Soccer Fields as Rainfall Detectors using Machine Learning

The case of Ghana

D.T. Touloumidis

TU Delft

Soccer Fields as Rainfall Detectors using Machine Learning

The case of Ghana

by

D.T. Touloumidis

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Monday, December 6, 2021 at 16:30 AM.

Student number:	5161975	
Project duration:	February 1, 2021 – December 6, 2021	
Thesis committee:	Prof. dr. ir. N. van de Giesen,	TU Delft, Supervisor
	Prof. dr. ir. M. C. ten Veldhuis,	TU Delft
	Prof. dr. ir. S. Steele-Dunne,	TU Delft
	MSc. M. E. Camarena,	TU Delft

This thesis is confidential and cannot be made public until December 6, 2022.

Preface

The graduation thesis research presented in that report incorporates the last piece of work to obtain my Master of Science degree in Water Management, which is hosted by the Faculty of Civil Engineering and Geosciences at TU Delft. This work was carried out in the context of the Horizon 2020 project *"Transforming Water, weather, and climate information through In situ observations for Geo-services in Africa (TWIGA)"*. The contribution of the non-profit organization Trans-African Hydro-Meteorological Observatory (TAHMO) was essential for this study that took place between February 2021 and November 2021. During these months, I put much effort and passed through multiple challenges that appeared. Undeniably, it was a thrilling experience that helped me evolve both professionally and personally and proved me that I love doing research.

First of all, I would like to thank the members of my graduation committee Marie-Claire ten Veldhuis, Susan Steele-Dunne and Monica Camarena. You always provided me with constructive comments and valuable recommendations during our meetings.

Furthermore, I would like to thank Rebecca Hochreutener, Rick Hagenaars and the whole team of TAHMO organization that were always willing and helpful proving me data.

Nick van de Giesen, my daily supervisor, thank you for our long discussions and for sharing with me your knowledge and your passion that gave me motivation to keep working. You were always finding the time for questions and you were open to consider and discuss my suggestions.

Finally, I could not do it without the endless support of my family and my friends who never left by my side. And of course Despoina who gave me a warm shelter during the COVID-19 pandemic era. Thank you all for being besides me throughout this very special part of my life!

D.T. Touloumidis
Delft, December 2021

Abstract

Agriculture is an important source of income for many countries in the Global South, where it may account for as much as 25% of GDP. Precipitation is crucial for agriculture in countries like Ghana, where ∼95% of farming is rainfed. Accurate rainfall observations are limited in Ghana. The sparse rain gauge network and the lack of weather radars make remote sensing methods a potentially attractive alternative source of rainfall data. Radar satellites, such as Sentinel-1, emit radiation that passes through the atmosphere and is scattered back to the satellite by the Earth's surface. The backscatter measured by the satellite is correlated with the wetness of the soil but the existence of vegetation hinders straightforward quantification of soil moisture. By choosing sites with a simple and, more or less, constant phenology, it may be possible to eliminate the effect of vegetation on backscatter. Soccer field may qualify as sites with such a simple and constant phenology. The main objective of this study is to use the Sentinel-1 data over soccer fields and assess them as rainfall detectors. A machine learning approach will be used to reach this objective.

This research assessed the stability and the generalization capabilities of a classification model (rain/no rain). The model was trained with and applied to different locations and periods (2019 & 2020). Ground observations from 53 Ghanaian (TAHMO) and 1 Greek stations were used. Soccer fields in Ghana and Greece were selected and their suitability as rainfall detectors was checked based on the correlation between modeled soil moisture and backscatter strength.

The rain/no rain classification of the soccer fields was made with a stacked classifier that was trained and validated with both spaceborne and ground data. The classifier was tested on six different datasets from Greece and Ghana 2019 and 2020. The stability of the model was assessed by a Leave-p out cross-validation approach. The generalization in space was tested by using different environments. The generalization in time was tested by using different time periods. The results showed that the classification was stable. The minimum and maximum performances for the different testing datasets were 0.43 to 0.85. The median performance of the algorithm in Ghana for 2020 is 67%. The stacked classifier was found to have the best performance compared to other classifiers. Finally, the performance of the stacked classifier was competitive in comparison with the performance of the well-known IMERG algorithm.

The study showed that there is a potential for using radar backscatter from suitable fields to detect rainfall. The classifier is stable and can be generalized in time and space under certain conditions.

Contents

List of Figures

List of Tables

List of Abbreviations

AI	Artificial Intelligence
AMSR-E	Advanced Microwave Scanning Radiometer for EOS
ANN	Artificial Neural Network
ASCAT	The Advanced Scatterometer
AUTH	Aristotle University of Thessaloniki
CART	Classification And Regression Trees
CCN	Cloud Condensation Nuclei
CTT	Cloud Top Temperature
DSD	Drop Size Distribution
ESA	European Space Agency
EW	Extra Wide Swath
GDP	Gross Domestic Product
GEE	Google Earth Engine
GPCC	Global Precipitation Climatology Centre
GPCP	Global Precipitation Climatology Project
GPM	Global Precipitation Measurement
GRD	Ground Range Detected
H	Horizontal polarization
IW	Interferometric Wide Swath
KNN	K-Neighbors Classifier
LULC	Land Use/Land Cover
ML	Machine Learning
MLP	Multi-Layer Perceptron
MSWEP	Multi-Source Weighted-Ensemble Precipitation
NASA	National Aeronautics and Space Administration
OCN	Ocean
OSM	Open Street Maps
PMW	Passive Microwave
RF	Random Forest
RS	Remote Sensing
SAR	Synthetic Aperture Radar
SL	Supervised Learning
SLC	Single Look Complex
SM	Strip Map
SM2Rain	SM2Rain
SMOS	Soil Moisture and Ocean Salinity
SRTM	Shuttle Radar Topography Mission
SVM	Support Vector Machine
TAHMO	Trans-African Hydro-Meteorological Observatory
TRMM	Tropical Rainfall Measuring Mission
V	Vertical polarization
WAM	West African Monsoon
WV	Wave

1 Introduction

Recent information from the World Bank (2020) claims that the global economy is closely connected to the agriculture production which stands for about 4% of the global GDP. This percentage is up to 25% for many developing countries (e.g., Somalia, Kenya, Nigeria, Uzbekistan), many of them are located in Africa (World Bank, 2020). Specifically in Ghana, 69 per cent of the land is used for agriculture which accounts for the 17.3 per cent of the country's GDP (2019) (World Bank, 2020). When it comes to the climate, West African climate highly depends on the monsoons (WAM) that are occurring from June to October and account for more than seventy per cent (70%) of the annual precipitation (Sultan and Gaetani, 2016). Agriculture in Ghana is closely connected to weather since more than 95% of the cultivated area is rainfed (Le Coz and van de Giesen, 2019). The spatial and temporal variability and the intensity of these extreme precipitation phenomena have a significant impact on the social and economic development of the country, a fact that makes rainfall vital for the well-being and the development of the society (Maranan et al., 2018).

1.1. Literature review

1.1.1. Weather data availability

Precipitation is a very crucial factor for the growth of countries with high rainfed agriculture-based economies. Three prevalent methods to estimate rainfall are i) in-situ observations, ii) weather radar estimates and iii) remote sensing imagery. Apart from the direct methods, rainfall can also be estimated by numerical weather models.

Rain gauge instruments are used for the in-situ observation of precipitation, the most common type is the tipping bucket. Undeniably, the ground observations achieve the most accurate estimates. The limitation of that method is the distribution of ground stations in the world. According to Kidd et al. (2017), it is estimated that the Earth is covered with ∼100,000 not equally distributed stations. The observations of each station correspond to a maximum radius of 5-km and therefore it is determined that only ∼1% of the total Earth is covered. The distribution is dense over Europe and America but sparse over S. America, Africa and Australia (Figure 1.1). The reason behind this unequal distribution is probably the high installation and maintenance cost of the stations. It should be mentioned that rainfall observations using rain gauges are prone to error in areas where convective rainfall occurs. This type of precipitation is short, typically less than 60 minutes, but intense with a spatial range of 5 kilometers (Cristiano et al., 2017). West Africa is an area with low coverage of stations and frequent occurrence of convective rainfall events with high spatial variability (Maidment et al., 2017). Thus, different sources of data are used to enhance the performance of rainfall estimations over Africa.

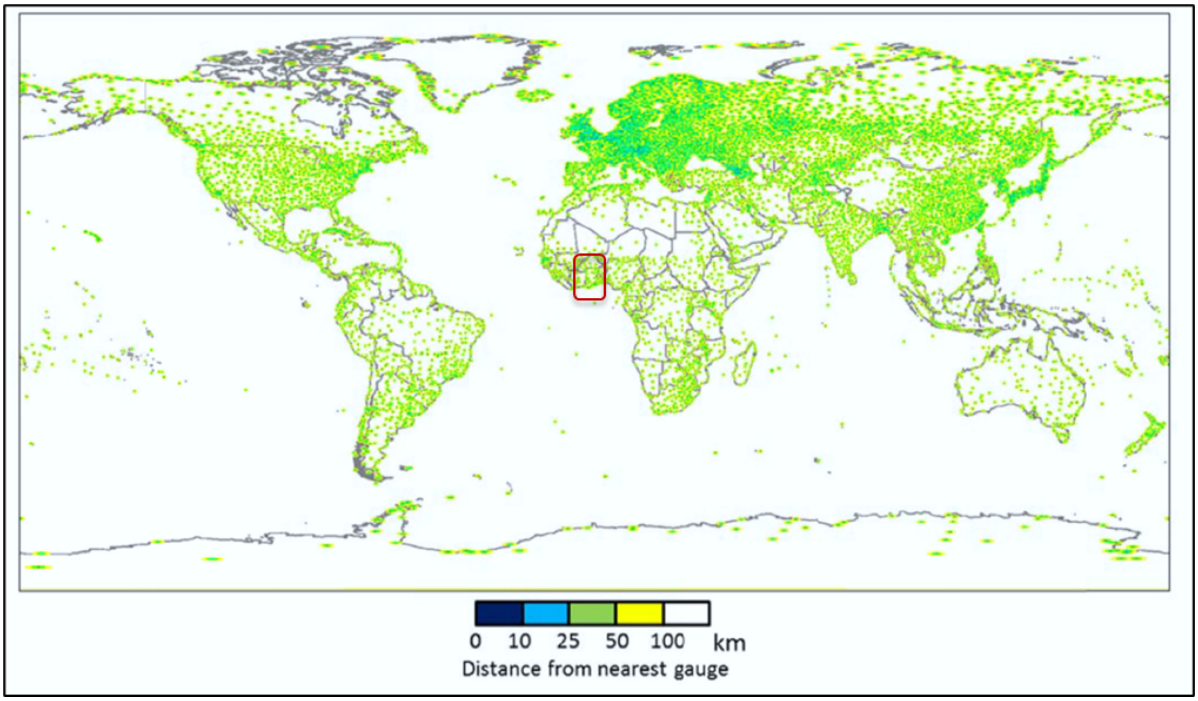

Figure 1.1: Spatial distribution of the rain gauges network across Earth (Kidd et al., 2017). The red circled region is Ghana.

Weather radars are ground-based instruments that transmit radar pulses to estimate precipitation. The high temporal and spatial resolution of the radar instruments gives an important potential for the estimation of precipitation and the improvement of ground based rainfall observations. The principle of the radar method is the quantification of rainfall intensity based on the reflection of radar signal. This method is widely used for hydrological and meteorological applications (Ochoa-Rodriguez et al., 2019; Sokol et al., 2021). In contrast to many benefits, there are also limitations. The most considerable problems stem from i) radar calibration, ii) wavelength errors (when precipitation particles are small), and iii) wave propagation, wave range and wave blockage (Hunter, 1996).

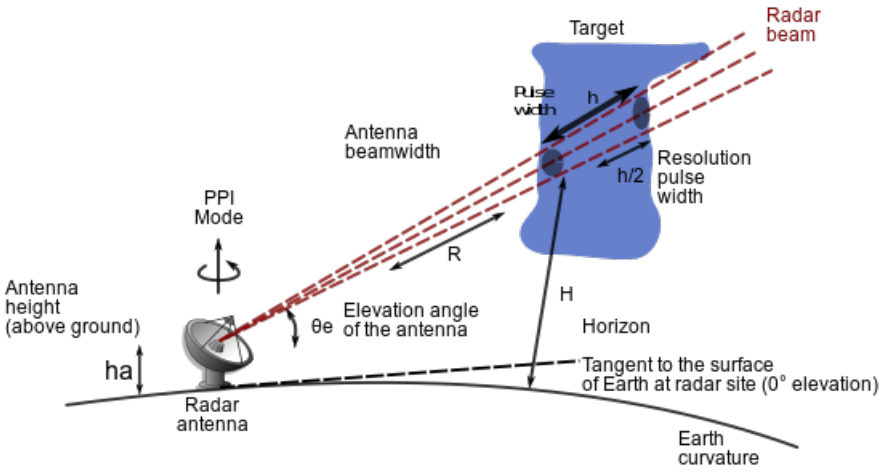

Figure 1.2: Graphical representation of the weather radar rainfall retrieval based on the reflection of radar signal on clouds. (source: Wikipedia)

Remote Sensing (RS) is the approach with the lowest accuracy and the highest potential. The most significant advantages of that method are the global coverage and non-affected by weather, measurements. In addition to these, the rapid development during the last years improved both the temporal and spatial resolution of the observations. Actually, rainfall is not directly measured by the satellite,

there are different approaches to convert satellite signal into rainfall. Some well-known approaches are the estimation of cloud properties, lightning observations, the inversion of soil moisture products and weather models(Le Coz and van de Giesen, 2019).

Apart from the aforementioned direct methods, there are numerical weather models that are used for the prediction of rainfall. These models are either combining past observations and today's weather models to reanalyze historical data and predict rainfall or they are trying to catch rainfall's mechanism (Le Coz and van de Giesen, 2019). Undoubtedly, satellite products are suffering from low performance and Alazzy et al., (2017) addresses the problem with the following factors: i) observations, ii) sampling, iii) retrieval algorithms, and iv) bias correction processes.

The performance of remote sensing rainfall products is noticeably lower (compared to, e.g. Europe). In the early 2000s, the study of Mccollum et al., (2000) tried to name possible reasons for the observed satellite rainfall overestimation over Africa. Except for the sparse rain gauge network, the drop size distribution (DSD) in the clouds is a significant reason. In general, the smaller the drops are in the cloud, the more difficult is to trigger the precipitation. The cloud condensation nuclei (CCN) over an area is closely connected to the DSD of the clouds above it. The concentration of CCN can be changed either by anthropogenic factors (e.g., fire) or by natural factors such as wind blown dust. The Sahara Desert, which is the world's biggest desert is a massive source of aerosols and CCN and affects the surrounding areas (including West Africa).

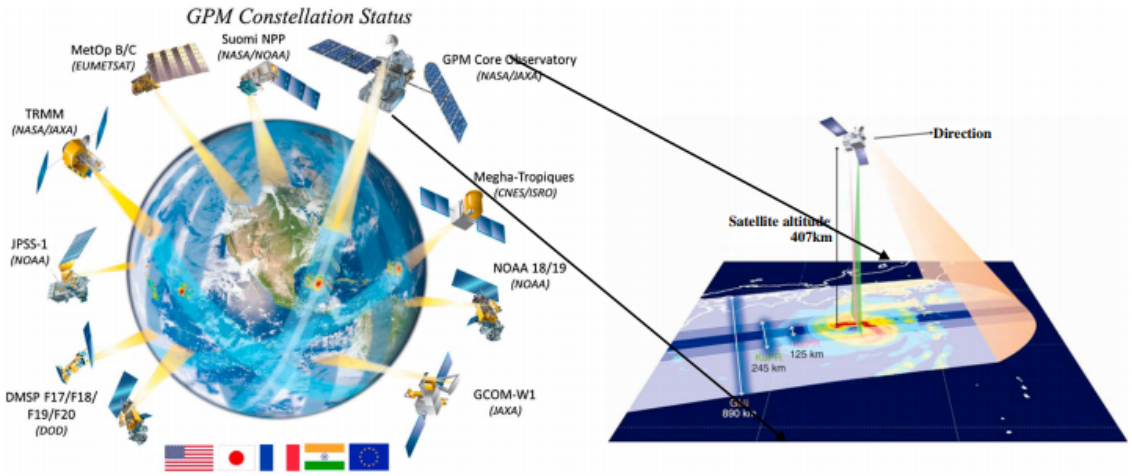

Figure 1.3: Precipitation measurement from satellites that comprise the GPM constellation. (source: Qiaohong et al., 2017)

Taking a closer look to the African continent, the study of Nicholson et al. (2018), analyzes the development of the rain gauge network over Africa since the 19^{th} century. It is clear that even now, the continent of Africa suffers from a dearth of ground based data. Furthermore, the convective rainfall events over West Africa increases the spatial variability of the events and discourages the interpolation of the observed rainfall over larger areas (country-level). Finally, according to Le Coz and van de Giesen (2019), the weather radar network over Africa is not dense enough. Also, the weather models cover the whole Earth, but they have low spatial resolution and accuracy. Therefore, the power of remote sensing for rainfall retrieval should be further investigated.

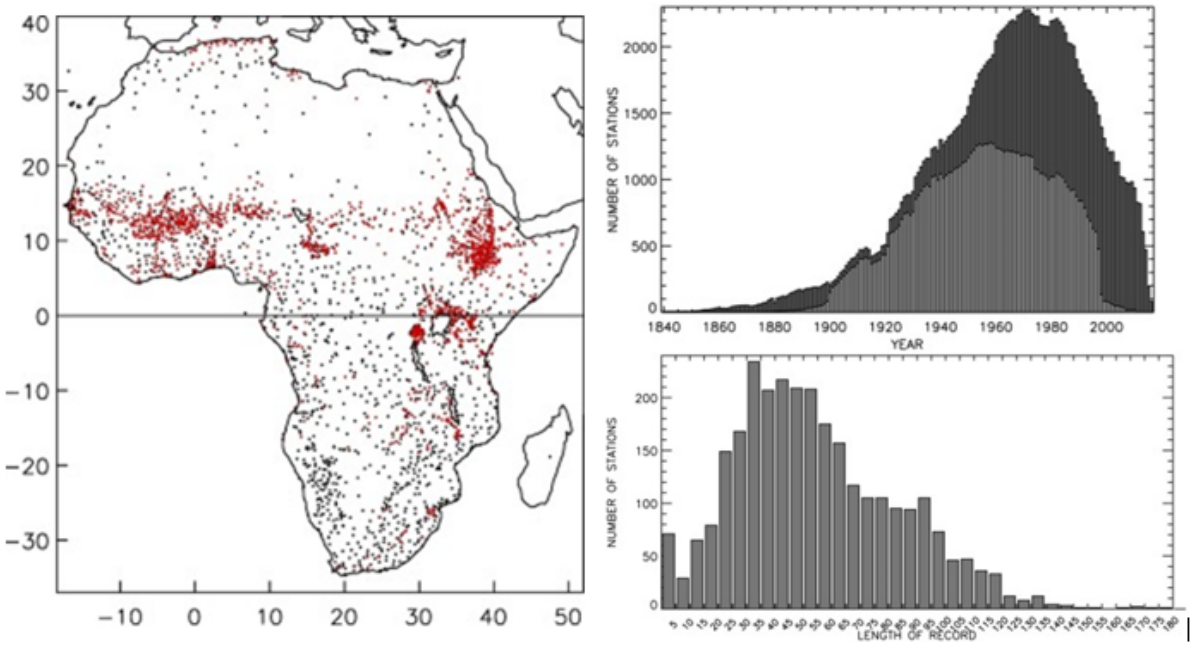

Figure 1.4: Map of the African rain gauges and their development during the last 180 years (Nicholson et al., 2018).

1.1.2. The theory of microwave remote sensing

Remote sensors typically observe the electromagnetic radiation that comes from the surface of the Earth; microwaves are part of the electromagnetic spectrum. Microwave remote sensing can be divided into two subcategories, passive and active, based on the type of satellite sensor. The microwaves are part of the long wave range of the spectrum (1cm to 1m), thus they are sensitive to specific physical properties. Long wave microwaves are able to pass small particles in the air (dust, clouds, etc.) but not the heavy rain drops. Longwave signal (microwave) is not affected by the atmosphere and the weather (Canada Centre for Remote Sensing, 2019). Since the 80s, many studies investigated the correlation of the backscatter signal from microwave remote sensing with soil moisture content (Jackson, 1993; Schmugge, 1984). This approach opened the road to correlate microwave remote sensing with rainfall directly.

An important parameter of the microwave theory is the polarization mode of the signal. The microwave signal can be transmitted and received horizontally (H), vertically (V), or in a combination of them. Two letters are used to characterize the polarization of an imagery, the first one corresponds to the polarization of the transmission and the second to the polarization of the receiver (e.g. polarization mode 'VV' which means vertically transmitted and vertically received).

Radiometers (passive sensors) are instruments that measure the microwave radiation that is naturally emitted from the surface of the Earth and the atmosphere. The concept of passive remote sensing is similar to thermal remote sensing. Objects continuously emit a low amount of energy which is related to the temperature and the wetness condition of the object. The passive sensors obtain signals from the atmosphere, the outer surface, and the inner part of the soil. The problem with this sensor is that it is not sensitive enough to receive low values of signals that come from the earth, which results in low spatial resolution (Canada Centre for Remote Sensing, 2019).

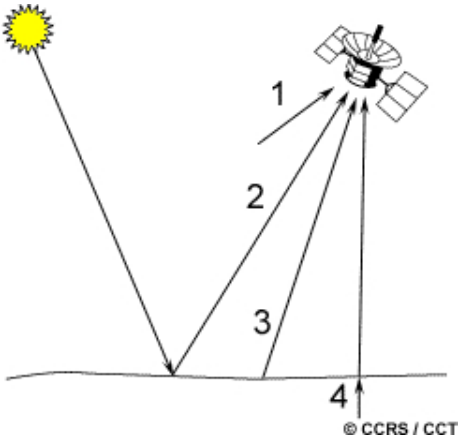

Figure 1.5: Graphical representation of the passive microwave remote sensing. Sensor observes signal that is naturally emitted from Earth's surface. (source: Canada Centre for Remote Sensing, 2019)

Active sensors, transmit a microwave signal and measure the backscatter reflection of it (Woodhouse, 2017). They can be categorized as imaging and non imaging. The main distinction of these two types of sensors is the dimensions of the results because non-imaging sensors cannot provide a two-dimensional image. Similarly to the passive sensors, the active ones are not affected by the weather or the time of day. In other words, they can operate during all the day, under any weather conditions or under any cloud coverage (Canada Centre for Remote Sensing, 2019).

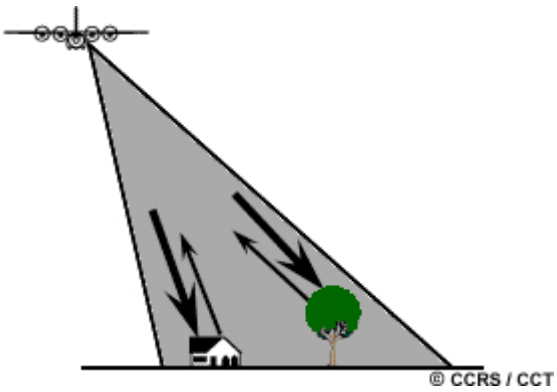

Figure 1.6: Graphical representation of the active microwave remote sensing. Sensor emits signal and measures its reflection from Earth's surface. (source: Canada Centre for Remote Sensing, 2019)

The first meteorological satellites for weather observations were put in orbit in 1960 by NASA. No later than 1970, satellites were equipped with microwave radiometers for remote sensing. At the same time, the first missions equipped with active radars were launched into space and the first satellite imagery was obtained by the end of the 1970s (Barret, 2001). Microwave remote sensing is widely used for agriculture, flooding and water resources management (Mohan, 2013). In the field of hydrology, there are many remote sensing products for precipitation amounts (or intensity of precipitation) that are generated by different approaches.

1.1.3. Overview of satellite-based rainfall products

Satellite-based rainfall products are produced from spaceborne data. There are numerous rainfall products with different spatial and temporal coverage, spatial and temporal resolution, and data source. The catalog of these products is large, so an overview of the most important products/approaches is presented. The main source of data for the satellite-based rainfall products are based on infrared and microwave signals. Apart from that, different products are using secondary data sources such as in-situ observations, weather radar estimates, and numerical weather models to enhance the performance of the product. The thermal infrared signal is used to invert cloud top temperature (CTT) to rainfall amount. In contrast to the CTT method which indirectly quantifies rainfall, the passive remote sensing

backscatter signal is more direct. Generally, only three missions (Tropical Rainfall Measuring Mission (TRMM), the Global Precipitation Measurement (GPM) and the CloudSat) are equipped with active sensors(Le Coz and van de Giesen, 2019). The infrared signal which is essential for the CTT approach gives high temporal products with a large coverage but it is prone to bias related to the cloud model and the variability of the rainfall events. On the other hand, the lower frequency signal that is transmitted during passive remote sensing can have better results. Kidd, 2001 claimed that the performance of the passive data is considerably low over snow and desertous areas. The following table contains the most important products with some useful characteristics (Aonashi et al., 2009; Ashouri et al., 2015; C. Funk et al., 2015; C. C. Funk et al., 2015; Hsu and Sorooshian, 2009; G. J. Huffman and Bolvin, 2011; G. J. Huffman et al., 2018; Maidment et al., 2017; Mega et al., 2019; Novella and Thiaw, 2009; Xie and Arkin, 1996; Xie et al., 2003). Most of the products (Table 1.1) combine a number of different data sources or different products to estimate rainfall in order to outbalance the drawbacks of the individual sensors (Le Coz and van de Giesen, 2019). Apart from these methods, different approaches have been developed the last years. The MSWEP (Beck et al., 2017) concept combines different precipitation products and not directly the satellite measurements. In the following table, IR stands for thermal infrared, PMW stands for passive microwave, sat-radar for satellite radar, and RT for real time.

Table 1.1: Major satellite rainfall products and their main characteristics that cover African continent. (source: Le Coz and van de Giesen, 2019)

Product	Data		Coverage	Spatial Resolution	Temporal Resolution
GPCP-SG	IR, PMW, gauges	Global	1979 -	2.5° x 2.5°	Monthly
CMAP	IR, PMW, gauges (R1 in enhanced version)	Global	1979 -	2.5° x 2.5°	Monthly
GPCP-pentad	GPCP-SG, CMAP	Global	1979 -	2.5° x 2.5°	Pentad
TMPA 3B43	IR, PMW, sat-radar, gauges	50°S - 50°N	1998 -	0.5° x 0.5°	Monthly
PERSIANN-CDR	IR, GPCPv2.2 (PMW)	60°S - 60°N	1983 -	0.25° x 0.25°	Daily
PERSIANN	IR, PMW	60°S - 60°N	2000 -	0.25° x 0.25°	Hourly
ARC2	IR, gauges	20°W - 55°E	1983 -	0.10° x 0.10°	Daily
GSMaP standard	IR, MWP, (sat-radar, model)	60°S - 60°N	2000 -	0.10° x 0.10°	Hourly
GSMaP-gauge	IR, MWP, gauges (sat-radar, model)	60°S - 60°N	2000 -	0.10° x 0.10°	Hourly
IMERG–final	IR, PMW, sat-radar, gauges	60°S - 60°N	2014 -	0.10° x 0.10°	Half-Hour
CHIRPS	IR, TMPA 3B42, gauges, CFSv2, CHPclim	50°S - 50°N	1981 -	0.05° x 0.05° & 0.25° x 0.25°	Half-Hour
PERSIANN-CCS	IR (PMW)	60°S - 60°N	2003 -	0.04° x 0.04°	Hourly
TARCAT	IR (gauges)	Africa	1983 -	0.0375° x 0.0375°	Daily

The SM2Rain is a novel algorithm introduced by Brocca et al., (2014) that investigated the relation of microwave remote sensing observations with precipitation by inversing a soil moisture satellite product. It is a bottom-up algorithm which quantifies precipitation using soil moisture satellite products such as ASCAT (Wagner et al., 2013), AMSR-E (Owe et al., 2008) and SMOS (Kerr et al., 2012); the product is daily with spatial resolution of 12.5 km x 12.5 km. The main principle of that approach is the expression of precipitation as a function of relative soil moisture. The function also includes three parameters (Z, α and b); parameters α and b represent the non linearity between soil saturation and drainage rate and Z is used to model layer depth and varies regarding to the land use (Brocca et al., 2019; Brocca et al., 2013; Brocca et al., 2017). The algorithm follows five steps: Rescale of soil moisture products to 1^o resolution, classification of area based on the historical rainfall, calibration of Z, α and b with objective functions, retrieve rainfall and finally application of filter in extreme conditions (e.g. freeze). In the study of Brocca et al., (2014), the land classification (based on rainfall amount) was applied using a GPCP product as a benchmark. At a further step, the calibration of the parameters was performed for the first two years and the validation for the next two. The results proved a significant correlation between soil moisture and precipitation with a global median R of 0.542. The three major limitations of the aforementioned algorithm are i) the fact that the backscatter signal corresponds to the upper surface of the ground (< 7 cm.), ii) the spatial resolution of satellite's data (the resolution the pixels) which are about 20 kilometers and iii) the heterogeneity of the soil (Brocca et al., 2017). The term heterogeneity corresponds to the wetness of the soil (dielectric constant) and the geometry (vegetation

stage/phenology). In other words, the behavior of the backscatter signal varies for different conditions. The study of Muhammad et al., (2020) concludes that SM2Rain algorithm is not accurate in really humid (Muhammad et al., 2020; Satge et al., 2020; Tang et al., 2020) and mountainous areas but outperforms other satellite-based products (e.g., IMERG, TRMM), especially for rain detection.

1.1.4. Machine Learning and Remote sensing

For the first time, Samuel, (1959) introduced the concept of Machine Learning (ML) which is a discipline of Artificial Intelligence (AI) by developing two ML algorithms. The next decades, many studies were published about the ML methodology; unfortunately, the implementation and use of these algorithms was hindered by the massive computational time needed (James et al., 2013). The study of Hinton et al. (2006), retriggered the Machine Learning and Deep Learning approach and since then numerous studies have been published. According to Samuel, (1959), Machine Learning is *"the field of study that gives computers the ability to learn without being explicitly programmed"*. The ML algorithms can be distinguished in two basic categories based on the learning methodology the supervised and the unsupervised learning. The Supervised Learning (SL) is trained using the reality (labels) that is included in the training data to achieve fitting and later feed the ML with future data (unlabeled) to predict the result. On the other hand, the labels (truth) are not used by the unsupervised learning. Both types of learning can be used for classification and regression learning. Classifiers are algorithms that are using the input variables to predict a discrete result (e.g., 0, 1, 2, etc.).

Many studies are proving the positive contribution of Machine Learning in the field of Earth Observation; the main advantage of the ML is the successful handle of data with high-dimensionality. The most popular machine learning algorithms in remote sensing are Support Vector Machine (SVM), Decision Tree (CART) and Artificial Neural Network (ANN) (Maxwell et al., 2018). Lamovec et al., (2013) and Lary et al., (2018) published their study in the detection of flooding and dust sources respectively. The metanalysis performed by Ma et al., (2019) showed the successful land use/land cover classification (but still not so accurate) using deep learning which outperforms the conventional classifiers. The recent study of 2021, Song et al. presented the detection of air pollution during Covid19 era using a random forest algorithm and Xu et al., (2021, managed to achieve more than 95% accuracy in oil detection. The correlation of remote sensing with soil moisture using machine learning is assessed in different studies. 2010, (Ahmad et al.) used an SVM classifier to calculate soil moisture content in the ground; the performance of the classifier ranged up to 80%. In addition to that, 2020, (Adab et al.) used different machine learning algorithms to estimate soil moisture reaching an accuracy of about 75%. In the field of rainfall detection and quantification, there are many studies that are using ML algorithms for remote sensing rainfall products (and many times also in situ data) (Chen et al., 2019; Kühnlein et al., 2014; Moraux et al., 2019; Shin et al., 2021; Stampoulis et al., 2019; Turini et al., 2021).

1.1.5. Problem Statement

According to Le Coz and van de Giesen (2019), the performance of the satellite rainfall products is highly connected to the temporal and spatial resolution, the in-situ stations, the geomorphology of the area and the seasonality of the rainfall.

As it was mentioned in Section 1.1.3, there are multiple satellite rainfall products that use data from different sources (PMW, sat-radar, IR, rain-gauges, a combination of the previous and other). After a further investigation in studies which compare the performance of satellite rainfall products over Africa, it can be seen that the most accurate satellite rainfall products use also ground observations as input data (e.g., CHIRPS) (Nkunzimana et al., 2020, Dinku et al., 2018, Dembélé and Zwart, 2016, Roy et al., 2020). Products dependent on rain gauges' data, suffer from a significant limitation that they are highly affected by the quality of the ground observations. Moreover, the performance and the reliability of ground observations are not perfect and often coverage of ground stations is sparse in West Africa. Undeniably, a quality control system for ground stations in low dense areas is crucial to enhance the performance of satellite rainfall products and ensure the regular operation of ground stations. In 2014, Brocca et al. in their study, developed the SM2Rain product, which is a bottom-up algorithm, that inverts satellite soil moisture products to precipitation and is only depended on satellite observations of active remote sensing. The performance of the SM2Rain for rainfall estimation was found competitive to well known products; though, the recent study of Duan et al. (2021), proved that the performance of the SM2Rain approach was low to assess rain/no rain. The presence of dense vegetation over the soil is one of the most significant factors behind that low performance (Brocca et al., 2017). For that reason,

inspired by the study of Brocca et al. (2014), the idea behind our study was to overcome the limitation of the vegetation by selecting homogeneous patches (e.g. soccer fields) with simple vegetation cover to use them as rainfall detectors.

When it comes to soccer fields in Ghana, most of them are not irrigated and the vegetation on them is low with small fluctuations during the year (simple phenology). Undoubtedly, it is not fully correct to assume that the same density and seasonality/pattern of vegetation occurs on all Ghanaian soccer fields. Some soccer fields are vegetated only during the wet season, other are bare only during the dry season and other are bare or vegetated during the whole year. For that reason, a criterion was set to narrow down the list and select the soccer fields that could potentially be used as rainfall detectors. The criterion of that selection was the performance of Pearson's correlation between the backscatter signals of the soccer field and its modeled soil moisture content. Fields with a good Pearson correlation coefficients were considered to have consistent backscatter signals.

Microwave remote sensing is known to have a good sensitivity to soil moisture. Passive soil moisture satellites cannot be use for monitoring of soccer fields due to their coarse resolutions. Instead, active radar, especially Synthetic aperture Radar (SAR) was chosen as it is has a high enough reolution to monitor soccer fields. Observations from Sentinel-1 were used in our study; specifically, the imaging radar mission Sentinel-1, which was launched by 2016, is equipped with Synthetic Aperture Radar (SAR) and provides C-band imagery with a resolution of 5x20m (ESA, 2020) and a temporal resolution of six days. Sentinel's-1 high spatial resolution makes the retrieval of spaceborne observations of soccer fields (small patches) possible, while its high temporal resolution (frequency of observations) achieves the retrieval of the soil dielectric constant of these fields.

1.2. Research Objective

The main aim of the study is to detect rainfall occurrence (rain/no rain), by applying a machine learning approach to satellite data retrieved from selected, homogeneous land cover, patches.

During the last decades, Machine Learning is widely used for Earth Observation. Supervised Learning classifiers use known input data and predictions to train a classifier and then make predictions for new input data. The main concept of this study is to deploy a Machine Learning classifier that would assess rain/no rain (quality control) over specific soccer fields, based on the backscatter signal from Sentinel-1. The following figure visualizes the reflection of the backscatter signal over fields with different soil wetness conditions.

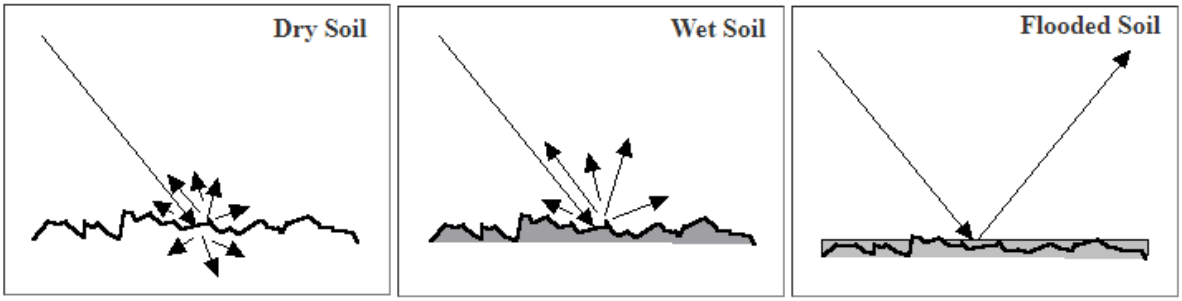

Figure 1.7: The reflection of satellite signal (emitted from active remote sensors) on the soil. The lowest values of backscatter signal correspond to the driest soil condition. (CRISP, 2001)

At this point, the main question and seven subquestions, regarding this study are put forward. Answering the subquestions, the main question of that research will be covered also. These subquestions are the guideline of this thesis.

Could we use the satellite's backscatter signal of soccer fields (with simple phenology) to assess rain/no rain using a machine learning approach?

– ***Question 1:*** How can we retrieve a big number of soccer fields with simple and stable vegetation (simple phenology) during the year and similar shape and size?

- ***Question 2:*** How to select which of the soccer fields are suitable to be used as rainfall detectors?
- ***Question 3:*** Which variables are important to assess rain/no rain?
- ***Question 4:*** How can I combine different Machine Learning classifiers to assess rain/no rain based on satellite data?
- ***Question 5:*** What is the stability and the generalization capabilities of the model when trained with and applied to different locations and times?
- ***Question 6:*** What is the performance of different classifiers?
- ***Question 7:*** Is the performance of the model competitive with other satellite rainfall products (e.g. IMERG)?

1.3. Outline

The manuscript is structured in five chapters. Chapter 2 consists of two main sections, the materials and the methodology of the study. The material part describes the satellite and the ground-based data observations that are used. The methodology part briefly presents the preprocessing procedure, the preparation of the data and the Machine Learning (classification) algorithm that is used. In addition to these, the same chapter also mentions the different datasets that are used for the training and testing of the classification and the metrics of classification. Chapter 3 presents the results of the methodology followed in the previous chapter. At next, the limitations and assumptions of the methodology and their effects on the results are further discussed in Chapter 4. Finally, Chapter 5, concludes with the major outcomes of the study and suggests parts where further research is necessary.

2

Methods and Materials

The main objective of this study is to develop a machine learning algorithm to assess rain/no rain over specific soccer fields that can potentially be used as a quality control tool for the in-situ network. Sentinel-1 is equipped with a SAR instrument, which means that the provided backscatter signal is affected by the vegetation but not affected by the weather. In addition to that, the phenology of the vegetation of soccer fields over developing countries (e.g. Ghana) is simple. Considering also, that the shape and the size of soccer fields is relatively the same, they are suitable to correlate the backscatter signal with the dielectric constant (wetness of soil). Firstly, Section 2.1 describes the location and climate of the study area. Section 2.2 presents the main sources of data and the preprocessing procedure. The last part is the methodology to explain the procedure to reach the objectives (Section 2.3).

2.1. Study area and climate

2.1.1. Republic of Ghana

The Republic of Ghana is a sub-Saharan country in West Africa with a total population of about 31 million people. It is confined between 4°45' and 11° North, and between 1°15' East and 3°15' West and its size is ~238.540 km^2 (FAO, 2018). In the northern part of the country, the mean elevation is higher and the amount of precipitation is lower than in the southern part. According to FAO (2018), the prevailing climate of northern and southern Ghana is Savannas and tropical respectively and in general, the dominant climate of Ghana is Tropical Savanna. The mean monthly temperature in Ghana does not show significant fluctuations (20^{o}C-27^{o}C). Ghana receives the African monsoon that takes place from June to October and is responsible for more than the 70% of the annual precipitation.

In 2018, the main land use in Ghana is agricultural. According to FAO, (2018), the agricultural land is the 62% of the total land (about 14.8 million ha). Half of the cultivated area is land under permanent meadows and pastures for at least 5 years, the 32% is arable land, which is land under temporal crops, and the rest (18%) are permanent crops, which consist of crops that are not replanted for a long period.

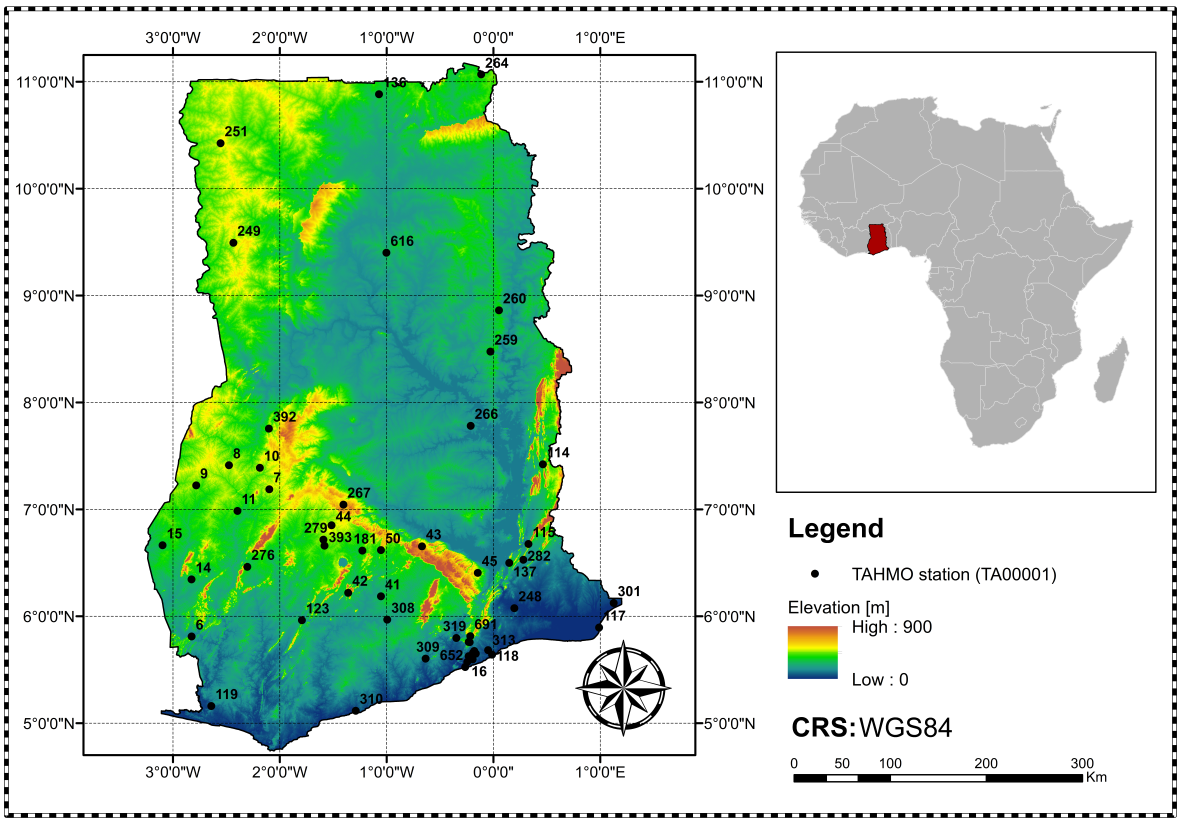

Figure 2.1: Elevation map of Ghana based on NASA Shuttle Radar Topography Mission (SRTM with spatial resolution 30x30m). Also, the location of Ghana in African continent is depicted.

2.1.2. Hellenic Republic

Greece is a Mediterranean country, located in southern Europe with a total population of about 10.7 million (2010). The climate of Greece is Mediterranean with an average temperature of 28^{o}C. According to FAO, (2018) about 47% of the total area of Greece is agricultural (more than 80% of the country's water goes to agriculture). Finally, agriculture corresponds to about 4% of the country's GPD.

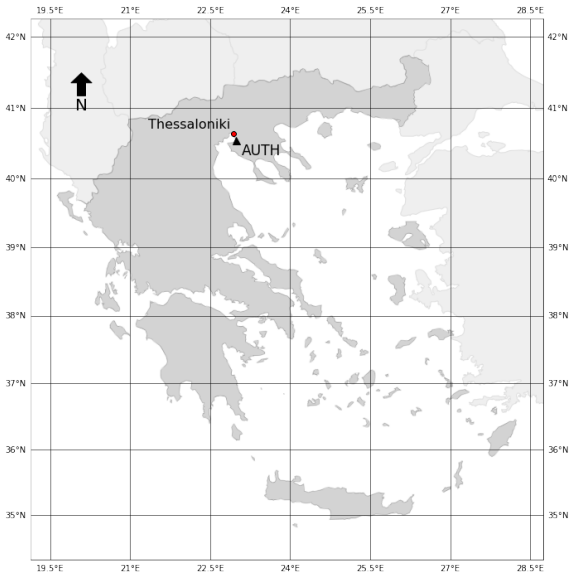

Figure 2.2: The location of AUTH station in Northern Greece, outer of Thessaloniki.

2.2. Data

2.2.1. TAHMO Network

The Trans-African Hydro-Meteorological Observatory (TAHMO) network (https://tahmo.org/) is operating since 2014 and it is a non-profit organization (Van De Giesen et al., 2013). The idea behind TAHMO is the development of a dense network with a density of one station per 30 kilometers (totally 20,000 stations) throughout the whole of Africa, specifically in the sub-Saharan region. Nowadays, 560 installed stations over the whole continent can be found; they are taking frequent measurements (5-minutes) of meteorological variables. Specifically, 101 of them are located in Ghana, and they are mostly distributed in the southern part of the country and near urban areas. TAHMO database consists of 5-minutes hourly and daily timeseries for 8 variables (Lightning distance, Precipitation, Relative humidity, Shortwave radiation, Soil electrical conductivity, Soil moisture content, Surface air temperature and Wind speed). Unfortunately, variables such as soil moisture content and soil electrical conductivity are available only for specific stations. The hourly timeseries of the Ghanaian stations were downloaded both for 2019 and 2020. The following figure represents the distribution of the mean daily temperature per month in Ghana during 2020.

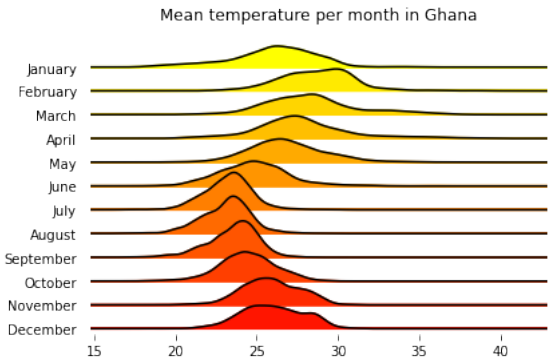

Figure 2.3: The distribution of monthly temperature observed from Ghanaian TAHMO stations in 2020

The following figure presents the annual rainfall, the mean monthly rainfall and the mean monthly temperature for 2020 in Ghana.

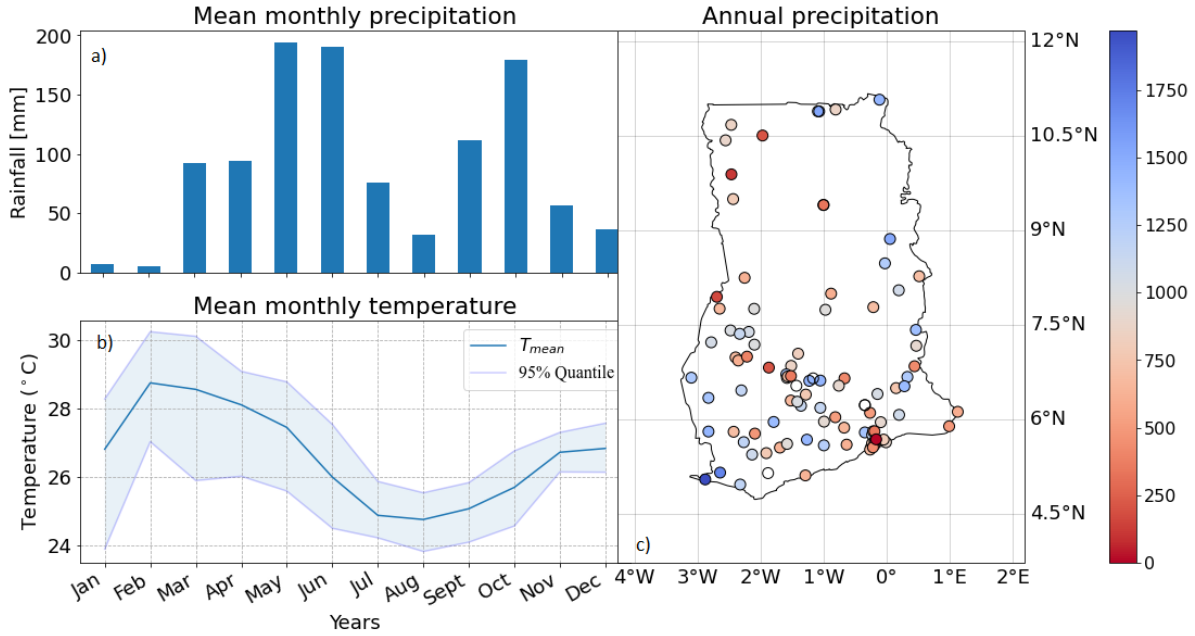

Figure 2.4: a) Mean monthly precipitation, b) Mean monthly temperature and c) Annual rainfall during 2020 of Ghanaian TAHMO stations. The colored stations are representing the annual summary precipitation.

2.2.2. AUTH Station

The AUTH station is located in Northern Greece, outside of Thessaloniki (Longitude = 40°32', Latitude = 22°00' and elevation = 15 meters) and was installed in 2007 in the context of the "MEDDMAN" project. Since then, it measures 6 meteorological variables (temperature, solar radiation, wind speed, relative humidity, rainfall and soil temperature) and stores them per hour.

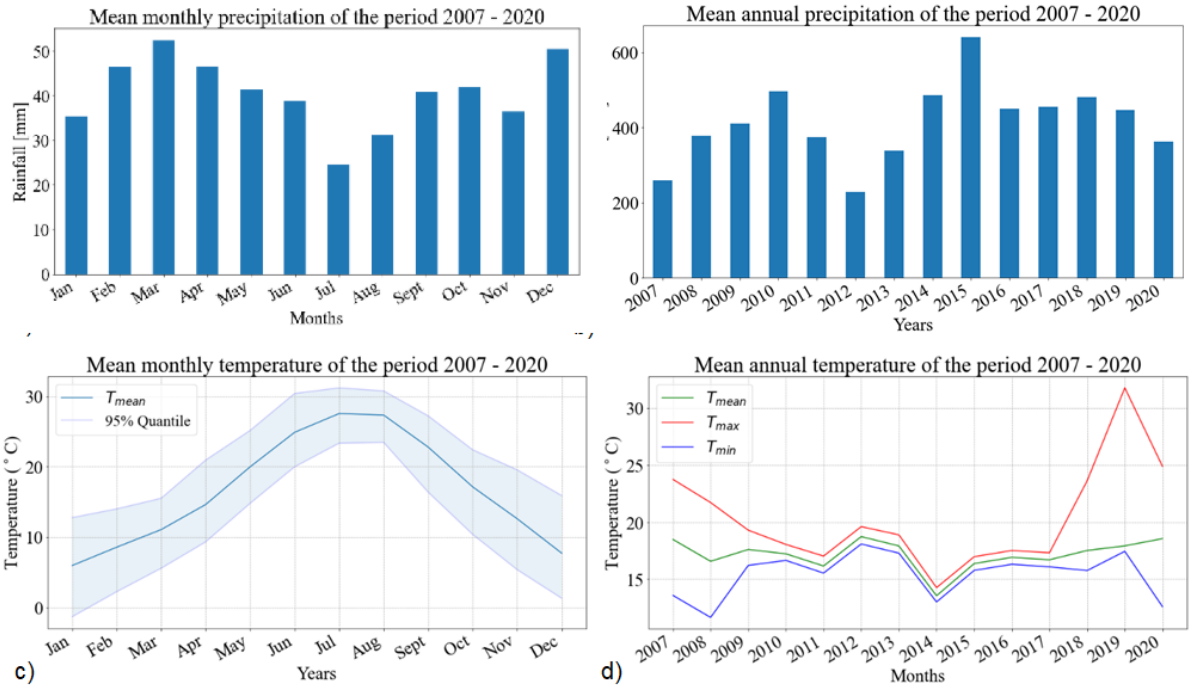

Figure 2.5: a) Mean monthly precipitation, b) Mean annual precipitation c) Mean monthly temperature and d) mean annual temperature for the period 2007-2020 of the AUTH station

2.2.3. Satellite products & Google Earth Engine

2.2.3.1. Copernicus Sentinel-1 mission

In the context of the European Space Agency's (ESA) program, the satellites Sentinel 1A and Sentinel 1B were launched in 2014 and 2016 to contribute to the monitoring of land and earth observation. Both satellites are equipped with a C-Band Synthetic Aperture Radar (SAR) instrument which transmits at 5.405 GHz. Furthermore, according to ESA (ESA, 2020), the radar can send out the signal both horizontally (H) and vertically (V) and can receive both polarizations. The availability of dual polarization data is helpful for earth observation applications.

The Sentinel-1 constellation generates data in four different modes; Strip Map (SM), Interferometric Wide Swath (IW), Extra Wide Swath (EW) and Wave (WV) Mode. Also, there are four types of data products, Level 0 - Raw data, Processed Level 1 - Single Look Complex (SLC) data, Level 1 - Ground Range Detected (GRD) data and Level 2 - Ocean (OCN) data. The GRD multilook product is produced by several Singe Look images, it is projected to the ground range using an Earth ellipsoid model and contains information about the amplitude.

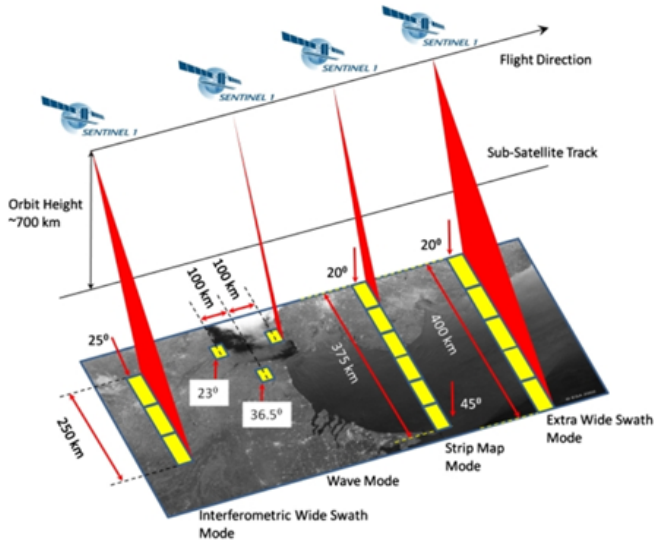

Figure 2.6: The four different acquisition modes of Sentinel-1 products. The products of each mode are provided in three different levels (SAR-Level0, Level1-SLC, Level1-GRD & Level2-OCN). (source: ESA, 2020)

The operation of this active SAR instrument is not affected by the weather, the time of the day (day/night), and the cloud cover.The high spatial and temporal resolution of the products (up to 5 meters) in combination with their rapid product availability, make the specific satellite a suitable choice for earth observation applications.

There are many studies that are showing correlations between soil moisture and backscatter signal's amplitude (Mirsoleimani et al., 2019, Bauer-Marschallinger et al., 2019). Based on ESA's recommendations and the aforementioned literature review, the IW GRD product which is providing high resolution (5x20 m), six-day temporal resolution and 250 km swath range was selected to investigate the behavior of the backscatter coefficient σ^0 after precipitation events. It should be mentioned that the motivation to select GRD product instead of SM (which has higher spatial resolution) is the size of the data (SM product is much bigger that a GRD of the same area), which leads to a significant reduced processing time.

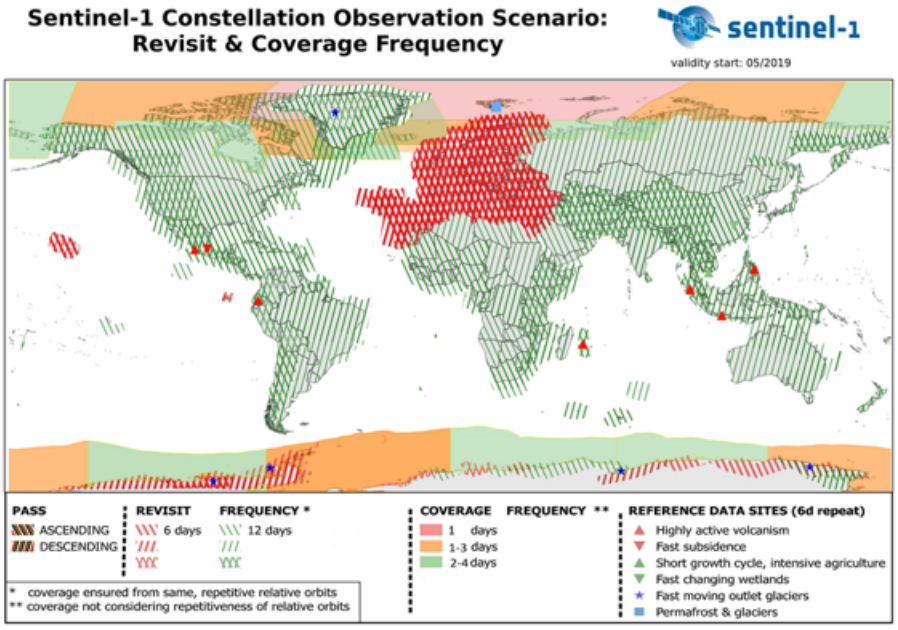

Figure 2.7: Revisit and coverage frequency of Sentinel-1 products. (source: ESA, 2020)

When it comes to polarisation mode, Sentinel's-1 mission generates both HH-HV (double polarization) and HH polarisation bands in polar environments and VV-VH and VV polarisation for the rest of the world (apart from Baltic sea). According to Hong and Wdowinski (2012), the most suitable band for water change detection is HH and the next is VV because VH is more sensitive to vegetation. Thus, for the current study, timeseries of the median VV band, its variation, and the approximate viewing incidence angle θ where exported for specific areas.

2.2.3.2. Global Precipitation Measurement (GPM)
Global Precipitation Measurement, is an international satellite mission to provide next-generation observations of rain and snow worldwide every three hours. The IMERG algorithm combines data from satellite passive microwave (PMW) precipitation estimates and infrared (IW) satellite rainfall estimates. Then, these estimates are reanalyzed with monthly ground precipitation (GPCC) and calibrated with other models (GPM-TRMM) (G. Huffman et al., 2019). The resolution of the product is 0.1°x 0.1° grid within 60°N/S every thirty minutes.

2.2.3.3. Google Earth Engine database
Google Earth Engine (GEE) (Gorelick et al., 2017) is a platform that contains numerous satellite images and geospatial datasets that are callable using an open source (for research/educational purposes) Python API.

Sentinel-1 SAR GRD
For this study, "*Sentinel-1 SAR GRD: C-band Synthetic Aperture Radar Ground Range Detected*" preprocessed collection, which provides the backscatter signal σ_0 (VV) and the incident angle (θ) sampled on a 10x10m grid, is selected. The recommended procedure to preprocess the backscatter signal (Filipponi, 2019) has already been applied in Google Earth Engine aforementioned dataset.

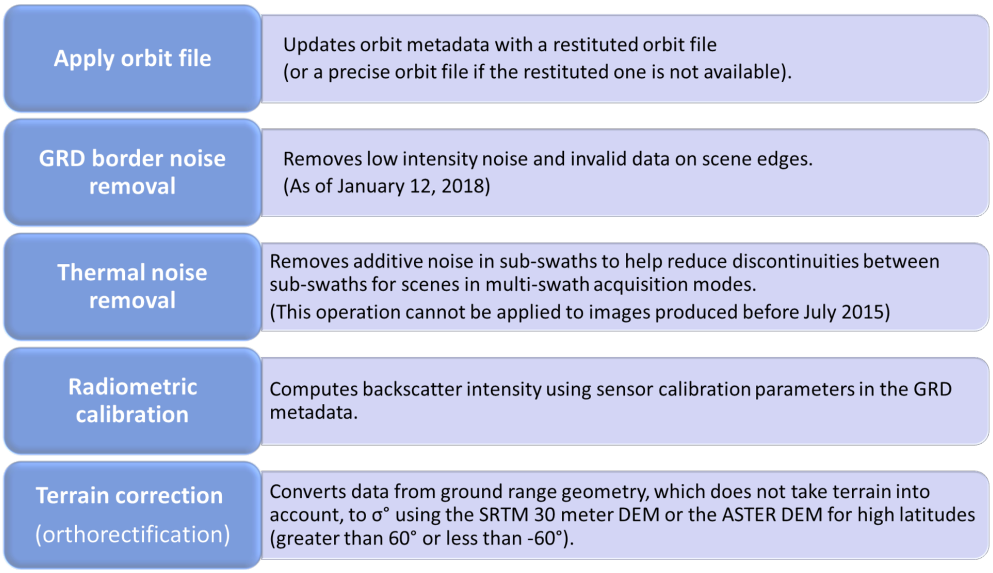

Figure 2.8: Google Earth Engine's preprocessing procedure of the backscatter Sentinel-1 signal. (source: Gorelick et al., 2017)

The final step is the rescaling of raw values to decibels using a logarithmic function:

$$\sigma^0 = 10 \cdot log_{10}(VV) \tag{2.1}$$

where VV is the preprocessed raw signal obtained by Sentinel-1.

Global Precipitation Measurement (GPM)
The GPM v.6 database contains multiple bands for precipitation and each band uses different data source to retrieve precipitation. For this study, the band "precipitation_cal" was used; that band provides the rainfall intensity (mm/hr) every 3-hours, with merged precipitation data from many sources (PMW, IR & Gauges).

2.2.4. OpenStreetMap

In 2004, Steve Coast introduced the OpenStreetMap (OSM) project. Its main purpose was the creation of open source geographical data for the whole world. The development of the project was rapid, in just two years, it reached about 2 million users. The non-profit OpenStreetMap Foundation, which is located in the UK, supports the website. The available data is structured with nodes, ways, relations, and tags and it provides global coverage. The Python module ”OSMnx” (Boeing, 2017) gives the ability of querying and retrieving specific features based on tagging. This allows us to search for features such as soccer fields.

2.2.5. Filtering & Preprocessing

2.2.5.1. Ground stations

During the preprocessing of the ground stations' data, some anomalies were detected and filtered. To be more specific, there are stations with problematic data; in some cases, the timeseries of one or more variables were missing for the whole years (2019 & 2020) or a part of them. Four rules were defined in order to clean the data.

- To assume a day as valid, station should have observations of at least 8 hours.
- The station's timeseries covers more than the 80% of the year.
- If rainfall is missing (NaN) for more than three consecutive days, it is assigned as unknown (NaN) else, the days are assumed as dry (0 mm).
- For the data of 2020: The timeseries of each station start the first day of January 2020.

2.2.5.2. Minimum amount of precipitation

Considering that rainfall data of 2019 and 2020 is available for TAHMO and AUTH stations, labeling is necessary for supervised learning classification. The fact that the observed rainfall is a continuous variable could complicate the classification. For example, a day with observed rainfall of 0.1 mm/6d (which is very low to be detected by the signal) could be considered as wet. Thus, it is important to set an amount of precipitation as a threshold; days with rainfall under that amount will not be considered as wet. After a trial and error method, the number of that threshold was set to 5mm/6d. The same amount was set also for the IMERG precipitation.

2.3. Methodology

A general overview of the methodology that was followed for the study is described here. The following paragraphs are related to the flowchart of the Figure 2.9.

The first goal of the study was to achieve an automated retrieval of soccer fields. In the scope of this objective, a Python tool was developed to automatically extract soccer fields from Open Street Maps using the OSMnx Python module.

The second goal was to assess which of them are suitable to be used as rainfall detectors. The simple phenology and the shape and size of soccer fields make it possible to correlate directly the fields' dielectric constant with the satellite's signal. Therefore, the criterion to select suitable soccer fields was the correlation of its modelled soil moisture (of the closest TAHMO station) with its backscatter signal. For the cases were observed soil moisture was missing, a prior procedure was essential to estimate soil moisture in order to answer the second objective. At first, the daily reference evaporation was calculated using the ASCE standarized method. The input values of the model were the daily timeseries of precipitation, wind speed, temperature, relative humidity and solar radiation at each TAHMO station. In a following step, a simple lumped hydrological model was implemented. The model used daily precipitation and evaporation to estimate the daily modelled soil moisture content based on the leaky bucket approach. In order to calibrate the model's parameters S_{in} (initial storage) and Su_{max} (maximum storage), the backscatter signal was retrieved for a number of soccer fields. The maximum value of S_{in}, Su_{max} was set to 10cm. A trial and error method was applied to correlate backscatter signal with modelled soil moisture of the closest TAHMO station for different model parameters. The modelled soil moisture was calculated for all TAHMO stations using the optimum value of Su_{max}. The soil moisture

of each field was assumed to be equal to the soil moisture of its closest station. At a further step, the backscatter signal was retrieved for all soccer fields and the correlation of the signal with the modelled soil moisture of the field was calculated. Finally, for almost all TAHMO stations, the soccer field with the higher correlation was selected, which resulted in a number of soccer fields in Ghana.

When it comes to Greece, the case was much simpler. The modelled soil moisture of the Greek station was calculated using the same approach, the available soccer fields were detected manually and the most suitable one was also selected based on the Pearson correlation (of its backscatter signal with the modelled soil moisture of the Greek station).

The third research question asks for suitable data for binary rain/no-rain classification. Generally, data from two different sources is available for each soccer field: satellite observations and in-situ precipitation. Aiming to detect rainfall based on satellite's observation changes over a specific field (between two consecutive measurements), it was necessary to prepare the data in a suitable format. Each row of the data was contained data from two consecutive dates. The desired format is depicted in the Table 2.1.

Table 2.1: Desired format of the ground-based, the spaceborne data and the additional variables for a specific field. Column *angle* includes the incident angle of measurement for the date of index. *VV_db* and *VV_var_db* is the observed backscatter signal and its variation over the field for the date of index. *P6d* is the observed accumulated rainfall of the closest to the field TAHMO station for the date of index. Columns with suffix -2 correspond to next date's observations. *Q* is the classification of signals *VV_db* and *VV_db2* of each row, based on the signal of all soccer fields. *P6d_int* is the interpolation of the observed rainfall from the TAHMO network to the specific field.

	angle	VV_db	VV_db2	VV_var_db	VV_var_db2	P6d	Q	P6d_int
Date1	angle1	VV_db1	VV_db21	VV_var_db1	VV_var_db21	P6d1	Q1	P6d_int1
Date2	angle2	VV_db2	VV_db22	VV_var_db2	VV_var_db22	P6d2	Q2	P6d_int2
Date3	angle3	VV_db3	VV_db23	VV_var_db3	VV_var_db23	P6d3	Q3	P6d_int3
Date4	angle4	VV_db4	VV_db24	VV_var_db4	VV_var_db24	P6d4	Q4	P6d_int4
Date5	angle5	VV_db5	VV_db25	VV_var_db5	VV_var_db25	P6d5	Q5	P6d_int5
...	...	...	...	...	...	...	...	...
Date2891	angle2891	VV_db2891	VV_db22891	VV_var_db2891	VV_var_db22891	P6d2891	Q2891	P6d_int2891
Date2892	angle2892	VV_db2892	VV_db22892	VV_var_db2892	VV_var_db22892	P6d2892	Q2892	P6d_int2892
Date2893	angle2893	VV_db2893	VV_db22893	VV_var_db2893	VV_var_db22893	P6d2893	Q2893	P6d_int2893
Date2894	angle2894	VV_db2894	VV_db22894	VV_var_db2894	VV_var_db22894	P6d2894	Q2894	P6d_int2894
Date2895	angle2895	VV_db2895	VV_db22895	VV_var_db2895	VV_var_db22895	P6d2895	Q2895	P6d_int2895

The main source of data comes from the observations of satellites (Sentinel-1) and ground stations (TAHMO and AUTH). In addition to that, two more variables were calculated and used as input data. The columns *angle-P6d* of the Table 2.1 contained both spaceborne and ground-based observed data. The Sentinel-1 constellation consists of two satellites that pass over Ghana and Greece once every 12-days, thus they provide one measurement per 6-days. In order to explain the desired format of the dataframe, the first row of the table above will be used. When it comes to the time, columns *angle*, *VV_db* and *VV_var_db* were observed at *Date1* and the columns *VV_db2* and *VV_var_db2* were observed at the next date which is six days later (*Date2*). The accumulated rainfall *P6d* was the rainfall of 6-days that was observed between *Date1* and *Date2*. In that way, each row contained spaceborne data for two consecutive observations and the observed rainfall between them. This dataframe was built for each soccer field. The measurement of the backscatter signal, corresponded to a field with a size of about 20 m^2. The column *VV_db* contained the median backscatter signal, the column *VV_var* the signal's variation and the column θ the incident angle of measurement. Generally, the columns with a name that ended with 2 (e.g. *VV_db2*), corresponded to the observations of the next Date (e.g. value of $VV\_db2_{(1)}$ at $\text{Date}_{(1)}$ is the same to the $VV\_db_{(2)}$ at $\text{Date}_{(2)}$). Then, the column *P6d* was the 6-days accumulated precipitation of the closest in-situ station was assigned between the two consecutive days of each row. Two more attributes were added in the final data, the interpolated rainfall $P6d_{int}$ and the quadrant of each row Q to enhance the model. The $P6d_{int}$ (see 2.3.4.3) was the estimation of rainfall

on the soccer field and the quadrant of day Q (see 2.3.4.2) was a classification of each *VV_db-VV_db2* pair based on the data retrieved for all soccer fields.The outcome of that research question was a dataframe with the desired format to be used as input for the machine learning classifier.

The fourth goal is to select a combination of multiple machine learning algorithms to develop a quality control tool, (predict rain/no rain), based on spaceborne data. Stacked classifier gives the ability to use multiple classifiers as voters and predict based on the most votes. Each individual machine learning classifier has specific powers and limitations. Using the stacked classifier, the consistency of the model was achieved by averaging the bias introduced by the different classifiers. For the following paragraphs, the performance of the different cases was assessed in terms of classification and the F1 score (Subsection for further explanation) was used as a metric.

Then, in order to evaluate the stability and the generalization capability of the model, the same architecture was trained and tested on different locations (Ghana and Greece) and years. Three of the training sets used data from Ghana and one used data from Greece. At the same time, different test datasets were defined for each training set. In total, seven test sets were defined; six of them were comprised of Ghanaian data during 2019 and 2020 and one was about Greek data from 2020. The performance of the model in the different training and testing gives answer to the fifth research question.

The performance of different classifiers was assessed in order to approach the sixth research question. The stacked classifier was replaced with i) an individual classifier (RF, CART, etc.,) and ii) a different stacked classifier (excluding one base classifier at each time). For the specific assessment, Ghanaian data of 2020 was used.

Finally, to address the last research question, the performance of the stacked classifier was compared to that of the IMERG product. The comparison was performed in terms of rain/no rain binary classification.

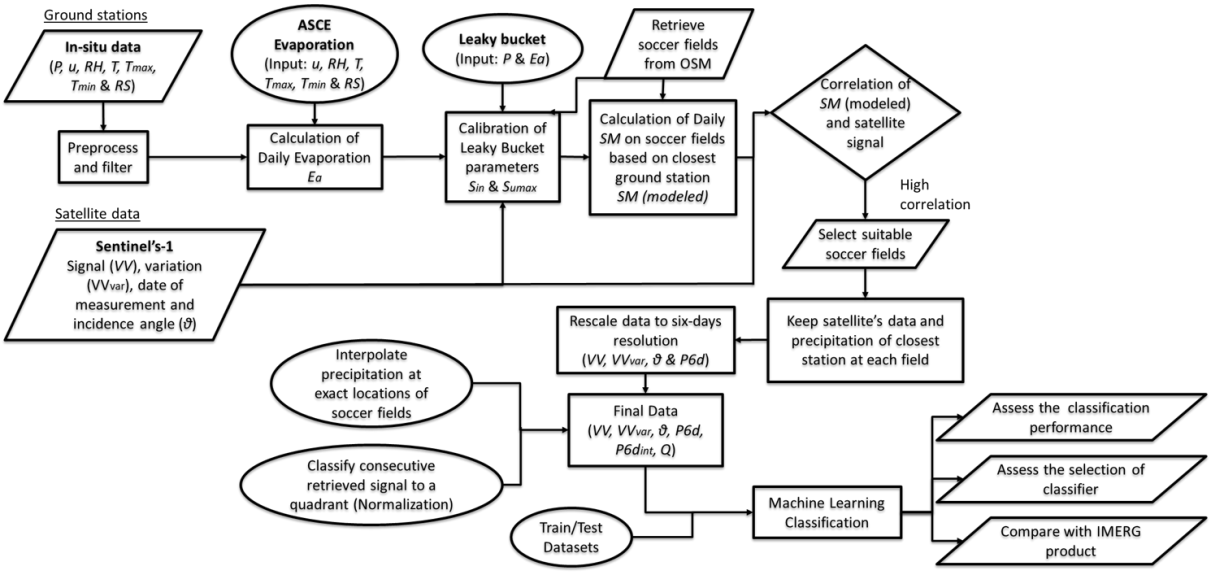

Figure 2.9: Flowchart with the procedures that were applied in this study.

2.3.1. Soccer Fields in Ghana

The first research question of the study regarded the retrieval of soccer fields.

For the detection of soccer fields in Ghana, a Python tool was developed. This module, which used the library ”OSMnx”, gave the opportunity to retrieve data from Open Street Maps based on a specific tag. Soccer fields are assigned in the category of ”Leisure” and subcategories ”soccer”, ”pitch”, etc. Therefore, the first part was to obtain them based on specific tags and then to apply a filter in terms of size and shape. Here the steps are given:

- Retrieve shapes in Ghana with the tag ”Leisure”.
- Filter the obtained shapes with further tags ”soccer”, ”stadium”, ”pitch” and ”stadium”.
- Two filters for shape and size:
  1. The size of the geometries: between 8,500 and 10,000 m^2.
  2. The shape of the pitch: rectangle with 75% tolerance.

The output was a large number of soccer fields over Ghana. At a further step, we will choose specific soccer fields to use as rainfall detectors. The criterion to choose the soccer field is the correlation of the modelled soil moisture (of the closest to the field, TAHMO station) with the backscatter signal from Sentinel-1.

2.3.2. Modelled Soil Moisture

This subsection is a middle step to answer the second research question. Soil moisture is one of the parameters that TAHMO Network monitors but unfortunately, the observations are available only for a limited number of Ghanaian stations. The Pearson correlation between these two variables (SM and backscatter signal) will be used to select suitable soccer fields, thus the daily soil moisture of each station should be modelled.

Generally, the main principle of the specific model is the representation of soil moisture content as a (interception) bucket of fluctuating storage based on specific processes (evaporation, precipitation, runoff, groundwater etc.) (Hrachowitz et al., 2014). In order to eliminate the uncertainty of soil properties, landscape effects, and etc., the storage of the bucket was assumed to depend only on the daily rainfall, recharge and reference evaporation E. At first, evaporation using the ASCE standarized method was calculated.

2.3.2.1. Reference evaporation

The daily reference evaporation of each station was calculated using the ASCE-standardized short reference crop evaporation E_r (Allen et al., 2005) method. The short reference crop (ASCE-short), corresponds to clipped grass of 12 cm height with a surface resistance of 70 $s \cdot m^{-1}$. The input data is the mean, maximum and minimum temperature, the elevation and the location of each station. In cases where the calculation of daily evaporation was not possible because of missing data, the daily evaporation was assumed equal to the mean annual evaporation at the specific station. The formula of the short reference evaporation is:

$$E_r = \frac{(0.408 \cdot \Delta \cdot (R_n - G) + (\gamma \cdot \frac{C_n}{(T_{mean} + 273.16)} \cdot u_2 \cdot (e_s - e_a))}{\Delta + \gamma \cdot (1 + C_d \cdot u_2)} \tag{2.2}$$

where E_r: is the reference short crop evaporation ($mm \cdot d^{-1}$), Δ: is the slope of the saturation vapour pressure-temperature curve ($kPa \cdot °C^{-1}$), R_n: is the net radiation at the crop surface ($MJ \cdot m^{-2} \cdot d^{-1}$), G: is the soil heat flux density at the soil surface ($MJ \cdot m^{-2} \cdot d^{-1}$), γ: is the psychrometric constant ($kPa \cdot °C^{-1}$), u_2: is the wind speed at 2 m height above the soil surface ($m \cdot s^{-1}$), e_s: is the saturation vapour pressure (kPa), e_a: is the actual vapour pressure (kPa), T_mean: is the mean daily air temperature ($°C$), C_n and C_d: are equal to 900 ($K \cdot mm \cdot s^3 Mg^{-1} \cdot d^{-1}$ or $K \cdot mm \cdot s^3 Mg^{-1} \cdot h^{-1}$) and 0.34 ($m \cdot s^{-1}$) respectively, which varies according to the time step and the reference crop type and describe the bulk surface resistance and aerodynamic roughness (Allen et al., 2005).

The daily evaporation, which is the result of the ASCE method, will be used as input in the Leaky Bucket Model together with the daily precipitation.

2.3.2.2. Leaky bucket model

The daily precipitation and evaporation of each station are used to model the soil moisture content of the station. The modelled soil moisture of each soccer field was assumed to be equal to the modelled soil moisture of the closest TAHMO station.

The interception bucket fills with precipitation while rainfall occurs, changing the initial state of storage (S_i). When the amount of rainfall ($P \cdot \Delta t$) exceeds the maximum initial storage S_{max} at a specific time step, the effective rainfall (P_e), which corresponds to the additional amount of water, flows out the bucket. When it comes to evaporation, it is assumed as zero during the wet days. Actual evaporation E_a, is equal to reference evaporation E_r times the ratio between the actual $S(t)$ and the maximum S_{max} storage. Finally, the interception storage is the the initial water minus the amount of the evaporation. The initial (S_{in}) and the maximum storage (Su_{max}) of the soil are the two parameters of the model. In cases when the precipitation of a day is observed, NaN (which means that more than three days were missing), the interception bucket's storage S_i, stays the same. The mathematical description is the following:

$$S_{i[t+dt]} = S_i + P \cdot dt - P_e \cdot dt - E_a \cdot dt \tag{2.3}$$

where:

$$S_e = \begin{cases} max(0, S_{i[i]} - S_{max}), & \text{if } P \cdot dt > 0 \\ 0, & \text{else} \end{cases} \tag{2.4}$$

and:

$$E_a = S_{i[t+dt]}/Su_{max} \cdot E_r \tag{2.5}$$

where P_e is the daily excess rainfall [mm], S_i is the initial condition of the bucket [mm] and E_i is the evaporation [mm]

The modelled soil moisture played a key role for the selection of soccer fields. As it was mentioned, the soil moisture of each soccer field was assumed equal to the soil moisture of the closest ground station. The correlation of the modelled soil moisture with the retrieved backscatter signal at each field will be investigated.

2.3.3. Selection of suitable soccer fields/rainfall detectors

In order to select suitable fields, which is actually the second research question, the timeseries of the backscatter signal of each soccer field were retrieved. The criterion to assess the suitability was the correlation of the modelled soil moisture content in patches (based on the model of the closest TAHMO station) with the retrieved backscatter signal σ_o.

Using the same criterion, the selection of an appropriate rainfall detector in Greece was done by a manual search. The AUTH station is located outside the city, near the airport area where there is plenty of empty space.

The selection of fields that can be used as rainfall detectors was achieved both in Ghana and Greece. There are two sources of data, satellite data and in-situ observations. The temporal resolution of these sources is different, the lower one belongs to the satellite signal (6-days). Before using the data of each field, as an input of the stacked classifier, it was essential to rescale them to have the same temporal resolution. Also, two more attributes were added to enhance the performance of the stacked classifier.

2.3.4. Preparation of dataframe

The preparation of the final data corresponds to the third research question. The goal of that part is to prepare the final data of each soccer field that will be used as input for the stacked classifier. The goal of the classifier is to detect rainfall based on the change of backscatter signal. Therefore, the data of each soccer field should have the backscatter signal observation and its variation (over the field) for two consecutive days, such as the satellite's incident angle of measurement and the rainfall between them for the whole year. Additionally, two more attributes were added to improve the performance of classification, the quadrant of each row and the interpolated rainfall at each field. Therefore, this section is structured in three parts, the first contains the observed data and the rest two describe the two additional variables that were used as input.

2.3.4.1. Spaceborne & in-situ data

The temporal resolution of each individual satellite (Sentinel-1A and Sentinel-1B) over Ghana is twelve days, so the resolution of the final product is 6-days. In order to investigate the relations between the backscatter signal and the observed precipitation, it was necessary to merge the data from both different sources to a final dataframe. The available TAHMO data are from 2019 and 2020, and the following procedure was done separately for both years.

The first step was to rescale the daily timeseries (TAHMO observations) to fit the dates that the satellite has taken measurements. The index of the merged dataframe is the list of the dates that Sentinel-1 performed observations each year. Each row (date) of the dataframe contains two backscatter signal values (*VV_db* and *VV_db2*) and their variation (*VV_var_db* and *VV_var_db2*), the incident angle of measurement (θ) and one accumulated precipitation value (*P6d*) between the two consecutive dates. The *VV_db* signal is the observed signal which corresponds to the date of the index and the second one (*VV_db2*) corresponds to the next week's observation (six days later). The same holds for the variations *VV_var_db* and *VV_var_db2*. The accumulated rainfall (*P6d*) is the amount of rainfall between the measurement of *VV_db* and *VV_db2*. Sentinel-1 passes over Ghana between 18:00 to 19:00 (as a function of latitude). For that reason, each day was assumed to run from 18:00 until 18:00. Moreover, each satellite has a swath width equal to 250 km, so, the whole territory of Ghana is covered in three different swaths (as a function of latitude again). That means that the overlapping area between two consecutive swaths contains data with double size (two timeseries with six-day interval, one from Sentinel-1A and one from Sentinel-1B).

The same procedure was followed in Greece with the difference that Sentinel passes at 16:30, thus the base time of resampling was different.

2.3.4.2. Quadrant of row

This part also answers to the third research question. After the retrieval of the backscatter signal for each field, one more attribute was added to the merged dataframe, the quadrant (*Q*). Each row of the final dataframe contained two signal measurements, one for the date of the index and one for the next date. The median value of these two signals was calculated using the data of all retrieved soccer fields. Then, each row of final data was categorized into four quadrants based on the two consecutive signals and the median value of the total retrieved fields. The *Q* parameter is categorical (1-4) and represents an early prediction of rain/no rain based on the two consecutive signals.

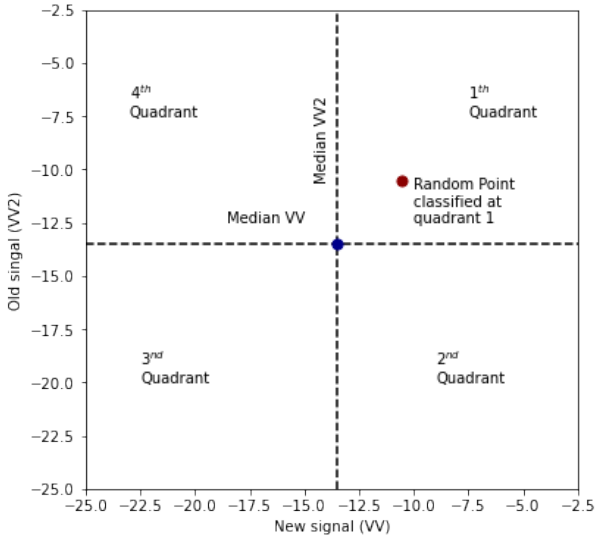

Figure 2.10: Classification of each row of each soccer field at a quadrant. The blue dot represents the median values of *VV_db* and *VV_db2* based on the observations of all soccer fields. The red dot is placed based on the *VV_db* and *VV_db2* of each row of each soccer field.

2.3.4.3. Interpolated rainfall

Finally, this is the last variable to answer the third research question. Unfortunately, TAHMO stations are not always located inside or even really close (< 1km) to the soccer fields. In order to estimate the rainfall at a specific point (field), an interpolation method was performed based on the rainfall measured by the surrounding stations (TAHMO stations). Two interpolation methods were used, the cubic and the nearest neighbor. The cubic method was the default one, the nearest neighbor was selected for locations outside the interpolation grid.

The output of that part is the interpolated rainfall $P6d_{int}$ of six days at a soccer field, based on the observations of specific stations.

To sum up, each row of the final dataframe of each soccer field contained seven columns. The first one (the index), is the date that the satellite passed over the soccer field. The rest are the backscatter signal and its variation for two consecutive dates (the date of the index and the next date), the accumulated rainfall between them, the incident angle of measurement, the quadrant of the row (based on the two signal values) and the interpolated rainfall to the soccer field. The final dataframe will be the input data of the stacked classifier.

2.3.5. Machine Learning

2.3.5.1. Stacked Classifier

The selection of a stacked classifier corresponds to the fourth research question. The approach of generalization (stacked classifier) is firstly introduced by Wolpert (1992). It is an ensemble technique that combines results from multiple classifiers to do the final decision/prediction. According to Wolpert (1992), the advantage of that method is that it reduces the error (bias) of the individual classifiers and in many cases, it outweighs the use of a single classifier (Džeroski and Ženko, 2004).

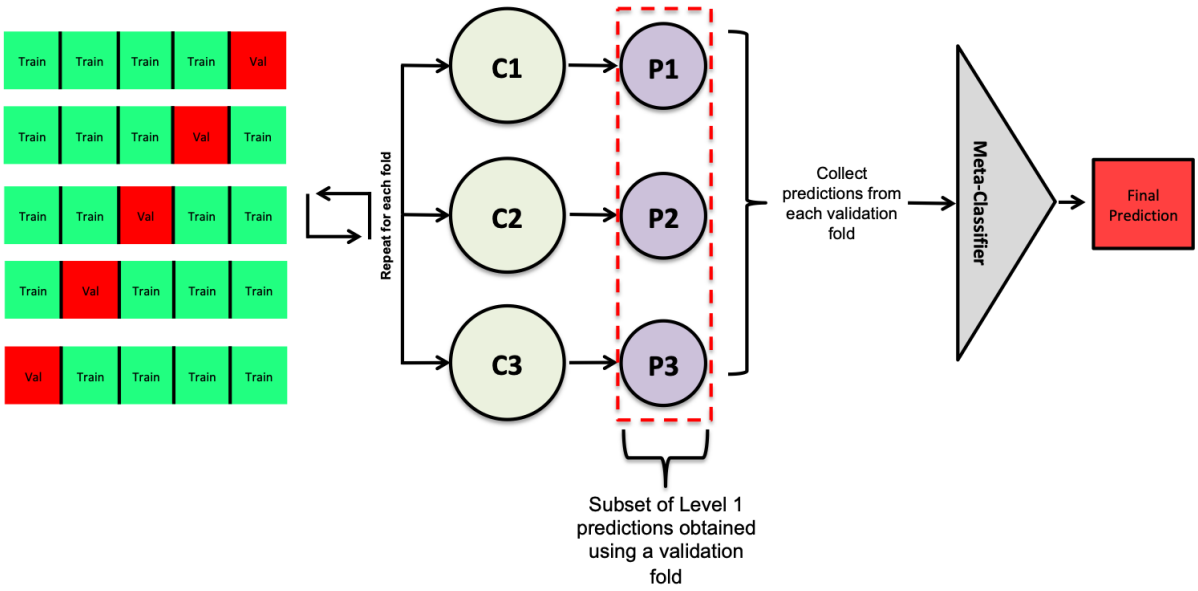

Figure 2.11: The structure of a stacked classifier. The matrix before the classifiers represent the Stratified k-fold procedure to train each level0 classifier. (source:Towards science)

For each stacked classifier, there are two classifiers, level-0 and the meta-classifier. It can be seen in the Figure 2.11 that the level-0 classifiers (C1, C3, etc.,) are doing inner predictions and then the level-1 classifier (meta-classifier) uses these inner predictions (ensemble method) to generate a final one. Data used for the prediction of level-1 was excluded from the training of the classifiers. A Stratified k-fold (stratified stands for the fact that the folds are made by preserving the percentage of samples for each class) cross validation methodology was applied. The training data of each classifier is split in a number of parts, and then, all parts except for one are used for the training and this part is used for the prediction (times the number of splits). This procedure is done for every classifier, so, the timeseries of the predictions of each level-0 classifier were fed to the meta-classifier to reach the final predictions.

For the specific study, five base classifiers and a meta-classifier were used. All of them are following the principles of Supervised Learning (SL).

- Logistic Classifier (Level-1)

1. K-Neighbors Classifier (Level-0)
2. C Support Vector Machine Classifier (Level-0)
3. Decision Tree Classifier (Level-0)
4. Random Forest Classifier (Level-0)
5. Multi-Layer Perceptron Classifier (Level-0)

2.3.5.2. Logistic Classifier

The Logistic classifier is a probabilistic classifier and the main principle behind it is the fact that a sigmoid curve (range 0-1) is fitted on the data.

$$y_{sigmoid}(x) = \frac{1}{1 + e^{-x}} \tag{2.6}$$

At the next step, the probability (y value) of the new data was calculated. Values with a probability higher or equal to a threshold (e.g., 0.5) were assigned with 1 and less than the same threshold are assigned with 0, thus the logistic classifier is calculated by the logistic regression (Bharadwaj et al., 2021).

In order to achieve the best fitting of the data while avoiding overfitting, the cost function also contains a term for regularization (minimum parameter values). Eq.2.7 represents the cost function and the first term is the ridge regularization which is necessary to prevent the overfit. The selection of this criterion instead of the typical ordinary least squares (OLS) lays on the fact that the ridge regularization can face the possible problem of co-linearity of the variables (Hastie, 2020). For the optimization of the cost function, the Broyden – Fletcher – Goldfarb – Shanno methodology is applied (”lbfgs”). The ”lbfgs” method is part of quasi-Newton method and it is robust for small and unscalled datasets (Daume III, 2004).

$$E = min_{w,c} \cdot \frac{1}{2} \cdot w^T \cdot w + C \cdot \sum_{i=1}^{n} log(e^{-y_i \cdot (X_i^T \cdot w + c)} + 1) \tag{2.7}$$

where E: cost, w: weight of the sigmoid parameters’ fit and C: inverse regularization strength (C = 1).

2.3.5.3. K-Neighbors Classifier

The K-Neighbors Classifier (KNN) is called a lazy learner algorithm, which means that it does not generate a model based on training data but it stores the training data. In the prediction part, the algorithm searches and matches the input data with similar/same instances of training data and then it predicts based on the votes of a number of neighbors. The most significant parameter of K-Neighbors Classifier is the parameter k, which is an integer which corresponds to the number of voters (neighbors) that the classifier will take into account for a prediction. The higher the k value, the more probable is the underfitting and the higher the value is, the more probable is the overfitting (Bharadwaj et al., 2021). For the specific study, the number of the voters was set to 5.

2.3.5.4. C Support Vector Machine Classifier

The Support vector machine (SVM) classifier is also characterized as an instance algorithm (lazy learner). This algorithm fits a hyperplane that achieves the maximum margin between the unique classes of the training data. More analytically, a hyperplane is a subspace of n-1 dimensions (n are different labels) based on a n-dimensional space (Alex J. Smola and Bernhard Scholkopf, 2004). In a given dot product space, a hyperplane can be described as:

$$[w] \cdot [x] + b = 0 \tag{2.8}$$

where: $[w]$ is the orthogonal vector to the hyperplane.

The goal of the SVM algorithm is the minimization of the orthogonal vector between all the dimensions:

$$E = \frac{1}{2}|||[w]|||^2 \tag{2.9}$$

The implementation of this algorithm includes a regularization process using the cost function Eq.2.7 with inverse regularization strength C = 1. The training of the algorithm is using the kernel Radial Basis Function ('rbf') with a gamma value dependent on the scale. A high value of C leads to detailed surfaces that are increasing the probability of overfitting, thus a C > 0 SVM classifier is also called a soft margin SVM; the gamma value expresses the weight of each individual training.

$$k_{rbf} = exp(-\gamma \cdot ||x - x'||^2) \tag{2.10}$$

where $||x - x'||^2$: is the euclidean distance between two vectors and

$$\gamma = \frac{1}{n \cdot var(x)} \tag{2.11}$$

where n: is the unique number of the labels and $var(x)$: is the variance of the input data

2.3.5.5. Decision Tree Classifier

The Decision Tree is a decision tree algorithm and is a very common concept in cases where classification is needed. The methodology that this algorithm follows is simple and can be represented as a flowchart (in the shape of a tree) consisting of nodes that are the classification conditions/features. The first node of the tree is called the root node and contains the most significant feature of the data (is selected based on the information gain criterion that will be explained further) and ending up with nodes that are called leaf nodes and are representing the different classes (the results). When it comes to the internal nodes, they are features created based on the information gain criterion. Considering that different conditions/features will lead to different results and different classification, for the selection of a feature at each step, the impurity of the results (using that specific feature) is calculated. The impurity of the feature is quantified using the Gini index (2.12) which expresses the proportion of right answers based on a specific feature. In cases where there are multiple subsets, the Gini index of the row is the weighted index of the particular subset. During the training, the decision tree is split and expanding the internal nodes (features) until a node is pure (based on the labels of the data) (Bharadwaj et al., 2021).

$$G = 1 - \sum_{i=1}^{n} p^2 \tag{2.12}$$

where p: is the proportion of samples that are contained in a specific class

2.3.5.6. Random Forest Classifier

Random Forest (RF) is characterized as an ensemble classifier; the methodology of that classifier is to combine multiple and independently trained decision tree classifiers. The result of the random forest classifier is the average probabilistic prediction of the individual decision tree classifiers. There is the option to define the number of the trees of the "forest", the more trees, the more probable is the overfitting. A trial and error procedure for the performance of the classification set was used to select 1000 trees for the random forest classifier. It is important to mention that the bootstrap option is enabled for the building of the decision trees. This option means that a different part of the whole dataset was used to train each decision tree (Bharadwaj et al., 2021).

2.3.5.7. Supervised Artificial Neural Network/Multi-Layer Perceptron Classifier

An Artificial Neural Network (ANN) algorithm mimics the structure and the operation of a biological neural system in the brain of organisms. The three basic parts of the structure of a biological neuron are the dendrites that are providing a neuron with data, the neuron's body which is processing the information and then, the axis which transmits the signal to a next neuron. When it comes to the ANN, it is structured by three layers, the input layer, the hidden layers and the output layer. The input layer contains the different features that are connected to the neuron of the hidden layer. Each connection has a specific weight (w); the neuron of the hidden layer is fed with the summary (z) which is calculated by multiplying the value of each feature with its weight (McCulloch and Pitts, 1943):

$$z = \sum_{i=1}^{n} X_i \cdot w_i \tag{2.13}$$

where n: is the number of the features connected to a neuron, X_i: is the value of each feature and w_i: is its predefined weight

In order to reach to a result, an activation function is applied to the summary z. For the specific study, the logistic function (2.7) is used:

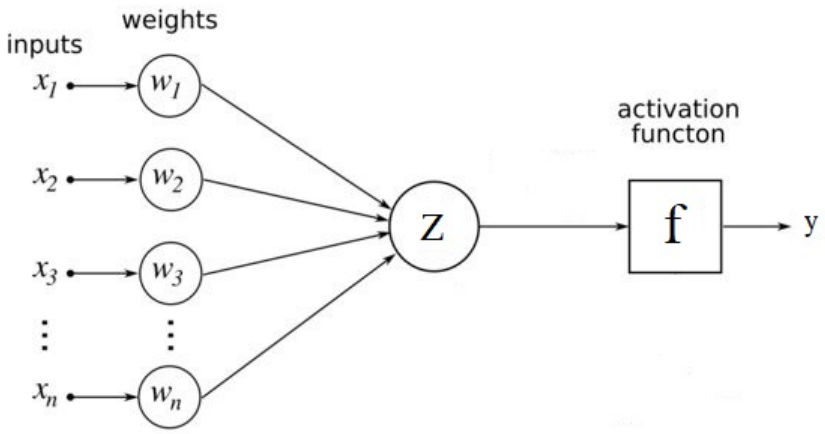

Figure 2.12: Graphical representation of a Multi-layer Perceptron (ANN) with one hidden layer. The features is the array x, the weights is the array w, the bias is the Z, the f is the activation function and y is the output. (Manzini, 2017)

For the evaluation of the result, a cost function is applied. The cost function is the Cross-Entropy which is optimized by a stochastic gradient descent (Bharadwaj et al., 2021).

$$E = -y \cdot ln(y_T) - (1 - y) \cdot ln(1 - y_T) + a \cdot ||W||_2^2 \tag{2.14}$$

where: $||W||_2^2$ is an L2-regularization term (Eq. 2.7) that penalizes complex models, y: is the estimation of the model and y_T: is the real value.

The stochastic gradient descent is an algorithm which is being used in order to find the lower possible value of the cost function. Thus, the steps of that algorithm are the following:

- Random selection of initial weights of the connections (low, close to zero)
- Calculation of the summary (z), application of the activation function and prediction of the result.
- Evaluation of the result based on the Cross-Entropy cost function
- Adjustment of the weights (back-propagation procedure) and recalculate error.

- Repeat the above steps, until the optimization of the cost function (based on the stochastic gradient descent criterion) is achieved.

The aforementioned methodology can also be extended for more than one hidden layer (Multi-Layer Perceptron Classifier (MLP)) which is a Deep Learning concept. In that case, the connection between hidden layers is set, thus each hidden layer is the input layer of the next hidden layer. The back-propagation of the error then adjusts the weight of the the connections between all hidden layers and the input layer with the hidden ones.

Finally, the structure of the stacked classifier is the following.

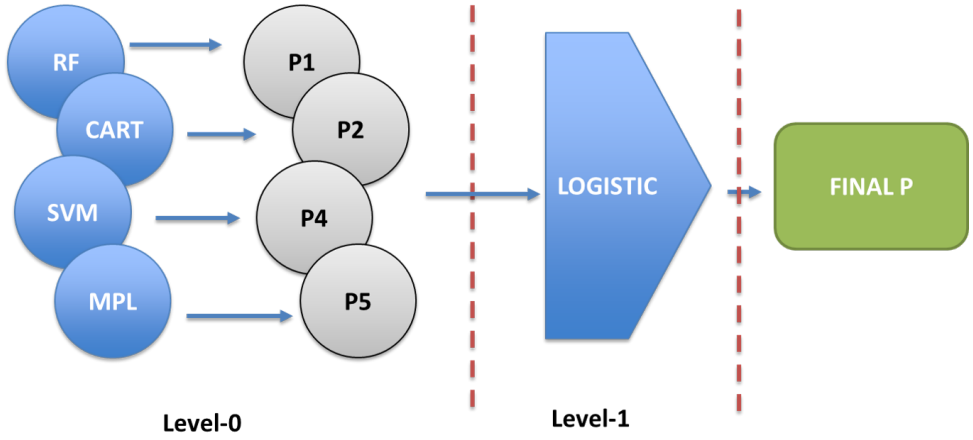

Figure 2.13: The structure of the stacked classifier used in that study. The blue bubbles contain the different level-0 machine learning classifiers and the silver bubbles their inner predictions. The big blue array contains the meta-classifier and the blue rectangle is the final prediction of the stacked classifier.

2.3.6. Different training & test datasets

The performance of the model trained on different training datasets and applied to different test datasets helped assessing the potential and the limitations of the stacked classifier. For that reason, four different training sets were used. The available data was both for the selected Ghanaian soccer fields and the Greek field for 2019 and 2020. The different training and test dataset selection will contribute answering the fifth and sixth research questions.

2.3.6.1. Leave-p-out

Before moving to the different combinations of datasets, it was worth to elaborate further on the Leave-p-out methodology.

A general assumption of the classification algorithms mentioned in Section 2.3.5 is the independence of the variables of the input data. The spatial variability of meteorological data is actually correlated. According to Pohjankukka et al. (2017), this spatial autocorrelation can cause biased results. In order to overcome that, a cross-validation procedure, in combination with the k-fold training of the classifier can improve the performance of the model.

The cross validation procedure followed the Leave-p-out cross-validation approach (Airola et al., 2019). In more detail, a p-number out of the total number of soccer fields (about 10%) was excluded from the training of the classification and afterwards, these n fields were used for the validation part. Also, the n fields were not used for the calculation of the interpolated rainfall.

In order to make the different validation datasets perfectly comparable, the same combination of excluded patches was used in the cases were this cross-validation method was implemented.

To achieve a product independent to ground observations or external data, the most training sets were containing data strictly from Sentinel-1. The standard parameters that were used as an input for the different training sets is: *VV_db, VV_db2, VV_var_db, VV_var_db2, Q and* θ. The last training set also considered the interpolated rainfall as an input.

2.3.6.2. Dataset 1
Multiple Ghanaian soccer fields (~50) were selected to be tested as rainfall detectors. The Leave-p-out cross-validation approach was applied to data from Ghanaian soccer fields during 2020. As a result of that, a stacked classifier was trained multiple times and then was checked on three test sets:

1. The excluded fields of each iteration (in Ghana) for 2020,
2. The total data for all soccer fields in Ghana during 2019 and
3. A random 30% proportion of the data for the Greek field

At each iteration, the F1 score was calculated individually for the excluded fields but also in total for all excluded fields (aggregated confusion matrix). This value was assigned as "All" and was calculated once for each combination.

The fact that there are multiple different combinations of Leave-p-out fields means that many of them were excluded more than once. The aim of the first testing set was to assess the performance of each trained stacked classifier and the overall classification in Ghana but also to check the stability of the performance of each field through the differently trained classifiers. For the second one, the generalization of each trained classifier to a different period (2019) assessed. Finally, the third dataset was used to evaluate the generalization of each training set to a different environment/space (Greece).

2.3.6.3. Dataset 2
At a further step, the classifier was trained with the total data from all Ghanaian soccer fields during 2020. Then, the performance of the stacked classifier was assessed at each single Ghanaian soccer field using the data from 2019.

The goal of that case was to evaluate the performance of each Ghanaian soccer field and assess the potential of using a specific soccer field as a rainfall detector in a different time period.

2.3.6.4. Dataset 3
The performance of the stacked classifier was also assessed for the Greek (AUTH) station. The data of the rainfall detector near AUTH station was split randomly to 70% and 30%. The stacked classifier was trained for the 70% and validated for the 30%. It is important to mention this specific testing set was the same used to validate the 1^{st} training set.

Using the Dataset 1, the stacked classifier was trained in Ghana and tested in Greece with a specific 30% of the total data. In this case, the stacked classifier was tested with the same 30% but it was trained in Greece with the rest 70%. The comparison of these two performances was used to evaluate the generalization of a pretrained stacked classifier in space.

2.3.6.5. Dataset 4
Apart from using exclusively airborne data, it was interesting also to investigate the performance of the stacked classifier including ground observation (interpolated rainfall) $P6d_{int}$. This parameter was calculated according to the methodology explained in Subsection 2.3.4.3. The Leave-p-out field approach was also used in that training set, thus different stacked classifiers were trained (based on different combinations) using data from 2020 for the Ghanaian stations and tested with the excluded data of each combination. It is important to mention that the excluded stations didn't participate to the calculation of the interpolated rainfall to them.

The training and testing datasets are summarized in the following table.

Table 2.2: Overview of the different training and test sets and their basic characteristics that were used to evaluate the performance of the stacked classifier

Training set	Training				Test		
	Variables	Area	Year	Training Fields	Area	Year	Test Fields
1	Standard *	Ghana	2020	Combinations of Leave-p-out	Ghana Ghana Greece	2020 2019 2020	Excluded-p All fields Greek 30% **
2	Standard	Ghana	2020	All fields in Ghana	Ghana	2019	Each field in Ghana
3	Standard	Greece	2020	One field in Greece 70%	Greece	2020	The rest 30% of the Greek
4	Standard & $P6d_{int}$	Ghana	2020	Combinations of Leave-p-out	Ghana	2020	Excluded-p

* Standard: VV, VV2, VV_var, VV_var2, Q and θ

** This Greek 30% is same to the validation part of the Training set 3

2.3.7. Assessment of the Stacked classifier

Regarding the sixth research question, the performance of different stacked or individual classifiers for rain/no rain prediction was assessed. The aforementioned procedure was performed using the 1^{st} training and testing sets.

Machine learning was used in different Earth Observation applications (fire detection, dust detection, etc.) by deploying single classifiers (SVM, CART and ANN). The combination of multiple single classifiers (5 base classifiers) into a stacked one was assessed in this study. Aiming to evaluate the performance of a stacked classifier compared to single ones, the stacked classifier was replaced by a single classifier. Then, training set 1 was used to train multiple times (based on the defined combinations) the individual classifiers (for Ghana 2020) and then validate them for the excluded Ghanaian stations during 2020. The whole procedure was followed for the specific classifiers: RF, CART, LOGISTIC, KNN, SVM and ANN.

Then, the selection of the specific combination of base classifiers was assessed based on the same training and testing sets as above.

Each time, one of the base classifiers was excluded and the performance of the validation was calculated. As stacked classifier's performance, it was taken the F1 score of the aggregated confusion matrix at each iteration/combination.

2.3.8. Comparison with IMERG product

This part covered the last research question by comparing the results of the implemented stacked classifier with the performance of the IMERG (satellite rainfall) product. IMERG has a temporal resolution of 3-hours. Timeseries of IMERG rainfall for 2020 were retrieved for all soccer fields. This data was rescaled in 6-days resolution based on Sentinel's-1 signal for each field. Then, the performance of IMERG product to detect rainfall was evaluated.

2.3.9. Validation metrics

2.3.9.1. Linear regression

Pearson's correlation coefficient was used to evaluate linear regression models. The best value of that metric is 1 and the worst is -1.

$$R_{Sq} = \left(\frac{\sum_{i=1}^{N} \left(yb_i - \overline{yb_i} \right) \bullet \left(y_i - \overline{y_i} \right)}{\sqrt{\sum_{i=1}^{N} \left(yb_i - \overline{yb_i} \right)^2 \bullet \sum_{i=1}^{N} \left(y_i - \overline{y_i} \right)^2}} \right)^2 \tag{2.15}$$

where yb: is the observed value, y: is the modelled value and $\overline{y}$ and $\overline{yb}$ are the median values of the predicted and observed values respectively.

2.3.9.2. Classification

In order to assess the performance of the classification on the validation part, F1 score statistical criterion introduced by Olson and Delen (2008) is applied. The F1 score considers the overall results of the classification in means of the total number of positive and negative predictions based on true and false events. This classification leads to four different classes that construct a confusion matrix. The first word of each class corresponds to the evaluation of the prediction (right or wrong). The second word of each class is about the reality (happened/not happened). The following figure, graphically represents the classes of the confusion matrix.

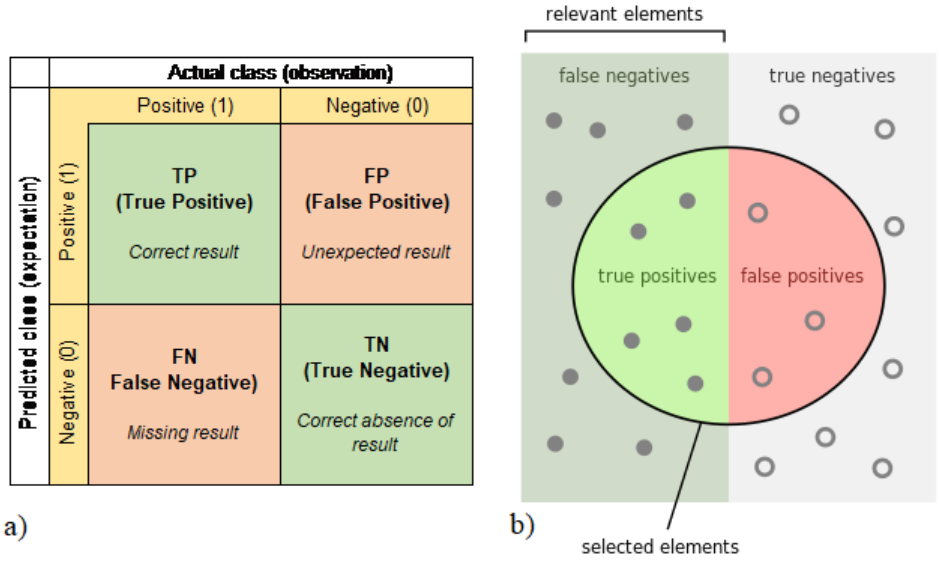

Figure 2.14: a) Classes of classification performance based on reality and the prediction. True and false correspond to the reality and positive and negative to the prediction (source:scikit-learn documentation) and b) Visual representation of the four recall classes in figure (source: Wikipedia)

The recall value of a classification is using the true positives (rainfall occurred and predicted) and the false negatives (rainfall occurred but not predicted).

$$recall = \frac{t_p}{t_p + f_n} \tag{2.16}$$

where t_p is the number of true positives and f_n the number of false negatives.

The precision value of a classification is using the true positives (rainfall occurred and predicted) and the false positives (rainfall occurred but predicted).

$$precision = \frac{t_p}{t_p + f_p} \tag{2.17}$$

where t_p is the number of true positives and f_p the number of false positives.

The F1 score, also known as balanced F-score or F-measure is calculated by the following equation. It can be seen that for its calculation, recall and precision values are necessary.

$$F1 = 2 \cdot \frac{precision \cdot recall}{precision + recall} \tag{2.18}$$

where t_p is the number of true positives and f_p the number of false positives.

3

Results

The results chapter contains four sections. The first one, Section 3.1, declares the valid TAHMO stations that successfully ”passed” the preprocessing procedure. The next one (Section 3.2) presents the rainfall detectors that were retrieved in Ghana and then Section 3.3 mentions the selected rainfall detectors in Ghana and Greece and the final, merged, data of each field. Afterwards, Section 3.4 analyses the results of the different train/test datasets assessed with the stacked classifier. Moreover, Section 3.5 evaluates the selection of the specific classifier instead of individual ones and instead of a different stacked classifier and finally, Section 3.6 checks the performance of the stacked classifier with the IMERG final product.

3.1. Filtering TAHMO stations

The Ghanaian TAHMO network consists of 101 stations. However, many of them malfunctioned for long (weeks, months, etc.) or short periods (hours, days, etc.). A preprocessing procedure of filtering the problematic TAHMO stations out was necessary to answer the research questions.

After the preprocessing procedure described in Section 2.2.5, the 48 out of 101 stations were dropped and 53 stations were used for the study. Most of the stations are located in the southern part of Ghana; their spatial distribution follows the distribution of the population throughout the country. The valid stations are presented in the Figure 3.1.

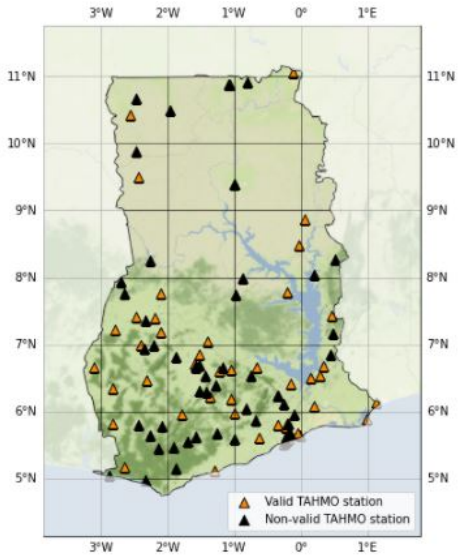

Figure 3.1: The map depicts all the Ghanaian TAHMO stations. The orange ones are the valid and the black ones are the non-valid.

Not only the selection of suitable soccer fields as rainfall detectors, but also the validation of the rain/no rain algorithm will use these valid stations.

3.2. Soccer fields in Ghana

To cover the first research question, the selection and retrieval of soccer fields was necessary. More analytically, a Python tool was implemented to reach the soccer fields over Ghana (using the Python library ”OSMnx”).

In total, 1450 fields were identified and retrieved across Ghana; their locations can be seen in the following figure. The most of the pitches are identified inside and near urban/populated areas, thus they are not equally distributed in Ghana.

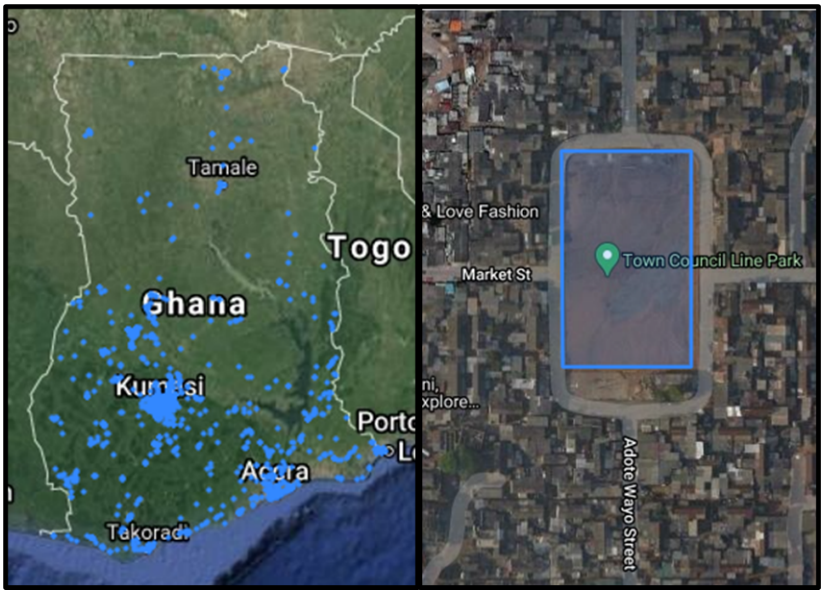

Figure 3.2: The total fields retrieved over the country of Ghana using the Open Street Map (OSM) API and an example of a random soccer field.

In the next step, the suitability of the soccer fields over Ghana to be used as rainfall detectors was tested. The condition of suitability of each station was determined using the Pearson correlation between its modelled soil moisture (of the closest to soccer field TAHMO station) and its backscatter signal from Sentinel-1. Regarding the lack of observed soil moisture both for the Ghanaian TAHMO stations and the Greek station, a simple lumped model was implemented to estimate daily soil moisture based on daily rainfall and evaporation.

3.3. Selection of rainfall detectors and data preparation

3.3.1. Calibration of hydrological model

In the context of the second research question, the estimation of modelled soil moisture was essential for the selection of suitable soccer fields.

The lumped hydrological model described in Section 2.3.2 was taking as input the daily precipitation and reference evaporation to estimate the daily soil moisture content. The specific model was applied at each TAHMO station. The soil moisture of each field was assumed equal to the soil moisture of the closest in-situ network. For the calibration of the initial and maximum bucket storage parameters, a trial and error method was performed (based on the correlation of modelled soil moisture with the retrieved backscatter signal). The calibrated values were equal to S_{in} = 25 mm and S_{umax} = 100 mm. Figure 3.3 depicts the modelled soil moisture against the rainfall of station TA00260; it can be seen that the behavior of the modelled soil moisture is sensitive to rainfall events.

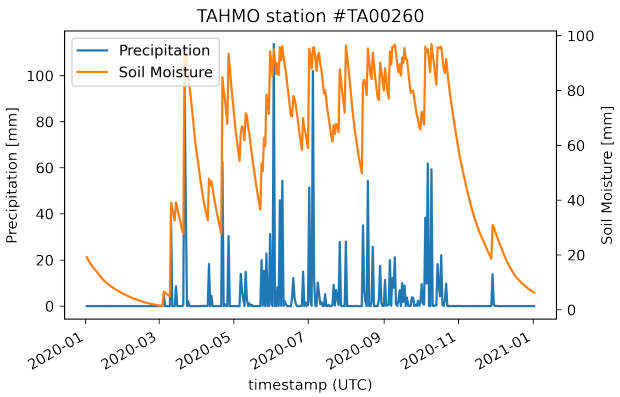

Figure 3.3: The modelled soil moisture against the observed precipitation of station TA000260 during 2020.

The same hydrological model was implemented for the Greek station, in order to calculate the modelled soil moisture of the station and select a suitable field in Greece.

The modelled soil moisture was calculated for all ground stations. Then, the soil moisture of each soccer field was assumed as equal to the modelled soil moisture of the closest in-situ station (both in Ghana and Greece). At the next step, the backscatter signal was retrieved for all soccer fields and correlated with their modelled soil moisture. That was the condition to select the suitable soccer fields.

3.3.2. Selection of rainfall detectors and preparation of the final data

This part is the direct answer to the second and third research questions regarding the selection of suitable soccer fields and the preparation of the final data of each field.

3.3.2.1. Greek rainfall detector

The case of Greece was simple, since observations from only one station were available, thus, one field was identified and selected. The suitable field, that is located near AUTH station, was manually detected based on the the aforementioned correlation of its modelled soil moisture and its backscatter signal. It is located about 1 km from the station, near the airport ”Macedonia”.

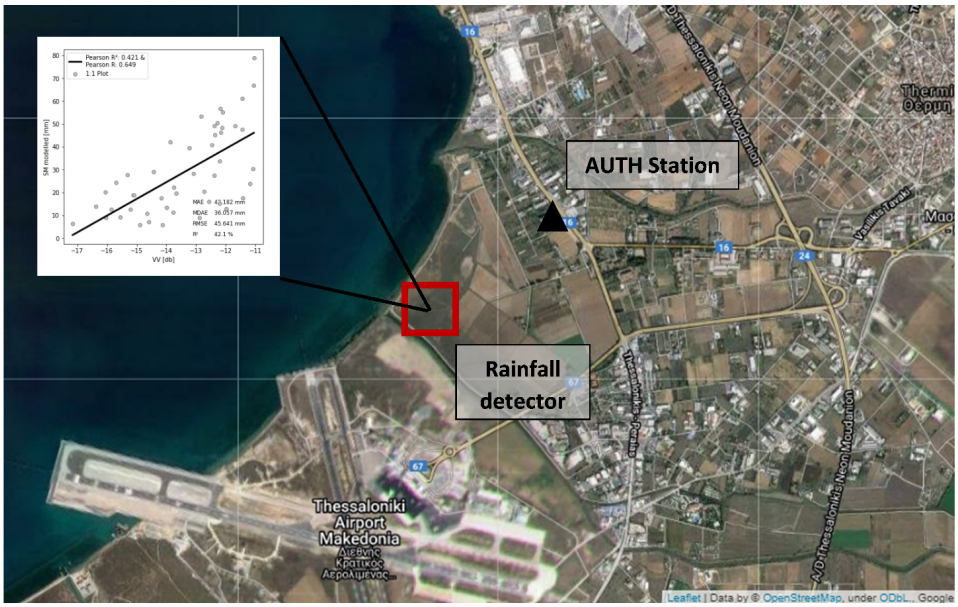

Figure 3.4: The modelled soil moisture against the backscatter signal of the selected field near station AUTH and its position relatively to the AUTH station

3.3.2.2. Ghanaian rainfall detectors

For the case of Ghana, about 50 soccer fields were selected since there were 53 valid TAHMO stations over the country. The selection of soccer fields is not a simple matter. A general problem is that the response/behavior of the backscatter signal over these fields, after rainfall events, varies from place to place. Also, many of the stations are located in towns and villages and the retrieved signal of urban areas is affected by the built-up areas and different land uses (Signal resolution is 5x20 m). For that reason, it was necessary to manually narrow down the retrieved soccer fields and select the suitable ones based on the Eq. 2.15. The threshold of the minimum R^2 was set manually to 50%, after a trial and error procedure.

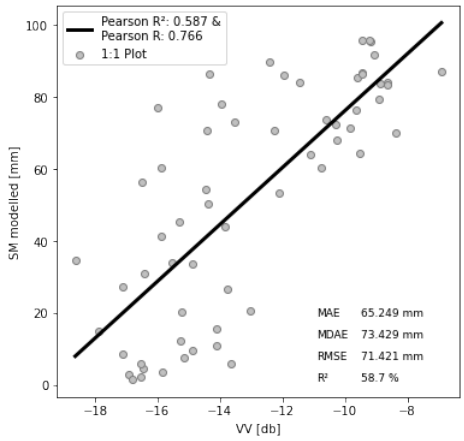

Figure 3.5: The modelled soil moisture against the obtained backscatter signal for a field near station TA000260 during 2020. The modelled soil moisture corresponds to the soil moisture of the closest TAHMO station.

The location of the soccer fields was also an important factor for their selection in order to achieve a well distributed rainfall detectors' network. After a careful inspection, 50 different fields across Ghana were selected as rainfall detectors. The distance of each soccer field to the closest TAHMO station is included in the right part of the Figure 3.6. The median distance is 3.8 km, the minimum is 2.5 m and the maximum is almost 76 km. The location of the fields and the location of TAHMO stations are included at the left part of same figure.

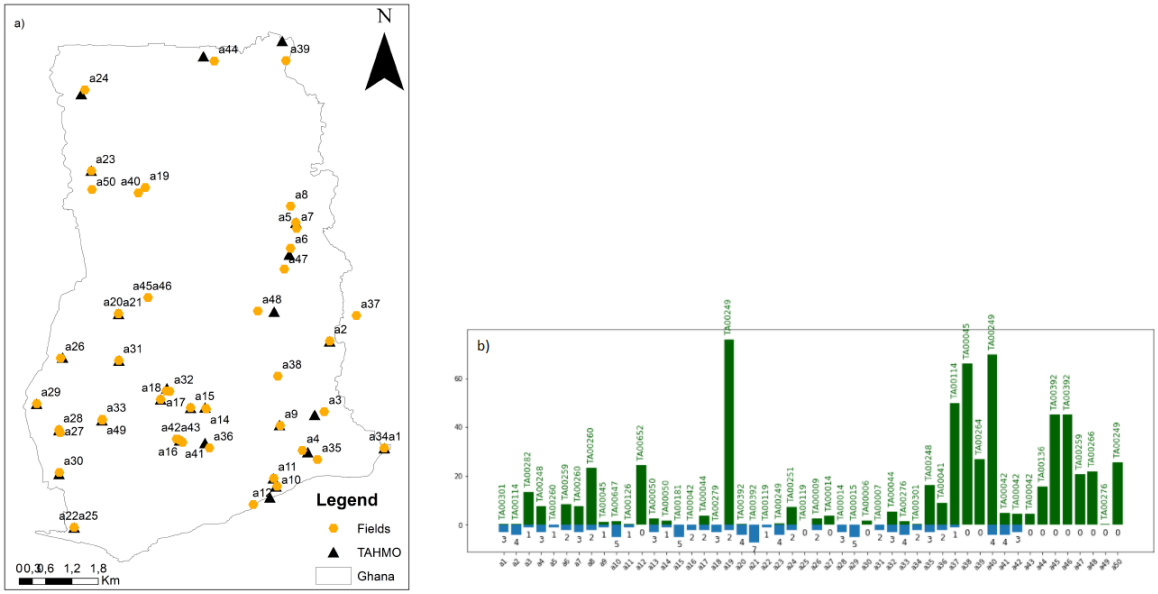

Figure 3.6: a) Location of selected soccer fields in Ghana and TAHMO network and b) the distance of each field to the closest TAHMO station

All the selected Ghanaian and Greek fields contained data from two different sources (satellite data with 6-days temporal resolution and in-situ data with 1-day temporal resolution). It was necessary to rescale the daily data to fit the low temporal resolution. At the next step, the merged data together with two new calculated variables will be presented.

3.3.2.3. Preparation of the final data

The final data of each soccer field was consisted of seven columns; the index contained the dates that satellite passed over the field. The first five columns were the spaceborne data (backscatter signal and its variation and the incident angle of measurement) corresponded to the date of the index and the next week's data. The last two columns were additional variables (quadrant of measurement and interpolated rainfall) to enhance the performance of the classification.

More analytically, the backscatter signal, its variation and the incident angle of measurement were retrieved for each field. The goal of the classifier was to assess rain/no rain based on consecutive satellite measurements, so, the data was formatted in a more convenient format. To be more specific, each field's row of data contained spaceborne data for two consecutive dates (current and next week) and the accumulated rainfall between them. The following figure contains the modelled soil moisture, the observed rainfall and the measured backscatter signal for a soccer field close to station TA000260. It can be seen that during and after a rainy season, the backscatter signal was increased and then, after significant rainfall events, it dropped.

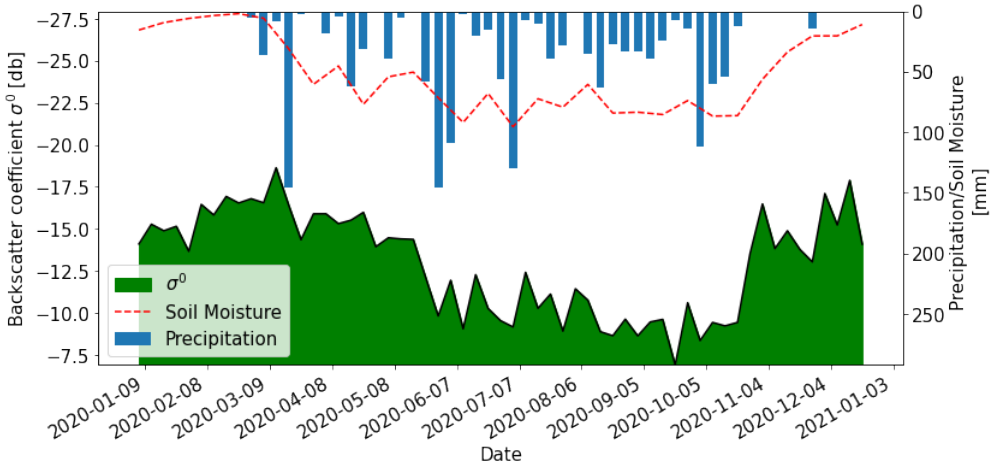

Figure 3.7: The modelled soil moisture, the precipitation and the backscatter signal of a field near station TA000260 in 2020. The soil moisture corresponds to the modelled of the closest TAHMO station.

Apart from the spaceborne data, two more variables were calculated for each suitable soccer field, the quadrant of measurement and the interpolated rainfall. Using the two backscatter signals of each row of all soccer fields, their median values were calculated. Then, based on these median values, each pair of backscatter signal values at each row of the data was classified into a quadrant. The described classification was applied for the whole timeseries of all soccer fields both in Ghana and Greece. An example is following; the gray scatter represents the present week's (x-axis) and previous week's timeseries (y-axis) of all 50 fields. The red scatter is one sample point that was classified in quadrant 1. The density of the points is higher in the 1^{st} and 3^{rd} quadrants.

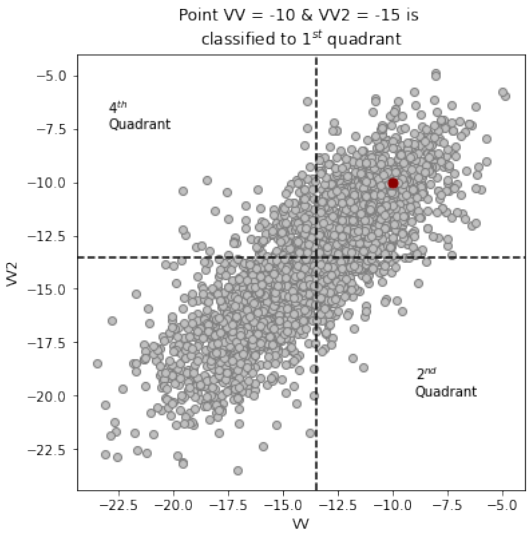

Figure 3.8: Visual representation of the selection of the quadrant of two consecutive weeks' signal (red dot) based on historical timeseries of the present and previous week's backscatter signal of all 50 soccer fields (gray dots). The black cross represent the median value of VV_db and VV_db2 of all fields.

Generally, the location of TAHMO stations is not always very close/inside the closest soccer field. For that reason, an interpolated rainfall was also calculated at each field based on the observations of the surrounding TAHMO stations. The interpolated rainfall was calculated differently for different training data. Thus it was calculated multiple times for the different training fields' combinations of the last training set.

The outcome of that part was used as input data for the stacked classifier. The performance of different validation sets based on different training sets was assessed.

3.4. Performance of the different training/test datasets

This section corresponds to the research questions 5. The performance of a stacked classifier to predict rain/no rain based on data from Sentinel-1 was assessed together with the potential of generalizing the algorithm both in time and in space.

Before training and testing the stacked classifier, it was necessary to define the different combinations of soccer fields for the Leave-p-out cross-validation approach. Regarding this approach, the stacked classifier is using 45 out of 50 stations for training and the rest five for testing. Ideally, the performance of each soccer field should be computed based on all different combinations of 45 soccer fields. The fact that 50 soccer fields can be split into million combinations of train-test datasets means that a massive computing power is needed to check them all. Concerning to the computational time, 20 random different combinations of soccer fields were generated and they are included in the following table. In cases where Leave-p-out fields cross validation approach was applied, except for the F1 score of the individual excluded stations, one aggregated F1 score for all the excluded stations was calculated for each combination. This score was called 'All' for simplicity.

Table 3.1: The 20 different combinations of Leave-p-out fields selected to evaluate the stacked classifier models. The five fields at each row, are the (excluded) fields that were used for testing the stacked classifier. The training test is contained of the rest 45 stations.

Combination	Excl. Field 1	Excl. Field 2	Excl. Field 3	Excl. Field 4	Excl. Field 5
1	a8	a10	a15	a18	a29
2	a15	a21	a33	a36	a37
3	a1	a2	a6	a31	a32
4	a4	a6	a13	a28	a40
5	a10	a15	a17	a26	a34
6	a7	a16	a20	a29	a35
7	a26	a32	a33	a36	a42
8	a4	a21	a23	a29	a41
9	a19	a24	a28	a33	a40
10	a2	a18	a31	a33	a41
11	a5	a7	a21	a41	a42
12	a10	a14	a16	a20	a21
13	a9	a13	a18	a23	a29
14	a1	a19	a21	a34	a40
15	a8	a13	a20	a23	a40
16	a2	a7	a10	a21	a23
17	a1	a2	a4	a17	a20
18	a3	a15	a21	a24	a35
19	a10	a22	a35	a41	a42
20	a11	a15	a28	a29	a32

3.4.1. Dataset 1

The first dataset contains data for all 50 soccer fields in Ghana. Following the combinations mentioned in 3.1, the stacked classifier was trained 20 different times with 45 soccer fields each time.

The first test dataset contains the excluded soccer fields of each combination. Many of the stations participated more than once in combinations, thus their performance is calculated more than once. The performance (F1 score) of each different soccer field is depicted in boxplots. The median 'All' (aggregated) F1 score of all different combinations was about 65%. Also, there are individual fields with median performance more than 95% and other with median F1 score close to 0%.

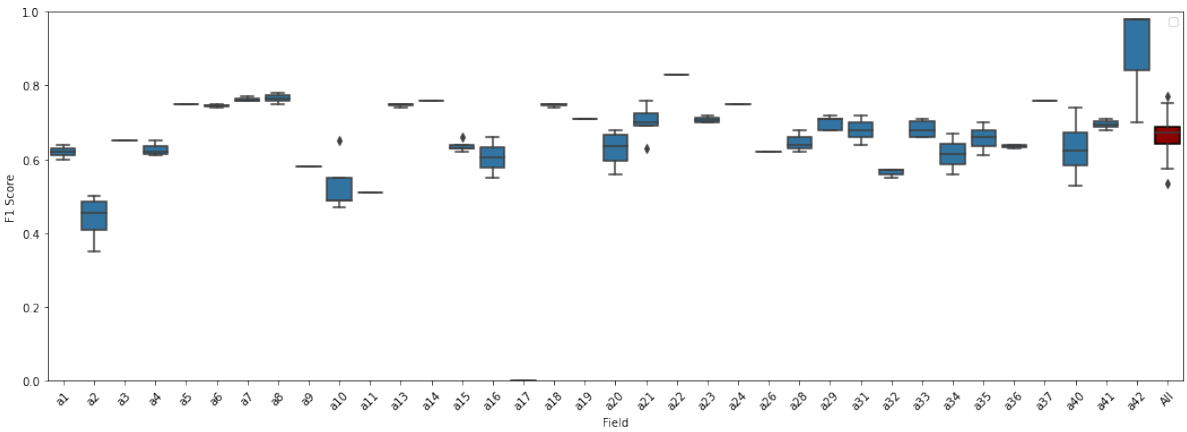

Figure 3.9: Performance (F1 score) of Ghanaian fields trained/tested with the different Leave-p out combinations for 2020 data. Many of the fields are participating only in one combination, thus their performance is one number (no boxplot).

The following three figures depict the minimum, maximum and median F1 score of each station for the 1^{st} validation dataset. The color-range of each figure is determined by the minimum and the maximum value in each case in order to provide a better visualization of the spatial variability. The results of the first validation dataset (p-fields excluded from each combination), showed that the performance

of the classification was better in the southern part of the country.

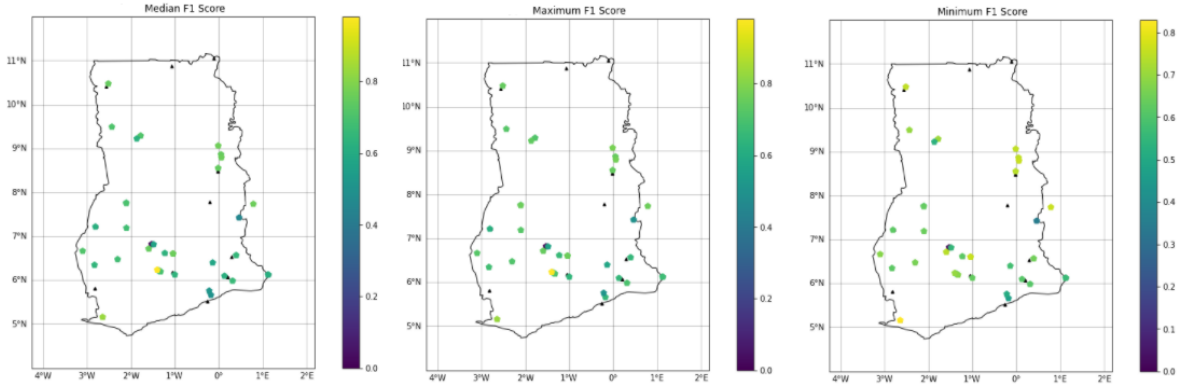

Figure 3.10: Spatial distribution of the performance of Ghanaian fields trained/tested with the different Leave-p out combinations for 2020 data. The closer to yellow are the best values

Except for the excluded stations, the stacked classifier is tested with two more datasets. For each Leave-p-out fields training, the stacked classifier is also tested on a specific proportion of data from the Greek station and on the data of all 50 soccer fields in Ghana during 2019. The second boxplot contains the performance of the stacked classifier trained on the Greek data and the third boxplot the performance to the Ghanaian data in 2019, after all 20 combinations. The first boxplot represents the aggregated ('All') F1 score of the stacked classifier tested in the excluded stations, after all 20 combinations (same to the red box of Figure 3.9). It can be seen that the stacked classifier that was tested on the excluded Ghanaian soccer fields during 2020 had the best performance and the testing in a totally different environment (Greece) had the lowest performance with the largest distribution.

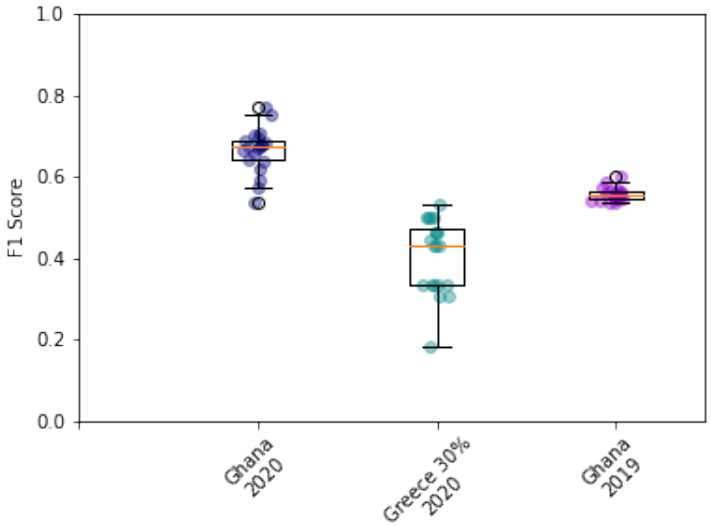

Figure 3.11: The performance of the stacked classifier trained with the 20 different combinations of Training set 1 and tested on i) The excluded p-stations in Ghana at 2020, ii) The same 30% of Greek station's data and iii) Data from all Ghanaian station at 2019

3.4.2. Dataset 2

The second training dataset consisted of data for all Ghanaian fields during 2020. The performance of the classifier trained on the 2^{nd} dataset was assessed for the data of each Ghanaian soccer field in 2019. The median performance of classification is about 60%, but there are also outliers close to 0%. When it comes to the spatial distribution of the performance, no evident pattern was detected.

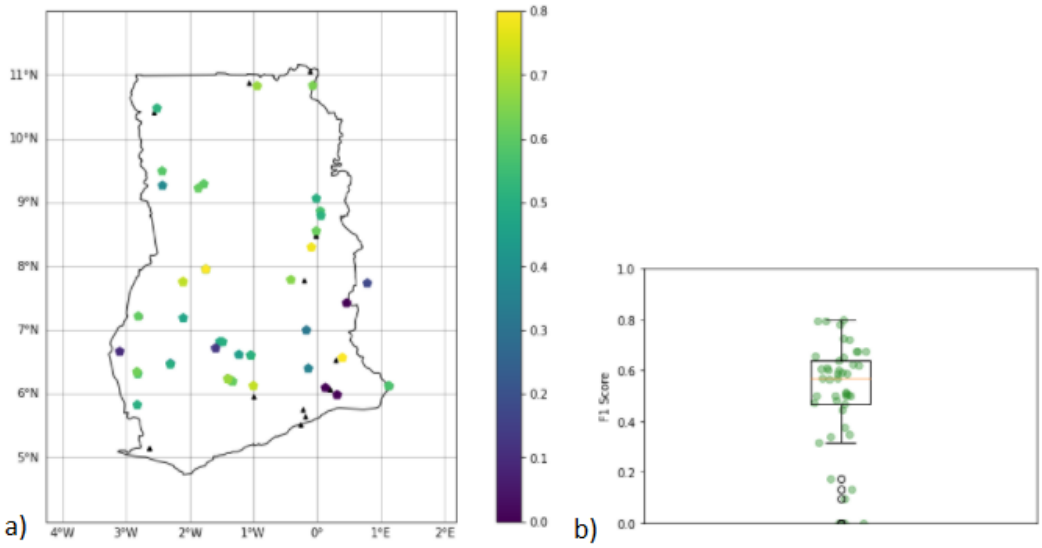

Figure 3.12: a) Spatial distribution and b) Boxplot of the F1 scores of the model trained with the 2^{nd} dataset and b) tested on each Ghanaian field in 2019

3.4.3. Dataset 3

The stacked classifier was assessed in a totally different environment, which is Greece. Keeping the validation set that was used for the validation of the training dataset 1 (30% Greek), the classifier was trained on the rest 70%. The F1 score of that case was about 57%.

3.4.4. Dataset 4

The sparse in-situ network in Ghana gives daily, very frequent measurements. The observations at specific locations, can be interpolated to estimate the rainfall on different locations. The effect of the interpolated rainfall (on soccer fields), on the performance of the stacked classifier was assessed. The training/testing datasets are the same as the first training/test datasets that were used in the Section 3.4.1. The classifier in this case was trained and tested 20 times based on the Leave-p-out field combinations of the Table 3.1. The only difference in that case is that each field contains an additional parameter, the interpolated rainfall.

So, as it was mentioned, the preliminary work for that model was the estimation of the interpolated rainfall for each excluded station. The specific calculation was done also for every combination of excluded stations because the training set was different each time. The interpolated rainfall of the excluded soccer fields was calculated without considering the observed rainfall on them.

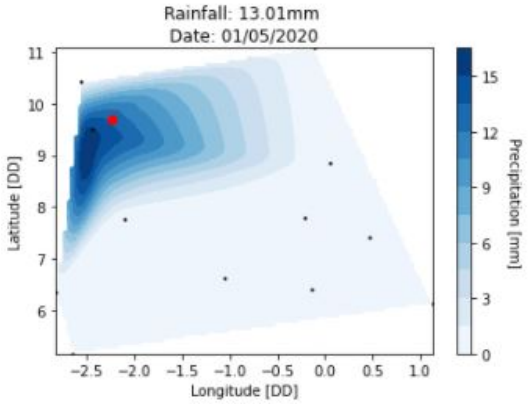

Figure 3.13: The cubic interpolated rainfall amount in Ghana. Date: 01/05/2020. The interpolated rainfall (13.01 mm) corresponds to the red dot. The black dots are the stations that are training the interpolation. (The figure is not scaled or transformed to any coordinate system)

The median F1 score of the different combinations is about 85%, there are values at 100% and other at ∼50%. The results showed that the performance of the stacked classifier, at each soccer field, trained with the additional parameter of the interpolated rainfall are better for most of the fields. The disadvantage though is the fact that the performance of each field is more sensitive compared and develops larger distribution.

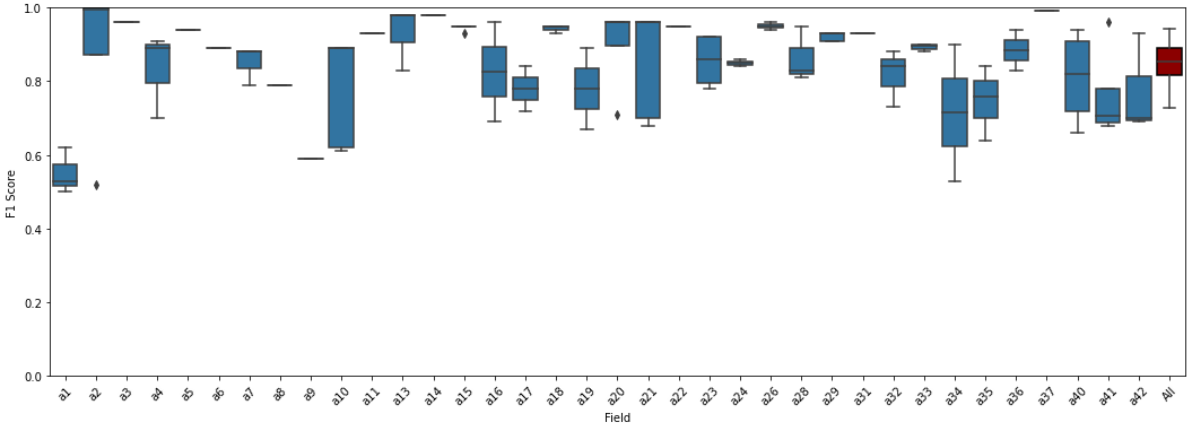

Figure 3.14: Performance (F1 score) of Ghanaian fields trained/tested with the different Leave-p out combinations for 2020 data including the interpolated rainfall. Many of the fields are participating only in one combination, thus their performance is one number (no boxplot).

The following figure illustrates the spatial distribution of the F1 score at each soccer field. Comparing with the Figure 3.10, it can be seen that the performance of the stacked classifier that used the interpolated rainfall is more equally distributed in Ghana,

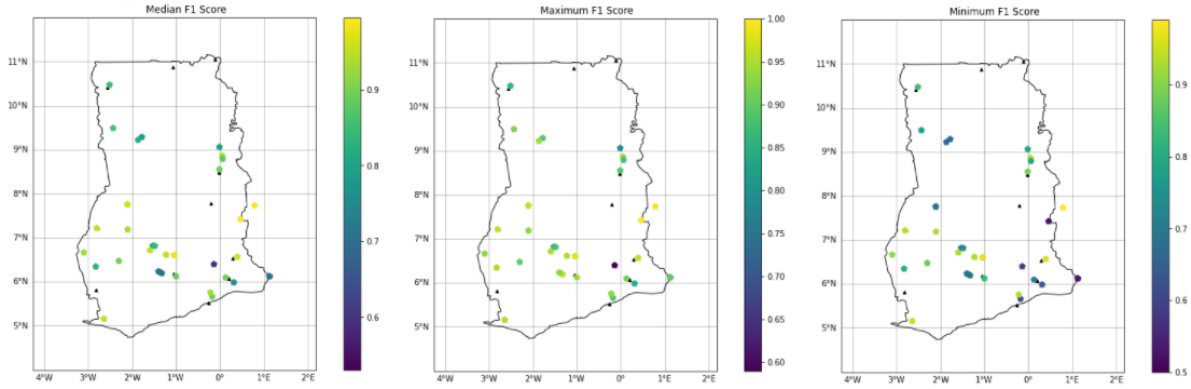

Figure 3.15: Spatial distribution of the performance of the model, trained with the 4^{th} dataset and tested with the excluded fields at each combination.

The following table contains information about the input, the calibration data and the results of each case. In the case where stacked classifier is trained and tested with the Leave-p-out cross-validation approach, the F1 score that is contained in the table is the median 'All' F1 score of all combinations.

Table 3.2: Overview of the different training and test sets and their basic characteristics that were used to evaluate the performance of the stacked classifier. Also, the performance of the different test datasets.

Training set	Training				Test			Median F1 Score **
	Variables	Area	Year	Training Fields	Area	Year	Test Fields	
1	Standard *	Ghana	2020	Combinations of Leave-p-out	Ghana	2020	Excluded-p	0.67
					Ghana	2019	All fields	0.43
					Greece	2020	Greek 30% **	0.55
2	Standard	Ghana	2020	All fields in Ghana	Ghana	2019	Each field in Ghana	0.57
3	Standard	Greece	2020	One field in Greece 70%	Greece	2020	The rest 30% of the Greek	0.57
4	Standard & $P6d_{int}$	Ghana	2020	Combinations of Leave-p-out	Ghana	2020	Excluded-p	0.85

* Standard: VV, VV2, VV_var, VV_var2, Q and θ

** This Greek 30% is same to the validation part of the Training set 3

*** The F1 Score is the median of "All"

3.5. Assessment of the Stacked classifier

Afterwards, the selection of the specific classifier was assessed (Research question 6). The model was trained and tested with specific datasets and the stacked classifier was replaced with an individual at each time. Next, the stacked classifier was replaced with a different stacked classifier, excluding one base classifier.

For that assessment, the first training/test datasets, with the Leave-p-out cross-valuation approach, were used. The 'All' F1 score of the 20 combinations is illustrated in the Figure 3.16a. The F1 score of the stacked classifier is almost 70% and it is very close to the performance of the single ones. Figure 3.16b contains the boxplots of the F1 score after removing one base layer classifier from the stacked one, at each time. The stacked classifier with all five base classifiers gave the best results. Apart from that, improvement was also found when investigating the base classifiers individually. For example, the F1-score of the field 'a17' without including random forest is 0% and using the stacked classifier is 30%.

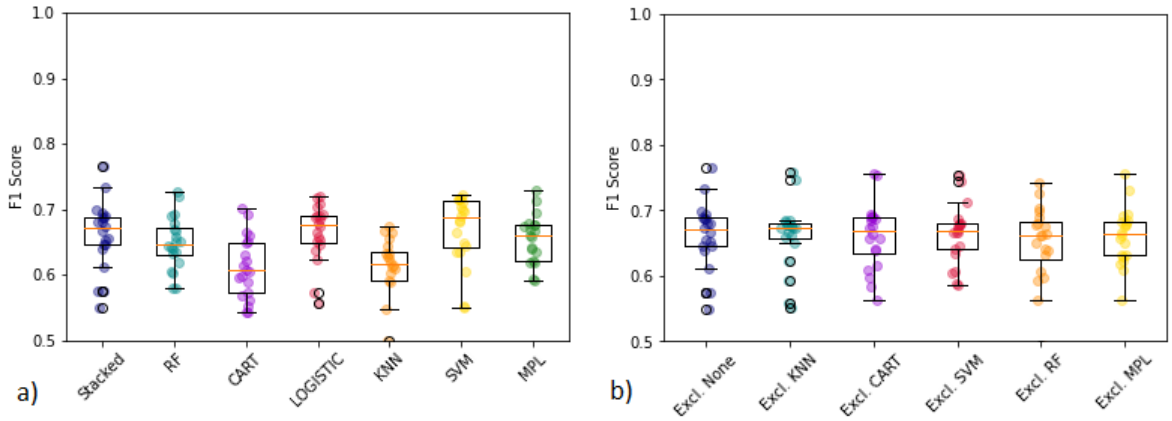

Figure 3.16: a) Boxplots of the total F1-score of each combination using a different single classifier at each time and b) Boxplots of the total F1-score of each combination excluding a different base layer classifier each time.

3.6. Comparison with IMERG product

Finally, the comparison of the performance of the stacked classifier with the IMERG product is evaluated (Research question 7).

Rainfall data from IMERG in 3-hours temporal resolution was retrieved for each soccer field in Ghana and Greece. Then, the data was resampled in 6-days temporal resolution to fit to the final data of each soccer field (that follows the date satellite observed). A threshold of 5 mm/6d was set as wet/dry day and the performance of prediction was calculated for IMERG product based on the ground-based observations. The median value of IMERG in Ghana is 72% and in Greece 80% (in Greece there is only one field, thus only one F1 score).

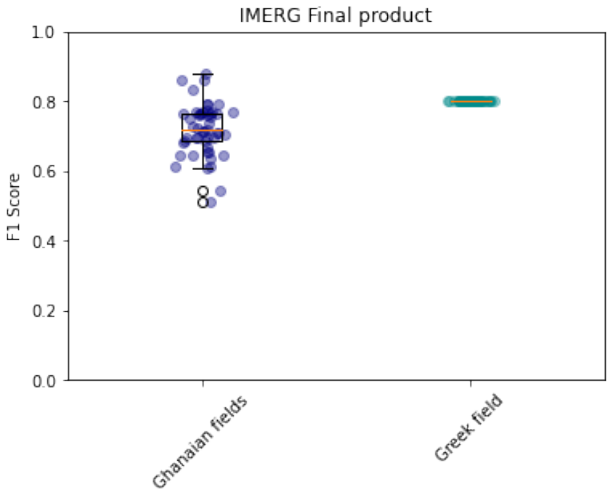

Figure 3.17: The performance of the IMERG product in the 50 rainfall detectors retrieving data of 2020 for Ghanaian and Greek fields.

4

Discussion

The methodology that was followed in Chapter 3, includes assumptions and uncertainties. The major limitations are discussed below. First, Section 2.2 describes the limitations of the two data sources, reliability of TAHMO network (Subsection 4.1.1) and satellite image resolution (Subsection 4.1.3. Then, the difficulties of rainfall detectors selection (Section 4.2) and the uncertainties of the hydrological model's calibration (Section 4.3) are analyzed. In the next Section 4.4, factors that account for the suitability of a soccer field are addressed. Section 4.5 concerns the limitations of the different train/test datasets and discusses about their results. In the last two sections, (Section 4.6 & Section 4.7) the selection of the specific classifier and the performance of the IMERG final product are discussed.

4.1. Data

4.1.1. TAHMO Network data

The TAHMO Network over Ghana consists of 101 stations and it is still growing aiming at reaching a density of one station per 1000 km^2. Observed (or real) precipitation is essential for the training of the stacked classifier, so, the measurements of TAHMO network during 2020 are further investigated.

4.1.2. Reliability of data

The operation of more than half of these stations is not regular. To be more specific, many stations are missing observations for more than 15 hours per day or even for many days, thus the problematic stations/days were excluded from the final dataset. Apart from that, there are stations with non-realistic observations. The next figure depicts the precipitation measured by stations TA00128 and TA00249, after filtering the bad observations. It can be seen that except for the deficient data, there are also non-realistic observations.

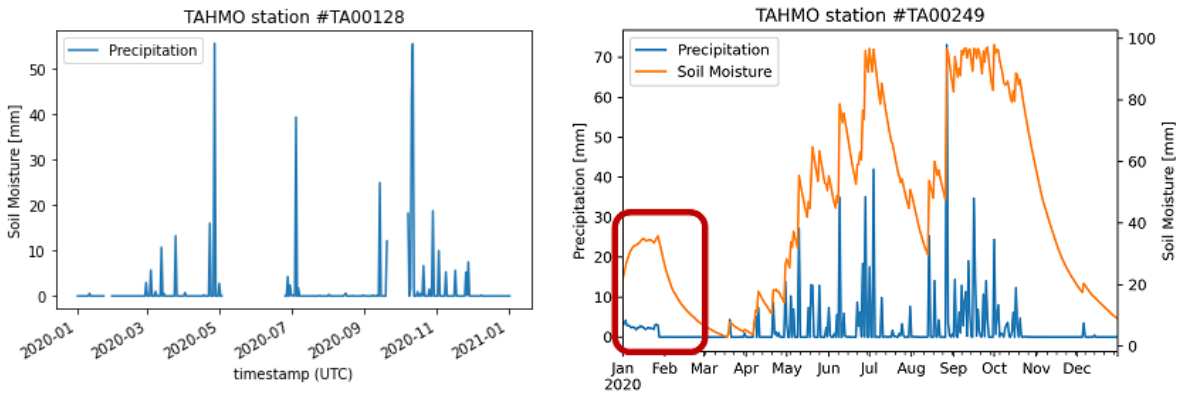

Figure 4.1: Timeseries of the observed precipitation of the TAHMO stations TA00128 and TA00249 during 2020

The problem with this malfunction is that it is difficult to detect using filters and needs manual investigation. The number of stations in combination with the ignorance of the reality make it impossible to manually check all the available stations/data efficiently. Of course, that leads us into using problematic data not only for the training but also for the test of the classification, which means that the performance of the model is probably biased.

4.1.3. Sentinel'1 limitations

4.1.3.1. Spatial Resolution

The Interferometric Wide Swath (IW) Ground Range Detected (GRD) Level-1 High-resolution (HR) product has 5x20 m (or 100 m^2) resolution and GEE has 10x10 m grid spacing. On the other hand, the size of a regular soccer field is about 600 m^2 which is of course bigger than the area that each pixel covers. The problem here is that the signal retrieved for each field may also contain the signal from the surroundings of the field. The Strip Map mode of Sentinel-1 could improve that problem (corresponds to 100 m^2) but the size of the imagery would inefficiently (and significantly) increase the calculation time.

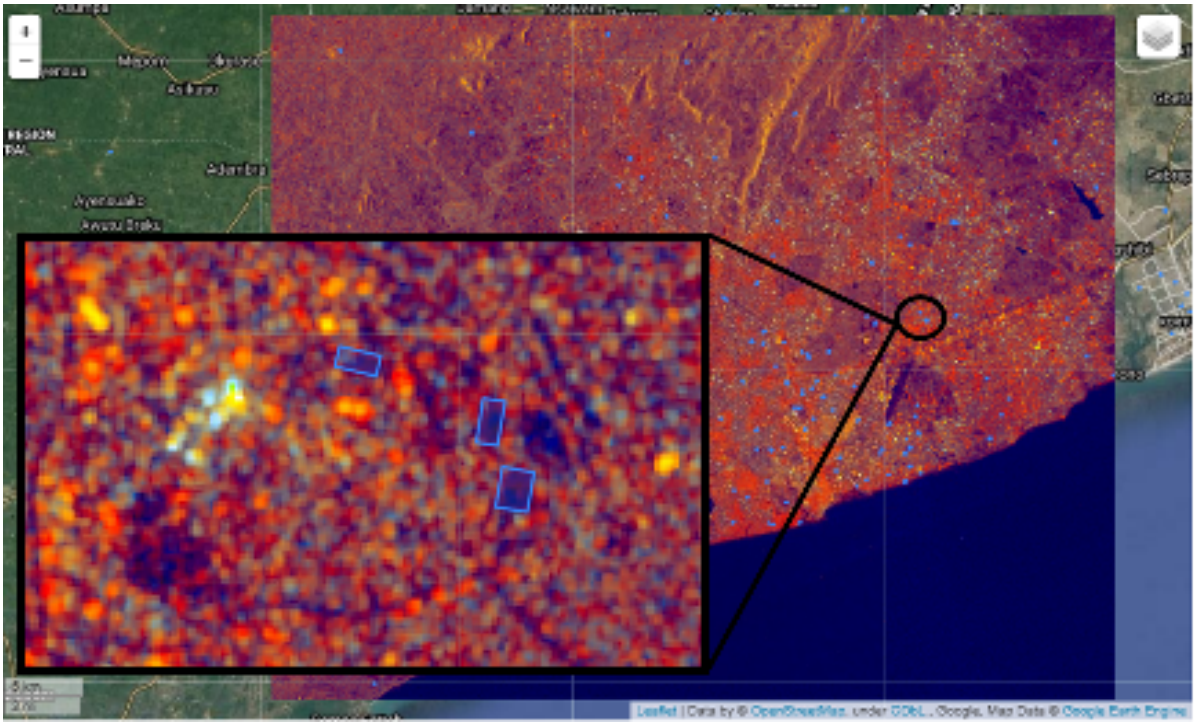

Figure 4.2: Sentinel's 1 GRD imagery over Accra, Ghana at 25/12/2019. The retrieved soccer fields are also included in the map. It can be seen that the pixels are not compatible with the borders of the fields.

4.1.3.2. Temporal Resolution

It can be seen in Figure 2.7 that the temporal resolution of Sentinel-1 constellation varies from one to 6 six or 12 days, which, compared to the previous ones, is considerably higher. The problem that arises is that even the 6-days high resolution is not able to follow the dynamics of rainfall. Thus, the product of that methodology assesses rain/no rain for a period of almost one week, which is not sufficient.

4.2. Selection of Rainfall Detectors and LULC

According to Ampim et al. (2021), the agricultural land use of Ghana increases rapidly. During 1975 and 2013, the agricultural area increased by 19% (from 13% to 32%). That change led to decrease of other land uses. For example, bare area reduced by about 93%, grass-land by about 50% and vegetation by about 40%. It is clear that there are changes in LULC even at annual level.

The direction of these changes is an inverse of the agricultural use, which accounts for the fact that the economy of the country is relying on farming. In this study, I tried to detect homogeneous patches

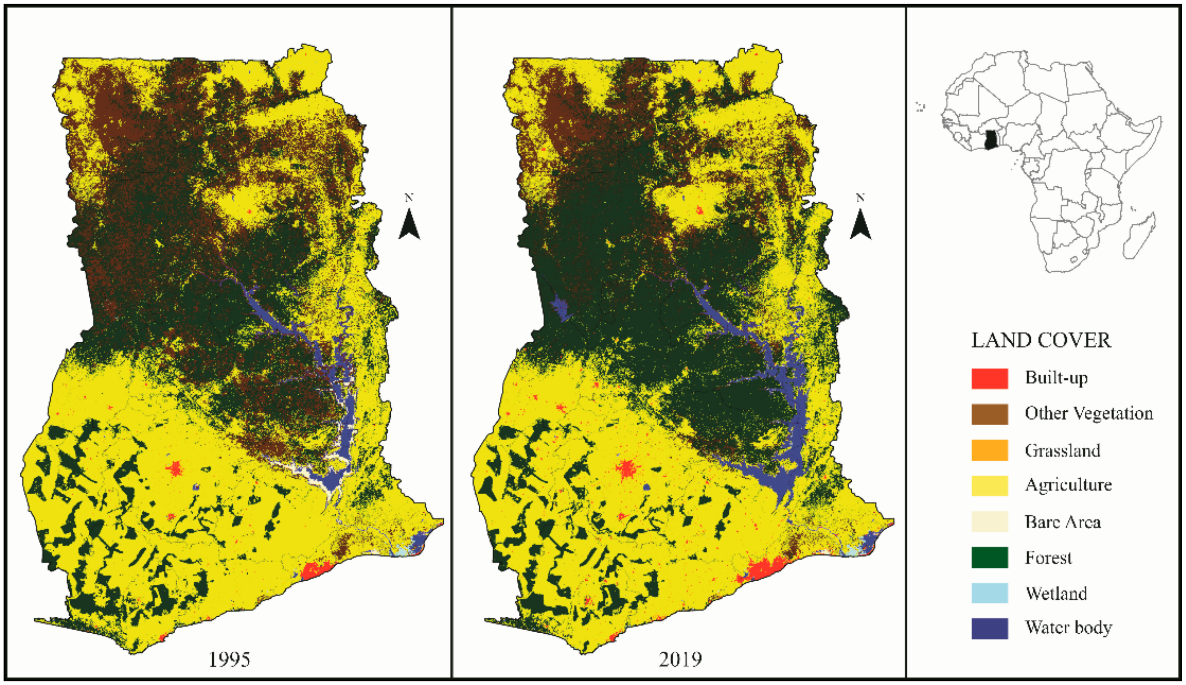

Figure 4.3: Land use change over Ghana during 1995 - 2019. (Ampim et al., 2021)

(mostly bare soil or low vegetation areas) that are not cultivated to retrieve the backscatter signal from Sentinel-1 and use these patches as rainfall detectors for future applications. The rapid LULC change, in combination with the rapid urbanization means that the patches I detected using data of 2020 may not be applicable even for the previous year.

4.3. Modelled soil moisture

Benninga et al., (2020) in their study assessed the uncertainty of soil moisture content based on observations from Sentinel-1. The penetration of the microwave signal in soil varies in different soil types and is estimated to range from 1 to 10 cm. Disregarding the soil properties of rainfall detectors, the calibration of the hydrological model was achieved by a trial and error calibration. Specifically, the calibration of the simple lumped model was done through two variables, S_{in} and S_{umax} that represent the initial storage and the maximum capacity of the soil respectively. For the selection of the final parameters, the relation of the modelled soil moisture with the retrieved signal was assessed. Of course, due to high uncertainty, the parameters were considered stable for the whole Ghana. The integration of soil types should be included in the classification scheme.

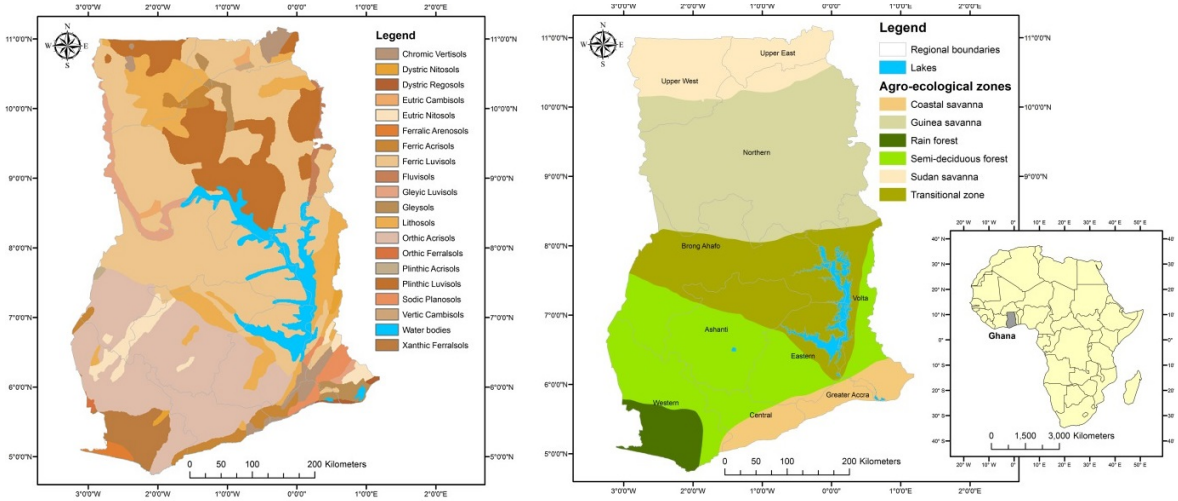

Figure 4.4: Soil units and agroecological zones in Ghana. (source: Rhebergen et al., 2016)

4.4. The suitability of a soccer field

In order to assess the suitability of the soccer fields as rainfall detectors, a manual inspection to specific fields will be assessed.

A noticeable result of the current study was the significant low performance of the soccer field 'a17', which was close to zero during all different training set combinations. Investigating further the specific field and its location, it was found that it is actually a soccer field, but a side area of an airway at Kumasi's International airport.

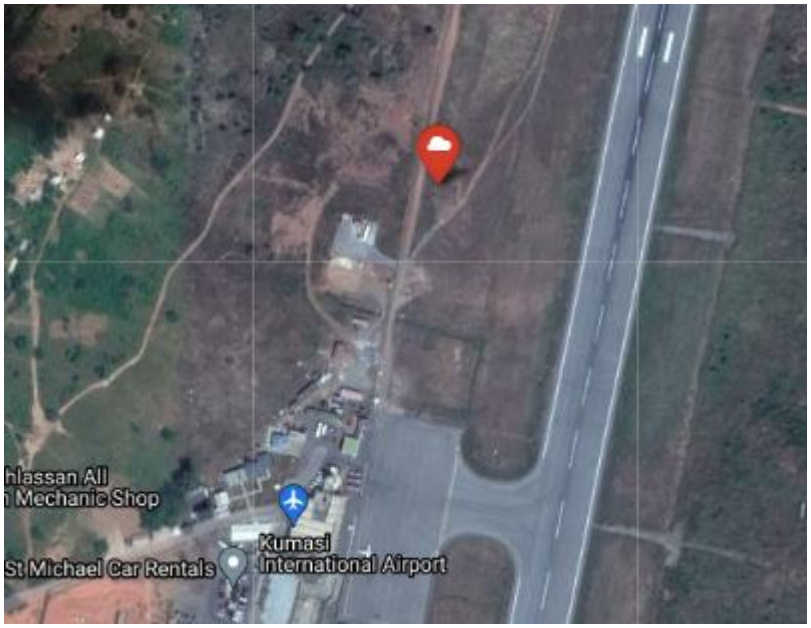

Figure 4.5: Exact location of the Ghanaian 'a17' soccer field, which is located close to the Kumasi International airport

Apart from 'a17', the rest of the median performance of the rest fields was not significantly different; it was close to the median 'All' F1 score of all combinations. A first attempt was to connect the performance of individual soccer fields with soil type classes in Ghana. Unfortunately, the distribution of the TAHMO network is not equal through Ghana and a big part of the country is not covered (middle to North-East). As a consequence, the detected soccer fields followed a similar spatial distribution. Therefore, the data corresponding to suitable soccer fields is not well-spread in the country and the approach of soil type classes cannot be achieved.

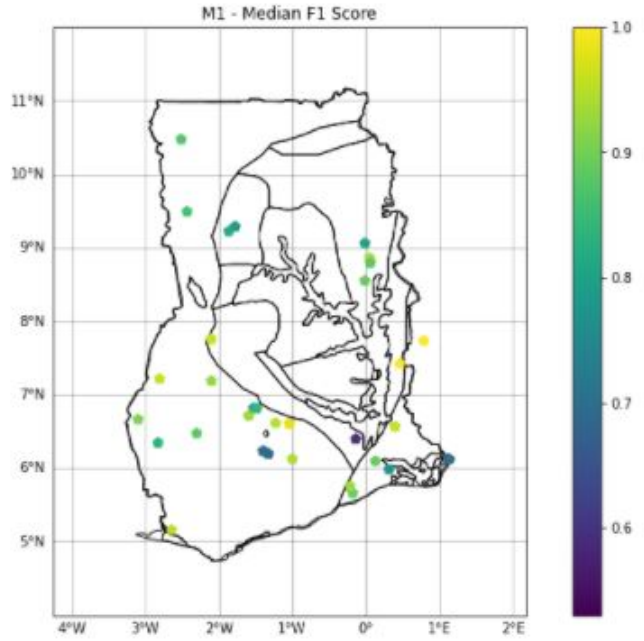

Figure 4.6: The median performance of the Ghanaian soccer fields trained, following Leave-p out method, on the 1^{st} dataset and tested on the excluded fields. The different soil type classes in Ghana are depicted with black lines. (source: Africa Groundwater Atlas (accessed Oct 2021))

After a careful inspection of the different selected soccer fields, a few common characteristics of them were detected:

- The location of the soccer fields is outside urban areas,
- There is low density of urban built-up constructions surrounding them,
- The built-up constructions are not located just outside the soccer field,
- The vegetation of the soccer fields is very low and some fields are even bare soil,
- The fluctuation of vegetation during the year are not high. (Figure 4.7).

The following figure depicts the fluctuation of NDVI per month for the different 50 fields.

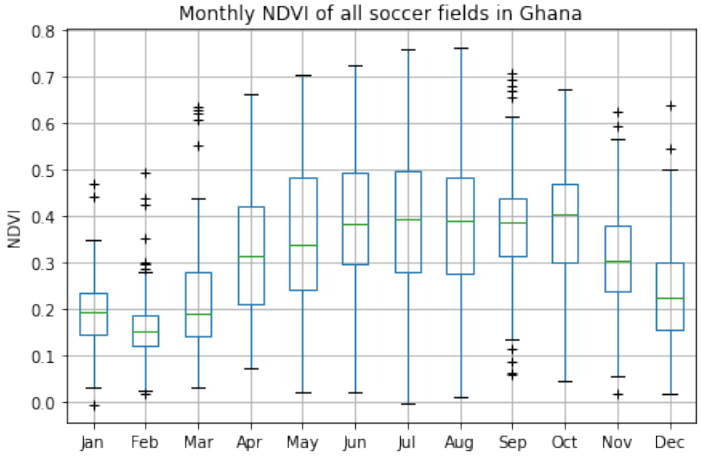

Figure 4.7: The boxplots of the NDVI retrieved in all soccer fields in Ghana in 2020, grouped by month.

4.5. Discussion on different training/test datasets

In order to assess the potential of a stacked classifier for rainfall detection and generalization in space and time, different train and test datasets were used.

4.5.1. Dataset 1

The first training set consists of 45 out of 50 soccer fields in Ghana. The stacked classifier was trained on these stations and then tested i) on the rest 5, ii) on data for all 50 fields in 2019 and iii) on Greek data. This procedure was done 20 times, for different combinations of training fields. Also, for each combination, a total F1 score of all excluded fields (at each combination) is calculated using an aggregated confusion matrix and it is called the 'All' F1 score (aggregated).

For the first testing dataset (excluded fields), many of the test fields are excluded more than once, thus, their performance is assessed many times by different testing tests. The performance of classification regarding all the different combinations in Ghana is encouraging with a median aggregated F1 score of about 67%. The robustness of the methodology was assessed by investigating the deviation of the performance of each soccer field through the different combinations. The results showed that the selected set of excluded fields does not have a large influence on the results. That finding can be proved by the Figure 3.9 where the range of the F1 score is not larger than 0.2. It is worth noting the soccer field 'a17', which haves the lower performance (close to 0) in different iterations, a fact which means that it is not a suitable rainfall detector.

The second testing dataset is the total data for 2019 of all 50 fields in Ghana. The median performance of the aggregated confusion matrix in that case was about 60% and developed very low deviations.

A possible reason of that drop may be the effect of the long-term rainfall variability. In 2017, Baidu et al. assessed the long-term rainfall variability over Ghana using data from the Global Precipitation Climatology Centre (GPCC). They performed the analysis of the rainfall anomaly both in the annual 5-year moving average and in decade scale proving an important decrement of precipitation through the years.

In the case of this study, one limitation is the absence of long (many years) historical rainfall to achieve a well-calibrated and robust machine learning model. Availability of longer timeseries would allow the model to catch the dynamics of the rainfall anomaly and could be more efficient for future predictions.

Finally, the trained classifier was also tested on a random 30% data of the Greek field. The median performance of the different combinations was about 50% and it developed larger deviations. The specific test datasets were used to assess the potential of the generalization of the stacked classifier model in a different environment such as Greece. A reason behind that drop could possibly be the different geology of the two countries. On the one hand, the soil type outside Thessaloniki is Neogene mixed phase fine deposits and on the other hand, soil type in Ghana is mainly Supracrustals, Intracrustals and Paleozoic. It was assumed that the retrieved backscatter signal of soccer fields could potentially represent their dielectric constant. Although, factors as vegetation, soil type, etc., can change the relation of the backscatter signal with the soil moisture of the upper soil. To sum up, it is possible that the successful generalization of the stacked classifier in space can be done in similar soil types.

One more reason of the low performance could be the different climatic and weather conditions that dominate in each area. The climate of Thessaloniki is Hot summer Mediterranean (Csa - Koppen) and the climate of Ghana is Tropical Savanna (Aw - Koppen). The regular conditions of the soil of an area (soil moisture fluctuations during the year) are highly depending on area's dominant climate. For example, the monsoon rainfall events in Ghana that last about 3 months and carry more than 70% of the annual precipitation. This means, that the soil moisture of the rest of the year is more dry compared to Greece that rainfall events are more equally distributed during the whole year.

4.5.2. Dataset 2

The second training set is the whole data from Ghanaian soccer fields during 2020. Then, the performance of that trained classifier is assessed individually on Ghanaian soccer fields for 2019 data. The median F1 score of the 50 soccer fields is close to 60% but there are also outliers close to 0%. An

important condition for the success of that methodology is that the rainfall detector (field) should not be cultivated or irrigated. Though, as it was explained also in Section 4.2, it is difficult to be sure that one field was kept untouched for many years.

4.5.3. Dataset 3

The stacked classifier of that case is trained and tested in Greece. The stacked classifier is tested with a 70% data of the Greek station and tested on the rest 30%. It is important to mention that the specific 30% is the same that was used for the 3^{rd} testing of the 1^{st} training set. The performance of the classification was almost 60%. The result shows that it is recommended that before the generalization of the stacked classifier in place, it should be calibrated in the new area (especially for different environments in terms of soil type and climate).

The limitation of that testing set is that it is randomly selected. It is possible that the selection of a different 30% part of the data would lead to different results but the specific part was selected to be perfectly comparable to the testing dataset of the 1^{st} training set.

4.5.4. Dataset 4

The interpolated rainfall of soccer fields was the extra attribute added in the 20 combinations of the first training set. In that case, the trained classifiers are only tested with the excluded stations of each combination.

A general remark of the comparison of these two models is that the median accuracy of M2 is improved (85%). The performance of the soccer field 'a17' is also significantly increased reaching almost 80%. A limitation of the model, according to results is the deviation of the performance of each field for different combinations. This shows that the model is not stable and the results are much affected by the accuracy of the interpolation. The general improvement of this model is yet not clearly understood because the convective rainfalls that do happen above Ghana are a discouraging factor to interpolate rainfall in a such sparse network.

4.6. Evaluation of Stacked classifier and selection of base classifiers

For the following assessments, the 1^{st} train/test datasets were used.

In this study, we tried to combine the power of the individual classifiers building a stacked one. The results of the individual classifiers compared to the stacked ones were investigated. They showed that the best classifier varied per iteration and per excluded field. The advantage of the stacked classifier is the fact that its performance mostly lays in the upper part of the boxplot.

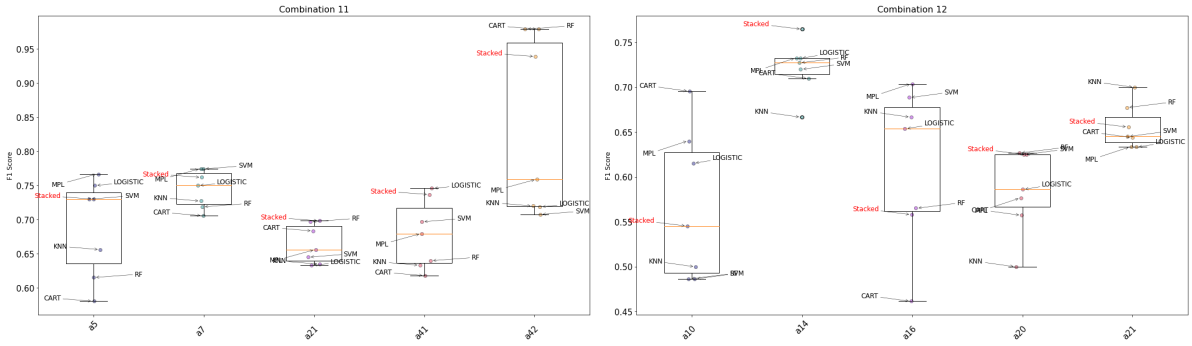

Figure 4.8: The performance of each classifier for each fields for combinations 11 & 12.

The same figure for the aggregated F1 score ('All') per combination shows similar results for the stacked classifier.

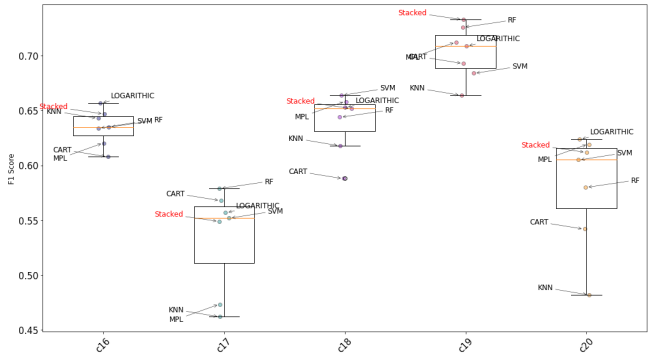

Figure 4.9: The performance of each classifier for all fields in combinations 16 - 20.

At a further step, the combination of base classifiers for the deployment of a stacked one was assessed. For that evaluation the classifier was applied by using a different stacked classifier (excluding one base classifier at each time). Figure 3.16b shows that the full combination of classifiers gave the best results.

4.7. Performance of the IMERG product

The last part was the comparison of the stacked classifier's performance with the performance of the final IMERG, which is a widely used satellite rainfall product.

The results showed that IMERG has a median F1 score for rainfall prediction in Ghana of about 74% and in Greece of about 80%. The IMERG product combines data from many sources that are working with different algorithms and recalibrated with ground data. On the other hand, the stacked classifier only uses spaceborne data from one single satellite. The prediction performance of both satellites is comparable since the IMERG catches about 74% and stacked classifier about 67% for the same soccer fields for 2020. On the other hand, the same comparison for Greece showed bigger differences since IMERG reached an F1 score of ~80% and the performance of the stacked classifier in Greece (Dataset 3 and 3^{rd} test set of Dataset 1) was lower than 60%. The simplicity of data that the stacked classifier uses to reach such high values proves the potential of that algorithm in rainfall prediction.

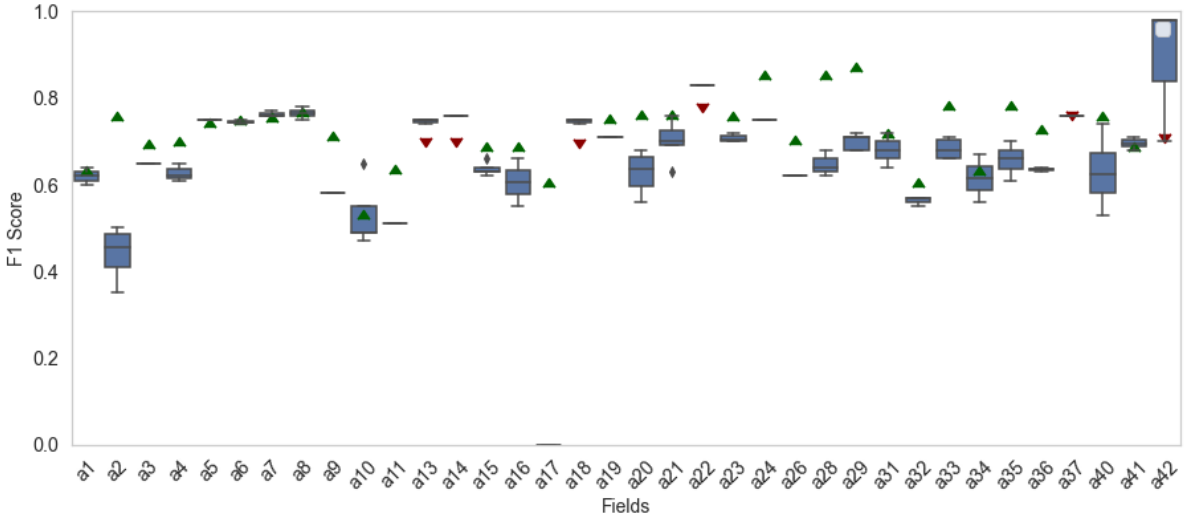

Figure 4.10: The performance of the IMERG algorithm at each soccer field compared with the performance of the stacked classifier trained with 1^{st} dataset and tested with 1^{st} test dataset

In the Figure 4.10, the performance of IMERG at each individual field was added to the results of the first test of the first train dataset. In cases where the performance of the IMERG was higher than the median performance of the box, the scatter of the IMERG was represented with a green arrow. For the other case, the arrow was colored dark red. For the most of the Ghanaian soccer fields in Ghana, the performance of the IMERG product has outweighted the one from the stacked classifier.

5

Conclusion & Future research

The primary goal of the study was to assess the potential of deploying (soccer) fields as rainfall detectors (rain/no rain prediction) using the backscatter C-band signal from Sentinel-1 in a Machine Learning approach. In other words, these rainfall detectors can potentially be used as a quality control tool for the ground stations. At first, the massive retrieval of soccer fields with a Python tool was achieved. In the next step, their soil moisture content was modelled and correlated with their backscatter signal; the high performing soccer fields were selected as rainfall detectors. Then, it was important to prepare the format of the data and select an accurate machine learning classifier to assess rain/no rain. The format of the final data was the same for all soccer fields and generally was the spaceborne data of two consecutive observations and the accumulated rainfall between them. In addition to that, two more attributes were added, the interpolated rainfall and an early prediction based on historical backscatter signal values (quadrant of measurement). In that context, meteorological data of 2020 from Ghana and Greece were used for the training of an ensemble stacked classifier that used multiple base layer classifiers (RF, SVM, CART, MPL and KNN). The stability and the generalization capabilities of the model when trained with and applied to different locations and times was assessed. The replacement of the stacked algorithm with different ones and its comparison with the IMERG product were assessed .

The study showed that:

- There is a considerable potential in the direct connection between backscatter signal and rainfall over homogeneous patches,
- An overall performance of classification using only spaceborne data is about 67% which is competitive with the well-known IMERG product,
- The behavior of the backscatter signal varies based on different factors such as climate and soil type,
- The role of an estimation of rainfall is important and can significantly improve the performance of the model.

The general recommendation for future studies is the evaluation of the Machine Learning methodology in a well-monitored area with multiple years of measurements (and no LULC change). Also, it would be useful for the rainfall detector selection to understand the behavior of different soil types to the backscatter signal. It would be very interesting to apply the same methodology in case of further (Sentinel's-1) temporal resolution improvement. Furthermore, this study assesses the performance of a stacked classifier, but it would be interesting to use individual classifiers. Lastly, it would be very interesting to incorporate the IMERG product in the stacked classifier for rainfall detection or even prediction.

Bibliography

Adab, H., Morbidelli, R., Saltalippi, C., Moradian, M., & Ghalhari, G. A. F. (2020). Machine learning to estimate surface soil moisture from remote sensing data. *Water*, *12*(11). https://doi.org/10.3390/w12113223

Ahmad, S., Kalra, A., & Stephen, H. (2010). Estimating soil moisture using remote sensing data: A machine learning approach. *Advances in Water Resources*, *33*(1), 69–80. https://doi.org/https://doi.org/10.1016/j.advwatres.2009.10.008

Airola, A., Pohjankukka, J., Torppa, J., Middleton, M., Nykänen, V., Heikkonen, J., & Pahikkala, T. (2019). The spatial leave-pair-out cross-validation method for reliable AUC estimation of spatial classifiers. *Data Mining and Knowledge Discovery*, *33*(3), 730–747. https://doi.org/10.1007/s10618-018-00607-x

Alazzy, A. A., Lü, H., Chen, R., Ali, A. B., Zhu, Y., & Su, J. (2017). Evaluation of Satellite Precipitation Products and Their Potential Influence on Hydrological Modeling over the Ganzi River Basin of the Tibetan Plateau (G.-F. Lin, Ed.). *Advances in Meteorology*, *2017*, 3695285. https://doi.org/10.1155/2017/3695285

Alex J. Smola, & Bernhard Scholkopf. (2004). Pattern Recognition and Machine Learning. *Statistics and Computing*, *14*, 199–222. https://doi.org/10.1007/978-3-030-57077-4_11

Allen, R. G., Walter, I. A., Elliott, R. L., Howell, T. A., Itenfisu, D., Jensen, M. E., & Snyder, R. L. (2005). *The asce standardized reference evapotranspiration equation*. American Society of Civil Engineers. https://doi.org/10.1061/9780784408056

Ampim, P. A. Y., Ogbe, M., Obeng, E., Akley, E. K., & MacCarthy, D. S. (2021). Land cover changes in ghana over the past 24 years. *Sustainability*, *13*(9). https://doi.org/10.3390/su13094951

Aonashi, K., Awaka, J., Hirose, M., Kozu, T., Kubota, T., Liu, G., Shige, S., Kida, S., Seto, S., Takahashi, N., & Takayabu, Y. N. (2009). Gsmap passive microwave precipitation retrieval algorithm : Algorithm description and validation. *Journal of the Meteorological Society of Japan. Ser. II*, *87A*, 119–136. https://doi.org/10.2151/jmsj.87A.119

Ashouri, H., Hsu, K.-L., Sorooshian, S., Braithwaite, D. K., Knapp, K. R., Cecil, L. D., Nelson, B. R., & Prat, O. P. (2015). Persiann-cdr: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies. *Bulletin of the American Meteorological Society*, *96*(1), 69–83. https://doi.org/10.1175/BAMS-D-13-00068.1

Baidu, M., Amekudzi, L. K., Aryee, J. N. A., & Annor, T. (2017). Assessment of long-term spatio-temporal rainfall variability over ghana using wavelet analysis. *Climate*, *5*(2). https://doi.org/10.3390/cli5020030

Barret, E. C. (2001). Satellite remote sensing of precipitation: progress and problems. *Remots Sensing and Hydrology*, 7–10.

Bauer-Marschallinger, B., Freeman, V., Cao, S., Paulik, C., Schaufler, S., Stachl, T., Modanesi, S., Massari, C., Ciabatta, L., Brocca, L., & Wagner, W. (2019). Toward global soil moisture monitoring with sentinel-1: Harnessing assets and overcoming obstacles. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(1), 520–539. https://doi.org/10.1109/TGRS.2018.2858004

Beck, H. E., van Dijk, A. I. J. M., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., & de Roo, A. (2017). Mswep: 3-hourly 0.25^{o} global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. *Hydrology and Earth System Sciences*, *21*(1), 589–615. https://doi.org/10.5194/hess-21-589-2017

Benninga, H.-J. F., van der Velde, R., & Su, Z. (2020). Sentinel-1 soil moisture content and its uncertainty over sparsely vegetated fields. *Journal of Hydrology X*, *9*, 100066. https://doi.org/https://doi.org/10.1016/j.hydroa.2020.100066

Bharadwaj, Prakash, K. B., & Kanagachidambaresan, G. R. (2021). *Pattern Recognition and Machine Learning*. https://doi.org/10.1007/978-3-030-57077-4_11

Boeing, G. (2017). Osmnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems*, *65*, 126–139. https://doi.org/https://doi.org/10.1016/j.compenvurbsys.2017.05.004

Brocca, L., Filippucci, P., Hahn, S., Ciabatta, L., Massari, C., Camici, S., Schüller, L., Bojkov, B., & Wagner, W. (2019). Sm2rain–ascat (2007–2018): Global daily satellite rainfall data from ascat soil moisture observations. *Earth System Science Data*, *11*(4), 1583–1601. https://doi.org/10.5194/essd-11-1583-2019

Brocca, L., Moramarco, T., Melone, F., & Wagner, W. (2013). A new method for rainfall estimation through soil moisture observations. *Geophysical Research Letters*, *40*(5), 853–858. https://doi.org/https://doi.org/10.1002/grl.50173

Brocca, L., Ciabatta, L., Massari, C., Camici, S., & Tarpanelli, A. (2017). Soil Moisture for Hydrological Applications: Open Questions and New Opportunities. *Water*, *9*(2). https://doi.org/10.3390/w9020140

Brocca, L., Ciabatta, L., Massari, C., Moramarco, T., Hahn, S., Hasenauer, S., Kidd, R., Dorigo, W., & Levizzani, V. (2014). Soil as a natural rain gauge: Estimating global rainfall from satellite soil moisture data. *Journal of Geophysical Research*, *119*, 5128–5141. https://doi.org/10.1002/2014JD021489

Canada Centre for Remote Sensing. (2019). *Fundamentals of Remote Sensing*. Canada Centre for Mapping; Earth Observation. https://www.nrcan.gc.ca/maps-tools-and-publications/satellite-imagery-and-air-photos/tutorial-fundamentals-remote-sensing/9309

Chen, H., Chandrasekar, V., Cifelli, R., & Xie, P. (2019). A Machine Learning System for Precipitation Estimation Using Satellite and Ground Radar Network Observations. *IEEE Transactions on Geoscience and Remote Sensing*, *PP*, 1–13. https://doi.org/10.1109/TGRS.2019.2942280

CRISP. (2001). Principles of Remote Sensing - Centre for Remote Imaging, Sensing and Processing, CRISP. Retrieved February 20, 2021, from https://crisp.nus.edu.sg/%7B~%7Dresearch/tutorial/sar%7B%5C_%7Dint.htm

Cristiano, E., ten Veldhuis, M.-C., & van de Giesen, N. (2017). Spatial and temporal variability of rainfall and their effects on hydrological response in urban areas – a review. *Hydrology and Earth System Sciences*, *21*(7), 3859–3878. https://doi.org/10.5194/hess-21-3859-2017

Daume III, H. (2004). Notes on CG and LM-BFGS Optimization of Logistic Regression. *MEGA Model Optimization Package*, 1–4. http://www.umiacs.umd.edu/$%5Csim$hal/docs/daume04cg-bfgs.pdf

Dembélé, M., & Zwart, S. J. (2016). Evaluation and comparison of satellite-based rainfall products in burkina faso, west africa. *International Journal of Remote Sensing*, *37*(17), 3995–4014. https://doi.org/10.1080/01431161.2016.1207258

Dinku, T., Funk, C., Peterson, P., Maidment, R., Tadesse, T., Gadain, H., & Ceccato, P. (2018). Validation of the chirps satellite rainfall estimates over eastern africa. *Quarterly Journal of the Royal Meteorological Society*, *144*(S1), 292–312. https://doi.org/https://doi.org/10.1002/qj.3244

Duan, Z., Duggan, E., Chen, C., Gao, H., Dong, J., & Liu, J. (2021). Comparison of traditional method and triple collocation analysis for evaluation of multiple gridded precipitation products across germany. *Journal of Hydrometeorology*, *22*(11), 2983–2999. https://doi.org/10.1175/JHM-D-21-0049.1

Džeroski, S., & Ženko, B. (2004). Is combining classifiers with stacking better than selecting the best one? *Machine Learning*, *54*, 255–273. https://doi.org/10.1023/B:MACH.0000015881.36452.6e

ESA. (2020). User Guides - Sentinel-1 SAR - Level-1 Ground Range Detected - Sentinel Online - Sentinel. Retrieved February 20, 2021, from https://sentinel.esa.int/web/sentinel/user-guides/sentinel-1-sar/resolutions/level-1-ground-range-detected

FAO. (2018). FAOSTAT. Retrieved February 20, 2021, from http://www.fao.org/faostat/en/#country/81

Filipponi, F. (2019). Sentinel-1 GRD Preprocessing Workflow. *Proceedings*, *18*, 6201. https://doi.org/10.3390/ECRS-3-06201

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data*, *2*(1), 150066. https://doi.org/10.1038/sdata.2015.66

Funk, C. C., Verdin, A. P., Michaelsen, J. C., Pedreros, D., Husak, G. J., & Peterson, P. (2015). A global satellite assisted precipitation climatology. *Earth System Science Data*, *8*, 401–425. https://doi.org/10.5194/essdd-8-401-2015

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google earth engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*. https://doi.org/10.1016/j.rse.2017.06.031

Hastie, T. (2020). Ridge regularization: An essential concept in data science. *Technometrics*, *62*(4), 426–433. https://doi.org/10.1080/00401706.2020.1791959

Hinton, G. E., Osindero, S., & Teh, Y.-W. (2006). A fast learning algorithm for deep belief nets. *Neural computation*, *18*, 1527–1554. https://doi.org/10.1162/neco.2006.18.7.1527

Hong, S.-H., & Wdowinski, S. (2012). Revising vegetation scattering theories : Adding a rotated dihedral double bounce scattering to Explain cross polarimetric SAR observations over wetlands. *1*(1), 19–23.

Hrachowitz, M., Fovet, O., Ruiz, L., Euser, T., Gharari, S., Nijzink, R., Freer, J., Savenije, H. H. G., & Gascuel-Odoux, C. (2014). Process consistency in models: The importance of system signatures, expert knowledge, and process complexity. *Water Resources Research*, *50*(9), 7445–7469. https://doi.org/https://doi.org/10.1002/2014WR015484

Hsu, K., & Sorooshian, S. (2009). Satellite-Based Precipitation Measurement Using PERSIANN System (Water Scie), 27–48. https://doi.org/10.1007/978-3-540-77843-1_2

Huffman, G. J., & Bolvin, D. T. (2011). *GPCP Version 2.2 Combined Precipitation Data Set Documentation* (tech. rep.). Laboratory for Atmospheres, NASA Goddard Space Flight Center, Science Systems, and Applications, Inc. USA.

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K., Joyce, R., & Xie, P. (2018). GPM Integrated Multi-Satellite Retrievals for GPM (IMERG) Algorithm Theoretical Basis Document (ATBD) v5.2, 35.

Huffman, G., Stocker; E., Bolvin; D., Nelkin; E., & Jackson, T. J. (2019). GPM IMERG Final Precipitation L3 Half Hourly 0.1 degree x 0.1 degree V06. https://doi.org/10.5067/GPM/IMERG/3B-HH/06

Hunter, S. (1996). WSR-88D radar rainfall estimation: Capabilities, limitations and potential improvements. *Natl. Wea. Dig.*, *20*, 26–38.

Jackson, T. J. (1993). Iii. measuring surface soil moisture using passive microwave remote sensing. *Hydrological Processes*, *7*(2), 139–152. https://doi.org/https://doi.org/10.1002/hyp.3360070205

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning with Applications in R*. Springer. https://doi.org/10.1007/978-1-4614-7138-7

Kerr, Y. H., Waldteufel, P., Richaume, P., Wigneron, J. P., Ferrazzoli, P., Mahmoodi, A., Al Bitar, A., Cabot, F., Gruhier, C., Juglea, S. E., Leroux, D., Mialon, A., & Delwart, S. (2012). The smos soil moisture retrieval algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, *50*(5), 1384–1403. https://doi.org/10.1109/TGRS.2012.2184548

Kidd, C. (2001). Satellite rainfall climatology: A review. *International Journal of Climatology*, *21*, 1041–1066. https://doi.org/10.1002/joc.635

Kidd, C., Becker, A., Huffman, G. J., Muller, C. L., Joe, P., Skofronick-Jackson, G., & Kirschbaum, D. B. (2017). So, how much of the Earth's surface is covered by rain gauges? (2017/01/23). *Bulletin of the American Meteorological Society*, *98*, 69–78. https://doi.org/10.1175/BAMS-D-14-00283.1

Kühnlein, M., Appelhans, T., Thies, B., & Nauss, T. (2014). Improving the accuracy of rainfall rates from optical satellite sensors with machine learning — a random forests-based approach applied to msg seviri. *Remote Sensing of Environment*, *141*, 129–143. https://doi.org/https://doi.org/10.1016/j.rse.2013.10.026

Lamovec, P., Velkanovski, T., Mikos, M., & Osir, K. (2013). Detecting flooded areas with machine learning techniques: case study of the Selška Sora river flash flood in September 2007. *Journal of Applied Remote Sensing*, *7*(1), 1–13. https://doi.org/10.1117/1.JRS.7.073564

Lary, D. J., Zewdie, G. K., Liu, X., Wu, D., Levetin, E., Allee, R. J., Malakar, N., Walker, A., Mussa, H., Mannino, A., & Aurin, D. (2018). Machine learning applications for earth observation. In P.-P. Mathieu & C. Aubrecht (Eds.), *Earth observation open science and innovation* (pp. 165–218). Springer International Publishing. https://doi.org/10.1007/978-3-319-65633-5_8

Le Coz, C., & van de Giesen, N. (2019). Comparison of Rainfall Products over Sub-Saharan Africa. *Journal of Hydrometeorology*, *21*. https://doi.org/10.1175/JHM-D-18-0256.1

Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., & Johnson, B. A. (2019). Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, *152*, 166–177. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2019.04.015

Maidment, R. I., Grimes, D., Black, E., Tarnavsky, E., Young, M., Greatrex, H., Allan, R. P., Stein, T., Nkonde, E., Senkunda, S., & Alcántara, E. M. U. (2017). A new, long-term daily satellite-based rainfall dataset for operational monitoring in Africa. *Scientific Data*, *4*(1), 170063. https://doi.org/10.1038/sdata.2017.63

Manzini, N. (2017). Single hidden layer neural network. Retrieved, from https://www.nicolamanzini.com/single-hidden-layer-neural-network/

Maranan, M., Fink, A., & Knippertz, P. (2018). Rainfall types over southern West Africa: Objective identification, climatology and synoptic environment. *Quarterly Journal of the Royal Meteorological Society*, *144*. https://doi.org/10.1002/qj.3345

Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. *International Journal of Remote Sensing*, *39*(9), 2784–2817. https://doi.org/10.1080/01431161.2018.1433343

Mccollum, J., Gruber, A., & Ba, M. (2000). Discrepancy between gauges and satellite estimates of rainfall in equatorial africa. *Journal of Applied Meteorology - J APPL METEOROL*, *39*, 666–679. https://doi.org/10.1175/1520-0450-39.5.666

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, *5*(4), 115–133. https://doi.org/10.1007/BF02478259

Mega, T., Ushio, T., Takahiro, M., Kubota, T., Kachi, M., & Oki, R. (2019). Gauge-adjusted global satellite mapping of precipitation. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(4), 1928–1935. https://doi.org/10.1109/TGRS.2018.2870199

Mirsoleimani, H. R., Sahebi, M. R., Baghdadi, N., & Hajj, M. E. (2019). Bare Soil Surface Moisture Retrieval from Sentinel-1 SAR Data Based on the Calibrated IEM and Dubois. (1992), 1–12.

Mohan, S. (2013). Potentials and applications of microwave remote sensing.

Moraux, A., Dewitte, S., Cornelis, B., & Munteanu, A. (2019). Deep learning for precipitation estimation from satellite and rain gauges measurements. *Remote Sensing*, *11*(21). https://doi.org/10.3390/rs11212463

Muhammad, E., Muhammad, W., Ahmad, I., Muhammad Khan, N., & Chen, S. (2020). Satellite precipitation product: Applicability and accuracy evaluation in diverse region. *Science China Technological Sciences*, *63*(5), 819–828. https://doi.org/10.1007/s11431-019-1457-3

Nicholson, S. E., Funk, C., & Fink, A. H. (2018). Rainfall over the african continent from the 19th through the 21st century. *Global and Planetary Change*, *165*, 114–127. https://doi.org/https://doi.org/10.1016/j.gloplacha.2017.12.014

Nkunzimana, A., Bi, S., Alriah, M., Zhi, T., & Kur, N. (2020). Comparative analysis of the performance of satellite□based rainfall products over various topographical unities in central east africa: Case of burundi. *Earth and Space Science*, *7*. https://doi.org/10.1029/2019EA000834

Novella, N. S., & Thiaw, W. (2009). Validation of satellite-derived rainfall products over the sahel.

Ochoa-Rodriguez, S., Wang, L.-P., Willems, P., & Onof, C. (2019). A review of radar-rain gauge data merging methods and their potential for urban hydrological applications. *Water Resources Research*, *55*(8), 6356–6391. https://doi.org/https://doi.org/10.1029/2018WR023332

Olson, D., & Delen, D. (2008). *Advanced data mining techniques*. https://doi.org/10.1007/978-3-540-76917-0

Owe, M., de Jeu, R., & Holmes, T. (2008). Multisensor historical climatology of satellite-derived global land surface moisture. *Journal of Geophysical Research: Earth Surface*, *113*(F1). https://doi.org/https://doi.org/10.1029/2007JF000769

Pohjankukka, J., Pahikkala, T., Nevalainen, P., & Heikkonen, J. (2017). Estimating the prediction performance of spatial models via spatial k-fold cross validation. *International Journal of Geographical Information Science*, *31*(10), 2001–2019. https://doi.org/10.1080/13658816.2017.1346255

Qiaohong, S., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., & Hsu, K. (2017). A review of global precipitation data sets: Data sources, estimation, and intercomparisons. *Reviews of Geophysics*, *56*. https://doi.org/10.1002/2017RG000574

Rhebergen, T., Fairhurst, T., Zingore, S., Fisher, M., Oberthür, T., & Whitbread, A. (2016). Climate, soil and land-use based land suitability evaluation for oil palm production in ghana. *European Journal of Agronomy*, *81*, 1–14. https://doi.org/https://doi.org/10.1016/j.eja.2016.08.004

Roy, D., Hassan, S. M. Q., & Sultana, S. S. (2020). An Assessment of Spatial Distribution of Four Different Satellite-Derived Rainfall Estimations and Observed Precipitation over Bangladesh. *Journal of Agricultural Chemistry and Environment*, *09*(04), 195–205. https://doi.org/10.4236/jacen.2020.94016

Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development*, *3*(3), 210–229. https://doi.org/10.1147/rd.33.0210

Satge, F., Hussain, Y., Molina Carpio, J., Zolá, R., Laugner, C., Akhter, G., & Bonnet, M.-P. (2020). Reliability of sm2rain precipitation datasets in comparison to gauge observations and hydrological modelling over arid regions. *International Journal of Climatology*. https://doi.org/10.1002/joc.6704

Schmugge, T. (1984). Microwave remote sensing of soil moisture. *481*. https://doi.org/10.1117/12.943093

Shin, K., Song, J. J., Bang, W., & Lee, G. (2021). Quantitative precipitation estimates using machine learning approaches with operational dual-polarization radar data. *Remote Sensing*, *13*(4). https://doi.org/10.3390/rs13040694

Sokol, Z., Szturc, J., Orellana Alvear, J., Popova, J., Jurczyk, A., & Celleri, R. (2021). The role of weather radar in rainfall estimation and its application in meteorological and hydrological modelling—a review. *Remote Sensing*, *13*(3). https://doi.org/10.3390/rs13030351

Song, Z., Bai, Y., Wang, D., Li, T., & He, X. (2021). Satellite retrieval of air pollution changes in central and eastern china during covid-19 lockdown based on a machine learning model. *Remote Sensing*, *13*(13). https://doi.org/10.3390/rs13132525

Stampoulis, D., Ghasemi Damavandi, H., Boscovic, D., & Sabo, J. (2019). Using satellite remote sensing and machine learning techniques towards precipitation prediction and vegetation classification. *Journal of Environmental Informatics*. https://doi.org/10.3808/jei.202000427

Sultan, B., & Gaetani, M. (2016). Agriculture in West Africa in the Twenty-First Century: Climate Change and Impacts Scenarios, and Potential for Adaptation. *Frontiers in plant science*, *7*, 1262. https://doi.org/10.3389/fpls.2016.01262

Tang, G., Clark, M. P., Papalexiou, S. M., Ma, Z., & Hong, Y. (2020). Have satellite precipitation products improved over last two decades? a comprehensive comparison of gpm imerg with nine satellite and reanalysis datasets. *Remote Sensing of Environment*, *240*, 111697. https://doi.org/https://doi.org/10.1016/j.rse.2020.111697

Turini, N., Thies, B., Horna, N., & Bendix, J. (2021). Random forest-based rainfall retrieval for ecuador using goes-16 and imerg-v06 data. *European Journal of Remote Sensing*, *54*(1), 117–139. https://doi.org/10.1080/22797254.2021.1884002

Van De Giesen, N., Hut, R., Andreini, M., & Selker, J. S. (2013). Trans-African Hydro-Meteorological Observatory (TAHMO): A network to monitor weather, water, and climate in Africa. *AGU Fall Meeting Abstracts*, *2013*, Article H52C-04, H52C–04.

Wagner, W., Hahn, S., Kidd, R., Melzer, T., Bartalis, Z., Hasenauer, S., Figa-Saldaña, J., de Rosnay, P., Jann, A., Schneider, S., Komma, J., Kubu, G., Brugger, K., Aubrecht, C., Züger, J., Gangkofner, U., Kienberger, S., Brocca, L., Wang, Y., … Steinnocher, K. (2013). The ascat soil moisture product: A review of its specifications, validation results, and emerging applications. *Meteorologische Zeitschrift*, *22*(1), 5–33. https://doi.org/10.1127/0941-2948/2013/0399

Wolpert, D. H. (1992). Stacked generalization. *Neural Networks*, *5*(2), 241–259. https://doi.org/https://doi.org/10.1016/S0893-6080(05)80023-1

Woodhouse, I. H. (2017). Introduction to Microwave Remote Sensing. https://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=5475678

World Bank. (2020). Agriculture Overview. Retrieved February 20, 2021, from https://www.worldbank.org/en/topic/agriculture/overview

Xie, P., & Arkin, P. A. (1996). Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions. *Journal of Climate*, *9*(4), 840–858. https://doi.org/10.1175/1520-0442(1996)009<0840:AOGMPU>2.0.CO;2

Xie, P., Janowiak, J. E., Arkin, P. A., Adler, R., Gruber, A., Ferraro, R., Huffman, G. J., & Curtis, S. (2003). Gpcp pentad precipitation analyses: An experimental dataset based on gauge observations and satellite estimates. *Journal of Climate*, *16*(13), 2197–2214. https://doi.org/10.1175/2769.1

Xu, K., Qian, J., Hu, Z.-Y., Duan, Z., Chen, C., Liu, J., Sun, J., Wei, S., & Xing, X. (2021). A new machine learning approach in detecting the oil palm plantations using remote sensing data. *Remote Sensing*, *13*, 236. https://doi.org/10.3390/rs13020236