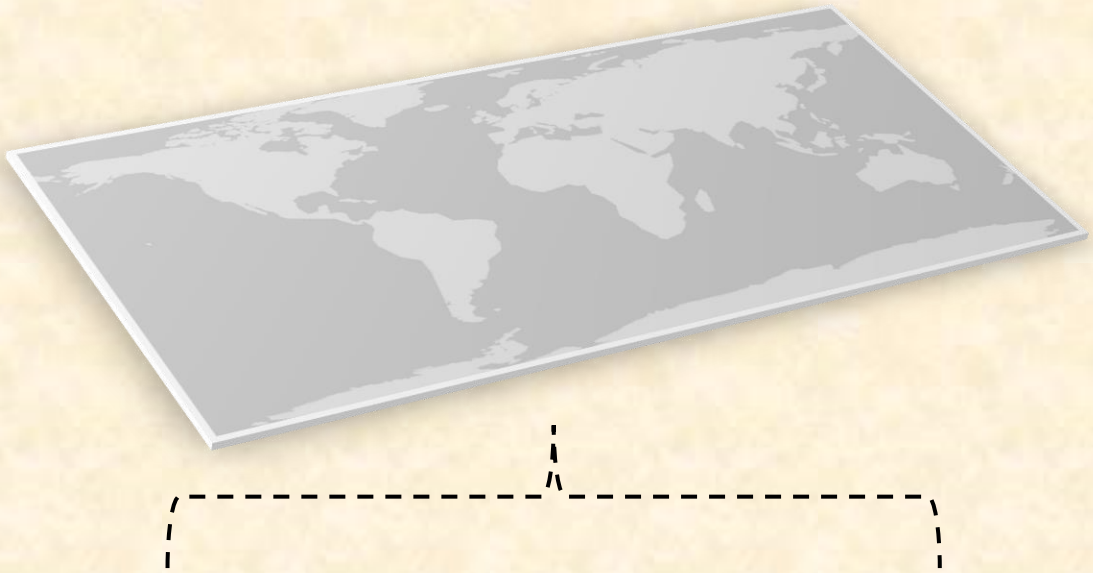

An Integrated Framework for Global Crop Mapping

A case study on mapping plantation and smallholder banana

Victor Bonekamp

March 2022

SGL-80436

An Integrated Framework for Global Crop Mapping

A case study on mapping plantation and smallholder banana

V. (Victor) Bonekamp

971026-093-050

Master thesis

Earth and Environment: Soil Geography and Earth Surface Dynamics

Soil Geography and Landscape (SGL) group

Wageningen University and Research

Supervisors:

Dr. ir. J.J. (Jetse) Stoorvogel

Soil Geography and Landscape group

Wageningen University and Research

L.A.Q.M. (Nard) Onderwater

Soil Geography and Landscape group

Wageningen University and Research

Dr. J.M. (Jeroen) Schoorl

Soil Geography and Landscape group

Wageningen University and Research

Abstract

Global crop maps are essential tools for agricultural policy making. Opportunities for global crop mapping have emerged due to developments in computer science, GIS and remote sensing. This study presents an integrated framework for global crop mapping that combines different approaches. The framework was specifically applied to banana at a fifteen arcsec (~ 500 m) resolution. It was argued that the coexistence of different production systems cannot be neglected. This is why a clear distinction between large-scale and smallholder production systems was made. Plantation banana was detected based on polarimetric radar remote sensing data. Smallholder banana was mapped through the disaggregation of national production statistics with fuzzy land suitability. Outcomes were validated with satellite imagery and evaluated with existing global banana maps. The radar analysis showed that part of the global banana surface (6.5%) can be mapped with remote sensing. The disaggregation with fuzzy suitability gave mixed results for mapping smallholder banana, depending on the specific country. The resulting map visualized the general distribution of banana at a finer spatial resolution than existing maps. The study meanwhile illustrates an approach for global crop mapping based on available spatial data. Moreover, the study indicates how global crop mapping can benefit from the integration of remote sensing, production statistics and land suitability concepts.

Acknowledgements

I would like to take this opportunity to thank a number of people for their support during my thesis. First of all, I have to mention my supervisor Jetse Stoorvogel. It's Jetse who introduced me to the world of global crop mapping in the first place. Our meetings were always fruitful and stimulated me to get the most out of this thesis project. Jetse's affinity with the topic has definitely enabled me to bring this study to a higher level. I would also like to say a few words to my second supervisor Nard Onderwater. I could always reach out to him for advice and feedback on my work. Nard was always ready to share his insights with me to improve the results of this study. It has been an honor to build further on his previous work. I could not have wished for more engaged supervisors.

To conclude, I would like to thank all professors of Wageningen University who introduced me to the worlds of geography, GIS and remote sensing. In particular, great thanks to Adugna Mullissa for his introduction to polarimetric SAR remote sensing. This earth observation technique turned out extremely useful for my thesis project. In addition, thanks to Dainius Masiliunas for the inspiring course on geo-scripting. I could never have done this research without the skills that I learned during his course.

Table of contents

List of figures

List of tables

1. Introduction

1.1. Global crop mapping

Future global agriculture will face major challenges. Agricultural demands will increase substantially under influence of economic growth and population growth. At the same time, climate change will hamper the agricultural production (FAO, 2017). This will cause an increased pressure on both global food security and the environment (FAO et al., 2020). Another development is the ongoing dispersion of crop pests and pathogens around the world (Bebber et al., 2014). This has increased the risk of crop disease pandemics and associated crop failures. All these challenges call for informed agricultural decision making across different scales. Effective agricultural policies require insights on crop production. The spatial distribution of crops is consequently crucial information (You et al., 2014). In this context, global crop maps are essential tools to manage the projected agricultural challenges in the years to come.

Global crop mapping made significant progress since the advancements in GIS and remote sensing technology at the end of the twentieth century (Giri, 2012). The original focus in global crop mapping was more on landcover classes than on specific crops (Mora et al., 2014). All arable land was often grouped together in a single landcover class. Defries and Townshend (1994) produced one of the first global landcover maps based on spectral characteristics. Not much later, Ramankutty and Foley (1998) mapped the global distribution of croplands with satellite and census data at a five arcmin resolution. Similar methodologies turned out successful for mapping major crop types (Still et al., 2003, Leff et al., 2004). Monfreda et al. (2008) thereafter distinguished 175 individual crops at a five arcmin resolution by using crop cover data from Ramankutty et al. (2008). This study can be considered the birthplace of contemporary global crop mapping.

Crops are cultivated in a variety of production systems around the world. Commonly a distinction is made between large-scale and smallholder systems (Lansing et al., 2008, Brüntrup et al., 2018, Tiffen and Mortimore, 1990). Those systems differ from each other in several aspects. Most obviously, large-scale systems are large in size while smallholder systems typically don not exceed a few hectares (Fox and Castella, 2013). Another characteristic of large-scale systems are large homogenous cropping surfaces in the form of monocultures. Traditional smallholder systems, however, are often polycultures such as intercropping systems (Ntumngia, 2010). Finally, large landowners often have the means to regulate environmental conditions while most smallholder farmers are resource-constrained (Rurinda, 2014, Hebbar et al., 2016). Most smallholder systems therefore rely on the prevailing natural conditions. The challenge of global crop mapping is to cover the full spectrum of production systems, despite their different characteristics.

1.2. Global crop mapping approaches

The spatial distribution of crops can be mapped in various ways. Local and regional studies often rely on local data acquisition and data collection. Field work, aerial photography and cadastral maps are common data sources in such studies (Ahani and Noshadi, 2019, Jurado-Expósito et al., 2019, Handique et al., 2017, Gomez Selvaraj et al., 2020). The opportunities for global crop mapping are however more

limited. Field work at the global scale is simply not possible. Furthermore, creating global crop maps by compiling national crop maps is constrained. Only some countries have such maps at their disposal and the willingness to share data is generally low. Global crop mapping therefore asks for a generic approach that employs existing and available data sources.

Crop mapping can be done with direct and indirect mapping approaches. The high quality of remote sensing products nowadays provides opportunities for the direct detection of specific crops. Different products provide data at various spatial, temporal and spectral resolutions. Multispectral sensors provide reflectance data for multiple wavelengths. The full spectral signal is often transformed to spectral indices (Lu and Weng, 2007). A well-known index for crop mapping is the normalized difference vegetation index (NDVI). Several studies already used seasonal fluctuations in NDVI and other indices to map crops at a regional level (Heupel et al., 2018, Peña-Barragán et al., 2011, Zhang et al., 2011). Another common type of remote sensing is radar. Radar sensors actively transmit microwaves and measure the returned signal. Radar data can be acquired during both day and night-time. In addition, radar signals are insensitive to interference by clouds (Chen et al., 2018, Lu and Weng, 2007).

There are several approaches for indirect crop mapping. Those approaches are not mutually exclusive and often complementary to remote sensing (Zhang et al., 2017). A good example is the use of statistical data. Whereas crop maps are scarce, agricultural census data is widely available (Kluger et al., 2021). Several studies used agricultural census data as a starting point for crop mapping at a global scale (You et al., 2014, Monfreda et al., 2008, Onderwater, 2020). Another interesting approach involves the notion of land suitability. Land suitability concepts originate from agroecology and establish a link between crop occurrence and environmental conditions. A conventional application of land suitability for spatial planning is the FAO framework for land evaluation (FAO, 1976). In addition, several land suitability studies used fuzzy classification techniques as an alternative framework (Atijosan et al., 2015, Ali et al., 2010, Joss et al., 2008). Land suitability concepts have also been used to map land use and land use change explicitly. The CLUE modelling framework by Verburg et al. (1999) provides an early example of such an approach. Finally, machine learning techniques are increasingly used for crop mapping. A variety of crop classification algorithms exists. For instance, Sonobe et al. (2014) applied Random Forest and CART algorithms to map crop types in Japan. In addition, Miettinen and Liew (2011) separated plantation crops in Southeast Asia by using statistical distance techniques.

1.3. The demand for a global banana map

Banana is an important crop for global agriculture. Banana cultivation and consumption dominates the livelihoods of 400 million people around the world, especially in the tropics (García-Bastidas et al., 2019). Its production amounted to 117 million tonnes in 2019, making banana the fourth most important global food crop (FAO, 2020, Churchill, 2011). The majority of banana yields are used for local consumption, less than one fifth is exported to international markets (Evans et al., 2020). In other words, banana plays an important role in local food security and rural economies. Banana is a perennial crop that is traditionally cultivated in home gardens or multi-crop fields (Robinson and Galán Saúco, 2011). These traditional production systems are still found in smallholder agriculture. For example, Kenyan farmers often plant banana at a considerable part of their smallholder plots (Wahome et al., 2021). The average size of banana plantations in a country like Kenya is 0.32 hectares (Obaga and Mwaura, 2018). On the other hand, monoculture banana plantations larger than 100 hectares are found in export-oriented regions like Latin-America (Stoorvogel et al., 2004). It should be noted that

smallholder systems and banana plantations generally coexist in banana producing countries (Bijker, 2014).

The global banana production relies for 95% on cultivars with strong genetic ties. The single Cavendish cultivar even accounts for 50% of the global banana production as well as nearly all export production (Bakry and Horry, 2016, Ploetz et al., 2015b). This poor genetic diversity makes banana vulnerable to a variety of diseases (Jones, 2018). For instance, Xanthomonas bacterial wilt forms a serious threat for banana production in East and Central Africa (Ocimati et al., 2019). Also fungal leaf-spot diseases like black Sigatoka affect banana production at a global level (Jesus Júnior et al., 2008, Jones, 2018). Banana is currently severely threatened by Panama disease. This soil-borne disease is caused by the fungus Fusarium Tropical Race 4 (Foc TR4) (Ploetz, 2015a). Predecessors of Foc TR4 had already devastating effects in the past (Ordóñez R et al., 2018). Originating from Indonesia (Maryani et al., 2019), Foc TR4 has recently spread all over the world (Zheng et al., 2018, O'Neill et al., 2016, Ploetz et al., 2015a, Thangavelu et al., 2019, Viljoen et al., 2020, García-Bastidas et al., 2020). The disease has severe consequences for banana production given that effective management options do not exist (Ploetz, 2015b). This calls for adequate measures to prevent the spread of the disease.

Banana diseases spread through landscapes via different pathways. There is often a clear relation between the disease and the dominant spreading mechanism (Blomme et al., 2013, Mengesha et al., 2018). The spatial distribution of host plants is certainly a crucial factor (Otten and Gilligan, 2006). In this context, Bijker (2014) applied a multi-scale approach to identify risk factors for the spread of Foc TR4 in Costa Rica. An interesting observation was the abundancy of individual banana plants between large plantations. Those plants connected the Costa Rican banana landscape by serving as steppingstones for the spread of diseases. Her work underlines that plantations, smallholder cultivations and wild plants all have an impact on landscape connectivity. Knowing where banana grows is therefore essential for disease monitoring and risk evaluation. Despite its importance, however, no high-resolution global banana map is currently available.

1.4. Towards a global banana map

Monfreda et al. (2008) produced the first global banana map at a five arcmin resolution. The authors achieved this milestone by using global landcover data for the disaggregation of production statistics. More recently, Onderwater (2020) produced a global banana map at a thirty arcsec resolution. Onderwater increased the resolution substantially by using both remote sensing and land suitability for the disaggregation of production statistics. Both studies indicate the added value of census data for global crop mapping. At the same time, both studies took only limited advantage of the opportunities that remote sensing has to offer. Neither study used remote sensing to map banana through direct crop mapping. Furthermore, neither study incorporated the notion of coexisting production systems in their methodological frameworks. These observations indicate crucial focus points for further research.

Remote sensing has been used extensively for banana mapping at the regional level (Johansen et al., 2009, Handique et al., 2017, Gomez Selvaraj et al., 2020). Previous work on radar remote sensing deserves special attention. Xie et al. (2015) used polarimetric SAR to separate banana from rice, eucalyptus and sugarcane in southern China. Furthermore, Wang et al. (2010) used multitemporal SAR to distinguish banana from sugarcane, water and cities. Both studies linked the good separability of banana to the distinct shape of the plant. Both studies also mentioned the advantage of radar for crop

mapping in the cloudy tropics. The generic character of either method indicates good prospects for global crop mapping. Literature also shows interesting applications of land suitability for crop mapping. For instance, land suitability for banana was mapped through a fuzzy classification method in the Philippines (Salvacion et al., 2019, Salvacion, 2021). A similar approach could be useful for banana mapping at a global scale.

This study presents an integrated framework for global crop mapping. The framework is hereby applied to banana specifically. The study considers differences between production systems and includes a strong remote sensing component. The research expands the previous work on global banana mapping by Monfreda et al. (2008) and Onderwater (2020). In this context, a global banana map is created at a fifteen arcsec resolution.

2. Materials & Methods

2.1. General Framework

This study presents a methodological framework for global crop mapping (Figure 1). The framework consists of two methods that each focus on a different production system. The first method was developed to detect large-scale production systems through the analysis of remote sensing data. The second method was developed to predict the location of smallholder systems based on land suitability and census data. In this way, the framework disconnects the occurrence of large-scale systems from their biophysical environment. This decision reflects the considerable impact of management practices on growing conditions in large-scale systems. The general framework consists of the four steps below. The dashed line indicates that the results of step two are required as input for step three.

1. Preparation of input data
2. Mapping plantation banana
3. Mapping smallholder banana
4. Validation and evaluation

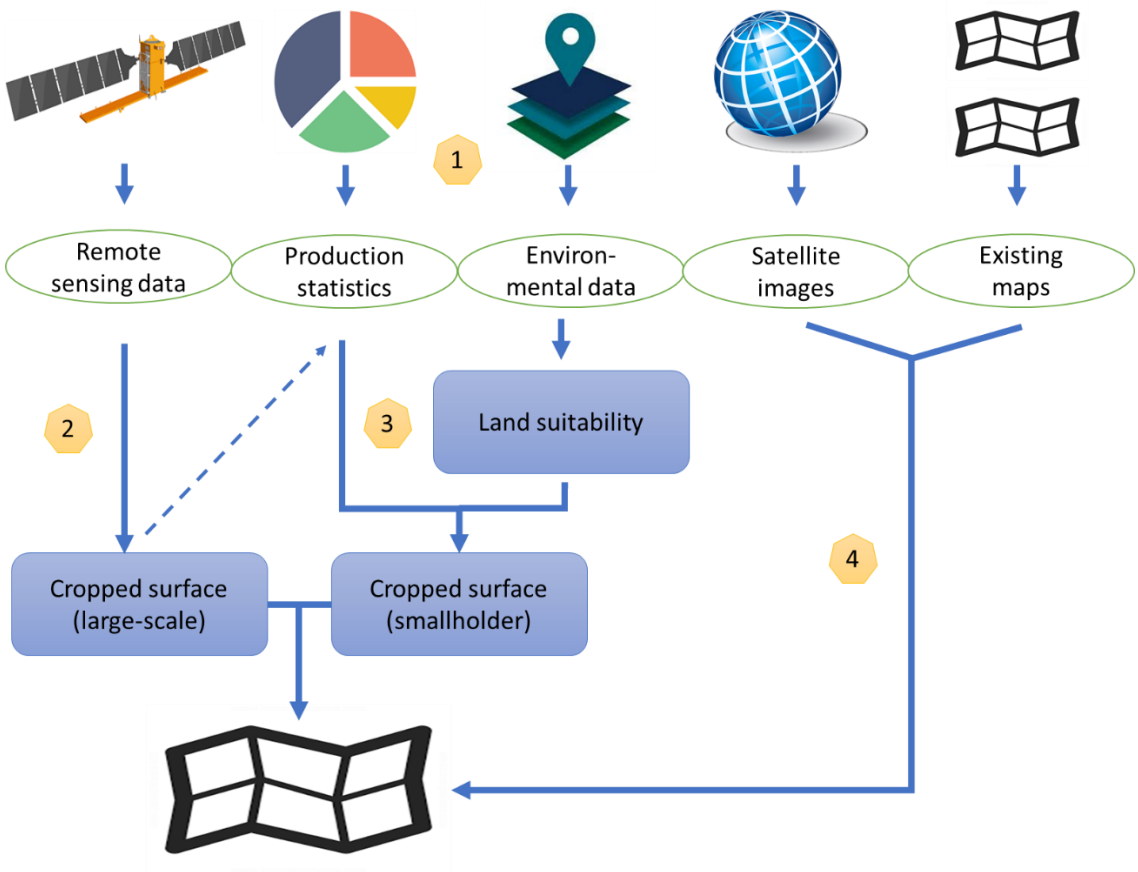

Figure 1: The general framework for global crop mapping. (The numbers in the Figure refer to the different steps in the text)

2.2. Materials

2.2.1. Data overview

Global radar, landcover and slope data were used for mapping large-scale systems. In addition, available statistics and land use data were collected for calibration purposes. Climate, soil and water table depth data were used to map land suitability. National production statistics were disaggregated to map smallholder banana based on the suitability results. Finally, two existing global banana maps were collected to allow for a comparison of methods and results. Table 1 provides an overview of the data that was used in this study. Each dataset is described in more detail in Chapter 2.2.2.

Table 1: Data overview. (See Annex A for data access)

Data	Details	Research component	Temporal extent	Spatial resolution	Source
Synthetic Aperture Radar (SAR)	Sentinel-1 product from ESA C-band, VV and VH polarization	Plantation banana	2018-2019	10 / 25 / 40 m (depending on scene)	(ESA, 2020)
	PALSAR-2 product from JAXA L-band, HH and HV polarization	Plantation banana	2018-2019	25 m	(Shimada et al., 2014)
Landcover	Copernicus global land service (CGLS) product	Plantation banana + Pre-processing	2018	100 m	(Buchhorn et al., 2020)
	GlobCover product	Plantation banana	2009	300 m	(Arino, 2010)
Elevation	NASADEM product	Plantation banana	2000	30 m	(NASA-JPL, 2020)
Calibration data	Costa Rican banana map	Plantation banana	2019	-	(Corbana, 2019)
	Mexican land use map	Plantation banana	2017	-	(INEGI, 2017)
	Philippine banana statistics (regional)	Plantation banana	2020	-	(Salvacion, 2020)
Climate	WorldClim V1 Bioclim product	Smallholder banana + Pre-processing	1960-1991	≈ 1000 m	(Hijmans et al., 2005)
Soil	SoilGrids product	Smallholder banana	-	250 m	(Poggio et al., 2021)
Water table depth	EartH2Observe product	Smallholder banana	-	≈ 1000 m	(Fan et al., 2013)
National production statistics	FAOSTAT product	Smallholder banana	2019	-	(FAO, 2021b)
Existing global banana maps	Global banana map	Plantation + Smallholder banana	2000	10000 m	(Monfreda et al., 2008)
	Global banana map	Plantation + Smallholder banana	2015	1000 m	(Onderwater, 2020)

2.2.2. Data collection and justification

Synthetic Aperture Radar (SAR)

An interesting radar technique is polarimetric synthetic aperture radar (SAR). Polarimetric SAR signals reflect the structure and geometry of ground targets (Lopez-Martinez and Lopez-Sanchez, 2017). In addition, they reflect several crop characteristics such as ground cover, biomass, moisture content and crop height (Srikanth et al., 2016). Polarimetric SAR was used in several regional crop mapping and landcover studies (Tao et al., 2017, Ullmann et al., 2017, Zakeri et al., 2017).

Polarimetric radar data from the Sentinel-1 and ALOS PALSAR-2 satellites were downloaded for 2018 and 2019. The temporal extent of two years ensured a global data coverage for both platforms. The two satellites provided complementary data because of their different polarization channels. Using both datasets yielded VV, VH, HH and HV backscatter data. In addition, the sensors of Sentinel-1 (C-band) and PALSAR-2 (L-band) measure in different parts of the electromagnetic spectrum. Radar signals are backscattered by elements that have dimensions similar to the wavelength of the signal (Le Toan, 2007). In other words, Sentinel-1 and PALSAR-2 sense different parts of ground objects (Figure 2). Combining both datasets consequently provides more information on vegetation and crop structures.

Different types of SAR products exist. A distinction is made between Single Look Complex (SLC) and Ground Range Detected (GRD) products (Holloway, 2018). However, the exact naming of products varies per satellite platform (Airbus, 2015). In short, SLC contains information on signal strength and phase while GRD contains signal strength only (Holloway, 2018). This makes that GRD products are significantly smaller in size than SLC products. This study used GRD products for this reason, allowing for data analysis on a global scale. The SAR images have been aggregated temporally for each polarization channel.

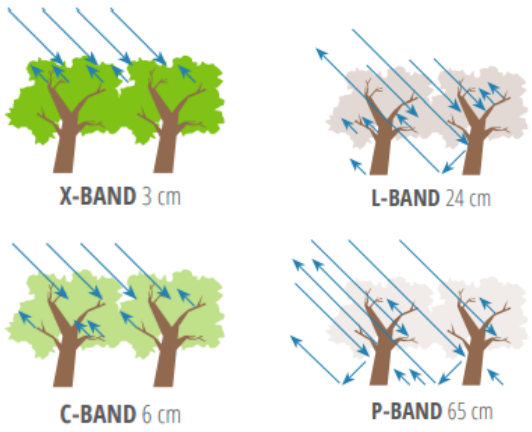

Figure 2: Penetration of radar signal into tree objects for different radar wavelengths (Sassan, 2019)

Landcover

The optimal landcover product would separate plantation agriculture from mixed cropping systems. Such a landcover product was however not available for this study. Instead, the usability of two common global landcover products was evaluated: GlobCover and Copernicus global land service (CGLS). A visual comparison with satellite imagery showed that GlobCover covers prominent banana

production regions by a few classes only. In this way, the potential banana area could be reduced more substantially compared to CGLS. GlobCover was therefore used as the main source of landcover data in this study. The 2009 product was considered representative for contemporary landcover. Besides GlobCover, urban cover data from CGLS was used as an additional input. This dataset provided information on urban cover fraction, complementary to the discrete GlobCover classes. Both GlobCover and CGLS have a finer resolution than the level of detail in this study.

Elevation

A large variety of Digital Elevation Models (DEM) exists. In this study, the NASADEM product was selected arbitrarily. NASADEM is an improved digital elevation model that combines data from the Shuttle Radar Topography Mission (SRTM) with auxiliary datasets (Buckley et al., 2020). Its spatial resolution of 30 m largely meets the requirements for this study. The NASADEM product was converted to a slope map for further analysis. This was done in GEE to benefit from the strong processing power that the platform offers. The NASADEM product was aggregated to a 500 m resolution, after which slope was calculated with a neighborhood of four grid cells. The resulting slope map displayed the general gradient in landscapes. Such a general slope map had sufficient detail for this study.

Calibration data

Two appropriate country maps were found for calibration purposes. The first map was a banana plantation map of north-east Costa Rica from the country's national banana cooperation (Corbana). The map delineates the majority of banana plantations in the region. The second map was a Mexican land use map from the national institute for statistics and geography (INEGI). The map separates agricultural land use classes based on water supply and the crop's life cycle. Finally, regional banana statistics of the Philippines were used as an additional source of calibration data.

Climate

Climate gradients normally arise at the landscape scale. This allowed for the use of climate data with a coarser spatial resolution than the fifteen arcsec output resolution. In addition, major climate trends occur at timescales beyond years or decades. The 2005 WorldClim product with a thirty arcsec resolution was therefore considered appropriate for this study. The WorldClim subproduct BioClim contains bioclimatic variables derived from monthly temperature and rainfall data. BioClim hereby offered biologically meaningful climate data that was required for this study.

Soil

Soil variability occurs at finer spatial scales than climate. The SoilGrids product was selected because its spatial resolution matches the requirements of this study. SoilGrids contains information on several textural, structural and chemical soil properties on a global level. The data is grouped into six vertical intervals, together covering a soil depth of two meters in total. In this study, variability within soil profiles has not been considered. Instead, average soil property values were calculated for the rootzone of banana. The idea behind this approach is that soil conditions affect banana growth most in the rootzone. Banana roots are concentrated in the top 30 cm of the soil (Robinson and Galán Saúco, 2011). The soil data has therefore been aggregated over this soil depth.

Water table depth

Banana requires well-drained soils (Robinson and Galán Saúco, 2011). In this study, water table depth was used as an indicator for drainage conditions. Data on water table depth with a 1000 m spatial resolution turned out available. The EartH2Observe product was selected because it provides water table data on a global level.

National production statistics

National production statistics summarize the agricultural production levels of countries. FAOSTAT provides an overview for all countries based on national surveys. The database includes production statistics for a large number of crops, including banana and plantain. The plantain statistics refer to the Musa paradisiaca species specifically (FAO, 2021d). The banana statistics cover all other species of the genus Musa that produce edible fruits. The distinction made by the FAO is however arbitrary. The Musa paradisiaca group has strong genetic ties with other Musa groups. Robinson and Galán Saúco (2011) therefore argue that the name Musa paradisiaca cannot be used to distinguish banana from plantain. Moreover, spatial interactions between banana and plantain exist. The work of Bijker (2014) for example shows how both groups influence landscape connectivity. It was argued that the inclusion of all Musa species yields the most informative results within this study. The banana and plantain statistics were therefore merged for mapping smallholder banana.

The FAOSTAT database contains three production indicators: harvested area, yield and production quantity. This study focused more on the global banana distribution than on the global banana production. This is why harvested area was selected as the most appropriate production indicator. Harvested area statistics for perennial crops like banana are generally equivalent to the total planted area (FAO, 2021d). Most recent statistics were available for the year 2019. The statistics were made spatially explicit by joining the national production statistics to the QGIS World Map. Eleven regions were thereafter removed from the QGIS world map to reduce processing time. Among those regions were features with erroneous geometry or zero production statistics. In addition, some remote parts of American and French territory were excluded. The excluded areas were either very small or too cold for banana.

Existing global banana maps

Two global banana maps were collected for validation purposes. First of all, Monfreda et al. (2008) mapped the global banana distribution for the year 2000 at a five arcmin resolution. Secondly, Onderwater (2020) mapped the global banana distribution for the year 2015 at a thirty arcsec resolution. Both studies applied a disaggregation method based on production statistics.

2.2.3. Data handling and pre-processing

Open-source software was used for data processing, including QGIS, Python and R software. Some models have been executed in batch mode to avoid computer memory problems. This implied that models were run for different parts of the world separately and merged afterwards. For the same reason, intermediate results were rescaled to Byte data type whenever possible. Most datasets in Table 1 were accessed through Google Earth Engine (GEE). GEE is a cloud-based platform that contains

a wide variety of geospatial data. The Python API was used to export the GEE data to GeoTIFF format. Pre-processing was performed in GEE before exporting, unless stated differently. The different pre-processing steps are discussed below.

Temperature mask

Banana has a low tolerance to cold climate conditions. A temperature mask was therefore applied to filter out locations that are too cold for banana. The crop is severely damaged when temperatures drop below 6 °C (Robinson and Galán Saúco, 2011). Banana occurrence at locations with such temperatures during part of the year is very unlikely. Regions with average cold quarter temperatures below 6 °C were consequently excluded from further analysis.

Landcover mask

A general landcover mask was applied to the data. The purpose of the landcover mask was to exclude particular landcover types from further analysis. A distinction was made between large-scale systems and smallholder systems. For large-scale systems, built-up areas and water bodies were masked out to improve the separability of SAR signals. Additional landcover filters for mapping plantation banana are described in Chapter 2.3. For smallholder systems, water bodies were excluded to ensure that banana was disaggregated over land surface only. Contrary to large-scale systems, built-up area was not masked out while smallholder banana is often cultivated in home gardens.

Further pre-processing

Data subsets were merged if required to obtain inputs with a global coverage. In addition, all datasets were aligned in terms of coordinate reference system, extent and resolution. After alignment all datasets were in EPSG:4326 with a global coverage and a fifteen arcsec spatial resolution. It was decided to work in EPSG:4326 because most datasets already had this coordinate reference system. The alignment implied that datasets with finer resolutions were aggregated to fifteen arcsec. Analogously, datasets with coarser resolutions were interpolated to fifteen arcsec. The alignment of GEE datasets was done automatically when exporting to GeoTIFF format. The alignment with other datasets was done in QGIS manually.

2.3. Mapping plantation banana

2.3.1. Overview

A method was developed to map large banana surfaces with remote sensing. A schematic overview of the method is provided in Figure 3. The method consists of three steps that are explained in subchapters 2.3.2. - 2.3.4. respectively.

1. **Classification of satellite imagery**

   Synthetic Aperture Radar (SAR) data was classified through a multivariate statistical distance analysis. The satellite imagery was hereby compared to characteristic backscatter signals of banana. This step yielded a fuzzy banana map, indicating the similarity between grid cell values and characteristic values for banana.

2. **Filtering of results**

   Areas where banana occurrence is unlikely based on supplementary data were filtered out from the dataset. Filtering was done based on four characteristics, namely: landcover, urban cover fraction, slope and the SAR signal of the VH polarization channel. The filtering improved the results by excluding areas that produced significant noise.

3. **Conversion to plantation banana map**

   The filtered map was thereafter converted to a final banana map. A calibration procedure was developed to do this conversion as unambiguously as possible.

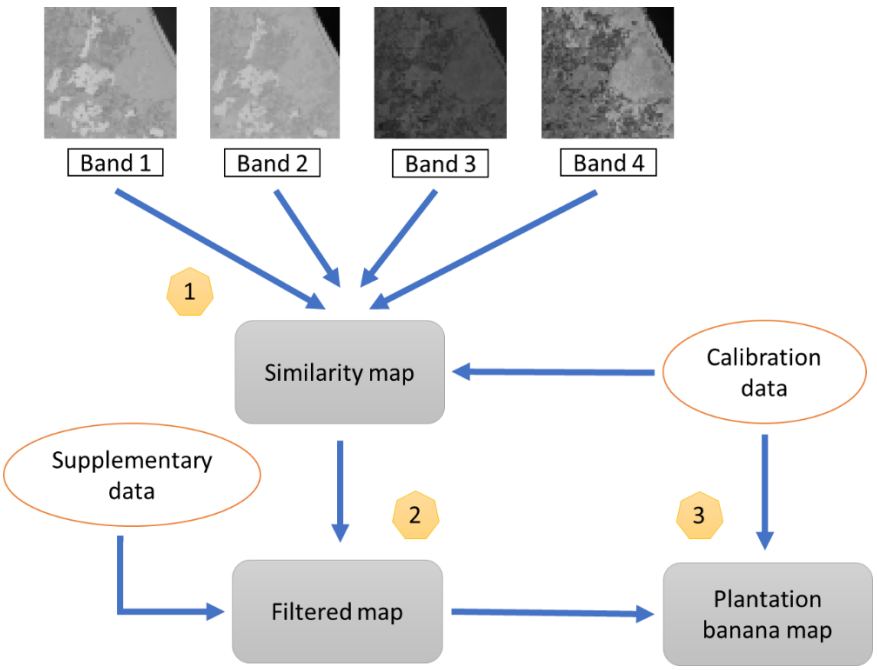

Figure 3: The method for mapping plantation banana.
(The numbers in the Figure refer to the different steps in the text)

2.3.2. Classification of satellite imagery

A comparison was made between SAR data and characteristic backscatter values of banana. All four polarization channels were involved in this comparison to make optimal use of the data. This required a multivariate approach that allowed for the combination of four inputs. Similarity between the SAR data and characteristic banana values was synthesized into a single value. In literature, such a measure of similarity is often called a distance (Brereton, 2011, Xu and Tian, 2015). Analogously, algorithms to calculate distances are referred to as distance functions. In the context of this study the following applied: the smaller the distance, the more likely banana is present.

Several multivariate distance functions exist (Xu and Tian, 2015), but most common are the Euclidean distance and Mahalanobis distance (De Maesschalck et al., 2000). Contrary to the Euclidean distance, the Mahalanobis distance accounts for correlation between variables and differences in measurement scale (McLachlan, 1999). The Mahalanobis distance function was therefore selected to measure similarity in this study. Conveniently, Harasymczuk (2016) developed a QGIS plugin for calculating the Mahalanobis distance. This plugin was used after modification and rewriting to a QGIS tool compatible with the QGIS model builder.

Characteristic backscatter values for banana were derived from SAR imagery of banana plantations in north-east Costa Rica. The available Costa Rican banana map was overlayed with the SAR imagery of all four polarization channels. In this way, the average backscatter signal from Costa Rican banana plantations was determined (see Table 2). The observed values were thereafter compared with SAR data from banana production regions in northern Colombia and southern Mindanao, Philippines. This confirmed that similar backscatter values were found for banana plantations in other parts of the world. It was therefore argued that the values in Table 2 can be used for plantation banana in general.

Table 2: The characteristic backscatter values of Costa Rican banana plantations for different SAR polarization channels. (Backscatter values of PALSAR-2 and Sentinel-1 are unitless, values refer to the specific measurement scale of either platform)

Polarization channel	Characteristic value
HH	6793
HV	2962
VV	- 3.93
VH	-10.90

2.3.3. Filtering of results

Landcover filter

The landcover filter masked out landcover classes that were irrelevant for mapping plantation banana. The banana production zones of the world are covered by several different landcover classes (Chapter 2.2.2). For example, Costa Rican banana plantations are either classified as mosaic cropland, mosaic vegetation or broadleaved evergreen forest. Similar observations were done for banana regions elsewhere, indicating the need for a lenient landcover filter. It turned out necessary to preserve seven

landcover classes in total (see Table 3). Those seven classes together covered the major banana production regions of Latin-America, Sub-Saharan Africa and South/South-East Asia.

Table 3: The GlobCover classes that together cover major banana production regions.

GlobCover Value	Description
11	Post-flooding or irrigated croplands
14	Rainfed croplands
20	Mosaic cropland / vegetation (grassland, shrubland, forest)
30	Mosaic vegetation (grassland, shrubland, forest) / cropland
40	Closed to open broadleaved evergreen and/or semi-deciduous forest
50	Closed broadleaved deciduous forest
60	Open broadleaved deciduous forest

Urban cover fraction filter

The urban filter excluded locations that were considered too urbanized for large-scale banana production. Peri-urban environments turned out an important source of erroneous results. It was observed that backscatter signals similar to banana emerged from the outskirts of towns. An explanation could be the mixture of built-up area and agriculture at such locations. On the one hand, built-up areas are characterized by high backscatter in several polarization channels (Ban et al., 2015). On the other hand, non-banana agriculture has relatively low backscatter for most polarization channels. The combined effect of high and low scatterers in peri-urban areas could explain the confusion with signals from banana. Peri-urban areas were therefore removed to avoid confusion by applying an arbitrary threshold of 20% for urban cover fraction.

Slope filter

The slope filter excluded locations that are too steep for banana plantations. This was done to eliminate confusing signals originating from mountainous areas. Large plantations are generally not constructed at steep terrain for two principal reasons. First of all, terrains with a steep slope have a high risk of surface erosion (Robinson and Galán Saúco, 2011). Secondly, steep slopes constrain the operation of banana transport systems such as cableways (Thompson and Burden, 1995). Previous studies applied slope filters successfully in the context of urban mapping (Ban et al., 2015, Gamba and Lisini, 2013). According to Delvaux (1995), terrains with slopes below 8% are agronomically suitable for banana production. Furthermore, Onderwater (2020) and Thompson and Burden (1995) mention slope thresholds of 0.5% and 0.2% for the operation of cableways. In this study, an intermediate slope threshold of 2% has been applied to filter out mountainous areas while preserving banana plantations on gently sloping terrain.

SAR-VH filter

The SAR-VH filter excludes locations that are most likely palm oil plantations. The filter was needed while SAR signals of palm oil and banana were confused. For instance, Indonesian palm oil plantations on Sumatra and Kalimantan were consistently classified as banana. However, it was observed that palm oil plantations returned VH backscatter signals smaller than -13 on the Sentinel-1 measurement

scale. Contrarily, banana plantations returned higher VH backscatter signals. A SAR-VH filter with a threshold of -13 on the Sentinel-1 measurement scale was therefore applied to eliminate the confusion.

2.3.4. Conversion to plantation banana map

Conversion parameters

The improved fuzzy map was finally converted to a plantation banana map. A distance threshold was introduced for this purpose. The distance threshold was the maximum Mahalanobis distance for which locations were still classified as banana. In this way, the distance threshold allowed for the separation of plantation banana from non-banana surfaces. Conversion based on only the distance threshold yielded results with significant noise. This is why an additional focal filter was applied. The underlying argument was that banana grid cells are generally surrounded by other banana grid cells. In other words, isolated banana grid cells have a high risk of being erroneous results. A neighborhood threshold was introduced to indicate the minimum required number of banana neighbors. The focal filter excluded all potential banana grid cells that had less banana neighbors than the neighborhood threshold.

A formal calibration procedure was developed to determine the optimal distance and neighborhood threshold values. In this way, the risk of biased results was reduced and reproducibility of the research was secured. It was decided to work with a small neighborhood size of 3 by 3 grid cells. Larger frames were not considered for practical reasons. Different combinations of neighborhood and distance thresholds were tested and results were compared with one another. The neighborhood threshold was varied from zero neighbors to a full encirclement of eight neighbors. The distance threshold was varied from one to fifty on a Byte scale. This yielded a total of 450 different combinations for distance and neighborhood thresholds. An overview of the calibration parameters is provided in Table 4.

Table 4: The ranges of conversion parameter values that were used for calibration. (The combination of eight different neighborhood thresholds with fifty different Mahalanobis distance thresholds yielded 450 calibration runs in total)

Parameter	**Fixed value**	**Lower value**	**Higher value**	**Step size**
Neighborhood size *(x by x grid cells)*	3	-	-	-
Neighborhood threshold *(nr. of neighbors)*	-	0	8	1
Distance threshold *(Byte scale)*	-	0	50	1

False negatives vs. false positives

Calibration largely reduced false positive and false negative results. False positives were defined as non-banana surfaces that had been classified as banana. Analogously, false negatives were defined as actual banana surfaces that had been overlooked in the classification. The relative impact of false positives and false negatives on the quality of results is usage-dependent (Connors et al., 2014). In this study, the impacts of false positives were considered larger than those of false negatives. It was argued that particularly false positives lead to deceptive global patterns. False positives were therefore

arbitrarily assigned a four times larger weight than false negatives. The occurrence of false positives and false negatives was assessed based on the calibration data. Calibration scores were calculated for each run by multiplying the total number of false positives and false negatives with their respective weight factors. The false positive and false negative scores were thereafter summed to obtain an overall score. The distance and neighborhood threshold giving the lowest score were thereafter selected as the optimal conversion parameters.

Calibration data

Calibration data turned out scarce for this study. Apart from Costa Rica, banana plantation maps were not available. Nevertheless, an alternative procedure was developed that made use of available data from other sources. Results were in this way calibrated for Costa Rica, Mexico, Nigeria and the Philippines. Each country involved a slightly different calibration method.

Costa Rica

A Costa Rican banana plantation map from 2019 was used to assess false negatives. The map was considered appropriate for calibration because its temporal extent roughly matched the SAR data in this study. False negatives were defined as Costa Rican banana plantations that were overlooked in this study. The map was only used for the assessment of false negatives, not for false positives. Google Earth imagery from 2019 indicated that not all plantations were covered by the map. Calibrating on false positives would consequently give distorted results.

Mexico

A Mexican land use map from 2016 was used to assess false positive results. The map distinguished different types of agriculture based on the crop's life cycle and its water supply. Any banana grid cell outside the perennial and semi-perennial land use classes was considered a false positive. The 2016 land use map was considered representative for the years 2018 and 2019 from which dated the SAR data in this study.

Nigeria

Nigeria is known for its large production of cooking banana for local consumption (Bifarin et al., 2010). The strong focus on production for the own market was confirmed by trade statistics on the country. According to FAOSTAT, Nigeria exported only 3 tons of banana and 11 tons of plantain in 2019 (FAO, 2021c). Compared to a neighboring country like Cameroon (≈ 190,000 tons), the export of Nigeria is negligible. Literature furthermore indicates that cooking banana in the country is cultivated by smallholder farmers (Akinyemi et al., 2010). This suggests that large-scale plantations associated with export banana are absent. Any banana grid cell in Nigeria was therefore considered a false positive.

Philippines

Regional production statistics of different banana cultivars were available for the Philippines. As a reminder, nearly all export banana in the world is Cavendish (Ploetz et al., 2015b). It was therefore argued that large banana plantations only occurred in regions with Cavendish banana. Banana grid cells were considered false positives when located in a region with zero Cavendish production. In this way, regional Cavendish banana statistics were used to assess false positives in the Philippines.

2.4. Mapping smallholder banana

2.4.1. Overview

A method was developed to map smallholder banana with land suitability and production statistics. A schematic overview of the method is provided in Figure 4. The method consists of two steps that are further explained in subchapters 2.4.2. and 2.4.3..

1. **Land evaluation**

   Land suitability was mapped at a fifteen arcsec resolution based on climate, soil and water table depth. A fuzzy classification was applied to the data. Suitability values were calculated for each environmental factor and thereafter combined into a single suitability map.

2. **Disaggregation of production statistics**

   National production statistics were disaggregated according to the land suitability map. This implied that statistics were distributed over locations with highest suitability. Furthermore, the maximum banana cover in each grid cell was made dependent on land suitability.

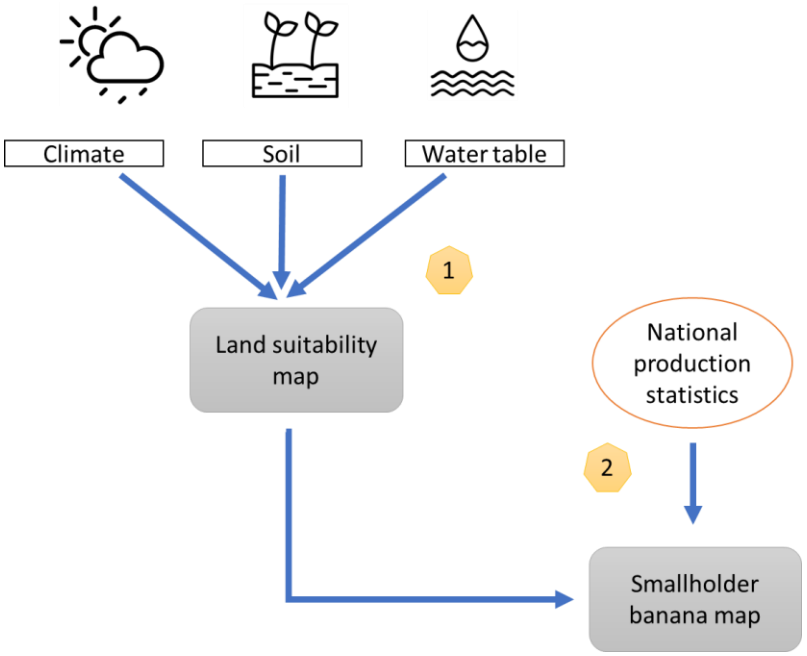

Figure 4: The method for mapping smallholder banana.
(The numbers in the Figure refer to the different steps in the text)

.

2.4.2. Land evaluation

Plantations have intensive management that allows for the adaptation of local conditions. On the contrary, smallholder systems strongly rely on natural environmental conditions (Rurinda, 2014). Small-scale farmers often lack the resources to regulate land characteristics such as moisture conditions, organic matter fractions and groundwater levels. The natural conditions consequently have to match the land use requirements in smallholder systems. In this context, the FAO framework for land evaluation introduced the concept of land quality. Land qualities are defined as complex land attributes that influence land suitability for a specific usage (FAO, 1976). A wide variety of land qualities exists, ranging from availability of inputs to absence of natural disasters. Most land qualities are described as a function of one or multiple land characteristics. Interestingly, land characteristics sometimes have an opposite impact on different land qualities. For instance, shallow water tables add to moisture availability while reducing oxygen availability. The relative importance of different land qualities depends on the specific land use requirements of a certain production system or crop.

Land qualities and land characteristics

In this study, six land qualities were considered for mapping land suitability (see Table 5). Those specific land qualities were selected based on their relevance for smallholder banana. High temperatures and solar irradiance are essential to sustain banana growth. Furthermore, high moisture availability is required to meet the crop water requirements of banana throughout the year. High oxygen availability is necessary to ensure the aeration of the soil. To conclude, high nutrient availability stimulates banana growth and high resistance to soil erosion limits the loss of nutrients, seeds and other inputs.

The land qualities were described by five distinct land characteristics (see Table 5). Cold quarter temperatures were used to assess temperature requirements of banana. Annual precipitation was used as an indicator for cloudiness and therefore solar irradiation (Robinson and Galán Saúco, 2011). The land quality moisture availability was described by four land characteristics in total: dry quarter precipitation, annual precipitation, water table depth and soil organic matter. Water table depth was used to assess oxygen availability in the rootzone. To conclude, soil organic carbon was used as an indicator for both nutrient availability and resistance to erosion.

The general relationships between land qualities and their associated land characteristics have been listed in Table 5. These relationships basically summarize the aforementioned. For instance, high annual precipitation increases cloudiness and thus decreases solar irradiation. This implies a negative relationship between annual precipitation and solar irradiation.

Table 5: The six land qualities and their associated land characteristics that were used for land suitability mapping. (The third column indicates the general relationship between the qualities and their associated land characteristics).

Land quality	**Land characteristic(s)**	**Relationship**
Temperature	Temperature (*average for the coldest quarter of the year*)	Positive
Solar irradiance	Annual precipitation	Negative
Moisture availability	Precipitation (*average for the driest quarter of the year*) Annual precipitation Water table depth Soil organic carbon	Positive Positive Negative Positive
Oxygen availability	Water table depth	Positive
Nutrient availability	Soil organic carbon	Positive
Resistance to erosion	Soil organic carbon	Positive

Fuzzy classification of land suitability

A rule-based model was developed for the fuzzy classification of land suitability. The rule-based nature of the model allowed to derive suitability thresholds from literature. This was convenient given that global training data on smallholder banana occurrence was not available. Fuzzy classifications move beyond the rigid classification schemes of traditional land evaluation methods. Instead of discrete classes, fuzzy classifications produce membership values that reflect how well an object fits into a class (Joss et al., 2008). Previous studies showed several advantages of fuzzy classification for crop mapping (Qiu et al., 2014). Most importantly, fuzzy classifications represent the continuous gradients in landscapes better than discrete classifications. In this study, fuzzy logic was applied to compare land characteristics with land use requirements of banana. This produced continuous suitability values for smallholder banana.

Fuzzy suitability values were calculated for each of the five land characteristics in Table 5. This step is called the *fuzzification* of input data. Membership functions were formulated to describe the relation between land characteristics and land suitability. The shape of membership functions reflected the correlation between a land characteristic and its corresponding land qualities. For example, the membership function for cold quarter temperature increased with temperature. This corresponds to the positive correlation between cold quarter temperature and the temperature land quality. Some land characteristics asked for more complex membership functions. For instance, annual precipitation has a contrasting influence on moisture availability and solar irradiation. This relationship was modelled by an increase in suitability at low precipitation levels and a decrease in suitability at high precipitation levels.

Two types of membership functions were used in this study. Those functions were referred to as type 1 and type 2 (Figure 5 & 6). The type 1 function applied to land characteristics that exclusively had a positive correlation with suitability. Those were cold quarter temperature, dry quarter precipitation and soil organic carbon. The type 2 function applied to land characteristics that had a positive correlation with suitability for low values and a negative correlation with suitability for high values. Annual precipitation fell into this category.

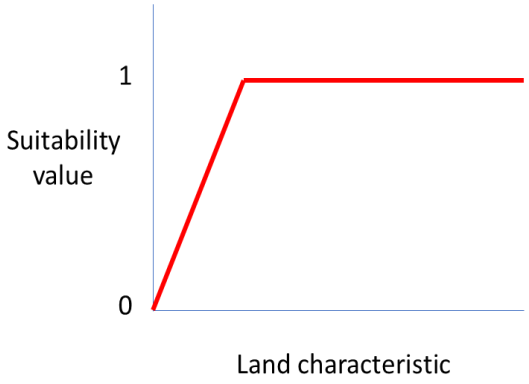

Figure 5: The type 1 membership function for the relationship between land characteristics and suitability values

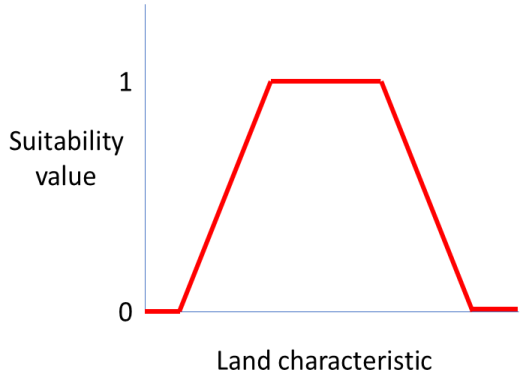

Figure 6: The type 2 membership function for the relationship between land characteristics and suitability values

The land characteristic water table depth formed a special case while both membership functions were used. It was argued that the prevalence of either membership function depended on climate conditions. The type 1 function was considered appropriate for wet climate conditions and the type 2 function for dry climate conditions. Under wet climate conditions, suitability increases with water table depth due to the increased oxygen availability. Under dry climate conditions, however, this positive correlation was considered only relevant up to a certain depth. At greater water table depths, suitability decreases due to constraints in moisture availability. To characterize climate wetness, annual precipitation was used as an indicator. The relevance of each membership function was expressed by a weight factor based on annual precipitation. Appropriate weights were derived by distinguishing three scenarios (Figure 7). First of all, there is a precipitation level below which only the membership function for dry climate conditions was considered relevant (Scenario A). Secondly, there is a precipitation level above which only the membership function for wet climate conditions was considered relevant (Scenario C). Thirdly, any precipitation level in-between resulted in a suitability value that was a linear combination of the two membership functions (Scenario B).

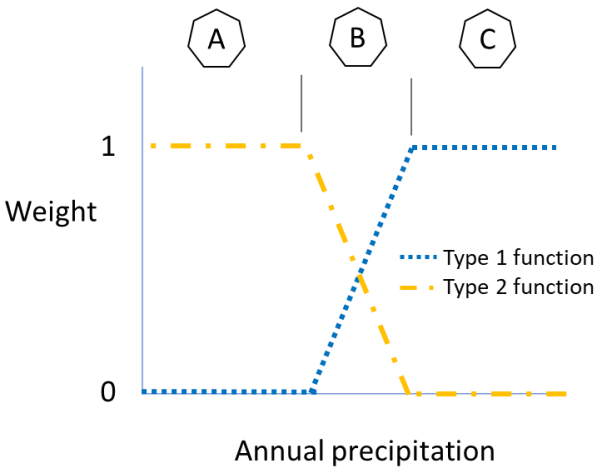

Figure 7: The relationship between annual precipitation and the relative weight of the type 1 and type 2 membership function for water table depth. (Scenario A, B and C correspond to the scenarios explained in the text)

The fuzzy membership values of each land characteristic were thereafter combined into a single suitability map. In literature, this step is often called *fuzzy inference*. Several different methods for fuzzy inference exist (Qiu et al., 2014). In this study, fuzzy inference was performed by calculating the average fuzzy membership value at each location. No distinction was made between the different variables by assigning an equal weight to all of them. This decision reflected the absence of field data on smallholder banana for the calibration of those weights.

Suitability thresholds

The suitability thresholds that were derived from literature are listed in Table 6. In line with Figure 5 and 6, variables with a trapezoidal membership function have four different thresholds and variables with a linear membership function have two thresholds. The choice for each threshold value is substantiated below.

Table 6: The suitability thresholds for each land characteristic. (Land characteristics with a type 1 membership function have two thresholds, land characteristics with a type 2 membership function have four thresholds)

Variable	**Unit**	**Threshold 1:** ***Minimum***	**Threshold 2:** ***Lower optimum***	**Threshold 3:** ***Higher optimum***	**Threshold 4:** ***Maximum***
Temperature (cold quarter)	°C	14	22	-	-
Precipitation (annual)	mm	650	1400	4000	5000
Precipitation (dry quarter)	mm	200	350	-	-
Water table (wet climate)	cm *(below surface)*	30	120	-	-
Water table (dry climate)	cm *(below surface)*	30	60	100	500
Soil organic carbon	%	0	1.5	-	-

Cold quarter temperature

Banana suffers from cold damage when temperatures are too low. For this reason, all locations with a cold quarter temperature below 6 °C were masked out as described in Chapter 2.2.3. Those locations were not considered for further analysis, neither for mapping plantation banana nor smallholder banana. Additional suitability thresholds for cold quarter temperature were defined for mapping smallholder banana. This implies that locations could score low on temperature within the land evaluation, despite passing the hard 6 °C threshold. According to Robinson and Galán Saúco (2011), banana growth ceases below 14 °C. Lower temperatures provoke a discontinuity in the dry matter accumulation. When mean temperatures exceed 14 °C the whole year, banana growth is not interrupted during the coldest months. This is why a minimum threshold of 14 °C was chosen for cold

quarter temperature. The same authors reported optimum temperatures for the net assimilation rate of banana at 22°C. The 22 °C is therefore used as the lower optimum threshold for cold quarter temperature.

Annual precipitation

Banana is characterized by a high water demand. According to literature, annual precipitation levels of at least 1300 to 1500 mm are required to meet the crop water requirements of banana (Delvaux, 1995, Robinson and Galán Saúco, 2011). A value of 1400 mm was therefore chosen as the lower optimum threshold. With respect to solar irradiation, Robinson and Galán Saúco (2011) observed predominantly overcast conditions at annual precipitation levels of 4000 mm. It was argued that the availability of photosynthetically active radiation becomes a limitation to banana growth at such high precipitation levels. For this reason, this value was used as a higher optimum threshold. To conclude, multiple land suitability studies used 650 and 5000 mm as absolute precipitation thresholds for banana (Salvacion et al., 2019, Onderwater, 2020). These values were therefore adopted as the minimum and maximum thresholds for annual precipitation in this study.

Dry quarter precipitation

Banana is sensitive to drought stress caused by temporary water shortage. High precipitation levels are required throughout the year to avoid constraints on banana growth and development (Robinson and Galán Saúco, 2011, Daniells, 1984). In this context, Onderwater (2020) introduced a minimum threshold of 200 mm for dry quarter precipitation. The same value was adopted in this study. As a reminder, the lower optimum threshold for annual precipitation was set to 1400 mm. It was argued that dry quarter precipitation should be at least one-fourth of this 1400 mm to ensure optimal conditions throughout the year. This is why a lower optimum threshold of 350 mm was used for the dry quarter precipitation.

Water table depth

Water table depth has an effect on both drainage and moisture availability. Deep water tables result in favorable drainage conditions while shallow water tables result in high moisture availability (Zhu et al., 2013). The need for a deep or shallow water table depends on the location. In (semi-)arid regions, ecosystems often rely on groundwater to meet their moisture requirements (Eamus et al., 2016). On the other hand, good drainage is the principal requirement in humid regions. Banana is susceptible to waterlogging in the rootzone under wet climate conditions. According to Robinson and Galán Saúco (2011), the vast majority of banana roots is within the top 30 cm of the soil. It was therefore argued that water tables shallower than 30 cm result in highly unsuitable conditions. In addition, the same authors stated that water tables deeper than 120 cm are most suitable for banana in the humid tropics. This is why a minimum threshold of 30 cm and a lower optimum threshold of 120 cm were selected for water table depth in a wet climate.

Meanwhile, shallow water tables can be favorable to water availability under dry climate conditions. Regardless of the climate, however, water table depths larger than 30 cm are considered a precondition to avoid waterlogging. The previously defined minimum threshold of 30 cm was therefore also used for water table depth in semi-arid areas. In addition, banana roots penetrate below 60 cm only occasionally (Robinson and Galán Saúco, 2011). It was reasoned that consequently the risk of waterlogging in semi-arid areas is negligible if water table depths exceed 60 cm. This is why a lower optimum threshold of 60 cm was applied in this study. Furthermore, literature on the functional relationship between groundwater levels and soil moisture was used to select the remaining thresholds. Kollet and Maxwell (2008) conceptualized the interconnection between groundwater and

shallow soil moisture into different cases. They argued that the influence of groundwater on soil moisture starts to decline for water table depths deeper than 100 cm. In addition, they stated that the upward redistribution of soil moisture becomes negligible for water table depths larger than 500 cm. The higher optimum and maximum thresholds were therefore set to 100 and 500 cm respectively.

As discussed before, annual precipitation was used to assess the relative importance of the type 1 and type 2 membership functions (Figure 5 & 6). In line with Figure 7, the calculation of weight factors required two thresholds for annual precipitation. Pfadenhauer and Klötzli (2020b) associated the arid (sub)tropics with precipitation levels below 500 to 600 mm. Furthermore, the same authors defined precipitation levels over 2000 mm for the humid tropics (Pfadenhauer and Klötzli, 2020c). This is why the threshold between scenario A and B was set to 550 mm and the threshold between scenario B and C to 2000 mm (see Table 7).

Table 7: The annual precipitation thresholds to differentiate between dry, intermediate and wet climate conditions. (Threshold A-B and B-C refer to scenario A and B and scenario B and C in Figure 7 respectively)

Threshold name	Threshold value	Definition
Threshold A-B	550 mm	Tipping point between arid and moderate climate conditions
Threshold B-C	2000 mm	Tipping point between moderate and humid climate conditions

Soil organic carbon (SOC)

Banana benefits from high SOC for multiple reasons. The positive relationship between SOC and soil quality has a continuous nature (Sparling et al., 2003, McCallister and Chien, 2000). In this study, it was therefore assumed that suitability for crop production increases gradually from zero SOC onwards. Furthermore, a wide variety of optimal SOC values was found in literature. Values ranging from 0.6% and 3% have been reported by different studies depending on their approach and study extent (Hijbeek et al., 2017). In this study, SOC values below 1.5% were considered suboptimal for crop production. The threshold of 1.5% falls within the range observed by Hijbeek et al. (2017) but was further chosen arbitrarily. The selected suitability thresholds of 0% for the minimum and 1.5% for the lower optimum reflect the findings above.

2.4.3. Disaggregation of production statistics

The smallholder banana map was created through the disaggregation of production statistics. The suitability map and plantation banana map were transformed to the coordinate reference system EPSG:3857 before disaggregation. This was done to obtain data with linear units of measure for coordinates. A nearest neighbor resampling was applied to preserve the original grid values.

Adjustment of production statistics

The FAO production statistics on banana and plantain were used in this study (see Chapter 2.2.2). An important preparatory step was the reduction of the statistics based on the results for plantation banana. Plantation banana was mapped independent from the production statistics as explained in Chapter 2.3.. The production statistics from FAOSTAT however make no distinction between production systems and therefore cover both plantation and smallholder banana. This means that the plantation banana surface had to be subtracted from the production statistics. The locations were also removed from the suitability map to avoid the allocation of banana to the same grid cell twice.

For each country, the number of banana grid cells (n_{cell}) was multiplied with the cell size (A_{cell}). The cell size and cropping surface are often considered the same in disaggregation studies (Khan et al., 2010). However, the crop coverage at banana plantations is lower than the cell size due to the presence of infrastructure and buildings. This is why a crop coverage factor of 0.9 was applied (f_{cover}). Equation 1 summarizes the procedure for the adjustment of production statistics.

In which:

$Stats_{new}$ *= Harvested area production statistics after adjustment for plantation banana*

$Stats_{old}$ *= Harvested area production statistics as reported by the FAO (plantain + banana)*

n_{cell} *= Number of banana grid cells*

A_{cell} *= Cell size*

f_{cover} *= Correction factor for crop coverage*

Disaggregation procedure

Locations with high suitability had prevalence over locations with low suitability during the disaggregation. This implied that smallholder banana was distributed over locations with highest suitability first, locations with second highest suitability second and so on. Grid cells were basically filled up with banana from high to low suitability until all statistics were distributed. Most grid cells got therefore either zero banana coverage or maximum banana coverage. The only exception were grid cells for which the suitability value coincided with the suitability threshold. The suitability threshold was defined as the lowest suitability level over which statistics were disaggregated. The statistics were generally not sufficient to obtain the maximum banana coverage in these grid cells. Statistics were therefore equally distributed over all grid cells with this suitability level. The suitability threshold differed per country and was inferred automatically from the data.

The specific characteristics of smallholder systems were taken into account in the disaggregation process. Contrary to plantation banana, smallholder banana is often cultivated in multi-cropping systems (Robinson and Galán Saúco, 2011). It was therefore argued that smallholder banana coverage is significantly lower than plantation banana coverage. Furthermore, it was reasoned that locations with high suitability have larger banana coverage than locations with low suitability. The smallholder banana coverage was therefore explicitly linked to suitability. Maximum banana coverage was arbitrarily set to 0.4 for locations with optimal suitability (f_{max}). From there, banana coverage decreased

linearly to zero for lower suitability values (S_{act} / S_{pot}). This relationship between suitability and crop cover fraction is summarized in Equation 2.

In which:

f_{act} *= Maximum banana coverage of a grid cell after the correction based on suitability*

f_{max} *= Maximum banana coverage for grid cells with optimal suitability*

S_{act} *= Actual suitability value of a grid cell*

S_{pot} *= Potential (maximum) suitability value of a grid cell*

2.5. Validation of results

Validation was done by the visual comparison of results with satellite imagery and existing global banana maps. The visual comparison with Google Earth satellite imagery was performed to assess differences between the banana map and actual crop cover. The validation was done for several regions in different parts the world. The exact validation locations are specified in Table 8. None of the validation locations were previously used for calibration purposes.

A visual comparison with two other global banana maps was made to compare spatial banana patterns at a global scale. Both maps involved a certain procedure for the disaggregation of production statistics. Monfreda et al. (2008) disaggregated production statistics based on global landcover data. Onderwater (2020) disaggregated production statistics with a combination of landcover and environmental data. Neither study mapped large-scale systems and smallholder systems differently.

Table 8: The locations for validation of the results with Google Earth imagery. (The region and location for Tanzania are missing because the whole country was considered for validation)

Country	Region	Location
Venezuela	Zulia state	Lake Maracaibo
Tanzania	-	-
India	Maharashtra state	Jalgaon
Australia	Queensland	Tully & Innisfail

3. Results & Discussion

3.1. Plantation banana map

3.1.1. Fuzzy plantation banana map

The classification of radar satellite data resulted in a fuzzy banana map (Figure 8). This map gives a first impression of banana occurrences. The values indicate how similar the measured radar signals are to banana. In Figure 8, locations with high similarity are represented by light colors. These are the locations where banana would be expected based on the radar signal. The similarity is clearly higher in the tropics than in desert zones. Furthermore, the map indicates higher similarity outside the tropical rainforest zones around the equator.

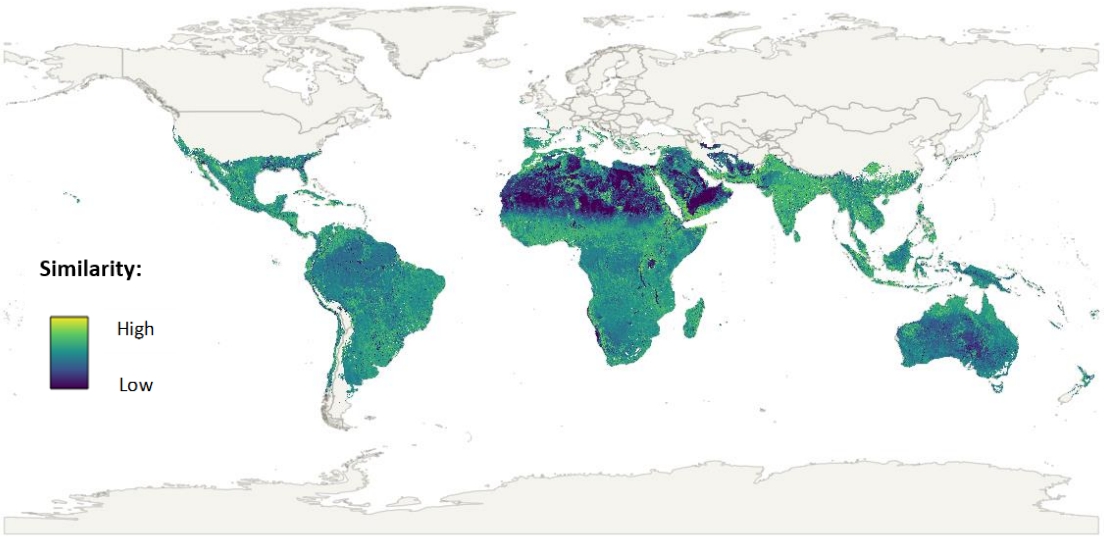

Figure 8: The fuzzy plantation banana map of the world. (Yellow colors indicate high similarity between the measured radar signal and the characteristic banana signal)

The map was created using the Mahalanobis distance algorithm. The Mahalanobis distance algorithm is commonly used for classification of satellite imagery. Several studies compared the performance of Mahalanobis distance with other mapping techniques. For instance, Hossen et al. (2018) compared unsupervised clustering, maximum likelihood and statistical distance techniques for landcover monitoring. The Mahalanobis distance was among the best performing algorithms in their study. In addition, Shi et al. (2019) assessed the performance of different classification techniques for crop mapping. Mahalanobis distance yielded high prediction accuracies similar to other algorithms.

Global crop mapping asks for simple classification algorithms like the Mahalanobis distance to limit processing time. Moreover, data requirements should not exceed available computation power. This is why GRD radar products were used in this study (Chapter 2.2.2.). It should be noted that the processing opportunities for GRD products are more limited than for SLC products. For instance, SLC products can be processed through polarimetric decomposition (López-Martínez and Pottier, 2021,

Raney, 2014). Polarimetric target decomposition provides insight in the dominant scattering mechanisms behind the measured radar signal (Srikanth et al., 2016). In this way, SLC products would allow for more informed crop classifications than GRD data (Xie et al., 2015).

Figure 9 shows the fuzzy plantation banana map for Costa Rica. The yellow-colored banana plantations in the northeastern part of the country clearly stand out. This makes sense given that a banana plantation map of this region was used for calibration of the conversion parameters. The sharp-eyed reader may observe a diagonal artifact line in the results, indicated by the green arrow in Figure 9. This line forms the boundary between two neighboring flight paths of the PALSAR-2 satellite. Since 2019, JAXA doesn't balance the backscatter intensity of neighboring PALSAR-2 paths in mosaic images anymore (JAXA, 2021). Images from different flight paths are now only orthorectified and corrected for slope. This decision has direct effect on the fuzzy map in Figure 9 because PALSAR mosaics for 2018/2019 were used. Radiometric discontinuities between PALSAR images with a different acquisition time can be attributed to seasonality and local weather conditions (Grandi et al., 2011). Different SAR signals may emerge from associated changes in vegetation cover and moisture conditions over the year. The results however indicate that the artifact line only impacts the classification of plantation banana marginally. Figure 9 shows that the banana plantations in the region stand out regardless of the flight path that covers them. This could reflect the regulated vegetation cover and moisture levels at banana plantations during the year.

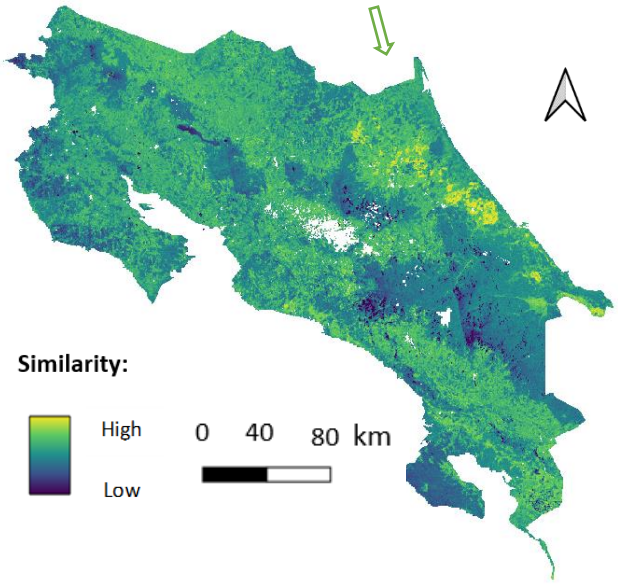

Figure 9: The fuzzy plantation banana map of Costa Rica. (Yellow colors indicate high similarity between the measured radar signal and the characteristic banana signal. The green arrow points to an artifact line that originates from the PALSAR-2 data)

3.1.2. Filtered plantation banana map

Non-banana surfaces were filtered out from the fuzzy plantation banana map. Besides a general landcover filter, specific filters were applied for urban cover fraction, slope and VH backscatter. The filters reduced the potential banana area and moreover reduced the confusion with peri-urban areas, mountainous areas and palm oil plantations. The four filters together reduced the potential area to 260245 km^2, unevenly distributed over the world. The effect of each separate filter is illustrated in more detail on the basis of specific regions.

The effect of the landcover filter is clearly visible in Egypt. The country mainly consists of dessert areas that are too arid for large-scale banana production (Figure 10). Alongside the Nile River, however, a band of irrigated agricultural land can be found. The landcover filter masked out the dessert areas while preserving the agricultural lands. This reduced the potential banana area in Egypt drastically.

The filter for urban cover fraction improved results significantly for several countries, including Nigeria. The filter allowed for the identification of (peri-)urban areas that remained invisible on the landcover map. This was needed because a large number of erroneous results appeared to originate from the outskirts of towns. In the case of Nigeria, fifteen times more urban area was filtered out by the filter for urban cover fraction compared to the landcover filter (Figure 11).

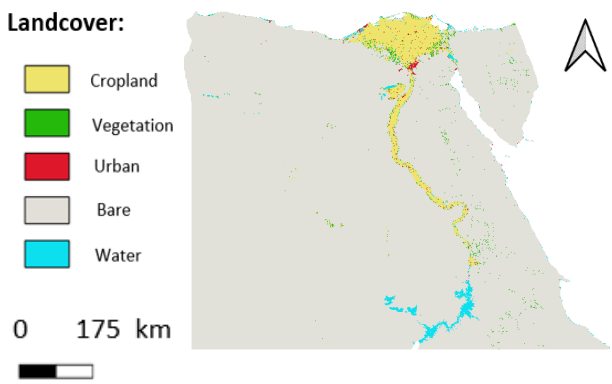

Figure 10: The landcover map of Egypt. The landcover filter eliminated bare areas while preserving cropland

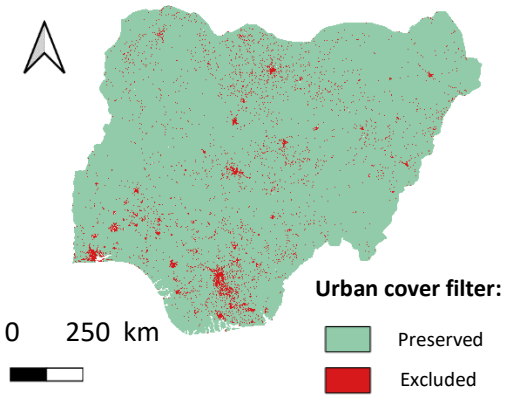

Figure 11: The urban cover fraction filter for Nigeria. (Green areas were preserved, red areas excluded)

Mountainous areas turned out another source of errors. The slope filter made it possible to exclude steep terrains that are unsuitable for large-scale plantation agriculture. This reduced the noise in the results significantly. A good example of the effect of the slope filter is the Dominican Republic (Figure 12). Satellite imagery indicates that banana production on the island exclusively takes place in gently sloping valleys. The slope filter successfully filtered out the coastal and central mountain ranges while preserving the valleys in between.

Lastly, the SAR signal from oil palm plantations caused significant confusion with the signal of banana. It was observed, however, that oil palm has generally lower VH backscatter signals compared to banana. The additional VH filter allowed for the separation of oil palm and banana by putting a hard boundary on the VH signal. For example, most large oil palm plantations in southern Borneo, Indonesia were filtered out in this way (Figure 13).

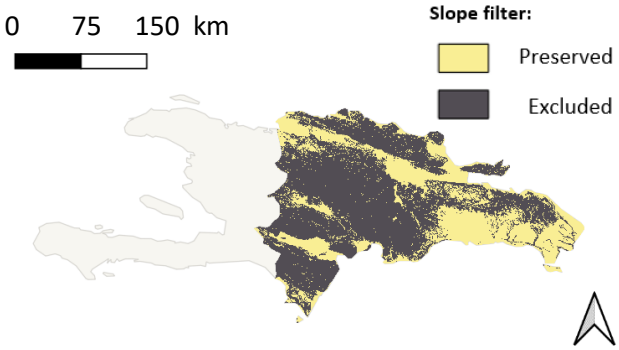

Figure 12: The slope filter for the Dominican Republic. (Yellow areas were preserved, black areas excluded)

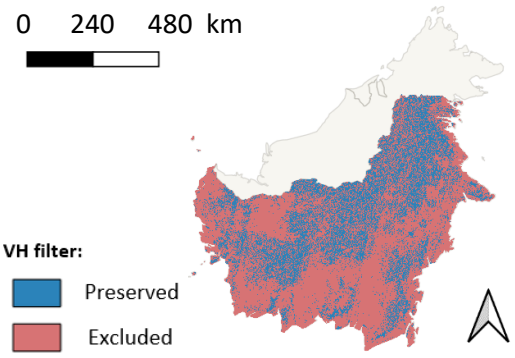

Figure 13: The VH filter for Borneo, Indonesia. (Blue areas were preserved, red areas were excluded)

3.1.3. Final plantation banana map

The filtered fuzzy banana map was finally converted into a plantation banana map. Best results were produced with a distance threshold of 25 combined with at least 4 neighboring banana grid cells in a 3 by 3 neighborhood. Figure 14 presents the resulting map, showing the global distribution of plantation banana in the world. The plantation banana surface in Figure 14 corresponds to approximately 6.5 % of the global banana surface according to the FAO (FAO, 2021b).

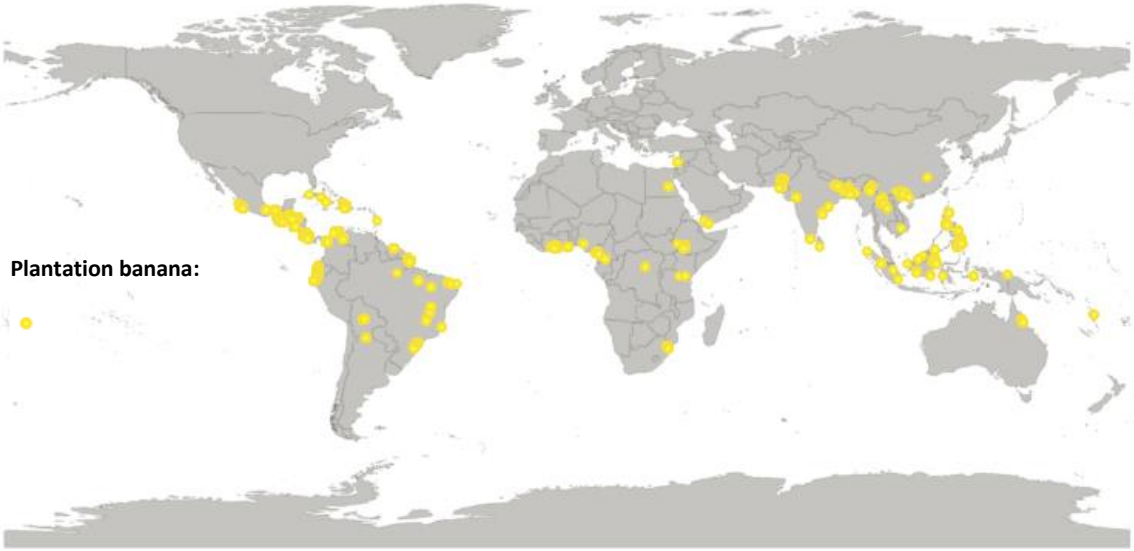

Figure 14: The final global plantation banana map. (The banana grid cells are represented by points for visualization purposes)

The results clearly reflect the climatic requirements of banana. Banana occurrence is concentrated in the frost-free (sub)tropics, roughly situated between the tropic of Cancer and the tropic of Capricorn (Figure 15). Most locations are also in areas with high annual precipitation levels, such as Central and South America, West Africa and South-East Asia (Figure 16). Meanwhile several exceptions to those

spatial patterns can be observed. Plantation banana was found under more temperate conditions up to Israel in the northern hemisphere. Analogously, plantation banana was observed up to South-Africa and southern Brazil in the southern hemisphere. Banana also appears present in (semi)arid environments such as Pakistan and Yemen. These kind of azonal banana occurrences may be attributed to favorable local climate conditions that are different from the macro-climate (Pfadenhauer and Klötzli, 2020a). In addition, they may be a result of intensive management practices to improve the growing conditions for banana.

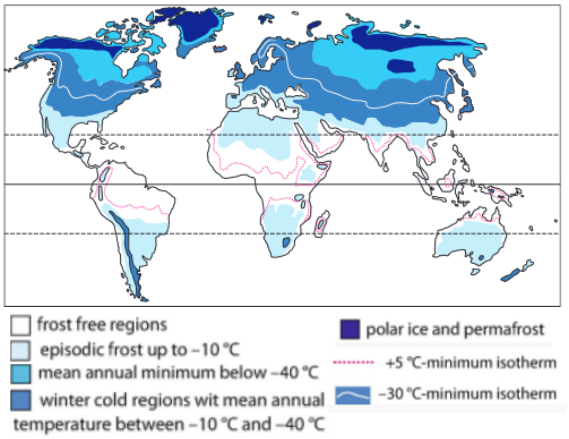

Figure 15: Global frost map (Pfadenhauer and Klötzli, 2020a)

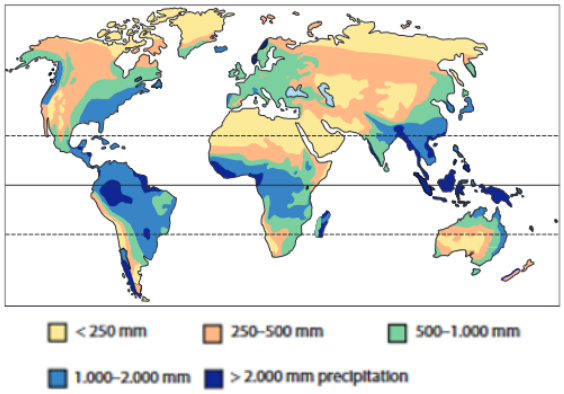

Figure 16: Global precipitation map (Pfadenhauer and Klötzli, 2020a)

The calibration procedure was performed based on available data of four different countries. The number of false negatives was minimized for Costa Rica and the number of false positives for Nigeria, Mexico and the Philippines. In ecology, false positives are traditionally considered absent from species distribution mapping (Pillay et al., 2014). This has however changed with the emergence of satellite-based approaches (Chambert et al., 2015, Fretwell et al., 2014). In this study, false positives were considered more disruptive than false negatives. It was argued that particularly false positives affect the general global patterns as displayed by the map. The results showed that four times more false negatives persisted after calibration compared to false positives. This reflects the decision to give false positives a four times larger weight than false negatives. It should however be noted that the specific one to four ratio was chosen arbitrarily. Other ratios would be equally suitable, depending on the preferences of users and the intended map usage (Connors et al., 2014).

Figure 17 to 22 show the final results of the four countries that were used for calibration. In Costa Rica, over eighty percent of the calibration data was mapped accurately (Figure 17). This producer accuracy should be seen as a good classification score, especially for a crop map with global extent (Vintrou et al., 2012). Furthermore, several plantations that are not present on the existing map clearly appear on the global map (Figure 18). This affirms the decision to use the Costa Rican map to minimize false negative results only. It turned out that the Costa Rican map did not cover all plantations in the region. Using this map to also minimize false positives would consequently have led to biased results.

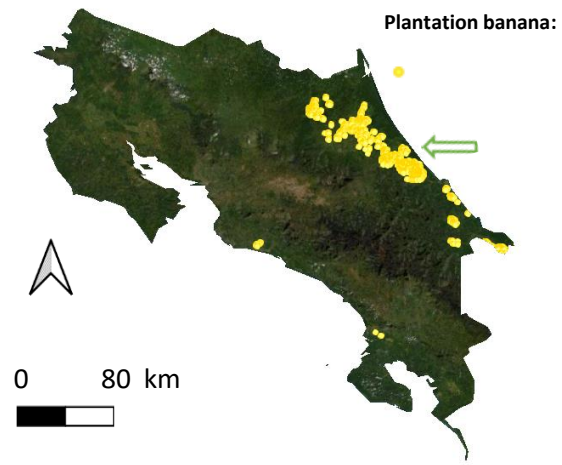

Figure 17: The final plantation banana map of Costa Rica. (The green arrow shows the location of Figure 18)

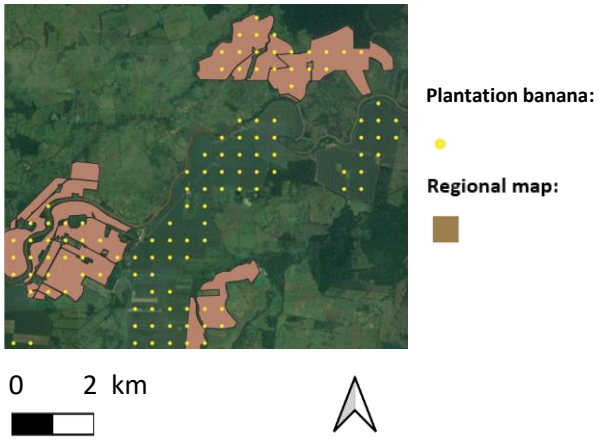

Figure 18: Close-up of Costa Rican banana fields. (The regional map and study results have been mapped on top)

The results for Mexico show a concentration of plantation banana in the southern part of the country (Figure 19). The satellite imagery confirms that this is the humid, tropical side of the country where indeed banana would be expected. The number of remaining false negatives after calibration is negligible, being two observations only (i.e., 0.25% of the total). Even better, those few false negatives closely resemble actual banana plantations on satellite imagery (Figure 20). The false negatives in Mexico consequently seem to originate from inaccuracies in the calibration data. This suggests a perfect user accuracy for Mexico: every observation on the map is an actual banana plantation in the field.

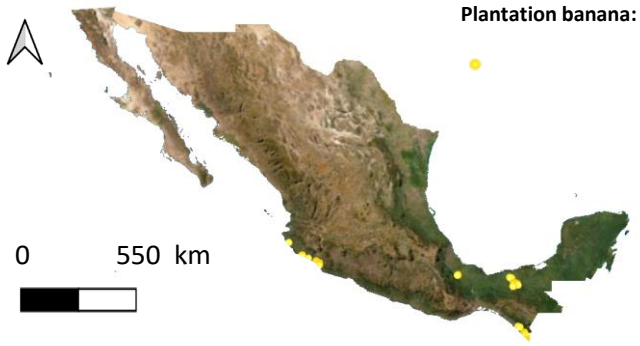

Figure 19: The final plantation banana map of Mexico.

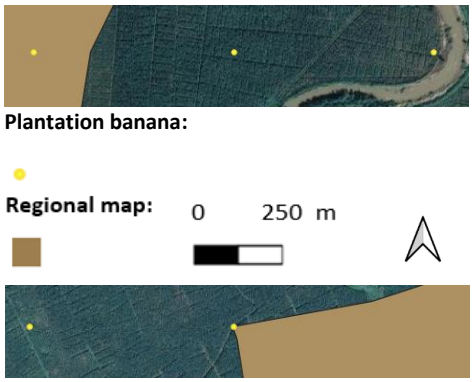

Figure 20: Close-up of false negative results in Mexico. (The two locations actually seem banana plantations)

In Nigeria, two false negatives remained after calibration. Both observations originate from the outskirts of a town called Akure in the southwestern part of the country (Figure 21). The filter for urban cover fraction failed to eliminate those two locations from the results. A significantly larger number of 67 false negatives (i.e. 5.01% of total) persisted in the results of the Philippines (Figure 22). The calibration for the Philippines was done based on regional Cavendish production statistics. All observations in regions with zero Cavendish production levels were classified as false positives. The principal argument for this approach was that large-scale banana production revolves around the

Cavendish cultivar (Ploetz, 2015b). However, the Saba and Lakatan cultivars are cultivated in large quantities in some northern provinces of the Philippines (Figure 23). Satellite data confirms that the alleged false negatives here are actually banana. On the one hand, this observation means that the calibration procedure for the Philippines was not perfectly accurate. The Saba and Lakatan banana plantations in the northern Philippines were unrightfully treated as false negatives. On the other hand, this observation suggests that the SAR-based method in this study provides robust results for different cultivars. This provides good prospects for future work on global banana mapping with similar methods.

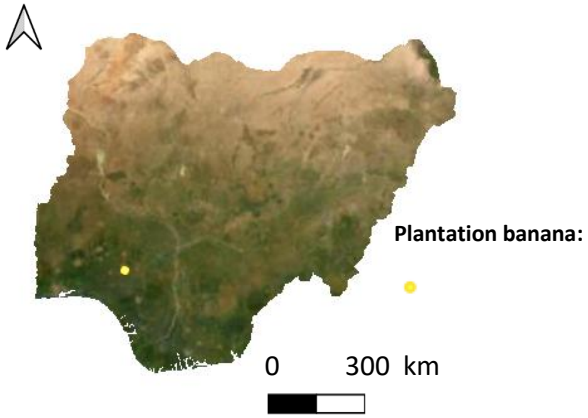

Figure 21: The final plantation banana map of Nigeria.

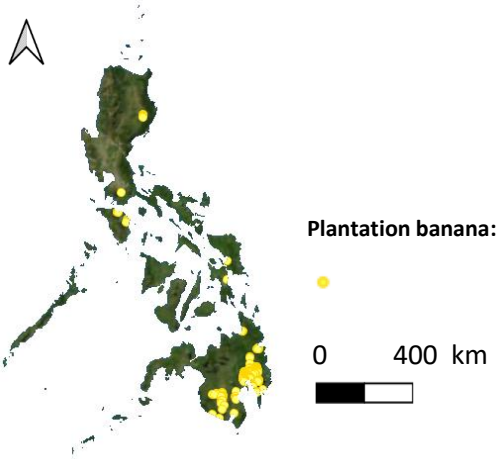

Figure 22: The final plantation banana map of the Philippines.

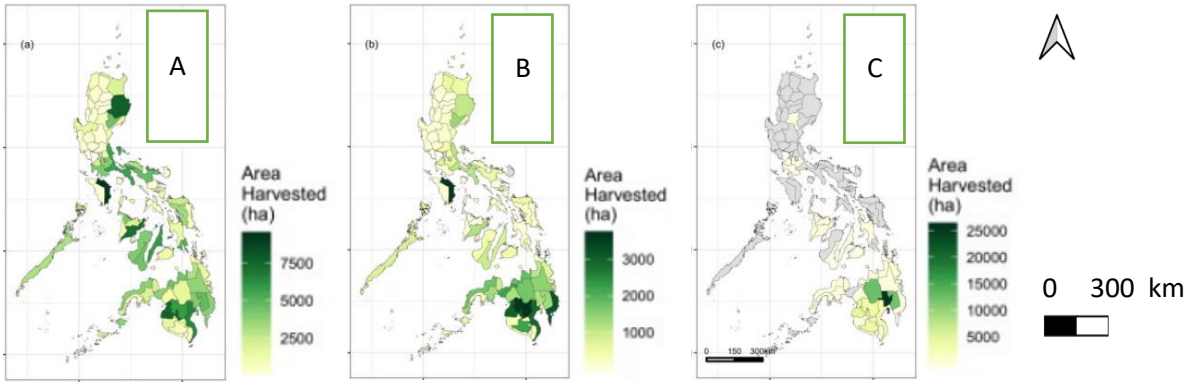

Figure 23: Regional banana production statistics of the Philippines for the Saba (A), Lakatan (B) and Cavendish (C) cultivars. (Salvacion, 2020)

3.2. Smallholder banana map

3.2.1. Fuzzy land suitability maps

Fuzzy land suitability maps were created for the five land characteristics that were used in this study (Figure 24). The fuzzy values express the suitability for banana cultivation on a continuous scale that runs from zero to one. The closer to one, the higher the suitability. Suitability values for annual precipitation and cold quarter temperature are highest in the tropics. This is in correspondence with the high water and temperature requirements of banana. Dry quarter precipitation turns out to be the most limiting environmental factor. The suitability values for this land characteristic are zero in a large part of the word. Soil organic carbon (SOC) appears to be the least limiting factor. Most locations on earth have a suitability value for SOC larger than zero. The global SOC pattern roughly resembles the annual precipitation pattern. This relationship reflects the impact of precipitation on plant productivity and SOC storage consequently (Saiz et al., 2012). To conclude, two suitability maps were created for water table depth. The first map shows suitability for dry climate conditions, the second map for wet climate conditions. The two maps roughly show a reversed pattern. Predominantly low values were mapped for dry conditions whereas predominantly high values were modelled for wet conditions. These patterns show that most locations on earth have water table depth larger than five meters.

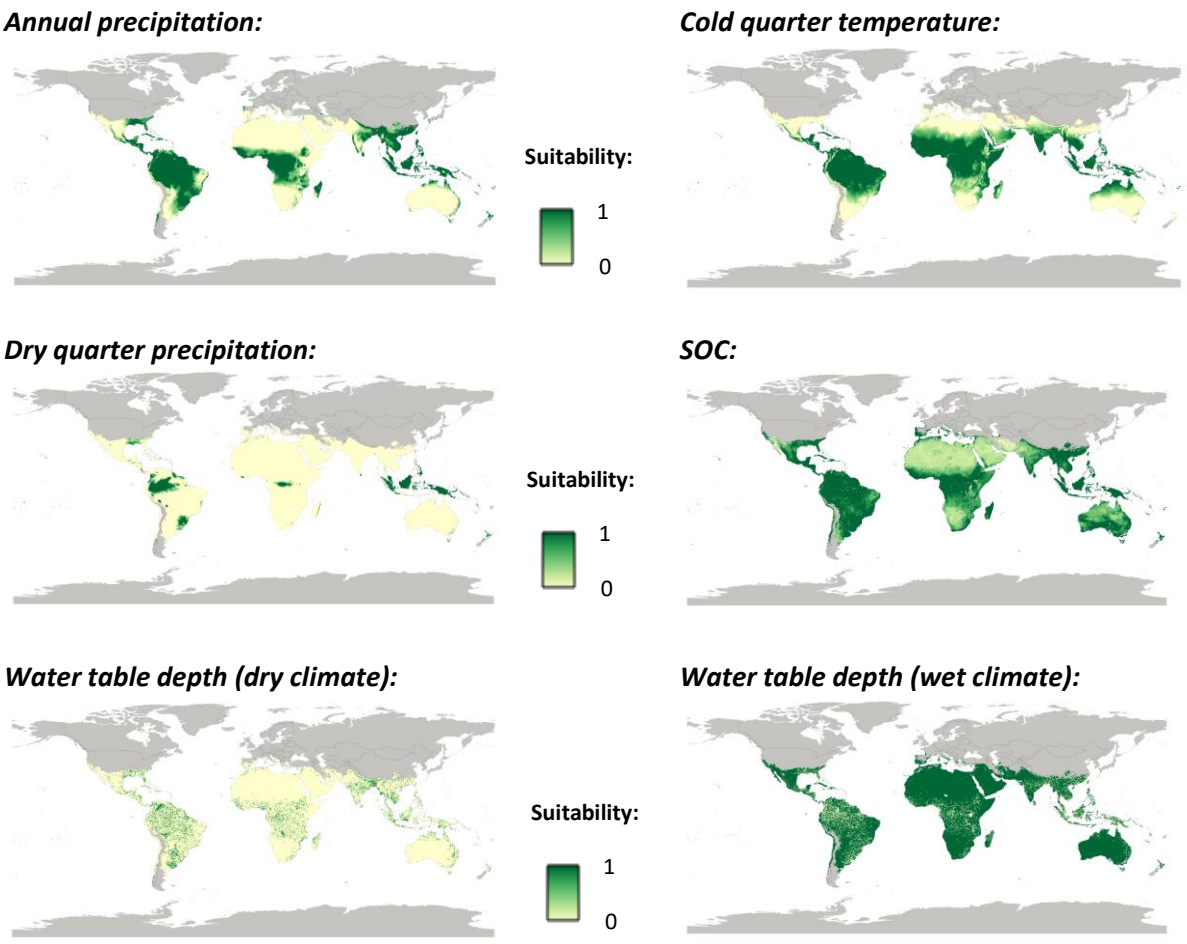

Figure 24: The fuzzy suitability maps for the five land characteristics that were used in this study.

The two suitability maps for water table depth were combined based on annual precipitation (see Chapter 2.4.2.). The global distribution of different rainfall zones is shown in Figure 25. The rainfall zones have a clear impact on the combined suitability map for water table depth (Figure 26). For example, the suitability values in the Sahara reflect the dry climate. Analogously, the suitability values in Indonesia reflect the wet climate. Finally, the suitability values in northern Australia are a linear combination of both maps in Figure 24. The results for Ethiopia show how this approach works out at regional scales. A distinct east-west precipitation gradient runs through this country (Figure 27). This gradient is well visible in the combined suitability map (Figure 28). Suitability values are lower in the (semi)arid eastern part of the country compared to the humid highlands in the west. Moreover, the relative location of high suitability values differs per rainfall zone (Figure 29). In the arid parts, highest values are found close to river streams. In the humid parts, highest values are found further away from streams. These patterns reflect the spatial differences in land use requirements of banana with respect to water table depth.

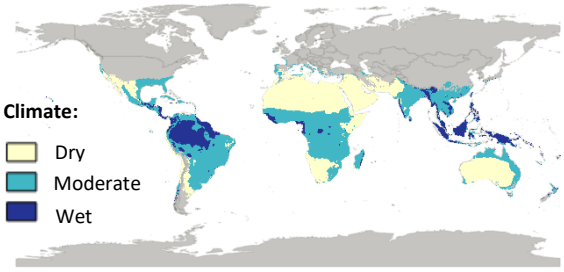

Figure 25: The global distribution of rainfall zones. (The zones reflect differences in annual precipitation)

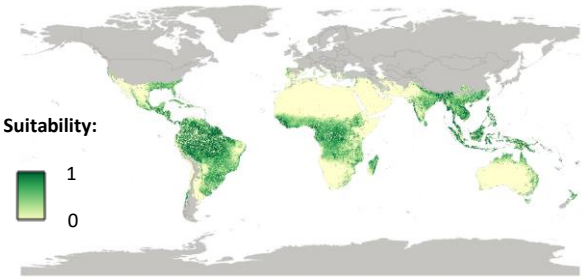

Figure 26: The combined suitability map for water table depth.

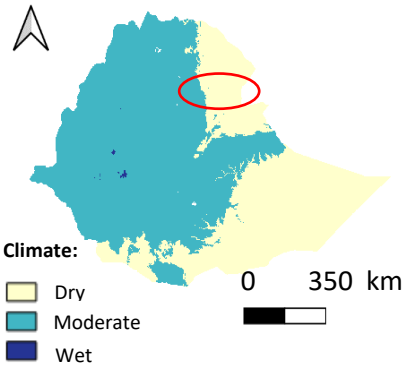

Figure 27: The distribution of rainfall zones in Ethiopia. (The red oval shows the extent of Figure 29)

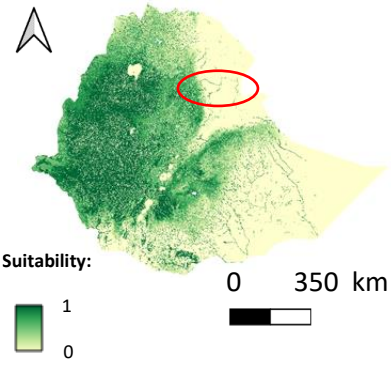

Figure 28: The combined suitability map for water table depth in Ethiopia. (The red oval shows the extent of Figure 29)

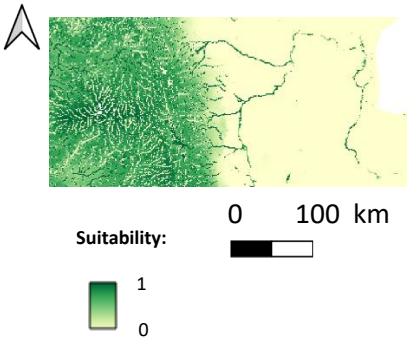

Figure 29: Close-up of the combined suitability map for water table depth in north-east Ethiopia.

This study acknowledged the complex interaction of the land characteristic water table depth with oxygen availability and moisture availability. Its impact on these two land qualities was made dependent on another land characteristic, namely annual precipitation. Previous studies followed a similar logic. For instance, Dubovyk et al. (2016) performed a fuzzy land suitability study for two different watersheds in Uzbekistan. They varied the parameters for membership functions according

to the land characteristics in the region. Furthermore, Juhos et al. (2019) did a fuzzy suitability study for arable soils in Hungary. Suitability values were determined based on quality indicators consisting of multiple variables. Texture suitability was for example calculated based on clay content and water table depth. The authors argued that texture suitability is only meaningful if water table depth is considered. In other words, they stated that intercorrelations between land characteristics should not be neglected. A similar argument was made in this study. Water table data was considered meaningful only when interpreted with precipitation data.

Precipitation data was used specifically to model the complex relationship between water table depth and land qualities. This decision was convenient because it required no additional data collection. There are however alternative options based on different indicators. For example, the climate moisture index (CMI) is a common measure that considers both precipitation and evapotranspiration (Malisawa and Rautenbach, 2012). The CMI consequently provides more information on the local water balance that precipitation data alone. This in turn gives a better idea of whether moisture availability or oxygen availability is the most limiting land quality at a certain location.

Fuzzy approaches are common practice for land suitability modelling. Several studies found that fuzzy logic simulates the continuous nature of landscape variation well (Ali et al., 2010, Atijosan et al., 2015, Joss et al., 2008, Salvacion, 2021). Another advantage is the resilience of fuzzy approaches to measurement uncertainty. The risk of error propagation with fuzzy approaches is low compared to discrete classification approaches (Qiu et al., 2014). In this study a rule-based fuzzy model was created with threshold values from literature. Many alternative approaches exist, however most of them require training data. For example, several studies did crop mapping by establishing a relationship between environmental conditions and field observations. The induced relationship was thereafter applied to all locations for which environmental data was available (Heumann et al., 2011, Ray et al., 2016). Such approaches are widespread in ecology and called niche-based models (Zhang et al., 2020, Panda and Behera, 2019, Phillips et al., 2006). Similarly, Mertens et al. (2021) and Borborah et al. (2020) recently mapped wild banana species with niche-based models in South-East Asia and India. The global applicability of niche-based modelling is however low due to the high training data requirements.

3.2.2. Final land suitability map

The final suitability map was constructed from the five individual suitability maps for annual precipitation, cold quarter temperature, dry quarter precipitation, soil organic matter and water table depth. Figure 30 shows the resulting map displayed at the same color scale as its constituents. In general, the final suitability map provides a smoother global pattern than the fuzzy maps for each land characteristic. Nevertheless, several distinctive features can still be observed. Dry quarter precipitation has a profound effect in the tropics. Locations with high dry quarter precipitation have significantly higher suitability values compared to neighboring regions. Soil organic matter has a strong effect in the (semi)arid regions of Australia and North-Africa. Suitability values larger than zero can be found here despite the unfavorable water table and climate conditions.

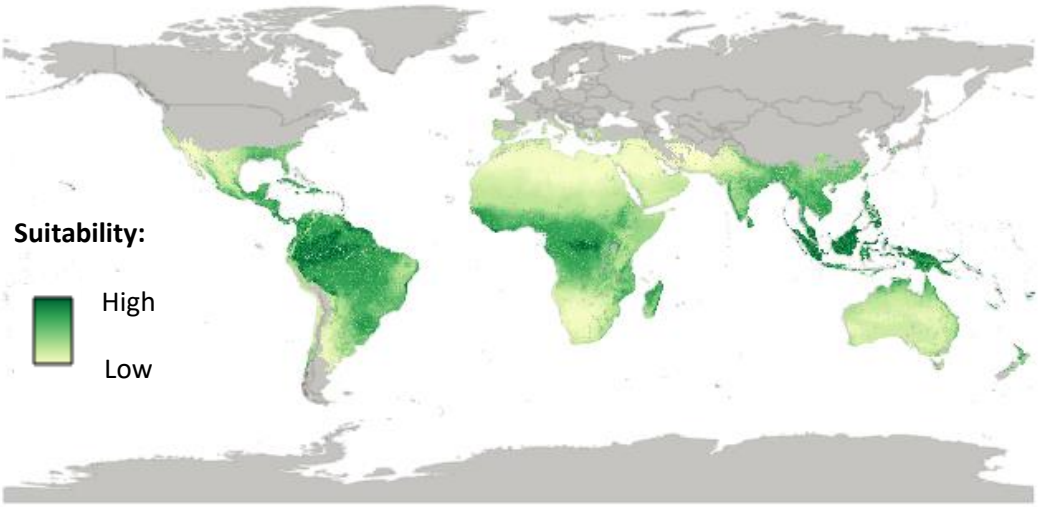

Figure 30: The final land suitability map. (All five constituent suitability maps were given an equal weight)

Each fuzzy suitability map of the five land characteristics was considered equally important. This is why the final suitability map in Figure 30 was created by the calculation of average grid cell values. Several alternative approaches for fuzzy inference exist (Qiu et al., 2014). For instance, one could aggregate different suitability values by taking the minimum or maximum value at each location. Such approaches could be appropriate in case of highly risk-averse or risk-tolerant suitability studies. A more balanced suitability map is obtained by calculating a weighted average. This approach considers all inputs but meanwhile provides flexibility to make certain inputs more dominant (Yager, 1988). A major challenge that comes with this approach is the definition of appropriate weight values. Kapoor et al. (2020) developed an inventive framework for the determination of weight factors in land suitability assessments. The authors calculated the percentage influence of different land characteristics on planning decisions based on literature and expert consultations. The data was thereafter translated to weighting factors through an analytic hierarchy process. In a similar way, weight factors could be derived from data on decision making by smallholder banana farmers. This would lead to a better understanding of farmer's behavior and thus the dominant factors that govern global banana patterns. At the same time, the limitations of such an approach are evident for studies on a global scale.

3.2.3. Disaggregated smallholder banana map

Finally, FAO production statistics were disaggregated according to the suitability results. Both banana and plantain statistics were used for this purpose. Figure 31 shows the harvested area of banana for different countries of the world. For some European countries the banana production turned out larger than expected. This can be explained by the formal definition of national territories. For example, the French statistics include production on the Caribbean islands Guadeloupe and Martinique (FAO, 2021a). The same applies for Spain, having its banana production concentrated in the Canaries (Cakmak et al., 2019, FAO, 2021a). The harvested area of plantain is displayed in Figure 32. The maps indicate that more countries have large banana surfaces than large plantain surfaces. Furthermore, several countries have either zero banana statistics or zero plantain statistics. The first case for example applies to Peru, Nigeria and Myanmar, the second case to Brazil, Angola and Indonesia. Such

a clear distinction is not realistic and originates from differences in data collection techniques among national surveys (FAO, 2021d). This observation reaffirms the decision to merge the FAO statistics for banana and plantain in this study.

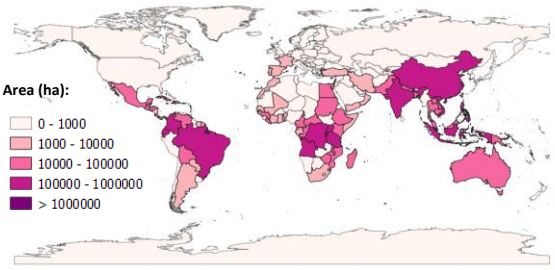

Figure 31: The harvested areas of banana for different countries in the world (FAO, 2021b)

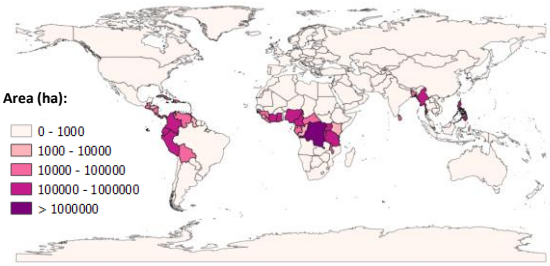

Figure 32: The harvested areas of plantain for different countries in the world (FAO, 2021b)

The disaggregated smallholder banana map is presented in Figure 33. The plantation banana map has also been projected on top to show both results together in one map (Figure 34) . The potential banana area with temperatures above 6 °C was added as a background layer. The smallholder banana map reveals the global distribution of banana at a fifteen arcsec (~ 500m) resolution. A strong concentration of smallholder banana around the equator is observed. In South America, smallholder banana is strongly concentrated between minus five and plus five degrees latitude. This is where dry quarter precipitation is highest and thus the growing conditions for banana are most suitable. Similar patterns appear for the rest of the world. Smallholder banana is strongly present in the coastal regions of West-African countries. Those regions have high water availability and high soil organic carbon compared to other parts of the countries. The distribution of smallholder banana in Southeast Asia reflects similar environmental gradients. Smallholder banana is predominantly found in the low-latitude coastal parts of countries within this region.

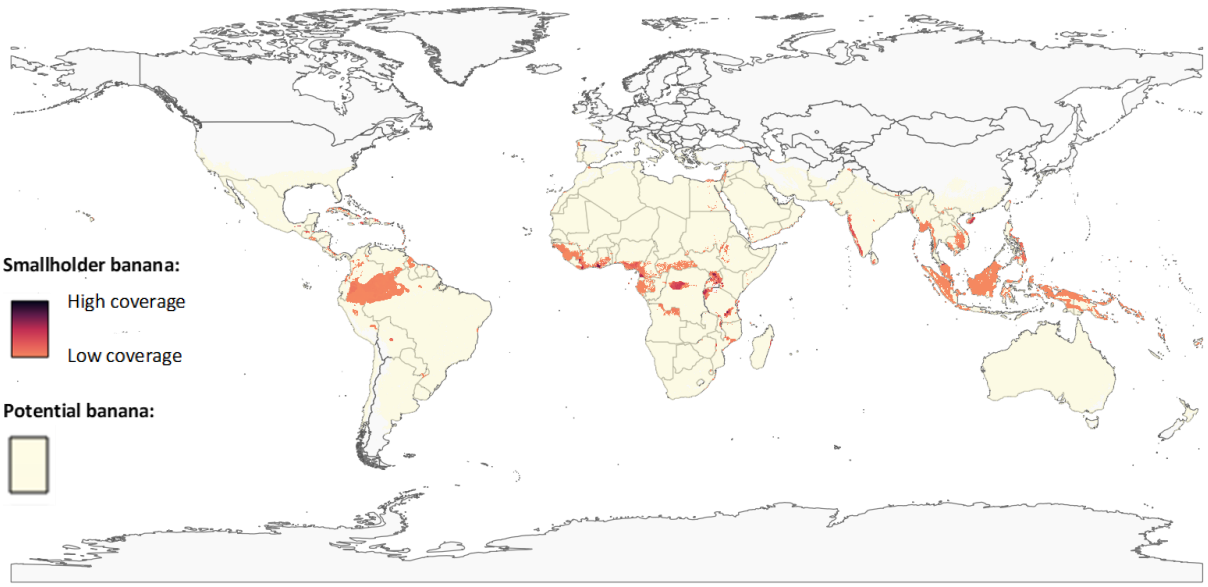

Figure 33: The global smallholder banana map (The yellowish background layer shows the world's potential banana area with cold quarter temperatures above 6 °C. The smallholder banana results have been aggregated to a 10km resolution for visualization purposes)

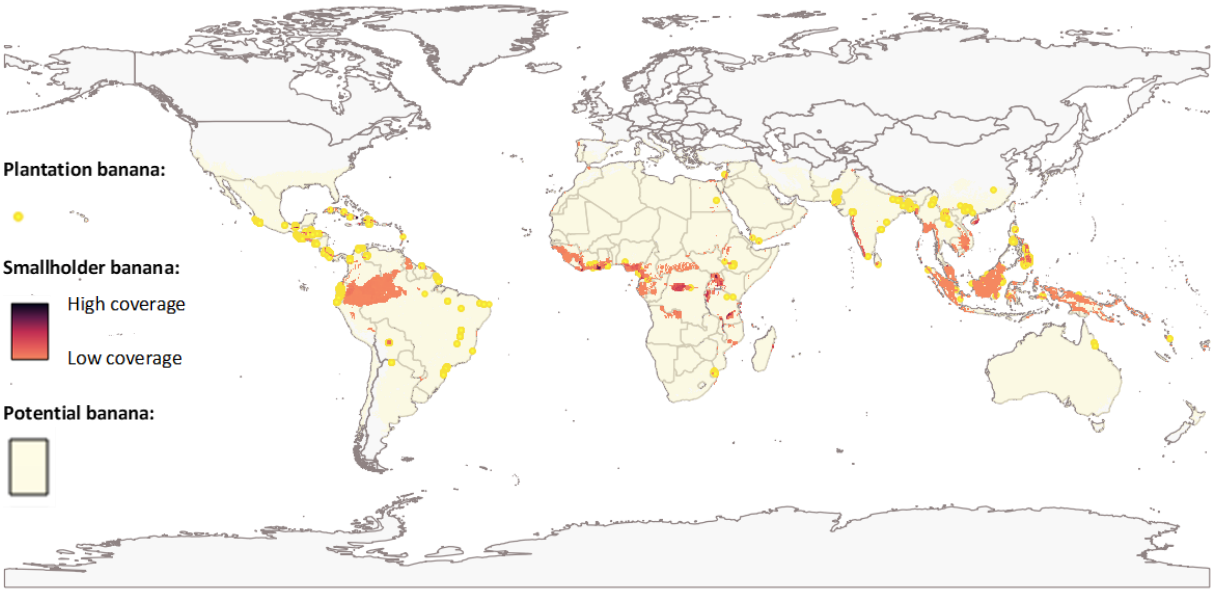

Figure 34: The global smallholder banana map with plantation banana mapped on top. (The yellowish background layer shows the world's potential banana area with cold quarter temperatures above 6 °C. The smallholder banana results have been aggregated to a 10km resolution for visualization purposes)

The FAO production statistics could be disaggregated for the majority of countries in the world. However, problems emerged for the nine island states listed in Table 9. The suitable area in these states was smaller than the banana surface according to the statistics. Moreover, no statistics at all were disaggregated for eight of the nine states. Further examination revealed that data on water table depth was missing for each of these eight states. The missing data caused NoData values in the final suitability map and statistics could therefore not be disaggregated. The island state of Kiribati forms a special case. Data on water table depth was missing for only part of the archipelago. Data was available around the other islands but did not align exactly with their extent. A comparison with satellite data showed that the extent of Kiribati was not presented well on the QGIS world map. This caused only marginal overlap with the environmental data and consequently a poor data coverage for Kiribati. These findings emphasize the strong impact of data quality on the final results.

Table 9: The island states for which no or not all productions statistics could be disaggregated.

Island states	Banana + plantain statistics (ha)	Final smallholder banana area (ha)
Cabo Verde	258	0
Cook Islands	5	0
French Polynesia	26	0
Kiribati	1629	551
Maldives	8	0
Mauritius	508	0
Niue	41	0
Samoa	5206	0
Tonga	858	0

The fuzzy suitability map captured subtle environmental gradients in landscapes. For some countries, however, this strongly reduced the extent of the smallholder banana area. A comparison between Indonesia and China illustrates this. The Indonesian production statistics were distributed over large parts of the country. This resulted in an extensive smallholder area with low banana coverage. The results correspond well to reality given that banana is cultivated across the country (Hermanto et al., 2011). Contrarily, the Chinese production statistics were entirely allocated to the Hainan province in the south. This is not realistic while actual banana occurrence extends to the whole subtropical part of South China (Thiers et al., 2019, Xiao-Lan et al., 2007). The fuzzy suitability map provided too much detail for a realistic disaggregation of production statistics in this case.

In this study, locations were only used for disaggregation if all locations with higher suitability were already satisfied. This means that grid cells were filled with smallholder banana from high to low suitability until no statistics were left. Grid cells got assigned a smallholder banana coverage between zero and forty percent, depending on their suitability. This specific disaggregation procedure sometimes produced very small banana areas, as was the case for China. Larger areas would be attained by lowering the banana coverage parameter. Banana coverage could also be linked to the prevailing smallholder production systems in different countries. Alternatively, a different disaggregation procedure could be applied to improve the results. A fixed suitability threshold would separate banana presence from banana absence independent of national production statistics. Such an approach is often referred to as defuzzification (Qiu et al., 2014). Production statistics could thereafter be distributed over locations with higher suitability than the threshold. Another idea would be to group suitability values into discrete classes before the disaggregation of production statistics. This would produce results with more general patterns, similar to conventional suitability mapping.

3.3. Validation of results

The results were validated through a visual comparison with satellite imagery and existing global banana maps. The visual comparison with satellite imagery was done for Venezuela, Tanzania, India and Australia. Those four countries were selected because they are part of different continents. In addition, neither of the countries were used for calibration. The comparison with existing global banana maps was done with the maps of Monfreda et al. (2008) and Onderwater (2020).

3.3.1. Comparison with satellite imagery

The results identify banana in north-west Venezuela (Figure 35). Satellite imagery reveals that the banana occurrences are located in the Maracaibo Lake basin (Figure 36). The basin is a wide depression in the landscape with extensive alluvial plains (Morales et al., 2001). Literature confirms the presence of banana plantations in this region (Gámez and Acconcia, 2009). A closer look on satellite imagery shows how the results clearly demarcate individual plantations (Figure 37). More detailed close-ups indicate that those plantations are banana (Figure 38). Although the imagery becomes blurry at high resolutions, a high resemblance with Costa Rican banana plantations is observed (Figure 39). The method presented in this study turned out suitable for the identification of plantation banana in Venezuela. Moreover, smallholder banana is projected in the same area based on land suitability

(Figure 36). This indicates favorable banana growing conditions in the Maracaibo Lake basin and consequently explains the presence of banana. A large share of smallholder banana is also placed in south Venezuela (Figure 35). It is however questionable whether smallholder banana is abundant in this region. Satellite imagery shows that south Venezuela is mainly covered by tropical rainforest. Agricultural activity is therefore most likely limited.

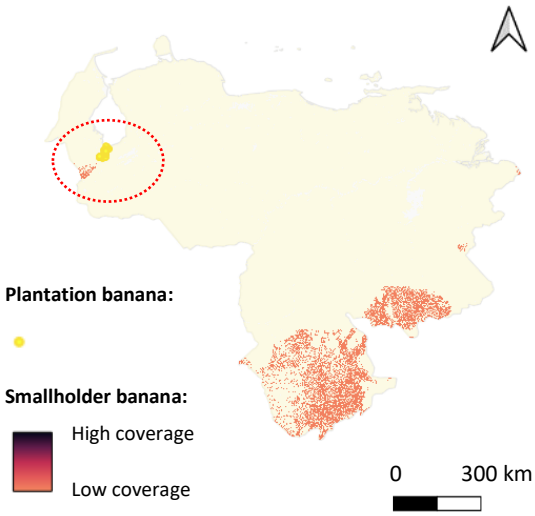

Figure 35: The final banana map of Venezuela. (The red oval indicates the extent of Figure 36)

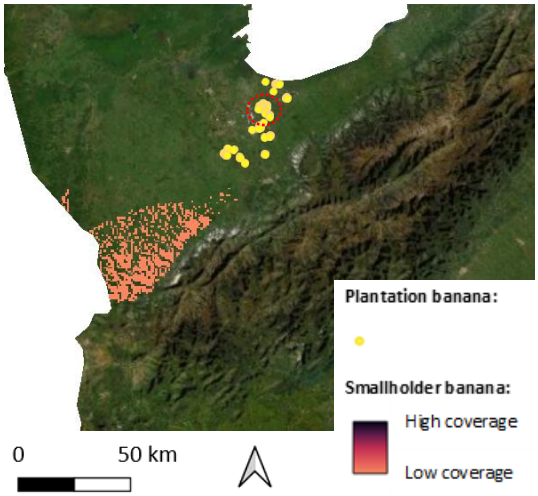

Figure 36: The results for the Maracaibo Lake basin in Venezuela. (The red oval indicates the extent of Figure 37)

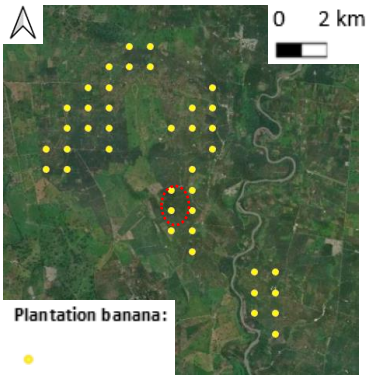

Figure 37: Plantation banana in the Maracaibo Lake basin. (The red oval indicates the extent of Figure 38)

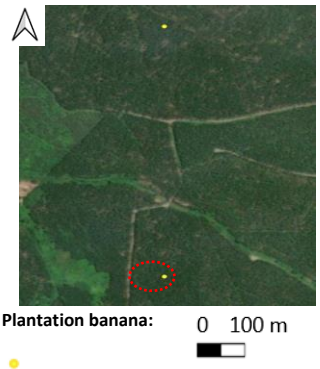

Figure 38: Close-up of plantation banana in the Maracaibo Lake basin. (The red oval indicates the extent of Figure 39a)

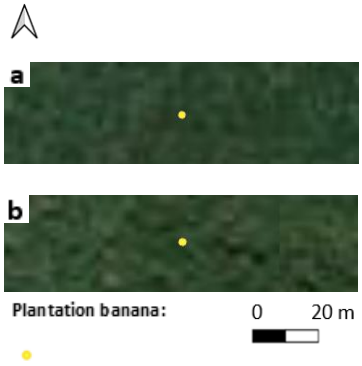

Figure 39: Comparison between a plantation in the Maracaibo Lake basin (a) and northeast Costa Rica (b).

The results for Tanzania are dominated by smallholder systems (Figure 40). Only one observation for plantation banana was made, located just north of Lake Manyara. Satellite imagery confirms that this is indeed a banana plantation. Smallholder banana is allocated to different regions across the country. In general, the resulting patterns match regional production statistics (Figure 41). Smallholder banana

is rightfully projected in most production regions. At the same time, smallholder banana remains indeed absent from the central and south-eastern part of the country. Banana occurrence in the northern Kilimanjaro region remains unfortunately undetected. This is a pity given that the region is among the top-three banana producing regions of the country (Suleiman, 2018). Environmental conditions on the slopes of Mt. Kilimanjaro are sufficiently suitable for the cultivation of banana. The high population pressure in the area have resulted in intensive intercropping systems with coffee and banana (Soini, 2005). The example of the Kilimanjaro region illustrates that population dynamics also have an influence on crop production, in addition to biophysical suitability.

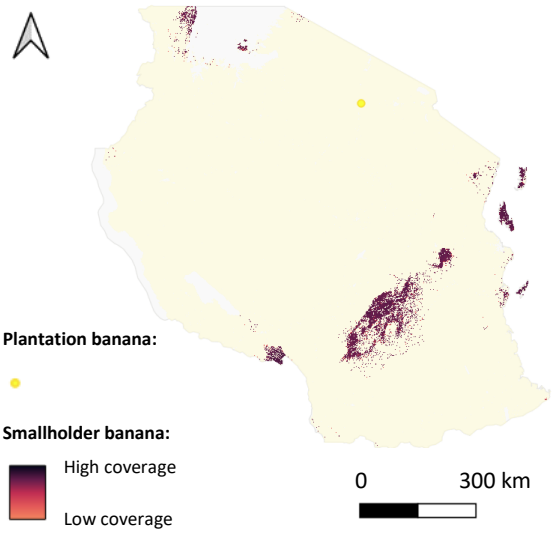

Figure 40: The final banana map of Tanzania.

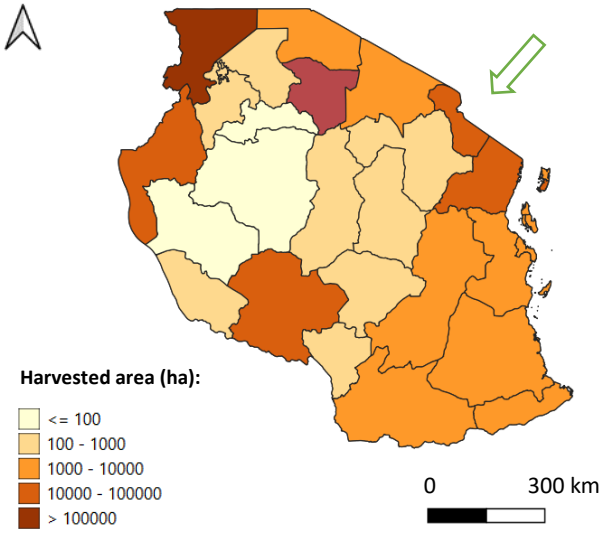

Figure 41: Regional banana production statistics of Tanzania (NBS Tanzania, 2021). (The green arrow point to the Kilimanjaro region)

India has the largest banana surface of all countries in the world (FAO, 2021b). The main banana producing regions are found in the southern and western part of the country (Singh, 2010). The results indicate banana occurrence in these regions, but also in the eastern part of the country (Figure 42). Satellite imagery confirms that plantation banana was classified correctly. However, the plantation banana surface is quite low given the high production statistics of India. Especially in the production zones of south India more observations were expected. For example, the surroundings of Jalgaon in Maharashtra state are known for their high production (Singh, 2021). Satellite imagery indeed shows a belt of intensive agriculture north-east of the city (Figure 43a). In addition, the SAR data indicates a backscatter signal that is highly similar to banana (Figure 43b). Most potential banana locations were however removed after application of the VH filter (Figure 43c). The remaining locations thereafter got lost during the conversion process due to a lack of neighbors. Satellite imagery reveals that banana plots in the region are abundant but small in size. The region is characterized by alternating spatial patterns of banana and other crops. This results in mixed SAR signals with VH backscatter values exceeding the threshold. The example of Maharashtra state points out a common limitation of global crop mapping. The VH filter strongly improved results for Indonesia by reducing the confusion with

palm oil. The same filter simultaneously decreased the results for India. Global decision rules often have different, if not contrasting effects in different parts of the world.

The Jalgaon region is also absent on the smallholder banana map. The majority of the production statistics is disaggregated to the coastal regions of Western India (Figure 42). This area has high suitability values due to high annual precipitation. Analogously, the Jalgaon district has low suitability values due to constraints on water availability. The suitability analysis however ignores that irrigation is common practice among smallholder farmers in Northern Maharashtra (Gorain et al., 2020, Gorain et al., 2018). This makes that the actual land suitability in Jalgaon district is underestimated. The example indicates that intensive smallholder production systems are currently not detected.

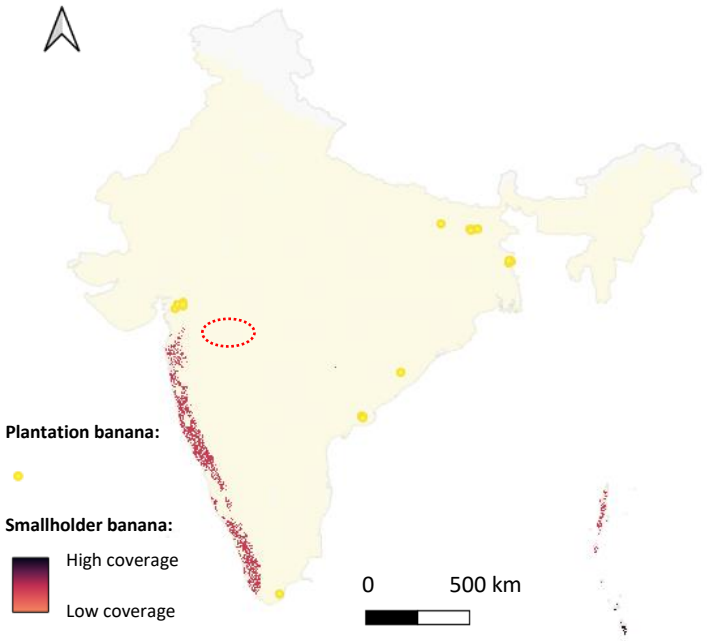

Figure 42: The final banana map of India. (The red oval indicates the extent of Figure 44)

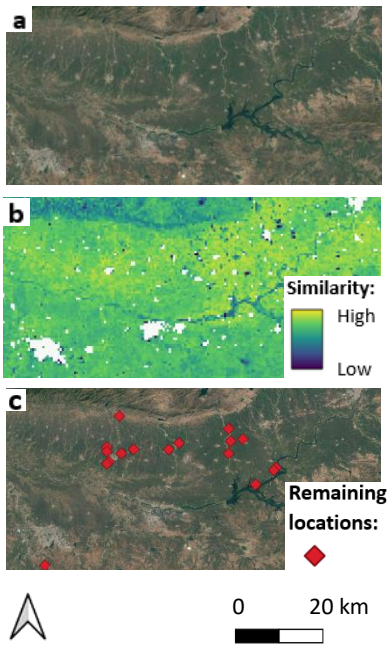

Figure 43: The results for the Jalgaon area in India. Subfigures: a) Satellite image, b) SAR similarity map, c) filtered SAR similarity map.

The Australian banana production is concentrated around the cities Tully and Innisfail in north-east Queensland (Daniells, 1984). This is clearly reflected by the results of this study (Figure 44). Plantation banana was detected in the Tully region (Figure 45 & 46). Furthermore, all smallholder banana is allocated to the coastal area between Tully and Innisfail (Figure 45). This is remarkably accurate given the large size of Australia. In some cases, the predicted smallholder banana areas even overlap with actual banana plantations (Figure 47). Banana production in Australia appears strongly connected to land suitability. The results show that the methods for mapping plantation and smallholder banana were both very accurate for Australia.

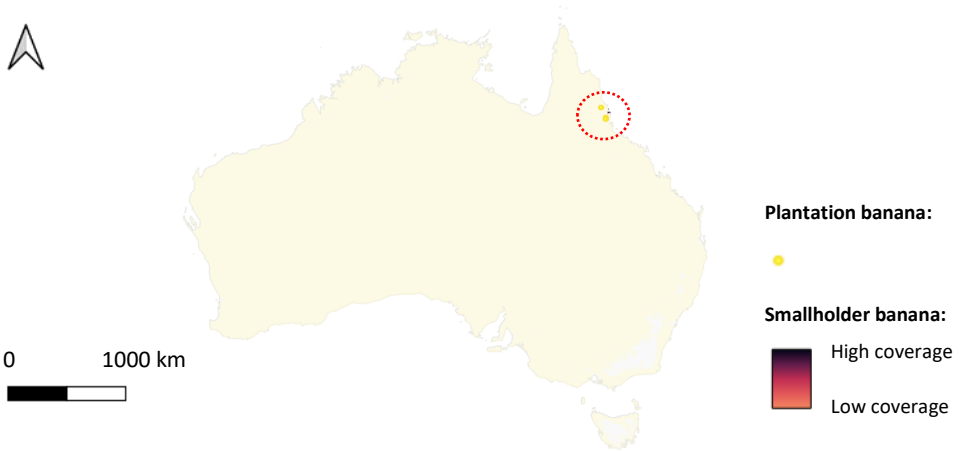

Figure 44: The final banana map of Australia. (The red oval indicates the extent of Figure 46)

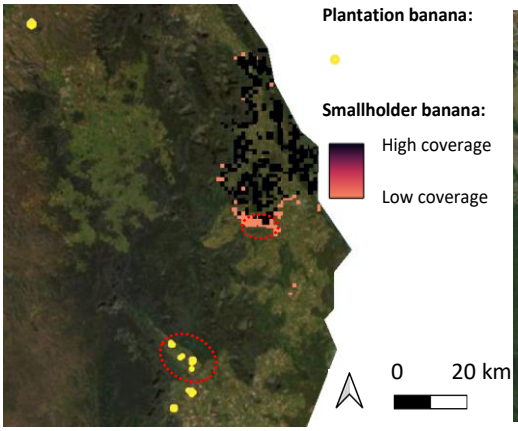

Figure 45: The results for the Tully region, north-east Queensland. (The red ovals indicate the extents of Figures 46 and 47)

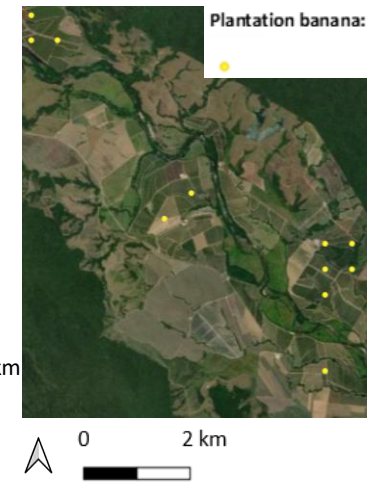

Figure 46: Plantation banana near Tully.

Figure 47: Comparison between the results (a) and satellite imagery (b) for smallholder banana near Tully.

3.3.2. Comparison with global banana maps

The map by Monfreda et al. (2008) was one of the first global banana map ever made. The authors disaggregated agricultural inventory data based on a global cropland map at a five arcmin (~ 10km) resolution. The global cropland map was created by Ramankutty et al. (2008) and showed which fraction of grid cells was covered by cropland. The map was created through a regression analysis between agricultural census data and landcover products from the MODIS and SPOT satellites. Monfreda et al. (2008) used subnational statistics to calculate the banana fraction in different administrative units with respect to total cropland in the same area. The fractions were thereafter multiplied with the percentage agricultural cover in each grid cell. Within provinces or states, grid cells with a large percentage of agriculture consequently got assigned a large banana area. This approach resulted in an extensive banana area with small banana coverage in the order of magnitude from 10^{-6} to 10^{-4} for most parts of the world (Figure 48).

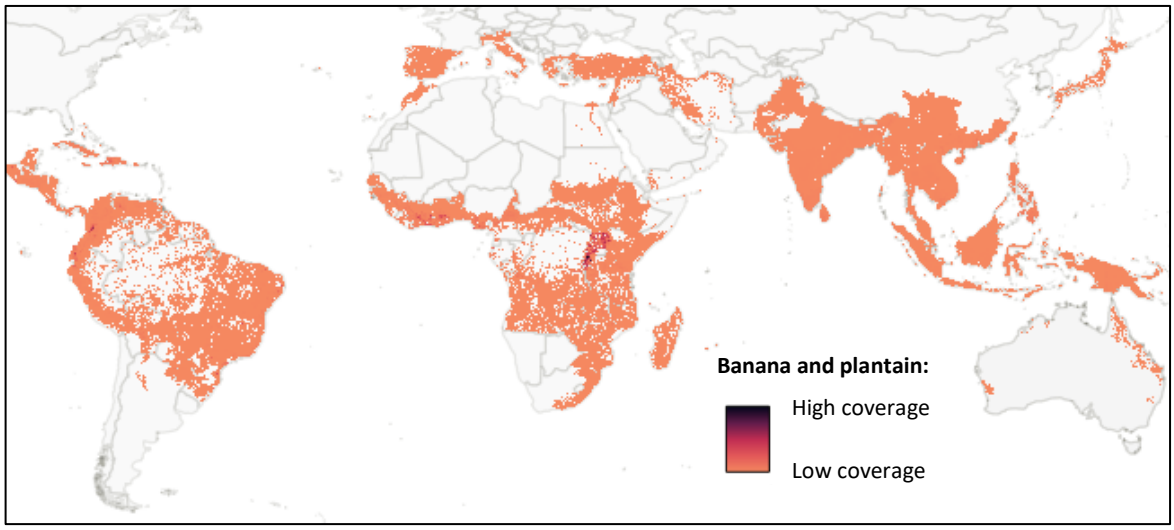

Figure 48: The global banana map by Monfreda et al. (2008). (Values range between 10^{-6} to 10^{-4} for most parts of the world)

Onderwater (2020) refined the disaggregation process by introducing land suitability concepts. The author first narrowed the potential banana area through a land use analysis based on landcover, NDVI and canopy height. A land suitability map was thereafter created according to the FAO guidelines for land evaluation. The suitability analysis was based on several biophysical variables, comprising climate, slope, drainage and soil type. The traditional land evaluation approach yielded a suitability map with four discrete classes. National production statistics were disaggregated according to this map. The disaggregation procedure was highly similar to the procedure applied in this study. In short, locations with superior suitability classes were considered first before moving to less suitable locations. The results show a global banana distribution confined to a smaller part of the earth's surface compared to Monfreda et al. (2008). In this way, Onderwater (2020) produced more informative spatial banana distribution patterns at a thirty arcsec (~ 1km) resolution (Figure 49).

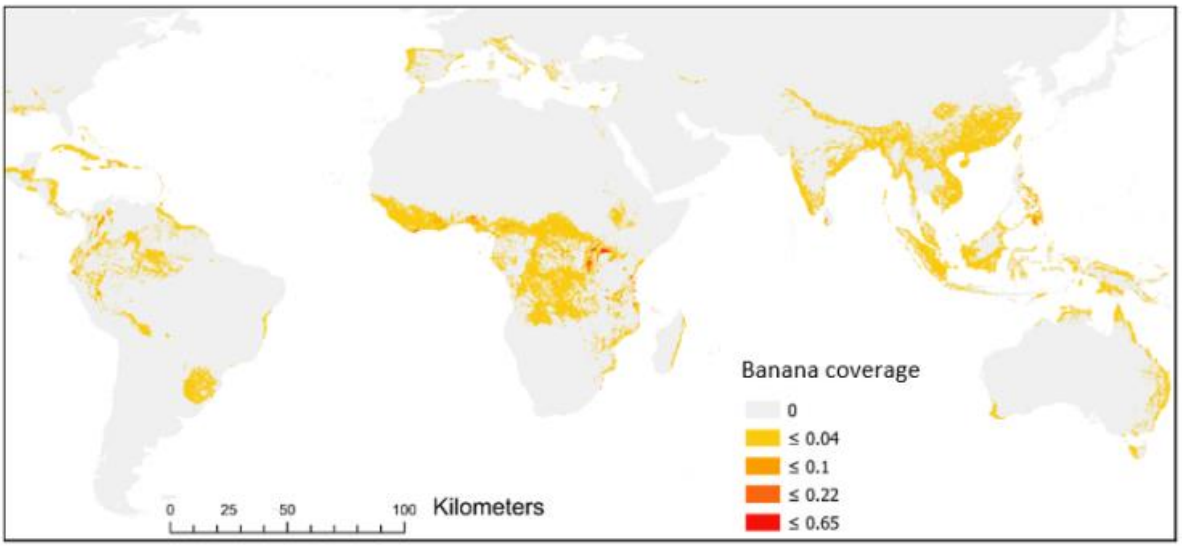

Figure 49: The global banana map by Onderwater (2020). (Values represent banana coverage on a 10km aggregated grid).

Our study differs in several aspects from its predecessors. First of all, the spatial resolution was increased to fifteen arcsec (~ 500m). Secondly, different methods were developed to map banana in large-scale and smallholder production systems. Onderwater (2020) already acknowledged the impact of production characteristics on crop distributions. His work involved different suitability thresholds for countries based on their agricultural production intensity. Such an approach however neglects the coexistence of different production systems within countries (Bijker, 2014). This study therefore developed a framework that subdivides banana occurrence into plantation banana and smallholder banana. In this way, different methods could be applied that match the specific characteristics of either system. For instance, about 6.5 % of the global banana surface was detected with the satellite-based method for plantation banana. This added great detail to the global banana map (Figure 50).

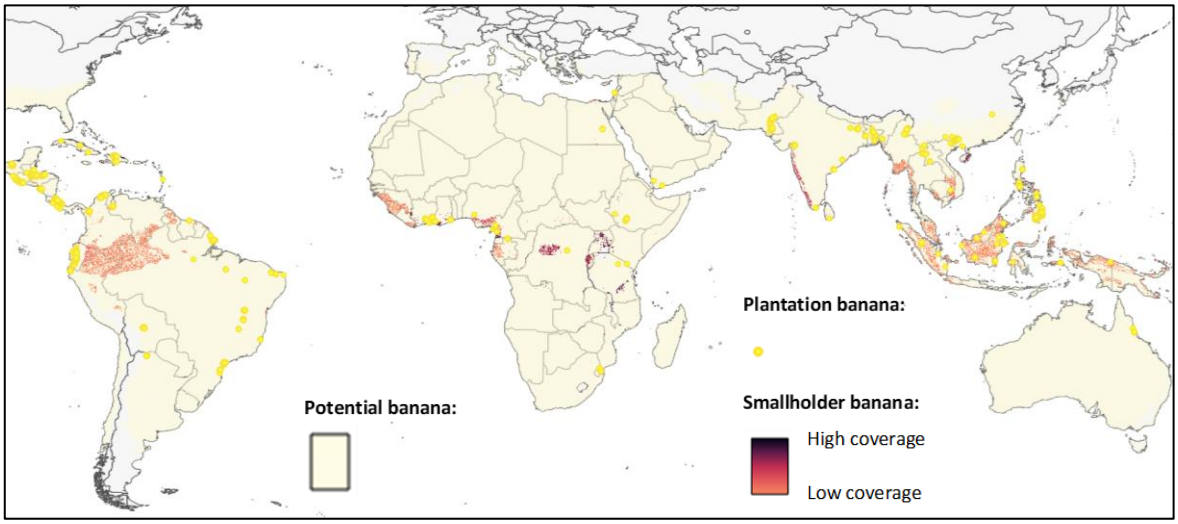

Figure 50: The global banana map in this study. (See Figure 34 for the full global extent. The yellowish background layer shows the world's potential banana area with cold quarter temperatures above 6 °C. The plantation banana grid cells are represented by points for visualization purposes)

Furthermore, our study applied a fuzzy approach for mapping land suitability based on environmental data only. Contrary to Onderwater (2020), the potential banana area was not reduced based on remote sensing data. Such a filter was considered inappropriate given the heterogeneity in smallholder landscapes. The fuzzy approach resulted in very detailed distribution patterns for smallholder banana. The patterns cover only a fraction of the extent predicted by Monfreda et al. (2008). This reaffirms the added value of disaggregation through suitability as already shown by Onderwater (2020). Compared to the latter, the smallholder banana extent was much smaller in this study. The fuzzy approach captured the continuous suitability gradients in landscapes. For some countries like Australia this resulted in highly accurate results (see Chapter 3.3.1.). For other countries like China, however, the patterns are questionable (see Chapter 3.1.3.). It is therefore difficult to say whether the fuzzy approach outperformed the conventional land evaluation.

Further research is needed to evaluate the global applicability of fuzzy land suitability for the disaggregation of production statistics. The challenge would be to generate more reasonable patterns for some countries while preserving the detailed patterns for others. This may be achieved by adjusting several model parameters or the disaggregation procedure as a whole (see Chapter 3.2.3). To conclude, Monfreda et al. (2008) used subnational production statistics one or two administrative levels below the national whenever available. Such subnational data could improve the results of a disaggregation with fuzzy suitability, especially within large countries.

3.4. General discussion

Knowing which crops are grown where is crucial information (Giri, 2012). Ongoing developments in GIS and remote sensing have expanded the opportunities for crop mapping over the past decades. Satellite data at large spatiotemporal resolutions is nowadays available for global applications. This has enabled object-oriented land use inventories way beyond the subnational zones proposed by Huising (1993). Global crop mapping with remote sensing has however its limitations. The detection of specific crop types relies on the detection of unique reflectance characteristics. This requires homogenous crop surfaces with dimensions that exceed the spatial scale of analysis (Schmidtlein et al., 2014). At a global level, most crops are not exclusively cultivated in homogenous systems. For most crops a wide variety of production systems exists, ranging from large monocultures to smallholder polycultures. It may be questioned whether direct crop mapping with satellite data will ever work for all farming systems in the world (Huising, 1993). Global crop mapping consequently asks for a more complex, integrated approach.

This study presented a framework for mapping plantation and smallholder crops with different methods. The agroecological interactions between people and their environment differ substantially in large-scale and smallholder production systems (Dalgaard et al., 2003). The aforementioned distinction between large-scale and smallholder was therefore considered crucial to map global crop distributions (Rabbinge, 1997). Large cropping surfaces can be mapped through remote sensing. This study presented a method for the processing of SAR satellite data through statistical distance analysis. Some supplementary satellite products were required to filter the results and obtain satisfactory results. The method was successfully applied to banana and designed for the application to other crops. Mapping smallholder banana turned out more challenging. A method was developed for the disaggregation of production statistics through fuzzy land suitability. The underlying assumption was that crops are cultivated in areas with high biophysical suitability. Indeed, several studies confirm the general relationship between biophysical land suitability and crop occurrence (Yengoh and Ardö, 2014, Jamil et al., 2018, Heumann et al., 2013).

A suitability-based approach has however several limitations. First of all, land suitability is not mutually exclusive among different crops. Areas that are suitable for crop A may be equally suitable for crop B. This poses a challenge for mapping specific crop types. Secondly, biophysical suitability explains only part of smallholder farmers' decision-making behavior. Møller et al. (2021) observed large discrepancies between biophysical suitability and prevailing land use patterns in Denmark. The authors call for a proper definition of land suitability, covering both ecological and socioeconomic aspects. Several studies incorporated socioeconomic factors for crop mapping, ranging from household characteristics and assets to market accessibility and prizes (Heumann et al., 2013, You et al., 2014). Similar approaches could improve the results for intensive smallholder production systems like those in India (Chapter 3.3.1.). Thirdly, suitability assessments neglect human presence as the ultimate precondition for agricultural activity. The results for Venezuela and Tanzania in this study provide a good example (see Chapter 3.3.1.). The disregard of demographics is inconvenient given that population density has a major influence on both land use and production intensity (Letourneau et al., 2012). At the same time, it should be noted that the inclusion of socioeconomics and demographics in crop mapping may have undesirable side effects. In this way, occurrences of crop wild relatives will no longer appear in the results. This could be a disadvantage for crop mapping studies, depending on the user application and crop type. In this study, a broad focus on domesticated and wild banana was preferred given the relevance of both for the transmittance of diseases.

A major challenge in contemporary crop mapping is to develop methods for the exploitation of big earth observation data (Corbane et al., 2017). Satellite platforms nowadays offer a wealth of information at various spatiotemporal scales. This study presents a straightforward method for mapping large-scale cropping surfaces with open-source data and software. Further research should confirm the applicability of the method to crop mapping in general. Implementations at higher spatial resolutions may greatly improve the results. In addition, the benefits of integrating multi-temporal and multi-source data could be explored (Khosravi and Alavipanah, 2019). The method for mapping large-scale cropping surfaces in this study is rule-based and requires only few calibration data. The choice to develop such a method was made intentionally given that calibration data for global crop mapping is often scarce. Future studies could benefit from more data-driven approaches if extensive calibration data becomes available. In this context, the results of this study could even be used to find appropriate training data at a global level.

Remote sensing will become ever more important for global crop mapping when processing opportunities expand. At the same time, land suitability assessments will remain a reasonable alternative for mapping mixed smallholder systems. The method for mapping smallholder cropping surfaces in this study leaves room for improvement. The recalibration of model parameters may result in more informative global patterns. In addition, the effect of including socio-economic and demographic variables could be explored. Another focus point would be the use of subnational statistics within the disaggregation process. To conclude, universal applicability should always be the cornerstone of global crop mapping. Future studies should explicitly consider the trade-off between model complexity and global applicability. Methodological frameworks should be as generic as possible, albeit not oversimplifying the complex global system. In this study, a balance was found by integrating methods for large-scale and smallholder systems. Future crop mapping studies may benefit from a similar approach.

4. Conclusion

- This study demonstrates how global spatial patterns in agriculture can be mapped at a fifteen arcsec resolution. An integrated framework for global crop mapping was developed that combines remote sensing, production statistics and land suitability concepts.

- The framework was applied to banana and yielded a global banana map with applications for monitoring the spread of diseases. The methodological concepts are generic and could be applied to other crops.

- Special attention was paid to the coexistence of different crop production systems within landscapes. It was argued that fundamentally different systems ask for different mapping approaches.

- The study illustrates how large cropping surfaces can be mapped with remote sensing at a global level. Polarimetric SAR data was processed through a statistical distance method. The method leaves room for the adjustment of input parameters according to users' intentions.

- The results of this study indicate both opportunities and challenges of a fuzzy suitability approach for mapping smallholder crops. Further research should reveal how the inclusion of demographic and socio-economic variables could improve the results.

- Developments in GIS and remote sensing expand the processing opportunities for global crop mapping. Open source geodata at fine spatiotemporal resolutions is nowadays available. Advances in computational power make GIS applications accessible to all. Future studies should develop methods that exploit the wealth of data while serving a broad audience.

References

AHANI, H. & NOSHADI, M. 2019. Application of Cadastre Maps, Agricultural Database and MODIS Satellite Images for Monitoring Cultivated Areas. *Iranian Journal of Science and Technology, Transactions of Civil Engineering,* 43**,** 179-192.

AIRBUS 2015. TerraSAR-X Image Product Guide - Basic and Enhanced Radar Satellite Imagery. Airbus Defence and Space.

AKINYEMI, S. O. S., AIYELAAGBE, I. O. O. & AKYEAMPONG, E. PLANTAIN (MUSA SPP.) CULTIVATION IN NIGERIA: A REVIEW OF ITS PRODUCTION, MARKETING AND RESEARCH IN THE LAST TWO DECADES. 2010. International Society for Horticultural Science (ISHS), Leuven, Belgium, 211-218.

ALI, K., SARMADIAN, F., HEIDARI, A. & OMID, M. 2010. Land Suitability Evaluation Using Fuzzy Continuous Classification (A Case Study: Ziaran Region). *Modern Applied Science,* 4**,** 72-81.

ARINO, O., RAMOS, J., KALOGIROU, V., DEFOURNY, P., ACHARD, F. 2010. GlobCover 2009. *ESA Living Planet Symposium.* Bergen, Norway.

ATIJOSAN, A., BADRU, R., OGUNYEMI, S. & A, A. 2015. Agricultural Land Suitability Assessment using Fuzzy Logic and Geographic Information System Techniques. 1**,** 113-118.

BAKRY, F. & HORRY, J. P. Advances in genomics: applications to banana breeding. 2016. International Society for Horticultural Science (ISHS), Leuven, Belgium, 171-180.

BAN, Y., JACOB, A. & GAMBA, P. 2015. Spaceborne SAR data for global urban mapping at 30m resolution using a robust urban extractor. *ISPRS Journal of Photogrammetry and Remote Sensing,* 103**,** 28-37.

BEBBER, D. P., HOLMES, T. & GURR, S. J. 2014. The global spread of crop pests and pathogens. *Global Ecology and Biogeography,* 23**,** 1398-1407.

BIFARIN, J. O., ALIMI, T., BARUWA, O. I. & AJEWOLE, O. C. DETERMINANT OF TECHNICAL, ALLOCATIVE AND ECONOMIC EFFICIENCIES IN THE PLANTAIN (MUSA SPP.) PRODUCTION INDUSTRY, ONDO STATE, NIGERIA. 2010. International Society for Horticultural Science (ISHS), Leuven, Belgium, 199-209.

BIJKER, F. 2014. *Describing land use variability : using a multi-scale approach to analyse the banana production landscape of Costa Rica in the light of the threat of Panama disease.* Wageningen University.

BLOMME, G., PLOETZ, R., JONES, D., DE LANGHE, E., PRICE, N., GOLD, C., GEERING, A., VILJOEN, A., KARAMURA, D., PILLAY, M., TINZAARA, W., TEYCHENEY, P.-Y., LEPOINT, P., KARAMURA, E. & BUDDENHAGEN, I. 2013. A historical overview of the appearance and spread of Musa pests and pathogens on the African continent: highlighting the importance of clean Musa planting materials and quarantine measures. *Annals of Applied Biology,* 162**,** 4-26.

BORBORAH, K., DEKA, K., SAIKIA, D., BORTHAKUR, S. K. & TANTI, B. 2020. Habitat distribution mapping of Musa flaviflora Simmonds - a wild banana in Assam, India. *Acta Ecologica Sinica,* 40**,** 122-127.

BRERETON, R. G. 2011. One-class classifiers. *Journal of Chemometrics,* 25**,** 225-246.

BRÜNTRUP, M., SCHWARZ, F., ABSMAYR, T., DYLLA, J., ECKHARD, F., REMKE, K. & STERNISKO, K. 2018. Nucleus-outgrower schemes as an alternative to traditional smallholder agriculture in Tanzania – strengths, weaknesses and policy requirements. *Food Security,* 10**,** 807-826.

BUCHHORN, M., LESIV, M., TSENDBAZAR, N.-E., HEROLD, M., BERTELS, L. & SMETS, B. 2020. Copernicus Global Land Cover Layers—Collection 2. *Remote Sensing,* 12**,** 1044.

BUCKLEY, S. M., AGRAM, P. S., BELZ, J. E., CRIPPEN, R. E., GURROLA, E. M., HENSLEY, S., KOBRICK, M., LAVALLE, M., MARTIN, J. M., NEUMANN, M., NGUYEN, Q. D., ROSEN, P. A., SHIMADA, J. G., SIMARD, M. & TUNG, W. W. 2020. NASADEM. NASA-JPL.

CAKMAK, T., ANA PIEDRA-BUENA, D., ESTRELLA HERNÁNDEZ, S. & CARLOS ALVAREZ, A. 2019. Chrysodeixis chalcites (Esper) (Lepidoptera: Noctuidae) oviposition preferences on different

growing stages of banana (Musa acuminate Colla, Musaceae) plants. *Phytoparasitica,* 47**,** 485-498.
CHAMBERT, T., MILLER, D. A. W. & NICHOLS, J. D. 2015. Modeling false positive detections in species occurrence data under different study designs. *Ecology,* 96**,** 332-339.
CHEN, S.-W., WANG, X.-S., XIAO, S.-P. & SATO, M. 2018. Fundamentals of Polarimetric Radar Imaging and Interpretation. *Target Scattering Mechanism in Polarimetric Synthetic Aperture Radar : Interpretation and Application.* Singapore: Springer Singapore.
CHURCHILL, A. C. L. 2011. Mycosphaerella fijiensis, the black leaf streak pathogen of banana: progress towards understanding pathogen biology and detection, disease development, and the challenges of control. *Molecular Plant Pathology,* 12**,** 307-328.
CONNORS, B. M., COOPER, A. B., PETERMAN, R. M. & DULVY, N. K. 2014. The false classification of extinction risk in noisy environments. *Proceedings: Biological Sciences,* 281**,** 1-10.
CORBANA 2019. Banana Map of North-East Costa Rica.
CORBANE, C., PESARESI, M., POLITIS, P., SYRRIS, V., FLORCZYK, A. J., SOILLE, P., MAFFENINI, L., BURGER, A., VASILEV, V., RODRIGUEZ, D., SABO, F., DIJKSTRA, L. & KEMPER, T. 2017. Big earth data analytics on Sentinel-1 and Landsat imagery in support to global human settlements mapping. *Big Earth Data,* 1**,** 118-144.
DALGAARD, T., HUTCHINGS, N. & PORTER, J. 2003. Agroecology, scaling and interdisciplinarity. *Agriculture, Ecosystems & Environment,* 100**,** 39-51.
DANIELLS, J. W. 1984. The banana industry in North Queensland. *Queensland Agricultural Journal,* September-October**,** 282-290.
DE MAESSCHALCK, R., JOUAN-RIMBAUD, D. & MASSART, D. L. 2000. The Mahalanobis distance. *Chemometrics and Intelligent Laboratory Systems,* 50**,** 1-18.
DEFRIES, R. & TOWNSHEND, J. 1994. NDVI-Derived Land Cover Classification at a Global Scale. *International Journal of Remote Sensing - INT J REMOTE SENS,* 15**,** 3567-3586.
DELVAUX, B. 1995. Soils. *In:* GOWEN, S. (ed.) *Bananas and Plantains.* Dordrecht: Springer Science+Business Media.
DUBOVYK, O., MENZ, G. & KHAMZINA, A. 2016. Land Suitability Assessment for Afforestation with Elaeagnus Angustifolia L. in Degraded Agricultural Areas of the Lower Amudarya River Basin. *Land Degradation & Development,* 27**,** 1831-1839.
EAMUS, D., FU, B., SPRINGER, A. E. & STEVENS, L. E. 2016. Groundwater Dependent Ecosystems: Classification, Identification Techniques and Threats. *In:* JAKEMAN, A. J., BARRETEAU, O., HUNT, R. J., RINAUDO, J.-D. & ROSS, A. (eds.) *Integrated Groundwater Management: Concepts, Approaches and Challenges.* Cham: Springer International Publishing.
ESA 2020. Copernicus Sentinel Data. *In:* ESA (ed.).
EVANS, E. A., BALLEN, F. H. & SIDDIQ, M. 2020. Banana Production, Global Trade, Consumption Trends, Postharvest Handling, and Processing. *Handbook of Banana Production, Postharvest Science, Processing Technology, and Nutrition.*
FAN, Y., LI, H. & MIGUEZ-MACHO, G. 2013. Global patterns of groundwater table depth. *Science,* 339**,** 940-943.
FAO 1976. *A framework for land evaluation,* Rome, Food and Agriculture Organization of the United Nations.
FAO 2017. The future of food and agriculture – Trends and challenges. Rome.
FAO 2020. FAOSTAT - Crops. 22/12/2020 ed.
FAO 2021a. Banana market review – Preliminary results 2020. Rome.
FAO 2021b. FAOSTAT - Production - Crops and livestock products.
FAO 2021c. FAOSTAT - Trade - Crops and livestock products.
FAO 2021d. Methodology - Crops and Livestock. FAOSTAT.
FAO, IFAD, UNICEF, WFP & WHO 2020. Progress towards hunger and food insecurity targets. *The state of food security and nutrition in the world 2020: Transforming food systems for affordable healthy diets.* Rome.

FOX, J. & CASTELLA, J.-C. 2013. Expansion of rubber (Hevea brasiliensis) in Mainland Southeast Asia: what are the prospects for smallholders? *The Journal of Peasant Studies,* 40**,** 155-170.

FRETWELL, P. T., STANILAND, I. J. & FORCADA, J. 2014. Whales from Space: Counting Southern Right Whales by Satellite. *PLOS ONE,* 9**,** e88655.

GAMBA, P. & LISINI, G. 2013. Fast and Efficient Urban Extent Extraction Using ASAR Wide Swath Mode Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* 6**,** 2184-2195.

GÁMEZ, J. & ACCONCIA, R. 2009. Ecological informations on Coprophanaeus (Coprophanaeus) gamezi Arnaud (Coleoptera: Scarabaeidae: Phanaeini) in a rural agricultural area in the depression of Maracaibo, Zulia State, Venezuela. *Acta zoológica mexicana,* 25**,** 387-396.

GARCÍA-BASTIDAS, F. A., KEMA, G. H. J. P. D., ARANGO-ISAZA, R. D., GARCÍA-BASTIDAS, F. A., KEMA, G. H. J. P. D. & ARANGO-ISAZA, R. D. 2019. *Panama disease in banana : spread, screens and genes.* Wageningen University.

GARCÍA-BASTIDAS, F. A., QUINTERO-VARGAS, J. C., AYALA-VASQUEZ, M., SCHERMER, T., SEIDL, M. F., SANTOS-PAIVA, M., NOGUERA, A. M., AGUILERA-GALVEZ, C., WITTENBERG, A., HOFSTEDE, R., SØRENSEN, A. & KEMA, G. H. J. 2020. First report of Fusarium Wilt Tropical race 4 in Cavendish Bananas Caused by Fusarium odoratissimum in Colombia. *Plant Disease,* 104**,** 994.

GIRI, C. P. 2012. *Remote sensing of land use and land cover : principles and applications,* Boca Raton (FL), USA, CRC Press.

GOMEZ SELVARAJ, M., VERGARA, A., MONTENEGRO, F., ALONSO RUIZ, H., SAFARI, N., RAYMAEKERS, D., OCIMATI, W., NTAMWIRA, J., TITS, L., OMONDI, A. B. & BLOMME, G. 2020. Detection of banana plants and their major diseases through aerial images and machine learning methods: A case study in DR Congo and Republic of Benin. *ISPRS Journal of Photogrammetry and Remote Sensing,* 169**,** 110-124.

GORAIN, S., SINGH, D. R., KUMAR, P., VENKATESH, P. & JHA, G. K. 2018. Social Costs and Benefits Analysis of Drip Irrigation System in Northern Maharashtra. *Economic Affairs,* 63**,** 1061-1065.

GORAIN, S., SINGH, D. R., KUMAR, P., VENKATESH, P. & JHA, G. K. 2020. Economics of Sugarcane and Banana Cultivation under Drip Irrigation System: A Case Study of Northern Maharashtra. *Economic Affairs,* 65**,** 151-159.

GRANDI, G. D., BOUVET, A., LUCAS, R., SHIMADA, M., MONACO, S. & ROSENQVIST, A. 2011. The K&C PALSAR mosaic of the African continent: processing issues and first thematic results. *IEEE Transactions on Geoscience and Remote Sensing,* 49**,** 3593-3610.

HANDIQUE, B. K., KHAN, A. Q., GOSWAMI, C., PRASHNANI, M., GUPTA, C. & RAJU, P. L. N. 2017. Crop Discrimination Using Multispectral Sensor Onboard Unmanned Aerial Vehicle. *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences,* 87**,** 713-719.

HARASYMCZUK, D. M. 2016. Mahalanobis distance plugin [Add-on QGIS software].

HEBBAR, K. B., BERWAL, M. K. & CHATURVEDI, V. K. 2016. Plantation crops: climatic risks and adaptation strategies. *Indian Journal of Plant Physiology,* 21**,** 428-436.

HERMANTO, C., SUTANTO, A., JUMJUNIDANG, EDISON, H. S., DANIELLS, J. W., O'NEILL, W. T., SINOHIN, V. G. O., MOLINA, A. B. & TAYLOR, P. INCIDENCE AND DISTRIBUTION OF FUSARIUM WILT DISEASE OF BANANA IN INDONESIA. 2011. International Society for Horticultural Science (ISHS), Leuven, Belgium, 313-322.

HEUMANN, B. W., WALSH, S. J. & MCDANIEL, P. M. 2011. Assessing the application of a geographic presence-only model for land suitability mapping. *Ecological Informatics,* 6**,** 257-269.

HEUMANN, B. W., WALSH, S. J., VERDERY, A. M., MCDANIEL, P. M. & RINDFUSS, R. R. 2013. Land Suitability Modeling Using a Geographic Socio-Environmental Niche-Based Approach: A Case Study from Northeastern Thailand. *Annals of the Association of American Geographers,* 103**,** 764-784.

HEUPEL, K., SPENGLER, D. & ITZEROTT, S. 2018. A Progressive Crop-Type Classification Using Multitemporal Remote Sensing Data and Phenological Information. *PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science,* 86**,** 53-69.

HIJBEEK, R., ITTERSUM, M. K. V. P. D. & BERGE, H. F. M. T. D. 2017. *On the role of soil organic matter for crop production in European arable farming.* Wageningen University.

HIJMANS, R. J., CAMERON, S. E., PARRA, J. L., JONES, P. G. & JARVIS, A. 2005. Very High Resolution Interpolated Climate Surfaces for Global Land Areas. *International Journal of Climatology,* 25**,** 1965-1978.

HOLLOWAY, K. 2018. A Layman's Interpretation Guide to L-band and C-band Synthetic Aperture Radar Data. Committee on Earth Observation Satellites (CEOS).

HOSSEN, H., IBRAHIM, M. G., MAHMOD, W. E., NEGM, A., NADAOKA, K. & SAAVEDRA, O. 2018. Forecasting future changes in Manzala Lake surface area by considering variations in land use and land cover using remote sensing approach. *Arabian Journal of Geosciences,* 11**,** 93.

HUISING, J. 1993. *Land use zones and land use patterns in the Atlantic zone of Costa Rica : a pattern recognition approach to land use inventory at the sub-regional scale, using remote sensing and GIS, applying an object-oriented and data-driven strategy.* s.n.].

INEGI - INSTITUTO NACIONAL DE ESTADÍSTICA Y GEOGRAFÍA 2017. Conjunto de datos vectoriales de la carta de Uso del suelo y vegetación. Escala 1:250 000. Serie VI. Conjunto Nacional.

JAMIL, M., SAHANA, M. & SAJJAD, H. 2018. Crop Suitability Analysis in the Bijnor District, UP, Using Geospatial Tools and Fuzzy Analytical Hierarchy Process. *Agricultural Research,* 7**,** 506-522.

JAXA 2021. Global 25 m Resolution PALSAR-2/PALSAR Mosaic and Forest/Non-Forest Map (FNF) Dataset Description. JAXA.

JESUS JÚNIOR, W. C. D., VALADARES JÚNIOR, R., CECÍLIO, R. A., MORAES, W. B., RIBEIRO DO VALE, F. X., ALVES, F. R. & PAUL, P. A. 2008. Worldwide geographical distribution of Black Sigatoka for banana: predictions based on climate change models. *Scientia Agricola,* 65**,** 40-53.

JOHANSEN, K., PHINN, S., WITTE, C., PHILIP, S. & NEWTON, L. 2009. Mapping Banana Plantations from Object-oriented Classification of SPOT-5 Imagery. *Photogrammetric Engineering & Remote Sensing,* 75**,** 1069-1081.

JONES, D. R. 2018. *The handbook of diseases of banana, abacá and enset,* Boston, MA, CABI.

JOSS, B. N., HALL, R. J., SIDDERS, D. M. & KEDDY, T. J. 2008. Fuzzy-logic modeling of land suitability for hybrid poplar across the Prairie Provinces of Canada. *Environmental Monitoring and Assessment,* 141**,** 79-96.

JUHOS, K., CZIGÁNY, S., MADARÁSZ, B. & LADÁNYI, M. 2019. Interpretation of soil quality indicators for land suitability assessment – A multivariate approach for Central European arable soils. *Ecological Indicators,* 99**,** 261-272.

JURADO-EXPÓSITO, M., DE CASTRO, A. I., TORRES-SÁNCHEZ, J., JIMÉNEZ-BRENES, F. M. & LÓPEZ-GRANADOS, F. 2019. Papaver rhoeas L. mapping with cokriging using UAV imagery. *Precision Agriculture,* 20**,** 1045-1067.

KAPOOR, N., JAIN, M. & BANSAL, V. K. 2020. A methodological approach for weighting factors in land suitability assessment: a tool for facilitating spatial planning. *Journal of Mountain Science,* 17**,** 724-739.

KHAN, M. R., DE BIE, C. A. J. M., VAN KEULEN, H., SMALING, E. M. A. & REAL, R. 2010. Disaggregating and mapping crop statistics using hypertemporal remote sensing. *International Journal of Applied Earth Observation and Geoinformation,* 12**,** 36-46.

KHOSRAVI, I. & ALAVIPANAH, S. K. 2019. A random forest-based framework for crop mapping using temporal, spectral, textural and polarimetric observations. *International Journal of Remote Sensing,* 40**,** 7221-7251.

KLUGER, D. M., WANG, S. & LOBELL, D. B. 2021. Two shifts for crop mapping: Leveraging aggregate crop statistics to improve satellite-based maps in new regions. *Remote Sensing of Environment,* 262**,** 112488.

KOLLET, S. J. & MAXWELL, R. M. 2008. Capturing the influence of groundwater dynamics on land surface processes using an integrated, distributed watershed model. *Water Resources Research,* 44.

LANSING, D., BIDEGARAY, P., HANSEN, D. O. & MCSWEENEY, K. 2008. Placing the plantation in smallholder agriculture: Evidence from Costa Rica. *Ecological Engineering,* 34**,** 358-372.

LE TOAN, T. 2007. Introduction to SAR Remote Sensing. ESA.
LEFF, B., RAMANKUTTY, N. & FOLEY, J. A. 2004. Geographic distribution of major crops across the world. *Global Biogeochemical Cycles,* 18.
LETOURNEAU, A., VERBURG, P. H. & STEHFEST, E. 2012. A land-use systems approach to represent land-use dynamics at continental and global scales. *Environmental Modelling & Software,* 33**,** 61-79.
LOPEZ-MARTINEZ, C. & LOPEZ-SANCHEZ, J. M. 2017. Special Issue on Polarimetric SAR Techniques and Applications. *Applied Sciences,* 7**,** 768.
LÓPEZ-MARTÍNEZ, C. & POTTIER, E. 2021. Basic Principles of SAR Polarimetry. *In:* HAJNSEK, I. & DESNOS, Y.-L. (eds.) *Polarimetric Synthetic Aperture Radar: Principles and Application.* Cham: Springer International Publishing.
LU, D. & WENG, Q. 2007. A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing,* 28**,** 823-870.
MALISAWA, M. S. & RAUTENBACH, C. J. D. W. 2012. Evaluating water scarcity in the Southern African Development Community (SADC) region by using a climate moisture index (CMI) indicator. *Water Supply,* 12**,** 45-55.
MARYANI, N., LOMBARD, L., POERBA, Y. S., SUBANDIYAH, S., CROUS, P. W. & KEMA, G. H. J. 2019. Phylogeny and genetic diversity of the banana Fusarium wilt pathogen Fusarium oxysporum f. sp. cubense in the Indonesian centre of origin. *Studies in Mycology,* 92**,** 155-194.
MCCALLISTER, D. L. & CHIEN, W. L. 2000. Organic carbon quantity and forms as influenced by tillage and cropping sequence. *Communications in Soil Science and Plant Analysis,* 31**,** 465-479.
MCLACHLAN, G. J. 1999. Mahalanobis distance. *Resonance,* 4**,** 20-26.
MENGESHA, G. G., YETAYEW, T. H. & SAKO, K. A. 2018. Spatial distribution and association of banana (Musa spp.) Fusarium wilt (Fusarium oxysporum f.sp. cubense) epidemics with biophysical factors in southwestern Ethiopia. *Archives of Phytopathology and Plant Protection,* 51**,** 575-601.
MERTENS, A., SWENNEN, R., RØNSTED, N., VANDELOOK, F., PANIS, B., SACHTER-SMITH, G., VU, D. T. & JANSSENS, S. B. 2021. Conservation status assessment of banana crop wild relatives using species distribution modelling. *Diversity and Distributions,* 27**,** 729-746.
MIETTINEN, J. & LIEW, S. C. 2011. Separability of insular Southeast Asian woody plantation species in the 50 m resolution ALOS PALSAR mosaic product. *Remote Sensing Letters,* 2**,** 299-307.
MØLLER, A. B., MULDER, V. L., HEUVELINK, G. B. M., JACOBSEN, N. M. & GREVE, M. H. 2021. Can We Use Machine Learning for Agricultural Land Suitability Assessment? *Agronomy,* 11**,** 703.
MONFREDA, C., RAMANKUTTY, N. & FOLEY, J. A. 2008. Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global Biogeochemical Cycles,* 22.
MORA, B., TSENDBAZAR, N.-E., HEROLD, M. & ARINO, O. 2014. Global Land Cover Mapping: Current Status and Future Trends. *In:* MANAKOS, I. & BRAUN, M. (eds.) *Land Use and Land Cover Mapping in Europe: Practices & Trends.* Dordrecht: Springer Netherlands.
MORALES, J. A., SANCHEZ, L., VELASQUEZ, H., DE BORREGO, B., DE NAVA, M., PORTILLO, D., CANO, Y., MORILLO, A., ALBORNOZ, A. & SOCORRO, E. 2001. Nutrient Loading by Precipitation in the Maracaibo Lake Basin, Venezuela. *Water, Air and Soil Pollution,* 130**,** 511-516.
NASA-JPL 2020. NASADEM Merged DEM Global 1 arc second V001. *In:* EOSDIS (ed.). Land Processes DAAC.
NBS TANZANIA 2021. National Sample Census of Agriculture 2019/20 National Report.
NTUMNGIA, R. N. 2010. *Dangerous assumptions : the agroecology and ethnobiology of traditional polyculture cassava systems in rural Cameroon and implications of green revolution technologies for sustainability, food security, and rural welfare.* s.n.].
O'NEILL, W. T., HENDERSON, J., PATTEMORE, J. A., O'DWYER, C., PERRY, S., BEASLEY, D. R., TAN, Y. P., SMYTH, A. L., GOOSEM, C. H., THOMSON, K. M., HOBBS, R. L., GRICE, K. R. E., TREVORROW, P., VAWDREY, L. L., PATHANIA, N. & SHIVAS, R. G. 2016. Detection of Fusarium oxysporum f.

sp. cubense tropical race 4 strain in northern Queensland. *Australasian Plant Disease Notes* [Online], 11.
OBAGA, B. & MWAURA, F. 2018. Impact of farmers' participation in banana value addition in household welfare in Kisii Central Sub-County. *International Academic Journal of Social Sciences and Education,* 2**,** 25-46.
OCIMATI, W., BOUWMEESTER, H., GROOT, J. C. J., TITTONELL, P., BROWN, D. & BLOMME, G. 2019. The risk posed by Xanthomonas wilt disease of banana: Mapping of disease hotspots, fronts and vulnerable landscapes. *PLOS ONE,* 14**,** e0213691.
ONDERWATER, N. 2020. *A global banana map : disaggregating national production statistics through land use analysis and land suitability evaluation.* Wageningen University.
ORDÓÑEZ R, N., KEMA, G. H. J. P. D., SEIDL, M. F. D., MEIJER, H. J. G. D., ORDÓÑEZ R, N., KEMA, G. H. J. P. D., SEIDL, M. F. D. & MEIJER, H. J. G. D. 2018. *A global genetic diversity analysis of Fusarium oxysporum f.sp. cubense : the Panama disease pathogen of banana.* Wageningen University.
OTTEN, W. & GILLIGAN, C. A. 2006. Soil structure and soil-borne diseases: using epidemiological concepts to scale from fungal spread to plant epidemics. *European Journal of Soil Science,* 57**,** 26-37.
PANDA, R. M. & BEHERA, M. D. 2019. Assessing harmony in distribution patterns of plant invasions: a case study of two invasive alien species in India. *Biodiversity and Conservation,* 28**,** 2245-2258.
PEÑA-BARRAGÁN, J. M., NGUGI, M. K., PLANT, R. E. & SIX, J. 2011. Object-based crop identification using multiple vegetation indices, textural features and crop phenology. *Remote Sensing of Environment,* 115**,** 1301-1316.
PFADENHAUER, J. S. & KLÖTZLI, F. A. 2020a. Fundamentals towards Understanding Global Vegetation. *Global Vegetation: Fundamentals, Ecology and Distribution.* Cham: Springer International Publishing.
PFADENHAUER, J. S. & KLÖTZLI, F. A. 2020b. Zonal Vegetation of the Tropical-Subtropical Dry Zone. *Global Vegetation: Fundamentals, Ecology and Distribution.* Cham: Springer International Publishing.
PFADENHAUER, J. S. & KLÖTZLI, F. A. 2020c. Zonal Vegetation of the Tropical Zone with Year-Round Rain. *Global Vegetation: Fundamentals, Ecology and Distribution.* Cham: Springer International Publishing.
PHILLIPS, S. J., ANDERSON, R. P. & SCHAPIRE, R. E. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling,* 190**,** 231-259.
PILLAY, R., MILLER, D. A. W., HINES, J. E., JOSHI, A. A. & MADHUSUDAN, M. D. 2014. Accounting for false positives improves estimates of occupancy from key informant interviews. *Diversity and Distributions,* 20**,** 223-235.
PLOETZ, R., FREEMAN, S., KONKOL, J., AL-ABED, A., NASER, Z., SHALAN, K., BARAKAT, R. & ISRAELI, Y. 2015a. Tropical race 4 of Panama disease in the Middle East. *Phytoparasitica,* 43**,** 283-293.
PLOETZ, R. C. 2015a. Fusarium Wilt of Banana. *Phytopathology,* 105**,** 1512-21.
PLOETZ, R. C. 2015b. Management of Fusarium wilt of banana: A review with special reference to tropical race 4. *Crop Protection,* 73**,** 7-15.
PLOETZ, R. C., KEMA, G. H. J. & MA, L.-J. 2015b. Impact of Diseases on Export and Smallholder Production of Banana. *Annual Review of Phytopathology,* 53**,** 269-288.
POGGIO, L., DE SOUSA, L. M., BATJES, N. H., HEUVELINK, G. B. M., KEMPEN, B., RIBEIRO, E. & ROSSITER, D. 2021. SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. *SOIL,* 7**,** 217-240.
QIU, F., CHASTAIN, B., ZHOU, Y., ZHANG, C. & SRIDHARAN, H. 2014. Modeling land suitability/capability using fuzzy evaluation. *GeoJournal,* 79**,** 167-182.
RABBINGE, R. 1997. Integrating policy and technical issues for international research on agriculture and the environment using systems approaches. *In:* TENG, P. S., KROPFF, M. J., TEN BERGE, H. F. M., DENT, J. B., LANSIGAN, F. P. & VAN LAAR, H. H. (eds.) *Applications of Systems*

Approaches at the Farm and Regional Levels Volume 1: Proceedings of the Second International Symposium on Systems Approaches for Agricultural Development, held at IRRI, Los Baños, Philippines, 6–8 December 1995. Dordrecht: Springer Netherlands.

RAMANKUTTY, N., EVAN, A. T., MONFREDA, C. & FOLEY, J. A. 2008. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Global Biogeochemical Cycles,* 22.

RAMANKUTTY, N. & FOLEY, J. A. 1998. Characterizing patterns of global land use: An analysis of global croplands data. *Global Biogeochemical Cycles,* 12**,** 667-685.

RANEY, K. 2014. Radar, Synthetic Aperture. *In:* NJOKU, E. G. (ed.) *Encyclopedia of Remote Sensing.* New York, NY: Springer New York.

RAY, D., BEHERA, M. D. & JACOB, J. 2016. Predicting the distribution of rubber trees (Hevea brasiliensis) through ecological niche modelling with climate, soil, topography and socioeconomic factors. *Ecological Research,* 31**,** 75-91.

ROBINSON, J. C. & GALÁN SAÚCO, V. C. 2011. Bananas and plantains. 2nd ed. ed. Wallingford: CABI.

RURINDA, J. 2014. *Vulnerability and adaptation to climate variability and change in smallholder farming systems in Zimbabwe.* Wageningen University.

SAIZ, G., BIRD, M. I., DOMINGUES, T., SCHRODT, F., SCHWARZ, M., FELDPAUSCH, T. R., VEENENDAAL, E., DJAGBLETEY, G., HIEN, F., COMPAORE, H., DIALLO, A. & LLOYD, J. 2012. Variation in soil carbon stocks and their determinants across a precipitation gradient in West Africa. *Global Change Biology,* 18**,** 1670-1683.

SALVACION, A. R. 2020. Effect of climate on provincial-level banana yield in the Philippines. *Information Processing in Agriculture,* 7**,** 50-57.

SALVACION, A. R. 2021. Mapping land limitations for agricultural land use planning using fuzzy logic approach: a case study for Marinduque Island, Philippines. *GeoJournal,* 86**,** 915-925.

SALVACION, A. R., CUMAGUN, C. J. R., PANGGA, I. B., MAGCALE-MACANDOG, D. B., CRUZ, P. C. S., SALUDES, R. B., SOLPOT, T. C. & AGUILAR, E. A. 2019. Banana suitability and Fusarium wilt distribution in the Philippines under climate change. *Spatial Information Research,* 27**,** 339-349.

SASSAN, S. 2019. SAR Methods for Mapping and Monitoring Forest Biomass. *In:* FLORES, A., HERNDON, K., THAPA, R. & CHERRINGTON, E. (eds.) *SAR Handbook: Comprehensive Methodologies for Forest Monitoring and Biomass Estimation.* NASA.

SCHMIDTLEIN, S., FAUDE, U., STENZEL, S. & FEILHAUER, H. 2014. Remote Sensing of Vegetation for Nature Conservation. *In:* MANAKOS, I. & BRAUN, M. (eds.) *Land Use and Land Cover Mapping in Europe: Practices & Trends.* Dordrecht: Springer Netherlands.

SHI, Y., LI, J., MA, D., ZHANG, T. & LI, Q. 2019. Method for crop classification based on multi-source remote sensing data. *IOP Conference Series: Materials Science and Engineering,* 592**,** 012192.

SHIMADA, M., ITOH, T., MOTOOKA, T., WATANABE, M., SHIRAISHI, T., THAPA, R. & LUCAS, R. 2014. New global forest/non-forest maps from ALOS PALSAR data (2007–2010). *Remote Sensing of Environment,* 155**,** 13-31.

SINGH, H. P. Country Paper on Banana in India. *In:* MOLINA, A. B., BAROÑA, M. L. J., SINOHIN, V. G. O. & GENEROSO, J. D., eds. Advancing Banana and Plantain R&D in Asia and The Pacific: Proceedings of the 7th BAPNET Steering Committee meeting, 2010 Hanoi, Vietnam. Bioversity International.

SINGH, S. 2021. *How Jalgaon, "Banana City of India," Became 7th Biggest Banana Producer in World* [Online]. Krishi Jagran. Available: https://krishijagran.com/agriculture-world/how-jalgaon-banana-city-of-india-became-7th-biggest-banana-producer-in-world/ [Accessed 9 December 2021].

SOINI, E. 2005. Changing livelihoods on the slopes of Mt. Kilimanjaro, Tanzania: Challenges and opportunities in the Chagga homegarden system. *Agroforestry Systems,* 64**,** 157-167.

SONOBE, R., TANI, H., WANG, X., KOBAYASHI, N. & SHIMAMURA, H. 2014. Random forest classification of crop type using multi-temporal TerraSAR-X dual-polarimetric data. *Remote Sensing Letters,* 5**,** 157-164.

SPARLING, G., PARFITT, R. L., HEWITT, A. E. & SCHIPPER, L. A. 2003. Three Approaches to Define Desired Soil Organic Matter Contents. *Journal of Environmental Quality,* 32**,** 760-766.
SRIKANTH, P., RAMANA, K. V., DEEPIKA, U., KALYAN CHAKRAVARTHI, P. & SESHA SAI, M. V. R. 2016. Comparison of Various Polarimetric Decomposition Techniques for Crop Classification. *Journal of the Indian Society of Remote Sensing,* 44**,** 635-642.
STILL, C. J., BERRY, J. A., COLLATZ, G. J. & DEFRIES, R. S. 2003. Global distribution of C3 and C4 vegetation: Carbon cycle implications. *Global Biogeochemical Cycles,* 17**,** 6-1-6-14.
STOORVOGEL, J. J., BOUMA, J. & ORLICH, R. A. 2004. Participatory Research for Systems Analysis: Prototyping for a Costa Rican Banana Plantation. *Agronomy Journal,* 96**,** 323-336.
SULEIMAN, R. 2018. *Local and regional variations in conditions for agriculture and food security in Tanzania: A review*.
TAO, C., CHEN, S., LI, Y. & XIAO, S. 2017. PolSAR Land Cover Classification Based on Roll-Invariant and Selected Hidden Polarimetric Features in the Rotation Domain. *Remote Sensing,* 9**,** 660.
THANGAVELU, R., MOSTERT, D., GOPI, M., DEVI, P. G., PADMANABAN, B., MOLINA, A. B. & VILJOEN, A. 2019. First detection of Fusarium oxysporum f. sp. cubense tropical race 4 (TR4) on Cavendish banana in India. *European Journal of Plant Pathology,* 154**,** 777-786.
THIERS, R., LIU, J., NAVAS, G., SCHEIDEL, A. & MOU, H. 2019. A 'political fungology' of China's booming banana sector. 56.
THOMPSON, A. K. & BURDEN, O. J. 1995. Harvesting and fruit care. *In:* GOWEN, S. (ed.) *Bananas and Plantains.* Dordrecht: Springer Science+Business Media.
TIFFEN, M. & MORTIMORE, M. 1990. *Theory and practice in plantation agriculture : an economic review,* London, Overseas Development Institute.
ULLMANN, T., BANKS, S. N., SCHMITT, A. & JAGDHUBER, T. 2017. Scattering Characteristics of X-, C- and L-Band PolSAR Data Examined for the Tundra Environment of the Tuktoyaktuk Peninsula, Canada. *Applied Sciences,* 7**,** 595.
VERBURG, P. H., DE KONING, G. H. J., KOK, K., VELDKAMP, A. & BOUMA, J. 1999. A spatial explicit allocation procedure for modelling the pattern of land use change based upon actual land use. *Ecological Modelling,* 116**,** 45-61.
VILJOEN, A., MOSTERT, D., CHICONELA, T., BEUKES, I., FRASER, C., DWYER, J., MURRAY, H., AMISSE, J., MATABUANA, E. L., TAZAN, G., AMUGOLI, O. M., MONDJANA, A., VAZ, A., PRETORIUS, A., BOTHMA, S., ROSE, L. J., BEED, F., DUSUNCELI, F., CHAO, C.-P. & MOLINA, A. B. 2020. Occurrence and spread of the banana fungus Fusarium oxysporum f. sp. cubense TR4 in Mozambique. *South African Journal of Science,* 116**,** 1-11.
VINTROU, E., DESBROSSE, A., BÉGUÉ, A., TRAORÉ, S., BARON, C. & LO SEEN, D. 2012. Crop area mapping in West Africa using landscape stratification of MODIS time series and comparison with existing global land products. *International Journal of Applied Earth Observation and Geoinformation,* 14**,** 83-93.
WAHOME, C. N., MAINGI, J. M., OMBORI, O., KIMITI, J. M. & NJERU, E. M. 2021. Banana Production Trends, Cultivar Diversity, and Tissue Culture Technologies Uptake in Kenya. *International Journal of Agronomy,* 2021**,** 6634046.
WANG, D., LIN, H., CHEN, J., ZHANG, Y. & ZENG, Q. 2010. Application of multi-temporal ENVISAT ASAR data to agricultural area mapping in the Pearl River Delta. *International Journal of Remote Sensing,* 31**,** 1555-1572.
XIAO-LAN, W., CHIANG, T.-Y., ROUX, N., HAO, G. & XUE-JUN, G. 2007. Genetic diversity of wild banana (Musa balbisiana Colla) in China as revealed by AFLP markers. *Genetic Resources and Crop Evolution,* 54**,** 1125-1132.
XIE, L., ZHANG, H., LI, H. & WANG, C. 2015. A unified framework for crop classification in southern China using fully polarimetric, dual polarimetric, and compact polarimetric SAR data. *International Journal of Remote Sensing,* 36**,** 3798-3818.
XU, D. & TIAN, Y. 2015. A Comprehensive Survey of Clustering Algorithms. *Annals of Data Science,* 2**,** 165-193.

YAGER, R. R. 1988. On ordered weighted averaging aggregation operators in multicriteria decisionmaking. *IEEE Transactions on Systems, Man, and Cybernetics,* 18**,** 183-190.
YENGOH, G. T. & ARDÖ, J. 2014. Crop Yield Gaps in Cameroon. *Ambio,* 43**,** 175-190.
YOU, L., WOOD, S., WOOD-SICHRA, U. & WU, W. 2014. Generating global crop distribution maps: From census to grid. *Agricultural Systems,* 127**,** 53-60.
ZAKERI, H., YAMAZAKI, F. & LIU, W. 2017. Texture Analysis and Land Cover Classification of Tehran Using Polarimetric Synthetic Aperture Radar Imagery. *Applied Sciences,* 7**,** 452.
ZHANG, K., LIU, H., PAN, H., SHI, W., ZHAO, Y., LI, S., LIU, J. & TAO, J. 2020. Shifts in potential geographical distribution of Pterocarya stenoptera under climate change scenarios in China. *Ecology and evolution,* 10**,** 4828-4837.
ZHANG, S., LEI, Y., WANG, L., LI, H. & ZHAO, H. 2011. Crop classification using MODIS NDVI data denoised by wavelet: A case study in Hebei Plain, China. *Chinese Geographical Science,* 21**,** 322-333.
ZHANG, X., WU, S., YAN, X. & CHEN, Z. 2017. A global classification of vegetation based on NDVI, rainfall and temperature. *International Journal of Climatology,* 37**,** 2318-2324.
ZHENG, S.-J., GARCÍA-BASTIDAS, F. A., LI, X., ZENG, L., BAI, T., XU, S., YIN, K., LI, H., FU, G., YU, Y., YANG, L., NGUYEN, H. C., DOUANGBOUPHA, B., KHAING, A. A., DRENTH, A., SEIDL, M. F., MEIJER, H. J. G. & KEMA, G. H. J. 2018. New Geographical Insights of the Latest Expansion of Fusarium oxysporum f.sp. cubense Tropical Race 4 Into the Greater Mekong Subregion. *Frontiers in Plant Science,* 9.
ZHU, Y., REN, L., LÜ, H., DRAKE, S., YU, Z., WANG, Z., FANG, X. & YUAN, F. 2013. Effect of Water Table Depth on Growth and Yield of Soybean Yudou 16. *Journal of Hydrologic Engineering,* 18**,** 1070-1076.

Appendix

Data overview with instructions for access:

Data	Details	Source	Access (+ *GEE snippet*)
Synthetic Aperture Radar (SAR)	Sentinel-1 product from ESA C-band, VV and VH polarization	(ESA, 2020)	Google Earth Engine *ee.Image("COPERNICUS/S1_GRD")*
	PALSAR-2 product from JAXA L-band, HH and HV polarization	(Shimada et al., 2014)	Google Earth Engine *ee.Image("JAXA/ALOS/PALSAR/YEARLY/SAR")*
Landcover	Copernicus global land service (CGLS) product	(Buchhorn et al., 2020)	Google Earth Engine *ee.Image("COPERNICUS/Landcover/100m/Proba-V-C3/Global")*
	GlobCover product	(Arino, 2010)	Google Earth Engine *ee.Image("(ESA/GLOBCOVER_L4_200901_200912_V2_3)")*
Elevation	NASADEM product	(NASA-JPL, 2020)	Google Earth Engine *ee.Image("NASA/NASADEM_HGT/001")*
Calibration data	Costa Rican banana map	(Corbana, 2019)	Through Corbana
	Mexican land use map	(INEGI, 2017)	https://www.inegi.org.mx/app/biblioteca/ficha.html?upc=889463598459
	Philippine banana statistics (regional)	(Salvacion, 2020)	Other Crops: Area Planted / Harvested, by Region and by Province, by Semester, 2010-2020 https://openstat.psa.gov.ph/
Climate	WorldClim V1 Bioclim product	(Hijmans et al., 2005)	Google Earth Engine *ee.Image("WORLDCLIM/V1/BIO")*
Soil	SoilGrids product	(Poggio et al., 2021)	Google Earth Engine *ee.Image("projects/soilgrids-isric/soc_mean")*
Water table depth	EartH2Observe product	(Fan et al., 2013)	THREDDS Data Server of Plymouth Marine Laboratory: https://wci.earth2observe.eu/thredds/catalog/usc/water-table-depth/catalog.html (HTTPServer)
National production statistics	FAOSTAT product	(FAO, 2021b)	http://www.fao.org/faostat/en/#data/QCL
Existing global banana maps	Global banana map	(Monfreda et al., 2008)	http://www.earthstat.org/harvested-area-yield-175-crops/
	Global banana map	(Onderwater, 2020)	Data not available