Delft University of Technology

MASTER PROJECT

Identification of Rice Fields in Rwanda with Sentinel-1 in Google Earth Engine

An exploration of remote sensing and rice fields

Author: Shilian Jiang (5248272)

Committee Chair: Dr. ir. Maurits W. Ertsen Prof. dr. ir. Nick van de Giesen Uwacu Alban Singirankabo(MSc) Dr. Stef Lhermitte

October 18, 2022



Abstract

This article uses Sentinel-1 satellite images to identify rice fields in Rwanda from 2017 to 2021 in an attempt to derive the current status of rice in Muvumba catchment, in the northern part of Rwanda. The timing of rice cultivation in each season in that area was identified as not homogeneous, but generally aligned with the local rainy season. The results identified after late 2019 show a large change, with flooding from extreme rainfall as a possible cause. The paddy fields may have been completely flooded and the infrastructure destroyed. Policy changes in Rwanda's agriculture may also be a contributing factor. Among them, the implementation of land consolidation policies can influence some of the farmers' options, for example, by withdrawing from the rice planting program. The absence of official data and field data makes this project makes it exploratory and offers the possibility for potential follow-up studies. The uncertainty of the results also suggests that more attention needs to be focused on these topics. However, what is certain is that the use of remote sensing images to monitor rice has the potential to be cost-effective.

Contents

1	Introduction 3					
2	Research Context 2.1 Research Objectives 2.2 Hydrology and geology characteristics of Muvumba 2.3 Rice field characteristics 2.4 Remote sensing products learning and selection 2.5 Data selection: a brief introduction to selected products and reasons for choosing them	5 5 6 8 9				
3	Methodology 3.1 Data processing platform introduction 3.2 Methodology Overview 3.3 The process of rice identification 3.3.1 Image access in GEE: Preprocessing 3.3.2 Reasons for the selection of frequency and time range of the data set 3.3.3 Supervised classification in GEE	12 12 13 17 17 18 19				
4	Results 4.1 Common patterns	 22 22 23 24 24 				
5	Discussion and Conclusion 5.1 Limitations 5.2 Potential explanations for big changes 5.2.1 Extreme weather 5.2.2 Policy and farmer's choice 5.3 Conclusion	30 30 30 30 31 32				
Ι	Code link	36				
II	I Less important or abandoned methods 37					
III	III Result maps and plots 42					

1 Introduction

This study aims to identify and track rice fields in Muvumba catchment, in the northern part of Rwanda through remote sensing in order to explore the informative potential of remotely monitoring small regional wetland rice fields with an online geo-data processing platform to facilitate understanding of local agriculture. The first step is to find out the location of rice planted in each season in a rice growing area, and the timing and area of different batches of rice planted in each season, by the pattern of backscattering from rice fields. The second step is to compare and analyze the results obtained to try to identify the cause of this rice growing pattern and try to trace this cause back to farmers or policy linkages.

Agriculture in Africa needs attention because food insecurity is still a subject that Africa needs to face. A better understanding of local farming situations can potentially improve the well-being of people. The Republic of Rwanda is a landlocked country in the Great Rift Valley. In this country, agriculture remains the backbone for sustained economic growth, providing direct employment to 62.3% of the population(Ministry of Agriculture and Animal Resources, 2021). It is a typical example of a country that is undergoing land reform, from land registration to land consolidation policies that address land fragmentation. Like more than 70 percent of other countries in sub-Saharan Africa (Green, 2019), rain-fed agriculture without or with supplementary irrigation systems implies a high degree of environmental sensitivity and dependence. Both policy and environment affect agriculture, which is fragile but a major source of income for the general public. In particular, Rwandan agriculture is dominated by smallholder farming, which due to its scattered, diverse, and small-scale properties makes it an easily neglected group in agricultural research. For most smallholder families, agriculture is the only or most important means of subsistence.

The initial plan of the project was to be able to go into the rice fields and conduct field research to understand the current state of Rwanda's rice fields and the plight of smallholder farmers. Rice is a crop that has been seen as promising in recent years, since Rwanda has started to reclaim its rich marsh resources suitable for rice cultivation. It has a high demand as a food in the country and also has a high commercial value. But in March of 2020, the sudden outbreak of the worldwide infectious disease, COVID-19, was a reminder of the high threshold of field research. Field trips to Rwanda were no longer feasible with countries blocked down, and thus using remote sensing to observe Rwanda became the best option for Muvumba.

After choosing the remote sensing approach, the P8 rice cultivation scheme in Nyagatare, which is large scale and has relatively more data for easier verification, was selected as the study area. In 2011, an irrigation system was established at the site to meet the needs of water during the non-rainy season, but the literature shows that the irrigation system did not function efficiently (Green, 2019). Monitoring the farmland at this site provides clues to understanding water use in rain-fed and low-tech rice fields.

The selection of the satellite product was also an important step in the project, as it needed to be able to identify rice and respond to rice growth conditions, such as sensitivity to water or to crop structure. The C-band Sentinel-1 image collection was therefore chosen, in addition to its freedom from clouds and rain (non-extreme rainfall) and its high spatial and temporal resolution, which offered the possibility of time series analysis (Global Forest Observations Initiative, 2018). The water environment of rice fields makes this satellite considered the best product for analyzing rice fields.

In this study, based on the pattern of rice backscatter, using the online geographic information platform Google Earth Engine's fast access to rich data sets, Sentinel-1 images of five years were used to capture the differences between the growth of cycle of rice and other crops based on the principle of multi-temporal RGB synthesis, thereby identifying the location of rice. The backscatter values of the rice field area also provide clues to its growth status. Ten seasons from five years of data from 2017 to 2021 were analyzed. The results of the monitored rice fields are presented as maps of rice planting areas, maps of rice planting times, and backscatter time series plots. The first half of the five year period and the second half show large differences in the results. The results of the first half of the five years showed similar patterns in terms of timing and area of planting. The gradual regularization of rice areas planted in the same batch may be the effectiveness of land consolidation in Rwanda. A slight gradual pattern of delayed planting can also be found in this time period. The results of the second period do not show such a clear pattern. The planting area showed an expansion and a shift in time and place. Combining news, amount of rainfall, and backscatter characteristics of other crops, it could be suggested that the region suffered from extreme rainfall-induced flooding in the latter half of 2019 to 2020, with consequent damage to crops and infrastructure. The area may have remained affected in 2021 as well. In addition to this, Rwandan policies for rice fields and farmers' corresponding responses could also be responsible for the fluctuations of the rice area and timing of cultivation.

The project used Google Earth Engine, a promising online geographic information platform, to remotely identify and inspect rice fields in Rwanda with satellite products, giving results that outline the rice fields in a relatively clear manner. Results partially reflect the growth of the rice. It proves the efficiency of remote low-cost inspection of farmland by remote sensing products. At the same time, it provides a basic for developing a farm-based perspective on smallholder farmers, reflecting the (limited) options farmers may have in the face of natural disasters and changing policies. Below, after describing the background of the study in more detail, the methodology will be explained. The results will be discussed in detail, after which possible explanation will be evaluated.

2 Research Context

After giving the research objectives of this study, this section will present the background of the research area and the remote sensing products that are used. The study area was chosen to be Muvumba P8, a paddy area in the north-central part of Nyagatare in the Muvumba catchment, as this area has a concentrated rice cultivation method supported by an irrigation system and the yields are outstanding in the country (Baikirize, 2019). It is a representative area of concentrated rice cultivation for smallholder farmers in Rwanda. The main satellite product used in this project is Sentinel-1. Independence from clouds, high sensitivity to rice structure and water, and the fit with the temporal-spatial resolution needed to investigate rice growth are its strengths. It is also a product supported by Google Earth Engine, a promising online geographic information processing platform.

2.1 Research Objectives

This project is exploratory and the objective is to use remote sensing images to detect and understand rice fields in Muvumba. Two aspects are included. The first aspect is to explore the potential of satellite images for monitoring rice fields, that is, whether remote sensing can be used to efficiently detect rice fields. The second aspect is to understand the rice fields in Muvumba, trying to understand the growth of local rice through satellites and, accordingly, to interpret the local rice cultivation.

2.2 Hydrology and geology characteristics of Muvumba

The Muvumba P8 rice scheme is part of the Muvumba catchment within Nyagatare province, located near the center of Nyagatare, as shown in figure 1. This area is chosen as the study area for several reasons. Firstly, this catchment is important in the country, with its river sides rich in marsh resources, and is one of the representative places for hydroponic rice in the region. Secondly, a rice cultivation scheme and an irrigation system were constructed in 2011 to increase rice yields supporting food security (Green, 2019). Thirdly, according to Niyonkuru et al. (2020), the soil in the area is ideal for rice growth. Fourthly, this region, the southeast of Rwanda, has the least amount of rainfall in the country (REMA, 2011). As rain-fed agriculture is highly vulnerable to climate change, monitoring the rice fields at this site can provide more insights into climate- and hydrology-related agricultural dilemmas. Weather data scarcity is a general problem which creates a limitation when choosing a study area(Franiriantsoa in green, 2019). More data are available for this region. In addition to the P8 area, this study also covers the rice fields near the Colline Kiboga area in the southern part of the region. Since this area is identified as an area of concentrated rice cultivation in the data for 2020 and 2021, which was not the case in previous years, as will be discussed in detail in of the results section.

Rwanda has a temperate tropical highland climate, with lower temperatures than are typical for equatorial countries because of its high elevation (U.S.Department, n.d.). Figure 2 is a map of the SRTM digital elevation model, with the elevation of the study site at 1200-1400 m. Kigali, in the center of the country, has a typical daily temperature range between 12 and 27 °C (54 and 81 °F), with little variation throughout the year. There is little difference in temperature between provinces, so the temperature pattern is almost the same in Nyagatare (WMO, 2022). According to Green (2019), Rwanda is rich in surface water resources, such as lakes, wetlands, and rivers. But Rwanda is still facing water stress because of its population size. Nyagatare's drought vulnerability was evaluated as high(Ministry of Environment , 2018). The areas downstream along the Muvumba river have already been suffering from long periods of drought and these water shortages



Figure 1: Map of Nyagatare District showing the location of the study area, the location of the irrigation dam and the photo from local researcher (?Africana Hills, 2022)

can potentially get worse in the future. Therefore, as an important source of water for agriculture, rainfall patterns are very important for studying local agriculture. Figure 3 and 4 show the Muvumba monthly rainfall for 2017-2022 and the daily rainfall comparison for 2017-2019, respectively. Rwanda has two rainy seasons per year, bridged by two relatively dry periods. Both 2018 and 2020 have a significant peak in rainfall around March compared to the rest of the years. Based on these two maps and the aforementioned temperature patterns, it is easy to see that there are no distinct spring, summer, fall, and winter seasons in Rwanda. The local distinction between the seasons of the year is based on rainfall, with four climatic seasons and two corresponding agricultural seasons. The country's four climatic seasons are represented through the long rainy season (March to May) and the short rainy season (September to November). These seasons alternate with the long dry season (June to August) and the short dry season (December to February). The periods for the two agricultural seasons (sometimes referred to as Season A and Season B) in the region include the first cultivable season (from September to January) and the second cultivable season (from February to June). This again reflects the high dependence of agriculture on rainfall in the area.

2.3 Rice field characteristics

According to the Ministry of Agriculture and Animal Resources (2021); Ministry of Environment (2018); Green (2019), agriculture remains the backbone for sustained economic growth, providing direct employment to 62.3% of the population. In Muvumba catchment, this figure can reach to 90%. About 58% of Rwanda's land is arable, of which about 5% is potentially marshlands. Currently only half of Rwanda's potential marshlands, about 63,742 ha, is actually used for agriculture. In Muvumba catchment, 22% of wetland utilization is rice. However, Rwanda's domestic rice production does not meet even half of the national rice demand, and the rest of the demand is met by importing. The demand for rice as an easy-to-cook cash crop is increasing in Rwanda, where the population is rising rapidly. This suggests that rice is a promising and popular crop in Rwanda (or the Muvumba region), both as a food and as an economic product.



Figure 2: SRTM digital elevation map in the Muvumba area in Rwanda (2000)

The rice project in the research area is called Muvumba P8(Figure 1). The rice fields are mainly rain-fed, but is supplemented by a simple irrigation system in the dry season to enable two seasons of rice in one year(Baikirize, 2019; Green, 2019; Ministry of Environment, 2018; Ministry of Agriculture and Animal Resources, 2021). The dark green dot the figure 1 shows the irrigation dam (30.281465, -1.313913) at the head of the irrigation district. A photo of the dam is shown in the lower left corner (Africana Hills, 2022). A simplified layout of the irrigation system with the diversion dam, the main channel, the reservoirs and the Muvumba river is shown in Figure 5. The three reservoirs in the main study area and the two reservoirs in the secondary study area are shown in the black circle. Two recent studies of the area claim that there is 1750 ha of rice fields in the area (Green, 2019; Niyonkuru et al., 2020), but an official press report states that the farmland area is only 1050 ha (Minagri, 2022). Figure 1 is used to estimate the scale of study areas rather than to validate the results. As farmers have the freedom to participate and withdraw from leasing paddy fields from the government for cultivation, the area of paddy fields is variable. An accurate annual registered area is not available.

On Muvumba P8, the farmers are growing rice twice a year. The varieties of rice are Japonica (*Oryza sativa japonica*) and Indica (*Oryza sativa*). The yield is about 6t/ha/season and 12t/ha/year (Green, 2019). Wetland resources are state-owned land in Rwanda, and the farmers who live near these lands pay a small fee to the government for the right to use wetland areas. Farmers are each allocated 0.1-1.0 hectares of land and share the agricultural infrastructure (Niyonkuru et al., 2020; Green, 2019).

In 2020, Niyonkuru et al. (2020) selected three plots in the area for soil experiments. A total of 45 cm thickness of soil was divided equally into three layers and tested for soil type, bulk density, soil water content, average infiltration rate, and infiltration rate. This article concluded that soil there was able to provide sufficient moisture and support for rice. There is no uniform timing of rice cultivation in the region. Rice in the study area is grown twice a year, but the time of the start of each season provided by local farmers can vary by three months (Green, 2019). This is characteristic of smallholder farming. In this study, rice fields are studied based on different planting periods.



Figure 3: Average monthly rainfall data in muvumba area in Rwanda drived from CHIRPS Pendted data



Figure 4: Daily rainfall date in 2017, 2018 and 2019 in Muvumba area that drived from CHIRP DAILY data

2.4 Remote sensing products learning and selection

Remote sensing of rice growing areas helps to accurately map rice areas and monitor paddies. It is increasingly used by scientists around the world to trace rice, and this type of information has proven to be reliable, cost-effective, quantitative, instantaneous, and non-destructive to the observed objects (Kuenzer & Knauer, 2012). The lack of hydrological data in Sub-Saharan Africa has been mentioned(Green, 2019; Ministry of Environment, 2018), and this situation is also applicable to Rwanda. The scope of this missing data is very broad. For example, there is no timely information on land classification, and locating farmland is relatively difficult. Likewise, it is difficult to know the location and structure of local irrigation systems. Green (2019) found it hard to give accurate predictions of climate and water demand changes for Nyagatare watershed because of missing hydrological, soil, and climate data. Therefore, in the absence of sufficient field data, remote sensing data becomes an essential source to study agriculture and water in Africa, hence it is used in this study.



Figure 5: A simplified layout of the irrigation system with the diversion dam, the main channel, the reservoir and the Muvumba river (Green, 2019) The three reservoirs in the main study area and the two reservoirs in the secondary study area are shown in the black circle in the small figure on the right.

2.5 Data selection: a brief introduction to selected products and reasons for choosing them

Given the wide variety of remote sensing data, an important part of this study was to compare different remote sensing products and filter the ones that can effectively identify rice fields and give rice growth information. Therefore, this subsection will introduce the different satellite products and provide reasons for the selection.

This study will combine Synthetic Aperture Radar (SAR) and optical satellite data to identify and track rice conditions. It will be supported by other auxiliary data such as Shuttle Radar Topography Mission (SRTM) digital elevation data, Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) rainfall data and data from literature or government documents on the target irrigation areas and local water resources.

• Synthetic Aperture Radar (SAR)

SAR is short for Synthetic Aperture Radar. There are many products covered under SAR, and this paragraph will describe the definition, working principle and product classification of SAR. SAR is a type of active data collection where a sensor produces energy and then records the amount of that energy reflected after interacting with Earth. The signal it received is the backscatter. The synthetic aperture is a sequence of acquisitions from a shorter antenna that is combined to simulate a much larger antenna so that the synthetic aperture can greatly improve the spatial resolution of radar (*Synthetic Aperture Radar — Earthdata*, n.d.). Sentinel-1A was launched in March 2014. This mission was the first C-band mission that has the goals of global coverage, continuity, temporally dense coverage and free and open data access built-in (Clauss et al., 2017). Theoretically, data before 2014 are not available. In practice, data before 2017 were found to be missing during the experiment. Therefore, in this study, rice fields from 2017 to 2021 were observed.

SAR products have a wide range of options, depending on the wavelength of their emitted signal, the direction of the emitted and received signal (polarization), the method of picture collection, the level of picture processing, etc. The strength of the backscatter varies with the different options.

The commonly used bands can be divided into P-band (~69.0 cm), L-band (~23.5 cm), S-band (~9.4cm), C-band (~5.6cm) and X-band (~3.1cm) according to the wavelength from long to short. The strength of the backscatter also depends on the direction of signal emission and reception (polarization), vertical or horizontal. This polarization is commonly denoted by HH, HV, VH, and VV, in which H means horizontal and V means Vertical. Compared to optical sensors, radar sensors utilize longer wavelengths at the centimeter to meter, which enable them to see through clouds. This advantage of freedom from cloud interference is often used in agricultural research. The selection of the final product depends on the characteristics of the study object. This will be explained in the next paragraph.

Sentinel-1 C-band GRD data is the product that is used in this research, and the instrument mode is Interferometric Wide Swath (IW). VH and VV images are selected. SAR backscatter is a function of a target's (the plant itself and the soil or water in which it is planted) moisture content and structural characteristics, and also to standing open water and below-canopy standing water. It is also influenced by interactions between vegetation and underlying soil (Steele-Dunne et al., 2017). So, the radar signal is sensitive to the morphological changes in the plant canopy and the changes in water content within the plant and in the soil (Kuenzer & Knauer, 2012).

The choice of the waveband depends on the size of the object to be observed. This is because radar waves can pass through objects that are shorter than the starting band. The C-band (about 5.6 cm) matches the morphological size of rice. Figure 6 is a comparison of L-band and C-band effects on an irrigated rice field in Vietnam. Rice at different stages of cultivation can be distinguished when the C-band is applied. Therefore, this band is also described as probably the most suitable for irrigated rice fields (Global Forest Observations Initiative, 2018). VH and VV polarizations are used in this project. The choice of polarization was based on the experience of other scholars. In a review paper of Steele-Dunne et al. (2017), many studies have confirmed that cross-polarization (VH or HV) is the single most important polarization for the identification of most crops. However, the addition of a second polarization would give increased accuracy in identification and an improvement in the quality of results for some crops. Nguyen & Wagner (2017); Csorba et al. (2019) demonstrated the potential of VH polarization to establish a continental-scale framework for rice field observations. Some other experiments using VV and VH polarization demonstrate that this method outperforms optical picture-based models in detecting crops, including yield prediction (Veloso et al., 2017; Ranjan & Parida, 2020). Bazzi et al. (2019) also used VV and VH polarization images to build a rice identification model, which proved to be a reliable and operational approach. As for the GRD, it is short for Ground Range Detection. Level-1 data can be processed into either Single Look Complex (SLC) and/or Ground Range Detected (GRD) products. SLC products preserved phase information and are processed at the natural pixel spacing whereas GRD products contained the detected amplitude and are multi-looked to reduce the impact of speckle (European Space Agenc, n.d.). GRD is the only available collection in GEE.

The Sentinel-1 mission has a twelve-day roving cycle, which means that under normal conditions, a new satellite image is produced every twelve days. This temporal resolution allows for the successful capture of changes in crop phenology in the life cycle of rice during time series analysis, which is necessary for crop classification.



Figure 6: The different effects of C-band and L-band Sentinel-1 applied to rice fields (Global Forest Observations Initiative, 2018)

3 Methodology

Rice fields from 2017 to 2021 for a total of five years and ten growing seasons using multiple S1 images were identified with GEE. Rice can be distinguished from other crops by the trend of its backscattering during the planting stage due to its submerged character. The RGB (red, green, blue color model) image composition is used to assist in capturing representative pixels at the stage when the rice is just planted. The remaining areas are classified using supervised classification. Maps and backscatter time series plots of different batches of rice were thus obtained for further analysis and discussion. Method construction is the core of this project. This project is of the exploration type, in other words, it explores how much information is given by remote sensing observation of rice fields in Rwanda. Remote sensing being a field unfamiliar to the author, the method used was not established until after testing methods established from other literature. Therefore, the detailed steps are presented in this section to give the reader as much clarity as possible and to enable the possibility of reproducing the method. The earlier methods used for exploration are presented in the appendix, as they are not the primary step in deriving the results.

3.1 Data processing platform introduction

Geodata is data that also contain information about geographic locations and that can be handled a geographic information system (GIS) (Lhermitte, 2021c). The geographic information used in this study was accessed, processed and operated by the online platform, GEE. It combines a multi-petabyte catalog of satellite imagery and geospatial datasets with planetary-scale analysis capabilities (Google, n.d.). The advantage of this platform over traditional geographic information processing is that the information is stored and computed on the server, and the client only needs to program the web to request the processing results from the server. Figure 7 illustrates the logic between clients and GEE server. The efficiency gap between this platform and traditional methods is obvious. The size of a SAR image is typically 1G, and the application of time analysis usually involves many images (about 150 images were used in this project). This step is done in GEE on the server side instead of the user side. Therefore, the researcher is not limited to the storage size of the computer and does not spend time on filtering and downloading data. This also makes it easy to handle large amounts of data simultaneously. There is also increased fault tolerance, as the client only uses modified code to modify the information. There are Javascipt Application Programming Interface (API), Python interface and Qgis plugin to get access to the GEE web server. Javascript API is the most mature and easiest to get started in these three methods, and this study is based on the principle of this method. However, compared with Python, its data export and processing and plotting functions are much weaker, so this study used Python and Qgis to complete the plotting and data scientific analysis (Lhermitte, 2021b).



Figure 7: A picture showing the logic of the GEE side and the client's access to information(Lhermitte, 2021b).

3.2 Methodology Overview

The supervised classification performed in this thesis is based on the different characteristics of rice from other crops as reflected in remote sensing satellite images. This is the technique most often used for the quantitative analysis of remote sensing image data (Richards, 2022).

The core of this method is the segmentation of the spectral domain into regions that can be associated with classes of land surface cover. The spectral domain is divided into regions that can be associated with the ground cover category of interest for a certain purpose. Many algorithms exist for this task, and this study applies the Random Forest. In practice, ground cover may be misclassified because some features are similar to other categories, so there are requirements on the type, quantity and quality of the input images (Richards, 2022). Figure 8 shows the working steps of general supervised classification. The first three parts, inputting images, collecting training data and training a classifier are done by the operator. The later steps are done by the computer, hence the method is also called machine learning. Figure 9a shows the flow chart for the identification of rice fields in this project. This figure shows more details of the method and completes the methodological structure of this study.



Figure 8: work flow of supervised classification(Lhermitte, 2021a)



(a) Flow chart for identification of paddy fields



(b) Agriculture RGB composite s2 map of research district

Figure 9: RGB composite of late March in 2017 using VH or VV; The magenta part is the area where the rice is located, where more details are shown in VH and the spectrum is more continuous, but the rice part in VV is more striking

The input image-collections selected for this study were S1 and S2. The time series analysis of S1 provide the backscattering characteristics that distinguish rice fields from other vegetation species. S2 images provide a visually clearer view of the farmland area by combining images from different bands.

S1 is a synthetic aperture radar whose received signal strength is very sensitive to the morphological structure of water and crops. In figure 10, rice growth can be divided into three stages, planting, growing and harvesting. Rice in the planting stage is submerged by water and its backscattering is similar to standing water. During this rice growing phase, water is the dominant signal characteristic received by S1. Therefore, the backscatter of the rice field during this period will be significantly lower than other plants and close to the value of the water surface. The backscatter value of the rice field gradually increases as the rice plants grow, and the water surface is becoming covered by rice plants. Thus backscatter has a variation from 'high' to 'low' and then 'higher' during the growth cycle of rice. Since the representation of backscattering is negative, its value shows an increase from cultivation to growth. To better capture this feature, this study uses RGB composite images to render visually distinguishable features. (The basics of how multichannel SAR imagery is displayed in figure 11 (OpenGeo Lab, 2022).)



Figure 10: Backscatter mechanisms in dependence of rice crop growth (Clauss et al., 2017).

1 st Date (Red)	2 nd Date	3 rd Date (Blue)	FCC Colour
Value	(Green) Value	Value	
Low	Medium	High	
(~80)	(~140)	(~180)	
High	Low	Medium	
(~180)	(~80)	(~120)	
High	High	Low	
(~200)	(~180)	(~70)	

Figure 11: RGB synthesis logic schematic (OpenGeo Lab, 2022).

Three grey SAR images are assigned to Red/Green/Blue band, and when the values (backscattering) of the images present high (white in pixel), low(black in pixel) and medium (white in pixel), the resulting color will be magenta. When the values of the three images are trending from high (white in pixel) to low (black), the composite image display cyan. When the value of the pictures shows a trend of gradually decreasing, the resulting color of the composite will be yellow. Therefore, by capturing the magenta pixel of the composite image, the rice-related pattern selected. The description of the trend alone seems too abstract, Box I following visualizes the specific steps that the operator did in the GEE.

Rearly February + Glass February + Bearly May	Three VH S1 images from different times are assigned as red, green and blue channels.
(A) (B) (C)	 Collect three points (A, B and C) in the composite image as sample points. The backscattering trends corresponding to the sample points (backscattering at different times) show the lowest middle value at pointd B and C (backscattering is negative, and longer bars indicate lower values). This is the visualization of "a variation from 'high' to 'low' and then 'higher' during the growth cycle of rice". In the composite map, points B and C appear magenta due to this trend and are considered to be successful in capturing the rice cultivation stage. Therefore these two pixels are classified as rice as training data.
Rine February * Gearly March * Bearly May Gearly March * Bearly May (A) (B) (C) (D)	Three images from different times are assigned as red, green and blue channels. Collect four points (A, B C and D) in the composite image as sample points. The backscattering trends corresponding to the sample points show the lowest middle value at point D. In the composite map, point D appear magenta due to this trend and is considered to be successful in capturing the rice cultivation stage. Therefore this one pixel is classified as rice as training data.
Rearly March + Glane March + Bearly Aged	Three images from different times are assigned as red, green and blue channels.Collect four points (A, B C and D) in the composite image as sample points.The backscattering trends corresponding to the sample points show the lowest middle value at point D.In the composite map, point D appear magenta due to this trend and is considered to be successful in capturing the rice cultivation stage. Therefore this one pixel is classified as rice as training data.



Box I: detailed step-by-step approach to identify rice fields.

Detecting rice areas by capturing the change in backscatter values from the flooding period after rice planting to the stage of harvesting has proven to be a simple and effective means of rice area detection (Kuenzer & Knauer, 2012; Global Forest Observations Initiative, 2018). Multitemporal images combined in the Red/Green/Blue bands are used to capture the low-lying values of backscatter during rice growth and thus to determine the location of the rice field.

The role of S2 is to assist in determining the area of the rice. The combination of bands 11 (SWIR 1), 8 (NIR), 2 (Blue), which is also known as the agricultural RGB combination, is used in this study. This combination is mostly used to monitor crop health, as both short-wave and near infrared bands are particularly good at highlighting dense vegetation, which appears dark green in the composite. SWIR measurements can be helpful to reflect water which is present in plants and soil, as water reflects SWIR light (Sinergise, n.d.; *IRS Lab 2 - GEARS - Geospatial Ecology and Remote Sensing*, n.d.). The effect of agricultural RGB on the study area is shown in figure 9b. Farmland areas (mostly rice fields and forests) are distinguished in figure 9b as being greener than other land uses.

The reference used in this study to construct the methodology to identify and analyze rice paddy are mainly publicly available online GEE learning materials (OpenGeo Lab, 2022; GEARS, n.d.; Gandhi, n.d.) and TU Delft related courses (CIE4616, CIE5401). The code is also partially written with reference to the answers of some google search results. The addresses of all codes used in this project are listed in the Appendix A.

3.3 The process of rice identification

3.3.1 Image access in GEE: Preprocessing

The Level 1 S1 data used in this study had to be preprocessed for its use in the analysis to be meaningful (Esri, n.d.). Pre-processing deals with noise and distortion caused by sensors, terrain, radiation, heat, etc. It is also influenced by the purpose of use and scope of study, so preprocessing is not a uniform fixed step. Some common pre-processing operations are listed below.

- Apply orbit file
- Remove thermal noise
- Apply radiometric calibration
- Terrain correction
- GRD border noise removal
- Despeckle
- Apply geometric terrain correction

The first five of these processing operations are done automatically in GEE, of which GRD boundary noise removal was terminated on January 12, 2018 (Google Developers, 2022). Compared with traditional methods, Earth Engine provides a simpler mode of image collection operation, but with less flexibility. The user cannot independently choose the preprocessing operation to be performed. In addition to this, the preprocessing of S1 images by Earth Engine also varies with time changes, which can cause systematic errors to the results of this experiment. To eliminate the scattering effect of terrain on microwave radiation, GEE also automatically performs a unitable transformation of its images. The S1 image collection in the Earth Engine consists of a level 1 ground sounding (GRD) scene with a processed scattering coefficient (σ°) in decibels (dB). The backscatter coefficient represents target backscattering area (radar cross-section) per unit ground area. Because it can vary by several orders of magnitude, it is converted to dB as $10*\log 10^{\circ}$. It measures whether the radiating terrain preferentially scatters the incident microwave radiation away from the SAR sensor (dB<0) or towards the SAR sensor (dB>0). This scattering behavior depends on the physical properties of the terrain, mainly the geometry of the terrain elements and their electromagnetic properties.

3.3.2 Reasons for the selection of frequency and time range of the data set

This section will give explanations of the reasons for the choice of the time period for the data. A time series analysis of each rice season will be conducted separately in this project. Depending on the time of year when the rice is planted and grown, the rice in the region is divided into short rainy season (season A) and long rainy season (season B), with a predicted growth cycle of four to six months for each season (Green, 2019). The minimum time unit used for this thesis was chosen to be half a month, with the exact number of days varying with the month, a guarantee that two observations would be made each month for short rainy season (January 1 to the end of July) and long rainy season (July 1 and to the end of January of the next year) for year 2017 to 2021. Since the cycle of S1 at research area is usually 12 days, sometimes there may be more than one image in a half month. In this case, the average of all the pictures of this period is combined into a new image. For example, the image called 'early March' in this study is the average of all S1 images within the first half of March of that year. Fourteen averaged radar images are available for almost every season (Late January data are missing in 2017, Early November data are missing in 2021).

- (1) One of the reasons for this timing determination is that the timing of rice planting in this study area is not uniform. Only small and fragmented areas of land are allocated to each Rwandan farmer, and local farmers the freedom to choose when to plant. Therefore, a planting window that may be up to several months long is expected to be observed in this study area. Thus, the study time window for the two seasons overlaps for a total of two months throughout the year, January and July. This is to prevent potential planting times from being skipped.
- (2) The temporal frequency and year of the acquired satellite images are due to the accessibility of the data(temporal resolution) and the features of the agricultural observations(for example, number of crops, consistency of planting practices, presence of intercropping, and number of cropping seasons per year (Steele-Dunne et al., 2017). The minimum cycle in this study is one growth cycle of rice, which is a growing season of about half a year. Satellite images with a temporal resolution of half a month matches the temporal resolution of crop growth, implying that changes in crop phenology can be captured. The twelve-day revisit time of S-1's out-of-Europe region image collection enables this analysis.
- (3) The time span selected for this study is 2017-2021, which is subject to the limited amount of available data. The time frame for the original plan was chosen to be 2015-2019, for a total of five years and ten crop growing seasons. Since the purpose of this study is to analyze the reasons why farmers make agricultural strategies, unexpected social factors can be a distracting factor in the analysis. The years 2020 and 2021 were considered to be potentially affected by the COVID-19, and it was difficult to weight the factors. However, continuous S1(launched in 2014) images covering the present study area were not available in GEE until after 2017. Overall, five years of data compared to three years of data are more reliable in the operation of conducting time series analysis to guarantee the credibility of the results. Therefore, the final time period chosen for the study are 2017-2021.

3.3.3 Supervised classification in GEE

In this study, the RGB composite S1 images were used as the references for supervised classification of the experimental area to identify the area in which the rice was located, the results were compared accordingly. The introduction of the method has been mentioned at the beginning of this chapter, and the steps of the method will be described in this section.

(1) Load the boundary of the research area

The first step was to introduce the second level of FAO's Global Administrative Unit Layer (GAUL2015) map in GEE and filter the 'Nyagatare' area by the name. The study area for this project was rice grown within the marshes along the river in the region, so this step screened out a much larger area than the actual site. The step of refining the boundaries will be performed later.

(2) Load and stack the images

In GEE, need to filter the required images by the code made by GEE library. The required indices for filtering images are time, signal acceptance polarization, instrumentation pattern, orbit passing properties and location. The 14 filtered VH images containing only backscatter are then stacked and labeled in order to facilitate the combination of the requested images.

(3) Classify and label representative pixels with rice paddy features (Collect training data)

This step labels and classifies the representative pixel points that represent rice, forest, town, surface water and other surface types, based on the spectral information of the satellite images, in order to enable the computer to classify all the pixels remaining according to this classification logic. This method is commonly called 'collecting training data'. 'Training' refers to the estimation of the parameters that the classifier needs in order to be able to recognise and label unseen pixels. Therefore, the different types of land use should have visually distinguishable features, and each spectral class is required to have enough training pixels to allow reasonable estimation of the elements of the class conditional mean vector and covariance matrix (Richards, 2022).

Generally, the principle is to identify the location of rice by using the color synthesized from the trend of the values of three different pictures to capture the dipping part of the rice life cycle. The actual process is as follows. The images potentially located in the rice planting period are sequentially used as the middle image of the RGB composite image, traversing all of them. After taking into account the upper and lower bounds of backscattering, the magenta pixels in all composites are captured and classified by the time in which their middle pictures lie, indicating that this is the time from which the captured rice was planted. The operation is repeated for each season. The next paragraph is a practical example.

Screenshots of the classification process are shown in figure 12. The three RGB channels are from early March, late March and early May. As can be seen in Fig 12a, the RGB combination of VH is sensitive to the different growth states of rice. The magenta gradient in this figure is relatively smooth and rich, which can show the relatively continuous spectrum of rice. The saturation of the magenta in the same region in image 12b has a much higher contrast compared with other areas, but there is relatively less detail in the picture. Thus VH and VV work together to improve the quality of classification results by improving the accuracy of target pixels.



Figure 12: RGB composite of late March in 2017 using VH or VV; The magenta part is the area where the rice is located, where more details are shown in VH and the spectrum is more continuous, but the rice part in VV is more striking

(4) Classify the image

Phenological parameters derived from the S-1 VH and VV backscatter time series were used as an input to a classifier in order to classify the input data into rice and other areas. The trained classifier is used to label every pixel in the image, then the whole image is classified. The algorithm adopted in this study is Random Forest (Richards, 2022). This classifier was chosen because the RF classifier provides performance improvements relative to the decision tree classifier (H.McNairn, 2012; Bazzi et al., 2019). Random Forest consists of a large number of individual decision trees that operate as an ensemble. Each individual tree in the random forest spits out a class prediction and the class with the most votes becomes our model's prediction (figure 13a) (Yiu, 2021). Because of the presence of multiple decision trees, the distribution of classification results, where each decision is considered as independent, is more uniform, meaning higher probability of success of the classification. The better classification performance is the characteristic of this method, but according to Richards (2022), there are requirements of data quality for the RF methodology to work well. The selected data should be sufficiently different from each other, and there should be enough data to ensure a good classifier. To guarantee this, the classification operation is repeated three times. Each time more than one-third of the training data that can be recognized in the map is captured.

(5) Results Processing and presentation

Before the final rice map was generated, some slight image operations were used to optimize the results.

1) The first thing to do is to limit the range. The study site for this research is in the irrigation area on both sides of the Nyagatare River. Scattered paddy fields of small farmers outside the marshes or pixels that were misclassified as paddy fields were removed by delineating the



Tally: Six 1s and Three 0s Prediction: 1

 $(a) \ Visualization \ of \ a \ random \ forest \ model \ making \ a \ prediction$



boundaries of the area. The extent of the study area was delineated manually on the map for each season, based on the agricultural RGB composite map of S-2 for the rice maturation time period, leaving only the dark green (considered as vegetation) parts of the river banks in the picture. This ensures that the final selected area must be vegetation. It is then required to consider whether to extend the study area based on the results and objectives. Considering the focus of the study, for the maps up to 2020, only the part downstream of the Muvumba irrigation dam will be shown. For the two years 2020 and 2021, since the area where it shows the concentration of rice results is in the Mimuli part more upstream, this part will be included in the map as well. Some scattered rice fields with no more than three pixels are not removed as noise because the edges of the unnormalized image are more natural.

- 2) The second step of the operation to be considered is the visualization of the results. Different batches of rice starting to be planted were the focus of this project, and the visible spectrum is divided into seven sections (rainbow colors) as visualization parameter to distinguish the different planting times. From red to purple, the planting time of the rice in its specific area is gradually delayed. Since there are five batches of rice in most seasons, most of the maps eventually show the color from red to blue (cyan is used for blue in this experiment). The planting order is accomplished by sequentially displaying a map of the rice that has been planted before a certain time point and the rice that was planted at that time. As shown in Figure 15, the red color indicates rice planted at the current time period (within half a month), and the cyan color indicates rice that has been planted before the current time period.
- 3) The final step is to create plots. These plots show the time series of backscattering for different batches of rice. Land use types classified as water and forest are also shown as comparisons in the plots, where other types of crops appear in the plots for 2020 and 2021. The colors of the lines used for different batches of rice are consistent with those on the rice map. Dashed lines are used for other land use types. The time of the horizontal coordinate corresponding to the lowest point of each line is the time when the batch of rice is considered to be seeded, and rice growth starts from this point. These lines have not been mathematically fitted so as to reflect the direct data.

4 Results

This chapter presents the results of this study, which consists of maps showing the location of rice, schematic maps of rice planting orders, and plots of time series analysis of backward scattering. Additional image results are in the appendix. The first five growing seasons, 2017 to the first half of 2019, showed a similar pattern, reflected in the areas where rice was planted and the timing of planting in different areas. This would be a reflection of the relative consistency of farmers' practices in growing rice. The results after this period differed considerably from the previous one, and were particularly different in 2020. The backscatter time series for other land use types for this year partially overlap with those for rice. Possible reasons for this result will be discussed in the next chapter.

4.1 Common patterns

Of the data being analyzed from 2017 to 2021, the data from 2017 to the first half of 2019 show relatively similar paddy patterns.

4.1.1 Relatively consistent patterns of rice area



Figure 14: Map of rice from the first half of 2017 to the first half of 2019 for a total of five growing seasons, with red to blue color blocks indicating planting batches from early to late.

The results of rice identification from the first half of 2017 to the first half of 2019 are shown in Figure 14. The different colors indicate different planting batches of rice in the same season, presented in the order of the visible spectrum, with red being the earliest planting batch of the season and blue being the latest. The earliest batch of rice planted during the long rainy season starts in late February, with planting ending in late April. The first batch of rice planted in the short rainy season starts in August, and the last batch ends in late October. The number of batches of rice planted per season (5 batches) and the locations where the rice was planted were almost the same. The locations where the same batches were planted were similar. The areas where rice was first planted are all located in the south-central part of the map. Rice planting ends in the middle of the area.

Looking from left to right from 2017 to 2019 in Figure 14, the maps show a decrease in red and an increase in blue, suggesting an overall gradual delay in rice planting. In addition to this, the

timing of planting on adjacent land tends to be the same, with a large amount of blue and green appearing in the graph during the first half of 2019. Such a shift indicates that the farmers' rice planting strategy started to change. This trend may serve as a basis for the change in results in the latter two years and can partially explain the change. However, these links will be explained in the discussion section.



2017 Late February

2017 Early March

2017 Late March

2017 Early April

2017 Late April

Figure 15: Map of rice from the late February of 2017 to the late April of 2017. The red area indicates rice planted within the period to which it belongs, and the cyan color indicates rice planted before this period.

Figure 15 shows a map of the rice planting order during the long rainy season in 2017. The graphs for the rest of the period are in the Appendix. The red color of the graph indicates rice that was started during that period, and the cyan color indicates rice that was already planted before that period. The graph shows that March 2017 was the time when the largest area of rice was planted. This time is usually the beginning of the long rainy season in Rwanda. Relatively less rice was planted in April. The distribution of rice planted in the same batch shown in the figure is relatively concentrated, and farmers living closer to each other may choose to refer to each other for planting strategies. However, the graph does not show why the farmers chose to plant rice in this order. This may be related to water resources.

4.1.2 Relatively consistent patterns of backscatter value

Figure 16 shows a time series plot of backscatter from 2017 to 2019, with different colors representing the same as in the previous map of rice distribution (figure 14). Higher backscatter values occur at the maturing stage of crop, when the crop is growing and water is shaded by the plants. Excluding the second half of 2019 (bottom right), more standard low-lying values are captured for each batch in each season, with minimum values ranging from -22 to -25. The maximum values are distributed in January and July, with values ranging from -15 to -16. Among all batches, the backscatter values for the first batch of each season are relatively less stable, with the occurrence of two minimum values and longer periods of duration of the minimum values. However, the overall trend is similar for each batch. The second half of 2019 starts to show a clear inconsistency. The minimum values of the captured batches are not sufficiently low and there is high variability between batches. Its trend is also different from other seasons of rice. This part will be discussed later.



Figure 16: Time series of rice field backscatter values for 2017 to 2019. The different colored lines indicate the rice planted at different periods. The dotted lines indicate the water and the forest respectively.

4.2 Different patterns

4.2.1 From late 2019 to 2021: harder to detect rice

As previously mentioned, the 2019 rice detection results were unexpected. Rice failed to be captured at the sowing stage, but pixels matching the backscattering pattern of rice were captured at the rice maturity stage. Similar to the situation in late 2019, the results for the area of rice fields and their cropping patterns derived in 2020 and 2021 are very different from the previous three years. It is nearly impossible to directly attribute the reasons for such a large change, and several potential causes are discussed in the next section. These potential impacts could be the effects of flooding from extreme rainfall, the effects of policies, or changes in farmers' choices. The next paragraph will describe specifically the differences that appear in the results.

For 2019(figure 17), the rice areas identified have little similarity to the previous three years. The areas identified as most intensively cultivated with rice in the previous years were hardly observed with rice in later two years. Instead, the wetlands near Colline Kiboga (another river in the direction upstream of the study area) appear to be full with rice. This area showed little rice in 2017-2019 and its not in the scope of the P8 program (Figure 18). Therefore it does not attract attention in the first three years. Figure 19 shows that the 2019 situation is also visible in 2020 and 2021. In addition to this, even if there is some similarity in the planting areas from the second half of 2019 to 2021, there is a large variation in the colors of the same areas in the graph, which indicates that there is no regularity in rice planting times.



Late 2019 (Short rainy season)

Figure 17: 2019 Results Map. A zoomed-out version is shown on the left. The concentrated area identified for rice cultivation can be seen on the map on the left in the Colline Kiboga area in the upstream direction of the Nyagatare area. In the Nyagatare region in the right map, no rice was identified on a large scale.

This result has been noted in the process of rice identification. Figure 20 shows a composite RGB image with late August 2019 as the middle image; Figure 20a is a zoomed out image, and Figure 20b is a screenshot of a zoomed-in image of the river. For the past few years, this period was the first rice planting time of the short rainy season: the combination of marshes and seedlings would have made the pixels representing the rice fields on the image to be magenta. However, there is almost no magenta in Figure 20a, and the various crops on this map are all a similar purple or pale pink color. When the image is scaled up, in Figure 20b, only scattered magenta appears along the river, which may represent low-lying lands along the river that were flooded at that time of year. Therefore, the lack of success in identifying rice in this region was not due to operational errors.

Even in the areas where rice was identified, its backscatter values were somewhat higher than the values of the previous years. In addition to this, the areas identified as other crops and mountainous (ridge topography) in the second half of 2020 show almost the same trend as the rice planted starting in late September. Figure 20 shows the same color on all the land during the experiment, thus the rice area does not show distinguishable blocks of color. Therefore, the lines labeled as other land use types were added to the time series analysis of rice in Figure 21. The dotted black line is the line for the other land use types. Since most of the site is agricultural land, this line also partially represents the backscatter time series of other crops, such as soybeans and maize.

In Figure 21, the remarkable situation is that only three batches of rice were identified in each season in 2020 and their trends were different even compared to 2021. The minimum values of the lines in the short rainy season of 2020 were even higher than -20. The maximum values also varied widely and fluctuated considerably around the maximum values. Another noteworthy aspect is the other land use types in 2020. The black dashed line is similar to or even coincides with the trend of rice in both seasons of 2020 (the black dashed line almost coincides with the solid orange line in the long rainy season), and the low-lying points of the black dashed line are observed in early March and late September 2020. This phenomenon occurred in that year only. In 2021, the backscatter values are relatively similar, although the location of rice cultivation is different from the previous years. The difference is higher for the rice planted after the middle of the two seasons in 2021, where the minimum value is higher than -20. However, the backscatter values for the other land types are different from rice, and the line tends to be straight.



2017 Long rainy season 2017 Short rainy season 2018 Long rainy season 2018 Short rainy season 2019 Long rainy season

Figure 18: Rice paddy map from 2017 to 2019 in a larger scope.



2020 Long rainy season 2020 Short rainy season 2021 Long rainy season 2021 Short rainy season

Figure 19: Rice paddy map from 2020 to 2021. Different colors indicate rice planted at different times.

Year-to-year and batch-to-batch time series of backscatter are compared in figure 22 and figure 23. Figure 22 shows rice planted in March and September, with similar trends for each of the five years. This shows a relatively stable growth of rice planted at that time of the year, coinciding with the beginning of the two rainy seasons in Rwanda. Both plots show that the backscatter in 2020 and 2021 are less variable over the rice growth cycle than in the previous three years. The dark blue line and the light blue line are almost always above the other lines, no matter what the planting stage or maturity stage of rice are.



(a) Almost no magenta in the image.



(b) Only the pixels by the river appear magenta

Figure 20: In late August 2019, which should have been the time when the first rice plantings began at the site, the rice can only barely be captured on the map (pixels showing a contrasting magenta color).



Figure 21: Time series of backscatter values for paddy fields from 2020 to 2021. The different colored lines indicate the rice planted at different periods. The dotted lines indicate water and forest respectively, but the black dashed lines indicate all other types of cover (mostly other kinds of crops, so they can also be seen as crops other than rice).



Figure 22: Backscatter from the same planting batch in different years.



Figure 23: Comparison of backscatter of all batches of rice grown in different years.

29

5 Discussion and Conclusion

This discussion section aims to explain in the results, and discuss limitations of the present experimental methodology. This project did not produce conclusive results, which is an inevitable consequence when purely remote monitoring experiments are conducted and field data are not available. In order to expand the significance of this study, the discussion section gives the fullest possible explanation of the causes of the phenomena in the results. Open conclusions enable further research. Flooding due to extreme rainfall, policy changes and farmers' own strategy changes are the more likely causes, and data support for this attribution includes meteorological data, news reports, government documents and articles written by local scholars. Limitations of this experimental approach include, but are not limited to, the lack of local data support for validation and the lack of long time series due to the lack of satellite data.

5.1 Limitations

No definitive results are derived in this study, which is a result of the absence of field data. The speculations provided in the above discussion can only be considered as possibilities.

If time permits, other products of remote sensing can be used to confirm or determine the extent of flooding. This thesis briefly describes this method to facilitate potential subsequent studies. Gandhi (2020) describes a method for mapping flooded areas inside the GEE using the same S1 product as this study. (OpenGeo Lab, 2022)'s approach uses SNAP in conjunction with QGIS to pre-process the S1 image set and present the results. This method is not limited to S1 products in GEE and has the flexibility to add pre-processing steps on demand. QGIS also provides a more convenient way to create maps and thus better display the results. However, since the products usually used to detect flooding and the principles are the same as the ones used to capture rice in this project, a partial overlap results are expected. For policy-related discussions, verification is much more difficult, and fieldwork is very helpful.

Moreover, this project covers a limited time period and does not reveal a long term pattern. Any changes that may occur are difficult to consider in terms of historical causes. This is limited by the development of satellite technology, which was launched late and earlier data were not available. Therefore future studies using satellites are to be expected to provide longer term result.

5.2 Potential explanations for big changes

5.2.1 Extreme weather

A reason for the inconsistency of the results in late 2019 and 2020 with earlier years could be flooding. News reports, rainfall data, values of backscatter and distribution areas support this statement.

- (1) Rwanda has increasingly experienced heavy rains, affecting most parts of the country. Rwanda was hit by floods at the end of 2019 due to huge rainfall, but the study area was not the worst hit at this time. Between 03 and 05 March 2020, extensive rainfall was recorded, leading to flooding and windstorms, in Gisagara and Gasabo districts in Kigali City, and Ngoma and Nyagatare districts in Eastern Province. Between 07 and 09 May 2020, more heavy rainfall caused flooding and landslides in seceral areas in Rwanda(Reliefweb, 2021; IFRC, 2021).
- (2) Combining the Rwanda monthly precipitation plots (figure 3), the monthly rainfall in March 2020 was much higher than other study periods, exceeding 350 mm/month, which indicates that Rwanda suffered from floods in both the long and short rainy seasons in 2020.



Figure 24: Farmers try to salvage some of their rice produce that had been washed away by floods in Nyagatare District (Nkurunziza, 2021)

- (3) Also able to support the argument are the time series plots of backscattering in 2020 and 2021 (figure 21). The other land use types distributed around the swampy rice fields (mostly non-rice crops), represented by the black dotted line, showed a minimum value in early March, which is usually considered to be the part of the low value captured by the rice cultivation stage that has a higher reflection value because it presents an agricultural flooded state. This scenario only occurs in 2020. The other land use types typically do not show much fluctuation in backscatter values throughout the year, as shown in figure 21 for 2021. The occurrence of a low value means that there is a large amount of water in the region. The explanation is consistent with flooding following extreme rainfall inundating large areas, including areas slightly higher in elevation than the rice fields. Accordingly, if almost all low-lying areas in the area were flooded, then the identified rice field may not be a rice field at that time, but an area that was inundated due to flooding, as shown in figure 24.
- (4) Farmland after the flood transit may not be able to support cultivation due to damage to infrastructure and changes in the nature of the land, requiring a recovery period, which is consistent with the result that only a limited amount of rice can be observed in 2021.

In summary, flooding is a likely cause. However, strictly speaking, the information only mentions that the province of Nyagatare was severely affected, without mentioning the specific areas of the province that were affected. So it is not possible to conclude that the study site was definitely affected. The use of other remote sensing products to detect the actual impact areas of the 2019 and 2020 floods could increase the credibility of this assertion.

5.2.2 Policy and farmer's choice

This section discusses policy-induced changes in rice fields. Policy determines the flow of funds and technical support, which partly determines the direction of agricultural development and cannot be ignored as a factor that also triggers a change in Rwanda's rice fields. The Rwandan government promotes swamp reclamation of rice fields to increase rice production, but in the face of Rwanda's agricultural transformation, i.e. land consolidation policy, growing rice is not necessarily the only option for local farmers.

The government of Rwanda recognizes the economic benefits provided by rice. According to its policy, the gradual expansion of rice fields, and even their extension in other areas, is predictable. Officially published yields of about 12 t/ha per year in Rwanda's rice fields also prove the usefulness of growing rice in Rwanda. Plans for rice development are supported by the National Agricultural

Policy (2019). Rice is also one of the eight major crops of the Crop Intensification Program (CIP), a program that promotes land consolidation. The National Rice Development Strategy-II (2020-2030) plan mentions the expansion of rice fields. At least 9,000 ha of new marshlands will be developed and at least 3,000 ha of existing marshland will be rehabilitated for rice cultivation by 2024. An additional 4,600 ha of new marshlands will be developed for rice cultivation by 2030 (Ministry of Agriculture and Animal Resources, 2021; Green, 2019; Nilsson, 2018).

Rwanda's agricultural reforms demand new strategies for farmers in terms of labour and planning. For some farmers, this policy has limited their options for farmland use. The reclaimed rice fields are owned by government cooperatives that provide the supporting infrastructure, and the farmers pay for the lease of the fields themselves. Influenced by commercial monocropping, farmers do not have the freedom to choose the source and price of seeds and fertilizers, nor do they have access to markets. Because of the uneven quality of field reclamation, farmers' income varies. The foreseeable option for farmers under the policy may be to withdraw from the rice growing program, because they cannot afford the capital and risk. As a result some farmers have quit rice cultivation after receiving negative agricultural consequences (Huggins, 2022). Similarly, the slightly increasing concentration of planted rice observed for the first three years may be the result of land consolidation, and the delay in planting rice could be attributed to the time constrains of farmers (Huggins, 2022).

In summary, policies set the rules of the game and farmers add or withdraw based on their individual capabilities.

5.2.3 Conclusion

The remote sensing approach provide successfully in identify rice. In the first 5 seasons, a consistent pattern of growth in Nyagatare could be identified. Planting times were slightly delayed our times, which is possibly related to government polices and associated farmers' options. In the second 5 seasons that were analysed, flooding seems to have disrupted the status, slightly changing rice growth patterns. It goes beyond this study to exactly explain observations further, nor speculate on the longer term effects of disruptions.

References

(n.d.).

- Africana Hills. (2022, 7). HYDRAULIC STRUCTURES-MUVUMBA P8 IRRIGATION SCHEME — NYAGATARE AGRICULTURE — IRRIGATION-RWANDA — WEIR. Retrieved from https://www.youtube.com/watch?v=ri9qc8b7FG8
- Baikirize, M. (2019, 9). Effect of irrigation on rice production in Rwanda a case study of Muvumba irrigation scheme nyagatare district. *Global Scientific Journals*, 7(9).
- Bazzi, H., Baghdadi, N., El Hajj, M., Zribi, M., Minh, D. H. T., Ndikumana, E., ... Belhouchette, H. (2019). Mapping Paddy Rice Using Sentinel-1 SAR Time Series in Camargue, France. *Remote Sensing*, 11(7), 887. doi: 10.3390/rs11070887
- Clauss, K., Ottinger, M., & Kuenzer, C. (2017). Mapping rice areas with Sentinel-1 time series and superpixel segmentation. *International Journal of Remote Sensing*, 39(5), 1399–1420. doi: 10.1080/01431161.2017.1404162
- Csorba, A., Szegi, T., Fodor, H., Bukombe, B., Uwiragiye, Y., Naramabuye, F. X., & Michéli, E. (2019). Characterization of rice agriculture in the Southern Province of Rwanda by means of microwave remote sensing. *Physics and Chemistry of the Earth, Parts A/B/C*, 112, 58–65. doi: 10.1016/j.pce.2019.02.002
- Esri. (n.d.). Analysis-ready Sentinel-1 GRD data generation—ArcGIS Pro Documentation. Retrieved from https://pro.arcgis.com/en/pro-app/latest/help/analysis/image-analyst/ analysis-ready-sentinel-1-grd-data-generation.htm
- European Space Agenc. (n.d.). Sentinel-1 SAR Technical Guide Ground Range Detected. Retrieved from https://sentinels.copernicus.eu/web/sentinel/technical-guides/ sentinel-1-sar/products-algorithms/level-1-algorithms/ground-range-detected

European Space Agency. (2015). Sentinel-2 user handbook. European Space Agency.

- Gandhi, U. (n.d.). End-to-End Google Earth Engine (Full Course Material). Retrieved from https://courses.spatialthoughts.com/end-to-end-gee.html#working-with -image-collections
- Gandhi, U. (2020, 12). Spatial Thoughts: Flood Mapping Part1 (Earth Engine Guided Project). Retrieved from https://www.youtube.com/watch?v=tT9iD9wRzUo
- GEARS. (n.d.). *GEARS Geospatial Ecology and Remote Sensing*. Retrieved from https://www.gears-lab.com/
- Global Forest Observations Initiative. (2018). A layman's interpretation guide to l-band and c-band synthetic aperture radar data (2nd ed.). GFOI.
- Google. (n.d.). Google Earth Engine. Retrieved from https://earthengine.google.com/
- Google Developers. (2022). Sentinel-1 Algorithms Google Earth Engine . Retrieved from https://developers.google.com/earth-engine/guides/sentinel1#:%7E:text=Sentinel% 2D1%20is%20a%20space,variety%20of%20polarizations%20and%20resolutions.
- Green, M. (2019, 8). Water management for agriculture under a changing climate: case study of Nyagatare watershed in Rwanda (Tech. Rep.). Retrieved from http://uu.diva-portal.org/ smash/record.jsf?pid=diva2%3A1344783&dswid=-3685

- H.McNairn, X., J.Shang. (2012, 1). Towards operational radar-only crop type classification: comparison of a traditional decision tree with a random forest classifier. *Canadian Journal of Remote Sensing*, 38(1), 60–68. Retrieved from http://dx.doi.org/10.5589/m12-012 doi: 10.5589/m12-012
- Huggins, C. (2022, 9). CURBS ON LAND RIGHTS IN RWANDA: THE 'BUNDLE OF RIGHTS' IN CONTEXT (Tech. Rep.).
- IFRC. (2021, 2). Rwanda Floods and Windstorms (Tech. Rep. No. MDRRW019).
- IRS Lab 2 GEARS Geospatial Ecology and Remote Sensing. (n.d.). Retrieved from https://
 www.gears-lab.com/intro_rs_lab2/
- Kuenzer, C., & Knauer, K. (2012). Remote sensing of rice crop areas. International Journal of Remote Sensing, 34(6), 2101–2139. doi: 10.1080/01431161.2012.738946
- Lhermitte, S. (2021a). *CIE4616: Remote sensing Big data Classification (bis)*. Retrieved from https://brightspace.tudelft.nl/d21/le/content/399347/viewContent/2497739/View
- Lhermitte, S. (2021b). *Remote sensing Big data: intro 2 GEE.* Retrieved from https:// brightspace.tudelft.nl/d21/le/content/399347/viewContent/2484874/View
- Lhermitte, S. (2021c). *Remote sensing Big data: introduction*. Retrieved from https:// brightspace.tudelft.nl/d21/le/content/399347/viewContent/2478149/View
- Minagri. (2022). Muvumba rice growers upbeat as they acquire tractors worth about frw 174 million. Retrieved from https://www.minagri.gov.rw/updates/news-details/muvumba-rice -growers-upbeat-as-they-acquire-tractors-worth-about-frw-174-million
- Ministry of Agriculture and Animal Resources. (2021, 7). National Rice Development Strategy(2021-2030) (Tech. Rep.).
- Ministry of Environment . (2018, 10). Muvumba Catchment Management Plan (2018-2024) (Tech. Rep.).
- Nguyen, D. B., & Wagner, W. (2017). European Rice Cropland Mapping with Sentinel-1 Data: The Mediterranean Region Case Study. *Water*, 9(6), 392. doi: 10.3390/w9060392
- Ni, R., Tian, J., Li, X., Yin, D., Li, J., Gong, H., ... Wu, D. (2021). An enhanced pixelbased phenological feature for accurate paddy rice mapping with Sentinel-2 imagery in Google Earth Engine. *ISPRS Journal of Photogrammetry and Remote Sensing*, 178, 282–296. doi: 10.1016/j.isprsjprs.2021.06.018
- Nilsson, P. (2018). The Role of Land Use Consolidation in Improving Crop Yields among Farm Households in Rwanda. The Journal of Development Studies, 55(8), 1726–1740. doi: 10.1080/ 00220388.2018.1520217
- Niyonkuru, R., M, S., Pande, K, S., A, H., & C, K. (2020). ASSESSING THE SUITABILITY OF RICE CULTIVATION IN MUVUMBA p-8 MARSHLAND OF RWANDA USING SOIL PROPERTIES. Rwanda journal of agricultural sciences, 2(1). Retrieved from https://www .ajol.info/index.php/rjeas/article/view/200852
- Nkurunziza, M. (2021). Economic impact of disasters in rwanda over the past 10 months. Retrieved from https://www.newtimes.co.rw/business/economic-impact-disasters-rwanda -over-past-10-months

- OpenGeo Lab. (2022, 04). Google Earth Engine Rice/Paddy Crop Classification using Sentinel-1 SAR data. Retrieved from https://www.youtube.com/watch?v=Ex544uYJRYw
- Ranjan, A. K., & Parida, B. R. (2020). Predicting paddy yield at spatial scale using optical and Synthetic Aperture Radar (SAR) based satellite data in conjunction with field-based Crop Cutting Experiment (CCE) data. International Journal of Remote Sensing, 42(6), 2046–2071. doi: 10.1080/01431161.2020.1851063
- Reliefweb. (2021, 02). Rwanda: Floods Dec 2019. Retrieved from https://reliefweb.int/ disaster/fl-2019-000170-rwa
- REMA. (2011). Atlas of Rwanda's Changing Environment Implications for Climate Change Resilience (Tech. Rep.). Retrieved from https://wedocs.unep.org/20.500.11822/9539
- Richards, J. (2022). Remote Sensing Digital Image Analysis: An Introduction (5th ed.). Springer Verlag.
- Sinergise, S.-H. B. (n.d.). Simple RGB Composites (Sentinel-2). Retrieved from https://custom -scripts.sentinel-hub.com/custom-scripts/sentinel-2/composites/
- Son, N.-T., Chen, C.-F., Chen, C.-R., & Guo, H.-Y. (2020). Classification of multitemporal Sentinel-2 data for field-level monitoring of rice cropping practices in Taiwan. Advances in Space Research, 65(8), 1910–1921. doi: 10.1016/j.asr.2020.01.028
- Steele-Dunne, S. C., McNairn, H., Monsivais-Huertero, A., Judge, J., Liu, P.-W., & Papathanassiou, K. (2017). Radar Remote Sensing of Agricultural Canopies: A Review. *IEEE Journal* of Selected Topics in Applied Earth Observations and Remote Sensing, 10(5), 2249–2273. doi: 10.1109/jstars.2016.2639043
- Synthetic aperture radar earthdata. (n.d.). Retrieved from https://earthdata.nasa.gov/ learn/backgrounders/what-is-sar
- Veloso, A., Mermoz, S., Bouvet, A., Le Toan, T., Planells, M., Dejoux, J.-F., & Ceschia, E. (2017). Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications. *Remote Sensing of Environment*, 199, 415–426. doi: 10.1016/ j.rse.2017.07.015
- Weichun Zhang, W. W. L. Z. W. J., Hongbin Liu. (2020, 5). Mapping Rice Paddy Based on Machine Learning with Sentinel-2 Multi-Temporal Data: Model Comparison and Transferability. *remote* sensing. doi: 10.3390/rs12101620
- WMO. (2022). world weather information service. Retrieved from https://worldweather.wmo .int/en/city.html?cityId=254
- Yiu, T. (2021, 12). Understanding Random Forest Towards Data Science. Retrieved from https://towardsdatascience.com/understanding-random-forest-58381e0602d2
- Zhang, M., Lin, H., Wang, G., Sun, H., & Fu, J. (2018). Mapping Paddy Rice Using a Convolutional Neural Network (CNN) with Landsat 8 Datasets in the Dongting Lake Area, China. *Remote Sensing*, 10(11), 1840. doi: 10.3390/rs10111840

I Code link

The addresses of all codes used in this project are listed below. Code list for rice identification:

- https://code.earthengine.google.com/?scriptPath=users\%2Fjiangshilian\%2Fseeworld\%3Abishe1_201901
- https://code.earthengine.google.com/?scriptPath=users\%2Fjiangshilian\%2Fseeworld\%3Abishe2_201902
- https://code.earthengine.google.com/?scriptPath=users\%2Fjiangshilian\%2Fseeworld\%3Abishe3_201801
- https://code.earthengine.google.com/?scriptPath=users\%2Fjiangshilian\%2Fseeworld\%3Abishe4_201802
- https://code.earthengine.google.com/?scriptPath=users\%2Fjiangshilian\%2Fseeworld\%3Abishe5_201701
- https://code.earthengine.google.com/?scriptPath=users\%2Fjiangshilian\%2Fseeworld\%3Abishe6_201702
- https://code.earthengine.google.com/?scriptPath=users\%2Fjiangshilian\%2Fseeworld\%3Abishe7_202001
- https://code.earthengine.google.com/?scriptPath=users\%2Fjiangshilian\%2Fseeworld\%3Abishe8_202002
- https://code.earthengine.google.com/?scriptPath=users\%2Fjiangshilian\%2Fseeworld\%3Abishe9_202102
- https://code.earthengine.google.com/?scriptPath=users\%2Fjiangshilian\%2Fseeworld\%3Abishe10_202101

Code list for land classification:

 https://code.earthengine.google.com/?scriptPath=users\%2Fjiangshilian\%2Fseeworld\%3Aclassification\%203\ %20(copy)

II Less important or abandoned methods

• Methodology exploration: Test if Sentinel-2 alone can identify rice fields

Sentinel-2 Multispectral Instrument (MSI) is one of the passive optical image sources from the European Space Agency, and the characteristics of this mission are wide-swath, high-resolution, multi-spectral imaging. The interpreting method of this optical image is similar to interpreting a photograph (*Synthetic Aperture Radar — Earthdata*, n.d.). Optical sensors collect data in the visible, near-infrared, and short-wave infrared portions of the electromagnetic spectrum (European Space Agency, 2015). The launch of this satellite took place on June 23, 2015, so the time of this study will be limited by this product.

The reflectance characteristics of rice in optical remote sensing images also make medium-resolution optical images one of the methods used to detect rice. Authors Ni et al. (2021); Son et al. (2020) explored the potential of tracking and mapping rice based on rice phenological characteristics using only Sentinel-2 images. Both experiments successfully identified rice simply and efficiently. In this project, however, using Sentinel-2 was less successful. The results of rice identification using this method are given in the fourth section of this study, but their accuracy does not meet the requirements of this project. Therefore, the Sentinel 2 pictures will not be used as the most dominant information in the subsequent sections of this project.

The method of making decision trees to identify rice fields using parameters derived from optical images (Sentinel-2) was mentioned in several studies (Kuenzer & Knauer, 2012; Weichun Zhang, 2020; Son et al., 2020; Ni et al., 2021) and tested to be effective. Therefore this method was tried during the exploratory period of this study. A land classification was performed through the use of S2 as a preliminary investigation into the land use of the study site. To verify the accuracy of this method, both methods using only S2 and NDVI, and also additionally using LSWI and EVI, were tested in this study.

The decision tree references the study of Son et al. (2020). The land was identified as rice field when EVI was greater than or equal to 0.5 and less than 0.7, LSWI was greater than 0 and less than 0.5, and NDVI was less than 0.64.

Supervised classification is also the method used for the first land classification method, and its brief step-by-step description is listed below, with the first five steps for method 1 and an additional sixth step added for method 2. Figure 25 shows the location of some of the training data. Figure 26 shows the land classification map of Nyagatare region obtained using method I. Rice in the area is grown in the lowlands along the river, and this part of the area is potentially swampy. This part of the farmland is separated from the city by a narrow strip of forest. Such a result is consistent with the study of Green (2019). The overall accuracy is higher than 90% (table 1).

Figure 27 shows the two results obtained using different classification trees and only the part classified as agricultural land (rice fields) is shown. The left figure shows the results of supervised classification by selecting training data to train the classifier (the dark green part is the rice field). The right figure shows the filtered part based on the decision tree from the literature (white part).

The overall accuracy is high, but the results are not ideal. The red circled parts in both figures are ridge-like types that are misclassified as rice fields. This misclassification can be explained by the study of Zhang et al. (2018) in Bazzi et al. (2019). In their study that rice and other crops are often misclassified when using only the spectral information of optical images. This shows that the rice fields in the study area cannot be effectively identified using Sentinel-2 alone and therefore is not used as the main method for this project.

Steps:

1) Filter the area of interest:

Using the Second-Level Administrative Units from FAO to select the area of interest by specifying the area name, Nyagatare.

2) Filter sentinel-2 image collection

Filter the image collection by the number of clouds, area of interest and the time (March 1, 2021 to the end of 2021).

3) NDVI mapping and unit specification and training data selection

Add NDVI using the map function and normalize its units. Add feature collections, classify into land use types of interest (agricultural, water, urban, forest and others) and select representative points and representative geometries.

4) Add the required bands and build the classifier

5) Classification is performed using CART and random forests, respectively, and points are reselected for validation and accuracy is calculated.

6) In addition to NDVI, LSWI and EVI indicators are added for classification to compare the result maps showing only agricultural areas.



Figure 25: Classification of land into agricultural, water, urban, forest and bareland (other) based on the type of land of interest; Selection of representative points and representative geometries (training data collecting)

In order to estimate the time of the start of rice cultivation, time series plots of NDVI (figure 28) were produced for rice fields from 2017 to 2019 (the results of Method 2 used). Due to the cloud cover, the map shows significant data missing in different years, which is an unavoidable drawback of optical images. However, the map can still give some valid information. In combination with the figure4, it can be found that NDVI, an indicator showing status of vegetation, has a clear positive correlation with rainfall. When it comes to transplanting time, the first season (season A) starts around late February and the second season (season B)in early August. This can be used as a reference for subsequent selection of time periods for remote sensing images.

• Discarded methods: using Sentinel-1 VH, VV and their mathematically fitted parameters as decision trees for identifying rice fields

This study decided to incorporate radar images after an initial exploration of using only optical images.Prior to understanding the RGB synthesis method, mathematical fitting parameters of the



Figure 26: Land classification map of study sites using randomly selected points combined with machine learning methods

Classifier	Overall accuracy	Cohen's kappa
Cart	0.900	0.822
Random Forest	0.933	0.881

Table 1: Accuracy of two different machine learning methods for land classification

Sentinel I time series were used to identify the decision tree of rice, where both VH and VV are used. The principle and steps of this method are mentioned in detail in the literature of Bazzi et al. (2019). It is verified that this method can be used to identify rice fields to obtain an overall accuracy of 96%.

The principle of this method is that the temporal behavior of the Sentinel 1 backscattering coefficient of rice is significantly different from that of other crops at VV, VH as well as the ratio of VV/VH polarizations. To standardize their temporal behavior, metrics are calculated for use in classification. Gaussian fits and variance fitting are performed for the time series of VV/VH, and linear fits are performed for the VH signal. Finally a decision tree made from the above metrics is used for classification. After this step is completed, the implications of the water resources aspects of all signal responses are then analyzed. The workflow is showed in figure 29.

This method was abandoned mainly because it was time-consuming and difficult. There are also more steps and techniques to be used. This was not in line with the time schedule of this project and the aim of finding a simple and efficient method.



Figure 27: Two different methods were used to try to identify rice fields



Figure 28: NDVI in 2017, 2018 and 2019 in a rice field that drived from Sentinel-2



Figure 29: work flow of the discard method

III Result maps and plots

Result maps and plots



2017 Late August

2017 Early September

2017 Late September

2017 Early October

2017 Late October

Figure 30: Map of rice from the late August of 2017 to the late October of 2017. The red area indicates rice planted within the time period to which it belongs, and the cyan color indicates rice planted before this time period.



2018 Late February

2018 Early March

2018 Late March

2018 Early April

2018 Late April

Figure 31: Map of rice from the late February of 2018 to the late April of 2018. The red area indicates rice planted within the time period to which it belongs, and the cyan color indicates rice planted before this time period.



Figure 32: Map of rice from the late August of 2018 to the late October of 2018. The red area indicates rice planted within the time period to which it belongs, and the cyan color indicates rice planted before this time period.



2019 Late February

2019 Early March

2019 Late March

2019 Early April

2019 Late April

Figure 33: Map of rice from the late February of 2019 to the late April of 2019. The red area indicates rice planted within the time period to which it belongs, and the cyan color indicates rice planted before this time period.