University of Fort Hare
Together in Excellence
Department of Geography and Environmental Science

INVESTIGATING THE FEASIBILITY OF USING REMOTE SENSING IN INDEX-BASED CROP INSURANCE FOR SOUTH AFRICA'S SMALLHOLDER FARMING SYSTEMS

A THESIS SUBMITTED IN FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY DEGREE IN GEOGRAPHY

BY

WONGA MASIZA

SUPERVISOR: DR. J.G CHIRIMA

CO-SUPERVISORS: DR. A.M. KALUMBA, PROF. H. HAMANDAWANA, DR. H.B MAGAGULA

DATE: OCTOBER 2021

DECLARATION

I, the undersigned, declare that this thesis titled "Investigating the feasibility of using remote sensing in index-based crop insurance for South Africa's smallholder farming systems" is my own original work except where stated, and that it has not been submitted to any other University.

Name: Wonga Masiza

Signature: _____________

Date: _________________

PLAGIARISM STATEMENT

I, Wonga Masiza, student number: 200901062 hereby declare that I am fully aware of the University of Fort Hare's policy on plagiarism and I have taken every precaution to comply with the regulations.

Signature.. Date....................................

DISCLAIMER AND STATEMENT OF CONTRIBUTIONS

This thesis is composed of seven chapters, five of which are manuscripts; however, all the references are presented as a consolidated list after the last chapter. The first and the last chapters follow the traditional format, whereas the other chapters are presented as standalone manuscripts, which are either published, under review, or to be submitted for publication. Consequently, there could be some information overlaps between chapters due to the adopted format. It is also important to note that as the principal author, I conceptualised, conducted the fieldwork, analysed and drafted these manuscripts under the guidance of the supervisors with whom I co-authored the manuscripts.

Published manuscripts

- **Masiza, W.,** Chirima, J. G., Hamandawana, H., and Pillay, R. (2020). Enhanced mapping of a smallholder crop farming landscape through image fusion and model stacking. *International Journal of Remote Sensing*, 41(22), 8739-8756. https://doi.org/10.1080/01431161.2020.1783017
- **Masiza, W.,** Chirima, J. G., Hamandawana, H., Kalumba, A. M., and Magagula, H. B. (2021). Linking Agricultural Index Insurance with Factors That Influence Maize Yield in Rain-Fed Smallholder Farming Systems. *Sustainability*, 13(9), 5176. https://doi.org/10.3390/su13095176

- **Masiza, W**., Chirima, J. G., Hamandawana, H., Kalumba, A. M., and Magagula, H. B. (2022). Do Satellite Data Correlate with In Situ Rainfall and Smallholder Crop Yields? Implications for Crop Insurance. Sustainability, 14(3), 1670. https://doi.org/10.3390/su14031670

Manuscripts under review

- **Masiza, W**., Chirima, J.G., Kalumba, A.M., Magagula, H.B., Hamandawana, H. (2022). Remote sensing for smallholder crop insurance in Africa: A review. *Remote Sensing Applications: Society and Environment*.

- Masiza, W., Chirima, J.G., Kalumba, A.M., Magagula, H.B., Hamandawana, H. (2022). A proposed satellite-based crop insurance system for smallholder maize farming. *Remote Sensing.*

ACKNOWLEDGEMENTS

I thank God for the gift of life and the opportunity He gives us to study in order to understand Him, His creation and ourselves. I am grateful for the guidance and support provided by my supervisors; Johannes George Chirima, Ahmed Mukalazi Kalumba, Hamisai Hamandawana, Hezekiel Bheki Magagula, and Rajendran Pillay, until the completion of this thesis.

My sincere and heartfelt gratitude to the following colleagues who contributed to this work in different ways: Pathutshedzo Eugene Ratshiedana, Kgaogelo Mogano, Siphokazi Gcayi, Bonolo Mosuwe, Zinhle Mashaba-Munghemezulu, Reneilwe Maake, Deon Du Toit, Richard Tswai, Khaled Abutaleb, Ahmed Mohamed Abd Elbasit, Derick Vermaak, Gert de Nysschen, Lindumusa Myeni, Mxolisi Mtyobile, and the support staff at the Agricultural Research Council.

I wish to acknowledge the Agricultural Research Council as the leader of South Africa's Crop Estimate Consortium as well as the SIQ and GeoterraImage.

Special thanks to the farmers in Mthatha, Coffee Bay, Ngqeleni, Libode, Tsolo, Qumbu, Mqanduli, Xhorha, Port St Johns, and Qunu who made it possible for us to do these investigations by granting us access to their farms.

I greatly appreciate the support I received from my family and friends; Anelisa Masiza, Chwayita Kolela, Mfundo Masiza, Xola Ngceni, Loyiso Macgyva Malgas, Nombini Tingo, Babatope Akinyemi, Odwa Mtatambi, Bongo Mkiva, Olwethu Loki, Lutic Molo Mosoane, and Patrick Nemububuni.

ABSTRACT

Crop farming in Sub-Saharan Africa (SSA) is largely practiced by resource-poor farmers under rain-fed and unpredictable weather conditions. Since agriculture is the mainstay of SSA's economy, the lack of improved and adapted agricultural technologies in this region sets back economic development and the fight against poverty. Overcoming this constraint and achieving the sustainable development goal to end poverty, requires innovative tools that can be used for weather risk management. One tool that has been gaining momentum recently is index-based crop insurance (IBCI). Since the launch of the first IBCI program in Africa around 2005, the number of IBCI programs has increased. Unfortunately, these programs are constrained by poor product design, basis risk, and low uptake of contracts. When these issues were first pointed-out in the earliest IBCI programs, many reports suggested satellite remote sensing (RS) as a viable solution.

Hence, the first objective of this study was to assess how RS has been used in IBCI, the challenges RS faces, and potential contributions of RS that have not yet been meaningfully exploited. The literature shows that IBCI programs are increasingly adopting RS. RS has improved demarcation of unit areas of insurance and enabled IBCI to reach inaccessible areas that do not have sufficient meteorological infrastructure. However, the literature also shows that IBCI is still tainted by basis risk, which emanates from poor contract designs, the influence of non-weather factors on crop yields, imperfect correlations between satellite-based indices and crop yields, and the lack of historical data for calibration. Although IBCI reports cover vegetation and crop health monitoring, few to none cover crop type and crop area mapping. Furthermore, areas including high-resolution mapping, data fusion, microwave RS, machine learning, and computer vision have not been sufficiently tested in IBCI.

The second objective of this study was to assess how RS and machine learning techniques can be used to enhance the mapping of smallholder crop farming landscapes. The findings show that machine learning ensembles and the combination of optical and microwave data can map a smallholder farming landscape with a maximum accuracy of 97.71 %. The third objective was to identify factors that influence crop yields and crop losses in order to improve IBCI design. Results demonstrated that the pervasive notion that low yields in smallholder farms are related to rainfall is an oversimplification. Factors including fertilizer use, seed variety, soil properties, soil moisture, growing degree-days, management, and socioeconomic conditions are some of the most important

factors influencing crop yields and crop losses in smallholder farming systems. This shows why IBCI needs to be part of a comprehensive risk management system that understands and approaches smallholder crop farming as complex by linking insurance with advisories and input supplies. Improved inputs and good farming practices could reduce the influence of non-weather factors on crop losses, and thereby reduce basis risk in weather-based index insurance (WII) contracts.

The fourth objective of this study was to assess how well the combination of synthetic aperture radar (SAR) and optical indices estimate soil moisture. As stated earlier, soil moisture was found to be one of the most important factors affecting crop yields. Although this method better estimated soil moisture over the first half of the growing season, estimation accuracies were comparable to those found in studies that had used similar datasets (RMSE = 0.043 $m^3 m^{-3}$, MAE = 0.034 $m^3 m^{-3}$). Further interrogation of interaction effects between the variables used in this study and consideration of other factors that affect SAR backscatter could improve the method. More importantly, incorporating high-resolution satellite-based monitoring of soil moisture into IBCI could potentially reduce basis risk.

The fifth objective of this study was to develop an IBCI for smallholder crop farming systems. The proposed IBCI scheme covers maize and derives index thresholds from crop water requirements and satellite-based rainfall estimates. It covers rainfall deficits over the vegetative, mid-season, and late-season stages of maize growth. The key contribution of this system is the derivation of index thresholds from CWR and site-specific rainfall conditions. The widely used approach, which calibrates IBCI by correlating yields and rainfall, exposes contracts to basis risk because, by simply correlating yield and rainfall data, it overlooks the influence of non-weather factors on crop yields and losses. The proposed system must be linked or bundled with non-weather variables that affect crop yields. Effectively, this means that the insurance must be linked or bundled with advisories and input supplies to address the influence of non-weather factors on crop losses. This system also incorporates a crop area-mapping component, which was found to be lacking in many IBCI systems. In conclusion, an IBCI that is based on crop water requirements, which incorporates crop area mapping and links insurance with non-weather crop yield-determining factors, is potentially capable of improving crop insurance for smallholder farming systems.

Keywords: Index insurance; crop yields; smallholder; maize; remote sensing; machine learning

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1 : Introduction

1.1. Background

Crop farming in smallholder-dominated areas like Sub-Saharan Africa (SSA) is largely practiced by resource-poor farmers under rain-fed and unpredictable weather conditions (Wani et al., 2009; Buhaug et al., 2015). The lack of access to improved inputs such as climate-adapted seeds, fertilizers, and irrigation, as well as the absence of early warning systems and good-quality advisory services, leaves agriculture in these areas poorly protected and vulnerable to adverse weather. This lack of capacity to manage weather risks often results in production fluctuations and low yields, which threaten food security, increase poverty, and hamper economic development. In South Africa, for example, recent droughts resulted in the abandonment of farm activities, loss of employment, and increased poverty (Manderson et al., 2016; Nembilwi et al., 2021). Similar scenarios are reported in East Africa where the occurrences of agricultural droughts between 2008 and 2018 resulted in low production, reduced incomes, increased food prices, unemployment, famine, and migration (Gebremeskel et al., 2019). In West Africa, reports show that droughts reduce the quality of grain and increase food scarcity, malnutrition, diseases, and mortality (Gautier et al., 2016).

Although perils such as rainfall floods affect agriculture in many SSA countries (Musyoki et al., 2016; Henry et al., 2018; Kilavi et al., 2018; Mavhura, 2019), droughts are more catastrophic than other natural disasters (Carter et al., 2014). To minimize the adverse effects of weather shocks, farmers have traditionally employed crop diversification and changing of planting dates (Akinnagbe and Irohibe, 2014). Unfortunately, these risk-avoidance approaches are constraints because, by managing risks in this manner, farmers do not make profitable investments in improved agricultural technologies such as fertilizers, climate-adapted seeds, irrigation systems, and other agronomic inputs (Hazell, 1992; Alderman and Haque, 2007; Dercon and Christiaensen, 2011). Farmers cannot afford these technologies because of limited access to capital, credit, and insurance. On the other hand, the traditional informal risk-sharing schemes are often ineffective in the face of covariate catastrophic events (Will et al., 2021).

These challenges have necessitated the exploration of various risk-management strategies and tools, one of which is index-based crop insurance (IBCI). IBCI is regarded as being capable of counteracting the limitations of traditional claim-based insurances (Barnett and Mahul, 2007). Traditional insurance schemes involve on-field damage assessments that come with high administrative costs, moral hazard and adverse selection. IBCI, on the other hand, is less susceptible to these shortcomings because it indemnifies clients based on objective indices that are known to be associated with crop losses (Barnett and Mahul, 2007; Carter et al., 2014). These indices may be derived from weather parameters such as rainfall, temperature, soil moisture, and weather-related vegetation indices which can be obtained from weather station records and remote sensing data (Ntukamazina et al., 2017; Di Marcantonio and Kayitakire, 2017; IFAD, 2015). By using these indices, IBCI has less administrative costs because it does not involve on-field surveys of individual farms (Carter et al., 2017).

IBCI was conceptualized by Halcrow, (1948, 1949) and advanced by Dandekar, (1985) and Agarwal, (1979). The earliest attempts to incorporate satellite remote sensing in IBCI were in the mid-1970s and early-1980s (Towery et al., 1975, 1980). To date, satellite-based IBCI has been implemented in several African countries and other parts of the world (Di Marcantonio and Kayitakire, 2017; Ntukamazina et al., 2017; Gaurav and Chaudhary, 2020; Ye et al., 2020). Recent studies have been exploring the feasibility of using satellite data because of the advantages that satellites have over weather stations. These advantages include the ability of satellites to cover wide and inaccessible geographic areas, and freely available datasets that are downloadable from open access archives. In addition to providing information on agrometeorological parameters such as temperature, rainfall, clouds and soil moisture, satellite datasets enable mapping of cropped areas and monitoring of vegetation health. Despite the advantages that remote sensing provides to IBCI, uptake of insurance is still low due to basis risk which is associated with poor product design, overlooking of non-weather factors that influence crop yields, lack of calibration data, and imperfect correlations between satellite-based indices and crop yields (Clement et al., 2018; Ntshakira-Rukundo et al., 2021).

The primary goal of this study is to explore the feasibility of using remote sensing in enhancing the design of IBCI for smallholder farming systems (SFS). By undertaking this investigation, this study complements the efforts which researchers and insurers are making to improve the existing satellite-based IBCI products in Africa. Furthermore, satellite-based IBCI has not yet been tested in South Africa's SFS.

1.2. Problem statement

Crop farming in South Africa, just like in other SSA countries, is vulnerable to production fluctuations and losses because of SSA's unpredictable weather. To counteract this challenge, researchers see agricultural insurance as one of the viable tools that can protect farmers against the adverse impacts of weather. Although many African countries are piloting insurance schemes, there has been little to no research investigating the feasibility of satellite-based IBCI in South Africa. The few studies that have been published so far report that South Africa's smallholder farmers do not afford insurance premiums and are unaware of the existing traditional insurance products (Partridge and Wagner, 2016; Elum et al., 2017; Oduniyi et al., 2020). However, the inclusive literature covering the rest of Africa and other parts of the world shows that uptake of insurance, particularly IBCI, is low, mainly because of poor product design and poor product quality. This often manifests as the lack of correspondence between the indices used to quantify losses and the actual losses in the farms, which results in farmers not getting proper compensation. This study, therefore, identified the lack of insurance in South Africa's SFS and the poor design of IBCI as gaps for innovation, and determined that there is an urgent need to develop a sustainable satellite-based IBCI in order to ensure food security under the current unpredictable climate scenarios.

1.3. Research questions

- What developments has RS introduced in IBCI and what are its limitations and untapped opportunities?
- How can IBCI optimally use RS data and machine learning techniques to enhance the mapping of smallholder crop farming landscapes?
- What are the major factors contributing to crop yields and crop losses which remote sensing is not able to detect, which should be linked or bundled with IBCI?
- Can higher resolution SAR data and optical indices be used to estimate soil moisture in smallholder croplands?
- Is a satellite-based IBCI that incorporates crop water requirements, non-weather yield-determining factors, and crop area mapping feasible?

1.4. Aim

- The aim of this study was to investigate the feasibility of using remote sensing in the design of IBCI for South Africa's SFS and improve current IBCI designs by providing empirically informed insights about rain-fed SFS.

1.5. Objectives

To achieve the aim of the study, the objectives were to:

- Review the contributions of remote sensing in the development of IBCI in Africa by tracking progress and identifying gaps and opportunities
- Examine the performance of different machine learning and remote sensing techniques in the mapping of smallholder crop farming areas
- Identify crop yield and crop loss-determining factors that need to be linked or bundled with IBCI.
- Investigate the potential of retrieving soil moisture measurements from optical and SAR data
- Develop a satellite-based IBCI for SFS

1.6. Justification of the study

The initiatives taken by insurers in partnership with governments to provide IBCI in Africa and other parts of the world demonstrate the urgency of the fight against poverty. The overarching purpose of these efforts is to achieve food security, which is one of the targets of the sustainable development goals (SDG) of the United Nations. To support these efforts, there has been an increasing use of geospatial technology in agriculture, which has resulted in major developments such as precision, smart, and digital farming. However, there are still gaps and room for innovation, especially in agricultural insurance and particularly in South Africa where smallholder and emerging farmers do not have access to insurance and credit. This study, therefore, is strategically aligned with the SDG and the efforts to achieve sustainable agriculture. It is hoped that the findings of this research will reach all the relevant stakeholders including governments, insurers, bankers, farmers, and the research community for policy formulation and implementation.

1.7. Overview of the study area

All investigations were conducted in O.R Tambo District Municipality (ORTDM), which is one of the major smallholder crop farming areas in the Eastern Cape Province and South Africa at large (Figure 1.1).

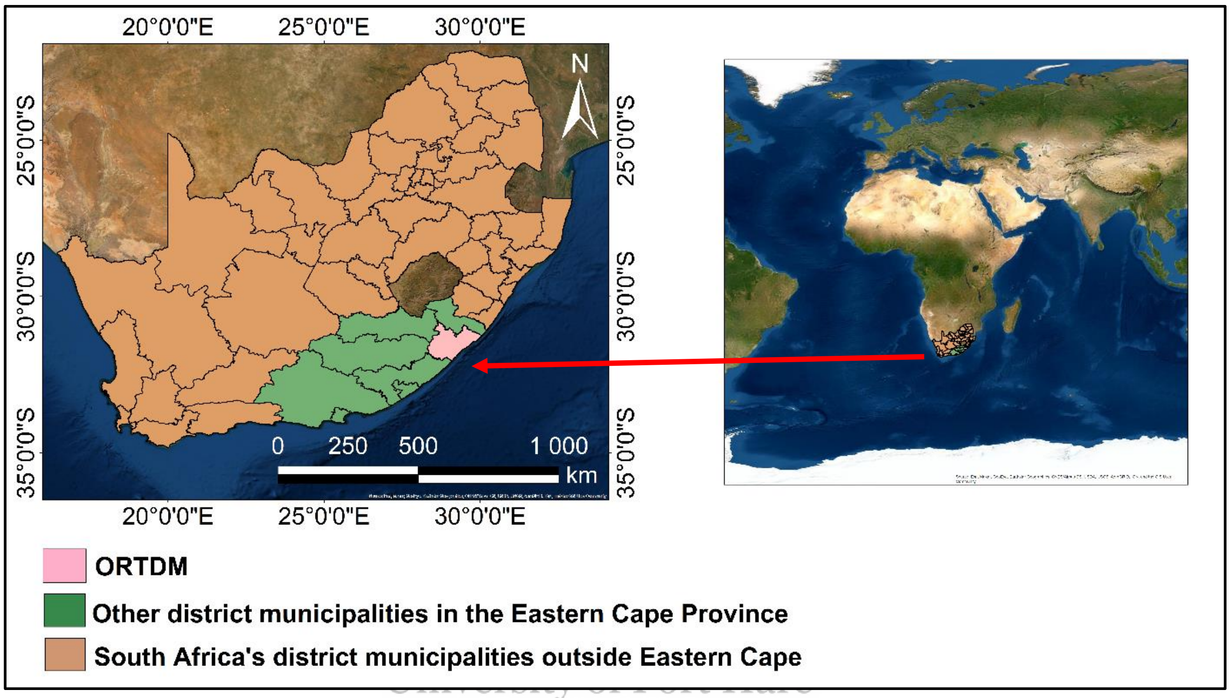


Figure 1.1: Location of ORTDM

ORTDM is the core of the former Transkei, which was one of South Africa's homelands for black people under the apartheid government. The entire district municipality consists of five local municipalities covering a total area of 12 096 km². The majority (94%) of its 1.47 million population (projected to be 1.55 million by 2021) lives in villages (ECSECC, 2017). The South African Local Government Association (SALGA) reports that ORTDM has an Unemployment rate of 41 % (SALGA, 2020). The major sources of livelihood are government welfare grants and employment in community services, trade, and agriculture. Agriculture is mostly livestock and small-scale maize cultivation. Yellow maize is sold to animal feed retailers, villagers, and informal traders, and some is used as feed for domestic animals. White maize is sold to maize milling plants and also used for subsistence. The growing season starts in October and the harvest season in June.

Although the farmers do get support from the government through cropping programs, they do not have any formal agricultural insurance.

Agriculture in this area is practiced under warm oceanic, sub-tropical, and semi-arid climatic conditions. Average rainfall per year ranges between 900 mm and 1300 mm, with summer minimum and maximum temperatures of 14-19 °C and 14-27 °C, respectively. The coastal areas consist of densely populated natural vegetation, open grasslands and an undulating terrain with a maximum elevation of 500 m. The interior areas are mostly open grasslands with a gentle-to-moderate-sloping terrain. The northern areas consist of savannas and a maximum elevation of 1500 m. The geology of the area comprises sedimentary rocks; mainly sandstones, mudstones and shales, which are intruded by dolerite dykes and sills (KSD, 2021; Sibanda et al., 2016). The soils are largely dominated by sandy loams, sandy clay loams, and clays that are yellow to black in colour and slightly acidic (Eta and Grace, 2013; Nyandeni Local Municipality, 2018). ORTDM was selected because it typifies South Africa's smallholder maize farming areas.

1.8. Methodological approach

The study adopted a mix of methods comprising a systematic review, quantitative statistical analysis, and machine learning. Systematic literature review collects information from publications that meet defined criteria in order to answer specific questions (Mengist et al., 2020). This approach is widely used in agriculture and environmental sciences to identify gaps in the existing literature in order to guide new research (Koutsos et al., 2019; Mengist et al., 2020). This study conducted a systematic literature review to scrutinize how remote sensing has been used in crop insurance in order to identify untapped opportunities. The study also used conventional statistics and machine learning, which overlap but differ in their purpose. Although incorporated into many machine-learning algorithms, conventional statistical methods focus on inferring relationships between variables. Although machine learning can also relate variables, it compromises model interpretability and focuses more on optimising prediction rather than making inference about relationships between variables (Bzdok et al., 2018). In other words, machine learning focuses on model performance rather than understanding the underlying statistical mechanisms, whereas traditional statistical methods focus on understanding the underlying model, its mechanisms and assumptions. Machine learning seeks to make the most accurate prediction by finding patterns and learning from the data, whereas statistics is concerned about the inference that can be drawn from the model. Compared to conventional statistics, machine-learning methods can handle a large number of input variables and they work best with large sample sizes (Ley et al., 2022). Conventional statistical approaches are easier to interpret and simple to understand, while most machine learning methods are difficult to interpret (Rajula et al., 2020). However, ML techniques are flexible and, unlike conventional statistical approaches, are free of statistical assumptions (Rajula et al., 2020). This study took advantage of the abilities of both conventional statistics and machine learning by employing each of the two approaches where it was more suitable.

1.8.1. Sampling

Since it is costly, time-consuming and almost impossible to collect information about everyone, everything or every individual in a population, research studies and surveys collect samples that are representative of population groups. Sampling, therefore, must be guided by a sampling plan that determines the sampling strategy (e.g. random, judgmental, stratified, etc.), the type of sampling (e.g. grab, composite, etc.), the appropriate times to collect the samples, and the minimum number of samples required (Mitra et al., 2018). This study used different sampling methods each of which was selected according to the research question, the conditions of the environment from which samples were collected, the nature of the sampled data, and the techniques used to analyse the data. These sampling procedures are described in the main chapters of the thesis.

1.8.2. Data collection

The sources of data for this study were online databases, governmental agriculture institutions, and field surveys. Since most IBCI schemes protect farmers against crop losses caused by adverse weather events, the study used satellite data that are capable of providing crop and weather-related information at the level of smallholder crop fields. Crop losses, however, can also be induced by other numerous inputs including certain biological, edaphic, socioeconomic, and managerial factors, some of which are not measurable with satellites. Therefore, it was necessary to account for these factors in order to highlight the limitations of remote sensing and its untapped opportunities in IBCI. Information about these non-weather factors was collected through field surveys by following the sampling techniques specified in the main chapters of the thesis.

1.9. Ethical clearance

The University of Fort Hare's Research Ethics Committee granted ethics approval for this study under the project number: KAL021MAS01. The student understood that all research should be conducted following the ethical standards of research. The student obtained permission to conduct research in the study area following a signed consensual agreement, which entailed a confidentiality statement and an explanation of risks and benefits of this research. Datasets obtained from various institutions were utilized in a manner that is consistent with the institutions' data-use policies. The research was undertaken with scientific rigour and social and moral responsibility.

1.10. Outline of the thesis

The thesis is organized according to the five objectives outlined in section 1.4. Each of the main chapters is an independent manuscript either published, under peer-review or to be submitted to an accredited journal for publication. The main chapters are preceded by chapter 1, which outlines the background and basis for this research.

Chapter 2 is a review of literature whose objective was to highlight the contributions, limitations, and opportunities that exist in IBCI for satellite remote sensing. The chapter does this by reviewing major IBCI programs in Africa, the types of contracts provided by these programs, the datasets they use in their systems, the challenges they face and untapped opportunities. The chapter also reviews peer-reviewed studies that investigated the usability of satellite data in IBCI and the data analysis methods these studies used. The review shows that rainfall is not the only major limiting factor on yields as often assumed by most IBCI approaches. The review also shows that remote sensing-based IBCI is exposed to basis risk. The basis risk is associated with non-weather crop inputs, the lack of calibration data, poor contract design, and imperfect correlations between satellite-based indices and crop yields. Among the different agrometeorological variables that are often used as drought indicators, the review highlights soil moisture as one of the most important factors affecting crop yields. Lastly, the review shows that many remote sensing datasets and data analysis techniques, such as mapping of farm parcels and crop types, machine learning and computer vision, have not been sufficiently applied in IBCI.

The objective of chapter 3 was to compare the performances of combined Sentinel-1 and Sentinel-2 data as well as the performances of different machine learning algorithms in the mapping of smallholder crop farming areas. This study was prompted by the lack of IBCI reports that cover crop mapping. The findings show that the combination of optical and Synthetic Aperture Radar (SAR) data and ensemble learning improve mapping accuracy. The chapter concludes that insurers can use these methods to map the exact locations, distributions and sizes of maize fields.

The objective of chapter 4 was to identify yield-determining factors that can be bundled or linked with IBCI. This chapter used multi-source data to identify critical yield-determining factors that may not be measurable with satellites. Results show that the influence of agro-meteorological variables on yields is significant even under optimal weather conditions, but this influence is largely associated with planning and sowing dates. Results also show that non-weather factors such as agronomic inputs and socioeconomic conditions affect crop yields significantly. The chapter then concludes that a holistic scheme that links IBCI with non-weather crop inputs is potentially the best approach to risk management and risk reduction. The chapter also concludes that linking IBCI with advisory services and input supplies to address the impact of non-weather variables on crop losses could enable remote sensing to accurately model, quantify or isolate weather-induced crop losses.

The objective of chapter 5 was to explore the feasibility of estimating soil moisture from a combination of optical and SAR data. This chapter was motivated by the findings of chapters two and four, which showed that soil moisture content is one of the important yield-determining factors. Results show that the methods used in this study better estimate soil moisture when leaf area and plant height are low than when they are high. Another important finding is that the combination of optical indices and SAR polarimetric channels performs better than the standalone use of SAR channels. This chapter concludes by providing recommendations that could be used to improve the method used in this investigation.

The objective of chapter 6 was to use the findings of chapters two to five to develop a feasible prototype IBCI model for SFS. The chapter does this by first comparing the performances of two satellite rainfall datasets in the mapping of rainfall patterns at different spatial and temporal scales. The study then takes the better of the two datasets and uses it in conjunction with other climate data to determine crop water requirements of maize in the study area. Results show that the mid-season stage of maize growth is the most critical phase, followed by the development stage, and the late-season stage. This information is then used to determine index thresholds and an IBCI payout structure. The proposed IBCI also incorporates crop area mapping and must be linked or bundled with factors that influence crop yields. .

Chapter 7 links and synthesizes the findings of the preceding chapters, highlights the limitations of the thesis, and provides recommendations for future research and policy formulation.

Chapter 2 : Literature review

Abstract

Index-based crop insurance (IBCI) has the potential to reduce the devastating impacts of adverse weather on smallholder agriculture. For more than a decade, research and innovation have sought to address the challenges confronting traditional insurance systems by exploring the feasibility of using IBCI to hedge Africa’s smallholder farming systems against weather risks. This study is a review of how remote sensing (RS) has been used in Africa to develop IBCI. It reviews scientific articles and other published reports and focuses on the achievements, challenges and opportunities that have arisen due to the use of RS in IBCI. This review shows that, since the introduction of IBCI in Africa around 2005, there has been an increasing trend of studies and insurance schemes incorporating RS in IBCI. RS has enabled improved demarcations of unit areas of insurance and enabled coverage of inaccessible areas. The literature also shows that IBCI schemes and research studies continue to leverage satellite based estimates of rainfall, soil moisture, vegetation vigour, and evapotranspiration to improve IBCI. However, there is still persistence of basis risk that is associated with non-weather yield-determining factors, the lack of calibration data, imperfect correlations between indices and crop yield data, and poor contract design. Despite these challenges, there are still untapped opportunities that merit further testing of RS in IBCI. The study discusses these in detail and provides recommendations for the way forward.

2.1. Background

Since agriculture is the main source of livelihood in the developing countries (Mehta et al., 2010; Asfaw et al., 2015), the nexus between agriculture, food security, and the uncertainty of weather has become a topical issue (FAO, 2008; Sasson, 2012; Adhikari et al., 2015; Sultan et al., 2019). About 95% of the farmland in Sub-Saharan Africa (SSA) is rain-fed (Wani et al., 2009) and vulnerable to extreme weather events (van Asten et al., 2011; Masih et al., 2014; Mpandeli et al., 2015; Delacote et al., 2019; Sultan et al., 2019; Frischen et al., 2020). Farmers in this region hardly invest in improved agricultural technologies because of undercapitalization and the lack of access to credit (Langyintuo, 2020). Lack of collateral and the high susceptibility of SSA to adverse weather expose lenders to credit risk which makes it difficult for farmers to access loans (von Negenborn et al., 2018; Möllmann et al., 2020). Consequently, farmers tend to minimise risks by opting for low-cost, less risky, and less profitable crops, low-use of fertilizers, cultivation of smaller areas, and abandonment of crop production for livestock farming (Alem et al., 2010; Manyevere et al., 2014; Salazar-espinoza et al., 2015). These risk-avoidance strategies together with the lack of investment in improved agronomic inputs often translate into low yields and chronic poverty. In the recent years, researchers, insurers, and governments have been exploring the feasibility of reducing the impacts of adverse weather on agriculture by piloting index-based crop insurance (IBCI) (Di Marcantonio and Kayitakire, 2017; Ntukamazina et al., 2017; Singh and Agrawal, 2020; Wang et al., 2020).

IBCI is potentially better than traditional claim-based insurance because it is less costly and less susceptible to moral hazard and adverse selection (Barnett and Mahul, 2007; Miranda and Farrin, 2012). Moral hazard is a situation where the insured parties deliberately expose themselves to risks in order to increase the chances of receiving payouts (Carter et al., 2014). Adverse selection is a situation where farmers who are exposed to high-risk take up insurance more frequently than others because they perceive more profits from insurance (Carter et al., 2014; He et al., 2019). Furthermore, traditional insurance schemes have high administrative expenses and high premiums because they involve strenuous and time-consuming on-field surveys (de Leeuw et al., 2014). IBCI, on the other hand, is less susceptible to moral hazard and adverse selection because it calculates compensation from objectively derived indices so that neither the farmer nor the insurer is able to manipulate the system (Barnett and Mahul, 2007). These indices may be derived from weather variables like rainfall, temperature, soil moisture, and weather-related vegetation indices or from average yields of a defined area. This makes IBCI less costly because the data from which indices are calculated come from satellites, weather stations, and areal yield statistics (Barnett and Mahul, 2007).

Despite these announced advantages, IBCI is impaired by “basis risk”. Basis risk which arises when farm losses as measured by the index and the real losses on the farm do not correlate (Collier et al., 2009). There are three types of basis risk, which are temporal, spatial, and product basis risk. Temporal basis risk arises when the index fails to predict potential losses by virtue of being calculated outside the crop’s critical growth stages (Dalhaus et al., 2018). Spatial basis risk arises when the index fails to capture geographical variation in the parameter measured by the index (Hazell et al., 2010). Product basis risk arises when the selected index is poorly correlated with yields / losses due to poor product design and inappropriate index selection (Boyd et al., 2019). Given this scenario, insured farmers may end up being uninsured because of the index’s failure to identify afflicted clients. Likewise, insurers also face the same fate because they can pay unaffected farmers due to chance inclusion in the areas affected by adverse weather. Basis risk may be unavoidable when IBCI depends on weather station data, especially if the weather station network is sparse (Barnett and Mahul, 2007). However, several techniques have been proposed to minimise basis risk, one of which is the use of satellite remote sensing (RS) data in lieu of weather station data (Black et al., 2015, 2016; Kölle et al., 2021).

Although the applicability of RS in IBCI was first explored in the 1970s (Towery, 1980), pilot IBCI projects in Africa were implemented around 2004 / 2005 (Makaudze, 2005; Makaudze and Miranda, 2008; Di Marcantonio and Kayitakire, 2017). However, the developments, challenges, and opportunities that have arisen due to the use of RS in Africa’s IBCI systems are not well documented. Alderman and Haque, (2007) reviewed pilot projects in East Africa by evaluating how insurance protects low-income countries from weather shocks. Tadesse et al., (2015) reviewed conceptual issues around weather index insurance by drawing lessons from challenges faced by pilot projects in Ethiopia, Kenya, and Malawi. Ntukamazina et al., (2017) provided a generic review of insurance products available in SSA, factors that influence farmers to buy insurance, challenges they faced in accessing these products, and interventions that could improve uptake of the products. Di Marcantonio and Kayitakire, (2017) also adopted a similar approach by reviewing pilot insurance projects across Africa in order to identify factors that constrain the scalability of IBCI. Ntshakira-Rukundo et al., (2021) provided a review of factors that influence the uptake of insurance and how uptake may be improved. The commonality of these reviews is the absence of a thorough review of the role of RS despite its documented potential.

Identifying gaps that need to be filled and successful initiatives that require reinforcement will be a big step towards making insurance available, affordable, and more effective. For RS to make a meaningful contribution to IBCI, there is need to better understand what it has done, what it has not done and what it needs to do. This need explains why this review was conceived and undertaken. This review is an inventory of readily accessible peer-reviewed studies that applied remote sensing in crop insurance around Africa, but also includes other relevant publications and reports. The aim of this review is to identify gaps that need to be bridged and opportunities that merit further exploration in the use of RS in IBCI.

2.2. Review methods

To identify relevant publications for review, the author systematically queried online science databases including Scopus, Web of Science, ScienceDirect, Google scholar, and websites of the organizations involved in the development and administration of IBCI in Africa. Boolean searches were used to retrieve peer-reviewed papers and other reports containing the following key terms: "remote sensing", "satellite", "remotely sensed", "agriculture", "agricultural", "smallholder", "Africa", "index insurance", "basis risk", and names of countries and organizations providing IBCI. Search results were filtered and relevant articles were selected by assessing article title, abstract, methods, and other contents of interest.

2.3. Results

The purpose of this review was to highlight achievements, challenges and opportunities that are relevant to the use of remote sensing in IBCI. Therefore, exhaustive reporting of all the projects and organizations providing IBCI in Africa is not within the scope of this review. Section 2.3.1 summarizes some of the major programs that are/were providing IBCI in Africa.

2.3.1. IBCI projects in Africa

Between 2004 and 2005, MicroEnsure partnered with the World Bank and the World Food Program (WFP) to develop Africa's first agricultural index insurance (Mapfumo, 2008; World

Bank, 2011; Makaudze, 2018). Although the contracts were initially provided in Malawi, the scheme was later extended to Kenya, Ethiopia, and Tanzania (Tadesse et al., 2015). In this program, IBCI was a component of a holistic risk management package because it covered specific risks, particularly drought, and specific crops which were maize and groundnuts (Osgood et al., 2007). Indices were derived from rainfall data collected through weather stations. The scheme covered different stages of the crop growth cycle. Index thresholds were based on approximations of crop water requirements. The program was discontinued because the insured farmers did not understand how insurance operates, premiums were unaffordable, weather station networks were poor, and there was insufficient historical data (Tadesse et al., 2015; Miranda and Mulangu, 2016; Makaudze, 2018).

- *The R4 Rural Resilience Initiative*

The R4 Rural Resilience Initiative (R4) was established by the WFP and Oxfam America in 2009 under the name, Horn of Africa Risk Transfer for Adaptation (HARITA). By 2020, the R4 had administered three crop insurance schemes, namely; Weather-Index Insurance (WII), Area-Yield Index Insurance (AYII), Hybrid-Index insurance (HII). The WII covers multiple rainfall-caused perils including late onset, early cessation, drought, dry spells, and floods. It uses rainfall indices derived from weather station and satellite data. The AYII uses a sampling method at the end of the season to determine losses by measuring and comparing the realized average yield of an area with the historical average yield for that area. The HII covers multiple perils, including pests and diseases by combining WII and AYII. These products use weather station data, the Africa Rainfall Climatology version 2.0 (ARC2) data, the Normalised Difference Vegetation Index (NDVI), and yield data. Insured crops include wheat, maize, teff, millet, beans, cowpeas, and sorghum. The R4 continually develops its insurance products through consultations with various stakeholders including farmers, funders, research institutes, reinsurers, humanitarian organizations, extension service providers, input suppliers, and governments (Sharoff et al., 2015). In 2020, the program covered nearly 180 000 farmers in Ethiopia, Burkina Faso, Senegal, Kenya, Madagascar, Malawi, Mozambique, Zambia, Zimbabwe, and Bangladesh (Sharoff et al., 2015; WFP, 2020). The use of satellite data has helped the R4 to reach secluded areas. Challenges to the project have been (1) limited records of historical rainfall and yield data, (2) basis risk, (3) lack of local capacity for index design, (4) lack of data management tools, (5) lack of distribution channels and, (6) lack of financial education.

- *Agriculture and Climate Risk Enterprise (ACRE)*

The ACRE was launched in 2009 under the pilot name, Kilimo Salama. It provides WII for drought, floods, and storms in Kenya, Rwanda, Tanzania, Uganda, Ghana, Malawi, Senegal, and Mozambique (Bulte et al., 2020). ACRE is funded by the Syngenta Foundation and the Global Index Insurance Facility. The sources of data for this program are weather stations, satellites, and area yield statistics from governments. The program has recently introduced AYII to minimise basis risk. It also provides a hybrid solution that covers idiosyncratic perils such as hail, storms, frosts, pests and diseases. Efforts are underway to incorporate soil moisture measurements in ACRE's IBCI products. Insured crops include maize, beans, wheat, sorghum, millet, soybeans, sunflower, coffee, and potatoes (Greatrex et al., 2015). By 2018, ACRE had insured a cumulative 1.7 million farmers in Africa. The program ascribes its achievements to the use of different datasets and the partnerships that the project has built with different stakeholders including research institutes, mobile technology service providers, input suppliers, farmer associations, governments and advisories. Sibiko and Qaim (2020) report that insured farmers are producing higher crop yields because insurance has enabled them to adopt improved inputs. Despite these achievements, it is reported that lack of historical data has negatively affected product design and accuracy (ACRE, 2020).

- *African Risk Capacity*

Although the African Risk Capacity does not particularly offer insurance directly to smallholder farmers, it is one of the biggest index insurance schemes in Africa. The African Union founded the African Risk Capacity in 2012. In 2013, the African Risk Capacity established ARC Ltd to provide insurance to participating African states. ARC Ltd groups countries into risk pools that consist of countries with diverse climatic characteristics. The program uses Africa Riskview as its main software system and the Water Requirements Satisfaction Index (WRSI) as its main drought index. This system translates satellite-based rainfall estimates into real impacts of drought on agriculture. It uses multiple datasets including Rainfall Estimates version 2.0 (RFE2), ARC2, Tropical Applications of Meteorology using SATellite data (TAMSAT), Climate Hazards Group InfraRed Precipitation with Station (CHIRPS), NDVI from the Moderate Resolution Imaging

Spectroradiometer (MODIS), Fraction of Absorbed Photosynthetically Active Radiation (Fapar) and Fraction of Vegetation Cover (FCOVER) from Proba-V, evapotranspiration, and Standardized Precipitation Index (SPI) (African Risk Capacity, 2019). By 2020, ARC Ltd had paid out indemnities to several countries including Madagascar, Zimbabwe, Cote d'Ivoire, Mauritania, Malawi and Niger, benefiting millions of people. However, Awondo, (2019) reports that the ARC Ltd's risk pools need further refinement to minimise systemic risk. There are also reported instances of rainfall overestimation by the ARC Ltd system which, in the past, have resulted in the scheme failing to indemnify some of its client states (Johnson, 2021).

- *PULA*

PULA collaborates with governments and input suppliers to provide satellite-based WII and AYII in Kenya, Uganda, Tanzania, Rwanda, Malawi, Nigeria, Zambia, Ethiopia, Senegal, and Mali (Raithatha and Priebe, 2020). By 2020, the program had reached a cumulative 3.4 million clients. Despite this achievement, reports show that the AYII is costly and logistically burdensome compared to the WII which PULA now considers more suitable for its clients (Hernandez et al., 2018). Recently, the program assessed the performances of different indices derived from MODIS, Sentinel, Landsat, and CHIRPS in explaining rice and maize yields (Hernandez et al., 2018). PULA reports that although the tested indices could not explain farm-level crop yields, they performed better at coarser spatial scales. Informed by these findings, the program was able to reduce the cost of yield sampling by defining new unit areas of insurance (UAI) (Raithatha and Priebe, 2020).

- *Other IBCI projects in Africa*

Between 2012 and 2016, Japan International Cooperation Agency (JICA) and Oromia Insurance Company introduced a satellite-based WII in Ethiopia (JICA, 2018). The index is based on a 1-km NDVI product measured over 10-day periods (Belissa et al., 2019). It insures maize, wheat, barley, and teff against drought over the seedling and flowering stages of crop growth (Belissa et al., 2019). Belissa et al., (2019) ascribe low uptake of this insurance to contract ambiguity and basis risk.

Mayfair Insurance and Risk Shield Ltd provide satellite-based WII for cotton and maize in Zambia (Arce, 2016). The crops are insured against drought and flood-prone periods of the crop growth cycle. Key issues that have been reported in this program are basis risk and field surveys that raise operational costs and premiums. Information about other projects can be found in the following publications; Greatrex et al., (2015); Miranda and Mulangu (2016); Di Marcantonio and Kayitakire (2017); Ntukamazina et al., (2017); Ntshakira-Rukundo et al., (2021).

2.3.2. Geographic units of insurance

Most of the IBCI contracts piloted in Africa insure farmers against weather risks. This requires the contract design to factor geographical variability of weather to ensure that the measured index reflects the growing conditions around the insured farms. A weather station-based contract, for example, covers farms that lie within a certain distance from a reference weather station. This type of contract assumes that geographical variability of weather within the defined distance is insignificant if there is variability at all. For example, some schemes delineate each UAI as a group of farms that lie within 20 to 30 km from the reference weather station (Osgood et al., 2007; Mapfumo, 2007, 2008; Chen et al., 2017). Other schemes delineate UAI based on agro-ecological characteristics where a UAI, for example, is an area with a homogenous topography, soil, production output, vegetation or climate (Hernandez et al., 2018; De Oto et al., 2019; Valverde-Arias et al., 2019).

PULA defines UAI based on agro-ecological characteristics by using yield and satellite data coupled with machine learning (ML). They report that using this approach to define UAI reduces administrative costs (Hernandez et al., 2018). Alternatively, IBCI can be provided at the village, administrative, and national scales. The R4 program provides village-level contracts that are calibrated based on crop calendars and rainfall patterns of insured villages (WFP, 2020). Other IBCI schemes are provided at the level of government administrative units such as sub-national regions or districts (ACRE, 2020). ARC Ltd offers insurance at the national level by grouping countries into risk pools. This system leverages Africa's diverse environments such that the occurrences of catastrophic events in a given year do not affect every country in a risk pool. This approach enables ARC Ltd to manage risk with less funds than countries would on their own (African Risk Capacity, 2017; Awondo, 2019). The geographic scale at which IBCI is provided has implications on the spatial resolution and spatial coverage of the datasets used in measuring the index. Therefore, the RS datasets used must be able to provide measurements at the required geographic scales to avoid or minimize spatial basis risk.

2.3.3. Times of the season over which insurance indices are measured

The insurance index can be measured as a seasonal average, seasonal total or as an accumulated amount over a specific stage of crop growth. The contract has to specify the measurement period, which must be agreed upon by both the insurer and the insured. An index that is based on seasonal totals or seasonal averages may fail to capture adverse events during the critical growth stages of the crop, which may expose farmers to basis risk. The crop may receive the desirable amount of cumulative rainfall over the entire season while receiving no rainfall during the crucial growth periods, resulting in the index failing to trigger payouts. Some of the IBCI schemes foresaw this and elected to measure rainfall over the crucial stages of crop growth. For example, the WFP identified the emergence and the tasseling stages of maize as the most vulnerable to drought and assigned weights to the different growth stages of maize based on their importance (Hess and Syroka, 2005). The HII and the WII of the R4 cover germination failure and dry spells (WFP, 2020). ACRE reported that they would introduce flexibility to their WII products by aligning the products with crop phenology and by adding extra days to the index measurement periods to reduce basis risk (ACRE, 2020). In the final analysis, IBCI must be based on reliable datasets that provide measurements at the required temporal frequencies to avoid or minimise temporal basis risk.

2.3.4. Remote sensing data and methods used in IBCI

- *Satellite-based rainfall estimates*

Since rainfall is spatially variable, satellite-based rainfall estimates (SRFE) provide better spatial coverage than ground-based rain gauges. Although rain gauges take direct measurements of rainfall, they can only take these measurements within their immediate surroundings. Farming areas with sparsely distributed gauges may not be well represented by a rain gauge-based index. As a result, rainfall in areas with no rain gauges is usually estimated from nearby gauges through interpolation, which may produce inaccurate estimations (Mendelsohn et al., 2007). However, SRFE can provide spatially continuous measurements covering inaccessible areas. A number of SRFE products have been developed over the years; however, this section focuses on the ones that have been tested in IBCI and over Africa. These include CHIRPS, ARC2, TAMSAT, RFE2, and the Global Precipitation Climatology Project (GPCP) (Table 2.1).

Table 2.1: SRFE products used in Africa for IBCI

Data	Temporal coverage	Time step (s)	Spatial resolution	IBCI articles *
CHIRPS	1981 – present	daily, pentadal, monthly	0.05°	1; 2; 3; 4; 5; 6; 7; 8
TAMSAT	1983 – present	daily, pentadal, monthly,	0.0375°	3; 9
ARC2	1983 – present	daily	0.1°	1; 3; 5; 10; 11
RFE2	2001 – present	daily	0.1°	12
GPCP	1979 – present	Monthly	2.5°	13

* [1]Enenkel et al., (2019); [2]Eze et al., (2020a); [3]Tarnavsky et al., (2018); [4]Enenkel et al., (2017); [5]Osgood et al., (2018); [6]Enenkel et al., (2018); [7]Eze et al., (2020b); [8]Blakeley et al., (2020); [9]Black et al., (2016); [10]Awondo, (2019); [11]Awondo et al., (2020); [12]African Risk Capacity (2019); [13]Siebert (2016).

Table 2.1 shows that CHIRPS is the most researched rainfall product, followed by ARC2, TAMSAT, and then RFE2 and GPCP. Table 2.2 below shows some of the highlights reported about these datasets in the African IBCI literature.

Table 2.2: Highlights from studies that used SRFE

Article	Analysis	Highlights
Enenkel et al., (2019)	Assessed the ability of CHIRPS, ARC2, CCI-SM, and ESI to represent agricultural drought conditions over Africa.	CHIRPS showed better performance than ARC2. Both performed poorer than CCI-SM.
Tarnavsky et al., (2018)	Assessed the sensitivity of WRSI to different rainfall datasets	CHIRPS performed better than ARC2 and TAMSAT. WRSI with CHIRPS explained 52 to 61 % of maize yield variation.
Enenkel et al., (2017)	Assessed SWC and rainfall’s relationship with vegetation health (NDVI) for drought assessment.	Stronger correlation of NDVI with CCI-SM (0.76 and 0.82) than with CHIRPS (0.54 and

		0.64) shows the importance of SWC in IBCI.
Enenkel et al., (2018)	Assessed correlations between CCI-SM, CHIRPS rainfall, ESI, and their ability predict crop yield.	The datasets show matching spatial / temporal patterns. Yield prediction accuracy was better when the data were combined (R^2 = 0.43).
Black et al., (2016)	Assessed the relationship of TAMSAT with crop loss, the relationship of rainfall with SWC, and the effect of spatial and temporal aggregation of satellite data on local conditions of rainfall.	TAMSAT explained 65% of yield losses. Relationship between meteorological (rainfall) and agricultural drought (SWC) was significant. Spatial and temporal aggregation must be performed at appropriate scales.
Siebert, (2016)	Assessed relationships between agricultural production and different indices including GPCP.	Correlation between the GPCP and agricultural production was weak (figures are not reported by the author).

- *Satellite-based soil moisture estimates*

One of the most important agrometeorological variables is soil water content (SWC). SWC is one of the most important factors affecting the growth and development of crops and natural vegetation. Soil water is one of the basic requirements for seed germination and plant emergence (Kozlowski, 1972; Itabari et al., 1993; Shaban, 2013). Likewise, the vegetative stages of crop growth need water for leaf area development, stem elongation, and biomass accumulation (Çakir, 2004). The reproduction stages of the maize including grain filling, kernel and ear development, are greatly influenced by the amount of available soil water (Yang et al., 1993; Ge et al., 2012; Mi et al., 2018). This means that SWC is a critical determinant of crop yield especially in Sub-Saharan Africa where crop farming is predominantly rain-fed. SWC has to be constantly monitored at appropriate spatial scales to enable mitigation of the adverse effects of unpredictable soil water deficits because SWC is the major indicator of agricultural drought. Agricultural drought is defined as the lack of sufficient SWC to support the development of crops (Mannocchi et al., 2004; Łabędzki and Bąk, 2015; Martínez-Fernández et al., 2016).

Like gauge-based measurements of rainfall, ground-based measurements of SWC do not accurately represent spatial variations in SWC because they are point-based. As a result, satellite-based SWC products have been developed over the years and research continues to explore ways to improve the available SWC products (Petropoulos et al., 2015; Karthikeyan et al., 2017; Sharma et al., 2018; Myeni et al., 2019; Rodríguez-Fernández et al., 2019). The main advantages of satellite data are cost effectiveness and their ability to cover large spatial scales including inaccessible places at the required temporal scales. Passive microwave sensors measure naturally emitted radiation from the surface of the earth in the form of radiance temperatures, while active microwave sensors emit focused radiation towards the earth's surface which they then sense back as backscatter signal (Karthikeyan et al., 2017a; Sharma et al., 2018). SWC is measurable with microwave sensors because the dielectric properties of dry and wet soil influence the emitted radiation received by passive sensors and the backscattered signal received by active sensors. Unlike optical radiation, microwaves can penetrate clouds and are less affected by atmospheric effects (Petropoulos et al., 2015).

A number of studies in Africa have investigated the usefulness of the Climate Change Initiative Soil Moisture (CCI-SM) dataset for IBCI. The CCI-SM dataset provides daily measurements of SWC at a spatial resolution of 0.25° (~28km), covering the period from 1978 to 2020 (Dorigo et al., 2017; Gruber et al., 2019).Enenkel et al., (2017) assessed the relationship between CCI-SM and vegetation health. They found correlations of 0.38 to 0.76 between SWC and vegetation anomalies, while correlations were 0.33 to 0.82 with lagged vegetation response. Based on these findings, the study recommended the use of satellite SWC data to complement other datasets that are used in IBCI. Enenkel et al., (2018) assessed spatial and temporal correlations between CCI-SM, CHIRPS rainfall, and the evaporative stress index (ESI), and the ability of these variables to predict crop yield. Correlations between the datasets were temporally and spatially variable, while their ability to predict yield depended on the time of the season and on how the datasets were combined. Enenkel et al., (2019) assessed the performance of the ESI, CCI-SM, CHIRPS, ARC2 in detecting agricultural drought in Africa. They found higher correlations between SWC and vegetation vigour, concluding that satellite-measured SWC is a potential indicator of agricultural drought.

Osgood et al., (2018) used the CCI-SM with other datasets to predict historical drought events in Ethiopia. These data were compared with farmer recollections of historical droughts. In the final analysis, the study recommended the incorporation of SWC in IBCI. In other studies, Black et al., (2016) used the Joint UK Land Environment Simulator (JULES) and TAMSAT rainfall ensembles at 0.5° spatial resolution to demonstrate how the relationship between rainfall and soil moisture reflects the development from meteorological to agricultural drought in Zambia, while Tarnavsky et al., (2018) found low correlations between WRSI-simulated SWC and maize yields in Tanzania.

- *Satellite-based mapping and monitoring of crops and vegetation*

In addition to rainfall and SWC, satellite sensors can also measure crop health. Spectral wavebands and derivative indices are sensitive to plant parameters including leaf area index, canopy cover, chlorophyll content, canopy height, canopy water content, biomass and grain yield (Xue and Su, 2017; Karthikeyan et al., 2020; Sishodia et al., 2020). Table 2.3 below shows RS datasets that have been used to measure vegetation parameters for IBCI.

Table 2.3: Satellite data used to measure vegetation parameters

Sensor	Spatial resolution	Time step (s)	Index	IBCI articles *
AVHRR	8km, 4km	Bimonthly, 10 days	NDVI, TCI, VHI	1; 3; 9; 6; 13
MODIS	250m, 500m, 1km, 0.05°	8 days,16 days	EVI, NDVI, GNDVI, LAI, GPP, NDMI, FPAR	1; 2; 5; 7; 10
Sentinel-2	10m, 20m	5 days	VIS, IR, NDVI, NDBI, SAVI, EVI2	4; 11; 12
Sentinel-1	5m by 20m	6 to 12 days	VV, VH	11
Worldview-1 and 2	0.5m, 0.46m	1.1 days	VIS, IR	8

* [1]Enenkel et al., (2019); [2]Eze et al, (2020); [3]Siebert (2016); [4]Okeyo and Mulaku, (2020); [5]Enenkel et al., (2017); [6]Möllmann et al., (2020); [7]Osgood et al., (2018); [8]Neigh et al., (2018);

[9]Hochrainer-Stigler et al., (2014); [10]Kenduiywo et al., (2020); [11]Masiza et al., (2020); [12]Masiza et al., (2021); [13]Makaudze and Miranda, (2010)

Table 2.3 shows that the most used satellite datasets for vegetation assessment are MODIS and AVHRR spectral indices, followed by Sentinel-2, then Sentinel-1 and Worldview. Table 2.4 below summarizes some of the findings and highlights reported in the IBCI studies.

Table 2.4: Highlights from studies that measured vegetation parameters using satellite data

Article	Analysis	Highlights
Makaudze and Miranda, (2010)	Assessed the feasibility of IBCI for maize and cotton using AVHRR NDVI and rainfall data.	NDVI-based indemnities (0.40 to 0.90) were better correlated with yield losses than rainfall-based indemnities (0.25 to 0.70).
Enenkel et al., (2019)	Assessed the ability of MODIS and AVHRR NDVI, SWC, and ESI to represent agricultural drought conditions over Arica.	NDVI and SWC correlated by > 0.8. The authors concluded that the data are capable of capturing agricultural drought conditions over Africa
Eze et al, (2020a)	Estimated crop yield and designed IBCI using 250 m MODIS NDVI and CHIRPS rainfall.	Correlations between NDVI and yield were between 0.21 and 0.64 depending on the time of the season. NDVI was the closest proxy to crop losses. The authors recommended the use of NDVI in IBCI design.
Siebert, (2016)	Assessed the relationship between crop yield and different indices including the 8 km AVHRR NDVI.	Correlations between NDVI and crop yield were weak and not reported by the author.
Okeyo and Mulaku, (2020)	Mapped crops and designed IBCI using Sentinel-2.	69.3% classification accuracy. Insurance design produced promising results, but it still requires further refinement.
Enenkel et al., (2017)	Assessed the relationship between 500 m MODIS EVI, rainfall and SWC.	Stronger correlation of EVI with CCI-SM than with CHIRPS. Maximum

		correlations between SWC and EVI were 0.76 and 0.82 depending on the time of the year.
Möllmann et al., (2020)	Assessed the performance of AVHRR vegetation indices in explaining credit risk.	Correlations of credit risk with VHI, TCI, and VCI were between 0.06 and 0.52 depending on the loan repayment delay time. Largely, the indices explained credit risk.
Osgood et al., (2018)	Compared farmer-reported historical drought events with 250 m MODIS NDVI and EVI and other satellite datasets.	Farmer-reported historical drought events were reflected by NDVI and EVI.
Neigh et al., (2018)	Mapped smallholder crop area with Worldview-1and2.	Mapping accuracies were superior to those obtained in studies using moderate resolution data.
Hochrainer-Stigler et al., (2014)	Assessed the relationship between 8 km AVHRR VHI and yield. Insurance design.	VHI explained 60% of teff yield variation. A real world IBCI could not be constructed on these results. Model fit required improvement by including other variables related to yield.
Kenduiywo et al., (2020)	Estimated yield using 500 m MODIS indices including NDVI, GNDVI, LAI, GPP, NDMI, FPAR	The models explained 69% and 70% of maize yield variation.
Masizet al., (2020)	Assessed the combination of Sentinel-1 and Sentinel-2 and model ensembles in crop mapping	97.71% accuracy with sensor fusion and 96.06% with model ensembles. Combining optical and microwave data improves accuracy. Combining classifiers also improves accuracy.
Masiza et al., (2021)	Predicted maize yield using Sentinel-2 MSI, NDVI, EVI2 and other non-RS data.	MSI was one of the most important yield-determining factors.

2.3.5. Data analysis methods used in IBCI

Data analysis methods play an important role in determining the degree to which datasets reflect facts in the real world. The reviewed publications show that the IBCI programs and the research studies use traditional statistics, ML, and process-based physical models. There is an overlap between ML and statistics; however, traditional statistics mainly involves making inferences about relationships between variables, whereas ML focuses on making optimum predictions (Bzdok et al., 2018). Physical models are simulations of physical processes, characteristics, and laws of the modelled system (Sparling, 2016). Traditional statistical techniques including correlation and linear regression analysis are the most used techniques in the literature and in the IBCI schemes. The review shows that these methods are used to relate indices such as rainfall, soil moisture, evapotranspiration and vegetation indices with crop yields, crop losses, and drought. Correlation analyses are also used to assess the degree of correspondence between different datasets such as gauge-based rainfall data and SRFE and between different variables such as SWC, rainfall, and vegetation vigour (see Table 2 and Table 4).

The most common applications of ML in RS and agriculture include classification, clustering, regression, and dimensionality reduction (Holloway and Mengersen, 2018). Raithatha and Priebe, (2020) report that PULA uses ML to demarcate UAI according to agro-ecological characteristics. Research studies used ML techniques comprising Binary Logistic Regression, Artificial Neural Networks, Random Forest, Support Vector Machines, Naïve Bayes, and Extreme Gradient Boosting. These methods were used to identify crop yield-determining factors (Masiza et al., 2021) and to predict variables including crop yields (Kenduiywo et al., 2020), cropped areas (Masiza et al., 2020; Okeyo and Mulaku, 2020), and frost zones (Kotikot et al., 2018, 2020).

Some of the process-based models that have been tested in IBCI include WRSI, CROPWAT, ALEXI, and JULES. The WRSI is a crop-specific water balance model used for monitoring crop growing conditions based on the availability of water to the crop and is widely used to monitor droughts (McNally et al., 2015). CROPWAT is a computer system developed by the FAO which uses soil, climate, and crop data to determine crop water requirements (Smith et al., 2002). CROPWAT was used in a research study to estimate crop yields in Ethiopia (Eze et al., 2020). ALEXI is a two-source energy balance model that calculates ET by partitioning radiometric temperature into soil and vegetation temperatures (Anderson et al., 1997, 2011). ALEXI was used in research studies to estimate ESI for drought monitoring and crop yield prediction (Enenkel et al., 2018, 2019). JULES is a land surface model developed by the UK Meteorological Office used both as a standalone model and as a component in the Unified Model, and is fully described by Best et al., (2011). Black et al., (2016) used a SRFE-driven JULES to calculate soil moisture and to describe the progression from meteorological to agricultural drought in Zambia.

2.4. Discussions and conclusions

This review study highlighted the developments, challenges, and opportunities brought about by the use of RS in Africa's IBCI. The objective of this review was to track progress, highlight obstacles, and identify untapped opportunities RS can exploit to improve IBCI in Africa. The findings show that, since the inception of IBCI in Africa around 2005, there has been an increasing trend of studies and insurance schemes incorporating RS in IBCI. The use of RS in IBCI has improved the way insurers define UAI. Notwithstanding this development, some IBCI schemes still define UAI according to administrative borders, which reportedly hikes insurance costs and negatively affects the performance of the product (De Oto et al., 2019; Hernandez et al., 2018). In addition to the improvements achieved by PULA in defining UAI, an index based livestock insurance study proposed an ecological stratification approach by which demarcations of UAI could be improved (De Oto et al., 2019). Additional research is required to improve demarcations of UAI as this could reduce spatial basis risk. Improvements could be made by categorizing cropped areas according to crop yield patterns, rainfall patterns derived from SRFEs, and vegetation and topographic characteristics.

The review also shows that RS has enabled IBCI coverage of areas with poor distributions of weather stations. SRFEs, especially CHIRPS, perform well in the mapping of spatial and temporal patterns of rainfall. The accuracy of rainfall mapping could be further improved through the synergistic use of SRFEs and rain gauge data. Many organisations including ACRE, R4, South Africa's Agricultural Research Council to name a few, are continually installing gauges and weather stations to improve the spatial coverage of weather data in Africa. On the other hand, spatial downscaling is also potentially capable of improving SRFEs to accommodate small-scale agriculture (Ceccherini et al., 2015; Gebremedhin et al., 2021).

RS's ability to provide measurements of different parameters such rainfall, SWC, vegetation health, and temperature, which are associated with crop performance, has enabled insurance schemes and research studies to leverage multiple RS datasets to improve the effectiveness of IBCI. For example, the use of multiple datasets by ARC Ltd and ACRE is intended to enhance the ability of IBCI to predict the impacts of adverse weather on agriculture. These developments have been realized because most of the satellite data used in IBCI are cost-free and easily downloadable from the internet, enabling researchers to run quick experiments, compared to other datasets that require field surveys or routine requests from agencies.

Although the use of RS in IBCI has resulted in all the developments highlighted above, there are persistent challenges. The most important of these challenges is basis risk, which is associated with (1) the disregard of non-weather perils and other non-weather yield-determining factors that are not measurable with satellites but contribute crop losses, (2) imperfect correlations between satellite-based indices and crop yields/losses, (3) lack of calibration data and (4) poor product design. Non-weather factors affecting yields are not only pests and diseases, which are covered by some insurance programs, but also include inputs such as fertilizer, seed variety, soil and topographic properties, management and socioeconomic factors (Masiza et al., 2021). It is because of these non-weather factors that satellite data are, in some cases, unable to explain yields and yield losses accurately. This is especially the case when rainwater is not the only major limiting factor. To address this limitation, IBCI programs and other stakeholders involved in smallholder farming need to investigate factors that influence crop yields for specific localities.

Black et al., (2016) pointed out that a strong influence of non-weather-related factors on yield will cause high basis risk in WII contracts. The influence of non-weather factors on yields can be quantified through variable ranking algorithms or other multivariate techniques. Information about non-weather factors could also assist insurers who might want to link or bundle insurance with inputs and advisory services. Adoption of improved inputs and good farming practices could reduce the influence of non-weather factors on crop losses, and thereby reduce basis risk in WII contracts. However, this requires reliable data on the current farming practices and crop yields. The lack thereof will necessitate collection of new data. For example, Hernandez et al., (2018) report that the PULA insurance program continues to collect new data to recalibrate and improve its insurance contracts. However, collecting new yield data using crop cut methods can be time consuming and costly. Insurers need to collaborate with governments, input suppliers, research institutions, and farmer organizations to achieve this. In South Africa, public–private partnerships are already underway. Public extension services, private farmer organizations, input suppliers, research institutions, and milling plants are all involved in smallholder cropping programs (Iortyom et al., 2018; Masiza et al., 2021). Such programs could be exploited to establish data standards and data sharing platforms.

As an alternative to calibrating insurance contracts by simply correlating crop yields with weather indices, IBCI schemes could use crop water requirements models. For example, the CROPWAT system can model water-deficit-caused yield reductions (Muhammad, 2009; Kumar et al., 2019). It is a water-driven model that focuses on crop water needs and does not require a large number of input data from non-weather yield-determining factors. It categorizes maize growth stages into initial (20, 25, or 30 days), development (30 days), mid-season (40 days), and late-season (30 days). Using this crop water requirements approach, IBCI would issue payouts when water requirements are not met over these 20-to-40-day periods. SRFE could be used as input rainfall data in the model. The applicability of the CROPWAT model in IBCI, for example, has been tested by a few studies. It was used by Eze et al., (2020) for crop yield estimation in Ethiopia and by Zendera et al., (2018) to compute evapotranspiration in Zimbabwe.

The review also shows that the IBCI schemes and the research studies are overwhelmingly dominated by medium and coarse resolution datasets. This is understandable because these are ready-to-use products that require minimal processing; they are derived from instruments with high revisit frequencies and are freely available. Besides, crop statistics for contract design and model calibration are often absent or available at coarse spatial scales. Although various techniques including spatial downscaling and the use of high-resolution (HR) and multi-sensor data have the potential to address basis risk, only a few studies have tested the effectiveness of these techniques in Africa. In India, Shirsath et al., (2020) show how RS can be used to downscale coarse yield statistics to local scales for use in IBCI. In Germany, Kölle et al., (2021) investigated the impact of HR satellite data on hedging winter wheat yields. They report that HR satellite data do reduce basis risk but not at the crop field level.

The review also shows that most studies and other IBCI reports focus more on vegetation and crop health monitoring while fewer reports cover crop area mapping. The use of RS in crop area mapping has been widely researched outside IBCI. More research on crop mapping is needed because insurers need to know the geographic distributions of farms and the amount of cropped area to determine payouts. RS can be a more objective and cost-effective source of this information than farmer reports and costly and time-consuming on-field surveys. Masiza et al., (2020) showed that smallholder crop farming areas could be mapped with high accuracy by combining SAR, optical data and machine learning ensembles. Furthermore, a similar study used these mapping techniques and demonstrated that they are very accurate ($p < 0.05$, R = 0.84) in estimating the amount of cropped area (Mashaba-Munghemezulu et al., 2021)

For SWC retrieval, the CCI-SM dataset which has been tested widely in Africa has a coarse spatial resolution, which limits its applications in smallholder crop farming. Follow-up initiatives stand to benefit by exploring other datasets like SMAP-Sentinel-1 (1-3km) which have been developed in the recent years (Lievens et al., 2017; Bai et al., 2019; Das et al., 2019). Sentinel-1 also has the potential to provide SWC measurements at high spatial resolution, with revisit frequencies of 6 and 12 days. Additional research is needed to explore how different optical datasets like MODIS, Landsat, Sentinel-2, Sentinel-3 and others can be synergistically used with Sentinel-1 and other SAR datasets to provide SWC measurements at high spatial and high temporal resolutions. This review showed that SWC data could provide a more robust index than the other indices used in IBCI.

Lastly, techniques like ML, deep learning, and computer vision are used less frequently than traditional statistics. These are robust data processing and data analysis techniques that have been shown to be potentially capable of improving agricultural index insurance products (Djerriri et al., 2018; Hernandez et al., 2018; Hobbs and Svetlichnaya, 2020; Masiza et al., 2020, 2021). Overall, progress is being made, but IBCI is still tainted by basis risk. For RS to contribute effectively in IBCI, future initiatives need to investigate and quantify the impact of non-weather crop-yield determining factors on basis risk. This could enable remote sensing to accurately isolate and quantify the impact of weather on crop losses. To improve the quality of calibration data, as stated earlier, efforts including large-scale campaigns must be undertaken to collect more crop statistics. These should be seen as opportunities that need to be exploited in order to improve the currently existing IBCI products, develop new products, and extend insurance to uninsured farmers.

Chapter 3 : Mapping of a smallholder crop farming landscape[1]

Abstract

Globally, Smallholder farming systems (SFS) are recognized as one of the most important pillars of rural economic development and poverty alleviation because of their contribution to food security. However, support for this agricultural sector is hampered by the lack of reliable information on the distributions and acreage of smallholder farms. This information is essential in not only monitoring food security and informing the markets, but also in ascertaining the amount of support that individual farmers need from governments and other service providers. There is urgent need for robust techniques that can be used to cost-effectively and time-efficiently map smallholder farms especially in Sub Saharan Africa and Asia. This study attempts to do this by systematically classifying combined optical and Synthetic Aperture Radar (SAR) data using Extreme Gradient Boosting (Xgboost). Furthermore, the study investigates the effectiveness of model stacking in improving classification accuracy. To achieve this, Xgboost is combined with Random Forest (RF), Support Vector Machine (SVM), Artificial Neural Networks (ANN), and Naïve Bayes (NB). The combined use of multi-temporal Sentinel-2 bands, spectral indices, and Sentinel-1 polarimetric channels, produced better results than the exclusive use of optical data ($p = 0.0005$). Moreover, stacking of classification algorithms based on model comparisons achieved higher accuracy than indiscriminate stacking ($p = 0.0100$). Through systematic fusion of SAR and optical data and hyper-parameter tuning of Xgboost, the study achieved a maximum classification accuracy of 97.71% and a maximum of 96.06% with model stacking. This highlights the importance of combining multi-sensor data and the robustness of combining multiple classifiers when mapping fragmented smallholder agricultural landscapes.

[1] Masiza, W., Chirima, J. G., Hamandawana, H., and Pillay, R. (2020). Enhanced mapping of a smallholder crop farming landscape through image fusion and model stacking. International Journal of Remote Sensing, 41(22), 8739-8756. https://doi.org/10.1080/01431161.2020.1783017

3.1. Background

Smallholder farming systems (SFS) dominate the 570 million farms around the world (Graeub et al., 2016; Lowder et al., 2016; Samberg et al., 2016), and are increasingly seen as a pivotal niche for economic development and food security (Aliber and Hart, 2009; Rapsomanikis, 2015; UNCTAD, 2015). However, the capacity of farmers to advance the realisation of this objective is undermined by numerous constraints. These include extreme weather events (Manderson et al., 2016; Chapagain and Raizada, 2017; Mugambiwa and Tirivangasi, 2017; Harvey et al., 2018), ineffective response to shocks (Devereux, 2007), infrastructural bottlenecks and limited capital (ASFG, 2013; Khapayi and Celliers, 2016; Von Fintel and Pienaar, 2016), poor extension services (Akpalu, 2013; Fanadzo and Ncube, 2018), inappropriate land tenure arrangements (Bembridge, 2000), and lack of good quality data among many (Carletto et al., 2013; Lowder et al., 2016; Samberg et al., 2016). Among these, data quality is very critical because of the central role of data in science. Governments, creditors, policymakers, and insurers also rely on data when allocating resources and providing services to farmers (Carletto et al., 2013; Samberg et al., 2016). The most basic data that is needed from crop farmers is spatial data, which is information on the amount of area under cultivation and the locations of crop fields.

Methods that have been used to locate and measure cropped areas include soliciting of information from farmers (FAO, 2016), use of Global Positioning Systems (GPS) (Keita et al., 2010), on-screen digitization of high-resolution satellite and aerial images (FAO, 2016) , pixel-based and object-based image classification (Dhumal et al., 2013). Amongst these, image classification is regarded as the least costly, time-efficient, and most objective technique. However, there are still unresolved complications with using this method because of the difficulties involved in spectrally discriminating different crops and vegetation (Rao, 2008). This difficulty is made worse by landscape heterogeneity in smallholder farming areas (Debats et al., 2016), irregular shapes and sizes of fields (Persello et al., 2019) and the inter-annual fluctuation of cultivated land (Nieuwoudt and Groenewald, 2003). To address these challenges, researchers are riding the wave of "big data" and machine learning exploring ways by which multi-sensor and multi-temporal data can assist in producing reliable crop maps. Since multi-temporal imagery can model phenological differences between plants, they enhance the ability to discriminate different plant types (Brisco et al., 1984; Debats et al., 2016; Zurita-Milla et al., 2017; Aguilar et al., 2018; Useya and Chen, 2019).

Combining data from multiple sensors can improve prediction or classification accuracy by taking the advantage provided by each sensor. For example, SAR is sensitive to texture and structural information of the target feature and it is less limited by atmospheric conditions than optical sensors (Kulkarni and Rege, 2020). On the other hand, optical sensors provide spectral information about the target feature over a wide range of spectral channels (Kulkarni and Rege, 2020). Studies have shown that complementary use of these datasets improves crop mapping accuracy (Forkuor et al., 2014; Dimov et al., 2017; Zhou et al., 2017; Van Tricht et al., 2018;Abubakar et al., 2020; Qadir and Mondal, 2020; Adrian et al., 2021).

However, mapping SFS does not only require optimum combinations of datasets, but it also requires optimization of mapping techniques. Although most research has been comparing capabilities of different machine learning classifiers, there is also a growing trend of studies that combine different classifiers. The practice of combining classification algorithms is predicated on the hypothesis that integrating multiple classification algorithms should produce a more powerful model than the use of standalone classifiers. However, very few studies have tested this approach in mapping SFS. Furthermore, Wu et al., (2012) noted that model stacking in particular is less popular in remote sensing applications. Few notable studies that have used this approach include Aguilar et al., (2018) who combined of Random Forest (RF), Maximum Entropy, and Support Vector Machines (SVM), and Sonobe et al., (2018), who combined SVM and RF. Useya and Chen, (2018) combined Maximum Likelihood Classification, SVM, and Spectral Information Divergence, while Salas et al., (2019) combined Generalized Linear Model, RF, Boosted Regression Trees, Maximum Entropy, and Multivariate Adaptive Regression Splines. All these studies report that combining different classifiers improves classification accuracy compared to using a single classifier. However, these studies do not specify whether the algorithms were combined systematically or not.

This study investigates the effectiveness of image fusion and model stacking in improving mapping accuracy of a smallholder farming landscape. From a stack of 40 multi-temporal Sentinel-2 bands, the study filters out unimportant bands using a variable importance algorithm. The remaining bands are then combined with different spectral indices and the vertical transmit, horizontal receive (VH) and vertical transmit, vertical receive (VV) polarimetric bands of Sentinel-1 in a series of classification trials. The study uses the Extreme Gradient Boosting (Xgboost) classifier partly because it is recently developed and has been outperforming other algorithms in machine learning competitions (Nielsen, 2016). Model stacking is performed by combining Xgboost with RF, SVM, Artificial Neural Networks (ANN), and Naïve Bayes (NB). These classifiers are first combined indiscriminately and then combined systematically. The second stacking approach is based on the theory that model stacking should be an ensemble of sub-models that are weakly correlated (Džeroski and Ženko, 2004; Merz, 1999). Essentially, this investigation is a search for optimum results through fusion of SAR and optical data and through model stacking.

3.2. Materials and methods

3.2.1. Study area

The O.R. Tambo District Municipality (ORTDM) is in the Eastern Cape Province of South Africa (Figure 3.1).

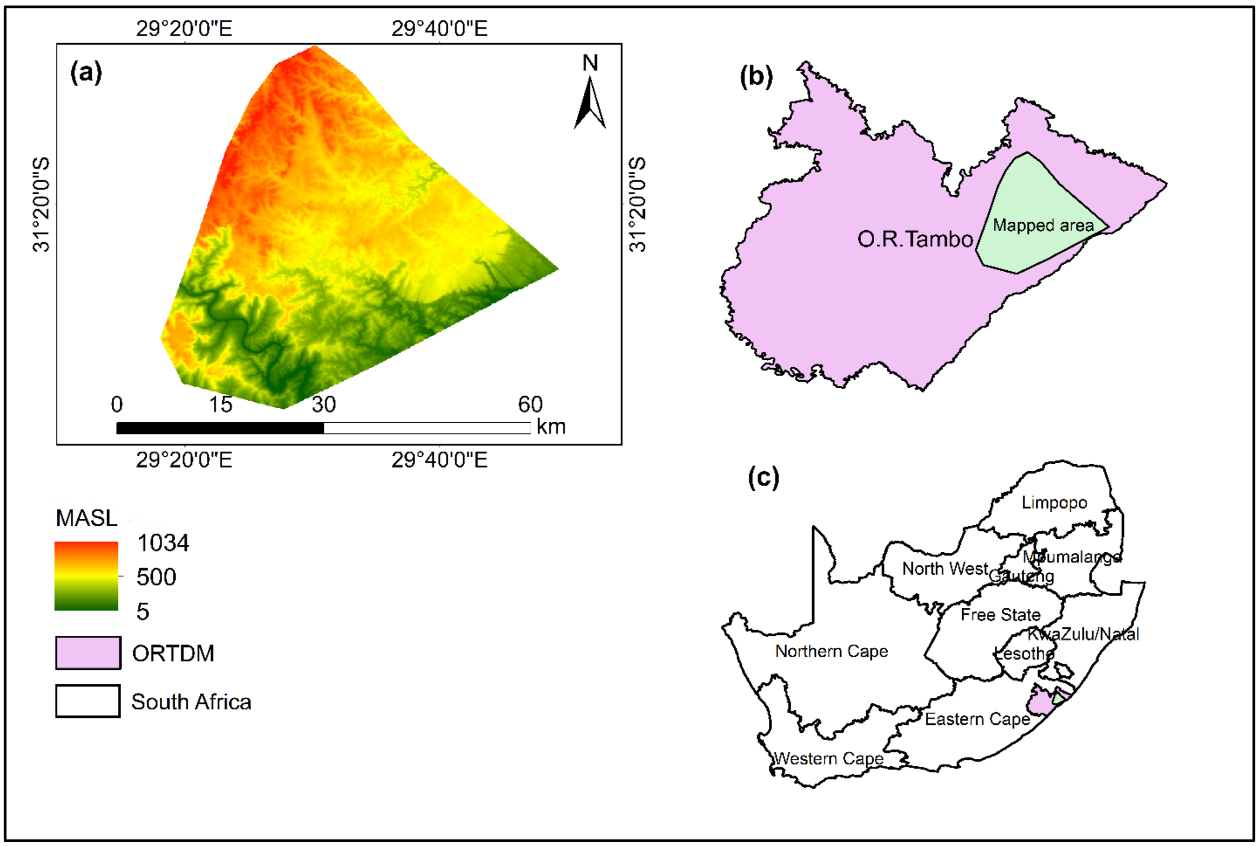


Figure 3.1: Location and elevation of the mapped area (a) in ORTDM (b), South Africa (c)

The district is part of the former Transkei homelands and the majority (94%) of its population lives in villages. The major sources of livelihood include social grants from the government, livestock farming, and crop production in the form of maize, tea, cannabis, and vegetables (ECSECC, 2014; Municipality, 2017). Most farms produce only maize, partly because they get inputs from the government and also because of the mealie-meal producing RED Hub project, which has had a significant impact on the rural income of this area (Iortyom et al., 2018). The maize crop fields range in size from a minimum of one-hectare plots to bigger collective farms in which individual smallholders cultivate their own sub-plots. Crop production intensifies between October and April with planting dates ranging between November and January depending on the commencement of the rainfall season. Although most of the farmers sell their produce to local supermarkets, animal feed retailers, and milling plants, part of this is usually destined for domestic consumption (DRDLR, 2016). Vegetables and cannabis are grown at a smaller scaler than maize, with individual fields not exceeding two hectares in size. There are also two large-scale tea farms publicly known as Majola and Magwa Tea Estates which are owned by the government. ORTDM's mean annual rainfall ranges between 900 mm and 1300 mm, while summer minimum and maximum temperatures range between 14 ºC and 19 ºC and 14 ºC and 27 ºC, respectively (Jordaan et al., 2017). The soils are largely dominated by sandy loams, sandy clay loams, and clays that are yellow to black in colour and slightly acidic (Eta and Grace, 2013; Sibanda et al., 2016). The physiography is characterized by densely vegetated undulating landscapes of 5 m to 500 m elevation along the wild coast, gentle-to moderate sloping grasslands in the interior and elevated uplands of the savannas and forests in the northern areas where elevation rises to 1500 m. The area covered in this study was purposefully selected as it captures the heterogeneity of ORTDM's physiography (Figure ***). The mapped area covers the low-lying, uneven, and densely vegetated areas of the wild coast around Port St Johns and Libode and the smooth-sloping grassy lands and savannas between Lusikisiki and Flagstaff. Unlike the other parts of ORTDM which grow only maize, the mapped area grows a variety of crops including maize, tea, vegetables, and cannabis.

3.2.2. Image compilation and pre-processing

Multi-date Sentinel-2 Level-1C and Sentinel-1 Level-1 Ground Range Detected (GRD) images were obtained from the European Space Agency's Copernicus Hub. Table 3.1 provides the details of these images.

Table 3.1: List of images providing footprint coverages of the study area

Sentinel - 2	
▪ L1C_T35HPE_A013245_20180104T080546	▪ L1C_T35JPF_A013245_20180104T080546
▪ L1C_T35JPF_A005123_20180228T075917	▪ L1C_T35HPE_A013717_20180206T081553
▪ L1C_T35HPE_A014432_20180328T081650	▪ L1C_T35JPF_A014432_20180328T081650
▪ L1C_T35HPE_A005881_20180422T081053	▪ L1C_T35JPF_A005881_20180422T081053
Sentinel - 1	
▪ S1A_IW_GRDH_1SDV_20180113T164515_20180113T164540_020138_022584_6743	
▪ S1A_IW_GRDH_1SDV_20180218T164515_20180218T164540_020663_02363B_4B89	
▪ S1A_IW_GRDH_1SDV_20180326T164515_20180326T164540_021188_0246E4_A133	
▪ S1A_IW_GRDH_1SDV_20180419T164516_20180419T164541_021538_0251D4_89FD	

Sentinel-2 images consisted of four cloud-free pairs of tiles (Table 3.1), each of which provided a footprint coverage of the selected crop producing areas. The lower-resolution bands of Sentinel-2, which are bands 1, 9, and 10, were purposefully excluded and only the 20 m and 10 m resolution bands were used. The images were acquired on 04 January, 28 February, 28 March, and 22 April 2018. These dates were selected because these were the only dates with cloud-free images and because they coincides with the growing season. As stated earlier, maize, which is the most cultivated crop, is planted between November and January and it reaches maturity between March and April. These Sentinel-2 images were pre-processed using the Sen2Cor 2.5.5 plugin to correct Top-Of-Atmosphere (TOA) Level-1C reflectance to Bottom-Of-Atmosphere (BOA) surface reflectance. Thereafter, individual pairs of Sentinel-2 images were mosaicked in QGIS. The Sentinel-2 product specification document provides more details about Sentinel-2 data (Gatti and Bertolini, 2015).

Each of the Sentinel-1 images were big enough to cover the study area. Therefore, there were only four tiles used, each with the VV and VH polarimetric channels. Sentinel-1 acquires this data in a pre-programmed conflict-free Interferometric Wide swath mode, which provides a swath width of 250 km at a ground resolution of 5 m by 20 m (Prats-Iraola et al., 2015). Pre-processing techniques applied on Sentinel-1 included: (1) orbit file application, (2) radiometric calibration, (3) speckle filtering, and (4) geometric correction. The orbit file updates the image metadata by providing an accurate position of the Sentinel-1 image. Radiometric calibration reduces radiometric bias by providing pixel values that can be directly related to the radar backscatter of the scene. Radiometric calibration is necessary for the comparison of images acquired by different sensors or with the same sensor at different times. Speckle filtering reduces the inherent image speckles, which degrade the quality of the image and make interpretation of features more difficult. Geometric correction reduces geometric distortions in the image. These distortions are caused by the topographic variations of the scene and the tilt of the satellite sensor. After these pre-processing techniques were applied, both the Sentinel-1 and Sentinel-2 bands were co-registered to WGS84 UTM Zone 35S and resampled to 10 m spatial resolution.

3.3.3. Field-data collection

Training data were collected between May and July 2018. A field guide map was prepared from crop cultivation records of the 2017 to 2018 season acquired from extension officers of the Department of Agriculture, Land Reform and Rural Development (DALRRD). These records included geocoded lists of crop fields with contact details of all smallholder farmers who had planted maize, cannabis, and vegetables in the 2017 to 2018 season. Since there were very few legal cannabis farms at the time, the illegal cannabis farms were accessed with the assistance of the farmers. Although most of the maize and vegetable farms were visited, additional information about unvisited farms was solicited through telephone interviews. The tea farms are well known and easily identifiable in Sentinel-2 images and Google Earth. GPS coordinates of each crop field were recorded and then edited in Google Earth and ArcGIS to generate training polygons. GPS coordinates of other land use and land cover types were also collected in order to enhance confident identification of crop fields. Training samples for each land use and land cover types were uniformly distributed across the study area. Table 3.2 summarizes the information classes for which detailed information was compiled.

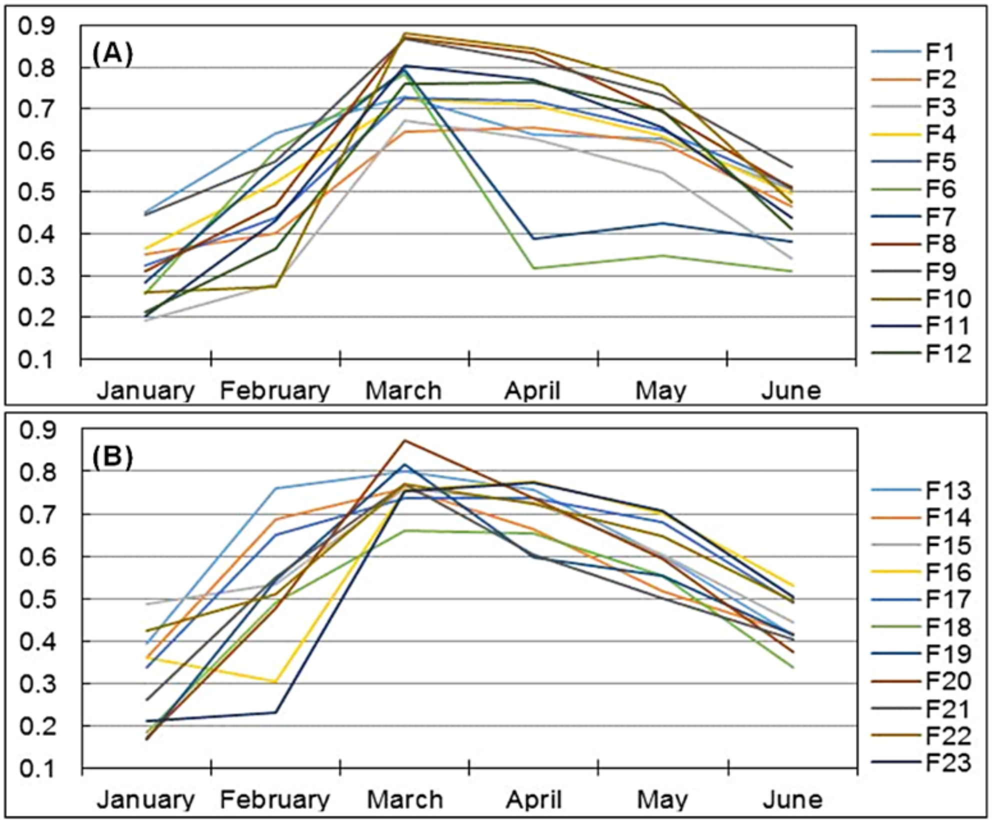
Table 3.2: Information classes identified during field investigation

Name	Description	Number of training polygons

Maize	Maize-cultivated crop fields	188
Vegetables	Intercropping of cabbage, beetroot, and spinach	55
Cannabis	Cannabis crop fields	62
Tea	Magwa and Majola Tea Estates	98
Planted forest	Forest plantations	106
Grassland	Open grassland with < 10% trees and canopy cover	146
Urban	Villages, buildings, and other man-made structures	98
Bare surfaces	Non-vegetated areas with bare soil and rock outcrops	108
Clear water	Surface water bodies with clear water	82
Turbid water	Surface water bodies with suspended sediment	96
Dense natural trees	Densely distributed trees ≥ 2 m tall	118
Sparse natural trees	Sparsely distributed trees ≥ 2 m tall	93
Shrub	Areas with herbaceous plants ≤ 1 m tall	75

3.3.4. Image analysis

The image analysis procedures that were applied include: (1) compilation of relevant predictor variables from Sentinel-2 data, (2) combination of Sentinel-1 and Sentinel-2 through a systematically determined number of classification trials, and (3) classification through model stacking. To identify the optimum period to produce crop maps during the crop development cycle, the study initially used six Sentinel-2 images covering the period from January to June 2018 to compute multi-temporal NDVI profiles of maize. Maize was selected because it is the most cultivated crop over the growing season. Figure 3.2 shows temporal variations in maize NDVI profiles over this 6-month period. NDVI values of maize ranged from approximately 0.2 to 0.5 in January. Thereafter, NDVI values in most of the fields increased and reached maximum in March, with the exception of a few that were planted in January whose profiles marginally increased to reach maximum in April. All the profiles then rapidly declined up to June as the crop transitioned into senescence. These results were the basis upon which January to April images were selected.



Field	Planting Date	Field	Planting Date	Field	Planting Date	Field	Planting Date
1	2017/10/15	7	2017/11/29	13	2017/12/24	19	2018/01/15
2	2017/11/14	8	2017/12/03	14	2017/12/26	20	2018/01/15
3	2017/11/22	8	2017/12/05	15	2018/01/01	21	2018/01/15
4	2017/11/25	10	2017/12/14	16	2018/01/02	22	2018/01/18
5	2017/11/26	11	2017/12/15	17	2018/01/05	23	2018/01/26
6	2017/11/29	12	2017/12/20	18	2018/01/08		

Figure 3.2: Temporal variations in NDVI profiles for maize

- *Variable importance analysis*

Variable importance analysis (VI) was used to rank the multi-temporal bands of Sentinel 2. The Mean Decrease Accuracy (MDA) function embedded in the RF algorithms was used to rank the explanatory power of each band. The MDA is computed as shown in Equation 1 by calculating and averaging predictor errors in out-of-bag (OOB) data before and after permutation (Souri and Vajedian, 2015; Hur et al., 2017).

$$\mathrm{VI}_{(x_j)}^{t} = \frac{\sum_{i \in (\mathrm{OOB})^t} \mathrm{I}(y_i = f(x_i))}{|(\mathrm{OOB})^t|} - \frac{\sum_{i \in (\mathrm{OOB})^t} \mathrm{I}\left(y_i = f\left(x_i^j\right)\right)}{|(\mathrm{OOB})^t|} \quad (1)$$

In equation 1, x_j is variable j in tree t, $y_i = f(x_i)$ is the predicted class for observation i before permuting x_j, and; $y_i = f\left(x_i^j\right)$ is the predicted class for observation i after permuting variable x_j. This technique improves prediction by aiding the identification and elimination of redundant variables. The study used the rfPermute function in RStudio to compute p values for the important variables. The top 10 variables out of the 40 had significant p-values ($p < 0.05$). The top 10 Sentinel-2 bands ranked as the most important were then used in the main analyses. Four spectral indices were selected because of their different sensitivities to different surface features. The Soil Adjusted Vegetation Index (SAVI) was computed using the Sentinel-2 mosaic of January. SAVI minimizes the effects of soil reflectance on vegetation and is sensitive to leaf area index in the early stages of crop growth (Hatfield and Prueger, 2010). SAVI is calculated according to equation 2.

$$\text{SAVI} = \frac{(\text{NIR})-(\text{RED})}{(\text{NIR})+(\text{RED})+L\times(1+L)} \tag{2}$$

RED and NIR are the red and near-infrared wavelengths respectively, and L is the canopy background (soil) adjustment factor. Normalised Difference Vegetation Index (NDVI) was computed using the Sentinel-2 mosaic of February. NDVI performs best when vegetation density is low (Phadikar and Goswami, 2016; Xue and Su, 2017) and is calculated according to equation 3.

$$\text{NDVI} = \frac{(\text{NIR})-(\text{RED})}{(\text{NIR})+(\text{RED})} \tag{3}$$

The Two Band Enhanced Vegetation Index (EVI2) was computed for March and April. EVI2 is more effective than NDVI when vegetation density is high and is less affected by canopy background signal and atmospheric influences (Jiang et al., 2008). It is calculated according to equation 4.

$$\text{EVI2} = 2.4\,\frac{(\text{NIR})-(\text{RED})}{(\text{NIR})+(\text{RED})+1} \tag{4}$$

The Normalized Difference Built-up Index (NDBI) was computed using the Sentinel-2 mosaic of March. NDBI includes the Short Wave Infrared (SWIR) band because urban surfaces have high reflectance in SWIR than other wavebands (Zha et al., 2003), and is calculated according to equation 5.

$$NDBI = \frac{(SWIR)-(NIR)}{(SWIR)+(NIR)} \quad (5)$$

NDBI was selected because crop farming is practices in immediate vicinities of urban features. From Sentinel-1, both the VH and VV polarimetric channels were used. These variables were selected in order to facilitate the identification of combinations potentially capable of improving mapping accuracy. Different combinations of these datasets were used as input variables in different classification trials (Table 3.3).

Table 3.3: Data combinations as used in six different classification trials

Trial number	Input variable	Period	Total
1	S2 bands	January - April	10
2	S2 indices	January - April	*5
3	S2 bands, S2 indices	January - April	15
4	S2 bands, S2 indices, VH	January - April	19
5	S2 bands, S2 indices, VV	January - April	19
6	S2 bands, S2 indices, VH, VV	January - April	23
	S1 = Sentinel-1, S2 = Sentinel-2, * EVI2 was computed for March and April		

3.3.5. Classification

Xgboost is an implementation of gradient boosting that uses an additional regularization term in the objective function (Chen and Guestrin, 2016). As in other ensemble tree-boosting algorithms, models are added in a sequential manner with the next predictor correcting errors made by previous predictors (weaker learners) until the training data is accurately predicted and no further improvements are made. This iterative process uses a gradient descent algorithm to optimise the loss function when adding new models by minimizing loss. The regularization term helps to avoid overfitting by controlling the complexity of the model. The study used the caret package in R-Studio to fine-tune seven Xgboost hyper-parameters for optimal performance. The advantage of this method is that it often outperforms other machine learning algorithms. It also has several hyperparameter tuning options. However, just like any other machine learning algorithms, Xgboost has disadvantages, which include, being sensitive to outliers and overfitting if the hyperparameters are not tuned properly (Shi et al., 2021).

Model stacking is an ensemble technique that takes predictions generated by multiple machine learning algorithms and uses them as inputs in a second level learning classifier (Wolpert, 1992). Unlike other ensemble methods like bagging and boosting, which only combine algorithms of the same type, model stacking can combine different types of algorithms through a meta-model to maximize the prediction accuracy. Using this technique, this study combined Xgboost with RF (Breiman, 2001), SVM (Cortez and Vapnik, 1995), ANN (Ripley, 1996), and NB (Solares and Sanz, 2007). Correlations between the predictions generated by these classifiers were calculated in R-Studio. Model stacking was performed in two in different ways. The first approach stacked all the five algorithms and the second approach stacked only weakly correlated algorithms (Džeroski and Ženko, 2004; Merz, 1999). These two approaches were then compared based on classification accuracy.

3.3.6. Training and performance evaluation

The training data was randomly split into three independent segments. Firstly, the data was split into 70% and 30% for model building and model evaluation, respectively. Following Géron, (2019), the model-building segment was further split into two subsets; the first subset to train the individual classifiers, and the second subset to generate new predictions to train the level two model. The models were evaluated using different performance metrics, namely overall accuracy, and model sensitivity and specificity (Rwanga and Ndambuki, 2017). The study ran significance tests to evaluate whether data fusion and model stacking produced real improvements. These tests were computed in R-Studio using model comparison functions that were developed from the works of Hothorn et al., (2005) and Eugster et al., (2008).

3.4. Results

The results of this investigation are presented under three sub-sections: sub-section 3.3.1 presents importance rankings of the Sentinel-2 bands, sub-section 3.3.2 summarises the results of data fusion, and sub-section 3.3.3 summarises the results of model stacking.

3.4.1. Variable importance

To reduce data redundancy and noise, the Sentinel-2 images were subjected to VI analysis after which the top 10 multi-date image bands were selected. Figure 3.3 shows the top 10 out of the 40 Sentinel-2 bands.

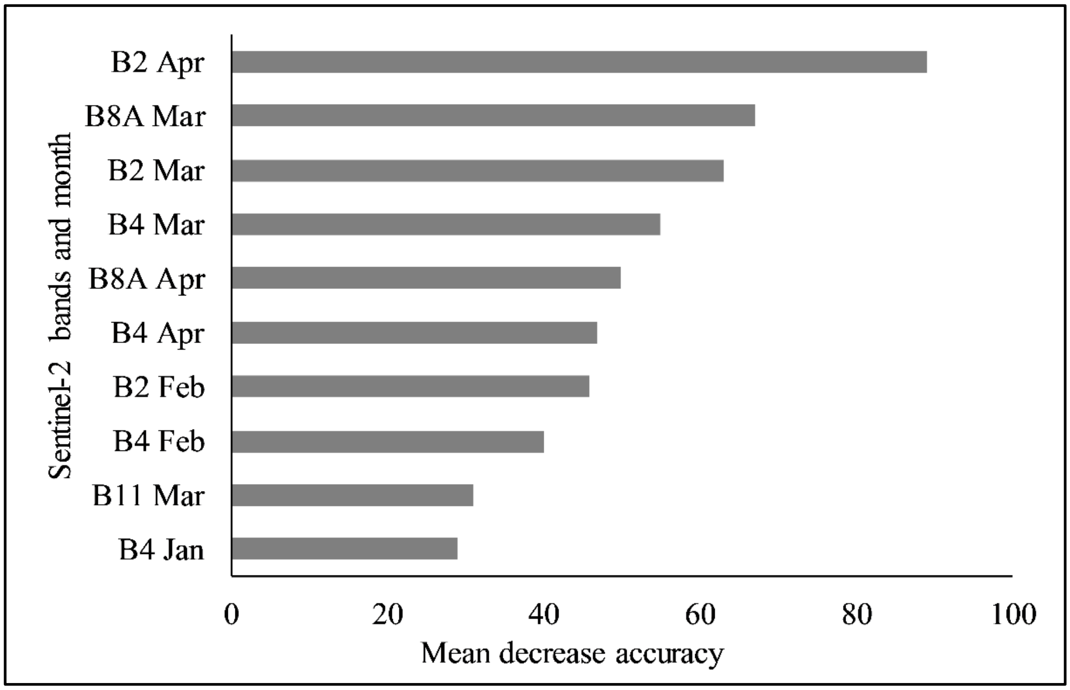


Figure 3.3: Top 10 of the 40 Sentinel-2 bands

The top 10 bands include four images from March, three from April, two from February, and one from January, with bands B2, B8A, B4 and B11 being the most important.

3.4.2. Results of image fusion

Table 3.4 shows overall accuracies (OA), per-class Sensitivity (SE) and Specificity (SP) that were obtained by using Sentinel-2 data in the first three trials.

Table 3.4: Results obtained from Sentinel-2 data

Information class	Trial 1		Trial 2		Trial 3	
	SE	SP	SE	SP	SE	SP
Maize	100.00	99.98	100.00	99.93	100.00	99.98
Vegetables	4.55	100.00	18.18	100.00	9.09	100.00
Cannabis	44.44	100.00	16.67	99.99	22.22	100.00
Tea	99.76	99.84	96.48	97.02	99.76	99.80
Planted forest	90.53	99.90	87.96	99.80	90.20	99.90
Grassland	88.90	98.48	91.11	96.40	87.91	98.63
Urban	96.36	90.15	91.14	91.26	96.86	88.82
Bare soil	43.86	99.96	11.07	99.96	42.11	99.98
Clear water	95.00	100.00	95.00	99.96	95.00	100.00
Dense natural trees	85.92	99.92	78.40	99.71	88.03	99.90
Sparse natural trees	90.78	99.59	31.36	99.10	86.96	99.64
Turbid water	100.00	100.00	100.00	100.00	100.00	100.00
Shrubs	12.10	99.60	0.00	99.97	8.87	99.94
OA (%)	89.85		89.82		89.09	

The 10 multi-temporal wavebands of Sentinel-2 in trial 1 produced the highest OA (89.85 %). However, the spectral indices in trial 2 were more sensitive to vegetables (18.18 %) than the wavebands in trial 1 (4.55 %) and the combination thereof in trial 3 (9.09 %). The wavebands in trial 1 were more sensitive to cannabis (44.44 %) than the spectral indices in trial 2 (16.67 %) and the combination thereof in trial 3 (22.22 %). The spectral indices were less sensitivity to tea (96.48 %) than the wavebands (99.76 %) and the combination thereof (99.76 %). Sensitivity to maize remained at 100 % in all the Sentinel-2 classification trials. Figure 3.4 shows the classification maps obtained from the exclusive use of Sentinel-2 data.

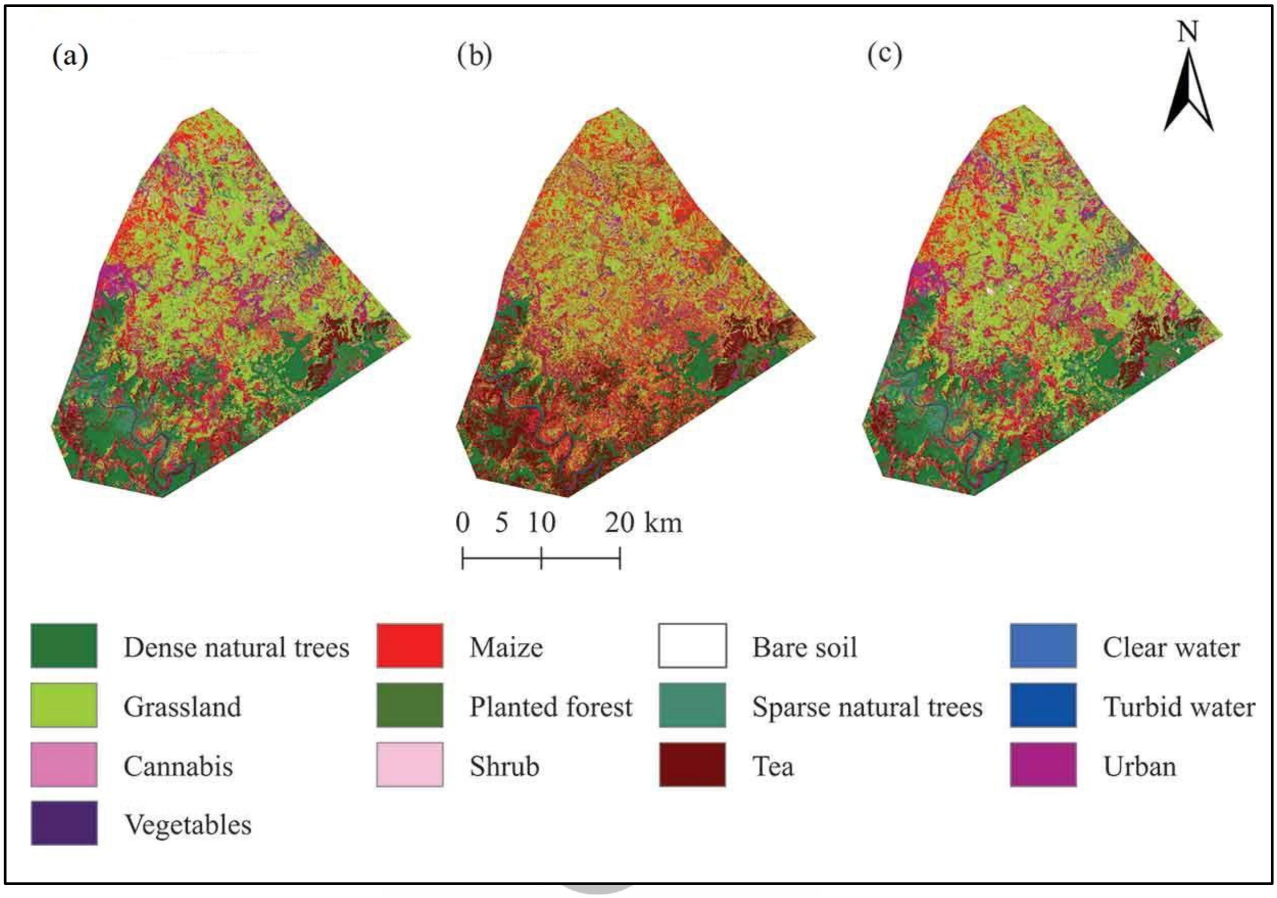


Figure 3.4: Results of trial 1 (a), trial 2 (b) and trial 3 (c)

Table 3.5 shows the results obtained from the combination of optical and SAR datasets in trials 4 to 6.

Table 3.5: Results of combining optical and SAR data

Class	Trial 4		Trial 5		Trial 6	
	SE	SP	SE	SP	SE	SP
Maize	100.00	99.99	100.00	99.98	100.00	99.99
Vegetables	0.00	100.00	13.64	100.00	0.00	100.00
Cannabis	33.33	100.00	22.22	99.97	11.11	99.84
Tea	99.88	99.80	99.84	99.90	99.98	99.88
Planted forest	90.03	99.90	90.03	99.88	90.12	100.00
Grassland	98.15	99.24	93.45	99.17	98.28	99.33
Urban	97.87	98.27	97.99	93.62	97.98	98.38
Bare soil	85.20	99.91	44.52	99.97	85.31	99.90
Clear water	95.00	100.00	95.00	100.00	95.00	100.00
Dense natural trees	89.44	99.88	86.15	99.90	88.26	99.89

Sparse natural trees	85.24	99.67	84.98	99.65	87.74	99.66
Turbid water	100.00	100.00	100.00	100.00	100.00	100.00
Shrubs	7.26	99.94	8.87	99.94	7.26	99.96
OA (%)	97.58		93.46		97.71	

Explanation: Highest OA with S2 = 89.85 %, Highest OA with S1-S2 fusion = 97.71 % Difference = 7.86 % , p = 0.0005

Combining optical data with both the VV and VH bands of SAR in trial 6 produced the highest OA (97.71 %) than all the other different combinations of SAR and optical data. Sensitivity to maize remained at 100 % in all the trials. The exclusive use of optical data in trial 2 showed more sensitivity to vegetables (18.18 %) and cannabis (44.44 %) compared to the combinations of optical and SAR data in Table 3.5. Of all the optical and SAR combinations, the combination of optical data and the VV channel in trial 5 was the most sensitive to vegetables (13.64 %), while the combination of optical data and the VH channel in trial 4 was the most sensitive to cannabis (33.33 %). The combination of optical data with both VV and VH in trial 6 was more sensitive to tea (99.98 %) than the combination of Optical data and VH in trial 4 (99.88 %), the combination of optical data and VV in trial 5 (99.84 %), and the exclusive use of optical data in trial 3 (99.76 %). Statistical significance tests were performed to ascertain if combining optical and SAR data really improved classification accuracy or it was just a chance event. Table 3.5 shows that combining optical and SAR data produced a statistically significant improvement compared to the exclusive use of optical data (p = 0.0005). Figure 3.5 shows the maps obtained from the combinations of optical and SAR data.

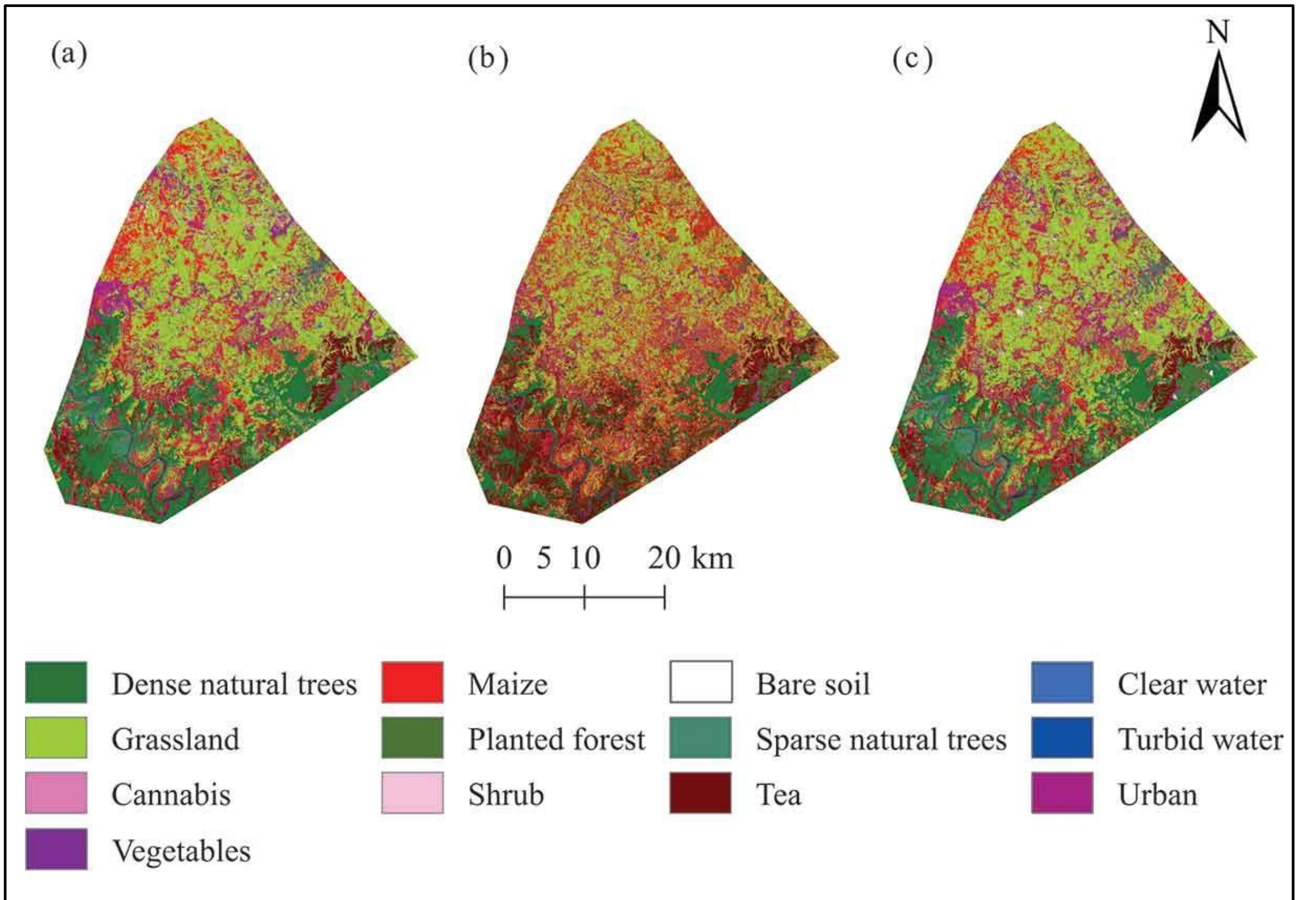


Figure 3.5: Results of trial 4 (a), trial 5 (b), and trial 6 (c)

3.4.3. Results of model stacking

The study also tested the mapping effectiveness of model stacking. The stacked algorithms were first trained individually and their predictions were compared using correlation. Model stacking was tested on the combined Sentinel-2 variables only (trial 3). Table 3.6 shows how the models correlated.

Table 3.6: Model correlation matrix

	Xgboost	RF	SVM	ANN	NB
Xgboost	1.00	0.55	0.49	0.71	0.42
RF	0.55	1.00	0.94	0.73	0.69
SVM	0.49	0.94	1.00	0.58	0.79

ANN	0.71	0.73	0.58	1.00	0.67
NB	0.42	0.69	0.79	0.67	1.00

Correlations between Xgboost and the other algorithms were not as strong (0.55, 0.49, 0.71, and 0.42). SVM correlated strongly with RF (0.94) and NB (0.79). Xgboost had stronger correlations with ANN (0.71) than with RF (0.55), SVM (0.49), and NB (0.42). Based on this correlation matrix, it was concluded, in the final analysis, that the least correlated models were Xgboost, RF, and NB. The combination of these three models was then compared with the combination of all the five models (Table 3.7).

Table 3.7: Model-stacking results

Method	OA (%)
[a] Xgboost	89.09
ANN	86.88
RF	85.03
SVM	84.92
NB	84.37
Model stacking	
[b] Xgboost, RF, SVM, ANN, NB	93.97
[c] Xgboost, SVM, ANN	96.06
c - b = 2.09 %, p = 0.0242	
c - a = 6.97 %, p = 0.0100	

Of the standalone models (Table 3.7), Xgboost achieved the highest overall accuracy (89.09 %). Stacking the algorithms without ANN and SVM produced higher classification accuracy (96.06 %) than stacking all the algorithms (93.97 %). The improvement attained by stacking the models based on correlation tests was statistically significant (p = 0.0242). Model stacking also produced better results than using a single algorithm (p = 0.0100). Table 3.8 summarises of the results obtained through model stacking.

Table 3.8: Results of model stacking

Information class	**Xgboost**	**Xgboost, RF, SVM, ANN, NB**	**Xgboost, RF, NB**

	SE	SP	SE	SP	SE	SP
Maize	100.00	99.98	100.00	99.98	100.00	99.99
Vegetables	9.09	100.00	0.00	100.00	9.09	100.00
Cannabis	22.22	100.00	22.22	100.00	27.71	100.00
Tea	99.76	99.80	99.76	99.90	99.76	99.80
Planted forest	90.00	100.00	91.11	99.93	92.28	99.91
Grassland	87.91	98.63	99.17	89.71	99.35	90.01
Urban	96.86	88.82	79.47	98.34	79.53	98.70
Soil	42.11	99.98	24.67	99.99	43.20	99.98
Clear water	95.00	100.00	95.00	100.00	95.00	100.00
Dense natural trees	88.03	99.90	85.79	99.95	86.62	99.97
Sparse natural trees	86.96	99.64	86.61	99.64	93.15	66.69
Turbid water	100.00	100.00	100.00	100.00	100.00	100.00
Shrubs	8.87	99.94	7.25	99.99	8.87	99.99
OA (%)	89.09		93.97		96.09	

Sensitivity of the stacked models to maize and tea did not change compared to the exclusive use of Xgboost. However, sensitivity to vegetables decreased to 0.00% with the 5-model ensemble, while remaining at 9.09% with the 3-model ensemble. The 5-model ensemble's sensitivity to cannabis remained at 22.22%, while increasing to 27.71% with the 3-model ensemble. The differences in OA between the single-model classification and the ensembles are mainly due to increases in sensitivity and specificity in the non-crop classes. Figure 3.6 shows the maps obtained from model stacking.

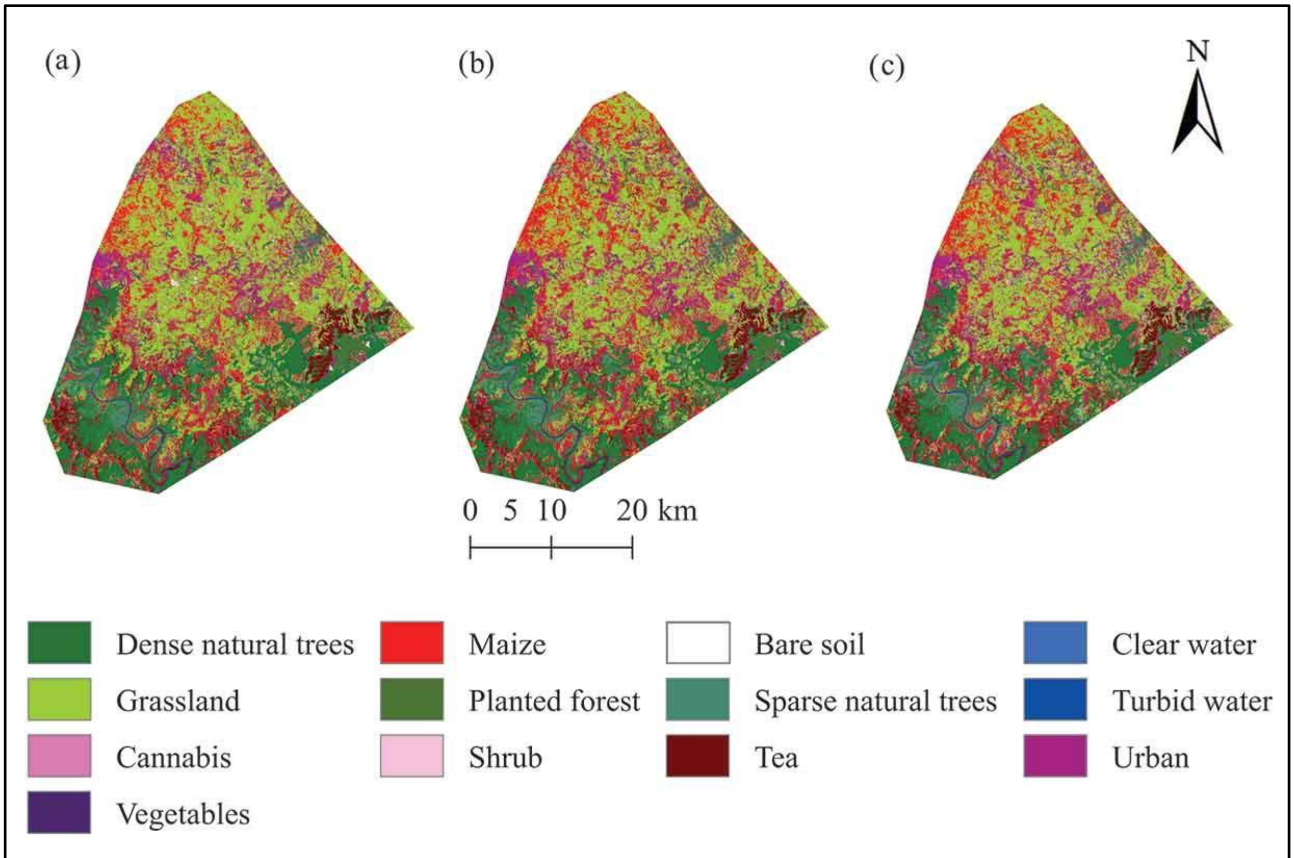


Figure 3.6: Xgboost (a), All algorithms combined (b), Xgboost, RF, NB (c)

3.5. Discussions and conclusions

Optical and SAR data were combined to map a smallholder crop farming landscape. Of the optical data, Sentinel-2 bands 2, 4, 8A, and 11 were ranked as the most important for discriminating surface features and were found to be equally effective in other studies (Santos et al., 2019; Zhang et al., 2019). This finding demonstrates the importance of combining NIR, red and red edge bands, which are sensitive to vegetation dynamics, with the SWIR, which is sensitive to urban features. It also shows the efficacy of the SWIR when crop fields are close to built-up areas. Other studies have also demonstrated the SWIR band's ability to enhance vegetation mapping in urban and heterogeneous landscapes (Hartling et al., 2019; Sidike et al., 2019).

This study elected to systematically add SAR data to optical data because most studies have found the latter to be more accurate than the former (Fontanelli et al., 2014; Clerici et al., 2017; Denize et al., 2019; Mercier et al., 2019; Tavares et al., 2019). All the combinations of SAR and optical data used in this study mapped maize and tea with very high accuracy. Although most of the maize fields were smaller than two hectares, maize was the majority class in the training data and the tea fields were big in size and easily identifiable. The high misclassification of vegetables and cannabis can be attributed to various interacting factors including the small sizes of both cannabis and vegetable fields, most of which ranged between about 0.01 and 1.00 hectare (Figure 3.7).

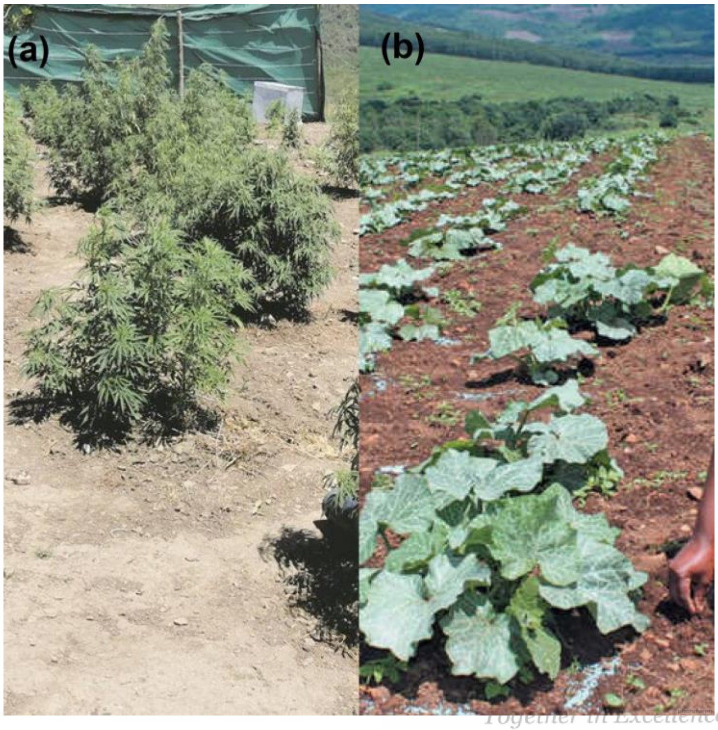


Figure 3.7: Photos showing cannabis (a) and vegetable fields (b)

Secondly, there was poor plant spacing and irregular planting patterns particularly in the cannabis fields, and big patches of bare soil in both cannabis and vegetable fields (Figure 3.7). Although the combination of the optical data with the VH band improved overall accuracy and sensitivity to tea, this combination reduced sensitivity to vegetables and cannabis. Background interference from bare soils, which has its own dynamic spectral properties, usually distorts plant canopy reflectance (Prudnikova et al., 2019). Thirdly, due to the rarity of vegetable farming in the area, the number of training points for vegetables was small compared to the other classes. Imbalanced training data often leads to imbalanced learning, which affects classification accuracy (Douzas et al., 2019). Future research can improve this technique by using data balancing techniques which have been shown to be effective (Douzas et al., 2019; Xia et al., 2019). Fourthly, the cannabis fields are located in obscure and remote locations because most of the cannabis farming in the area is unlicensed. As a result, cannabis was also a minority class in the training data.

Concerning SAR, the VH polarized band proved more important than the VV band. Although some studies report marginal differences between these bands (Abdikan et al., 2016; Inglada, Vincent, Arias, and Marais-Sicre, 2016; Van Tricht et al., 2018; Whelen and Siqueira, 2018), the VH band has higher sensitivity to vegetation and other land cover types than the VV band (Ban, 2016; Dimov et al., 2017; Rajah et al., 2018). This can be explained by the fact that VV backscatter is negatively affected by temporal changes in the scattering mechanism (Xu et al., 2019), whereas VH backscatter is a better characterizer of vegetation because of its sensitivity to vegetation volume scattering (Brisco et al., 1992; Vreugdenhil et al., 2018). Regardless, the joint-use of these two bands outperformed the single-use of either bands.

Model stacking was computed to improve classification accuracy through the combination of different classifiers. Of the base-classifiers, Xgboost achieved higher classification accuracy, outperforming the second best performing classifier (ANN) by 2.21%. Model correlations between Xgboost and the other classifiers were low to moderate, warranting stacking of Xgboost with any of the other four classifiers. However, SVM's correlations with RF (0.94) and NB (0.79) were strong. Xgboost also had stronger correlations with ANN than with RF, SVM, and NB. Model stacking does not add much value when the base models have strongly correlated predictions. It is for this reason that the models were stacked analytically. Although stacking of all the models increased accuracy from 89.09% to 93.97%, exclusion of ANN and SVM increased classification accuracy to 96.06% (p = 0.04).

This study demonstrated that successful mapping of a fragmented agricultural landscape is a function of objectively derived datasets, which are adapted to the geographic context of the mapped area, and an informed optimization of mapping algorithms. Although both image fusion and model stacking improved classification accuracy, adding SAR data to the optical data produced higher classification accuracy than merely applying model stacking to optical data. Future investigations could apply model stacking to fused SAR and optical data for even better results. Since this method produced a very high mapping accuracy of maize, the author recommended that it be tested in the calculation of maize field sizes in other smallholder farming areas. A study to this effect was undertaken by Mashaba-Munghemezulu et al., (2021) who achieved good results ($p < 0.05$, R = 0.84). This approach is objective and not limited by study area in the sense that the datasets were adapted to the physiographic conditions of the study area through variable ranking, while model stacking was informed by model comparison tests. Future research could test different image fusion techniques, SAR techniques, and different sets of machine learning/deep learning algorithms. Follow-up studies could also test the techniques provided in this study by mapping the distributions of smallholder crop fields in other places. Successful mapping of smallholder farms could improve farmer support from insurers, creditors and governments. Insurers can use these methods to locate and quantify sizes of insured farms rather than relying on farmers' reports and costly on-field surveys.

Chapter 4 : Factors that influence maize-yields in rain-fed smallholder farms

Abstract

Weather extremes pose substantial threats to food security in areas where the main source of livelihood is rain-fed crop production. In most of these areas, index-based crop insurance (IBCI) is recognized as being capable of securitizing food production because it provides safety nets against weather-induced crop losses. Unfortunately, however, IBCI does not indemnify farmers for non-weather-related crop losses. This study investigates how this gap can be filled by exploring how IBCI can be linked with non-weather factors that influence crop production. The study uses an improvised variable ranking methodology to identify these factors in the O.R. Tambo District Municipality, South Africa. Results show that key agrometeorological variables comprising surface moisture content, growing degree-days, and precipitation influence maize yield even under optimal weather conditions due to planning and sowing dates, while seed variety, fertilizer application rate, soil pH, and ownership of machinery also play a significant role on crop yields. This finding is important because it demonstrates that although most IBCI systems focus more on weather-induced losses, there are non-weather factors that may expose farmers to production risks under optimal weather conditions. As such, linking IBCI with critical non-weather yield-determining factors can improve risk management. Furthermore, reducing the influence on non-weather variables on crop losses could enable remote sensing to accurately isolate and quantify the influence of weather on crop losses.

4.1. Background

Smallholder farming around the world contributes substantially to economic growth and food security, especially in rural areas (UNCTAD, 2015; Fan and Rue, 2020). However, the increasing occurrence of weather shocks threatens agriculture, especially in Sub-Saharan Africa (SSA) where 95% of farmland is rain-fed (Wani et al., 2009; Masih et al., 2014; Buhaug et al., 2015). In the past 10 to 15 years, attempts were made to support farmers through index-based crop insurance (IBCI), which acts as a safety net against the adverse effects of weather-induced crop failures (Michael Carter et al., 2017; Di Marcantonio and Kayitakire, 2017; Ntukamazina et al., 2017). Recent studies show that insurance encourages farmers to take risks and make more investments in productive inputs. In Bangladesh, for example, purchasing insurance led to the expansion of agricultural land and more investment in fertilizer, labour, irrigation, and pesticides (Hill et al., 2019). In Kenya, an uptake of insurance was significantly associated with increased use of fertilizer and expenditures on seeds by 50% and 65%, respectively, and a corresponding increase in maize yields by 60% (Sibiko and Qaim, 2020).

This shows that farmers will allocate their resources in a manner that maximizes returns if they are assured of financial compensation for losses arising from factors beyond their control (Shashi et al., 2012). Other studies show that insurance has the potential to unlock credit for farmers because IBCI reduces credit risk for lenders (von Negenborn et al., 2018; Möllmann et al., 2020). These scenarios suggest that linking insurance with credit and agronomic inputs could improve insurance uptake and enhance the capacities of farmers to manage risks. In Nigeria, for example, farmers were more willing to take up insurance because of the anticipated benefits of bundling IBCI with agricultural inputs (Adeoti et al., 2020). Similarly, Awondo et al., (2020) found that bundling IBCI with drought-tolerant seed varieties, grown in suitable environments, could result in lower premiums and higher guaranteed returns. Linking insurance with credit, inputs, market opportunities, management advisories, and training was identified to be a major contributor to the success of pilot IBCI projects (Hellin et al., 2015).

A holistic risk management approach that bundles or links IBCI with inputs and credit could help farmers focus on the critical factors that undermine crop production even under moderate weather shocks and normal conditions. For example, IBCI does not payout indemnity if crop failures are associated with factors such as fertilizer use, cultivars, pests, and diseases because payouts are based on weather and rainfall-dependent vegetation indices rather than actual yield losses. Since a wide range of factors influence crop production in smallholder farming systems (Banerjee et al., 2014; Tamene et al., 2016; Djurfeldt et al., 2018; Abdulai et al., 2020; Dutta et al., 2020), knowing which factors pose significant risks could enable farmers to improve management practices by focusing on the most important yield-determining factors. Moreover, linking and bundling IBCI with non-weather, crop yield-determining factors could attract more farmers to take up insurance contracts (Syll et al., 2017; Adeoti et al., 2020).

Although most of what is in the literature suggests that linking IBCI with different production factors leads to higher productivity, this reasoning is not adequately supported by any universally agreed list of these factors because they vary over time and from place to place. This justifies why it is necessary to optimise methods that can be used to reliably identify these factors for specific localities. This study aims to investigate maize yield-determining factors that can be linked with IBCI to reduce production risk and improve productivity in the O.R Tambo District Municipality (ORTDM), South Africa. This was done using an improvised variable ranking technique that could also be applied to other crops under different environmental settings.

4.2. Materials and methods

4.2.1. Study area

The O.R. Tambo District Municipality is in the northeastern part of South Africa's Eastern Cape Province (Figure 4.1). This district municipality is sub-divided into five local municipalities covering an area of 12,096 km^2. It is the second poorest of the seven district municipalities in the Eastern Cape Province (World Bank, 2018). About 94% of the population in ORTDM are rural dwellers whose main sources of livelihood include livestock farming, rain-fed maize production and other crops, and government social grants (Eastern Cape Socio Economic Consultative Council (ECSECC)., 2017). This investigation focused on three of ORTDM's local municipalities whose identification was guided by maize yield records obtained from the Department of Agriculture, Land Reform and Rural Development (DALRRD). These records consist of GPS locations of maize fields and contact details of farmers. The maize fields included small plots that ranged in size from one hectare to slightly more and outfield collective farms (≤80 ha) in which individual farmers cultivate specific plots. Maize was selected partly because it is the primary staple food crop which is widely produced under rain-fed conditions in SSA (Ekpa et al., 2018;

Ayanlade and Radeny, 2020; Santpoort, 2020), including South Africa, where white maize is destined for human consumption and yellow maize for animal feed (DALRRD, 2018).

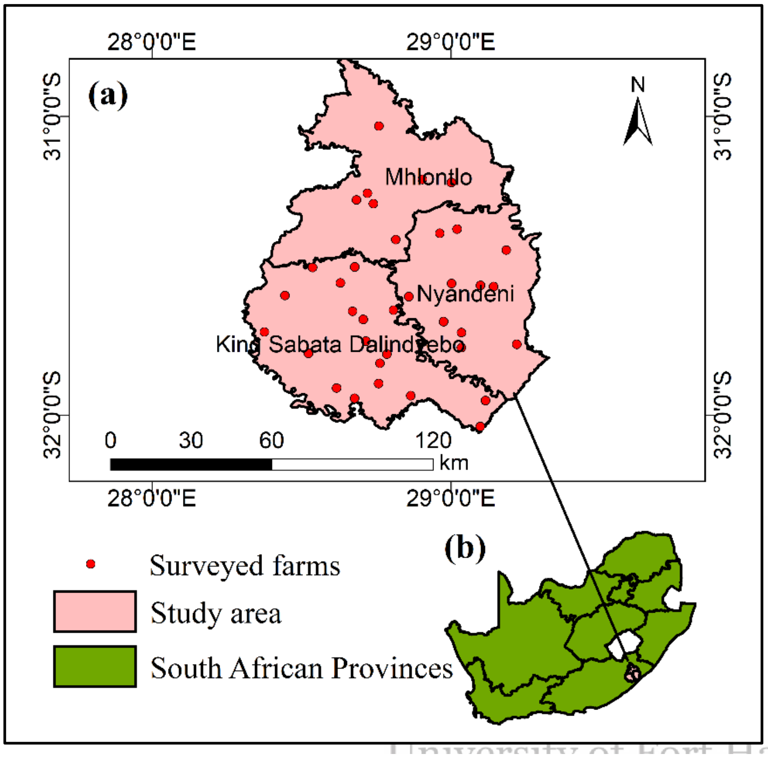


Figure 4.1: Location of the study area (a) in South Africa (b).

The selected farms were evenly distributed in a landscape that is characterized by (a) low-lying densely vegetated areas along the Wild Coast where elevation ranges from 5 m to 500 m, (b) gentle-to-moderate-sloping grasslands in the interior, and (c) savannas and forests in the northern areas where elevation extends to 1500 m. The area has a warm oceanic climate, which includes the humid sub-tropical climate of the northeastern peripheries and the semi-arid climates of the southwestern parts (Beck et al., 2018).

Mean annual rainfall ranges from 900 mm to 1300 mm, with summer minimum and maximum temperatures of 14–19 °C and 14–27 °C, respectively (Jordaan et al., 2017). The soils are largely dominated by sandy loams, sandy clay loams, and clays that are yellow to black in colour and slightly acidic (Eta and Grace, 2013; Sibanda et al., 2016). In the past, the farmers used to begin planting maize in the first dekad of October, but planting now begins in mid-November, often extending up to late-January. The harvest season is usually from June to August.

4.2.2. Data compilation and pre-processing

Datasets used include maize yield data; remote sensing indices; and socio-economic, agronomic, soil, and meteorological variables for 65 farms. These variables were selected because (1) the literature shows that smallholder farm yields are generally influenced by wide-ranging factors and (2) the main objective of this study was not only to identify yield-determining factors, but to also rank them in the order of their importance. This is important because most index insurance schemes calibrate insurance contracts by correlating crop yields or crop losses with weather variables; however, the extent to which non-weather variables influence crop yields and losses is unknown. Therefore, knowing how influential weather variables are in comparison to non-weather variables could help in addressing the problem of basis risk and help insurers who might want to bundle or link insurance with inputs and advisory services. Table 4.1 describes all the variables that were investigated in this study.

Table 4.1: Variables investigated in the study

Variable	**Description**
Gender	Male or female
Age	Farmer's age in years
Education	Level of education (Primary , Secondary , Tertiary)
Farm ownership	Individual or collective ownership
Farming experience	Number of years since the farmer started growing maize
Farm labour	Hired or sourced from family members
Planting date	Number of days from sowing to the day of yield survey
Target market	Village, commercial retailers, Subsistence
Machinery	Owned or hired
Field size	Crop field size in hectares
Seed variety	Pannar, Monsanto, Pioneer
Fertilizers application rate	Quantity of fertilizer (kg) applied per hectare
Soil properties	CEC, EC, K, Na, P, Ca, Mg, pH, OM, Total N, Total C, *Sand, *Silt, *Clay, *Bulk density
Spectral indices	NDVI, EVI2, MSI
Growing Degree Days (GDDs)	Accumulated heat units
Precipitation	Total precipitation from planting to survey date (mm)
Yield	Estimate of the amount of maize harvested (kg/ha)

Key to soil properties:
Cation exchange capacity (CEC), Electric conductivity (EC), Potassium (K), Sodium (Na), Phosphorus (P), Calcium (Ca), Magnesium (Mg), Acidity / alkalinity (pH), Organic matter (OM), Total Nitrogen (Total N), Total Carbon (Total C). *As described

- *Socioeconomic and agronomic data*

Baseline socio-economic and agronomic information was solicited through a pilot survey and a follow-up semi-structured interview with the farmers (Table 4.1). Questions related to intercropping, usage of manure, and other inputs were omitted because the farmers were practicing monoculture and using chemical fertilizers provided by DALRRD and Grain South Africa (GrainSA). Information about income was also omitted because the farmers were reluctant to disclose their off-farm sources of livelihood and their annual and monthly incomes.

- *Soil data*

The maize fields have gentle to flat slopes and homogenous vegetation, which allowed the surveyors to collect composite soil samples. The soils were collected from ground level to a depth of 30 cm using a soil auger. Thereafter, all the soil samples were taken to South Africa's Agricultural Research Council (ARC) laboratories for chemical and physical analyses of the parameters listed in Table 4.1.

- *Meteorological data*

Daily precipitation and maximum and minimum temperature records were obtained from the ARC's agro-climate databank, which continuously receives data from automatic weather stations distributed across the study area. These data were used to compute accumulated growing degree days (GDD) and total precipitation from the first day of planting up to the day of yield survey. GDD are used as an agrometeorological index to model the rate at which crops develop from one stage to another in their lifecycle (Ahmad et al., 2017). GDD are recognized as a more accurate estimate of plant physiological development than calendar days because a crop plant develops when the temperature is above a specific base temperature and below a certain upper threshold (Ahmad et al., 2017). For maize, the lower limit/base temperature (T_{base}) is 10 °C and the upper limit is 30 °C. Thus, GDD were calculated using the following equation 1:

$$\mathrm{GDD} = \frac{\mathrm{T_{min}} + \mathrm{T_{max}}}{2} - \mathrm{T_{base}} \quad (1)$$

where T_{min} and T_{max} are the daily minimum and maximum temperatures, respectively.

- *Remotely sensed spectral indices*

A Garmin Montana 650 GPS was used to geo-locate all the investigated maize fields during the farm surveys. Atmospherically corrected Sentinel-2 images acquired from November 2017 to June 2018, from November 2018 to June 2019, and from November 2019 to June 2020, were downloaded from the European Space Agency's Copernicus Hub and used to compute spectral indices. These indices were times-series maps of the Normalized Difference Vegetation Index (NDVI), the Two-band Enhanced Vegetation Index (EVI2), and the Moisture Stress Index (MSI). NDVI was selected because of its established ability to provide reliable results in modelling vegetation health and crop yield (Lewis et al., 1998; Mkhabela et al., 2005; Ngie and Ahmed, 2018). EVI2, which is also recognized as useful for estimating crop yield, is more effective than NDVI when vegetation density is high (Jiang et al., 2008; Gao et al., 2018; Liu et al., 2019; Bai et al., 2019). MSI was selected due to its sensitivity to leaf and soil water content, which are related with grain yield (Hunt and Rock, 1989; Tahara et al., 1990). These indices are calculated according to equations 2 to 4;

$$\mathrm{NDVI} = \frac{\mathrm{NIR} - \mathrm{RED}}{\mathrm{NIR} + \mathrm{RED}} \quad (2)$$

$$\mathrm{EVI2} = 2.4\,\frac{\mathrm{NIR} - \mathrm{RED}}{\mathrm{NIR} + \mathrm{RED} + 1} \quad (3)$$

$$\mathrm{MSI} = \frac{\mathrm{SWIR}}{\mathrm{NIR}} \quad (4)$$

where RED is the red band of Sentinel-2, and NIR and SWIR are the near and shortwave infrared bands, respectively. We initially regressed multi-temporal NDVI and EVI2 against grain yield to identify the period within which these two indices were best related to yield and proceeded to use peak values of these indices and the seasonal average of MSI after observing that they were closely related to yield. These processes were finalized by converting all the categorical independent variables listed in Table 4.1 to dummy variables.

- *Maize yield data*

Yield surveys were conducted at the beginning of the harvest season in 2018, 2019, and 2020. Prior to this period, farmers were interviewed to ascertain whether they or the DALRRD apply any methods to quantify yield. About 80% of the farmers reported that although they get inputs from DALRRD and GrainSA, no one conducts yield surveys in their farms. Therefore, the study adopted the objective yield survey method that is used by South Africa's Crop Estimates Committee to estimate yields. Detailed information about this method is provided in FAO's 2016 crop yield forecasting report (FAO, 2016).

4.2.3. Data analysis

To rank the independent variables, the study used percent increase in mean squared error (%IncMSE), which is a variable importance measure embedded in the random forest (RF) regression algorithm (Breiman, 2001). This method was selected because it is a model-based approach with the ability to order independent variables according to their relative importance. RF is an ensemble learning technique that works by constructing a number of decision trees and computing the mean prediction of the individual trees. This algorithm trains each decision tree on a different sample of the training set, where sampling is performed with replacement. The importance $VI_{perm}^{(k)}(x_j)$ is calculated in the kth tree for variable x_j according to equation 5 (Aldrich, 2020).

$$VI_{perm}^{(k)}(x_j) = \frac{1}{n_{OOB}} \sum_{i=1}^{n_{OOB}} (y_i^{(k)} - \hat{y}_i^{(k)})^2 - \frac{1}{n_{OOB}} \sum_{i=1}^{n_{OOB}} (y_i^{(k)} - \hat{y}_{j,i}^{(k)})^2 \tag{5}$$

where $y_i^{(k)}$ is the observation of dependent variable in tree k, $\hat{y}_i^{(k)}$ is the prediction by tree k, and $\hat{y}_{j,i}^{(k)}$ is the prediction by tree k when the jth variable is permuted. n_{OOB} is the number of samples in the out of bag (*OOB*) data seen by each of the trees in the forest. In order to identify the most important variables, different RF models were trained sequentially. In this process, the least important variables were removed until a model with the optimum number of variables and the lowest root mean squared error (RMSE = 656.62 kg/ha) was achieved. The study also used the rfPermute function in RStudio to compute p values for the important variables (Archer, 2020).

4.2.4. Partial dependence plots

The study used partial dependence plots (PDPs) to assess the relationships between the independent variables and yield. PDPs were developed to interpret complicated regression models (or black boxes) (Friedman, 2001). PDPs are derived by using Equations (6) and (7):

$$\hat{f}_{x_S}(x_S) = E_{x_C}\left[\hat{f}(x_S, x_S)\right] = \int \hat{f}(x_S, x_C) d\mathbb{P}_{(x_C)} \qquad (6)$$

where x_S is the variable for which partial dependence is assessed, x_C are the other independent variables in the model $\hat{f}$, and $\hat{f}_{x_S}$ is calculated from a set of training data;

$$\hat{f}_{x_S}(x_S) = \frac{1}{n}\sum_{i=1}^{n} \hat{f}\left(x_S, x_C^{(i)}\right) \qquad (7)$$

where $x_C^{(i)}$ are the values of the other variables – the x_c variables, and n is the number of instances in the data.

4.3. Results

The results are presented in the form of descriptive statistics of factors that influence maize yield (sub-section 4.3.1), and rankings of these factors (sub-section 4.3.2).

4.3.1. Descriptive statistics of factors influencing maize yields in ORTDM

Table 4.2 shows descriptive statistics of factors that influence maize yield in ORTDM.

Table 4.2: Descriptive statistics of factors influencing maize yields

Important variables	**Description**	**Mean**	**Std dev**	**Min**	**Max**
Yield	Kilograms per hectare (kg/ha)	3259.16	1815.93	367.10	7449.13
Seed	Pioneer PHB3356BR	0.54	0.51	0	1
	Monsanto 7674BR	0.20	0.41	0	1
	Pannar PAN14	0.26	0.44	0	1
MSI	Moisture stress index	0.81	0.12	0.60	1.09
Fertilizer	Application rate (kg/ha)	166.67	26.49	100	200
GDD	Heat units	1531.95	231.64	948.64	2013.16
Precipitation	Rainfall in mm	484.30	142.70	110.15	709.50
pH	Soil alkalinity/acidity	5.55	0.40	4.89	6.57

Machinery	Owned	0.41	0.50	0	1
	Hired	0.59	0.50	0	1

Over the three-year period investigated in this study, the farms produced between 367.10 and 7449.13 kg/ha of maize with an average yield of 3259.16 kg/ha. Average yields in 2018, 2019, and 2020 were 2946.40, 3067, and 3509 kg/ha, respectively. Fifty-four percent (54 %) of the farmers planted Pioneer PHB3356BR, while 20 % and 26 % planted Monsanto 7674BR and Pan-14 seeds, respectively (hereafter referred to as Pioneer, Monsanto, and Pannar). Moisture stress ranged between 0.60 and 1.09 with a mean of 0.81 (MSI values typically range between 0 and >3). Most farmers applied 150 and 200 kg/ha of the same NPK fertilizer that was provided by the DALRRD in partnership with Grain South Africa (GrainSA). Only one farmer used a fertilizer application rate of 100 kg/ha. Accumulated GDD ranged between 948.64 and 2013.16 with a mean of 1531.95, while precipitation ranged between 110.15 and 709.50 mm with a mean of 484.30 mm. Soil pH ranged between 4.89 and 6.57 with a mean of 5.55. Fifty-nine percent (59 %) of the farmers used hired machinery and 41 % used their own.

4.3.2. Variable importance

Figure 4.2 shows variables that were ranked as the most important.

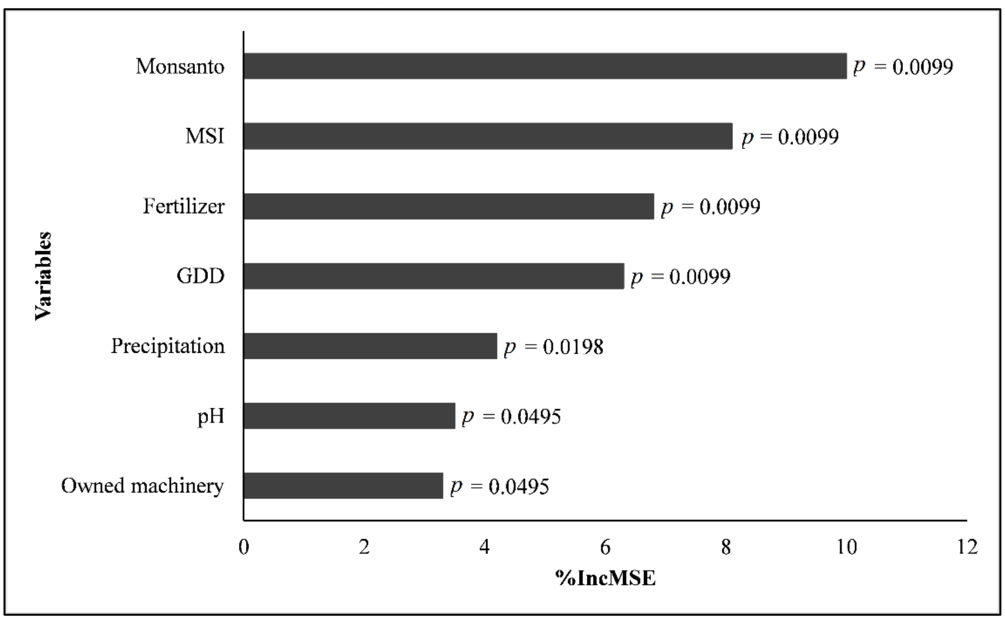


Figure 4.2: Important variables affecting maize yields

Results of variable ranking in Figure 4.2 show that maize yield was highly dependent on seed variety, surface water content as measured by MSI, fertilizer application rate, and GDD ($p < 0.01$). Yield was also dependent on precipitation, soil pH, and ownership of machinery ($p < 0.05$). Figure 4.3 shows how these variables were associated with maize yield. Seed variety was the most important variable. With yields above 3000 kg/ha, the Monsanto seed (1.00) performed better than the other seeds (0.00, Figure 4.3a). Surface water content was the second most important variable, with Figure 4.3b showing that yields between 2750 and 3500 kg/ha generally decreased as MSI increased. The third most important factor was fertilizer application rate as seen in Figure 4.3c, which shows that above 3000 kg/ha of yield, 100 and 150 kg/ha of fertilizer generally performed poorer than 200 kg/ha. The fourth most important factor was GDD in Figure 4.3d, which shows that yields above 2750 kg/ha generally increased with the amount of accumulated GDD. The fifth most important factor was total precipitation.

There was no notable response of yield to precipitation below 400 mm; however, there was a drastic increasing trend of yield between 3000 and 3500 kg/ha, which was as a function of precipitation above 400 mm (Figure 4.3e). The sixth most important factor was soil pH, as shown in Figure 4.3f that farms with soil pH above 5.0 produced higher maize yields. Lastly, Figure 4.3g shows that above 3000 kg/ha of yield, farmers who owned machinery (1.00) achieved higher yields than those who hired machinery (0.00).

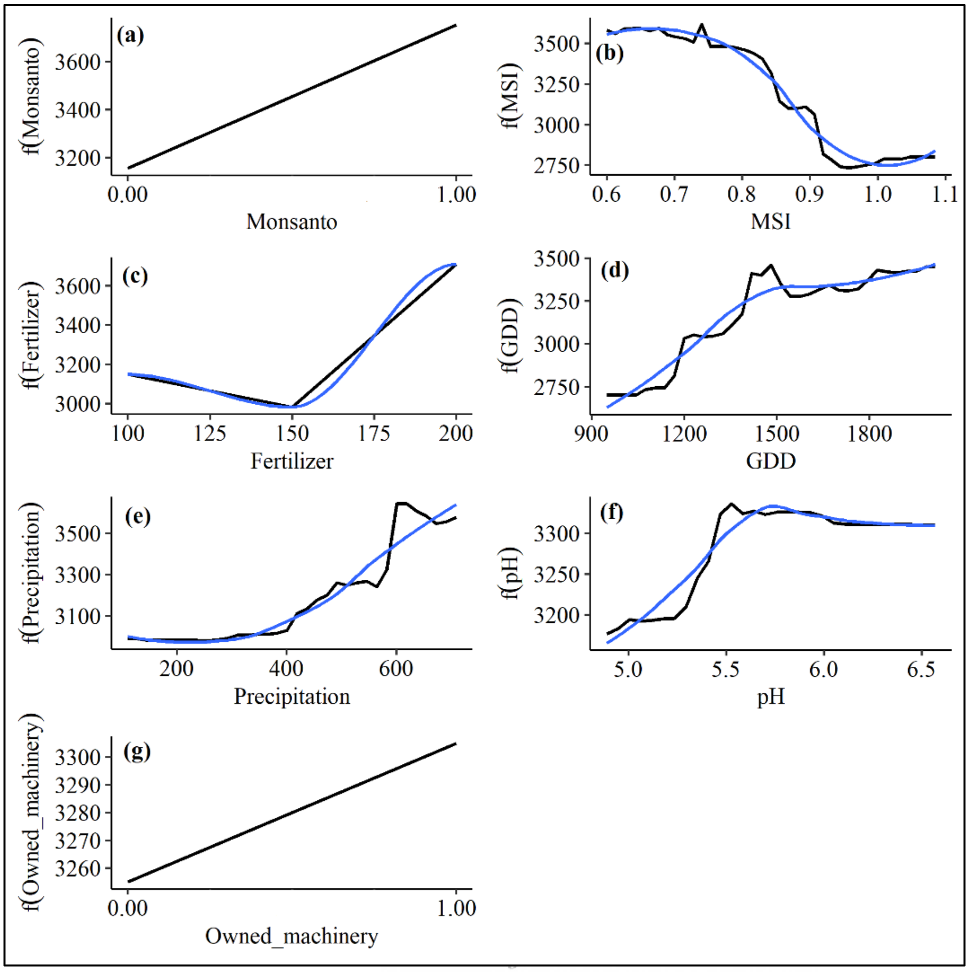


Figure 4.3: Partial dependence plots (PDPs) showing effects of variables on yield. Monsanto (1.00) vs other seeds (0.00) (a) MSI (b) Fertilizer (kg/ha) (c) GDD (d) Precipitation (mm) (e) Soil pH (f) Owned machinery (1.00) vs hired machinery (0.00) (g)

There were notable correlations between the independent variables, as shown in Table 4.3.

Table 4.3: Notable correlations between the independent variables

Correlated variables	Correlation coefficient (r)
MSI, EVI2	-0.52
GDD, Planting date	0.85
pH, Ca	0.63
pH, Mg	0.60
pH, K	0.56
pH, Na	0.55
Owned machinery, Hired machinery	-1.00
Owned machinery, Male farmers	0.67

Owned machinery, Community farms	-0.55
Owned machinery, Individual-owned farms	0.55

Some of the independent variables with lower importance scores were significantly correlated with the highly important variables (Table 4.3). MSI correlated with EVI2, while GDD correlated strongly with planting date. Ownership of machinery correlated with the hiring of machinery, male farmers, collective farms, and individually owned farms. Soil pH correlated with Ca, Mg, K, and Na.

4.4. Discussions and conclusions

This study investigated factors that are affecting maize yields in ORTDM, South Africa. The purpose of this investigation was to identify critical factors that influence maize yield and inputs that IBCI and creditors could provide or assist farmers to invest in. Over the 3-year period between 2018 and 2020, maize yields ranged between 367.10 and 7449.13 kg/ha with an average of 3259.16 kg/ha, which reveals that most farmers produced below the national average, which often ranges from 4700 to more than 7000 kg/ha (USDA, 2018; Greyling and Pardey, 2019; Haarhoff et al., 2020). Recent studies also reported that maize yields in ORTDM and the Eastern Cape province at large are low and less than the potential productivity of the area (Chimonyo et al., 2019, 2020; Kambanje et al., 2020). The most important input was seed variety. The farmers were receiving extension services and input recommendations from the DALRRD and GrainSA (a private association of grain farmers). GrainSA recommended 200 kg/ha of fertilizer and supplied the farmers with a Monsanto seed variety, whereas the DALRRD recommended 150 kg/ha of fertilizer and supplied the other farmers with Pannar and Pioneer seed varieties.

Although 200 kg/ha of fertilizer was associated with higher yields, the fertilizer application rates were less than what is generally recommended (>250 kg/ha) for maize in many parts of South Africa (van Averbeke et al., 2013; Lotriet et al., 2017; Diko and Jun, 2020). The low usage of fertilizer among South Africa's smallholder farmers is partly due to the perception that the widely recommended fertilizer application rates are unrealistic, risky, expensive, and meant for resource-rich farmers (Minde et al., 2008; Chimonyo et al., 2020). Therefore, the reason Pannar and Pioneer seed varieties produced lower yields compared to Monsanto could be due to the lower fertilizer application rates. Smallholder farmers also have limited experience with new seed varieties. In

2019, most of the farmers complained that the Pioneer seed was new and that it reached senescence prematurely.

The second most important factor was moisture stress, which was significantly associated with EVI2. MSI is sensitive to surface water content and EVI2 is sensitive to plant chlorophyll content, which influences the rate of photosynthesis and crop yield (Ghimire et al., 2015). Interestingly, MSI was more important than precipitation. Recent studies on IBCI explored the feasibility of using soil moisture rather than rainfall indices (Black et al., 2016; Enenkel et al., 2017; Osgood et al., 2018) because surface water content, rather than precipitation, is a better indicator of water availability to plants and a better measure of agricultural drought. Although there is evidence that remotely sensed MSI can estimate soil moisture and vegetation water content (Benabdelouahab et al., 2015; Welikhe et al., 2017), no studies in Africa's IBCI systems have investigated the effectiveness of MSI and other moisture indices that are derived from optical data.

The most important edaphic factor was soil pH, which was associated with base saturation (Table 15). In areas like ORTDM, where soils are largely acidic, an IBCI bundle would need to include lime to help neutralise soil pH, improve nutrient availability, minimise production risk, and enhance crop productivity. Lastly, ownership of machinery was associated with males who cultivate maize in individually owned farms. This shows that cooperatives and female-owned farms are more vulnerable to production risk. Farmers without equipment tend to cultivate smaller areas, delay the application of agronomic inputs, and lose portions of their harvest, whereas equipped farmers produce higher yields because of timely operations and improved labour productivity (Sims and Kienzle, 2016). Encouraging farmers and the government to invest in affordable implements instead of focusing on large machines, which are not only expensive but also unsuitable for farmers with small fields, could improve access to equipment (Van Loon et al., 2020). Alternatively, insurance could unlock credit for farmers so that they get the money needed for timely hiring of equipment.

The factors influencing maize yield in ORTDM are some of the common factors affecting crop yields in Africa's SFS. However, the exact nature and importance of these factors may vary from one place to another depending on the socioeconomic and environmental conditions. Studies in other localities could provide more insight on some of the most important factors that need to be addressed in SFS to minimise production risk and to improve agricultural productivity. More research is also needed to find efficient ways by which IBCI and agricultural inputs can be systematically packaged into comprehensive risk management portfolios. Although linking IBCI with factors that influence crop yield may attract farmers to take up insurance, more work still needs to be carried out to reduce basis risk, which is one of the reasons why insurance uptake remains low (Clement et al., 2018). Addressing non-weather factors by bundling or linking insurance with input supplies and advisories could enable IBCI to accurately isolate quantify the influence of weather on crop losses.

In conclusion, ORTDM experienced no weather shocks over the three seasons investigated in this study; therefore, weather conditions were suitable for maize cultivation. The below-average yields demonstrate that maize production in ORTDM could plummet even further in the event of a moderate weather shock (e.g. a mild drought). Producing low yields under optimal weather conditions also shows that non-weather variables are in play. For instance, the importance of agrometeorological factors on yield was largely associated with planning and planting dates. In countries like South Africa where farmers delay planting because of the late delivery of inputs and weather risks, linking IBCI with inputs, advisories, and credit would ensure that farmers have timely access to mechanization and that they use the appropriate seed varieties and fertilizers. This approach has worked well in other countries like India and Kenya, for example, where insured farmers also get advisory services and weather information via mobile phones (Mukherjee et al., 2017; Raithatha and Priebe, 2020). Accomplishing this requires a strategic synchronization of efforts. Acknowledging the influence of non-weather elements on crop yields will potentially enhance risk management and risk reduction. Future research could focus on how these non-weather elements can be incorporated into the design and packaging of IBCI. Lastly, additional research could explore moisture indices like MSI, which could potentially reduce basis risk. This study measured MSI from Sentinel-2 images, which are not available at the desirable time steps because of clouds. The next chapter investigated the feasibility of retrieving soil moisture content by combining Sentinel-2 and Sentinel-1.

Chapter 5 : Estimating soil water content using SAR, optical data, and support vector regression

Abstract

The amount of water in the soil during the growing season affects the development of crop plants. Therefore, it is necessary to monitor soil water content (SWC) throughout the crop growth cycle. One of the techniques used to estimate SWC is satellite remote sensing, which has the advantage over other methods because it provides spatially continuous measurements over large geographic areas. However, the currently available SWC products derived from satellite data have limited applications in small-scale crop farms because of their coarse spatial resolutions. This study investigates how SWC in small-scale maize fields can be estimated with Support Vector Regression (SVR) and a combination of Sentinel-1 data and Sentinel-2 derived spectral indices. Results show that the modelling approach used in this study better estimates SWC when leaf area and plant height are low compared to when they are high. Furthermore, the combination of Sentinel-1 and optical indices performs better than the standalone use of Sentinel-1. The highest estimation accuracy achieved is a RMSE of 0.043 m^3 m^{-3} and a corresponding MAE of 0.034 m^3 m^{-3}. This demonstrates that the methods used in this study are capable of measuring SWC in small-scale crop fields although further improvements can be achieved by investigating the effects of variable interactions.

5.1. Background

Soil water content (SWC) is one of the most important factors affecting the growth and development of crops and natural vegetation. Soil water is one of the basic requirements for seed germination and plant emergence (Kozlowski, 1972; Itabari et al., 1993; Shaban, 2013). Likewise, the vegetative stages of crop growth need water for leaf area development, stem elongation, and biomass accumulation (Çakir, 2004). In the reproductive stages of maize growth, SWC influences grain filling and the number of kernels and ear weight (Yang et al., 1993; Ge et al., 2012; Mi et al., 2018). This means that SWC is a critical determinant of crop yield especially in Sub-Saharan Africa where crop farming is predominantly rain-fed. SWC has to be constantly monitored at appropriate spatial scales to enable mitigation of the adverse effects of unpredictable soil water deficits.

Over the years, different methods of measuring SWC have been developed. Each of these approaches has its advantages and disadvantages depending on the needs and objectives of the project. Sharma et al., (2018) reviewed the strengths and weaknesses of these methods. Myeni et al., (2019) reviewed SWC estimation methods that have been used in Africa. These studies report that although gravimetric techniques and point-based soil moisture probes have been routinely used to accurately measure SWC, these in-situ methods are costly and incapable of providing timely information at large spatial scales. To address these limitations, researchers have strived to develop reliable ways to retrieve SWC from optical and microwave satellite data (Petropoulos et al., 2015; Karthikeyan et al., 2017a, 2017b; Rodríguez-Fernández et al., 2019). The main advantages of using satellite data are (1) cost effectiveness and (2) their ability to cover large spatial scales including inaccessible places at appropriate temporal scales.

Passive microwave sensors measure naturally emitted radiation from the surface of the earth in the form of radiance temperatures, while active microwave sensors emit focused radiation towards the earth's surface which they then sense back as backscatter signal (Karthikeyan et al., 2017a; Sharma et al., 2018). SWC is measurable with microwave sensors because the dielectric properties of dry and wet soil influence the emitted radiation and the backscatter signal. Unlike optical radiation, microwaves can penetrate clouds and are less affected by atmospheric effects (Petropoulos et al., 2015). Although several studies have explored the applicability of the European Space Agency's Climate Change Initiative Soil Moisture product (CCI-SM) in index-based crop insurance (IBCI)

(Enenkel et al., 2017, 2018, 2019; Osgood et al., 2018), this product and other frequently used satellite-based SWC products have limited applications in small-scale agriculture because of their coarse spatial resolutions. Also, the backscatter signal measured by Synthetic Aperture Radars (SAR) is affected by soil surface roughness, vegetation cover, and sensor configurations including polarization, frequency, and incidence angle (Sharma et al., 2018). This complicates the retrieval of SWC from SAR because the effects of these factors can be accounted for by complex physical models which require in-situ data (Karthikeyan et al., 2017a).

However, several studies report that combining SAR data and optical data can reduce the effects of vegetation cover and soil physical properties. In Spain for example, Gao et al., (2017) combined Sentinel-1 and Sentinel-2 imagery using a change detection method which does not require site calibrations and achieved estimation accuracies of 0.087 m^3 m^{-3} and 0.059 m^3 m^{-3}. In Ethiopia, Ayehu et al., (2020) achieved accuracies of 0.058 m^3 m^{-3} and 0.097 m^3 m^{-3} by applying a stepwise cluster analysis on Sentinel-1, MODIS Normalized Difference Vegetation Index (NDVI) and a digital elevation model. In an experimental study conducted in several countries including Tunisia, Niger, Benin, and France, Foucras et al., (2020) achieved estimation accuracies of less than 6 %Vol by applying a change detection method on different combinations of Sentinel-1, Sentinel-2, and MODIS data.

This study investigates the feasibility of estimating SWC in small-scale maize fields by using different combinations of Sentinel-1 polarimetric bands and Sentinel-2 derived indices. SWC is estimated over the first half of the growing season when leaf area is low to moderate, over the second half of the season when leaf area is high, and over the entire season. The data are fed into the support vector regression (SVR) trained with a polynomial kernel, which permits interactions between input variables. The choice of this algorithm was based on the results of preliminary experimenting with different algorithms. The objectives of this investigation were to; (1) assess the performance of combined SAR, NDVI and Normalised Difference Infrared Index (NDII) in estimating SWC in smallholder maize fields and (2) assess their performance over the first half of the growing season, over second half and over the entire season.

5.2. Materials and methods

5.2.1. Study area

The research was conducted in O.R Tambo District Municipality (ORTDM), which is situated in the Eastern Cape Province of South Africa (Figure 5.1).

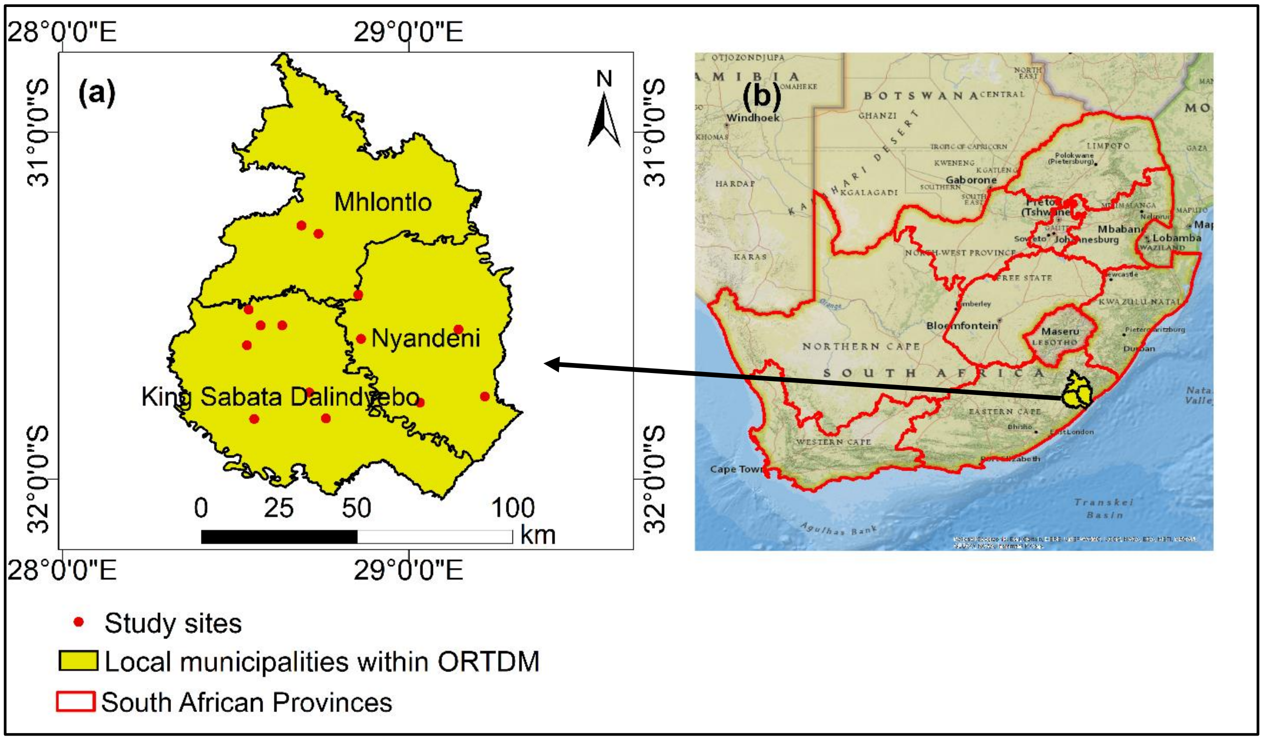


Figure 5.1: Location of ORTDM (a) in South Africa (b)

The most cultivated crop in the area is maize. Planting starts from mid-November to late-January and harvest season is usually between June and August when kernel moisture is low. The area has a warm oceanic climate; the humid subtropical climate of the northeastern parts and the semi-arid climates of the southwestern areas (Beck et al., 2018). Mean annual rainfall ranges from 900 mm to 1300 mm. Growing season minimum and maximum temperatures range from 14 °C to 19° C and 14 °C to 27 °C, respectively (Jordaan et al., 2017). The growing season is from November to April and it coincides with summer, which is also the rainy season. The physiography consists of densely vegetated and undulating landscapes along the coast where elevation is 5 m to 500 m, gentle-to-moderate sloping grassland in the interior, and savannas and forests in the northern areas where maximum elevation is 1500 m. The soils are sandy loams, sandy clay loams, and clays that are yellow to black in colour and slightly acidic (Eta and Grace, 2013; Sibanda et al., 2016).

5.2.2. In situ soil moisture measurements

In-situ soil moisture data were collected from 14 farms (Figure 5.1) over the 2019-2020 maize growing season. A combination of judgmental and convenience sampling were used to identify farms that are representative of the study area. These arms were selected on the basis that they were far part from each other, distributed across the study area, and safe enough to keep SWC probes from being vandalized. A minimum of two Dirk Friedhelm Mercker (DFM) capacitance probes were installed in each farm immediately after planting (Figure 5.2). The DFM probes have multiple thermal and capacitance sensors that measure temperature and dielectric permittivity at various depths from 10 cm to 80 cm and within a radius of 10 cm from the probe. The uppermost sensor was configured to take readings in the top-most layer (0 cm < depth <10 cm). For sensor calibration, soil samples were collected over a 15-day period during which soil water conditions were controlled to simulate different levels of water saturation.

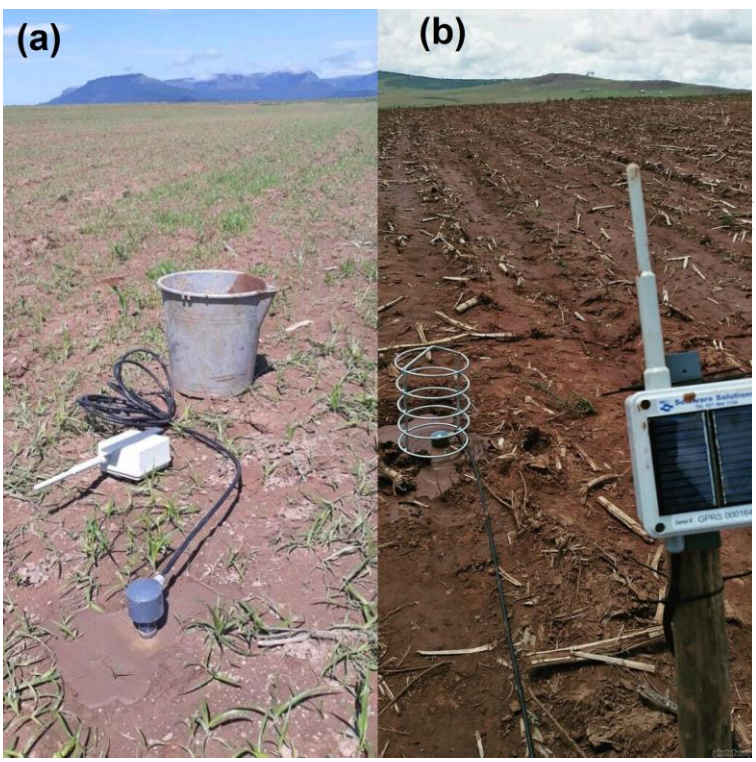


Figure 5.2: Photos showing soil moisture probes installed in two different fields

These samples were weighed immediately and then weighed again for dry mass after oven drying to derive gravimetric SWC. The gravimetric SWC was multiplied by soil bulk density to derive volumetric SWC. The probes were then calibrated using linear regression by relating dielectric permittivity to volumetric SWC. R^2 values ranged from 0.92 to 0.98 indicating that the probes were sensitive to SWC. DFM probes have been tested and validated in other studies as well (Zerizghy et al., 2013; Mjanyelwa et al., 2016; Myeni et al., 2021).

5.2.3. Satellite data

Sentinel-1 and Sentinel-2 images covering the 2019-2020 growing season were downloaded from the European Space Agency's Copernicus Hub (Table 5.1). Interpolation was applied to align the Sentinel-2 data with Sentinel-1 where the difference between the satellite acquisition dates was more than one day.

Table 5.1: Acquisition dates satellite images

Sentine1-1	Sentinel-2
22 December 2019	-
15 January 2020	17January 2020
27 January 2020	29 January 2020
08 February 2020	13 February 2020
20 February 2020	26 February 2020
03 March 2020	04 March 2020
15 March 2020	14 March 2020
27 March 2020	-
08 April 2020	08 April 2020
20 April 2020	18 April 2020
02 May 2020	03 May 2020
14 May 2020	16 May 2020

The Sentinel-1 data were level-1 Ground Range Detected (GRD) images with a 250 km Interferometric Wide (IW) swath. Pre-processes applied to Sentinel-1 were orbit file application, thermal noise removal, radiometric calibration, speckle filtering and multi-looking, terrain correction, and conversion of sigma nought backscatter to decibels. The Sentinel-2 data were level-1 multispectral bands, free of cloud cover over the areas that were targeted for investigation. These multispectral bands were atmospherically corrected from Top-Of-Atmosphere (TOA) to Bottom-of-Atmosphere (BOA) reflectance. The red (RED), near infrared (NIR), and short-wave infrared (SWIR) bands were then used to calculate the NDVI and the NDII according to equations 1 and 2.

$$\text{NDVI} = \frac{\text{NIR}-\text{RED}}{\text{NIR}+\text{RED}} \quad (1)$$

$$\text{NDII} = \frac{\text{NIR}-\text{SWIR}}{\text{NIR}+\text{SWIR}} \quad (2)$$

NDVI was selected because of the strong dependence of plants on SWC (Denmead and Shaw 1960; Çakir 2004; Engstrom et al., 2008; Chen et al., 2014; Boke-Olén et al., 2018) and because studies have observed good relationships between SWC and NDVI (Engstrom et al., 2008; Boke-Olén et al., 2018; Zhang et al., 2018). NDII was selected because the strong absorption of the SWIR by water molecules makes it a good indicator of SWC (Tian and Philpot, 2015). A number of studies report that NDII is capable of mapping temporal and spatial variations in SWC and leaf water content (Wang et al., 2008; Sriwongsitanon et al., 2015; Zhang et al., 2018 Xu et al., 2020). NDII is related to the Moisture Stress Index (MSI), which was identified as one of the most important yield-determining factors in chapter 4. Both indices are based on the same wavebands; MSI is a simple ratio, whereas NDII is based on normalized difference formulation.

5.2.4. Data analysis

A series of preliminary tests were performed to estimate SWC using different machine leaning algorithms with the in-situ SWC data as the response variable and the satellite data as input data. Ultimately, support vector regression (SVR) produced the lowest RMSEs and was used in the main analysis. The SVR used in this study was trained with the polynomial kernel because it performed better than the linear and the radial basis function kernels in the preliminary tests. Another advantage with the polynomial kernel is that it also accounts for interactions between input variables (Morita et al., 2009). The main analysis involved different regression trials in which different variables derived from the satellite data were investigated (Table 5.2).

Table 5.2: Input variables in different regression trials

Variable Identifier (ID)	Variables
V1	VH
V2	VV
V3	VH, NDVI
V4	VH, NDVI, NDII
V5	VV, NDVI
V6	VV, NDVI, NDII
V7	VV, VH, NDVI, NDII

The data was organized into three groups covering the periods, 22/12/2019 to 03/03/2020, 14/03/2020 to 14/05/2020, and 22/12/2019 to 14/05/2020.

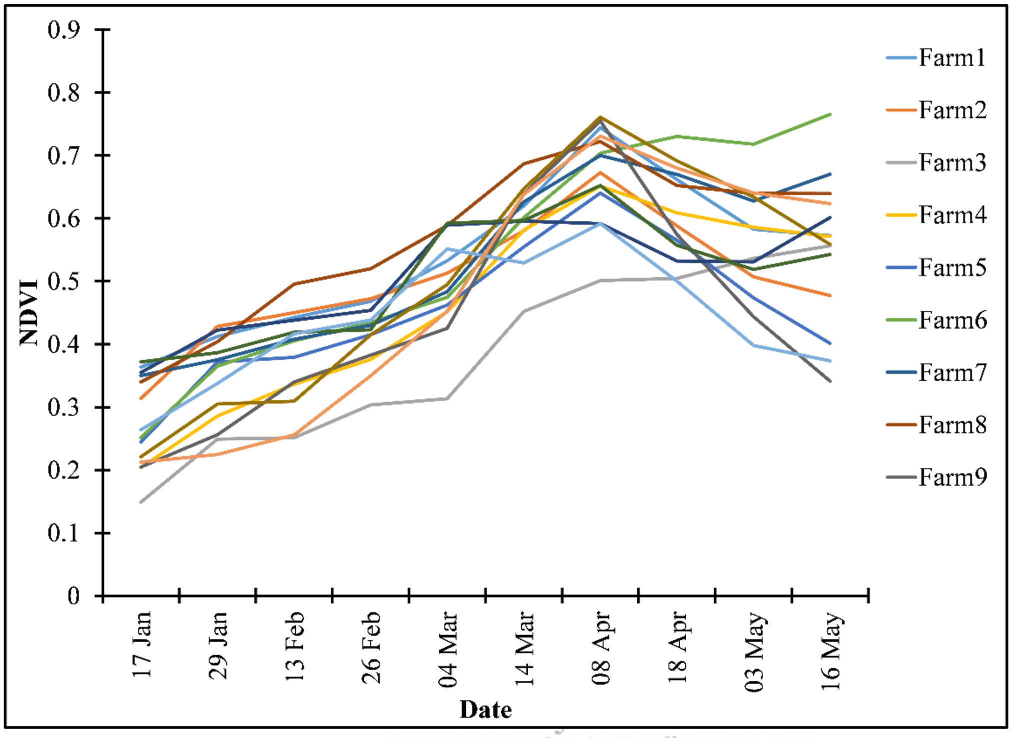


Figure 5.3: Seasonal NDVI profiles of the 14 farms investigated in the study

The first period corresponded to the phase during which leaf area and plant height are low to medium. NDVI values for this period were 0.13 to 0.59. The second period corresponded to the phase during which plants were fully-grown. NDVI values for this period were 0.30 to 0.77. The third period covered the entire season.

5.2.5. Model performance assessment

Root mean square error (RMSE) (equation 3) and mean absolute error (MAE) (equation 4) were used as performance metrics to assess the accuracy SVR and the satellite data in estimating SWC.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\theta_i - \hat{\theta}_i)^2} \tag{3}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|\theta_i - \hat{\theta}_i| \tag{4}$$

In equations 3 and 4, θ and $\hat{\theta}$ are the in-situ measured and predicted SWC, respectively, n is the number of the observations and i the ith observation (Chai and Draxler, 2014).

5.3. Results

Figure 5.3 shows the results of SWC estimation with satellite data over the first half of the growing season.

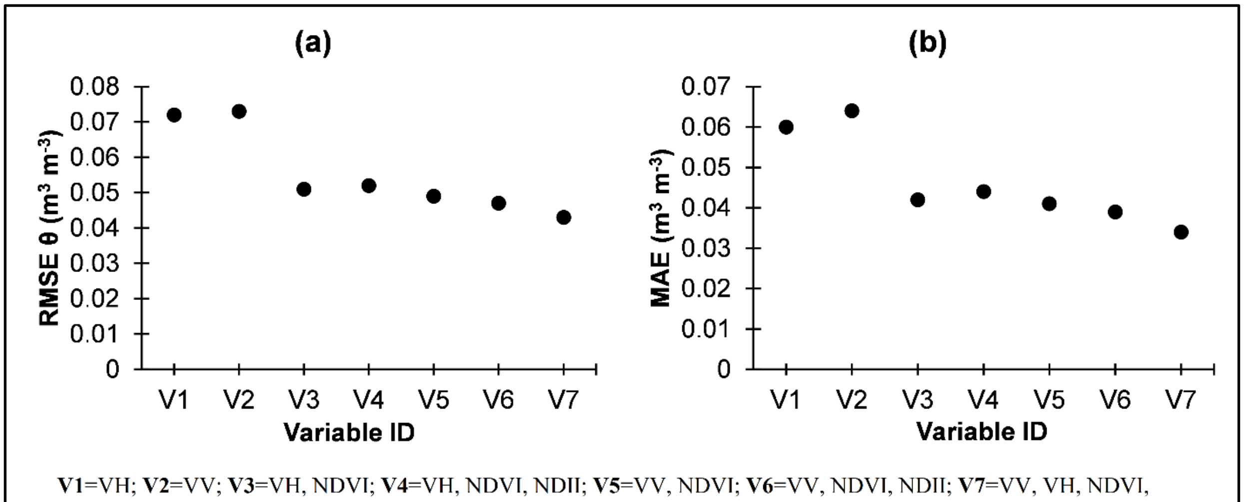


Figure 5.4: Satellite-based SWC estimation over the first half of the growing season

Over the first phase of the growing season, SWC was predicted with RMSEs of 0.043 $m^3\,m^{-3}$ to 0.073 $m^3\,m^{-3}$ (Figure 5.3). The variable that produced the highest prediction accuracy (RMSE = 0.043 $m^3\,m^{-3}$, MAE = 0.034 $m^3\,m^{-3}$) was the V8 combination of all the predictor variables. The second highest accuracy (RMSE = 0.047 $m^3\,m^{-3}$, MAE 0.039 $m^3\,m^{-3}$) was achieved by the V6 combination of VV, NDVI, and NDII. The standalone use of Sentinel-1 achieved low accuracies, with the VV channel producing a RMSE of 0.073 $m^3\,m^{-3}$ (MAE = 0.064 $m^3\,m^{-3}$), and the VH channel producing a RMSE of 0.072 $m^3\,m^{-3}$ (MAE = 0.060 $m^3\,m^{-3}$). Figure 5.4 shows the results of SWC estimation with satellite data over the second half of the growing season.

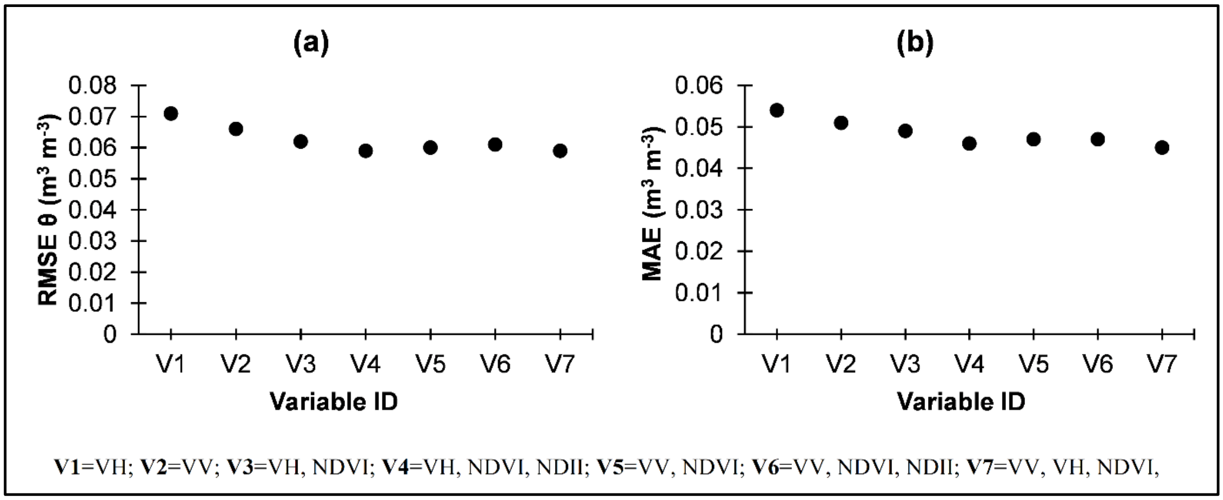


Figure 5.5: Satellite-based SWC estimation over the second half of the growing season

Over the second phase of the growing season, estimation accuracies ranged between RMSE = 0.059 $m^3 m^{-3}$ and RMSE = 0.071 $m^3 m^{-3}$ (Figure 5.4). The combination of all the variables (V8) produced the highest prediction accuracy (RMSE = 0.059 $m^3 m^{-3}$, MAE = 0.045 $m^3 m^{-3}$). The second highest accuracy (RMSE = 0.059 $m^3 m^{-3}$, MAE = 0.046 $m^3 m^{-3}$) was achieved by the combination of VH, NDVI, and NDII (V4). The standalone use of Seninel-1 produced the lowest accuracies, with the VH channel producing a RMSE of 0.071 $m^3 m^{-3}$ (MAE = 0.054 $m^3 m^{-3}$) and the VV channel producing a RMSE of 0.066 $m^3 m^{-3}$ (MAE = 0.051 $m^3 m^{-3}$). Figure 5.5 shows the results of SWC estimation with satellite data over the entire growing season.

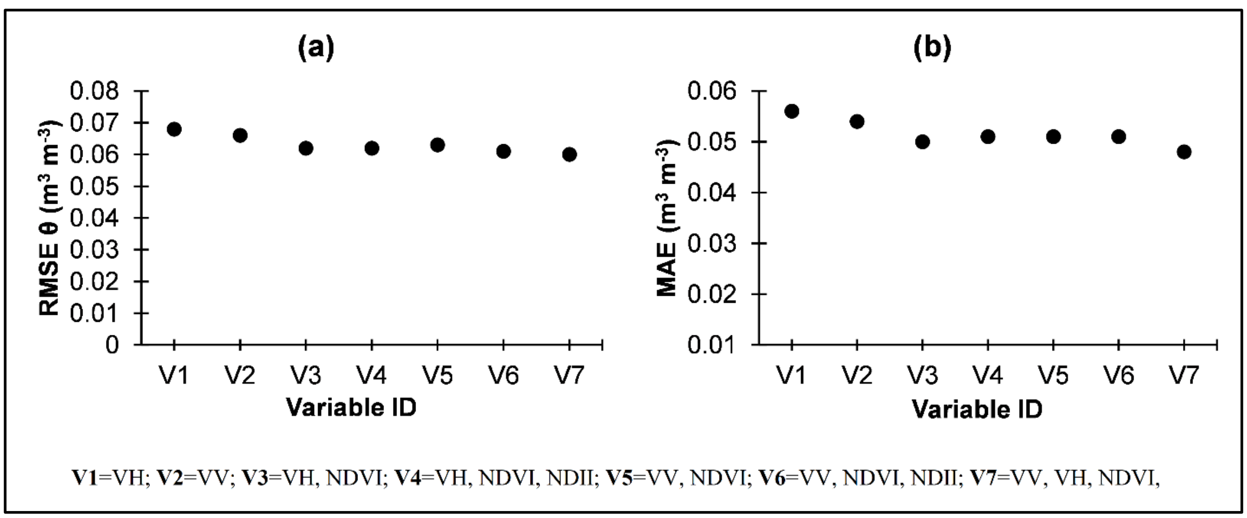


Figure 5.6: Satellite-based SWC estimation over the entire growing season

Over the entire season, estimation accuracies ranged between RMSE = 0.060 $m^3 m^{-3}$ and RMSE = 0.068 $m^3 m^{-3}$ (Figure 5.5) with the combination of all the variables (V8) achieving the highest accuracy (0.060 $m^3 m^{-3}$, MAE = 0.048 $m^3 m^{-3}$). The V3 combination of VH and NDVI achieved the second highest accuracy (RMSE = 0.062 $m^3 m^{-3}$, MAE = 0.050 $m^3 m^{-3}$). The standalone use of Sentinel-1 produced the lowest accuracies, with the VH channel producing a RMSE of 0.068 $m^3 m^{-3}$ (MAE = 0.056 $m^3 m^{-3}$) and the VV channel producing a RMSE of 0.066 $m^3 m^{-3}$ (MAE = 0.054 $m^3 m^{-3}$).

5.4. Discussions and conclusions

SWC was estimated in smallholder maize fields over the growing season using different combinations of SAR polarimetric channels, NDVI and NDII. The findings show that the combination of SAR and optical indices outperforms the standalone use of SAR in estimating SWC. More importantly, the addition of the NDII to the typical combination of SAR and NDVI further improves SWC estimation accuracy. This finding demonstrates that combining complementary datasets that are associated with the target minimizes prediction error. However, further investigation into feature interactions could provide more insight on how the combination of these datasets could be optimized because estimation accuracy was higher over the first half of the growing season than the second half. This highlights the influence of the plant growth stages on the radiation received by the satellite sensors. Over the first half of the growing season, NDVI values ranged from 0.13 to 0.59, while ranging from 0.30 to 0.77 (Figure 5.6) over the second half due to canopy cover (the maximum values for the second half could have been higher with a vegetation index that does not saturate as NDVI does).

Despite these discrepancies, the spectral indices were still able to boost the estimation accuracies of SWC in all three phases. The poor performance of the standalone use of SAR may be attributed to the omission of factors such as surface roughness and incidence angle, which contribute to SAR backscatter. Further improvements could be achieved by either accounting for these factors or by reducing their effects on the SAR backscatter. Although some studies account for these factors by applying physical models and Neural Networks (Mirsoleimani et al., 2019; Ayehu et al., 2020), Gao et al., (2017) argue that, since Sentinel-1 has a high spatial resolution, resampling to a lower resolution that is still suitable for agricultural applications could minimise the effects of surface roughness and vegetation heterogeneity.

The novelty of the approach used in this study was the 1) addition of the NDII to the combination of SAR data and NDVI and 2) the insights drawn from combining these datasets under different levels of vegetation cover. Although the approach can still be improved, the levels of accuracy achieved, especially over the first phase of the growing season, are comparable to those achieved by other researchers (Gao et al., 2017; Mirsoleimani et al., 2019; Adab et al., 2020; Ayehu et al., 2020; Xu et al., 2020). However, more research is needed to further reduce estimation errors by addressing the shortcomings highlighted above. Additional limitations include the revisit frequencies of Sentinel-1 and Sentinel-2 which make it difficult to monitor temporal variations in SWC over shorter time periods to meet the requirements of crop insurance. Additional research is needed to explore how different optical datasets like MODIS, Landsat, Sentinel-3 and others can be synergistically used with SAR data to bridge this gap. It is due to this limitation that the design of crop insurance in next chapter was based on rainfall rather than soil moisture.

Chapter 6 : A proposed model for a satellite-based crop insurance for South Africa's smallholder maize farming

Abstract

Crop farming in Sub-Saharan Africa is constantly confronted by adverse weather events. Researchers have been developing different tools that can be used to reduce the impacts of adverse weather on agriculture. Index-based crop insurance (IBCI) has emerged to be one of the tools that could potentially hedge farmers against weather-related risks. However, IBCI is still con-strained by poor product design and basis risk. This study complements the efforts to improve the design of IBCI by evaluating the performances of the Tropical Applications of Meteorology using SATellite data and ground-based observations (TAMSAT) and Climate Hazards Group In-fraRed Precipitation with Station data (CHIRPS) in estimating rainfall at different spatial scales over the maize-growing season in a smallholder farming area, South Africa. Results show that CHIRPS outperforms TAMSAT and produces better results at 20-day and monthly time steps. The study then develops an IBCI payout model by deriving index thresholds using CHIRPS and a crop water requirements (CWR) model. Results of CWR modeling show that IBCI can cover the development, mid-season, and late-season stages of maize growth in the study area. The study then uses this information to calculate the weight, trigger, exit, and tick of each growth stage. The proposed IBCI also incorporates crop area mapping and links insurance with factors that influence crop yields and crop losses.

6.1. Background

Smallholder farming is an integral part of Sub-Saharan Africa's (SSA) economies(Akinnagbe and Irohibe, 2014; Langyintuo, 2020). However, frequent occurrences of extreme weather events in this region disrupt crop farming and hamper economic development (Buhaug et al., 2015; Gebremeskel et al., 2019; Sultan et al., 2019). Drought is the most prevalent source of risk in Sub-Saharan Africa (Carter et al., 2014; Masih et al., 2014) and it threatens food security, increases poverty, and hampers economic development by causing production fluctuations and low yields (Manderson et al., 2016; Gautier et al., 2016; Gebremeskel et al., 2019; Nembilwi et al., 2021). These devastating impacts of drought need to be addressed with enhanced risk management and risk reduction strategies (Kahan, 2013; Hansen et al., 2019; Komarek et al., 2020). Reports show that index-based crop insurance (IBCI) is potentially capable of hedging SSA's farmers against weather-related risks (Sharoff et al., 2015; ACRE, 2020; WFP, 2020). IBCI is a type of insurance that calculates compensation from a predetermined index rather than a direct proof of the incurred loss (Barnett and Mahul, 2007). The index must be associated with crop health, crop yields and crop losses (Barnett and Mahul, 2007; Ntukamazina et al., 2017). The commonly used indices are rainfall, temperature, soil moisture, evapotranspiration and crop yield-related indicators like the Normalised Difference Vegetation Index (NDVI) and area yields (Ntukamazina et al., 2017; Tadesse et al., 2015). IBCI then issues payouts when these indices deviate from the normal levels associated with healthy crops because the deviations cause crop losses.

By using indices, IBCI avoids moral hazard, adverse selection and high insurance expenses, which plague traditional claim-based insurance systems. IBCI is able to avoid moral hazard and adverse selection because the indices are objectively measured and cannot be easily manipulated by the farmer. In addition, using indices rather than direct loss assessment reduces administrative costs and premiums. Although IBCI is widely considered to be better than traditional claim-based insurance, it is constrained by basis risk (Clement et al., 2018). Basis risk arises when there is a mismatch between the insurance index and the losses experienced by the farmer, which exposes the farmer to the risk of not receiving fair compensation or the insurer to the risk of overcompensating (Clement et al., 2018). Some studies attempt to reduce basis risk by optimizing the relationship between crop yields and the indices mentioned above (Choudhury et al., 2016; Enenkel et al., 2019; Eze et al., 2020). However, these indices are not always the most important factors associated with crop yields and losses.

Masiza et al., (2021) demonstrated that non-weather factors including seed variety, fertilizer application rate, mechanization, and soil pH have a significant influence on the maize yields of O.R Tambo District Municipality (ORTDM), South Africa. Furthermore, they (Masiza et al., 2021) showed that the relationships between maize yield and indices like rainfall and soil moisture were nonlinear and observable on yields above 3000 kg/ha. Other studies conducted elsewhere in Africa attribute yield variability to various weather and non-weather factors (Assefa et al., 2020; Beza et al., 2017; Gornott et al., 2015; Kihara et al., 2015; MacCarthy et al., 2018). Outside Africa, in India for example, Dutta et al., (2020) attributed smallholder maize yield variability to complex interplays between soil factors, seeds, fertilizer, and farm labour, while Banerjee et al., (2014) identified socioeconomic, agronomic, and weather conditions as the major determinants. These findings demonstrate intricate influences of non-weather factors, which make it difficult to determine the exact contribution of weather to yield losses. In addition, the relationship between rainfall and yield does not always follow a linear pattern as often assumed in most cases (Masiza et al., 2022). It is even difficult to calibrate yield-rainfall models in SSA where reliable yield records are lacking (Carletto et al., 2013; Djurfeldt et al., 2018; Osgood et al., 2018).

This study systematically manipulates this complexity by deriving IBCI thresholds and an insurance payout structure from crop water requirements (CWR) and satellite-based rainfall estimates (SRFEs). CWR is the amount of water needed by the crop for optimal growth and development under field conditions at a given place (Pereira and Alves, 2013). A CWR approach was chosen because it is water-driven; it focuses on crop water needs and does not require a large number of input data from non-weather yield-determining factors. In addition, since SSA's smallholder crop farming is rain-fed and vulnerable to water stress (Buhaug et al., 2015; Wani et al., 2009), water deficit (deficit being the difference between actual rainfall and CWR) is the most used parameter in most SSA IBCI programs (ACRE, 2020; ARC, 2020; Belissa et al., 2019; WFP, 2020; Worldbank, 2011). In other words, most often, rainfall-related yield reductions and crop losses result from the failure of rainfall to meet CWR in the water-sensitive stages of the crop (Alcaide et al., 2019; Butts-Wilmsmeyer et al., 2019; Çakir, 2004; Song et al., 2019). SRFEs were selected because satellite data have the advantage over ground-based data to provide spatially continuous measurements covering large geographic areas, which could reduce spatial basis risk. The Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) and Tropical Applications of Meteorology using SATellite data (TAMSAT) were selected because they have higher spatial resolutions than other SRFEs covering Africa (Gebremicael et al., 2017; Le Coz and Van De Giesen, 2020). Therefore, the aims of this study were to (1) compare the performances of CHIRPS and TAMSAT data in estimating rainfall at different spatial and temporal scales (2) design an IBCI payout system that is based on maize CWR and SRFE and incorporates crop area mapping and non-weather factors that influence yields.

6.2. Materials and methods

6.2.1. Study area

The study focused on the ORTDM in the Eastern Cape Province of South Africa. ORTDM's average annual rainfall ranges between 900 mm and 1300 mm. Summer temperatures range between 14°C and 27°C (Jordaan et al., 2017). The physiography is characterized by densely vegetated undulating landscapes of 5 m to 500 m elevations along the wild coast, gently-sloping grasslands in the interior and savannas and forests in the northern areas where elevation rises to 1500 m. The soils are sandy loams, sandy clay loams, and clays that are slightly acidic and yellow to black in colour (Eta and Grace, 2013; Sibanda et al., 2016).

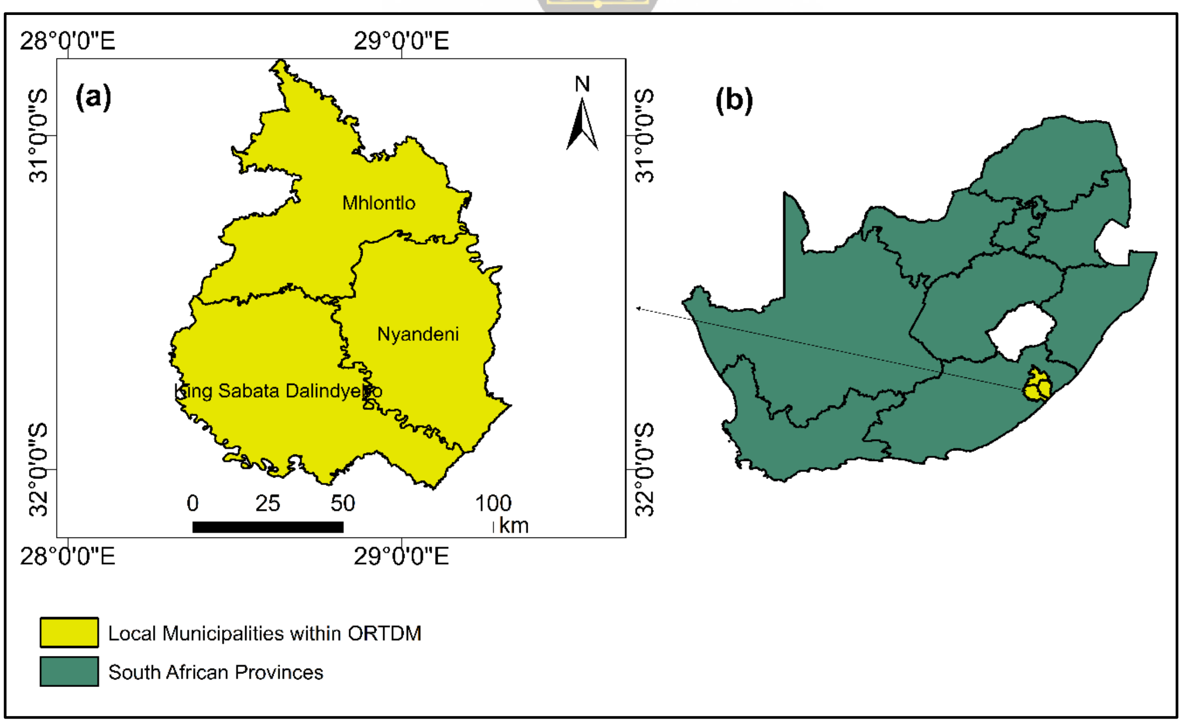


Figure 6.1: Study area (a) within South Africa (b)

The study sites were located in ORTDM's three local municipalities including King Sabata Dalindyebo (3027 km^2), Nyandeni (2474 km^2) and Mhlontlo (2826 km^2) (Figure 6.1a). In these local municipalities, the study was limited around five weather stations, which are distributed within an area of approximately 70 km × 50 km (Figure 6.2). More information about the study area is provided in section 6.2.3. and Figure 6.2. The choice of this area was based on a list of maize producing farms that we obtained from the Department of Agriculture, Land Reform and Rural Development (DALRRD). Maize is the most important grain crop in South Africa; white maize is the major staple food and yellow maize is mostly used for feeding animals (DALRRD, 2018). It is sown from October to early-January and harvested between June and August; however, some of the farmers plant late because of delayed delivery of inputs. Farmers and DALRRD officials indicated that the optimum planting window in the study area is 15 November to 31 December. Therefore, this study focused on the maize growing period which coincides with the rainy season and spans between November and April. The cultivation records we obtained from the DALRRD show that target yield is five tonnes per hectare (t/ha). However, studies show that maize yields in ORTDM are low and often below five t/ha (Chimonyo et al., 2019, 2020; Masiza et al., 2021)). Currently, all of ORTDM's maize farmers, including the majority of South Africa's smallholder farmers, do not have any type of formal insurance (Partridge and Wagner, 2016; Elum et al., 2017; NAMC, 2020).

6.2.2. Data

The rainfall data used in this study include (1) in-situ rainfall records from automatic weather stations (WSs), (2) TAMSAT and (3) CHIRPS data. Long-term averages of minimum and maximum temperatures, humidity, wind speed, and sunshine hours were obtained from the CLIMWAT database.

- *In-situ rainfall data*

In-situ rainfall records were obtained from the Agricultural Research Council's agro-climate databank, which receives weather data from automatic WSs situated in Tsolo, Libode, Ross Mission, Qunu and Mthatha (Table 6.1). The study focused on the critical maize growing period, which is from November to April and covered the 18 seasons between 2002 and 2019. This period was selected because there were numerous missing data in the rainfall records of the manual WSs that were operating prior to 2002. The five WS listed in Table 6.1 were selected because they are situated within the area that is suitable for growing maize. Although maize is also grown in the areas around the other three WS shown in Figure 2, these areas have high rainfall and uneven terrain and are suitable for vegetables, fruits and forestry (Nyandeni Local Municipality, 2018)

Table 6.1: Locations of the weather stations

Station name	Latitude	longitude
Tsolo	-31.2923	28.7627
Libode	-31.4481	28.9430
Ross Mission	-31.5427	28.6153
Qunu	-31.8060	28.6161
Mthatha	-31.5803	28.7754

- *TAMSAT*

TAMSAT version 3.1 data covering 2002 to 2019 were downloaded from https://www.tamsat.org.uk/. This dataset has a resolution 0.0375° and includes daily, pentadal, monthly, and seasonal rainfall products developed thermal infrared satellite data and gauge-based rainfall data by the University of Reading. The approach used to estimate TAMSAT rainfall is based on the understanding that rainfall predominantly comes from convective cumulonimbus storm clouds (Maidment et al., 2020). Precipitation occurs when these clouds reach a certain temperature. TAMSAT infers from the thermal infrared imagery. The threshold temperature and the regression coefficients that estimate rainfall from the thermal imagery are determined by relating gauge observations to Cold Cloud Duration (CCD). TAMSAT was originally developed to specifically monitor the impacts of rainfall deficits on crop yields over the Sahel but it now covers the rest of Africa.

- *CHIRPS*

CHIRPS data covering 2002 to 2019 were downloaded from https://climateserv.servirglobal.net/. CHIRPS includes daily, pentadal, and monthly precipitation observations at 0.05° spatial resolution. It is developed by the Climate Hazard Group (CHG) and the United States Geological Survey (USGS) Earth Resources Observation and Science Centre (EROS). It is a product of multiple satellite datasets and gauge-based rainfall observations (Funk et al., 2015). The data is produced in two phases in which; (1) World Meteorological Organisation's Global Telecommunication System gauge data are blended with CCD rainfall estimates at every pentad and, (2) the best available monthly and pentadal station data are combined with CCD-based rainfall estimates. CHIRPS was developed to support drought monitoring and trend analyses.

6.2.3. Climate data used for calculating CWR

Long-term averages of minimum and maximum temperatures, humidity, wind speed, and sunshine hours were obtained from the CLIMWAT database. CLIMWAT is climatic database that is used with the CROPWAT model to calculate CWR (Smith et al., 2002). Of all the WSs in ORTDM, the CLIMWAT database only has climatic data of the Mthatha WS (denoted as Umtata).

6.2.4. Data analysis

- *Preprocessing of in-situ rainfall data*

WS data were preprocessed and transformed to daily, 20-day cumulative, and monthly cumulative measurements covering November to April for each of the 18 seasons between 2002 and 2019. The 20-day and monthly time steps correspond to the lengths of the growth stages of maize as classified by the Food and Agriculture Organization (FAO) in the CROPWAT model (Smith et al., 2002). These stages are the plant's initial, development, mid-season and late growth phase.

- Preprocessing of SRFEs

A python script was used to extract time series rainfall values from CHIRPS images to comma separated values (CSV) files. TAMSAT data is provided in the form of source-compiled CSV files. Studies report that satellite-based rainfall estimates are better when spatially aggregated than in their native spatial resolutions (Black et al., 2016; Enenkel et al., 2019). Therefore, both CHIRPS and TAMSAT were spatially aggregated (Figure 6.2).

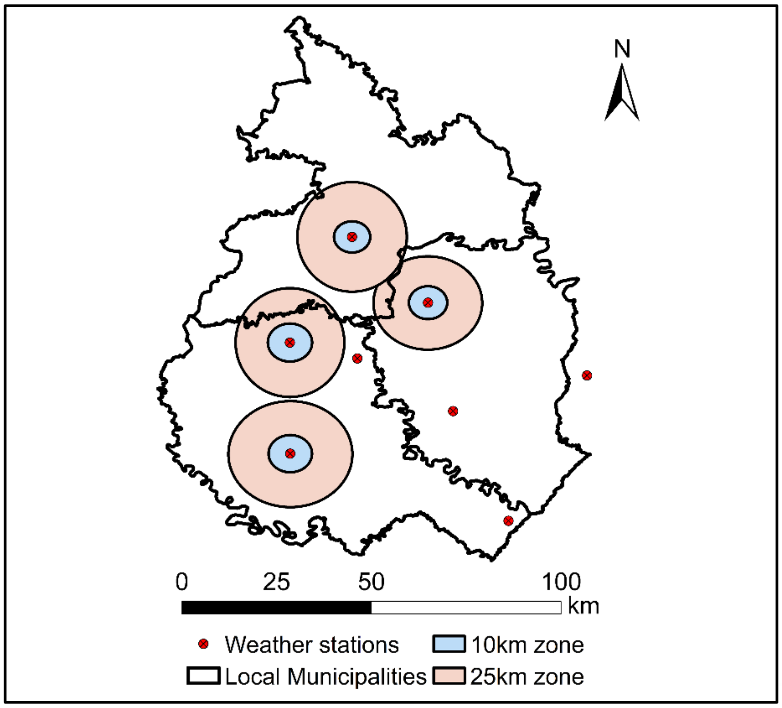


Figure 6.2: The spatial scales at which satellite data were tested

CHIRPS data were aggregated by averaging pixel values that lie within 5 km and 12.5 km from the reference WS (Figure 6.2). TAMSAT data were aggregated to the same spatial scales using the sub-setting tool available in the TAMSAT website, which allows users to download spatially averaged values within a specified boundary. Both datasets were daily, 20-day cumulative, and monthly cumulative records covering November to April for each of the 18 seasons between 2002 and 2019.

- Evaluating TAMSAT and CHIRPS against in-situ rainfall data

The SRFEs were evaluated against the WS data using Pearson's correlation. Correlation analysis was purposefully selected because it has been used to validate satellite-based rainfall products (Caroletti et al., 2019; Dinku et al., 2018; Tarnavsky et al., 2014) and to assess different satellite data in IBCI (Black et al., 2016; Enenkel et al., 2019; Enenkel, Osgood, and Powell, 2017; Eze et al., 2020; Makaudze and Miranda, 2010). Of all the WSs in ORTDM, the Mthatha WS is the only one whose data are used for calibration by the developers of CHIRPS and TAMSAT (Table 6.1). Therefore, the SRFEs were evaluated against the other four WSs excluding the Mthatha WS.

However, after evaluating and validating the SRFEs against these four WSs, we used the Mthatha WS for the next objective, which was calculate CWR and develop the IBCI payout structure.

- *Crop water requirements*

Stagewise assessments of crop water requirements (CWR) were conducted using FAO's CROPWAT 8.0 model (Smith et al., 2002). CROPWAT is computer model for calculating CWR and irrigation requirements based on soil, climate, and crop data (Muhammad, 2009; Smith et al., 2002). CROPWAT can also be used to estimate crop performance under both rain-fed and irrigated conditions. This model has been used in South Africa by several studies to calculate evapotranspiration and CWR (Worldbank, 2007; Dabrowski et al., 2009; Nyambo and Wakindiki, 2015; Singo et al., 2016). The model calculates growing season CWR or crop evapotranspiration ET_c from reference evapotranspiration (ET_0) and crop coefficients (K_c) according to equation 1.

$$\mathrm{ET_c} = \mathrm{K_c} \times \mathrm{ET_0} \tag{1}$$

The calculation of ET_0 was based on the Penman-Monteith method and long-term average minimum and maximum temperatures, humidity, wind speed, and sunshine hours. Rainfall data were extracted from CHIRPS because CHIRPS performed better than TAMSAT. CWR were calculated for each of the earlier-stated four growth stages of maize for a 120-day maturity variety. In South Africa, maize takes 120 or more days to mature (du Plessis, 2003; Frost et al., 2013; Masupha and Moeletsi, 2020). Crop coefficients K_c were based on FAO's recommendations, which are available in the maize file provided in CROPWAT's crop folder. Effective rainfall was calculated using the USDA soil conservation service method (Ali and Mubarak, 2017; Bokke and Shoro, 2020). Lastly, rainfall deficits (RD) or irrigation water requirements, which give an indication of whether CWR are met or not, were calculated for each stage. These calculations were carried out in order to (1) explore the possibility of using the CWR of maize as the trigger threshold for insurance, and (2) identify the most water critical of the four growth stages of maize by comparing their CWR and RD.

- *IBCI development and payout thresholds*

Designing an IBCI involves determination of key thresholds, which are, (1) the trigger and exit thresholds, (2) the crop growth period over which the index is measured, (3) the tick, which is the payout frequency between the trigger and the exit, and (4) the insured amount. The study adopted the approach used by Choudhury et al., (2016) as well as Eze et al., (2020) for the IBCI payout structure. However, instead of correlating crop yields and rainfall, we used CWR and long-term SRFEs to derive the trigger and exit thresholds. The payout structure is based on equation 2;

$$\text{Payout} = \begin{cases} \mathrm{IA_i}, & \text{if } \mathrm{A_i} \leq \mathrm{E_i} \\ \mathrm{IA_i}\left(\frac{\mathrm{T_i}-\mathrm{A_i}}{\mathrm{T_i}-\mathrm{E_i}}\right), & \text{if } \mathrm{E_i} < \mathrm{A_i} \leq \mathrm{T_i} \\ 0, & \text{if } \mathrm{A_i} > \mathrm{T_i} \end{cases} \tag{2}$$

where,

- IA_i is the insured amount at growth stage i, which is a portion of the costs spent on inputs (seeds, fertilizers, pesticides, herbicides, land preparations etc.),
- A_i is the actual index value over growth stage i of the insured year,
- T_i is the trigger point below which payout starts and
- E_i is the exit threshold below which full payout is given.

When A_i is below E_i, the farmer gets the entire insured amount for that growth stage. If A_i is between T_i and E_i, the scheme pays out a proportion of the amount based on $IA_i\left(\frac{T_i-A_i}{T_i-E_i}\right)$. When $A_i$is above T_i, the scheme pays nothing.

Following Chen et al., (2017), the tick was derived according to equation 3 by dividing the amount insured by the difference of the trigger and the exit.

$$\text{Tick} = \frac{\mathrm{IA_i}}{\mathrm{T_i}-\mathrm{E_i}} \tag{3}$$

The insured amount is equal to the sum of money spent on inputs and land preparation, which, according to the farmers and the DALRRD, is approximately R10 000 ($623.22 by 02 December 2021) per hectare. In an earlier study, Masiza et al., (2020) demonstrated that information on hectarage and distributions of farms can be accurately derived by optimizing machine learning and remote sensing methods. A similar study used the same techniques and demonstrated that the approach is very accurate (p < 0.05, R = 0.84) in estimating crop hectarage (Mashaba-Munghemezulu et al., 2021).

To calculate the insurable amount for each growth stage, the stages were weighted according to RD. CROPWAT reports RD as irrigation water requirements. The weightings of the growth stages were based on RD and not on CWR because the calculation of RD takes into account both CWR and the actual rainfall received by the crop (i.e. RD = CWR – Effective Rainfall). The weight of growth stage i, therefore, is given by the RD of growth stage i divided by the total RD of the four stages (equation 4).

$$\mathrm{Weight_i} = \frac{\mathrm{RD_i}}{\sum_1^n \mathrm{RD}} \tag{4}$$

The insurable amount for each stage is then calculated according to equation 5.

$$\mathrm{IA_i} = \mathrm{Weight_i} \times \mathrm{R10\,000} \tag{5}$$

6.3. Results

The results are presented under three sub-sections (subsections 6.3.1 – 6.3.2). The first subsection (subsection 6.3.1) presents composite graphs that show correlations between satellite and WS data (Figure 6.3). The second subsection (subsection 6.3.2) presents a set of tables (Tables 6.1 to 6.4) that show the CWR of maize planted on 21 November, 1 December, 11 December and 21 December respectively. The last subsection (subsection 6.3.3) presents a table (Table 6.5) that shows the final insurance payout thresholds.

6.3.1. Correlations between satellite and WS data at different spatial and temporal scales

Figure 6.3 shows how WS and SRFEs (TS = TAMSAT, CH = CHIRPS) correlated at different spatial and temporal scales.

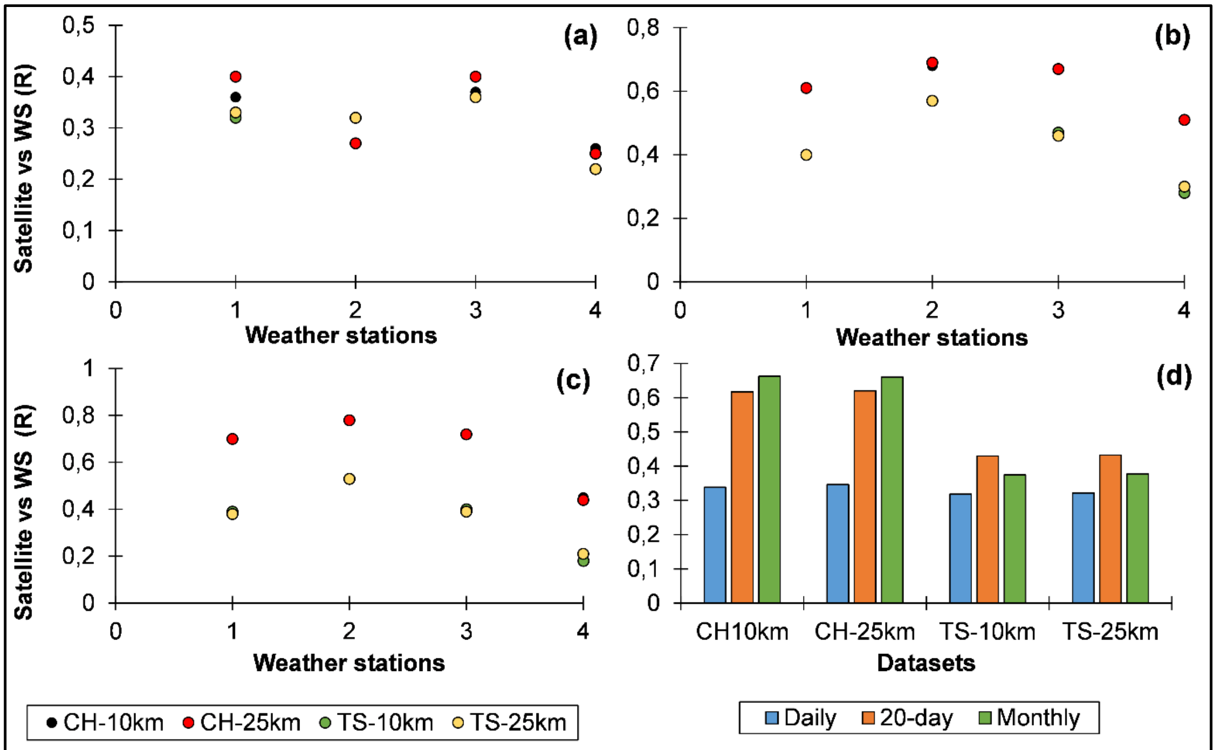


Figure 6.3: Correlations of WS and satellite data at daily (a), 20-day (b), and monthly (c) intervals. Mean correlations between WS and satellite data (d).

All correlations were statistically significant at the 95% confidence level with all p-values less than 0.05 except for the lowest correlations ($R \leq 0.21$) between TAMSAT and WS data in Figure 6.3c. Correlations between daily CHIRPS and WS data were weak ($R \leq 0.40$) and slightly weaker between daily TAMSAT and WS data ($R \leq 0.37$) (Figure 6.3a). However, the results improved at the 20-day time step, with CHIRPS and WS data correlating by 0.51 to 0.68 at the 10 km scale and by 0.51 to 0.69 at the 25 km scale. Correlations between TAMSAT and WS data were lower with a maximum correlation of 0.57 (Figure 6.3b). At the monthly time step, CHIRPS agreed strongly with WS data by achieving maximum correlations of 0.78 at both spatial scales, while TAMSAT reached a moderate maximum of 0.53 (Figure 6.3c). Figure 6.4d summarizes Figures 6.4a, 6.4b and 6.4c by averaging the correlations for the four weather stations. Overall, CHIRPS outperformed TAMSAT at all the spatial and temporal scales and performed best at the monthly time step.

6.3.2. Crop water requirements based on different planting dates

Tables 6.2, 6.3, 6.4, and 6.5 show results of CWR assessment for maize planted on different dates. The metric unit for rainfall is millimeters (mm). Across all the four planting dates, the initial growth stage has the lowest CWR and the mid-season stage has the highest (Tables 6.2 to 6.5). The late-season and development stages have the second-lowest and second-highest CWR, respectively. For all the four planting dates, the initial stage gets enough effective rainfall and no RD. When planting is on 21 November or 01 December, the late-season period also has no RD, whereas planting on 11 December or 21 December results in late-season RD. The mid-season period has the highest RD followed by the development stage. The initial stage has the lowest rainfall across all planting dates while the mid-season period has the highest. Planting on 21 November or 01 December results in the late-season period receiving the second-highest rainfall, whereas planting on 11 December or 21 December results in the development stage receiving the second-highest rainfall

Table 6.2: CWR when planting is on 21 November.

Stage	No. of days	ET_c / CWR	Effective rainfall	RD	Mean rainfall	Mean minimum rainfall
Initial	20	25.20	50.40	0.00	67.07	37.79
Development	30	102.30	83.70	18.60	90.22	63.92
Mid-season	40	201.00	115.10	85.90	141.98	109.63
Late season	30	73.50	86.40	0.00	94.58	26.84
Total	**120**	**402.00**	**335.60**	**104.50**	**393.82**	**238.18**

Table 6.3: CWR of when planting is on 01 December

Stage	No. of days	ET_c / CWR	Effective rainfall	RD	Mean rainfall	Mean minimum rainfall
Initial	20	25.20	52.80	0.00	62.23	47.76
Development	30	103.20	86.50	16.70	92.03	66.59
Mid-season	40	184.20	113.10	71.10	146.74	108.17
Late season	30	78.20	81.30	0.00	96.45	75.74
Total	**120**	**390.80**	**333.70**	**87.80**	**397.45**	**298.26**

Table 6.4: CWR when planting is on 11 December

Stage	No. of days	ET_c / CWR	Effective rainfall	RD	Mean rainfall	Mean minimum rainfall
Initial	20	26.80	54.80	0.00	54.51	26.60
Development	30	105.30	87.90	17.40	100.62	84.21
Mid-season	40	170.70	112.90	57.80	137.10	109.03
Late season	30	74.20	72.10	2.10	87.88	25.93
Total	**120**	**377.00**	**327.70**	**77.30**	**380.11**	**245.77**

Table 6.5: CWR when planting is on 21 December

Stage	No. of days	ET_c / CWR	Effective rainfall	RD	Mean rainfall	Mean minimum rainfall
Initial	20	27.20	56.70	0.00	65.22	26.53
Development	30	102.40	87.30	15.10	96.27	52.40
Mid-season	40	163.00	114.20	48.80	140.87	73.55
Late season	30	70.30	58.20	12.10	67.52	24.94
Total	**120**	**362.90**	**316.40**	**76.00**	**369.88**	**177.42**

6.3.3. Index threshold values

The index thresholds presented in Table 6.6 are based on the results presented in section 3.2 and on equations 2 to 5. The thresholds are based on the optimum planting date, which was 21 December. The optimum planting date may change from year to year; the optimum planting date in this case is based on long-term data. The selection of this date was based on the following considerations:

1. First, we observed that total seasonal and mid-season CWR decrease with delayed planting. In other words, planting on 21 December results in less CWR than planting on 11 December, 01 December and 21 November;
2. Second, total seasonal and mid-season RD also decrease with delayed planting. In other words, planting on 21 December results in less RD than planting earlier;
3. Third, the mid-season stage is given more weight because it has the highest CWR and RD, and it is the most sensitive growth stage;

4. Fourth, a planting date that evenly and proportionately distributes CWR and RD across multiple stages is less risky
5. Fifth, the farmers' experiences and historical planting dates were considered.

Based on these factors, 21 December was selected as the optimum planting date because it meets the first two of the considerations listed above. 21 December is also the date on which the most important stage (i.e. the mid-season) stage has the lowest RD. Lastly, 21 December meets the fourth consideration because it evenly and proportionately distributes CWR and RD across all the growth stages. This date was then used to develop the proposed payout structure. However, instead of using CWR, the study used long-term mean-total and mean-minimum rainfall as trigger and exit points (Table 5). The choice of these thresholds is explained in the discussion section. The most crucial growth stages for a planting date of 21 December are the development, mid-season and late-season stages. Since the initial growth stage has enough effective rainfall and zero RD (Tables 1 to 4), this stage has zero weight. The development stage has a weight of 20%, the mid-season stage has a weight of 64% and the late-season stage has a weight of 16%. For each millimeter of rainfall below the trigger point, amounts of R45.59, R95.07, and R37.58 are paid out for the development, mid-season and late-season stages, respectively (Table 6.6).

Table 6.6: Index threshold values for maize planted on 21 December

Growth stages	Trigger (mm)	Exit(mm)	Tick (R/mm)	Weight	Amount (R)
Development	96.27	52.40	45.59	0.20	2000
Mid-season	140.87	73.55	95.07	0.64	6400
Late-season	67.52	24.94	37.58	0.16	1600

6.4. Discussions and conclusions

6.4.1. Satellite data for IBCI design

The first objective of this study was to assess the performances of CHIRPS and TAMSAT datasets in estimating rainfall at different spatial and temporal scales in ORTDM. These datasets performed better at 20-day and monthly time-steps, with CHIRPS consistently performing better than

TAMSAT. The superiority of CHIRPS over TAMSAT is also reported by other studies (Dinku et al., 2018; Kimani et al., 2017; Tarnavsky et al., 2018). The different spatial scales at which the analyses were carried out did not influence the performances of these datasets. This means that, CHIRPS data, aggregated to any spatial scale between 10 km × 10 km and 25 km × 25 km, can fairly represent the local rainfall conditions in ORTDM. This finding is supported by DuPlessis and Kibii, (2021) who found that monthly CHIRPS data estimated rainfall very well in 46 locations across South Africa, with an average R^2 of 0.60. In the Eastern Cape, Mahlalela et al., (2020) also found significant correlations ranging between 0.61 and 0.90 between CHIRPS and WS data . However, these findings must be interpreted with caution as SRFEs' ability to estimate rainfall depends on the climate and physiography of the area. It is reported that SRFEs tend to misestimate rainfall in coastal and mountainous areas (Dinku et al., 2018; Gebremicael et al., 2017; Kimani et al., 2017; Le Coz and Van De Giesen, 2020). Therefore, since insurance requires reliable and good quality data to minimize basis risk, SRFEs must be validated and compared with in-situ rainfall and other rainfall-related datasets (Enenkel et al., 2017, 2018, 2019). Nevertheless, since CHIRPS performed well in ORTDM, this dataset can be used in conjunction with agro-ecological information to demarcate unit areas of insurance. In addition, since CHIRPS performed well at 20-day and monthly time-steps, this dataset can also be used to develop indices for the different growth stages of maize, which range in length between 20 and 40 days. The study, therefore, proceeded to use CHIRPS in the assessment of CWR and the design of the IBCI payout structure.

6.4.2. Maize water requirements and IBCI design

It is evident from the results that, regardless of the planting date, the average seasonal CWR exceed seasonal rainfall, resulting in RD. These RD are minimal when planting is on 21 December. Although RD pose a challenge to rain-fed maize production, the initial stage, which corresponds to germination, emergence, and the early vegetative stages, receives sufficient rainfall to meet the CWR. This scenario implies that the initial stage is less prone to water deficits and has a weight of zero. It also validates that 21 November to 21 December is a suitable planting window for ORTDM. Therefore, the index is not measured over the initial stage. Although IBCI programs such as the R4 rural resilience initiative, ACRE Africa and others cover germination failure (Arce, 2016; WFP, 2020), RD-induced germination failure is unlikely to occur in ORTDM if maize is planted between 21 November and 21 December. However, since germination failure is compensated for by replanting, farmers can still be indemnified with the amount of money needed for replanting when RD-induced germination failure occurs.

The development and mid-season stages' RD decrease with delayed planting, whereas the late-season stage's RD increase with delayed planting. Remember, rainfall deficit is the difference between CWR and rainfall. If RD are likely to occur every year as the results show, insurance cannot compensate for these RD. This means that CWR cannot be used as the trigger threshold as this study initially intended. However, studies report that the weather conditions of ORTDM can produce good yields when inputs are managed properly (Masiza et al., 2021; Sibanda et al., 2016; Sibanda et al., 2016)]. This is supported by Osgood et al.,(2007) who observed that crops that are adapted to the prevalent local climate conditions have low vulnerability during dry spells. They (Osgood et al., 2007) also point out that the parameters of CWR models are not a definite predictor of crop behavior. It is for these reasons that long-term mean and mean-minimum rainfall were used as trigger and exit thresholds, respectively. Therefore, the proposed system uses long-term mean and mean-minimum rainfall as thresholds; the rainfall data is derived from CHIRPS; the system measures the index over the development, mid-season, and late-season stages and it operates on a linear payout structure as shown in Figure 6.4.

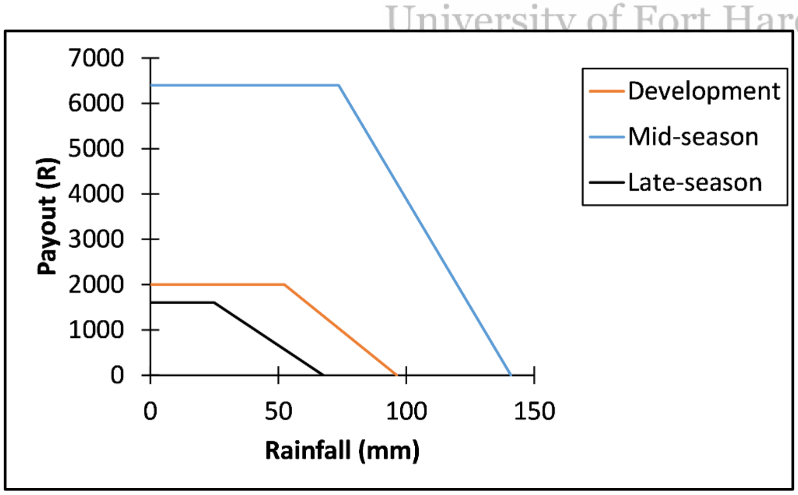


Figure 6.4: Development stage (trigger = 96.27mm, exit = 52.40mm), mid-season (trigger = 140.87mm, exit = 73.55mm) late-season (trigger = 67.52mm, exit = 24.94mm)

The model gives more weight to the mid-season stage, which corresponds to the late vegetative, tasseling, silking, pollination, and blister stages. It is over these growth periods that water and nutrient demands are high (du Plessis, 2003; Masupha and Moeletsi, 2017; Udom and Kamalu, 2019). The late-season stage has the lowest weight because it has lower CWR than the development and mid-season stages. Measuring an insurance index over critical crop growth stages, as demonstrated in this study, rather than measuring it once at the end of the crop-growing season, potentially reduces temporal basis risk. By taking into account the findings of chapters 3, this proposed IBCI system must also map cropped area from satellite data rather than relying on farmer reports and on-field surveys. The findings of chapter 3 showed that crop area mapping can be achieved with the fusion of SAR data and optical data and machine learning ensembles. Furthermore, a similar study used these mapping techniques and demonstrated that they are very accurate ($p < 0.05$, R = 0.84) in estimating the amount of cropped area (Mashaba-Munghemezulu et al., 2021). Incorporating crop area mapping into IBCI would enable insurers to locate, verify, and quantify cropped areas.

This proposed IBCI must also be linked or bundled with input supplies and advisory services to address non-weather factors that influence crop losses. This would enable insurers to isolate and accurately quantify the impact of weather on crop losses, which could reduce basis risk. Key points that differentiate the approach used in this study from other approaches are (1) the incorporation of crop area mapping (2) the linking or bundling of IBCI with input supplies and advisories to address non-weather factors that influence crop losses, and (3) the use of CWR and site-specific rainfall conditions to derive index thresholds. This approach objectively derives IBCI thresholds from CWR instead of calculating the thresholds by using empirical models that simply relate yields and rainfall. The widely used approach, which calibrates IBCI by correlating yields and rainfall exposes contracts to basis risk because, by simply correlating yield and rainfall data, it overlooks the influence of non-weather factors on smallholder crop yields and losses (Masiza et al., 2022). However, the approach used in this study is informed by the fact that, in most cases, rainfall-related yield reduction results from rainfall deficit, which is the failure of rainfall to meet CWR. Although the approach proposed in this study is premised on the prevailing conditions in ORTDM and South Africa, it can be applied in other areas with different growing conditions. More research can improve this approach by investigating the relationship between RD and yield losses under controlled and well-managed non-weather factors.

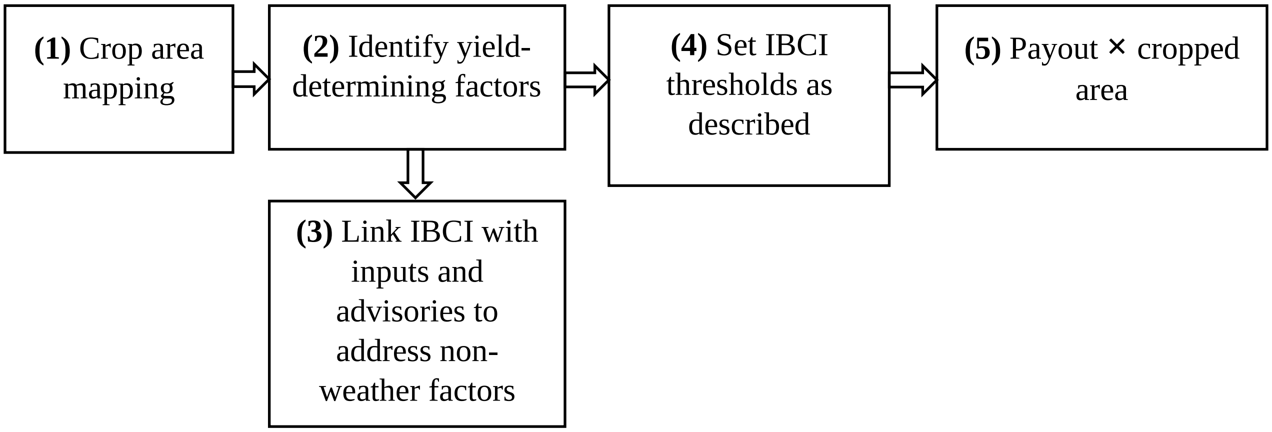


Figure 6.5: Flow chart of the proposed IBCI system

Chapter 7 : Synthesis and conclusions

7.1. Introduction

The unpredictable weather conditions of Sub-Saharan Africa threaten agriculture, especially smallholder crop farming. Since agriculture is the biggest economic sector in SSA, the inability of farmers to adopt improved and adapted agricultural technologies is a setback to the region's economic development and the fight against poverty. The need for comprehensive risk management and risk reduction strategies cannot be overstated. Research shows that among the various initiatives that have been implemented to address production risks, IBCI promises to address some of the challenges confronting crop farmers. This progress is illustrated by the implementation of IBCI programs, which have encouraged farmers to take risks and invest in improved technologies and management practices.

Although IBCI provides numerous benefits, its implementation is confronted by numerous challenges that require innovative solutions. One of these challenges is low-uptake, which is often associated with poor design and poor quality of contracts. When some of these design-related challenges were pointed-out in the earliest IBCI programs, many reports recommended the use of satellite remote sensing. This study investigated the feasibility of using remote sensing in IBCI by reviewing what remote sensing has contributed to IBCI, what its limitations have been, and what opportunities it still needs to explore. The findings showed that non-weather factors that are not measurable with remote sensing also have influence on crop yields and crop losses. Furthermore, the use of remote sensing has been focused more on the monitoring of vegetation health, the prediction of crop yields, and the estimation of rainfall, and less on the mapping of crop area and soil moisture. The study then attempted to address some of the gaps through empirical experimentation. The following sub-sections provide a summary and synthesis of the findings of this undertaking. The chapter concludes by outlining the limitations of the study and giving recommendations for future research.

7.2. Chapter 2: Progress, gaps and opportunities

Chapter 2 was a literature review, which sought to provide a comprehensive overview of what remote sensing has contributed to Africa's IBCI, what it has not done, and what it needs to do. Relevant information was collected from peer-reviewed articles and other published reports. Results showed that researchers have investigated the applicability of remote sensing in IBCI and insurers have been and are still using remote sensing datasets for rainfall, vegetation, soil moisture, and evapotranspiration mapping. It also revealed that the use of remote sensing data has improved delineation of unit areas of insurance and enabled coverage of inaccessible farms and areas that do not have weather stations. However, the review showed that remote sensing-based IBCI is tainted by basis risk that is associated with poor product design, non-weather yield-determining factors, imperfection correlations between satellite-based indices and crop yield data, and the lack of historical data for contract calibration.

IBCI schemes need to survey specific localities for major yield-limiting factors because yield variability in smallholder-farming areas is a product of complex interplay between different environmental, biological, managerial, and socioeconomic factors. Although some of the IBCI projects have embarked on large-scale field campaigns to collect yield data for calibration, these efforts have not yet been able to deliver the desired results. This shortfall requires government extension services to contribute by collecting reliable and accurate data because they are tasked with collecting information from farms and providing advice to farmers. With regards to remote sensing, opportunities to provide viable IBCI still exist in various untapped areas that include data fusion, high-resolution mapping, and microwave remote sensing. On the other hand, data enhancement and data analysis techniques such as machine / deep learning and computer vision have been under-used.

7.3. Chapter 3: Mapping of a smallholder crop farming area

Chapter 3 was conceived because the literature review showed that although plant parameters such crop yield and vegetation vigour were often mentioned in the IBCI literature, publications on crop mapping are few and sketchy. The geographic distributions of farms and the amount of cultivated area is key information that enables insurers to locate and map crop fields, which is critical for the verification and quantification of crop losses. Multiple machine learning classifiers and a combination of optical and SAR datasets were used to enhance the mapping of smallholder-farming areas. Maize, which was the crop type of interest in this thesis, was mapped with very high levels of accuracy reaching 100%. A similar study which used the same mapping approach demonstrated that these methods are very accurate ($p < 0.05$, R = 0.84) in estimating the amount of cropped area (Mashaba-Munghemezulu et al., 2021). This demonstrates the robustness of machine learning and remote sensing, and the ability of these techniques to provide objective information rather than relying on farmers' reports and costly on-field surveys that are less capable of providing practically usable information.

7.4. Chapter 4: Factors that influence crop yields

The review of literature in chapter 2 also showed that some of the factors that make it difficult to model the relationship between weather indices and crop losses are non-weather inputs. As a result, chapter 4 employed a machine learning technique to identify factors that influence maize yield in the study area. The findings demonstrated why IBCI needs to be part of a comprehensive system that understands and approaches smallholder crop farming as complex and multifaceted. The consistently low yields of smallholder farming systems show that rainfall is not the only limiting factor in the same way that it may not always be the major limiting factor. The influence of agrometeorological variables such as rainfall and growing degree-days were largely associated with planning and planting dates. Inputs such as fertilizer, seeds, soil conditions, management, mechanization and socioeconomic conditions were equally important. Linking or bundling IBCI with advisories and input supplies could address the impacts of non-weather factors on yield and also reduce basis risk by enhancing remote sensing's ability to quantify and isolate the impact of weather on crop losses.

7.5. Chapter 5: Estimating soil water content from satellite data

Chapters 2 and 4 showed that soil moisture stress is one of the most important yield-determining factors. However, the currently available and frequently used satellite-based products of soil moisture have coarse spatial resolutions and limited applicability in small-scale farms. Moreover, the satellite-derived moisture index used in chapter 4 was derived from optical data with low temporal frequency. Chapter 5 then attempted to address these limitations by investigating the feasibility of estimating soil moisture from a combination of high-resolution optical and SAR datasets. The results of this improvised combination showed that the estimation accuracies of soil water content were comparable to those in studies that had used similar datasets but different methods. However, the modelling approach used in this study can be further improved by investigating interaction effects between NDVI, NDII, and the SAR bands. Improvements can also be achieved by either taking into account factors such as surface roughness and incidence angle or by using methods that reduce their effect on backscatter. Additional research is also needed to explore how microwave-based soil moisture estimates and optical datasets like MODIS, Landsat, Sentinel-3 and others can be combined to estimate soil moisture content at high-spatial resolutions and high temporal frequencies that will be suitable for IBCI.

7.6. Chapter 6: A proposed IBCI model for South Africa's smallholder farmers

The review of literature demonstrated that the contribution of remote sensing in IBCI lies in its ability to reduce basis risk and to improve contract design by accurately modelling crop losses that are due to weather. Chapters 2 and 4 also revealed that studies have achieved mixed and unconvincing results from the yield-rainfall empirical models because of the influence of the non-weather factors on crop yields. Chapter 6, therefore, proposed an IBCI system that derives index thresholds from site-specific crop water requirements of maize. Rainfall observations were taken from CHIRPS, which performed better than TAMSAT in estimating rainfall. This approach was based on the understanding that rainfall-related crop losses result from rainfall's failure to meet crop water requirements. The proposed IBCI measures rainfall over the vegetative, mid-season, and late-season stages of maize, which are 30, 40, and 30 days long, respectively. It assigns crop water requirement-derived weights to these stages, with the mid-season and late-season stages being assigned the highest and lowest weights, respectively.

The findings of these experiments demonstrated that the weightings assigned to these respective growth stages are consistent with the actual water requirements of maize. This chapter also emphasized that crop area mapping must be incorporated in IBCI because crop maps provide important spatial information that insurers need when determining insurance payouts. Furthermore, the IBCI system must be linked or bundled with advisory services and input supplies to address non-weather inputs that influence crop yields and crop losses. IBCI programs and other stakeholders involved in smallholder farming need to investigate factors that influence crop yields for specific localities. Helping farmers to manage production risks in this holistic manner could also enhance remote sensing's ability to accurately quantify and isolate the influence of weather on crop losses.

Key contributions that differentiate the approach used in this study from other approaches are (1) the incorporation of crop area mapping which enables insurers to locate, verify, and quantify cropped areas, (2) the use of CWR and site-specific rainfall conditions to derive index thresholds, and (3) the linking of IBCI with input supplies and advisories to address non-weather yield-determining factors. The approach objectively derives IBCI thresholds from CWR and site-specific rainfall conditions instead of calculating the thresholds by using empirical models that simply relate yields and rainfall. Overall, this is a more holistic approach that is potentially capable of improving the design and uptake of IBCI.

7.6. Limitations and recommendations for future research

The main limitation of this study was the lack of reliable long-term historical yield records. As a result, the investigations were limited to one district with cross-sectional data collected over three growing seasons. The study also relied much on the IBCI literature and experiences of other countries because South Africa does not have IBCI schemes yet. Future studies will have to include other smallholder farming areas in other districts and provinces. Cooperation between researchers, extension services, input suppliers, and farmers could improve the yield records of smallholder farming systems. Research institutions such as South Africa's Dohne Agricultural Development Institute, the Agricultural Research Council (ARC), the universities and others, are capacitated with research and technical expertise. Through these institutions, the Crop Estimates Committee has been collecting yield and crop hectarage statistics from large-scale farms for years.

The survey techniques, like the objective yield survey (OYS) method and other skills and techniques, should be transferred to the government extension officers who work with smallholder farmers in specific localities. For example, data quality and record keeping can be improved by taking advantage of the smart mobile phones that the government provides to extension officers and other agricultural officials. The research institutions have the expertise to develop smart mobile applications that can be used for data collection. The ARC for example has already developed a couple of smart mobile applications for agriculture-related inquiries. A mobile application for collecting data could feature user login details and be linked to a GPS system (similar to ESRI's Survey 123) for accountability and data quality assurance. For reliable record keeping, the application could be linked to an online database similar to the ARC's agro-climate databank. Input suppliers could contribute by regularly soliciting feedbacks about inputs and yields from farmers. Co-operation and information sharing between insurers, research institutes, governments, input suppliers and other stakeholders, could reduce unnecessary field-surveys and costs.

Since management and agronomic factors proved to have a significant influence on yield, this may have imposed limitations on remote sensing's ability to isolate the effect of rainfall on crop yields.

Non-weather factors such as seeds, fertilizers, soil conditions and others, could be influencing crop yields in other smallholder farming areas as well. Research in other localities could provide more insight on some of the most important factors that need to be addressed to minimise production risk and basis risk. The fact that surface moisture stress was one of the critical yield-determining factors requires researchers to determine how to combine rainfall and soil moisture data and how to improve the retrieval of soil moisture from satellites for a more robust weather index. This study attempted to measure soil moisture using a combination of Sentinel-1 and Sentinel-2 data. However, improvements can be made by incorporating other high-resolution sensors and by further exploring interaction effects and the influence of vegetation on the datasets used in this study for soil moisture estimation. Furthermore, improvements can also be made by minimizing the influence of surface roughness on microwave backscatter.

The highest spatial resolution at which the CHIRPS dataset was investigated was 10 km. There may be geographic variations of rainfall at higher spatial scales that were not captured at the 10 km resolution. Future research could investigate the sensitivity of CHIRPS to rainfall at spatial resolutions higher than 10 km. More research is also needed in other areas to assess actual rainfall patterns against crop water requirements. This may be supplemented with crop suitability analyses because public reports show that some of the areas in ORTDM that are designated for vegetables, fruits, and forestry are instead growing maize (Nyandeni Local Municipality, 2018). These are areas with very high rainfalls, cloud cover, and uneven terrains. This is important because insurers would want to cover areas that grow crops that are for those areas.

References

Abdikan, S., Sanli, M., Ustuner, M., and Calo, F. (2016). Land cover mapping using Sentinel-1 SAR data. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 757–761. https://doi.org/DOI: 10.5194/isprs-archives-XLI-B7-757-2016

Abdulai, I., Hoffmann, M. P., Jassogne, L., Asare, R., Graefe, S., Tao, H. H., Muilerman, S., Vaast, P., Van Asten, P., Läderach, P., and Rötter, R. P. (2020). Variations in yield gaps of smallholder cocoa systems and the main determining factors along a climate gradient in Ghana. *Agricultural Systems*, *181*(October 2019), 102812. https://doi.org/10.1016/j.agsy.2020.102812

Abubakar, G. A., Wang, K., Shahtahamssebi, A., Xue, X., Belete, M., Gudo, A. J. A., Shuka, K. A. M., and Gan, M. (2020). Mapping Maize Fields by Using Multi-Temporal Sentinel-1A and Sentinel-2A Images in Makarfi, Northern Nigeria, Africa. *Sustainability (Switzerland)*, *12*(6), 1–18. https://doi.org/10.3390/su12062539

ACRE. (2020). *3-D client value assessment for ACRE Rwanda maize and livestock insurance products*. https://doi.org/Available online:https://acreafrica.com/ (accessed on 26 May 2021)

Adab, H., Morbidelli, R., Saltalippi, C., Moradian, M., and Ghalhari, G. A. F. (2020). Machine learning to estimate surface soil moisture from remote sensing data. *Water (Switzerland)*, *12*(11), 1–28. https://doi.org/10.3390/w12113223

Adeoti, S. O., Moses, A., and Adedamola, A. M. (2020). Bundling of insurance with agricultural inputs : a better way of introducing insurance product to farmers. *AshEse Journal of Agricultural Science*, *3*(2), 107–113. https://doi.org/Available online: http://ashese.co.uk/agricultural-science1/blog (accessed on 11 December 2020)

Adhikari, U., Nejadhashemi, A. P., and Woznicki, S. A. (2015). Climate change and eastern Africa: A review of impact on major crops. *Food and Energy Security*, *4*(2), 110–132. https://doi.org/10.1002/fes3.61

Adrian, J., Sagan, V., and Maimaitijiang, M. (2021). Sentinel SAR-optical fusion for crop type mapping using deep learning and Google Earth Engine. *ISPRS Journal of Photogrammetry and Remote Sensing*, *175*(February), 215–235. https://doi.org/10.1016/j.isprsjprs.2021.02.018

African Risk Capacity. (2017). *African Risk Capacity: Sovereign disaster risk solutions*. African Risk Capacity. https://doi.org/Available online: https://www.africanriskcapacity.org/wp-content/uploads/2017/07/ARC-and-ARC-LTD-EN.pdf (Accessed on 29 January 2021)

African Risk Capacity. (2019). *Africa RiskView software suite: A collection of software tools for monitoring natural hazards and preparing for effective disaster response*. African Risk Capacity. https://doi.org/Available online: https://www.africanriskcapacity.org/wp-content/uploads/2019/05/Africa-RiskView-Research-brochure-April-2019.pdf (Accessed on 5 April 2020)

African Smallholder Farmers Group (ASFG). (2013). *Supporting smallholder farmers in Africa: A framework for an enabling environment*. https://doi.org/Available online: http://www.asfg.org.uk/framework-report/overview (Accessed on 18 March 2020)

Agarwal, A. (1979). Crop insurance : Need for a new approach. *Economic and Political Weekly*, *14*(39), 39–42. https://doi.org/https://www.jstor.org/stable/4364738

Aguilar, R., Zurita-Milla, R., Izquierdo-Verdiguier, E., and de By, R. A. (2018). A cloud-based multi-temporal ensemble classifier to map smallholder farming systems. *Remote Sensing*, *10*(5). https://doi.org/10.3390/rs10050729

Ahmad, L., Habib Kanth, R., Parvaze, S., Sheraz Mahdi, S., Ahmad, L., Habib Kanth, R., Parvaze, S., and Sheraz Mahdi, S. (2017). Growing Degree Days to forecast crop stages. In L. Ahmad, R. Habib Kanth, S. Parvaze, S. Sheraz Mahdi, L. Ahmad, R. Habib Kanth, S. Parvaze, and S. Sheraz Mahdi (Eds.), *Experimental Agrometeorology: A Practical Manual* (2017 Editi, pp. 95–98). Springer International Publishing. https://doi.org/10.1007/978-3-319-69185-5_14

Akinnagbe, O. M., and Irohibe, I. J. (2014). Agricultural adaptation strategies to climate change impacts in africa: a review a. *Bangladesh Journal of Agriculture*, *39*(3), 407–418. https://doi.org/https://doi.org/10.3329/bjar.v39i3.21984

Akpalu, D. (2013). Agriculture extension service delivery in a semi-arid rural area in South Africa: The case study of Thorndale in the Limpopo Province. *African Journal of Food, Agriculture, Nutrition and Development*, *13*(4), 8058–8076. https://doi.org/http://hdl.handle.net/1807/56061

Alcaide, K. D. R., Buncag, M. J. J., Mendoza, R. E. S., and Santos, L. B. U. (2019). Developing a Rainfall-Based Index for Corn Crop Insurance in Isabela , Philippines. *International Journal of Science and Management Studies*, *2*(6), 77–84. https://doi.org/DOI: 10.51386/25815946/ijsms-v2i6p109

Alderman, H., and Haque, T. (2007). Insurance against covariate shocks. In H. Alderman and T. Haque (Eds.), *World Bank* (1st ed.). The World Bank. https://doi.org/10.1596/978-0-8213-7036-0

Aldrich, C. (2020). Process variable importance analysis by use of random forests in a shapley regression framework. *Minerals*, *10*(5), 1–17. https://doi.org/10.3390/min10050420

Alem, Y., Bezabih, M., Kassie, M., and Zikhali, P. (2010). Does fertilizer use respond to rainfall variability ? Panel data evidence from Ethiopia. *Agricultural Economics*, *41*, 165–175. https://doi.org/10.1111/j.1574-0862.2009.00436.x

Ali, M., and Mubarak, S. (2017). Effective rainfall calculation methods for field crops: An overview, analysis and new formulation. *Asian Research Journal of Agriculture*, *7*(1), 1–12. https://doi.org/10.9734/arja/2017/36812

Aliber, M., and Hart, T. G. B. (2009). Should subsistence agriculture be supported as a strategy to address rural food insecurity? *Agrekon*, *48*(4), 434–458. https://doi.org/10.1080/03031853.2009.9523835

Anderson, M., Kustas, W. P., Norman, J. M., Hain, C. R., Mecikalski, J. R., Schultz, L., González-Dugo, M. P., Cammalleri, C., D’Urso, G., Pimstein, A., and Gao, F. (2011). Mapping daily evapotranspiration at field to continental scales using geostationary and polar orbiting satellite imagery. *Hydrology and Earth System Sciences*, *15*(1), 223–239. https://doi.org/10.5194/hess-15-223-2011

Anderson, M., Norman, J. M., Diak, G. R., Kustas, W. P., and Mecikalski, J. R. (1997). A two-source time-integrated model for estimating surface fluxes using thermal infrared remote sensing. *Remote Sensing of Environment*, *60*(2), 195–216. https://doi.org/10.1016/S0034-4257(96)00215-5

ARC. (2020). *Africa RiskView: End of season report | Malawi (2020/21 season).*

Arce, C. (2016). *Comparative assessment of weather index insurance strategies in Sub-Saharan Africa*. https://doi.org/Available online: http://www.vuna-africa.com (Accessed on 19 August 2019)

Archer, E. (2020). Package ‘ rfPermute .’ *The R Journal*. https://doi.org/Available online: https://cran.r-project.org/web/packages/rfPermute/index.html (Accessed on 18 October 2020)

Assefa, B. T., Chamberlin, J., Reidsma, P., Silva, J. V., and van Ittersum, M. K. (2020). Unravelling the variability and causes of smallholder maize yield gaps in Ethiopia. *Food Security*, *12*(2), 489–490. https://doi.org/10.1007/s12571-019-00998-9

Awondo, S., Kostandini, G., Setimela, P., and Erenstein, O. (2020). Multi-site bundling of drought tolerant maize varieties and index insurance. *Journal of Agricultural Economics*, *71*(1), 239–259. https://doi.org/10.1111/1477-9552.12344

Awondo, S. N. (2019). Efficiency of region-wide catastrophic weather risk pools: Implications for African Risk Capacity insurance program. *Journal of Development Economics*, *136*(September 2018), 111–118. https://doi.org/https://doi.org/10.1016/j.jdeveco.2018.10.004

Ayanlade, A., and Radeny, M. (2020). COVID-19 and food security in Sub-Saharan Africa: implications of lockdown during agricultural planting seasons. *Npj Science of Food*, *4*(1), 1–6. https://doi.org/10.1038/s41538-020-00073-0

Ayehu, G., Tadesse, T., Gessesse, B., Yigrem, Y., and M.melesse, A. (2020). Combined use of sentinel-1 sar and landsat sensors products for residual soil moisture retrieval over agricultural fields in the upper blue nile basin, ethiopia. *Sensors (Switzerland)*, *20*(11), 1–25. https://doi.org/10.3390/s20113282

Bai, J., Cui, Q., Zhang, W., and Meng, L. (2019). An approach for downscaling SMAP soil moisture by combining Sentinel-1 SAR and MODIS data. *Remote Sensing*, *11*(23), 1–20. https://doi.org/10.3390/rs11232736

Ban, Y. (2016). Multitemporal remote sensing: methods and applications. In *Springer* (Issue November 2017). https://doi.org/10.1007/978-3-319-47037-5

Banerjee, H., Goswami, R., Chakraborty, S., Dutta, S., Majumdar, K., Satyanarayana, T., Jat, M. L., and Zingore, S. (2014). Understanding biophysical and socio-economic determinants of maize (Zea mays L.) yield variability in eastern India. *NJAS - Wageningen Journal of Life Sciences*, *70*, 79–93. https://doi.org/10.1016/j.njas.2014.08.001

Barnett, B. J., and Mahul, O. (2007). Weather index insurance for agriculture and rural areas in lower-income countries. *American Journal of Agricultural Economics*, *89*(5), 1241–1247. https://doi.org/10.1111/j.1467-8276.2007.01091.x

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., and Wood, E. F. (2018). Present and future köppen-geiger climate classification maps at 1-km resolution. *Scientific Data*, *5*, 1–12. https://doi.org/10.1038/sdata.2018.214

Belissa, T., Bulte, E., Cecchi, F., Gangopadhyay, S., and Lensink, R. (2019). Liquidity constraints, informal institutions, and the adoption of weather insurance: A randomized controlled trial in Ethiopia. *Journal of Development Economics*, *140*(September 2018), 269–278. https://doi.org/10.1016/j.jdeveco.2019.06.006

Belissa, T. K., Lensink, R., and van Asseldonk, M. (2019). Risk and ambiguity aversion behavior in index-based insurance uptake decisions: experimental evidence from Ethiopia. *Journal of Economic Behavior and Organization*. https://doi.org/10.1016/j.jebo.2019.07.018

Bembridge, T. J. (2000). Guidelines for rehabilitation of small-scale farmer irrigation schemes in South Africa. In *Water Research Commission* (Issue 891). https://doi.org/Available online: http://www.wrc.org.za/wp-content/uploads/mdocs/891-1-00.pdf (Accessed on 15 June 2018)

Benabdelouahab, T., Balaghi, R., Hadria, R., Lionboui, H., Minet, J., and Tychon, B. (2015). Monitoring surface water content using visible and short-wave infrared SPOT-5 data of wheat plots in irrigated semi-arid regions. *International Journal of Remote Sensing*, *36*(15), 4018–4036. https://doi.org/10.1080/01431161.2015.1072650

Best, M., Clark, D. B., Mercado, L. M., Sitch, S., Jones, C. D., Gedney, N., Best, M. J., Pryor, M., Rooney, G. G., Essery, R. L. H., Blyth, E., Boucher, O., Harding, R. J., and Cox, P. M. (2011). The Joint UK Land Environment Simulator (JULES), model description – Part 2: Carbon fluxes and vegetation. *Geoscientific Model Development Discussions*, *4*(1), 641–688. https://doi.org/10.5194/gmdd-4-641-2011

Beza, E., Silva, J. V., Kooistra, L., and Reidsma, P. (2017). Review of yield gap explaining factors and opportunities for alternative data collection approaches. *European Journal of Agronomy*, *82*, 206–222. https://doi.org/10.1016/j.eja.2016.06.016

Black, E., Tarnavsky, E., Greatrex, H., Maidment, R., Mookerjee, A., Quaife, T., and Price, J. (2015). Exploiting satellite-based rainfall for weather index insurance: the challenges of spatial and temporal aggregation. In I. Gitas (Ed.), *1st International Electronic Conference on Remote Sensing*. https://doi.org/10.3390/ecrs-1-f002

Black, E., Tarnavsky, E., Maidment, R., Greatrex, H., Mookerjee, A., Quaife, T., and Brown, M. (2016). The use of remotely sensed rainfall for managing drought risk: A case study of weather index insurance in Zambia. *Remote Sensing*, *8*(4). https://doi.org/10.3390/rs8040342

Blakeley, S. L., Sweeney, S., Husak, G., Harrison, L., Funk, C., Peterson, P., and Osgood, D. E. (2020). Identifying precipitation and reference evapotranspiration trends in West Africa to support drought insurance. *Remote Sensing*, *12*(15). https://doi.org/10.3390/RS12152432

Boke-Olén, N., Ardö, J., Eklundh, L., Holst, T., and Lehsten, V. (2018). Remotely sensed soil moisture to estimate savannah NDVI. *PLoS ONE*, *13*(7), 1–11. https://doi.org/10.1371/journal.pone.0200328

Bokke, A. S., and Shoro, K. E. (2020). Impact of effective rainfall on net irrigation water requirement: The case of Ethiopia. *Water Science*, *34*(1), 155–163. https://doi.org/10.1080/11104929.2020.1749780

Boyd, M., Porth, B., Porth, L., and Turenne, D. (2019). The impact of spatial interpolation techniques on spatial basis risk for weather insurance: an application to forage crops. *North*

American Actuarial Journal, *23*(3), 412–433. https://doi.org/10.1080/10920277.2019.1566074

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

Brisco, B., Brown, R. J., Gairns, J. G., and Snider, B. (1992). Temporal ground-based scatterometer observations of crops in Western Canada. *Canadian Journal of Remote Sensing*, *18*(1), 14–21. https://doi.org/10.1080/07038992.1992.10855138

Brisco, B., Ulaby, F. T., and Protz, R. (1984). Improving crop classification through attention to the timing of airborne radar acquisitions. *Photogrammetric Engineering and Remote Sensing*, *50*(6), 739–745. https://doi.org/Available online: https://www.asprs.org/wp-content/uploads/pers/1984journal/jun/1984_jun_739-745.pdf (Accessed on 19 May 2018)

Buhaug, H., Benaminsen, T. A., Sjaastad, E., and Magnus Theisen, O. (2015). Climate variability, food production shocks, and violent conflict in Sub-Saharan Africa. *Environmental Research Letters*, *10*(12). https://doi.org/10.1088/1748-9326/10/12/125015

Bulte, E., Cecchi, F., Lensink, R., Marr, A., and van Asseldonk, M. (2020). Does bundling crop insurance with certified seeds crowd-in investments? Experimental evidence from Kenya. *Journal of Economic Behavior and Organization*, *180*, 744–757. https://doi.org/10.1016/j.jebo.2019.07.006

Butts-Wilmsmeyer, C. J., Seebauer, J. R., Singleton, L., and Below, F. E. (2019). Weather during key growth stages explains grain quality and yield of maize. *Agronomy*, *9*(1), 1–15. https://doi.org/10.3390/agronomy9010016

Bzdok, D., Altman, N., and Krzywinski, M. (2018). Statistics versus machine learning. *Nature Methods*, *15*(4), 233–234. https://doi.org/10.1038/nmeth.4642

Çakir, R. (2004). Effect of water stress at different development stages on vegetative and reproductive growth of corn. *Field Crops Research*, *89*(1), 1–16. https://doi.org/10.1016/j.fcr.2004.01.005

Carletto, C., Jolliffe, D., and Banerjee, R. (2013). The Emperor has no data! Agricultural statistics in Sub-Saharan Africa. In *World Bank Working Paper*. https://doi.org/Available online: http://mortenjerven.com/wp-content/uploads/2013/04/Panel-3-Carletto.pdf (Accessed on 16 March 2018)

Caroletti, G. N., Coscarelli, R., and Caloiero, T. (2019). Validation of satellite, reanalysis and RCM data of monthly rainfall in Calabria (Southern Italy). *Remote Sensing*, *11*(13). https://doi.org/10.3390/rs11131625

Carter, M., De Janvry, A., Sadoulet, E., and Sarris, A. (2014). *Index-based weather insurance for developing countries: A review of evidence and a set of propositions for up-scaling*. https://doi.org/Available online: https://ferdi.fr/dl/df-kLB5hYBaUiR2Vysm2cH5DfCP/ferdi-p111-index-based-weather-insurance-for-developing-countries-a-review.pdf (Accessed on 18 January 2018)

Carter, Michael, De Janvry, A., Sadoulet, E., and Sarris, A. (2017). Index insurance for developing country agriculture: A reassessment. *Annual Review of Resource Economics*, *9*, 421–438. https://doi.org/10.1146/annurev-resource-100516-053352

Ceccherini, G., Ameztoy, I., Hernández, C. P. R., and Moreno, C. C. (2015). High-resolution precipitation datasets in South America and West Africa based on satellite-derived rainfall, enhanced vegetation index and digital elevation model. *Remote Sensing*, *7*(5), 6454–6488. https://doi.org/10.3390/rs70506454

Chai, T., and Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, *7*(3), 1247–1250. https://doi.org/10.5194/gmd-7-1247-2014

Chapagain, T., and Raizada, M. N. (2017). Impacts of natural disasters on smallholder farmers: Gaps and recommendations. *Agriculture and Food Security*, *6*(1), 1–16. https://doi.org/10.1186/s40066-017-0116-6

Chen, T., de Jeu, R. A. M., Liu, Y. Y., van der Werf, G. R., and Dolman, A. J. (2014). Using satellite based soil moisture to quantify the water driven variability in NDVI: A case study over mainland Australia. *Remote Sensing of Environment*, *140*, 330–338. https://doi.org/10.1016/j.rse.2013.08.022

Chen, T., and Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the*

ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, *13-17-Augu*, 785–794. https://doi.org/10.1145/2939672.2939785

Chen, W., Hohl, R., and Tiong, L. K. (2017). Rainfall index insurance for corn farmers in Shandong based on high-resolution weather and yield data. *Agricultural Finance Review*, *77*(2), 337–354. https://doi.org/10.1108/AFR-10-2015-0042

Chimonyo, V. G. P., Mutengwa, C. ., Chiduza, C., and Tandzi, L. (2019). Participatory variety selection of maize genotypes in the Eastern cape province of South Africa. *South African Journal of Agricultural Extension*, *47*(n1), 103–117. https://doi.org/http://dx.doi.org/10.17159/2413-3221/2019/v47n1a493

Chimonyo, V. G. P., Mutengwa, C. ., Chiduza, C., and Tandzi, L. (2020). Characteristics of maize growing farmers, varietal use, and constraints to increase productivity in selected villages in the Eastern Cape province of South Africa. *South African Journal of Agricultural Extension*, *48*(62), 71–77. https://doi.org/http://dx.doi.org/10.17159/2413-3221/2020/v48n2a538

Choudhury, A., Jones, J., Okine, A., and Choudhury, R. L. (2016). Drought-triggered index insurance using cluster analysis of rainfall affected by climate change. *Journal of Insurance Issues*, *39*(2), 169–186. https://doi.org/Available online: https://www.jstor.org/stable/43921956 (Accessed on 16 February 2020)

Clement, K. Y., Wouter Botzen, W. J., Brouwer, R., and Aerts, J. C. J. H. (2018). A global review of the impact of basis risk on the functioning of and demand for index insurance. *International Journal of Disaster Risk Reduction*, *28*(January), 845–853. https://doi.org/10.1016/j.ijdrr.2018.01.001

Clerici, N., Valbuena Calderón, C. A., and Posada, J. M. (2017). Fusion of sentinel-1a and sentinel-2A data for land cover mapping: A case study in the lower Magdalena region, Colombia. *Journal of Maps*, *13*(2), 718–726. https://doi.org/10.1080/17445647.2017.1372316

Collier, B., Skees, J., and Barnett, B. (2009). Weather index insurance and climate change: Opportunities and challenges in lower income countries. *Geneva Papers on Risk and*

Insurance: Issues and Practice, *34*(3), 401–424. https://doi.org/10.1057/gpp.2009.11

Cortez, C., and Vapnik, V. (1995). Support-Vector Networks. *Machine Learning*, *20*, 273–297. https://doi.org/10.1109/64.163674

Dabrowski, J. M., Murray, K., Ashton, P. J., and Leaner, J. J. (2009). Agricultural impacts on water quality and implications for virtual water trading decisions. *Ecological Economics*, *68*(4), 1074–1082. https://doi.org/10.1016/j.ecolecon.2008.07.016

Dalhaus, T., Musshoff, O., and Finger, R. (2018). Phenology information contributes to reduce temporal basis risk in agricultural weather index insurance. *Scientific Reports*, 1–10. https://doi.org/10.1038/s41598-017-18656-5

DALRRD. (2018). Trends in the agricultural sector. In *DALRRD*. https://doi.org/Available online: https://www.dalrrd.gov.za/Portals/0/Statistics and Economic Analysis/Statistical Information/Trends in the Agricultural Sector 2017.pdf (Accessed on 15 November 2018)

Dandekar, V. (1976). Crop insurance in India. *Economic and Political Weekly*, *11*(26). https://doi.org/Available online: https://www.jstor.org/stable/4364738 (Accessed on 18 January 2018)

Das, N. N., Entekhabi, D., Dunbar, R. S., Chaubell, M. J., Colliander, A., Yueh, S., Jagdhuber, T., Chen, F., Crow, W., O'Neill, P. E., Walker, J. P., Berg, A., Bosch, D. D., Caldwell, T., Cosh, M. H., Collins, C. H., Lopez-Baeza, E., and Thibeault, M. (2019). The SMAP and Copernicus Sentinel 1A/B microwave active-passive high resolution surface soil moisture product. *Remote Sensing of Environment*, *233*(February). https://doi.org/10.1016/j.rse.2019.111380

de Leeuw, J., Vrieling, A., Shee, A., Atzberger, C., Hadgu, K. M., Biradar, C. M., Keah, H., and Turvey, C. (2014). The potential and uptake of remote sensing in insurance: A review. *Remote Sensing*, *6*(11), 10888–10912. https://doi.org/10.3390/rs61110888

De Oto, L., Vrieling, A., Fava, F., and de Bie, K. (C A. J. M.). (2019). Exploring improvements to the design of an operational seasonal forage scarcity index from NDVI time series for livestock insurance in East Africa. *International Journal of Applied Earth Observation and Geoinformation*, *82*(January), 101885. https://doi.org/10.1016/j.jag.2019.05.018

Debats, S. R., Luo, D., Estes, L. D., Fuchs, T., and Caylor, K. (2015). *A generalized computer vision approach to mapping crop fields in heterogeneous agricultural landscapes*. 1–34. https://doi.org/10.7287/peerj.preprints.63v1

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

Delacote, P., Girard, J., and Leblois, A. (2019). *Agricultural households' adaptation to weather shocks in Sub-Saharan Africa: what implications for land-use change and deforestation?* (Issue September). https://doi.org/Available online: http://www.chaireeconomieduclimat.org/RePEc/cec/wpaper/19-09-18-Cahier-R-2019-02-Delacote-Girard-Leblois.pdf (Accessed on 15 July 2020)

Denize, J., Hubert-Moy, L., Betbeder, J., Corgne, S., Baudry, J., and Pottier, E. (2019). Evaluation of using sentinel-1 and -2 time-series to identify winter land use in agricultural landscapes. *Remote Sensing*, *11*(1). https://doi.org/10.3390/rs11010037

Denmead, O. T., and Shaw, R. H. (1960). The effects of soil moisture stress at different stages of growth on the development and yield of corn. *Agronomy Journal*, *52*(5), 272–274. https://doi.org/10.2134/agronj1960.00021962005200050010x

Department of Agriculture Land Reform and Rural Development (DALRRD). (2016). *Master agri-park business plan: Eastern Cape: O.R. Tambo District business plan* (Issue April). https://doi.org/Available online: http://www.ruraldevelopment.gov.za/phocadownload/Agri-parks/Business-Plans/OR-Tambo/o r tambo dm final master agri-park business plan section 1.pdf (Accessed on 18 July 2018)

Dercon, S., and Christiaensen, L. (2011). Consumption risk, technology adoption and poverty traps: Evidence from Ethiopia. *Journal of Development Economics*, *96*(2), 159–173. https://doi.org/http://dx.doi.org/10.1016/j.jdeveco.2010.08.003

Devereux, S. (2007). The impact of droughts and floods on food security and policy options to alleviate negative effects. *Agricultural Economics*, *37*(S1), 47–58. https://doi.org/10.1111/j.1574-0862.2007.00234.x

Dhumal, R., Rajendra, Y., Kale, K. V, and Mehrotra, S. C. (2013). Classification of crops from remotely sensed images using fuzzy classification approach. *International Journal of Advanced Research in Computer Science and Software Engineering*, *3*(7), 758–761. https://doi.org/Available online: www.ijera.com (Accessed on 16 March 2019)

Di Marcantonio, F., and Kayitakire, F. (2017). Review of pilot projects on index-based insurance in Africa: Insights and lessons learned. In M. Tiepolo, A. Pezzoli, and V. Tarchiani (Eds.), *Renewing Local Planning to Face Climate Change in the Tropics* (pp. 323–341). Springer International Publishing. https://doi.org/10.1007/978-3-319-59096-7_16

Diko, A., and Jun, W. (2020). Influencing factors of maize production in South Africa: the case of Mpumalanga, Free State and North West Provinces. *Asian Journal of Advances in Agricultural Research*, *14*(1), 25–34. https://doi.org/10.9734/ajaar/2020/v14i130121

Dimov, D., Low, F., Ibrakhimov, M., Stulina, G., and Conrad, C. (2017). SAR and optical time series for crop classification. *International Geoscience and Remote Sensing Symposium (IGARSS)*, *2017-July*(July), 811–814. https://doi.org/10.1109/IGARSS.2017.8127076

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

Djerriri, K., Ghabi, M., Karoui, M. S., and Adjoudj, R. (2018). Palm trees counting in remote sensing imagery using regression convolutional neural network. *International Geoscience and Remote Sensing Symposium (IGARSS)*, *2018-July*, 2627–2630. https://doi.org/10.1109/IGARSS.2018.8519188

Djurfeldt, G., Hall, O., Jirström, M., Archila Bustos, M., Holmquist, B., and Nasrin, S. (2018). Using panel survey and remote sensing data to explain yield gaps for maize in sub-Saharan Africa. *Journal of Land Use Science*, *13*(3), 344–357. https://doi.org/10.1080/1747423X.2018.1511763

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P. D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., Lahoz, W., Liu, Y. Y., Miralles, D., … Lecomte, P. (2017). ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions. *Remote Sensing of Environment*, *203*, 185–215. https://doi.org/10.1016/j.rse.2017.07.001

Douzas, G., Bacao, F., Fonseca, J., and Khudinyan, M. (2019). Imbalanced learning in land cover classification: Improving minority classes' prediction accuracy using the geometric SMOTE algorithm. *Remote Sensing*, *11*(24). https://doi.org/10.3390/rs11243040

du Plessis, J. (2003). Maize production. In *Agricultural Research Council*. https://doi.org/10.1787/8cb9bd05-en

DuPlessis, K., and Kibii, J. (2021). Applicability of CHIRPS-based satellite rainfall estimates for South Africa. *Journal of the South African Institution of Civil Engineering*, *63*(3), 43–54. https://doi.org/10.17159/2309-8775/2021/v63n3a4

Dutta, S., Chakraborty, S., Goswami, R., Banerjee, H., Majumdar, K., Li, B., and Jat, M. L. (2020). Maize yield in smallholder agriculture system—An approach integrating socio-economic and crop management factors. *Plos One*, *15*(2), e0229100. https://doi.org/10.1371/journal.pone.0229100

Džeroski, S., and Ženko, B. (2004). Is combining classifiers with stacking better than selecting the best one? *Machine Learning*, *54*(3), 255–273. https://doi.org/10.1023/B:MACH.0000015881.36452.6e

Eastern Cape Socio Economic Consultative Council (ECSECC). (2017). *O.R. Tambo District Municipality socio economic review and outlook, 2017*. https://doi.org/Available online: http://www.ecsecc.org (17 March 2019)

Eastern Cape Socio Economic Consultative Council (ECSECC). (2014). *Socio-economic profile of Eastern Cape province*. https://doi.org/Available online: http://www.ecsecc.org (17 March 2019)

Ekpa, O., Palacios-Rojas, N., Kruseman, G., Fogliano, V., and Linnemann, A. R. (2018). Sub-Saharan African maize-based foods: Technological perspectives to increase the food and nutrition security impacts of maize breeding programmes. *Global Food Security*, *17*(January), 48–56. https://doi.org/10.1016/j.gfs.2018.03.007

Elum, Z. A., Modise, D. M., and Marr, A. (2017). Farmer's perception of climate change and responsive strategies in three selected provinces of South Africa. *Climate Risk Management*, *16*, 246–257. https://doi.org/10.1016/j.crm.2016.11.001

Enenkel, M., Farah, C., Hain, C., White, A., Anderson, M., You, L., Wagner, W., and Osgood, D. (2018). What rainfall does not tell us — enhancing financial instruments with satellite-derived soil moisture and evaporative stress. *Remote Sensing*, *10*(11). https://doi.org/10.3390/rs10111819

Enenkel, M., Osgood, D., Anderson, M., Powell, B., McCarty, J., Neigh, C., Carroll, M., Wooten, M., Husak, G., Hain, C., and Brown, M. (2019). Exploiting the convergence of evidence in satellite data for advanced weather index insurance design. *Weather, Climate, and Society*, *11*(1), 65–93. https://doi.org/10.1175/WCAS-D-17-0111.1

Enenkel, M., Osgood, D., and Powell, B. (2017). The added value of satellite soil moisture for agricultural index insurance. In G. P. Petropoulos and T. Islam (Eds.), *Remote Sensing of Hydrometeorological Hazards* (pp. 69–83). CRS Press. https://doi.org/https://doi.org/10.1201/9781315154947

Engstrom, R., Hope, A., Kwon, H., and Stow, D. (2008). The relationship between soil moisture and NDVI near Barrow, Alaska. *Physical Geography*, *29*(1), 38–53. https://doi.org/10.2747/0272-3646.29.1.38

Eta, N. E., and Grace, G. (2013). Investigation of some physicochemical charactyeristics / prperties of geophagic soil in the Oliver Tambo District Munucipality in the Eastern cape. *Academic Journal of Science*, *2*(2), 465–471.

Eugster, M. J. A., Hothorn, T., and Leisch, F. (2008). Exploratory and inferential analysis of benchmark experiments. *Ludwigs-Maximilians-Universitat Munchen, Department of Statistics, Technical Report, Number 030*, *030*. https://doi.org/Available online: http://www.stat.uni-muenchen.de (Accessed on 11 March 2019)

Eze, E., Girma, A., Zenebe, A. A., and Zenebe, G. (2020). Feasible crop insurance indexes for drought risk management in Northern Ethiopia. *International Journal of Disaster Risk Reduction*, *47*(February), 101544. https://doi.org/10.1016/j.ijdrr.2020.101544

Eze, E., Girma, A., Zenebe, A., Kourouma, J. M., and Zenebe, G. (2020). Exploring the possibilities of remote yield estimation using crop water requirements for area yield index insurance in a data-scarce dryland. *Journal of Arid Environments*, *183*. https://doi.org/https://doi.org/10.1016/j.jaridenv.2020.104261

Fan, S., and Rue, C. (2020). The Role of Smallholder Farms in a Changing World. In S. y Paloma, L. Riesgo, and K. Louhichi (Eds.), *The Role of Smallholder Farms in Food and Nutrition Security* (pp. 13–28). Springer International Publishing. https://doi.org/10.1007/978-3-030-42148-9_2

Fanadzo, M., and Ncube, B. (2018). Challenges and opportunities for revitalising smallholder irrigation schemes in South Africa. *Water SA*, *44*(3), 436–447. https://doi.org/10.4314/wsa.v44i3.11

FAO. (2008). Climate change and food security: A framework document. In *Food and Agriculture Organization of the United Nations*. https://doi.org/Available online: www.fao.org/clim/index_en.htm.(Accessed on 17 June 2018)

FAO. (2015). *Livelihood diversification and vulnerability to poverty in rural Malawi* (Issue 15). https://doi.org/Available online: http://www.fao.org/3/a-i4887e.pdf (Accessed on 18 June 2018)

FAO. (2016). *Crop yield forecasting: Methodological and institutional aspects*. https://doi.org/Available online: www.fao.org/publications (Accessed on 12 February 2018)

Fontanelli, G., Crema, A., Azar, R., Stroppiana, D., Villa, P., and Boschetti, M. (2014). Agricultural crop mapping using optical and SAR multi-temporal seasonal data: A case study in Lombardy region, Italy. *International Geoscience and Remote Sensing Symposium (IGARSS)*, *July*, 1489–1492. https://doi.org/10.1109/IGARSS.2014.6946719

Forkuor, G., Conrad, C., Thiel, M., Ullmann, T., and Zoungrana, E. (2014). Integration of optical and synthetic aperture radar imagery for improving crop mapping in northwestern Benin, West Africa. *Remote Sensing*, *6*(7), 6472–6499. https://doi.org/10.3390/rs6076472

Foucras, M., Zribi, M., Clément, A., Baghdadi, N., Calvet, J.-C., and Pellarin, T. (2020). Estimating 500-m resolution soil moisture using Sentinel-1 and optical data synergy. *Water*, *12*(3). https://doi.org/doi:10.3390/w12030866

Friedman, J. (2001). Greedy function approximation : a gradient boosting machine. *The Annals of Statistics*, *29*(5), 1189–1232. https://doi.org/10.1214/009053606000000795

Frischen, J., Meza, I., Rupp, D., Wietler, K., and Hagenlocher, M. (2020). Drought risk to agricultural systems in Zimbabwe: A spatial analysis of hazard, exposure, and vulnerability. *Sustainability (Switzerland)*, *12*(3). https://doi.org/10.3390/su12030752

Frost, C., Thiebaut, N., and Newby, T. (2013). Evaluating Terra MODIS satellite sensor data products for maize yield estimation in South Africa. *South African Journal of Geomatics*, *2*(2), 106-119–119.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., and Michaelsen, J. (2015). The climate hazards infrared precipitation with stations - A new environmental record for monitoring extremes. *Scientific Data*, *2*, 1–21. https://doi.org/10.1038/sdata.2015.66

Gao, F., Anderson, M., Daughtry, C., and Johnson, D. (2018). Assessing the variability of corn and soybean yields in central Iowa using high spatiotemporal resolution multi-satellite imagery. *Remote Sensing*, *10*(9). https://doi.org/10.3390/rs10091489

Gao, Q., Zribi, M., and Escorihuela, M. J. (2017). *Synergetic use of Sentinel-1 and Sentinel-2 data for soil moisture mapping at 100 m resolution*. https://doi.org/10.3390/s17091966

Gatti, A., and Bertolini, A. (2015). Sentinel-2 Products Specification Document Sentinel-2. *European Space Agency*. https://doi.org/Available online:https://sentinel.esa.int/documents/247904/349490/S2_MSI_Product_Specification.pdf (Accessed on 22 February 2022)

Gaurav, S., and Chaudhary, V. (2020). Do farmers care about basis risk? Evidence from a field experiment in India. *Climate Risk Management*, *27*(November 2018), 100201. https://doi.org/10.1016/j.crm.2019.100201

Gautier, D., Denis, D., and Locatelli, B. (2016). Impacts of drought and responses of rural populations in West Africa: a systematic review. *Wiley Interdisciplinary Reviews: Climate Change*, *7*(5), 666–681. https://doi.org/https://doi.org/10.1002/wcc.411

Ge, T., Sui, F., Bai, L., Tong, C., and Sun, N. (2012). Effects of water stress on growth, biomass partitioning, and water-use efficiency in summer maize (Zea mays L.) throughout the growth cycle. *Acta Physiologiae Plantarum*, *34*(3), 1043–1053. https://doi.org/10.1007/s11738-011-0901-y

Gebremedhin, M. A., Lubczynski, M. W., Maathuis, B. H. P., and Teka, D. (2021). Novel approach to integrate daily satellite rainfall with in-situ rainfall, Upper Tekeze Basin, Ethiopia. *Atmospheric Research*, *248*(June 2020), 105135. https://doi.org/10.1016/j.atmosres.2020.105135

Gebremeskel, G., Tang, Q., Sun, S., Huang, Z., Zhang, X., and Liu, X. (2019). Droughts in East Africa: Causes, impacts and resilience. *Earth-Science Reviews*, *193*(June 2018), 146–161. https://doi.org/10.1016/j.earscirev.2019.04.015

Gebremicael, T., Mohamed, Y., van der Zaag, P., Berhe, A., Haile, G., Hagos, E., and Hagos, M. (2017). Comparison and validation of eight satellite rainfall products over the rugged topography of Tekeze-Atbara Basin at different spatial and temporal scales. *Hydrology and Earth System Sciences Discussions*, *August*, 1–31. https://doi.org/10.5194/hess-2017-504

Géron, A. (2019). *Hands-On Machine Learning with Scikit-Learn and TensorFlow*. O'Reilly Media. https://doi.org/Available online: http://oreilly.com/safari (Accessed on 18 January 2020)

Ghimire, B., Timsina, D., and Nepal, J. (2015). Analysis of chlorophyll content and its correlation with yield attributing traits on early varieties of maize (Zea mays L.). *Journal of Maize Research and Development*, *1*(1), 134–145. https://doi.org/10.3126/jmrd.v1i1.14251

Gornott, C., Hattermann, F., and Wechsung, F. (2015). Yield gap analysis for Tanzania – the impacts of climate, management, and socio-economic impacts on maize yields. *Procedia Environmental Sciences*, *29*(Agri), 231. https://doi.org/10.1016/j.proenv.2015.07.287

Graeub, B. E., Chappell, M. J., Wittman, H., Ledermann, S., Kerr, R. B., and Gemmill-Herren, B. (2016). The state of family farms in the world. *World Development*, *87*, 1–15. https://doi.org/10.1016/j.worlddev.2015.05.012

Greatrex, H., Hansen, J., Garvin, S., Diro, R., Blakeley, S., Guen, M. Le, Rao, K., Osgood, D., Hansen, J., and Garvin, S. (2015). *Scaling up index insurance for smallholder farmers : Recent evidence and insights* (Issue 14). https://doi.org/Available online: www.ccafs.cgiar.org%0AContact (Accessed on 5 may 2020)

Greyling, J. C., and Pardey, P. G. (2019). Measuring maize in South Africa: The shifting structure of production during the twentieth century, 1904–2015. *Agrekon*, *58*(1), 21–41. https://doi.org/10.1080/03031853.2018.1523017

Gruber, A., Scanlon, T., Van Der Schalie, R., Wagner, W., and Dorigo, W. (2019). Evolution of the ESA CCI Soil Moisture climate data records and their underlying merging methodology. *Earth System Science Data*, *11*(2), 717–739. https://doi.org/10.5194/essd-11-717-2019

Haarhoff, S. J., Kotzé, T. N., and Swanepoel, P. A. (2020). A prospectus for sustainability of rainfed maize production systems in South Africa. *Crop Science*, *60*(1), 14–28. https://doi.org/10.1002/csc2.20103

Halcrow, H. (1948). *The theory of crop insurance* [University of Chicago]. https://doi.org/Available online: https://ufdc.ufl.edu/UF00055285/00001 (Accessed on 16 January 2018)

Halcrow, H. (1949). Actuarial structures for crop insurance. *Journal of Farm Economics*, *31*(3), 418–443. https://doi.org/https://doi.org/10.2307/1232330

Hansen, J., Hellin, J., Rosenstock, T., Fisher, E., Cairns, J., Stirling, C., Lamanna, C., van Etten, J., Rose, A., and Campbell, B. (2019). Climate risk management and rural poverty reduction. *Agricultural Systems*, *172*(January 2018), 28–46. https://doi.org/10.1016/j.agsy.2018.01.019

Hartling, S., Sagan, V., Sidike, P., Maimaitijiang, M., and Carron, J. (2019). Urban tree species classification using a worldview-2/3 and liDAR data fusion approach and deep learning. *Sensors (Switzerland)*, *19*(6), 1–23. https://doi.org/10.3390/s19061284

Harvey, C. A., Saborio-Rodríguez, M., Martinez-Rodríguez, M. R., Viguera, B., Chain-Guadarrama, A., Vignola, R., and Alpizar, F. (2018). Climate change impacts and adaptation among smallholder farmers in Central America. *Agriculture and Food Security*, *7*(1), 1–20. https://doi.org/10.1186/s40066-018-0209-x

Hatfield, J. L., and Prueger, J. H. (2010). Value of using different vegetative indices to quantify agricultural crop characteristics at different growth stages under varying management practices. *Remote Sensing*, *2*(2), 562–578. https://doi.org/10.3390/rs2020562

Hazell, P., Anderson, J., Balzer, N., Clemmensen, A. H., Hess, U., and Rispoli, F. (2010). *The potential for scale and sustainability in weather index insurance for agriculture and rural livelihoods*. https://doi.org/Available online: http://www.ifad.org/ruralfinance/pub/weather.pdf (Accessed on 15 February 2018)

Hazell, P. B. R. (1992). The appropriate role of agricultural insurance in developing countries. *Journal of International Development*, *4*(6), 567–581. https://doi.org/https://doi.org/10.1002/jid.3380040602

He, J., Zheng, X., Rejesus, R. M., and Jr, J. M. Y. (2019). Moral hazard and adverse selection effects of cost-of-production crop insurance : evidence from the Philippines. *Australian Journal of Agricultural and Resource Economics*, 166–197. https://doi.org/10.1111/1467-8489.12290

Hellin, J., Hansen, J., and Araba, D. (2015). *Evidence-based insurance development for Nigeria's Farmers : Briefing paper for Nigerian Federal Ministry of Agriculture and Rural Development (FMARD)-CCAFS Knowledge-sharing workshop, London, 27-28 January 2015* (Issue January 2015). https://doi.org/https://hdl.handle.net/10568/68139

Henry, N. B., Lee, M., and Richard, G. (2018). Enhancing local livelihoods resilience and food security in the face of frequent flooding in Africa: A disaster management perspective. *Journal of African Studies and Development*, *10*(7), 85–100. https://doi.org/10.5897/jasd2018.0510

Hernandez, E., Goslinga, R., and Wang, V. (2018). Using satellite data to scale smallholder agricultural insurance. In *CGAP Brief* (Issue August). https://doi.org/Available online: https://www.cgap.org/sites/default/files/Brief-Using-Satellite-Data-Smallholder-Agricultural-Insurance-Aug-2018.pdf (Accessed on 18 June 2019)

Hess, U., and Syroka, J. (2005). Weather-based insurance in Southern Africa: the Case of

Malawi. In *Production*. https://doi.org/Available online: https://documents.worldbank.org/en/publication/documents-reports/documentdetail/818731468055789892/weather-based-insurance-in-southern-africa-the-case-of-malawi (Accessed on 9 March 2019)

Hill, R. V., Kumar, N., Magnan, N., Makhija, S., de Nicola, F., Spielman, D. J., and Ward, P. S. (2019). Ex ante and ex post effects of hybrid index insurance in Bangladesh. *Journal of Development Economics*, *136*(8), 1–17. https://doi.org/10.1016/j.jdeveco.2018.09.003

Hobbs, A., and Svetlichnaya, S. (2020). Satellite-based Prediction of Forage Conditions for Livestock in Northern Kenya. *Internation Conference On Learning Representations*, *2019*, 1–6.

Hochrainer-Stigler, S., van der Velde, M., Fritz, S., and Pflug, G. (2014). Remote sensing data for managing climate risks: Index-based insurance and growth related applications for smallhold-farmers in Ethiopia. *Climate Risk Management*, *6*, 27–38. https://doi.org/10.1016/j.crm.2014.09.002

Holloway, J., and Mengersen, K. (2018). Statistical Machine Learning Methods and Remote Sensing for Sustainable Development Goals : A Review. *Remote Sensing*, *10*(9). https://doi.org/10.3390/rs10091365

Hothorn, T., Leisch, F., Zeileis, A., and Hornik, K. (2005). The design and analysis of benchmark experiments. *Journal of Computational and Graphical Statistics*, *14*(3), 675–699. https://doi.org/10.1198/106186005X59630

Hunt, E. R., and Rock, B. N. (1989). Detection of changes in leaf water content using Near- and Middle-Infrared reflectances. *Remote Sensing of Environment*, *30*(1), 43–54. https://doi.org/10.1016/0034-4257(89)90046-1

Hur, J. H., Ihm, S. Y., and Park, Y. H. (2017). A variable impacts measurement in random forest for mobile cloud computing. *Wireless Communications and Mobile Computing*, *2017*. https://doi.org/10.1155/2017/6817627

IFAD. (2015). *Remote sensing for index insurance*. International Fund of Agricultural Development (IFAD). https://doi.org/Available online:

https://www.ifad.org/en/web/knowledge/publication/asset/39552352 (Accessed on 12 March 2018)

Inglada, J., Vincent, A., Arias, M., and Marais-Sicre, C. (2016). Improved early crop type identification by joint use of high temporal resolution sar and optical image time series. *Remote Sensing*, *8*(5). https://doi.org/10.3390/rs8050362

Iortyom, E. T., Mazinyo, S. P., and Nel, W. (2018). Analysis of the economic impact of Rural Enterprise Development Hub Project on maize farmers in Mqanduli, South Africa. *Indian Journal of Agricultural Research*, *52*(3), 243–249. https://doi.org/10.18805/IJARe.A-319

Itabari, J. K., Gregory, P. J., and Jones, R. K. (1993). Effects of temperature, soil water status and depth of planting on germination and emergence of maize (zea mays) adapted to semi-arid eastern kenya. *Experimental Agriculture*, *29*(3), 351–364. https://doi.org/10.1017/S0014479700020913

Jiang, Z., Huete, A. R., Didan, K., and Miura, T. (2008). Development of a two-band enhanced vegetation index without a blue band. *Remote Sensing of Environment*, *112*(10), 3833–3845. https://doi.org/10.1016/j.rse.2008.06.006

JICA. (2018). *Rural Resilience Enhancement Project in the Federal Democratic Republic of Ethiopia*. https://doi.org/Available online: https://www.jica.go.jp/jica-ri/index.html (Accessed on 10 December 2020)

Johnson, L. (2021). Rescaling index insurance for climate and development in Africa. *Economy and Society*, *50*(2), 248–274. https://doi.org/https://doi.org/10.1080/03085147.2020.1853364

Jordaan, A. ., Makate, D., Mashego, T., Ligthelm, M., Malherbe, J., Mwaka, B., Olivier, W., Symington, W., and Zyl, V. (2017). Vulnerability , adaptation to and coping with drought : the case of commercial and subsistence rain fed farming in the Eastern Cape. In *Volume II: Vol. I* (Issue May). Water Research Commission. https://doi.org/Available online: www.wrc.org.za (Accessed on 6 June 2018)

Kahan, D. (2013). Farm management extension guide: managing risks in farming. In D. Kahan (Ed.), *Food and Agriculture Organization of the United Nations* (2013th ed., Vol. 6). FAO.

https://doi.org/Available online: http://www.fao.org/uploads/media/3-ManagingRiskInternLores.pdf (Accessed on 14 July 2020)

Kambanje, A., Taruvinga, A., Mushunje, A., Mutengwa, C., and Ngarava, S. (2020). Determinants of food security status amongst smallholder farmers utilizing. *The Journal of Social Sciences Research*, *6*(2), 133–139. https://doi.org/Available online: https://www.arpgweb.com/pdf-files/jssr6(2)133-139.pdf (Accessed on 11 July 2019)

Karthikeyan, L., Chawla, I., and Mishra, A. K. (2020). A review of remote sensing applications in agriculture for food security: Crop growth and yield, irrigation, and crop losses. *Journal of Hydrology*, *586*(March), 124905. https://doi.org/10.1016/j.jhydrol.2020.124905

Karthikeyan, L., Pan, M., Wanders, N., Kumar, D. N., and Wood, E. F. (2017a). Four decades of microwave satellite soil moisture observations: Part 1. A review of retrieval algorithms. *Advances in Water Resources*, *109*, 106–120. https://doi.org/10.1016/j.advwatres.2017.09.006

Karthikeyan, L., Pan, M., Wanders, N., Kumar, D. N., and Wood, E. F. (2017b). Four decades of microwave satellite soil moisture observations: Part 2. Product validation and inter-satellite comparisons. *Advances in Water Resources*, *109*, 236–252. https://doi.org/10.1016/j.advwatres.2017.09.010

Keita, N., Carfagna, E., and Mu'Ammar, G. (2010). Issues and guidelines for the emerging use of GPS and PDAs in agricultural statistics in developing countries. *The Fifth International Conference on on Agricultural Statistics (ICAS V)*, 12–15. https://doi.org/Available online: http://www.fao.org/fileadmin/templates/ess/documents/meetings_and_workshops/ICAS5/PDF/ICASV_2.2_097_Paper_Keita.pdf (Accessed on 13 January 2018)

Kenduiywo, B. K., Ghosh, A., Hijmans, R., and Ndungu, L. (2020). Maize yield estimation in kenya using modis. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *5*(3), 477–482. https://doi.org/10.5194/isprs-Annals-V-3-2020-477-2020

Khapayi, M., and Celliers, P. R. (2016). Factors limiting and preventing emerging farmers to progress to commercial agricultural farming in the King William's town area of the Eastern

Cape Province, South Africa. *South African Journal of Agricultural Extension*, *44*(1), 1–10. https://doi.org/http://dx.doi.org/10.17159/2413-3221/2016/v44n1a374

Kihara, J., Tamene, L. D., Massawe, P., and Bekunda, M. (2015). Agronomic survey to assess crop yield, controlling factors and management implications: a case-study of Babati in northern Tanzania. *Nutrient Cycling in Agroecosystems*, *102*(1), 5–16. https://doi.org/10.1007/s10705-014-9648-3

Kilavi, M., MacLeod, D., Ambani, M., Robbins, J., Dankers, R., Graham, R., Helen, T., Salih, A. A. M., and Todd, M. C. (2018). Extreme rainfall and flooding over Central Kenya Including Nairobi City during the long-rains season 2018: Causes, predictability, and potential for early warning and actions. *Atmosphere*, *9*(12). https://doi.org/https://doi.org/10.3390/atmos9120472

Kimani, M. W., Hoedjes, J. C. B., and Su, Z. (2017). An assessment of satellite-derived rainfall products relative to ground observations over East Africa. *Remote Sensing*, *9*(5). https://doi.org/10.3390/rs9050430

Kölle, W., Buchholz, M., and Musshoff, O. (2021). Do high-resolution satellite indices at field level reduce basis risk of satellite-based weather index insurance? *Agricultural Finance Review*, *ahead-of-p*(ahead-of-print). https://doi.org/10.1108/afr-12-2020-0177

Komarek, A. M., De Pinto, A., and Smith, V. H. (2020). A review of types of risks in agriculture: What we know and what we need to know. *Agricultural Systems*, *178*(October 2019), 102738. https://doi.org/10.1016/j.agsy.2019.102738

Kotikot, S. M., Flores, A., Griffin, R. E., Nyaga, J., Case, J. L., Mugo, R., Sedah, A., Adams, E., Limaye, A., and Irwin, D. E. (2020). Statistical characterization of frost zones: Case of tea freeze damage in the Kenyan highlands. *International Journal of Applied Earth Observation and Geoinformation*, *84*(August 2019), 101971. https://doi.org/10.1016/j.jag.2019.101971

Kotikot, S. M., Flores, A., Griffin, R. E., Sedah, A., Nyaga, J., Mugo, R., Limaye, A., and Irwin, D. E. (2018). Mapping threats to agriculture in East Africa: Performance of MODIS derived LST for frost identification in Kenya's tea plantations. *International Journal of Applied*

Earth Observation and Geoinformation, *72*(October 2017), 131–139. https://doi.org/10.1016/j.jag.2018.05.009

Koutsos, T. M., Menexes, G. C., and Dordas, C. A. (2019). An efficient framework for conducting systematic literature reviews in agricultural sciences. *Science of the Total Environment*, *682*, 106–117. https://doi.org/10.1016/j.scitotenv.2019.04.354

Kozlowski, T. (1972). *Plant responses and control of water balance* (T. Kozlowski (ed.); Volume 3). Academic Press.

KSD. (2021). *King Sabata Dalindyebo Municipality IDP 2021/2022*. https://doi.org/Available online: https://ksd.gov.za/documents (Accessed on 15 June 2021)

Kulkarni, S. C., and Rege, P. P. (2020). Pixel level fusion techniques for SAR and optical images: A review. *Information Fusion*, *59*(November 2019), 13–29. https://doi.org/10.1016/j.inffus.2020.01.003

Kumar, A. A., Akula, B., and Mani, A. (2019). Estimation of water requirement in maize using CROPWAT model. *Journal of Pharmacognosy and Phytochemistry*, *8*(6), 2533–2537. https://doi.org/Available online: https://www.phytojournal.com/archives/2019/vol8issue6/PartAO/9-1-72-831.pdf (Accessed on 15 November 2021)

Łabędzki, L., and Bąk, B. (2015). Meteorological and agricultural drought indices used in drought monitoring in Poland: a review. *Meteorology Hydrology and Water Management*, *2*(2), 3–14. https://doi.org/10.26491/mhwm/34265

Langyintuo, A. (2020). Smallholder farmers' access to inputs and finance in Africa. In S. Paloma, L. Riesgo, and K. Louhichi (Eds.), *The Role of Smallholder Farms in Food and Nutrition Security* (pp. 133–152). Springer International Publishing. https://doi.org/10.1007/978-3-030-42148-9

Le Coz, C., and Van De Giesen, N. (2020). Comparison of rainfall products over sub-saharan africa. *Journal of Hydrometeorology*, *21*(4), 553–596. https://doi.org/10.1175/JHM-D-18-0256.1

Lewis, J. E., Rowland, J., and Nadeau, A. (1998). Estimating maize production in kenya using ndvi: Some statistical considerations. *International Journal of Remote Sensing*, *19*(13), 2609–2617. https://doi.org/10.1080/014311698214677

Ley, C., Martin, R. K., Pareek, A., Groll, A., Seil, R., and Tischer, T. (2022). Machine learning and conventional statistics: making sense of the differences. *Knee Surgery, Sports Traumatology, Arthroscopy*, 1–5. https://doi.org/10.1007/s00167-022-06896-6

Lievens, H., Reichle, R. H., Liu, Q., De Lannoy, G. J. M., Dunbar, R. S., Kim, S. B., Das, N. N., Cosh, M., Walker, J. P., and Wagner, W. (2017). Joint Sentinel-1 and SMAP data assimilation to improve soil moisture estimates. *Geophysical Research Letters*, *44*(12), 6145–6153. https://doi.org/10.1002/2017GL073904

Liu, J., Shang, J., Qian, B., Huffman, T., Zhang, Y., Dong, T., Jing, Q., and Martin, T. (2019). Crop yield estimation using time-series MODIS data and the effects of cropland masks in Ontario, Canada. *Remote Sensing*, *11*(20). https://doi.org/10.3390/rs11202419

Lotriet, R., Bisschoff, C., and Kole, A. (2017). Assessing fertilizer buying behavior of emerging farmers in a south african grain producing area. *Problems and Perspectives in Management*, *15*(2), 456–467. https://doi.org/10.21511/ppm.15(2-2).2017.14

Lowder, S. K., Skoet, J., and Raney, T. (2016). The number, size, and distribution of farms, smallholder farms, and family farms worldwide. *World Development*, *87*, 16–29. https://doi.org/10.1016/j.worlddev.2015.10.041

MacCarthy, D. S., Adiku, S. G., Freduah, B. S., Kamara, A. Y., Narh, S., and Abdulai, A. L. (2018). Evaluating maize yield variability and gaps in two agroecologies in northern Ghana using a crop simulation model. *South African Journal of Plant and Soil*, *35*(2), 137–147. https://doi.org/10.1080/02571862.2017.1354407

Mahlalela, P. T., Blamey, R. C., Hart, N. C. G., and Reason, C. J. C. (2020). Drought in the Eastern Cape region of South Africa and trends in rainfall characteristics. *Climate Dynamics*, *55*(9–10), 2743–2759. https://doi.org/10.1007/s00382-020-05413-0

Maidment, R., Black, E., Greatrex, H., and Young, M. (2020). TAMSAT. In V. Levizzani, C. Kidd, D. Kirschbaum, C. Kummerow, K. Nakamura, Turk, and F (Eds.), *Satellite Precipitation Measurement*. Springer. https://doi.org/https://doi.org/10.1007/978-3-030-

24568-9_22

Makaudze, E. M. (2005). *Do seasonal climate forecasts and crop insurance matter for smallholder farmers in Zimbabwe? Using contingent valuation method and Remote Sensing applications* [The Ohio State University]. https://doi.org/Available online: https://etd.ohiolink.edu/apexprod/rws_olink/r/1501/10?clear=10andp10_accession_num=osu1110389049 (Accessed on 16 April 2019)

Makaudze, E. M. (2018). Malawi's Experience with weather index insurance as agricultural risk mitigation strategy against extreme drought events. In P. J. Sallis (Ed.), *Extreme Weather* (p. 13). IntechOpen. https://doi.org/http://dx.doi.org/10.5772/intechopen.77106

Makaudze, E. M., and Miranda, M. (2008). Using Remotely-Sensed Vegetation Condition Index to investigate the feasibility of drought insurance in Zimbabwe. *University of Oxford Conference on Economic Develoment in Africa*. https://doi.org/10.1017/CBO9781107415324.004

Makaudze, E. M., and Miranda, M. J. (2010). Catastrophic drought insurance based on the remotely sensed normalised difference vegetation index for smallholder farmers in Zimbabwe. *Agrekon*, *49*(4), 418–432. https://doi.org/10.1080/03031853.2010.526690

Manderson, A., Kubayi, N., and Drimie, S. (2016). The impact of the South africa drought as experienced by smallholder farmers over June 2015-February 2016 period in the Mopani district of Limpopo, South Africa. In *Southern Africa Food Lab* (Issue February). https://doi.org/Available online: http://www.southernafricafoodlab.org/wp-content/uploads/2016/06/SAFL-Drough-Impact-Assessment-2016-Final.pdf (Accessed on 10 May 2019)

Mannocchi, F., Todisco, F., and Vergni, L. (2004). Agricultural drought: indices, definition and analysis. *Proceedings of Die UNESCO/IAHS/IWIIA Symposium Held in Rome. December 2003*, *286*, 246–254. https://doi.org/Available online: http://hydrologie.org/redbooks/a286/iahs_286_0246.pdf (Accessed on 13 July 2020)

Manyevere, A., Muchaonyerwa, P., Laker, M., and Mnkeni, P. (2014). Farmers ' perspectives with regard to crop production : an analysis of Nkonkobe Municipality , South Africa.

Journal of Agriculture and Rural Development in the Tropics and Subtropics, *115*(1), 41–53. https://doi.org/Available online: https://kobra.uni-kassel.de/handle/123456789/2014020344887 (Accessed on 15 May 2019)

Mapfumo, S. (2007). Weather index insurance: the case for South Africa. In *Weather* (Issue February). Micro Insurance Agency Holdings LLC. https://doi.org/Available online: https://www.findevgateway.org/sites/default/files/publications/files/mfg-en-case-study-weather-index-insurance-the-case-for-south-africa-feb-2007.pdf (Accessed on 15 January 2018)

Mapfumo, S. (2008). *Weather index crop insurance: implementation, product design, challenges and successes - lessons learned in the field* (Issue November). https://doi.org/Available online: https://www.findevgateway.org/paper/2008/11/weather-index-crop-insurance-implementation-product-design-challenges-and-successes (Accessed on 16 January 2018)

Martínez-Fernández, J., González-Zamora, A., Sánchez, N., Gumuzzio, A., and Herrero-Jiménez, C. M. (2016). Satellite soil moisture for agricultural drought monitoring: Assessment of the SMOS derived Soil Water Deficit Index. *Remote Sensing of Environment*, *177*, 277–286. https://doi.org/10.1016/j.rse.2016.02.064

Mashaba-Munghemezulu, Z., Chirima, G. J., and Munghemezulu, C. (2021). Delineating smallholder maize farms from sentinel-1 coupled with sentinel-2 data using machine learning. *Sustainability (Switzerland)*, *13*(9). https://doi.org/10.3390/su13094728

Masih, I., Maskey, S., Mussá, F. E. F., and Trambauer, P. (2014). A review of droughts on the African continent: A geospatial and long-term perspective. *Hydrology and Earth System Sciences*, *18*(9), 3635–3649. https://doi.org/10.5194/hess-18-3635-2014

Masiza, W, Chirima, J. ., Hamandawana, H., Kalumba, A. M., and Magagula, H. B. (2021). Linking agricultural index insurance with factors that influence maize yield in rain-fed smallholder farming systems. *Sustainability*, *13*(9). https://doi.org/https://doi.org/10.3390/su13095176

Masiza, W, Chirima, J., Hamandawana, H., and Pillay, R. (2020). Enhanced mapping of a smallholder crop farming landscape through image fusion and model stacking. *International*

Journal of Remote Sensing, *41*(22). https://doi.org/http://dx.doi.org/10.1080/01431161.2020.1783017

Masiza, Wonga, Chirima, G. J., Hamandawana, H., Kalumba, A. M., and Magagula, H. B. (2022). Do satellite data correlate with in-situ rainfall and smallholder crop yields? Implications for crop insurance. *Sustainability*, *14*(3). https://doi.org/https://doi.org/10.3390/su14031670

Masiza, Wonga, Chirima, J. G., Hamandawana, H., Kalumba, A. M., and Magagula, H. B. (2021). *Linking Agricultural Index Insurance with Factors That Influence Maize Yield in Rain-Fed Smallholder Farming Systems*.

Masupha, Teboho E., and Moeletsi, M. E. (2020). The use of Water Requirement Satisfaction Index for assessing agricultural drought on rain-fed maize, in the Luvuvhu River catchment, South Africa. *Agricultural Water Management*, *237*(March), 106142. https://doi.org/10.1016/j.agwat.2020.106142

Masupha, Teboho Elisa, and Moeletsi, M. E. (2017). Use of standardized precipitation evapotranspiration index to investigate drought relative to maize, in the Luvuvhu River catchment area, South Africa. *Physics and Chemistry of the Earth*, *102*, 1–9. https://doi.org/10.1016/j.pce.2017.08.002

Mavhura, E. (2019). Analysing drivers of vulnerability to flooding: a systems approach. *South African Geographical Journal*, *101*(1), 72–90. https://doi.org/https://doi.org/10.1080/03736245.2018.1541020

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

McNally, A., Husak, G. J., Brown, M., Carroll, M., Funk, C., Yatheendradas, S., Arsenault, K., Peters-Lidard, C., and Verdin, J. P. (2015). Calculating crop water requirement satisfaction in the West Africa sahel with remotely sensed soil moisture. *Journal of Hydrometeorology*, *16*(1), 295–305. https://doi.org/10.1175/JHM-D-14-0049.1

Mehta, S., Rupela, O., Sonali, B., Nayak, A., and Hedge, N. (2010). Improving the livelihoods of the resource-poor smallholder farmers and producers in developing countries : an urgent appeal for action by GCARD. *Global Conference on Agricultural Research for Development*, 1–29. https://doi.org/Available online: http://www.fao.org/docs/eims/upload/273459/White%20Paper%205Mar2k10.pdf (Accessed 15 January 2018)

Mendelsohn, R., Kurukulasuriya, P., Basist, A., Kogan, F., and Williams, C. (2007). Climate analysis with satellite versus weather station data. *Climatic Change*, *81*(1), 71–83. https://doi.org/10.1007/s10584-006-9139-x

Mengist, W., Soromessa, T., and Legese, G. (2020). Method for conducting systematic literature review and meta-analysis for environmental science research. *MethodsX*, *7*, 100777. https://doi.org/10.1016/j.mex.2019.100777

Mercier, A., Betbeder, J., Rumiano, F., Baudry, J., Gond, V., Blanc, L., Bourgoin, C., Cornu, G., Ciudad, C., Marchamelo, M., Poccard-Chapuis, R., and Hubert-Moy, L. (2019). Evaluation of Sentinel-1 and 2 time series for land cover classification of forest–agriculture mosaics in temperate and tropical landscapes. *Remote Sensing*, *11*(8), 1–20. https://doi.org/10.3390/rs11080979

Merz, C. J. (1999). Using correspondence analysis to combine classifiers. *Machine Learning*, *36*(1), 33–58. https://doi.org/10.1023/A:1007559205422

Mi, N., Cai, F., Zhang, Y., Ji, R., Zhang, S., and Wang, Y. (2018). Differential responses of maize yield to drought at vegetative and reproductive stagess. *Plant, Soil and Environment*, *64*(6), 260–267. https://doi.org/10.17221/141/2018-PSE

Minde, I., Pedzisa, T., and Dimes, J. (2008). *Improving access and utilization of fertilizers by smallholder farmers in the Limpopo Province of South Africa*. https://doi.org/Avilable online: http://www.icrisat.org/Publications/EBooksOnlinePublications/Publications-2008/IDRC%20Report%203_Improving%20Final.pdf (Accessed on 19 March 2018)

Miranda, M. J., and Farrin, K. (2012). Index insurance for developing countries. *Applied*

Economic Perspectives and Policy, *34*(3), 391–427. https://doi.org/10.1093/aepp/pps031

Miranda, M. J., and Mulangu, F. M. (2016). *Index insurance for agricultural transformation in Africa*. https://doi.org/Available online: https://www.jica.go.jp/jica-ri/index (Accessed on 16 March 2019)

Mirsoleimani, H. R., Sahebi, M. R., Baghdadi, N., and El Hajj, M. (2019). Bare soil surface moisture retrieval from sentinel-1 SAR data based on the calibrated IEM and dubois models using neural networks. *Sensors (Switzerland)*, *19*(14), 1–12. https://doi.org/10.3390/s19143209

Mitra, S., Patnaik, P., and Kebbekus, B. B. (2018). Environmental chemical analysis. In S. Mitra, P. Patnaik, and B. B. Kebbekus (Eds.), *Environmental Chemical Analysis* (2nd ed.). CRC Press. https://doi.org/10.1201/b15740

Mjanyelwa, N., Bello, Z. A., Greaves, W., and van Rensburg, L. D. (2016). Precision and accuracy of DFM soil water capacitance probes to measure temperature. *Computers and Electronics in Agriculture*, *125*, 125–128. https://doi.org/10.1016/j.compag.2016.04.027

Mkhabela, M. S., Mkhabela, M. S., and Mashinini, N. N. (2005). Early maize yield forecasting in the four agro-ecological regions of Swaziland using NDVI data derived from NOAA's-AVHRR. *Agricultural and Forest Meteorology*, *129*(1–2), 1–9. https://doi.org/10.1016/j.agrformet.2004.12.006

Möllmann, J., Buchholz, M., Kölle, W., and Musshoff, O. (2020). Do remotely-sensed vegetation health indices explain credit risk in agricultural microfinance? *World Development*, *127*. https://doi.org/10.1016/j.worlddev.2019.104771

Morita, H., Takamura, H., and Okumura, M. (2009). Structured output learning with polynomial kernel. In G. Angelova and R. Mitkov (Eds.), *International Conference Recent Advances in Natural Language Processing, RANLP* (pp. 281–286). Association for Computational Linguistics.

Mpandeli, S., Nesamvuni, E., and Maponya, P. (2015). Adapting to the impacts of drought by smallholder farmers in Sekhukhune District in Limpopo Province, South Africa. *Journal of Agricultural Science*, *7*(2). https://doi.org/10.5539/jas.v7n2p115

Mugambiwa, S. S., and Tirivangasi, H. M. (2017). Climate change: A threat towards achieving "sustainable development goal number two" (end hunger, achieve food security and improved nutrition and promote sustainable agriculture) in South Africa. *Jamba: Journal of Disaster Risk Studies*, *9*(1), 1–6. https://doi.org/10.4102/jamba.v9i1.350

Muhammad, N. (2009). Simulation of maize crop under irrigated and rainfed conditions with CROPWAT model. *Journal of Agricultural and Biological Science, 4*(2), 68–73. https://doi.org/Available online: http://www.arpnjournals.com/jabs/research_papers/rp_2009/jabs_0309_123.pdf (Accessed on 18 November 2021)

Mukherjee, P., Pandey, M., and Prashad, P. (2017). *Bundling to make agriculture insurance work* (Issue 47). https://doi.org/Available online: http://www.impactinsurance.org/publications/mp47 (Accessed on 18 January 2021)

Musyoki, A., Thifhulufhelwi, R., and Murungweni, F. M. (2016). The impact of and responses to flooding in Thulamela Municipality, Limpopo Province, South Africa. *Jamba: Journal of Disaster Risk Studies*, *8*(2), 1–10. https://doi.org/http://dx.doi. org/10.4102/jamba.v8i2.166

Myeni, L., Moeletsi, M. E., and Clulow, A. D. (2019). Present status of soil moisture estimation over the African continent. *Journal of Hydrology: Regional Studies*, *21*(May 2018), 14–24. https://doi.org/10.1016/j.ejrh.2018.11.004

Myeni, L., Moeletsi, M. E., and Clulow, A. D. (2021). Field calibration of dfm capacitance probes for continuous soil moisture monitoring. *Water SA*, *47*(1), 88–96. https://doi.org/10.17159/wsa/2021.v47.i1.9448

NAMC. (2020). *Agripreneur: Agriculture insurance* (Issue 20). https://doi.org/Available online: https://www.namc.co.za/wp-content/uploads/2020/04/Agrepreneur-Issue-20-March-2020.pdf (Accessed on 13 February 2022)

Neigh, C. S. R., Carroll, M. L., Wooten, M. R., McCarty, J. L., Powell, B. F., Husak, G. J., Enenkel, M., and Hain, C. R. (2018). Smallholder crop area mapped with wall-to-wall WorldView sub-meter panchromatic image texture: A test case for Tigray, Ethiopia. *Remote Sensing of Environment*, *212*(April), 8–20. https://doi.org/10.1016/j.rse.2018.04.025

Nembilwi, N., Chikoore, H., Kori, E., Munyai, R. B., and Manyanya, T. C. (2021). The occurrence of drought in Mopani District Municipality, South Africa: Impacts, vulnerability and adaptation. *Climate*, *9*(4), 61. https://doi.org/https://doi.org/10.3390/cli9040061

Ngie, A., and Ahmed, F. (2018). Estimation of Maize grain yield using multispectral satellite data sets (SPOT 5) and the random forest algorithm. *South African Journal of Geomatics*, *7*(1), 11. https://doi.org/10.4314/sajg.v7i1.2

Nielsen, D. (2016). Tree boosting with XGBoost [Norwegian University of Science and Technology]. In *Master's Thesis, Norweigian University of Science and Technology*. https://doi.org/10.1111/j.1758-5899.2011.00096.x

Nieuwoudt, L., and Groenewald, J. (2003). *The challenge of change: Agriculture, land and the South African economy* (L. Nieuwoudt and J. Groenewald (eds.); 1st ed.). University of KwaZulu-Natal Press.

Ntshakira-Rukundo, E., Kamau, J., and Baumüller, H. (2021). Determinants of uptake and strategies to improve agricultural insurance in Africa : a review. *Environment and Development Economics*, 1–27. https://doi.org/10.1017/S1355770X21000085

Ntukamazina, N., Onwonga, R. N., Sommer, R., Rubyogo, J. C., Mukankusi, C. M., Mburu, J., and Kariuki, R. (2017). Index-based agricultural insurance products: Challenges, opportunities and prospects for uptake in sub-Sahara Africa. *Journal of Agriculture and Rural Development in the Tropics and Subtropics*, *118*(2), 171–185. https://doi.org/Available online: https://www.jarts.info/index.php/jarts/article/view/2017042052372 (Accessed on 13 May 2019)

Nyambo, P., and Wakindiki, I. I. C. (2015). Water footprint of growing vegetables in selected smallholder irrigation schemes in South Africa. *Water SA*, *41*(4), 571–578. https://doi.org/10.4314/wsa.v41i4.17

Nyandeni Local Municipality. (2018). *Nyandeni Local Municipality: local economic development strategy review* (Issue November). https://doi.org/Available online: https://www.nyandenilm.gov.za/wp-content/uploads/2019/08/Final-Nyandeni-LM-LED-

Strategy-Review-pdf (Accessed on 18 November 2020)

Oduniyi, O. S., Antwi, M. A., and Tekana, S. S. (2020). Farmers'willingness to pay for index-based livestock insurance in the North West of South Africa. *Climate*, *8*(3). https://doi.org/10.3390/cli8030047

Okeyo, S. A., and Mulaku, G. C. (2020). Crop insurance based on satellite remote sensing: case of Migori county, Kenya. *Advances in Remote Sensing*, *09*, 1–11. https://doi.org/10.4236/ars.2020.91001

ORTDM. (2017). Annual report 2016/2017. In *O.R. Tambo District Municipality*. https://doi.org/Available online:https://www.ortambodm.gov.za (Accessed on 13 May 2018)

Osgood, D., Mclaurin, M., Carriquiry, M., Mishra, A., Fiondella, F., Hansen, J., Peterson, N., and Ward, N. (2007). Designing weather insurance contracts for farmers in Malawi, Tanzania, and Kenya. In *Final Report to the Commodity Risk Management Group, ARD, World Bank*. https://doi.org/Available online: https://iri.columbia.edu/~deo/IRI-CRMG-Africa-Insurance-Report-6-2007/IRI-CRMG-Kenya-Tanzania-Malawi-Insurance-Report-6-2007.pdf (Accessed on 15 March 2018)

Osgood, D., Powell, B., Diro, R., Farah, C., Enenkel, M., Brown, M. E., Husak, G., Blakeley, S. L., Hoffman, L., and Mccarty, J. L. (2018). Farmer perception , recollection , and remote sensing in weather index insurance : An Ethiopia case study. *Remote Sensing*, *10*(12). https://doi.org/10.3390/rs10121887

Partridge, A. G., and Wagner, N. J. (2016). Risky business: agricultural insurance in the face of climate change: Elsenburg journal. *Agriprobe*, *13*(3), 49–53. https://doi.org/https://hdl.handle.net/10520/EJC195182

Pereira, L. S., and Alves, I. (2013). Crop Water Requirements. In *Reference Module in Earth Systems and Environmental Sciences* (Issue May). Elsevier Inc. https://doi.org/10.1016/b978-0-12-409548-9.05129-0

Persello, C., Tolpekin, V. A., Bergado, J. R., and de By, R. A. (2019). Delineation of agricultural fields in smallholder farms from satellite images using fully convolutional networks and combinatorial grouping. *Remote Sensing of Environment*, *231*(April). https://doi.org/10.1016/j.rse.2019.111253

Petropoulos, G. P., Ireland, G., and Barrett, B. (2015). Surface soil moisture retrievals from remote sensing: Current status, products and future trends. *Physics and Chemistry of the Earth, 83–84*, 36–56. https://doi.org/10.1016/j.pce.2015.02.009

Phadikar, S., and Goswami, J. (2016). Vegetation indices based segmentation for automatic classification of brown spot and blast diseases of rice. *2016 3rd International Conference on Recent Advances in Information Technology, RAIT 2016*, 284–289. https://doi.org/10.1109/RAIT.2016.7507917

Prats-Iraola, P., Nannini, M., Scheiber, R., De Zan, F., Wollstadt, S., Minati, F., Vecchioli, F., Costantini, M., Borgstrom, S., De Martino, P., Siniscalchi, V., Walter, T., Foumelis, M., and Desnos, Y. L. (2015). Sentinel-1 assessment of the interferometric wide-swath mode. *International Geoscience and Remote Sensing Symposium (IGARSS)*, *2015-Novem*, 5248–5251. https://doi.org/10.1109/IGARSS.2015.7327018

Prudnikova, E., Savin, I., Vindeker, G., Grubina, P., Shishkonakova, E., and Sharychev, D. (2019). Influence of soil background on spectral reflectance of winter wheat crop canopy. *Remote Sensing*, *11*(16). https://doi.org/10.3390/rs11161932

Qadir, A., and Mondal, P. (2020). Synergistic use of radar and optical satellite data for improved monsoon cropland mapping in India. *Remote Sensing*, *12*(3). https://doi.org/10.3390/rs12030522

Raithatha, R., and Priebe, J. (2020). *Agricultural insurance for smallholder farmers Digital innovations for scale*. https://doi.org/Available online: https://www.gsma.com/mobilefordevelopment/wp-content/uploads/2020/05/Agricultural_Insurance_for_Smallholder_Farmers_Digital_Innovations_for_Scale.pdf (Accessed on 1 September 2021)

Rajah, P., Odindi, J., and Mutanga, O. (2018). Feature level image fusion of optical imagery and Synthetic Aperture Radar (SAR) for invasive alien plant species detection and mapping. *Remote Sensing Applications: Society and Environment*, *10*, 198–208.

https://doi.org/10.1016/j.rsase.2018.04.007

Rajula, H. S. R., Verlato, G., Manchia, M., Antonucci, N., and Fanos, V. (2020). Comparison of conventional statistical methods with machine learning in medicine: Diagnosis, drug development, and treatment. *Medicina (Lithuania)*, *56*(9), 1–10. https://doi.org/10.3390/medicina56090455

Rao, N. R. (2008). Development of a crop-specific spectral library and discrimination of various agricultural crop varieties using hyperspectral imagery. *International Journal of Remote Sensing*, *29*(1), 131–144. https://doi.org/10.1080/01431160701241779

Rapsomanikis, G. (2015). *The economic lives of smallholder farmers*. https://doi.org/Available online: http://www.fao.org/3/i5251e/i5251e.pdf (Accessed on 7 April 2018)

Ripley, B. D. (1996). Pattern recognition and Neural Networks. In *Cambridge University Press*. https://doi.org/https://doi.org/10.1017/CBO9780511812651

Rodríguez-Fernández, N., Al Bitar, A., Colliander, A., and Zhao, T. (2019). Soil moisture remote sensing across scales. *Remote Sensing*, *11*(2), 1–5. https://doi.org/10.3390/rs11020190

Rwanga, S. S., and Ndambuki, J. M. (2017). Accuracy Assessment of Land Use/Land Cover Classification Using Remote Sensing and GIS. *International Journal of Geosciences*, *08*(04), 611–622. https://doi.org/10.4236/ijg.2017.84033

Salas, E. A. L., Subburayalu, S. K., Slater, B., Zhao, K., Bhattacharya, B., Tripathy, R., Das, A., Nigam, R., Dave, R., and Parekh, P. (2019). Mapping crop types in fragmented arable landscapes using AVIRIS-NG imagery and limited field data. *International Journal of Image and Data Fusion*, *11*(1), 33–56. https://doi.org/10.1080/19479832.2019.1706646

Salazar-espinoza, C., Jones, S., and Tarp, F. (2015). Weather shocks and cropland decisions in rural Mozambique. *Food Policy*, *53*, 9–21. https://doi.org/10.1016/j.foodpol.2015.03.003

SALGA. (2020). *O . R Tambo District Municipality: the district development model* (Issue December). https://doi.org/Available online:https://www.salga.org.za/Documents/ (Accessed on 9 July 2021)

Samberg, L. H., Gerber, J. S., Ramankutty, N., Herrero, M., and West, P. C. (2016). Subnational distribution of average farm size and smallholder contributions to global food production. *Environmental Research Letters*, *11*(12). https://doi.org/10.1088/1748-9326/11/12/124010

Santos, C. L. M. de O., Lamparelli, R. A. C., Figueiredo, G. K. D. A., Dupuy, S., Boury, J., Luciano, A. C. dos S., Torres, R. da S., and le Maire, G. (2019). Classification of crops, pastures, and tree plantations along the season with multi-sensor image time series in a subtropical agricultural region. *Remote Sensing*, *11*(3). https://doi.org/10.3390/rs11030334

Santpoort, R. (2020). The drivers of maize area expansion in sub-Saharan Africa. How policies to boost maize production overlook the interests of smallholder farmers. *Land*, *9*(3). https://doi.org/10.3390/land9030068

Sasson, A. (2012). Food security for Africa: An urgent global challenge. *Agriculture and Food Security*, *1*(1), 1–16. https://doi.org/10.1186/2048-7010-1-2

Shaban, M. (2013). Effect of water and temperature on seed germination and emergence as a seed hydrothermal time model. *International Journal of Advanced Biological and Biomedical Research*, *1*(12), 1686–1691. https://doi.org/Available online: http://www.ijabbr.com/article_7963.html (Accessed on 6 March 2018)

Sharma, P. K., Kumar, D., Srivastava, H. S., and Patel, P. (2018). Assessment of different methods for soil moisture estimation : A review. *Journal of Remote Sensing and GIS*, *9*(1), 57–73.

Sharoff, J., Diro, R. L., McCarney, G., and Norton, M. (2015). R4 Rural Resilience Initiative in Ethiopia. In *Climate Services Partnership* (pp. 1–7). Climate Services Partnership. https://doi.org/Available online: https://www.climate-services.org/wp-content/uploads/2015/09/R4_Ethiopia_Case_Study.pdf (Accessed on 18 January 2021)

Shashi Kiran, A. S., and Umesh, K. B. (2012). Crop insurance- strategy to minimize risk in agriculture. *International Association of Agricultural Economics (IAAE) Triennial Conference*, 18–24. https://doi.org/10.22004/ag.econ.126734

Shi, Z., Chu, Y., Zhang, Y., Wang, Y., and Wei, D. Q. (2021). Prediction of blood-brain barrier permeability of compounds by fusing resampling strategies and extreme gradient boosting. *IEEE Access*, *9*, 9557–9566. https://doi.org/10.1109/ACCESS.2020.3047852

Shirsath, P. B., Sehgal, V. K., and Aggarwal, P. K. (2020). Downscaling regional crop yields to local scale using remote sensing. *Agriculture*, *10*(3), 1–14. https://doi.org/10.3390/agriculture10030058

Sibanda, M, Mushunje, A., and Mutengwa, C. . (2016). Factors influencing the demand for improved maize open pollinated varieties (OPVs) by smallholder farmers in the Eastern Cape Province, South Africa. *Journal of Cereals and Oilseeds*, *7*(2), 14–26. https://doi.org/10.5897/jco2015.0142

Sibanda, Melusi, Mushunje, A., and Mutengwa, C. S. (2016). An evaluation on the profitability of growing improved maize open pollinated varieties in the Eastern Cape Province, South Africa. *Journal of Development and Agricultural Economics*, *8*(1), 1–13. https://doi.org/10.5897/jdae2014.0713

Sibiko, K. W., and Qaim, M. (2020). Weather index insurance, agricultural input use, and crop productivity in Kenya. *Food Security*, *12*(1), 151–167. https://doi.org/10.1007/s12571-019-00987-y

Sidike, P., Sagan, V., Maimaitijiang, M., Maimaitiyiming, M., Shakoor, N., Burken, J., Mockler, T., and Fritschi, F. B. (2019). dPEN: deep Progressively Expanded Network for mapping heterogeneous agricultural landscape using WorldView-3 satellite imagery. *Remote Sensing of Environment*, *221*(April 2018), 756–772. https://doi.org/10.1016/j.rse.2018.11.031

Siebert, A. (2016). Analysis of index insurance potential for adaptation to hydroclimatic risks in the west African Sahel. *Weather, Climate, and Society*, *8*(3), 265–283. https://doi.org/10.1175/WCAS-D-15-0040.1

Sims, B., and Kienzle, J. (2016). Making mechanization accessible to smallholder farmers in sub-Saharan Africa. *Environments - MDPI*, *3*(2), 1–18. https://doi.org/10.3390/environments3020011

Singh, P., and Agrawal, G. (2020). Development, present status and performance analysis of agriculture insurance schemes in India: Review of evidence. *International Journal of Social Economics*, *47*(4), 461–481. https://doi.org/10.1108/IJSE-02-2019-0119

Singo, L. R., Kundu, P., and Mathivha, F. (2016). Spatial variation of reference evapotranspiration and its influence on the hydrology of Luvuvhu River Catchment. *Research Journal of Agriculture and Environmental Management*, *5*(December 2017), 187–196. https://doi.org/Available online: https://univendspace.univen.ac.za/handle/11602/1265 (Accessed on 10 february 2022)

Sishodia, R. P., Ray, R. L., and Singh, S. K. (2020). Applications of remote sensing in precision agriculture: A Review. *Remote Sensing*, *12*(19), 3136. https://doi.org/10.3390/rs12193136

Smith, M., Kivumbi, D., and Heng, L. . (2002). Use of the FAO CROPWAT model in deficit irrigation studies. *FAO*, 17–27. https://doi.org/Available online: https://agris.fao.org/agris-search/search.do?recordID=XF2002407869 (Accessed on 6 February 2019)

Solares, C., and Sanz, A. (2007). Bayesian networks in the classification of multispectral and hyperspectral remote sensing images. *3rd WSEAS International Conference on Remote Sensing, June*, 83–86. https://doi.org/Available online: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.511.2596andrep=rep1andtype=pdf (Accessed on 16 August 2019)

Song, L., Jin, J., and He, J. (2019). Effects of severe water stress on maize growth processes in the field. *Sustainability (Switzerland)*, *11*(18). https://doi.org/10.3390/su11185086

Sonobe, R., Yamaya, Y., Tani, H., Wang, X., Kobayashi, N., and Mochizuki, K. (2018). Crop classification from Sentinel-2-derived vegetation indices using ensemble learning. *Journal of Applied Remote Sensing*, *12*(02), 1. https://doi.org/10.1117/1.jrs.12.026019

Souri, A. H., and Vajedian, S. (2015). Dust storm detection using random forests and physical-based approaches over the Middle East. *Journal of Earth System Science*, *124*(5), 1127–1141. https://doi.org/10.1007/s12040-015-0585-6

Sparling, D. W. (2016). Modeling in Ecotoxicology. *Ecotoxicology Essentials*, 361–390. https://doi.org/10.1016/b978-0-12-801947-4.00012-3

Sriwongsitanon, N., Gao, H., Savenije, H. H. G., Maekan, E., Saengsawang, S., and Thianpopirug, S. (2015). The Normalized Difference Infrared Index (NDII) as a proxy for soil moisture storage in hydrological modelling. *Hydrology and Earth System Sciences Discussions*, *12*(8), 8419–8457. https://doi.org/10.5194/hessd-12-8419-2015

Sultan, B., Defrance, D., and Iizumi, T. (2019). Evidence of crop production losses in West Africa due to historical global warming in two crop models. *Scientific Reports*, *9*(1), 1–15. https://doi.org/10.1038/s41598-019-49167-0

Syll, M., Ly, A., Diouf, I., Weingartner, L., and Mbaye, A. (2017). *Bundling weather index insurance and credit in Senegal*. https://doi.org/Available online: https://developmentevidence.3ieimpact.org/search-result-details/impact-evaluation-repository/bundling-weather-index-based-insurance-with-loans-in-senegal-evaluating-impact-on-insurance-take-up-and-farmers-well-being/8948 (Accessed on 8 March 2015)

Tadesse, M. A., Shiferaw, B. A., and Erenstein, O. (2015). Weather index insurance for managing drought risk in smallholder agriculture: lessons and policy implications for sub-Saharan Africa. *Agricultural and Food Economics*, *3*(1). https://doi.org/10.1186/s40100-015-0044-3

Tahara, M., Carver, B. F., Johnson, R. C., and Smith, E. L. (1990). Relationship between relative water content during reproductive development and winter wheat grain yield. *Euphytica*, *49*(3), 255–262. https://doi.org/10.1007/BF00036297

Tamene, L., Mponela, P., Ndengu, G., and Kihara, J. (2016). Assessment of maize yield gap and major determinant factors between smallholder farmers in the Dedza district of Malawi. *Nutrient Cycling in Agroecosystems*, *105*(3), 291–308. https://doi.org/10.1007/s10705-015-9692-7

Tarnavsky, E., Chavez, E., and Boogaard, H. (2018). Agro-meteorological risks to maize production in Tanzania: Sensitivity of an adapted Water Requirements Satisfaction Index (WRSI) model to rainfall. *International Journal of Applied Earth Observation and Geoinformation*, *73*(June), 77–87. https://doi.org/10.1016/j.jag.2018.04.008

Tarnavsky, E., Grimes, D., Maidment, R., Black, E., Allan, R. P., Stringer, M., Chadwick, R., and Kayitakire, F. (2014). Extension of the TAMSAT satellite-based rainfall monitoring over Africa and from 1983 to present. *Journal of Applied Meteorology and Climatology*, *53*(12), 2805–2822. https://doi.org/10.1175/JAMC-D-14-0016.1

Tavares, P. A., Beltrão, N. E. S., Guimarães, U. S., and Teodoro, A. C. (2019). Integration of sentinel-1 and sentinel-2 for classification and LULC mapping in the urban area of Belém, eastern Brazilian Amazon. *Sensors (Switzerland)*, *19*(5). https://doi.org/10.3390/s19051140

Tian, J., and Philpot, W. D. (2015). Relationship between surface soil water content, evaporation rate, and water absorption band depths in SWIR reflectance spectra. *Remote Sensing of Environment*, *169*, 280–289. https://doi.org/10.1016/j.rse.2015.08.007

Towery, N., Eyton, J., Stanley, A., Changnon, J., and Dailey, C. (1975). *Remote sensing of crop hail damage*. https://doi.org/Available online: https://www.ideals.illinois.edu/bitstream/handle/2142/55723/ISWSCR-165.pdf?sequence=2 (Accessed on 20 March 2018)

Towery, N. G. (1980). *Some applications of remote sensing of crop-hail damage in the insurance industry*. https://doi.org/Available online: https://www.isws.illinois.edu/pubdoc/C/ISWSC-143.pdf (Accessed on 20 March 2018)

Udom, B. E., and Kamalu, O. J. (2019). Crop water requirements during growth period of maize (Zea mays L.) in a moderate permeability soil on coastal Plain sands. *International Journal of Plant Research*, *2019*(1), 1–7. https://doi.org/10.5923/j.plant.20190901.01

UNCTAD. (2015). The role of smallholder farmers in sustainable commodities production and trade. In *United nations conference on trade and development*. United Nations Conference on Trade and Development.

USDA. (2018). Commodity intelligence report. In *United States Department of Agriculture Foreign Agricultural Service Report*. https://doi.org/Available online: https://ipad.fas.usda.gov/highlights/2012/08/Mexico_corn/ (Accessed on 13 June 2020)

Useya, J., and Chen, S. (2018). Comparative Performance Evaluation of Pixel-Level and Decision-Level Data Fusion of Landsat 8 OLI, Landsat 7 ETM+ and Sentinel-2 MSI for Crop Ensemble Classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*(11), 4441–4451. https://doi.org/10.1109/JSTARS.2018.2870650

Useya, J., and Chen, S. (2019). Exploring the Potential of Mapping Cropping Patterns on Smallholder Scale Croplands Using Sentinel-1 SAR Data. *Chinese Geographical Science*, *29*(4), 626–639. https://doi.org/10.1007/s11769-019-1060-0

Valverde-Arias, O., Garrido, A., Saa-Requejo, A., Carreño, F., and Tarquis, A. M. (2019). Agro-ecological variability effects on an index-based insurance design for extreme events. *Geoderma*, *337*(March 2018), 1341–1350. https://doi.org/10.1016/j.geoderma.2018.10.043

van Asten, P. J. A., Fermont, A. M., and Taulya, G. (2011). Drought is a major yield loss factor for rainfed East African highland banana. *Agricultural Water Management*, *98*(4), 541–552. https://doi.org/10.1016/j.agwat.2010.10.005

van Averbeke, W., Ralivhesa, K., Mbuli, S., Khosa, T. B., and Manyelo, K. W. (2013). *Growing green maize on canal schemes in Vhembe: production guidelines*. https://doi.org/Available online: http://www.wrc.org.za/wp-content/uploads/mdocs/TT 567-13.pdf (Accessed on 17 September 2020)

Van Loon, J., Woltering, L., Krupnik, T. J., Baudron, F., Boa, M., and Govaerts, B. (2020). Scaling agricultural mechanization services in smallholder farming systems: Case studies from sub-Saharan Africa, South Asia, and Latin America. *Agricultural Systems*, *180*. https://doi.org/10.1016/j.agsy.2020.102792

Van Tricht, K., Gobin, A., Gilliams, S., and Piccard, I. (2018). Synergistic use of radar sentinel-1 and optical sentinel-2 imagery for crop mapping: A case study for Belgium. *Remote Sensing*, *10*(10), 1–22. https://doi.org/10.3390/rs10101642

Von Fintel, D., and Pienaar, L. (2016). Small-scale farming and food security: The enabling role of cash transfers in South Africa’s former homelands. *IZA Discussion Paper*, *10377*(November). https://doi.org/10.22004/ag.econ.211916

von Negenborn, F., Weber, R., and Musshoff, O. (2018). Explaining weather-related credit risk with evapotranspiration and precipitation indices. *Agricultural Finance Review*, *78*(2), 246–261. https://doi.org/10.1108/AFR-07-2017-0058

Vreugdenhil, M., Wagner, W., Bauer-Marschallinger, B., Pfeil, I., Teubner, I., Rüdiger, C., and Strauss, P. (2018). Sensitivity of Sentinel-1 backscatter to vegetation dynamics: An Austrian case study. *Remote Sensing*, *10*(9), 1–19. https://doi.org/10.3390/rs10091396

Wang, H. H., Liu, L., Ortega, D. L., Jiang, Y., and Zheng, Q. (2020). Are smallholder farmers willing to pay for different types of crop insurance? An application of labelled choice experiments to Chinese corn growers. *Geneva Papers on Risk and Insurance: Issues and Practice*, *45*(1), 86–110. https://doi.org/10.1057/s41288-019-00153-7

Wang, L., Qu, J. J., Hao, X., and Zhu, Q. (2008). Sensitivity studies of the moisture effects on MODIS SWIR reflectance and vegetation water indices. *International Journal of Remote Sensing*, *29*(24), 7065–7075. https://doi.org/10.1080/01431160802226034

Wani, S. P., Sreedevi, T. K., Rockström, J., and Ramakrishna, Y. S. (2009). Rainfed agriculture - past trends and future prospects. In S. P. Wani, T. K. Sreedevi, J. Rockström, and Y. S. Ramakrishna (Eds.), *Rainfed Agriculture: Unlocking the potential* (1st Editio, pp. 1–35). CABI. https://doi.org/10.1079/9781845933890.0001

Welikhe, P., Quansah, J. E., Fall, S., and McElhenney, W. (2017). Estimation of soil moisture percentage using Landsat-based moisture stress index. *Journal of Remote Sensing and GIS*, *06*(02). https://doi.org/10.4172/2469-4134.1000200

WFP. (2020). *R4 Rural resilience initiative: Annual report*. https://doi.org/Available online: https://www.wfp.org/publications/r4-rural-resilience-initiative-2020-annual-report (Accessed on 4 February 2021)

Whelen, T., and Siqueira, P. (2018). Time-series classification of Sentinel-1 agricultural data over North Dakota. *Remote Sensing Letters*, *9*(5), 411–420. https://doi.org/10.1080/2150704X.2018.1430393

Will, M., Groeneveld, J., Frank, K., and Muller, B. (2021). Informal risk-sharing between smallholders may be threatened by formal insurance: Lessons from a stylized agent-based model. *PLoS ONE*, *16*(3 March), 1–18. https://doi.org/https://doi.org/10.1371/journal.pone.0248757

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

World Bank. (2011). Weather index insurance for agriculture: guidance for development practitioners. In *Agriculture and rural development discussion paper* (Issue 50). https://doi.org/Available online:

https://documents1.worldbank.org/curated/en/590721468155130451/pdf/662740NWP0Box30or0Ag020110final0web.pdf (Accessed on 28 April 2019)

World Bank. (2018). Overcoming poverty and inequality in South Africa. In *World Bank Publications: Vol. World Bank*. https://doi.org/Available online: http://documents.worldbank.org/curated/en/530481521735906534/pdf/124521-REV-OUO-South-Africa-Poverty-and-Inequality-Assessment-Report-2018-FINAL-WEB.pdf (Accessed on 3 June 2019)

Worldbank. (2007). Actual crop water use in project countries - A synthesis at the regional level. In *Policy Research Working Paper* (No. 2488; Vol. 4288). https://doi.org/Available online: https://elibrary.worldbank.org/doi/abs/10.1596/1813-9450-4288 (Accessed on 10 February 2022)

Worldbank. (2011). Weather index insurance for agriculture: guidance for development practitioners. In *Agriculture and rural development discussion paper* (No. 50; Issue 50). https://doi.org/Available online: https://documents1.worldbank.org/curated/en/590721468155130451/pdf/662740NWP0Box30or0Ag020110final0web.pdf (Accessed on 11 February 2022)

Wu, Q., Wang, L., and Wu, J. (2012). Ensemble learning on hyperspectral remote sensing image classification. *Advanced Materials Research*, *546–547*, 508–513. https://doi.org/10.4028/www.scientific.net/AMR.546-547.508

Xia, W., Ma, C., Liu, J., Liu, S., Chen, F., Yang, Z., and Duan, J. (2019). High-resolution remote sensing imagery classification of imbalanced data using multistage sampling method and deep neural networks. *Remote Sensing*, *11*(21). https://doi.org/10.3390/rs11212523

Xu, C., Qu, J. J., Hao, X., and Wu, D. (2020). Monitoring surface soil moisture content over the vegetated area by integrating optical and SAR satellite observations in the permafrost region of Tibetan Plateau. *Remote Sensing*, *12*(1). https://doi.org/10.3390/RS12010183

Xu, L., Zhang, H., Wang, C., Zhang, B., and Liu, M. (2019). Crop classification based on temporal information using Sentinel-1 SAR time-series data. *Remote Sensing*, *11*(1), 1–18. https://doi.org/10.3390/rs11010053

Xue, J., and Su, B. (2017). Significant remote sensing vegetation indices: A review of developments and applications. *Journal of Sensors*, *2017*. https://doi.org/10.1155/2017/1353691

Yang, C.-M., Fan, M., and Hsiang, W.-M. (1993). Growth and yield respones of maize (Zea mays L.) to soil water Deficits. II. Effects of water deficit timing and strength. *Jour. Agric. Res. China*, *42*(2), 173–186. https://doi.org/10.29951/JARC.199306.0008

Ye, T., Hu, W., Barnett, B. J., Wang, J., and Gao, Y. (2020). Area Yield Index Insurance or Farm Yield Crop Insurance? Chinese Perspectives on Farmers' Welfare and Government Subsidy Effectiveness. *Journal of Agricultural Economics*, *71*(1), 144–164. https://doi.org/10.1111/1477-9552.12326

Zendera, W., Mhlanga, S., and Mutetwa, M. (2018). Weather index based crop insurance using Artificial Neural Networks. *Elixir Agriculture*, *119*, 51107–51113.

Zerizghy, M. G., van Rensburg, L. D., and Anderson, J. J. (2013). Comparison of neutron scattering and DFM capacitance instruments in measuring soil water evaporation. *Water SA*, *39*(2), 183–190. https://doi.org/10.4314/wsa.v39i2.1

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

Zhang, H., Chang, J., Zhang, L., Wang, Y., Li, Y., and Wang, X. (2018). NDVI dynamic changes and their relationship with meteorological factors and soil moisture. *Environmental Earth Sciences*, *77*(16), 1–11. https://doi.org/10.1007/s12665-018-7759-x

Zhang, L., Meng, Q., Yao, S., Wang, Q., Zeng, J., Zhao, S., and Ma, J. (2018). Soil moisture retrieval from the chinese GF-3 satellite and optical data over agricultural fields. *Sensors (Switzerland)*, *18*(8). https://doi.org/10.3390/s18082675

Zhang, T. X., Su, J. Y., Liu, C. J., and Chen, W. H. (2019). Potential bands of Sentinel-2A satellite for classification problems in precision agriculture. *International Journal of Automation and Computing*, *16*(1), 16–26. https://doi.org/10.1007/s11633-018-1143-x

Zhou, T., Pan, J., Zhang, P., Wei, S., and Han, T. (2017). Mapping winter wheat with multi-temporal SAR and optical images in an urban agricultural region. *Sensors (Switzerland)*, *17*(6), 1–16. https://doi.org/10.3390/s17061210

Zurita-Milla, R., Izquierdo-Verdiguier, E., and De By, R. A. (2017). Identifying crops in smallholder farms using time series of WorldView-2 images. *2017 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images, MultiTemp 2017*, 17–19. https://doi.org/10.1109/Multi-Temp.2017.8035246

APPENDIX

ETHICAL CLEARANCE

University of Fort Hare
Together in Excellence

ETHICS CLEARANCE REC-270710-028-RA Level 01

Project Number:	KAL021SMAS01
Project title:	Investigating the feasibility of using remote sensing in index-based crop insurance for South Africa's smalholder farming.
Qualification:	PhD in Geography
Student name:	Wonga Masiza
Registration number	200901062
Supervisor:	Dr A.M Kalumba
Department:	Geography
Co-supervisor:	Dr H.B Magagula Dr J.G Chirima

On behalf of the University of Fort Hare's Research Ethics Committee (UREC) I hereby grant ethics approval for KAL021SMAS01. This approval is valid for 12 months from the date of approval. Renewal of approval must be applied for BEFOREtermination of this approval period. Renewal is subject to receipt of a satisfactory progress report. The approval covers the undertakings contained in the above-mentioned project and research instrument(s). The research may commence as fromthe 15/06/21, using the reference number indicated above.

Note that should any other instruments be required or amendments become necessary, these require separate authorisation.
Please note that UREC must be informed immediately of

- Any material changes in the conditions or undertakings mentioned in the document;
- Any material breaches of ethical undertakings or events that impact upon theethical conduct of the research.

The student must report to the UREC in the prescribed format, where applicable, annually, and at the end of the project, in respect of ethical compliance.

UREC retains the right to

- Withdraw or amend this approval if
    - Any unethical principal or practices are revealed or suspected;
    - Relevant information has been withheld or misrepresented;
    - Regulatory changes of whatsoever nature so require;
    - The conditions contained in the Certificate have not been adhered to.

- Request access to any information or data at any time during the course orafter completion of the project.

Your compliance with Department of Health 2015 guidelines and any other applicable regulatory instruments and with UREC ethics requirements as containedin UREC policies and standard operating procedures, is implied.

UREC wishes you well in your research.

Yours sincerely

Dr N Taole-Mjimba

Chairperson: University Research Ethics Committee

29 July 2021

University of Fort Hare
Together in Excellence