegetation Cover (FCover). LAI denotes the total leaf area for a given ground area. Changes in LAI over time are closely linked to the phenological stages of crop growth, making LAI time series particularly relevant for distinguishing between different types of crops. FAPAR corresponds to the fraction of photosynthetically active radiation absorbed by the canopy (Weiss et al., 2020). FCover corresponds to the fraction of ground covered by green vegetation. Practically, it quantifies the spatial extent of the vegetation (Sinergise, 2024).

Sonobe et al. (2018) calculated 82 indices and evaluated their importance for classifying crop types on a single image from S2 MSI. for that purpose, they used the 2 most common classifications algorithms for creating crop types, support vector machine (SVM) and Random Forest (RF). They also used the stacking methodology, which is an ensemble method where the output from multiple classifiers is passed as an input to a meta-classifier for the task of the final classification. Their results showed that vegetation indices had the greatest contributions in identifying specific crop types as the original reflectance value possessed importance values of $< 1\%$.

2.3 Crops classification

To effectively address the complexities inherent in remote sensing-based crop land classification, it is needed to acknowledge that distinct crop types exhibit maximum discrimination at various stages of their growth cycle. This variability underscores the limitations of relying on single-date image captures for comprehensive crop type differentiation. As elucidated by Hlavka et al. (1980), leveraging multi-date or multi-temporal image data significantly enhances the discrimination capability, advocating for a more dynamic approach to capturing the nuanced differences across crop types. This foundational understanding sets the stage for exploring advanced classification algorithms that

can adeptly navigate the temporal dimensions of crop growth and development

2.3.1 Supervised classification

2.3.1.1 Maximum likelihood

The Maximum Likelihood Classification (MLC) is a parametric supervised classification method. It stands as a fundamental benchmark for evaluating subsequent algorithms developed for image data analysis. At the heart of MLC is the principle that a pixel is classified into a class ω_i based on the highest probability of class membership, formalized by the decision rule:

$$\mathbf{X} \in \omega_i \text{ if } p(\omega_i|\mathbf{X}) > p(\omega_j|\mathbf{X}) \text{ for all } j \neq i, \tag{8}$$

where $\mathbf{X}$ represents a column vector of the pixel's brightness values across different measurement bands, essentially positioning the pixel within a spectral space. The probability $p(\omega_i|\mathbf{X})$ indicates the likelihood of class ω_i being the correct classification for the pixel at $\mathbf{X}$ (Richards, 2013).

However, directly determining $p(\omega_i|\mathbf{X})$ for classification is challenging as these probabilities are not readily available. Instead, what can be more easily obtained are the class conditional probabilities $p(\mathbf{X}|\omega_i)$, which estimate the likelihood of observing a pixel with spectral signature $\mathbf{X}$ in class ω_i. These probabilities are derived from labeled training data.

The relationship between the desired $p(\omega_i|\mathbf{X})$ and the available $p(\mathbf{X}|\omega_i)$ is established through Bayes' theorem, leading to the equation:

$$p(\omega_i|\mathbf{X}) = \frac{p(\mathbf{X}|\omega_i) \cdot p(\omega_i)}{p(\mathbf{X})}, \tag{9}$$

where $p(\omega_i)$ represents the prior probability of class ω_i occurring anywhere in the image, and $p(\mathbf{X})$ is the probability of observing the pixel's spectral data. Consequently, the decision rule in Equation 8 is practically applied as:

$$\mathbf{X} \in \omega_i \text{ if } p(\mathbf{X}|\omega_i) \cdot p(\omega_i) > p(\mathbf{X}|\omega_j) \cdot p(\omega_j) \text{ for all } j \neq i. \tag{10}$$

The efficacy of MLC hinges on accurate estimates of the mean vector and covariance matrix for each class, which in turn relies on having a sufficient number of training samples. Insufficient training data can lead to inaccurate parameter estimates, adversely affecting classification accuracy. utilizes only the mean spectral class positions for classification, may provide a more effective solution. This approach classifies a pixel based

on its proximity to class means, circumventing the complexities associated with covariance estimation in MLC.

2.3.1.2 Support vector machine

The original Support Vector Machine (SVM) algorithm, introduced by Vapnik, operates as a linear classifier designed to handle multidimensional data by mapping it into a higher-dimensional space. Within this expanded space, SVM seeks to identify a hyperplane that linearly separates the data, aiming to maximize the margin between distinct classes (Vapnik, 1999). This approach contrasts with traditional classifiers, which often face challenges in high-dimensional spaces due to the Hughes effect—a scenario where an increase in feature dimensions, without a corresponding rise in training data, diminishes classification accuracy. SVMs address this issue by constructing one or more hyperplanes in a space that can be high or even infinite-dimensional.

At the core of the SVM algorithm is the concept of the separating hyperplane, designed to differentiate between classes by enlarging the margin between the nearest data points of each class, termed support vectors. These vectors play a crucial role as they directly impact the hyperplane's placement. The primary goal of SVMs is to bolster model generalization to new data by prioritizing the maximization of this margin.

A key feature of SVMs is their use of kernel functions, such as linear, polynomial, and radial basis functions (RBF), which enable the algorithm to operate in a transformed feature space without directly computing the transformation (Gualtieri et al., 1999). This capability allows SVMs to adeptly manage non-linear classification tasks by identifying linear separating planes in the high-dimensional space, which correspond to complex, non-linear decision boundaries in the original input space. Through these mechanisms, SVMs maintain effective performance across a range of classification scenarios, including those challenged by high-dimensional data.

2.3.1.2 Random forest

Random forest algorithm is based on ensemble methods. they are based on the hypothesis that combining multiple models together can often produce a much more powerful model. A random forest (RF) classifier is an ensemble classifier that produces multiple decision trees, using a randomly selected subset of training samples and variables. This classifier has become popular within the remote sensing community due to the accuracy of its classifications (Belgiu and Drăguţ, 2016) (Rodriguez-Galiano et al., 2012). The performance of the random forest classifier is often compared with support vector machines in terms of classification accuracy, training time and user-defined parameters (Sheykhmousa et al., 2020).

2.3.1.2 Gradient boosting trees - Extreme gradient boosting

Gradient Boosting Trees (GBT) is an ensemble machine learning technique that builds upon the concept of decision trees by sequentially combining weak learners to create a

predictive model. Unlike random forests, which build trees in parallel, GBT constructs trees iteratively, with each new tree aiming to correct the errors of the previously combined ensemble. The principle of gradient boosting is to minimize a loss function by adding weak learners (typically shallow decision trees) in a stage-wise manner. Each iteration focuses on the residual errors of the previous ensemble. This approach allows the model to capture non-linear relationships in the data while maintaining a degree of interpretability inherited from its constituent decision trees (Elith et al., 2008).

Extreme Gradient Boosting (XGBoost), introduced by Chen and Guestrin, is an implementation of gradient boosting that has been used in machine learning competitions and practical applications (Chen and Guestrin, 2016). XGBoost modifies the traditional gradient boosting framework through several features. It incorporates L1 (lasso) regularization, which encourages sparsity in the model by penalizing the absolute values of the model parameters, and L2 (ridge) regularization, which prevents overfitting by penalizing the squared values of the model parameters. To handle large datasets, XGBoost employs a tree-building algorithm that uses histogram-based approximate greedy splitting. The algorithm can handle sparse data by learning how to manage missing values. XGBoost is designed to use multi-core processors and distributed computing environments, enhancing its computational efficiency. For datasets that don't fit into memory, XGBoost provides out-of-core computing capabilities. These features make XGBoost effective in handling structured or tabular data. In remote sensing applications, XGBoost has been applied to tasks such as land cover classification, vegetation mapping, and change detection (Georganos et al., 2018). The features of XGBoost, including its ability to handle missing data and provide feature importance rankings, make it a tool to consider in remote sensing, particularly when working with high-dimensional datasets typical in Earth observation studies.

2.3.1.3 Neural Networks and Deep Learning

Neural networks and deep learning algorithms have gained attention in remote sensing applications due to their ability to handle complex data structures and extract meaningful features automatically. These methods can be particularly useful when dealing with high-dimensional data and when temporal or spatial correlations are present. Convolutional Neural Networks (CNNs) have shown effectiveness in capturing spatial patterns in remote sensing imagery, making them suitable for tasks such as vegetation property extraction. The multi-layer architecture of CNNs allows for the identification of higher-level features and abstract concepts, potentially revealing complex relationships within the data (Zhang et al., 2016). However, the increased depth of these models often requires larger training datasets and more computational resources.

It is important to note that neural networks, especially deep learning models, are computationally intensive and complex to work with. The training process for these networks

can be time-consuming and may require specialized hardware such as GPUs or TPUs to handle the large number of parameters and the extensive computations involved. This computational complexity also extends to tasks like hyperparameter tuning and model optimization, which can be more demanding compared to traditional machine learning algorithms.

While CNNs excel at spatial pattern analysis, they may not be as well-suited for managing temporal dependencies in data. To address this limitation, Recurrent Neural Networks (RNNs) have been developed. RNNs are designed to handle sequential data by maintaining an internal state that can capture information from previous time steps. This makes them particularly useful for tasks involving time series data in remote sensing (Mou et al., 2019). A specific type of RNN, known as Long Short-Term Memory (LSTM) networks, has been introduced to address the challenge of learning long-term dependencies. LSTMs are designed to mitigate the vanishing gradient problem that can occur in standard RNNs, allowing them to capture and utilize information over longer sequences.

It's important to note that the performance of neural networks and deep learning methods can vary depending on the specific application and dataset. Some studies have found that these methods may not always outperform traditional machine learning approaches such as SVM or RF in terms of accuracy or robustness (Ma et al., 2019). The choice of classifier should be based on careful consideration of the specific remote sensing task, the characteristics of the available data, and the computational resources at hand.

When implementing neural networks for remote sensing applications, several technical considerations are important. Data preprocessing, including centering and scaling, can significantly impact model performance. Hyperparameter tuning, often done through grid search, is crucial for optimizing model architecture. The choice of activation functions can affect the network's ability to learn complex patterns. Techniques such as bagging and mini-batch gradient descent can improve model training and generalization. Additionally, the initial weight initialization can influence the convergence and final performance of the model. While neural networks and deep learning offer powerful tools for remote sensing analysis, their application should be approached with careful consideration of their strengths, limitations, and the specific requirements of the task at hand.

2.3.2 Unsupervised classification

Unsupervised classification does not require labeled training samples. It instead uses the data's inherent statistical properties to automatically categorize it into classes sharing similar spectral characteristics. This process facilitates the preliminary stages of remote sensing analysis by directly grouping the data into spectrally homogeneous clusters.

This is particularly beneficial in cases with limited prior knowledge about the dataset. Unsupervised classification streamlines the organization of data into clusters, aiding in the efficient preprocessing of imagery for further analysis with supervised classification techniques if necessary. Furthermore, the advantage of unsupervised classification is its capability to organize and label large datasets relying on the intrinsic features of the data for class prediction. This makes it a useful tool for initial data examination and segmentation. Combining unsupervised and supervised classification methods can enhance the accuracy of the analysis, leading to the development of semi-supervised frameworks that leverage the strengths of both approaches

2.3.2.1 K-means clustering

The k-means clustering algorithm categorizes data into a specified number of clusters, denoted as 'k'. This method operates through iterative steps aimed at minimizing the Euclidean distance between the samples within each cluster and the cluster's centroid. Initially, the algorithm requires the assignment of measurement vectors to these clusters, based on an arbitrarily chosen set of initial cluster centers or means. Following this preliminary grouping, the means of the clusters are recalculated, and the samples are reassigned to the cluster with the nearest mean. This process of reassignment and recalculation is repeated until there is no significant movement of samples between clusters, indicating stability within the cluster assignments.

The quality of the clustering outcome is influenced by the choice of initial samples and centroids. Furthermore, like many unsupervised classification algorithms, k-means is sensitive to the presence of outliers, which can significantly affect the final clustering result. The inherent randomness in the initial selection of cluster centers means that the algorithm's outcome can vary, underlining the importance of employing methods such as cross-validation to verify and enhance the reliability of the clustering results. This approach ensures that the algorithm's application, especially in complex datasets, achieves an optimal organization of data into meaningful clusters that accurately reflect the underlying patterns within the data.

Unsupervised classification algorithm such as k-means can be used jointly with supervised ones. He et al. (2014) obtained an overal accuracy of 97,75% using both SVM and a fuzzy k-means algortihm (FKM). That accuracy was higher than using SVM alone.

2.3.3 Hughes phenomenon

In the domain of supervised machine learning, training a classifier requires an adequate number of samples for each class. The ability of a classifier to apply learned patterns to unseen data, known as generalization, hinges on a substantial dataset. A key premise is that a robust dataset is necessary for the accurate estimation of parameters critical for reliable predictions. Insufficient data leads to undertraining and suboptimal performance

when the classifier encounters unlabelled data.

The interplay between the number of features, or dimensions, and the volume of training samples is crucial. More features can initially seem beneficial by providing additional information, potentially improving classifier performance. However, this benefit exists only to a certain limit. As feature space dimensionality expands without an equivalent increase in training samples, parameter estimation becomes more complex. This complexity adversely affects both the training efficacy and the generalization capacity of the classifier.

Hughes phenomenon encapsulates this challenge (see figure 4), illustrating the trade-off between the number of features and the size of the training set. This phenomenon has been observed in many applications, where adding features incrementally improves classification accuracy only up to a certain point. Beyond this point, the reliability of parameters estimated through maximum likelihood deteriorates, diminishing classification accuracy. This highlights the importance of maintaining a balance between feature dimensionality and training data availability in designing classifiers.

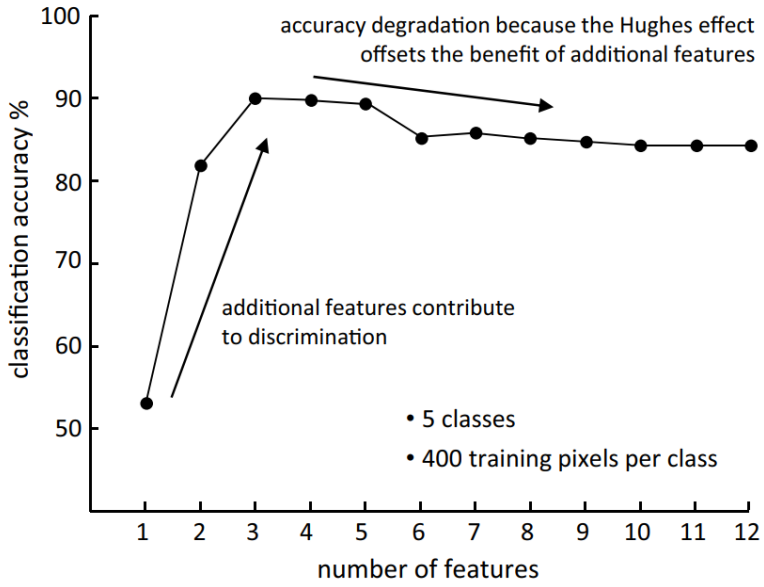


Figure 4: Actual experimental classification results showing the Hughes phenomenon—the loss of classification accuracy as the dimensionality of the spectral domain increases (Richards, 2013)

2.3.4 Data distribution and resampling methods

The distribution of training data across different classes influences the performance of machine learning models for land cover classification. Imbalanced datasets, where certain classes are underrepresented, can affect model predictions. This issue is particularly relevant in remote sensing applications, as land cover classes often have natural imbalances (Krawczyk, 2016).

To address this challenge, various resampling techniques have been developed. These methods adjust the class distribution in the training set by either reducing the number of samples in majority classes (under-sampling) or increasing the number of samples in minority classes (oversampling). Some approaches combine both techniques.

One widely used oversampling method is the Synthetic Minority Oversampling Technique (SMOTE) (Chawla et al., 2002). SMOTE generates synthetic samples for minority classes by interpolating between existing instances in the feature space. This approach balances the dataset without simply duplicating existing samples, which can help mitigate over-fitting. Several variants of the SMOTE algorithm have been developed to address different aspects of the imbalanced data problem:

- ADASYN (Adaptive Synthetic Sampling Approach): this extension of SMOTE considers density distribution to determine the number of synthetic instances created for challenging samples (He et al., 2008).
- SMOTE + Tomek Links: this hybrid approach combines SMOTE oversampling with Tomek links under-sampling to remove overlapping data points and enhance class separation (Swana et al., 2022).
- SMOTE + Edited Nearest Neighbors (ENN): another hybrid method that uses ENN to remove misclassified samples from both classes after SMOTE oversampling, resulting in clearer decision boundaries (Hai Ly et al., 2022).

Figure 5 provides a visual comparison of these SMOTE variants. The figure illustrates how different SMOTE algorithms affect the distribution of minority class samples. We can observe that while all methods increase the number of minority samples, they do so in distinct ways. SMOTE + Tomek Links and SMOTE + ENN show how the hybrid approaches not only generate new samples but also clean the decision boundaries between classes.

These resampling techniques can enhance the performance of machine learning models in remote sensing applications, especially when dealing with imbalanced land cover classes. However, the choice and application of these methods require careful consideration to avoid introducing artificial bias or over-fitting (Johnson and Iizuka, 2016). The effectiveness of each method may vary depending on the specific characteristics of the

dataset and the classification task at hand.

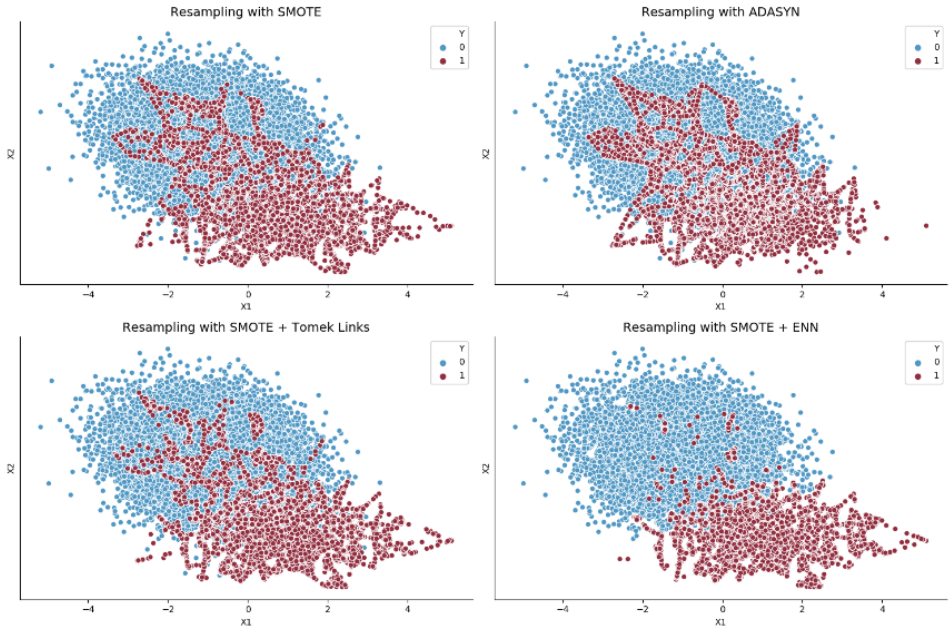


Figure 5: Overview of different SMOTE algorithms results (SATPATHY, 2020)

2.3.5 Assessing accuracy - Performance metrics

Evaluating the performance of a classifier involves several metrics, most of which stem from confusion matrices. These matrices provide a clear depiction of how well a classifier predicts across different classes, with the actual classes of the samples displayed in rows and the predicted classes in columns.

		Class	
		Positive	Negative
Prediction	Positive	TP (True Positive)	FP (False Positive)
	Negative	FN (False Negative)	TN (True Negative)
		P = TP + FN	N = FP + TN

Table 2: Performance metrics : confusion matric

Accuracy, which is the percentage of correctly classified samples :

$$Accuracy = \frac{TP+FN}{TP+TN+FP+FN} = \frac{TP+TN}{P+N} \tag{11}$$

It is an appropriate metrics when classes are not unbalanced and that the errors FP and FN have the same importance. But accuracy doesn't totally translate the behaviour of classifiers.

Precision, which gives the percentage of samples classified as positives that are truly positives.

$$Precision = \frac{TP}{TP+FP} \tag{12}$$

Recall, which gives the percentage of positive samples that are correctly classified as positives.

$$Recall = \frac{TP}{TP+FN} = \frac{TP}{P} \tag{13}$$

Precision and recall, while informative, may not fully capture a classifier's behavior, especially in cases of class imbalance. These metrics, when considered independently, can be misleading as they reflect only partial aspects of performance. For instance, a classifier could achieve high precision by making very few positive predictions, or high recall by classifying most instances as positive, neither of which may be desirable in practice. The F1-score addresses this limitation by combining precision and recall into a single, balanced metric. It is particularly useful for imbalanced datasets and can be weighted to prioritize either precision or recall based on the specific requirements of the classification task. This provides a more comprehensive evaluation of classifier performance, especially when the consequences of false positives and false negatives differ in importance.

$$F1 = 2.\frac{Precision.Recall}{Precision+Recall} \tag{14}$$

Finally, we have the overall accuracy (OA):

$$OA = \frac{\sum_i (TP_i + TN_i)}{\sum_i (TP_i + TN_i + FP_i + FN_i)} \tag{15}$$

2.4 Crop classification in smallholder West African agriculture: challenges and existing approaches

Crop type maps derived from satellite data are used for quantifying crop growth, health, and yields. However, such maps are often unavailable for many parts of Africa, where smallholder farming is prevalent. Several studies have examined the challenges of crop

mapping in African contexts, particularly in West Africa, where conditions similar to the study area in Senegal are observed.

Ibrahim et al. (2021) identified challenges in creating crop type maps for sub-Saharan Africa, including cloud cover, small farm sizes, and mixed cropping systems. In their study on the Jos Plateau, Nigeria, Sentinel-2 time series and high-resolution SkySat data were employed to map main crops and intercropping systems with a random forest classifier. Spectral-temporal metrics and a hierarchical random forest classification were utilized, resulting in an overall accuracy of 84% for cropland identification and 72% for crop type mapping. The effectiveness of their classification system was demonstrated for intercropped classes, which are characterized by temporal and spatial heterogeneity.

In West Africa, Forkuor et al. (2014) worked with high spatial resolution multi-temporal optical (RapidEye) and dual polarized SAR (TerraSAR-X) data were integrated to map crops in northwestern Benin. Using the random forest classification algorithm. The overall goal was to ascertain the contribution of the SAR data to crop mapping in the region. The integration of optical and SAR data was found to improve classification accuracy by 10-15% compared to using optical data alone, with accuracies of up to 75% achieved when both data types were available for the entire cropping season.

A review by Choukri et al. (2024) noted increasing interest in combining optical and radar data to improve crop classification accuracy. The complementary nature of these data sources was highlighted, with optical imagery providing spectral information and radar offering penetration capabilities in challenging environmental conditions. However, further research is needed to refine and adapt remote sensing image combination methods and data classification strategies for the African context, taking into account the characteristics and spectral signatures of African crops and the heterogeneity of agricultural practices across the continent.

Inglada et al. (2015) conducted a global study to assess the potential of high-resolution multi-temporal optical imagery, specifically using SPOT4 (Take5) and Landsat 8 data, for crop type mapping using state-of-the-art supervised classification methods. Their research, which included test sites in Burkina Faso, revealed significant challenges specific to West African contexts. While their random forest classifier achieved overall accuracies above 80% for most global sites, the Burkina Faso site showed notably lower performance, with an overall accuracy of approximately 55% for crop-type classification. This poor performance was attributed to the presence of mixed tree-crop systems in the fields, which complicated the spectral signatures. Importantly, the study highlighted the difficulty in discriminating between corn and sorghum in the West African context. Their confusion matrix for the Burkina Faso site showed misclassification between these two crop types, underscoring a challenge in crop mapping for the region.

Despite these advancements, there remains a gap in the literature regarding the specific

Table 3: Confusion matrix for Random Forest classifier on the Burkina Faso test site (Inglada et al., 2015)

Class	Maize	Sorghum	Cow Peas	Cotton	Other Crops
Maize	**155**	21	3	56	1
Sorghum	62	**9**	3	38	1
Cow Peas	30	9	**12**	21	1
Cotton	44	10	3	**262**	0
Other Crops	48	11	15	25	**3**

challenges of differentiating between millet, sorghum, and corn using optical imagery in West African contexts.

2.5 Literature review conclusion

The literature review has revealed a diverse array of methodologies for classification using remote sensing data. It has been observed that effective results can be achieved through various approaches, encompassing both a wide range of algorithms and diverse data types.

In the context of crop type classification, significant emphasis has been placed on the computation of specific features that enable discrimination between different crop types. The review highlighted the importance of spectral indices, such as NDVI, EVI, and GNDVI, in capturing vegetation characteristics. Additionally, the potential of red-edge bands in enhancing classification accuracy was noted, particularly through indices like RZP and NDVIre2.

The limitations of optical remote sensing, primarily due to cloud cover and atmospheric interference, were discussed. This led to the exploration of complementary data sources, such as S1 SAR data, which can provide valuable information even in cloudy conditions. The integration of optical and radar data has shown promise in improving classification accuracy and robustness.

To address the common issue of class imbalance in remote sensing datasets, SMOTE was identified as a potential solution. This method, along with its variants, offers a way to improve classifier performance by balancing the representation of different crop classes in the training data.

Among the various classification algorithms reviewed, the XGBoost classifier was selected for this study due to its high performance in handling complex, high-dimensional datasets typical in remote sensing applications.

While the literature provides a strong foundation for crop type classification using re-

mote sensing data, it also highlights areas for further exploration. These include the optimization of feature selection for specific crop types, the development of more sophisticated data fusion techniques for optical and radar data, and the investigation of advanced deep learning approaches that can better leverage the spatial and temporal aspects of satellite image time series.

Chapter 3: objectives

3.1 General Context

Senegal faces challenges in accurately mapping and monitoring its agricultural lands. In Senegal, the Direction de l'Analyse, de la Prévision et des Statistiques Agricoles (DAPSA) conducts annual field campaigns to collect data from farmers. However, these campaigns are time-consuming and logistically challenging due to the extensive territory and limited road network.

As explained is section 2.2.1, S2 is an ideal data source for developing crop mapping tools that can potentially streamline and complement traditional field surveys. The availability of this satellite data offers an opportunity to enhance agricultural monitoring in Senegal, potentially reducing the reliance on extensive field campaigns while maintaining or improving the quality of crop-type information.

3.2 Objectives

In this context, this master's thesis focuses on the Kolda region of Senegal, with the overall objective of developing two classification approaches to create an accurate crop-type map using S2 time series imagery. The study aims to assess whether optical data from S2 alone is sufficient for this task or if additional data sources are needed.

To achieve this goal, the research will first pre-process and analyze both in-situ data from DAPSA and S2 imagery according to state-of-the-art standards. This step ensures the quality and reliability of the input data for subsequent analysis.

A part of this study is to compare the performance of the XGBoost classifier for different set of features and different data resampling algorithms. XGBoost has shown promising results in handling complex, high-dimensional datasets typical in remote sensing applications, and its performance will be evaluated.

Given the class imbalance present in the in-situ data collected by DAPSA, this research will assess and compare various SMOTE algorithms. The aim is to address the imbalance issue effectively and improve the overall classification accuracy.

The study will also assess the accuracy and reliability of crop-type maps produced using only S2 optical data, on cloud free with a focus on spectral information in opposition with temporal information and will investigate the spectral signature of the received in-situ data. This evaluation will help determine if this data source alone is sufficient for accurate crop-type mapping in the Kolda region, or if future studies should consider integrating additional data sources such as SAR data.

Finally, based on the findings, this research aims to provide recommendations for future crop-type mapping efforts in Senegal.

Chapter 4: materials and method

This chapter presents a detailed description of the materials used and the methodological approach employed in this study to develop and evaluate crop classification models in Kolda, Senegal. It begins by defining the region of interest and describing the in-situ data collection process, including the challenges encountered during field campaigns. The chapter then outlines the extensive pre-processing steps applied to both the in-situ data and S2 satellite imagery. Following this, the feature extraction process and the development of various classification scenarios are explained, including the application of different SMOTE algorithms to address class imbalance issues. The chapter also details the methods used for spectral signature analysis and the exploration of classification without time-series data. Finally, it describes the evaluation metrics used to assess model performance and the strategy employed for scaling up the classification to the entire region of interest.

4.1 Region of interest

The region of interest (ROI) is located in south Senegal, in the Kolda region. More specifically, the ROI is located under 3 different S2 tiles (28PDV, 28PEV and 28PFV) each one covering 100 km^2 (figure 6). This region is characterized by its mains crops which are namely peanut, rice, corn, millet and sorghum (Knoema, 2024).

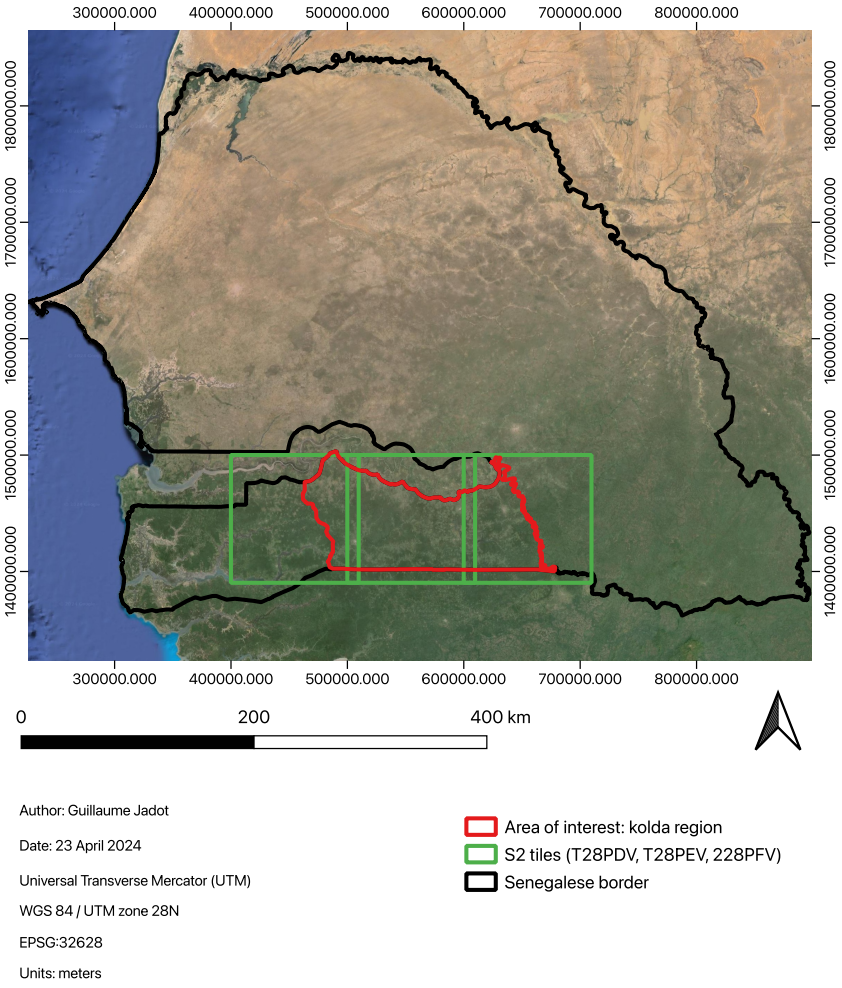


Figure 6: ROI location: Kolda region in Senegal under 3 Sentinel 2 tiles

4.2 In-situ data preparation

4.2.1 Field campaign for in-situ data collection

In the framework of the Sen4Stats project, that aims to facilitate the uptake of earth observation information by the national statistical offices supporting the agricultural statistics (Sen4Stat, 2024), a field campaign was conducted in October 2023 in collaboration with DAPSA in 5 districts (administrative units) including Kolda. While October marks the end of the crop calendar in the region, the timing still allowed for observation of distinguishing characteristics among different crops within the Tambacounda region (Figure 7).

The campaign provided firsthand experience of the challenges involved in collecting agricultural data in the region. Traveling between villages on poorly maintained road networks proved to be arduous, and accessing individual fields often required walking considerable distances under intense heat. These difficult conditions not only tested the resolve of the data collection team but also highlighted potential sources of bias in the collected data.

The physical constraints of the environment sometimes made it impractical to reach certain crop fields, even when they were deemed important for the campaign's objectives. This limitation could potentially lead to under-representation of some crop types or growing conditions in the final dataset.

Communication presented another significant challenge during the campaign. In smaller villages, linguistic diversity often necessitated the use of interpreters or reliance on local guides. Obtaining permission from village chiefs to visit crop parcels sometimes involved complex negotiations, further complicating the data collection process.

These observations underscore the complexity of conducting comprehensive agricultural surveys in regions like Kolda. They also highlight the potential value of remote sensing techniques in complementing traditional field surveys, particularly in areas where physical access or communication barriers may limit the scope of on-the-ground data collection.

The experiences and challenges encountered during this in-situ campaign will be discussed in more detail in subsequent sections, providing context for the interpretation of the collected data and the results of the crop classification model.

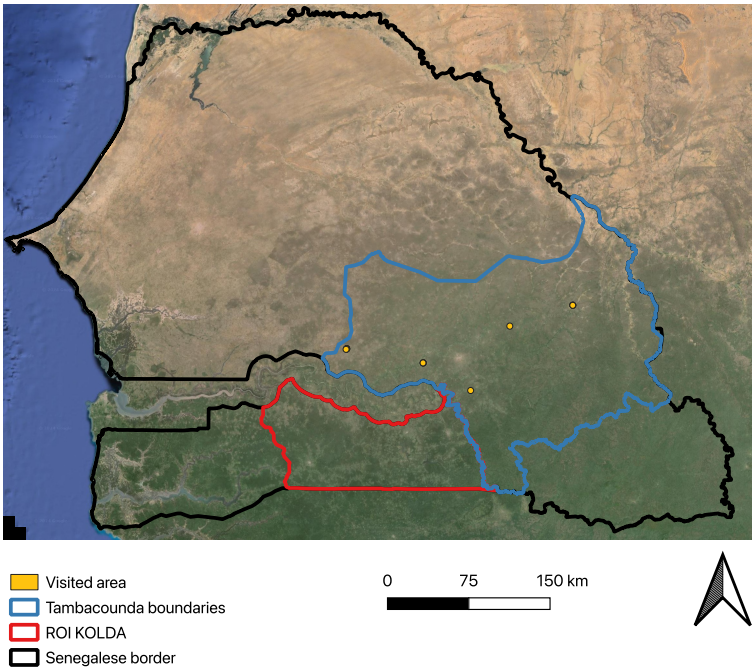


Figure 7: Area visited during the field campaign in October

4.2.2 Exploratory analysis

DAPSA provided a file containing information about each field visited during their 2023 in-situ campaign, along with corresponding GPX tracks for each crop polygon, which delineate the boundaries of each field (i.e., parcel).

The initial data analysis revealed inconsistencies between the crop ID labels in the file and the corresponding GPX tracks. These inconsistencies included duplicate entries and mismatches between the IDs, leading to potential data loss and reduced reliability of the dataset. To address this issue, a spatial join was conducted between the polygons defined by the GPX tracks and the center point coordinates provided for each crop in the attribute data. This approach involved linking each crop's center (as indicated by the file) with the nearest GPX polygon, effectively matching the field records with their spatial boundaries. Although this method introduced some uncertainty, it was the most practical solution given the data constraints.

The dataset was then filtered to include only polygons within the Kolda region. This initial processing resulted in 445 polygons, as shown in Figure 8.

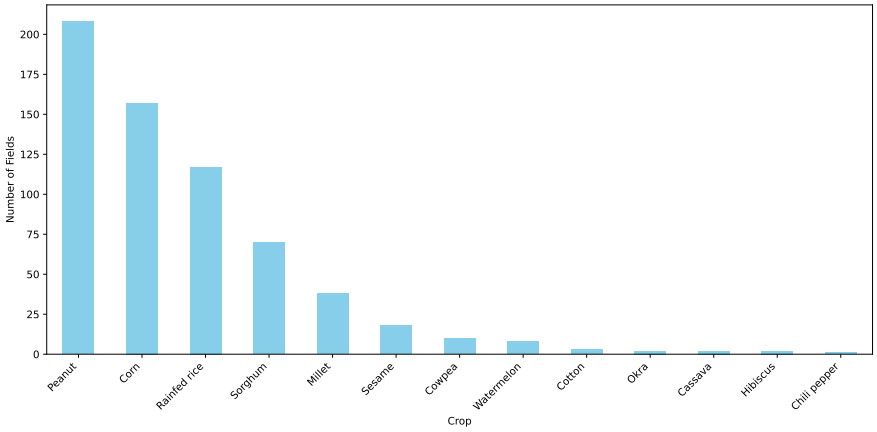


Figure 8: Initial distribution of crop polygons for each class in the Kolda region

Further pre-processing steps were applied to ensure data quality. An inner buffer of 10 meters was applied to each polygon to mitigate edge effects and avoid contamination from border pixels. This step reduced the number of polygons from 445 to 424, as some polygons became too small after buffering. Polygons that were too small to accommodate the inner buffer were removed to maintain data integrity. Only classes with more than 6 polygons were retained to ensure a minimum amount of data for the classifier. This step resulted in the removal of 7 crop classes that had insufficient representation: cowpea, watermelon, cotton, okra, cassava, hibiscus, and chili pepper.

It's worth noting that these pre-processing steps significantly reduced the number of polygons to ensure better data quality and representativeness for each class. This reduction, while necessary for robust classification, also highlights the challenges in maintaining sufficient sample sizes for all crop types in the region.

4.2.3 Non-crop classes

To complement the crop data, non-crop classes were added through photo-interpretation of Google Satellite images. The selection of non-crop classes was inspired by the World Cover classification scheme and includes:

- Trees
- Shrubland
- Grassland
- Urban
- Bare
- Water
- Herbaceous wetland

For these non-crop classes, point data were collected and a buffer of 15 meters was applied to each point. This buffer size was chosen to increase the number of pixels for each non-crop class while ensuring that the spectral signature remained pure and representative of the class.

After integrating the non-crop classes, the total number of polygons in the dataset increased to 1066. However, further filtering based on a minimum pixel count of 3 per polygon and a minimum of 6 polygons per class was applied to ensure robust representation. This final pre-processing step resulted in a set of 1054 polygons of well-represented classes for the classification process.

The final distribution of crop polygons and the number of pixels for each class after pre-processing are illustrated in Figure 9.

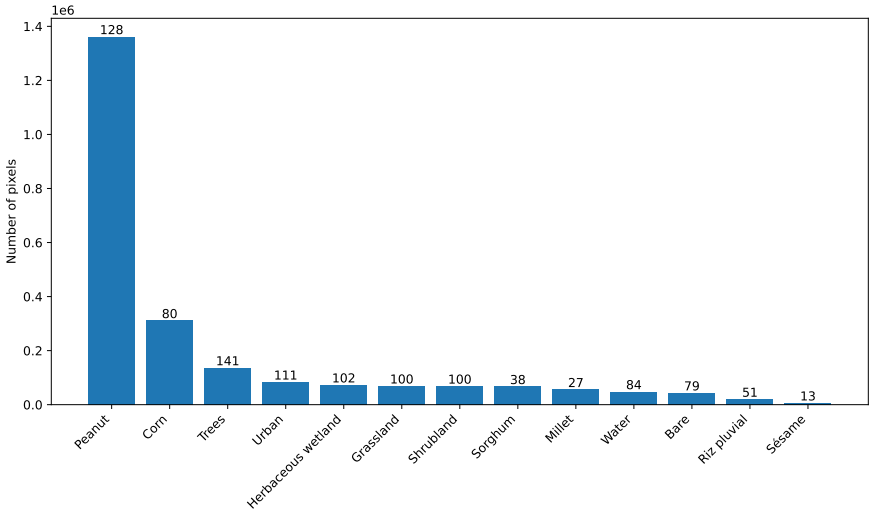


Figure 9: Count of crop polygons and number of pixels for each class after pre-processing

4.2.4 Calibration and validation datasets split

The final pre-processed dataset resulted in a Shapefile containing 1054 polygons, categorized into 13 distinct classes. These classes include agricultural crops such as peanut, maize, millet, rainfed rice, sorghum, and sesame. Non-crop classes were also represented, comprising trees, shrubland, grassland, urban areas, bare land, water bodies, and herbaceous wetland.

To prepare for the classification process, this dataset was divided into calibration and validation subsets. The splitting procedure was designed to maintain the proportional representation of each class in both subsets, ensuring that the model would be trained

and evaluated on representative samples of all land cover types. A target proportion of 75% for calibration and 25% for validation was established.

The splitting process began by calculating the total pixel count for each class. Based on the 75% target, pixel count targets for the calibration set were determined for each class. To avoid any spatial or temporal biases in the selection process, the dataset was randomly shuffled. Polygons were then iteratively assigned to either the calibration or validation set. For each polygon, if adding its pixels to the calibration set for its class would not exceed the target, it was assigned to calibration; otherwise, it was assigned to validation. This process continued until all polygons were assigned, ensuring that the calibration set contained approximately 75% of the pixels for each class, with the remaining polygons forming the validation set.

This stratified splitting approach ensures that both the calibration and validation sets contain representative samples of all land cover types, which is needed for developing a robust and generalizable classification model. By maintaining the class proportions in both sets, the method helps to mitigate potential biases that could arise from imbalanced representation in the calibration or validation phases. The entire pre-processing workflow, from initial data cleaning to the final dataset splitting, is illustrated in Figure 10. This data preparation and splitting process forms the foundation for the subsequent classification tasks.

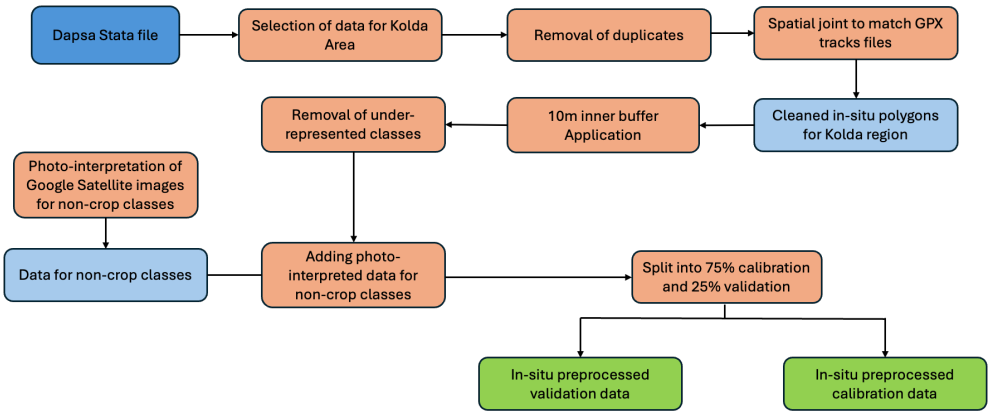


Figure 10: Overview of the preparation and splitting of the in-situ data

4.2.5 Dealing with unbalanced classes

To address the issue of class imbalance in the dataset, two methods discussed in Chapter 2 were implemented in this study. The first approach involved under-sampling the majority class (Peanut), while the second method utilized the SMOTE algorithm to create synthetic data for minority classes.

As illustrated in Figure 11, the application of these resampling methods resulted in a more balanced distribution of pixels across crop classes. The oversampling strategy increased the representation of minority crop classes to match the count of the largest minority class. Concurrently, the under-sampling strategy reduced the majority class to at most the same count as the largest minority class.

This balanced approach allows the classifier to learn from a new distribution of in-situ data, mitigating potential bias that could arise from a disproportionately large majority class. By adjusting the proportion of each class, the model is given an opportunity to develop a more robust understanding of the distinguishing features across all crop types, including those originally underrepresented in the dataset.

It is worth noting that the non-crop classes were not subjected to this resampling process. This decision was made due to the inherent uncertainty in the purity of spectral signatures for these classes. For instance, areas classified as bare soil might inadvertently include patches of trees or shrubs. While this uncertainty is relatively small, it was deemed significant enough to warrant caution. Creating synthetic data based on potentially mixed spectral signatures could lead to the generation of inaccurate or misleading calibration samples. Therefore, the non-crop classes were left in their original state to maintain the integrity of their spectral information.

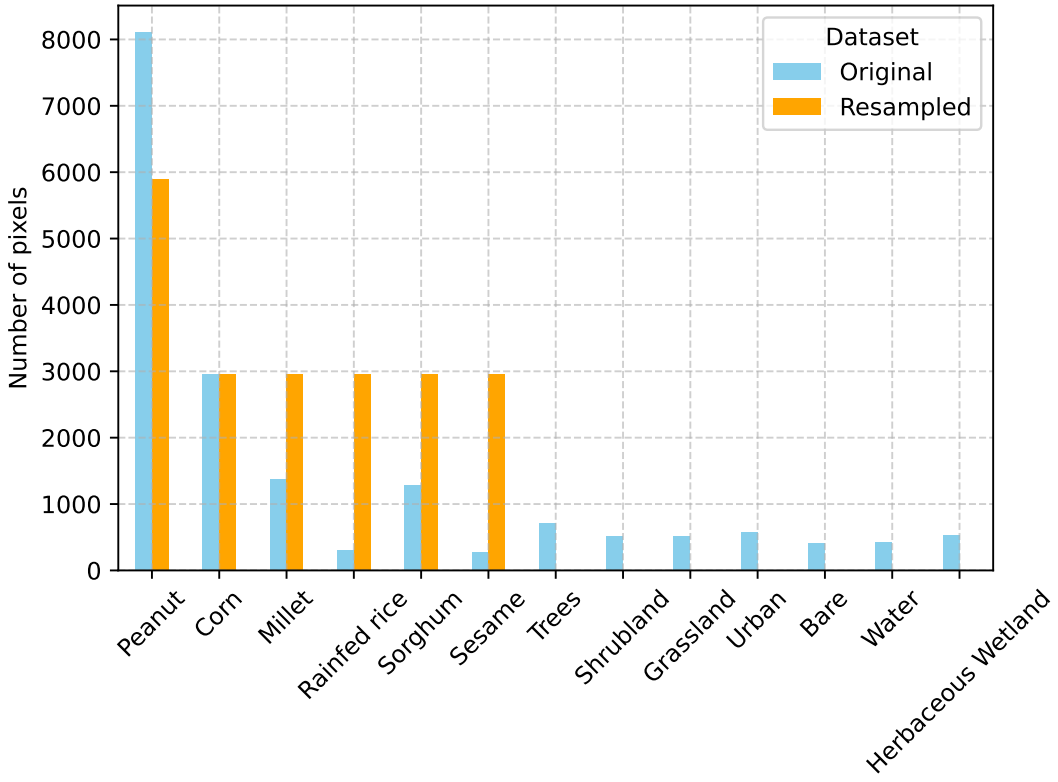


Figure 11: Distribution of data samples after SMOTE application

4.3 Sentinel-2 data pre-processing

The selection and preparation of appropriate features is essential for classification tasks. In this study, features are derived from time series of S2 images, requiring effective pre-processing to optimize data quality and enhance classification accuracy.

Raw satellite data present several challenges, including the presence of clouds and their shadows, saturated pixels, and residual atmospheric effects. To address these issues, the initial optical data used in this study consists of L2A S2 images that have already undergone atmospheric correction. As a result, the pixel values correspond to bottom-of-atmosphere calibrated reflectances.

All S2 bands were utilized in this study, with the exception of bands 1 and 8a. To ensure consistency in spatial resolution across all data, the reflectance bands originally at 20-meter resolution were resampled to 10 meters using bilinear interpolation. This method computes each new pixel value as a weighted average of the nearest four pixels, resulting in smoother transitions between different regions of the image compared to the

nearest neighbor method.

The Scene Classification Layer (SCL) bands were treated differently due to their categorical nature. These bands were resampled using the nearest neighbor method, where each new pixel in the enlarged image adopts the value of the nearest original pixel. This approach preserves the discrete classification values essential for accurate scene interpretation.

4.3.1 Clouds masking, time-series, temporal resampling and gap-filling

The dataset comprised 292 S2 images from three tiles (T28PFV, T28PEV, and T28PDV) covering the ROI for the year 2023. Due to orbital differences, the number of acquisition dates varied between tiles:

- T28PFV, which encompasses two orbits, had 146 acquisition dates.
- T28PEV had 73 acquisition dates.
- T28PDV, similar to PEV, also had 73 acquisition dates.

The date range spanned from early January (January 1st for PFV, January 4th for PEV and PDV) to December 30, 2023. For the detection and masking of cloud cover, the Sen2cor algorithm was employed. This processing tool from the European Space Agency, categorizes scenes by assigning cloud probabilities at three distinct levels. In addition to cloud identification, Sen2cor also distinguishes other surface features such as snow and water bodies, while flagging any invalid data points. Following cloud masking, images were stacked for each band as separate rasters:

- T28PFV: 146 bands (one per acquisition date)
- For T28PEV and T28PDV: 73 bands each (one per acquisition date)

To mitigate the impact of cloud-masked pixels and ensure uniform temporal sampling across all tiles, the dataset underwent further processing. A series of 37 synthetic observation dates was established, spaced at regular 10-day intervals spanning from January 4th to December 30th, 2023. For individual pixels within the study area, spectral reflectance time series were constructed for each spectral band. Where cloud masking had created data gaps, values were estimated through linear interpolation, utilizing the nearest available cloud-free observations on either side of the masked date.

Although S2 imagery for the entire year 2023 was available, the analysis focused on the period from June to December. This time frame was selected based on the crop calendar of the Kolda region, which encompasses the primary growing season for the diverse array of crops present in the area. By concentrating on these months, the study captures the seasonal variations and phenological characteristics needed for accurate crop type

discrimination. This approach enhances the classifier's ability to distinguish between different crop types and especially avoids potential saturation with less informative data from the non-growing season.

4.3.2 Mosaic

To work on the entire ROI over the different Sentinel-2 tiles, a mosaicking step was required. This was achieved by concatenating together images from each tile that had the same date and band, creating a seamless and comprehensive image for the entire region of interest.

The overall approach for Sentinel-2 images pre-processing can be seen in Figure 12

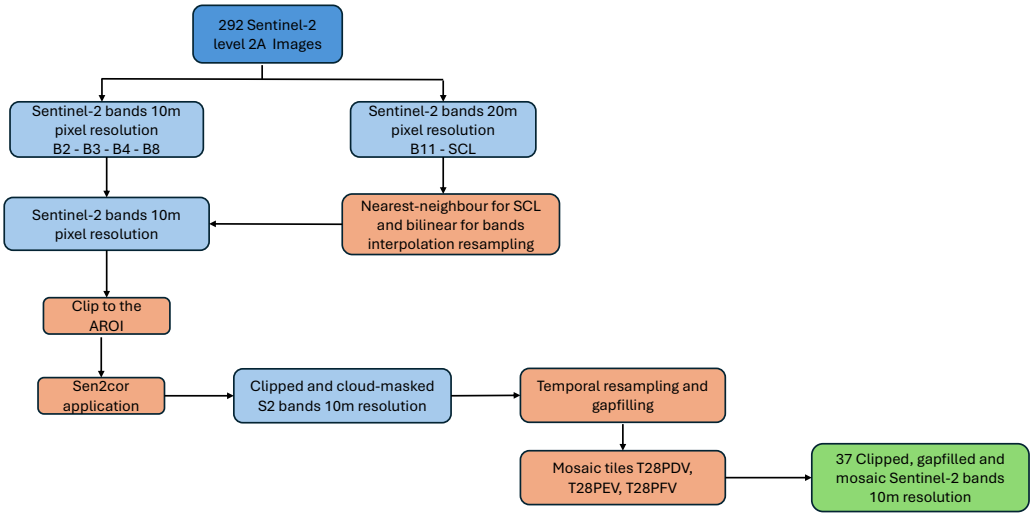


Figure 12: Overview of the pre-processing steps applied to the Sentinel-2 images

4.4 Spectral indices

Following the completion of pre-processing steps, spectral indices time-series were computed based on the mosaiced images. These time-series of indices can be directly utilized as input features for classification algorithms, and they also form the basis for

computing crop-specific metrics. The spectral indices employed in this study, having been previously detailed in Chapter 2, will not be reiterated here. For each of the 37 images in the time series, the following indices were derived:

- BSI
- NDVI
- NDVIre2
- REP
- GNDVI

4.5 Data transformation and feature extraction

The preparation of data for the classifier required a strategic approach to handle the large size of the ROI and the mosaiced S2 imagery. Given the computational constraints, a method was developed to extract features only for areas corresponding to the in-situ polygons, avoiding the need to load entire images into memory.

The in-situ data was first transformed from polygonal data to a raster whose pixel values were given by the factored crop type attribute of the polygons, transforming the crop types classes to numbers.

This process involved sequential processing of each input raster, including S2 spectral bands and derived indices such as NDVI, BSI, NDVIre2, REP, GNDVI, and EVI. The rasterized in-situ data served as a mask to identify and extract relevant pixels. This approach ensured that only the necessary data was processed, significantly reducing memory requirements.

For efficient storage and retrieval, the extracted samples were stored in HDF5 format. This method allowed for flexibility in the workflow, enabling the loading of previously extracted samples when needed, thus saving computational time in subsequent analyses.

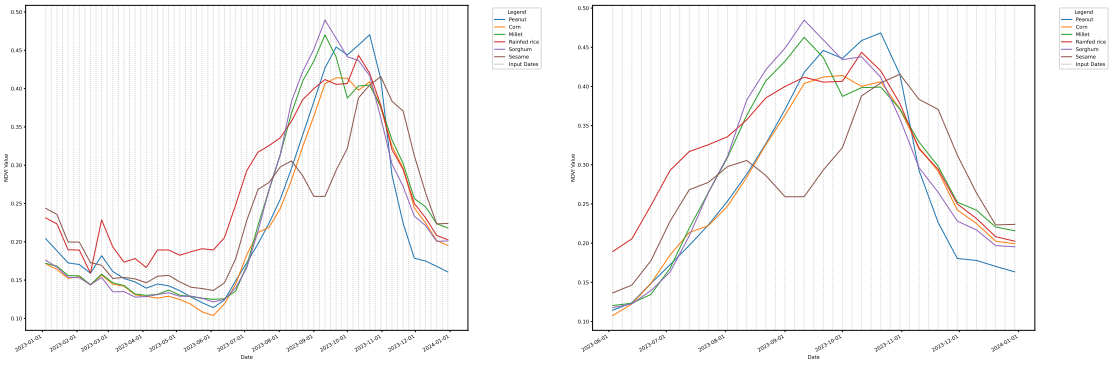
The feature extraction process produced two 3-dimensional matrices—one for the calibration dataset and one for the validation dataset—each containing spectral bands and computed metrics across multiple time steps. For both datasets, the matrices have dimensions corresponding to the number of pixels (17978 for calibration and 6819 for validation), 15 spectral features (9 spectral bands and 6 computed indices), and 22 selected time steps. These 22 time steps were chosen based on crop calendars and NDVI profiles generated from 400 random pixels per crop class (as shown in figure 13). Analysis of these profiles revealed that dates before June contributed minimal useful information, leading to their exclusion.

To prepare the data for classification, the 3D matrices were reshaped into 2D arrays. The spectral bands were flattened into arrays with dimensions (17978, 198) for calibration and (6819, 198) for validation, where 198 is the product of 9 spectral bands and 22 time

steps. Similarly, the computed metrics were reshaped into (17978, 110) and (6819, 110), reflecting 5 metrics over 22 time steps. These reshaped arrays were then concatenated along the feature axis, resulting in final input matrices with dimensions (17978, 308) and (6819, 308) for calibration and validation, respectively. Corresponding label arrays, with shapes (17978,) and (6819,), contain the encoded crop type classes for each pixel, ensuring correct alignment between features and labels.

For the labels, corresponding arrays were created to match the extracted features. The calibration label array, for instance, has a shape matching the number of pixels in the calibration dataset. After filtering out no-data values, the resulting calibration labels array contains 6819 samples, each representing a specific crop type. These labels serve as the ground truth for evaluating the classifier's performance on unseen data.

To address the class imbalance issue, a combination of oversampling and under-sampling techniques was implemented as discussed previously. Custom sampling strategies were defined for both SMOTE and random under-sampling. For oversampling, all minority classes were increased to match the count of the largest minority class. For under-sampling, the majority class was reduced to at most the same count as the largest minority class. The resampling was implemented using a pipeline that first applies SMOTE and then random under-sampling, resulting in a balanced dataset.



(a) NDVI mean curves (out of 400 random pixels) for each crop class for the entire year

(b) NDVI mean curves (out of 400 random pixels) for each crop class for the selected dates

Figure 13: NDVI mean curves for crop classes

4.6 Algorithm training, model and SMOTE algorithms comparison

4.6.1 Model and scenarios

The classifier used in this study is an XGBoost classifier, selected for its performance and efficiency in training and prediction. The model was configured with 300 estimators, setting the number of boosting rounds or trees to be built. This configuration allows

the model to capture complex patterns in the data. A learning rate of 0.02 was chosen to enable more conservative boosting, potentially improving generalization by reducing the impact of individual trees and allowing the model to learn refined features over multiple iterations. The maximum depth of each tree was limited to 5 to help prevent overfitting by restricting tree complexity. This parameter balance aims to create a model capable of capturing important data patterns while maintaining generalization to unseen examples.

To assess the importance of features, metrics and various SMOTE algorithms, different scenarios were tested:

- Scenarios 1-3: Spectral bands only, Metrics only, and Combined (Spectral + Metrics), using SMOTE
- Scenarios 4-6: Same as 1-3, but using ADASYN
- Scenarios 7-9: Same as 1-3, but using SMOTEENN

For each scenario, the model was trained and tested on the same calibration dataset (17978 pixels) and validation dataset (6819 pixels). The class imbalance issue was addressed using a combination of oversampling (SMOTE variants) and under-sampling techniques, as discussed in previously. This approach aims to provide the classifier with a more balanced dataset, potentially improving its ability to correctly identify all crop types, including those with fewer samples in the original dataset.

It should be noted that cross-validation was not performed in this study, but it could have provided valuable insights into model stability and generalization. Cross-validation, especially k-fold cross-validation, involves dividing the dataset into k subsets and training the model k times, each time using a different subset as the validation set and the remaining k-1 subsets as the training set. This approach helps to ensure that the model's performance is not overly dependent on a specific split of the data and reduces the risk of overfitting. Implementing cross-validation could offer a more robust evaluation by providing a better estimate of how the model is likely to perform on unseen data. It would also allow for tuning hyperparameters more effectively, optimizing the model for improved generalization across different scenarios.

4.6.2 Training with Sen4Sats cropland mask

A final scenario was evaluated to assess the classifier's ability to differentiate between crop types based on their spectral bands and/or metrics. This scenario utilized a cropland mask obtained from the Sen4Stats project to mitigate potential confusion with non-crop classes. The training of the model was conducted using only crop classes, employing the combination of features and SMOTE algorithm that demonstrated the best performance in previous scenarios. This approach involves a significant assumption: the model that

performs best for distinguishing between crop and non-crop classes may not necessarily be the most effective for differentiating between crop classes alone.

The cropland mask was used to focus the classification task on agricultural areas only. This approach aimed to improve the model's ability to distinguish between different crop types. By removing non-crop classes from the training data, the classifier could potentially better capture the differences in spectral and metric characteristics among crops. The results of this scenario were expected to show how well the model could differentiate between crop types when only considering agricultural land. A more detailed analysis of the spectral signatures for each crop type was planned for later sections of the study.

4.6.3 Resampling with SMOTE algorithms

To visualize the impact of different SMOTE algorithms on the time-series dataset, Principal Component Analysis (PCA) was performed on augmented and non-augmented datasets, reducing the high-dimensional data to two principal components. Figures 14 and 15 illustrate the effects of different SMOTE algorithms on the spectral and metrics datasets, respectively for crop classes. These classes include peanut, corn, millet, rainfed rice, sorghum and sesame. Each figure presents the original data distribution alongside the results of four SMOTE variants: standard SMOTE, ADASYN, SMOTEENN, and SVMSMOTE.

To enhance the visibility of the class distribution, outliers were removed using a statistical approach. Points lying beyond three standard deviations from the median in either principal component were considered outliers and excluded from the visualization. This process facilitated a clearer observation of the core data structure and the effects of resampling. The original data plots (top row) in both figures show a clear class imbalance, with some classes dominating the distribution while others are underrepresented. The subsequent rows demonstrate how each SMOTE algorithm attempts to balance the class distribution by generating synthetic samples for minority classes. All SMOTE variants successfully increased the representation of minority classes, with the synthetic data generally aligning well with the clusters formed by the original data, indicating that the generated samples are consistent with the underlying class structure in the dataset. However, each algorithm showed slight variations in its approach to synthetic sample generation. For instance, ADASYN tended to focus more on boundary regions between classes, while SMOTEENN produced cleaner separation by removing some overlapping samples. It's worth noting that even though the first two principal components capture the most variance, clear separation among classes is not evident, particularly in the spectral dataset. This could be attributed to the PCA being computed on the entire dataset of spectral time-series, where spectral values change considerably across time as vegetation varies. When applying PCA to time-series data, what we of-

ten see is that the components primarily reflect overall trends or seasonal patterns rather than class-specific differences, making it challenging to distinguish between classes that have similar growth cycles or phenological stages.

In contrast, the PCA on metrics shows better separation, likely because some metrics (i.e. vegetation indices) remain more stable over time, focusing on specific vegetation properties that are more distinctive across different crop types. As a result, the reduced variance in these metrics allows for clearer class separation, highlighting key features more relevant to classification.

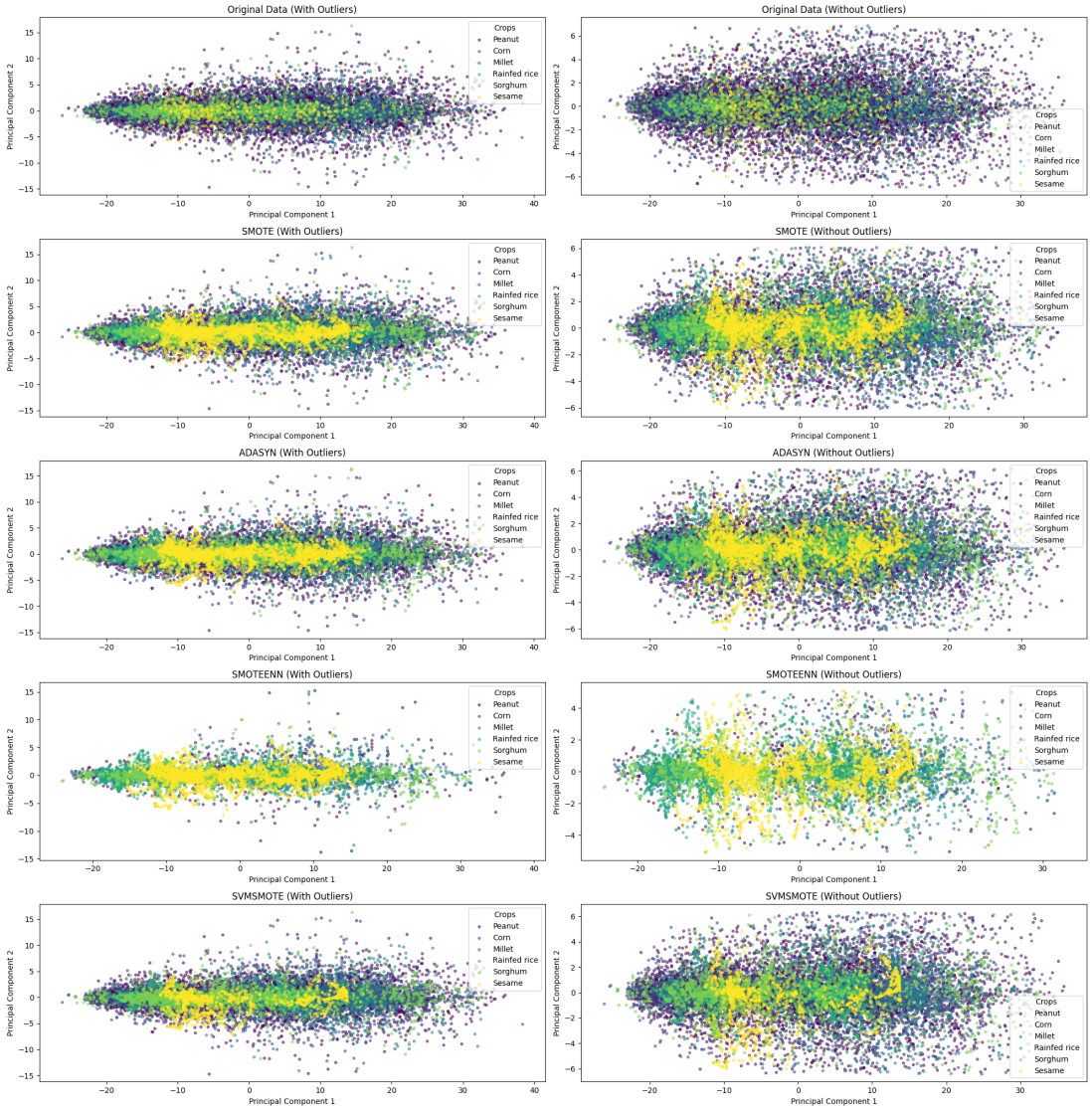


Figure 14: Effect of different SMOTE algorithms on the spectral dataset for crops classes

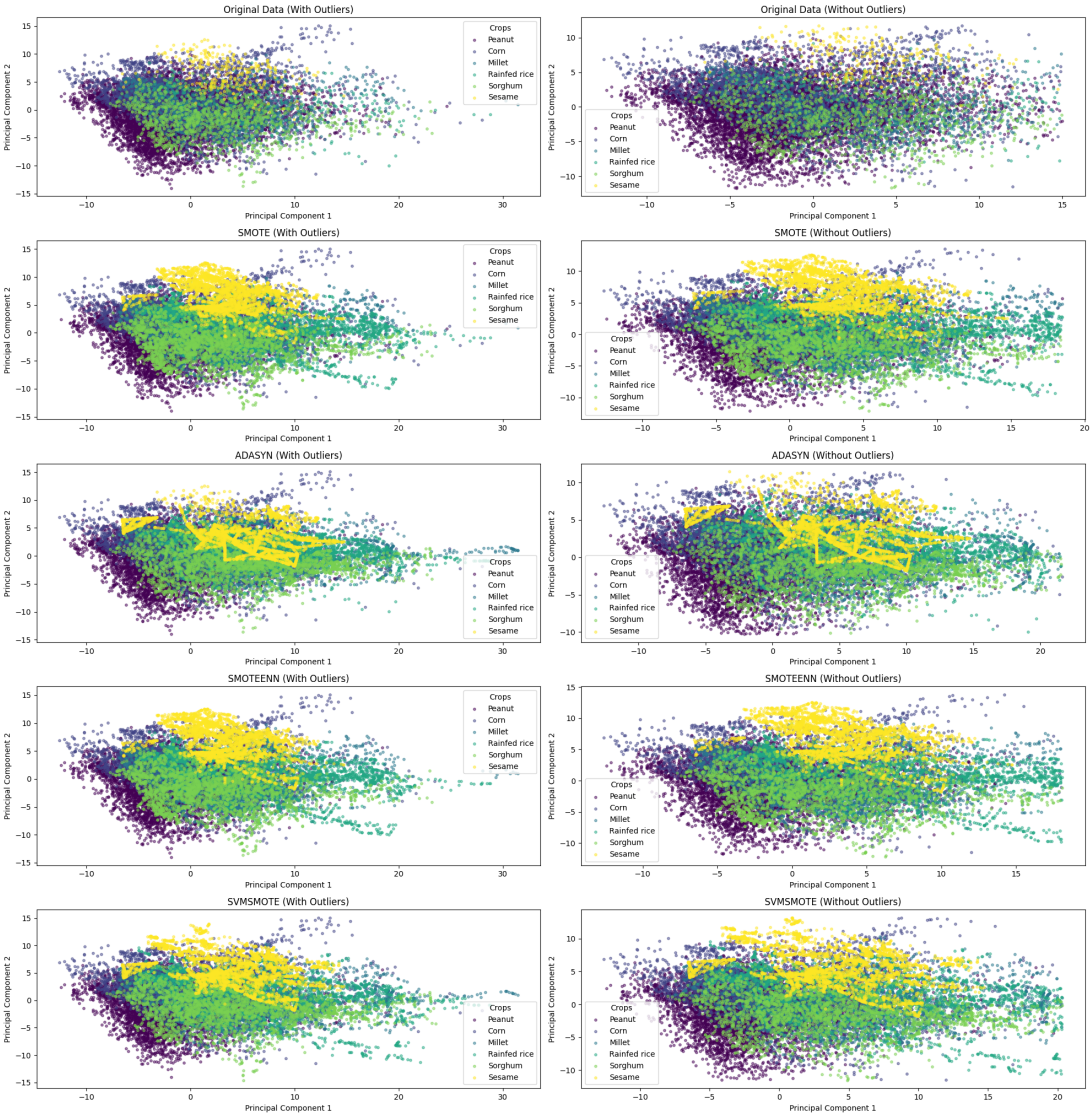


Figure 15: Effect of different SMOTE algorithms on the metrics dataset for crop classes

Figures 17 and 16 illustrate the effects of different SMOTE algorithms on the spectral and metrics datasets, respectively, for non-crop classes. These classes include trees, shrubland, grassland, urban, bare, water, and herbaceous wetland.

Notably, for the metrics dataset of non-crop classes, the separation between different classes is clearly visible. This distinct separation could be attributed to the fact that non-crop classes tend to be more stable over time compared to crop classes. The relative constancy of non-crop land cover types results in more consistent spectral and metric signatures throughout the year, leading to better differentiation in the PCA visualization.

This clear separation in the non-crop classes underscores the potential of using metrics derived from time series data for land cover classification, particularly for distinguishing between different types of non-agricultural land use. The stability of these classes over time appears to translate into more distinct clusters in the feature space, which could contribute to improved classification accuracy for these land cover types.

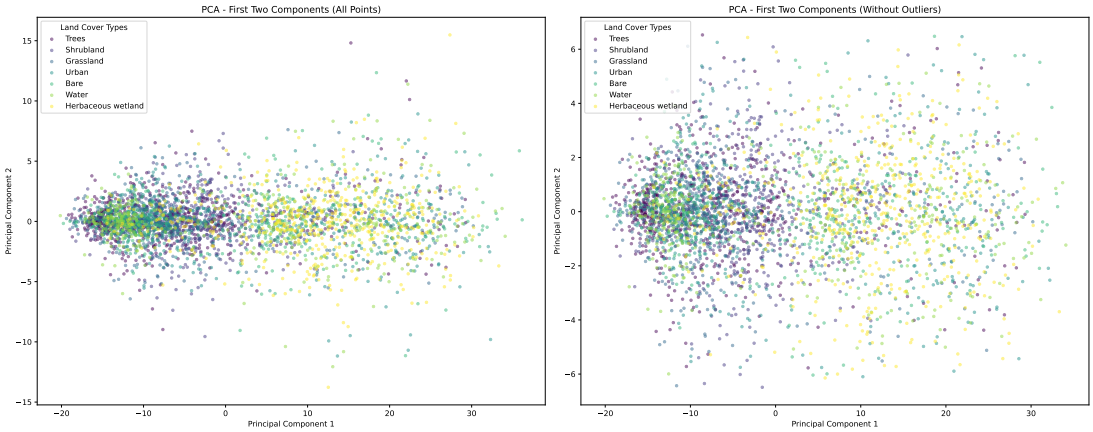


Figure 16: PCA on the spectral dataset for non-crop classes

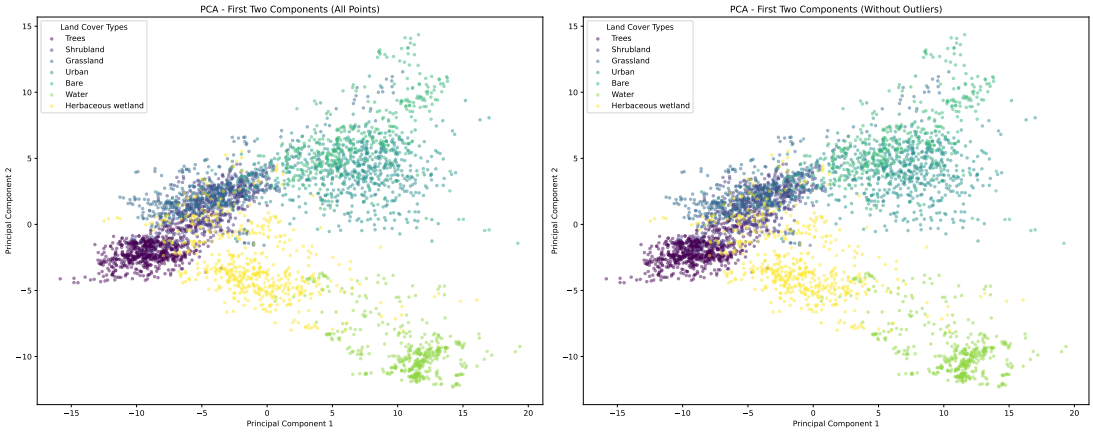


Figure 17: PCA on the metrics dataset for non-crop classes

4.7 Spectral domain

4.7.1 Spectral signature of each crop class

The ability of the classifier to effectively distinguish between different crop types was further investigated through an analysis of their spectral signatures.

Crop plots under the S2 PEV tile for each class was examined. All bands used in previous sections were selected (B02, B03, B04, B05, B06, B07, B08, B11, and B12). Quantitative analysis was carried out through the generation and comparison of spectral curves, histogram and probability density function for each crop class. These provided a statistical representation of the reflectance values across different spectral bands, allowing for a comparison between crop types. The analysis focused on the six crop classes: peanut, maize, millet, rainfed rice, sorghum sesame.

This process involved masking the pre-processed S2 imagery for the 31/10/2023 with the in-situ training data to ensure only relevant pixels were included in the analysis.

4.7.2 Classification without time-series

4.7.2.1 Classification on cloud-free images

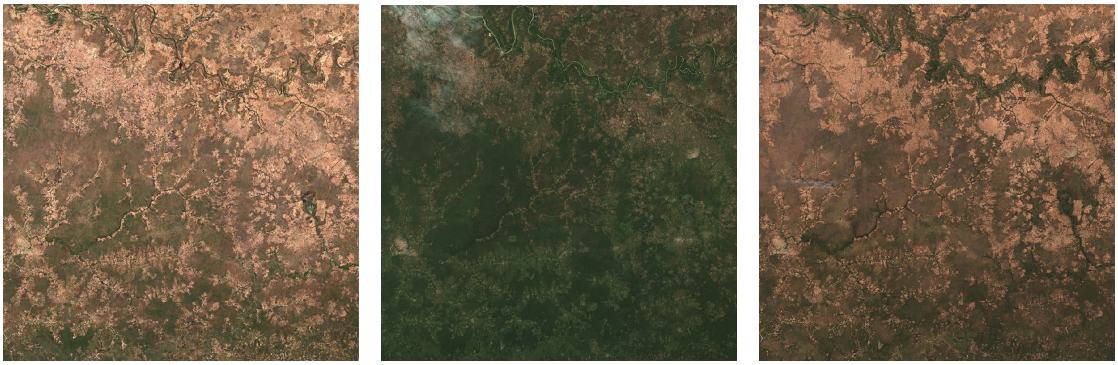
Kolda region experiences significant cloud cover during the rainy season from July to November, posing challenges for using a complete time-series dataset in classification tasks. Utilizing the entire time-series might introduce unnecessary noise, particularly due to the interpolation required during the gap-filling process.

The spectral signatures of crops are expected to differ across various phenological stages during the growing season. It is reasonable to assume that capturing these distinct stages, even with just three images, could provide sufficient information for effective crop differentiationa nd it may be possible to achieve a reliable classification without the need for a complete time series.

Three S2 images were selected based on the best compromise between the good phonological dates and the ones were S2 cloud-free images were available. those dates are June 8, October 31, and December 30, 2023.

Figure 18 displays these three selected S2 images for the PEV tile. It should be noted that the June and December images fall slightly outside the main growing season. However, these were the only cloud-free images available for those periods.

The feature structure for this classification approach remained consistent with that used in the time-series classification. Spectral features comprised 9 bands for each of the 3 selected dates, resulting in a 27-band matrix. The computed metrics remained the same as in the previous approach, including NDVI, BSI, NDVIre2, REP, GNDVI, and EVI.



(a) S2 tile PEV, 08/06 (b) S2 tile PEV, 31/10 (c) S2 tile PEV, 30/12

Figure 18: S2 tile PEV cloud-free images for different dates across the growing season

4.8 Results evaluation

The performance of the classifiers was evaluated using confusion matrices and their derived metrics. Overall accuracy provided a baseline for comparison across different scenarios. To analyze the prediction performance both within and between classes, individual class metrics were examined, including f1-score, recall, and accuracy.

4.9 Scale-up and strategy for prediction on the entire ROI

The ROI covers three S2 tiles, forming a matrix of 9994 by 21720 pixels. Predicting on this full area presents computational challenges, as flattening the data to match the classifier's input format would yield over 217 million pixels multiplied by 22 features, far exceeding available memory and CPU capacity.

To overcome this, a chunking strategy is applied, dividing the ROI into manageable blocks of 2048 by 2048 pixels. For each chunk, relevant data, such as spectral bands and indices, are loaded and processed to match the training format. A mask is then applied to exclude no-data and zero values, leaving only valid pixels (i.e. within the ROI). The features for each chunk are fed into the pre-trained XGBoost classifier for prediction, and the results are stored in a pre-allocated array representing the entire ROI. This process repeats until all chunks are classified, with periodic checkpoints saved for recovery in case of interruptions.

The complexity of this approach, which involved careful memory management, frequent garbage collection, and optimized variable handling after each iteration, demanded significant time and effort. Despite the challenges, the process was executed and once all chunks were processed, the complete classified image is saved as a GeoTIFF file.

Chapter 5 : results

5.1 Time-series classification results

The OA results of the XGBoost classifier for each feature composition scenario are displayed in Table 4. Overall accuracy as a sole performance metric must be interpreted cautiously when evaluating models that combine crop and non-crop classifications. High overall accuracy may be achieved despite misclassification of crop pixels if non-crop pixels, often more numerous, are correctly classified. To draw robust conclusions about model performance, class-specific metrics and confusion matrices should be analyzed in addition to overall accuracy. This approach provides a more comprehensive assessment of the model's effectiveness in distinguishing between different crop types.

Table 4: XGBoost performance for different feature combinations and sampling techniques

Scenario	Features	XGBoost OA
1	Spectral bands + SMOTE	0.6812
2	Metrics + SMOTE	0.6915
3	Spectral bands + metrics + SMOTE	0.7118
4	Spectral bands + ADASYN	0.6596
5	Metrics + ADASYN	0.6885
6	Spectral bands + metrics + ADASYN	0.7000
7	Spectral bands + SMOTEENN	0.6741
8	Metrics + SMOTEENN	0.6900
9	**Spectral bands + metrics + SMOTEENN**	**0.7337**

5.1.1 Classification with SMOTE algorithm

Table 5: Performance comparison of different feature combinations with SMOTE

Features	Precision	Recall	F1-score
Spectral bands + SMOTE	0.66	0.68	0.65
Metrics + SMOTE	0.67	0.69	0.68
Spectral bands + metrics + SMOTE	**0.69**	**0.71**	**0.69**

Table 5 presents the performance comparison of different feature combinations using the SMOTE algorithm for balancing the dataset. The results indicate that the combination of spectral bands and computed metrics yielded the best performance for the model, with precision, recall, and F1-score of 0.69, 0.71, and 0.69 respectively. However, it

Table 6: Confusion matrix for SC3_spectral_metrics_smote_strategy

Predicted class / True class	Peanut	Corn	Millet	Rainfed rice	Sorghum	Sesame	Trees	Shrubland	Grassland	Urban	Bare	Water	Herbaceous W.	Omission
Peanut	**2485**	52	18	30	9	0	0	0	10	0	0	0	0	4.55%
Corn	548	**491**	97	46	59	1	4	1	1	1	0	0	0	61.24%
Millet	7	253	**409**	15	76	0	0	0	0	0	0	0	0	45.74%
Rainfed rice	187	45	4	**29**	5	13	0	0	0	0	0	0	0	89.68%
Sorghum	5	147	51	22	**224**	0	2	8	10	0	0	0	1	53.14%
Sesame	0	32	76	0	0	**94**	0	0	0	0	0	0	0	53.47%
Trees	2	0	0	2	0	0	**252**	1	1	0	0	0	0	2.33%
Shrubland	0	7	0	0	0	0	13	**141**	7	0	0	0	0	16.07%
Grassland	15	5	0	1	4	0	0	0	**152**	0	0	0	1	14.61%
Urban	2	3	0	4	2	0	0	0	0	**170**	0	0	0	6.08%
Bare	0	3	2	0	0	4	0	0	12	16	**102**	0	0	26.62%
Water	0	0	0	0	0	0	0	0	0	0	0	**141**	4	2.76%
Herbaceous Wetland	0	0	0	0	0	0	1	0	6	0	0	11	**164**	10.38%
Commission Error	23.47%	52.25%	37.76%	80.54%	40.67%	16.07%	7.35%	6.62%	23.62%	9.09%	0.00%	7.89%	4.12%	

is important to note that the performance differences between the feature combinations are not substantial.

It can also be observed that the differences between precision and recall values for each feature combination are minimal. This balance suggests that the model performs consistently in terms of both correctly identifying positive instances (precision) and capturing a high proportion of actual positive instances (recall) across different classes. Such equilibrium is desirable in classification tasks with over-prediction and under-prediction. However, it should be noted that the omission and commission error rarely compensate for each other. For instance, analysis of omission and commission errors of the third scenario (table 6) reveals that these errors do not generally compensate for each other, particularly for crop classes. For instance, peanut shows a relatively low omission error (4.55%) but a higher commission error (23.47%), indicating some over-prediction of this class. Conversely, rainfed rice exhibits both a very high omission error (89.68%) and a high commission error (80.54%), suggesting severe classification difficulties for this crop. Corn also shows significant imbalance with high omission (61.24%) and commission (52.25%) errors. This lack of compensation between omission and commission errors points to systematic biases in the model's predictions, rather than random misclassifications. Non-crop classes generally show better balance between these errors, with trees (2.33% omission, 7.35% commission) and water (2.76% omission, 7.89% commission) demonstrating particularly reliable classification. This pattern indicates that while the model performs well for non-crop categories, it struggles with accurately distinguishing between different crop types

The overall performance scores (all below 0.70 for F1-score) reflects spectral similarities between certain crop types or the influence of factors such as varying growth stages and environmental conditions, it will be investigate in further section.

The classification results reveal varying performance across different crop and land cover classes. Non-crop classes (Trees, Shrubland, Grassland, Urban, Bare, Water, and Herbaceous Wetland) consistently show high performance metrics (F1-scores >0.80) across all feature combinations, indicating robust classification for these categories.

Table 7: Performance comparison of different feature combinations with SMOTE across various classes

Classes	Spectral bands + SMOTE			Metrics + SMOTE			Spectral bands + metrics + SMOTE			Samples
	Precision	Recall	F1-score	Precision	Recall	F1-score	Precision	Recall	F1-score	
Peanut	0.71	**0.98**	0.83	**0.81**	0.93	**0.86**	0.76	0.95	0.85	2604
Corn	0.39	0.32	0.35	0.43	0.37	0.40	**0.47**	**0.39**	**0.43**	1249
Millet	**0.60**	0.32	0.42	0.46	**0.58**	0.51	**0.62**	0.54	**0.58**	760
Rainfed rice	0.18	0.08	0.12	0.13	0.07	0.09	**0.19**	**0.10**	**0.13**	283
Sorghum	**0.69**	**0.49**	**0.58**	0.60	0.41	0.49	0.59	0.48	0.53	470
Sesame	0.83	0.40	0.54	0.70	**0.48**	0.56	**0.84**	0.47	**0.60**	202
Trees	0.91	0.97	0.94	0.92	0.95	0.93	**0.93**	**0.98**	**0.95**	258
Shrubland	0.86	0.74	0.80	0.88	0.76	0.82	**0.93**	**0.84**	**0.88**	168
Grassland	**0.85**	**0.85**	**0.85**	0.75	0.71	0.73	0.76	**0.85**	0.81	178
Urban	0.85	**0.96**	0.90	**0.91**	0.90	0.90	**0.91**	0.94	**0.92**	181
Bare	**1.00**	0.71	0.83	0.98	**0.83**	**0.90**	**1.00**	0.73	0.85	139
Water	**0.95**	0.99	**0.97**	0.94	**1.00**	**0.97**	0.93	0.97	0.95	145
Herbaceous Wetland	**0.99**	**0.90**	**0.94**	0.98	0.87	0.92	0.96	**0.90**	0.93	182

Among crop classes, peanut demonstrates the best performance, with F1-scores above 0.83 for all feature combinations. However, other crop types show lower classification accuracy. Notably, rainfed rice exhibits the poorest performance, with F1-scores below 0.13 across all scenarios, suggesting significant challenges in its identification.

Table 8: Accuracy comparison for crop and non-crop classes across different SMOTE models

Model	**Crop Classes**	**Non-Crop Classes**
SC1_spectral_SMOTE	63.58%	88.33%
SC2_metrics_SMOTE	65.30%	86.25%
SC3_spectral_metrics_SMOTE	67.03%	89.69%

It can clearly be noted that the non-crop classes performed better that the crop ones, meaning that the problem is probably inherent to the data for crop classes and not inherent to the model. Further investigation will be discussed in further section.

5.1.2 Classification with ADASYN algorithm

Table 9: Performance comparison of different feature combinations with ADASYN

Features	**Precision**	**Recall**	**F1-score**
Spectral bands + ADASYN	0.63	0.66	0.63
Metrics + ADASYN	0.67	0.69	0.68
Spectral bands + metrics + ADASYN	0.68	0.70	0.68

The ADASYN resampling algorithm shows similar trends to those observed with SMOTE, albeit with slightly lower overall performance. The combination of spectral bands and

metrics again yields the best results, though the improvement over using metrics alone is marginal. Notably, the spectral bands alone perform worse with ADASYN compared to SMOTE, suggesting that ADASYN might be less effective in handling the spectral data's complexity.

Table 10: Performance comparison of different feature combinations with ADASYN across various classes

Classes	Spectral bands + ADASYN			Metrics + ADASYN			Spectral bands + metrics + ADASYN			Samples
	Precision	Recall	F1-score	Precision	Recall	F1-score	Precision	Recall	F1-score	
Peanut	0.73	**0.96**	0.83	**0.81**	0.92	**0.86**	0.76	0.95	0.84	2604
Corn	0.35	0.36	0.35	**0.46**	0.38	0.42	**0.46**	**0.44**	**0.45**	1249
Millet	0.47	0.23	0.31	0.48	**0.57**	0.52	**0.60**	0.49	**0.54**	760
Rainfed rice	0.20	0.11	0.15	0.12	0.10	0.11	**0.21**	**0.14**	**0.17**	283
Sorghum	0.58	0.31	0.40	0.48	**0.37**	0.42	**0.66**	0.36	**0.47**	470
Sesame	0.68	0.40	0.50	**0.73**	**0.53**	**0.61**	0.71	0.36	0.47	202
Trees	**0.95**	**0.97**	**0.96**	0.94	**0.97**	0.95	0.91	**0.97**	0.94	258
Shrubland	0.82	**0.82**	0.82	0.87	0.73	0.79	**0.98**	0.77	**0.87**	168
Grassland	0.83	0.81	0.82	0.73	0.65	0.69	**0.87**	**0.85**	**0.86**	178
Urban	0.85	**0.94**	0.90	**0.92**	0.91	**0.92**	0.83	0.93	0.88	181
Bare	**1.00**	0.70	0.82	0.93	**0.82**	**0.87**	**1.00**	0.74	0.85	139
Water	0.94	0.97	0.96	0.95	**1.00**	**0.98**	0.94	0.98	0.96	145
Herbaceous Wetland	0.96	**0.95**	**0.95**	**0.98**	0.89	0.93	0.96	0.92	0.94	182

The ADASYN results largely mirror the patterns observed with SMOTE, with some minor differences. Non-crop classes continue to show high performance (F1-scores >0.80), while crop classes, especially rainfed rice, corn, and millet remain challenging to classify accurately. The accuracy comparison (Table 11) reinforces the significant disparity between crop and non-crop classification performance, consistent with the SMOTE results. Peanut maintains high performance across all feature combinations, similar to the SMOTE results. These observations suggest that while ADASYN provides comparable results to SMOTE, it may be slightly less effective for this particular dataset and classification task, especially when using spectral bands alone.

Table 11: Accuracy comparison for crop and non-crop classes across different ADASYN models

Model	Crop Classes	Non-Crop Classes
SC2_spectral_ADASYN	60.79%	89.00%
SC2_metrics_ADASYN	65.05%	85.77%
SC2_spectral_metrics_ADASYN	65.72%	89.05%

5.1.3 Classification with SMOTEENN algorithm

The SMOTEENN algorithm shows overall improved performance compared to SMOTE and ADASYN. The combination of spectral bands and metrics yields the highest scores (Precision: 0.71, Recall: 0.73, F1-score: 0.71), outperforming both SMOTE and ADASYN

Table 12: Performance comparison of different feature combinations with SMOTEENN

Features	Precision	Recall	F1-score
Spectral bands + SMOTEENN	0.65	0.67	0.65
Metrics + SMOTEENN	0.67	0.69	0.67
Spectral bands + metrics + SMOTEENN	0.71	0.73	0.71

Table 13: Confusion matrix for SC9_spectral_metrics_smoteENN_strategy

True class \ Predicted class	Peanut	Corn	Millet	Rainfed rice	Sorghum	Sesame	Trees	Shrubland	Grassland	Urban	Bare	Water	Herbaceous W.	Omission
Peanut	**2519**	35	15	16	4	2	0	0	13	0	0	0	0	3.26%
Corn	506	**559**	65	65	42	2	3	0	5	2	0	0	0	55.24%
Millet	5	269	**393**	15	78	0	0	0	0	0	0	0	0	48.29%
Rainfed rice	197	47	4	**29**	1	5	0	0	0	0	0	0	0	89.75%
Sorghum	7	119	37	23	**258**	0	3	8	14	0	0	0	1	45.11%
Sesame	1	34	76	0	0	**91**	0	0	0	0	0	0	0	54.95%
Trees	3	0	0	2	0	1	**252**	0	0	0	0	0	0	2.33%
Shrubland	0	3	2	0	0	0	21	**135**	7	0	0	0	0	19.64%
Grassland	17	3	1	1	4	0	0	0	**151**	0	0	0	1	15.17%
Urban	0	7	0	2	3	0	0	0	0	**169**	0	0	0	6.63%
Bare	1	3	1	0	0	3	0	0	3	12	**116**	0	0	16.55%
Water	0	0	0	0	0	0	0	0	0	0	0	**145**	0	0.00%
Herbaceous Wetland	0	0	0	0	0	0	5	1	1	0	0	11	**164**	9.89%
Commission Error	22.61%	48.41%	33.86%	81.05%	33.59%	12.50%	11.41%	6.25%	22.22%	7.65%	0.00%	7.05%	1.81%	

in this configuration. This suggests that SMOTEENN's approach of combining oversampling with noise reduction is particularly effective for this dataset.

The confusion matrix for scenario 9 reveals some interesting patterns and improvements compared to scenario 3. The performance for peanut classification has improved slightly, with the omission error decreasing from 4.55% to 3.26% and the commission error decreasing from 23.47% to 22.61%, indicating better overall accuracy for this class. There's also a notable improvement in corn classification, with the omission error decreasing from 61.24% to 55.24% and the commission error improving from 52.25% to 48.41%. This suggests that the SMOTEENN strategy has helped in better distinguishing corn from other crops. However, rainfed rice continues to be challenging, with only a marginal improvement. The omission error slightly increased from 89.68% to 89.75%, while the commission error remained very high at 81.05% (compared to 80.54% previously). This indicates that rainfed rice remains difficult to classify accurately. On the other hand, there's a significant improvement in sorghum classification, with the omission error decreasing from 53.14% to 45.11% and the commission error improving from 40.67% to 33.59%, suggesting that the model is now better at identifying sorghum. The performance for non-crop classes remains generally good, with some variations. Trees maintain the same low omission error (2.33%) but show a slight increase in commission error (from 7.35% to 11.41%). Water classification has improved, with the omission error dropping to 0% (from 2.76%) and a slight improvement in commission error (7.89% to 7.05%). Urban areas show a slight increase in omission error (from 6.08% to 6.63%) but an improvement in commission error (9.09% to 7.65%). All confusion matrices are available in the annexes.

Table 14: Performance comparison of different feature combinations with SMOTEENN across various classes

Classes	Spectral bands + SMOTEENN			Metrics + SMOTEENN			Spectral bands + metrics + SMOTEENN			Samples
	Precision	Recall	F1-score	Precision	Recall	F1-score	Precision	Recall	F1-score	
Peanut	0.75	0.96	0.84	**0.81**	0.93	**0.87**	0.77	**0.97**	0.86	2604
Corn	0.39	0.36	0.37	0.43	0.38	0.40	**0.52**	**0.45**	**0.48**	1249
Millet	0.60	0.36	0.45	0.45	**0.53**	0.49	**0.66**	0.52	**0.58**	760
Rainfed rice	0.14	**0.10**	0.11	0.13	0.07	0.09	**0.19**	**0.10**	**0.13**	283
Sorghum	0.51	0.36	0.42	0.58	0.43	0.50	**0.66**	**0.55**	**0.60**	470
Sesame	**0.89**	0.39	0.54	0.67	**0.48**	0.55	0.88	0.45	**0.59**	202
Trees	0.93	**0.98**	**0.95**	**0.94**	0.95	0.94	0.89	**0.98**	0.93	258
Shrubland	0.69	0.76	0.72	0.91	0.79	0.85	**0.94**	**0.80**	**0.87**	168
Grassland	**0.84**	**0.84**	**0.84**	0.72	0.65	0.68	0.78	**0.85**	0.81	178
Urban	0.84	**0.94**	0.89	**0.92**	0.91	0.91	**0.92**	0.93	**0.93**	181
Bare	**1.00**	0.72	0.84	0.92	0.82	0.87	**1.00**	**0.83**	**0.91**	139
Water	0.93	**1.00**	**0.96**	0.93	**1.00**	**0.96**	0.93	**1.00**	**0.96**	145
Herbaceous Wetland	**1.00**	0.86	0.92	0.99	0.87	0.92	0.99	**0.90**	**0.94**	182

Class-specific performance with SMOTEENN follows similar patterns to SMOTE and ADASYN, with non-crop classes consistently outperforming crop classes. However, SMOTEENN shows slight improvements in F1-scores for several crop classes, particularly when using both spectral bands and metrics. Rainfed rice remains the most challenging crop to classify accurately across all feature combinations.

Table 15: Accuracy comparison for crop and non-crop classes across different SMOTEENN models

Model	Crop Classes	Non-Crop Classes
SC3_spectral_SMOTEENN	62.81%	87.93%
SC3_metrics_SMOTEENN	64.92%	85.93%
SC3_spectral_metrics_SMOTEENN	69.13%	90.49%

The accuracy comparison (Table 15) reveals that SMOTEENN achieves the highest accuracy for both crop (69.13%) and non-crop (90.49%) classes when combining spectral bands and metrics. However, it can be seen that the accuracy for crop classes are still below 70% and are still far below non-crop classes. An analysis on crops data in further section will try to explain this result.

5.1.4 Prediction over the entire ROI

The best-performing model among those previously presented was selected to generate a prediction over the entire ROI (figure 19). The resulting classification map reveals several noteworthy aspects. The structure of crop fields is clearly visible in the prediction, which aligns well with expected agricultural patterns in the region, suggesting that the model has successfully captured some of the spatial characteristics of crop distribution. However, those results have to be interpreted in light of the classification performance

metrics discussed earlier. The model's limitations, particularly in distinguishing certain crop types, should be considered when analyzing the prediction map.

Furthermore the classification appears to overestimate the presence of certain crop types, with areas that visually look like shrubland or grassland often classified as crop types such as sesame, corn, sorghum, or millet. This tendency towards overestimation of crops highlights the challenges in accurately differentiating between crops and natural vegetation in this landscape, this point will be discussed in later section. These observations underscore both the potential and the limitations of the current classification approach.

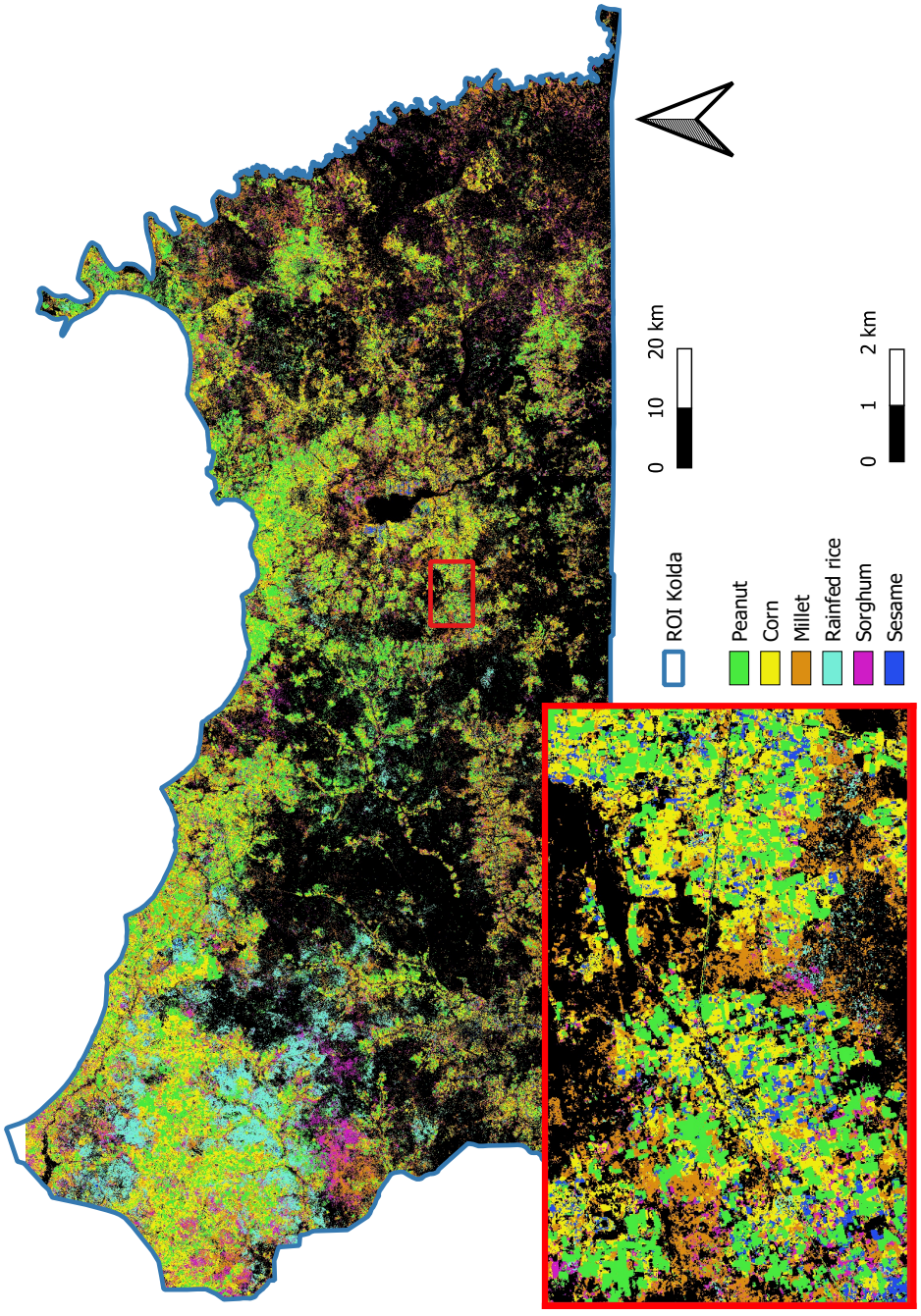


Figure 19: Prediction over the entire ROI with the model trained with scenario 9 with non-crop classes masked

5.1.5 Classification with Sen4Stats cropland mask

The application of the Sen4Stats cropland mask in combination with spectral bands, metrics, and the SMOTEENN algorithm was evaluated for crop classification. Table 16 presents the overall performance metrics, which indicate moderate effectiveness with precision, recall, and F1-score values of approximately 0.68-0.69. The accuracy of 0.69 is comparable to the previously discussed SMOTEENN model (0.71) that utilized both spectral bands and metrics, especially if we look at the accuracy of 0.69 for crop classes for this model (14).

Table 16: Performance comparison for crop classes with cropland mask

Features	**Precision**	**Recall**	**F1-score**	**Accuracy**
Spectral bands + metrics + SMOTEENN	0.68	0.69	0.67	0.69

Analysis of individual crop classes (Table 17) reveals variations in classification performance across different crop types. Peanut demonstrates higher performance metrics with precision (0.85), recall (0.99), and F1-score (0.91). This represents a slight increase from the SMOTEENN model's F1-score of 0.86 for peanut. Millet, rainfed rice and sesame exhibit lower performance, with F1-scores ranging from 0.35 to 0.43. Rainfed rice presents a case of low precision (0.24) but higher recall (0.66), indicating a tendency for comission errors. This imbalance suggests that while the model identifies many instances of rainfed rice, it also incorrectly classifies many non-rainfed rice samples as this class. Sesame shows low recall (0.31), implying frequent misclassification as other crop types.

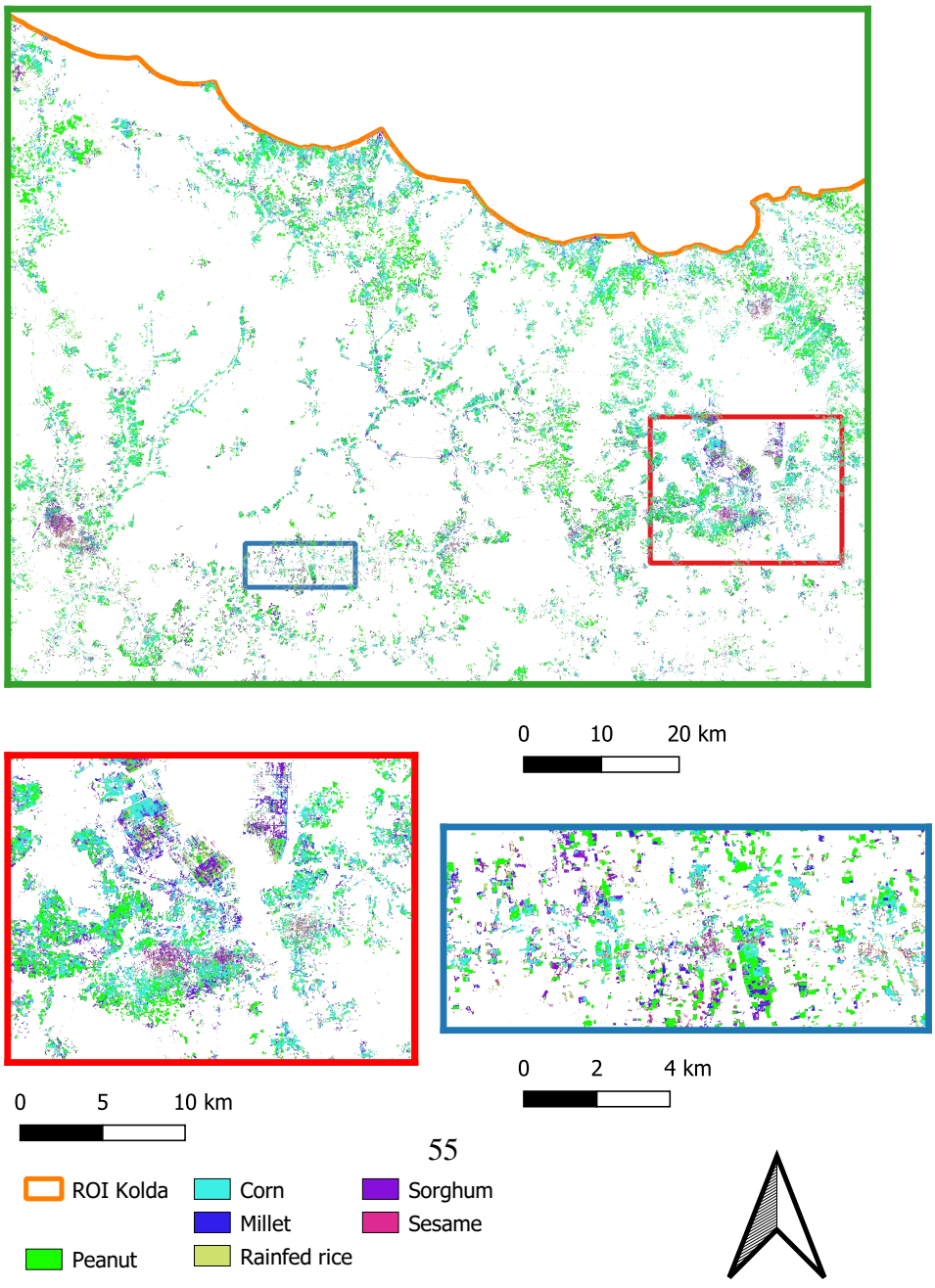
In comparison to the SMOTEENN model using spectral bands and metrics with time-series, the cropland mask approach does not significantly enhance the classification.

It's important to note that these metrics are contingent on the accuracy of the cropland mask. The in-situ data used for this analysis were limited to areas where the cropland mask indicated valid crop values. Consequently, if the cropland mask itself is inaccurate, it could lead to a biased selection of in-situ data points, potentially affecting the reliability of the classification results.

Figure 20 shows the results of the classification over the PEV tile. 2 areas have been zoomed out to better visualize the classification pattern.

Table 17: Classification performance metrics by class for classification with cropland mask

Class	Precision	Recall	F1-Score	Support
Peanut	0.85	0.99	0.91	1597
Corn	0.60	0.40	0.48	908
Millet	0.46	0.40	0.43	226
Rainfed rice	0.24	0.66	0.35	47
Sorghum	0.46	0.55	0.50	398
Sesame	0.54	0.31	0.40	199



5.2 Qualitative assessment of class separability in the spectral domain

5.2.1 Spectral signature of each crop classes

5.2.1.1 Spectral curves from in-situ data

The S2 spectral signatures of different crop classes were analyzed using various visualization techniques by analysing pixels values for in-situ polygons for each crop class at the date of the 31 october 2023 over the PEV tile 18b. Figure 21 illustrates the 95th percentile value of each spectral band for each crop class, along with the standard deviation. This approach (i.e. taking the 95th percentile) was adopted to mitigate the influence of potential outliers while having better discrimination than the mean values. Analysis of these spectral signatures reveals that bands 4, 6, 7, 8, 11, and 12 exhibit the most pronounced differences between classes, suggesting their potential significance in crop discrimination.

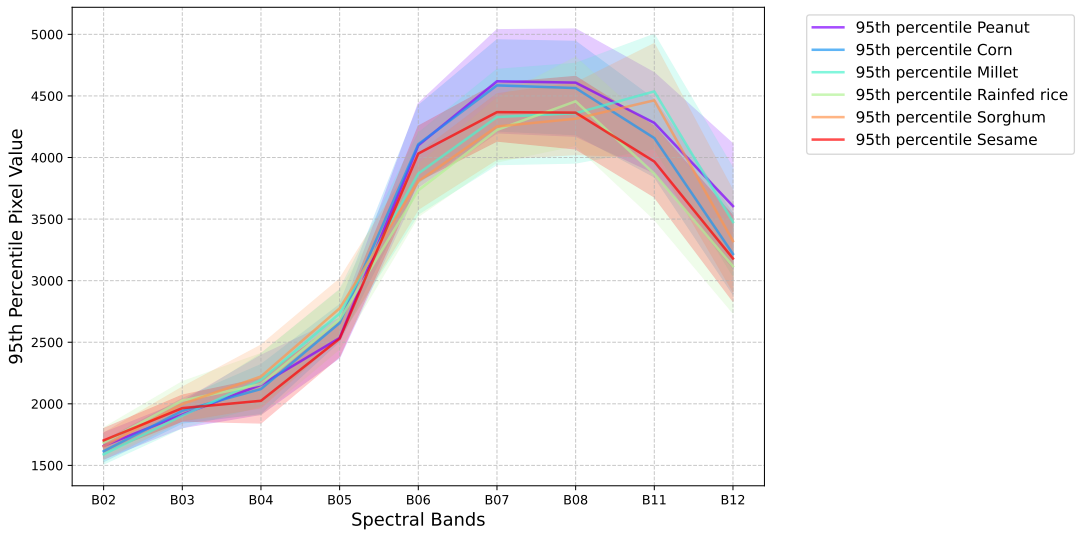


Figure 21: Spectral signatures curves with 95percentile values and standard deviation for crops classes

Further investigation of spectral properties was conducted through probability density functions of spectral signatures for each class, as depicted in Figure 22. Bands 6 and 7 demonstrate particular efficacy in distinguishing millet and sesame classes. Although the sorghum class exhibits higher peaks in these bands, the presence of other classes with similar pixel values may pose challenges for classification algorithms. Bands 2 and 3 show significant overlap between classes, corroborating observations from Figure

21 and indicating limited utility for class differentiation. Despite overlapping in bands 4, 5, 11, and 12, these bands may prove valuable in discriminating the corn class due to its slightly offset peak. Notably, the peanut class peak consistently coincides with either sesame or millet across all bands. The only classes presenting isolated peaks across various bands are millet (bands 8, 6, 7) and corn (bands 11, 12, 5, and 4).

It is noteworthy that the rainfed rice class does not exhibit any isolated peaks, which may account for its poor classification performance. Similarly, the consistently similar pixel values across all bands for corn and sorghum classes could explain their suboptimal classification results. These observations underscore the challenges in spectral-based discrimination of certain crop types and highlight the potential limitations of relying solely on spectral information for accurate crop classification.

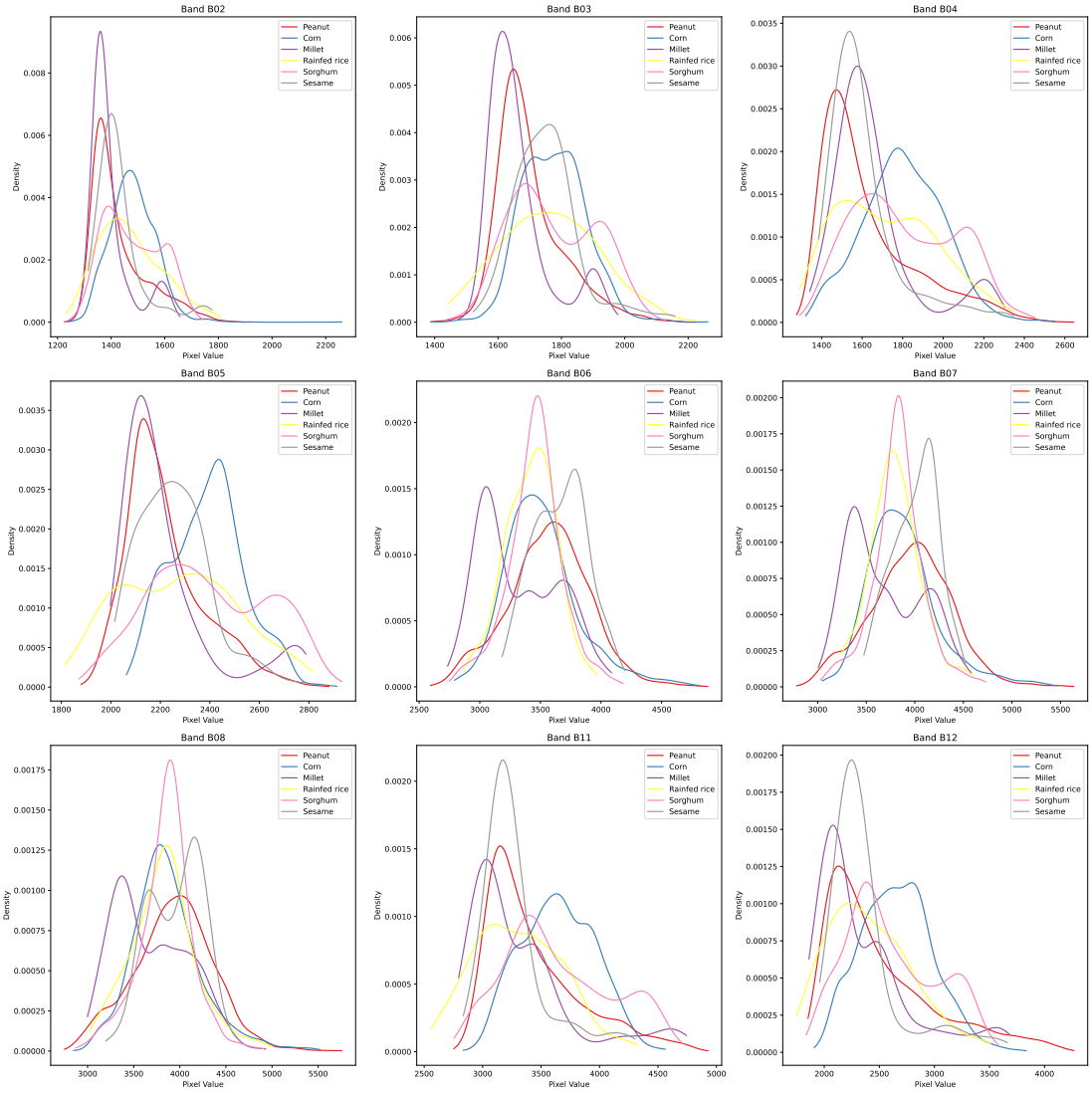


Figure 22: Spectral signatures probability density function for crop classes

Histograms of spectral band values for crop classes, presented in Figure 23, provide additional insights. While confirming the high values of corn in bands 12 and 4, they also reveal similarly high values for the sorghum class in these bands. The previously noted discrimination of millet in band 8 is evident, albeit with subtle differences that may be challenging for classifiers to detect amidst large datasets. Comparison of crop class spectral signatures with those of non-crop classes (Figure 24) generally indicates distinct patterns, with the exception of trees, shrubland, and grassland, which exhibit relatively similar spectral characteristics to crop classes.

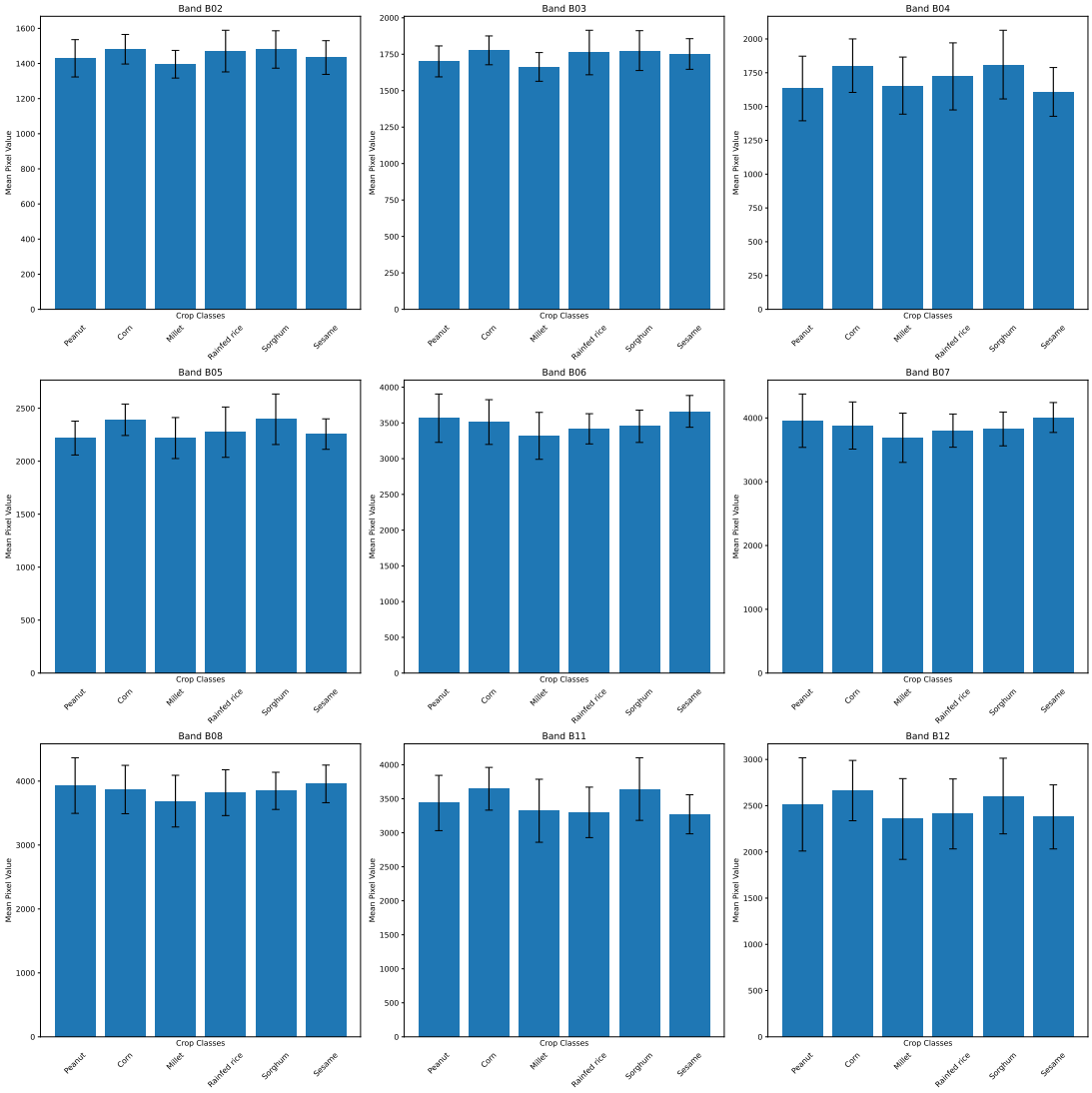


Figure 23: Spectral values histogram for each spectral band and each crop class

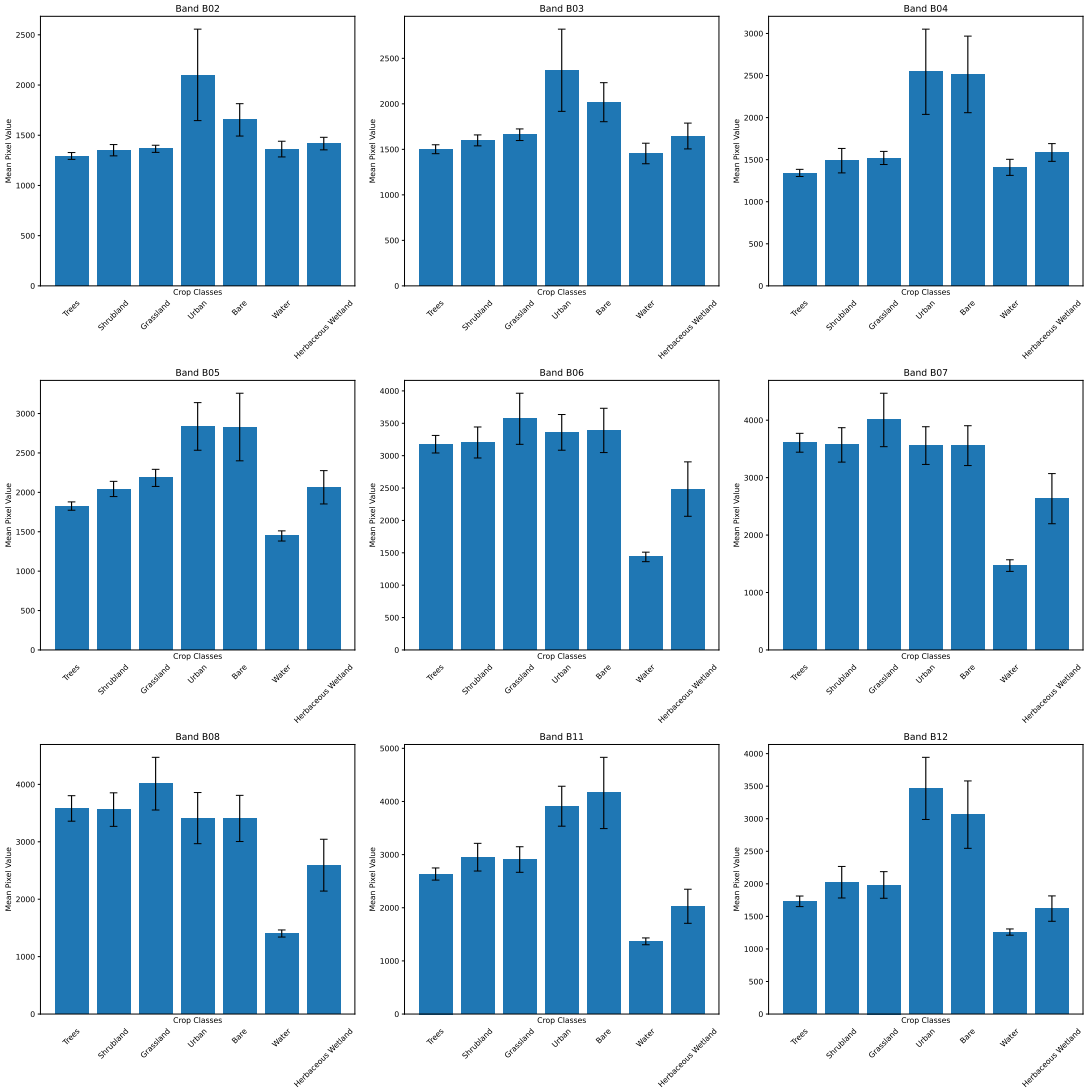


Figure 24: Spectral values histogram for each spectral band and each non-crop class

5.2.1.2 In-situ campaign observation

The field campaign conducted in October 2023 revealed variability within and among different crop types, as illustrated in Figure 25. The peanut plots, shown in Figure 25a, demonstrate a high presence of adventitious plants. This mixed vegetation composition add noise to the peanut fields, potentially leading to misclassification. In cases where the classifier is uncertain, it may erroneously categorize these plots as grassland or other vegetation types due to the presence of non-crop plants as the vegetation signature is not related to one crop ony which would be the ideal case.

Corn fields exhibit considerable heterogeneity, as evidenced in Figures 25c and 25d. The plot in Figure 25c shows only a few corn plants, with the majority of the area covered by grass, adventitious species and bare soil. This sparse corn coverage means that the spectral signature of this plot will be dominated by non-crop vegetation, potentially leading to misclassification. Furthermore, Figure 25d depicts a field where the corn has already been harvested, highlighting the temporal variability within the maize class. This temporal inconsistency adds another layer of complexity to the classification process, as some October imagery will capture standing corn while other images will show harvested fields, both labeled as maize.

In contrast, the sorghum plot (Figures 25e and 25f) appears dense and uniform for that plot, suggesting that its spectral signature should be more distinct and easier for the classifier to recognize accurately.

These images underscore the challenges in crop classification due to intra-class variability. For instance, an arachide plot without adventitious plants may have a significantly different spectral signature compared to one with trees and weeds, despite both being classified as arachide. This variability extends to other factors such as planting density, crop health, and harvest timing, all of which influence the spectral characteristics of the fields.



(a) Peanut



(b) Sesame



(c) Corn (plot 1)



(d) Corn (plot 2)



(e) Sorghum



(f) Sorghum (different angle)

Figure 25: Crop plots visited during the field campaign in October 2023

5.2.2 Classification without time-series

5.2.2.1 Classification on cloud-free images

The classification utilizing few cloud-free images at specific dates demonstrates promising results, with overall performance metrics marginally surpassing those achieved by the optimal set of features with SMOTEENN model. As illustrated in Table 18, the cloud-free classification attains a precision of 0.73, recall of 0.74, and F1-score of 0.73, compared to SMOTEENN's best performance of 0.71, 0.73, and 0.71 respectively. This modest improvement suggests that the utilization of only three cloud-free images may enhance overall classification performance. The enhancement is likely attributable to reduced noise in the dataset. The majority of images in the original dataset exhibit significant cloud cover, necessitating interpolation during the gap-filling process. This interpolation, particularly when applied over extended periods between valid pixels, may introduce values that lack physical significance. Consequently, the use of solely cloud-free images potentially mitigates these data quality issues, resulting in a more robust classification outcome.

Table 18: Performance for the classification on cloud-free images

Features	**Precision**	**Recall**	**F1-score**
Spectral bands + metrics	0.73	0.74	0.73

The confusion matrix reveals some interesting patterns when compared to the scenario 9 with SMOTEENN. Overall, there are notable improvements in certain classes, while others show decreased performance. Peanut classification has shown an improvement, with the omission error decreasing from 3.26% to just 0.99%. This suggests that this strategy is better at identifying peanut crops. Similarly, the commission error for peanuts has decreased from 22.61% to 19.27%, indicating fewer false positives. Corn classification has also improved, with the omission error decreasing from 55.24% to 52.20%. However, the commission error has decreased from 48.41% to 36.39%, suggesting a better balance between precision and recall for this class. Interestingly, rainfed rice shows an improvement, with the omission error decreasing from 89.75% to 55.32%. This is a significant enhancement in the model's ability to correctly identify rainfed rice, which was previously one of the most challenging classes. The commission error for rainfed rice has also improved, decreasing from 81.05% to 68.89%, though it remains relatively high. Sorghum classification shows a slight decline in performance, with the omission error increasing from 45.11% to 57.66%. However, the commission error has improved from 33.59% to 45.83%, indicating a trade-off between precision and recall for this class. Sesame classification has worsened, with the omission error increasing from 54.95% to 79.21%. This suggests that the classification with cloud-free images struggles more with correctly identifying sesame crops. Non-crop classes show mixed

Table 19: Confusion matrix for the classification on few clouds free images

Predicted class / True class	Peanut	Corn	Millet	Rainfed rice	Sorghum	Sesame	Trees	Shrubland	Grassland	Urban	Bare	Water	Herbaceous W.	Omission
Peanut	**1719**	2	0	5	6	0	0	4	0	0	0	0	0	0.99%
Corn	308	**597**	4	1	85	2	0	0	0	0	0	0	0	52.20%
Millet	1	119	**83**	16	25	0	0	0	0	0	0	0	0	66.12%
Rainfed rice	10	14	0	**21**	0	2	0	0	0	0	0	0	0	55.32%
Sorghum	19	169	30	21	**199**	6	2	4	20	0	0	0	0	57.66%
Sesame	42	30	75	0	13	**42**	0	0	0	0	0	0	0	79.21%
Trees	0	0	0	1	2	0	**121**	7	0	0	0	0	0	7.63%
Shrubland	0	0	0	0	11	0	0	**64**	0	0	0	0	0	14.67%
Grassland	14	1	6	0	8	0	0	0	**14**	0	0	0	0	67.44%
Urban	7	0	0	1	3	1	0	0	0	**152**	0	0	0	7.32%
Bare	4	5	1	0	7	0	0	0	0	8	**114**	0	0	17.99%
Water	0	0	0	0	0	0	0	0	0	0	0	**25**	0	0.00%
Herbaceous Wetland	0	0	0	0	7	0	0	0	1	0	0	0	**134**	5.63%
Commission Error	19.27%	36.39%	58.78%	68.89%	45.83%	20.75%	1.63%	18.99%	60.00%	5.00%	0.00%	0.00%	0.00%	

results. Trees and shrubland have slightly higher omission errors (7.63% and 14.67% respectively, compared to 2.33% and 19.64% in the SMOTEENN strategy). However, urban areas and water bodies show improved classification with lower omission errors.

Table 20: Performance metrics by class for the classification on cloud-free images

Class	Precision	Recall	F1-score	Samples
Peanut	0.81	0.99	0.89	1736
Corn	0.64	0.60	0.62	997
Millet	0.42	0.34	0.37	244
Rainfed rice	0.32	0.45	0.37	47
Sorghum	0.54	0.42	0.48	470
Sesame	0.79	0.21	0.33	202
Trees	0.98	0.92	0.95	131
Shrubland	0.81	0.85	0.83	75
Grassland	0.40	0.33	0.36	43
Urban	0.95	0.93	0.94	164
Bare	1.00	0.82	0.90	139
Water	1.00	1.00	1.00	25
Herbaceous Wetland	1.00	0.94	0.97	142

Examining class-specific performance in Table 20, it can be observed the results varied compared to SMOTEENN. Some crop classes, such as peanut, corn, and rainfed rice show improved F1-scores (0.89, 0.62, and 0.37 respectively). However, millet and sorghum have slightly lower F1-scores (0.37 and 0.48), while sesame's performance decreased significantly (F1-score: 0.33). Most non-crop classes maintain high performance, with water achieving perfect classification (F1-score: 1.00). Notably, grassland shows a decrease in performance (F1-score: 0.36) compared to SMOTEENN.

Table 21 reveals that crop classes achieve an accuracy of 71.99%, higher than SMOTEENN's best result (69.13%). This indicates improved crops classification with cloud-

free images. However, non-crop classes show a slightly lower accuracy (86.78%) compared to SMOTEENN (90.49%).

Table 21: Accuracy comparison for crop and non-crop classes for the classification on cloud-free images

Class Type	**Accuracy**	**Samples**
Crop Classes	71.99%	3696
Non-Crop Classes	86.78%	719

The results indicate that classification performance improved slightly when utilizing cloud-free images. This improvement is likely attributable to noise reduction, as previously discussed, in the context of the dataset employed. However, it should be noted that this classification was limited to the PEV S2 tile. Obtaining cloud-free images at identical dates across all three tiles covering the entire ROI proved unfeasible with the available imagery. Consequently, despite demonstrating good performance with less computationally intensive classification, this model's applicability for predictions over the entire ROI using the previously presented scale-up strategy is limited. The model's training on only one-third of the area would inevitably introduce bias when attempting full prediction across the ROI.

Figure 26 shows the results of the classification over the PEV tile. 2 areas have been zoomed out to better visualize the classification pattern.

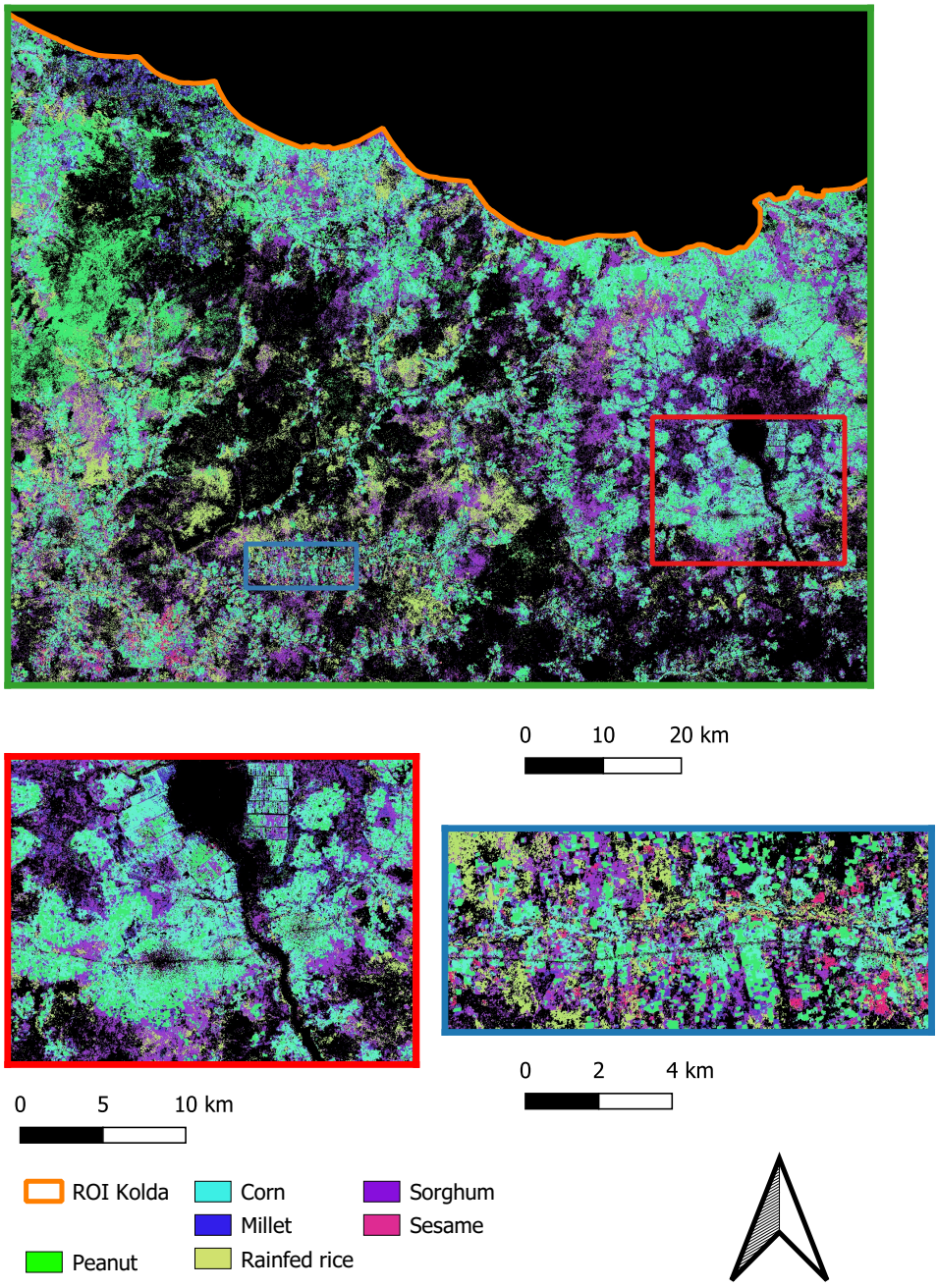


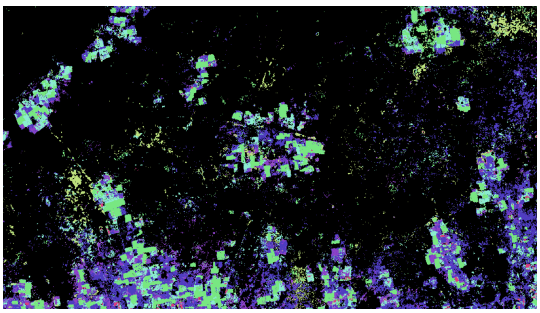
Figure 26: Prediction over the PEV tile with the model trained with 3 cloud-free images with spectral bands and metrics (non-crop classes masked)

5.3 Qualitative comparison between the classification results

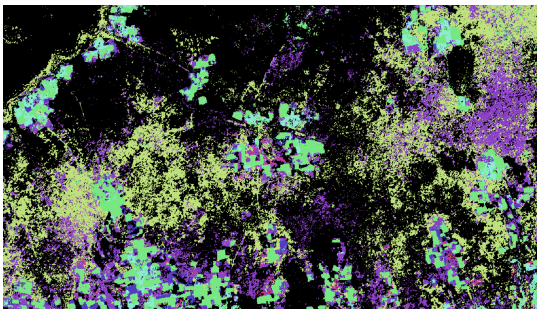
The best feature set with SMOTEENN and the model utilizing cloud-free images demonstrated approximately equivalent quantitative performance. However, qualitative inspection of the two classifications revealed notable differences. The classification derived from cloud-free images exhibited a tendency to overestimate cropland area. As illustrated in Figure 27, areas comprising a mixture of shrubland and trees were classified as non-crop in the time-series classification results (Figure 27b), whereas the same areas were identified as rainfed rice in the cloud-free classification results (Figure 27c). Generally, the time-series classification tended to overestimate the millet, sorghum and corn classes, extending it into non-crop land cover areas. One potential explanation for this phenomenon is that millet, sorghum, and corn fields often exhibit lower crop density. This reduced density allows for the presence of bare soil, grass, and adventitious species, which can introduce noise into the spectral signature, as previously discussed in section 5.2.1.2. Conversely, in the cloud-free classification, the Sorghum and rainfed rice classes displayed a similar tendency to encroach upon non-crop areas.



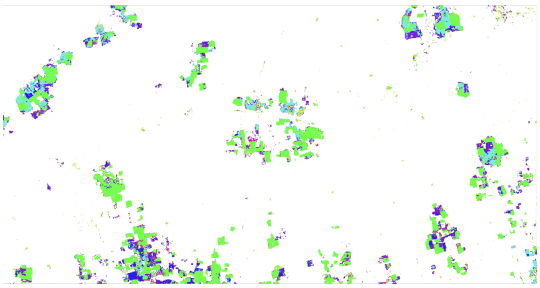
(a) Google satellite image of the comparative area for the different classification



(b) Classification results of time-series dataset (SMOTEENN + spectral bands + metrics



(c) Classification results of cloud-free images dataset with spectral bands + metrics



(d) Classification results of time-series dataset (cropland mask + SMOTEENN + with spectral bands + metrics

Figure 27: Comparison between the four classification models

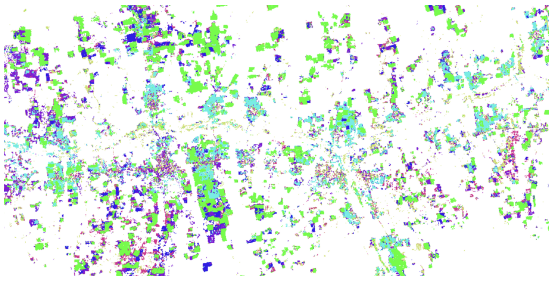
Figure 27d, obtained using a cropland mask, shows therefore no confusion with non-crop classes. However, a slight underestimation of crop plots is observed.

Figure 28 further illustrates this underestimation of crop classes. Some plots classified as sorghum or millet (pink and purple) in Figure 28c are not present in Figure 28b.

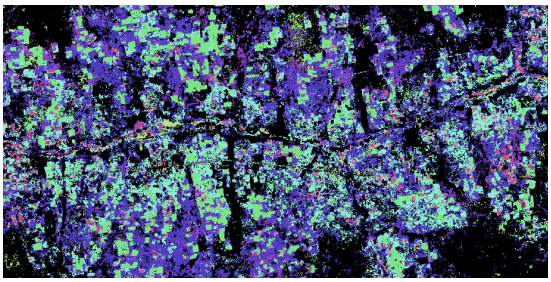
While this difference is visually apparent, the precision and recall values for the time-series classification (Table 14) suggest that this discrepancy may be partially attributed to the model misclassifying these crops as non-crop classes such as shrubland or grassland.



(a) Google satellite image of the comparative area for the different classification



(b) Classification results of time-series dataset (cropland mask + SMOTEENN + with spectral bands + metrics)



(c) Classification results of time-series dataset (SMOTEENN + with spectral bands + metrics)

Figure 28: Comparison between classification with and without cropland mask

Chapter 6 : Discussion and conclusion

This master's thesis explored two approaches for crop-type mapping using S2 data in Kolda: a data-intensive time series approach and a selective cloud-free image approach. Both methods revealed strengths and limitations, providing insights into the challenges of crop classification in this region.

The time series approach, utilizing both spectral bands and computed metrics, demonstrated varying performance across different classes. Non-crop classes consistently achieved higher accuracy compared to crop classes, with overall accuracy for non-crop classes reaching 90.49% and 69,13% frop crop classes using the SMOTEENN algorithm with the spectral and chosen metrics data. This result aligns with the findings of Ibrahim et al. (2021), who reported 72% accuracy for the crop type map identification in Nigeria. However, this is slightly lower than the 75% reported by Forkuor et al. (2014) in Benin that used both optical and SAR data. The disparity in performance between crop and non-crop classes and differences in performance between several crop classes indicates inherent challenges in crop discrimination using the selected features and classification approach for those hard conditions.

Analysis of spectral signatures and probability density functions revealed limited separability among certain crop classes, particularly for rainfed rice, corn, and sorghum. This limited separability likely contributed to the lower classification accuracy for these crops, highlighting a gap in the literature regarding the differentiation of these specific crop types using optical imagery in West African contexts. Field campaign observations further highlighted the complexity of the classification task, revealing significant intra-class variability and the presence of adventitious plants within crop fields. These factors introduce noise into the spectral signatures, potentially leading to misclassification.

The classification using selected cloud-free images showed a slight improvement in overall performance compared to the time series approach, with an increase in F1-score from 0.71 to 0.73. This improvement may be attributed to reduced noise in the dataset, as the cloud-free images eliminated the need for extensive interpolation required in the time series approach. This finding presents an interesting contrast to the general assumption that more temporal data leads to better classification results, suggesting that in regions with high cloud cover and significant interpolation requirements, carefully selected cloud-free images might offer a viable alternative to extensive time series data. However, this comparison should be interpreted cautiously, as the time series approach models were trained on the entire region of interest, while the cloud-free classification was limited to a single S2 tile (PEV). This discrepancy in training data coverage may introduce bias and limit the direct comparability of the two approaches when applied to the entire region.

Several methodological aspects influenced the study outcomes. The use of spatial join-

ing to recover in-situ data may have introduced biases and potential misclassification of polygons, affecting the quality of the training and validation datasets. The prevalence of cloudy images necessitated significant interpolation during preprocessing, potentially introducing errors, particularly in areas with prolonged periods between valid observations or in fields with high variability. The choice of cloud masking algorithm (Sen2Cor) may have impacted the quality of the input data, with other algorithms like Fmask potentially offering improved results.

The application of resampling techniques, specifically SMOTE algorithms, to address class imbalance introduces another layer of complexity. While necessary to mitigate the effects of imbalanced data, these techniques may have inadvertently introduced bias, particularly given the noisy nature of the data as demonstrated by the qualitative assessment of spectral separability and field observations. The current two-level resampling strategy could be counterproductive for classes other than peanut. Future work should explore more nuanced resampling approaches that consider the specific characteristics and challenges of each crop class, potentially employing different resampling techniques or ratios for different classes based on their representation in the dataset and their spectral separability.

The study's methodology also lacked cross-validation for hyperparameter tuning, which may have led to suboptimal model configurations. The XGBoost classifier used in this study was configured with 300 estimators, a learning rate of 0.02, and a max depth of 5. While these parameters were chosen based on initial experiments and literature review, a more comprehensive hyperparameter tuning process, including cross-validation, could potentially improve model performance and should be considered in future work to enhance the robustness and generalizability of the classification models.

The integration of the Sen4Stats cropland mask in the classification process revealed both benefits and limitations. While it effectively eliminated confusion with non-crop classes, it also led to a slight underestimation of crop areas. This underestimation may be attributed to the model's optimization for both crop and non-crop classes, which may not be optimal for crop-only classification, and to potential errors in the Sen4Stats mask itself propagating through the classification process.

Qualitative comparison between the time series and cloud-free image classifications revealed notable differences despite similar quantitative performance. The cloud-free image classification tended to overestimate cropland area, misclassifying areas of mixed shrubland and trees as crops (particularly rainfed rice). The time series classification showed a tendency to overestimate millet, sorghum, and corn classes, extending into non-crop areas. These misclassifications may be attributed to the lower crop density in these fields, allowing for the presence of bare soil, grass, and adventitious species that introduce noise into the spectral signatures.

These findings underscore the complexity of developing robust crop classification systems in regions with challenging environmental conditions and unbalanced ground truth data. While both the data quantity (time series) and data quality (cloud-free) approaches showed comparable overall performance, they exhibited different patterns of misclassification, highlighting the nuanced challenges in crop type discrimination. The slight improvement observed with carefully selected cloud-free images challenges the assumption that more temporal data always leads to better classification results, especially in regions with high cloud cover. However, this observation requires further investigation across larger areas, particularly given the discrepancy in training data coverage between the two approaches.

Future research should focus on refining feature selection, exploring advanced machine learning techniques, incorporating additional data sources such as SAR, and developing region-specific phenological models to improve crop type discrimination in similar challenging environments. These efforts could help address the specific difficulties in differentiating between crops like millet, sorghum, and corn using earth observation data in West African smallholder agriculture contexts. Additionally, further investigation into optimal resampling strategies for imbalanced datasets in this specific context and more comprehensive hyperparameter tuning could potentially yield improvements in classification accuracy and model robustness. Future studies should also ensure consistent training data coverage when comparing different classification approaches to enable more direct and unbiased comparisons.

References

Abramova, I. O. (2022). The Population of Africa under the Conditions of Transformation of the World Order. *Herald of the Russian Academy of Sciences*, 92(14):S1306–S1315.

Belgiu, M. and Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114:24–31.

Chavez, P. S. et al. (1996). Image-based atmospheric corrections-revisited and improved. *Photogrammetric engineering and remote sensing*, 62(9):1025–1035.

Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16:321–357.

Chen, T. and Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *ACM Digital Library*, pages 785–794.

Choukri, M., Laamrani, A., and Chehbouni, A. (2024). Use of Optical and Radar Imagery for Crop Type Classification in Africa: A Review. *Sensors*, 24(11):3618. Number: 11 Publisher: Multidisciplinary Digital Publishing Institute.

Eastwood, R., Lipton, M., and Newell, A. (2010). Chapter 65 Farm Size. In *Handbook of Agricultural Economics*, volume 4, pages 3323–3397. Elsevier.

Elith, J., Leathwick, J. R., and Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, 77(4):802–813. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2656.2008.01390.x.

FAO AgriSurvey (2023). Senegal published the Annual Agricultural Survey 2019-2020 report, providing data and information on the country's agricultural sector. | Agricultural Integrated Survey Programme - AGRISurvey | Food and Agriculture Organization of the United Nations. `https://www.fao.org/in-action/agrisurvey/news-and-events/detail-events/en/c/1372696/`. Accessed November 11, 2023.

FAO World Census of Agriculture (2023). WCA 2020 | World Programme for the Census of Agriculture | Food and Agriculture Organization of the United Nations. `https://www.fao.org/world-census-agriculture/wcarounds/wca2020/en/`. Accessed November 11, 2023.

Flores-Anderson, A. I., Herndon, K. E., Thapa, R. B., and Cherrington, E. (2019). The SAR Handbook: Comprehensive Methodologies for Forest Monitoring and Biomass Estimation. Accessed November 20, 2023.

Forkuor, G., Conrad, C., Thiel, M., Ullmann, T., and Zoungrana, E. (2014). Integration of Optical and Synthetic Aperture Radar Imagery for Improving Crop Mapping in Northwestern Benin, West Africa. *Remote Sensing*, 6(7):6472–6499. Number: 7 Publisher: Multidisciplinary Digital Publishing Institute.

Fukushima, H., Higurashi, A., Mitomi, Y., Nakajima, T., Noguchi, T., Tanaka, T., and Toratani, M. (1998). Correction of atmospheric effect on ADEOS/OCTS ocean color data: Algorithm description and evaluation of its performance. *Journal of Oceanography*, 54(5):417–430.

Gascon, F., Bouzinac, C., Thépaut, O., Jung, M., Francesconi, B., Louis, J., Lonjou, V., Lafrance, B., Massera, S., Gaudel-Vacaresse, A., Languille, F., Alhammoud, B., Viallefont, F., Pflug, B., Bieniarz, J., Clerc, S., Pessiot, L., Trémas, T., Cadau, E., De Bonis, R., Isola, C., Martimort, P., and Fernandez, V. (2017). Copernicus Sentinel-2A Calibration and Products Validation Status. *Remote Sensing*, 9(6):584.

Georganos, S., Grippa, T., Vanhuysse, S., Lennert, M., Shimoni, M., and Wolff, E. (2018). Very high resolution object-based land use–land cover urban classification using extreme gradient boosting. *IEEE Geoscience and Remote Sensing Letters*, 15(4):607–611. Conference Name: IEEE Geoscience and Remote Sensing Letters.

Geudtner, D., Torres, R., Snoeij, P., Davidson, M., and Rommen, B. (2014). Sentinel-1 System capabilities and applications. In *2014 IEEE Geoscience and Remote Sensing Symposium*, pages 1457–1460. Accessed November 20, 2023.

Gu, J.-Y., Ryzhkov, A., Zhang, P., Neilley, P., Knight, M., Wolf, B., and Lee, D.-I. (2011). Polarimetric Attenuation Correction in Heavy Rain at C Band. *Journal of Applied Meteorology and Climatology*, 50(1):39–58.

Gualtieri, J. A., Chettri, S. R., Cromp, R. F., and Johnson, L. (1999). Support vector machine classifiers as applied to aviris data. In *Proc. Eighth JPL Airborne Geoscience Workshop*, pages 8–11.

Hai Ly, N., Nguyen, H. D., Loubiere, P., Van Tran, T., Șerban, G., Zelenakova, M., Brețcan, P., and Laffly, D. (2022). The composition of time-series images and using the technique smote enn for balancing datasets in land use/cover mapping. *Acta Montan Slovaca*, 27(2):342–359.

He, H., Bai, Y., Garcia, E. A., and Li, S. (2008). ADASYN: Adaptive synthetic sampling approach for imbalanced learning. In *2008 IEEE International Joint Conference*

on Neural Networks (IEEE World Congress on Computational Intelligence), pages 1322–1328. ISSN: 2161-4407.

He, T., Sun, Y.-J., Xu, J.-D., Wang, X.-J., and Hu, C.-R. (2014). Enhanced land use/-cover classification using support vector machines and fuzzy k-means clustering algorithms. *Journal of Applied Remote Sensing*, 8(1):083636.

Hlavka, C. A., Haralick, R. M., Carlyle, S. M., and Yokoyama, R. (1980). The discrimination of winter wheat using a growth-state signature. *Remote Sensing of Environment*, 9(4):277–294.

Huang, H., Gong, P., Clinton, N., and Hui, F. (2008). Reduction of atmospheric and topographic effect on Landsat TM data for forest classification. *International Journal of Remote Sensing*, 29(19):5623–5642.

Ibrahim, E. S., Rufin, P., Nill, L., Kamali, B., Nendel, C., and Hostert, P. (2021). Mapping Crop Types and Cropping Systems in Nigeria with Sentinel-2 Imagery. *Remote Sensing*, 13(17):3523. Number: 17 Publisher: Multidisciplinary Digital Publishing Institute.

Inglada, J., Arias, M., Tardy, B., Hagolle, O., Valero, S., Morin, D., Dedieu, G., Sepulcre, G., Bontemps, S., Defourny, P., and Koetz, B. (2015). Assessment of an Operational System for Crop Type Map Production Using High Temporal and Spatial Resolution Satellite Optical Imagery. *Remote Sensing*, 7(9):12356–12379. Number: 9 Publisher: Multidisciplinary Digital Publishing Institute.

Johnson, B. A. and Iizuka, K. (2016). Integrating OpenStreetMap crowdsourced data and landsat time-series imagery for rapid land use/land cover (LULC) mapping: Case study of the laguna de bay area of the philippines. *Applied Geography*, 67:140–149.

Knoema (2024). Superficie, Rendement et Production agricoles - Senegal Data Portal. `https://senegal.opendataforafrica.org//ekihmme/superficie-rendement-et-production-agricoles?region=1000300-region-de-kolda`. Accessed June 19, 2024.

Krawczyk, B. (2016). Learning from imbalanced data: open challenges and future directions. *Progress in Artificial Intelligence*, 5(4):221–232.

Lowder, S. K., Skoet, J., and Raney, T. (2016). The Number, Size, and Distribution of Farms, Smallholder Farms, and Family Farms Worldwide. *World Development*, 87:16–29.

Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., and Johnson, B. A. (2019). Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 152:166–177.

Mandal, D., Hosseini, M., McNairn, H., Kumar, V., Bhattacharya, A., Rao, Y. S., Mitchell, S., Robertson, L. D., Davidson, A., and Dabrowska-Zielinska, K. (2019). An investigation of inversion methodologies to retrieve the leaf area index of corn from C-band SAR data. *International Journal of Applied Earth Observation and Geoinformation*, 82:101893.

Mou, L., Bruzzone, L., and Zhu, X. X. (2019). Learning spectral-spatial-temporal features via a recurrent convolutional neural network for change detection in multispectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 57(2):924–935. Conference Name: IEEE Transactions on Geoscience and Remote Sensing.

Richards, J. A. (2013). Supervised Classification Techniques. In Richards, J. A., editor, *Remote Sensing Digital Image Analysis: An Introduction*, pages 247–318. Springer, Berlin, Heidelberg.

Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M., and Rigol-Sanchez, J. P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 67:93–104.

SATPATHY, S. (2020). SMOTE for Imbalanced Classification with Python. Accessed June 27, 2024.

Sen4Stat (2024). Overview of Sen4Stat – SEN4STAT. Accessed August 16, 2024.

Sentinel, C. (2023). Illustration of ascending and descending satellite orbits. https://www.researchgate.net/figure/Illustration-of-ascending-and-descending-satellite-orbits_fig1_310799598. Accessed November 14, 2023.

Shen, Y., Wu, L., Di, L., Yu, G., Tang, H., Yu, G., and Shao, Y. (2013). Hidden Markov Models for Real-Time Estimation of Corn Progress Stages Using MODIS and Meteorological Data. *Remote Sensing*, 5(4):1734–1753.

Sheykhmousa, M., Mahdianpari, M., Ghanbari, H., Mohammadimanesh, F., Ghamisi, P., and Homayouni, S. (2020). Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13:6308–6325.

Siachalou, S., Mallinis, G., and Tsakiri-Strati, M. (2017). Analysis of Time-Series Spectral Index Data to Enhance Crop Identification Over a Mediterranean Rural Landscape. *IEEE Geoscience and Remote Sensing Letters*, 14(9):1508–1512.

Sinergise, S.-H. (2024). FCOVER (Fraction of green Vegetation Cover). `https://custom-scripts.sentinel-hub.com/custom-scripts/sentinel-2/fcover/`. Assessed February 9, 2024.

Skriver, H., Svendsen, M., and Thomsen, A. (1999). Multitemporal C- and L-band polarimetric signatures of crops. *IEEE Transactions on Geoscience and Remote Sensing*, 37(5):2413–2429.

Song, C., Woodcock, C. E., Seto, K. C., Lenney, M. P., and Macomber, S. A. (2001). Classification and Change Detection Using Landsat TM Data: When and How to Correct Atmospheric Effects? *Remote Sensing of Environment*, 75(2):230–244.

Sonobe, R., Yamaya, Y., Tani, H., Wang, X., Kobayashi, N., and ichiro Mochizuki, K. (2018). 20. Crop classification from Sentinel-2-derived vegetation indices using ensemble learning. *Journal of Applied Remote Sensing*, 12(2):026019.

Steinhausen, M. J., Wagner, P. D., Narasimhan, B., and Waske, B. (2018). Combining sentinel-1 and sentinel-2 data for improved land use and land cover mapping of monsoon regions. *International Journal of Applied Earth Observation and Geoinformation*, 73:595–604.

Swana, E. F., Doorsamy, W., and Bokoro, P. (2022). Tomek Link and SMOTE Approaches for Machine Fault Classification with an Imbalanced Dataset. *Sensors*, 22(9):3246. Number: 9 Publisher: Multidisciplinary Digital Publishing Institute.

Ulaby, F. (1975). Radar Response to Vegetation. *IEEE Transactions on Antennas and Propagation*, 23(1):36–45.

Vapnik, V. (1999). An overview of statistical learning theory. *IEEE Transactions on Neural Networks*, 10(5):988–999.

Weiss, M., Baret, F., and Jay, S. (2020). S2toolbox level 2 products: Lai, fapar, fcover. *Institut National de la Recherche Agronomique (INRA)*.

Zhang, L., Zhang, L., and Du, B. (2016). Deep learning for remote sensing data: A technical tutorial on the state of the art. *IEEE Geoscience and Remote Sensing Magazine*, 4(2):22–40. Conference Name: IEEE Geoscience and Remote Sensing Magazine.

Appendix A: Confusion matrices for all scenarios tested

Table 22: Confusion matrix for SC1_spectral_smote_strategy (Accuracy: 68.12%)

True class \ Predicted class	Peanut	Corn	Millet	Rainfed rice	Sorghum	Sesame	Trees	Shrubland	Grassland	Urban	Bare	Water	Herbaceous W.	Omission
Peanut	**2559**	21	2	12	1	7	0	1	1	0	0	0	0	1.74%
Corn	731	**399**	32	51	30	3	2	1	0	0	0	0	0	68.20%
Millet	69	368	**246**	8	61	0	0	8	0	0	0	0	0	67.81%
Rainfed rice	202	50	0	**24**	1	2	0	4	0	0	0	0	0	91.47%
Sorghum	20	134	52	19	**232**	2	0	5	6	0	0	0	0	51.36%
Sesame	0	32	76	0	0	**80**	0	0	0	14	0	0	0	60.40%
Trees	1	0	0	6	1	1	**249**	0	0	0	0	0	0	3.49%
Shrubland	0	16	2	2	0	0	16	**125**	7	0	0	0	0	25.60%
Grassland	16	0	2	3	5	0	0	0	**152**	0	0	0	0	14.61%
Urban	1	4	0	1	2	0	0	0	0	**173**	0	0	0	5.43%
Bare	0	6	0	2	2	1	0	0	12	17	**99**	0	0	28.78%
Water	0	0	0	0	0	0	0	0	0	0	0	**143**	2	1.38%
Herbaceous Wetland	0	0	0	3	0	0	6	2	0	0	0	7	**164**	10.38%
Commission Error	29.03%	70.70%	40.43%	81.68%	31.18%	16.67%	8.79%	14.38%	14.61%	15.20%	0.00%	4.67%	1.20%	

Table 23: Confusion matrix for SC2_metrics_smote_strategy (Accuracy: 69.15%)

True class \ Predicted class	Peanut	Corn	Millet	Rainfed rice	Sorghum	Sesame	Trees	Shrubland	Grassland	Urban	Bare	Water	Herbaceous W.	Omission
Peanut	**2428**	78	11	47	7	15	0	2	16	0	0	0	0	7.17%
Corn	352	**460**	341	21	58	6	0	8	1	2	0	0	0	63.40%
Millet	8	258	**438**	0	52	1	0	3	0	0	0	0	0	42.19%
Rainfed rice	183	63	4	**19**	4	7	0	2	1	0	0	0	0	93.22%
Sorghum	11	155	62	26	**195**	3	2	2	12	0	0	0	2	59.13%
Sesame	2	28	76	0	0	**96**	0	0	0	0	0	0	0	52.48%
Trees	0	0	0	8	0	6	**244**	0	0	0	0	0	0	5.43%
Shrubland	2	1	5	6	1	0	20	**128**	5	0	0	0	0	23.81%
Grassland	19	15	8	0	7	0	0	1	**126**	1	0	0	1	29.21%
Urban	3	1	0	11	2	0	0	0	0	**162**	2	0	0	11.05%
Bare	2	5	0	0	0	4	0	0	0	13	**115**	0	0	17.27%
Water	0	0	0	0	0	0	0	0	0	0	0	**145**	0	0.00%
Herbaceous Wetland	0	0	0	7	0	0	0	0	6	0	0	10	**159**	13.19%
Commission Error	19.39%	63.41%	52.45%	86.21%	40.12%	30.51%	8.27%	12.33%	24.55%	9.00%	1.71%	6.45%	1.86%	

Table 24: Confusion matrix for SC3_spectral_metrics_smote_strategy (Accuracy: 71.18%)

True class \ Predicted class	Peanut	Corn	Millet	Rainfed rice	Sorghum	Sesame	Trees	Shrubland	Grassland	Urban	Bare	Water	Herbaceous W.	Omission
Peanut	**2485**	52	18	30	9	0	0	0	10	0	0	0	0	4.55%
Corn	548	**491**	97	46	59	1	4	1	1	1	0	0	0	61.24%
Millet	7	253	**409**	15	76	0	0	0	0	0	0	0	0	45.74%
Rainfed rice	187	45	4	**29**	5	13	0	0	0	0	0	0	0	89.68%
Sorghum	5	147	51	22	**224**	0	2	8	10	0	0	0	1	53.14%
Sesame	0	32	76	0	0	**94**	0	0	0	0	0	0	0	53.47%
Trees	2	0	0	2	0	0	**252**	1	1	0	0	0	0	2.33%
Shrubland	0	7	0	0	0	0	13	**141**	7	0	0	0	0	16.07%
Grassland	15	5	0	1	4	0	0	0	**152**	0	0	0	1	14.61%
Urban	2	3	0	4	2	0	0	0	0	**170**	0	0	0	6.08%
Bare	0	3	2	0	0	4	0	0	12	16	**102**	0	0	26.62%
Water	0	0	0	0	0	0	0	0	0	0	0	**141**	4	2.76%
Herbaceous Wetland	0	0	0	0	0	0	1	0	6	0	0	11	**164**	10.38%
Commission Error	23.47%	52.25%	37.76%	80.54%	40.67%	16.07%	7.35%	6.62%	23.62%	9.09%	0.00%	7.89%	4.12%	

Table 25: Confusion matrix for SC4_spectral_ADASYN_strategy (Accuracy: 65.96%)

True class \ Predicted class	Peanut	Corn	Millet	Rainfed rice	Sorghum	Sesame	Trees	Shrubland	Grassland	Urban	Bare	Water	Herbaceous W.	Omission
Peanut	**2507**	46	8	11	21	6	0	1	4	0	0	0	0	3.72%
Corn	655	**444**	38	49	24	12	2	25	0	0	0	0	0	64.45%
Millet	24	480	**177**	25	53	1	0	0	0	0	0	0	0	76.71%
Rainfed rice	193	46	0	**32**	2	9	0	0	1	0	0	0	0	88.69%
Sorghum	16	197	71	25	**145**	6	1	3	5	0	0	0	1	69.15%
Sesame	0	35	76	0	0	**80**	0	0	0	11	0	0	0	60.40%
Trees	0	0	0	7	0	0	**250**	0	0	0	0	0	1	3.10%
Shrubland	0	14	0	0	0	0	10	**137**	6	0	0	0	1	18.45%
Grassland	16	5	4	1	3	2	0	2	**145**	0	0	0	0	18.54%
Urban	0	5	0	3	2	0	0	0	0	**171**	0	0	0	5.52%
Bare	0	6	0	3	0	1	0	0	14	18	**97**	0	0	30.22%
Water	0	0	0	0	0	0	0	0	0	0	0	**141**	4	2.76%
Herbaceous Wetland	0	0	0	1	0	0	0	0	0	0	0	9	**172**	5.49%
Commission Error	26.43%	64.98%	52.67%	79.62%	42.00%	31.62%	4.94%	18.45%	17.14%	14.50%	0.00%	6.00%	4.44%	

Table 26: Confusion matrix for SC5_metrics_ADASYN_strategy (Accuracy: 68.85%)

True class \ Predicted class	Peanut	Corn	Millet	Rainfed rice	Sorghum	Sesame	Trees	Shrubland	Grassland	Urban	Bare	Water	Herbaceous W.	Omission
Peanut	**2404**	82	24	55	7	11	0	2	19	0	0	0	0	7.68%
Corn	347	**474**	296	34	80	11	0	4	2	1	0	0	0	62.05%
Millet	7	215	**435**	35	64	0	0	4	0	0	0	0	0	42.76%
Rainfed rice	184	51	4	**27**	9	7	0	0	1	0	0	0	0	90.46%
Sorghum	4	173	59	37	**175**	1	2	4	12	0	0	0	3	62.77%
Sesame	1	16	76	0	1	**107**	0	0	1	0	0	0	0	47.03%
Trees	0	0	0	3	0	5	**249**	1	0	0	0	0	0	3.49%
Shrubland	3	1	5	12	12	0	13	**122**	0	0	0	0	0	27.38%
Grassland	19	10	3	3	14	1	0	4	**116**	0	7	0	1	34.83%
Urban	1	4	0	7	2	0	0	0	1	**165**	1	0	0	8.84%
Bare	1	6	0	0	1	4	0	0	0	13	**114**	0	0	17.99%
Water	0	0	0	0	0	0	0	0	0	0	0	**145**	0	0.00%
Herbaceous Wetland	0	0	0	7	0	0	0	0	6	0	0	7	**162**	10.99%
Commission Error	19.21%	59.79%	52.43%	87.50%	52.00%	26.71%	5.68%	13.36%	30.19%	7.69%	7.02%	4.61%	2.41%	

Table 27: Confusion matrix for SC6_spectral_metrics_ADASYN_strategy (Accuracy: 70.00%)

True class \ Predicted class	Peanut	Corn	Millet	Rainfed rice	Sorghum	Sesame	Trees	Shrubland	Grassland	Urban	Bare	Water	Herbaceous W.	Omission
Peanut	**2463**	58	17	19	24	20	0	0	3	0	0	0	0	5.41%
Corn	549	**546**	60	62	24	3	4	0	0	1	0	0	0	56.29%
Millet	24	313	**369**	27	27	0	0	0	0	0	0	0	0	51.45%
Rainfed rice	187	46	3	**39**	6	2	0	0	0	0	0	0	0	86.22%
Sorghum	14	166	84	27	**170**	2	2	2	3	0	0	0	0	63.83%
Sesame	2	36	76	0	0	**72**	0	0	0	16	0	0	0	64.36%
Trees	0	0	0	6	0	0	**251**	0	0	0	0	0	1	2.71%
Shrubland	1	14	0	0	2	0	20	**130**	1	0	0	0	0	22.62%
Grassland	12	3	2	0	4	2	0	0	**152**	0	0	0	3	14.61%
Urban	1	5	0	4	2	0	0	0	0	**169**	0	0	0	6.63%
Bare	1	3	0	3	0	1	0	0	10	18	**103**	0	0	25.90%
Water	0	0	0	0	0	0	0	0	0	0	0	**142**	3	2.07%
Herbaceous Wetland	0	0	0	0	0	0	0	0	6	0	0	9	**167**	8.24%
Commission Error	24.36%	54.44%	39.20%	78.62%	34.37%	29.41%	9.35%	1.52%	13.14%	17.16%	0.00%	11.26%	4.02%	

Table 28: Confusion matrix for SC7_spectral_smoteENN_strategy (Accuracy: 67.41%)

True class \ Predicted class	Peanut	Corn	Millet	Rainfed rice	Sorghum	Sesame	Trees	Shrubland	Grassland	Urban	Bare	Water	Herbaceous W.	Omission
Peanut	**2506**	53	7	18	13	4	0	1	2	0	0	0	0	3.76%
Corn	599	**447**	59	86	37	3	1	16	0	1	0	0	0	64.21%
Millet	26	318	**271**	24	98	0	0	23	0	0	0	0	0	64.34%
Rainfed rice	201	47	0	**27**	2	1	0	2	3	0	0	0	0	90.46%
Sorghum	16	213	36	24	**168**	1	1	10	1	0	0	0	0	64.26%
Sesame	0	35	76	0	0	**78**	0	0	0	13	0	0	0	61.39%
Trees	0	0	0	3	0	0	**252**	3	0	0	0	0	0	2.33%
Shrubland	0	18	0	1	0	0	9	**127**	13	0	0	0	0	24.40%
Grassland	14	3	2	5	5	0	0	0	**149**	0	0	0	0	16.29%
Urban	0	4	0	2	4	0	0	0	0	**171**	0	0	0	5.52%
Bare	0	6	2	1	1	1	0	0	10	18	**100**	0	0	28.06%
Water	0	0	0	0	0	0	0	0	0	0	0	**145**	0	0.00%
Herbaceous Wetland	0	0	0	7	0	0	7	1	0	0	0	11	**156**	14.29%
Commission Error	25.43%	61.04%	40.10%	86.82%	49.12%	11.36%	6.67%	30.60%	16.29%	15.79%	0.00%	7.05%	0.64%	

Table 29: Confusion matrix for SC8_metrics_smoteENN_strategy (Accuracy: 69.00%)

Predicted class / True class	Peanut	Corn	Millet	Rainfed rice	Sorghum	Sesame	Trees	Shrubland	Grassland	Urban	Bare	Water	Herbaceous W.	Omission
Peanut	**2419**	76	11	51	8	20	0	1	18	0	0	0	0	7.10%
Corn	348	**473**	329	25	58	5	1	5	4	1	0	0	0	62.13%
Millet	10	285	**403**	0	60	0	0	2	0	0	0	0	0	46.97%
Rainfed rice	174	58	6	**21**	9	13	0	1	1	0	0	0	0	92.58%
Sorghum	7	150	61	29	**203**	1	2	4	11	0	0	0	2	56.81%
Sesame	2	28	74	0	1	**96**	0	0	1	0	0	0	0	52.48%
Trees	0	0	0	7	0	6	**245**	0	0	0	0	0	0	5.04%
Shrubland	2	1	4	10	1	0	13	**133**	4	0	0	0	0	20.83%
Grassland	18	16	8	2	8	0	0	0	**115**	2	9	0	0	35.39%
Urban	2	1	0	11	1	0	0	0	0	**165**	1	0	0	8.84%
Bare	1	8	0	0	1	3	0	0	0	12	**114**	0	0	17.99%
Water	0	0	0	0	0	0	0	0	0	0	0	**145**	0	0.00%
Herbaceous Wetland	0	0	0	7	0	0	0	0	6	0	0	11	**158**	13.19%
Commission Error	19.14%	59.80%	54.79%	86.21%	40.98%	33.33%	6.12%	9.59%	30.63%	8.33%	8.06%	7.05%	2.47%	

Table 30: Confusion matrix for SC9_spectral_metrics_smoteENN_strategy (Accuracy: 73.05%)

Predicted class / True class	Peanut	Corn	Millet	Rainfed rice	Sorghum	Sesame	Trees	Shrubland	Grassland	Urban	Bare	Water	Herbaceous W.	Omission
Peanut	**2519**	35	15	16	4	2	0	0	13	0	0	0	0	3.26%
Corn	506	**559**	65	65	42	2	3	0	5	2	0	0	0	55.24%
Millet	5	269	**393**	15	78	0	0	0	0	0	0	0	0	48.29%
Rainfed rice	197	47	4	**29**	1	5	0	0	0	0	0	0	0	89.75%
Sorghum	7	119	37	23	**258**	0	3	8	14	0	0	0	1	45.11%
Sesame	1	34	76	0	0	**91**	0	0	0	0	0	0	0	54.95%
Trees	3	0	0	2	0	1	**252**	0	0	0	0	0	0	2.33%
Shrubland	0	3	2	0	0	0	21	**135**	7	0	0	0	0	19.64%
Grassland	17	3	1	1	4	0	0	0	**151**	0	0	0	1	15.17%
Urban	0	7	0	2	3	0	0	0	0	**169**	0	0	0	6.63%
Bare	1	3	1	0	0	3	0	0	3	12	**116**	0	0	16.55%
Water	0	0	0	0	0	0	0	0	0	0	0	**145**	0	0.00%
Herbaceous Wetland	0	0	0	0	0	0	5	1	1	0	0	11	**164**	9.89%
Commission Error	22.61%	48.41%	33.86%	81.05%	33.59%	12.50%	11.41%	6.25%	22.22%	7.65%	0.00%	7.05%	1.81%	

Table 31: Confusion matrix for the classification on few clouds free images

Predicted class / True class	Peanut	Corn	Millet	Rainfed rice	Sorghum	Sesame	Trees	Shrubland	Grassland	Urban	Bare	Water	Herbaceous W.	Omission
Peanut	**1719**	2	0	5	6	0	0	4	0	0	0	0	0	0.99%
Corn	308	**597**	4	1	85	2	0	0	0	0	0	0	0	52.20%
Millet	1	119	**83**	16	25	0	0	0	0	0	0	0	0	66.12%
Rainfed rice	10	14	0	**21**	0	2	0	0	0	0	0	0	0	55.32%
Sorghum	19	169	30	21	**199**	6	2	4	20	0	0	0	0	57.66%
Sesame	42	30	75	0	13	**42**	0	0	0	0	0	0	0	79.21%
Trees	0	0	0	1	2	0	**121**	7	0	0	0	0	0	7.63%
Shrubland	0	0	0	0	11	0	0	**64**	0	0	0	0	0	14.67%
Grassland	14	1	6	0	8	0	0	0	**14**	0	0	0	0	67.44%
Urban	7	0	0	1	3	1	0	0	0	**152**	0	0	0	7.32%
Bare	4	5	1	0	7	0	0	0	0	8	**114**	0	0	17.99%
Water	0	0	0	0	0	0	0	0	0	0	0	**25**	0	0.00%
Herbaceous Wetland	0	0	0	0	7	0	0	0	1	0	0	0	**134**	5.63%
Commission Error	19.27%	36.39%	58.78%	68.89%	45.83%	20.75%	1.63%	18.99%	60.00%	5.00%	0.00%	0.00%	0.00%	

Sentinel-2 satellites for crop-type mapping: use case in Kolda, Senegal, using spectral, temporal augmented and non-augmented data

Guillaume Jadot

Smallholder agriculture is crucial for food production in Africa, particularly in Senegal, but faces challenges that create an urgent need for accurate crop monitoring and mapping systems. Remote sensing data and advanced machine learning techniques offer promising solutions to map smallholder agricultural landscapes, serving as decision-making tools and providing comprehensive land-use change assessments.

While crop monitoring and mapping research has expanded to include sub-Saharan Africa, results remain variable, and accurate classification of certain crop types continues to pose challenges. This variability in outcomes underscores the need for further investigation and method refinement. Therefore, the objective of this master's thesis is to design and compare two classification approaches using Sentinel-2 imagery: one utilizing time series data and another using selected cloud-free images.

To meet these objectives, an extreme gradient boosting classifier was employed to test nine different scenarios for the time series classification. These scenarios explored various SMOTE algorithms to resample the in-situ data. The same classifier was also used for the classification on selected cloud-free images. Both approaches utilized spectral bands and computed metrics specific to crop temporal patterns, aiming to optimize performance in the challenging context of Senegalese smallholder agriculture.

The time series approach using both spectral bands and computed metrics showed varying performance across different classes, with non-crop classes consistently achieving higher accuracy than crop classes. The best overall performance was achieved using the SMOTEENN algorithm, combining spectral bands and metrics. The cloud-free image approach demonstrated a slight improvement in overall accuracy compared to the time series method, but its applicability was limited to a single Sentinel-2 tile. Both approaches faced challenges in accurately classifying specific crop types, particularly those with similar spectral signatures or grown in mixed or low-density fields.

The study concludes by discussing potential improvements to the classification approaches and proposing ideas for further development.

Université catholique de Louvain
Faculté des bioingénieurs
Croix du Sud, 2bte L7.05.01, 1348 Louvain-La-Neuve, Belgique | www.uclouvain.be/agro