UCLouvain

AGRO LOUVAIN

Faculté des bioingénieurs

Exploiting Multi-resolution Earth Observation Data and Intra-Parcel Heterogeneity Information for Improving Crop Type Mapping in Smallholder Cropping Systems in Mali

Author	Philip Milan Knirsch
Co-Supervisor(s)	Prof. Pierre Defourny, UCLouvain Dr. Sophie Bontemps, UCLouvain
Reader(s)	Prof. Charles Bielders, UCLouvain Dr. Julien Radoux, UCLouvain
Academic Year	2024–2025

Master thesis presented in fulfillment of the requirements for the degree Master SAIV – MSc Erasmus+ GEM: Geo-information Science and Earth Observation for Environmental Modelling and Management

Acknowledgements

My gratitude goes to my supervisors Dr. Sophie Bontemps and Professor Pierre Defourny as well as to Boris Nörgaard and Abdoulaye Guindo for their guidance and support in the completion of this thesis. I would also like to thank the staff in the department for creating such a welcoming environment.

I am grateful to Professors Evelyn Uuemaa and Alexander Kmoch at the University of Tartu for giving me a great start to my Master studies and for laying the foundation of my education in geoinformatics and remote sensing.

My appreciation also goes to my peers - Orsi, Arathi, Majedeh, Jorge, Moritz and Will - for their support and friendship during my studies, as well as to all the friends I have made in Kot Erasmus. I will cherish the memories we have had.

Finally, I thank my grandma and my parents for their unwavering support as well as my partner Liisi for her encouragement and patience throughout this year.

Table of Contents

List of Figures

List of Tables

List of Acronyms

BSI	Bare Soil Index
CDL	Cropland Data Layer
CV	Coefficient of Variation
DN	Digital Number
ESA	European Space Agency
EO	Earth Observation
FAO	Food and Agriculture Organization
LUCAS	Land Use and Cover Area Frame Survey
LULC	Land Use and Land Cover
MGRS	Military Grid Reference System
MSI	Multispectral Imager
NDBI	Normalized Difference Built-up Index
NDVI	Normalized Difference Vegetation Index
NIR	Near-Infrared
PAN	Panchromatic
PSRI	Plant Senescing Reflectance Index
RF	Random Forest
RGB	Red, Green, Blue
SAR	Synthetic Aperture Radar
SAVI	Soil-Adjusted Vegetation Index
SMOTE	Synthetic Minority Oversampling Technique
SWIR	Shortwave Infrared
UAV	Unmanned Aerial Vehicle
UTM	Universal Transverse Mercator
VHR	Very High Resolution
VIS	Visible spectrum

1. Introduction

Food security is one of the world's major pressing challenges as the global population grows and climate change threatens agricultural production (Abawiera Wongnaa et al., 2024; Mbow et al., 2019; Sutton, 2022). Agricultural productivity has increased far more slowly in sub-Saharan Africa than in other regions of the world. Smallholders, typically farming on less than 10 ha of land, collectively manage roughly 80 % of the farmland in the region (Giller et al., 2021 ; FAO, n.d).

Mali provides a compelling regional context to examine smallholder cropping. Agriculture employs about two-thirds of Mali's labour force and contributes over one-third of the national GDP (ELD Initiative, 2020; World Bank, n.d.). The countries' cotton belt, a key agricultural region in the southern part of the country, is dominated by rainfed smallholder farming. Smallholders cultivate staple cereals such as sorghum and millet alongside cash crops like cotton and maize.

Smallholder agriculture in the region faces numerous challenges. Farm productivity is constrained by factors such as depleted soil fertility and highly variable rainfall from year to year. Spatial variability in inputs, management practices and environmental conditions leads to strong fluctuations in crop production across time and space (Lambert et al., 2018; Traore et al., 2013). In this context, improving the resilience and productivity of Mali's smallholder sector is critical for both local welfare and broader food security.

Ensuring food security under accelerating climate and demographic pressures hinges on better information about what crops are grown, where, and when. Traditional, labour-intensive methods of collecting these agricultural statistics are costly, slow, and often inaccurate. (Karmakar et al., 2024) Remote sensing offers a promising alternative for monitoring agricultural systems at scale by providing frequent measurements across large areas. Recent missions such as ESA's Sentinel-2 have markedly improved spatial, spectral and temporal detail for crop type mapping. In smallholder landscapes this however remains challenging due to small parcel sizes, irregular field boundaries and strong within-field heterogeneity (Akinyemi et al., 2023; Delrue et al., 2013; Lambert et al., 2018)

A growing literature has adapted classification frameworks to smallholder contexts. Yet how to address and leverage intra-parcel heterogeneity itself, rather than treat it only as a confounding factor, remains underexplored. This thesis aims to address that gap by systematically leveraging multi-resolution Earth Observation data (Sentinel-2, PlanetScope & Pléiades) together with field observations from Mali's cotton belt in order to develop and evaluate strategies for improving crop type mapping in heterogeneous smallholder landscapes.

2. Literature Review

Remote Sensing has emerged as a major asset for mapping and monitoring cropland, offering possibilities to observe agricultural areas cost-effectively, frequently and at scale. It today underpins a large variety of applications, such as for cropland mapping, yield forecasting, pest and disease detection and timely monitoring of environmental impacts on farms (Becker-Reshef et al., 2023). However, in regions dominated by smallholder agriculture, such as Mali's cotton belt, the complexities inherent within its cropping systems, pose particular challenges for leveraging remote sensing for crop type classification. This literature review examines three key aspects: the current use of remote sensing for cropland and crop type mapping, focusing on global and regional scales, the diversity of smallholder cropping systems in the Malian cotton belt and related challenges for crop type classification, and classification approaches in smallholder cropland systems and strategies for addressing intra-parcel heterogeneity in crop type classification frameworks. Together, these subchapters highlight current knowledge and gaps in classifying crops in smallholder farming systems in Mali using remote sensing.

2.1. Remote Sensing for Cropland and Crop Type Mapping

Remote sensing has emerged as an important domain in agriculture, offering regular, non-contact measurements of radiation reflected from crop canopies and soils. Since its inception in the early 1970s, remote sensing technology has evolved significantly, enabling more effective analysis of crop variability and advancing precision agriculture practices (Goffart et al., 2022). Continuous technological innovations have driven improvements in spatial, temporal and spectral resolution, leading to increased data availability, making it feasible to monitor cropland frequently at finer scales.

Cropland and crop type maps can provide data of use for agricultural statistics and food security planning. In countries such as Mali, agricultural statistics have, and so far, mainly still rely on labour-intensive field surveys that are constrained in terms of representativeness and reliability (Lambert et al., 2018). These conventional methods involve manual field inspections that are time-consuming, resource-intensive, prone to human error, and restricted by limited sample sizes that can be realistically evaluated within a given timeframe (Karmakar et al., 2024). Remote sensing offers a scalable and efficient alternative: satellite and aerial imagery can cover large areas with repetitive coverage, allowing data collection at short intervals and the creation of time-series datasets for effective crop monitoring over time. This data provides a scalable and efficient alternative by enabling manual analysis and digitization of cropland based on visual interpretation of the data as well as through automated classification of landscape elements (Karmakar et al., 2024). The later is commonly done through Land Use and Land Cover (LULC) classifications that assign classes to cropland and non-cropland elements in the landscape, allowing to derive cropland masks of cultivated areas. Classifications can be further refined for crop type classification, offering insights into crop type-specific distributions.

Cropland extent maps derived from remote sensing data have proliferated in recent years. Global products range from comprehensive land cover maps that include cropland classes (such as ESA WorldCover 2021 at 10m resolution) to dedicated cropland-specific products like the Global Cropland Extent Map (GCEM) at 30m resolution (Thenkabail et al., 2021). Potapov et al. (2022) created a time series of global cropland extent maps and change maps between the period 2000-2019 in 4-year intervals at 30m resolution based on Landsat data. ESA's World Cereal project created cropland extent maps and season-specific maps on active cropland and selected crop types at 10m resolution, derived from Sentinel-1 and Sentinel-2 data for the year 2021 and are currently developing a scalable open-source system for producing globally consistent yet locally tuneable cropland masks (Van Tricht et al., 2023).

The performance of global cropland products can vary significantly across different regions (Pérez-Hoyos et al., 2017) and have shown overall low accuracies in Africa (Nabil et al., 2020). Further, large discrepancies between products have been reported. Kerner et al. (2024) conducted a study evaluating and intercomparing 11 publicly available global and regional cropland extent and land cover maps with a cropland class (Digital Earth Africa Cropland Extent, Dynamic World, Esri LULC, ESA WorldCover, CCI Land Cover Africa, GFSAD Global Cropland Extent, 30 m African cropland layer for 2016 by Nabil et al. (2022), GLAD, Copernicus Land Cover, ESA GlobCover, and ASAP Crop Mask) to assess their accuracy in cropland classification. Key contributions of their study included an accuracy assessment using reference datasets in eight countries in Sub-Saharan Africa including Mali, quantification of cropland classification consensus across all maps and pairwise agreement between individual maps as well as an evaluation of accuracy against spatial resolution. They found very low consensus across the 11 compared maps when predicting cropland, with the maps only unanimously agreeing on a cropland prediction in fewer than 0.5% of pixels in each of the 8 countries studied. Their accuracy assessment against cropland reference datasets showed large disparities in performance, exhibiting F1 scores ranging from 0.21 for Mali to 0.71 for Rwanda. The authors further showed that the performance of the cropland products are overall weakly correlated with their spatial resolution, noting that their findings suggest that finer spatial resolution satellite data reduces false positives (other land cover types incorrectly classified as cropland) but not false negatives (cropland incorrectly classified as other land cover).

Crop type maps go beyond the foundational delineation provided by cropland masks, offering detailed insights into the specific crops cultivated within agricultural areas. They play a vital role globally, but their development still faces significant challenges. Historically, insufficient satellite data and computational limitations posed barriers to developing these maps, but contemporary challenges revolve around the scarcity of high-quality ground reference data needed for calibrating and validating crop type classifications (Becker-Reshef et al., 2023). Currently at the global level, datasets like IFPRI SPAM2010, M3-Crops, and MIRCA2000 provide broad overviews of crop type distributions. However, these global datasets rely on spatial allocation models and subnational statistics rather than the spectral signals of crops, limiting their resolution (~10 km) and accuracy, as well as their temporal relevance (Becker-Reshef et al., 2023). For the European Continent, the European Crop Type Map for 2018 and 2022 provides comprehensive crop type classification at 10m resolution for 19 crop types in 30 countries based on Sentinel-1 and Sentinel-2 satellite data. This was achieved through leveraging an extensive collection of field data, generated through the European Land Use and Cover Area Survey (LUCAS) (Ghassemi et al., 2024). A similar product, the Cropland Data Layer (CDL) exists for the United States, which provides annual crop type classifications for the entire Continental United States since 2008 at 30 m resolution and since 2024 at 10 m resolution. The current CDL is based on a combination of Landsat and Sentinel-2 data, and also leverages extensive field data collected through the National Agricultural Statistics Service (USDA NASS, 2024).

Becker-Reshef et al. (2023) produced the first globally harmonized crop type product named GEOGLAM-BACS (Global Best Available Crop Specific Masks) for Maize, Soybean, Rice, Winter Wheat, and Spring Wheat by leveraging a collection of national, regional and global crop-specific masks and harmonizing them to a common 0.05-degree resolution. The best available crop-specific masks were selected, including the USDA NASS CDL for the US, and the European Crop Type Map. For Africa (excluding South Africa), the product however still relies on the IFPRI SPAM2010 statistical allocation model.

Africa currently lacks a comprehensive, operational crop type mapping system with continental coverage based on EO data. Efforts are being made under the Digital Earth Africa initiative umbrella to develop a crop type classification system for Africa, with pilot projects in several countries, though operational products remain in development (FAO, n.d.)

LULC mapping at large spatial extents, whether at country, continental, or global scales, requires automated classification approaches, as manual methods become prohibitively unfeasible and resource-intensive (Gilić et al., 2023). The approaches to LULC can be broadly categorized into pixel-based and object-based methods, as well as supervised and unsupervised classification techniques. Most classification approaches are based on per-pixel information, in which each pixel is classified into exclusive classes based on the spectral information contained in each pixel. Supervised pixel-based methods require training data where the class labels are known, enabling the algorithm to learn the relationship between spectral signatures and LULC classes. Common supervised algorithms include Maximum Likelihood, Random Forests, Support Vector Machines (SVM), and more recently, Deep Learning models (Talukdar et al., 2020). Unsupervised pixel-based methods, such as K-means clustering and ISODATA, group pixels with similar spectral characteristics without requiring training data, though these clusters must subsequently be labelled by analysts. Object-based approaches first segment the image into homogeneous objects or regions based on spatial, spectral, and textural characteristics, and then classify these objects rather than individual pixels. This method better represents landscape elements like fields or housing blocks and can reduce the "salt-and-pepper" effect often observed in pixel-based classifications for high resolution imagery (Blaschke, 2010). The selection of appropriate classification methods depends on several factors including the spatial resolution of the imagery, the complexity of the landscape, the availability of training data, and the specific objectives of the mapping exercise. In the context of crop type classification, temporal information of changes in spectral signatures over time, capturing phenological cycles, are also commonly being integrated into classification methods, aiming to increase the accuracy of crop type differentiation (Waldner et al., 2015; Zhang & Li, 2022).

2.2. Cropland Diversity in Smallholder Cropping Systems and Challenges for Classification

In the Malian cotton belt, as across much of Africa, smallholder farming systems represent the predominant agricultural production model. These systems are characterized by their complexity and diversity, involving various cropping patterns, including intercropping and agroforestry elements (Lambert et al., 2018). Understanding and effectively capturing the processes and diversity within these farming systems is crucial for improving crop type classification.

One way to conceptualize the complexity of smallholder cropping is by breaking it into key dimensions. Bégué et al. (2018) use three fundamental dimensions to describe cropping systems, which can also guide remote sensing analyses in agriculture:

- **Crop succession:** The sequence of crops (and fallow periods) on a given field over multiple growing seasons. This captures the temporal aspect of how agricultural land use changes between years

- **Cropping pattern:** The spatial arrangement and combination of crops and fallows within a field in a growing season, including whether the field is monocropped or hosts multiple crops simultaneously. This dimension addresses the spatial complexity of what is grown where.

- **Crop management techniques:** The set of agricultural practices applied to the field, such as tillage, irrigation, fertilization, or pest management and harvesting

Together, these dimensions provide a framework for capturing the temporal, spatial, and management complexities of smallholder agriculture. In Mali's cotton belt farmers employ a variety of cropping patterns. While monocropping (cultivating a single crop on a field at a time) is common, other practices like intercropping are also widespread. Intercropping involves growing two or more crop species simultaneously on the same parcel of land, often in a mixed or alternating arrangement (Mahlayeye et al., 2022). Trees are also commonly interspersed within fields in smallholder cropping systems (Lungu et al., 2024; Soumaré, 2008). Further, smallholder crop-livestock systems are prevalent, characterized by an interdependence between crop cultivation and livestock keeping (Dissa et al., 2024). Livestock interacts with the cropping system in various ways, such as providing manure and grazing on crop residues.

There is also diversity in crop succession across Mali's cotton belt. While two-year or three-year rotations of cotton-cereal rotations are common, with cotton typically alternating with cereals like maize, millet or sorghum, the specific sequence and duration vary considerably between farms (Dissa et al., 2024; Soumaré, 2008). Soumaré (2008) describes the diversity of crop management practises in the Malian cotton belt, noting that farmers employ a range of techniques tailored to local conditions, technological means and crop types, and that dates for sowing and harvesting and the timing and intensity of management practises can vary between farmers.

Smallholder farming environments in sub-Saharan Africa are further characterized by variable soil fertility at short distances. Within farms, the interplay of soil-forming processes and preferential management have resulted in fields with very different conditions (Soumaré, 2008; Vanlauwe et al., 2017).

These factors constitute considerable challenges for cropland and crop type classification with remote sensing. One bottleneck is the small parcel size. The majority of Africa's agricultural land falls within small (<1.5 ha) and very small (<0.15 ha) parcel sizes (Pérez-Hoyos et al., 2017). Smallholder parcels further typically feature irregular boundaries and shapes that don't align with pixel grids. To detect pure class signatures, pixels need to be smaller than the mean parcel size (Delrue et al., 2013). Delrue et al., (2013) highlight that this can achieved in many regions of the world with medium to high resolution imagery, referring specifically to Landsat 5 at 30 metre resolution, but showed that this resolution was not sufficient for classification in smallholder agricultural regions. Resolutions of around 10m or higher are therefore a prerequisite for addressing crop type classification in these contexts (Lambert et al., 2018).

But challenges go beyond spatial resolution. As highlighted, parcels in smallholder systems show large presence of trees within fields. Meaning that even at resolutions sufficient to capture small fields, the spectral signatures of crops are obscured by the presence of trees interspersed within the parcel. Intercropping presents a further unique challenge, as the crop diversity introduced by intercropping means that the spectral signature of pixels within a field can be a complex composite of spectral signatures from different crop types. Further the variability in management practices and their timing leads to a temporal misalignment of spectral signatures of crops.

In this study, the term heterogeneity is used to refer to the inherent variability of agricultural parcels within smallholder cropping systems. Two distinct forms of heterogeneity can be conceptualized within these systems: inter-parcel heterogeneity, which refers to variability in characteristics, management and yield between different plots, and intra-parcel heterogeneity, which refers to variability within a single plot, including mixed cropping, presence of trees, soil fertility gradients and vegetation health. (Sida et al., 2021). These agricultural realities directly translate into specific remote sensing classification challenges. As Huang et al. (2025) note, smallholder systems create a classification environment characterized by two problematic patterns: lower inter-class variability, where parcels of different crop types appear spectrally more similar due to confounding factors, and higher intra-class variability, where the same crop type exhibits diverse spectral signatures across pixels of different fields and within the same field. Together, these forms of heterogeneity fundamentally challenge conventional classification approaches that assume spectrally distinct classes with internal homogeneity. Both inter-parcel and intra-parcel heterogeneity must therefore be effectively addressed to achieve accurate crop type mapping in smallholder landscapes.

2.3. Crop Type Classification Approaches in Smallholder Cropping Systems

Akinyemi et al. (2023) note that most current Earth Observation approaches for crop type mapping were predominantly developed for large-scale monocultures in industrialized economies, creating methodological challenges when applied to smallholder agricultural contexts. The inherent complexity of smallholder systems, characterized by small parcel sizes, heterogeneous field conditions, extended and variable sowing windows, and the wide presence of non-crop features within fields necessitates classification methodologies capable of handling multidimensional complexity across temporal, spatial, and spectral domains

The advent of Sentinel-2, with it's 10 m resolution in the VIS and NIR domain and temporal revisit time of 5 days has provided the capabilities to address some of these challenges. Studies by for example Gumma et al. (2022) in India, Ibrahim et al. (2021) in Nigeria, Mazarire et al., (2020) in South Africa and Lambert et al., 2018 in Mali have performed crop type mapping in smallholder systems using Sentinel-2 time series. Lambert et al. (2018) mapped crop types using a supervised Random Forest classification approach in Mali's cotton belt for the crops cotton, maize, millet, sorghum and peanut achieving an overall accuracy of 80 %. Furthermore, they incorporated a tree mask derived from soil-specific NDVI thresholding of a VHR image and investigated its effect on crop classification performance. Their analysis revealed that removing trees in classification did not significantly improve accuracy, leading them to conclude that tree masking was not beneficial for crop type mapping at 10-meter spatial resolution within their study area.

The scarcity of cloud-free imagery in the growing season often necessitates combining data from multiple sensors to collect sufficient information of different phenological stages for crop type discrimination. The combination of optical and radar data is particularly compelling as each can capture complementary information with respect to structure of the canopy and biochemical properties of crops (Choukri et al., 2024). A study by Akinyemi et al. (2023) utilized optical and Synthetic Aperture Radar (SAR) imagery for crop type mapping of heterogeneous smallholder farming systems in southwest Nigeria with a focus specifically on capturing maize-cassava intercropping. They employed a combination of Sentinel-1 and Sentinel-2 time series data in different combinations at monthly and bimonthly intervals to classify cropland classes across two growing cycles. Their approach, employing the random forest algorithm, allowed for the identification of eight distinct farming system classes based on crop types and planting periods (Early Cassava, Late-Cassava, Early Maize, Late-Maize, Intercropped Cassava and Maize, Rice, Yam, and "Others"). While the classification model with only monthly Sentinel-1 time series data performed poorly with an accuracy of 0.50, the model combining monthly Sentinel-1 with bimonthly Sentinel-2 data performed better with an overall accuracy of 0.77 and effectively reducing cloud-related data gaps to just 0.2%. Classification utilizing only Sentinel-2 data performed similarly well, but with data gaps affecting over 10 % of the cropland area (Akinyemi et al., 2023). Of note here is that the authors also specifically targeted an intercropped class "maize-cassava" which they were able to identify with a user accuracy of 0.79 and a producer accuracy of 0.85.

Kpienbaareh et al. (2021) combined Sentinel-1 with Sentinel-2 and PlanetScope optical imagery for crop type and land cover mapping in Malawi. Their approach involved an optical image fusion approach of corresponding images of Sentinel-2 and PlanetScope to produce higher resolution Sentinel-2 multispectral images. The combination of the fused imagery with Sentinel-1 data achieved significantly higher accuracy (over 85% overall accuracy) compared to Sentinel-1 or fused Sentinel-2 - PlanetScope imagery alone.

Commercial very high-resolution (VHR) satellite imagery is sometimes employed for classification in smallholder agricultural system; however, such studies are typically constrained to limited geographic extents mainly due to the prohibitive acquisition costs of commercial VHR data (Ibrahim et al., 2021). Aguilar et al. (2018) used spatial and spectral features derived from seven multi-spectral Worldview-2 images (~2m spatial resolution) throughout the cropping season to classify crop types in southern Mali and employed ensemble classifiers to achieve a higher overall accuracy (75.9%) than a single classifier. They found that spatial features can enhance crop discrimination capabilities, particularly in heterogeneous agricultural landscapes where local variance patterns become increasingly pronounced and potentially informative when analysed through VHR imagery. Debats et al. (2016) highlight that features that capture textural and contextual information have been found to improve classification accuracy over spectral information alone, but their use has been limited by their complexity and high computational cost.

According to a review of Mahlayeye et al. (2022), there are few substantial studies so far that have classified crop-combinations within intercropped fields. Hegarty-Craver et al. (2020) mapped dominant crop types in intercropped Rwandan fields by combining unmanned aerial vehicle (UAV)-derived ground truth data with Sentinel-1 SAR and Sentinel-2 optical imagery. Their approach overlaid a 10x10 m satellite grid with UAV imagery, labelling each pixel according to its dominant crop type or landcover (≥75%) across seven classes. The Random Forest model achieved 83% overall accuracy (91% for maize, 84% for beans). Richard et al. (2017) utilized RapidEye data at 5 metre spatial resolution for classifying mono-and mixed maize cropping systems in Kenya. Using two images during maize stem elongation and flowering phenological crop development stages, they first created a crop mask to exclude non-crop features and then classified the cropland into mono- and mixed maize classes using a Random Forest classifier. Their approach achieved an overall accuracy above 85%. They highlight that the confusion patterns were primarily linked to weed infestation and the presence of trees within some fields (Richard et al. 2017). They concluded that explicit mapping of different cropping systems is feasible in smallholder agricultural landscapes if high resolution and multi-temporal satellite data is employed.

Mahlayeye et al., (2022) highlight the potential of VHR sensors for discriminating crop types in intercropped fields. Higer spatial resolution enables the differentiation of individual crops grown in proximity, though distinguishing intercropped species remains challenging due to overlapping spectral signatures and varying crop densities. Despite this potential, studies employing VHR imagery are typically restricted to local applications, with no research yet successfully mapping intercropping patterns or crop proportions across multiple crop types at scale (Mahlayeye et al., 2022). A common methodological approach in crop type mapping involves first creating a cropland mask before proceeding to crop type classification. However, the accuracy and spatial resolution of existing cropland products often hamper the precise identification of smallholder fields. This limitation frequently necessitates manual digitization of field boundaries to create cropland masks (as done by Debats et al., 2016; Lambert et al., 2018) or LULC classifications as precursors to crop type mapping (as done by Richard et al., 2017).

The effectiveness of crop type classification approaches in smallholder contexts is highly dependent on reliable ground data and an understanding of local agricultural dynamics. Heiss et al. (2025) note in their review of Earth Observation applications for mapping cropping systems in West Africa that numerous studies criticize the lack and quality of reference data for smallholder farming, with multiple researchers noting the scarcity of validation and training datasets. The lack of field data availability limits the calibration and validation methods used for mapping intercropping patterns. While not substituting field data, VHR imagery can be utilised to gather reference data for training a model and assessing its accuracy. Additionally, the increasing accessibility of UAV technology presents promising opportunities for collecting reference data at submeter resolutions (Mahlayeye et al., 2022).

3. Research Objectives

Crop type mapping in smallholder agricultural systems is an active research area, with regional and global efforts for improving mapping products being performed. However, few studies have systematically explored how intra-field heterogeneity can be leveraged to improve crop type mapping performance in smallholder systems. This study aims to narrow this gap by investigating the potential of intra-field heterogeneity to enhance crop type classification accuracy.

To explore this, this study will exploit data from three sensors, namely Sentinel-2, PlanetScope, and Pléiades, as well as field observations on crop types and intra-field heterogeneity recorded for cropland parcels in the study area. The study will focus and progress with the following objectives:

(1) Investigate correspondence of field-assessed heterogeneity classifications with spectral-temporal signatures and visible structural properties of parcels across crop types and different sensor resolutions.

(2) Compare the performance differences between Sentinel-2, PlanetScope, and Pléiades imagery for crop type classification, focusing on how spatial resolution affects classification performance across heterogeneous field conditions.

(3) Assess how the integration of intra-field heterogeneity information and satellite derived intra-field NDVI variability affects crop type classification accuracy across varying sensor resolutions.

(4) Evaluate how excluding non-crop features from training and validation datasets through VHR-derived tree and cropland masks influences crop type classification accuracy metrics.

4. Study Area and Data

4.1. Study Area

The study area comprises a 10 × 10 km region (12.64°N-12.73°N, -7.43°E to -7.34°E) located east of Bamako in the Dioïla Cercle, an administrative subdivision of Mali's Koulikoro Region (Figure 1). This area falls within Mali's cotton belt, a key agricultural zone characterized by smallholder farming systems where cotton is the primary cash crop.

Elevation ranges between 345 and 415 metres above sea level. The study area is located in the Sudano-Sahelian zone and is classified as Tropical Savanna Climate (Aw) according to the Köppen-Geiger climate classification. Annual mean temperature in the study area is 27.2°C, with considerable seasonal variation ranging from a minimum of 13.5°C in the coldest month to a maximum of 39.3°C in the warmest month. Precipitation totals approximately 805 mm annually (WorldClim - Fick & Hijmans, 2017). The region experiences a prolonged dry season from November to April with virtually no precipitation during these months. During the dry season the temperature and saturation vapour deficit increase and crop production is not possible without irrigation (Traore et al., 2013).

Soils in the study area are classified as lithic leptosols according to the FAO soil classification system. These are very fine soils developed on continuous rock with the parent rock appearing less than 10 cm from the soil surface and have an extremely rich content of coarse fragments (Anjos et al., 2015; Gómez-Escalonilla et al., 2022).

The study area is largely classified as an alternation of shrub- and grassland with cropland and some built-up area according to the ESA WorldCover land cover map for 2021. The route 6 national highway cuts through the study area in the northwest. The region widely features mixed agrosylvo-pastoral systems centred around cotton, the primary cash crop, rotated with cereals (sorghum, pearl millet, maize) and legumes. Cotton and to a lesser extent maize receive organic manure, chemical fertilizers, and pesticides, while other cereals rarely get any inputs, leading to soil nutrient depletion and declining soil organic matter. Cattle, goats, and sheep are held as livestock with agro-pastoralists practicing primarily sedentary farming, with transhumance taking place in the dry season (Traoré et al., 2013).

Climate variability significantly impacts agriculture in the Sudano-Sahelian region through heat waves, droughts, floods, and extreme weather events. A review by Traoré et al. (2007) of knowledge on regional climate in Sudano-Sahelian West Africa identified rainfall unpredictability as a major agricultural challenge. Farmers face difficult planting decisions, as early rains may not reliably indicate monsoon establishment, risking crop failure from dry spells after planting or waterlogged fields if planting is delayed (Traore et al., 2013).

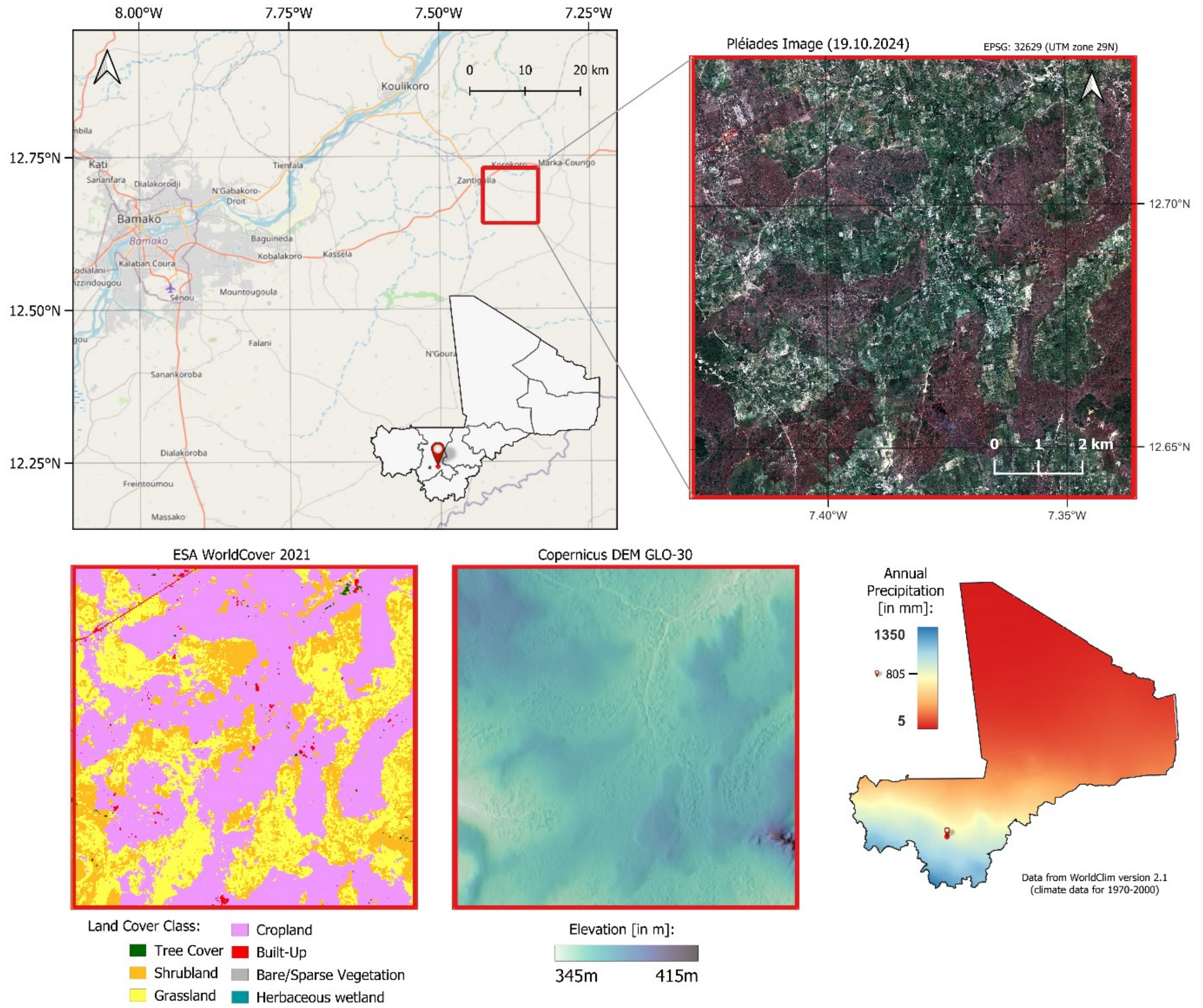

Figure 1: Study area in Mali shown with Pléiades imagery with contextual maps displaying location (Open Street Map), land cover (ESA WorldCover 2021), elevation (Copernicus DEM GLO-30), and annual precipitation (WorldClim v2.1).

4.1.1. Seasonal Pattern of Agricultural Activities and Crop Development

Figure 2 illustrates the agricultural calendar in Mali's cotton belt and describes the characteristic timing of key agricultural activities such as field preparation, sowing and harvesting for major crop types. This calendar (from Soumaré, 2008, see Figure 2) is crucial for understanding the phenological patterns of cultivating crops in the region and will aid to interpret the satellite imagery in subsequent chapters.

Cotton, the primary cash crop in the region, initiates the agricultural season, beginning with land preparation (field cleaning and organic manure application) as early as February-March and continuing into April. Sowing begins with the first significant rains in late May and extends through early June, closely followed by two key weeding periods in June and July, accompanied by urea fertilization (i.e. low-cost nitrogen fertilizer) and ridging. Cotton reaches its peak vegetative growth around September, with systematic insecticide treatments during flowering and boll-setting phases. Harvesting typically occurs between October and November.

Field preparation for maize starts slightly later than cotton, with cleaning activities in March and April. The main sowing period occurs from late May through June, with weeding and urea fertilization primarily in July, followed by ridging in August. Grain filling of the crop usually occurs through September, with harvesting beginning as early as the end of September lasting until the end of October. Post-harvesting processes such as threshing and storaging occur in December and January.

For Millet and Sorghum, the agricultural season begins with field cleaning in May, with sowing taking place between the end of May to early August. Weeding and fertilization takes place from July trough August, with harvesting typically occurring from late October into November, slightly later than for maize and cotton. Threshing and storaging occur thereafter, lasting into December and January.

For "petites cultures," which includes various legumes or vegetables, sowing typically occurs in July with harvesting taking place around the end of September and into October. Overall, the agricultural calendar in Mali's cotton belt is characterized by a clear sequence of activities that are closely tied to the seasonal rainfall patterns. Each crop type follows a slightly different temporal progression from land preparation and planting to harvest.

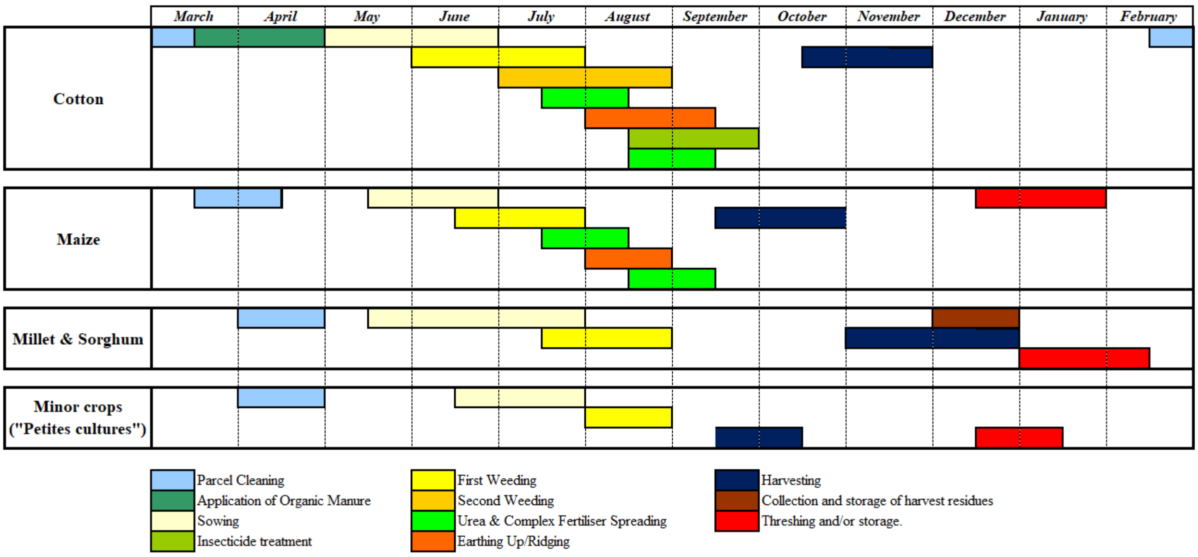

Figure 2: Agricultural calendar showing main crop cultivation periods and timing of agricultural activities in the region. Translated and adapted from Soumaré (2008, p.177).

4.2. In-Situ Data

The in-situ data for this study was collected during a comprehensive field campaign conducted in October and November 2024 under the FAO-EOSTAT programme, with the primary objective of acquiring field data for producing precise cotton maps across Mali's cotton belt. The campaign collected approximately 10,000 geolocated crop points via a windshield survey and 4332 fields were digitized and labelled by enumerators within Area Sample Frames (ASF). Data collection was synchronized with the acquisition of high-resolution Pléiades satellite imagery over selected 10 km x 10 km areas to capture a snapshot of the parcels that were inspected. While the campaign covered multiple sites, this thesis focuses on a single 10 km x 10 km study area. The dataset utilized herein is a subset of the larger campaign, comprising 825 labelled fields derived from the ASF approach for this specific study area. The data was initially recorded in French and translated to English for analysis purposes.

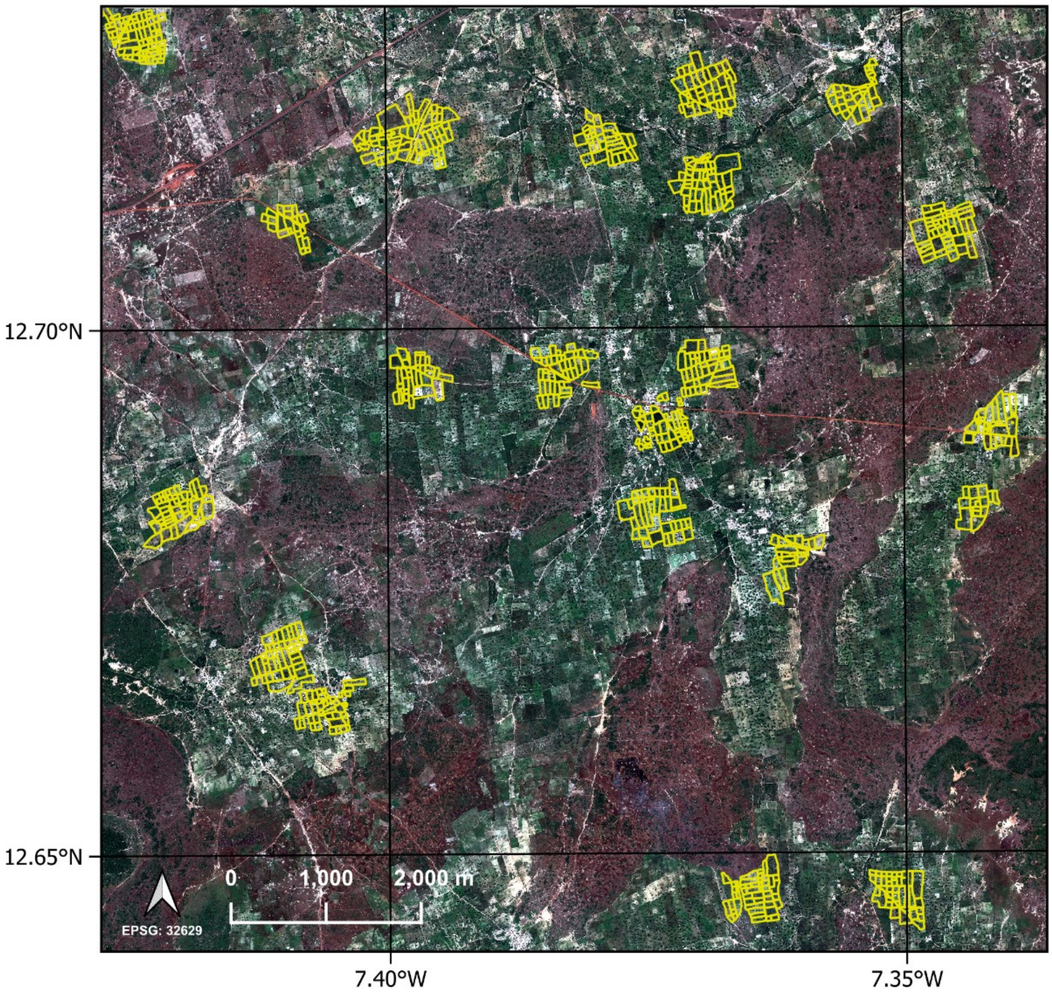

Figure 3: Parcels for which data was collected in the Study Area overlayed on pansharpened Pléiades VHR image taken on 19.10.2024

4.2.1. Overview of In-Situ Data Collection and Key Categories of Information

The data collection focused on the following key informational categories:

Table 1: Categories of Information recorded for the parcels

Category of Information	Description
Crop Type	The type of crop(s) grown in each parcel. Recorded across three columns: (1) the primary crop, (2) the second crop, and (3) the third crop (secondary and tertiary crop only if multiple crops are grown).
Crop System	Classification of cropping system: either (1) Unique, (2) Associated/Intercropped, (3) Mixed, (4) Empty Plot or (5) NA. Reflects monoculture vs. polyculture practices.
Heterogeneity Class	Ordinal classification of parcel-level heterogeneity: (1) Very Low, (2) Low, (3) Medium, (4) High, (5) Very High.

Enumerators recorded additional context or observations in free text fields during data collection. Further, photographs were taken for select parcels to visually document the field conditions. These photographs serve as a visual reference for the recorded observations and provide additional context to the data collected.

4.2.2. Classification Scheme of Intra-Field Heterogeneity

For the in-situ data collection campaign, heterogeneity is operationalized as an ordinal measure that considers structural and functional aspects of the parcels. While ordinal in nature these classes correspond to percentage intervals in bins of 20 percent. The considered elements include crop density, level of weed presence, gaps/bare patches as well as the general structural make-up of the parcel. The enumerators were trained to identify and classify these elements based on their visual observations. The heterogeneity class was then derived by a holistic assessment of these elements (see Table 2).

Based on this classification scheme, a field with lower heterogeneity is typically characterized by a uniform crop structure, with consistent crop cover and no significant bare patches or irregularities. The field appears well-maintained, with minimal weed presence and it shows little to no signs of stress or damage from external factors like flooding or pests. Such a field would be reflective of effective agricultural management and high yield potential, which is why it receives a good field rating.

Conversely, higher heterogeneity fields would display inconsistent vegetative structure with characteristics such as low crop density, visible gaps, patchy coverage, significant weed presence, or damage from environmental factors. Importantly, not all these elements must be present simultaneously. Fields may exhibit only some of these characteristics yet still receive a higher heterogeneity rating.

It is important to note that, according to the classification scheme, the presence of trees within a parcel was not considered a factor in this heterogeneity assessment. They are viewed as a separate element that is supposed to not directly influence the heterogeneity rating of a parcel.

Table 2: Criteria for heterogeneity assessment of parcels

Heterogeneity (%)	Characteristics	Field Rating
Very Low (0–20)	Very well-maintained field + homogeneous crop structure + high density + very few weeds + no significant damage.	Very good field
Low (20–40)	Well-maintained field + homogeneous crop structure + high density + few weeds + very little damage.	Good field
Medium (40–60)	Moderately maintained field + mixed crop structure + medium density + moderate weed presence + moderate damage (flooding, insects, cultivated and lost parts).	Average field
High (60–80)	Poorly maintained field + heterogeneous crop structure + low density + high weed presence + significant damage (flooding, insects, cultivated and lost parts).	Poor field
Very High (80–100)	Very poorly maintained field + heterogeneous crop structure + very low density + very high weed presence + very high damage (flooding, insects, cultivated and lost parts).	Very poor field

4.2.3. Strengths and Limitations of the In-Situ Heterogeneity Assessment

In this classification scheme, ultimately, the level of heterogeneity directly informs the field rating, equating the perceived structural and functional characteristics of the field with its agronomic performance. The classification system is designed to be intuitive and straightforward, allowing for quick assessments based on visual inspection and expert judgment by the enumerators. While this approach is practical for field teams, it does introduce subjectivity and inter-observer variability, especially in handling borderline cases. The scheme treats various factors as composite indicators of heterogeneity. These factors themselves are not directly quantified but are instead assessed qualitatively in an intermediate step, with the information then being combined into a single heterogeneity class. As the individual characteristics are not explicitly measured or reported, this leads to a loss of information and detail, potentially masking important variations within the data. This approach also removes the ability to derive the fundamental underlying causes of the observed heterogeneity in the study area. Further, no weighting is applied to the individual factors, meaning that all elements are treated as equally important in determining the overall heterogeneity class. This can lead to inconsistencies in how different enumerators interpret and apply the classification scheme, particularly in cases where the observed characteristics are not clear-cut.

Further this data lacks a temporal dimension, meaning that it does not capture potential changes in heterogeneity over time. While information on a parcel's crop type(s) can be expected to remain static throughout the growing season, the heterogeneity class is a snapshot of the field's condition at the time of data collection and does not account for how heterogeneity may have evolved throughout the season. This is a significant limitation, as management practices and conditions can vary widely and adversely or positively affect field conditions. For example, a field may have been well-maintained at the beginning of the season but may have deteriorated or even abandoned due to adverse weather conditions or pest outbreaks. Conversely, a field that appeared heterogeneous at the start of the season may have improved due to effective management practices. Acknowledging these constraints, particularly the static nature of the in-situ heterogeneity assessment, is an important consideration when working with this data.

4.3. Remote Sensing Data

Satellite imagery from Sentinel-2, PlanetScope and Pléiades, with varying spatial and spectral characteristics are utilized. The utilized spectral bands and their bandwidths are provided in the Annex (see Table 18).

The Sentinel-2 mission, part of the European Copernicus Programme, consists of three sun-synchronous satellites: Sentinel-2A, Sentinel-2B, and the recently launched Sentinel-2C which is set to replace Sentinel-2A (European Space Agency, 2025). The satellites carry a Multispectral Imager Sensor that captures thirteen spectral bands spanning from coastal blue ($\lambda = 443$ nm) over the visible, near infrared to shortwave infrared ($\lambda = 2190$ nm), at varying spatial resolutions of 10m, 20m, and 60m. In a two-satellite configuration it provides around a five-day revisit frequency near the equator with data available since 2016/2017. Sentinel-2 has since been widely employed for land-use landcover cover classification tasks (Phiri et al., 2020).

PlanetScope Analytics 8-band imagery, from commercial smallsat provider Planet Labs, offers enhanced spatial detail at a resampled 3 metre resolution with daily revisit capability. The PlanetScope constellation consists of "flocks" of Dove satellites. They are miniature CubeSats operating in sun-synchronous orbit, designed to work as a coordinated system for continuous Earth observation. The newest generation, called SuperDoves, was first launched in April 2019 and carries the 'PSB.SD' sensor capable of capturing eight spectral bands in the VIS and NIR domain. Imagery is available from mid-March 2020. Specified nadir ground sample distance varies between 3.7 meter to 4.2 meters (Planet, 2023). Despite its advantages, PlanetScope imagery presents technical challenges for scientific applications. As Frazier and Hemingway (2021) note, the constellation approach results in data acquisition from numerous different satellites rather than a single or dual sensor platform. Since individual PlanetScope scenes cover relatively small areas, multiple adjacent scenes must typically be mosaicked together to cover research areas. This creates substantial scene-to-scene variations in reflectance due to different acquisition times, illumination conditions, and subtle sensor differences among the flock. This process can introduce geometric and radiometric inconsistencies. The authors highlight that positional errors can be considerable in PlanetScope imagery. Nevertheless, the combination of high spatial resolution and daily revisit capability offers unique advantages for applications requiring frequent observations, particularly in agricultural monitoring contexts.

For the study area, a very high resolution Pléiades satellite image was specifically commissioned and acquired to enable detailed analysis at sub-meter resolution. The Pléiades constellation, operated by Airbus Defence and Space, consists of two identical satellites (1A and 1B) in sun-synchronous orbit that capture optical imagery with a 20 km swath width at nadir. The sensor system collects data across four multispectral bands - Blue, Green, Red, and Near-Infrared - at 2 m spatial resolution, along with a panchromatic band at 0.5 m resolution that enables pansharpening for enhanced spatial detail (Airbus Defence and Space Intelligence, n.d.). The imagery was provided in pre-processed format, already orthorectified and corrected for off-nadir viewing angles and topographic relief effects. The acquisition date of the image is the 19th of October 2024.

5. Processing of Remote Sensing Data

This chapter details the acquisition and processing workflows implemented for the three satellite imagery sources utilized in this study: Sentinel-2, PlanetScope, and Pléiades. For Sentinel-2 and PlanetScope, the processing workflow was designed to generate consistent time series spanning the agricultural growing season (May-November 2024), requiring procedures for cloud masking, compositing, and gap-filling to overcome the challenges of frequent cloud cover. For the Pléiades' single acquisition (October 19, 2024), pansharpening was performed to maximize spatial detail.

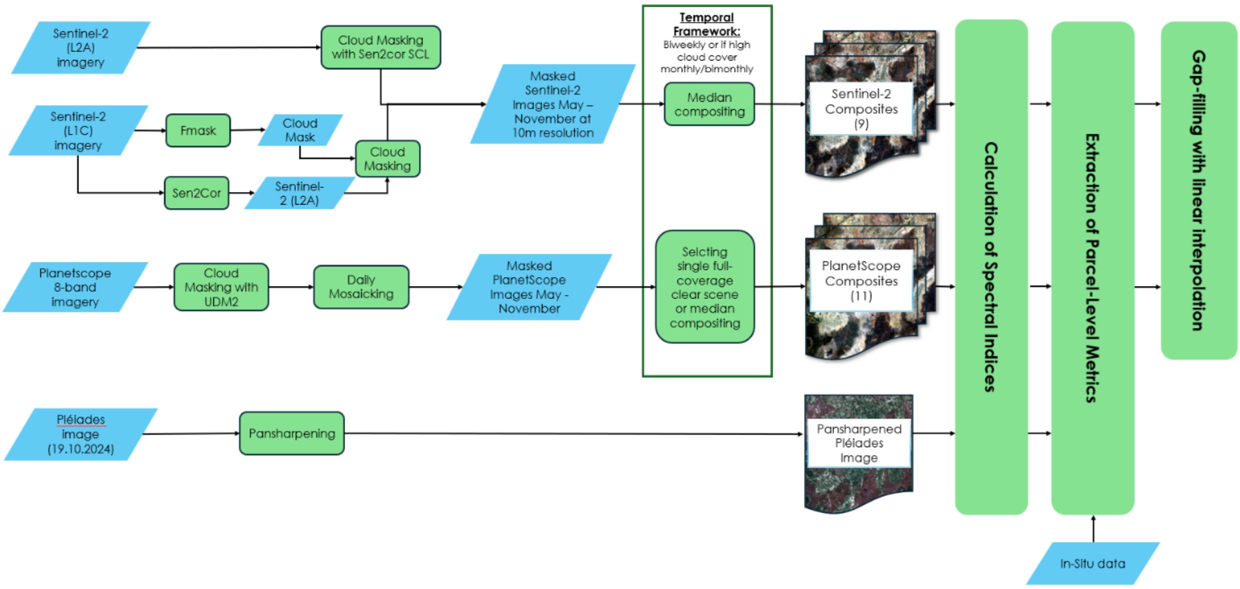

Figure 4: Workflow of EO Data Processing

5.1. Processing of Sentinel-2 Imagery

5.1.1. Data Acquisition

Sentinel-2 imagery was acquired through Google Earth Engine platform using the Earth Engine Python API. The Copernicus Sentinel-2 Collection 1 MSI Level-2A dataset was utilized, which provides orthorectified surface reflectance with atmospheric correction already performed using the Sentinel-2 Correction (Sen2Cor) algorithm (European Space Agency, 2021). A comprehensive image collection was assembled by applying multiple filtering criteria: including spatial intersection with the study area boundary, acquisition dates between May 1 and November 30, 2024, and Sen2Cor-based cloud cover assessment below 80%. This initial filtering retained only potentially usable scenes while excluding very heavily clouded acquisitions. The filtered collection was subsequently clipped to the precise extent of the study area.

Additionally, for the months with persistent cloud cover (June-August), the Sentinel-2 Level-1C Top-of-Atmosphere reflectance product was acquired from the Copernicus Data Space

Ecosystem using an institutional API access script. This approach targeted the MGRS tile covering the study area (29PPQ) and was implemented because Level-1C data enables the application of alternative cloud masking algorithms, specifically Function of Mask (Fmask) 4.3, which has shown to be effective in cases when Sen2Cor's cloud detection underperforms (Sanchez et al., 2020). The Level-1C images were processed and atmospherically corrected using Sen2Cor before applying the cloud masking.

5.1.2. Cloud Masking

Two complementary cloud masking approaches were implemented for Sentinel-2 imagery due to varying cloud conditions throughout the study period.

For the months with lower cloud prevalence (May, September-November), the Sen2Cor scene classification layer was utilized, excluding pixels classified as clouds, thin cirrus, cloud shadows, and defective/saturated pixels. This approach proved adequate for these months but failed to effectively mask clouds for the months of June to August. Alternative cloud masking was attempted using the s2cloudless cloud probability layer, generated by Sentinel Hub, and which is available through Google Earth Engine. This layer provides pixel-wise cloud probabilities (0-100%). Despite testing multiple probability thresholds for cloud detection, none significantly improved the masking effectiveness in these months.

This necessitated the implementation of the Fmask algorithm as a more robust alternative solution for these month (Qiu et al., 2019; Sanchez et al., 2020). Implementation involved processing the Level-1C imagery through Fmask 4.3 algorithm using multiple probability thresholds (10%, 20%, 30%, and 40%), which classified pixels into six categories (clear land, clear water, cloud shadow, snow, cloud, or no observation). Based on visual assessment, the 10% probability threshold provided reliable cloud identification. From these, binary masks were created that were subsequently applied to the Sen2Cor-processed Level-2A surface reflectance product, effectively only retaining pixels that were classified as clear land or clear water.

5.1.3. Resampling and Creation of Composite Images

The 20-meter resolution Sentinel-2 bands (B5, B6, B7, B8A, B11, B12) were resampled to a consistent 10m resolution using bilinear interpolation, which calculates distance weighted averages from the four nearest pixels (ESA, n.d.)

An adaptive temporal compositing framework was implemented to address varying cloud conditions throughout the study period. For months with low or moderate cloud prevalence (May, September, October, November), biweekly (15-day) composites were generated. For the months with extensive masking due to cloud prevalence (June, July and August) different approaches were trialled. The June data was ultimately excluded due to persistent cloud contamination preventing meaningful compositing, while July and August imagery was combined into a single two-month composite to accumulate sufficient valid observations.

All composites were generated calculating the pixel-wise median for the band values across their respective temporal windows. This resulted in 9 composite images distributed across the growing season, capturing different phenological stages. Figure 5 shows the composite images for the different time periods.

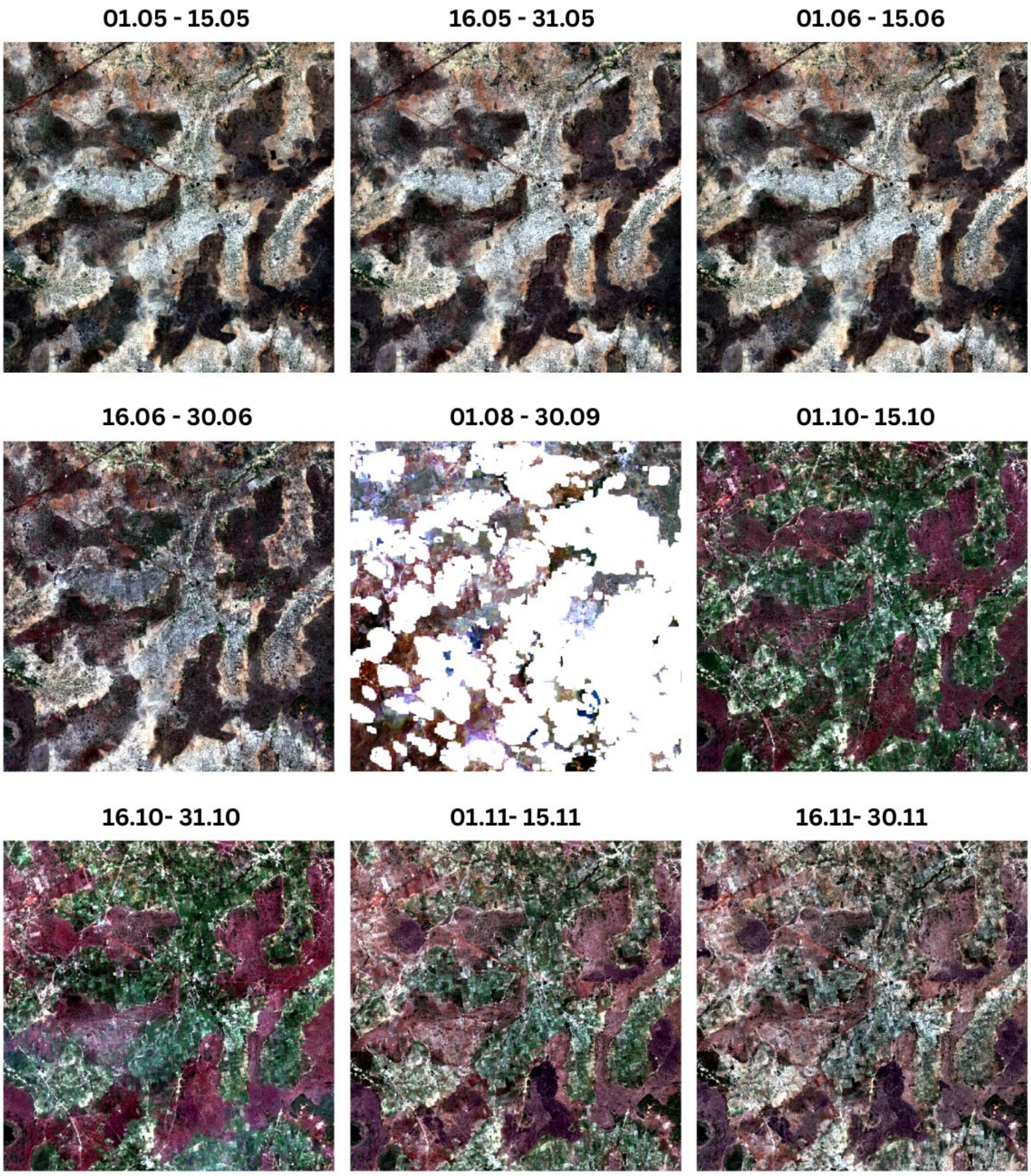

Figure 5: Visualization of 9 Sentinel-2 composite images across the growing season 2024

5.2. Processing of PlanetScope Imagery

5.2.1. Data Acquisition

PlanetScope imagery was obtained through the Planet Explorer platform. The PlanetScope orthorectified 8-band Surface Reflectance (SR) product was selected. The data is delivered in

UTM projection (UTM29N for the study area) with 16-bit radiometric resolution (Planet Labs PBC, 2022). The imagery was filtered for the time period from 1. May to 30. November 2024. The temporal resolution of the SuperDove constellation is approximately 1-4 days during the period of interest in the study area. Image tiles of specific dates were selected manually based on cloud cover and acquisition date, intending to have a balanced distribution of images across the study period.

5.2.2. Cloud Masking

Cloud and quality masking for PlanetScope imagery was implemented using the Usable Data Mask 2 (UDM2) file that accompanies the surface reflectance product. The UDM2 is an 8-band quality assessment layer providing per-pixel binary classifications and associated classification confidence scores (Planet, 2023). Bands represent specific quality attributes: clear pixels, clouds, cloud shadows, heavy haze, light haze, snow, and unusable pixels due to sensor or processing anomalies. The implemented masking procedure excluded all pixels from the scenes that were flagged with any non-clear condition in the mask.

5.2.3. Mosaicking and Creation of Composite Images

Since the PlanetScope imagery is provided in ortho tiles, the individual tiles covering the study area were mosaiced together to create a single image for each date. For each acquisition date, all available tiles were mosaicked using the "last" rule for overlap areas, which prioritizes pixel values from the most recent acquisition when multiple tiles cover the same location.

Following the temporal framework established for Sentinel-2, PlanetScope imagery was organized into biweekly (15-day) time spans for May, June, July, October, and November 2024. The compositing strategy employed a hierarchical approach: if a single cloud-free scene with complete study area coverage existed within a biweekly period, it was selected directly; otherwise, a pixel-wise median composite was generated from all available scenes within that period.

For August, characterized by extensive cloud cover, a monthly composite was created to have sufficient valid observations. September imagery was ultimately excluded for analysis due to persistent cloud contamination that prevented the creation of a reliable composite.

The band values of the images were then converted from Digital Numbers (DN) to reflectance values. Figure 16 (see in Annex) shows the composite images for the different time periods.

5.3. Pansharpening of Pléiades Image

Pansharpening was performed on the Pléiades image to enhance the spatial resolution of the multispectral bands from 2 m to 0.5 m. The BundleToPerfectSensor application of the Orfeo Toolbox (OTB) was used. The OTB is an open-source project for processing remote sensing data, and it provides a range of tools for image processing, including pansharpening (Inglada & Christophe, 2009).

For the pansharpening process, the 0.5 m panchromatic band (PAN) and 2 m multispectral bands (XS) were selected as inputs. Since the Pleiades bands are already coregistered, the application automatically switched to Pléiades High Resolution (PHR) mode, which relies on a simple affine transformation rather than complex sensor models. Linear interpolation was selected for resampling the multispectral (XS) bands to match the higher resolution. The Ratio Component Substitution (RCS) fusion method was employed, which applies a low-pass filter to the panchromatic band, creates a ratio between the multispectral bands and the filtered panchromatic image, and then multiplies this ratio by the original high-resolution panchromatic data (Mhangara et al., 2020; Orfeo ToolBox Team, 2018). The final output is a 0.5 m resolution four-band raster (RGB and NIR) that effectively combines the high spatial detail of the panchromatic band with the spectral characteristics of the multispectral bands (see Figure 3).

5.4. Calculation of Spectral Indices

A suite of spectral indices was calculated for Sentinel-2, PlanetScope, and Pléiades imagery to track crop phenology and enhance crop type classification capabilities. Primarily vegetation indices were selected that leverage reflectance information from visible and near-infrared regions, as these can exploit the chlorophyll absorption and plant structural features for assessing vegetation states. (Montero et al. 2023). These include the Normalized Difference Vegetation Index (NDVI) which is widely established in the field for quantifying vegetation greenness, supplemented by variants like the Soil-Adjusted Vegetation Index (SAVI) that incorporates soil brightness correction factors for soil interference, and indices like the Plant Senescing Reflectance Index (PSRI) that has advantages over NDVI in capturing plant senescence (Anderegg et al., 2019). Additionally, indices targeting non-vegetation features, including the Normalized Difference Water Index (NDWI) and Normalized Difference Built-up Index (NDBI) were included, for improved distinction between different landcover classes.

Table 19 (see Annex) lists the spectral indices, and the involved bands for calculation. Each sensor's specific capabilities informed the spectral index selection. Sentinel-2's broad spectral range allowed calculation of all indices, including those using red-edge bands for vegetation characterization and those utilizing shortwave infrared bands. PlanetScope supports most vegetation indices with its single red-edge band, while Pléiades' limited range (blue, green, red, NIR) only accommodated few indices. All calculations followed standardized formulations from the Awesome Spectral Indices catalogue (Montero et al., 2023).

5.5. Linear Gap-Filling of Satellite Time Series

To address data gaps in the satellite imagery time series resulting from cloud masking and insufficient coverage for complete compositing of scenes, a temporal gap-filling procedure was implemented for both Sentinel-2 and PlanetScope time series. This processing step was necessary for performing classification, since classifiers such as Random Forest require complete observations across all time points to construct the multidimensional feature space.

The gap-filling was performed using linear temporal interpolation, where missing pixel values were estimated by calculating a weighted average between the nearest valid observations before and after each gap. This weighting is proportional to the temporal distance between the target date and the available observations. The implementation utilized the interpolate_na() function from the xarray library. Prior to interpolation, the data was structured into a four-dimensional array with dimensions (time, band, x, y), with each composite assigned a representative timestamp corresponding to the midpoint of its compositing period. The interpolation was performed independently for each spectral band and derived index. Figure 17 and Figure 18 show the gap-filled NDVI time series for Sentinel-2 and PlanetScope.

5.6. Generation of a Tree Mask from Pléiades Image

A critical component in the analysis of agricultural landscapes with agroforestry elements is the accurate delineation of tree crowns within cropland parcels. To address this, a tree mask needed to be developed using the pansharpened 0.5 m Pléiades image. Recent studies have demonstrated significant potential for mapping individual trees using submeter resolution satellite imagery (Li et al., 2023; Wagner et al., 2018).

Initially, a pre-existing tree mask developed by Brandt et al. (2020) was evaluated for potential utilization. This product, created through deep-learning techniques, had successfully mapped billions of individual trees and shrubs across sub-Saharan desert and Sahelian savanna landscape types. However, upon assessment, several limitations were identified. The product was generated using imagery collected between 2005 and 2018, resulting in temporal misalignment with current field conditions. Furthermore, while tree crowns were indeed extensively mapped within the study area, the product exhibited insufficient accuracy for the purposes of this research. Specifically, smaller trees were not adequately captured, and significant locational mismatches were observed between the mapped crowns and the trees identified in the available satellite imagery.

Subsequently, a segmentation-classification approach was attempted using Trimble eCognition software. The methodology employed a multiresolution segmentation algorithm, which functions as a bottom-up segmentation process based on pairwise region merging. This algorithm iteratively merges pixels or existing image objects to minimize average heterogeneity while maximizing homogeneity across the resulting segments (Trimble Inc., n.d.). The segmentation parameters were calibrated using selected spectral bands and derived indices including NDVI, NDWI, and NDBI. Following segmentation, training samples were to be selected to train a supervised classifier for tree crown classification.

Despite trialling various parameter configurations, the segmentation results did not achieve satisfactory tree crown delineation. A primary limitation is the timing of the Pléiades image acquisition, which coincided with a period of active crop cover. This temporal characteristic contrasted significantly with the approach employed by Brandt et al. (2020), who strategically utilized dry season imagery when only woody vegetation remains photosynthetically active, creating a strong spectral contrast between trees crowns and the background. The presence of active crop cover interfered with the spectral separation between tree crowns and surrounding vegetation, hampering the segmentation process. Additionally, the computational demands of processing high-resolution imagery in eCognition presented practical implementation challenges. Alternative approaches using Python libraries such as scikit-image were also explored but similarly failed to produce acceptable results for operational use in this study.

The solution was ultimately provided by PhD student Abdoulaye Guindo, who successfully developed a tree mask using a comparable segmentation-classification approach based on the Pléiades image. While not achieving perfect crown delineation, this mask (see Figure 19 in Annex) captured tree crowns within the study area with sufficient accuracy for analysis. The mask was cropped to the in-situ parcel boundaries and served dual analytical purposes: it was integrated in the parcel-level analysis and enabled the systematic exclusion of tree-covered pixels using variable coverage thresholds, thereby allowing assessment of the removal of tree covered pixels on crop classification performance.

While the tree mask provides information on woody vegetation distribution within cropland, the methodological approach was subsequently expanded during crop type classification to more holistically address intra-parcel heterogeneity. Leveraging the same Pléiades image, a binary cropland/non-cropland mask was developed that identifies trees but also captures bare soil patches, built-up structures, and other non-crop elements within agricultural parcels. This very high-resolution mask was used for refining the validation methodology by enabling the quantification of cropland proportion within each Sentinel-2 and PlanetScope pixel, allowing for effectively isolating classification accuracy from the confounding influence of within-field non-crop elements. The methodology for creating this mask is expanded upon in Chapter 7.

5.7. Extracting Parcel-Level Metrics from Remote Sensing Data

Zonal statistics were calculated for each agricultural parcel using the exactextractr package in R (Baston, 2024) to extract spectral properties across all available imagery. Statistical calculations were performed on the time series composite images (non-gap-filled) of Sentinel-2 and PlanetScope, as well as the single pansharpened Pléiades image. To minimize edge effects and mixed pixels, sensor specific inward buffers were applied: 5m for Sentinel-2, 3m for PlanetScope, and 0.5m for Pleiades. Five meter was chosen for Sentinel-2, because it was deemed necessary for sustaining sufficient pixels per parcel at the time, however an inward buffer of 10 meters (which was applied for crop type classification) would have been more consistent in masking out mixels at the edges of the parcels. The process was conducted two times: once incorporating the high-resolution tree mask to exclude tree-covered pixels and again without tree masking to provide comparative statistics (see Table 20 in Annex for comparison of parcel-level NDVI for Pléiades). The tree-filtered data was used for the subsequent analysis to minimize the influence of tree vegetation on parcel spectral signatures.

For each parcel and image date, six statistical metrics were calculated per spectral band and derived index: mean, standard deviation, coefficient of variation (CV), median, minimum, and maximum values. The mean values represent average reflectance of parcels, while the CV (calculated as the ratio of standard deviation to mean) quantifies relative variability independent of magnitude.

In parallel, geometric and spatial metrics were calculated for all (unbuffered) parcels. The resulting data layers comprise, for each sensor, multi-temporal spectral statistics and spatial metrics for each cropland parcel in the in-situ dataset.

6. Parcel-Level Heterogeneity: An Integrated Analysis of In-Situ and Satellite Observations

This chapter presents a detailed parcel-level analysis integrating the in-situ data with multi-sensor satellite imagery to characterize the parcels and their inherent heterogeneity. It first provides a descriptive analysis of the in-situ dataset, examining the distribution of crop types, parcel sizes, and heterogeneity classes. These ground-based observations are then systematically integrated with satellite-derived spectral information to identify relationships between field-assessed heterogeneity and remotely sensed parcel metrics. The analysis explores how these relationships manifest across different spatial resolutions and temporal stages of the growing season. The chapter concludes with a critical discussion of the findings and their implications for informing subsequent crop type classification methodologies.

6.1. Descriptive Statistics of the In-Situ Dataset

The in-situ dataset comprises a total of 825 agricultural parcels of which 821 parcels have information on the primary crop type. The combined area of the parcels is approximately 498 hectares, with individual parcel sizes ranging from 0.03 to 5.49 hectares (see Table 3). The median parcel area is 0.49 ha, and the mean area is slightly higher at 0.60 ha. The distribution shows a right-skewed pattern, with a few parcels being significantly larger than the average. The data shows that typical farms in the region are smallholder in scale.

Table 3: Summary statistics on parcel area

Min Area (m^2)	Max Area (m^2)	Mean Area (m^2)	Median Area (m^2)	Standard Deviation
310.68	54 914.95	6 041.09	4 919.18	4 892.00

6.1.1. Crop Type Distribution and Cropping Practices

The dataset features a multitude of crop types, with a few being predominant (see Table 21 in Annex). Sorghum is the most prevalent primary crop, found on 339 parcels (including both monocropped fields and multi-cropped parcels where sorghum is the dominant crop), followed by maize (107 parcels), millet (93 parcels), cotton (77 parcels), sesame (46 parcels) and peanut (24 parcels). 108 parcels are classified as being cultivated primarily with "Other Crops", for which the specific crop type is not specified. Also, beans, okra, rice, cowpea, tomato and watermelon parcels are included in the data set, however only with a feature count of less than 10 parcels each.

The overwhelming majority of surveyed parcels are monocropped, i.e. planted with a single crop species. 59 parcels (~7%) were classified as mixed cropped and 20 as associated/intercropped (~2%). 44 parcels (~5%) were classified as empty plots, meaning that no crops were grown on these parcels at the time of data collection.

For the limited number of fields that are multi-cropped (intercropping or mixed cropping), a common combination is maize with sorghum, with a feature count of nine. Sorghum is commonly intercropped with "Other Crops" (15 parcels), so is maize and peanut. Other multi-crop pairings occurred in only a few cases. Overall, the data shows that while single-crop fields are the norm, a notable proportion of parcels possesses crop diversity at the parcel level.

6.1.2. Relationship between Crop Type and Parcel Area

Examining the distribution of parcel area for each primary crop reveals clear differences in typical field size between crop types (see Figure 6 and Table 22 in Annex). Cotton fields are the largest on average, with a mean parcel size of ~0.82 ha and a median of ~0.65 ha. Sorghum fields are smaller on average, with a mean area of ~0.60 ha and a median of ~0.46 ha. For both crops, a few larger parcels over 2 hectares significantly skew the mean. Maize and sesame display mean values of 0.65 ha and 0.60 ha respectively. "Other Crops" and millet parcels are on average under 0.55 ha and peanut fields are the smallest, with a mean area of under 0.44 ha. These differences indicate that crop cultivation choice in the region is related to the size of the parcel, with larger fields being more likely to be planted with cash crops like cotton, while smaller fields more commonly used for subsistence crops.

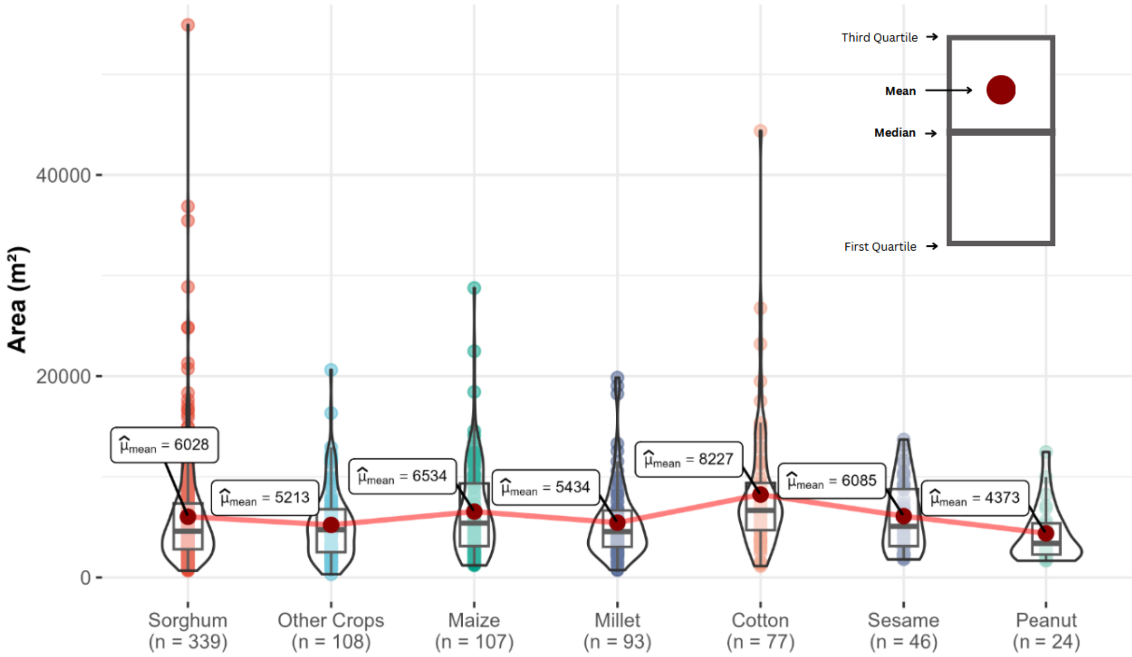

Figure 6: Violin plot of parcel area for main crop types

6.1.3. Distribution of In-Situ Heterogeneity Classes Across Crop Types

Analysing the heterogeneity classes assigned to the parcels reveals that the majority of fields are attributed as having a low to medium heterogeneity (see Table 4).

Table 4: Proportion of parcels in each heterogeneity class

Heterogeneity Level					
Very Low	Low	Medium	High	Very High	NA
6.3%	42.2%	24.1%	15.0%	7.3%	5.1%

Considering only monocropped parcels (see Table 23 in Annex), for most crop types, most parcels fall into the low or medium heterogeneity classes. For instance, 60% of maize parcels are rated as having a low degree of heterogeneity, followed by ~24 % of parcels rated as medium, ~8% for high and very high respectively, and none assigned as very low. Also, for sorghum, cotton, and peanut fields, most parcels are classified as low heterogeneity, on the order of 40-50% of the parcels. Sorghum and cotton also show a significant fraction of parcels rated with medium and high heterogeneity. The very low and very high heterogeneity class are proportionally less common across the major crops, generally accounting for single digits or low double-digit percentages. Some noteworthy exceptions are rice, of which 75% of parcels were rated as very low heterogeneity and beans for which one third of the parcels were rated as very high. Since the feature count for these two crops is very low, the results are however limited in their interpretability. The heterogeneity distribution of millet parcels is slightly more varied with a comparatively higher proportion of parcels rated as medium (~23%) and high (~28%) heterogeneity. The general "Other Crops" category also skews towards a higher heterogeneity: only ~19% of those fields are classified as low heterogeneity, while the largest share (~46%) are in the medium heterogeneity class. Low to medium heterogeneity classification appears to be the norm for most monocropped parcels.

According to the heterogeneity classification scheme, the multi-cropped parcels could be expected to tend to have higher heterogeneity ratings (since multiple crop types would present a less uniform crop structure). The data (see Table 24) however, suggests that this is not the case. For the maize-sorghum mixed parcels, ~89% of parcels were rated as low heterogeneity, with 11% rated as medium. For sorghum combined with "Other Crops", this is less pronounced, but also here 60 % of parcels were rated as low heterogeneity, and ~27% as medium and only ~13% as high. Only for the cotton-sorghum mixed parcels (only 3 parcels) a higher heterogeneity degree is recorded with two out of the three parcels rated as medium heterogeneity. This suggests that by the criteria used for the classification by the enumerators, the intercropped and mixed-crop parcels are not assessed as more heterogeneous than the monocropped ones. In fact, based on the limited data available on multi-cropped parcels, the data suggests that they are less heterogeneous than the monocropped ones.

6.1.4. Relationship between Parcel Area and Heterogeneity Level

Relating parcel area to the heterogeneity levels reveals a positive relationship between higher heterogeneity ratings and parcel size. Mean parcel area demonstrates an upward trend across heterogeneity classes, increasing from 5 466 m² in very low heterogeneity parcels to 7 067 m² in very high heterogeneity parcels. Similarly, median values follow this pattern, rising consistently from 4 272 m² to 5 726 m² across the heterogeneity levels. These findings suggest that parcel size may be an important determinant of assessed within-field heterogeneity.

Table 5: Parcel area statistics by heterogeneity level

		Area (m^2)				
Heterogeneity Level	**Parcel Count**	**Mean**	**Median**	**Min**	**Max**	**SD**
very low	52	**5 466**	4 272	739	20 618	4 618
low	346	**5 945**	4 805	660	54 915	5 282
medium	198	**6 287**	4 964	795	36 873	5 038
high	123	**6 126**	5 287	768	24 877	3 733
very high	55	**7 076**	5 726	1 346	35 466	5 576

6.2. Multi-Sensor Assessment of Parcel Characteristics by Heterogeneity and Crop Type

Following the characterization of parcels through the in-situ data, this section addresses the first research objective by investigating the correspondence between field-assessed heterogeneity classifications and satellite-derived spectral-temporal signatures across different crop types and sensor resolutions.

First, the NDVI signatures for the six predominant crops (cotton, millet, sesame, sorghum, maize, and peanut) are analysed, focusing on how these signatures evolve throughout the growing season and differ across heterogeneity levels. The analysis first establishes baseline NDVI patterns for each crop type by averaging values across all parcels, then stratifies these by heterogeneity class to identify potential relationships between field-assessed heterogeneity and spectral response. While the entire growing season is considered, particular emphasis is placed on late October 2024, which lies in temporal proximity to the in-situ data collection period.

Intra-parcel spectral variability is analysed using the CV calculated from NDVI pixel values within each parcel. This metric serves as an indicator of internal spectral heterogeneity, with higher CV values suggesting greater relative variability. The relationship between satellite-derived parcel level spectral variability and field-observed heterogeneity classification is examined across all three sensors.

The analysis is complemented by systematic visual examination of selected parcels with available satellite imagery from late October alongside available ground photographs. This component focuses on four crop types (cotton, sorghum, maize, and millet) that possess sufficient representation across heterogeneity classes to enable meaningful comparison. Although ground photographs (primarily from early November 2024) do not perfectly align with satellite acquisition dates, they provide valuable reference for field conditions. This visual assessment aims to establish how parcel characteristics vary with in-situ heterogeneity levels and how these manifestations differ across spatial resolutions, thereby providing a more comprehensive understanding of heterogeneity expression in remotely sensed data.

6.2.1. Crop-Specific Mean NDVI Signatures

Figure 7 and Figure 8 below show the mean NDVI signatures of the six main crops, with the parcel-level NDVI averaged across all monocropped parcels for each crop type. The temporal NDVI profiles of Sentinel-2 and PlanetScope show a clear phenological pattern with a trajectory of an increase of NDVI values from the beginning of June, reaching peak NDVI around October, and then declining subsequently towards the end of the growing season. While generally following the same pattern, the NDVI values for the different crops show some distinct differences in their peak NDVI values and the timing of these peaks. While Sentinel-2 provides clear temporal trends, PlanetScope data displays greater temporal variability, likely attributable to radiometric inconsistencies between scenes. While this can affect the general trajectory of NDVI, differences between the crop types can still be derived.

Cotton consistently demonstrates the highest NDVI values across sensors (peaking at approximately 0.68 in Sentinel-2 data), followed by sorghum and maize (0.63-0.64), while peanut and millet show values of approximately 0.62 for Sentinel-2. Sesame exhibits a distinct phenological pattern with a later peak in the second half of October. The period of greatest crop separability occurs during August-September and at peak NDVI in October. Analysis of late October data for Pléiades (see Figure 20 in Annex) confirms these crop-specific NDVI patterns, with the crops displaying similar relative positioning as seen with Sentinel-2 and PlanetScope.

Notably, substantial variability in NDVI values exists between parcels of each crop type throughout the growing season (see Figure 21 and Figure 22 in Annex), as evidenced by the high standard deviations (0.05-0.15) across all crops. Cotton demonstrates the most consistent conditions with a relatively narrow distribution of NDVI values, while sesame and maize exhibit considerably wider ranges, suggesting significant differences in field conditions or phenological stages among parcels of the same crop type.

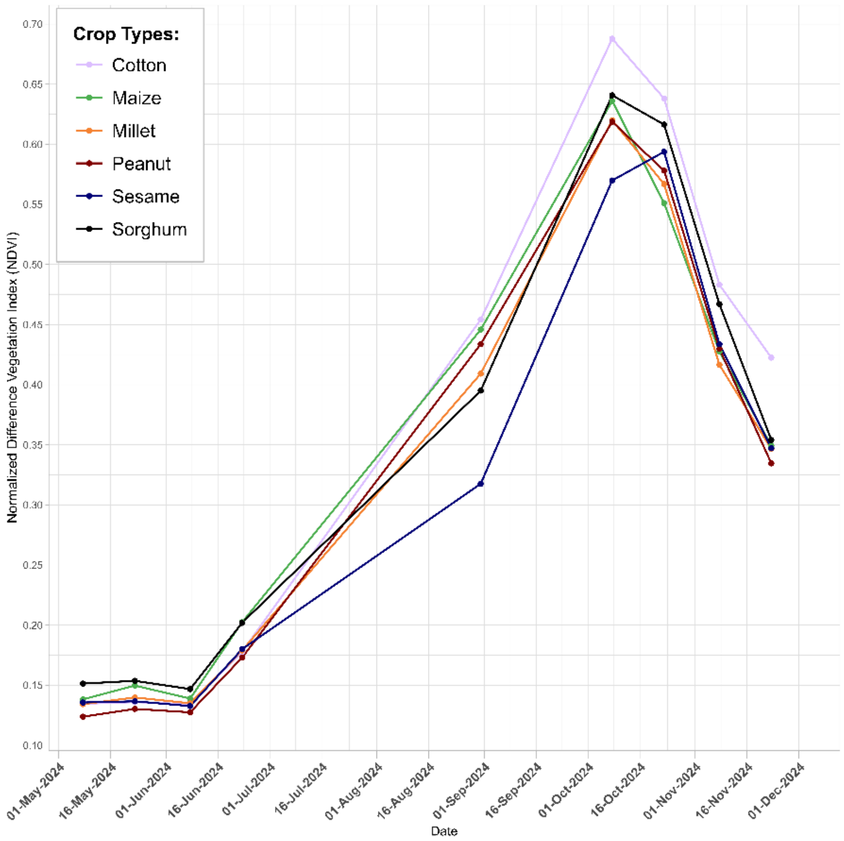

Figure 7: Sentinel-2 - Temporal profile of mean NDVI for main crop types

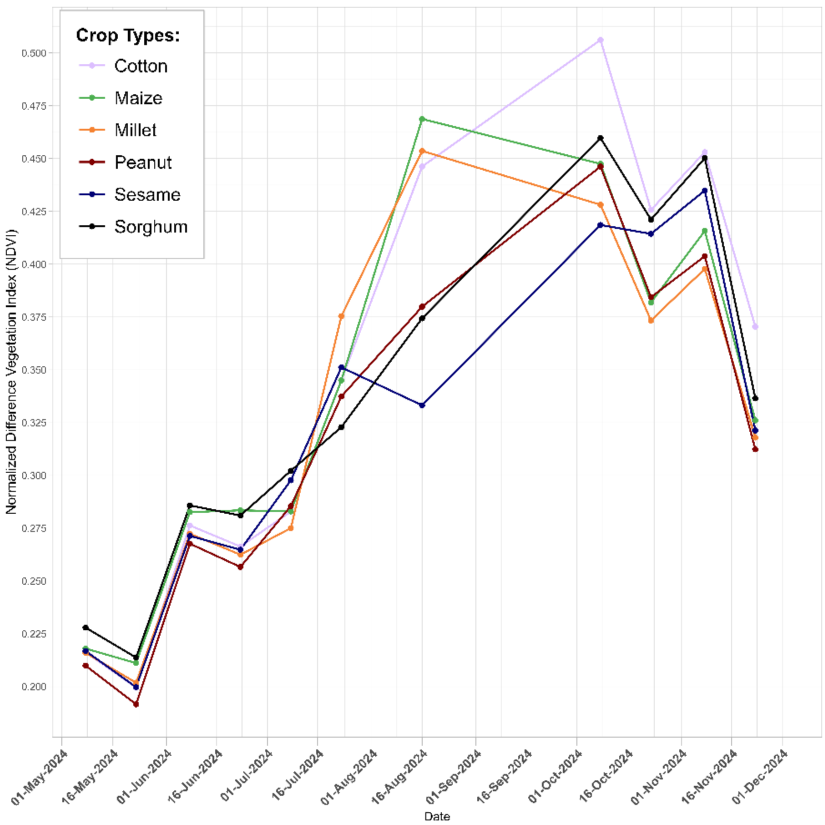

Figure 8: PlanetScope - Temporal profile of mean NDVI for main crop types

6.2.2. NDVI of Parcels across Heterogeneity Levels

The mean NDVI values of the six main crops in the different heterogeneity classes were analysed to understand if and how the NDVI values of the parcels vary with the level of intra-parcel heterogeneity recorded in the in-situ data. Representation across heterogeneity classes varies considerably (see Table 25 in Annex), with most crops having disproportionately more parcels in the "low" class, which necessitates cautious interpretation, particularly for classes with minimal representation. Considering these limitations, both Sentinel-2 and PlanetScope data reveal that mean NDVI values at different heterogeneity levels generally follow similar phenological patterns, though with variations in magnitude that are crop-specific and temporally dynamic.

Analysis of temporal trajectories (see Figure 9 for Sentinel-2 below and Figure 23 for PlanetScope in Annex) shows complex and inconsistent relationships between heterogeneity classes and NDVI values. For cotton, heterogeneity classes display similar NDVI values early in the growing season, with the very high heterogeneity class showing markedly higher values during August-September in Sentinel-2 data, though this pattern was not observed in the PlanetScope data. Millet exhibits an inverse relationship, with the lowest heterogeneity class showing the highest NDVI values at peak season. Sesame demonstrates the clearest stratification, with NDVI values consistently increasing from very low to very high heterogeneity classes during October-November across both sensors. Also, maize displays a consistent pattern of higher NDVI values in higher heterogeneity classes throughout the growing season. Sorghum, having the most balanced representation across heterogeneity classes, showed minimal overall NDVI differentiation between heterogeneity levels. For this crop the very low heterogeneity class shows slightly higher NDVI values in the beginning of October and the lowest relative value in November, where the very high heterogeneity class shows the highest NDVI values.

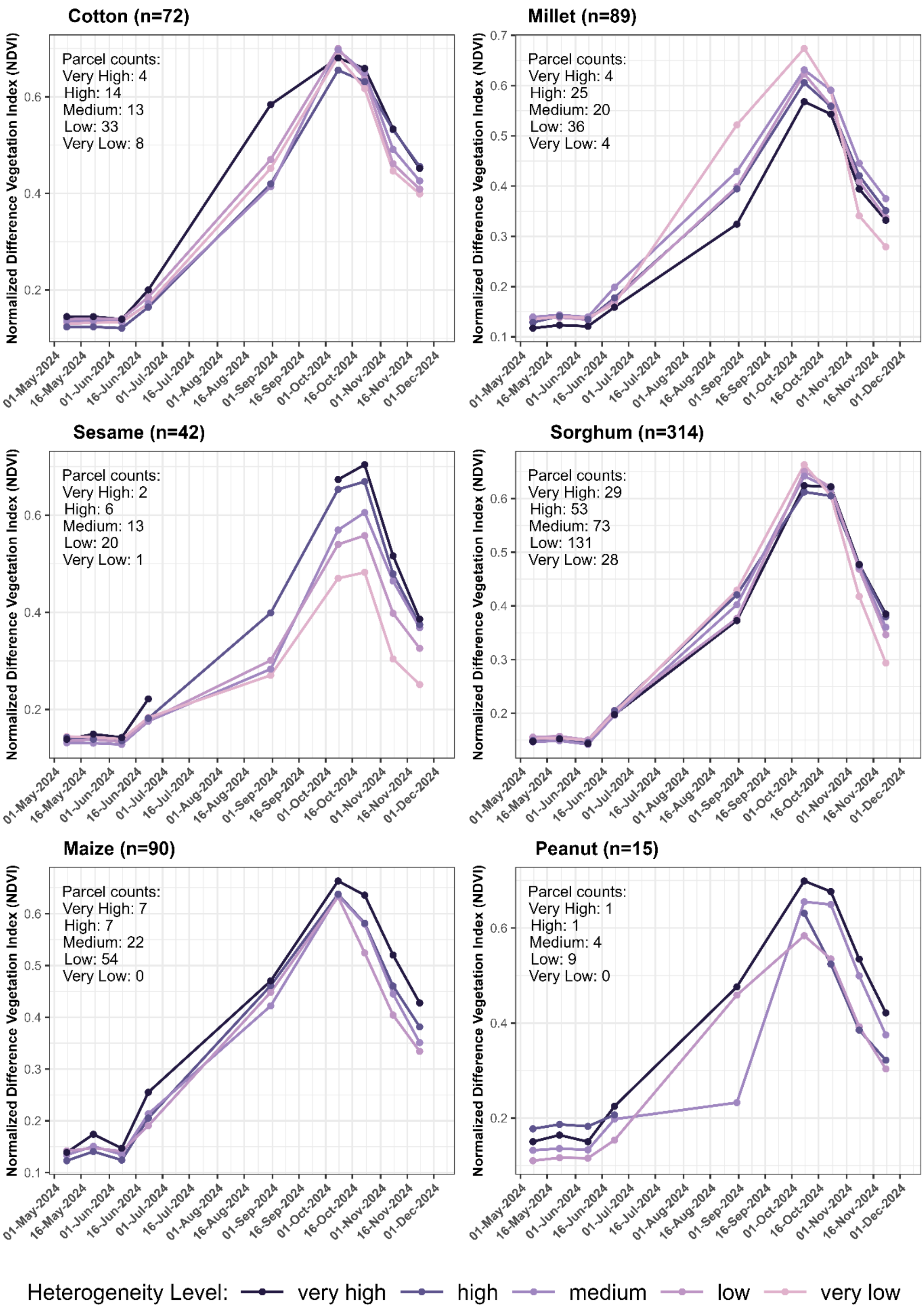

Figure 9: Sentinel-2 - Temporal Profile of NDVI by Heterogeneity Level for Main Crop Types

Late October analysis across all three sensors (for Sentinel-2 see Table 26, PlanetScope Table 27, and Pléiades Table 28 in Annex) provides a focused examination of parcel values shortly before the in-situ data collection period. Sesame and maize demonstrate the most consistent patterns, with mean NDVI steadily increasing from very low to very high heterogeneity classes across all sensors, suggesting greater green biomass in more heterogeneous parcels at this point in time. Cotton shows a similar but weaker trend with some sensor-dependent variations, particularly a non-linear response in Sentinel-2 and Pléiades data where medium heterogeneity parcels exhibit elevated values. For millet, the relationship varies between sensors, but with consistently high relative NDVI values in the medium heterogeneity class. Sorghum displays notably consistent NDVI values across all heterogeneity classes.

Analysis of standard deviations reveals additional patterns about the dispersion of NDVI values within heterogeneity classes. For cotton, the medium heterogeneity class showed the lowest standard deviation across sensors, while the very high heterogeneity class consistently demonstrated greater standard deviation, indicating more variable conditions. For sesame, parcels with very high heterogeneity showed comparatively low standard deviation, while the low heterogeneity class exhibited the highest variability. Sorghum displayed a clear upward trend in standard deviation with increasing heterogeneity, while maize shows the highest standard deviation in the low heterogeneity class and lowest in the very high heterogeneity class.

These findings demonstrate that the relationship between field-assessed heterogeneity and parcel-level NDVI is highly crop-specific, temporally dynamic, and in some cases inconsistent across sensors. While some crops (sesame, maize) exhibit clear patterns within specific timeframes, others (sorghum) show minimal NDVI differentiation between heterogeneity classes.

6.2.3. Coefficient of Variation of NDVI by Crop Type and Heterogeneity

Figure 10 for Sentinel-2 (see below) and Figure 24 for PlanetScope (see in Annex) show the temporal profile of the CV of NDVI for the different crop types and heterogeneity classes. The data for Sentinel-2 and to a lesser magnitude PlanetScope shows that there seems to be a clear seasonal pattern of the highest CV values at the beginning of the growing season around the end of May, with a decrease until September to October, and then again, an increase towards the end of the growing season. This pattern is consistent across the crop types, showing that the intra-parcel NDVI variability is highest at the beginning of the growing season, where the parcels NDVI values are low, and lowest around the time of peak NDVI.

The relationship between field-assessed heterogeneity and CV values varies substantially by crop type and changes dynamically throughout the growing season. For cotton, early-season data (May-June) shows a positive association, with the very high heterogeneity class exhibiting the highest CV values (>40% for Sentinel-2, >20% for PlanetScope). However, this pattern collapses by August-September, with heterogeneity classes showing minimal differentiation. Millet displays notably elevated CV values throughout most of the season for the very high heterogeneity class however does not show a clear overall stratification pattern by class. The data for sesame shows an inverse relationship by late season, with the very high heterogeneity class showing lower CV values than other classes but increasing again at the end of November.

For sorghum the very high heterogeneity class shows the lowest mean CV values in May and June, with the very low heterogeneity class showing the highest values. The very low heterogeneity class subsequently drops in mean CV values around August-September and then increases in relation to the other classes towards the end of the season. Despite these temporal shifts, no consistent stratification pattern is observed across heterogeneity classes, with mean CV values of most classes remaining close throughout most of the growing period. Maize similarly shows no clear stratification between the heterogeneity classes and the mean CV values. The very high heterogeneity class shows the comparatively lowest mean CV values in May and June, and at the end of the season. The high heterogeneity class however shows comparatively high mean CV values in the same time period.

Late October analysis across all three sensors (see Table 29 for Sentinel-2 and Table 30 for PlanetScope in Annex and Table 6 for Pléiades below) highlights the influence of spatial resolution on CV measurements. As could be expected, Pléiades (0.5m) consistently records the highest mean CV values, followed by PlanetScope (3m) and Sentinel-2 (10m), reflecting that higher spatial resolution captures finer-scale internal variability. This resolution effect was consistent across all crop types and heterogeneity classes, with Pléiades values ranging typically 3-5 percentage points higher than corresponding Sentinel-2 measurements.

Crop-specific patterns in late October further highlight the complex relationship between field-assessed heterogeneity and spectral variability. Cotton shows no clear trend in CV values across heterogeneity classes with Pléiades data, though Sentinel-2 and PlanetScope indicate highest variability in the medium heterogeneity class. Sesame exhibits a negative association, with CV values decreasing as heterogeneity increases. Maize shows relatively stable CV values across low to high heterogeneity classes, but with a marked decrease in the very high class. Sorghum displays no discernible pattern across heterogeneity levels in this time.

Additionally, the high standard deviations and ranges of CV values within heterogeneity classes for many crops (e.g. Sorghum) indicate that parcels with identical heterogeneity ratings often display substantially different degrees of internal spectral variability.

Most significantly, the analysis reveals that field-assessed "high" or "very high" heterogeneity does not consistently translate to higher pixel-level CV. For some crops (sesame, maize), the highest heterogeneity class exhibited the lowest CV values, suggesting these fields were paradoxically more spectrally uniform despite being visually assessed as highly heterogeneous.

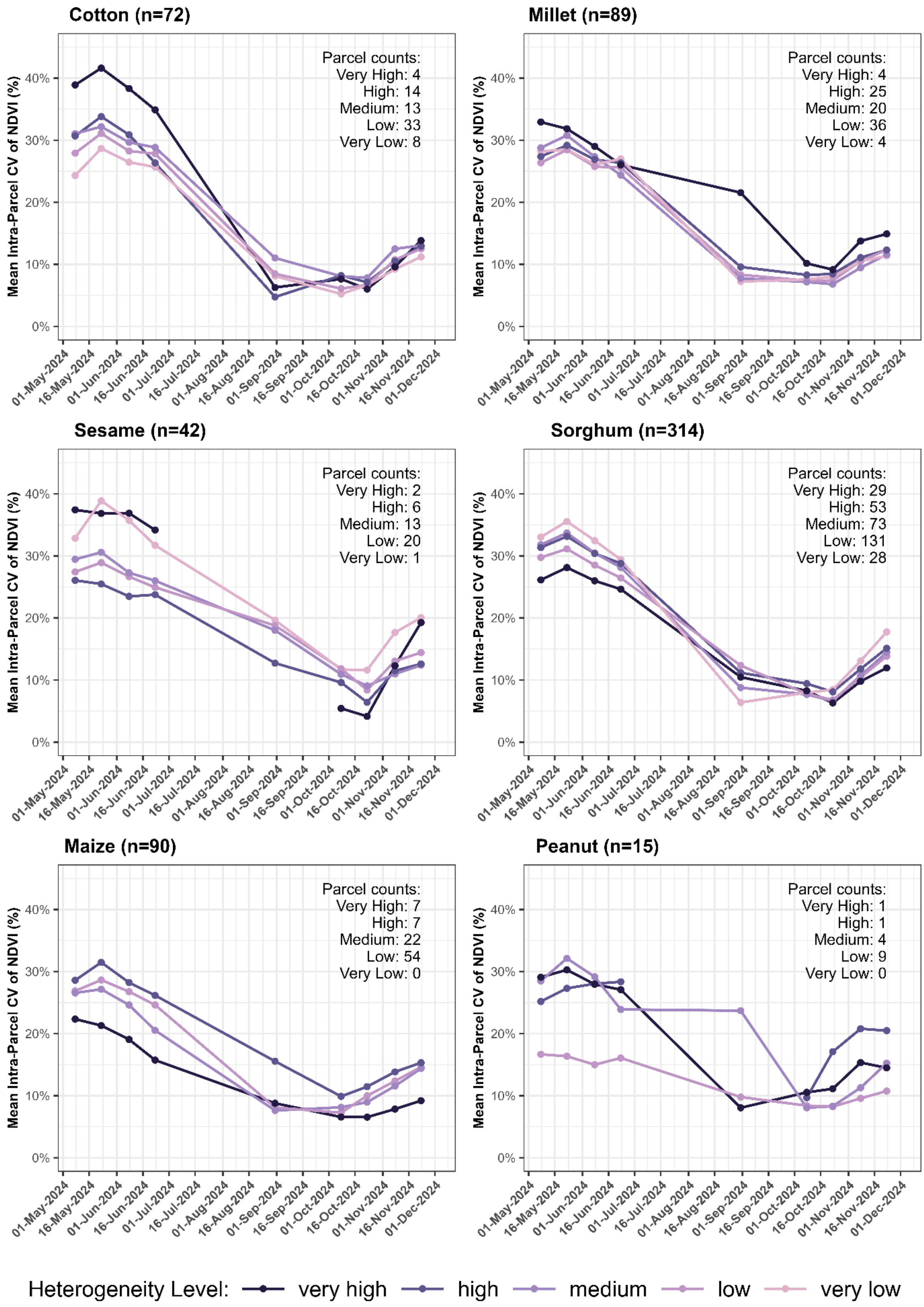

Figure 10: Sentinel-2 - Temporal Profile of Parcel Coefficient of Variation of NDVI by Crop Type and Heterogeneity Level

Table 6: Pléiades Data from Late October - Summary Statistics of Parcel Coefficient of Variation of NDVI by Crop Type and Heterogeneity Level

Pléiades data for October 19, 2024

	Cotton					Millet					Sesame				
	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max
very low	**11.1%**	11.2%	2.1%	7.0%	14.5%	**13.6%**	13.8%	2.7%	10.3%	16.5%	**15.0%**	15.0%	NA	15.0%	15.0%
low	**10.9%**	10.4%	3.0%	6.1%	17.0%	**12.4%**	12.3%	4.2%	4.8%	25.2%	**13.6%**	12.1%	5.1%	8.5%	31.8%
medium	**11.8%**	11.6%	4.5%	6.7%	25.1%	**11.5%**	11.4%	3.4%	5.7%	17.0%	**12.3%**	11.6%	3.9%	7.3%	20.4%
high	**11.3%**	10.7%	3.2%	6.4%	16.1%	**12.6%**	11.9%	4.2%	6.0%	22.8%	**9.9%**	10.1%	2.6%	6.3%	13.0%
very high	**11.5%**	11.7%	5.2%	5.1%	17.6%	**14.1%**	13.6%	4.3%	10.2%	19.3%	**9.6%**	9.6%	3.9%	6.8%	12.4%

	Sorghum					Maize					Peanut				
	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max
very low	**13.6%**	12.8%	5.5%	5.2%	25.2%	**NA**	NA	NA	NA	NA	**NA**	NA	NA	NA	NA
low	**11.2%**	10.4%	4.8%	1.4%	37.7%	**15.7%**	14.1%	5.6%	7.2%	33.4%	**11.6%**	10.5%	3.2%	8.3%	16.7%
medium	**10.8%**	10.1%	4.2%	5.1%	23.4%	**13.6%**	12.7%	3.9%	8.3%	22.0%	**11.2%**	10.8%	2.3%	9.0%	14.4%
high	**12.6%**	11.9%	4.7%	4.8%	24.0%	**14.9%**	13.3%	5.8%	7.9%	24.0%	**24.1%**	24.1%	NA	24.1%	24.1%
very high	**10.7%**	9.8%	4.3%	4.2%	25.3%	**10.4%**	9.3%	3.5%	6.0%	16.5%	**19.0%**	19.0%	NA	19.0%	19.0%

Metrics (Mean, Median, Standard Deviation, Min, Max) represent summary statistics of the parcel-level Coefficient of Variation of NDVI.

6.2.4. Visual Comparison of Parcels at different Heterogeneity Levels

To complement the spectral metrics with visual evidence, selected parcels were examined using satellite imagery alongside available ground photographs. Late October 2024 satellite imagery was selected to coincide with the availability of Pléiades VHR data, while acknowledging that ground photographs, primarily collected in early November 2024, represent a temporal offset that may include intervening harvest activities or other on-field changes. Despite this temporal misalignment, the photographs provide valuable contextual reference for field conditions during the in-situ assessment period. The investigation focuses on cotton, sorghum, maize, and millet, crops selected for their predominance in the dataset and sufficient representation across heterogeneity classes, enabling meaningful within-class comparisons of field conditions.

Figure 11 for sorghum (see below) and Figure 26 for cotton, Figure 28 for millet and Figure 30 for maize (in Annex) show the selected parcels for each crop type, with the corresponding satellite imagery from Sentinel-2, PlanetScope and Pléiades. The parcels are grouped by their assessed level of heterogeneity, ranging from very low to very high. For each level, if available, four parcels are presented. Random selections were constrained to be close to the median parcel size for each crop type and favour those with available photographic imagery for comparison.

Figure 27 (for cotton), Figure 29 (for millet), Figure 25 (for sorghum), Figure 31 (for maize) in the Annex show some of the corresponding ground photographs which can be referenced to the parcels visible in the satellite imagery by the parcel ID.

Spatial resolution modulates how intra-parcel variation is perceived. Pléiades imagery (0.5m) reveals detailed patterns of vegetation density and soil exposure, enabling clear delineation of row structures and patchy vegetation. However, a dense patch of vegetation appears mostly as a consistent, structurally uniform area that does not allow for further distinction of the vegetation cover. PlanetScope (3m) preserves broader tonal contrasts but smoothes finer patch boundaries, while Sentinel-2 (10m) further generalizes field characteristics, rendering most intra-parcel variability indistinguishable except for the largest bare soil areas. This resolution-dependent gradient demonstrates the substantial information loss that occurs with decreasing resolution, affecting how intra-parcel heterogeneity is perceived through remote sensing imagery.

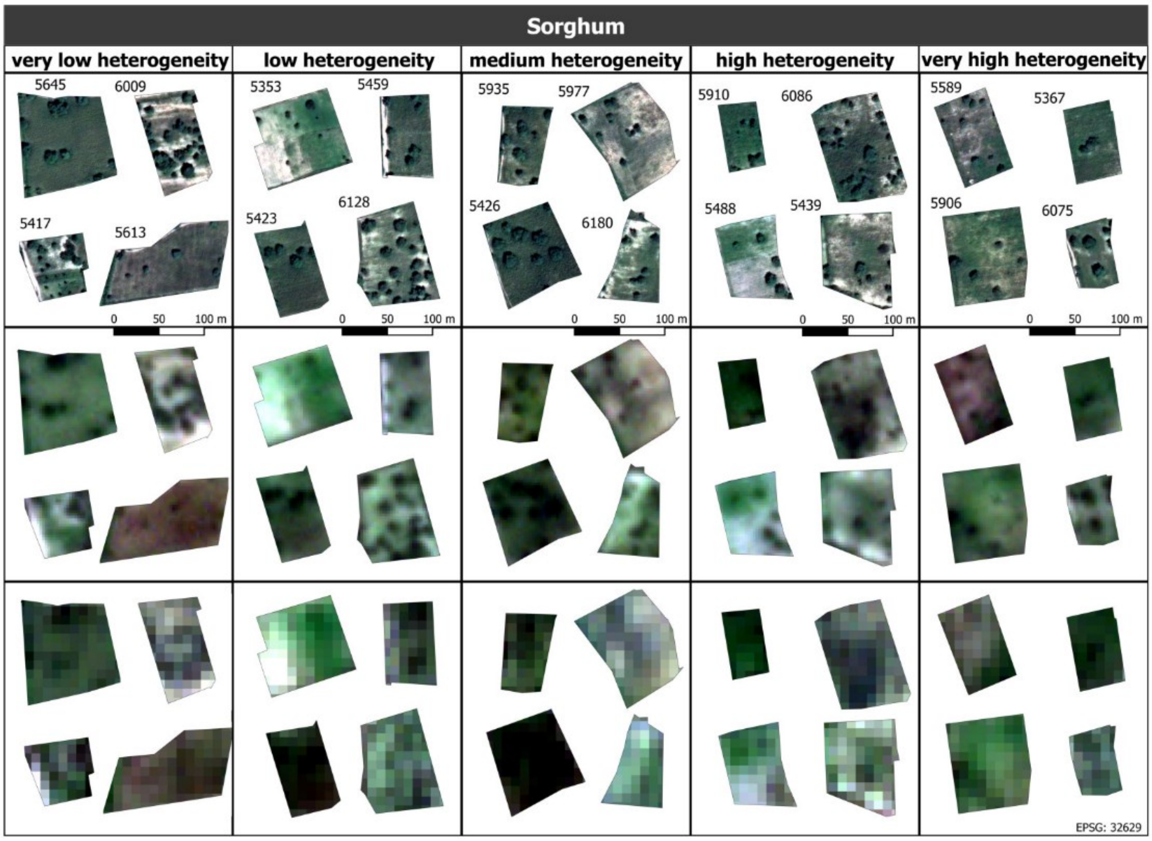

Figure 11: Sorghum parcels of different assessed heterogeneity levels on Pléiades, PlanetScope and Sentinel-2 imagery from late October

Figure 11 shows that Sorghum parcels are interspersed with trees across all heterogeneity levels. The visualized parcels display marked structural variability. Some very low heterogeneity fields (e.g., 5645, 5613) appear structurally homogeneous, whereas others (e.g., 6009, 5417) are seen as patchy in the Pléiades imagery. Likewise, some low heterogeneity parcels (e.g., 5353) are partially harvested which leads to a sharp contrast on the field, while some very high heterogeneity fields (e.g., 5367) appear structurally uniform. Parcels such as 5977 in the medium class exhibit pronounced patchiness and variations in sowing direction. Parcel photographs (see Figure 25 in Annex) in some cases contradict the satellite observations: for example, parcel 5417 looks uniform in the field photographs yet reveals diverse greenness and bare-soil patches in the Pléiades data. Overall, no consistent pattern emerges across heterogeneity levels on the satellite imagery. Photographs suggest that weed presence may increase with heterogeneity, but this remains tentative.

Figure 26 (see Annex) shows that across all heterogeneity classes, cotton parcels are uniformly interspersed with tree cover and exhibit relatively defined sowing rows. Pléiades imagery shows that even parcels rated with a very low heterogeneity nonetheless display pronounced intra-parcel variability in green biomass density, with some areas showing a dense foliage and others markedly sparse. In the high and very high classes, certain parcels (e.g., 5403) exhibit more uniformly dense biomass, while others (e.g., 5326 and 5368) reveal extensive bare-soil patches. The variations do not corroborate the in-situ ratings: dense foliage can occur in both "low" and "high" classes, and bare patches are found in parcels across classes. This indicates that observed foliage density and the extent of the presence of bare soil patches alone is not a reliable proxy for the level of heterogeneity. Parcel photographs (see Figure 27) fail to resolve this ambiguity, although a tentative observation is that weed incidence appears slightly elevated in the imagery of the higher heterogeneity class parcels.

Millet parcels (see Figure 28) are also interspersed with scattered trees. The parcels exhibit row-sowing patterns across all heterogeneity classes. However, some millet fields (e.g. parcel 5717 in the "low" class) appear already partially harvested. Individual parcels within the same heterogeneity class show significant patchiness, while some exhibit dense green foliage. Intra-parcel contrasts show no consistent trend from very low to very high heterogeneity. Ground photographs (see Figure 29) offer little further insights, also variations in weed presence cannot be distinguished. The millet observations closely mirror those for cotton, where neither tree cover nor row-sowing patterns align systematically with in-situ heterogeneity ratings, and bare-soil patches occur across all classes.

The visual assessment of maize parcels (see Figure 30) reveals some interesting findings. Tree presence is variable between individual parcels but occurs generally across all heterogeneity classes. Notably, some parcels in the low and medium heterogeneity classes (e.g., parcel 5926) show evidence of piled up crop stalks in the Pléiades imagery, a feature not observed in the high and very high classes. The presence of exposed soil varies significantly between the parcels within the same heterogeneity class, with some showing high bare soil presence, while others displaying dense green foliage. Ground photographs (see Figure 31) provide additional context. For low and medium heterogeneity parcels, photographs show structured row sowing patterns. On the photographs for these classes from around the beginning of November, the crops appear already are harvested. In contrast, photographs of high and very high heterogeneity parcels tend to depict less structural uniformity, higher patchiness and greater weed presence, with these parcels appearing as unharvested. The observations suggest a pattern of increasing weed presence and decreasing structural uniformity with increasing heterogeneity, however these characteristics are not consistently derivable from the satellite imagery alone, even at sub-metre resolution.

The visual exploration of parcels across different crops and heterogeneity classes reveals that satellite imagery at the given spatial resolutions does not enable the consistent capture of all visual indicators of intra-parcel heterogeneity, such as varying crop structure, foliage density, soil exposure, and signs of poor maintenance, that would allow for a clear differentiation between the different levels of heterogeneity. The findings show that field conditions vary significantly between parcels within the same heterogeneity class, to the extent that across class patterns are very hard to unveil. While high-resolution Pléiades imagery highlights patchiness within parcels, these contrasts showed no clear trend across heterogeneity levels, and distinguishing significant differences in crop uniformity or the presence of weeds remained infeasible even at this sub-metre resolution. Parcels with low or no patchiness largely appeared structurally uniform, but further differentiation on what constitutes the vegetation cover within the parcels were not possible. So, while parcel photographs suggest that weed presence plays an important role in how the heterogeneity classes were attributed, this could not reliably be observed or confirmed with satellite data. Across the crops, structural variability appeared unpredictable on satellite imagery and failed to align systematically with field-based heterogeneity ratings.

6.3. Synthesis of Parcel-Level Findings on Heterogeneity

This chapter examined the cropland parcels in the in-situ data and the relationship between field-assed parcel-level heterogeneity and satellite-derived spectral metrics across different spatial resolutions. Parcel-level NDVI and intra-parcel NDVI variability were related to the heterogeneity classes, as well as the structural appearance of these parcels on satellite imagery was examined.

Investigation of the in-situ dataset revealed that while most surveyed parcels are monocropped, a substantial subset exhibits crop diversity at the parcel level. Cotton fields stand out as larger than other crop types, reflecting their status as the region's primary cash crop, however, are still smallholder in scale. Analysis uncovered a positive relationship between parcel size and heterogeneity ratings, suggesting that larger fields tend to display greater heterogeneity. Overall, the majority of fields were classified as having a low to medium heterogeneity, which appears to be the norm for monocropped parcels across most crop types.

A multi-sensor comparison of time series revealed that the six main crops under study exhibited a coherent seasonal NDVI trajectory: values rise from mid-June, peak in early-to-mid October, and subsequently decline. Although the overall trajectory is shared, distinct crop-specific patterns were observed; some crops exhibit higher peak NDVI values than others, and separability appears possible at certain points in time. However, wide variability in NDVI was detected among individual parcels of the same crop type.

To explore this variability, parcels of the main six crop types were analysed by their assigned heterogeneity level, with mean NDVI values compared across heterogeneity classes. Two theoretical patterns had been posited to be plausible. First, higher heterogeneity might correspond to higher NDVI values (e.g., because weed presence might increase total biomass at particular stages). Conversely, higher heterogeneity might correspond to lower NDVI values if bare soil or sparse vegetation dominates the parcel. Either dynamic was expected to not be mutually exclusive as well as temporally non-stationary, i.e. one or the other could be more pronounced at specific points in the season.

Across the growing season, the relationship between in-situ heterogeneity and parcel-level NDVI proved to be complex, crop-specific, and temporally variable. For example, sesame and millet both displayed stratification by heterogeneity class between June and October; yet higher heterogeneity coincided with higher NDVI for sesame, whereas the opposite held for millet. Other crops showed no consistent stratification over time. Looking specifically at the second half of October, sesame and maize displayed the most consistent trends across all three sensors, with mean NDVI steadily increasing from very-low to very-high heterogeneity classes, suggesting that more heterogeneous parcels carried greater green biomass at this time. For the remaining crops, class-wise means were very close together, or unstratified, suggesting that the heterogeneity classes can not explain NDVI variability of parcels at that time.

The analysis followed up with an analysis of the intra-parcel variability, using the coefficient of variation of NDVI as a measure of internal spectral variability. The hypothesis that was postulated here was that higher levels of in-situ assessed heterogeneity would correspond to higher intra-parcel spectral variability, as more heterogeneous parcels would be expected to exhibit greater variability in vegetation density and soil exposure. However, the results showed that this relationship is not clear-cut and varies significantly between crops and sensors. Higher in-situ assessed heterogeneity can not be generalised to translate to greater intra-parcel spectral variability at the resolution of the sensors employed in this study.

Visual inspection confirmed that the structural composition of parcels varied widely. Pléiades imagery, and to a considerably lesser extent, PlanetScope and Sentinel-2, revealed parcels with strongly differing field conditions. Variations were observed in structural makeup, vegetation density and soil exposure; however, these variations did not consistently align with the in-situ heterogeneity classes. Field photographs suggested a possible association between weed presence and assigned heterogeneity yet weed presence could not be reliably inferred from satellite data alone, even at sub-metre resolution.

The analysis was limited by an imbalance in feature counts across crop types and, more acutely, across heterogeneity classes. Only sorghum possesses sufficient parcels in every heterogeneity class to permit robust conclusions. For sorghum, however, variation in parcel-level NDVI and intra-parcel spectral variability appear largely unexplained by the heterogeneity attributes. While certain crops (e.g., sesame, maize, millet) display NDVI stratification at specific times, the direction of the relationship is inconsistent between crops and variable for crop types at different times. Given these contradictions, the sparse sample sizes for most crops, and the ambiguous correspondence between visually perceived parcel characteristics and heterogeneity classes, the in-situ heterogeneity does not emerge as a reliable explanatory variable for parcel-level spectral variability or NDVI dynamics.

6.4. From Parcel-Level Findings to Classification Framework

A critical question addressed by this chapter was whether parcel-level heterogeneity can be a meaningful input attribute for crop type classification and if so, then how this could be operationalised.

The data shows that significant variability exists in the parcels' spectral characteristics, structural makeup, and vegetation density. It can be clearly observed that parcels of the same crop type differ substantially from one another. The analysis indicates, however, that field-assessed heterogeneity cannot be considered a dependable explanatory variable for understanding the variations in spectral characteristics or structural composition of parcels as observed in satellite imagery. The theoretical and practical constraints that may explain these patterns, as well as the limitations of the analysis are discussed in Chapter 8.

Heterogeneity is not dismissed here as a potentially useful input for crop-type classification; however, in its current operationalisation it has proved a poor predictor of parcel-level spectral variation. Consequently, classification will trial two alternative approaches: (i) simplifying the in-situ heterogeneity label to a binary distinction of "higher heterogeneity" versus "lower heterogeneity" and (ii) partitioning the crop parcels into two groups based on the coefficient-of-variation of NDVI distribution derived from Pléiades imagery.

The binary recategorization of field-assessed heterogeneity is predicated on the assumption that while fine gradations of heterogeneity may not translate to distinguishable spectral patterns, a more fundamental division between substantially heterogeneous and relatively homogeneous parcels might still improve crop type classification. The coefficient-of-variation approach aims to integrate an alternative objectively measured spectral metric that can be derived from satellite imagery itself.

Theoretically, the transition from parcel-level separability of crops to pixel-level classification presents substantial methodological challenges. While parcel-level analysis operates on aggregated statistics that smooth out internal variability, pixel-level classification must contend with each individual pixel's spectral signature. This fundamental difference can potentially lead to inconsistent classification outcomes within a single agricultural field - some pixels may be classified as a specific crop type while adjacent pixels may be classified as bare soil, despite collectively constituting a single crop parcel. Classification outcomes are expected to vary significantly between sensors due to the crucial role spatial resolution plays in capturing heterogeneous landscape elements. At Sentinel-2's resolution (10m), pixels in patchy fields represent an amalgamation of spectral reflectance from different field components, including vegetation, trees and bare soil. These mixed pixels would substantially influence or even define the spectral signatures associated with different crop types. Conversely, in very high-resolution imagery, individual pixels theoretically possess more specific spectral signatures as they are less likely to contain mixed signals from different parcel components.

This resolution-dependent dynamic interacts critically with training data selection. If a classifier is trained on data where parcels contain significant bare soil exposure, it may learn to associate the spectral signature of bare soil with a particular crop type, potentially leading to misclassification of pixels in fields with different soil exposure conditions. Furthermore, boundary pixels between parcels present significant classification challenges, as do trees and non-agricultural features such as buildings or paths that introduce additional spectral complexity.

Ultimately, only classification performance will reveal whether the parcel-level separability of crops can be successfully transferred to pixel-level classification of crop types and how this interacts with different spatial resolutions.

7. Crop Type Classification

Building upon the parcel-level analysis presented in Chapter 6, this chapter advances to pixel-based supervised crop type classification. While Chapter 6 questioned the utility of field-assessed heterogeneity as a direct predictor of spectral variation at the parcel level, it revealed distinct phenological patterns and crop-specific spectral signatures throughout the growing season. These spectral-temporal characteristics provide a foundation for distinguishing between crop types. This chapter develops and evaluates classification frameworks employing the imagery from Sentinel-2, PlanetScope, and Pléiades. Through systematic comparison of classification scenarios, it is assessed how effectively these sensors can discriminate between crop types in the study area. It further will be explored whether incorporating field-assessed heterogeneity and spectral variability of parcels operationalized as CV of NDVI can enhance classification performance.

The influence of tree cover on classification performance is assessed by applying the generated tree mask on the training data. Lastly, a more comprehensive approach is implemented to address the full spectrum of non-crop elements within cropland. This is achieved by generating a VHR-derived cropland mask to systematically refine the Sentinel-2 and PlanetScope validation datasets, allowing for an evaluation of how accuracy metrics are affected by the exclusion of these confounding features.

7.1. Classification Framework

A supervised classification approach is employed to distinguish between different crop types. The classification framework is built around the Random Forest algorithm, a robust ensemble learning method that combines multiple decision trees to improve prediction accuracy and reduce overfitting (Belgiu & Drăguţ, 2016; Breiman, 2001). Random Forest was selected for its ability to handle high-dimensional data with multiple spectral bands and vegetation indices across temporal sequences, and its demonstrated effectiveness in LULC and agricultural remote sensing applications (Belgiu & Drăguţ, 2016).

The classifier leverages the in-situ data on crop types, samples of non-cropland features and the spectral bands and indices for the multi-temporal imagery of Sentinel-2, PlanetScope, and the single date imagery of Pléiades. The Random Forest model is trained on a set of labelled samples, where each sample consists of spectral features extracted from the imagery at specific time points during the growing season. The model is expected to learn to associate these spectral features and their temporal dynamics with the corresponding crop types and non-cropland features based on the training data. The Random Forest classifier is implemented using the scikit-learn library in Python (Pedregosa et al., 2011).

7.2. Data Preparation

The data from the in-situ data collection solely comprises data on cropland. For the classification of the different elements that comprise the landscape, it is however also required to have additional samples on non-cropland features. This additional data is essential to provide a comprehensive classification framework that can distinguish between agricultural and non-agricultural features in the landscape. The data was collected through visual interpretation of the very high-resolution Pléiades image, through comparison of spectral signatures of Sentinel-2 and PlanetScope over time and with ancillary imagery from Google Earth. It is important to note here that the landscape elements in the study area are themselves highly heterogeneous. We mainly have a landscape matrix of bare-soil, shrubs, trees, grassland and cropland with patches of small settlements, small waterbodies and roads with different surface types. The classification framework therefore needs to account for this heterogeneity and the fact that the different land cover classes can be found in close proximity to each other. This proved a challenge in defining the classes to sample. After some classification trials and deliberations, it was decided to sample following non-cropland features:

Table 7: Non-cropland landcover samples

Class	Sample Polygons	Description
Built-Up Area	44	Residential compounds, buildings, and infrastructure with concrete or metal roofing materials
Asphalt Road	45	Roads paved with asphalt
Lateritic Rock Road	45	Unpaved roads constructed from compacted lateritic gravel and rock fragments
Clay Bare Soil	49	Exposed clay-rich soils, appearing predominantly white or light-coloured
Lateritic Bare Soil	45	Exposed lateritic soils, reddish-brown in appearance
Woody Vegetation	46	Dense tree and shrub cover including gallery forests and woodland patches
Forest Edge	47	Transitional zones between dense woody vegetation and open areas of bare soil
Grassland	42	Natural and semi-natural grass-dominated areas used for grazing or left fallow
Water	16	Permanent and seasonal water bodies including rivers, ponds, and reservoirs

It is important to note here that these classes were selected both based on their prevalence in the landscape and to reduce spectral confusion and misclassification with the crops. It is not a comprehensive land cover classification scheme but rather a selection of classes that are relevant for this application. The non-cropland samples were collected through a single digitization process, delineating feature boundaries using the Pléiades image as reference, while ensuring these samples maintained representative spectral characteristics across all three sensor resolutions.

The data on the crop parcels were filtered to remove parcels with more than one crop type, to try to ensure that the classifier is not trained on pixels of mixed cropland. Only the crops sorghum, maize, millet, cotton, sesame and peanut as the main six-crops as well as the low feature count crops bean, okra, rice, cowpea and tomato are selected for the classification. The "Other crops" class was deliberately excluded due to the ambiguity of the class. Further, inward buffers were applied to reduce the influence of edge effects, through mixels, which can introduce noise in the classification process. The parcels are buffered inward by 10 metres for Sentinel-2, and 3 metres for Pléiades and PlanetScope to remove pixels at the edges and to account for inaccuracies in parcel delineation.

7.3. Classification Scenarios and Methodological approach

The aim is to assess the classification performance of different crop types using the available sample data and utilizing spectral data from sensors with different spatial resolutions and temporal data depth. First, a baseline classification using solely crop type as the predictor is generated. Second and third, it is explored if field-assessed heterogeneity and parcel level spectral variability (CV of NDVI) can be a meaningful input variable to improve classification performance. While Chapter 6 demonstrated that field-assessed heterogeneity is not a strong predictor of spectral variation at the parcel level, it remains hypothesized that a simplified binary categorization might still impact classification performance. Two approaches are trialled here.

- For one, the in-situ assessed ordinal scaled data (from 1-very low to 5-very high) is recoded into two classes, based on the crop-specific median heterogeneity values. This means that parcels with a heterogeneity value higher than the median value for the respective crop type are classified as "high heterogeneity" and those below the median as "low heterogeneity". This creates paired classes such as "Sorghum - high heterogeneity" and "Sorghum - low heterogeneity" for each crop type, effectively doubling the number of main crop classes in the classification framework.

- A crop-specific classification based on NDVI variability derived from the Pléiades image. This approach uses the CV of NDVI calculated for each parcel. Parcels are categorized as "high CV" or "low CV" based on whether their value exceeds the crop-specific median CV.

This leads us to three classification scenarios that will be performed for each sensor:

Scenario 1. Baseline classification: A Random Forest classification using the crop type as the sole predictor class, without considering heterogeneity.

Scenario 2. **Classification incorporating field-assessed heterogeneity**: A Random Forest classification that incorporates a binary heterogeneity categorization as an additional input alongside crop type.

Scenario 3. Classification incorporating parcel-level CV of NDVI: A Random Forest classification that uses the NDVI-based coefficient-of-variation of parcels as an additional input alongside crop type

For Sentinel-2, three factors known to affect classification performance are systematically evaluated during the baseline classification (i.e. Scenario 1). First, class imbalance is addressed using Synthetic Minority Over-sampling Technique (SMOTE) at varying sampling ratios to improve minority crop as well as landcover representation. Second, different tree masking thresholds are tested to minimize spectral interference from tree cover within agricultural parcels. Third, the Random Forest algorithm is optimized by fine-tuning key hyperparameters including the number of trees and minimum samples required for node splitting. Scenarios 2 and 3 are then implemented using the optimized classification workflow identified in the baseline scenario.

The optimized classification workflow obtained with the Scenario 1 approach for Sentinel-2 data is subsequently applied to the PlanetScope and Pléiades imagery. Only the impact of tree masking is re-evaluated for each sensor since the different spatial resolution and characteristics of the sensors may yield different results. Scenarios 2 and 3 are then performed for each of the sensors.

7.4. Evaluation of Classification Performance

To evaluate the classification performance comprehensively, a range of metrics that provide insights into the model's performance will be reported. The primary assessment will focus on class-specific metrics for sorghum, maize, millet, cotton, sesame, and peanut, particularly by examining F1-scores and precision and recall to understand both commission and omission errors. The overall accuracy provides a useful summary statistic but can be misleading in the presence of class imbalance, which is significant in this dataset. Therefore, also the macro average F1-score will be reported to assess the model's performance, which averages all the per-class F1-scores, treating all classes equally regardless of their support size. The weighted average F1-score is further calculated, which takes the weighted mean of all per-class F1-scores, with the weights determined by the support size of the classes.

While the six major crop types remain the focus, performance on minor crops (combined as "Other crops" in confusion matrices) and non-cropland classes are monitored to ensure the classifier maintains reasonable discrimination across all the elements of the landscape. Confusion matrices are analysed to identify specific misclassification patterns between crops and between crop and non-crop classes.

7.5. Classification of Sentinel-2 time series

7.5.1. Methodology and Different Baseline Configurations

The Sentinel-2 classification begins with establishing a systematic baseline approach against which subsequent modifications can be evaluated. Initially, a reference Random Forest classification is implemented without sampling adjustments or tree masking, targeting the six primary crop types, five minor crops and nine non-cropland classes. Data preparation involves a 70/30 stratified split to maintain class distribution proportions. Table 31 shows the distribution of the sample data across the different classes and the pixel count of the training set after rasterization. The training data is then prepared by extracting the features for each sample raster cell, which includes the spectral bands B2, B3, B4, B5, B6, B7, B8, B8a, B11, B12 and the calculated indices (see Table 32 in Annex) for each of the 9 timesteps.

To address the pronounced class imbalance evident in the training data, SMOTE is implemented using the imblearn library in Python (Lemaitre et al., 2016). SMOTE is applied using varying sampling ratios (0.25, 0.5, 0.75, and 1.0) relative to the majority class (i.e. sorghum). This approach generates synthetic samples for minority classes with k=5 nearest neighbours, a value selected to accommodate the low sample count of the smallest class (i.e. tomato). The sampling strategy systematically evaluates how different balancing intensities affect classification performance, particularly for underrepresented crop types.

The influence of tree cover within cropland on classification performance is assessed through a threshold-based masking approach. Tree polygons are rasterized at 10-times the Sentinel-2 resolution, allowing calculation of percentage tree cover per pixel. Three threshold levels (20%, 50%, and 80%) are evaluated to determine optimal tree filtering intensity for both training and validation data. For each threshold, both training and validation datasets are filtered to exclude pixels exceeding the specified tree coverage percentage. Classification results are evaluated using both tree-filtered validation (representing idealized performance without tree interference) and non-tree-filtered validation (representing performance in the agricultural landscape where trees are common).

To optimize the Random Forest classifier, key hyperparameters are systematically tuned. Starting from a baseline configuration of 100 trees and a minimum of 2 samples per split, the number of trees is varied (100, 200, 500) along with the minimum samples required to split a node (2, 5, 10, 20). This process aims to identify the optimal configuration that balances classification accuracy with computational efficiency.

7.5.2. Scenario 1: Baseline Classification

7.5.2.1. Results of Baseline Classification

The first baseline classification achieved an overall accuracy of 71.3 % on the validation dataset, with a macro average F1-score of 0.54 and a weighted average F1-score of 0.69. The detailed classification report is shown in Table 8.

Table 8: Classification Report for Sentinel-2 Scenario 1 Baseline Classification

Class	Precision	Recall	F1-score	Support
Sorghum	0.77	0.81	**0.79**	3450
Maize	0.59	0.71	**0.64**	905
Millet	0.65	0.67	**0.66**	572
Cotton	0.67	0.74	**0.71**	1119
Sesame	0.38	0.16	**0.23**	411

Peanut	0.50	0.02	**0.05**	122
Bean	0.00	0.00	0.00	120
Okra	0.00	0.00	0.00	2
Rice	0.00	0.00	0.00	7
Cowpea	0.00	0.00	0.00	25
Tomato	0.00	0.00	0.00	8
asphalt road	0.97	1.00	0.98	58
built-up	0.79	0.70	0.74	37
clay bare soil	0.80	0.67	0.73	55
forest-edge	0.73	0.91	0.81	47
grassland	1.00	0.73	0.84	62
lateritic bare soil	0.87	0.95	0.91	80
lateritic rock road	1.00	0.79	0.88	48
water	1.00	0.88	0.94	17
woody vegetation	0.96	0.98	0.97	90
accuracy			**0.71**	**7235**
macro avg	0.58	0.54	**0.54**	7235
weighted avg	0.69	0.71	**0.69**	7235

Crop classification performance varies. Sorghum demonstrated the strongest performance with an F1-score of 0.79 followed by cotton (0.71), millet (0.66) and maize (0.64), while sesame and peanut showed low performance with F1-scores of 0.23 and 0.05, respectively. Overall, the recall is higher than the precision for the main crop types, indicating that the model is better at correctly identifying the crops than at avoiding false positives. The low feature count crops (bean, okra, rice, cowpea, tomato) exhibited complete classification failure with F1-scores of zero.

Non-cropland classification demonstrated superior performance compared to the crop type classes, with several classes achieving good results. Water bodies (F1: 0.94), woody vegetation (F1: 0.97), asphalt roads (F1: 0.98), and lateritic bare soil (F1: 0.91) showed strong classification accuracy. Good performance was also observed for grassland (F1: 0.84) and forest edge (F1: 0.81), while built-up areas (F1: 0.74) and clay bare soil (F1: 0.73) achieved moderate accuracy levels.

The macro average being lower than the overall accuracy indicates uneven performance across classes. The weighted average F1-score (0.69) being closer to overall accuracy confirms that performance is driven primarily by success on dominant classes, particularly Sorghum.

The confusion matrix (see Table 33 in Annex) reveals that there is substantial classification confusion among the primary crop types, with many maize, millet, cotton, sesame and peanut pixels being misclassified as sorghum. Also, sorghum is in many cases misclassified as one of the other main crop types except for peanut. The minority classes (bean, okra, rice, cowpea, tomato) were mostly misclassified as cotton. For the non-cropland classes, there is also some notable confusion of grassland being misclassified as either sorghum, maize or millet.

7.5.2.2. Impact of Implementing SMOTE

Table 34 in the Annex shows the results of the SMOTE application with a sampling ratio of 1. Here, the number of training pixels in each class is equal to the number of pixels in the sorghum class (i.e. 6467). The overall classification accuracy decreases slightly to 70.85%**,** with the macro average increasing to 0.58 and the weighted average remaining at 0.69. The F1-scores for the main crop types remain largely comparable to the baseline classification. Some low feature count crops improve. The most notable improvement is rice classification going from complete failure to a F1-score of 0.50. Also, some of the non-cropland classes show improved performance, with the F1-scores for built up areas (0.84), lateritic rock road (0.99), lateritic bare soil and water (0.97) increasing. However, most low feature count crops still exhibit complete classification failure and the classification performance for sesame and peanut remain low with F1-scores of 0.23 and 0.06, respectively. The confusion matrix (Table 35 in the Annex) reveals that the model still struggles with the low feature count crops, which are often misclassified as cotton or sorghum. The main crop types are still commonly misclassified as sorghum.

The results of the SMOTE application with a sampling ratio of 0.5 are shown in the Annex in Table 36. The overall classification accuracy lies at 71.17%, being slightly higher than the previous SMOTE scenario but slightly lower than the baseline scenario without SMOTE. The macro average F1-score increases slightly to 0.59. The F1-scores for the main crop types and for the non-cropland classes are largely comparable. The low feature count crops still struggle, but rice showed an improvement from 0.50 to 0.55. Minority crops such as bean, okra, cowpea, and tomato still exhibit complete classification failure. Comparing the confusion matrix (see Table 37 in Annex) we can see similar patterns of misclassification as in the previous SMOTE scenario, albeit with variations in the extent of specific misclassifications.

It can be concluded that applying SMOTE helped improve the classification performance for some minority classes but does not overcome the inherent challenges of class imbalance and spectral confusion in the dataset. The impact on the better performing crops sorghum, maize, millet and cotton are limited, while sesame and peanut and most low feature count crops still struggle with classification. We can observe an increase in F1 scores for most non-cropland classes. Due to this limited improvement and the higher macro average F1-score, the SMOTE scenario with a sampling ratio of 0.5 is implemented as the baseline for further analysis.

7.5.2.3. Impact of Tree Masking

To assess the impact of tree cover, pixels exceeding 20%, 50%, and 80% tree cover thresholds are masked from both training and validation data. The accuracy metrics on both the tree-filtered and non-tree filtered validation data are compared.

For the non-tree filtered validation set, the overall accuracy is at 71.13 % with 80 percent tree masking, 70.88 % at 50 percent tree masking and 71.49% with a 20 percent tree masking threshold applied. The macro averages are 0.58 (80 % threshold), 0.58 (50 %), and 0.57 (20 %) respectively. The F1-scores for the main crop types sorghum, maize, millet and cotton are similar across the different tree masking thresholds, deviating only by an extent of one or two percentage points to the baseline without tree masking. Sesame and peanut show no improvement in classification as well as the low feature count crops bean, okra, rice, cowpea and tomato. The non-cropland classes show variability in classification performance across the different tree masking thresholds, with some classes improving and some classes decreasing in classification performance depending on the chosen threshold.

With the 20 % threshold (see Table 39 in Annex) the non-cropland classes forest-edge and grassland F1-scores improved slightly, however with other non-cropland classes slightly decreasing. Comparing the tree-filtered and non-tree-filtered validation (Table 38 and Table 39 in Annex) , we can see that the tree-filtered validation achieved a higher overall accuracy of 72.56% compared to 71.49%. The macro average F1-score for the tree-filtered validation is 0.58, while it is 0.57 for the non-tree-filtered validation. The confusion matrix for the non-tree-filtered validation dataset (Table 40 in Annex) shows that the overall dynamics of misclassification did not change substantially. The proportion of misclassification of sorghum for example changed slightly, with misclassification as sesame and millet being more prevalent. Overall, no significant changes in the dynamics of misclassification can be observed.

The tree masking threshold of 20% demonstrates only a very small improvement in classification performance for sorghum, maize, and millet compared to the non-tree filtered scenario. Additionally, the macro average F1-score decreases with tree masking applied, indicating that tree masking negatively impacts the balanced performance across all classes.

Considering this, it is decided to exclude tree masking from further analysis for several reasons which will be expanded on in Chapter 8. From an operational perspective, implementing tree masking adds substantial processing complexity with minimal performance gain.

7.5.2.4. Impact of RF Model Calibration on Classification Performance

Different values are trialled of number of trees (100, 200, 500) and minimum samples split (2, 5, 10, 20). The results generally show minor variations in classification performance across different parameter settings. A slight decrease in accuracy is observed with higher values of minimum samples split, while increasing the number of trees improves performance. The configuration with 500 trees and a minimum sample split of 2 is reported on here.

The results (Table 41 in Annex) show that compared to the previous baseline model (with 0.5 SMOTE), the F1 scores of maize (0.67) and millet (0.68) as well as sesame (0.28) increase by 1 to 3 percentage points, while the F1-score for sorghum, cotton and peanut remain unchanged. The low feature count crops still struggle with classification, with F1-scores of zero for bean, okra, cowpea and tomato. Rice shows a slight improvement of F1-score to 0.61. The non-cropland classes show similar performance to the 0.5 SMOTE baseline scenario, with some classes improving slightly, with others unchanged or showing a slight decrease. The patterns of misclassification (see Table 42 in Annex) have not changed substantially.

The overall increase in classification performance (0.72), the stable macro average F1 (0.59), and the improved F1-scores for maize, millet, and sesame show that the model calibration to a higher number of trees has a positive impact on the classification performance. Therefore, this model is used as the baseline for further analysis.

7.5.3. Scenario 2: Incorporating Field-Assessed Heterogeneity

Building on the baseline classification approach, Scenario 2 explores whether incorporating field-observed heterogeneity through a binary categorization into “high heterogeneity” parcels and “low heterogeneity” parcels can improve crop type classification performance. The reporting highlights both the classification performance of the crop classes split by heterogeneity as well as when merging the subdivided classes back to the original crop classes for validation assessment.

Results show that incorporating field-observed heterogeneity improves overall accuracy to 76.27% with a macro average F1-score of 0.62. (see Table 10) Lower heterogeneity class subsets consistently exhibit higher F1-scores than their higher heterogeneity counterparts (see Table 9). The most substantial improvements over baseline classification occurred for cotton (F1: 0.80), sesame (F1: 0.59) and peanut (F1: 0.27), with also sorghum (F1: 0.81) and maize (F1: 0.68) showing improved F1-scored. The F1-score for millet decreases (F1: 0.64) while the low feature count crops still struggle with classification, and rice showing a slight performance decrease. The non-cropland classed show variability in classification performance, with some classes improving and some decreasing compared to the previous baseline scenario. The confusion matrix (see Table 43 in Annex) reveals reduced misclassification of sorghum as one of the other major crops, particularly maize. We can also see that the merged "Other crops" are here more commonly misclassified as sorghum instead of cotton, compared to the baseline scenario.

It can be concluded that the incorporation of in-situ heterogeneity information into the classification framework leads to an improvement in classification performance, reflected in higher overall accuracy and improved F1-scores of major crops.

Table 9: Validation metrics for heterogeneity subsets for Sentinel-2 Scenario 2 (Incorporating Field-assessed Heterogeneity)

Class	Subset	Precision	Recall	F1-score	Support
Sorghum	**higher** heterogeneity	0.47	0.35	0.40	2046
	lower heterogeneity	0.42	0.75	**0.54**	1281
Maize	**higher** heterogeneity	0.44	0.28	0.34	296
	lower heterogeneity	0.79	0.63	**0.70**	486
Millet	**higher** heterogeneity	0.75	0.33	0.45	243
	lower heterogeneity	0.64	0.59	**0.61**	361
Cotton	**higher** heterogeneity	0.54	0.68	0.61	586
	lower heterogeneity	0.69	0.58	**0.63**	762
Sesame	**higher** heterogeneity	0.34	0.41	0.37	217
	lower heterogeneity	0.62	0.45	**0.53**	216
Peanut	**higher** heterogeneity	0	0	0	20
	lower heterogeneity	0.77	0.50	**0.61**	20

Table 10: Classification report for Sentinel-2 Scenario 2 (Incorporating Field-assessed Heterogneity)

Class	Precision	Recall	F1-score	Support
Sorghum	0.76	0.87	**0.81**	3327
Maize	0.81	0.59	**0.68**	782
Millet	0.76	0.55	**0.64**	604
Cotton	0.79	0.81	**0.80**	1348
Sesame	0.60	0.58	**0.59**	433
Peanut	0.29	0.25	**0.27**	40
Bean	0.08	0.02	0.03	120
Okra	0.00	0.00	0.00	9
Rice	0.00	0.00	0.00	44
Cowpea	0.33	0.80	0.47	5
Tomato	0.00	0.00	0.00	8
asphalt road	1.00	1.00	1.00	58
built-up	0.80	0.89	0.85	37
clay bare soil	0.62	0.71	0.66	55
forest-edge	0.89	0.89	0.89	47
grassland	0.94	0.74	0.83	62
lateritic bare soil	0.88	0.95	0.92	80
lateritic rock road	1.00	0.98	0.99	48
water	1.00	0.94	0.97	17
woody vegetation	0.90	1.00	0.95	90
accuracy			**0.76**	**7214**
macro avg	0.62	0.63	**0.62**	7214
weighted avg	0.75	0.76	0.75	7214

7.5.4. Scenario 3: Incorporating Intra-Parcel Spectral Variability

The results of the Scenario 3 approach do not show a clear pattern of performance differences between the lower and higher CV subset classes (see Table 45 in Annex). Sorghum and maize show higher recall and F1-scores for higher CV subsets, while the lower CV subsets performed better for cotton and sesame. Millet showed equivalent F1-scores across both classes.

Overall accuracy decreases to 69.72% compared to the baseline, though with a slightly improved macro average F1-score of 0.60 (see Table 44). Performance varies significantly by crop type: millet (F1: 0.73), cotton (F1: 0.73), and sesame (F1: 0.46) improved, while sorghum (F1: 0.75) and maize (F1: 0.50) declined. Peanut and most low-feature crops show complete classification failure, with rice being a notable exception with a strong performance increase (F1: 0.93). The confusion matrix (see Table 46) reveals increased sorghum recall but also greater misclassification of other crops as sorghum, with maize particularly affected by this pattern.

Overall, it can be concluded that the incorporation of intra-parcel spectral variability of crop parcels does not lead to an improvement in classification performance for Sentinel-2.

7.5.5. Summary of Sentinel-2 Classification Results

The classification experiments yielded valuable insights into crop mapping capabilities with Sentinel-2. Neither SMOTE sampling, to address class imbalance, nor tree masking significantly improved overall performance. More substantial improvements emerged when incorporating field-observed heterogeneity in Scenario 2, which achieved the highest overall accuracy (76.27%) and macro-average F1-score (0.62). This approach revealed a consistent pattern: pixels from lower-heterogeneity parcels were more accurately classified than those from higher-heterogeneity parcels across most crop types. Conversely, Scenario 3's integration of Pléiades-derived spectral variability of parcels decreased overall accuracy to 69.72%, with no clear pattern across the CV subset classes.

Despite these improvements, classification performance remains suboptimal for operational crop mapping, with only sorghum and cotton achieving F1-scores above 0.80 and most minor crops failing completely. The persistent confusion between crop pairs indicates fundamental limitations in spectral-temporal separability using Sentinel-2 data in this complex agricultural landscape.

7.5.6. Applying the Scenario 2 Classifier on Study Area

Figure 12 shows the classification results of the Scenario 2 model applied to the entire study area. The classification shows a complex landscape with large extents of forest, bare soil and forest edges, interspersed with large patches of cropland. There appears to be a general dichotomy between cropland and non-cropland areas, with the non-cropland features appearing to be well captured by the classification and man-made features such as roads and built-up areas clearly identifiable. Sorghum (shown in red) clearly dominates the cropland matrix of the classified landscape, appearing in large continuous patches throughout. Other major crops appear in smaller clusters, with maize (yellow), millet (orange), and cotton (blue) forming distinct patterns resembling cropland parcels. Sesame (purple) and peanut (brown) are less prevalent.

The inset maps demonstrate the relationship between the classification results and the in-situ cropland parcels (outlined in black), labelled by their assessed crop type(s). The hashed parcels indicate those parcels on which the model was trained (after inward buffering), while the non-hashed parcels are either part of the validation set or constitute additional parcels with multiple recorded crops or where "Other Crops" were recorded as the primary crop type. Parcels labelled in-situ as sorghum largely align with the classified result, with most pixels within these parcels classified as sorghum. In some cases, individual pixels within these parcels are misclassified, and in some rare cases larger patches of pixels are classified as other crops. Cotton shows similar patterns but with partial misclassification primarily as sorghum. Maize and millet parcels exhibit more pronounced misclassification, predominantly as sorghum. The classification of peanut and "Other crops" performs poorly. For parcels with multiple recorded crops the classification shows limited success in reflecting this diversity, with some parcels showing a mix of crop types, although not always in correspondence with the in-situ data labels.

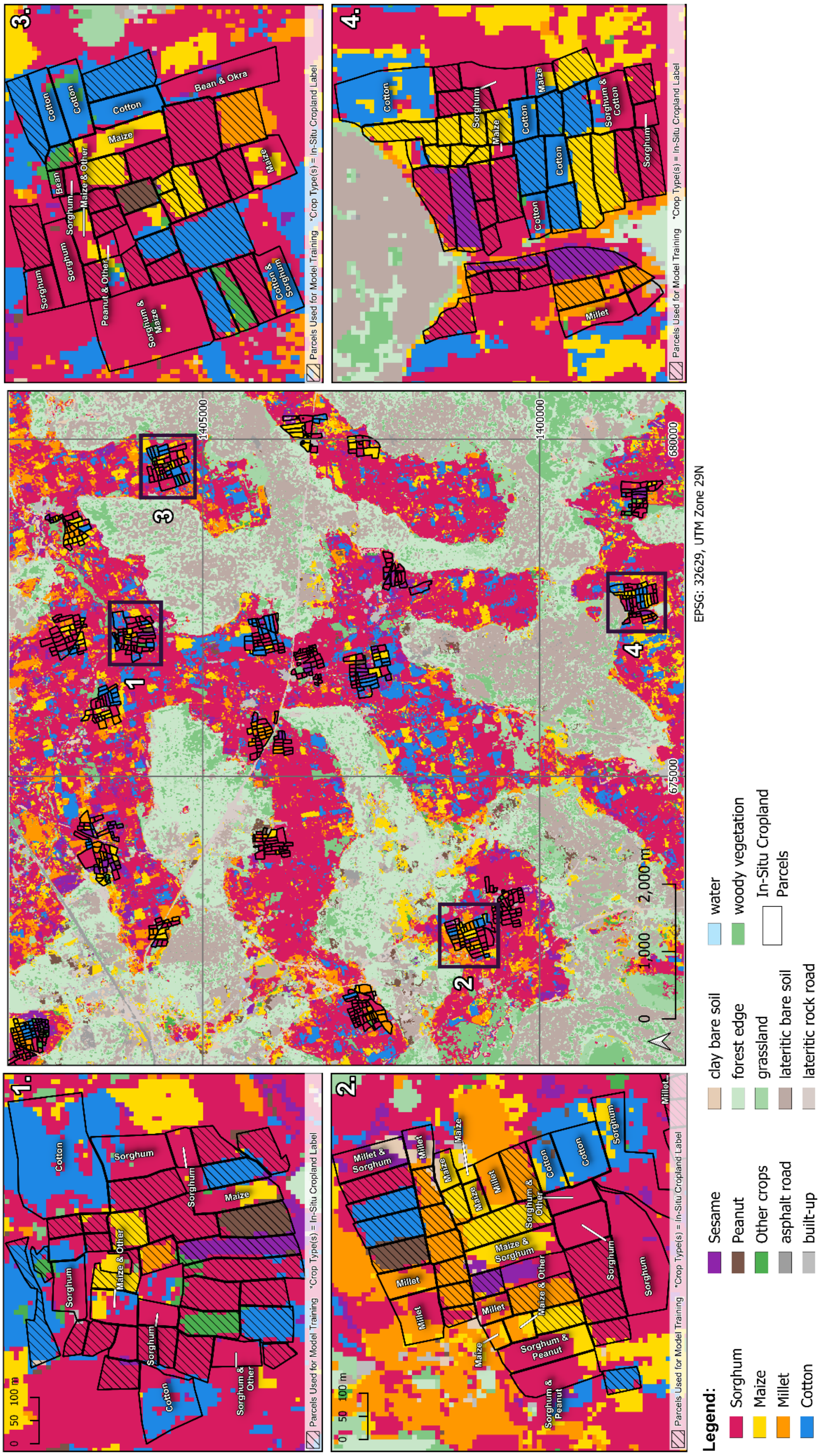

Figure 12: Classification of study area with Sentinel-2 Scenario 2 classifier

7.6. Classification of PlanetScope time series

7.6.1. Methodology and Configuration

The PlanetScope classification applies the optimized Sentinel-2 parameters (SMOTE ratio 0.5, Random Forest with 500 trees and a minimum samples split of 2) without repeating parameter tuning. The same processing methodology is followed, with adjustments for PlanetScope's higher spatial resolution, with inward buffers of 3 metres applied to remove edge pixels. The data undergoes the same 70/30 stratified training/validation split. Table 31 in the Annex shows the distribution of the sample data set across the different classes and the pixel count of the training data after rasterization.

The classification is run using the bands Blue, Green I, Green, Yellow, Red, Red Edge, and NIR and the calculated indices for PlanetScope (see Table 47 in Annex) for the 11 images in the time series. Initially, in the baseline scenario, the effect of tree masking with a 20% threshold is evaluated to assess whether excluding tree-covered pixels has an impact on classification accuracy.

7.6.2. Scenario 1: Baseline Classification

7.6.2.1. Results of Baseline Classification without Tree Masking

The PlanetScope baseline classification without tree masking achieves 70.18% accuracy with a macro average F1-score of 0.57 (see Table 48 in Annex). Performance forms a clear hierarchy: sorghum (F1: 0.79) performs the strongest, followed by cotton (0.72), millet (0.69), maize (0.63). Sesame (0.21) and peanut (0.11) perform poorly. Low-feature count crops show complete classification failure, while non-cropland classes performing generally well. The confusion matrix (see Table 49) reveals a systematic bias towards sorghum classification, which absorbs substantial portions of both minority crops and major crops including maize, millet, and cotton.

7.6.2.2. Results of Baseline Classification with Tree Masking

Tree masking at a 20% threshold shows minimal impact on PlanetScope classification performance (see Table 50 and Table 51 in Annex), with overall accuracy improving negligibly from 70.18% to 70.23% (70.73% on tree-filtered validation). The macro average F1-score remains unchanged at 0.57, with main crop types showing only minor fluctuations in performance metrics. Minority crops continue to exhibit complete classification failure, and fundamental misclassification patterns (see Table 52 in Annex) remain unchanged. These results show that integrating the tree mask provides no substantial benefit for crop type classification with PlanetScope.

7.6.3. Scenario 2: Incorporating Field-assessed Heterogeneity

Incorporating field-assessed heterogeneity into the PlanetScope classification reveals that lower heterogeneity subsets outperform higher heterogeneity subsets across most major crop types, except for sorghum (see Table 11). Overall classification accuracy increases to 74.35% with a macro average F1-score of 0.61 (see Table 12). Sorghum (F1: 0.81), cotton (0.79), millet (0.70), maize (0.65) show improved F1-scores compared to the baseline. Improvements are substantial for underperforming crops like sesame (0.50) and peanut (0.31). Rice improves from complete classification failure to an F1-score of 0.32, while other low-feature count crops continue to show classification failure. The confusion matrix (Table 53 in Annex) reveals reduced misclassification as sorghum across most main crop types, though minority crops and peanut and sesame remain strongly misclassified as sorghum.

Table 11: Validation metrics for heterogeneity subsets for PlanetScope Scenario 2 (Incorporating Field-assessed Heterogeneity)

Class	Subset	Precision	Recall	F1-score	Support
Sorghum	**higher** heterogeneity	0.48	0.62	0.54	28 025
	lower heterogeneity	0.51	0.54	**0.52**	25 008
Maize	**higher** heterogeneity	0.39	0.32	0.35	5391
	lower heterogeneity	0.65	0.58	**0.61**	8734
Millet	**higher** heterogeneity	0.68	0.35	0.46	4244
	lower heterogeneity	0.60	0.61	**0.61**	7277
Cotton	**higher** heterogeneity	0.64	0.57	0.60	9143
	lower heterogeneity	0.80	0.71	**0.75**	11 599
Sesame	**higher** heterogeneity	0.43	0.36	0.39	4771
	lower heterogeneity	0.47	0.34	**0.40**	2725
Peanut	**higher** heterogeneity	0.00	0.00	0.00	429
	lower heterogeneity	0.90	0.58	**0.71**	563

Table 12: Classification report for PlanetScope Scenario 2 (Incorporating Field-assessed Heterogeneity)

Class	Precision	Recall	F1-score	Support
Sorghum	0.75	0.89	**0.81**	53 033
Maize	0.71	0.61	**0.65**	14 125
Millet	0.77	0.63	**0.70**	11 521
Cotton	0.84	0.74	**0.79**	20 742
Sesame	0.56	0.44	**0.50**	7496
Peanut	0.29	0.33	**0.31**	992
Bean	0.01	0.00	0.00	1796
Okra	0.00	0.00	0.00	810
Rice	0.20	0.76	0.32	46
Cowpea	0.00	0.00	0.00	224
Tomato	0.00	0.00	0.00	246
asphalt road	0.99	1.00	0.99	215
built-up	0.98	0.78	0.87	152
clay bare soil	0.84	0.70	0.76	260
forest-edge	0.83	0.94	0.88	200
grassland	0.90	0.75	0.82	440
lateritic bare soil	0.86	0.99	0.92	426

lateritic rock road	1.00	0.98	0.99	209
water	0.97	0.93	0.95	97
woody vegetation	1.00	0.99	0.99	520
accuracy			**0.74**	113 550
macro avg	0.62	0.62	**0.61**	113 550
weighted avg	0.73	0.74	**0.73**	113 550

7.6.4. Scenario 3: Incorporating Intra-Parcel Spectral Variability

Scenario 3 shows no clear performance pattern between high and low CV subsets (see Table 54 in Annex). The overall accuracy reaches 71.12% with a macro average F1-score of 0.58 (see Table 55 in Annex). Performance varies by crop type: millet (F1-score: 0.72), cotton (0.79), and sesame (0.54) show improvement compared to the baseline, while sorghum (0.77) and maize (0.55) decline. Peanut and low-feature count crops show complete classification failure, being predominantly misclassified as Sorghum (see Table 56 in Annex).

The incorporation of spectral variability of parcels in Scenario 3 leads to a minor improvement in classification performance compared to the baseline scenario, mainly driven by improved performance of millet, cotton and sesame, while other main crops show a decrease in classification accuracy.

7.6.5. Summary of PlanetScope Classification Results

The baseline classification for PlanetScope achieved an overall accuracy of 70.18% and a macro average F1-score of 0.57, which is comparable to the Sentinel-2 baseline classification results. The main crops sorghum, cotton, and millet performed well, while minority crops such as sesame and peanut struggled significantly. Tree masking proved ineffective with negligible accuracy improvements (70.23%). The incorporation of in-situ heterogeneity assessments in Scenario 2 leads to improved classification performance, achieving an overall accuracy of 74.35%, a macro average F1-score of 0.61 and a substantial increase in F1-scores for crops such as sesame and peanut. Scenario 3's spectral variability approach produced more modest improvements (accuracy: 71.12%, macro F1: 0.58), with inconsistent effects across crop types.

7.6.6. Applying the Scenario 2 classifier on Study Area

Figure 13 shows the classification results of the Scenario 2 classifier applied to the study area. We again see a clear dichotomy between cropland and non-cropland, with cropland appearing in large patches in the landscape matrix. Sorghum clearly dominates the cropland matrix of the classified landscape, with other major crops such as maize, millet and cotton forming distinct patterns resembling cropland parcels. We also see that man-made features such as roads and built-up areas are clearly identifiable.

Examining validation parcels in the inset maps we can see that sorghum is well captured in the classification, with parcels being largely correctly classified with few misclassifications of individual pixels. Cotton shows less consistent classification, with some parcels being well captured and others showing a high number of misclassified pixels. Larger cotton fields (see inset map 1 and 4) show significant misclassification of pixels as sorghum or as maize. Maize and millet parcels exhibit mixed results, with many parcels correctly classified but others showing significant misclassifications with other main crop types. Sesame and peanut appear to be widely misclassified. In case of multiple crops being present in the same parcel, the classification only manages to capture this diversity in some cases (e.g. Sorghum & Cotton in inset map 4 and Sorghum & Maize in inset map 3) but fails to do so in other cases (e.g. Cotton & Sorghum in inset map 2).

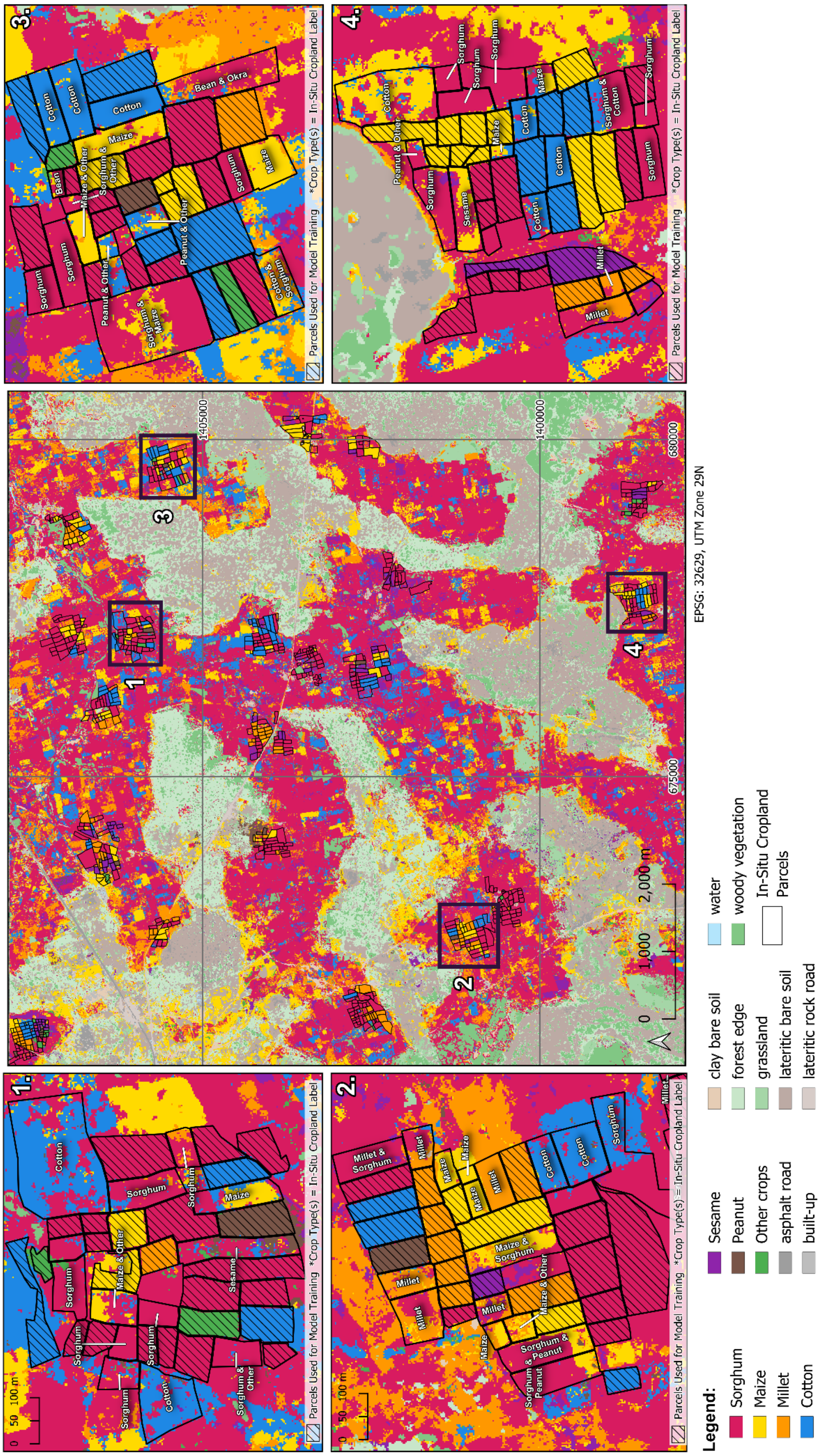

Figure 13: Classification of Study Area with PlanetScope Scenario 2 Classifier

7.7. Classification of single-date Pléiades Image

7.7.1 Methodology and Configuration

For the Pléiades classification, several methodological adaptations were necessary to accommodate its 0.5m resolution. A 3m inward buffer was applied to compensate for parcel delineation imprecisions. Data preparation followed the same stratified 70/30 training/validation split used with the other sensors. The pixel count of each class after rasterization can be seen in Table 31 in the Annex. To manage the computational cost of the significantly higher pixel count, a two-step sampling strategy was implemented: majority classes were capped at 500 000 samples, then minority classes were boosted to 50% of this threshold (250 000) using SMOTE oversampling. This balanced class representation while maintaining a large set of training samples for majority classes. The Random Forest classifier was configured with 100 trees to reduce computational demands.

The classifications are run using the bands Red, Green, Blue, NIR of Pléiades and the calculated indices (see Table 57 in Annex).

7.7.2. Scenario 1: Baseline classification

7.7.2.1. Results of Baseline Classification without Tree Masking

The baseline classification results of Pléiades exhibit very poor performance with only 25.07% overall accuracy and a macro average F1-score of 0.32 (see Table 58 in Annex). All crop types showed weak results, with even the best-performing crops achieving low F1-scores (sorghum: 0.33, cotton: 0.34, millet: 0.27, maize: 0.18). Minority crops show almost complete classification failure. Natural non-cropland classes (grassland, woody vegetation, forest edge) performed near zero, with the confusion matrix (see Table 59 in Annex) revealing systematic misclassification of these classes as cropland. Major crops show extensive confusion with each other.

7.7.2.2. Results of Baseline Classification with Tree Masking

The integration of tree masking using a 20% threshold shows minimal impact on classification performance, with overall accuracy showing negligible improvement from 25.07% to 25.15% accuracy on the non-tree filtered validation data (see Table 60 in Annex) and 25.63% on the tree-filtered validation data (see Table 61 in Annex). The macro average F1-score remains unchanged at 0.32 (on-tree filter validation data), indicating no improvement in balanced performance across all classes. The fundamental misclassification patterns remain largely unchanged (see Table 62 in Annex).

7.7.3. Scenario 2: Incorporating Field-assessed Heterogeneity

Incorporating recategorized field-assessed heterogeneity into the Pléiades classification framework yields modest improvements, with overall accuracy increasing to 26.75% and macro average F1-score rising to 0.34 (see Table 64 in Annex). Analysis of heterogeneity-specific subclasses (see Table 63 in Annex) reveals a consistent pattern: lower heterogeneity subsets generally achieve higher F1-scores than their higher heterogeneity counterparts across major crops. When merged to original crop classes, maize, cotton, and sesame show improved F1-scores compared to the baseline, while sorghum and millet performance slightly declines. Peanut and all low-feature count crops continue to exhibit near-complete classification failure. The confusion matrix (see Table 65 in Annex) reveals persistent widespread misclassification between major crop types, with no substantial changes in misclassification patterns compared to the baseline scenario.

7.7.4. Scenario 3: Incorporating Intra-Parcel Spectral Variability

Scenario 3 shows an improvement in classification performance with overall accuracy increasing to 27.94% and macro average F1-score reaching 0.34 (see Table 67 in Annex). Analysis of the CV subsets (see Table 66 in Annex) reveal inconsistent patterns, with sorghum, cotton, and sesame showing higher F1-scores in their lower CV subsets, while maize and millet do not. The main crop types sorghum, millet, cotton and sesame show improved F1-scores compared to the baseline scenario. Peanut and the low-feature count crops still show complete or close to complete classification failure in this scenario. Crops are extensively misclassified with each other (see Table 68 in Annex).

7.7.5. Summary of Pléiades Results

Pléiades yielded consistently poor classification results, with accuracy ranging between 25-28% across the scenarios. All crop types showed low F1-scores, with some major crops performing better than minor crops, of which many experienced almost complete classification failure. Natural non-cropland classes (grassland, forest edge, woodland) also exhibit very low F1-scores close to zero. Tree masking provided no meaningful improvement. Scenario 2 increased accuracy to 26.75% with improvements for maize, cotton and sesame, while Scenario 3 achieved slightly better results with 27.94% (macro average F1: 0.34).

7.7.6. Applying the Scenario 2 Classifier on Study Area

Figure 14 shows the classification results of the Pléiades Scenario 2 classifier applied to the study area. The map reveals a fragmented and patchy classification. Man-made features such as roads and built-up areas are identifiable. A dichotomy can still be observed to some extent between cropland and non-cropland, although cropland pixels, especially sorghum and cotton, appear widely interspersed in the landscape matrix.

The validation parcels show a strong variability in classified crop classes. All parcels lack a coherent classification, with classified pixels of different crop types scattered throughout. Cotton and sorghum parcels both show a big proportion of the parcel classified as the other class respectively. Also, woody vegetation appears commonly within the cropland parcels. No clear differences in classification results can be observed between those parcels with a single field-assessed crop type and those with multiple crop types.

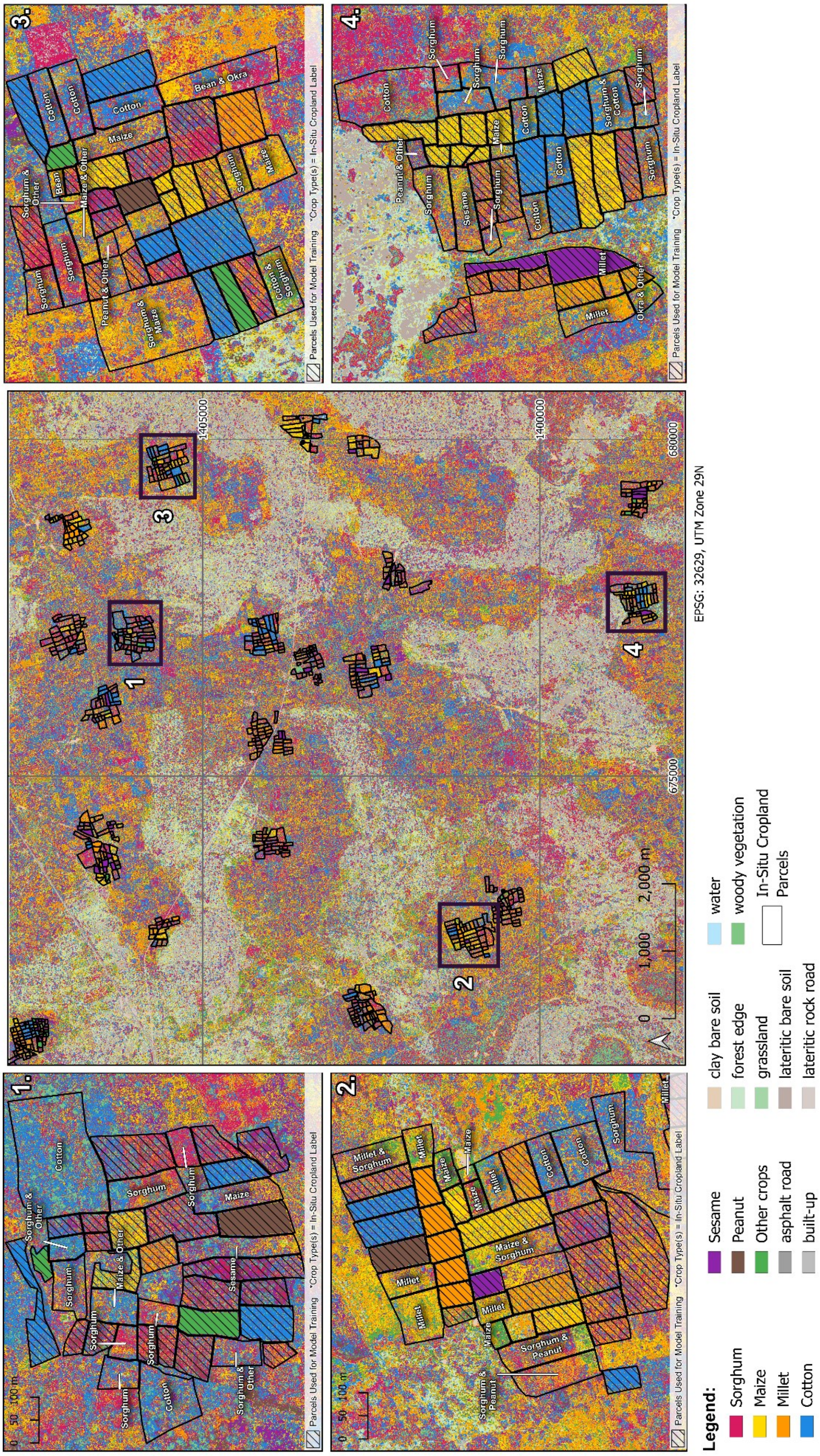

Figure 14: Classification of Study Area with Pléiades Scenario 2 Classifier

7.8. Refined Validation Approach of Sentinel-2 and PlanetScope Classifications Using VHR Cropland Mask

While single-date Pléiades imagery performed poorly for crop type discrimination, its superior spatial resolution (0.5m) offers a valuable methodological opportunity to refine the validation of Sentinel-2 and PlanetScope classifications.

While multi-temporal Sentinel-2 and PlanetScope classifications achieved substantially higher accuracy, validation of these results is confounded by the inherent heterogeneity of agricultural parcels. Non-crop features like trees and bare soil patches create a complex mixture of elements, introducing noise that affects validation. While a tree mask was previously implemented and tested, it did not fully account for this internal parcel complexity.

This section explores if a VHR-derived cropland mask enables a more precise evaluation of classification performance. By focusing validation on actual crop pixels by systematically excluding non-crop elements from parcels, this approach aims to reveal the classifiers' performance on cropland itself, free from the confounding effects of intra-parcel heterogeneity.

Methodologically the refined validation approach consists of three key steps:

(1) Generation of a VHR cropland/non-cropland mask: Using the Pléiades image (0.5m resolution), a binary cropland mask is created through supervised classification following a similar sampling approach to that used in the crop type classification but adapted to focus on the binary distinction between crop and non-crop elements.

(2) Masking validation data of Sentinel-2 and PlanetScope: The high-resolution mask is aggregated to Sentinel-2 (10m) and PlanetScope (3m) resolution, resulting in cropland proportion values for each sensors respective pixels. Different threshold approaches (e.g., 50% cropland proportion) are then applied to filter validation pixels, retaining only those which meet the cropland coverage threshold.

(3) Re-evaluation of classification accuracy: The Sentinel-2 and PlanetScope Scenario 2 classification is reassessed using the filtered validation sets, providing additional accurate measures of the classifier's ability to discriminate between crop types without the confounding influence of non-crop elements.

7.8.1. Generation of Very High-Resolution Cropland Mask

7.8.1.1. Methodology

For the cropland mask, the Pléiades imagery was classified using an adapted sample set to focus on the binary distinction between crop and non-crop elements. Cropland was manually sampled from within the in-situ data, ensuring that only pixels covered by dense crop vegetation were sampled, effectively isolating the spectral signatures of crops from within the heterogenous parcels.

This could only be done in sufficient quantities for the major crop types (sorghum, maize, millet, cotton, sesame, peanut) and okra so the sampling was limited to these crops. The non-cropland classes previously sampled were retained, but additional samples were added to the woody-vegetation class to increase the representation. The forest-edge class was removed after it was found to not be beneficial for the classification. No SMOTE oversampling was applied to the training data, as not to introduce synthetic pixels that could distort the spectral signatures of the crops. The classification was performed using a Random Forest classifier with 100 trees. The results were validated with a class-wise assessment evaluating the model's performance across all classes as well as a binary validation specifically focusing on the model's discriminative capacity between cropland and non-cropland.

7.8.1.2. Results

Binary cropland/non-cropland classification (see Table 13) shows an overall accuracy of 91.63%. The model exhibits high performance in cropland identification (F1-score: 95.05%, recall: 96.50% precision: 93.64%), effectively capturing most pixels covered by crops with minimal omission errors. However, non-cropland classification shows moderate performance (precision: 79.51%, recall: 67.48%, F1-score: 73.00%), indicating a bias of the classifier toward cropland prediction. The confusion matrix (see Table 70 in Annex) reveals that the main driver of the cropland prediction within non-cropland is misclassification of the grassland and woody-vegetation classes as cropland. The classifiers systemic bias toward cropland prediction represents a limitation for balanced land cover mapping, however it aligns with the methodological needs for creating a high-recall cropland mask that can be used to mask the parcel validation data. The model's ability to accurately identify cropland areas is deemed sufficient for this purpose and was used to predict the whole study area. An examination (see Figure 15) of the classification of the study area reveals that the prediction effectively captures the spatial distribution of cropland and is able to capture the presence of non-crop features such as trees and bare soil patches within cropland.

Table 13: Binary classification report for the Pléiades cropland/non-cropland classification

	precision	*recall*	*F1-score*	*support*
Cropland	0.9364	0.9650	**0.9505**	208588
Non-Cropland	0.7951	0.6748	**0.7300**	42039
accuracy			**0.9163**	250627
macro avg	0.8658	0.8199	0.8402	
weighted avg	0.9127	0.9163	0.9135	

Figure 15: Cropland/Non-Cropland classification derived from Pléiades image

7.8.2. Validation of Sentinel-2 and PlanetScope Classifications with different Cropland Proportion Thresholds

7.8.2.1. Methodology

A binary cropland mask is extracted from the Pléiades classification scene by isolating the crop class values. This binary mask is then spatially resampled to match both the 10m Sentinel-2 grid and the 3m PlanetScope grid using an area-weighted averaging approach, which calculates the percentage of cropland coverage within each pixel of the respective sensors. This resampling process preserves the sub-pixel spatial information during downscaling, resulting in continuous cropland percentage masks (0-100%) that quantify the proportion of each pixel covered by cropland at both 10m (Sentinel-2) and 3m (PlanetScope) resolutions.

The resampled cropland masks are then applied to the sensor validation data of the respective Scenario 2 classifier, retaining only validation pixels meeting defined thresholds (0, 10, 25, 50, 75, and 90%) of cropland coverage. The filtered validation datasets are then used to re-evaluate both the Sentinel-2 and PlanetScope classification accuracies.

7.8.2.2. Results

For Sentinel-2, as the threshold increases from 0% to 90 %, the decreasing support counts across thresholds confirm that mixed pixels are being systematically excluded. The overall validation accuracy of the classification improves from 76.3% to 77.5%. The macro-averaged F1-score (see Table 14) also increases from 0.62 to 0.65 with precision improving while recall remains stable. Sorghum, cotton and millet show modest performance gains (Table 15). Maize stands out as the only crop with a decrease in F1-score. Notably minority crops like peanut and rice show substantial relative improvements.

Table 14: Accuracy Metrics of Sentinel-2 (Scenario 2) Classification with different Cropland Validation Thresholds (Chapter 7.8.2.)

	Cropland Threshold					
	0 %	**10 %**	**25%**	**50%**	**75%**	**90%**
accuracy	0.763	0.765	0.768	0.770	0.770	0.775
macro avg	0.62	0.62	0.63	0.63	0.63	0.65

Table 15: Classification Performance of Sentinel-2 Scenario 2 classifier with 0 % and 90 % Cropland Validation Thresholds

	0% Cropland Threshold				**90% Cropland Threshold**			
<u>*Crop Class*</u>	**Precision**	**Recall**	**F1-Score**	**Support**	**Precision**	**Recall**	**F1-Score**	**Support**
Sorghum	0.764	0.872	**0.814**	3327	0.782	0.865	**0.821**	2174
Maize	0.809	0.590	**0.682**	782	0.767	0.562	**0.649**	452
Millet	0.774	0.550	**0.643**	604	0.791	0.614	**0.691**	370
Cotton	0.791	0.810	**0.801**	1348	0.780	0.841	**0.810**	769
Sesame	0.599	0.582	**0.590**	433	0.617	0.599	**0.608**	334
Peanut	0.294	0.250	**0.270**	40	0.345	0.286	**0.313**	35
Bean	0.080	0.017	**0.028**	120	0.118	0.025	**0.041**	80
Cowpea	0	0	**0**	9	0	0	**0**	5
Okra	0	0	**0**	44	0	0	**0**	8
Rice	0.333	0.800	**0.471**	5	0.667	0.800	**0.727**	5
Tomato	0	0	**0**	8	0	0	**0**	8
***macro avg**	0.62	0.63	0.62	7214	0.64	0.63	0.63	4734
***accuracy**			0.76	7214			0.78	4734

*Across all classes (Cropland & Non-Cropland)

For PlanetScope, the results show a decrease in support counts of around 25 % between the 0 % and 90 % threshold. The overall validation accuracy improves from 74.4 % to 75.8 %, with the macro-averaged F1-score increasing from 0.61 to 0.63 (see Table 16**)**, driven by both improvements in precision and recall. Major crops show varying degrees of improvement: Sorghum, millet, cotton and sesame show and increase in F1-scores. Maize, like for the Sentinel-2 results, shows a decrease in F1-score driven by a decrease in precision (see Table 17). Minor crops like sesame and peanut exhibit substantial relative gains, with rice showing the largest relative improvement.

Table 16: Accuracy metrics of all classes in PlanetScope (Scenario 2) Classification with different Cropland Validation Thresholds

	Cropland Threshold					
	0 %	**10 %**	**25%**	**50%**	**75%**	**90%**
accuracy	0.744	0.749	0.751	0.754	0.758	0.760
macro avg. F1	0.61	0.62	0.62	0.63	0.63	0.63

Table 17: Classification Performance of PlanetScope Scenario 2 classifier with 0 % and 90 % Cropland Validation Thresholds

	0% Cropland Threshold				**90% Cropland Threshold**			
<u>*Crop Class*</u>	**Precision**	**Recall**	**F1-Score**	**Support**	**Precision**	**Recall**	**F1-Score**	**Support**
Sorghum	0.748	0.891	**0.813**	53033	0.782	0.888	**0.832**	40443
Maize	0.709	0.607	**0.654**	14125	0.658	0.619	**0.638**	9545
Millet	0.771	0.634	**0.696**	11521	0.810	0.679	**0.739**	8263
Cotton	0.841	0.739	**0.786**	20742	0.835	0.766	**0.799**	14714
Sesame	0.559	0.445	**0.496**	7496	0.593	0.505	**0.545**	6085
Peanut	0.285	0.333	**0.307**	992	0.338	0.355	**0.346**	926
Bean	0.009	0.003	**0.004**	1796	0.010	0.004	**0.005**	1418
Cowpea	0	0	**0**	224	0	0	**0**	157
Okra	0	0	**0**	810	0	0	**0**	376
Rice	0.207	0.761	**0.326**	46	0.288	0.780	**0.421**	41
Tomato	0	0	**0**	246	0	0	**0**	239
***macro avg**	0.62	0.62	0.61	113550	0.64	0.63	0.63	84726
***accuracy**			0.74	113550			0.76	84726

*Across all classes (Cropland & Non-Cropland)

The results show that applying increasing cropland proportion thresholds to validation data, modestly but consistently improves classification performance across both sensors. Both sensors demonstrate improved overall accuracy and macro-averaged F1-scores. Most crop types show classification improvements, with particularly notable gains for rice, peanut, and sesame.

8. Discussion

8.1. Performance Differences across Sensors: Resolution and Temporal Data Depth

The different classification scenarios across the three sensors (Sentinel-2, PlanetScope, and Pléiades) reveal interesting insights into performance differences in crop type classification linked to sensor characteristics, temporal data depth, and spatial resolution. The baseline non-tree masked classification results showed that Sentinel-2 achieved the highest overall accuracy with a validation accuracy of 71.2 % and a macro average F1-score of 0.59, followed by PlanetScope with an overall accuracy of 70.2 % and a macro average F1-score of 0.57. Pléiades performed poorly with an overall accuracy of 25.1 % and a macro average F1-score of 0.32

This decreasing performance tangent is primarily explained by the critical role of temporal information in crop classification. Sentinel-2 and PlanetScope provided multi-temporal observations (9 and 11 images respectively) capturing the full phenological cycle from the beginning of May to the end of October, whereas Pléiades was limited to a single acquisition from October 19th. This temporal dimension allowed Sentinel-2 and PlanetScope to leverage distinctive phenological trajectories of bands and indices throughout the growing season, while Pléiades' single snapshot captured only a static view of the fields at a single point in time.

The similar performance between Sentinel-2 and PlanetScope is particularly noteworthy given their differences in both spatial resolution (10m vs 3m) and spectral capabilities. Sentinel-2's broader spectral range, including short wave infrared bands that enable integration of additional indices appears to not lead to major performance gains compared to PlanetScope, which is limited to bands in wavelengths from VIR and NIR domain. Conversely PlanetScope's finer spatial resolution (3m) also does not seem to translate into improved classification accuracy compared to Sentinel-2. These factors could potentially also counterbalance each other, leading to similar performance outcomes across the two sensors, however this was not explored in this study.

Cloud cover impacted the creation of time series for both Sentinel-2 and PlanetScope. The region is characterised by incessant cloud cover during the rainy season, which limited the number of usable images for classification. Biweekly composites could not be generated for key months. For Sentinel-2, June was excluded and August-September merged into one composite and for PlanetScope, September was excluded and August merged into one composite. This limitation in temporal depth of information, particularly in the key months of vegetative stage peak likely constrained the classification performance of both sensors. This highlights the need for integrating additional information sources in study areas with persistent cloud cover to fill these data gaps, such as SAR observations.

Across all sensors, a consistent hierarchy emerged in classification performance among crop types. More featured crops (sorghum, cotton, millet, maize) consistently outperformed minor crops (sesame, peanut) and "Other" crops (bean, okra, rice, cowpea, tomato), reflecting the class representation in the training data. This pattern persisted regardless of spatial resolution and sampling strategy, suggesting that class imbalance remains a fundamental challenge even when spatial detail and therefore pixel count of the training data is increased. For Sentinel-2 and PlanetScope, the performance hierarchy was remarkably similar. For both sensors, sorghum achieved the highest F1-scores (0.79), followed by cotton (0.72), millet (0.68-0.69), and maize (0.63-0.67). This hierarchy was also exhibited by Pléiades, but with substantially lower F1-scores across all crops. Misclassification patterns were also remarkably consistent across sensors, with sorghum frequently gaining misclassified pixels from other crops. This persistent pattern indicates that class imbalance effects on misclassification transcend resolution differences here.

The limited temporal data depth and underrepresentation of minority crops in the training data, constrain the interpretation of the classification results. It would be interesting to explore the impact of both these factors in a more rigorous manner, for example by systematically varying the temporal data depth and crop-specific samples to compare the impact on classification performance and to establish minimum requirements for reliable crop type mapping in this region. Further, a systematic investigation of feature importance across bands and derived indices would reveal which biophysical properties captured by optical sensors most effectively discriminate between crop types in this context, potentially simplifying operational implementations. Such a structured cross-comparison could potentially clarify whether higher classification performance stems from specific spectral capabilities (e.g., Sentinel-2's SWIR and red-edge bands), spatial resolution, or primarily from temporal information density, helping to optimize future sensor selection for similar applications.

8.2. Integrating Intra-Field Heterogeneity Information

The integration of heterogeneity information showed to have an impact on classification performance. Scenario 2, which incorporated recategorized in-situ heterogeneity assessments, showed the most substantial improvements of around 4 percent for both Sentinel-2 (overall accuracy increase to 76.3%) and PlanetScope (increase to 74.4%). This finding is particularly significant as it postulates that field-observed heterogeneity, despite its complex relationship with spectral variability and limited visual indicators of systematic parcel differences as established in Chapter 6, can provide a valuable input for discriminating crop types in classification.

Interestingly, the performance gains from heterogeneity integration per Scenario 2 were most pronounced for minority crops. For Sentinel-2, sesame’s F1-score improved substantially from 0.28 to 0.59, and peanut from 0.04 to 0.27. Similarly, PlanetScope also showed substantial improvements for sesame (0.21 to 0.50) and peanut (0.11 to 0.31) and for the minority crop rice. The observed pattern suggests that in-situ heterogeneity information particularly enhances classification discrimination of minority crops from majority crops. It was further found that the lower heterogeneity subsets of parcels were generally more accurately classified than higher heterogeneity parcels. The results point to potential benefits of leveraging field heterogeneity observations to enhance spectral separability between crop types.

Scenario 3, which used Pléiades derived parcel-level coefficient of variation of NDVI as a proxy for cropland heterogeneity had less of an impact on classification performance. For Sentinel-2, overall accuracy decreased from 71.7% to 69.7%, while PlanetScope showed only a slight improvement from 70.2% to 71.1%. For Pléiades, Scenario 3 produced more substantial improvements, with overall accuracy increasing from 25.1% to 27.9% as opposed to Scenario 2's 26.8%. Since Scenario 3 utilized CV of NDVI derived from the single October 19th Pléiades image, it captured only a static snapshot of spectral variability that likely reflected late-season conditions, including partial harvesting activities on parcels. This temporal limitation could explain the observed performance patterns across sensors. For Pléiades classification, which already operated with this single-date constraint, the CV metric naturally aligned with the imagery's existing spectral patterns, providing complementary information leading to a small improvement in classification accuracy. Conversely, for Sentinel-2 and PlanetScope, which leveraged multi-temporal data spanning the growing season, this static end-season parcel metric offered little discriminative value by imposing a temporally anchored metric on the classification of the entire time series.

8.3. Effects of Masking Trees from the Training Data

Tree masking of the training parcels yielded consistently minimal impact across all sensors, with accuracy improvements of less than 1 percentage point when validating both on tree-masked and non-masked data. This finding is surprising given the prevalence of trees in the agricultural parcels. Several factors may contribute to this result. For one, the random forest algorithm may have successfully learned to account for tree presence as part of the typical spectral signature of the crop type classes. For example, the classifier might learn that a pixel of "sorghum with scattered trees" has a particular composite spectral signature across the bands and indices that remains distinguishable from "millet with scattered trees." By training the algorithm on a diverse set of tree-crop spectral mixtures, the random forest classifier may have effectively incorporated these tree-crop spectral mixtures into its decision rule. Other explanations could be that the tree mask may have inadequately captured the trees in the study area, meaning that some tree-crop mixtures were nevertheless still present in the training data after masking, thereby diminishing the effect of trying to remove them. The results suggest that a tree mask for Sentinel-2 and PlanetScope may not be worth the additional processing complexity in operational crop mapping in this context and application. This aligns with the findings of Lambert et al. (2018) who similarly concluded that tree masking does not improve crop type classification accuracy with Sentinel-2 imagery in Mali's cotton belt.

A more definitive assessment of tree masking effects would have required a manually digitized high-precision tree mask. While labour-intensive, it would have provided a more definitive test of the hypothesis that tree presence confounds crop type classification. The mask employed in this study contained omission errors, potentially explaining the minimal difference observed between masked and unmasked classification results. A key limitation in creating an effective tree mask was the timing of the Pléiades image acquisition, which coincided with active crop cover hampering the segmentation of tree canopies. While acquiring additional VHR imagery during the non-growing season would be cost-prohibitive, PlanetScope imagery could have offered an alternative approach. Despite its lower resolution blurring precise tree boundaries, PlanetScope imagery from the dry season, when crop parcels feature predominantly bare soil, could potentially yield a tree mask with fewer omission errors, even if lacking the precision of masks derived from sub-metre imagery.

8.4. Pléiades for Cropland Masking and Effects of Integration in Validation

The Pléiades imagery was leveraged to create a binary cropland/non-cropland mask for the study area. Despite exhibiting some bias toward cropland prediction in the landscape, this mask effectively distinguished between crop and non-crop features within agricultural parcels. Although the processed mask could have been utilized in the training phase to exclude non-cropland elements from training samples and to evaluate the impact of all non-crop feature removal on classification performance, it was instead applied exclusively to validation data. In future studies, employing this mask during model training could provide valuable insights into how holistic filtering of non-cropland elements affects classification accuracy across different sensor platforms.

Using this mask to filter Sentinel-2 and PlanetScope validation pixels with increasing cropland proportion thresholds led to consistent modest improvements in classification performance with both sensors showing improved overall accuracy (by 1.2% and 1.4 % respectively) and macro-averaged F1-scores at the most stringent threshold of 90% cropland proportion. Practically, this provides a clearer view of classifier capability, as it demonstrates that validation approaches incorporating non-crop elements within agricultural parcels underestimate true crop classification performance. Focusing validation on areas with high crop proportion gives alternative metrics on the classifiers' actual capabilities.

8.5. Need for Adapting Sampling and Data Preparation to Spatial Detail

What became evident from the results is that classification at different resolutions requires different sampling approaches and data preparation. For Sentinel-2 and PlanetScope, correctly classified cropland generally covered entire parcels cohesively. Misclassifications occured between crop types, but classification results generally were able to exhibit parcel level coherence. This coherence can be largely attributed to the sensor resolution. Spectral mixing on a pixel level at these resolutions effectively aggregate within-field heterogeneity into classifiable, crop-specific signatures. The within-field landcover complexity manifests itself within the parcels as varying degrees of a mix of bare-soil, cropland, patches of grassland, trees, and in some cases small built-up structures on the field. At lower resolutions (10m for Sentinel-2 and 3m for PlanetScope), the mixed-pixel effect inherently generalizes the within-field landcover complexity, creating relatively homogeneous spectral signatures that correspond well to overall crop types at the parcel level. This natural aggregation effectively functions as an implicit spatial filter, where sub-pixel elements, including scattered trees, small bare soil patches, relief features, and isolated built-up structures, are integrated into composite signatures that can remain distinctive between crop types.

For example, a 10m Sentinel-2 pixel covering a sorghum field with 30% bare soil and 70% crop vegetation creates a characteristic “sorghum with bare soil” signature that can remain distinguishable from a “millet with bare soil” signature. Conversely, Pléiades’ 0.5m resolution fundamentally alters this dynamic by resolving the heterogeneity that exists within agricultural fields, essentially disaggregating the composite signatures into their constituent elements. At this resolution, a significant proportion of pixels within crop fields directly capture non-crop features rather than mixtures of crops and non-crop features. Tree canopies, bare soil patches, and small structures are represented as distinct constituent pixels rather than subpixel mixtures. For instance, a 4 m^2 clearing within a cotton field would occupy approximately 16 distinct “bare soil” pixels in Pleiades imagery, while being subsumed into the general cotton signature of a single pixel in Sentinel-2, of which it covers only 1/25th of the pixel area. The training data for Pleiades consequently contains a higher proportion of pixels representing entirely non-crop elements but that are labelled as one of the crop types, leading to increased within-class variance and reduced between-class separability. This is especially relevant since not only cropland, but also non-cropland land cover types were classified, such as bare soil, built-up structures, and forest. Therefore, the classifier can be expected to learn associations between the spectral signatures of non-crop features and crop type labels, rather than solely the actual crop signatures. This creates a fundamental challenge where the model learns that “Sorghum” encompasses not just the crop’s spectral signature but also distinct signatures of bare soil patches, small built-up structures, and scattered trees that happen to fall within sorghum-labelled parcels. This leads to a situation where the classifier is trained on a mixture of crop and non-crop signatures, resulting in a model that is less capable of distinguishing between these elements. Consequently, when encountering these non-crop pixels in the validation data, the classifier could reasonably assign them to the crop types with which they most frequently co-occurred during training. This confusion is evidenced by the poor performance of some natural landcover classes (e.g woody vegetation, forest edge) in the Pléiades classification, which are frequently misclassified as crops.

These patterns of misclassification demonstrate how high-resolution imagery can paradoxically lead to a reduction in classification performance by decomposing the integrated spectral signatures that characterize crop-specific parcels at lower resolutions. Addressing this would require fundamentally different sampling approaches tailored to the specific spatial resolution of the imagery. These could include segmenting of the cropland parcel and spectral thresholding approaches that employ vegetation indices to isolate crop pixels from non-crop pixels on parcels or integrating a cropland/non-cropland mask. Texture metrics that capture spatial relationships between neighbouring pixels could be beneficial in constructing contextual information (Ursani et al., 2012). Additionally, the sampling scheme of non-cropland would need to be adapted to account for the increased spatial detail of Pléiades imagery. The current sampling approach, with transition classes such as forest edge, which proved to be beneficial for PlanetScope and Sentinel-2 classification, showed to not be suitable for Pléiades where these transitions decompose into more distinct pixels such as bare soil, shadows, and trees.

Fundamentally this also ties in into the question of what we are trying to classify and why. The classification objective significantly influences methodological choices and determines what constitutes success in crop type mapping. If the goal is parcel-level crop type attribution, by for example, designating entire fields as “sorghum” or “cotton” or “sorghum-cotton mixed” for agricultural statistics or planning purposes, then the resolution of Sentinel-2 with its implicit spatial aggregation of heterogeneous elements may be preferable. Conversely, if the objective is precise surface area estimation of actual crop vegetation within parcels for yield prediction, then VHR imagery becomes necessary but requires very different methodological approaches. Both could also be achieved through a two-stage classification process: first classifying all constituent elements within parcels (crop vegetation, bare soil, trees, etc.) with appropriate sampling strategies for each element using VHR imagery and then aggregating these results to parcel-level labels using a rule-based pixel aggregation approach.

8.6. Study Limitations and Recommendations for Future Research

8.6.1. Parcel-Level Analysis

The parcel level analysis could have been enhanced through additional approaches that remained beyond the study's scope. While NDVI served as the primary vegetation index, incorporating specialized indices such as the Bare Soil Index (BSI), SAVI, and Enhanced Vegetation Index (EVI) would have provided complementary information on vegetation, soil exposure and spectral characteristics of parcels across heterogeneity classes. Such multi-index analysis could have potentially better isolated specific drivers of crop variability.

Additionally, the analysis primarily focused on monocropped parcels, as intercropped parcels lacked sufficient representation across heterogeneity classes for robust analysis. This limitation is particularly significant given that intercropping represents an important dimension of smallholder farming systems that likely exhibits distinctive spectral-temporal signatures and heterogeneity patterns. Future research with a more balanced representation of diverse cropping patterns could reveal important insights on heterogeneity and spectral characteristics of monocropped and intercropped parcels.

8.6.2. In-Situ Data Constraints

The composite nature of the in-situ heterogeneity assessment is viewed as a principal limitation in this study. By consolidating multiple factors, such as field maintenance, crop structure, density, and weed presence, into a single ordinal score, the assessment obscures potentially significant relationships between specific heterogeneity components and spectral properties. Parcels classified as having a very high heterogeneity may display higher NDVI than other parcels because of weed biomass, yet they could equally display lower NDVI if bare soil predominates or vegetation growth is otherwise constrained. The opposing factors could also offset one another. This is likely to be the primary reason for the observed intricate relationship between in-situ assessed parcel heterogeneity and parcel level spectral metrics, with no clear patterns emerging. Without disaggregated information on the individual components, such as a separate weed-presence category, we only see the composite heterogeneity label, which does not allow for a more nuanced understanding of the spectral characteristics of parcels. As a result, potential relationships are masked. If information on the individual components of the heterogeneity assessment were available, the hypothesized relationships, for example between higher weed presence and higher NDVI, could be verified.

A further constraint is the static nature of the heterogeneity assessment, which captured parcel conditions at a single point in the season. Whereas crop type can be assumed to remain stable for monocropped parcels throughout the growing period, parcel heterogeneity can potentially change substantially due to environmental changes and field management practises. It cannot be determined whether the heterogeneity label assigned to a parcel in late October/early November would have been the same had the parcel been assessed at a different time in the season.

Disentangling the drivers of parcel-level heterogeneity from its composite in-situ label using satellite imagery alone proved to be challenging. Crop and weed presence cannot be readily distinguished from each other, even on Pléiades imagery. Bare-soil detection is feasible in the late-October Pléiades image, but that date coincides with the onset of harvesting activities, so bare-soil exposure may be transient. Temporal misalignment between in-situ data collection and image acquisition, although only by days to weeks, complicates interpretation further. Sole reliance on the single Pléiades image for a more detailed parcel view is therefore problematic as it is not fully representative of field conditions at the time of the field inspection.

8.6.3. Recommendations for Future Research

Based on the findings of this study, some recommendations can be made for future field campaigns and studies on crop type mapping in smallholder cropping systems:

- **A more comprehensive assessment of heterogeneity components**: Future field campaigns should consider collecting more detailed information on the individual components of parcel heterogeneity, such as weed presence, crop structure, and bare soil exposure. This would allow for a more nuanced understanding of varying spectral characteristics of parcels of the same crop type and heterogeneity class and could potentially reveal relationships between specific components and spectral properties that are currently obscured by the composite nature of the assessment.
- **Integrating UAV imagery**: Incorporating UAV imagery alongside field campaigns could provide sub-decimetre resolution data capturing fine-scale parcel characteristics. This approach would help overcome the limitations of composite heterogeneity assessments by enabling the systematic evaluation of all individual heterogeneity components which was not possible with sub-metre satellite imagery and field photographs alone.
- **Utilizing larger sample sizes:** Future studies should aim to utilize larger sample sets from field campaigns to better represent the diversity of crop types and heterogeneity conditions and have a significant set for each crop-heterogeneity combination that

would allow for filtering out anomalous samples without compromising statistical robustness in classification.

- **Incorporating intercropping in classification:** When utilizing larger sample sets, future studies should incorporate intercropped parcels into the analysis, to assess if they can be discriminated from their monocropped counterparts and be successfully incorporated in classification frameworks spanning multiple crop types. Despite intercropping's prevalence in smallholder systems, no studies could be identified that successfully classified multiple intercropping combinations simultaneously, focusing mainly on single intercrop combinations (Akinyemi et al., 2023).
- **Leveraging dry season imagery for improved tree masking**: Future studies should consider utilising VHR imagery at a similar resolution to Pléiades or alternatively PlanetScope imagery from the dry season to create a tree mask with fewer omission errors. This would allow for a more comprehensive exclusion of tree covered pixels in the training data and a more accurate assessment of the impact of tree presence on classification performance.
- **Employing VHR imagery for comprehensive cropland masking**: A Pléiades-derived cropland mask could be adapted for filtering training samples to focus exclusively on pixels representing actual crop vegetation. This approach would require accounting for the temporal dynamics of non-crop features, particularly the timing and impact of harvesting activities on bare soil exposure as well as weed presence within fields.
- **Integrating SAR data to address cloud cover limitations:** Future studies should consider integrating SAR data, such as Sentinel-1, to complement optical imagery in areas with persistent cloud cover. Kpienbaareh et al. (2021) have demonstrated that integrating SAR with PlanetScope and Sentinel-2 imagery can improve crop type classification performance in smallholder cropping systems.
- **Multi-sensor fusion of optical data:** Future studies should explore the potential of multi-sensor fusion approaches of optical data to combine the strengths of different sensors to enhance crop type classification performance. For example, integrating Sentinel-2 and PlanetScope data could leverage the higher spatial resolution of PlanetScope while benefiting from the broader spectral range of Sentinel-2, potentially leading to improved classification accuracy (Heiss et al., 2025; Kpienbaareh et al., 2021; Samadzadegan et al., 2025).

9. Conclusion

This study set out to explore possibilities for improving crop type mapping in smallholder systems by leveraging intra-parcel heterogeneity information. The research utilized a unique dataset collected in Mali's cotton belt of field-observed heterogeneity for various crops and integrated it with multi-resolution satellite imagery from Sentinel-2, PlanetScope, and Pléiades. Analysis of this dataset revealed that while field-assessed heterogeneity levels correspond to some differences in spectral-temporal signatures of parcels, disentangling specific drivers of heterogeneity using satellite data alone proved to be difficult.

The comparative scenarios across Sentinel-2, PlanetScope and Pléiades indicate that classification performance is driven primarily by temporal data depth rather than varying spatial resolution. Time series for Sentinel-2 and PlanetScope captured different crop phenological stages and outperformed the Pléiades snapshot, which provided only a static late-season view of parcels.

Characterizing and leveraging intra-parcel heterogeneity offers a promising yet challenging avenue for improving crop type mapping in smallholder landscapes. Incorporating recategorized field-assessed heterogeneity descriptors into classification frameworks was shown to have an impact on classification performance with performance gains in overall accuracy of around 4 percent in Sentinel-2 and PlanetScope classifications. Conversely, leveraging parcel-level NDVI variability metrics derived from VHR imagery as an additional input yielded no improvements in multi-temporal classification. Also masking out trees within parcels using the VHR data resulted in only marginal changes to accuracy metrics.

While the study investigated excluding trees from the training data, a more comprehensive approach to non-crop feature exclusion should be explored. However, it can be tentatively said that misclassification largely arises from spectral-temporal overlap between crop classes and not merely from the presence of extraneous features. This highlights the need for leveraging denser time series that capture the full phenological cycle of crops to improve discrimination between crop types. Multi-sensor fusion approaches of optical sensors and leveraging SAR to overcome cloud cover limitations, should therefore be explored in future work. Achieving robust classification of single-crop fields is seen as a prerequisite before attempting to incorporate intercropping into classification frameworks.

The findings suggest that intra-parcel heterogeneity can influence crop mapping outcomes, but its effective integration requires more nuanced approaches. The limited gains from including a broad heterogeneity label imply that more detailed characterization of within-field variability is needed to substantially boost classification accuracy. Future research should pursue several directions to better exploit intra-parcel variability. More granular field data collection is recommended, breaking down heterogeneity into its components (e.g., separate indicators for weed presence, crop density, and bare soil) to allow clearer linkage with spectral signatures. Deploying UAVs in field campaigns could provide sub-decimetre resolution imagery, enabling the deconstruction of within-field heterogeneity by capturing field-variations that satellite sensors may miss.

References

Abawiera Wongnaa, C., Amoah Seyram, A., & Babu, S. (2024). A systematic review of climate change impacts, adaptation strategies, and policy development in West Africa. *Regional Sustainability*, *5*(2), 100137. https://doi.org/10.1016/j.regsus.2024.100137

Aguilar, R., Zurita-Milla, R., Izquierdo-Verdiguier, E., & A. de By, R. (2018). A Cloud-Based Multi-Temporal Ensemble Classifier to Map Smallholder Farming Systems. *Remote Sensing*, *10*(5), 729. https://doi.org/10.3390/rs10050729

Airbus Defence and Space Intelligence. (n.d.). *Pléiades Imagery User Guide*. https://storage.googleapis.com/p-ssp-iep-prod-8ff-strapi-uploads/210415_Airbus_Pleiades_Imagery_user_guide_be12b8f35b/210415_Airbus_Pleiades_Imagery_user_guide_be12b8f35b.pdf

Akinyemi, F. O., Rufin, P., Ibrahim, E. S., Hostert, P., Ogunsumi, L., Egbetokun, O., & Speranza, C. I. (2023). *Untangling intercropping in heterogeneous smallholder maize-cassava farming systems with remote sensing*. https://eartharxiv.org/repository/view/6422/

Anderegg, J., Yu, K., Aasen, H., Walter, A., Liebisch, F., & Hund, A. (2019). Spectral Vegetation Indices to Track Senescence Dynamics in Diverse Wheat Germplasm. *Frontiers in Plant Science*, *10*, 1749. https://doi.org/10.3389/fpls.2019.01749

Anjos, L., Gaistardo, C. C., Deckers, J., Dondeyne, S., Eberhardt, E., Gerasimova, M., Harms, B., Jones, A., Krasilnikov, P., Reinsch, T., Vargas, R., & Zhang, G.-L. (2015). *World reference base for soil resources 2014 International soil classification system for naming soils and creating legends for soil maps*. JRC Publications Repository. https://publications.jrc.ec.europa.eu/repository/handle/JRC91947

Baston, D. (2024). *exactextractr: Fast Extraction from Raster Datasets using Polygons* (Version R package version 0.10.0) [Computer software]. https://github.com/isciences/exactextractr

Becker-Reshef, I., Barker, B., Whitcraft, A., Oliva, P., Mobley, K., Justice, C., & Sahajpal, R. (2023). Crop Type Maps for Operational Global Agricultural Monitoring. *Scientific Data*, *10*(1), 172. https://doi.org/10.1038/s41597-023-02047-9

Belgiu, M., & 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. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, *65*(1), 2–16. https://doi.org/10.1016/j.isprsjprs.2009.06.004

Brandt, M., Tucker, C. J., Kariryaa, A., Rasmussen, K., Abel, C., Small, J., Chave, J., Rasmussen, L. V., Hiernaux, P., Diouf, A. A., Kergoat, L., Mertz, O., Igel, C., Gieseke, F., Schöning, J., Li, S., Melocik, K., Meyer, J., Sinno, S., … Fensholt, R. (2020). An unexpectedly large count of trees in the West African Sahara and Sahel. *Nature*, *587*(7832), 78–82. https://doi.org/10.1038/s41586-020-2824-5

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Choukri, M., Laamrani, A., & Chehbouni, A. (2024). Use of Optical and Radar Imagery for Crop Type Classification in Africa: A Review. *Sensors*, *24*(11), Article 11. https://doi.org/10.3390/s24113618

Debats, S. R., Luo, D., Estes, L. D., Fuchs, T. J., & Caylor, K. K. (2016). A generalized computer vision approach to mapping crop fields in heterogeneous agricultural landscapes. *Remote Sensing of Environment*, *179*, 210–221. https://doi.org/10.1016/j.rse.2016.03.010

Delrue, J., Bydekerke, L., Eerens, H., Gilliams, S., Piccard, I., & Swinnen, E. (2013). Crop mapping in countries with small-scale farming: A case study for West Shewa, Ethiopia. *International Journal of Remote Sensing*, *34*(7), 2566–2582. https://doi.org/10.1080/01431161.2012.747016

Dissa, A., Slingerland, M., Giller, K. E., & Descheemaeker, K. (2024). Navigating seasonality in cotton-based farming systems in southern Mali. *Climate Risk Management*, *46*, 100649. https://doi.org/10.1016/j.crm.2024.100649

Economics of Land Degradation (ELD) Initiative. (2020). *Mali case study: The economics of cotton production and land degradation in Mali. Policy brief.* https://www.eld-initiative.org/fileadmin/ELD_Filter_Tool/Case_Study_Mali_2020/Mali_2020_Cotton_ELD_Policy_Brief_EN.pdf#:~:text=Farming%20is%20the%20motor%20of,cotton%20production%20and%20is%20often

ESA. (n.d.). *Sentinel Toolbox Help—Resampling Methods*. Retrieved 13 August 2025, from https://seadas.gsfc.nasa.gov/help-9.2.0/general/overview/ResamplingMethods.html

European Space Agency. (2021, October 8). *Copernicus Sentinel-2 Collection 1 MSI Level-2A (L2A)*. Sentinel Online. https://sentinels.copernicus.eu/sentinel-data-access/sentinel-products/sentinel-2-data-products/collection-1-level-2a

European Space Agency. (2025). *S2 Mission*. https://sentiwiki.copernicus.eu/web/s2-mission

FAO. (n.d.). *FAO, Digital Earth Africa and Frontier SI's collaboration to enhance the use of Earth observations in Africa*. PRD-SDGHelpdesk. Retrieved 16 August 2025, from https://www.fao.org/sustainable-development-goals-helpdesk/champion/article-detail/fao--digital-earth-africa-and-frontier-si-s-collaboration-to-enhance-the-use-of-earth-observations-in-africa/en

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302–4315. https://doi.org/10.1002/joc.5086

Food and Agriculture Organization of the United Nations (FAO). (n.d.). *Smallholders and family farmers* (Sustainability Pathways – Fact Sheet). Retrieved 13 August 2025, from https://www.fao.org/fileadmin/templates/nr/sustainability_pathways/docs/Factsheet_SMALLHOLDERS.pdf#:~:text=Smallholders%20provide%20up%20to%2080,subsistence

Frazier, A. E., & Hemingway, B. L. (2021). A Technical Review of Planet Smallsat Data: Practical Considerations for Processing and Using PlanetScope Imagery. *Remote Sensing*, *13*(19), Article 19. https://doi.org/10.3390/rs13193930

Ghassemi, B., Izquierdo-Verdiguier, E., Verhegghen, A., Yordanov, M., Lemoine, G., Moreno Martínez, Á., De Marchi, D., van der Velde, M., Vuolo, F., & d'Andrimont, R. (2024). European Union crop map 2022: Earth observation's 10-meter dive into Europe's crop tapestry. *Scientific Data*, *11*(1), 1048. https://doi.org/10.1038/s41597-024-03884-y

Gilić, F., Gašparović, M., & Baučić, M. (2023). Current state and challenges in producing large-scale land cover maps: Review based on recent land cover products. *Geocarto International*, *38*(1), 2242693. https://doi.org/10.1080/10106049.2023.2242693

Giller, K. E., Delaune, T., Silva, J. V., van Wijk, M., Hammond, J., Descheemaeker, K., van de Ven, G., Schut, A. G. T., Taulya, G., Chikowo, R., & Andersson, J. A. (2021). Small farms and development in sub-Saharan Africa: Farming for food, for income or for lack of better options? *Food Security*, *13*(6), 1431–1454. https://doi.org/10.1007/s12571-021-01209-0

Gitelson, A. A., Gritz †, Y., & Merzlyak, M. N. (2003). Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *Journal of Plant Physiology*, *160*(3), 271–282. https://doi.org/10.1078/0176-1617-00887

Gitelson, A., Kaufman, Y., & Merzlyak, M. (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sensing of Environment*, *58*(3), 289–298. https://doi.org/10.1016/S0034-4257(96)00072-7

Gitelson, A., & Merzlyak, M. (1994). Quantitative estimation of chlorophyll-*a* using reflectance spectra: Experiments with autumn chestnut and maple leaves. *Journal of Photochemistry and Photobiology B: Biology*, *22*(3), 247–252. https://doi.org/10.1016/1011-1344(93)06963-4

Goffart, D., Dvorakova, K., Crucil, G., Curnel, Y., Limbourg, Q., Van Oost, K., Castaldi, F., Planchon, V., Goffart, J.-P., & van Wesemael, B. (2022). UAV Remote Sensing for Detecting within-Field Spatial Variation of Winter Wheat Growth and Links to Soil Properties and Historical Management Practices. A Case Study on Belgian Loamy Soil. *Remote Sensing*, *14*(12), Article 12. https://doi.org/10.3390/rs14122806

Gómez-Escalonilla, Diancoumba, O., Traoré, D. Y., Montero, E., Martín-Loeches, M., & Martínez-Santos, P. (2022). Multiclass spatial predictions of borehole yield in southern Mali by means of machine learning classifiers. *Journal of Hydrology: Regional Studies*, *44*, 101245. https://doi.org/10.1016/j.ejrh.2022.101245

Gumma, M. K., Tummala, K., Dixit, S., Collivignarelli, F., Holecz, F., Kolli, R. N., & Whitbread, A. M. (2022). Crop type identification and spatial mapping using Sentinel-2 satellite data with focus on field-level information. *Geocarto International*, *37*(7), 1833–1849. https://doi.org/10.1080/10106049.2020.1805029

Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J., & Strachan, I. B. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sensing of Environment*, *90*(3), 337–352. https://doi.org/10.1016/j.rse.2003.12.013

Hegarty-Craver, M., Polly, J., O'Neil, M., Ujeneza, N., Rineer, J., Beach, R. H., Lapidus, D., & Temple, D. S. (2020). Remote Crop Mapping at Scale: Using Satellite Imagery and UAV-Acquired Data as Ground Truth. *Remote Sensing*, *12*(12), Article 12. https://doi.org/10.3390/rs12121984

Heiss, N., Meier, J., Gessner, U., & Kuenzer, C. (2025). A Review: Potential of Earth Observation (EO) for Mapping Small-Scale Agriculture and Cropping Systems in West Africa. *Land*, *14*(1), 171. https://doi.org/10.3390/land14010171

Huang, X., Vrieling, A., Dou, Y., Li, X., & Nelson, A. (2025). Divergent crop mapping accuracies across different field types in smallholder farming regions. *International Journal of Applied Earth Observation and Geoinformation*, *139*, 1–13. https://doi.org/10.1016/j.jag.2025.104559

Huete, A. R., Liu, H. Q., Batchily, K., & van Leeuwen, W. (1997). A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sensing of Environment*, *59*(3), 440–451. https://doi.org/10.1016/S0034-4257(96)00112-5

Ibrahim, E. S., Rufin, P., Nill, L., Kamali, B., Nendel, C., & Hostert, P. (2021). Mapping Crop Types and Cropping Systems in Nigeria with Sentinel-2 Imagery. *Remote Sensing*, *13*(17), Article 17. https://doi.org/10.3390/rs13173523

Inglada, J., & Christophe, E. (2009). The Orfeo Toolbox remote sensing image processing software. *2009 IEEE International Geoscience and Remote Sensing Symposium*, *4*, IV-733-IV–736. https://doi.org/10.1109/IGARSS.2009.5417481

Karmakar, P., Teng, S. W., Murshed, M., Pang, S., Li, Y., & Lin, H. (2024). Crop monitoring by multimodal remote sensing: A review. *Remote Sensing Applications: Society and Environment*, *33*, 101093. https://doi.org/10.1016/j.rsase.2023.101093

Kerner, H., Nakalembe, C., Yang, A., Zvonkov, I., McWeeny, R., Tseng, G., & Becker-Reshef, I. (2024). How accurate are existing land cover maps for agriculture in Sub-Saharan Africa? *Scientific Data*, *11*(1), 486. https://doi.org/10.1038/s41597-024-03306-z

Kpienbaareh, D., Sun, X., Wang, J., Luginaah, I., Bezner Kerr, R., Lupafya, E., & Dakishoni, L. (2021). Crop Type and Land Cover Mapping in Northern Malawi Using the Integration of Sentinel-1, Sentinel-2, and PlanetScope Satellite Data. *Remote Sensing*, *13*(4), Article 4. https://doi.org/10.3390/rs13040700

Lambert, M.-J., Traore, P., Blaes, X., Baret, P., & Defourny, P. (2018). Estimating smallholder crops production at village level from Sentinel-2 time series in Mali's cotton belt. *Remote Sensing of Environment*, *216*. https://doi.org/10.1016/j.rse.2018.06.036

Lemaitre, G., Nogueira, F., & Aridas, C. K. (2016). *Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning* (arXiv:1609.06570). arXiv. https://doi.org/10.48550/arXiv.1609.06570

Li, S., Brandt, M., Fensholt, R., Kariryaa, A., Igel, C., Gieseke, F., Nord-Larsen, T., Oehmcke, S., Carlsen, A. H., Junttila, S., Tong, X., d'Aspremont, A., & Ciais, P. (2023). Deep learning enables image-based tree counting, crown segmentation, and height prediction at national scale. *PNAS Nexus*, *2*(4), pgad076. https://doi.org/10.1093/pnasnexus/pgad076

Lungu, G., Ndeke, C., Umer, Y., Mume, A. A., Hassen, A., & Chavula, P. (2024). An Overview of Trees in Smallholder Farmers' Agricultural Landscapes: A Case Study from Sub-Saharan Africa. *Asian Journal of Research in Agriculture and Forestry*, *10*(2), 119–129. https://doi.org/10.9734/ajraf/2024/v10i2291

Mahlayeye, M., Darvishzadeh, R., & Nelson, A. (2022). Cropping Patterns of Annual Crops: A Remote Sensing Review. *Remote Sensing*, *14*(10), Article 10. https://doi.org/10.3390/rs14102404

Mazarire, T. T., Ratshiedana, P. E., Nyamugama, A., Adam, E., & Chirima, G. (2020). Exploring machine learning algorithms for mapping crop types in a heterogeneous agriculture landscape using Sentinel-2 data. A case study of Free State Province, South Africa. *South African Journal of Geomatics*, *9*(2), 333–347. https://doi.org/10.4314/sajg.v9i2.22

Mbow, C., Rosenzweig, C., Barioni, L. G., Benton, T. G., Herrero, M., Krishnapilla, M., Liwenga, E., Pradhan, P., Rivera-Ferre, M. G., Sapkota, T., Tubiello, F. N., & Xu, Y. (2019). Chapter 5 "Food Security". In N. Benkeblia, A. Challino, A. Khan, & J. R. Porter (Eds.), *Climate Change and Land* (pp. 439–550). Cambridge University Press. https://doi.org/10.1017/9781009157988.007

McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, *17*(7), 1425–1432. https://doi.org/10.1080/01431169608948714

Merzlyak, M. N., Gitelson, A. A., Chivkunova, O. B., & Rakitin, V. YU. (1999). Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. *Physiologia Plantarum*, *106*(1), 135–141. https://doi.org/10.1034/j.1399-3054.1999.106119.x

Mhangara, P., Mapurisa, W., & Mudau, N. (2020). Comparison of Image Fusion Techniques Using Satellite Pour l'Observation de la Terre (SPOT) 6 Satellite Imagery. *Applied Sciences*, *10*(5), Article 5. https://doi.org/10.3390/app10051881

Montero, D., Aybar, C., Mahecha, M. D., Martinuzzi, F., Söchting, M., & Wieneke, S. (2023). A standardized catalogue of spectral indices to advance the use of remote sensing in Earth system research. *Scientific Data*, *10*(1), 197. https://doi.org/10.1038/s41597-023-02096-0

Nabil, M., Zhang, M., Bofana, J., Wu, B., Stein, A., Dong, T., Zeng, H., & Shang, J. (2020). Assessing factors impacting the spatial discrepancy of remote sensing based cropland products: A case study in Africa. *International Journal of Applied Earth Observation and Geoinformation*, *85*, 102010. https://doi.org/10.1016/j.jag.2019.102010

Nabil, M., Zhang, M., Wu, B., Bofana, J., & Elnashar, A. (2022). Constructing a 30m African Cropland Layer for 2016 by Integrating Multiple Remote sensing, crowdsourced, and Auxiliary Datasets. *Big Earth Data*, *6*(1), 54–76. https://doi.org/10.1080/20964471.2021.1914400

Orfeo ToolBox Team. (2018, June). *OTB CookBook Documentation—Release 6.6.0*. https://www.gisandbeers.com/GeoBazar/Tutoriales%20y%20manuales/Orfeo%20ToolBox%20Cook%20Book.pdf

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Müller, A., Nothman, J., Louppe, G., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). *Scikit-learn: Machine Learning in Python*. https://doi.org/10.48550/arXiv.1201.0490

Pérez-Hoyos, A., Rembold, F., Kerdiles, H., & Gallego, J. (2017). Comparison of Global Land Cover Datasets for Cropland Monitoring. *Remote Sensing*, *9*(11), Article 11. https://doi.org/10.3390/rs9111118

Phiri, D., Simwanda, M., Salekin, S., Nyirenda, V. R., Murayama, Y., & Ranagalage, M. (2020). Sentinel-2 Data for Land Cover/Use Mapping: A Review. *Remote Sensing*, *12*(14), 2291. https://doi.org/10.3390/rs12142291

Planet. (2023, December). *Planet Imagery Product Specifications*. https://assets.planet.com/docs/Planet_Combined_Imagery_Product_Specs_letter_screen.pdf

Planet Labs PBC. (2022, May). *Planet Imagery Product Specifications*. https://earth.esa.int/eogateway/documents/20142/37627/Planet_Combined_Imagery_Product_Specs_May_2022.pdf

Potapov, P., Turubanova, S., Hansen, M. C., Tyukavina, A., Zalles, V., Khan, A., Song, X.-P., Pickens, A., Shen, Q., & Cortez, J. (2022). Global maps of cropland extent and change show accelerated cropland expansion in the twenty-first century. *Nature Food*, *3*(1), 19–28. https://doi.org/10.1038/s43016-021-00429-z

Qi, J., Chehbouni, A., Huete, A. R., Kerr, Y. H., & Sorooshian, S. (1994). A modified soil adjusted vegetation index. *Remote Sensing of Environment*, *48*(2), 119–126. https://doi.org/10.1016/0034-4257(94)90134-1

Qiu, S., Zhu, Z., & He, B. (2019). Fmask 4.0: Improved cloud and cloud shadow detection in Landsats 4–8 and Sentinel-2 imagery. *Remote Sensing of Environment*, *231*, 111205. https://doi.org/10.1016/j.rse.2019.05.024

Richard, K., Abdel-Rahman, E. M., Subramanian, S., Nyasani, J. O., Thiel, M., Jozani, H., Borgemeister, C., & Landmann, T. (2017). Maize Cropping Systems Mapping Using RapidEye Observations in Agro-Ecological Landscapes in Kenya. *Sensors*, *17*(11), 2537. https://doi.org/10.3390/s17112537

Rikimaru, A., Roy, P., & Miyatake, S. (2002). Tropical forest cover density mapping. *Tropical Ecology*, *Vol 43*, 39–47.

Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1974, January 1). *Monitoring vegetation systems in the Great Plains with ERTS*. https://ntrs.nasa.gov/citations/19740022614

Samadzadegan, F., Toosi, A., & Dadrass Javan, F. (2025). A critical review on multi-sensor and multi-platform remote sensing data fusion approaches: Current status and prospects. *International Journal of Remote Sensing*, *46*(3), 1327–1402. https://doi.org/10.1080/01431161.2024.2429784

Sanchez, A. H., Picoli, M. C. A., Camara, G., Andrade, P. R., Chaves, M. E. D., Lechler, S., Soares, A. R., Marujo, R. F. B., Simões, R. E. O., Ferreira, K. R., & Queiroz, G. R. (2020). Comparison of Cloud Cover Detection Algorithms on Sentinel–2 Images of the Amazon Tropical Forest. *Remote Sensing*, *12*(8), Article 8. https://doi.org/10.3390/rs12081284

Sida, T. S., Chamberlin, J., Ayalew, H., Kosmowski, F., & Craufurd, P. (2021). Implications of intra-plot heterogeneity for yield estimation accuracy: Evidence from smallholder maize systems in Ethiopia. *Field Crops Research*, *267*, 108147. https://doi.org/10.1016/j.fcr.2021.108147

Soumaré, M. (2008). *Dynamique et durabilité des systèmes agraires à base de coton au Mali* (Mali) [Thesis, Université de Paris-Nanterre]. https://agritrop.cirad.fr/548140/

Sutton, W. R. (2022). *Climate Explainer: Food Security and Climate Change*. World Bank. https://www.worldbank.org/en/news/feature/2022/10/17/what-you-need-to-know-about-food-security-and-climate-change

Talukdar, S., Singha, P., Mahato, S., Shahfahad, Pal, S., Liou, Y.-A., & Rahman, A. (2020). Land-Use Land-Cover Classification by Machine Learning Classifiers for Satellite Observations—A Review. *Remote Sensing*, *12*(7), Article 7. https://doi.org/10.3390/rs12071135

Thenkabail, P. S., Teluguntla, P. G., Xiong, J., Oliphant, A., Congalton, R. G., Ozdogan, M., Gumma, M. K., Tilton, J. C., Giri, C., Milesi, C., Phalke, A., Massey, R., Yadav, K., Sankey, T., Zhong, Y., Aneece, I., & Foley, D. (2021). *Global cropland-extent product at 30-m resolution (GCEP30) derived from Landsat satellite time-series data for the year 2015 using multiple machine-learning algorithms on Google Earth Engine cloud* (Professional Paper). US Geological Survey. https://doi.org/10.3133/pp1868

Traore, B., Corbeels, M., van Wijk, M. T., Rufino, M. C., & Giller, K. E. (2013). Effects of climate variability and climate change on crop production in southern Mali. *European Journal of Agronomy*, *49*, 115–125. https://doi.org/10.1016/j.eja.2013.04.004

Traoré, P. C. S., Kouressy, M., Vaksmann, M., Tabo, R., Maikano, I., Traoré, S. B., & Cooper, P. (2007). Climate Prediction and Agriculture: What Is Different about Sudano-Sahelian West Africa? In M. V. K. Sivakumar & J. Hansen (Eds.), *Climate Prediction and Agriculture: Advances and Challenges* (pp. 189–203). Springer. https://doi.org/10.1007/978-3-540-44650-7_19

Trimble Inc. (n.d.). *Multiresolution segmentation*. ECognition Documentation: Reference Book – Algorithms & Processes – Segmentation Algorithms. Retrieved 4 August 2025, from https://docs.ecognition.com/eCognition_documentation/Reference%20Book/02%20Algorithms%20and%20Processes/2%20Segmentation%20Algorithms/Multiresolution%20Segmentation.htm

Ursani, A., Kpalma, K., Lelong, C., & Ronsin, J. (2012). Fusion of Textural and Spectral Information for Tree Crop and Other Agricultural Cover Mapping With Very-High Resolution Satellite Images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing - IEEE J SEL TOP APPL EARTH OBS*, *5*, 225–235. https://doi.org/10.1109/JSTARS.2011.2170289

USDA National Agricultural Statistics Service. (2024). *National Agricultural Statistics Service Cropland Data Layer*. https://nassgeodata.gmu.edu/CropScape/

Van Deventer, A. P., Ward, A. D., Gowda, P. M., & Lyon, J. G. (1997). Using thematic mapper data to identify contrasting soil plains and tillage practices. *Photogrammetric Engineering and Remote Sensing*, *63*(1), 87–93.

Van Tricht, K., Degerickx, J., Gilliams, S., Zanaga, D., Battude, M., Grosu, A., Brombacher, J., Lesiv, M., Bayas, J. C. L., Karanam, S., Fritz, S., Becker-Reshef, I., Franch, B., Mollà-Bononad, B., Boogaard, H., Pratihast, A. K., Koetz, B., & Szantoi, Z. (2023). WorldCereal: A dynamic open-source system for global-scale, seasonal, and reproducible crop and irrigation mapping. *Earth System Science Data*, *15*(12), 5491–5515. https://doi.org/10.5194/essd-15-5491-2023

Vanlauwe, B., AbdelGadir, A. H., Adewopo, J., Adjei-Nsiah, S., Ampadu-Boakye, T., Asare, R., Baijukya, F., Baars, E., Bekunda, M., Coyne, D., Dianda, M., Dontsop-Nguezet, P. M., Ebanyat, P., Hauser, S., Huising, J., Jalloh, A., Jassogne, L., Kamai, N., Kamara, A., … Mutsaers, H. J. W. (2017). Looking back and moving forward: 50 years of soil and soil fertility management research in sub-Saharan Africa. *International Journal of Agricultural Sustainability*, *15*(6), 613–631. https://doi.org/10.1080/14735903.2017.1393038

Wagner, F. H., Ferreira, M. P., Sanchez, A., Hirye, M. C. M., Zortea, M., Gloor, E., Phillips, O. L., de Souza Filho, C. R., Shimabukuro, Y. E., & Aragão, L. E. O. C. (2018). Individual tree crown delineation in a highly diverse tropical forest using very high resolution satellite images. *ISPRS Journal of Photogrammetry and Remote Sensing*, *145*, 362–377. https://doi.org/10.1016/j.isprsjprs.2018.09.013

Waldner, F., Lambert, M.-J., Li, W., Weiss, M., Demarez, V., Morin, D., Marais-Sicre, C., Hagolle, O., Baret, F., & Defourny, P. (2015). Land Cover and Crop Type Classification along the Season Based on Biophysical Variables Retrieved from Multi-Sensor High-Resolution Time Series. *Remote Sensing*, *7*(8), Article 8. https://doi.org/10.3390/rs70810400

World Bank. (n.d.). *Employment in agriculture (% of total employment) (modeled ILO estimate)—Mali*. World Bank Open Data. Retrieved 17 August 2025, from https://data.worldbank.org

Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, *27*(14), 3025–3033. https://doi.org/10.1080/01431160600589179

Zha, Y., Gao, J., & Ni, S. (2003). Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *International Journal of Remote Sensing*, *24*(3), 583–594. https://doi.org/10.1080/01431160304987

Zhang, C., & Li, X. (2022). Land Use and Land Cover Mapping in the Era of Big Data. *Land*, *11*(10), Article 10. https://doi.org/10.3390/land11101692

Annex

Table 18: Utilized spectral bands for Sentinel-2, PlanetScope and Pléiades with information on central wavelength and bandwidth of bands

	Utilized Bands of Sensors and Central Wavelength with Bandwidth		
EMS range:	**Sentinel-2**	**PlanetScope**	**Pléiades**
Visible	**B2**: 490nm (65nm)	**Blue**: 490nm (50nm)	**Blue**: 490-530nm
	B3: 560nm (35nm)	**Green I**: 531nm (36nm) **Green**: 565nm (36nm) **Yellow**: 610nm (20nm)	**Green**: 510-590nm
	B4: 665nm (30nm)	**Red**: 665nm (31nm)	**Red**: 620-700nm
Red Edge	**B5**: 705nm (15nm) **B6**: 740nm (15nm)	**Red Edge**: 705nm (15nm)	
Near Infrared	**B7**: 783nm (20nm) **B8**: 842nm (115nm) **B8a**: 865nm (20nm)	**NIR**: 865nm (40nm)	**NIR**: 775-915nm
Shortwave infrared	**B11**: 1610nm (90nm) **B12**: 2190nm (180nm)		

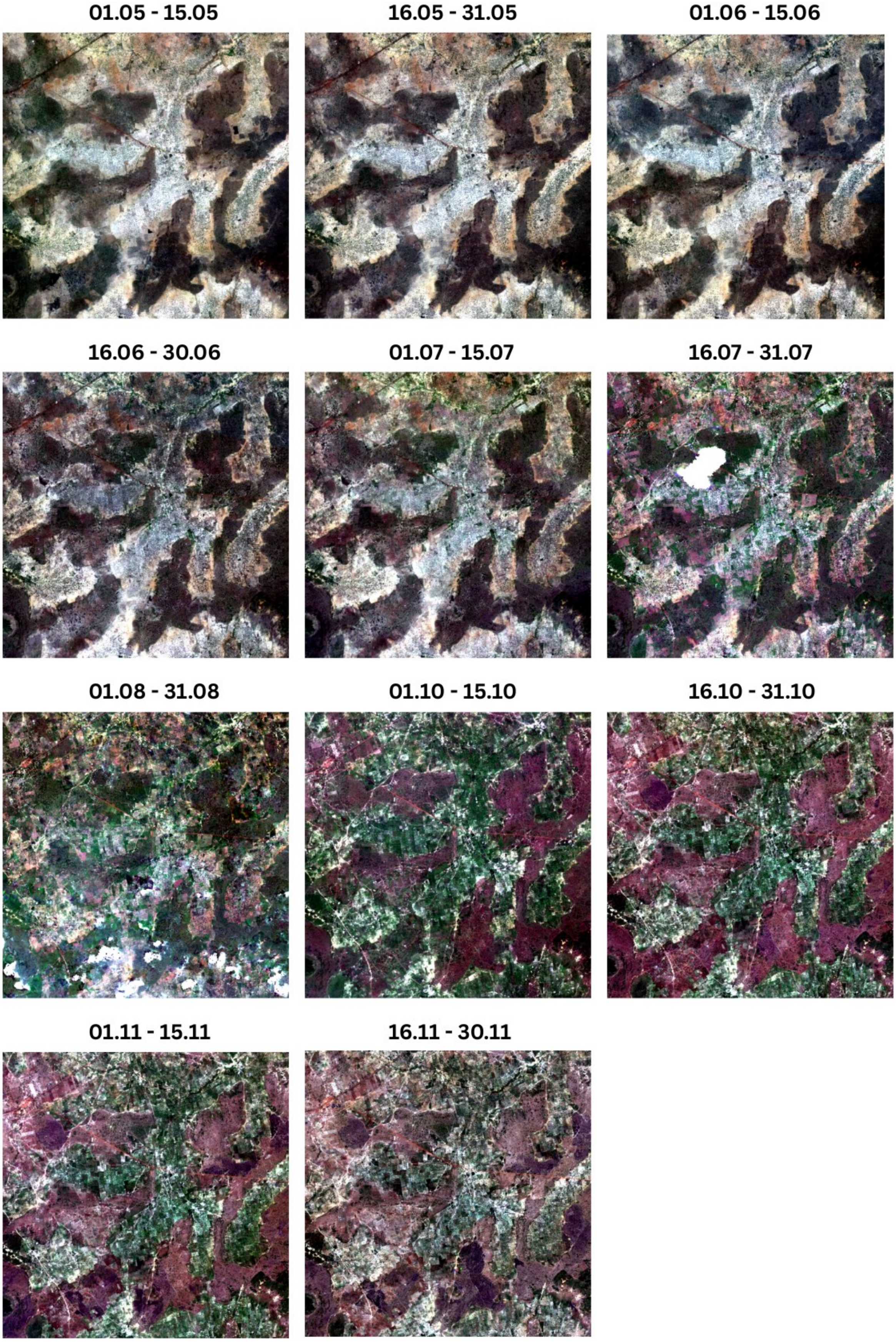

Figure 16: True colour visualization of 11 PlanetScope composite images distributed across the growing season 2024

Table 19: Calculated indices and involved bands for Sentinel-2, PlanetScope and Pléiades

Index	Abbr.	Sentinel-2 Bands	PlanetScope Bands	Pleiades Bands	Formula	Reference
Normalized Difference Vegetation Index	NDVI	B8, B4	nir, red	nir, red	$\frac{N-R}{N+R}$	(Rouse et al., 1974)
Green Normalized Difference Vegetation Index	GNDVI	B8, B3	nir, green	nir, green	$\frac{\mathrm{N-G}}{\mathrm{N+G}}$	(Gitelson et al., 1996)
Normalized Difference Red Edge Index	NDREI	B8, B5	nir, rededge	/	$\frac{\mathrm{N-RE1}}{\mathrm{N+RE1}}$	(Gitelson & Merzlyak, 1994)
Chlorophyll Index Green	CIG	B8, B3	nir, green	nir, green	$\frac{\mathrm{N}}{\mathrm{G}}-1$	(Gitelson et al., 2003)
Chlorophyll Index Red Edge	CIRE	B8, B5	nir, rededge	/	$\frac{\mathrm{N}}{\mathrm{RE1}}-1$	(Gitelson et al., 2003)
Modified Chlorophyll Absorption in Reflectance Index 2	MCARI2	B8, B4, B3	nir, red, green	nir, red, green	$\frac{1.5\,[2.5\,(\mathrm{N-R})-1.3\,(\mathrm{N-G})]}{\sqrt{(2\mathrm{N}+1)^2-(6\mathrm{N}-5\sqrt{\mathrm{R}})-0.5}}$	(Haboudane et al., 2004)
Plant Senescing Reflectance Index	PSRI	B4, B2 B6	/	/	$\frac{\mathrm{R-B}}{\mathrm{RE2}}$	(Merzlyak et al., 1999)
Enhanced Vegetation Index	EVI	B8, B4, B2,	nir, red, blue	nir, red, blue	$2.5\,\frac{\mathrm{N-R}}{\mathrm{N+6R-7.5B+1}}$	(Huete et al., 1997)
Soil Adjusted Vegetation Index	SAVI	B8, B4	nir, red	nir, red	$(1+0.5)\,\frac{\mathrm{N-R}}{\mathrm{N+R+0.5}}$	(Huete, 1988)
Modified Soil Adjusted Vegetation Index	MSAVI	B8, B4	nir, red	nir, red	$0.5\left(2N+1-\sqrt{(2N+1)^2-8(N-R)}\right)$	(Qi et al., 1994)
Normalized Difference Water Index	NDWI	B3, B8	green, nir	green, nir	$\frac{\mathrm{G-N}}{\mathrm{G+N}}$	(McFeeters, 1996)
Modified Normalized Difference Water Index	MNDWI	B3, B11	/	/	$\frac{\mathrm{G-S1}}{\mathrm{G+S1}}$	(Xu, 2006)
Normalized Difference Tillage Index	NDTI	B11, B12	/	/	$\frac{\mathrm{S1-S2}}{\mathrm{S1+S2}}$	(Van Deventer et al., 1997)
Bare Soil Index	BSI	B11, B4, B8, B2	/	/	$\frac{(\mathrm{S1+R})-(\mathrm{N+B})}{(\mathrm{S1+R})+(\mathrm{N+B})}$	(Rikimaru et al., 2002)
Normalized Difference Built-up Index	NDBI	B11, B8	/	/	$\frac{\mathrm{S1-N}}{\mathrm{S1+N}}$	(Zha et al., 2003)

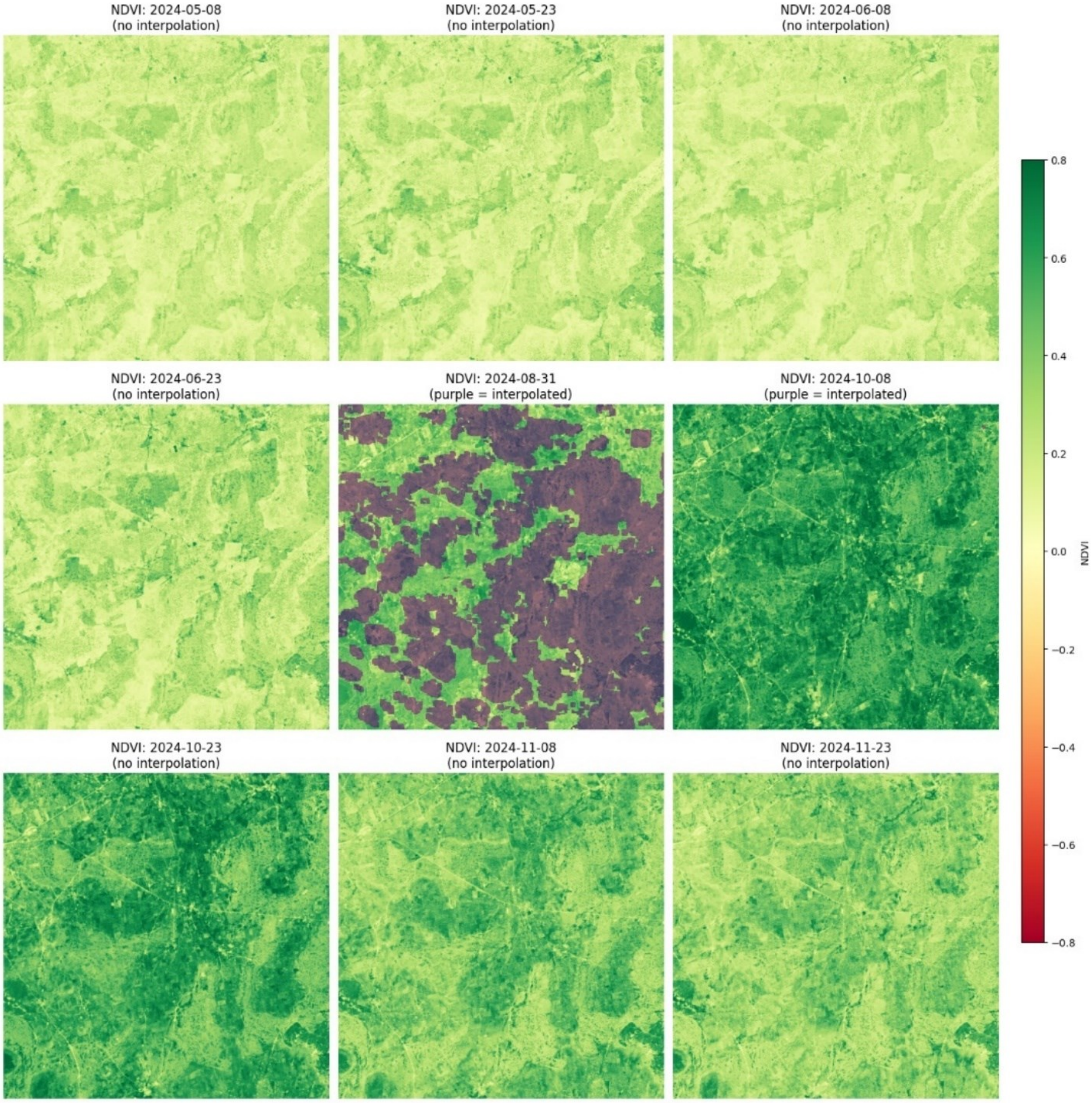

Figure 17: Gap-filled NDVI Sentinel-2 time series, with purple colouring highlighting interpolated pixels

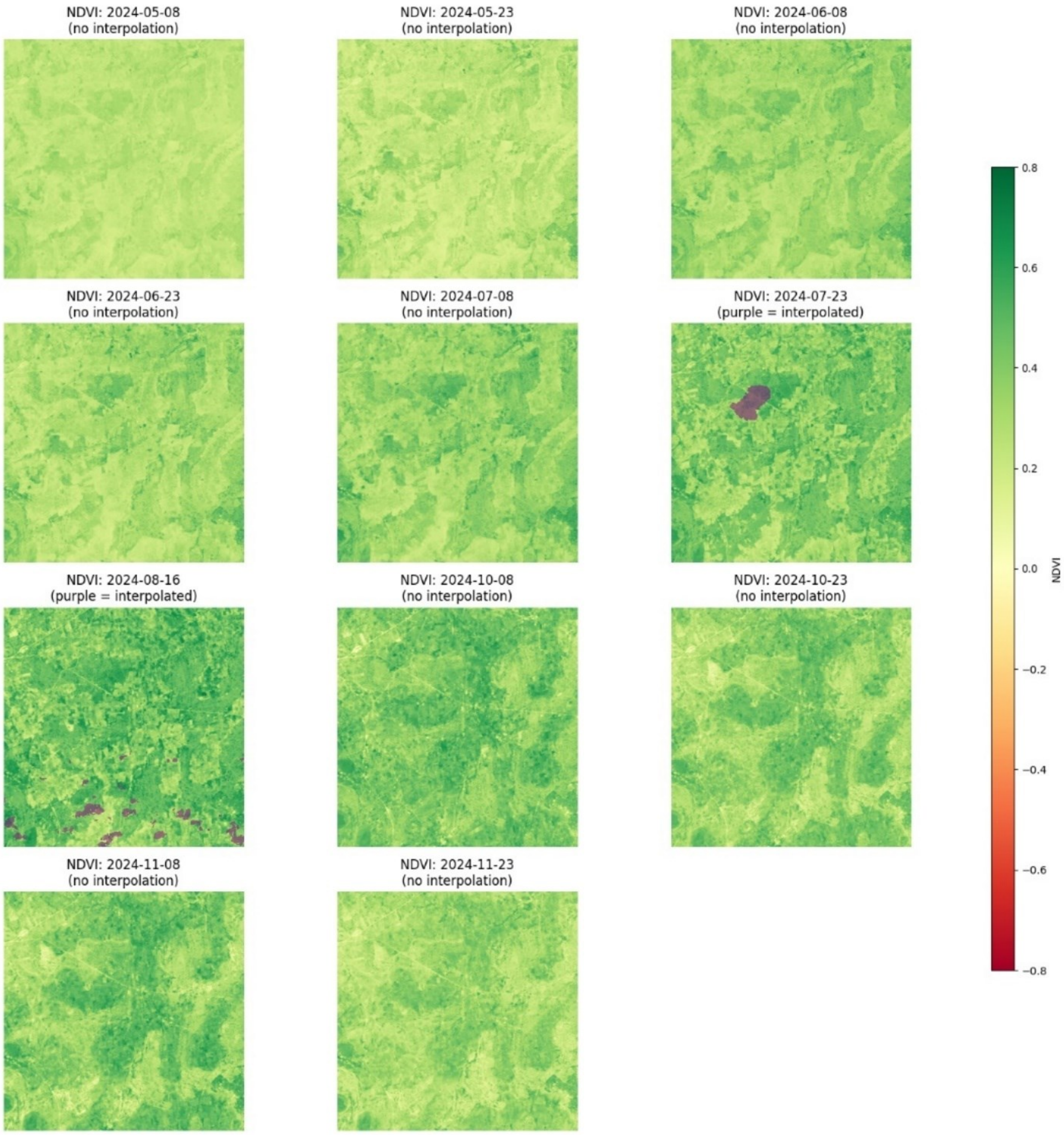

Figure 18: Gap-filled NDVI PlanetScope time series, with purple colouring highlighting interpolated pixels

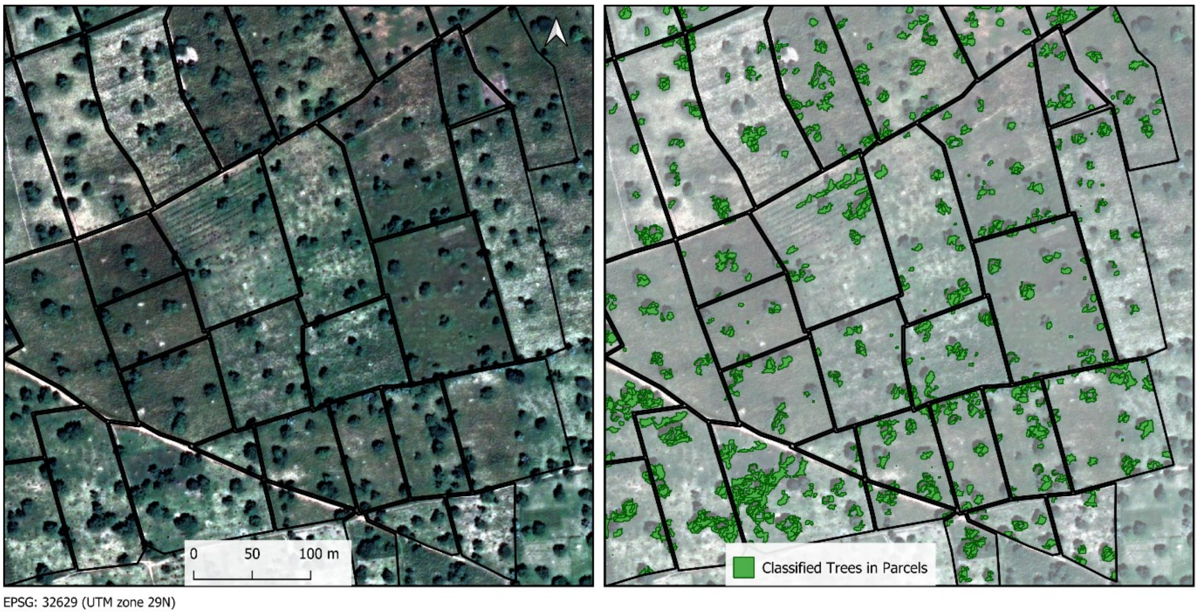

Figure 19: Example of Tree Mask generated by Abdoulaye Guindo depicting trees within parcels alongside Pléiades VHR image

Table 20: Mean NDVI of parcels on Pléiades image from 19.10.2024 for each main crop type with tree masking applied and without tree masking. Table shows that not-tree masked parcels show slightly higher NDVI than tree masked parcels

	Mean NDVI of Parcels		
Crop	**Tree-Masked**	**Not-Tree Masked**	Parcel Count
Cotton	0.567	0.577	72
Millet	0.508	0.515	89
Sesame	0.540	0.547	42
Sorghum	0.550	0.559	314
Maize	0.498	0.506	90
Peanut	0.521	0.525	16

Table 21: Count and area proportion of parcels by crop type(s) for single crop and multi-crop combinations

Crops	Count	Count Proportion	Total Area (ha)	Area Proportion
Single Crop				
Sorghum	314	38.25%	184.75	37.21%
Other Crops	97	11.81%	47.59	9.58%
Maize	91	11.08%	58.79	11.84%
Millet	88	10.72%	48.11	9.69%
Cotton	71	8.65%	55.80	11.24%
Sesame	42	5.12%	25.96	5.23%
Peanut	16	1.95%	6.43	1.29%
Bean	6	0.73%	3.21	0.65%
Okra	5	0.61%	2.36	0.48%
Rice	5	0.61%	1.83	0.37%
Cowpea	3	0.37%	1.59	0.32%
Tomato	2	0.24%	0.54	0.11%
Watermelon	1	0.12%	0.62	0.13%
Multi-crop (first listed attribute denotes primary crop)				
Sorghum-Other Crops	15	1.83%	8.32	1.68%
Maize-Sorghum	9	1.10%	7.08	1.43%
Other Crops-Other Crops	7	0.85%	5.98	1.20%
Maize-Other Crops	6	0.73%	3.59	0.72%
Peanut-Other Crops	6	0.73%	2.14	0.43%
Bean-Okra	3	0.37%	2.92	0.59%
Cotton-Sorghum	3	0.37%	4.08	0.82%
Millet-Sorghum	3	0.37%	1.69	0.34%
Sorghum-Maize	3	0.37%	4.83	0.97%
Cotton-Other Crops	2	0.24%	3.32	0.67%
Sorghum-Cotton	2	0.24%	2.67	0.54%
Sorghum-Millet	2	0.24%	1.29	0.26%
Sorghum-Peanut	2	0.24%	2.26	0.45%
Cotton-Sesame	1	0.12%	0.15	0.03%
Maize-Cowpea	1	0.12%	0.46	0.09%
Millet-Maize	1	0.12%	0.58	0.12%
Millet-Other Crops	1	0.12%	0.16	0.03%
Okra-Other Crops	1	0.12%	0.31	0.06%
Other Crops-Maize	1	0.12%	0.63	0.13%
Other Crops-Millet	1	0.12%	1.63	0.33%
Other Crops-Peanut	1	0.12%	0.31	0.06%
Other Crops-Sorghum	1	0.12%	0.16	0.03%
Peanut-Maize	1	0.12%	1.25	0.25%
Peanut-Sorghum	1	0.12%	0.68	0.14%
Sesame-Bean	1	0.12%	0.38	0.08%
Sesame-Okra	1	0.12%	0.65	0.13%
Sesame-Peanut	1	0.12%	0.69	0.14%
Sesame-Tomato	1	0.12%	0.32	0.06%
Sorghum-Sorghum	1	0.12%	0.23	0.05%
Tomato-Tomato	1	0.12%	0.24	0.05%

Table 22: Parcel area statistics by primary crop type

		Area (m^2)				
Primary Crop	Parcel Count	Mean	Median	Min	Max	SD
Sorghum	339	**6 028**	4 595	660	54 915	5 449
Other Crops	108	**5 213**	4 733	311	20 618	3 498
Maize	107	**6 534**	5 394	1 194	28 765	4 609
Millet	93	**5 434**	4 546	722	19 848	3 626
Cotton	77	**8 227**	6 654	1 121	44 381	6 417
Sesame	46	**6 085**	5 065	1 786	13 702	3 430
Peanut	24	**4 373**	3 376	1 650	12 467	2 827
Bean	9	**6 810**	3 944	1 750	16 701	5 974
Okra	6	**4 457**	3 712	1 839	8 673	2 481
Rice	5	**3 656**	2 323	739	8 551	3 142
Cowpea	3	**5 300**	6 052	2 605	7 242	2 409
Tomato	3	**2 607**	2 578	2 424	2 818	198
Watermelon	1	**6 250**	6 250	6 250	6 250	/

Table 23: Proportion of monocropped parcels in each heterogeneity class by crop type

	Heterogeneity Level					
Crop	Very Low	Low	Medium	High	Very High	Total Parcels
Sorghum	8.9%	**41.7%**	23.2%	16.9%	9.2%	314
Other Crops	7.0%	19.3%	**45.6%**	15.8%	3.5%	57
Maize	0.0%	**60.0%**	24.4%	7.8%	7.8%	90
Millet	4.5%	**40.4%**	22.5%	28.1%	4.5%	89
Cotton	11.1%	**45.8%**	18.1%	19.4%	5.6%	72
Sesame	2.4%	**47.6%**	31.0%	14.3%	4.8%	42
Peanut	0.0%	**56.2%**	25.0%	6.2%	6.2%	16
Bean	0.0%	**33.3%**	**33.3%**	0.0%	**33.3%**	6
Okra	0.0%	**60.0%**	40.0%	0.0%	0.0%	5
Rice	**75.0%**	25.0%	0.0%	0.0%	0.0%	4
Cowpea	0.0%	0.0%	**66.7%**	33.3%	0.0%	3
Tomato	**33.3%**	**33.3%**	0.0%	**33.3%**	0.0%	3
Watermelon	0.0%	**100.0%**	0.0%	0.0%	0.0%	1

Table 24: Proportion of multi-cropped parcels in each heterogeneity class

	Heterogeneity Level					
Crop Combination	Very Low	Low	Medium	High	Very High	Total Parcels
Sorghum-Other Crops	0.0%	**60.0%**	26.7%	13.3%	0.0%	15
Maize-Sorghum	0.0%	**88.9%**	11.1%	0.0%	0.0%	9
Other Crops-Other Crops	14.3%	**57.1%**	14.3%	14.3%	0.0%	7
Maize-Other Crops	0.0%	**50.0%**	33.3%	0.0%	16.7%	6
Peanut-Other Crops	0.0%	**50.0%**	33.3%	0.0%	16.7%	6
Bean-Okra	0.0%	**66.7%**	33.3%	0.0%	0.0%	3
Cotton-Sorghum	0.0%	33.3%	**66.7%**	0.0%	0.0%	3
Millet-Sorghum	0.0%	**33.3%**	**33.3%**	**33.3%**	0.0%	3
Sorghum-Maize	0.0%	**66.7%**	33.3%	0.0%	0.0%	3

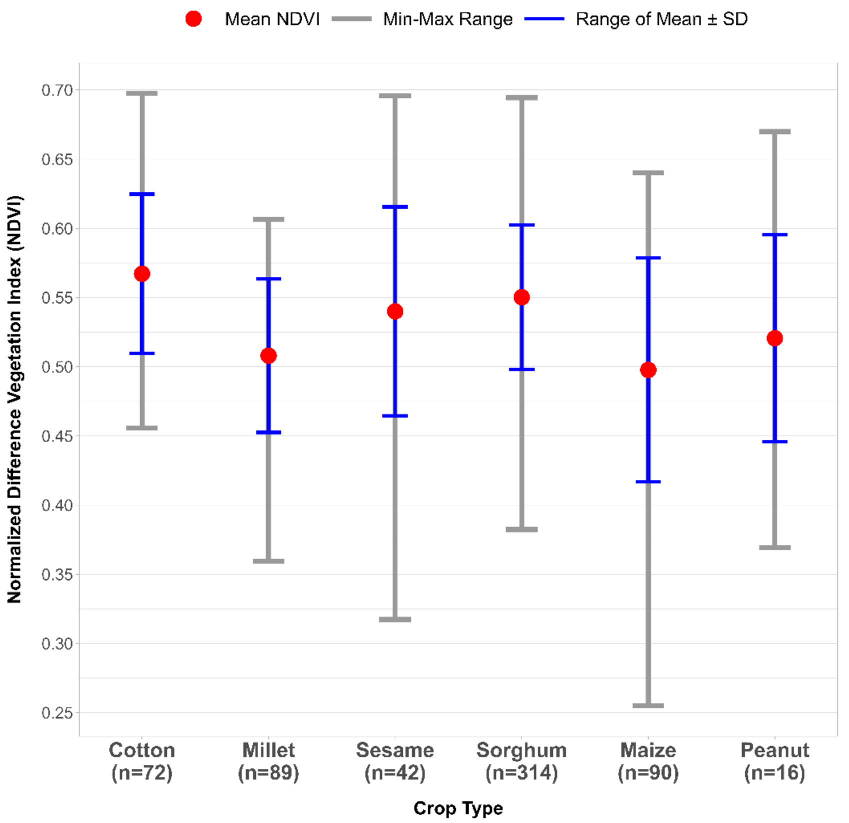

Figure 20: Pléiades- NDVI by main crop type for 19.10.2024 with mean, standard deviation and min-max range

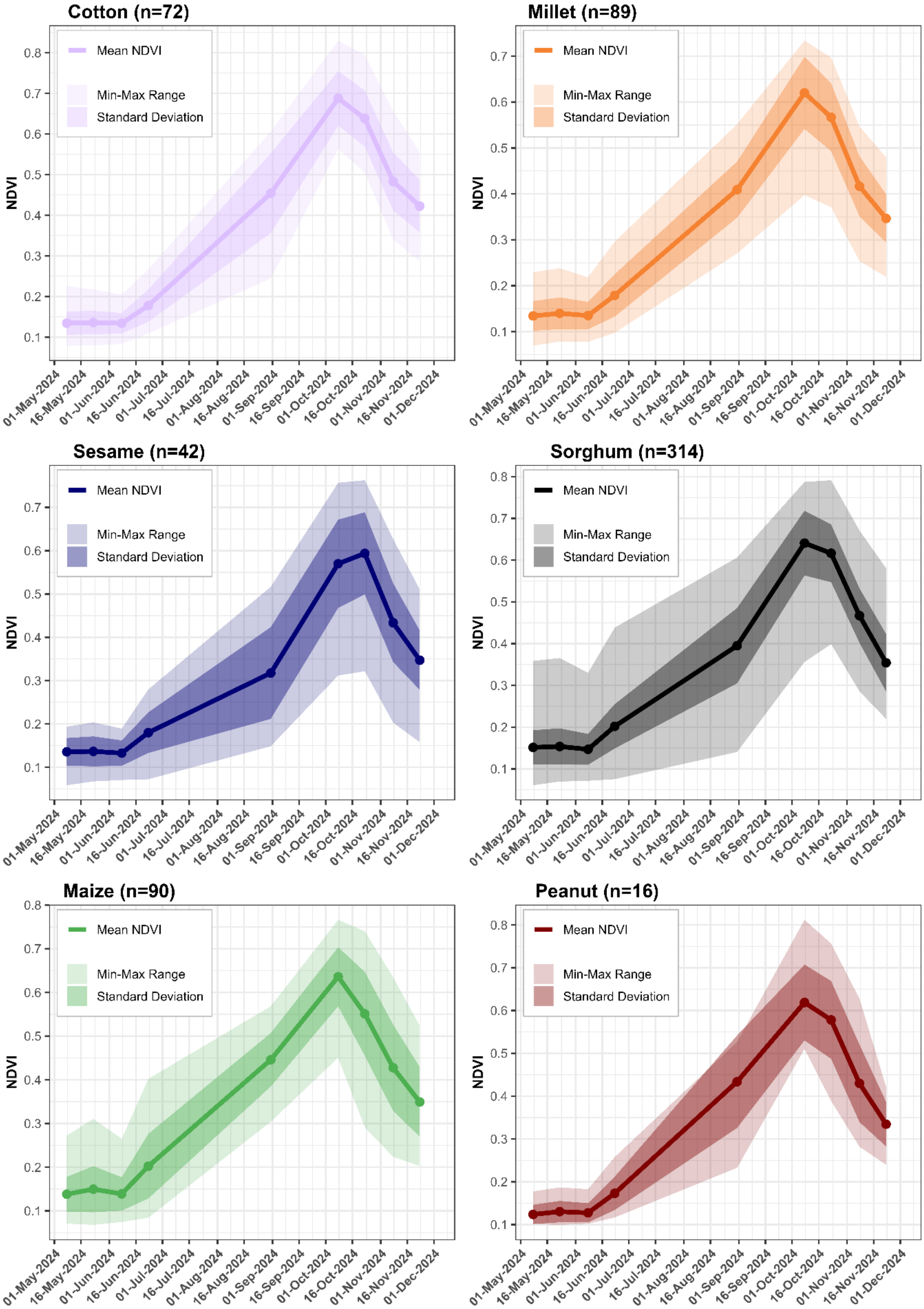

Figure 21: Sentinel-2: Temporal profile of NDVI by crop type with mean, standard deviation and min-max range of parcel values

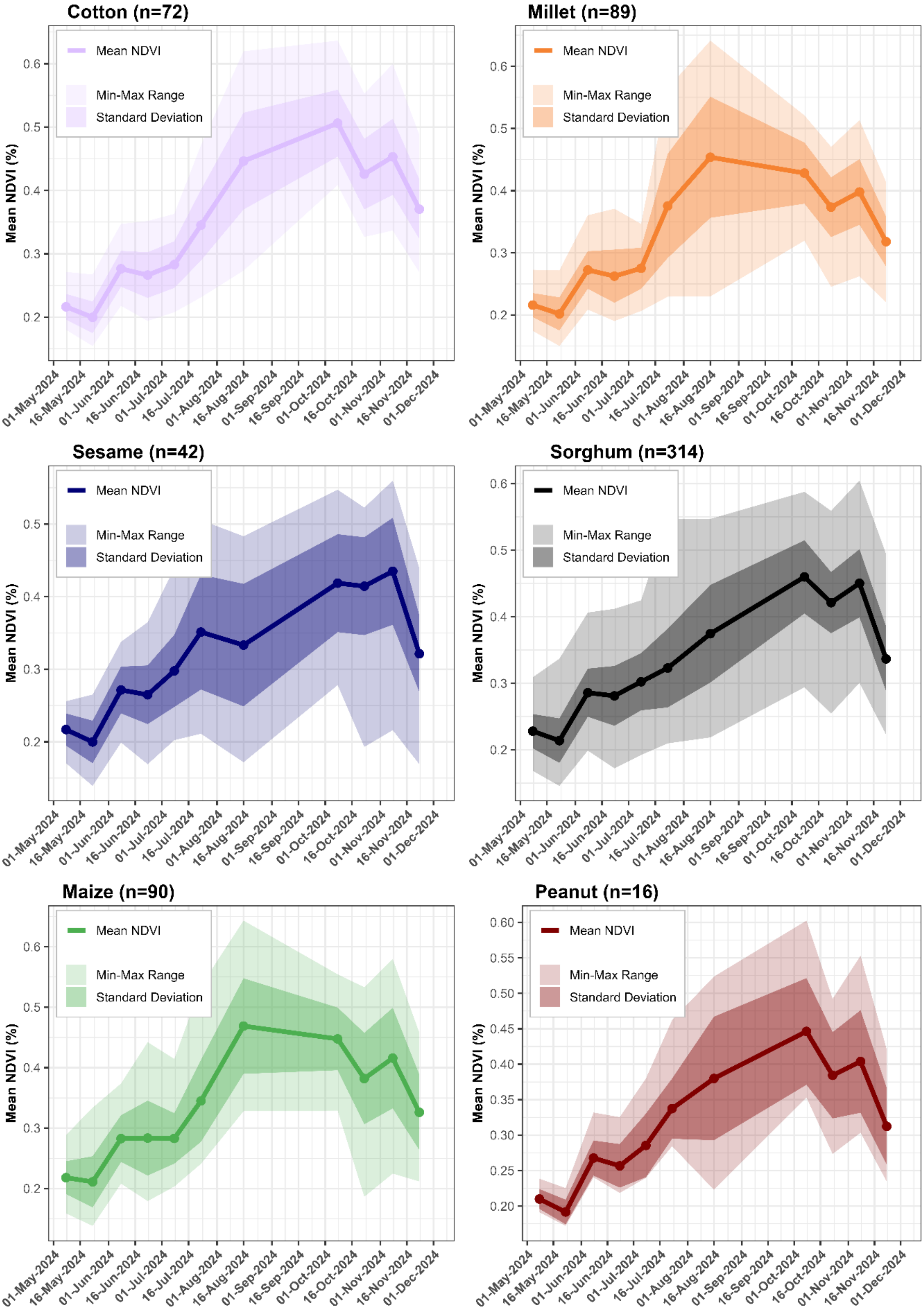

Figure 22: PlanetScope: Temporal profile of NDVI by crop type with mean, standard deviation and min-max range of parcel values

Table 25: Count of parcels in each heterogeneity class for main crop types

	Heterogeneity Level					
Crop	Very Low	Low	Medium	High	Very High	Total Monocropped Parcels
Sorghum	28	**131**	73	53	29	314
Maize	0	**54**	22	7	7	90
Millet	4	**36**	20	25	4	89
Cotton	8	**33**	13	14	4	72
Sesame	1	**20**	13	6	2	42
Peanut	0	**9**	4	1	1	16

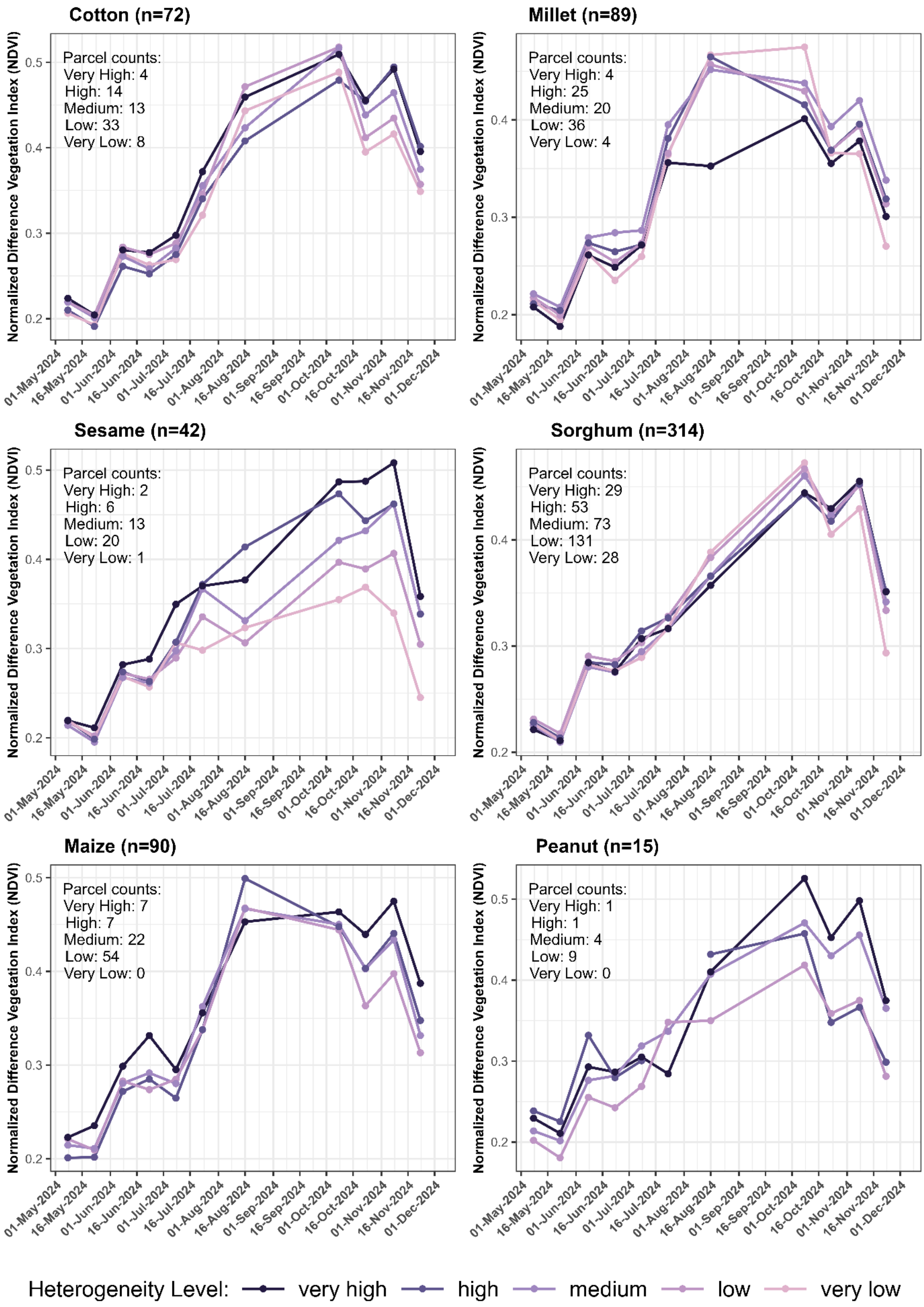

Figure 23: PlanetScope - Temporal profile of NDVI by heterogeneity level for main crop types

Table 26: Sentinel-2 data for late October - Summary stats of parcel NDVI by crop type and heterogeneity level

Sentinel-2 data for October 16 to October 23, 2024

	Cotton					Millet					Sesame				
	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max
very low	0.617	0.611	0.078	0.505	0.726	0.59	0.602	0.056	0.512	0.645	0.482	0.482	NA	0.482	0.482
low	0.639	0.62	0.078	0.506	0.794	0.56	0.552	0.074	0.417	0.697	0.558	0.582	0.1	0.322	0.688
medium	0.65	0.639	0.046	0.587	0.733	0.591	0.598	0.059	0.451	0.681	0.606	0.589	0.075	0.503	0.763
high	0.631	0.628	0.053	0.526	0.749	0.559	0.577	0.078	0.397	0.688	0.669	0.662	0.049	0.598	0.735
very high	0.659	0.631	0.094	0.578	0.794	0.543	0.567	0.124	0.371	0.667	0.704	0.704	0.023	0.688	0.72

	Sorghum					Maize					Peanut				
	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max
very low	0.606	0.596	0.059	0.513	0.713	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
low	0.621	0.627	0.062	0.421	0.758	0.524	0.539	0.099	0.291	0.702	0.535	0.529	0.076	0.389	0.683
medium	0.618	0.623	0.068	0.399	0.738	0.58	0.595	0.077	0.473	0.74	0.649	0.636	0.09	0.569	0.757
high	0.605	0.611	0.083	0.409	0.791	0.582	0.577	0.094	0.441	0.725	0.524	0.524	NA	0.524	0.524
very high	0.622	0.636	0.081	0.424	0.778	0.636	0.646	0.043	0.554	0.68	0.677	0.677	NA	0.677	0.677

Metrics (Mean, Median, Standard Deviation, Min, Max) represent summary statistics of the parcel-level Mean of NDVI.

Table 27: PlanetScope data for late October - Summary stats of parcel NDVI by crop type and heterogeneity level

PlanetScope data for October 16 to October 31, 2024

	Cotton					Millet					Sesame				
	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max
very low	0.395	0.378	0.052	0.328	0.477	0.366	0.368	0.036	0.321	0.408	0.369	0.369	NA	0.369	0.369
low	0.412	0.405	0.057	0.327	0.545	0.369	0.362	0.048	0.276	0.466	0.389	0.403	0.075	0.193	0.516
medium	0.438	0.433	0.036	0.394	0.5	0.393	0.398	0.041	0.303	0.47	0.432	0.424	0.05	0.359	0.523
high	0.455	0.448	0.051	0.332	0.553	0.369	0.383	0.048	0.245	0.437	0.443	0.453	0.052	0.341	0.483
very high	0.456	0.449	0.069	0.377	0.546	0.355	0.368	0.078	0.25	0.435	0.488	0.488	0.029	0.467	0.508

	Sorghum					Maize					Peanut				
	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max
very low	0.405	0.397	0.037	0.36	0.481	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
low	0.423	0.429	0.042	0.254	0.517	0.363	0.357	0.076	0.187	0.525	0.359	0.35	0.059	0.274	0.485
medium	0.424	0.434	0.044	0.281	0.517	0.403	0.412	0.073	0.27	0.533	0.43	0.428	0.049	0.372	0.492
high	0.418	0.415	0.054	0.283	0.559	0.403	0.381	0.07	0.322	0.505	0.348	0.348	NA	0.348	0.348
very high	0.429	0.428	0.055	0.322	0.547	0.439	0.431	0.039	0.392	0.49	0.453	0.453	NA	0.453	0.453

Metrics (Mean, Median, Standard Deviation, Min, Max) represent summary statistics of the parcel-level Mean of NDVI.

Table 28: Pléiades data for late October - Summary stats of parcel NDVI by crop type and heterogeneity level

Pléiades data for October 19, 2024

	Cotton					Millet					Sesame				
	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max
very low	0.536	0.526	0.062	0.456	0.627	0.535	0.541	0.037	0.487	0.57	0.476	0.476	NA	0.476	0.476
low	0.562	0.542	0.064	0.462	0.696	0.504	0.496	0.056	0.393	0.606	0.512	0.527	0.082	0.317	0.637
medium	0.586	0.582	0.03	0.546	0.638	0.525	0.524	0.042	0.433	0.591	0.55	0.529	0.061	0.465	0.696
high	0.574	0.574	0.051	0.466	0.672	0.501	0.509	0.06	0.374	0.597	0.598	0.59	0.038	0.546	0.653
very high	0.582	0.56	0.081	0.509	0.698	0.484	0.496	0.093	0.359	0.584	0.621	0.621	0.004	0.618	0.623

	Sorghum					Maize					Peanut				
	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max
very low	0.542	0.526	0.04	0.489	0.625	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
low	0.552	0.559	0.046	0.422	0.654	0.475	0.482	0.085	0.255	0.64	0.483	0.476	0.062	0.369	0.61
medium	0.554	0.561	0.051	0.39	0.644	0.527	0.536	0.063	0.415	0.637	0.577	0.564	0.073	0.511	0.67
high	0.54	0.543	0.065	0.395	0.694	0.525	0.524	0.07	0.422	0.637	0.492	0.492	NA	0.492	0.492
very high	0.559	0.559	0.063	0.382	0.684	0.557	0.565	0.038	0.482	0.599	0.58	0.58	NA	0.58	0.58

Metrics (Mean, Median, Standard Deviation, Min, Max) represent summary statistics of the parcel-level Mean of NDVI.

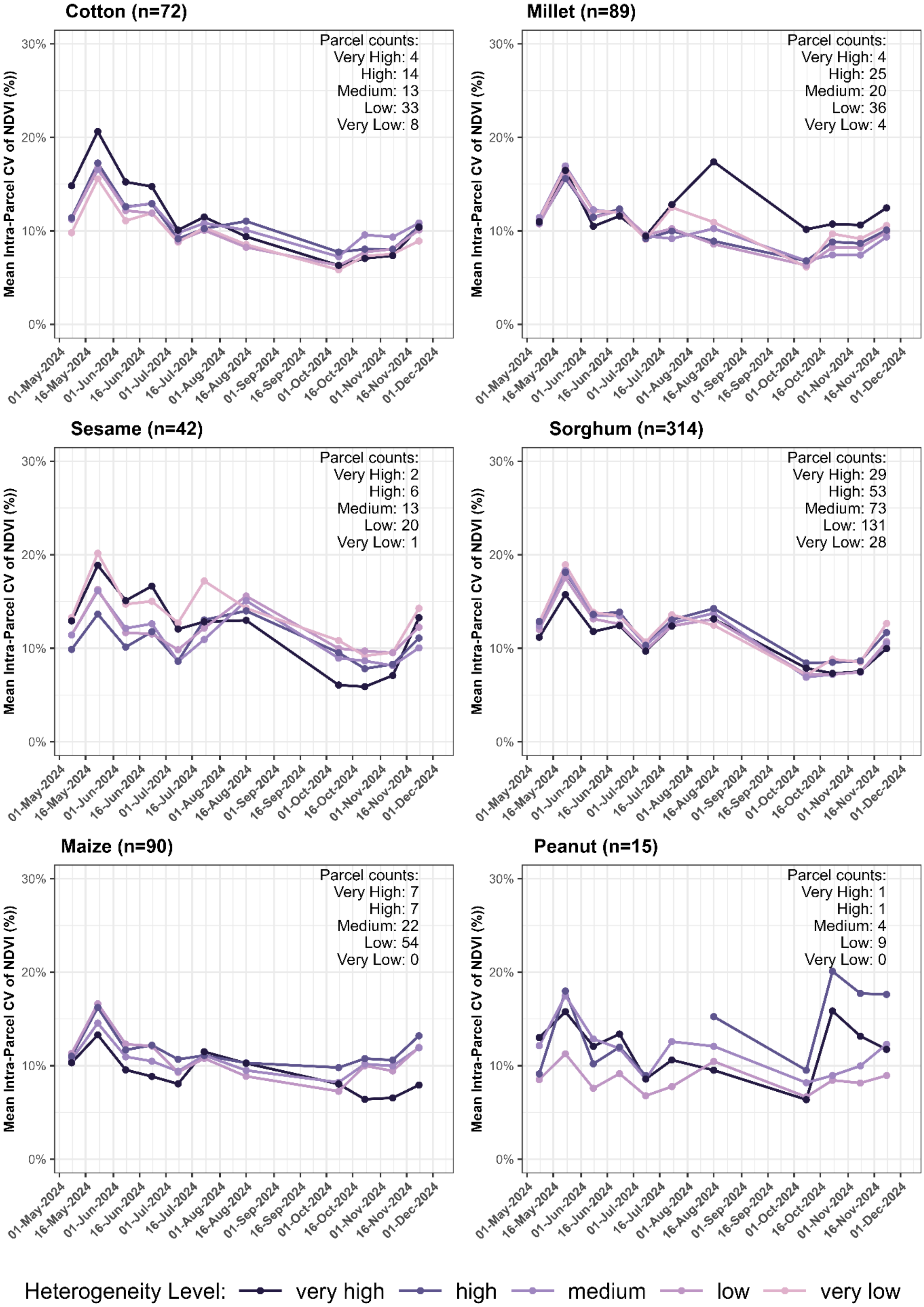

Figure 24: PlanetScope - Temporal profile of parcel coefficient of variation of NDVI by crop type and heterogeneity level

Table 29: Sentinel-2 data from late October - Summary statistics of parcel coefficient of variation of NDVI by crop type and heterogeneity level

Sentinel-2 data for October 16 to October 31, 2024

	Cotton					Millet					Sesame				
	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max
very low	**6.7%**	6.0%	2.4%	4.4%	11.9%	**8.0%**	7.6%	2.8%	5.5%	11.3%	**11.6%**	11.6%	NA	11.6%	11.6%
low	**6.7%**	6.5%	2.4%	3.0%	11.8%	**7.4%**	6.3%	3.4%	2.3%	16.6%	**8.4%**	7.7%	3.0%	5.0%	15.9%
medium	**7.8%**	6.7%	5.1%	3.8%	24.4%	**6.8%**	6.2%	2.9%	2.5%	13.9%	**9.1%**	9.4%	4.0%	3.5%	19.4%
high	**7.1%**	6.5%	3.0%	2.7%	13.7%	**8.5%**	8.1%	3.1%	4.4%	15.2%	**6.4%**	6.8%	2.4%	2.7%	9.4%
very high	**6.0%**	6.7%	2.4%	2.7%	8.0%	**9.2%**	9.3%	2.6%	6.1%	11.9%	**4.2%**	4.2%	1.6%	3.0%	5.3%

	Sorghum					Maize					Peanut				
	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max
very low	**8.5%**	7.4%	4.3%	3.3%	22.2%	**NA**	NA	NA	NA	NA	**NA**	NA	NA	NA	NA
low	**6.8%**	6.0%	3.6%	0.8%	29.7%	**10.0%**	9.6%	4.8%	3.5%	29.0%	**8.2%**	7.9%	3.0%	3.8%	12.3%
medium	**6.6%**	5.7%	3.2%	1.9%	17.6%	**9.0%**	8.5%	3.5%	3.2%	14.4%	**8.3%**	8.9%	3.3%	3.8%	11.7%
high	**8.1%**	7.4%	3.9%	2.5%	20.3%	**11.5%**	10.1%	6.1%	4.0%	22.3%	**17.1%**	17.1%	NA	17.1%	17.1%
very high	**6.3%**	6.3%	3.1%	1.5%	13.2%	**6.5%**	6.1%	1.7%	4.7%	9.5%	**11.1%**	11.1%	NA	11.1%	11.1%

Metrics (Mean, Median, Standard Deviation, Min, Max) represent summary statistics of the parcel-level Coefficient of Variation of NDVI.

Table 30: PlanetScope data from late October - Summary statistics of parcel coefficient of variation of NDVI by crop type and heterogeneity level

PlanetScope data for October 16 to October 31, 2024

	Cotton					Millet					Sesame				
	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max
very low	**7.3%**	7.0%	2.1%	4.5%	10.7%	**9.7%**	9.0%	4.2%	6.1%	14.5%	**9.2%**	9.2%	NA	9.2%	9.2%
low	**7.7%**	7.5%	2.1%	4.0%	11.9%	**8.2%**	7.6%	3.0%	3.7%	16.3%	**9.7%**	8.7%	3.9%	3.7%	16.5%
medium	**9.6%**	8.2%	4.7%	5.8%	23.9%	**7.4%**	7.3%	2.8%	3.6%	13.8%	**8.7%**	8.1%	4.0%	4.0%	18.8%
high	**8.0%**	7.0%	3.7%	4.0%	16.2%	**8.8%**	7.9%	3.5%	4.7%	19.7%	**7.8%**	7.6%	2.8%	3.7%	12.1%
very high	**7.1%**	6.9%	2.2%	4.6%	9.9%	**10.7%**	9.5%	4.6%	6.6%	17.3%	**5.9%**	5.9%	2.0%	4.5%	7.3%

	Sorghum					Maize					Peanut				
	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max
very low	**8.8%**	7.9%	4.0%	4.0%	16.3%	**NA**	NA	NA	NA	NA	**NA**	NA	NA	NA	NA
low	**7.2%**	6.3%	3.7%	1.8%	31.4%	**10.0%**	9.6%	4.0%	3.9%	22.4%	**8.5%**	8.5%	2.5%	4.8%	11.9%
medium	**7.2%**	6.3%	3.3%	2.8%	20.9%	**10.2%**	9.2%	6.4%	4.4%	34.4%	**9.0%**	8.8%	3.9%	4.9%	13.3%
high	**8.5%**	7.7%	3.9%	3.6%	24.1%	**10.8%**	8.1%	5.8%	4.2%	19.3%	**20.1%**	20.1%	NA	20.1%	20.1%
very high	**7.3%**	7.5%	3.0%	2.0%	13.5%	**6.4%**	6.2%	1.2%	5.4%	8.7%	**15.9%**	15.9%	NA	15.9%	15.9%

Metrics (Mean, Median, Standard Deviation, Min, Max) represent summary statistics of the parcel-level Coefficient of Variation of NDVI.

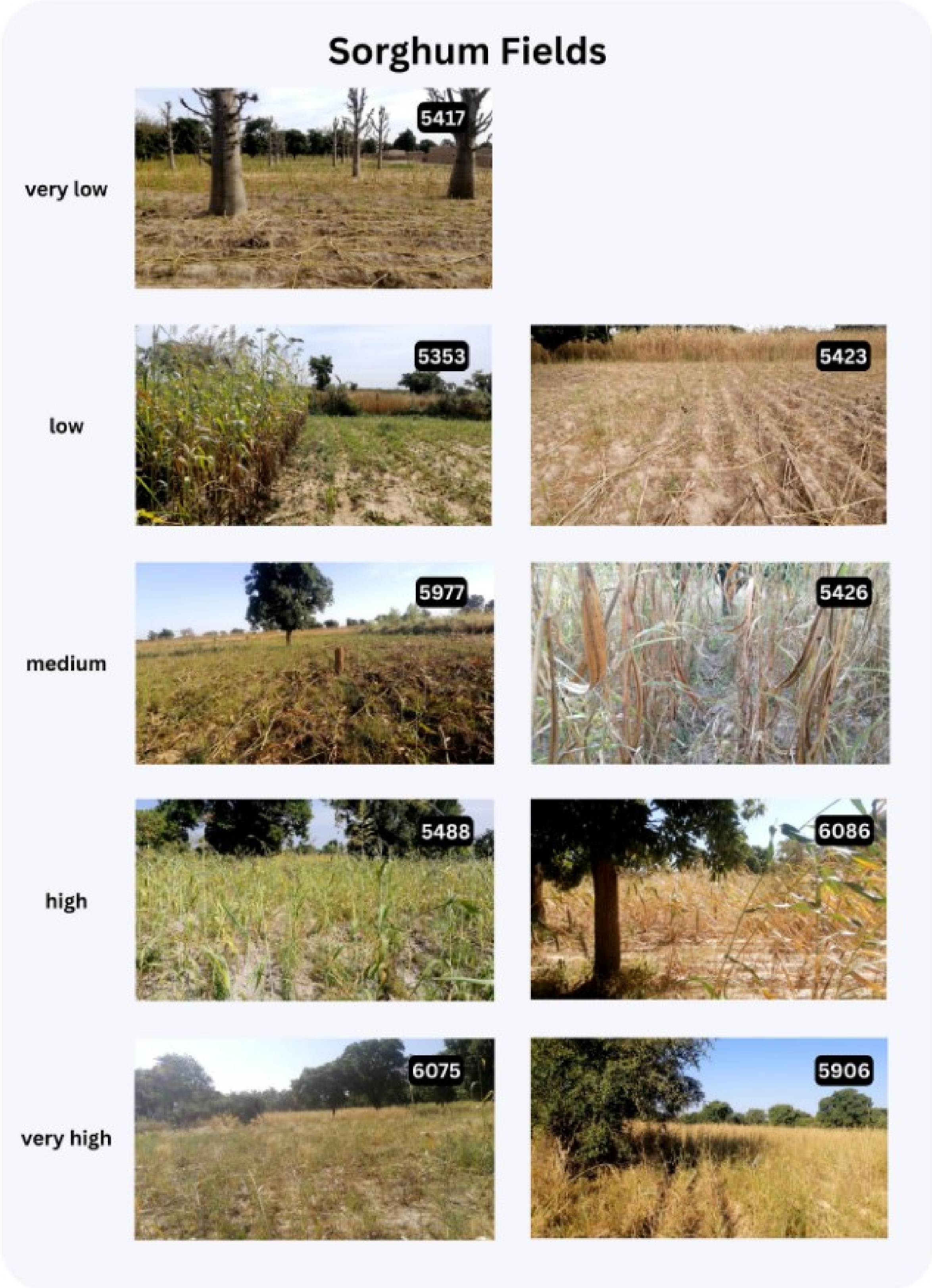

Figure 25: Photographs of sorghum parcels of different assessed levels of heterogeneity

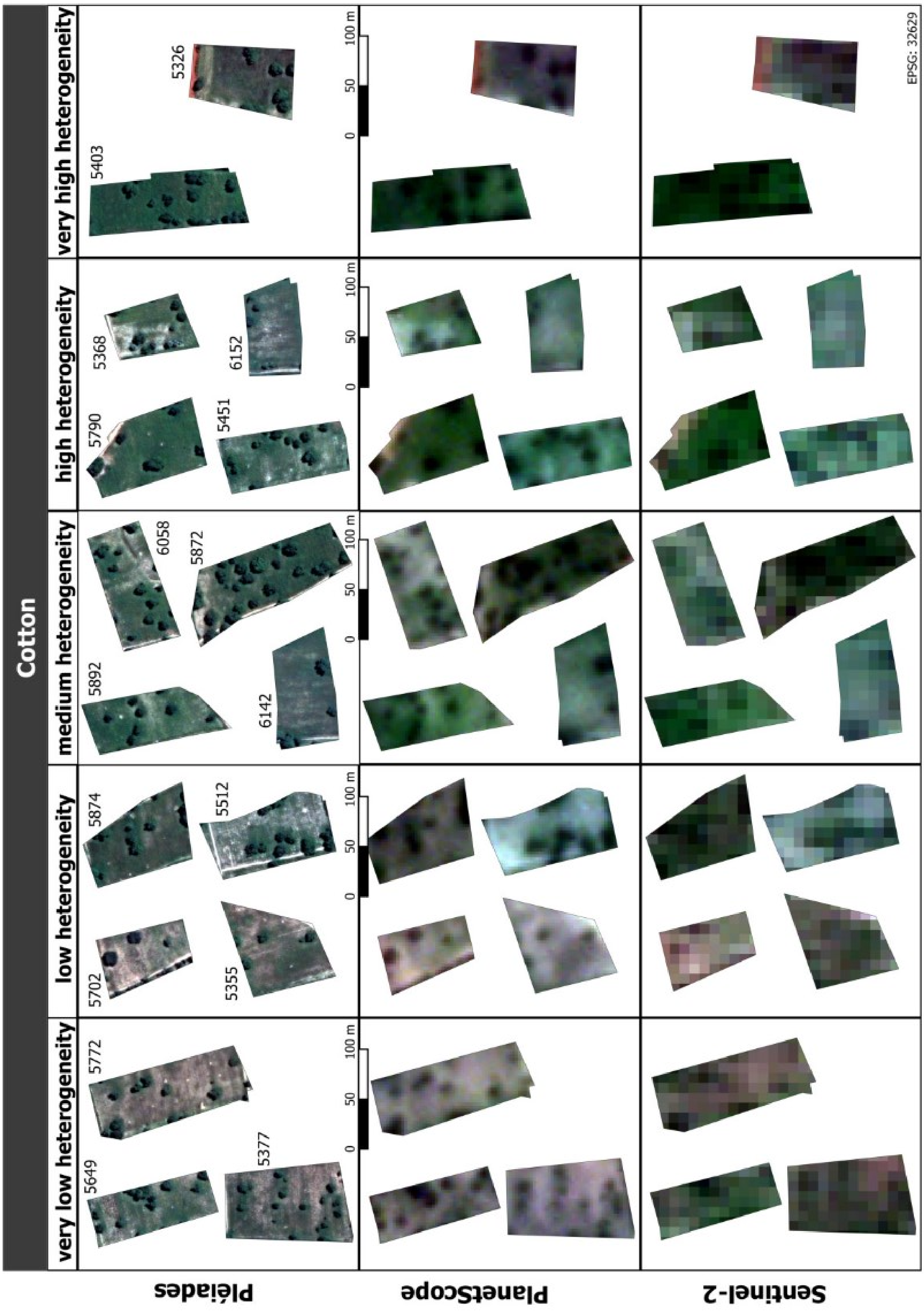

Figure 26: Cotton parcels of different assessed heterogeneity levels on Pléiades, PlanetScope and Sentinel-2 imagery from late October

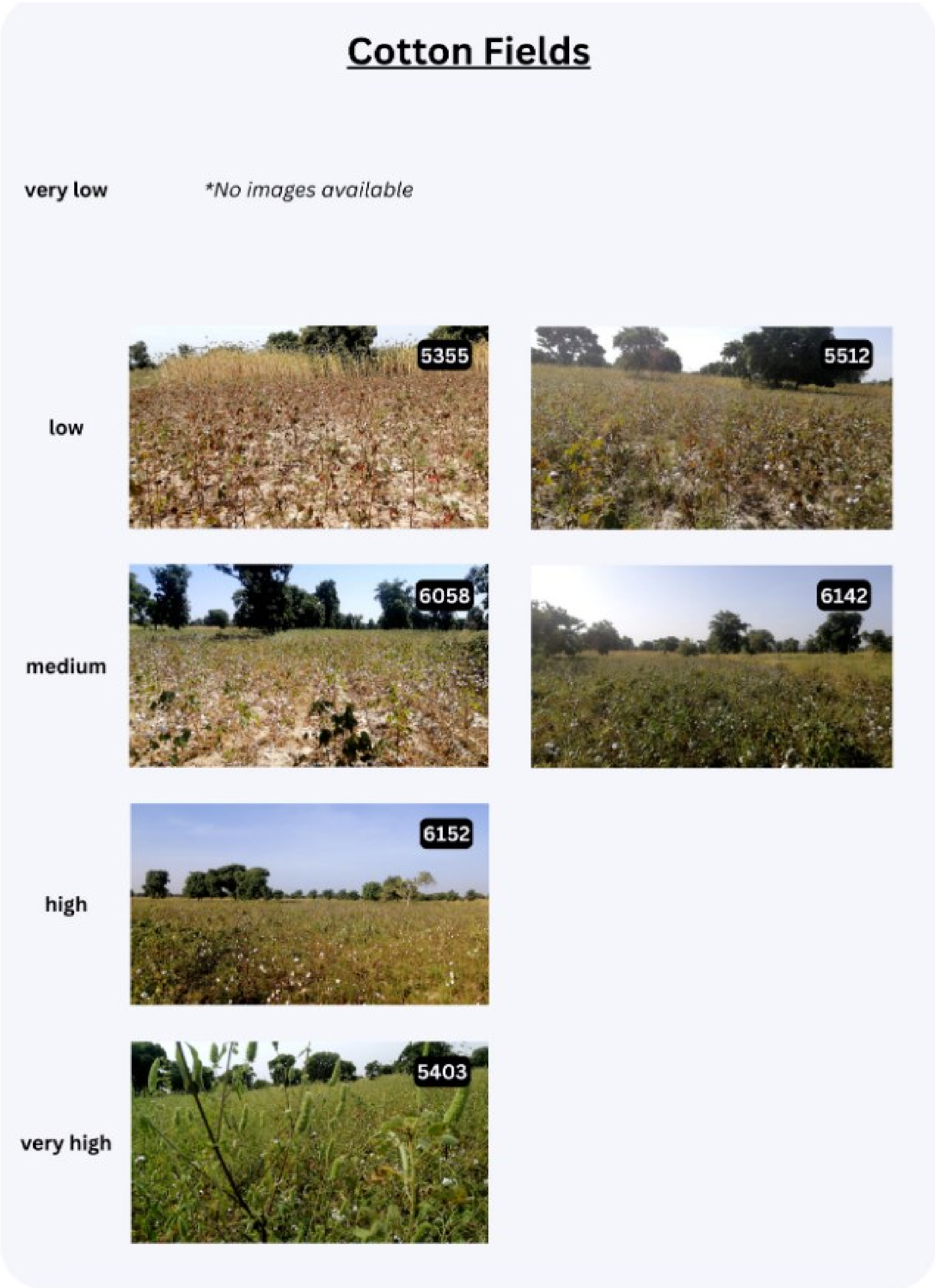

Figure 27: Photographs of cotton parcels of different assessed levels of heterogeneity

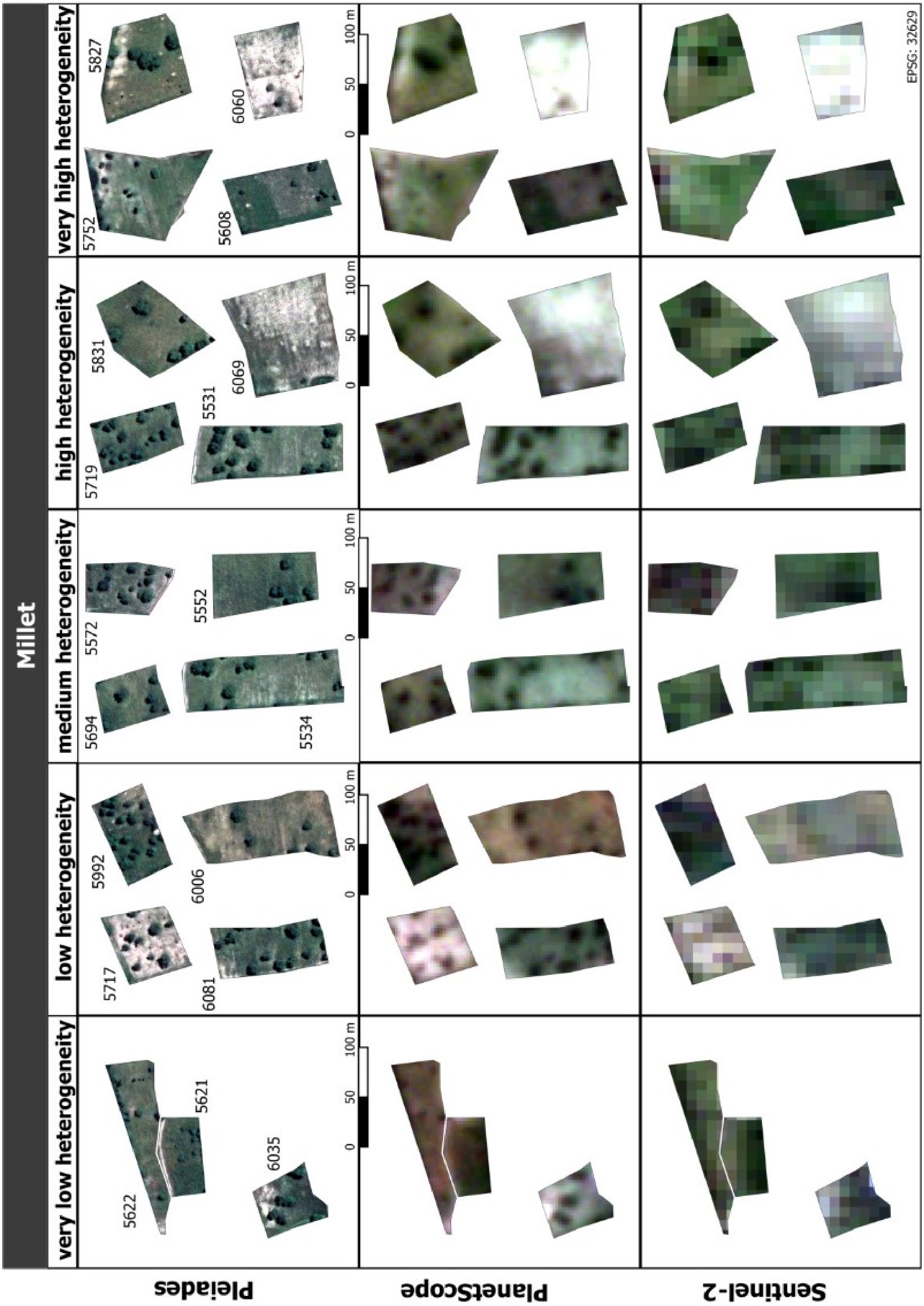

Figure 28: Millet Parcels of different assessed heterogeneity levels on Pléiades, PlanetScope and Sentinel-2 imagery from late October

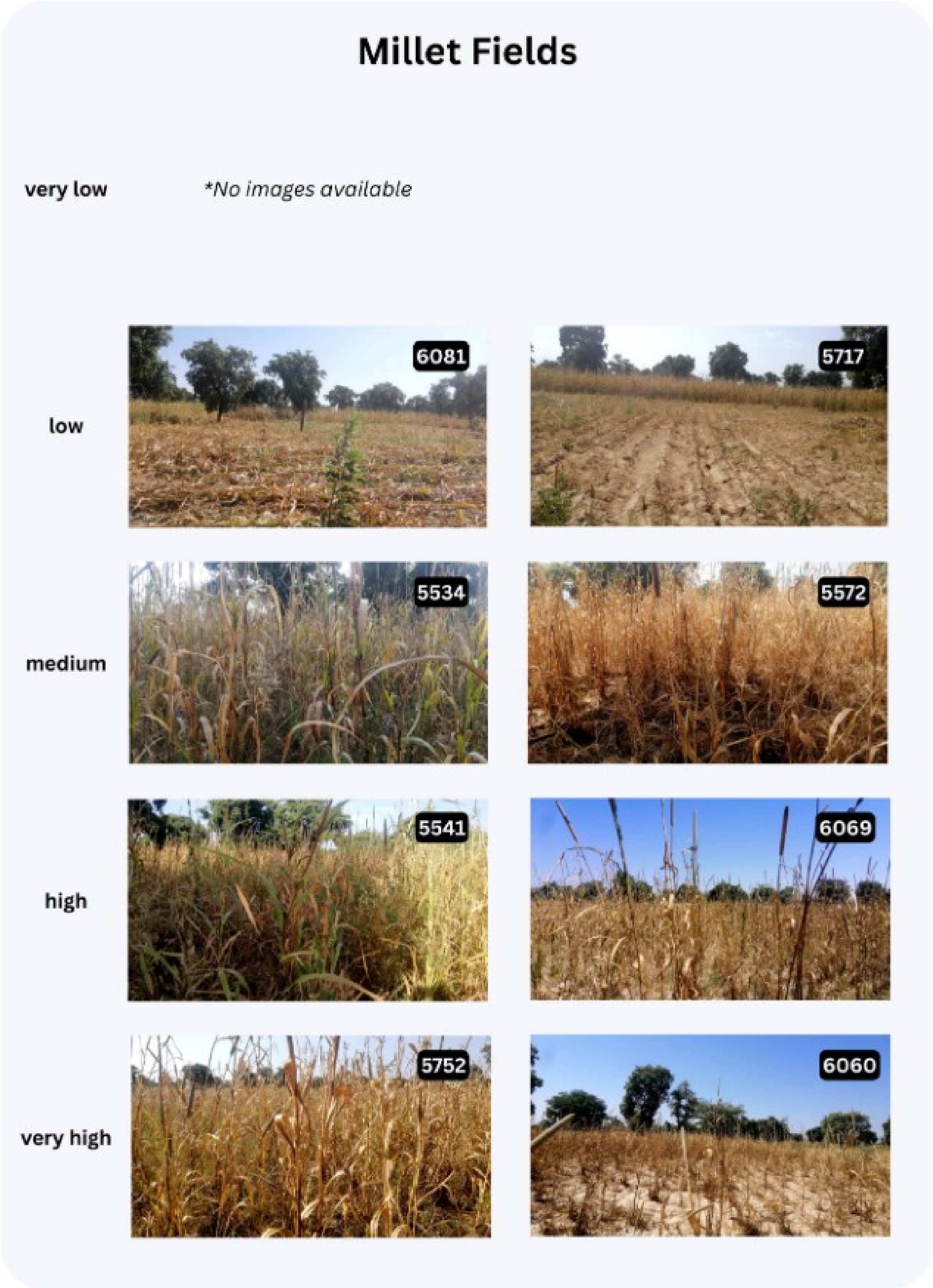

Figure 29: Photographs of millet parcels of different assessed levels of heterogeneity

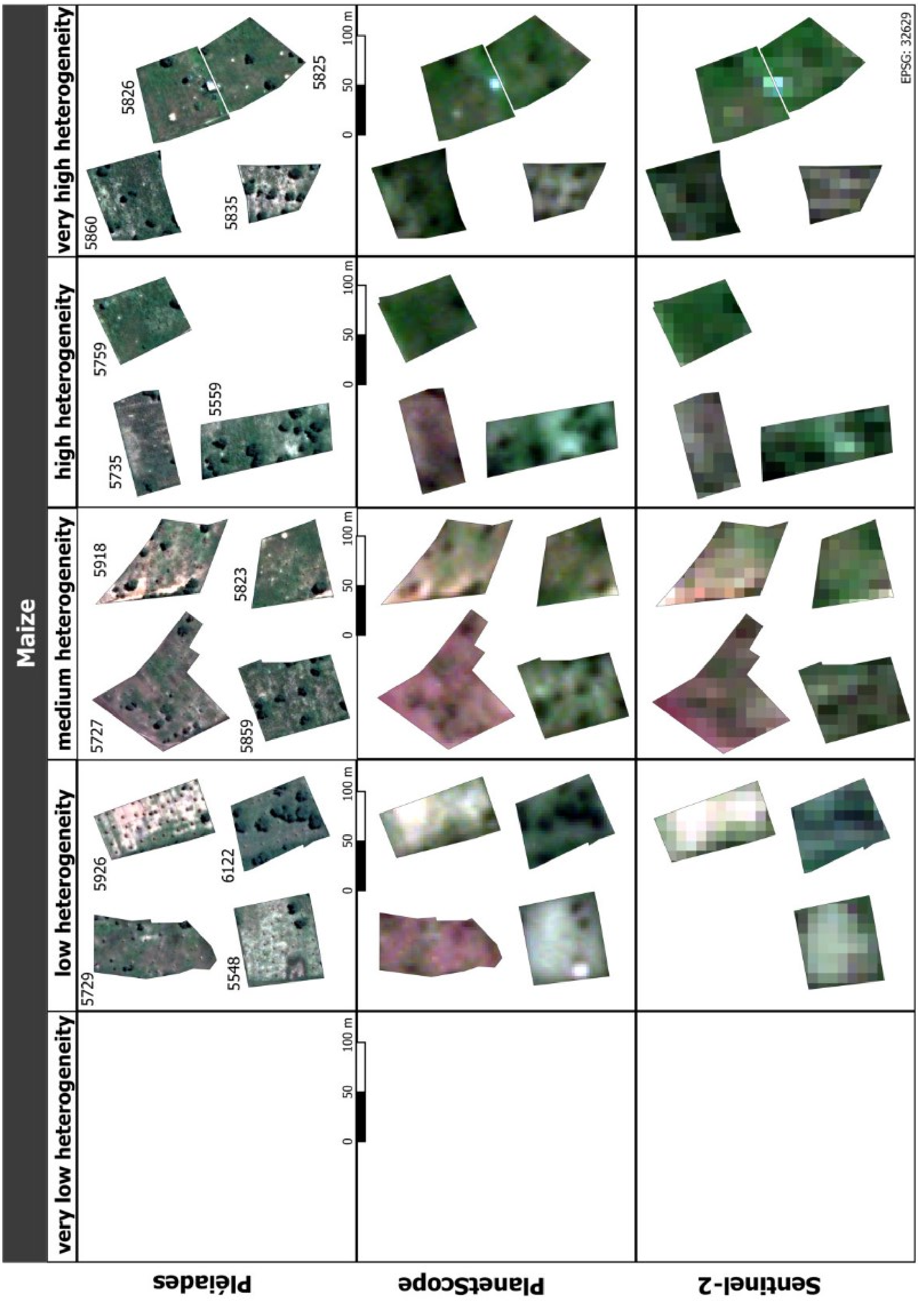

Figure 30: Maize parcels of different assessed heterogeneity levels on Pléiades, PlanetScope and Sentinel-2 imagery from late October

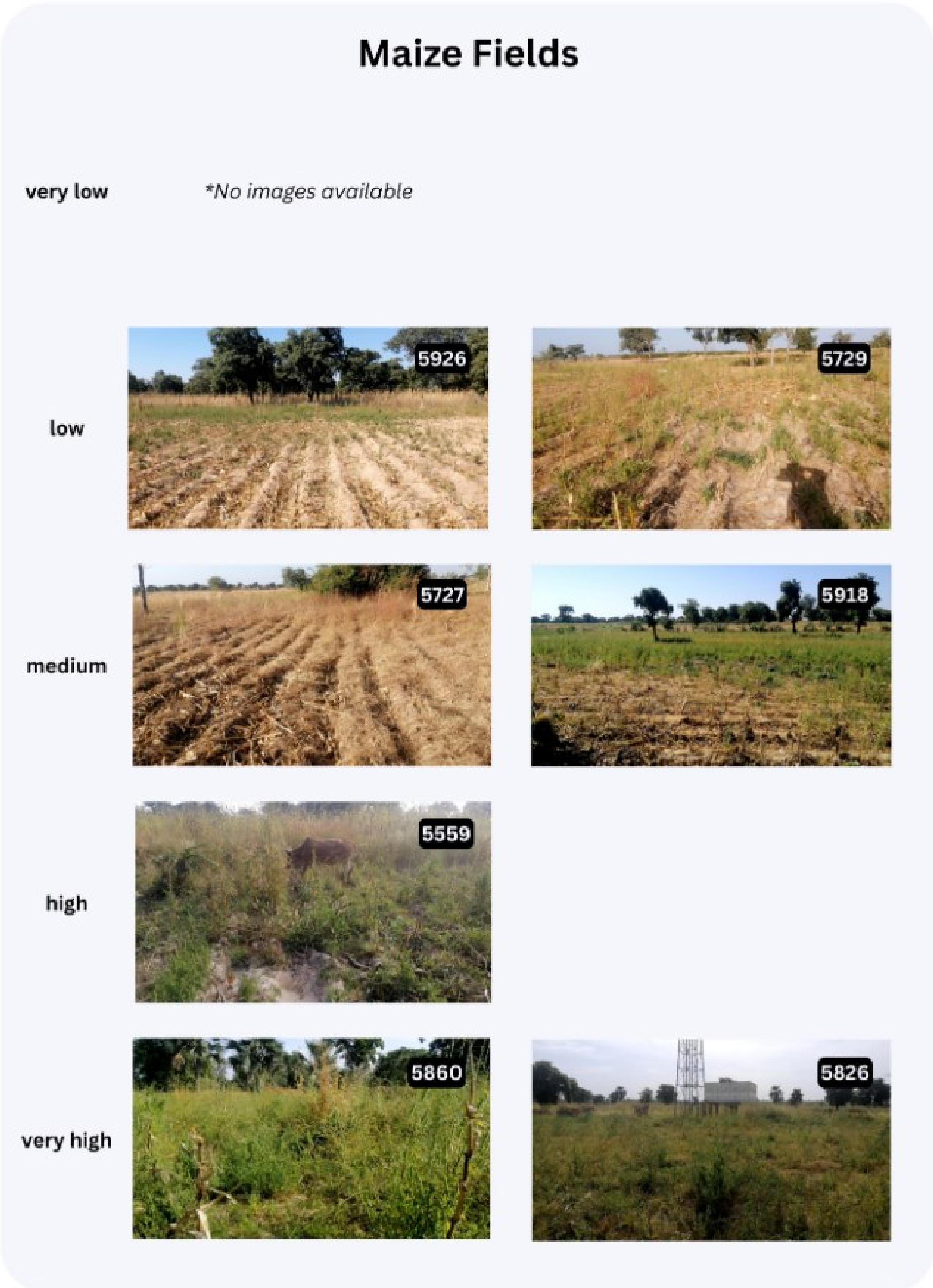

Figure 31: Photographs of maize parcels of different assessed levels of heterogeneity

Crop Type Classification

Table 31: Number of sample polygons and pixel count of training data for each class for Sentinel-2, PlanetScope and Pléiades classifications

		Pixel Count of Training Data Set		
Class	**Sample Polygons**	**Sentinel-2**	**PlanetScope**	**Pléiades**
Sorghum	313	6467	113 798	3 591 069
Maize	90	2274	39 027	1 256 794
Millet	88	1958	34 084	1 104 279
Cotton	71	2102	35 933	1 103 198
Sesame	42	921	16 689	523 323
Peanut	16	178	3518	113 207
Bean	6	55	1235	40 029
Okra	5	106	2006	61 733
Rice	3	26	385	12 457
Cowpea	3	52	917	26 143
Tomato	2	7	221	7518
Asphalt Road	45	155	621	10 888
Built-Up	44	87	299	5 087
Clay Bare Soil	49	115	535	10 653
Forest Edge	47	134	667	15 045
Grassland	42	127	944	22 822
Lateritic Bare Soil	45	198	992	22 747
Lateritic Rock road	45	139	608	12 040
Water	16	38	153	2825
Woody Vegetation	46	213	1307	32 954

Sentinel-2 Classifications

Table 32: Indices used in Sentinel-2 classification scenarios

Index	Abbreviation
Normalized Difference Vegetation Index	NDVI
Green Normalized Difference Vegetation Index	GNDVI
Normalized Difference Red Edge Index	NDREI
Chlorophyll Index Green	CIG
Chlorophyll Index Red Edge	CIRE
Modified Chlorophyll Absorption in Reflectance Index 2	MCARI2
Plant Senescence Reflectance Index	PSRI
Enhanced Vegetation Index	EVI
Soil Adjusted Vegetation Index	SAVI
Modified Soil Adjusted Vegetation Index	MSAVI
Normalized Difference Tillage Index	NDTI
Normalized Difference Water Index	NDWI
Modified Normalized Difference Water Index	MNDWI
Bare Soil Index	BSI
Normalized Difference Built-Up Index	NDBI

Table 33: Confusion matrix for Sentinel-2 Scenario 1 baseline classification (Chapter 7.5.2.1.)

Confusion Matrix with Recall & Precision

Rows=Actual, Columns=Predicted

	Sorghum	Maize	Millet	Cotton	Sesame	Peanut	asphalt road	built-up	clay bare soil	forest-baresoil interface	grassland	lateritic bare soil	lateritic rock road	water	woody vegetation	Other crops	Actual Total	Recall
Sorghum	**2803**	211	131	212	90	0	0	0	3	0	0	0	0	0	0	0	3450	0.812
Maize	180	**644**	34	19	5	0	0	0	0	13	0	1	0	0	0	9	905	0.712
Millet	134	55	**383**	0	0	0	0	0	0	0	0	0	0	0	0	0	572	0.670
Cotton	202	83	1	**833**	0	0	0	0	0	0	0	0	0	0	0	0	1119	0.744
Sesame	191	92	15	43	**67**	1	0	0	0	0	0	0	0	0	0	2	411	0.163
Peanut	86	2	1	19	11	**3**	0	0	0	0	0	0	0	0	0	0	122	0.025
asphalt road	0	0	0	0	0	0	**58**	0	0	0	0	0	0	0	0	0	58	1.000
built-up	2	0	3	0	0	0	1	**26**	5	0	0	0	0	0	0	0	37	0.703
clay bare soil	0	0	0	0	0	0	0	7	**37**	1	0	10	0	0	0	0	55	0.673
forest-baresoil interface	0	0	0	0	0	0	0	0	0	**43**	0	0	0	0	4	0	47	0.915
grassland	7	2	7	0	0	0	0	0	0	1	**45**	0	0	0	0	0	62	0.726
lateritic bare soil	0	0	3	0	0	0	0	0	0	1	0	**76**	0	0	0	0	80	0.950
lateritic rock road	2	0	8	0	0	0	0	0	0	0	0	0	**38**	0	0	0	48	0.792
water	0	0	0	0	0	0	1	0	1	0	0	0	0	**15**	0	0	17	0.882
woody vegetation	1	1	0	0	0	0	0	0	0	0	0	0	0	0	**88**	0	90	0.978
Other crops	33	8	0	117	2	2	0	0	0	0	0	0	0	0	0	**0**	162	0.000
Predicted Total	3641	1098	586	1243	175	6	60	33	46	59	45	87	38	15	92	11	7235	
Precision	0.770	0.587	0.654	0.670	0.383	0.500	0.967	0.788	0.804	0.729	1.000	0.874	1.000	1.000	0.957	0.000		

Table 34: Classification report for implementation of SMOTE with sampling ratio of 1 for Sentinel-2 Scenario 1 baseline classification (Chapter 7.5.2.2)

Class	Precision	Recall	F1-score	Support
Sorghum	0.78	0.77	0.78	3450
Maize	0.60	0.69	0.64	905
Millet	0.67	0.68	0.67	572
Cotton	0.67	0.81	0.73	1119
Sesame	0.29	0.19	0.23	411
Peanut	0.20	0.03	0.06	122
Bean	0.08	0.01	0.02	120
Okra	0.00	0.00	0.00	2
Rice	0.35	0.86	0.50	7
Cowpea	0.00	0.00	0.00	25
Tomato	0.00	0.00	0.00	8
asphalt road	1.00	1.00	1.00	58
built-up	0.79	0.89	0.84	37
clay bare soil	0.75	0.73	0.74	55
forest-edge	0.70	0.91	0.80	47
grassland	1.00	0.74	0.85	62
lateritic bare soil	0.87	0.95	0.91	80
lateritic rock road	1.00	0.98	0.99	48
water	1.00	0.94	0.97	17
woody vegetation	0.95	1.00	0.97	90
accuracy			**0.71**	**7235**
macro avg	0.59	0.61	0.58	7235
weighted avg	0.69	0.71	0.69	7235

Table 35: Confusion matrix for implementation of SMOTE with sampling ratio of 1 for Sentinel-2 Scenario 1 baseline classification (Chapter 7.5.2.2)

Confusion Matrix with Recall & Precision

Rows=Actual, Columns=Predicted

	Sorghum	Maize	Millet	Cotton	Sesame	Peanut	asphalt road	built-up	clay bare soil	forest-baresoil interface	grassland	lateritic bare soil	lateritic rock road	water	woody vegetation	Other crops	Actual Total	Recall
Sorghum	**2669**	230	120	241	159	4	0	1	4	0	0	0	0	0	1	21	3450	0.774
Maize	164	**628**	43	23	13	0	0	3	0	14	0	1	0	0	0	16	905	0.694
Millet	121	63	**387**	0	0	0	0	0	0	0	0	0	0	0	0	1	572	0.677
Cotton	152	62	3	**902**	0	0	0	0	0	0	0	0	0	0	0	0	1119	0.806
Sesame	181	62	10	51	**80**	8	0	1	5	0	0	0	0	0	0	13	411	0.195
Peanut	61	2	2	35	18	**4**	0	0	0	0	0	0	0	0	0	0	122	0.033
asphalt road	0	0	0	0	0	0	**58**	0	0	0	0	0	0	0	0	0	58	1.000
built-up	0	0	1	0	0	0	0	**33**	3	0	0	0	0	0	0	0	37	0.892
clay bare soil	0	0	0	0	0	0	0	4	**40**	1	0	10	0	0	0	0	55	0.727
forest-baresoil interface	0	0	0	0	0	0	0	0	0	**43**	0	0	0	0	4	0	47	0.915
grassland	6	0	8	0	0	0	0	0	0	2	**46**	0	0	0	0	0	62	0.742
lateritic bare soil	1	0	2	0	0	0	0	0	0	1	0	**76**	0	0	0	0	80	0.950
lateritic rock road	0	0	1	0	0	0	0	0	0	0	0	0	**47**	0	0	0	48	0.979
water	0	0	0	0	0	0	0	0	1	0	0	0	0	**16**	0	0	17	0.941
woody vegetation	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**90**	0	90	1.000
Other crops	46	4	0	98	3	4	0	0	0	0	0	0	0	0	0	**7**	162	0.043
Predicted Total	3401	1051	577	1350	273	20	58	42	53	61	46	87	47	16	95	58	7235	
Precision	0.785	0.598	0.671	0.668	0.293	0.200	1.000	0.786	0.755	0.705	1.000	0.874	1.000	1.000	0.947	0.121		

Table 36: Classification report for implementation of SMOTE with sampling ratio of 0.5 for Sentinel-2 Scenario 1 baseline classification (Chapter 7.5.2.2)

Class	Precision	Recall	F1-score	Support
Sorghum	0.78	0.80	0.79	3450
Maize	0.61	0.69	0.65	905
Millet	0.69	0.66	0.67	572
Cotton	0.68	0.77	0.72	1119
Sesame	0.31	0.21	0.25	411
Peanut	0.20	0.02	0.04	122
Bean	0.00	0.00	0.00	120
Okra	0.00	0.00	0.00	2
Rice	0.40	0.86	0.55	7
Cowpea	0.00	0.00	0.00	25
Tomato	0.00	0.00	0.00	8
asphalt road	1.00	1.00	1.00	58
built-up	0.75	0.89	0.81	37
clay bare soil	0.75	0.73	0.74	55
forest-edge	0.72	0.89	0.80	47
grassland	1.00	0.76	0.86	62
lateritic bare soil	0.87	0.95	0.91	80
lateritic rock road	1.00	0.98	0.99	48
water	1.00	0.94	0.97	17
woody vegetation	0.94	1.00	0.97	90
accuracy			**0.71**	**7235**
macro avg	0.58	0.61	0.59	7235
weighted avg	0.69	0.71	0.70	7235

Table 37: Confusion matrix for implementation of SMOTE with sampling ratio of 0.5 for Sentinel-2 Scenario 1 baseline classification (Chapter 7.5.2.2)

Confusion Matrix with Recall & Precision
Rows=Actual, Columns=Predicted

	Sorghum	Maize	Millet	Cotton	Sesame	Peanut	asphalt road	built-up	clay bare soil	forest-baresoil interface	grassland	lateritic bare soil	lateritic rock road	water	woody vegetation	Other crops	Actual Total	Recall
Sorghum	**2746**	206	100	206	160	5	0	3	4	0	0	0	0	0	1	19	**3450**	**0.796**
Maize	167	**620**	49	21	18	0	0	3	0	12	0	1	0	0	0	14	**905**	**0.685**
Millet	137	56	**378**	0	0	0	0	0	0	0	0	0	0	0	0	1	**572**	**0.661**
Cotton	204	50	4	**861**	0	0	0	0	0	0	0	0	0	0	0	0	**1119**	**0.769**
Sesame	170	77	8	45	**86**	4	0	1	5	0	0	0	0	0	0	15	**411**	**0.209**
Peanut	76	5	1	20	16	**3**	0	0	0	0	0	0	0	0	0	1	**122**	**0.025**
asphalt road	0	0	0	0	0	0	**58**	0	0	0	0	0	0	0	0	0	**58**	**1.000**
built-up	0	0	1	0	0	0	0	**33**	3	0	0	0	0	0	0	0	**37**	**0.892**
clay bare soil	0	0	0	0	0	0	0	4	**40**	1	0	10	0	0	0	0	**55**	**0.727**
forest-baresoil interface	0	0	0	0	0	0	0	0	0	**42**	0	0	0	0	5	0	**47**	**0.894**
grassland	7	1	5	0	0	0	0	0	0	2	**47**	0	0	0	0	0	**62**	**0.758**
lateritic bare soil	1	0	2	0	0	0	0	0	0	1	0	**76**	0	0	0	0	**80**	**0.950**
lateritic rock road	0	0	1	0	0	0	0	0	0	0	0	0	**47**	0	0	0	**48**	**0.979**
water	0	0	0	0	0	0	0	0	1	0	0	0	0	**16**	0	0	**17**	**0.941**
woody vegetation	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**90**	0	**90**	**1.000**
Other crops	31	1	0	120	1	3	0	0	0	0	0	0	0	0	0	**6**	**162**	**0.037**
Predicted Total	3539	1016	549	1273	281	15	58	44	53	58	47	87	47	16	96	56	7235	
Precision	**0.776**	**0.610**	**0.689**	**0.676**	**0.306**	**0.200**	**1.000**	**0.750**	**0.755**	**0.724**	**1.000**	**0.874**	**1.000**	**1.000**	**0.938**	**0.107**		

Table 38: Classification report for tree filtered validation data set with implementation of tree masking with a 20% Threshold for Sentinel-2 Scenario 1 baseline classification (Chapter 7.5.2.3.)

Class	Precision	Recall	F1-score	Support
Sorghum	0.79	0.81	**0.80**	2955
Maize	0.64	0.71	**0.67**	829
Millet	0.70	0.68	**0.69**	518
Cotton	0.68	0.78	**0.72**	800
Sesame	0.33	0.19	**0.24**	357
Peanut	0.14	0.03	**0.05**	109
Bean	0.09	0.02	0.03	55
Okra	0.00	0.00	0.00	2
Rice	0.31	0.57	0.40	7
Cowpea	0.00	0.00	0.00	18
Tomato	0.00	0.00	0.00	8
asphalt road	1.00	1.00	1.00	58
built-up	0.58	0.86	0.70	37
clay bare soil	0.74	0.67	0.70	55
forest-edge	0.73	0.91	0.81	47
grassland	1.00	0.77	0.87	62
lateritic bare soil	0.87	0.95	0.91	80
lateritic rock road	1.00	1.00	1.00	48
water	1.00	0.94	0.97	17
woody vegetation	0.96	1.00	0.98	90
accuracy			**0.73**	6152
macro avg	0.58	0.59	**0.58**	6152
weighted avg	0.70	0.73	0.71	6152

Table 39: Classification report for non-filtered validation data set with implementation of tree masking with a 20% threshold for Sentinel-2 Scenario 1 baseline classification (Chapter 7.5.2.3.)

Class	Precision	Recall	F1-score	Support
Sorghum	0.77	0.81	**0.79**	3450
Maize	0.63	0.69	**0.66**	905
Millet	0.71	0.65	**0.68**	572
Cotton	0.67	0.76	**0.71**	1119
Sesame	0.34	0.19	**0.24**	411
Peanut	0.14	0.02	**0.04**	122
Bean	0.08	0.01	0.02	120
Okra	0.00	0.00	0.00	2
Rice	0.31	0.57	0.40	7
Cowpea	0.00	0.00	0.00	25
Tomato	0.00	0.00	0.00	8
asphalt road	1.00	1.00	1.00	58
built-up	0.58	0.86	0.70	37
clay bare soil	0.74	0.67	0.70	55
forest-edge	0.73	0.91	0.81	47
grassland	1.00	0.77	0.87	62
lateritic bare soil	0.87	0.95	0.91	80
lateritic rock road	1.00	1.00	1.00	48
water	1.00	0.94	0.97	17
woody vegetation	0.92	1.00	0.96	90
accuracy			**0.71**	7235
macro avg	0.57	0.59	**0.57**	7235
weighted avg	0.69	0.71	0.70	7235

Table 40: Confusion matrix for non-filtered validation data set with implementation of tree masking with a 20% threshold for Sentinel-2 Scenario 1 baseline classification (Chapter 7.5.2.3.)

Confusion Matrix with Recall & Precision

Rows=Actual, Columns=Predicted

	Sorghum	Maize	Millet	Cotton	Sesame	Peanut	asphalt road	built-up	clay bare soil	forest-baresoil interface	grassland	lateritic bare soil	lateritic rock road	water	woody vegetation	Other crops	Actual Total	Recall
Sorghum	**2789**	205	84	221	119	6	0	13	5	0	0	0	0	0	2	6	3450	0.808
Maize	174	**623**	45	18	13	1	0	2	0	12	0	1	0	0	0	16	905	0.688
Millet	147	50	**373**	1	0	0	0	0	0	0	0	0	0	0	0	1	572	0.652
Cotton	232	32	0	**853**	0	0	0	0	0	0	0	0	0	0	2	0	1119	0.762
Sesame	184	69	11	44	**78**	7	0	1	3	0	0	0	0	0	0	14	411	0.190
Peanut	72	4	1	25	16	**3**	0	0	0	0	0	0	0	0	0	1	122	0.025
asphalt road	0	0	0	0	0	0	**58**	0	0	0	0	0	0	0	0	0	58	1.000
built-up	0	0	1	0	0	0	0	**32**	4	0	0	0	0	0	0	0	37	0.865
clay bare soil	0	0	0	0	0	0	0	7	**37**	1	0	10	0	0	0	0	55	0.673
forest-baresoil interface	0	0	0	0	0	0	0	0	0	**43**	0	0	0	0	4	0	47	0.915
grassland	7	0	5	0	0	0	0	0	0	2	**48**	0	0	0	0	0	62	0.774
lateritic bare soil	1	0	2	0	0	0	0	0	0	1	0	**76**	0	0	0	0	80	0.950
lateritic rock road	0	0	0	0	0	0	0	0	0	0	0	0	**48**	0	0	0	48	1.000
water	0	0	0	0	0	0	0	0	1	0	0	0	0	**16**	0	0	17	0.941
woody vegetation	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**90**	0	90	1.000
Other crops	37	4	0	106	5	5	0	0	0	0	0	0	0	0	0	**5**	162	0.031
Predicted Total	3643	987	522	1268	231	22	58	55	50	59	48	87	48	16	98	43	7235	
Precision	0.766	0.631	0.715	0.673	0.338	0.136	1.000	0.582	0.740	0.729	1.000	0.874	1.000	1.000	0.918	0.116		

Table 41: Classification report for RF model calibration with 500 trees for Sentinel-2 Scenario 1 baseline classification (Chapter 7.5.2.4)

Class	Precision	Recall	F1-score	Support
Sorghum	0.77	0.80	**0.79**	3450
Maize	0.64	0.69	**0.67**	905
Millet	0.70	0.66	**0.68**	572
Cotton	0.67	0.78	**0.72**	1119
Sesame	0.35	0.23	**0.28**	411
Peanut	0.19	0.02	**0.04**	122
Bean	0.00	0.00	0.00	120
Okra	0.00	0.00	0.00	2
Rice	0.44	1.00	0.61	7
Cowpea	0.00	0.00	0.00	25
Tomato	0.00	0.00	0.00	8
asphalt road	1.00	1.00	1.00	58
built-up	0.67	0.89	0.77	37
clay bare soil	0.74	0.73	0.73	55
forest-edge	0.70	0.91	0.80	47
grassland	1.00	0.76	0.86	62
lateritic bare soil	0.87	0.95	0.91	80
lateritic rock road	1.00	1.00	1.00	48
water	1.00	0.94	0.97	17
woody vegetation	0.96	1.00	0.98	90
accuracy			**0.72**	7235
macro avg	0.59	0.62	**0.59**	7235
weighted avg	0.69	0.72	0.70	7235

Table 42: Confusion matrix for RF model calibration with 500 trees for Sentinel-2 Scenario 1 baseline Classification (Chapter 7.5.2.4)

Confusion Matrix with Recall & Precision

Rows=Actual, Columns=Predicted

	Sorghum	Maize	Millet	Cotton	Sesame	Peanut	asphalt road	built-up	clay bare soil	forest-baresoil interface	grassland	lateritic bare soil	lateritic rock road	water	woody vegetation	Other crops	Actual Total	Recall
Sorghum	**2751**	181	99	232	144	4	0	8	5	0	0	0	0	0	0	26	**3450**	**0.797**
Maize	172	**626**	46	19	11	1	0	2	0	14	0	1	0	0	0	13	**905**	**0.692**
Millet	139	51	**380**	0	1	0	0	0	0	0	0	0	0	0	0	1	**572**	**0.664**
Cotton	198	44	2	**875**	0	0	0	0	0	0	0	0	0	0	0	0	**1119**	**0.782**
Sesame	178	66	6	41	**93**	4	0	2	5	0	0	0	0	0	0	16	**411**	**0.226**
Peanut	76	2	1	23	16	**3**	0	0	0	0	0	0	0	0	0	1	**122**	**0.025**
asphalt road	0	0	0	0	0	0	**58**	0	0	0	0	0	0	0	0	0	**58**	**1.000**
built-up	0	0	1	0	0	0	0	**33**	3	0	0	0	0	0	0	0	**37**	**0.892**
clay bare soil	0	0	0	0	0	0	0	4	**40**	1	0	10	0	0	0	0	**55**	**0.727**
forest-baresoil interface	0	0	0	0	0	0	0	0	0	**43**	0	0	0	0	4	0	**47**	**0.915**
grassland	7	0	6	0	0	0	0	0	0	2	**47**	0	0	0	0	0	**62**	**0.758**
lateritic bare soil	1	0	2	0	0	0	0	0	0	1	0	**76**	0	0	0	0	**80**	**0.950**
lateritic rock road	0	0	0	0	0	0	0	0	0	0	0	0	**48**	0	0	0	**48**	**1.000**
water	0	0	0	0	0	0	0	0	1	0	0	0	0	**16**	0	0	**17**	**0.941**
woody vegetation	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**90**	0	**90**	**1.000**
Other crops	33	1	0	117	0	4	0	0	0	0	0	0	0	0	0	**7**	**162**	**0.043**
Predicted Total	3555	971	543	1307	265	16	58	49	54	61	47	87	48	16	94	64	7235	
Precision	0.774	0.645	0.700	0.669	0.351	0.188	1.000	0.673	0.741	0.705	1.000	0.874	1.000	1.000	0.957	0.109		

Table 43: Confusion matrix for Sentinel-2 Scenario 2 (Incorporating Field-assessed Heterogneity) (Chapter 7.5.3)

Confusion Matrix with Recall & Precision
Rows=Actual, Columns=Predicted

	Sorghum	Maize	Millet	Cotton	Sesame	Peanut	Other crops	asphalt road	built-up	clay bare soil	forest-baresoil interface	grassland	lateritic bare soil	lateritic rock road	water	woody vegetation	Actual Total	Recall
Sorghum	**2902**	29	60	192	116	2	18	0	0	1	1	3	0	0	0	3	3327	0.872
Maize	259	**461**	33	11	5	4	9	0	0	0	0	0	0	0	0	0	782	0.590
Millet	230	16	**332**	0	8	0	4	0	3	11	0	0	0	0	0	0	604	0.550
Cotton	177	55	2	**1092**	5	1	14	0	0	0	0	0	0	0	0	2	1348	0.810
Sesame	134	6	2	21	**252**	9	1	0	0	8	0	0	0	0	0	0	433	0.582
Peanut	7	0	0	0	22	**10**	1	0	0	0	0	0	0	0	0	0	40	0.250
Other crops	91	3	0	64	13	8	**7**	0	0	0	0	0	0	0	0	0	186	0.038
asphalt road	0	0	0	0	0	0	0	**58**	0	0	0	0	0	0	0	0	58	1.000
built-up	0	0	0	0	1	0	0	0	**33**	3	0	0	0	0	0	0	37	0.892
clay bare soil	0	0	0	0	0	0	0	0	5	**39**	1	0	10	0	0	0	55	0.709
forest-baresoil interface	0	0	0	0	0	0	0	0	0	0	**42**	0	0	0	0	5	47	0.894
grassland	7	0	7	0	0	0	0	0	0	0	2	**46**	0	0	0	0	62	0.742
lateritic bare soil	1	0	2	0	0	0	0	0	0	0	1	0	**76**	0	0	0	80	0.950
lateritic rock road	0	0	1	0	0	0	0	0	0	0	0	0	0	**47**	0	0	48	0.979
water	0	0	0	0	0	0	0	0	0	1	0	0	0	0	**16**	0	17	0.941
woody vegetation	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**90**	90	1.000
Predicted Total	3808	570	439	1380	422	34	54	58	41	63	47	49	86	47	16	100	7214	
Precision	0.762	0.809	0.756	0.791	0.597	0.294	0.130	1.000	0.805	0.619	0.894	0.939	0.884	1.000	1.000	0.900		

Table 44: Classification report for Sentinel-2 Scenario 3 (Incorporating intra-parcel spectral variability) (Chapter 7.5.4.)

Class	Precision	Recall	F1-score	Support
Sorghum	0.68	0.84	**0.75**	3173
Maize	0.72	0.38	**0.50**	713
Millet	0.78	0.69	**0.73**	798
Cotton	0.76	0.70	**0.73**	1,219
Sesame	0.56	0.39	**0.46**	419
Peanut	0.00	0.00	**0.00**	134
Bean	0.00	0.00	0.00	128
Okra	0.00	0.00	0.00	9
Rice	0.00	0.00	0.00	44
Cowpea	0.88	1.00	0.93	7
Tomato	0.00	0.00	0.00	8
asphalt road	1.00	1.00	1.00	58
built-up	0.38	0.89	0.54	37
clay bare soil	0.72	0.71	0.72	55
forest-edge	0.91	0.91	0.91	47
grassland	1.00	0.76	0.86	62
lateritic bare soil	0.88	0.95	0.92	80
lateritic rock road	0.98	1.00	0.99	48
water	1.00	0.94	0.97	17
woody vegetation	0.96	1.00	0.98	90
accuracy			**0.70**	**7146**
macro avg	0.61	0.61	**0.60**	7146
weighted avg	0.68	0.70	**0.68**	7146

Table 45: Validation metrics for CV subsets for Sentinel-2 Scenario 3 (Incorporating intra-parcel spectral variability) (Chapter 7.5.4.)

Class	Subset	Precision	Recall	F1-score	Support
Sorghum	**higher** CV	0.58	0.75	0.65	2170
	lower CV	0.53	0.62	0.57	1003
Maize	**higher** CV	0.44	0.27	0.33	391
	lower CV	0.38	0.17	0.24	322
Millet	**higher** CV	0.51	0.50	0.50	470
	lower CV	0.59	0.44	0.50	328
Cotton	**higher** CV	0.48	0.49	0.48	570
	lower CV	0.65	0.56	0.60	649
Sesame	**higher** CV	0.20	0.19	0.20	197
	lower CV	0.42	0.21	0.28	222
Peanut	**higher** CV	0.00	0.00	0.00	71
	lower CV	0.00	0.00	0.00	63

Table 46: Confusion matrix for Sentinel-2 Scenario 3 (Incorporating intra-parcel spectral variability) (Chapter 7.5.4.)

Confusion Matrix with Recall & Precision
Rows=Actual, Columns=Predicted

	Sorghum	Maize	Millet	Cotton	Sesame	Peanut	Other crops	asphalt road	built-up	clay bare soil	forest-baresoil interface	grassland	lateritic bare soil	lateritic rock road	water	woody vegetation	Actual Total	Recall
Sorghum	**2676**	75	113	161	56	13	59	0	20	0	0	0	0	0	0	0	3173	0.843
Maize	309	**274**	14	44	16	3	27	0	26	0	0	0	0	0	0	0	713	0.384
Millet	229	8	**551**	0	6	0	1	0	2	0	0	0	0	1	0	0	798	0.690
Cotton	305	19	10	**859**	19	0	7	0	0	0	0	0	0	0	0	0	1219	0.705
Sesame	210	3	12	10	**165**	5	4	0	0	10	0	0	0	0	0	0	419	0.394
Peanut	89	1	0	18	25	**0**	1	0	0	0	0	0	0	0	0	0	134	0.000
Other crops	134	0	0	43	7	5	**7**	0	0	0	0	0	0	0	0	0	196	0.036
asphalt road	0	0	0	0	0	0	0	**58**	0	0	0	0	0	0	0	0	58	1.000
built-up	0	0	0	0	0	0	0	0	**33**	4	0	0	0	0	0	0	37	0.892
clay bare soil	0	0	0	0	0	0	0	0	5	**39**	1	0	10	0	0	0	55	0.709
forest-baresoil interface	0	0	0	0	0	0	0	0	0	0	**43**	0	0	0	0	4	47	0.915
grassland	6	0	6	0	0	0	1	0	0	0	2	**47**	0	0	0	0	62	0.758
lateritic bare soil	0	0	3	0	0	0	0	0	0	0	1	0	**76**	0	0	0	80	0.950
lateritic rock road	0	0	0	0	0	0	0	0	0	0	0	0	0	**48**	0	0	48	1.000
water	0	0	0	0	0	0	0	0	0	1	0	0	0	0	**16**	0	17	0.941
woody vegetation	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**90**	90	1.000
Predicted Total	3958	380	709	1135	294	26	107	58	86	54	47	47	86	49	16	94	7146	
Precision	0.676	0.721	0.777	0.757	0.561	0.000	0.065	1.000	0.384	0.722	0.915	1.000	0.884	0.980	1.000	0.957		

PlanetScope Classifications

Table 47: Indices used in PlanetScope classification scenarios

Index	Abbreviation
Normalized Difference Vegetation Index	NDVI
Green Normalized Difference Vegetation Index	GNDVI
Normalized Difference Red Edge Index	NDRE
Chlorophyll Index Green	CIG
Chlorophyll Index Red Edge	CIRE
Modified Chlorophyll Absorption in Reflectance Index 2	MCARI2
Enhanced Vegetation Index	EVI
Soil Adjusted Vegetation Index	SAVI
Modified Soil Adjusted Vegetation Index	MSAVI

Normalized Difference Water Index	NDWI

Table 48: Classification report for PlanetScope Scenario 1 baseline classification without tree masking (Chapter 7.6.2.1.)

Class	Precision	Recall	F1-score	Support
Sorghum	0.74	0.85	**0.79**	58 787
Maize	0.72	0.56	**0.63**	16 025
Millet	0.66	0.72	**0.69**	10 930
Cotton	0.71	0.73	**0.72**	17 499
Sesame	0.27	0.17	**0.21**	7497
Peanut	0.31	0.06	**0.11**	2302
Bean	0.13	0.00	0.00	1796
Okra	0.00	0.00	0.00	140
Rice	0.00	0.00	0.00	432
Cowpea	0.00	0.00	0.00	546
Tomato	0.00	0.00	0.00	246
asphalt road	0.99	1.00	0.99	215
built-up	0.79	0.76	0.77	152
clay bare soil	0.91	0.70	0.79	260
forest-edge	0.90	0.92	0.91	200
grassland	1.00	0.76	0.86	440
lateritic bare soil	0.85	0.99	0.92	426
lateritic rock road	1.00	0.99	0.99	209
water	0.97	0.93	0.95	97
woody vegetation	0.97	0.99	0.98	520
accuracy			**0.70**	118 719
macro avg	0.60	0.56	**0.57**	118 719
weighted avg	0.67	0.70	**0.68**	118 719

Table 49: Confusion matrix for PlanetScope Scenario 1 baseline classification without Tree Masking (Chapter 7.6.2.1.)

Confusion Matrix with Recall & Precision
Rows=Actual, Columns=Predicted

	Sorghum	Maize	Millet	Cotton	Sesame	Peanut	asphalt road	built-up	clay bare soil	forest-baresoil interface	grassland	lateritic bare soil	lateritic rock road	water	woody vegetation	Other crops	Actual Total	Recall
Sorghum	**49973**	1003	2489	2335	2475	202	0	4	8	3	0	1	0	0	12	282	**58787**	**0.850**
Maize	3885	**8928**	873	1512	636	1	0	26	0	8	0	4	0	0	0	152	**16025**	**0.557**
Millet	2395	591	**7905**	24	2	4	0	0	0	0	0	0	0	0	0	9	**10930**	**0.723**
Cotton	2960	1439	180	**12839**	47	29	0	0	0	0	0	0	0	0	0	5	**17499**	**0.734**
Sesame	4618	99	331	1054	**1259**	72	0	0	4	0	0	0	0	0	0	60	**7497**	**0.168**
Peanut	1408	36	43	285	276	**149**	0	0	0	0	0	0	0	0	0	105	**2302**	**0.065**
asphalt road	0	0	0	0	0	0	**215**	0	0	0	0	0	0	0	0	0	**215**	**1.000**
built-up	27	3	0	0	0	0	0	**115**	3	1	0	0	0	3	0	0	**152**	**0.757**
clay bare soil	1	3	7	0	0	0	0	1	**181**	0	0	67	0	0	0	0	**260**	**0.696**
forest-baresoil interface	8	1	3	0	0	0	0	0	0	**184**	1	2	0	0	1	0	**200**	**0.920**
grassland	87	0	11	0	0	0	0	0	0	8	**333**	0	0	0	1	0	**440**	**0.757**
lateritic bare soil	1	3	0	0	0	0	0	0	0	0	0	**422**	0	0	0	0	**426**	**0.991**
lateritic rock road	0	0	3	0	0	0	0	0	0	0	0	0	**206**	0	0	0	**209**	**0.986**
water	0	0	0	0	0	0	3	0	4	0	0	0	0	**90**	0	0	**97**	**0.928**
woody vegetation	6	0	0	0	0	0	0	0	0	0	0	0	0	0	**514**	0	**520**	**0.988**
Other crops	2590	270	82	152	34	26	0	0	0	0	0	0	0	0	0	**6**	**3160**	**0.002**
Predicted Total	67959	12376	11927	18201	4729	483	218	146	200	204	334	496	206	93	528	619	118719	
Precision	0.735	0.721	0.663	0.705	0.266	0.308	0.986	0.788	0.905	0.902	0.997	0.851	1.000	0.968	0.973	0.010		

Table 50: Classification report for tree filtered validation data set with implementation of Tree Masking with a 20% Threshold for PlanetScope Scenario 1 baseline classification (Chapter 7.6.2.2.)

Class	Precision	Recall	F1-score	Support
Sorghum	0.74	0.85	**0.79**	52 006
Maize	0.73	0.57	**0.64**	14 591
Millet	0.66	0.73	**0.69**	10 144
Cotton	0.70	0.73	**0.71**	13 988
Sesame	0.28	0.17	**0.21**	6711
Peanut	0.32	0.07	**0.12**	2099
Bean	0.11	0.00	0.00	1124
Okra	0.00	0.00	0.00	139
Rice	0.00	0.00	0.00	432
Cowpea	0.00	0.00	0.00	462
Tomato	0.00	0.00	0.00	238
asphalt road	0.99	1.00	0.99	215
built-up	0.81	0.76	0.78	152
clay bare soil	0.92	0.70	0.80	260
forest-edge	0.91	0.93	0.92	200
grassland	1.00	0.77	0.87	440
lateritic bare soil	0.85	0.99	0.91	426
lateritic rock road	1.00	0.98	0.99	209
water	0.97	0.94	0.95	97
woody vegetation	0.98	0.99	0.98	520
accuracy			**0.71**	104 453
macro avg	0.60	0.56	**0.57**	104 453
weighted avg	0.68	0.71	**0.69**	104 453

Table 51: Classification report for non-filtered validation data set with implementation of tree masking with a 20% Threshold for PlanetScope Scenario 1 baseline classification (Chapter 7.6.2.2.)

Class	Precision	Recall	F1-score	Support
Sorghum	0.73	0.86	**0.79**	58 787
Maize	0.73	0.55	**0.63**	16 025
Millet	0.67	0.72	**0.69**	10 930
Cotton	0.70	0.72	**0.71**	17 499
Sesame	0.28	0.17	**0.21**	7497
Peanut	0.31	0.06	**0.11**	2302
Bean	0.10	0.00	0.00	1796
Okra	0.00	0.00	0.00	140
Rice	0.00	0.00	0.00	432
Cowpea	0.00	0.00	0.00	546
Tomato	0.00	0.00	0.00	246
asphalt road	0.99	1.00	0.99	215
built-up	0.81	0.76	0.78	152
clay bare soil	0.92	0.70	0.80	260
forest-edge	0.90	0.93	0.91	200
grassland	1.00	0.77	0.87	440
lateritic bare soil	0.85	0.99	0.91	426
lateritic rock road	1.00	0.98	0.99	209
water	0.97	0.94	0.95	97

woody vegetation	0.94	0.99	0.96	520
accuracy			**0.70**	118 719
macro avg	0.59	0.56	**0.57**	118 719
weighted avg	0.67	0.70	**0.68**	118 719

Table 52: Confusion matrix for non-filtered validation data set with implementation of Tree Masking with a 20% Threshold for PlanetScope Scenario 1 baseline Classification (Chapter 7.6.2.2.)

Confusion Matrix with Recall & Precision
Rows=Actual, Columns=Predicted

	Sorghum	Maize	Millet	Cotton	Sesame	Peanut	asphalt road	built-up	clay bare soil	forest-baresoil interface	grassland	lateritic bare soil	lateritic rock road	water	woody vegetation	Other crops	Actual Total	Recall
Sorghum	**44377**	915	2352	1794	2153	206	0	2	8	1	0	1	0	0	7	190	52006	0.853
Maize	3439	**8283**	870	1268	576	1	0	23	0	9	0	4	0	0	2	116	14591	0.568
Millet	2135	549	**7423**	22	2	3	0	0	0	0	0	0	0	0	0	10	10144	0.732
Cotton	2384	1174	147	**10215**	42	23	0	0	0	0	0	0	0	0	2	1	13988	0.730
Sesame	4149	69	304	924	**1161**	64	0	0	4	0	0	0	0	0	0	36	6711	0.173
Peanut	1291	42	28	269	223	**149**	0	0	0	0	0	0	0	0	0	97	2099	0.071
asphalt road	0	0	0	0	0	0	**215**	0	0	0	0	0	0	0	0	0	215	1.000
built-up	27	2	0	0	0	0	0	**116**	2	2	0	0	0	3	0	0	152	0.763
clay bare soil	0	0	7	0	0	0	0	3	**183**	0	0	67	0	0	0	0	260	0.704
forest-baresoil interface	6	0	6	0	0	0	0	0	0	**185**	0	3	0	0	0	0	200	0.925
grassland	85	0	10	0	0	0	0	0	0	7	**337**	0	0	0	1	0	440	0.766
lateritic bare soil	2	2	0	0	0	0	0	0	0	0	0	**422**	0	0	0	0	426	0.991
lateritic rock road	0	0	4	0	0	0	0	0	0	0	0	0	**205**	0	0	0	209	0.981
water	0	0	0	0	0	0	3	0	3	0	0	0	0	**91**	0	0	97	0.938
woody vegetation	7	0	0	0	0	0	0	0	0	0	0	0	0	0	**513**	0	520	0.987
Other crops	1841	257	82	155	30	26	0	0	0	0	0	0	0	0	0	**4**	2395	0.002
Predicted Total	59743	11293	11233	14647	4187	472	218	144	200	204	337	497	205	94	525	454	104453	
Precision	0.743	0.733	0.661	0.697	0.277	0.316	0.986	0.806	0.915	0.907	1.000	0.849	1.000	0.968	0.977	0.009		

Table 53: Confusion matrix for PlanetScope Scenario 2 (Incorporating Field-assessed Heterogneity) (Chapter 7.6.3.)

Confusion Matrix with Recall & Precision
Rows=Actual, Columns=Predicted

	Sorghum	Maize	Millet	Cotton	Sesame	Peanut	Other crops	asphalt road	built-up	clay bare soil	forest-baresoil interface	grassland	lateritic bare soil	lateritic rock road	water	woody vegetation	Actual Total	Recall
Sorghum	**47254**	463	1121	1488	1776	321	561	0	0	0	17	31	1	0	0	0	53033	0.891
Maize	3911	**8577**	621	642	123	112	129	0	0	1	9	0	0	0	0	0	14125	0.607
Millet	3377	521	**7305**	134	120	30	15	0	0	16	1	2	0	0	0	0	11521	0.634
Cotton	3391	1846	58	**15322**	65	46	14	0	0	0	0	0	0	0	0	0	20742	0.739
Sesame	2529	500	367	412	**3334**	274	68	0	0	10	0	2	0	0	0	0	7496	0.445
Peanut	200	113	0	39	287	**330**	23	0	0	0	0	0	0	0	0	0	992	0.333
Other crops	2486	70	3	186	255	44	**78**	0	0	0	0	0	0	0	0	0	3122	0.025
asphalt road	0	0	0	0	0	0	0	**215**	0	0	0	0	0	0	0	0	215	1.000
built-up	19	3	0	0	2	0	0	0	**119**	4	2	0	0	0	3	0	152	0.783
clay bare soil	3	0	3	0	2	0	0	0	3	**182**	0	0	67	0	0	0	260	0.700
forest-baresoil interface	4	4	0	0	0	0	0	0	0	0	**187**	3	2	0	0	0	200	0.935
grassland	84	0	16	0	0	0	0	0	0	0	8	**331**	0	0	0	1	440	0.752
lateritic bare soil	0	2	0	0	0	0	2	0	0	0	0	0	**422**	0	0	0	426	0.991
lateritic rock road	1	0	3	0	0	0	0	0	0	0	0	0	0	**205**	0	0	209	0.981
water	0	0	0	0	0	0	0	3	0	4	0	0	0	0	**90**	0	97	0.928
woody vegetation	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**515**	520	0.990
Predicted Total	63264	12099	9497	18223	5964	1157	890	218	122	217	224	369	492	205	93	516	113550	
Precision	0.747	0.709	0.769	0.841	0.559	0.285	0.088	0.986	0.975	0.839	0.835	0.897	0.858	1.000	0.968	0.998		

Table 54: Validation metrics for CV subsets for PlanetScope Scenario 3 (Incorporating intra-parcel spectral variability) (Chapter 7.6.4.)

Class	Subset	Precision	Recall	F1-score	Support
Sorghum	**higher** CV	0.49	0.70	0.58	29 530
	lower CV	0.50	0.59	0.54	18 952
Maize	**higher** CV	0.44	0.28	0.34	7226
	lower CV	0.38	0.28	0.32	6071
Millet	**higher** CV	0.51	0.47	0.49	7944
	lower CV	0.72	0.49	0.58	6035
Cotton	**higher** CV	0.66	0.60	0.63	9081
	lower CV	0.68	0.59	0.63	9,897
Sesame	**higher** CV	0.18	0.27	0.22	3501
	lower CV	0.65	0.17	0.27	4091
Peanut	**higher** CV	0.00	0.00	0.00	1282
	lower CV	0.00	0.00	0.00	1117

Table 55: Classification report for PlanetScope Scenario 3 (Incorporating intra-parcel spectral variability) (Chapter 7.6.4.)

Class	Precision	Recall	F1-score	Support
Sorghum	0.68	0.90	**0.77**	48 482
Maize	0.67	0.46	**0.55**	13 297
Millet	0.80	0.66	**0.72**	13 979
Cotton	0.84	0.74	**0.79**	18 978
Sesame	0.60	0.49	**0.54**	7592
Peanut	0.00	0.00	**0.00**	2399
Bean	0.00	0.00	0.00	1828
Okra	0.00	0.00	0.00	1358
Rice	0.00	0.00	0.00	432
Cowpea	0.00	0.00	0.00	224
Tomato	0.00	0.00	0.00	246
asphalt road	0.99	1.00	0.99	215
built-up	0.86	0.74	0.79	152
clay bare soil	0.85	0.70	0.77	260
forest-edge	0.95	0.93	0.94	200
grassland	0.97	0.78	0.87	440
lateritic bare soil	0.86	0.99	0.92	426
lateritic rock road	0.98	0.97	0.97	209
water	0.97	0.93	0.95	97
woody vegetation	1.00	0.99	0.99	520
accuracy			**0.71**	111 334
macro avg	0.60	0.56	**0.58**	111 334
weighted avg	0.68	0.71	**0.69**	111 334

Table 56: Confusion matrix for PlanetScope Scenario 3 (Incorporating intra-parcel spectral variability) (Chapter 7.6.4.)

Confusion Matrix with Recall & Precision
Rows=Actual, Columns=Predicted

	Sorghum	Maize	Millet	Cotton	Sesame	Peanut	Other crops	asphalt road	built-up	clay bare soil	forest-baresoil interface	grassland	lateritic bare soil	lateritic rock road	water	woody vegetation	Actual Total	Recall
Sorghum	**43666**	763	1040	1110	1553	182	155	0	0	13	0	0	0	0	0	0	48482	0.901
Maize	5843	**6162**	802	270	127	66	10	0	17	0	0	0	0	0	0	0	13297	0.463
Millet	3856	575	**9238**	258	37	2	1	0	0	0	0	7	0	5	0	0	13979	0.661
Cotton	3685	976	13	**14112**	180	0	12	0	0	0	0	0	0	0	0	0	18978	0.744
Sesame	3199	161	240	167	**3732**	79	3	0	0	11	0	0	0	0	0	0	7592	0.492
Peanut	895	106	19	690	521	**0**	168	0	0	0	0	0	0	0	0	0	2399	0.000
Other crops	3085	431	160	245	110	36	**21**	0	0	0	0	0	0	0	0	0	4088	0.005
asphalt road	0	0	0	0	0	0	0	**215**	0	0	0	0	0	0	0	0	215	1.000
built-up	26	6	0	0	0	0	0	0	**112**	3	2	0	0	0	3	0	152	0.737
clay bare soil	3	3	4	0	0	0	0	0	1	**182**	0	0	67	0	0	0	260	0.700
forest-baresoil interface	8	0	1	0	0	0	0	0	0	0	**186**	2	3	0	0	0	200	0.930
grassland	80	0	7	0	0	0	0	0	0	0	8	**343**	0	0	0	2	440	0.780
lateritic bare soil	1	3	0	0	0	0	0	0	0	0	0	0	**422**	0	0	0	426	0.991
lateritic rock road	4	0	3	0	0	0	0	0	0	0	0	0	0	**202**	0	0	209	0.967
water	0	0	0	0	0	0	0	3	0	4	0	0	0	0	**90**	0	97	0.928
woody vegetation	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**515**	520	0.990
Predicted Total	64356	9186	11527	16852	6260	365	370	218	130	213	196	352	492	207	93	517	111334	
Precision	0.679	0.671	0.801	0.837	0.596	0.000	0.057	0.986	0.862	0.854	0.949	0.974	0.858	0.976	0.968	0.996		

Pléiades Classifications

Table 57: Indices used in Pléiades classification scenarios

Index	Abbreviation
Normalized Difference Vegetation Index	NDVI
Green Normalized Difference Vegetation Index	GNDVI
Normalized Difference Red Edge Index	NDREI
Chlorophyll Index Green	CIG
Modified Chlorophyll Absorption in Reflectance Index 2	MCARI2
Enhanced Vegetation Index	EVI
Soil Adjusted Vegetation Index	SAVI
Modified Soil Adjusted Vegetation Index	MSAVI
Normalized Difference Water Index	NDWI

Table 58: Classification report for Pléiades Scenario 1 baseline classification without Tree Masking (Chapter 7.7.2.1.)

Class	Precision	Recall	F1-score	Support
Sorghum	0.62	0.23	**0.33**	2 000 553
Maize	0.18	0.18	**0.18**	545 022
Millet	0.21	0.35	**0.27**	371 606
Cotton	0.31	0.37	**0.34**	594 879
Sesame	0.10	0.22	**0.14**	255 613
Peanut	0.06	0.08	**0.07**	78 510
Bean	0.01	0.02	0.01	60 958
Okra	0.00	0.02	0.00	4820
Rice	0.01	0.03	0.02	14 772
Cowpea	0.00	0.02	0.01	18 612
Tomato	0.00	0.00	0.00	8271
asphalt road	0.78	0.98	0.87	3697

built-up	0.67	0.92	0.78	2526
clay bare soil	0.54	0.63	0.58	5234
forest-edge	0.02	0.33	0.04	4164
grassland	0.01	0.13	0.02	10 742
lateritic bare soil	0.70	0.73	0.71	9664
lateritic rock road	0.94	0.95	0.95	3817
water	1.00	1.00	1.00	2066
woody vegetation	0.06	0.53	0.11	12 768
accuracy			**0.25**	4 008 294
macro avg	0.31	0.39	**0.32**	4 008 294
weighted avg	0.41	0.25	**0.28**	4 008 294

Table 59: Confusion matrix for Pléiades Scenario 1 baseline Classification without Tree Masking (Chapter 7.7.2.1.)

Confusion Matrix with Recall & Precision

Rows=Actual, Columns=Predicted

	Sorghum	Maize	Millet	Cotton	Sesame	Peanut	asphalt road	built-up	clay bare soil	forest-baresoil interface	grassland	lateritic bare soil	lateritic rock road	water	woody vegetation	Other crops	Actual Total	Recall
Sorghum	**457731**	268116	298636	311049	277616	50544	163	469	695	26493	69518	266	17	0	47800	191440	**2000553**	**0.229**
Maize	58840	**98314**	84336	64250	116634	26560	215	640	1958	12832	7165	614	186	0	7087	65391	**545022**	**0.180**
Millet	54309	55497	**129868**	38694	26922	8227	1	0	88	5770	17444	132	2	0	4287	30365	**371606**	**0.349**
Cotton	107473	57955	44445	**221525**	43572	8835	4	0	3	13108	5587	75	0	0	34978	57319	**594879**	**0.372**
Sesame	30934	46389	24275	39709	**57411**	11454	127	0	57	2944	2352	71	3	0	5321	34566	**255613**	**0.225**
Peanut	9366	9141	6824	7668	20796	**6567**	0	0	0	149	2806	5	0	0	3147	12041	**78510**	**0.084**
asphalt road	14	21	1	0	0	2	**3636**	13	0	4	0	1	1	0	0	4	**3697**	**0.984**
built-up	26	28	0	2	67	1	63	**2325**	1	0	0	0	0	0	0	13	**2526**	**0.920**
clay bare soil	15	101	6	14	0	3	102	0	**3308**	120	0	1555	10	0	0	0	**5234**	**0.632**
forest-baresoil interface	640	302	489	678	39	74	0	0	0	**1366**	148	268	1	0	0	159	**4164**	**0.328**
grassland	3548	1270	1386	679	647	287	0	0	0	647	**1346**	1	0	0	1	930	**10742**	**0.125**
lateritic bare soil	13	467	34	0	1	139	358	0	1	1568	0	**7028**	0	0	0	55	**9664**	**0.727**
lateritic rock road	1	6	57	45	0	0	10	0	0	22	0	25	**3641**	0	0	10	**3817**	**0.954**
water	3	0	0	0	0	0	1	5	0	0	0	0	0	**2057**	0	0	**2066**	**0.996**
woody vegetation	1732	290	501	1779	500	50	0	0	0	18	11	0	0	0	**6809**	1078	**12768**	**0.533**
Other crops	11454	22483	16691	24770	13284	2203	2	0	0	785	1005	7	0	0	2846	**11903**	**107433**	**0.111**
Predicted Total	**736099**	**560380**	**607549**	**710862**	**557489**	**114946**	**4682**	**3452**	**6111**	**65826**	**107382**	**10048**	**3861**	**2057**	**112276**	**405274**	**4008294**	
Precision	**0.622**	**0.175**	**0.214**	**0.312**	**0.103**	**0.057**	**0.777**	**0.674**	**0.541**	**0.021**	**0.013**	**0.699**	**0.943**	**1.000**	**0.061**	**0.029**		

Table 60: Classification report for non-filtered validation data set with implementation of Tree Masking with a 20% threshold for Pléiades Scenario 1 baseline Classification (Chapter 7.7.2.2.)

Class	Precision	Recall	F1-score	Support
Sorghum	0.62	0.23	**0.34**	2 000 553
Maize	0.18	0.18	**0.18**	545 022
Millet	0.22	0.34	**0.27**	371 606
Cotton	0.31	0.38	**0.34**	594 879
Sesame	0.10	0.22	**0.14**	255 613
Peanut	0.06	0.08	**0.07**	78 510
Bean	0.01	0.02	0.01	60 958
Okra	0.00	0.02	0.00	4820
Rice	0.01	0.03	0.02	14 772
Cowpea	0.00	0.02	0.00	18 612
Tomato	0.00	0.00	0.00	8271
asphalt road	0.78	0.98	0.87	3697
built-up	0.67	0.92	0.78	2526
clay bare soil	0.55	0.63	0.59	5234

forest-edge	0.02	0.32	0.04	4164
grassland	0.01	0.12	0.02	10 742
lateritic bare soil	0.70	0.73	0.71	9664
lateritic rock road	0.94	0.95	0.95	3817
water	1.00	0.99	1.00	2066
woody vegetation	0.06	0.64	0.10	12 768
accuracy			**0.25**	4 008 294
macro avg	0.31	0.39	**0.32**	4 008 294
weighted avg	0.41	0.25	**0.28**	4 008 294

Table 61: Classification report for tree filtered validation data set with implementation of Tree Masking with a 20% threshold for Pléiades Scenario 1 baseline classification (Chapter 7.7.2.2.)

Class	Precision	Recall	F1-score	Support
Sorghum	0.63	0.23	**0.34**	1 855 416
Maize	0.18	0.18	**0.18**	513 916
Millet	0.22	0.35	**0.27**	355 664
Cotton	0.31	0.39	**0.35**	518 007
Sesame	0.10	0.23	**0.14**	238 592
Peanut	0.06	0.08	**0.07**	74 420
Bean	0.01	0.02	**0.01**	44 457
Okra	0.00	0.02	**0.00**	4796
Rice	0.01	0.03	**0.02**	14 772
Cowpea	0.00	0.02	**0.00**	16 907
Tomato	0.00	0.00	**0.00**	8145
asphalt road	0.78	0.98	0.87	3697
built-up	0.67	0.92	0.78	2526
clay bare soil	0.55	0.63	0.59	5234
forest-edge	0.02	0.32	0.04	4164
grassland	0.01	0.12	0.02	10742
lateritic bare soil	0.70	0.73	0.72	9664
lateritic rock road	0.94	0.95	0.95	3817
water	1.00	0.99	1.00	2066
woody vegetation	0.10	0.64	0.18	12768
accuracy			**0.26**	3 699 770
macro avg	0.32	0.39	**0.33**	3 699 770
weighted avg	0.42	0.26	**0.29**	3 699 770

Table 62: Confusion matrix for tree-filtered validation data set with implementation of Tree Masking with a 20% threshold for Pléiades Scenario 1 baseline Classification (Chapter 7.7.2.2.)

Confusion Matrix with Recall & Precision
Rows=Actual, Columns=Predicted

	Sorghum	Maize	Millet	Cotton	Sesame	Peanut	asphalt road	built-up	clay bare soil	forest-baresoil interface	grassland	lateritic bare soil	lateritic rock road	water	woody vegetation	Other crops	Actual Total	Recall
Sorghum	**430970**	254084	276618	285948	262208	46825	139	475	671	24349	65191	256	17	0	32746	174919	**1855416**	**0.232**
Maize	54732	**93227**	80179	58227	112759	25175	202	650	1886	12216	6820	592	188	0	5132	61931	**513916**	**0.181**
Millet	51688	53656	**125926**	36609	25584	7785	2	0	88	5340	16494	138	1	0	3515	28838	**355664**	**0.354**
Cotton	95377	50616	38561	**202845**	36704	7523	3	0	2	12041	5195	65	0	0	22292	46783	**518007**	**0.392**
Sesame	28773	43851	22256	37330	**54839**	10886	128	0	50	2732	2183	72	2	0	3404	32086	**238592**	**0.230**
Peanut	9257	8757	6524	7026	20293	**6155**	0	0	0	140	2644	3	0	0	2259	11362	**74420**	**0.083**
asphalt road	5	23	1	0	0	4	**3633**	17	0	5	0	1	1	0	0	7	**3697**	**0.983**
built-up	38	17	1	1	65	0	65	**2325**	2	0	0	0	0	0	0	12	**2526**	**0.920**
clay bare soil	13	98	4	14	0	4	100	0	**3310**	124	0	1558	9	0	0	0	**5234**	**0.632**
forest-baresoil interface	638	308	452	723	50	61	1	0	0	**1347**	140	265	1	0	0	178	**4164**	**0.323**
grassland	3510	1321	1393	758	625	262	0	0	0	623	**1269**	5	0	0	1	975	**10742**	**0.118**
lateritic bare soil	8	408	121	0	4	98	385	0	0	1555	0	**7043**	0	0	0	42	**9664**	**0.729**
lateritic rock road	5	2	59	46	0	0	12	0	0	18	0	28	**3636**	0	0	11	**3817**	**0.953**
water	9	0	0	0	0	0	1	5	0	0	0	0	0	**2051**	0	0	**2066**	**0.993**
woody vegetation	1404	226	351	1396	387	45	0	0	0	21	11	0	0	0	**8168**	759	**12768**	**0.640**
Other crops	9505	18820	15705	19410	10730	1831	2	0	0	748	930	7	0	0	1038	**10351**	**89077**	**0.116**
Predicted Total	**685932**	**525414**	**568151**	**650333**	**524248**	**106654**	**4673**	**3472**	**6009**	**61259**	**100877**	**10033**	**3855**	**2051**	**78555**	**368254**	**3699770**	
Precision	**0.628**	**0.177**	**0.222**	**0.312**	**0.105**	**0.058**	**0.777**	**0.670**	**0.551**	**0.022**	**0.013**	**0.702**	**0.943**	**1.000**	**0.104**	**0.028**		

Table 63: Validation metrics for heterogeneity subsets for Pléiades Scenario 2 (Incorporating Field-assessed Heterogeneity) (Chapter 7.7.3)

Class	Subset	Precision	Recall	F1-score	Support
Sorghum	**higher** heterogeneity	0.27	0.09	0.13	953 891
	lower heterogeneity	0.36	0.18	**0.24**	850 672
Maize	**higher** heterogeneity	0.07	0.11	0.08	183 648
	lower heterogeneity	0.16	0.18	**0.17**	296 504
Millet	**higher** heterogeneity	0.11	0.13	0.12	144 688
	lower heterogeneity	0.15	0.20	**0.17**	247 422
Cotton	**higher** heterogeneity	0.16	0.19	0.18	311 011
	lower heterogeneity	0.27	0.30	**0.28**	393 568
Sesame	**higher** heterogeneity	0.10	0.11	0.10	162 832
	lower heterogeneity	0.06	0.17	**0.09**	92 887
Peanut	**higher** heterogeneity	0.01	0.03	0.01	14 438
	lower heterogeneity	0.03	0.14	**0.05**	19 417

Table 64: Classification report for Pléiades Scenario 2 (Incorporating Field-assessed Heterogeneity) (Chapter 7.7.3)

Class	Precision	Recall	F1-score	Support
Sorghum	0.57	0.23	**0.33**	1 804 563
Maize	0.19	0.24	**0.21**	480 152
Millet	0.22	0.29	**0.25**	392 110
Cotton	0.35	0.40	**0.37**	704 579
Sesame	0.14	0.24	**0.18**	255 719
Peanut	0.02	0.12	**0.04**	33 855
Bean	0.01	0.01	0.01	60 958
Okra	0.00	0.01	0.00	7658
Rice	0.01	0.03	0.01	27 578
Cowpea	0.00	0.11	0.01	1537
Tomato	0.00	0.00	0.00	8271
asphalt road	0.81	0.98	0.89	3697

built-up	0.93	0.90	0.92	2526
clay bare soil	0.75	0.63	0.68	5234
forest-edge	0.03	0.28	0.05	4164
grassland	0.02	0.10	0.03	10 742
lateritic bare soil	0.70	0.74	0.72	9664
lateritic rock road	0.97	0.95	0.96	3817
water	1.00	0.99	1.00	2066
woody vegetation	0.06	0.51	0.11	12 768
accuracy			**0.27**	3 831 658
macro avg	0.34	0.39	**0.34**	3 831 658
weighted avg	0.40	0.27	**0.29**	3 831 658

Table 65: Confusion matrix for Pléiades Scenario 2 (Incorporating Field-assessed Heterogeneity) (Chapter 7.7.3)

Confusion Matrix with Recall & Precision

Rows=Actual, Columns=Predicted

	Sorghum	Maize	Millet	Cotton	Sesame	Peanut	Other crops	asphalt road	built-up	clay bare soil	forest-baresoil interface	grassland	lateritic bare soil	lateritic rock road	water	woody vegetation	Actual Total	Recall
Sorghum	**419526**	274151	263469	324051	204928	83198	141914	58	98	191	18228	39476	252	9	0	35014	**1804563**	**0.232**
Maize	67058	**115107**	57293	79110	60056	24252	57031	183	19	594	6539	5696	522	36	0	6656	**480152**	**0.240**
Millet	67723	68116	**112428**	46858	19643	13915	42943	69	1	170	5302	9510	294	37	0	5101	**392110**	**0.287**
Cotton	128713	66795	44582	**281277**	50738	21836	54160	2	30	0	11003	6684	61	0	0	38698	**704579**	**0.399**
Sesame	28930	57439	24182	35905	**60364**	18811	21176	11	0	163	1098	2579	28	1	0	5032	**255719**	**0.236**
Peanut	2192	8272	2590	2763	10576	**3979**	2947	0	0	0	44	142	0	0	0	350	**33855**	**0.118**
Other crops	13022	21725	5465	27576	17700	6666	**8164**	0	0	0	423	291	30	2	0	4938	**106002**	**0.077**
asphalt road	16	8	0	0	15	1	8	**3632**	13	0	2	0	1	1	0	0	**3697**	**0.982**
built-up	58	71	4	0	39	1	8	62	**2282**	1	0	0	0	0	0	0	**2526**	**0.903**
clay bare soil	21	81	7	21	3	2	0	127	0	**3295**	119	0	1549	9	0	0	**5234**	**0.630**
forest-baresoil interface	696	478	395	797	33	54	169	0	0	0	**1154**	121	265	1	0	1	**4164**	**0.277**
grassland	3772	1194	1483	777	250	1259	446	0	0	0	534	**1022**	5	0	0	0	**10742**	**0.095**
lateritic bare soil	15	438	73	21	3	126	16	344	0	0	1472	0	**7156**	0	0	0	**9664**	**0.740**
lateritic rock road	9	44	40	55	0	0	2	9	0	0	10	0	21	**3627**	0	0	**3817**	**0.950**
water	2	0	0	0	0	0	0	1	10	0	0	0	0	0	**2053**	0	**2066**	**0.994**
woody vegetation	1817	363	421	2184	439	297	735	0	0	0	20	8	0	0	0	**6484**	**12768**	**0.508**
Predicted Total	**733570**	**614282**	**512432**	**801395**	**424787**	**174397**	**329719**	**4498**	**2453**	**4414**	**45948**	**65529**	**10184**	**3723**	**2053**	**102274**	**3831658**	
Precision	**0.572**	**0.187**	**0.219**	**0.351**	**0.142**	**0.023**	**0.025**	**0.807**	**0.930**	**0.746**	**0.025**	**0.016**	**0.703**	**0.974**	**1.000**	**0.063**		

Table 66: Validation metrics for CV subsets for Pléiades Scenario 3 (Incorporating intra-parcel spectral variability) (Chapter 7.7.4.)

Class	Subset	Precision	Recall	F1-score	Support
Sorghum	**higher** CV	0.32	0.10	0.15	1 003 970
Sorghum	**lower** CV	0.31	0.21	0.**25**	644 234
Maize	**higher** CV	0.11	0.12	0.11	245 721
Maize	**lower** CV	0.09	0.13	0.**11**	206 711
Millet	**higher** CV	0.21	0.21	0.21	270 588
Millet	**lower** CV	0.15	0.26	0.**19**	204 948
Cotton	**higher** CV	0.18	0.22	0.20	308 908
Cotton	**lower** CV	0.23	0.31	0.**26**	335 941
Sesame	**higher** CV	0.05	0.11	0.07	119 036
Sesame	**lower** CV	0.13	0.09	0.**11**	138 605
Peanut	**higher** CV	0.01	0.01	0.01	43 684
Peanut	**lower** CV	0.02	0.05	0.**03**	37 921

Table 67: Classification report for Pléiades Scenario 3 (Incorporating intra-parcel spectral variability) (Chapter 7.7.4.)

Class	Precision	Recall	F1-score	Support
Sorghum	0.55	0.25	**0.34**	1 648 204
Maize	0.17	0.20	**0.18**	452 432
Millet	0.27	0.36	**0.31**	475 536
Cotton	0.35	0.44	**0.39**	644 849
Sesame	0.15	0.22	**0.18**	257 641
Peanut	0.04	0.08	**0.05**	81 605
Bean	0.01	0.01	0.01	62 084
Okra	0.00	0.01	0.00	7658
Rice	0.02	0.02	0.02	46 329
Cowpea	0.01	0.03	0.02	14 772
Tomato	0.00	0.00	0.00	8271
asphalt road	0.81	0.98	0.89	3697
built-up	0.80	0.91	0.85	2526
clay bare soil	0.86	0.63	0.73	5234
forest-edge	0.02	0.29	0.05	4164
grassland	0.02	0.10	0.03	10742
lateritic bare soil	0.71	0.73	0.72	9664
lateritic rock road	0.95	0.95	0.95	3817
water	0.99	1.00	1.00	2066
woody vegetation	0.07	0.49	0.12	12768
accuracy			**0.28**	3 754 059
macro avg	0.34	0.39	**0.34**	3 754 059
weighted avg	0.37	0.28	**0.30**	3 754 059

Table 68: Confusion matrix for Pléiades Scenario 3 (Incorporating intra-parcel spectral variability) (Chapter 7.7.4.)

Confusion Matrix with Recall & Precision

Rows=Actual, Columns=Predicted

	Sorghum	Maize	Millet	Cotton	Sesame	Peanut	Other crops	asphalt road	built-up	clay bare soil	forest-baresoil interface	grassland	lateritic bare soil	lateritic rock road	water	woody vegetation	Actual Total	Recall
Sorghum	**411154**	229152	287072	291529	160783	68781	109935	225	166	104	21241	38589	320	6	13	29134	**1648204**	**0.249**
Maize	70181	**89280**	71500	91877	53285	23300	35397	12	399	135	5257	5191	247	27	0	6344	**452432**	**0.197**
Millet	79428	63816	**170002**	56372	28270	19562	34870	55	1	112	7345	10450	341	131	0	4781	**475536**	**0.357**
Cotton	110946	59629	41130	**283005**	43104	16585	41879	0	0	0	9770	3612	64	0	0	35125	**644849**	**0.439**
Sesame	32379	51656	25689	40980	**56267**	21415	21623	100	0	167	975	2610	64	0	0	3716	**257641**	**0.218**
Peanut	12029	12177	7399	13464	17721	**6596**	7104	0	0	0	94	2160	4	0	0	2857	**81605**	**0.081**
Other crops	21710	27104	16080	30414	20198	5258	**10895**	2	0	11	1151	1285	32	4	0	4970	**139114**	**0.078**
asphalt road	21	31	0	0	0	4	7	**3616**	15	0	2	0	1	0	0	0	**3697**	**0.978**
built-up	57	28	0	1	68	0	4	61	**2306**	1	0	0	0	0	0	0	**2526**	**0.913**
clay bare soil	37	121	12	14	0	3	0	73	0	**3301**	108	0	1555	10	0	0	**5234**	**0.631**
forest-baresoil interface	655	464	421	808	34	117	61	0	0	0	**1212**	123	266	1	0	2	**4164**	**0.291**
grassland	3513	1012	1880	997	649	474	420	0	0	0	682	**1097**	18	0	0	0	**10742**	**0.102**
lateritic bare soil	9	554	163	9	5	176	21	314	0	0	1398	0	**7015**	0	0	0	**9664**	**0.726**
lateritic rock road	10	36	43	43	0	0	0	9	0	1	17	0	18	**3640**	0	0	**3817**	**0.954**
water	0	0	0	0	0	0	0	1	6	0	0	0	0	0	**2059**	0	**2066**	**0.997**
woody vegetation	2071	404	527	2159	433	129	699	0	0	0	22	6	0	0	0	**6318**	**12768**	**0.495**
Predicted Total	**744200**	**535464**	**621918**	**811672**	**380817**	**162400**	**262915**	**4468**	**2893**	**3832**	**49274**	**65123**	**9945**	**3819**	**2072**	**93247**	**3754059**	
Precision	**0.552**	**0.167**	**0.273**	**0.349**	**0.148**	**0.041**	**0.041**	**0.809**	**0.797**	**0.861**	**0.025**	**0.017**	**0.705**	**0.953**	**0.994**	**0.068**		

Cropland / Non-cropland Classification with Pléiades

Table 69: Full classification report with all classes for cropland/non-cropland classification with Pléiades (Chapter 7.8.1)

Class	Precision	Recall	F1-score	Support
Sorghum	0.60	0.72	0.66	102848
Maize	0.20	0.10	0.13	20436
Millet	0.51	0.54	0.52	39776
Cotton	0.38	0.36	0.37	23207
Sesame	0.38	0.41	0.39	12716
Peanut	0.12	0.05	0.07	8858
Okra	0.00	0.00	0.00	747
built-up	0.97	1.00	0.98	1740
clay bare soil	0.87	1.00	0.93	2597
grassland	0.37	0.16	0.22	11507
lateritic bare soil	0.98	0.95	0.97	9049
lateritic rock road	0.99	1.00	0.99	5679
water	1.00	0.96	0.98	1398
woody vegetation	0.58	0.59	0.58	10069
accuracy			**0.56**	**250627**
macro avg	0.57	0.56	0.56	250627
weighted avg	0.52	0.56	0.53	250627

Table 70: Confusion Matrix for Pléiades cropland/non-cropland classification (Chapter 7.8.1)

	Sorghum	Maize	Millet	Cotton	Sesame	Peanut	built-up	clay bare soil	forest-baresoil interface	grassland	lateritic bare soil	lateritic rock road	water	woody vegetation	Other crops	Actual Total	Recall
Sorghum	**74764**	3121	11520	6088	2805	1000	0	0	342	1182	0	0	0	1880	146	**102848**	**0.727**
Maize	8541	**2055**	4727	2722	1722	249	0	0	46	213	4	0	0	89	68	**20436**	**0.101**
Millet	14696	1266	**21324**	1015	634	48	0	0	117	484	0	0	0	190	2	**39776**	**0.536**
Cotton	10164	1456	765	**8253**	1903	150	0	0	189	92	0	0	0	217	18	**23207**	**0.356**
Sesame	2652	1060	912	882	**5096**	1386	0	0	0	55	0	0	0	381	292	**12716**	**0.401**
Peanut	4458	133	242	2243	1068	**484**	0	0	0	19	0	0	0	206	5	**8858**	**0.055**
built-up	0	0	0	0	0	0	**1737**	0	1	0	0	0	2	0	0	**1740**	**0.998**
clay bare soil	0	0	0	0	0	0	0	**2597**	0	0	0	0	0	0	0	**2597**	**1.000**
forest-baresoil interface	1637	78	259	289	14	1	0	32	**2781**	214	563	2	0	44	0	**5914**	**0.470**
grassland	5205	309	2559	190	270	947	0	0	188	**1396**	6	0	0	39	138	**11247**	**0.124**
lateritic bare soil	0	0	0	0	0	0	4	1171	521	0	**7376**	0	0	0	0	**9072**	**0.813**
lateritic rock road	0	0	0	0	0	0	0	62	127	0	0	**3009**	0	0	0	**3198**	**0.941**
water	0	0	0	0	0	0	0	0	0	0	0	0	**1515**	0	0	**1515**	**1.000**
woody vegetation	2517	25	44	1539	43	22	0	0	4	5	0	0	0	**8312**	5	**12516**	**0.664**
Other crops	10	4	725	0	8	0	0	0	0	0	0	0	0	0	**0**	**747**	**0.000**
Predicted Total	**124644**	**9507**	**43077**	**23221**	**13563**	**4287**	**1741**	**3862**	**4316**	**3660**	**7949**	**3011**	**1517**	**11358**	**674**	**256387**	
Precision	**0.600**	**0.216**	**0.495**	**0.355**	**0.376**	**0.113**	**0.998**	**0.672**	**0.644**	**0.381**	**0.928**	**0.999**	**0.999**	**0.732**	**0.000**		

Exploiting Multi-resolution Earth Observation Data and Intra-Parcel Heterogeneity Information for Improving Crop Type Mapping in Smallholder Cropping Systems in Mali

Philip Milan Knirsch

Abstract

Smallholder cropping systems present a challenging scenario for crop type mapping with remote sensing approaches. Parcels are small and exhibit complex within-field heterogeneity, containing mixtures of crops and non-crop features such as trees, weeds and bare-soil. Leveraging a unique dataset of parcel-level heterogeneity ratings for various crop parcels in Mali's cotton belt, this study explored the potential of integration with multi-resolution satellite imagery from Sentinel-2, PlanetScope and Pléiades. Pixel-based classification was performed to compare the performance of the three sensors with varying spatial and spectral resolutions and temporal data depth. The study specifically investigated how incorporating intra-parcel heterogeneity information and excluding trees from training data influences classification accuracy. A Pléiades-derived cropland mask was further created to refine validation by accounting for non-crop elements within agricultural parcels.

A comprehensive classification framework was implemented, incorporating monocrop agricultural classes - focusing on cotton, sorghum, millet, maize, sesame and peanut - and various non-crop landcover types, creating a challenging multi-class problem representative of the smallholder landscape mosaic. Incorporating heterogeneity showed to improve classification performance: overall accuracy increased by ~4% for Sentinel-2 (from 71.2% to 76.3%) and PlanetScope (70.2% to 74.4%). Masking out tree-covered pixels within cropland had negligible effects (<0.5% change) across all sensors. Refined validation using a VHR cropland mask produced modest, consistent improvements, revealing that including non-crop elements within parcels in validation data underestimates crop classification performance.

Leveraging intra-parcel heterogeneity can enhance crop type mapping, but effective integration requires more detailed characterization of within-field variability. Relating field-assessed heterogeneity to parcel-level spectral metrics proved inconclusive, with no clear patterns across crop types. Future work should aim to decompose heterogeneity into its components (e.g., weed presence, crop density, bare soil) and create denser time series to capture full crop phenologies to improve discrimination between crop types and enable the incorporation of intercropping into classification frameworks.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
Faculté des bioingénieurs
Croix du Sud, 2 bte L7.05.01, 1348 Louvain-la-Neuve, Belgique | **www.uclouvain.be/agro**