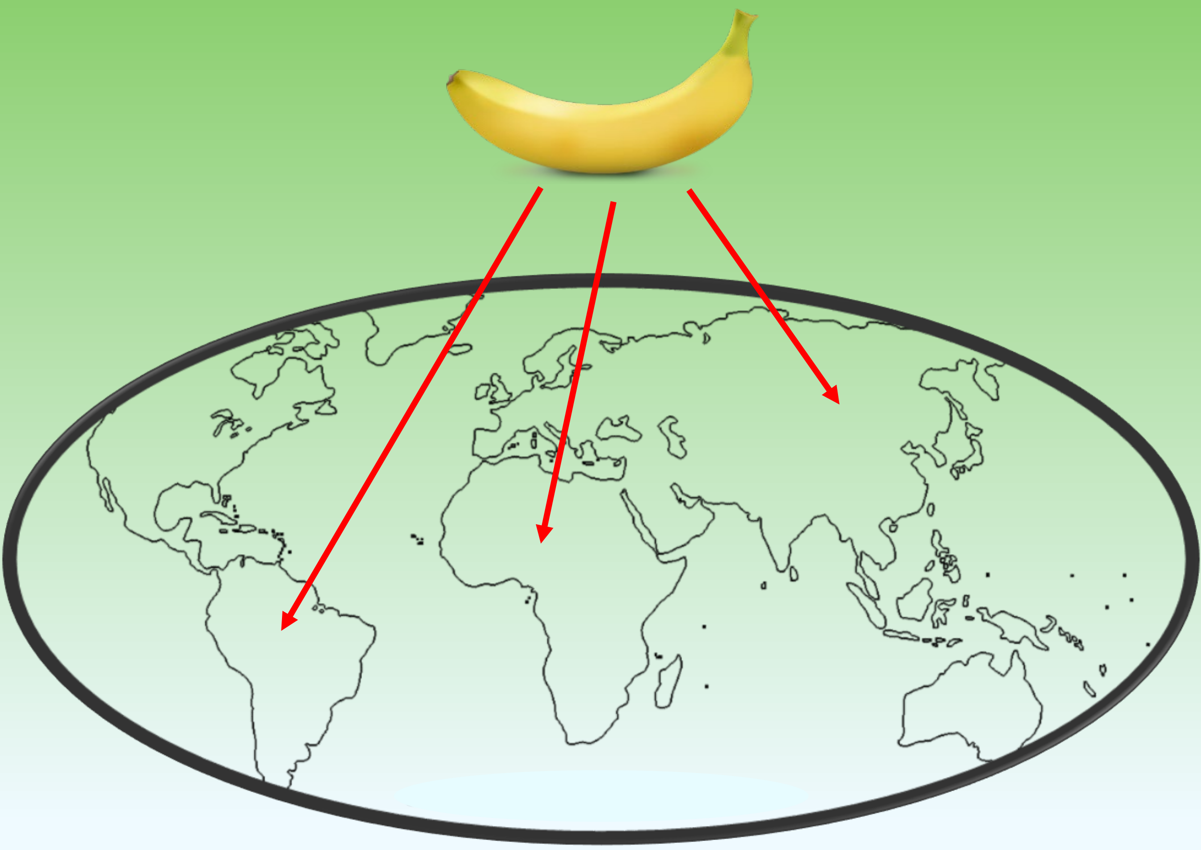
A Global Banana Map

Disaggregating National Production Statistics Through Land Use Analysis and Land Suitability Evaluation



Nard Onderwater

April, 2020

A Global Banana Map

Disaggregating National Production Statistics Through Land Use Analysis and Land Suitability Evaluation

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

951018621020

Master thesis

Earth and Environment: Soil Geography and Earth Surface Dynamics

April, 2020

Wageningen, The Netherlands

Soil Geography and Landscape group

Wageningen UR

Supervisors:

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

Soil Geography and Landscape group

Wageningen UR, Netherlands

L.A. (Lieke) Melsen PhD MSc

Hydrology and Quantitative Water Management group

Wageningen UR, Netherlands

Abstract

Sustainable land management is an important factor in securing enough food production for an ever-growing population. Agricultural land is lost because of erosion, nutrient depletion, contamination and salinization. Additionally, outbreaks of pests and diseases destroy harvests and infect soils. Information on crop distribution is essential for analysis on food security, land degradation and the spread of pests and diseases.

Bananas are an important staple food for nearly 400 million people in developing countries. The crop is currently threatened by Fusarium wilt, a fatal disease for banana plants, which makes commercial production impossible. To determine areas at risk and in order to control the spread of the disease, high-resolution global maps of banana production are required. Currently, no detailed information on its global distribution is available. New datasets open the opportunity to create high-resolution crop maps. A method was created to disaggregate national production statistics of a crop on a 30 arc seconds resolution. The disaggregation was done through land use analysis and land suitability analysis. Results of the two analyses were then combined to predict crop areas. The land use analysis was based on a global land cover map, NDVI images and a global canopy height map. Land suitability analysis was based on the following environmental properties: temperature, rainfall, the driest quarter, topography, water table depth and soil conditions. The method was applied on banana, resulting in a global map of banana. Calibration was performed by consulting experts on banana agriculture and through field visits to banana plantations, leading to improvements in the disaggregation method. Validation showed that the banana map represents the distribution of banana well. The map provides detailed information and will prove useful in studies into the spread of Fusarium wilt. Furthermore, in future research, the disaggregation method can be used to create detailed maps of other crops.

Acknowledgements

During the course of this thesis research I was supported by numerous people who I would like to thank here. First of all I would like to thank Jetse Stoorvogel and Lieke Melsen for guiding this study with their supervision, they would always be ready to help and give useful advice. Furthermore, I would like to thank Corbana for giving me the opportunity to step experience the banana world of Costa Rica at first. Some people in particular helped me a lot during my time at Corbana. Rafael Segura for supervising me during my field work in Costa Rica. He arranged plantation visits and set up meetings with experts. He even organised a meeting where I could present my work (in Spanish!) for the Corbana department in Guápiles. José Guzmán-Alvarez and Juan Samuels for taking me on several trips to plantations. José Guzmán-Alvarez also provided me a workplace in his office and he was always there to give advice on characteristics of banana agriculture in Costa Rica. Finally, on a more personal note, I would like to thank Rosemary Soto and Ricardo Morera who were very hospitable, arranging a room for me in their house and taking me on many trips in the weekends to show me their country.

CONTENTS

LIST OF FIGURES

List of tables

1. INTRODUCTION

Although food production has increased substantially in the past decades, hunger rates are increasing worldwide (Godfray et al., 2010, Mundial, 2008, Webb et al., 2018). Sustainable land management is an important factor in securing enough food production for an increasing population. Agricultural area is lost because of erosion, nutrient depletion, contamination and salinization. Additionally, outbreaks of pests and diseases destroy harvests and infect soils, making them unsuitable for agriculture. Information on crop distribution is essential for analysis on food security, land degradation, and the spreading of pests and diseases.

For nearly 400 million people in developing economies bananas are the key staple food and source of income (Ploetz, 1994, Karamura et al., 1998, García-Bastidas, 2019). Yet, the crop is underrepresented in agricultural research. A simple search on google scholar; "wheat distribution worldwide" gives 553.000 hits, while the search; "banana distribution worldwide" only gives 74.400 hits.

Moreover, banana cultivation is threatened by the spread of Fusarium wilt, a fatal disease for banana plants. Fusarium wilt tropical race 1 (TR1) has been a major problem for the banana industry in the first half of the 20th century, leaving almost none of the important production areas free of the disease (Stover and Simmonds, 1987). The introduction of a new cultivar, the fusarium TR1 resistant Cavendish banana, saved the large industries in 1960 (Ploetz, 1994). However, in the 1990's a new form of Fusarium wilt tropical race 4 (TR4) was discovered, already destroying ten thousands of hectares in Southeast Asia and Australia (fusariumwilt.org). In 2013, Fusarium TR4 was found on a commercial farm in Mozambique, which marked the start of spread over the African continent (rtb.cgiar.org). Recently, it has been encountered in Colombia, the path to spread over Central and South America is now open (Garcia-Bastidas et al., 2019). When Fusarium TR4 hits the large mono-crop plantations in Central and South America, many banana farms will be destroyed.

Knowing where banana is grown is the starting point of analysis on food security, land degradation and the spread of Fusarium wilt, . Despite the importance of the crop and the risks banana is facing, no detailed information on its global distribution is currently available. This can be partly attributed to the variability in the banana production systems in different parts of the world. Lack of high-resolution global data is another reason why such information is not yet available. These factors will now be further elaborated on.

Bananas are grown mainly in the developing countries of the tropics where 85 percent is used for own consumption or local trade (Gowen, 2012). Banana production is practiced in different ways. On one hand, there are large mono-crop plantations in Latin America, focussed on the export of the crop. On the other hand, subsistence farming in the form of small, multi-crop farms, is the dominant practice in Africa. Properties of mono-crop production systems are different from multi-cropping systems. For example, mono-crop plantations require a flat surface to support on-farm transportation of bunches, while small-scale farms can be found on steeper slopes (Karamura et al., 1998). Therefore, creating a uniform method to disaggregate banana production area is a challenging task.

Remote sensing provides data on a high spatial resolution, yet we are unable to identify specific crops on a global scale. Statistical data (FAO, 2019), rather holds crop-specific information, but has a poor spatial distribution. Combining remote sensing data with statistical data significantly improves the level of detail. Ramankutty and Foley (1998) were the first to relate satellite derived land cover data with national and subnational agricultural inventory data on a global scale. Working on a 5 arc minutes (~10 km by 10 km at the equator) resolution, they assigned fractions to each pixel, representing the part of the pixel that is covered by cropland. The results show cropland areas for the early 1990's, but which kind of crop is grown remains unknown. Some years later, Leff et al. (2004) managed to identify 18 different crops for the same period and spatial resolution globally. The grid cells are filled with fractions of each of the 18 crops. This dataset does not include bananas yet. Distinguishing between irrigated and rainfed areas, Portmann et al. (2010) presented a global dataset that also includes multi-cropping systems. However, this dataset consists of 26 crop classes and does not map banana separately.

Monfreda et al. (2008) were the first to publish a global banana map on a 5 by 5 arc minutes resolution. They combined a global gridded map of croplands for the year 2000, created by Ramankutty et al. (2008), with agricultural census data to produce a map representing 175 individual crops.

This study was aimed at developing a global banana map on a 30 arc seconds (~1 km by 1 km at the equator) resolution. Recent technological advances have produced high-resolution remote sensing and environmental data on a global scale. These auxiliary datasets allows for accurate mapping of specific crops, rather than multiple crops or groups of crops. In this study, a method was created to disaggregate national production statistics of a single crop. The method consists of a land use analysis and a land suitability evaluation. The method was applied on banana to create a global banana map. Until now crop distribution research has not specifically concentrated on mapping banana, let alone on a 30 arc seconds resolution. FAO-stat (Food and Agriculture Organization of the United Nations) provides data on production area at the country level (FAO-stat, 2019). Exact locations of banana cultivation within a country are unknown. Therefore, this study addresses the following research question:

> *How can we disaggregate national production statistics to create a map of global banana distribution using available remote sensing and environmental data?*

This question can be divided into the following sub-questions:

1. *How can we disaggregate national production statistics of banana through analysing land use and land suitability?*
2. *How can we predict banana areas and how can we calibrate the disaggregation method?*
3. *How valid is the resulting banana map?*

To answer these questions, a method was developed to disaggregate banana area. This method was calibrated through field visits to banana plantations and consults with experts on banana agriculture. The results were then compared to previous work and satellite imagery to assess the validity.

The next chapter will expand on the method of disaggregating national production statistics and application of the data. Section 3 will elaborate on the results of the disaggregation and calibration. In chapter 4 the method will be discussed, and the banana map will be validated. Finally, the conclusions of this study will be presented with recommendations for further research.

2. MATERIALS AND METHODS

2.1 OVERVIEW

In this study a method is created to map an individual crop based on available environmental data. This method can be adapted to identify other crops, based on the properties of the crop. Production statistics for most crops are available at a national level. Here, the method is applied on banana. The goal is to disaggregate banana production statistics to map banana globally. Land use data are used to narrow down the possible banana extent within a country. By adding more data, the potential banana area shrinks, and the percentage of that area used for banana production increases (see Figure 1). In parallel, a land suitability evaluation is carried out to find the most suitable areas for banana. Banana areas are then predicted by combining the results of the land use analysis and the land suitability evaluation. Figure 1 shows a schematic overview of the approach. The method is calibrated by field visits to banana plantations in Costa Rica and by consulting experts on banana agriculture.

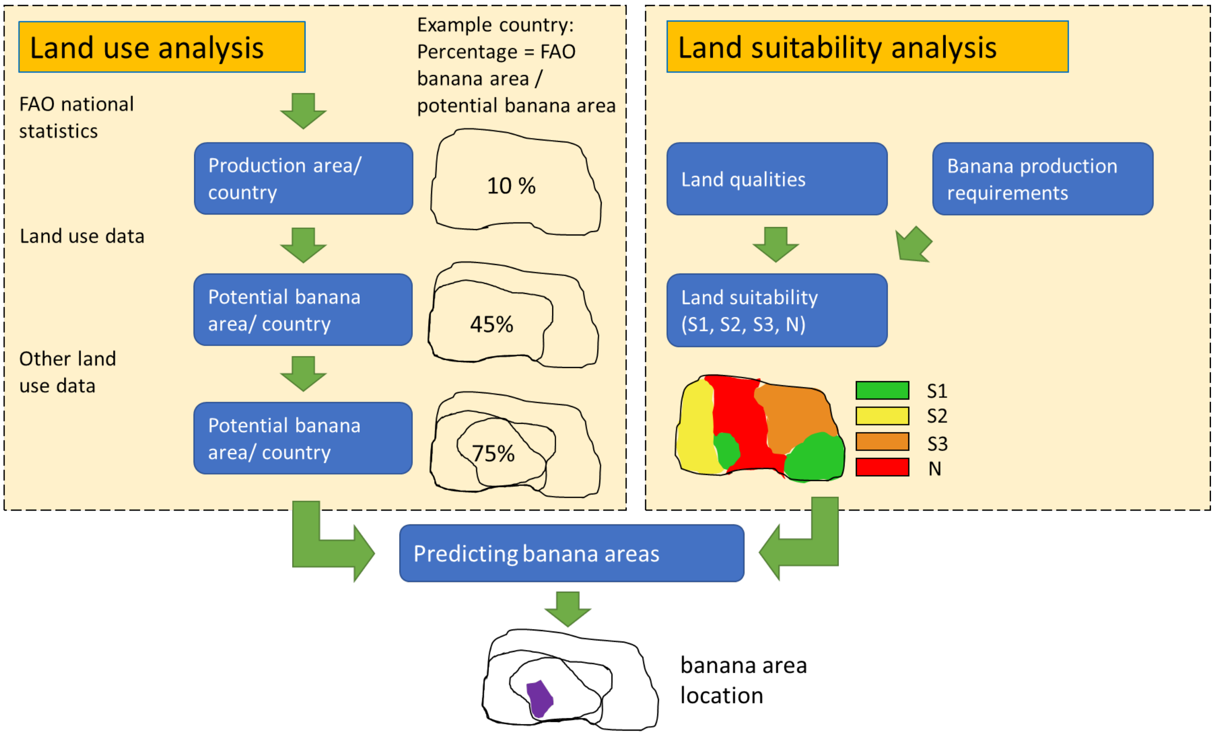


Figure 1. Overview of the disaggregation method. Left half depicts the land use analysis, right half the land suitability evaluation.

2.2 Disaggregating banana production statistics

A method is developed to determine where bananas are grown within a country. This section describes handling of the data in the development of the method. First, the land use analysis is explained followed by the land suitability evaluation and finally the prediction of banana areas. All data are for the year 2015.

Table 1. Datasets applied in the disaggregation method.

	Information on	Data source
National statistics	Harvested area and yield of banana	(FAO-stat, 2019)
Land use	Land cover	(ESA, 2017, Bontemps et al., 2010)
	Seasonal NDVI trend	(https://land.copernicus.eu/global/products/ndvi)
	Canopy height	(Simard et al., 2011)
Land suitability	Temperature, rainfall, rainfall driest quarter	(Fick and Hijmans, 2017)
	Slope	(Danielson and Gesch, 2011)
	Drainage	(Fan et al., 2013)
	Soil type	(Stoorvogel et al., 2017)

2.2.1 Land use analysis

The starting point of this study is the national production statistics from the FAO (FAO-stat, 2019). Using auxiliary data, a method is created to determine where bananas are grown within a country. A global land cover map is used to find agricultural areas that are potentially banana. Subsequently, areas with perennial crops are identified by analysing NDVI images. Banana is a perennial crop and will therefore appear green on NDVI images throughout the year. Lastly, a global canopy height map is used to detect and remove areas with tall trees. Intermediate maps have been created each time a new dataset was added. These maps show the areas that are potentially used to produce banana.

2.2.1.1 FAO national production statistics

The FAO statistics include data on production quantity, production area and total yield for 175 crops at a national, first administrative level. Although this dataset has limitations due to the different sources of information used to compile a global database of crop statistics, it is yet the only worldwide agricultural database. The FAO database is an open source global dataset used in many studies globally and is accepted as a reliable source. Harvested area for banana and plantain are downloaded for the year 2015. These data are combined to get the total area of banana per country. Dividing the area of banana by the total area of a country gives an impression of the extent of banana agriculture in the country. For the boundaries of countries the GADM database is used (www.gadm.org). The total area of a country is obtained by summing all the cells within the country's borders. Prior to calculating areas, cell sizes are recalculated to correct for the difference in cell area that originates from the long-lat projection. For example, in the projection, a degree longitude near the equator is approximately 110 km length while at the poles it is 0 km. This correction was done for all area calculations.

In formula, the percentage of banana agriculture in a country is calculated as:

$$Pcountry = \frac{Abx}{Ax} * 100 \qquad (Eq.\ 1)$$

In which:

Pcountry = Percentage of country area used for banana production

Abx = Area of banana in country x

Ax = Area of a country x

2.2.1.2 GLOBAL LAND COVER MAP

The Global CCI (Climate Change Initiative) land cover map, or GlobCov, is a global land use map developed by the European Space Agency. The original version has a 300m spatial resolution and provides annual maps from 1992 to 2015. The map distinguishes agricultural areas, urban areas and nature reserves. Furthermore, it shows rain-fed cropland areas and irrigated cropland areas. The land cover map is used to identify areas that are in agricultural use. Additionally, the map has within cell information in a class that is defined as mosaic agriculture; partly agriculture and partly natural vegetation. This additional information is implemented in the model by reclassifying these classes into a percentage. This percentage represents the chance that the pixel is used for banana agriculture. Completely agricultural areas in the GlobCov dataset are assigned a value of 100 percent. Classes which are partly mosaic got lower percentages. The complete reclassification of the GlobCov dataset can be found in appendix A. Noteworthy is the reclassification of forest classes into potential banana. Forest is included because banana plants are quite tall (up to 8 m) and the GlobCov dataset can see them as forest. The agricultural area of each cell is calculated by multiplying the percentage with the original value of the area of the cell. For example, a cell that has 40 percent agricultural area and originally had an area of 100 hectares will now have an area value of 40 hectares. Then, the area of all the cells are added up per country, resulting in total agricultural area of that country. Subsequently, the banana area is divided by the total agricultural area, giving a percentage of the agricultural area that is used for banana production. In formula:

$$Pagri = \frac{Abx}{Aax} * 100 \qquad (Eq.\ 2)$$

In which:

Pagri = Percentage of agricultural area per country that is used for banana production

Abx = Area of banana in a country x

Aax = Area of agriculture in country x

2.2.1.3 Normalized Difference Vegetation Index (NDVI)

NDVI images provide information about the greenness of the landscape. Since banana is a perennial crop, these images can be used to recognize banana by finding areas that are green throughout the year. Ten-daily global NDVI images for 2015 are downloaded from the Copernicus website (https://land.copernicus.eu/global/products/ndvi). From these 36 layers a new map is created by taking the lowest NDVI values on each pixel. The next step is to put a threshold on the minimum NDVI map so that areas with annual crops are discarded. According to a field study on a banana plantation in Costa Rica, NDVI values of bananas have a mean NDVI of 0.2 to 0.35 (Machovina et al., 2017). Johansen et al. (2009) reported mean NDVI values ranging from 0.1 to 0.573. Literature shows large differences in ranges of NDVI values. Therefore, a comparison of the minimum NDVI map with known banana plantation locations in Costa Rica is done to find a feasible NDVI threshold value. Most banana plantations show mean NDVI values of around 0.6. The threshold NDVI value for banana is set to 0.4 to make sure that annual crops are filtered out, while banana is included. Pixels with an NDVI below this threshold are reclassified to zero and pixels with a higher NDVI to one. Multiplying the NDVI map with the land cover map results in a map of the perennial agricultural areas. *Pagri* is updated with an NDVI filter, reducing the extent and increasing the percentage of potential banana area.

2.2.1.4 Canopy height

A global canopy height dataset is used to identify and filter out areas that have high canopies. The dataset gives a good overview of the global canopy height but is less accurate near the tropics. Nonetheless, the data is very useful to filter out rainforests. Banana plants are up to 8m tall, while trees in the tropical forests can be over 40 m tall. Logically, a threshold on canopies higher than 8 m would filter out the taller trees. Likewise, banana plants that produce banana are taller than 2 m so that a lower threshold could be applied to discriminate between pasture and banana. However, when comparing locations of known banana plantations in tropical countries to the same locations on the canopy height map, values up to 25 m are indicated on the canopy height map where 8 m is expected. On the other hand, there are countries that do not have any potential banana area left if a lower threshold of 2 m is applied. In Egypt for example, almost no area has canopies taller than 2 m, while there is banana according to the FAO statistics. This is because the resolution of the canopy height map is 1 km. The average canopy height in the cell is below 2 m but there are smaller plots with taller canopies. Therefore, only an upper threshold value of 25 m is used, to discard the taller canopies. This filter is added to the previous map, so that the map contains the areas with perennial agriculture with canopies below 25 m.

2.2.2 Land suitability evaluation

A land suitability evaluation is carried out in the form of a suitability analysis. The concept of the classic land evaluation developed by the FAO (FAO, 1976) is used and slightly adapted to find the most suitable areas for banana. Properties of the land are compared to production requirements of banana. The properties analysed in this suitability analysis are Temperature, Precipitation, the Driest Quarter, Drainage, Soil type and Slope. These properties are considered important factors for growing banana and are often referred to as limiting factors. Data on these properties are available through various sources which are explained in detail below. The requirements are ranked into the following suitability classes: S1 (highly suitable), S2 (moderately suitable), S3 (marginally suitable) and N (not suitable). The boundaries for these suitability classes are determined by contemplating literature and consulting experts. Accordingly, the properties are reclassified to these suitability classes. The separate suitability maps for each property are then combined into one suitability map by calculating a weighted average.

2.2.2.1 Production systems

Around the world there are different banana production systems, which have different banana production requirements. In some countries, banana is heavily irrigated to overcome the dry season, while in other countries fertilizer and pesticides are applied to guarantee high yields. Furthermore, large mono-crop banana cultivation has limitations on slopes steeper than 8 percent, while small scale multi-cropping systems can be found on slopes up to 30 percent. (Gowen, 2012). To account for these differences, a scale is defined that separates countries based on the average yield. This yield scale is a measure of intensity of the production system, differentiating High Intensity Production Systems (HIPS) from Low Intensity Production Systems (LIPS). The FAO production statistics include the average yield of each country, expressed in tonnes of fresh weight per hectare. These yields range from 0.8 to 77.7 fresh weight/ha. Three yield classes are defined based on countries for which the production systems and corresponding yield ranges are known. These classes are; low yield (1), medium yield (2) and high yield (3), with yield ranges of 7.936 - 25.000 hg/ha, 25.000 – 40.000 hg/ha and 40.000 - 776.610 hg/ha respectively. The reclassification of the properties into a suitability class differs depending on the country's average yield. For example, a low annual rainfall is reclassified as marginally suitable (S3) for a low-yield country, while the same annual rainfall is reclassified as moderately suitable (S2) for a high-yield country.

2.2.2.2 Temperature, precipitation and the driest quarter

From WorldClim2 monthly temperature, monthly precipitation and rainfall in the driest quarter are downloaded. Banana requirements are determined based on consultancy reports (Soto Ballestero, 2014; Elbehri et al., 2016; Sastry, 1988), the Ecocrop database (FAO, 2019) and expert consultations at The National Banana Corporation of Costa Rica (Corbana). Banana is very sensitive to these climate factors. Optimal yearly average temperature is between 23 °C and 30 °C and growth is severely hampered when the average temperature drops below 12 °C (Gowen, 2012). Annual precipitation amounts between 1200 mm and 3000 mm are required for banana to grow optimally (Ritung et al., 2007). Moreover, if the annual precipitation is enough but it all falls during one part of the year, the banana plants will experience drought stress. In this case banana needs to be irrigated. The minimum amount of precipitation during the driest quarter is based on the minimum yearly average. The data is reclassified according to these requirements. For LIPS countries, 300 mm was classified as highly suitable, for HIPS countries 150 mm. Temperature is not reclassified depending on yield scale, since it is not possible to increase or decrease this property by means of management practices. Precipitation and rainfall in the driest quarter do have different reclassification tables for each yield scale, because HIPS countries are more likely to apply irrigation than LIPS countries. Table 2, Table 3 and Table 4 show the complete reclassification schemes for the three properties, the green values represent the most suitable ranges.

Table 2. Temperature (°C) reclassification table

Suitability class	Lower limit	Upper limit
N	-54	12
S3	12	15
S2	15	23
S1	23	30
S2	30	35
S3	35	42

Table 3. Precipitation (mm/yr) reclassification table

	Yield scale 1 (low yield, LIPS)		Yield scale 2 (medium yield)		Yield scale 3 (high yield, HIPS)	
Suitability class	Lower limit	Upper limit	Lower limit	Upper limit	Lower limit	Upper limit
N	0	650	0	325	0	150
S3	650	900	325	650	150	325
S2	900	1200	650	900	325	650
S1	1200	3000	900	4000	650	5000
S2	3000	4000	4000	5000	5000	7000
S3	4000	5000	5000	7000	7000	10000
N	5000	12000	7000	12000	10000	12000

Table 4. Rainfall in the driest quarter (mm/0.25yr) reclassification table

	Yield scale 1 (low yield, LIPS)		Yield scale 2 (medium yield)		Yield scale 3 (high yield, HIPS)	
Suitability class	Lower limit	Upper limit	Lower limit	Upper limit	Lower limit	Upper limit
N	0	200	0	150	0	100
S3	200	250	150	187.5	100	125
S2	250	300	187.5	225	125	150
S1	300	600	225	600	150	600
S2	600	800	600	800	600	800
S3	800	1000	800	1000	800	1000
N	1000	1649	1000	1649	1000	1649

2.2.2.3 Drainage

Various sources indicate the importance of drainage for growing banana (Robinson and Bower, 1987); (Irizarry et al., 1980). Banana needs enough oxygen in the root zone to aerate the roots (Gowen, 2012). Production is optimal with a groundwater table depth of 1.2m (Ghavami, 1976). Information about the depth of the water table is available through the earth2observe website in the form of a global dataset (https://wci.earth2observe.eu/thredds/catalog/usc/water-table-depth/catalog.html). Areas with a shallow groundwater table are less suitable for banana. The Ecocrop database is used to assign suitability classes to different water table depths. In this process, often threshold values are interpolated to generate the reclassification table (Table 5). The yield scale is considered as well to represent the construction of ditches to improve drainage, which often occurs in HIPS countries.

Table 5. Water table depth reclassification table

	Yield scale 1 (low yield, LIPS)		Yield scale 2 (medium yield)		Yield scale 3 (high yield, HIPS)	
Suitability class	Lower limit	Upper limit	Lower limit	Upper limit	Lower limit	Upper limit
N	0	0.2				
S3	0.2	0.5	0	0.2		
S2	0.5	0.8	0.2	0.4	0	0.2
S1	0.8	2	0.4	4	0.2	6
S2	2	4	4	8	6	12
S3	4	8	8	16	12	24
N	8	999.999	16	999.999	24	999.999

2.2.2.4 Slope

Banana plantations are only found in areas with minimal slopes in order to be able to transport the banana bunches to the packaging site. Therefore, HIPS countries will not have banana in undulating areas. Slopes of up to 0.5% are considered suitable for HIPS countries. This threshold was obtained from field visits to plantations in Costa Rica. On the contrary, LIPS countries often produce banana on slopes, as it is also a means of countering soil erosion (Karamura et al., 1998). The GMTED2010 dataset provides a global Digital Elevation Model (DEM) with a standard deviation map of the elevation on a 30 arc seconds resolution. The standard deviation gives an idea of the slope in an area. This map is reclassified into suitability classes according to the yield scale; HIPS countries get a lower suitability in areas with more elevation differences compared to LIPS countries.

Table 6. Slope (m /1000 m) reclassification table

	Yield scale 1 (low yield, LIPS)		Yield scale 2 (medium yield)		Yield scale 3 (high yield, HIPS)	
Suitability class	Lower limit	Upper limit	Lower limit	Upper limit	Lower limit	Upper limit
S1	0	20	0	10	0	5
S2	20	40	10	20	5	10
S3	40	80	20	40	10	20
N	80	1408	40	1408	20	1408

2.2.2.5 Soil type

Soil types are defined based on the combination of the soil's chemical, physical and biological properties. The soil properties that are considered in the reclassification of the soils are the soil's texture, water availability and drainage, possible toxicity, and soil fertility. According to Purseglove (1988) bananas require a deep well drained loam soil with high humus content and good water holding capacity in order to grow effectively. An expert of Corbana, R. Bonilla, who is specialized in the role of soils in banana agriculture, also stressed the importance of drainage and water holding capacity. The S-world dataset provides a global map with soil types. The soil types are analysed and reclassified into suitability classes for the growth of banana. Management practices can improve the land's suitability for banana. The reclassification scheme is adapted in accordance with the yield scale of the country. Water availability and drainage were most determining for the suitability class of a soil. R. Bonilla stated that soil fertility was less limiting in HIPS countries. HIPS countries have the ability to improve the fertility of their agricultural land relatively easily by applying fertilizer and decrease toxicity by adding minerals. Nutrient-poor soils are assigned a low suitability for LIPS countries while the suitability is considered slightly better for HIPS countries. Furthermore, water availability can be controlled through irrigation and drainage networks in HIPS countries. Where heavy clay soils pose a great water logging issue in LIPS countries, HIPS countries will construct an extensive drainage network. On the other hand, sandy soils provide good aeration of the root zone but precipitation is discharged so quickly that the banana roots cannot absorb enough water and thus experience drought stress. In HIPS countries this problem is overcome by providing drip irrigation near the stems of the plants. In general, soils in HIPS countries were assigned higher suitabilities than soils in LIPS countries. The scheme below gives an overview of the properties and their suitability for banana. The complete reclassification table for all soil types per yield class can be found in appendix B.

Table 7. General overview of soil reclassification. The arrows indicate gradations of suitability in between.

	Texture	Water availability & drainage	Fertility & Toxicity
Not suitable	Heavy clays	Very dry	Toxic
	↑	↑	↑
Suitable	Loam soils	Sufficient water	Nutrient rich
	↓	↓	↓
Not suitable	Fine sands	Very wet	Nutrient poor

2.2.2.6 Combined suitability map

The six properties (temperature, precipitation, rainfall in the driest quarter, water table depth, slope and soil type) described in the previous sections are combined into one suitability map. In a classic land evaluation the suitabilities for each of the properties are added together and a suitability is assigned according to the most severe limitation. The approach in this study deviates from the standard approach by first computing a weighted average. Reason for this deviation is that the six properties are not considered to be equally important for the suitability of banana. Furthermore, the assignment of suitability values to land units is based on the average of all the properties, after filtering out the areas that are too cold. The choice to put temperature as a hard boundary before calculating the average suitability comes from the fact that temperature cannot be increased or decreased by means of management practices and is therefore considered most limiting. Weights are assigned to the properties depending on how much the property influences the suitability for banana (Table 8). Temperature, slope and the driest quarter are weighted equally. Water table depth, precipitation and soil type are assigned lower weights. Soil type is analysed mainly on the texture, water availability, drainage and toxicity. Texture, water availability and drainage show resemblance in the properties water table depth and precipitation. Therefore, their weights are lowered to make the other factors count more towards the combined suitability map.

Table 8. Relative weight of the properties in the suitability analysis.

Property	Weight
1. Temperature	1
2. Slope	1
3. Driest quarter	1
4. Water table depth	0.75
5. Precipitation	0.75
6. Soil type	0.5
Total	5

Then the combined suitability map is created. The suitability values (S1, S2, S3, N) are first converted to numbers (3, 2, 1, 0) in order to use them in calculations. These values are then multiplied by the weight of the property. Finally, all the suitability values are added up and divided by the total weight.

$$Sc = \frac{Sp1 * Wp1 + Sp2 * Wp2 + Sp3 * Wp3 + Sp4 * Wp4 + Sp5 * Wp5 + Sp6 * Wp6}{Wtot}$$

In which:

Sc = combined Suitability

Spn = Suitability of property n (n= 1-6)

Wpn = Weight of property n (n= 1-6)

Wtot = total Weight of all properties

The result is one map with suitability values between 0.00 and 3.00. These are converted back to suitability values (N = 0.00 -0.75, S3 = 0.75 -1.50, S2 = 1.50-2.25, S1 = 2.25 – 3.00).

2.2.3 Predicting banana distribution

Banana areas are predicted by adding the results of the land suitability evaluation to the land use analysis. Per country, the harvested area is distributed over the potential areas indicated on the land use map, while also considering the suitability map. The cells in the land use map show percentage values representing the area that is used to grow banana. The suitability map shows a suitability value, indicating how well banana can grow in that cell. The land use map is leading in the process, because it tells us what is observed in the pixel. The harvested area is first allocated to the areas in the land use map that are most suitable (S1). In case there is still banana area left after filling the most suitable areas, the area with a S2 suitability is filled, thereafter S3. This approach assumes that banana is produced in the most suitable areas. In countries that have more S1 area than banana area, banana area is divided equally over the S1 area. The same applies to S2 and S3 areas. The final banana map indicates for each cell the percentage of the cell used to grow banana. This percentage is calculated by dividing the banana area of the cell by the total area of the cell.

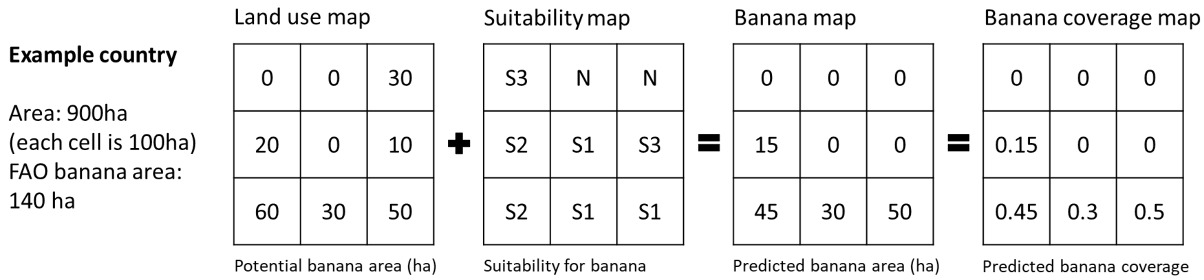


Figure 2. Method for predicting banana distribution.

Figure 2 shows a visual representation of the prediction method. The 140 ha is first divided over the S1 area, filling all the potential banana area. This adds up to a total of 80 ha. The remaining 60 ha (140 ha – 80 ha) is distributed over the S2 area. The total S2 area in this country is 80ha (20 ha + 60 ha), so the 60 ha is distributed proportionally. Resultingly, the S2 cells get a value of 20/80*60 = 15 ha and 60/80*60 = 45 ha. The banana coverage is then calculated by dividing the predicted banana area by the total cell area (divide by 100 ha).

The results are created using ArcGIS pro software (ESRI, 2019). The coordinate system on which all data is projected is the standard World Geodetic System 1984 (WGS84).

2.3 Method calibration

Plantations were visited in Costa Rica to get an idea of how banana is cultivated and to see if the method's decision rules are in line with the observations. Moreover, the disaggregation method is calibrated by consulting experts from Corbana.

In order to get a better understanding of large-scale banana production, several plantations were visited in Costa Rica. These field visits were aimed at identifying properties in the cultivation of the widely exported Cavendish variety that are important for the mapping of banana. The implementation of these properties in the model could then increase the accuracy of predicting banana distribution. Although the field visits were mainly focussed on the identification of intensive banana production areas, smallholder farms in the highlands were visited as well. These farms produce the Gros Michel variety, which is heavily affected by Fusarium TR1, leading to a different production system in which the plant dies within a few years and new plants are planted in a new part of the farm. These systems play an important role in the spreading of diseases and the observations were considered in the model as well. Table 3 shows the properties that were analysed during these field visits.

Table 9. Banana production properties analysed during field visits.

Property	*Method implementation*
Size of the plantation	Recognition possibility of banana on 30 arc seconds resolution
Height of the banana plants	Determine canopy height reclassification (Land use analysis)
Type of drainage system	Water table depth suitability classes (Land suitability evaluation)
Slope of the plantation	Standard deviation of height (Land suitability evaluation)

Furthermore, the maps were calibrated by consulting experts with knowledge on banana regions and production systems. The predicted banana map for Costa Rica was shown to the expert, leading to a discussion on the correctness of the map. In case of a mismatch, the next step was to identify the property leading to the discrepancy. The model could then be altered to account for this property, resulting in a new version of the map. In this iterative process, the new maps were subjected to an expert's eye each time the map was improved. The findings are discussed in the results section.

3. RESULTS

This section describes the results of the steps in the disaggregation method. The stepwise approach makes this study transparent in a way that the influence of each dataset on the result can be viewed separately.

3.1 LAND USE ANALYSIS

3.1.1 FAO NATIONAL PRODUCTION STATISTICS

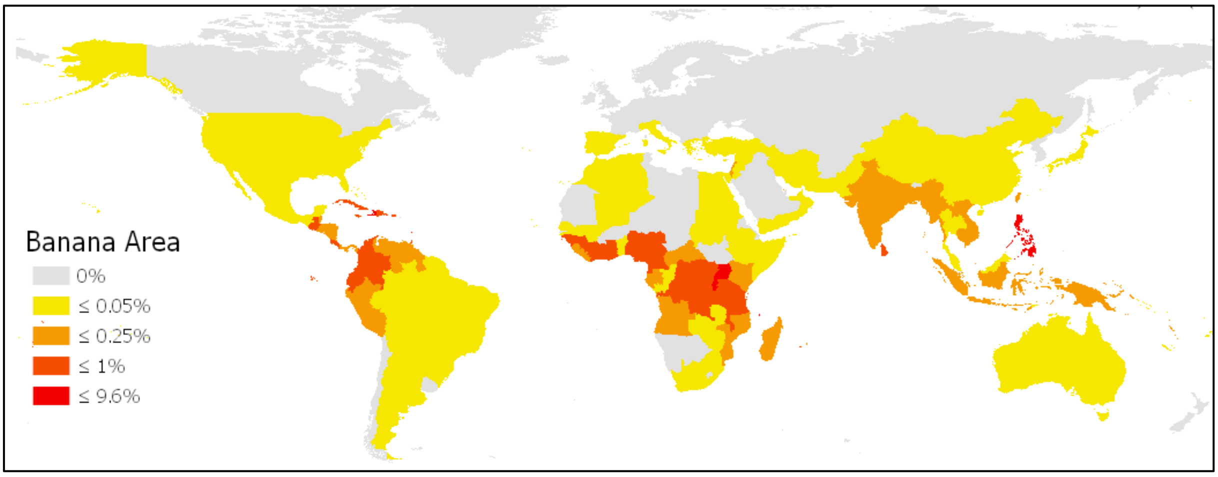


Figure 3. Percentage of country used for banana production (data from: FAO-stat, 2019).

Figure 3 shows the starting point of this study, the banana production statistics per country. Apparent is that most banana producing countries are located around the equator. The percentages represent *Pcountry* (*Eq. 1*). Some countries use almost 10 percent of their total area for banana production, while larger countries only reach 0.05 percent.

3.1.2 Land cover

Figure 4 shows the potential banana areas after applying the GlobCov dataset, the percentages represent *Pagri* (*Eq.* 2).

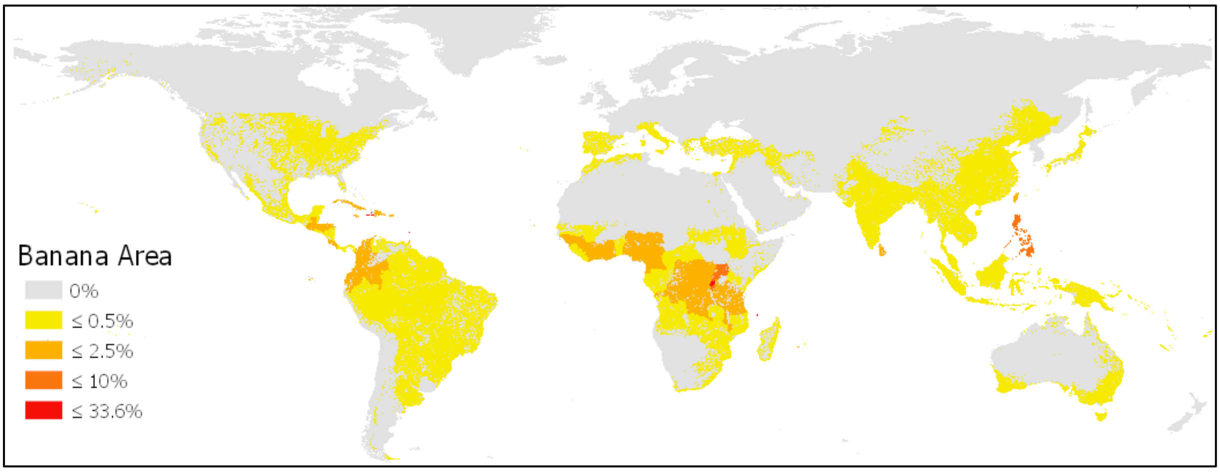


Figure 4. Percentage of agricultural area used for banana production per country.

The banana area percentages have increased significantly compared to Figure 3. It seems that for smaller countries the percentages go up fast, because of their relatively small area. The global extent has shrunk due to the removal of non-banana agricultural lands. Alaska and Australia for example, have significantly reduced potential banana area.

3.1.3 Normalized Difference Vegetation Index (NDVI)

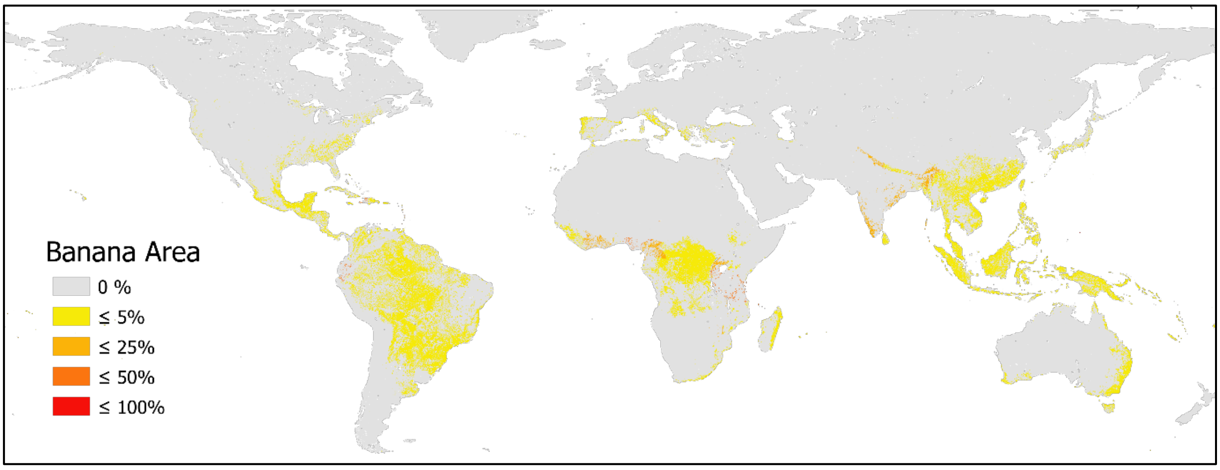


Figure 5. Percentage of perennial agricultural area per country that is used for banana production.

It is evident that the potential banana area is reduced further, and concentrates around the tropics. Notice that some very small countries (islands in the pacific) now have banana percentage values of 100. The map still shows potential banana in the Amazon and other tropical rainforests in Africa and Southeast Asia, due to the choice made in the reclassification of the land cover dataset.

3.1.4 Canopy height

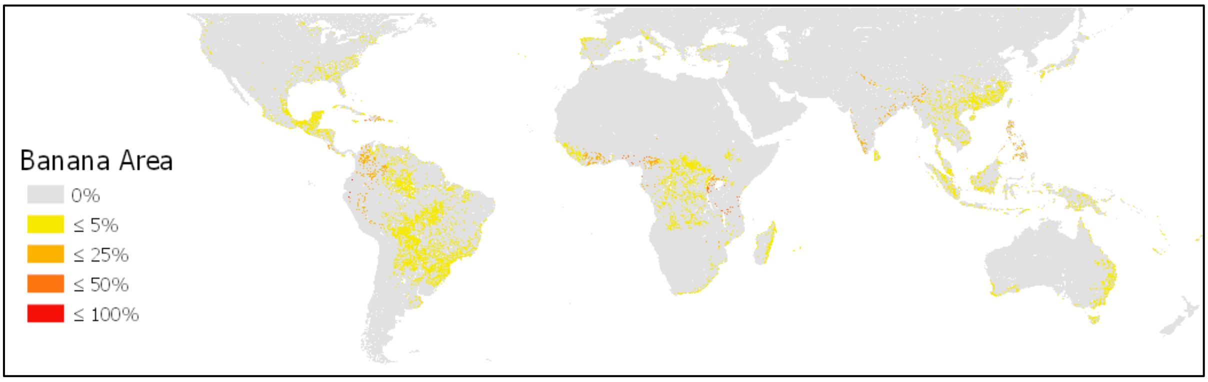


Figure 6. Percentage of perennial agricultural area with canopies below 25 m per country that is used for banana production.

The potential banana area has been narrowed down to small margins. There are large areas with up to 5 percent banana, indicating that 95 percent of this area is not used for banana production. These areas indicate smallholder production where banana is sparsely spread. On the other hand, there are countries that have up to 100 percent of their potential banana area in use for banana production. The next step in this study was to conduct a land suitability analysis to find the areas, within the potential banana areas, that are most suitable to grow banana.

3.2 Land suitability evaluation

Figure 7 shows the suitabilities of all individual properties. The maps show the suitabilities for the reclassification of the second yield scale, so the thresholds are not correct for all countries in these maps. However, suitable areas are already clear from these maps. The most constraining properties were temperature and precipitation. Most of China is too cold to grow banana and is therefore completely red. The cold Andes mountain range in Chile is also apparent from this property. Areas that receive insufficient precipitation can be observed in large parts of Australia, northern and southern Africa, and the US. These regions also cope with seasonality, which leads to unsuitable areas due to a precipitation deficit in the driest quarter of the year. Water table depth resembles the precipitation pattern, but shows more local differences, resulting in scattered areas of high suitability in, for example, the Amazon. Slope is often related to mountain ranges and mainly HIPS countries have large areas of low suitability due to this property (Figure 9).

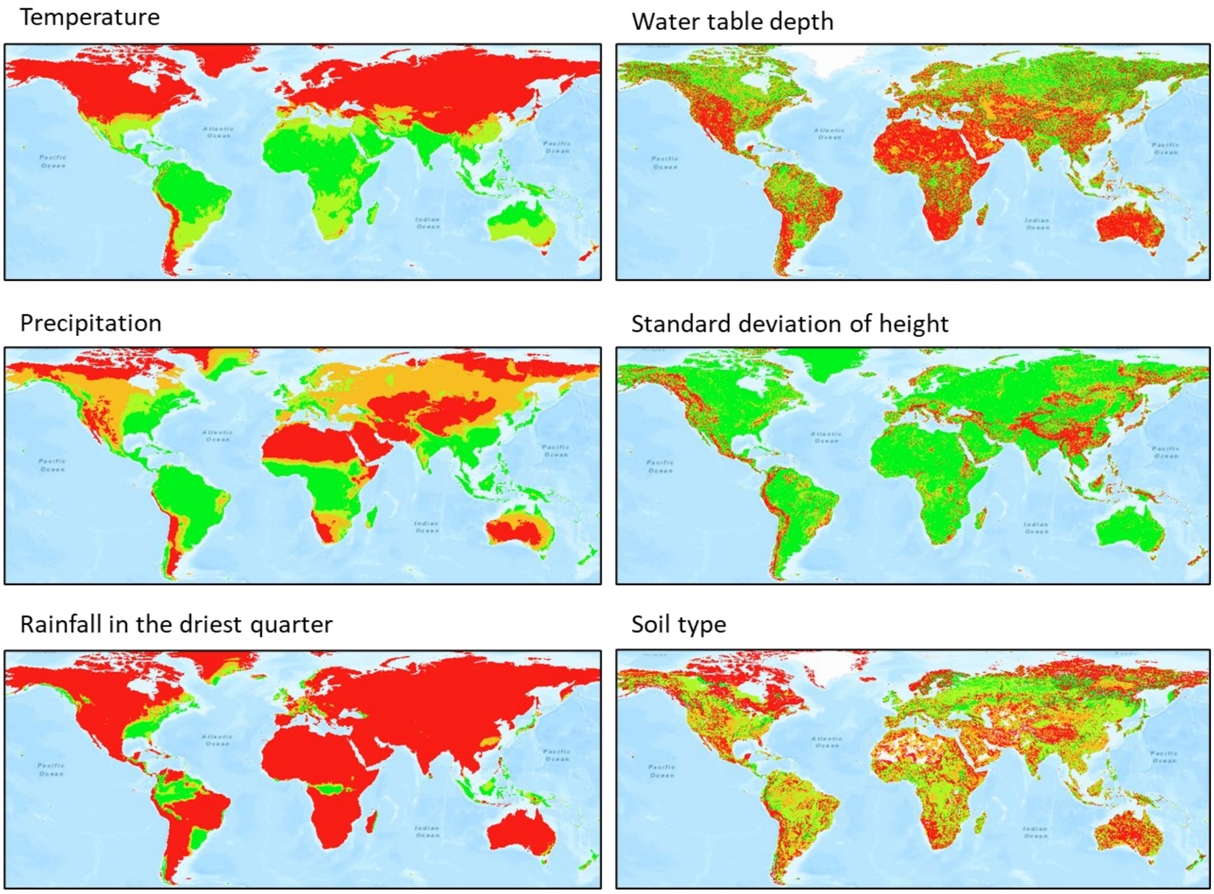


Figure 7. Suitability maps for all properties for yield scale 2. Green indicates suitable, red unsuitable.

After evaluation of land properties, one global suitability map was created (Figure 8). The suitability map gives a good overview of where banana can be grown. The map only shows the countries in which banana is produced. The most suitable areas are located around the equator.

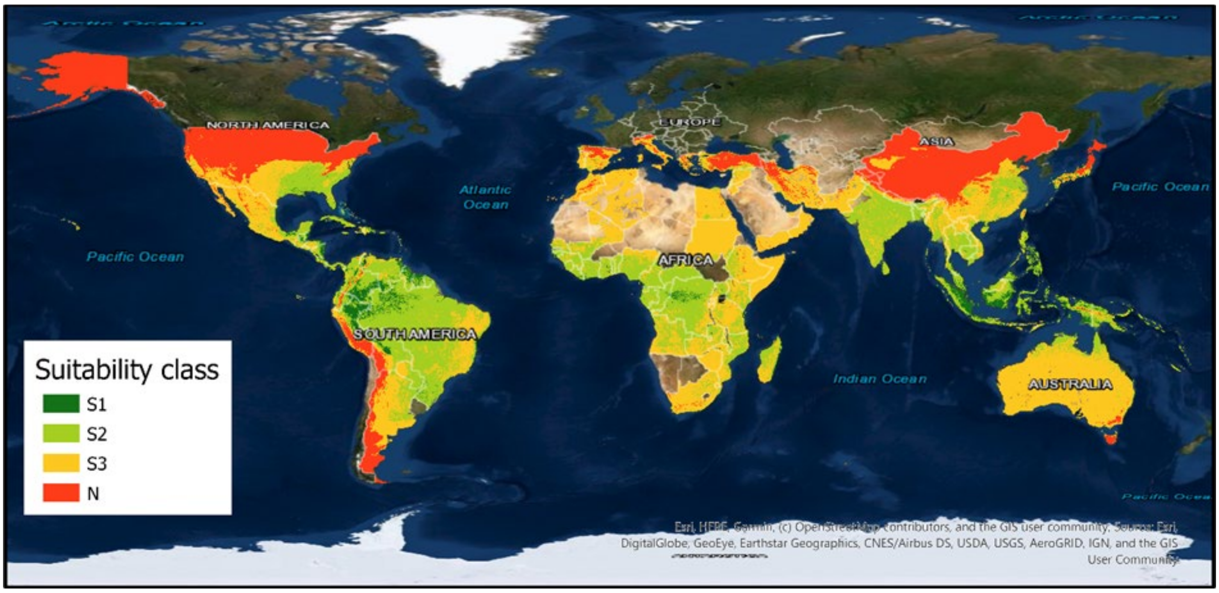


Figure 8. Global suitability map for banana. Suitabilities are only shown for countries that produce banana.

Identifying suitable areas at a national scale requires a zoomed in map. Figure 9 shows the suitability map for the country of Costa Rica. Here you can see the suitable areas in the east while the western part is less suitable. At this national scale, the driest quarter property plays a significant role in the suitability analysis. Costa Rica is located close to the equator, receives a lot of rain and average temperatures are high. Nonetheless, the western part of Costa Rica is very dry for a large part of the year as a result of seasonal fluctuation, while the eastern part has ample precipitation all year long. The two regions are split up by a mountain range in the middle of the country. This area is depicted by the red area in the map, the mountains are too cold and therefore non-suitable. The slopes of the mountains are too steep to construct a banana plantation and are therefore moderately suitable.

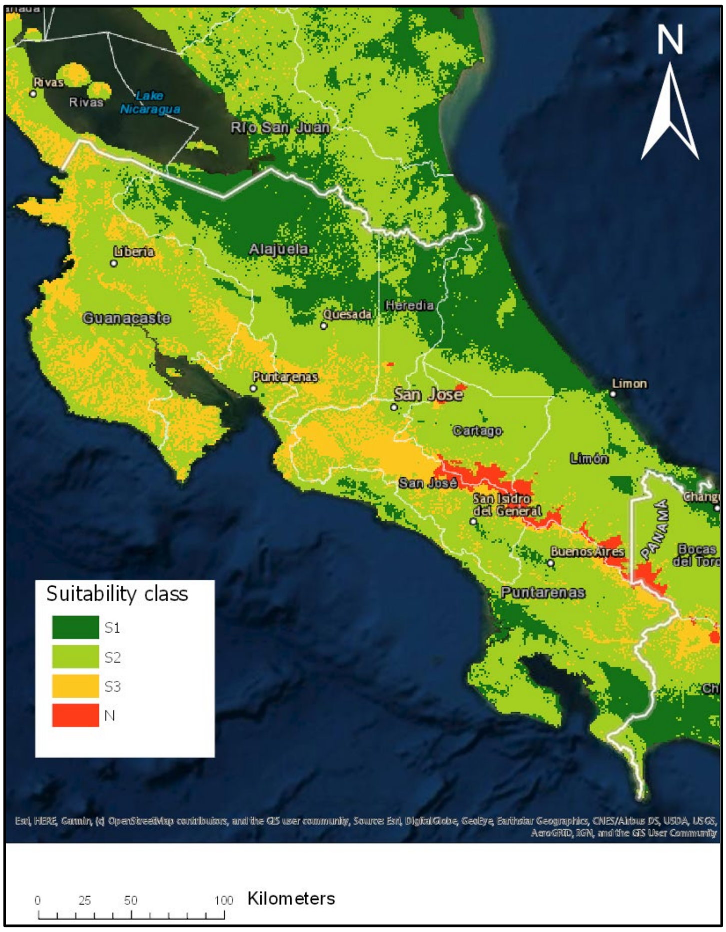


Figure 9. Suitability map of Costa Rica.

3.3 Predicting banana distribution

The results of the land use analysis (Figure 6) were combined with the results of the land suitability evaluation (Figure 8) to predict global banana distribution (Figure 10). The cells have been aggregated to a 10km resolution for visual interpretation. The values represent the banana coverage of a cell. The coverage values for most areas are rather low for most areas because they represent 1 cell of 1 km^2.

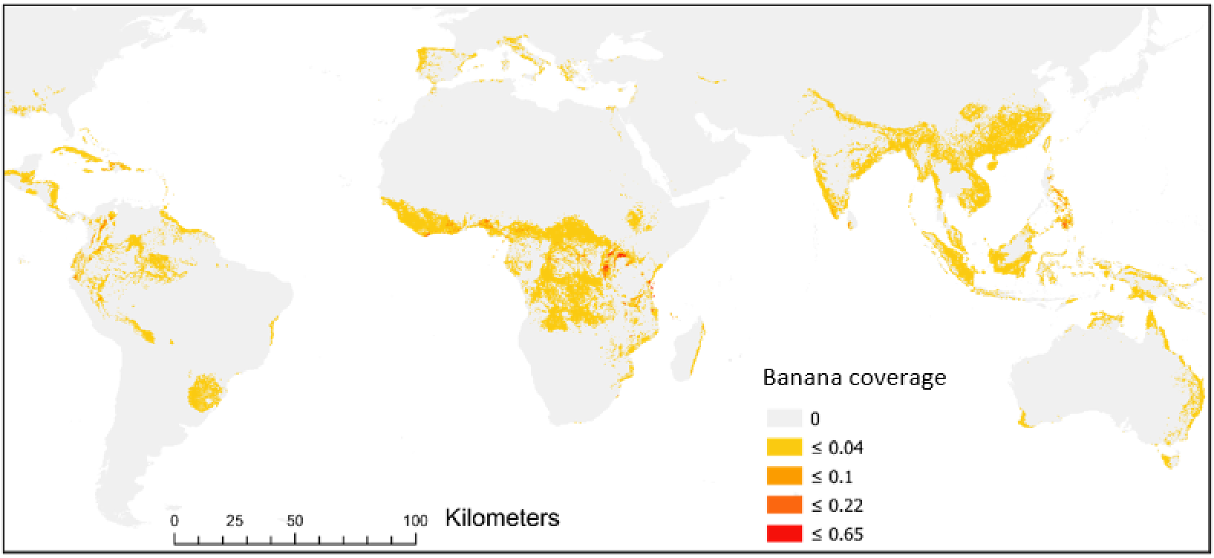


Figure 10. Global predicted banana map. The map has been aggregated to 10 km resolution for visual interpretation.

Figure 11 shows the predicted banana areas for Costa Rica. For Costa Rica, the model allocated all banana area to the most suitable area since there is more S1 area available than there is banana area to be distributed. The highest coverage for a cell in Costa Rica is below 0.33. The banana producing regions are clearly visible on the map.

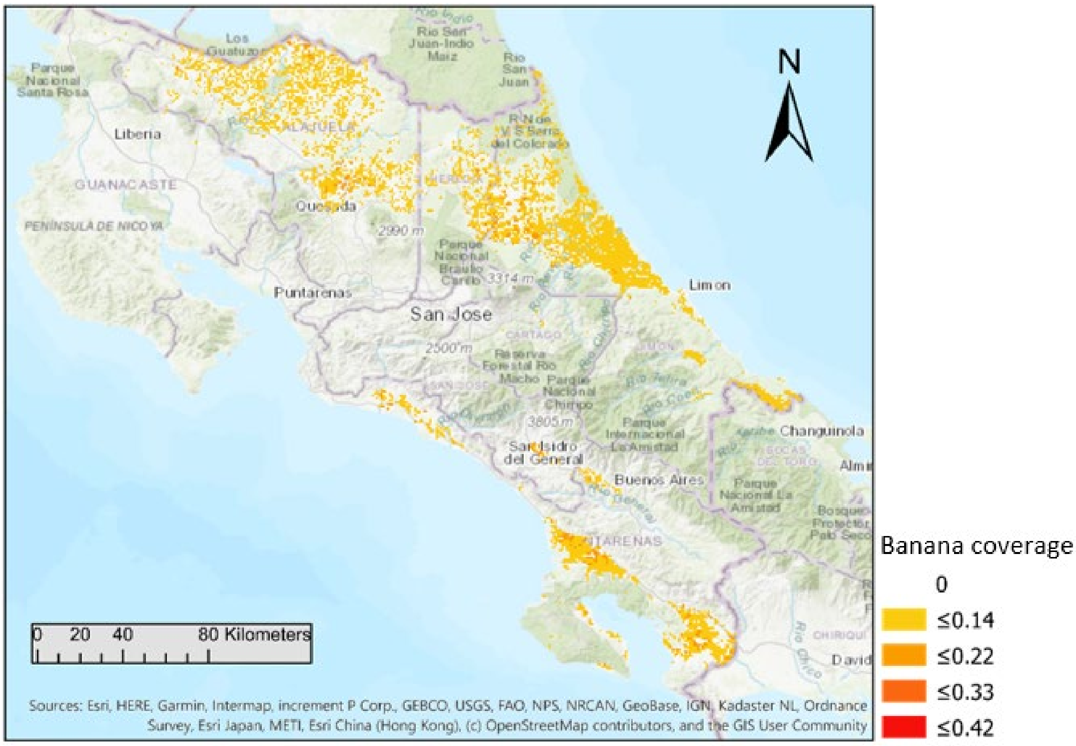


Figure 11. Predicted banana areas for Costa Rica, percentages indicate the banana coverage of a cell.

3.4 Method calibration

The most noteworthy observations of the field visits are described here. Additionally, outcomes of the consults with experts are described, including adaptations to the model. Elaborate reports on the farm visits and interviews can be found in appendix C. The changes made to the reclassification tables are already implemented in the tables of Appendix A (for the land use analysis) and B (for the land suitability evaluation).

3.4.1 Field visits

Plantations are usually 100ha (or 1 km^2) or larger and are therefore recognizable on the 1 km resolution datasets used in this study. However, many single banana plants and small plots of banana plants can be found in people's gardens for own consumption or for local trade. On a 1 km resolution these areas remain unidentified when looking at land use only, accentuating the importance of considering land suitability as well.

Plant height of the banana plants on the plantations is between 4-6 meters. This was compared to the global canopy height map on the same locations to define the threshold value of 25m for banana agriculture in the land use analysis (section 2.2.1.4).

In areas with high annual rainfall rates and a high water table often extensive drainage networks are present with ditches up to 6 m in depth, sometimes combined with a pump to get rid of excessive water. The reclassification of the water table depth dataset was adapted such that HIPS countries got a higher suitability class for areas that have a high water table and excessive rainfall. Appendix A represents the adapted reclassification tables.

Elevation differences on plantations are almost equal to zero, this is required in order to transport the heavy bunches from the plant to the packaging site in the centre of the plantation. Smallholder farms on the other hand, are found on slopes up to 45 degrees. These differences are all within the HIPS country of Costa Rica, making it hard to make a classification table that finds all banana.

3.4.2 Expert consultations

The consultations held with different experts led to changes in the prediction method. The suitability map was enhanced multiple times and the land use analysis was improved.

Most noteworthy is the discrepancy in the central area of Costa Rica, where the map did not show banana while multiple experts knew there was banana. Next to the observation of banana on steeper slopes in the field, the consultations showed that banana is used as a shadowing plant to produce coffee. On one hand there is large scale banana production in the flat lowlands of Costa Rica, while on the other hand banana is grown as shadowing for coffee on steep slopes in the mountainous inlands. The coexistence of these different production systems within one country, and thus in the same yield-scale, make it difficult to discriminate between them. To adapt to the circumstances of smallholder production, a second version of the model was created in which the slope was not implemented as a limiting factor. This gave somewhat better results for Costa Rica, as the output map shows the areas of smallholder production. However, the main areas were represented less accurately in this version. Figure 12 shows the difference, left with the slope considered and right without. The map on the right shows more banana in the centre, just east of San Jose. This is where the smallholder farms are on the steeper slopes. However, it also shows more banana in the other parts of the country where no banana is grown, especially in the south and north.

The decision was made to assign a low suitability to areas with steep slopes in HIPS countries, neglecting smallholder production. The model was aimed to identify the main banana areas in a country rather than small areas where banana is grown.

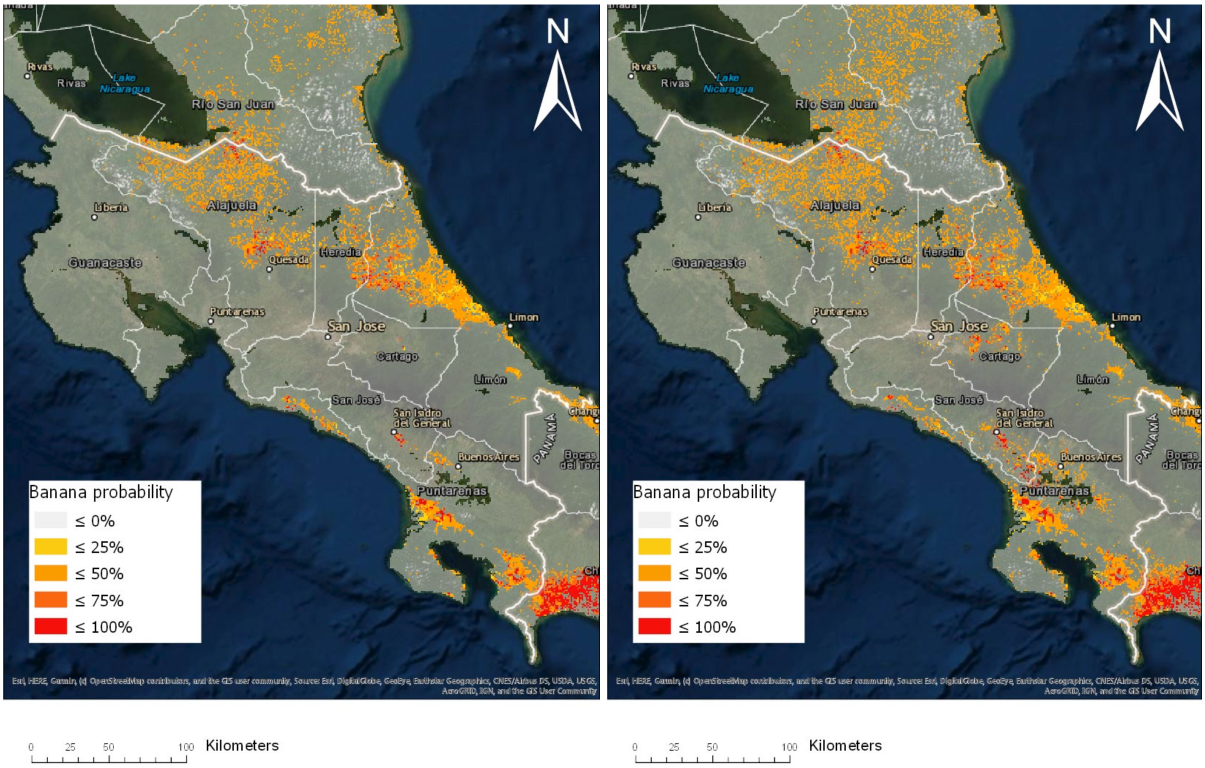


Figure 12. Left: predicted banana areas considering slope. Right: predicted banana areas without considering slope

Furthermore, the weight of the Driest Quarter rainfall property was increased from 0.5 to 1.0 in the weighted suitability analysis. J. Guzman-Alvarez from the BanaClima network of Corbana stressed the vulnerability of banana to drought stress.

Consult with R. Bonilla (Corbana, soil expert) lead to changes in the reclassification scheme of soil types. He emphasized the importance of texture and drainage, while fertility is less limiting. The soil classification was changed accordingly, assigning lower suitability classes to clayey soils and slightly increasing the suitability of soils with a low fertility.

4. DISCUSSION

This section starts with a validation, assessing the accuracy of the banana map. Then there will be a discussion on the method and the model output. Finally, the applicability of the map will be reviewed.

4.1 Validating the banana map

Validation of the banana map is required in order to give it value. Usually, validation is carried out by comparing the results of the model to a reference dataset. Other crop mapping studies compare results to previous research (You et al., 2014). This is challenging for a dataset on a global scale, especially because banana has not been mapped on such a high-resolution before. Global crop mapping does not have consensus yet for comparing results and according to Congalton et al. (2014) it is important that such guidelines are created in order to assess accuracy. True accuracy assessment is impossible as there is no true independent banana map available. However, there are other methods to give value to the map. Three methods are discussed here; (1) validation using satellite images, (2) comparison to a global map and (3) a comparison to local banana maps.

4.1.1 Validation using satellite imagery

The best way to determine the validity of the banana map is to assess the accuracy by comparing the predicted banana areas to satellite images. Satellite imagery generally has a high spatial resolution and extent. Since banana has quite a distinctive pattern from above, banana areas can be recognized on the images. This holds for areas in which banana is produced in a mono-cropping system, like Costa Rica. Complications arise when trying to identify banana areas for regions with multi-cropping systems. Furthermore, clouds form a major obstacle in the identification of banana. Clouds are a common issue in landcover mapping from satellite imagery, especially around the tropics (Mitchard et al., 2012, Lavreniuk et al., 2016). In these areas the satellite images are rather blurry, making it hard to distinguish banana from other plants or trees. Figure 13a, b and c give an idea of the issue.

Figure 13a. Mono-cropping system, banana plantation. Banana pattern clearly visible (source: google earth)

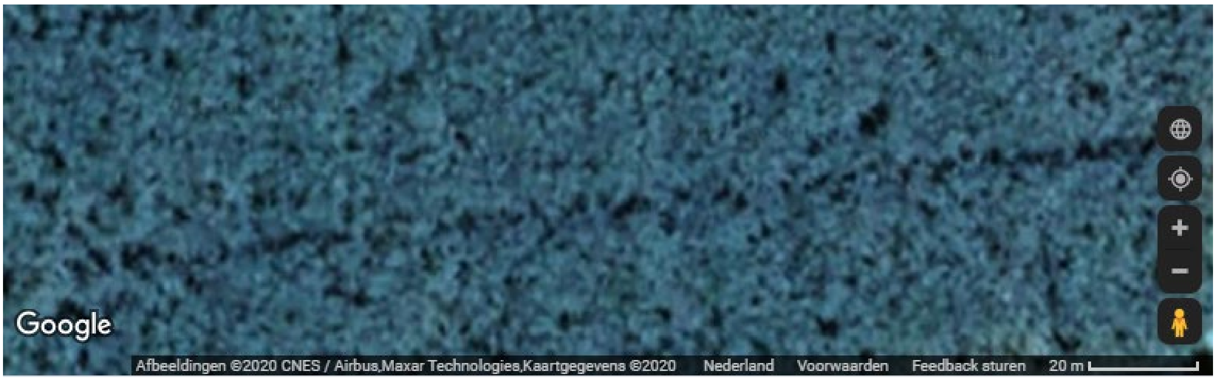


Figure 13b Mono-cropping system, difficult to distinguish banana due to blurry image (source: google earth)

Figure 13c Multi-cropping system, difficult to distinguish banana from other crops. (source: google earth)

Validation through comparison with satellite imagery is therefore not applicable on a global scale, for all production systems. Nonetheless, an assessment of the accuracy was carried out for Costa Rica, where banana is recognizable from satellite images due to the large size of the plantations. Accuracy can be measured from two perspectives, the producer's accuracy and the user's accuracy. As explained by Story and Congalton (1986), the producer's accuracy is the chance for any point on the map to be correctly classified, in this case either as "banana" or "not banana". The number of correctly classified banana pixels is divided by the total number of banana pixels that are in the whole study area. This requires a reference dataset that governs all banana pixels on a 1 km resolution, which is not available. The user's accuracy on the other hand, is a measure for the chance that a pixel without banana is wrongly identified as banana. The number of pixels correctly classified as banana is divided by the total number of pixels classified as banana, including pixels that are not banana but have been classified as such. The user's accuracy was calculated for Costa Rica by comparing the prediction map to satellite imagery from google earth. A higher number of points corresponding to the points on the satellite images indicates a better representation of reality and thus a more valuable map. The comparison was performed on 50 points, in which the minimum distance allowed between the points of the sample was set to 1 km. Table 10 shows the results of the validation. From the 50 locations, 23 were validated as banana areas.

Table 10. Validation results of comparing the banana map with satellite imagery

Predicted banana locations

Reality		Banana
	Banana	23
	Agriculture (not banana)	8
	Forest	7
	23	7
	30	5
	Total	50

The user's accuracy is then 23/50*100% = 46%. More than half of the pixels have been wrongly identified as banana. However, in this accuracy assessment all cells in the prediction map were assumed equally important; any pixel with a coverage larger than 0 was assumed to be banana. In Costa Rica, banana is produced in large scale systems, so in reality coverage of cells is higher for some cells while others have none. Finding banana in 46 percent of the cells that have a coverage of about 10-15 percent is significant. However, the sample was small in this approach and it might not represent the actual accuracy.

4.1.2 Qualitative comparison of banana area on global scale with crop-mapper.

The map was compared to a previously developed banana map, available on crop-mapper.org, to see if the results correspond to earlier studies. This map was created through a survey, asking experts to mark the main banana areas in their country (Crop-mapper, 2008). The level of detail varies significantly between regions, but it gives a general view of the regions where banana is grown globally. Snapshots for the three major banana zones, followed by a snapshot from the predicted banana map from the same area, are displayed in Figure 14-19. The distribution of predicted banana areas corresponds well to the survey-based areas, indicating that the model was able to detect the general banana areas. Most remarkable differences are the areas where polygons on the crop-mapper map are large, often the prediction map shows much smaller areas. The criteria that caused these areas to be excluded are discussed here.

The large banana areas in northern Peru in Figure 14 are not in the map due to high canopies in these areas (see Figure 6). In central Africa (D.R. Congo, Congo, southern Cameroon and eastern Gabon) and eastern Mozambique the crop-mapper map indicates banana (Figure 16). However, the model, removed these areas because the observed NDVI values are below the threshold of 04. The same holds for India, the crop-mapper map shows complete banana coverage (Figure 18), while the results of the model only show banana close to the coast and at the northern border. The inland areas do not show yearly high NDVI values and were removed in the land use analysis. The banana areas in south-east China are smaller on the map due to the low suitability for banana. Precipitation during the driest quarter is insufficient and the groundwater table is too deep. The crop-mapper map shows large areas while the prediction map has a lot more detail. Therefore, the added value of the prediction map is significant compared to the survey-based crop-mapper map.

Figure 14. Banana areas in Mid and South America, the colours represent different cultivars (source: http://www.crop-mapper.org/banana/mapper.php)

Figure 15. Banana areas in Mid and South America according to the prediction method.

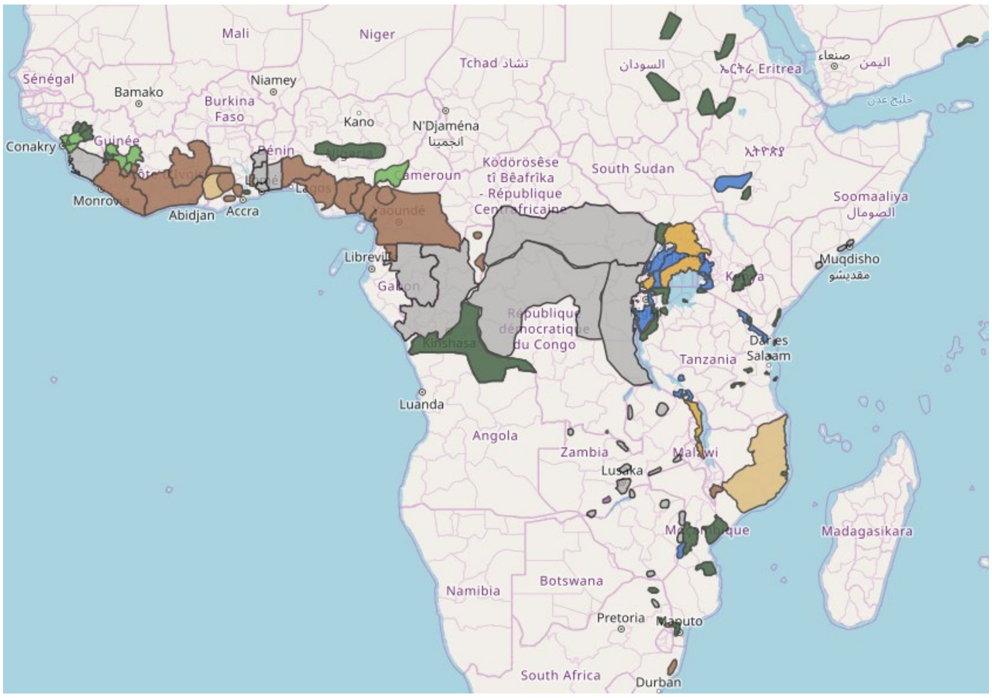


Figure 16. Banana areas in Africa, the colours represent different cultivars (source: http://www.crop-mapper.org/banana/mapper.php)

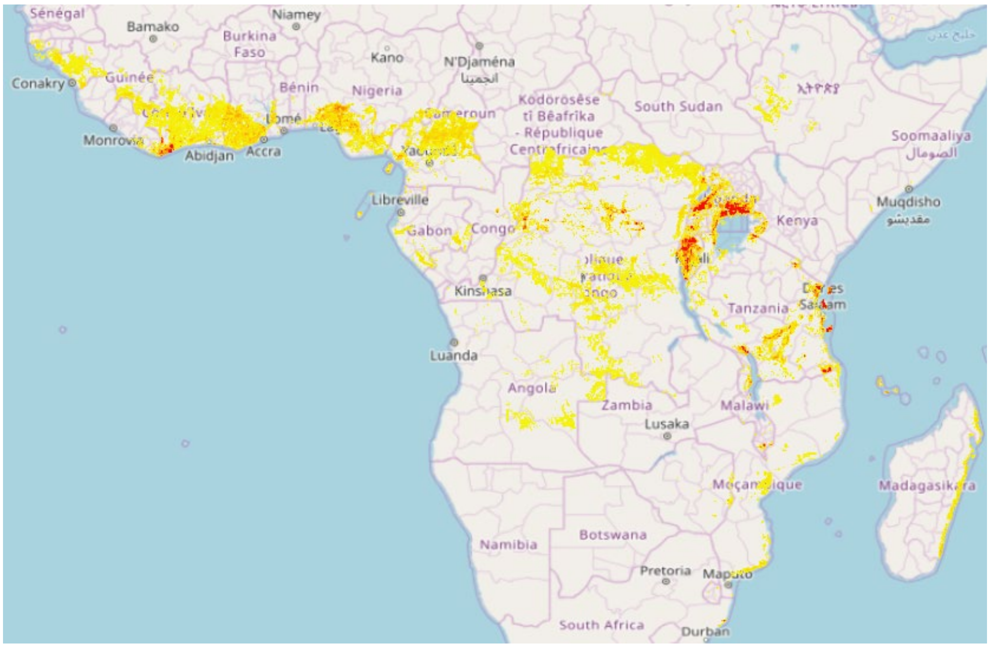


Figure 17. Banana areas in Africa according to the prediction method.

Figure 18. Banana areas in South-east Asia, the colours represent different cultivars (source: http://www.crop-mapper.org/banana/mapper.php)

Figure 19. Banana areas in South-east Asia, according to the prediction method.

4.1.3 Qualitative comparison of banana area on local scale

Finally, a comparison on a local/regional scale can be done for areas where this data is available. For three major banana producing countries, local information about the distribution of banana was collected. The countries are Costa Rica, Uganda and the Philippines. Costa Rica and the Philippines are more focussed on the export of the product, while Uganda produces banana for local trade. For each country the collected data was compared to the results from the prediction model.

Costa Rica

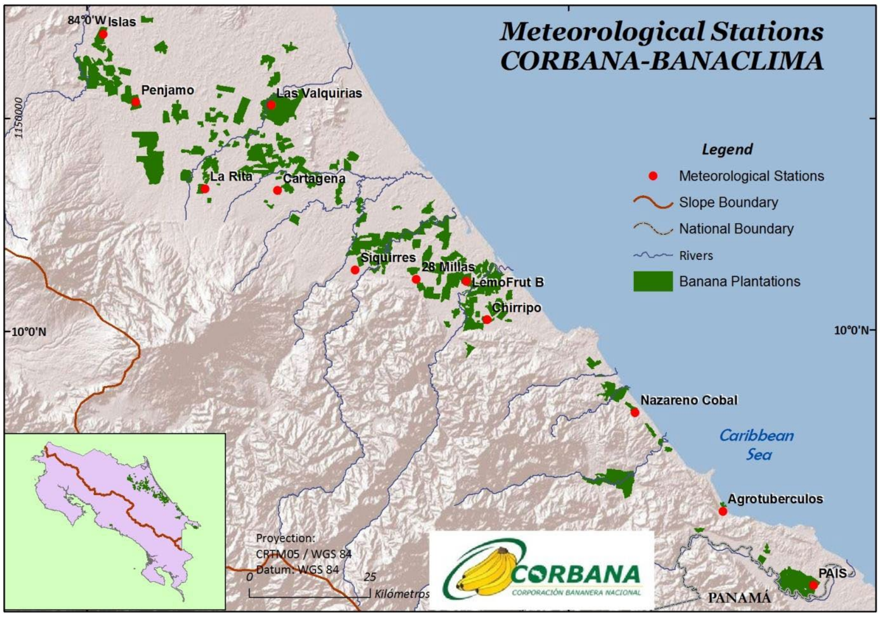


Figure 20. Main banana plantations in Costa Rica (source: Guzmán-Alvarez and Quesada (2014))

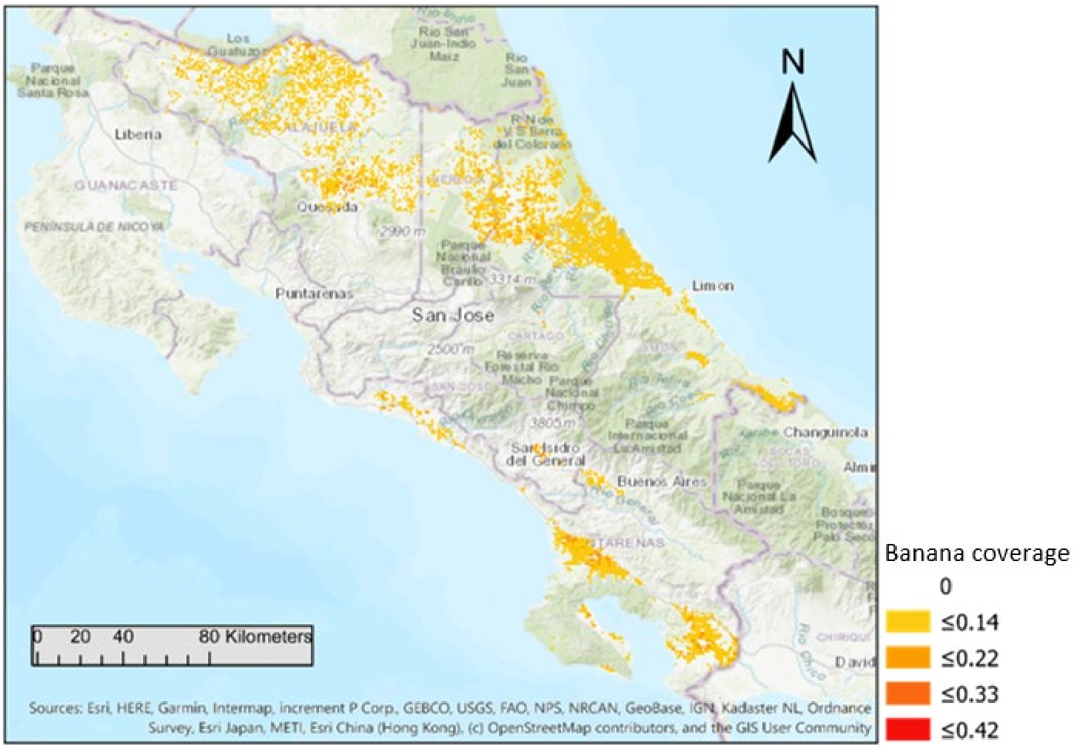


Figure 21. Banana areas in Costa Rica according to the prediction method.

The map of Figure 20 shows the large plantations in Costa Rica. Small farms are not in this map, rather, it represents the extensive banana production. There are some differences between the two maps. Especially in the northern part of Costa Rica, the model finds banana while the plantations map does not show banana. This area is very cloudy and therefore receives little sun throughout the year. Bananas require a lot of sun to grow and therefore are not grown in this area (FAO, 2019). Improvements to the model can be made by implementing sun hours as a factor for the suitability analysis. The banana distribution resembles the plantation locations, but the extent of the predicted areas is too large.

Uganda

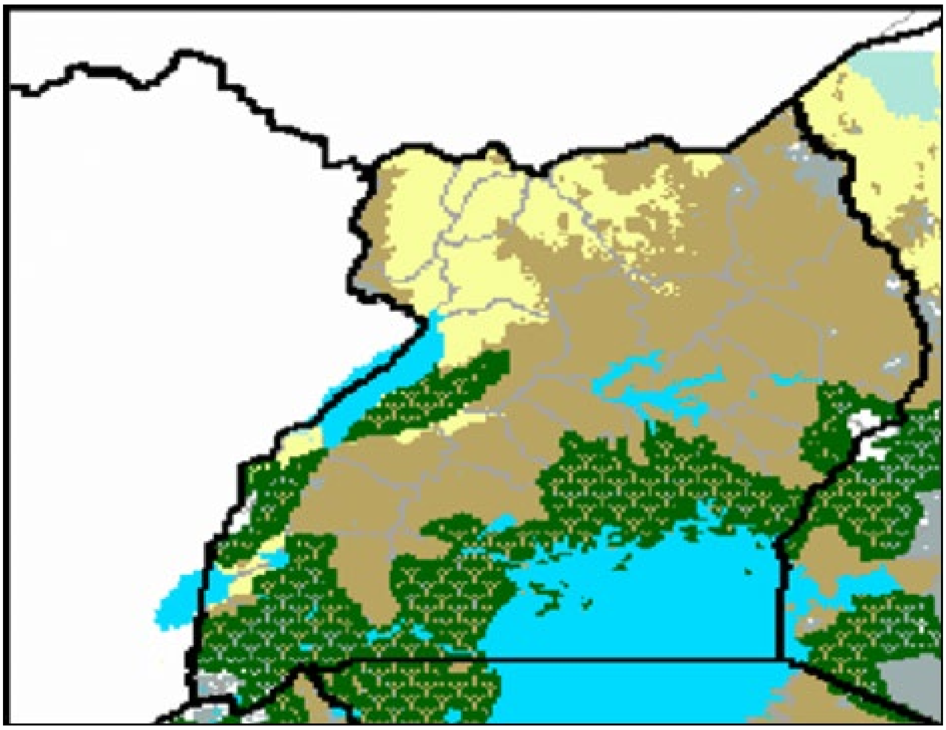


Figure 22. Banana areas in Uganda, dark green represents banana (from Bagamba (2007)).

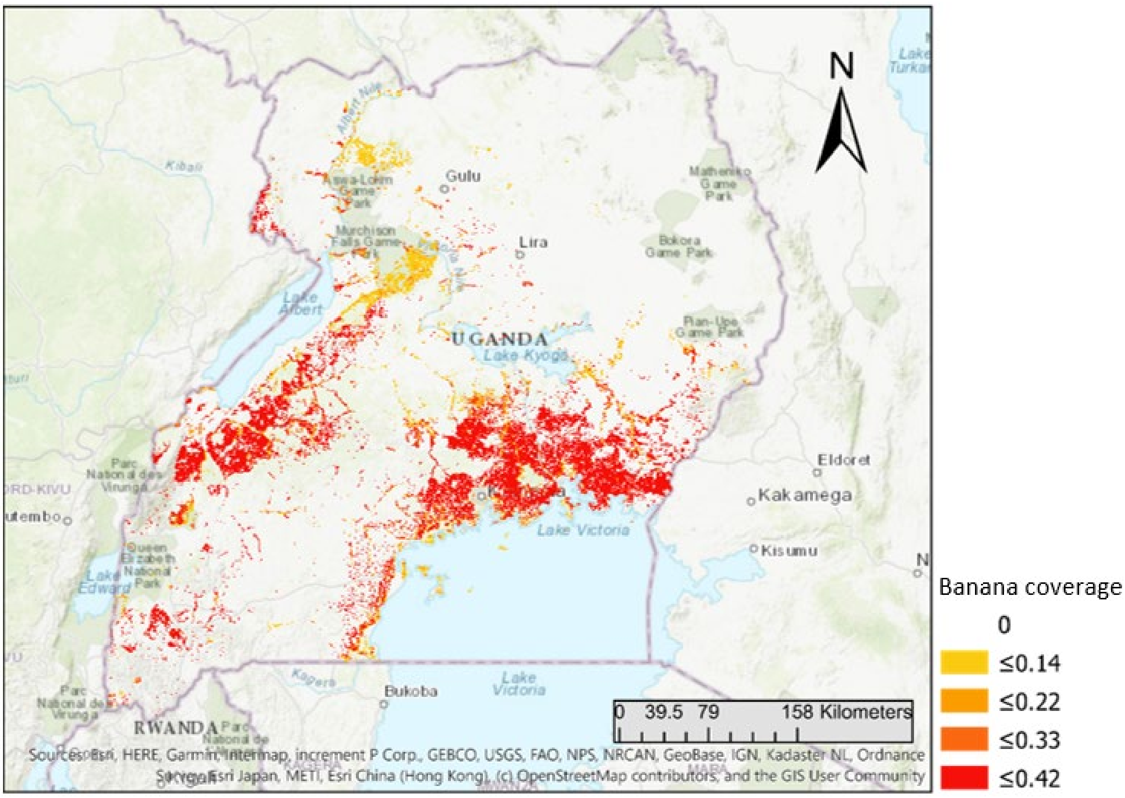


Figure 23. Banana areas in Uganda according to the prediction method.

The data from Uganda (Figure 22) is not as precise as the map for Costa Rica, but it gives an indication of the banana distribution in the country. In Figure 23, it seems that the model predicted the banana areas well for most of the country, with the largest discrepancies in the south-western part of the country. Moreover, the question remains whether Figure 22 is a valid map as it shows data from 2003 while the prediction was done for data of 2015. The suitability analysis shows that this area does not receive enough rainfall for a large part of the year, and therefore has a low suitability (S3). The changing climate could explain the difference between the map of 2015 and the map of 2003. Drought stress has become more problematic and banana is grown less in this area.

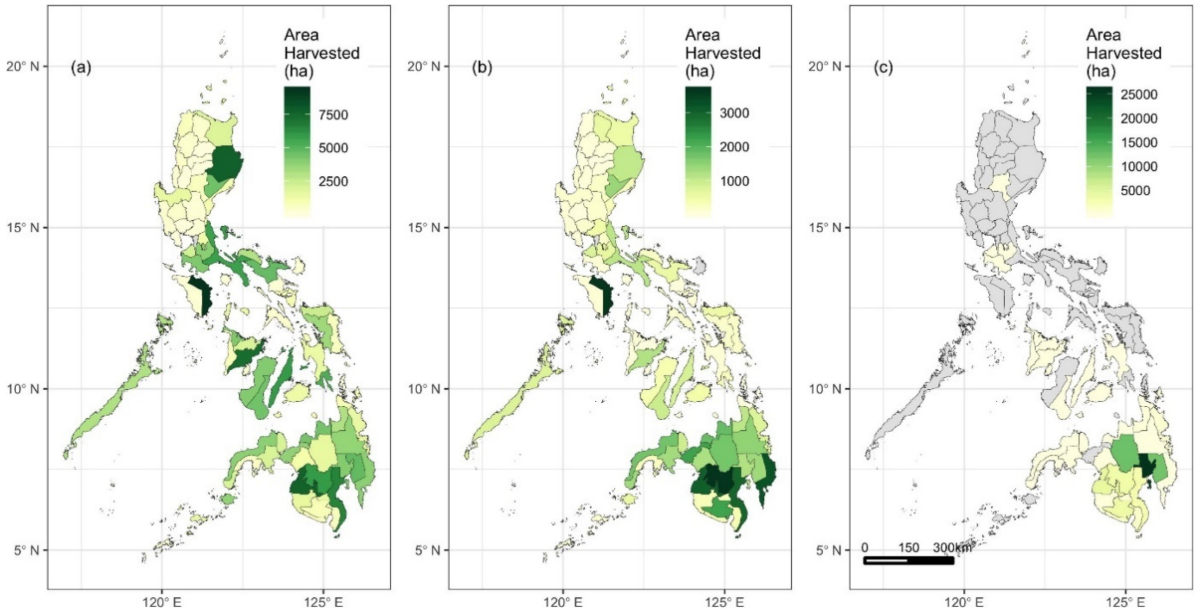


Figure 24. Banana production statistics for three different cultivars in the Philippines; Saba (a), Lakatan (b) and Cavendish (c), grey in the map means "No Data" (source: Salvacion (2019)). Note that the legend scales differ for the three cultivars.

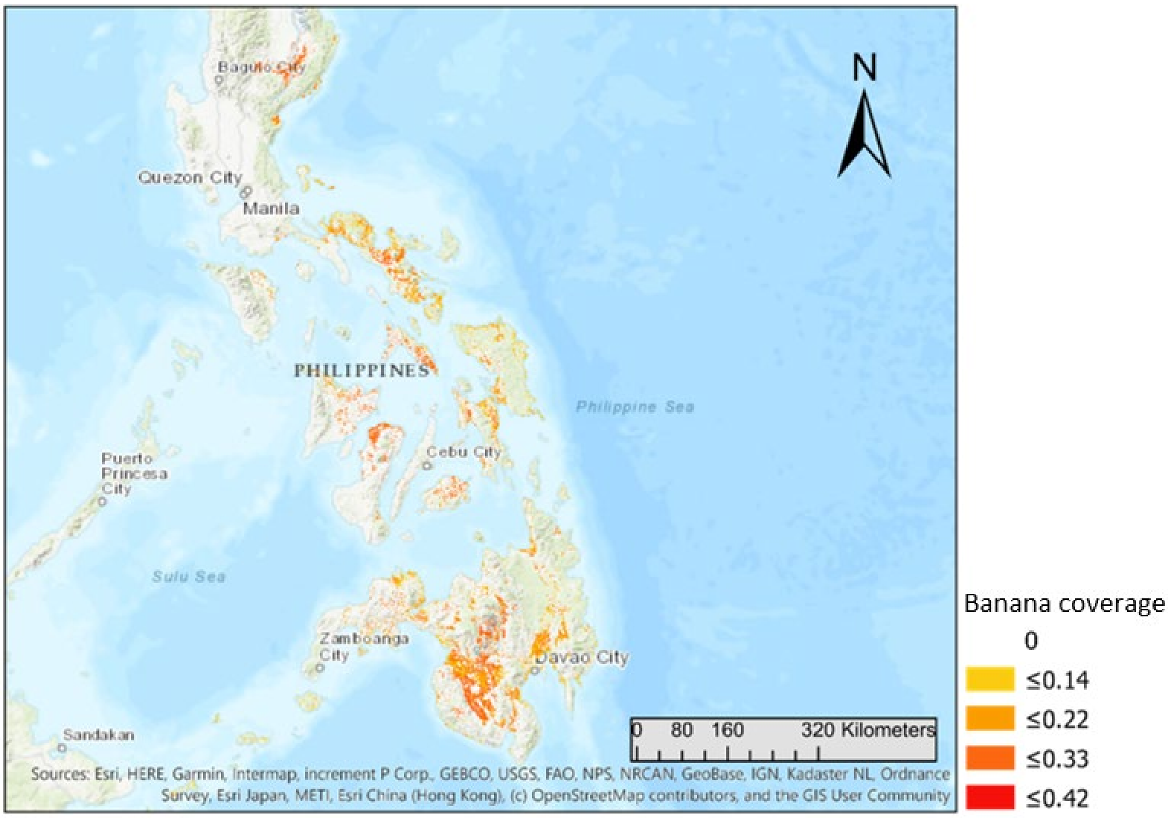


Figure 25. Banana areas in the Philippines according to the prediction method.

The reference map for three different cultivars show that Cavendish, which is the cultivar that is produced for export, is mainly grown in the southern part. The prediction map illustrates this production region well. The other two cultivars are less extensively cultivated, occupying less area. Nonetheless, these areas are detected by the prediction model as well.

The three validation techniques all have their advantages and disadvantages regarding scale and data availability. Yet, together they give a good overview of the validity of the banana map. According to the qualitative validation steps, the prediction model was able to detect the banana areas. The three regions in the local comparison were each from a different yield scale, and thus from a different production system. The correspondence of the prediction map to these reference maps shows that the model can disaggregate the banana production area accurately on a global scale. The quantitative approach showed that on small scale (for Costa Rica) the map was less precise, but that it does represent the banana distribution on country level. Improvements to the model are required to raise the precision of the maps on the smaller scale, if the goal is to identify single cells with banana.

4.2 Crop prediction method

This research shows that national production statistics can be disaggregated to a 30 arc seconds resolution using auxiliary datasets. Next to the high spatial resolution, novelty comes from the combination of a land use analysis with a land suitability evaluation to disaggregate national production statistics of a single crop. Previous studies focussed on disaggregating (sub-)national cropland area solely by land use analysis (Leff et al., 2004, Monfreda et al., 2008). Other research enhanced the prediction method by combining production system information, an irrigation map and a land suitability analysis to the disaggregation procedure (You et al., 2014). Their approach is comparable to the one presented here, integrating a suitability analysis into the prediction model, however, it was conducted at a 5 arc minutes resolution. The model presented here reaches a 30 arc seconds resolution. Furthermore, the model focusses on distributing production area of one single crop. Therefore, land use analysis and land suitability evaluation are specialized for the crop, allowing for more accurate mapping. Portmann et al. (2010) introduced 'cropping periods' to account for different growing seasons, these cropping periods represent the properties (e.g. NDVI, temperature, rainfall, drainage) for multiple crops. Extra uncertainty is introduced with the grouping of crops, compared to the classification for the specific properties in this study. A study in which different global land cover maps were reviewed on their differences, stressed the importance of the classification scheme that is applied in the model. In their research they compared four global landcover maps and found that the major differences sourced from the classification scheme that was applied (Congalton et al., 2014).

The model, specialized at identifying one single crop, can be used to map other crops as well, allowing to create global crop maps for single crops. The banana specific reclassification tables used in this study can be adapted to the characteristics of other crops. For the mapping of coffee for example, slope suitability classes will focus on finding areas with slopes instead of flat areas as is the case for banana. The production statistics can then be distributed according to the new land use analysis and land suitability evaluation.

4.2.1 Land use analysis

The approach in the land use analysis where the intermediate results are displayed, gives a clear view of the steps taken in the analysis. This way, when there is a disagreement with the map, the step where this disagreement originates can be easily traced back and adapted. The percentages of the area used to grow banana go up quickly by adding datasets, because of the shrinking of the potential area, especially for the smaller countries and often islands (Figure 3-5). In Figure 5, where the potential area has been reduced to only the agricultural areas with perennial crops, the map shows that some countries use 100% of their agricultural land to grow banana. This is not possible since there are no countries that only produce banana and no other crops. It indicates that either the FAO overestimates production statistics or that the model cuts out too much potential area, which is more likely. Additionally, the uncertainty from the other input datasets is automatically involved in the model output. Then, there is also the error introduced by the reclassification schemes (Appendix A). The GlobCov dataset for example, has very broad land cover classes (e.g. mosaic cropland >50%, mosaic natural vegetation >50% and different types of tree cover) making reclassification difficult and inevitably errors are introduced. In the reclassification of the NDVI images, the minimum threshold value of 0.4 reduces the potential banana area considerably, resulting in 100 percent banana area for some countries. Setting correct threshold values is difficult, especially in a study on a global scale for a crop with different production systems. Different classification schemes will result in significantly different maps, which will represent different parts of the world better. Improvements can be made by further studying the boundaries for the specific crop and production system, as was done for the land suitability evaluation. Next to specifying better boundaries for the classification schemes, more accurate input data will automatically result in better model results. An example is an improved map of canopy height in which the areas around the tropics are more accurate.

When such a map is available, the dataset can be reclassified to only keep areas with canopies between 2 – 6 m, the height range of banana plants. Another improvement to the land use analysis would be to introduce a dataset of protected nature areas. These areas are not used to grow banana and can therefore be removed.

4.2.2 Land suitability evaluation

The land suitability evaluation carried out in this study uses several properties to define suitability classes. The choice of these properties was based on the properties used in previous studies and the opinion of experts in banana agriculture from Corbana. The survey-based banana mapping approach in the crop-mapper project used temperature, precipitation and the driest quarter to define 'agro-climatic' zones. Although survey based, the unique banana areas were required to fall into the same agro-climatic zone on the map (Crop-mapper, 2008). In a way, this restriction works the same as a suitability analysis. These properties were also used in this study, together with slope, water table depth and soil type. The implementation of a yield scale to account for the different production systems made it possible to create a global suitability map. This approach is not perfect since there are often multiple production systems within one country. However, it gives a good representation of the suitable areas for the main production system, the system in which most of the banana is produced in that country. As discussed in the validation section, extra properties could be added to improve the suitability map. In the example of Costa Rica, the suitability map can be improved by adding sun hours as a property. Likewise, it is very well possible that other countries require different properties to be added to the equation in order to improve the suitability map. A possible downside of adding extra input to the suitability analysis is that the relative importance of the other properties decreases in the combined suitability map. The weight of each property when combining the individual maps needs to be examined thoroughly, this remains a difficult task as it is an expert's decision. Moreover, each input dataset adds more uncertainty to the model. There are possibilities to test each dataset's influence on the result of the model. This will be further evaluated in the recommendations section.

4.2.3 Predicting banana distribution

The banana prediction map shows the areas where banana is expected to be produced. Some countries, like the example of Costa Rica (Figure 11), are overestimated in their banana extent. The production area that is divided over the country is about one third of the potential area that is most suitable. The banana area is distributed equally over all potential, most suitable areas. These areas are often suitable for other crops as well, for example, pineapple also requires high temperatures and a lot of rainfall. As a result, there are no cells with 100 percent banana cover. Single banana plantations, which are sometimes larger than 1 cell on the map, are not identified. For Costa Rica, the extent can be reduced by sharpening the decision rules of the land use analysis and the land suitability evaluation, enhancing the accuracy and increasing the banana coverage of the pixels. This way individual banana plantations will become more apparent for Costa Rica. However, this will further reduce the potential extent for other countries as well. When sharpening the decision rules, the map will better represent Costa Rica, but the accuracy for these countries could decrease as the potential extent shrinks even further. Analysis into these decision rules is needed to find the right values.

The prediction method assumes that banana is produced in the most suitable areas. However, these suitable areas are suitable for other crops as well. Banana is a very profitable crop, and it is therefore likely that banana will be produced in regions where it is suitable. Nonetheless, there are other very profitable crops like pineapple and oil palm grown in these suitable areas making it hard to distinguish between them in a land suitability analysis. The discrimination must then come from the first step; the land use analysis. Pineapple could for example be excluded from potential banana area with the canopy height map by discarding areas with a canopy height below 2 m. When a more accurate map of global canopy height becomes available, such restrictions can further improve the model. Furthermore, there are socio-economic reasons, like risk minimalization, why a farmer might decide not to put the most profitable crop on his field (You and Wood, 2005). Therefore, the crop distribution method presented in this study is best applied to map high revenue crops, regarding the currently available data. When more detailed datasets become available, for example, a high-resolution canopy height map or a high-resolution NDVI map, better distinctions between crops can be made. Small patches of the crops can then be identified.

Remote sensing technology has accelerated rapidly in recent decades and will probably continue to do so. If future technological advancements allow for global satellite imagery on which individual leaves/crops can be distinguished, global crop mapping will become much easier. The identification of crops will be based on a single input image from which all crops can be determined directly rather than from a set of inputs. This would increase the accuracy immensely, both directly because of the higher resolution, and indirectly because there is just one source of uncertainty. However, such high-resolution satellite imagery is not available yet.

4.3 Applicability banana map

The banana map (Figure 10) shows the general distribution of banana areas globally. This is very useful for further research aiming to minimize the impact of Fusarium wilt. Knowing where the main banana areas are, indicates where the risk for loss of production is highest and thus extra precautions need to be taken. Additionally, the map can be used to identify areas that are at risk due to the impacts of climate change. Areas with more frequent droughts will not be suitable for growing banana on the long term. Especially for the low input production systems, which cannot rely on management practices to compensate for the changing conditions, the map helps in identifying areas that are prone to suffer in the case of drastic climate change. For other areas, that were previously not suitable to grow banana, the effects of climate change introduce better circumstances for the crop to grow as the average annual temperature rises or rainfall rates increase.

The spreading pathways of Fusarium wilt, however, are best identified by only looking at the suitability map (Figure 8, Figure 9). The suitability map shows the areas where banana can grow optimally. Even when banana is not grown for production in many of these areas, it is very likely that single plants or small plots of banana are scattered in these areas. Furthermore, smallholder farms that do not produce this year due to socio-economic reasons, might very well have banana the year after. These banana areas are not found on the banana map, but the suitability map shows that it can be grown here. The unsuitable areas form natural barriers for the fungus to disperse, while the suitable areas are quickly colonized through single plants.

5. CONCLUSIONS

A method was developed to disaggregate national crop production statistics. The method, in which the results of a land use analysis and a land suitability evaluation are combined, can be applied to many crops by implementing crop-specific thresholds on the input data. Applying the method on the production statistics of banana proved to be successful to create a global banana map. The land use analysis showed the areas where banana is possibly observed according to criteria related to agriculture, NDVI and the height of the canopy. The land suitability analysis showed the most suitable areas for banana. Banana areas were then predicted by combining the results of the two analyses. For the first time, a global banana map is presented on a 30 arc seconds resolution. Comparison to previous banana distribution studies showed that the map represents the banana distribution well on global and local administrative levels. Additionally, the map provides more detail compared to previous banana distribution research working on a 5 arc minutes resolution. The banana map depicts the vulnerable areas, while the suitability map provides information about the pathways for the spread of fusarium wilt TR4. The map can be improved further when new high accuracy datasets on a high-resolution become available. The disaggregation procedure itself is very suitable for application on other crops. This way crop maps of other single crops can be created.

6. RECOMMENDATIONS

Additional research is needed to further improve the land use analysis. Putting a very strict threshold on one of the datasets in the land use analysis causes the potential banana area to shrink immensely. Changing the threshold to be more lenient will result in less shrinking of the potential banana area. Research can be done on the relative influence of different threshold on the shrinking of the potential extent. It is still difficult to say whether the chosen threshold values represent the real situation optimally, as there is much uncertainty in literature for this relatively new field of science.

Furthermore, the reclassification tables in the land suitability evaluation need further examination. Similar to the land use analysis, the properties' thresholds influence the suitable extent. Studying the relative influence of the threshold gives insight in the importance of the thresholds. Next to studying the thresholds, extra properties could be analysed to improve the accuracy of the suitability analysis. However, the influence of each dataset on the results is also important; adding datasets inherently adds more uncertainty. Performing a sensitivity analysis to determine the impact of each dataset on the result will help in identifying which datasets are important and which are not. The datasets that do not affect the results much are better left out of the equation as they only add extra uncertainty.

7. WORKS CITED

BAGAMBA, F. 2007. *Market access and agricultural production: the case of banana production in Uganda*, Wageningen University.

BONTEMPS, S., DEFOURNY, P., VAN BOGAERT, E., ARINO, O. & KALOGIROU, V. 2010. GlobCover 2009: product description manual, version 1.0. *ESA and UCLouvain) Available at http://dup. esrin. esa. it/files/p68/GLOBCOVER2009_Product_Description_Manual_1. 0. pdf [Verified July 2012].*

CONGALTON, R. G., GU, J., YADAV, K., THENKABAIL, P. & OZDOGAN, M. 2014. Global land cover mapping: A review and uncertainty analysis. *Remote Sensing,* 6**,** 12070-12093.

CROP-MAPPER. 2008. *Spatial information on global banana production systems* [Online]. Available: http://www.crop-mapper.org/banana/ [Accessed].

DANIELSON, J. J. & GESCH, D. B. 2011. Global multi-resolution terrain elevation data 2010 (GMTED2010). US Geological Survey.

ELBEHRI, A., CALBERTO, G., STAVER, C., HOSPIDO, A. & SKULLY, D. 2016. Ecuador's banana sector under climate change: an economic and biophysical assessment to promote a sustainable and climate-compatible strategy.

ESA 2017. Land Cover CCI Product User Guide Version 2. Tech. Rep. (2017). Available at: maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf.

ESRI 2019. ArcGIS Dekstop Version 2.3.2.

FAN, Y., LI, H. & MIGUEZ-MACHO, G. 2013. Global patterns of groundwater table depth. *Science,* 339**,** 940-943.

FAO-STAT 2019. Crop statistics http://www.fao.org/faostat/en/#data/QC.

FAO 1976. A framework for land evaluation. *FAO Soils bulletin 32*.

FAO. 2019. *Banana facts and figures* [Online]. Available: http://www.fao.org/faostat/en/#data/QC [Accessed 12/5/2019 2019].

FICK, S. E. & HIJMANS, R. J. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International journal of climatology,* 37**,** 4302-4315.

FUSARIUMWILT.ORG 2019. What is Panama Disease? Available: https://fusariumwilt.org/index.php/en/the-problem/#what-is-panama-disease.

GARCIA-BASTIDAS, F., QUINTERO-VARGAS, C., AYALA-VASQUEZ, M., SEIDL, M., SCHERMER, T., SANTOS-PAIVA, M., NOGUERA, A. M., AGUILERA-GALVEZ, C., WITTENBERG, A. & SØRENSEN, A. 2019. First report of Fusarium wilt Tropical Race 4 in Cavendish bananas caused by Fusarium odoratissimum in Colombia. *Plant Disease*.

GARCÍA-BASTIDAS, F. A. 2019. *Panama disease in banana: Spread, screens and genes.* Wageningen University.

GHAVAMI, M. 1976. Banana plant response to water table levels. *Transactions of the ASAE,* 19**,** 675-0677.

GODFRAY, H. C. J., BEDDINGTON, J. R., CRUTE, I. R., HADDAD, L., LAWRENCE, D., MUIR, J. F., PRETTY, J., ROBINSON, S., THOMAS, S. M. & TOULMIN, C. 2010. Food security: the challenge of feeding 9 billion people. *science,* 327**,** 812-818.

GOWEN, S. 2012. *Bananas and plantains*, Springer Science & Business Media.

GUZMÁN-ALVAREZ, J. A. & QUESADA, M. G. 2014. CORBANA-BANACLIMA A weather monitoring system to support the banana production in Costa Rica. *CIGR Proceedings,* 1.

IRIZARRY, H., SILVA, S. & VICENTE-CHANDLER, J. 1980. Effect of water table level on yield and root system of plantains. *The Journal of Agriculture of the University of Puerto Rico,* 64, 33-36.

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.

KARAMURA, E., FRISON, E., KARAMURA, D. & SHARROCK, S. 1998. Banana production systems in eastern and southern Africa. *Bananas and food security. INIBAP, Montpellier*, 401-412.

LAVRENIUK, M., SKAKUN, S. V., SHELESTOV, A. J., YALIMOV, B. Y., YANCHEVSKII, S., YASCHUK, D. J. & KOSTECKIY, A. 2016. Large-scale classification of land cover using retrospective satellite data. *Cybernetics and Systems Analysis,* 52, 127-138.

LEFF, B., RAMANKUTTY, N. & FOLEY, J. A. 2004. Geographic distribution of major crops across the world. *Global Biogeochemical Cycles,* 18.

MACHOVINA, B. L., FEELEY, K. J. & MACHOVINA, B. J. 2017. UAV remote sensing of spatial variation in banana production. *Crop and Pasture Science,* 67, 1281-1287.

MITCHARD, E. T., SAATCHI, S. S., WHITE, L., ABERNETHY, K., JEFFERY, K. J., LEWIS, S. L., COLLINS, M., LEFSKY, M. A., LEAL, M. E. & WOODHOUSE, I. H. 2012. Mapping tropical forest biomass with radar and spaceborne LiDAR in Lopé National Park, Gabon: overcoming problems of high biomass and persistent cloud. *Biogeosciences,* 9, 179-191.

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.

MUNDIAL, B. 2008. World Development Report. Agriculture for Development. *Washington, DC: World Bank*.

PLOETZ, R. C. 1994. Panama disease: return of the first banana menace. *International journal of pest management,* 40, 326-336.

PORTMANN, F. T., SIEBERT, S. & DÖLL, P. 2010. MIRCA2000—Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. *Global biogeochemical cycles,* 24.

PURSEGLOVE, J. 1988. *Taming the flood: a history and natural history of rivers and wetlands,* Oxford University Press Oxford.

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.

RITUNG, S., AGUS, F. & HIDAYAT, H. 2007. Land suitability evaluation with a case map of Aceh Barat District.

ROBINSON, J. & BOWER, J. 1987. Transpiration characteristics of banana leaves (cultivar 'Williams') in response to progressive depletion of available soil moisture. *Scientia Horticulturae,* 30, 289-300.

RTB.CGIAR.ORG 2013. New banana disease found in Mozambique. Available from: https://www.rtb.cgiar.org/news/new-banana-disease-to-africa-found-in-mozambique/. *Joint statement issued by the Mozambique Department of Agriculture, Matanuska, IITA, Stellenbosch University and Bioversity International.*

SALVACION, A. R. 2019. Effect of climate on provincial-level banana yield in the Philippines. *Information Processing in Agriculture*.

SASTRY, P. 1988. Agrometeorology of the banana crop. *CAgM. Report/WMO*.

SIMARD, M., PINTO, N., FISHER, J. B. & BACCINI, A. 2011. Mapping forest canopy height globally with spaceborne lidar. *Journal of Geophysical Research: Biogeosciences,* 116.

SOTO BALLESTERO, M. 2014. Bananos: conceptos básicos. Tecnológica de CR.

STOORVOGEL, J. J., BAKKENES, M., TEMME, A. J., BATJES, N. H. & TEN BRINK, B. J. 2017. S-world: a global soil map for environmental modelling. *Land degradation & development,* 28, 22-33.

STORY, M. & CONGALTON, R. G. 1986. Accuracy assessment: a user's perspective. *Photogrammetric Engineering and remote sensing,* 52, 397-399.

STOVER, R. H. & SIMMONDS, N. W. 1987. *Bananas*, Longman Scientific and Technical.

WEBB, P., STORDALEN, G. A., SINGH, S., WIJESINHA-BETTONI, R., SHETTY, P. & LARTEY, A. 2018. Hunger and malnutrition in the 21st century. *bmj,* 361, k2238.

YOU, L. & WOOD, S. 2005. Assessing the spatial distribution of crop areas using a cross-entropy method. *International Journal of Applied Earth Observation and Geoinformation,* 7, 310-323.

YOU, L., WOOD, S., WOOD-SICHRA, U. & WU, W. 2014. Generating global crop distribution maps: From census to grid. *Agricultural Systems,* 127, 53-60.

8. APPENDIX

A. Decision rules for reclassification of land use input data

Landcover reclassification scheme of the classes used in the land cover analysis:

GlobCov class	Reclassification potential banana probability value (0-100%)
Cropland, rainfed	100
Herbaceous cover	100
Tree or shrub cover	100
Cropland, irrigated or post-flooding	100
Mosaic cropland (>50%) / natural vegetation (tree, shrub, herbaceous cover) (<50%)	80
Mosaic natural vegetation (tree, shrub, herbaceous cover) (>50%) / cropland (<50%)	40
Tree cover, broadleaved, evergreen, closed to open (>15%)	30
Tree cover, broadleaved, deciduous, closed to open (>15%)	30
Tree cover, broadleaved, deciduous, closed (>40%)	30
Tree cover, broadleaved, deciduous, open (15-40%)	30

NDVI reclassification scheme:

NDVI value	Reclassification value
0.0-0.4	0
0.4-1.0	1

Canopy height reclassification scheme:

Height of canopy	Reclassification value
0-25	1
25-73	0

B. Reclassification of soil types

Soil type reclassification scheme:

Soil ype	Yield scale 1 (low yield, LIPS)	Yield scale 2 (medium yield)	Yield scale 3 (high yield, HIPS)
Fluvisols	3	2	1
Eutric Fluvisols	3	2	1
Calcaric Fluvisols	3	2	1

Dystric Fluvisols	3	2	1
Mollic Fluvisols	3	2	1
Umbric Fluvisols	3	2	1
Thionic Fluvisols	3	2	1
Salic Fluvisols	3	2	1
Gleysols	3	2	2
Eutric Gleysols	3	2	2
Calcic Gleysols	3	2	2
Dystric Gleysols	3	2	2
Andic Gleysols	3	2	2
Mollic Gleysols	3	2	2
Umbric Gleysols	3	2	2
Thionic Gleysols	3	2	2
Gelic Gleysols	4	4	4
Acrisols	4	3	2
Haplic Acrisols	4	3	3
Ferric Acrisols	4	3	2
Humic Acrisols	4	3	2
Plinthic Acrisols	4	3	2
Gleyic Acrisols	4	3	2
Alisols	4	3	3
Haplic Alisols	4	3	3
Ferric Alisols	4	3	3
Humic Alisols	3	3	3
Plinthic Alisols	4	3	3
Stagnic Alisols	4	3	3
Gleyic Alisols	4	4	4
Andosols	1	1	1
Haplic Andosols	1	1	1
Mollic Andosols	1	1	1
Umbric Andosols	1	1	1
Vitric Andosols	2	1	1
Gleyic Andosols	1	1	1
Gelic Andosols	4	4	4
Arenosols	3	2	2
Haplic Arenosols	3	2	2
Cambic Arenosols	3	2	2
Luvic Arenosols	3	2	2
Ferralic Arenosols	3	2	2

Albic Arenosols	3	2	2
Calcaric Arenosols	3	3	2
Gleyic Arenosols	3	2	2
Anthrosols	3	3	3
Aric Anthrosols	3	3	3
Cumulic Anthrosols	3	3	3
Fimic Anthrosols	3	3	3
Urbic Anthrosols	4	4	4
Chernozems	2	2	1
Haplic Chernozems	1	1	1
Calcic Chernozems	2	2	2
Luvic Chernozems	2	1	2
Glossic Chernozems	1	1	1
Gleyic Chernozems	1	3	1
Calcisols	4	3	2
Haplic Calcisols	4	3	2
Luvic Calcisols	4	3	2
Petric Calcisols	4	3	2
Cambisols	2	2	2
Eutric Cambisols	2	2	2
Dystric Cambisols	2	2	2
Humic Cambisols	2	2	2
Calcaric Cambisols	3	3	2
Chromic Cambisols	2	2	2
Vertic Cambisols	4	4	4
Ferralic Cambisols	3	2	2
Gleyic Cambisols	3	3	2
Gelic Cambisols	4	4	4
Ferralsols	3	3	3
Haplic Ferralsols	3	3	3
Xanthic Ferralsols	3	3	3
Rhodic Ferralsols	3	3	3
Humic Ferralsols	3	3	3
Geric Ferralsols	3	3	3
Plinthic Ferralsols	3	3	3
Greyzems	2	2	1
Haplic Greyzems	2	2	1

Gleyic Greyzems	2	2	2
Gypsisols	4	3	3
Haplic Gypsisols	4	3	3
Calcic Gypsisols	4	3	3
Luvic Gypsisols	4	3	3
Petric Gypsisols	4	3	3
Histosols	3	3	2
Folic Histosols	3	2	2
Terric Histosols	3	3	2
Fibric Histosols	3	3	2
Thionic Histosols	4	4	3
Gelic Histosols	4	4	4
Kastanozems	1	1	1
Haplic Kastanozems	1	1	1
Luvic Kastanozems	1	1	1
Calcic Kastanozems	2	2	2
Gypsic Kastanozems	3	3	2
Leptosols	4	4	4
Eutric Leptosols	4	4	4
Dystric Leptosols	4	4	4
Rendzic Leptosols	4	4	4
Mollic Leptosols	4	4	4
Umbric Leptosols	4	4	4
Lithic Leptosols	4	4	4
Gelic Lepsols	4	4	4
Luvisols	2	2	2
Haplic Luvisols	2	2	2
Ferric Luvisols	3	3	2
Chromic Luvisols	3	3	2
Calcic Luvisols	3	3	2
Vertic Luvisols	3	3	3
Albic Luvisols	2	2	2
Stagnic Luvisols	3	3	3
Gleyic Luvisols	3	2	3
Lixisols	3	2	2
Haplic Lixisols	3	2	2
Ferric Lixisols	3	3	3

Plinthic Lixisols	3	2	2
Albic Lixisols	3	2	2
Stagnic Lixisols	4	3	2
Gleyic Lixisols	3	3	1
Nitisols	1	1	1
Haplic Nitisols	1	1	1
Rhodic Nitisols	1	1	1
Humic Nitisols	1	1	1
Podzoluvisols	3	3	3
Eutric Podzoluvisols	3	3	2
Dystric Podzoluvisols	3	3	3
Stagnic Podzoluvisols	4	3	2
Gleyic Podzoluvisols	3	2	2
Gelic Podsoluvisols	4	4	4
Phaeozems	2	2	1
Haplic Phaeozems	2	2	1
Calcaric Phaeozems	4	3	2
Luvic Phaeozems	3	2	2
Stagnic Phaeozems	4	3	2
Gleyic Phaeozems	3	3	3
Planosols	4	2	1
Eutric Planosols	4	2	1
Dystric Planosols	4	2	1
Mollic Planosols	4	2	1
Umbric Planosols	4	2	1
Gelic Planosols	4	4	4
Plinthosols	4	3	3
Eutric Plinthosols	4	3	3
Dystric Plinthosols	3	3	3
Humic Plinthosols	3	3	3
Albic Plinthosols	3	3	3
Podzols	4	4	2
Haplic Podzols	4	4	2
Cambic Podzols	4	4	2
Ferric Podzols	4	4	4
Carbic Podzols	4	4	2

Gleyic Podzols	4	4	4
Gelic Podzols	4	4	4
Regosols	4	4	4
Eutric Regosols	4	4	3
Calcaric Regosols	4	4	3
Gypsic Regosols	4	4	3
Dystric Regosols	4	4	3
Umbric Regosols	4	4	3
Gelic Regosols	4	4	3
Solonchaks	4	4	3
Haplic Solonchaks	4	4	3
Mollic Solonchaks	4	4	3
Gleyic Solonchaks	4	4	3
Calcic Solonchaks	4	4	4
Gypsic Solonchaks	4	4	4
Sodic Solonchaks	4	4	4
Gelic Solochaks	4	4	4
Solonetz	4	4	4
Haplic Solonetz	4	4	4
Mollic Solonetz	4	4	4
Calcic Solonetz	4	4	4
Gypsic Solonetz	4	4	4
Stagnic Solonetz	4	4	4
Gleyic Solonetz	4	4	4
Vertisols	4	4	2
Eutric Vertisols	4	3	2
Dystric Vertisols	4	4	3
Calcic Vertisols	4	4	3
Gypsic Vertisols	4	4	3
Dunes & Shift.sands	4	4	4
Rock outcrops	4	4	4
Water Bodies	4	4	4
Glaciers	4	4	4

C. Farm visits

The visits to the farms were meant to get a better understanding of the cultivation methods of banana. While visiting the farms, attention was paid to the following factors relating to the research:

- Size of the plantation: Recognition possibility on a 1km grid scale
- Drainage system: Depth of the drainage channels and water table depth of the plantation to classify global water table depth dataset.
- Canopy height: To discriminate banana from other crops and forests with the canopy height dataset.
- Slope of the plantation: To classify suitability of slopes from a Digital Elevation Model.

Next to these factors, other factors that seemed important were noted to possibly implement in the disaggregation method. A summary of each of the field visits is described here. Rainfall data on farms were obtained through José Antonio Guzmán-Alvarez, who is working for the climate network of Corbana (BanaClima). (https://www.corbana.co.cr/banaclima-2/)

Sahara farm (Lat: 10.164518, Lon: -83.334586)

The Sahara farm is a commercial plantation situated in the east of Costa Rica, near Puerto Limón. The annual average rainfall is around 3500mm. Although this is not the highest rainfall rate in the area, problems with drainage occur, making banana cultivation difficult in some parts of the farm. Pumps have been installed to get rid of the excess water, next to the extensive system of ditches. Three orders of ditches are present on the farm, the largest being 3-4m deep and 5m wide. Farm size is 120ha and the difference in height from one side to the other is not more than 2m, making the slope almost 0 degrees. The height of the banana canopy is 4-5m

Las Valquirias farm (Lat: 10.424284, Lon: -83.649735)

Situated in the north-eastern part of Costa Rica, Las Valquirias is a commercial plantation with high average annual rainfall, up to 5000 mm. Unlike the Sahara farm, there are less problems with drainage due to a more sandy soil texture. Extensive drainage channels are also present to aerate the soil. Drainage is organised with up to 3rd order channels, with channels depths up to 5-6 m. Farm size is 300ha, canopy height 4-5 m and a slope in the terrain is absent.

28 Millas (Lat: 10.093972, Lon: -83.376944)

A very small (2 ha), old banana farm which is now in use for research by Corbana. It is located between Siquirres and Matina, in the east of Costa Rica. Rainfall is comparable to that of Sahara farm, around 3500mm per year. Drainage is organised with primary and secondary channels, with a maximum depth of 2-2.5 m. No significant slope is present (0-2 m difference). At the moment a research on Fusarium TR1 is being established in which they try to harvest 3 generations of Gros Michel on the infected soils. The idea is to 'clean' the soil of the spores by putting plastic over it for 3-4 months and afterwards plant the banana and move the plastic to the plot next to it. When the banana has been successful for 3 generations the plant can be removed again and there will be new plastic applied to the soil.

Turrialba (Lat: 9.921562, Lon: -83.591661)

Turrialba is situated in the inlands of Costa Rica. The area is mountainous and there are no banana plantations like the ones in Limón province. The banana that is cultivated here is not meant for export and the farms are small, around 1-5 ha. The variety grown here is Gros Michel, heavily impacted by Fusarium TR1, they have to swift their banana farms occasionally. Most noteworthy however, is the fact that the slopes in this area are very steep. Height differences within one farm are large; slopes up to 45 degrees are common. I heard a story where a farmer has lost two horses already because they fell down the steep slope his banana plantation was on, while carrying the bunches he had just harvested.

San Pablo (Lat: 10.109814, Lon: -83.380104)

San Pablo is a commercial farm owned by Corbana, it is situated close to the research farm 28 Millas. Rainfall rates are similar, an average of 3500 mm per year. First order ditches are present with a maximum depth of 4 m. Secondary ditches reach up to 2-2.5 m depth. Farm size is 280 ha with a slope of almost 0 degrees. The height of the banana canopy is 4-5 m.

Consults with experts

As part of the validation a presentation was held at Corbana to a group of about 20 people to provoke conversation on the maps. The people were all working in different fields of banana science. In general, the reaction on the results was positive, indicating that the maps represent reality well. However, afterwards in 1 on 1 conversations it became clear that there were some discrepancies. To get a more clear idea of where these discrepancies come from, 1 on 1 conversations were held with different experts. Below are the summaries of these interviews.

José Antonio Guzmán Alvarez (Corbana, BanaClima network)

Jose knows a lot about the climatic conditions of banana agriculture in Costa Rica, therefore we were able to discuss the ranges of the rainfall suitability analysis. Optimal rainfall ranges for banana are between 1200mm and 4000mm per year, for production without irrigation or drainage. This corresponded to the ranges that were implemented in the model before quite well and since various sources recommend different ranges, this was not adapted further in the model. Moreover, he showed me the importance of the driest quarter in banana agriculture. Implementing a more strict classification for the driest quarter in the methods resulted in an improved map as areas in the Guanacaste peninsula (west Costa Rica) were correctly removed. Furthermore, he recommended me to look at the Icafe map (source Icafe) in which areas of coffee production for Costa Rica are shown. These areas have banana as well, because the banana is used as shadowing for the coffee plants. Later it became clear that this had to be achieved by removing the slope factor from the suitability equation.

David Brown (Bioversity International, GIS specialist) (https://www.bioversityinternational.org/)

David Brown from Bioversity international has worked on crop-mapper, an initiative to map banana productions globally from the international conference on banana and plantain in Africa, Kenya in 2008. (source: http://www.crop-mapper.org/banana/#history). This project, in which areas would be drawn manually by experts, is not updated anymore and he was interested in the approach of this study. Additions he pointed out for the map of Costa Rica were the smallholder farms in the highlands where the bananas are traded locally. This is an important factor for the spreading of Fusarium TR4. Following his advice two different maps have been created, one in which the slope is not taken into account as a factor for banana agriculture. This map shows not only large scale plantations but also areas with low density banana agriculture. Furthermore, he mentioned the importance of the rapid change in these smallholder production systems; one year a farmer might decide to plant banana while the next year other crops seem more profitable. Therefore, it is important to look at suitability maps to find areas where banana can be produced, if not this year, then another year. Lastly, he referred to his own (thesis? Brown, D., Jarvis, A., Ramirez, J. and Staver, C., Mapping Banana and Plantain Production Zones in Latin America and the Caribbean, Poster at ISHS-ProMusa Symposium - Bananas and plantains: Toward sustainable global production and improved uses Salvador, Bahia, Brazil. October, 2011) work on banana mapping where he came to the conclusion that the exchange of plant material by farmers, and the spreading via rivers were the most important factors in the spreading of Fusarium TR1.

RODDY ORTEGA BONILLA (CORBANA, SOIL EXPERT)

Roddy indicated that some areas that did show banana on the map, in the northern part of Costa Rica, do not have banana production. This is caused by the low amount of sun hours in the area due to clouds. Additionally, the high clay content forms a problem in these soils. He stressed the importance of soil texture and drainage with respect fertility. Poor drainage and a clayey soil texture are more problematic than a low soil fertility for the cultivation of banana. Therefore, the suitability of soil types with a clayey texture and poor drainage got lower suitability classes, while the suitability of the soil types with a low fertility was slightly raised. Furthermore, he mentioned the risk of flooding as being an important factor for suitability in low lying areas. The problem with this factor however, is that even though there is a risk of flooding every 2 to 3 years, people will still plant bananas. Thus flood risk is not implemented into the suitability analysis, as there will be banana in these areas nonetheless.