Assessing the forecast skill of agricultural drought forecast from satellite-derived products in the Lower Shire River Basin

Marijke Panis

Thesis committee

Dr. ir. H.C.W. (Hessel) Winsemius	TU Delft
Prof. dr. ir. P.H.A.J.M. (Pieter) van Gelder	TU Delft
Dr. ir. G.H.W. (Gerrit) Schoups	TU Delft
Dr. M. (Marc) van den Homberg	510
Ir. Aklilu Teklesadik	510

TU Delft

510 AN INITIATIVE OF THE NETHERLANDS RED CROSS

Assessing the forecast skill of agricultural drought forecast from satellite-derived products in the Lower Shire River Basin

by

M. Panis

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Tuesday March 19, 2019 at 10:00 AM.

Student number:	4107578	
Project duration:	April 24, 2018 – March 19, 2019	
Thesis committee:	Dr. ir. H.C.W. (Hessel) Winsemius,	Delft University of Technology
	Prof. dr. ir. P.H.A.J.M. (Pieter) van Gelder,	Delft University of Technology
	Dr. ir. G. H. W. (Gerrit) Schoups,	Delft University of Technology
	Dr. M. (Marc) van den Homberg,	510
	Ir. (Aklilu) Teklesadik,	510
Extra support	MSc. Mendy van der Vliet,	VanderSat
	Dr. Narcisa Nechita ,	VanderSat

Preface

This Master thesis is part of the MSc Watermanagement at Delft University of Technology. This research is carried out by 510 an initiative of the Netherlands Red Cross and Delft University of Technology, being part of two funded projects, The UK's Department for International Development funds the NERC-SHEAR[1] project *Improving Preparedness to Agro-Climatic Extremes in Malawi (IPACE-Malawi)* and the Forecast-based Financing (FbF) project funded by EU ECHO[2] II *Enhancing Resilience program in Malawi.* As part of my studies, I had the opportunity to work on a variety of (international) projects, which inspired me to combine my technical background with other disciplines. By reading, talking and discussing global water challenges I realised that water-related challenges are an essential factor in a lot of conflicts and in vulnerable areas which inspired me to use my technical background in a humanitarian context. This ambition is supported by 510 and the TU Delft, for which I'm very grateful.

First of all, I would like to thank Marc van den Homberg and Aklilu Teklesadik for giving me the opportunity to work with 510 and taking part in the Forecast-based Financing project. The variety within this project kept me motivated throughout the whole research. Within this research, I had the pleasure of seeing the whole process of implementing a forecast model. It was a great challenge combining the qualitative research of expert interviews and focus group discussions with the quantitative data analysis and a hydrological approach. The fieldwork gave a lot of inspiration and perspective in the essence of the research, while the data analysis and developing the forecast model gave me insights in a new field of research.

I'm sincerely thankful for the support of my TU Supervisors Hessel Winsemius, Pieter van Gelder and Gerrit Schoups. From the start, Hessel showed his enthusiasm in my ideas, research topic and the collaboration with 510. He was always willing to support me whenever needed. I valued the progress meetings together with Pieter van Gelder and Gerrit Schoups, who also shared their thoughts, comments, giving direction, and showing their interest. At the end of the meetings, I got even more energy and motivation to continue the work I was doing.

Furthermore, I would like to thank VanderSat, in particular, Mendy van der Vliet and Narcisa Nechita for giving me a sneak peek into the field of Remote Sensing. They helped me with creating an understanding in the satellite-product and implementing their data into my reserach,

Besides my Thesis Committee and extra support of VanderSat I had the pleasure of meeting a lot of inspiring people who are working in the same research area of using forecast systems for humanitarian purposes. Already at the beginning of this thesis project I got in contact with Ted Veldkamp and Gabriela Guimarães Nobre from the Vrije Universiteit Amsterdam, they welcomed me at their department and introduced me the Machine Learning Algorithm; Fast-and Frugal decision trees in drought forecast. Later on, Gabriela shared here knowledge working with this algorithm and supported me with the development of the drought forecast model, for which im very thankful.

This research includes a field visit to Malawi in collaboration with the University of Leeds. Therefore, I would like to thank Stephen Whitefield, Neha Mittal and Marta Bruno Soares from the University University of Leeds for this opportunity, the discussions, support and interest in my research. The time in Malawi gave me perspective, inspiration and food for thought in further elaborating my research. In Malawi, the Malawi Red Cross Society welcomed us with open arms, sharing their knowledge and introducing us to this wonderful country. The time in Malawi would not be this great experience without Gumbi Gumbi who joined me to the South of Malawi into the field.
Furthermore, I would like to thank the rest of the Malawi Red Cross Society for their contribution to this field

[1] Science for Humanitarian Emergencies & Resilience (SHEAR) project, Natural Environment Research Council (NERC) and the Economic & Social Research Council. This project is a collaboration between the University of Leeds, 510, Met Office and Lilongwe University of Agriculture & Natural Resources (LUANAR)
[2] European Civil Protection and Humanitarian Aid Operations (ECHO)

trip. Last, but not least, I will never forget Agatha Bucherie aka Nakabango for all the ’gezelligheid’ and motivation. Working together, during our time abroad, sharing our research struggles and funny moments.

All in all, I’m happy with the whole thesis trajectory; it gave me the realisation I would like to continue working in this area and contribute to developing these type of forecast models. I want to thank all the people who were open, enthusiastic and willing to take part in this research for interviews and discussions.

Additionally to the substantive support, I would like to thank my colleagues who showed their interest and made me feel welcome. In particular, I will thank Joris Westerveld who took the time to help me with my first steps in Machine Learning and Data Science. Furthermore, I would like to thank Dominik Semet with whom I started this research, combining our interviews in the qualitative research phase. Moreover, i’m grateful for the support of my friends Steffi Scheringa helping me with the preparation of the focus group discussions and the finishing touch of the cover and Kim Stolk who helped me with the design of the graphs.

Last, A special thanks to my family and friends who support me with every step I take. The last few years they spend hours listening to my crazy ideas, frustrations and challenges I had to face. My parents and sister with whom I shared my frustrations and best moments, they gave me the motivation and distraction when needed. My friends for always being there!

M. Panis
Delft, March 2019

Abstract

Worldwide 136 million people depend on humanitarian aid. Those people are either affected by conflicts or weather-related crises. Although the main driver of humanitarian aid is conflicts, natural disasters increase the number of people in need. The number of people affected by natural disasters is expected to increase with another 5% relative to the 136 million people in need in 2017. Which asks for more funding than ever before. To support these people, the humanitarian organisations are forced to respond more effectively and efficiently.

In 2008 the Red Cross Red Crescent (RCRC) started with Forecast-based Financing pilots to improve existing Early-Warning Early Action systems. Forecast-based financing is a new methodology to prepare, deliver and respond more effectively and efficiently, based on hazard forecasts. Actions are triggered when a forecast exceeds a danger level in a vulnerable intervention area. Forecast-based financing consists of several implementation steps, of which the first three aim at impact-based forecasting.

510, the data initiative of the Netherlands Red Cross and the Malawi Red Cross work on impact-based forecasting as part of several projects in Malawi, such as the EU ECHO II project and the Improving Preparedness to Agro-Climatic Extremes in Malawi (IPACE) project, which aims to: (1) identify critical agro-climatic indicators in central and southern Malawi; (2) test the skill of short term to seasonal forecast tools in simulating these indicators; and (3) co-design agricultural climate services and input into early warning early action systems based on these indicators/forecast tools. In the case-study area of Malawi droughts are the natural disaster leading to high agricultural damages and losses affecting the livelihood of smallholder farmers.

Floods, droughts and tropical storms are the main natural hazards affecting the agricultural sector. The impact of drought on agriculture is more severe than floods, due to the long duration and large spatial area influencing crop production and indirectly the livelihood of smallholder farmers. Although the impact of droughts is more significant than the impact of floods, drought impact is less developed and understood. The problem with drought is its complexity and slow onset with an increasing exponential impact. Another difficulty in drought analysis is the terminology, with four main drought types and definitions, either they look at precipitation deficits (meteorological drought), below-average water levels (hydrological drought), soil moisture water content (agricultural drought) or water demand for economic goods (socio-economical drought). In Malawi, smallholder farmers are highly dependent on rain-fed agriculture influenced by extreme weather conditions, for which agriculture drought is most severe in Malawi. Current agricultural drought forecasting procedures make use of mainly meteorological indices and therefore limited in their forecast skill. Agricultural droughts start when the available soil moisture content for plants drops to a certain level negatively affecting crop yield, making soil moisture the key driver of agricultural drought identification. Although soil moisture content is the primary variable in identifying agricultural drought, this indicator is rarely taken into account in current procedures.

Therefore, this study focuses on how the forecast skill of agricultural drought forecast can be improved. More specifically, the aim is to identify the contribution of machine learning and satellite-derived products on improving the forecast skill of agricultural drought forecasts in early warning early action systems in the Lower Shire River Basin in Malawi. The data on this set of predictors and predictands is obtained via exploratory sequential mixed methods, combining both qualitative and quantitative research. The methods used in this research are: desk research (open data available on geospatial data sharing platforms), semi-structured interviews, focus group discussions in Malawi and data analysis from satellite data company VanderSat. The final set of predictors and predictands is narrowed down based on which data is available and with which quality (timeliness, reliability, accuracy). Essential is also that both predictors and predictands are available at the same granularity. The machine learning model used is the Fast and Frugal decision tree, which is selected based on its simplicity and transparency. These characteristics are useful when a decision needs to be quickly understood and implemented.

The Fast-and Frugal decision tree is used to identify the combination and the optimal trigger threshold values corresponding to a chosen set of indicators. The data in the Fast and Frugal decision tree covers the growing season from November-March over the 2002-2017 period. As predictors, the following agro-climatic indices is used: cumulative precipitation, soil moisture anomalies, land surface temperature anomalies, El Niño Southern Oscillation in July and dry spells with a sequence of 3-4 dry days. As drought predictand, the normalised difference vegetation index (NDVI) and the vegetation optical depth (VOD) in March is used. Initial results show that the best combination of agro-climatic indicators depends on the lead time and spatial characteristics of the area. The crucial month (LT5: November and LT3: Janaury) in the growing season for drought forecast are used in the analysis. At the start of the growing season, initial states are essential, such as soil moisture, land surface temperature and precipitation. Throughout the season the vegetation indicators are increasingly important, whereas, mid-season dry spells return as an essential predictor. At lead time 3 the model performance shows a sudden drop for most areas, this drop could be interpreted as a signal that conditions change influencing drought prediction or they relate to noise in data. Machine learning techniques benefit from large datasets, for which this research should be repeated with a longer time serie. The best predictand at LT3 is VOD and at LT5 NDVI.

The outcome of the model supports humanitarian organisations to identify the lead time necessary to act upon a drought trigger and reduce the impact of such an event. The machine learning forecasting model cannot be used directly to provide forecasts at different intermittent timescales throughout the dry season; however, the predictors used to provide pointers for the agro-climatic indices will be used in the NERC SHEAR IPACE project to develop (mobile) climate services for farmers.

Contents

List of Figures

List of Tables

Acronyms

Acronyms	
ADD	Agricultural Development Division
AEDC	Agriculture Extension Development Coordinator
DADO	District Agricultural Development Officer
DAES	Department of Agricultural Extension Services
DCCMS	Department for Climate Change and Meteorological Services
DoDMA	Department of Disaster Management Affairs
DRR	Disaster Risk Reduction
DS	Dry spell
EAP	Early Action Protocol
ECHO	European Commission's Civil Protection and Humanitarian Aid Operations
ENSO	El Niño Southern Oscillation
EPA	Extention Planning Officer
ESA	European Space Agency
EU	European Unicon
EWEA	Early warning early action
EWS	Early - Warning System
FAO	Food and Agriculture Organization of the United Nations
FbF	Forecast-based Financing
FEWS NET	Famine Early Warning Systems Network
FGD	Focus Group Discussion
GDP	Gross Domestic Product
GoM	Government of Malawi
GVH	Group Village Head
IFRC	International Federation of Red Cross and Red Crsecent Societies
IPACE-Malawi	Improving Preparedness to Agro-Climatic Extremes in Malawi
ISOs	Intra-Seasonal Oscillations
ITCZ	Inter-Tropical Convergence Zone
LST	Land surface temperature
MLRC	Malawi Red Cros Society
NDVI	Normalised Difference Vegetation Index
NGO	Non-governmental organisation
ODSS	Operational Decision Support System
ONI	Oceanic Niño Index
RCRC	Red Cross Red Crescent
SHEAR	Science for Humanitarian Emergencies & Resilience
SM	Soil Moisture
SOP	Standard Operating Procedures
SPI	standardized precipitation index
SST	Sea Surface temperature
TA	Traditional Authority
USDM	US Drought Monitor
VdS	VanderSat
VOD	Vegetation Optical Depth

1
Introduction

Worldwide 136 million people depend on humanitarian aid. Those people are either affected by conflicts or weather-related crises. Although the main driver of humanitarian aid is conflicts, natural disaster increase the number of people in need, this increasing number of people in need enlarge the demand for humanitarian aid (Lowcock, 2018). This chapter first describes the necessitate of Early Warning, Early Action (EWEA) and Forecast-based Financing (FbF) concentrated on drought. In the case-study area of Malawi, droughts are leading to high agricultural damages and losses affecting the livelihood of smallholder farmers.

1.1. Introduction

The expectancy of the affected people by natural disasters in 2018 will increase even more with 5% relative to the number of people in need in 2017 (136 million people), which asks for more funding than ever before. Nowadays, the financial support from humanitarian organisations is only sufficient for 91 million people, resulting in a funding gap of 22 billion US dollar. In order to support these people in need, humanitarian agencies are forced to respond more effectively and efficiently (Lowcock, 2018).

Humanitarian aid aims to relieve suffering, save lives and to preserve human dignity in crises situations. In order to do so humanitarian organisations assist people with basic needs as food, shelter, clothing, water supplies, sanitation, protection and medical care when needed. This humanitarian assistance is coordinated by the European Commission's Civil Protection and Humanitarian Aid Operations (ECHO) funding UN agencies, international organisations and NGO's. EU ECHO provides humanitarian aid to 110 countries since 1992 to respond and mitigate disasters. Traditionally aid provided by humanitarian organisations was released after a disaster hits, today the release acts as a pre-disaster response, there is an increasing need for risk reduction and disaster preparedness after the Hyogo Framework for Action was signed in 2005 (Field et al., 2012).

The strategies and policies to reduce the impact of natural hazards fall under the term Disaster Risk Management (DRM). DRM exists of four phases illustrated in figure 1.1. *Mitigation* limit the adverse impacts of hazards and related disasters, *Preparedness* effectively anticipate, respond to and recover from the impacts of a natural hazard, *Response* provision of emergency services and public assistance during or immediately after a disaster and *Recovery* restoration and improvement of facilities, livelihoods and living conditions of disaster-affected communities (UNISDR, 2009). Drought early action takes place in the mitigation and preparedness phase before the disaster hit. The last few decades the focus shifted from the response phase to preparedness and mitigation asking for collaboration between forecasters and humanitarian decision makers. At the moment only 12 % of the total humanitarian budget is invested in risk reduction and disaster preparedness, reaching 24 million people in 2016 ((European Commission), (Enenkel et al., 2016) & (European Commission, 2018)). When a disaster hits the impact is not only related to the event itself, but also the vulnerability and the exposure of the affected area. DRM aims to reduce this vulnerability and exposure and at the same time to increase the resilience of extreme climate events (Field et al., 2012).

Risk = Exposure x Hazard x Vulnerability

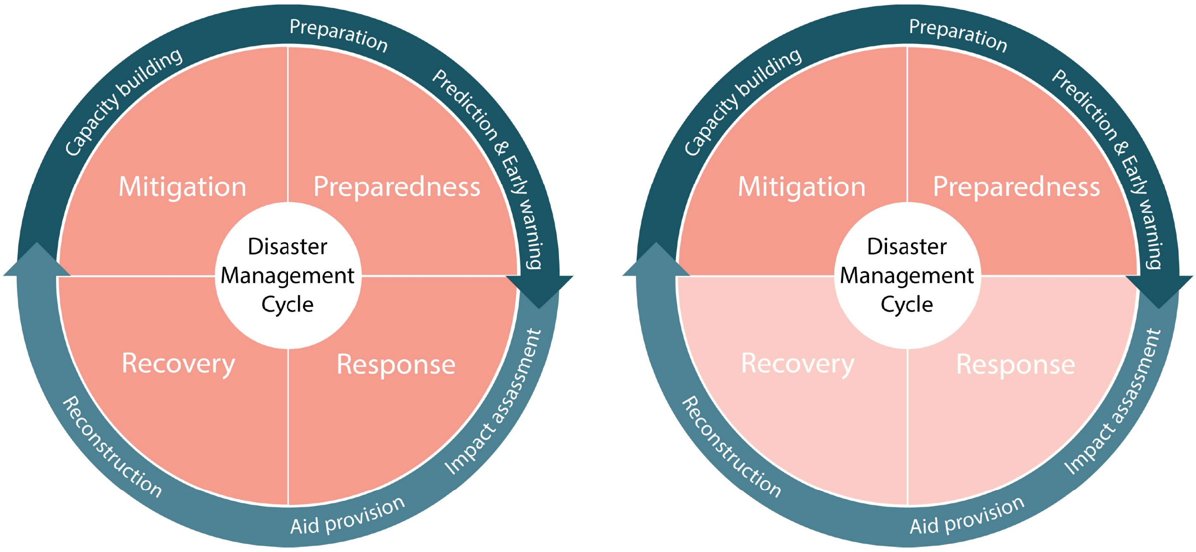


Figure 1.1: Disaster Management Cycle (design Elise Haak (2017) based on (Stojic, 2013))

1.2. Early Warning System and Forecast-based Financing

In order to respond in an early phase of disaster management, humanitarian organisations started with an EWEA approach. Climate-related disasters will cause an increasing number of people in need asking for a system to use financial resources more effectively. Operational forecast systems are used to estimate the probability of a natural disaster to happen in a certain time phase. The lead time of forecast can be in the order of years, seasonal, monthly, weekly or daily depending on the type of natural hazard and the type of forecast. Although a forecast predicted many natural disasters before they hit, the actions taken to reduce the impact stayed limited ((Rüth et al., 2017b), (IFRC, b)). This lead time between forecasts and the extreme events give humanitarian aid the opportunity to start EWEA and reduce impact (Coughlan De Perez et al., 2015). EWEA is defined by the International Federation of Red Cross Red Crescent Societies (IFRC) as *"routinely taking action before a disaster or health emergency happens, making full use of scientific information on all time scales"* (Rüth et al., 2017a). Nowadays, the forecast is improving and became more accurate by the implementation of science and technologies, such as computer models, expert meeting, information from national meteorological offices, field reports and local communities. Combining this knowledge results in better disaster preparedness, mitigation and anticipation to climate-related hazards (IFRC, 2008).

In 2008 the Red Cross Red Crescent (RCRC) started with FbF pilots to improve existing EWEA systems. FbF is a new methodology to prepare, deliver and respond more effectively and efficiently, based on hazard forecasts. Actions are triggered when a forecast exceeds a danger level in a vulnerable intervention area. FbF consists of several implementation steps, of which the first three aim at impact-based forecasting (IFRC, a). This approach contributes to several global goals set by the Sustainable Development Goals (SDGs) with the priority of investment in water and agriculture in climate change adaptation 2018, Sendai Framework of Disaster Risk Reduction, the World Humanitarian Summit, the Paris Agreement and the Agenda for Humanity. FbF does not only act as a disaster risk reduction and preparedness strategy; it also acts as a climate change adaptation strategy ((IFRC, c), (Rüth, 2015)). Important in FbF are the pre-agreed activities and roles of involved stakeholders allocated when a trigger is activated by reaching a certain threshold. The Early Action Protocols (EAPs) document these responsibilities and actions. The two main components of FbF are: the development of Standard Operating Procedures (SOPs/EAPs) that link specific forecast triggers to specific early actions and secondly, committing resources necessary to implement those actions when a triggering forecast is issued illustrated in figure (1.2) (IFRC, b). The success of FbF depends on the steps taken to implement this system: understanding risk scenarios, identify available forecasts, formulating early actions, identifying danger levels, creating a Standard Operating Procedure or early action guidelines and validate SOP with key actors (figure 1.2). Experiences of FbF pilots suggest to implemented FbF within a long-term structure coordinated by humanitarian actors and national authorities, and the FbF should be dynamic and give the ability to make adjustments over time (Rüth et al., 2017b).

Figure 1.2: Key aspects Forecast-based Financing (German Red Cross, 2016)

1.3. Problem Statement

The main natural hazards affecting the agricultural sector are floods, droughts and tropical storms. From these hazards, drought have the most impact on agriculture compared to floods, due to the long duration and large spatial area influencing crop production and indirectly the whole economy (Asa Giertz and German, 2015). Although the impact of the drought is more significant than the impact of floods, drought impact is less developed and understood (Stephens et al., 2015). The main problem with drought is its complexity and slow onset with an increasing exponential of impact ((Mishra and Singh, 2010)), (DoDMA, 2015))). Another difficulty in drought analysis is the terminology, with four main drought types and definitions, either they look at precipitation deficits (meteorological), below-average water levels (hydrological), soil moisture water content (agricultural) or water demand for economic goods (socio-economical) (Chapter 2).

From the different drought types, agriculture drought is most severe in Malawi. Smallholder farmers are highly dependent on rain-fed agriculture influenced by extreme weather conditions. Agricultural droughts start when the available soil moisture content for plants drops to a certain level negatively affecting crop yield, making soil moisture the critical driver of agricultural drought identification. Although soil moisture content is the primary variable in identifying agricultural drought, this indicator is rarely taken into account (Martínez-Fernández et al., 2016). Soil moisture is an important indicator because it balances water and energy fluxes between the atmosphere and land surface and creates interaction between soil, vegetation and climate. There are multiple ways to measure the soil moisture content, such as *In situ* measurements, using satellite-derived data and hydrological modelling. Combining *in situ* ground data and satellite imagery enhances knowledge of climate and vegetation variability, important in drought forecast (Brocca et al., 2011). In data scares regions remote sensing can be of added value in addition to *in situ* measurements. It provides the scientific community with many, precise and consistent spatial and temporal information (Scaini et al., 2015). Soil moisture satellites Soil Moisture and Ocean Salinity (SMOS) from ESA and Soil Moisture Active Passive (SMAP) from NASA gives an opportunity to implement remote sensing in agricultural drought monitoring and forecast. Meteorological and hydrological data most commonly classify agricultural drought and does not take into account variables, such as soil water content (Martínez-Fernández et al., 2016). Adding soil moisture as an indicator in drought predictions can increase the forecast skill of predicting agricultural drought. In drought predictions, the main challenge according to Rosso 2018 is detecting the severity and probability of drought based on limited data. Analysing the shift of climate parameters from normal conditions is used to identify the severity of a drought. Other aspects involved in the complexity of drought predictions are; limited data, single-value predictions and uncertainties over long lead times (Chen et al., 2012).

As Mylne (2002) mentioned: *"The concept behind forecast value is that forecasts only have value if a user takes action as a result, and the action saves the user money"*. This value of forecast requires information about the ability to predict a particular event and the relation between the result of a forecast and the costs-loss ratio when acting on the prediction. Identifying this relationship between the forecast and impact by developing danger levels is essential in EWEA systems. ((Lopez et al., 2018),(Coughlan De Perez et al., 2015)). This relation between using forecast information to reduce cost is one of the advantages of FbF. In FbF, there is an agreed threshold value to reduce avoidable losses and potential expected losses. Forecast based on threshold values is not that commonly used in the humanitarian sector (Lopez et al., 2018). In early warning drought forecast, there are several challenges to overcomes. One of the main barriers is the forecast uncertainty and data uncertainty with the consequence of 'acting in vain' and the confidant of determining if an action is worth taking. Another barrier is fund release in long-term project agreements. A third barrier is related to data availability, data sharing and data quality. As an example, in Malawi, data networks are inadequate in density, long-term records, data gaps, temporal and spatial hazard information and reliability on a seasonal timescale. Furthermore, government agencies and research institutions do not share all available data, and the data is often too technical and detailed for decision makers for direct implementation (Kita, 2017). These barriers translate into three challenges related to data collection, the forecast itself and the decision-making process.

1.4. Aim of this research

The objective of this research is to assess the forecast skill of agricultural drought forecasting from satellite-derived products at multiple lead time before the end of the growing season using a variety of agro-climatic indices. This section will present the main goal of the research and the questions to be answered.

1.4.1. Research Questions

The main research question is developed to overcome and create an understanding in the three barriers; data collection, the forecast model and the decision-making process. Together with the barriers, the objectives and challenges mentioned in the previous paragraphs result in the following research question:

> ***To what extent can machine learning and satellite-derived products be used to obtain agricultural drought forecasts with high forecast skill in early warning early action?***

This research question relates to a broad objective of investigating the added value of using machine learning techniques with satellite-derived products in agricultural drought forecast applied to the case-study area of the Lower Shire River Basin in Malawi. The answer to this question will follow from answering two sub-questions. The first sub-question focuses on creating an understanding of the information processing behaviour of the humanitarian sector and smallholder farmers. Whereas the second sub-question focuses on the agricultural drought forecast itself, developing a machine learning model, which identify the best combination of agro-climatic indicators in agricultural drought forecast. Overarching in this research is the practical implementation of a drought forecast model to reduce drought impact.

1. **How do humanitarian actors and smallholders farmers currently process information on drought forecast?**
   (a) What actors are involved in current drought forecast procedures in Malawi?
   (b) Which decision-making factors define the information processing behaviour of humanitarian organisations and smallholder farmers?
   (c) What spatial and temporal resolution is necessary for the optimisation of drought forecast at national- and smallholders farmer scale?
2. **What combination of agro-climatic features perform best in agricultural drought forecast in the Lower Shire River Basin?**
   (a) Which feature combination perform best in drought prediction at different spatial and temporal resolutions in the growing season?
   (b) How do the different model scenarios perform in terms of accuracy?

1.5. Thesis Outline

This report consists of seven chapters, the first two chapters (chapter 2 and 3) provide the reader with theoretical background information on the definition of drought, drought indicators, the principle of forecast models and introduce Malawi as case-study area. Chapter 4 discuss the methodology used in this research, followed by the results presented in chapter 5 and 6. Chapter 5 address the information processing behaviour of the key stakeholders in agricultural drought forecast, whereas chapter 6 present the outcomes of the Fast- and Frugal decision tree used to predict drought at the end of the growing season. The last chapter (chapter 7) summarises the conclusion and give recommendations for both 510 and future research.

2

Theoretical Background

This chapter provides theoretical background information for the study to understand the final results and the decisions made in the process. The chapter elaborates on the definition of drought, drought indicators and the principles of forecast models in the humanitarian decision-making process.

2.1. Definition of drought

Drought is a natural hazard affecting many sectors; it affects both surface and groundwater resources, crop failure, economic and social activities, water quality and ecological systems. Furthermore, drought impact livelihood security, personal security, access to education, industries, agriculture and food security. (Svoboda and Fuchs, 2017) The impact of drought on these different sectors, results in the use of different terminology. There is not one common expression of drought (Mishra and Singh, 2010). The World Meteorological Organisation (WMO) explains droughts as follows:

> *Droughts can arise from a range of hydrometeorological processes that suppress precipitation and/or limit surface water or groundwater availability, creating conditions that are significantly drier than normal or otherwise limiting moisture availability to a potentially damaging extent.*

Drought classification exists of five different categories: meteorological drought, hydrological drought, agricultural drought, social-economical drought and groundwater drought (figure 2.1). This last type is not officially included in the drought classification, while it highly impacts the hydrological cycle. (Mishra and Singh, 2010), (Hao et al., 2018). Figure F.7 summarise the five drought categories:

- **Meteorological drought**: Meteorological drought expresses the lack of precipitation over a certain period in time. This type of drought is region specific due to atmospheric patterns influences by seas surface temperatures (SST). Often a meteorological drought is driven by a sequence of events, rather than from one only single event ((Hao et al., 2018), (Mishra and Singh, 2010) and (Typ)).

- **Hydrological drought**: Hydrological drought is defined as the inadequate amount of precipitation to establish surface and subsurface water supply, which is highly influenced by local climate and catchment characteristics. Hydrological drought affects the whole river catchment or watershed.

- **Agricultural drought**: Agricultural drought is related to the period with declining soil moisture content resulting in crop failure. This type of drought links closely to meteorological, hydrological drought and agricultural practices. Agricultural drought is affected by both physical and biological properties. Most agricultural drought predictors include indicators derived from soil moisture indices. (Hao et al., 2018)

- **Socio-economical drought**: Socio-economical drought associated with the failure of supply and demand of water resources for economic goods. It shows the strong relationship between drought and human activities and emphasises the impact of drought.

- **Groundwater drought**: Groundwater drought is the fifth type of drought. It affects the recharge, discharge and the groundwater level table. This type of drought develops over months or years.

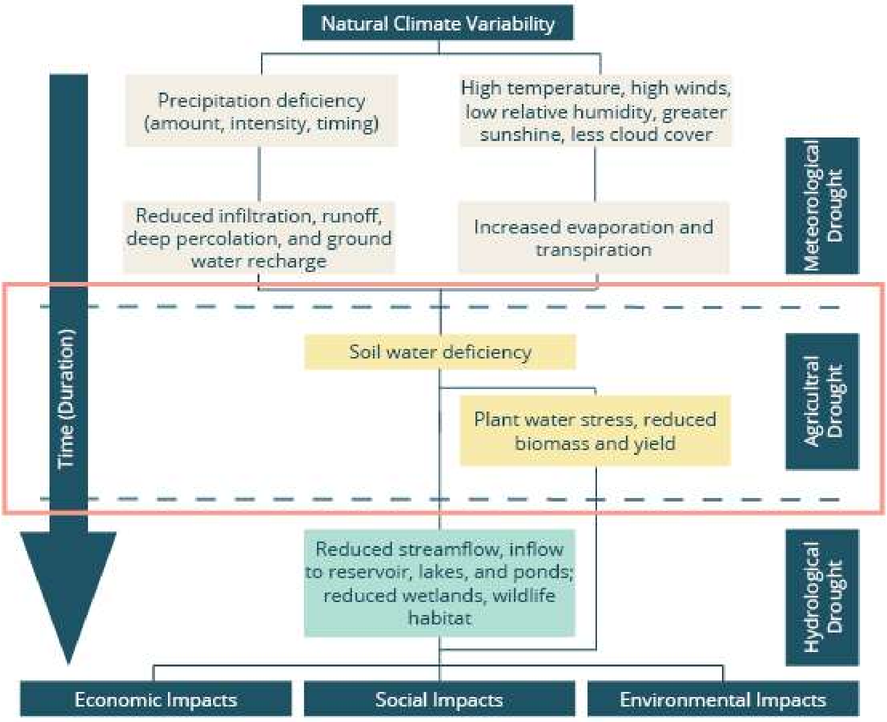


Figure 2.1: Overview types of droughts and their impact (NDMC),

Besides the drought classification, the severity, spatial distribution and duration of an event characterise droughts. Other characterisations are the return period, magnitude, predictability, timing and rate of onset (Zargar et al.). This variation within droughts shows its complexity.

2.2. Drought indices

In drought monitoring and EWEA systems, it is crucial to identify drought indicators according to the type of drought ((Barker et al., 2016), (Svoboda and Fuchs, 2017)). This matching of drought indicators to the correct type of drought requires knowledge about the hydrological cycle and how this affects the livelihood of people. This section presents an overview of the most commonly used drought indicators with their advantages and disadvantages per drought type. There is not one best drought indicator or combination of indicators; it is strongly dependent on specific characteristics of the area of interest. Chapter 4 (section 4.3.2) describes the selected drought indicators used in this master research.

2.2.1. Meteorological indicators

Precipitation or precipitation based indicators are the main feature of meteorological drought indicators, including the Standardised Precipitation Index (SPI), Standardised Precipitation and Evaporation Index (SPEI) and the number of days without rain. These indicators provide information about the duration and intensity of a drought event. Indicators based only on precipitation face challenges with a variety in spatial and temporal distribution, this variety can be reduced by using monthly averages or running means ((Mishra and Singh, 2010), (Bachmair et al., 2016a), (Wanders et al., 2010) and WMO 2006 report). Indicators completely based on precipitation underestimate the severity of a drought in regions with high evaporation rates (Barker et al., 2016). Combining precipitation with temperature and evapotranspiration rates improve in most cases the quality of the drought indicator. Important in the meteorological drought forecast is not only considering the local weather conditions but also include global weather patterns, such as the El Niño Southern Oscillation (ENSO). Table 2.1 presents the different meteorological indicators with their strengths and weaknesses.

Indicator	Description	Strength	Weaknesses	Reference
Precipitation	•Key component hydrological cycle •Create an understanding in the components of the water cycle. •Driving force water cyle is the energy cycle	•Detect possible trends and fluctuations. •Characterise global, regional and individual weather events	variety intensity, duration and spatial distribution	(Oreopoulos),
Dry Spells	•Sequence of dry days < 0.85 mm precipitation •dry periods affect all drought types	Link between precipitation and impact. Dry spells can cause reduced crop yields.	•Impact of a dry spell on different types of drought •Identifying the impact of a dry spell within a growing season.	(Wetterhall et al., 2015),
Standard Precipitation Index (SPI)	•Calculates precipitation deficit •Input: P	•Broadly Applicable •Easy to use •Comparable across regions	•Need complete datasets •Uses only P as input value	(Mckee et al., 1993),
Standardised precipitation Evaporation Index (SPEI)	•Uses SPI as basis •Input: P and T	••Comparable across regions •Suitable in different climate regimes	•Need complete datasets	(Vicente-Serrano et al., 2010)
Rainfall Deciles	•Distribution record over long period •Distinguish frequency and distribution precipitation	•Simple and flexible •Clear defined thresholds •Requires low amount of data	•Not suitable in strong seasonality climates •Long-term time series needed •Does not take into account other indicators	(Gibbs and Maher, 1967)
Rainfall Anomaly Index (RAI)	•Puts normalised precipitation values into historical perspective	•Easy to calculate •Applicable in multiple sectors	•Temporal variations need to be small •Requires complete datasets	(Svoboda and Fuchs, 2017)
Land Surface Tempearture (LST)	•Measures brightness temperature from IR spectrum. •mixture between vegetation and bare soil temperatures. •Depending on soil moisture, vegetation cover and albedo	•High resolution •Spatial coverage •Hourly data	•Cloud disturbance •Quick variations •Americas are no longer covered since Dec. 2017	(Freitas et al., 2013)
Multivariate ENSO Index (MEI)	•Global weather phenomena influencing wind and ocean currents. •Indicator of a El Niño La Niña •Input: six main variables in the Tropical Pacific seas surface temperature, surface air temperature, sea level pressure, surface winds and total coudiness fraction.	•Strong correlation with drought prediction in certain regions •One of the first indicators of drought	•Region specific influenced by ENSO	(Wolter and Timlin, 2011)
Oceanic Niño Index (ONI)	•Indicator of a El Niño La Niña •Sea surface temperature based (3-mont SST Anomaly) averaged over a 30-year period. •Indicator ocean conditions	•Most widely used in the US (NOAA standard) •Threshold broken down in severity classification	•Region specific influenced by ENSO	(NOAA's Climate Prediction)
Crop Moisture Index (CMI)	•Drought index specified on agricultural drought •Determine evapotransporation deficit by identifying the difference between Ep and Ea. •Input: Precipitation, temperature and previous CMI values.	•Response to rapidly changes •Comparable between climate regimes.	•Developed for grain-producing regions only	(Palmer, 1968)

Table 2.1: Meteorological indicators based on *Handbook of drought Indicators and Indices* (Svoboda and Fuchs, 2017), (Wanders et al., 2010), (Hao et al., 2018) and (Mishra and Singh, 2010))

2.2.2. Hydrological indicators

Hydrological indicators include streamflow and reservoir levels. These indicators do not have a direct relationship with the meteorological indices, which indicates that a meteorological drought does not directly influence a hydrological drought. A time lag between precipitation and recharge of (sub)surface water supplies results in a certain response time of the soil depending for example on the soil type and land use (Wong et al., 2013). Therefore, hydrological indicators provide information on water storage and water supply over a longer period. The indicators include available water content, groundwater levels, reservoir levels and streamflow (Typ). Table 2.2 list the widely used hydrological indicators.

Indicator	Description	Strength	Weaknesses	Reference
Palmer Hydrological Drought Index (PHDI)	•Based on the PDSI •Used to monitor the effect of drought on water resources •Calculate end of a drought •Input: Precipitation and temperature	•Water balance based, taking into account the whole water system.	•Human influences are not taken into account. •Variation in drought frequencies.	(Palmer, 1965)
Surface Water Supply Index (SWSI)	•Based on PDSI, adds water supply data. •Estimate drought frequency in four categories. •Input: Precipitation, reservoir storage, stream flow and snow pack	•Whole basin is taken into account •Indicate overall hydrological health of the basin.	•Comparison across basins is difficult due to inhomogeneity of the regions. •Basin specific	(Shafer and Dezman, 1982)
Standardised Water-level Index (SWI)	•Define impact of drought on groundwater recharge •Detect changes of soil moisture patterns •Input: Groudwater levels	•Ground waterlevels key component agricultural drought	•Looks at one parameter •Variations across region and climate regimes.	(Bhuiyan, 2004)

Table 2.2: Hydrological indicators based on *Handbook of drought Indicators and Indices* (Svoboda and Fuchs, 2017), ((Wanders et al., 2010), (Mishra and Singh, 2010) and (Hao et al., 2018)

2.2.3. Soil Moisture indicators

The soil moisture indicators indicate the amount of water present in the unsaturated zone; this amount of available water content is critical for plant growth and indirectly for crop yield. Soil moisture indicators include the features: precipitation, temperature, available water content, soil water deficit and storage data (Svoboda and Fuchs, 2017), (Mishra and Singh, 2010), (Bachmair et al., 2016a)). The decline of soil moisture is strongly related to both meteorological and hydrological factors.

Indicator	Description	Strength	Weaknesses	Reference
Palmer drought severity index (PDSI)	•Combines the water balance with precipitation and temperature to estimate soil moisture supply.	•Worldwide acknowledged •Robust system •Well tested and validated.	•Time lag in drought identification •Difficulties with strong seasonality regions •Does not account for snow.	(Palmer, 1965)
Soil Water Storage (SWS)	•Monitor soil water availability in the plant root zone storage. •Input: Soil type, soil water deficit, available water storage capacity.	•Simple, well-know method	•Inhomogeneous soil	(Wanders et al., 2010)
Soil Moisture Content	•Identify drought by using a soil moisture threshold value.	•Spatial coverage •High resolution	•Short term record •Cloud disturbance	

Table 2.3: Soil Moisture indicators based on *Handbook of drought Indicators and Indices* (Svoboda and Fuchs, 2017), (Wanders et al., 2010) and (Mishra and Singh, 2010)

2.2.4. Vegetation indicators

The last type of indicators are vegetation indicators. These type of indicators measure the vegetation conditions, which do not directly link them to drought. They measure vegetation dynamics influenced by weather conditions. In agricultural drought monitoring, this understanding between vegetation response and hydrological circumstances is crucial(Mishra and Singh, 2010). The vegetation indices measure the reluctance from canopies of electromagnetic waves. These remote sensing images show changes in the greenness of the vegetation (Xue and Su, 2017); these changes can identify for example the impact of weather conditions on crop production. The advantage of using remote sensing in drought monitoring is the high spatial resolution, cost-effectiveness and free access, a disadvantage for agricultural purposes is the pixel resolution and the revisit time of the satellite ((Mishra and Singh, 2010) and (Xue and Su, 2017)). Table 2.4 show the strength and weaknesses of the most commonly used vegetation indicators.

Indicator	Description	Strength	Weaknesses	Reference
Ratio Vegetation Index (RVI)	•One of the first Vegetation Indicators •Based on the absobtion of Near-infra led by vegetation •input: Near-Infra Red reflectance and Red band reflectance • RVI = $1\times\frac{\lambda R}{\lambda NIR}$	•Widely used •High densed vegetation	•Sensitive to vegetation •Sensitive to atmospheric effects	(Jordan, 1969)
Normalised Difference Vegetation Index (NDVI)	•Measure of greenness vegetation. •input: Visible and Near-Infra Red reluctance • NDVI = $1\times\frac{\lambda NIR-\lambda RED}{\lambda NIR-+\lambda RED}$	•Spatial coverage •High resolution •Takes into account canopy photosynthesis	•Short term record •sensitive to soil brightness, soil colour and leaf canopy shadow •Cloud disturbance	(Tarpley et al., 1984)
Vegetation Condition Index (VCI)	•Define duration of a drought •Compare vegetation changes with historical data. •Identify the severity of vegetation stress • VCI = $100\times\frac{NDVI_i-NDVI_{min}}{NDVI_{max}-NDVI_{min}}$	•Spatial coverage •High resolution	•Short term record •Cloud disturbance	(Kogan, 1995)
Vegetation Health Index (VHI)	•Identify vegetation stress. • VHI = $0.5\times VCI+0.5\times TCI$	•Short term record •Cloud disturbance	•Short term record •Cloud disturbance	(KOGAN, 1990)
Leaf Area Index (LAI)	•Interaction atmosphere and terrestrial environment •Amount of leaf material in ecosystem •Plant growth analysis	•Used in climate change studies •Regional and local scale •Indirect and direct measurement techniques.	•Underestimation in dense canopies.	'(Gobron)
Vegetation Optical Depth (VOD)	•Indicator of the total vegetation water content in the biomass •Equivalent of NDVI	•Sensitivity to the woody parts of the vegetation •Not disturbed by cloud cover •Canopy penetration is deeper , using longer wavelengths.	•Not yet been studied intensively •Open water bodies affect results	(Liu et al., 2011)
Normalised Difference Water Index (NDWI)	•Similar to NDVI •Monitor water content in vegetation. • NDWI = $1\times\frac{p(0.86\mu m)-p(1.24\mu m)}{p(0.86\mu m)+p(1.24\mu m)}$ $p(\mu)$ is the wavelength reflectance	•Less sensitive than NDVI to atmosphere scattering.		(Gao, 1996)
Water Requirement Satisfaction Index (WRSI)	•Monitor water availability for crops during the growing season •Input: Ea, WR, Ep, Kc •Monitoring and early warning tool used in food security preparedness	•Spatial coverage •High resolution	•Crop specific •Impact other	(Verdin and Klaver, 2002)

Table 2.4: Vegetation indicators based on *Handbook of drought Indicators and Indices* (Svoboda and Fuchs, 2017), (Wanders et al., 2010), (Xue and Su, 2017) and (**?**)

2.3. Forecast models

The difficulty in developing a robust weather forecast is the inherently variable nature of climate. Multiple factors influence the climate patterns, such as air-sea interaction, topography, internal dynamics and global weather systems. The effects of these factors can influence the climate system for months or years (for example ENSO). Also in drought forecast models, these climate systems play an important role ((NDMC). The fundamental parameters in drought forecast are precipitation and temperature influenced by climate conditions. The large-scale atmospheric circulations driven by SST anomalies and regional characteristics affect precipitation and temperature patterns. Those forcing factors have a certain level of drought predictability and can be of added value in drought prediction. On the region scale, land surface characteristics show their impact on the water and energy balance influencing temperature, precipitation variabilities and land-atmosphere interactions. At a seasonal time scale, these varying atmospheric boundary conditions are helpful. In other words, drought forecast models should incorporate different types of predictors: large-scale teleconnection patterns, which function as the driving force of drought. Secondly, the local climate variables and last the initial land conditions combining different types of indicators increase the persistence of the prediction. With these indicators, another problem occurs. There exists an inconsistency in the identification of drought by using different threshold values in the different models. Developing reliable drought predictions start with creating an understanding of which factors strengthen and weaken the severity of droughts and which threshold values can be determined. Over the past years, researchers developed ways to deal with the inconsistent threshold values in drought predictions. They developed scores, univariate or multivariate drought indicators based on a variety of hydroclimatic variables. Nonetheless, there is not yet a universally accepted drought characterisation (Hao et al., 2018). Besides the inconsistency in identifying drought, The US National Drought Mitigation Center mentions another problem in drought forecast; they mention the lack of knowledge by scientists to predict a drought a month or more ahead of the event.

When translating drought predictions to the forecast in the humanitarian sector, the lead time of the forecast and the mitigation time need to be balanced. The humanitarian organisation might find it useful to monitor drought predictions at multiple time scales, on the one hand, to prepare for short-term impacts and on the other hand to mitigate to the long-term changes within the natural climate. There are three-time scales used in the forecast. Long term forecast(decades to centuries), medium-term forecast (seasonal) and short term forecast (hours, days, weeks). Those different time scales have their advantages and disadvantages. An overview of these different forecast types is described in table 2.5. Besides the variety in time scale, forecasts differ in their decision triggers; a forecast can either be deterministic (yes or no drought) or probabilistic (there is a 60% chance a drought will occur in January) (table 2.5). In the decision-making process, it should be agreed, if the outcome should be deterministic or probabilistic, also a combined approach of the probabilistic and deterministic forecast can be used. The last partition within types of the forecast is the statistical, dynamical or ensemble forecast. An overview of the principle of all these different types of forecasts is presented in 2.5. This table combines the different forecast types mentioned by Potts (2012), Mwangi et al. 2014), Hao et al. (2018), FEWSNET (?), Seibert et al. (2017), Braman et al. (2013) and Lopez (2018).

2.3.1. Forecast skill

The forecast skill or skill score expresses the accuracy of a forecast, which is calculated by the difference between the observed and modelled forecast. Like mentioned before, there is not a fixed agreed threshold for a certain skill of the forecast ((NOAA, 2012), (Murphy, 1993), (FEWS NET, 2018)). There are many different methods available to validate a forecast. Murphy and Winkler ((Murphy and Winkler, 1987)) validate their forecast by *reliability, resolution, discrimination and sharpness*. These expressions do not yet decide on the 'best' forecast. Later Murphy (1993) uses three other validation types to evaluate a forecast: *consistency*, *quality* and *value*. This terminology to validate a model is broadly used; *Quality* refers to the accuracy or skill of a forecast, related to the relation between observations and forecast information (Potts, 2012)). *Consistency* is a measure corresponding to the judgement of the forecaster, the information contained by the forecaster should relate to the forecast itself, and the last expression *value*, formulate the benefits realised by the decision makers based on the forecast. Like the forecast itself, the forecast skill can be either deterministic or probabilistic, depending on its evaluation method. A deterministic validation method is a leave-one-out cross-validation, while a probabilistic method is an evaluation of the model by the ROC-score (Seibert et al., 2017).

2.3.2. Current drought forecast models

This section presents five agricultural drought forecast models. The models presented are: Famine Early Warning System Network (FEWS NET) developed by USAID, the Agricultural Stress Index System (ASIS) by the Food and Agriculture Organization of the United Nations (FAO), the Food Security Climate Resilience (FoodSECURE) from the World Food Program (WFP), the US Drought Monitor (USDM) and the Operational Decision Support System (ODSS) in Malawi developed by the Danish Hydrological Institute (DHI). The stakeholders mentioned these forecast models as a reference. The five agricultural drought forecast models presented in this section all have a slightly different focus, exposing different angles of agricultural drought and

Forecast	Description	Strength	Weaknesses
Long term forecast	•Predictions for 10-30 yr and to 2100 •Forecast general trends •Regional prediction	•Construct long-term visions •Reduce vulnerability of certain areas	•Prediction uncertainties •Specificity spatial area
Medium term forecast	•Seasonal forecast (covers 3 to 4 months period) •Monthly updated •Information about the coming season over large spatial area	•Information over large spatial area •Give information about above or below normal conditions •Anticipate and prepare for certain risks	•Coarse resolution forecast •Prediction uncertainties •Specificity spatial area
Short term forecast	Specific information about where and when an event might take place	•Forecast in context •Direct information about if 'extreme' events occur	•Minimal preparation time •Reliability forecast
Dynamic forecast	•Based on law of physics •Build from climate models and hydrological models •Seasonal time scale	•Enlighten multiple aspects of drought properties •Ability to capture non-linearity in a system •long-term lead time predictions	•Post-processing to improve coarse resolution climate models •Model biased •Computationally demanding
Statistical forecast	•Based on mathematical relationships •Uses a variety of predictors •obtained from historical observations •Aim to identify temporal and spatial dependency between predictors and predictand.	•Simple implementation •Low computational resources needed •Setting a baseline level of skill	•Involve large number of predictors •Risk of multicollinearity and overfitting •Rely on historical observations
Ensemble forecast	•Combination of dynamic and statistical models •Collection of multiple models •Compare spread an mean values of different models •Commonly merge: regression model, Bayesian posterior distribution and BMA.	Combine strengthens of both statistical and dynamic models	•Spread between forecast can be large
Deterministic forecast	•Special type of probabilistic forecast •Binary character, outcome expressed in zero or one	•unambiguously •Clear	•Not all outcomes can be expressed in a binary structure.
Probabilistic forecast	•select a probability to a set of outcome •Expressed in 'average', 'above-average' and 'below-average' values, related to a historical 30 year observed climate.	•Include uncertainties in decision making process. •Threshold for each forecast category •probabilities used to maximises value of information. •Suitable in wide range of users •In cost-loss problems	•Can be ambiguously •Agree upon a threshold probability in decision making process.

Table 2.5: Types of forecast (Potts, 2012), (Mwangi et al., 2014), (Hao et al., 2018), (FEWS NET, 2018), (Seibert et al., 2017), (Braman et al., 2013), (Lopez et al., 2018)

for different region/countries. FEWS NET focuses on the risks of food insecurity supporting decision-makers in the humanitarian sector, while ASIS provides risk managers halfway the season with a probability which drought level is expected at the end of the growing season, using the crop coefficient, crop type and water needs in the growing season. FoodSECURE mainly focuses on building climate resilience by community-centred actions, and the US DroughtMonitor is more a communication method advising multiple sectors about the current drought status, instead of actual predicting drought. The last model described is the ODSS, which the stakeholders in Malawi use for both flood- and drought predictions in the lower Shire river basin.

All models, except the US Drought monitor, can be used for predicting agricultural drought in Malawi, though at different granularity level. These four models form a baseline against which our machine learning model will have to be bench marked/tested. Only if our ML model performs better on the skill, accuracy, usability and accessibility will it be useful to adopt by decision makers.

FEWS NET

USAID develops the Famine Early Warning Systems Network (FEWS NET) in response to the famine in 1984 in East and West Africa. They aim to support early warning analysis and early warning forecast in order to reduce the impact of food insecurity. Together with their partners, the National Aeronautics and Space Administration (NASA), National Oceanic and Atmospheric Administration (NOAA), US Department of Agriculture (USDA) and US Geological Survey (USGS) they evaluate food security conditions (Whi). FEWS NET publishes weekly Global Weather Hazards and Seasonal Monitor reports sharing information on severe weather or climate conditions, impacts on production and gives a short-term forecast. The weekly Global Weather Hazards reports provide short- and medium forecast (up to 1 week) and include irregularity weather observations. Furthermore, FEWS NET give timely alerts on crisis and emergencies and publish reports with market and trade information, livelihood, nutrition and food assistant. Their partners have individual responsibilities, NOAA's Climate Prediction Center takes care of long-term seasonal outlooks, weather forecasts and analyses trends and USGS provide access to satellite and geospatial data. FEWS NET has its data portals and tools to monitor food security. Agroclimatology tools they use are Early Warning Explorer (EWX), FEWS NET Climate Engine, Drought Status Monitor (DSM), NASA Global Agricultural Monitoring, USDA Crop Explorer, NOA FEWS NET Data Portal and USGS FEWS NET Data Portal. The main products of these tools are NDVI, Rainfall Estimate (RFE), Land Surface Temperature (LST), Actual Evapotranspiration (ETa), Water Requirement Satisfaction Index (WRSI) and precipitation (Whi).

FAO Agricultural Stress Index System (ASIS)

The Agricultural Stress Index System (ASIS) combine the three characteristics of drought; intensity, temporal and spatial distribution using the VHI derived from NDVI and temperature anomalies. The method is used to monitor agricultural drought and to activate a trigger mechanism for agricultural insurance and implementation of mitigation activities. The tool monitors the crop cycle and provides information on the probability of a drought level at the end of the growing season. The drought classification exists of extreme, severe, moderate and slight drought. At the start of the season, 30 years of data used to identify the crop status at the start of the season. Throughout the season the tool gives a decadal update. Halfway the growing season the ASIS provides a certain probability of the drought classification. The ASIS uses the basic principles of absorption, transmission and reflection of solar energy to identify water stress, which is essential predicting the crop yield. Key products in the model are crop coefficient, water needs, VHI and Land Surface Temperature. If available they include yield data (Rojas).

Food Security Climate Resilience (FoodSECURE)

The Food Security Climate Resilience (FoodSECURE) tool is developed by the WFP reinforcing and building climate resilience by supporting community-centred actions. WFP aims to mitigate and manage risks by converting early warning to early action. FoodSECURE focuses on three windows: before, during and after a disaster strikes. Window 1) Climate forecasts stimulating trigger-based actions before shocks occur to increase community resilience (reducing the impact of a disaster); Window 2) complement existing early response instruments, and window 3) building resilience during post-disaster recovery periods by providing multi-year financing. This FoodSECURE model deploys forecast-based financing and link early warning indicators to pre-defined procedures in preparation for a disaster (WFP, 2016).

US Drought Monitor

The US Drought Monitor (USDM) is a product mapping current drought conditions in the US. The USDM is not a forecast model; it communicates the impact of the current drought conditions (Abo). The map integrates various climate indicators to a spatial extent, severity and time scale of a drought with corresponding thresholds ((Bachmair et al., 2016b), (Wahlström, 2009), (Hao et al., 2017)). Every week the USDM publishes a map presenting the impact and threats of the weather conditions in the different sectors (**?**). The drought monitor makes use of precipitation data, streamflow data, reservoir levels, temperature data, evaporation data, soil moisture data, media information and expert input of 400 observers. Key indicators are among others PDSI, SPI, SPEI, CMI, VHI, SM and streamflow (WMO, 2006), (Svoboda and Fuchs, 2017). The USDM classify drought in five categories, D0-D4 expressing the severity of a drought event. D4 means a 1 in 50-year exceptional drought. D0 is indicating an abnormally dry area, D1 Moderate drought, D2 severe drought and D3 extreme drought (figure F.7).

Operational Decision Support System

The Operational Decision Support System (ODSS) in Malawi is developed under the project *SRBMP/BC-19 Shire Basin Operational Decision Support System through Enhanced Hydrol-Meteorological Services*, with the purpose to support decision-makers and access information on targeted communities. The combination of ground monitoring, earth observation and weather and/or climate forecast is used to improve flood forecasting and warning systems. Components of the ODSS are flood- and flow forecasting, water infrastructure, a seasonal agricultural forecast (crop calendar), drought monitoring and flash flood forecasting. In this research, the seasonal forecast for agriculture and drought monitoring is of interest. The seasonal forecasts used in the ODSS are the Climate Forecast Systems [1] and the Global Forecast System [2] (MOAIWD, 2018). The

[1] CFS version 2, operational since March 2011 by the Environmental Modelling Center at NCEP.
[2] GFS by NCEP

ODSS component, agricultural forecast focusses on planting information, agricultural activities and satellite-based parameters important for water resource management. The cropping calendar provides information about the planting, growing and harvest period for each crop type. The drought monitor uses precipitation, SPI, NDVI, VHI, SWI, PET, ET and temperature data to classify drought. The ODSS applies the US Drought Monitoring drought classification.

3

Case study Malawi

The case study is located in Malawi, one of the African countries most affected by droughts (FAO, 2015b). Between 1979 and 2010, roughly 21.7 million people were affected by climate-related hazards. These hazards cause disruptions in food production and the reduction of community resilience. (Government of Malawi, 2015). This chapter will give a brief introduction of the study area by first discussing the country characteristics, climate and its vulnerability to natural hazards concerning the agriculture sector. The last section will specify field study locations.

3.1. Country characteristics

Malawi is a landlocked country in Southern Africa (lat. 9.5-17 °S and lon. 32-36 °E), bordered by Zambia, Tanzania and Mozambique (figure 3.1). The country has a total population of 19 million people from which 22% live in the Lower Shire River basin ((Coulibaly et al., 2015), (MoAIWD, 2016)). The country covers an area of 118,484 km^2 from which lakes and rivers cover 24,404 km^2. The Shire River, which is the only Southern outlet of Lake Malawi is flowing into the largest river basin of the country. The Shire River Basin is an essential vain of Malawi covering 16% of the total country area where 22% of the 19 million people live.(Coulibaly et al., 2015). This Shire River is a tributary of the larger trans-boundary Zambezi River meandering through the country eastwards to the Indian Ocean. Malawi is ranked number 171 on the world human development index (HDI) as one of the poorest countries in the world, with 71% (2010) of the population living under the poverty line of $1.90 a day (The World Bank).

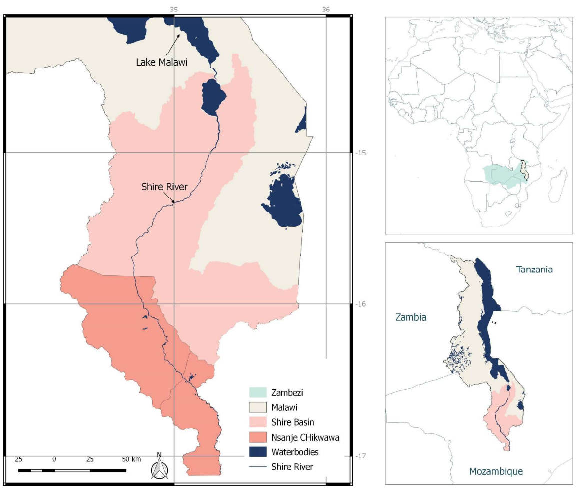


Figure 3.1: Geographical location Malawi

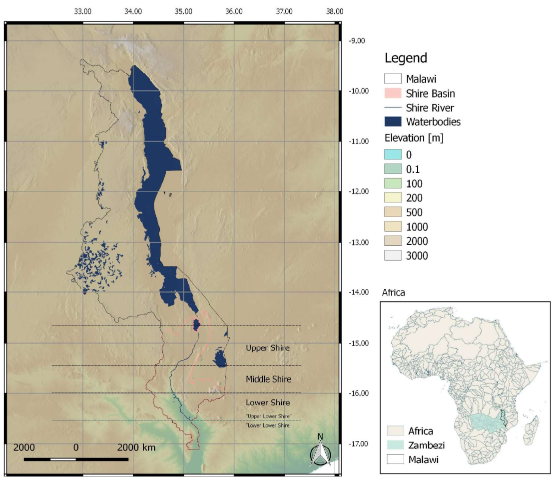


Figure 3.2: Elevation Malawi

3.1.1. Climate

The Great Rift Valley dominates the topography of Malawi stretching from North to South across the country. The valley exists of Lake Malawi and is surrounded by large plateaus with elevation heights between 914-1219 m and peaks in the Mulanje highlands of 2134-3048 m influencing Malawi's climate. Aside from the affect mountain riffs and Lake Malawi, the temperature and precipitation patterns fluctuate with the movement of the Inter-Tropical Convergence Zone (ITCZ). The ITCZ is responsible for the seasonality driven by the northeasterly monsoon and the southeasterly trade winds. Malawi experience a wet season from November-April and a dry season from May-October (Ngongondo et al., 2011). The intensity, the start and the end of the dry/wet season vary from year to year and vary between North and South Malawi. Another climate system influencing Malawi's climatic patterns is the Indian Ocean Sea Surface Temperature (SST) driven by the El Niño Southern Oscillation (ENSO). Malawi can experience either a decrease or an increase in precipitation response to El Niño, due to the location in a transition zone between two regions with opposite climate response ((Jury, 2014), (Vincent et al., 2014)). The effect of the SST and ENSO influence the inter-annual variability in the wet-season rainfall (Mcsweeney et al.). The annual average precipitation varies from 1600 mm along the shore of Lake Malawi and between 750-1000 mm in the rest of the country (GFDRR, 2011). In the Lower Shire basin, the annual mean is 1000 mm. The annual temperature varies between 18-27 °C with an average annual temperature of 24 °C. Between 1960-2003 the amount of hot days [1] increased by 30.5 days, this effect is largest in the DJF months. At the same time, the number of cold days [2] decreased, except for the SON months. An increase of 1.1-3 °C up to 1.5-5.0 °C is estimated for 2060 and 2090 respectively ((Mcsweeney et al.), (GFDRR, 2011))

3.1.2. Geology

This section describes five different soil types existing in the Lower Shire River Basin. The World Reference Base for Soil Resources (WRB) (FAO, 2015c) characterises *cambisols* as a soil type which makes productive agricultural land, with beginning soil formation. It experiences moderate weathering of parent material. *Vertisol* exists of heavy clay soils, with deep cracks during the dry season. This type is located in lower landscapes which periodically wet. The soil is used mainly for extensive grazing, wood chopping or charcoal burning. Although this soil type is less suitable for agricultural, it is used by smallholder farmers for millet, sorghum, cotton and chickpeas. It requires adaptive management to maintain production. The soil moisture range is small and turns into hard soil when dry and sticky in wet conditions (low infiltration rates). The third soil type discussed is *Lixisols*, which exists of high clay content in the subsoil and low clay content in the topsoils. This type is common in seasonal dry (sub) tropical climates and exists of low plant nutrients. A soil type common in flat lands in semi-arid climates is *Solonets* existing of high Na and Mg contents. The agricultural use of this type needs a deep (>25cm) humus-rich soil layer in order to create crop production, and therefore it is mostly used rather for grazing. The last soil type discussed is *Gleysols*, which is characterised by high groundwater tables. The parent material is mainly fluvial, marine and lacustrine sediment. This type occurs worldwide in various climates. The management of this type of soil exists of drainage systems needed to lower the groundwater tables.

Figure 3.3: Topographical impression TA Mitole

Figure 3.4: Topographical impression TA Zunde

[1] hot days are defined as the 10% temperature exceeding of the current climate.

[2] cold days are defined as the 10% of days below the recorded current climate.

3.2. Agricultural sector

The agricultural sector is the primary driver of Malawi's economy accounting for 80% of employment and one-third of the GDP (FAO, 2015a). As for most African countries, the agricultural sector is rain-fed making the country vulnerable to weather-related hazards (Coulibaly et al., 2015). Increasing temperatures and erratic rainfall patterns will increase the frequency and severity of drought influencing the agriculture production and the risk of food insecurity (Government of Malawi, 2015). In the case of drought, 80% of damages and losses hit the agricultural sector (FAO, 2015b).The Famine Early Warning Systems Network (FEWS NET) monitor the impact of drought on food insecurity and provide early warning supporting decision-makers in their response to humanitarian crises in more than 36 vulnerable, food-insecure countries. Malawi is one of these countries, and according to FEWS NET Food Security Outlook, the country will experience food stress developing towards food crises in Central and Southern Malawi during the lean season from October to January in 2018-2019 (FEWS NET, b). New predictions show below-average rainfall within the season resulting in limit agricultural activities.

In the Agricultural sector, 70% of agricultural labour is performed by women(Asa Giertz and German, 2015). Asa Giertz and German (2015) mention a higher vulnerability to environmental hazards, productivity and income for women-headed households compared to men-headed households. The aim for policymakers is mapping most vulnerable farmer communities, in most districts, these are female-headed households (Coulibaly et al., 2015). Women have less access to productive assets and resources. Furthermore, they cultivate a different type of crops. Men are engaged more to cash crops generating more incomes, which disproportionate male- or female-headed households (Asa Giertz and German, 2015). Another element is communication in case of a disaster; women tend to have less access to information resources than men in a male-dominated government. In the development of early-warning systems, the decision makers need to be aware of this inequality (Andrew Collings and van Aalst, 2009).

3.2.1. Growing season

In Malawi, the main food crops are maize, cassava and potatoes, from which maize is the main staple food. In the agricultural sector, droughts and pest generate the highest damaging production risks. The weather-related risks Malawi face are droughts, dry spells and erratic rains influencing the crop production. Drought occurs in different forms, covering extended dry spells, lower rainfall amounts or shortening of the growing season. Although droughts happen with a cycle of multiple years (for example driven by El Niño or La Niña), farmers experience drought as a shock. These different drought types impact the growing season in several ways. The main crop maize, show a high response to droughts reducing crop production. The crop calendar explains the impact of droughts on crop production (figure 3.5). Malawi follows a clear precipitation pattern with the rainy season from November to March and a dry season from May-September. Related to this climate pattern the crops follow a growing pattern from sowing, growing to harvest. Within these growing stages, there is a different crop water requirement. Table 3.1 present the crop water requirements for maize over the year using the Blaney-Criddle formula 3.1 and 3.2. Table 3.1 proved information on the length of each growing stage.

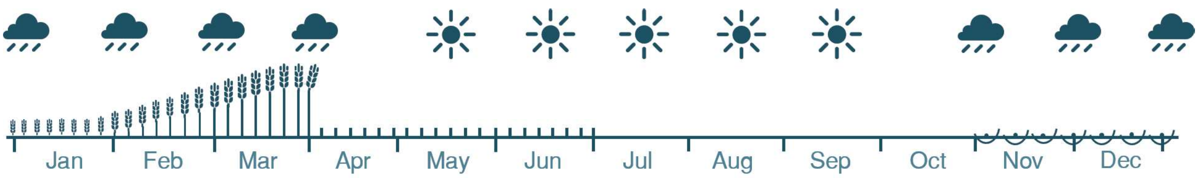


Figure 3.5: Seasonal crop calendar Lower Shire River Basin, based on USAID 2013

The climate factors influencing the crop water needs are solar radiation, temperature, humidity and wind speed. In a dry climate, the crop water need is higher compared to a more humid climate. Not only climate conditions indicate the crop water needs, but it also depends on the crop types. If we compare maize with sorghum, the length of the growing season for maize (sweet) covers 80-110 days and sorghum 120-130 days. The seasonal crop water needs vary between maize 500-800 mm and sorghum/millet 450-650 mm. Which makes maize highly sensitive to drought compared to sorghum (Brouwer, 1986) and (Critchley et al., 1991). Another critical factor of the crop water requirements is the growth stage. The crop water needs for a young plant is less than for a fully grown maize crop. The maximum crop water need takes place at the end of the crop development stage (beginning of the mid-season stage after +/- 55 days table 3.2). There are two ways of crop ripens. Maize is a dry harvested crop, which means that the crop is allowed to dry out within the late season stage (asking 25 % of the crop water requirements in the mid-season peak period). Table 3.1 present the crop factor of maize used in equation 3.1 and 3.2 to calculate the crop water requirements. This table also provided information about the length of each stage. The $ETcrop$ calculates the water requirements of the given crop and ETo calculates the evapotranspiration rate of a reference crop in mm per unit of time. In this equation p is the percentage of daytime hours depending on the month and latitude.

$$ETo = p(0.46Tmean + 8) \tag{3.1}$$

$$ETcrop = kc \cdot Eto \tag{3.2}$$

Crop	Initial stage	(days)	Crop dev. stage	(days)	Mid-season stage	(days)	Late season	(days)	Season average.
Maize	0.40	(20)	0.80	(35)	1.15	(40)	0.70	(30)	0.82

Table 3.1: Crop coefficient (Critchley et al., 1991)

	1	2	3	4	5	6	7	8	9	10	11	12
Eto	46.58	43.64	40.68	36.16	33.09	31.49	32.85	35.95	39.23	42.40	45.47	46.70
Tmean	299.06	298.82	298.42	297.05	295.34	293.78	293.09	295.18	298.49	300.43	301.48	299.84
											11	12
Initial stage	19	17	16	14	13	13	13	14	16	17	18	19
Crop development stage	37	35	33	29	26	25	26	29	31	34	36	37
Mid season	54	50	47	42	38	36	38	41	45	49	52	54
Late season	33	31	28	25	23	22	23	25	27	30	32	33
Seasonal average	38	36	33	30	27	26	27	29	32	35	37	38

Table 3.2: Maize crop water requirements in [mm/day] over the growing period of the crop. The growing season of Maize in Malawi is between November-March, with the harvesting from half of April until half of June. The values are measured at Lat=35 Height=10m (Tmean [K]), ETo=reference crop evapotranspiration [mm/day] calculated with the Blaney-Criddle method (Brouwer, 1986)

3.3. Natural Hazards

Malawi has to cope with multiple types of natural hazards, including floods, droughts and earthquakes. These natural hazards, floods and drought are strongly influenced by climate change and population growth(Coulibaly et al., 2015).

> "Climate change is often considered as the root cause of the increasing need for humanitarian aid, although climate change is a challenge on itself, land segregation, due to high population growth has a large influence on the impact of a natural hazard as well. The lack of resilience is a huge issue and should not be underestimated." -

In Malawi, the impact of weather-related disasters is high, because the majority of the population are smallholder farmers, who are depends on their own food production. This dependency makes them extremely vulnerable for natural hazards, even more, a sequence of such hazards (Coulibaly et al., 2015). Farmers do not have enough time between weather-related disasters to recover from a previous event and to prepare for a next disaster (Khamis, 2006). Among others, the changing rainfall patterns affect the length of the growing season and crop growth in the past years. Until 2006 Malawi experienced 40 weather-related disasters (Khamis, 2006). Most of them occurred after 1990. Since 2000 they have experienced much famine driven by El Niño in 2002, 2005 and again in 2012. Figure 3.6 shows the Oceanic Niño Index fluctuation, according to the index the season 91-92, 94-95, 97-98, 02-03, 09-10 and 15-16 experienced a strong El Niño, where 95-96, 89-99, 99-00, 07-08, 10-11 and 11-12 experienced a strong El Niña.

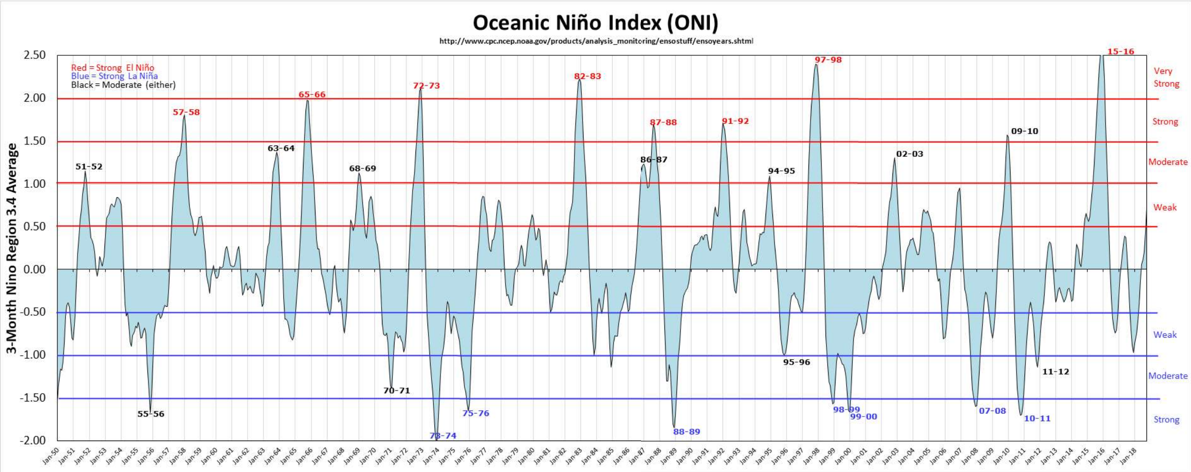


Figure 3.6: Ocean Niño Index 1950-present (Golden Gate Weather Services, 2019)

Droughts affect large parts of the country. Two databases are used to identify historical drought events. The first database is the Emergency Events Database (EM-DAT) and second the International Disaster Database and Global Drought Observatory (GDO). EM-DAT is used to support humanitarian organisations to prioritise and rationalise decision-making. They combine information from various sources, such as UN agencies, NGO's, insurance companies, press-agencies, governments, IFRC and research institutes (EM-). Data from UN Agencies, governments and the IFRC have priority, covering all disasters and contain less political limitations. At the country level, the database provides geographical, human, temporal and economic information on disasters. EM-DAT classify a when it encounters at least one of the following criteria ; i) More than 10 people lost their life [3]; ii) more than 100 people are affected[4], iii) a state of emergency is declared and iv) a call for international assistance is issued (possibly supported by an IASC humanitarian system-wide activation (IASC Working Group)) (Gui). The second database GDO is a database aiming at emergency response. This database analyses of a variety of drought indicators, such as precipitation, soil moisture, reservoir levels, river flow, groundwater levels and vegetation indicators associated with different drought types (Svoboda and Fuchs, 2017). Table (3.3) gives an overview of drought events mentioned by the two databases in combination with the Oceanic Niño Index.

Start	End	Duration	Severity	Intensity	Start	End	Affected	Duration	El Niño/ La Niña
GDO					EM-DAT				
Aug 2016	Dec 2016	5 months	6.26	1.25	Oct 2015	Dec 2017	6.700.000	26 months	Weak La Niña
Jun 2015	May 2016	12 months	17.45	1.45	-	-	-		Very strong El Niño
Feb 2012	May 2012	4 months	3.83	0.96	Aug 2012	2013	1.900.000		Very weak El Niño
Oct 2010	Jan 2011	4 months	4.24	1.06	-	-	-	-	Strong La Niña
Sep 2008	Apr 2009	8 months	4.20	0.52	-	-	-	-	Weak La Niña
Apr 2007	Sep 2007	6 months	5.58	0.93	Oct 2007	Dec 2007	520.000		Strong La Niña
Mar 2005	Mar 2006	13 months	15.62	1.20	Oct 2005	Mar 2006	5.100.000		Weak La Niña
Aug 2004	Oct 2004	3 months	4.55	1.52	-	-	-	-	Weak El Niño
Oct 2003	Apr 2004	7 months	8.8	1.18	-	-	-	-	Very weak La Niña
Jul 2002	Sep 2002	3 months	4.01	1.34	Feb 2002	2002	2.829.435		Moderate El Niño

Table 3.3: Overview historical drought events Malawi obtained from EM-DAT and the GDO database. The GDO data is based on the SPEI-3 indicator to predict drought. The severity is the (absolute) sum of the indicator values and the intensity is expressed in the ratio between the duration of the event and the severity.

[3]Total deaths is including both the number of people who lost their lives as a result of the event and the number of people whose whereabouts are not known, they are presumed dead in official figures.

[4]The sum of injured, affected and homeless people is the total number of affected presented in the table. EM-DAT explain the term affected as when people immediately require assistance during an emergency. The number of injured based on the need for immediate medical assistance as a direct result of a disaster and the number of houses destroyed or heavily damaged fall into the term homeless.

4

Methodology

This research applies an exploratory sequential mixed methods research design, combining semi-structured interviews, focus group discussions and data analysis(figure 4.1). The qualitative research phase explores the problem areas, create an understanding in the users (humanitarian organisations and smallholder farmers), review existing tools and limitations, and identify a possible list of indices to developed a new tool. The quantitative research phase implements the qualitative research results into building the drought forecast model, tested for the case study area of Malawi. This chapter gives a detailed description of the methodology applied in the different research phases.

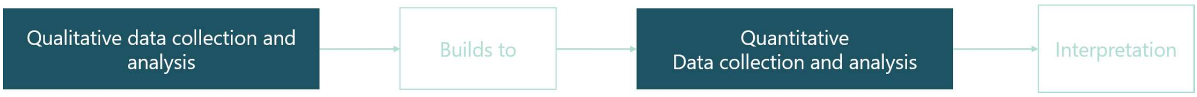


Figure 4.1: Exploratory sequential mixed method (Chapter 10, (Creswell))

4.1. Qualitative data collection and analysis

The qualitative data collection exists of semi-structured interviews and focuses group discussions (FGD) creating an understanding of the information processing behaviour of both the humanitarian organisations and the smallholder farmers. The conducted semi-structured interviews with international humanitarian organisations took place in collaboration with another MSc student at 510 Dominik Semet. Additionally to these interviews providing information about the decision-making process on a global scale, local stakeholders gave insights about the information processing behaviour in Malawi. The University of Leeds was involved in the FGD, section 4.1.2 elaborates on this collaboration.

4.1.1. Semi-structured interviews

The performed semi-structured interviews with humanitarian organisations that operate at both international and national level. The interviews give insight into the information processing behaviour of stakeholders involved in EWEA systems driven by natural disasters. Another aspect of the interviews is to identify the drought forecast systems currently in use. In total, nine international humanitarian organisations shared their experiences. The second part of the semi-structured interviews took place at the local community level were another ten interviews were held. These interviews aim to create an understanding of the information processing behaviour of national stakeholders and smallholder farmers in EWEA systems. The local interviews allowed to identify the current agricultural drought forecasting systems and agro-climatic indicators commonly used in Malawi. The stakeholders interviewed at local level exist of governmental organisations, national societies of international humanitarian organisations, such as WFP and RCRC, and agricultural officials in the research districts Nsanje and Chikwawa. During the interviews at the community level, a translator was available to help out with the language barrier. Appendix B present the protocols and questionnaires used in the qualitative research phase. As an addition to the nine international organisation, another four interviews took place in Malawi providing information about the decision-making procedures in Malawi. These four interviews exist of governmental institutes and national societies. The Department of Climate

Change and Meteorological Services (DCCMS) is responsible for the meteorological data and the weather forecast used by the Agricultural Officials to inform smallholder farmers with weather information. The Ministry of Agriculture, Irrigation and Water Development (MoAIWD) shared information on the current projects in Malawi.

	Type of Organisation	Organisation	Position/Role	Responsibilities	Region
1	Research	Bahir Dar University	Assistant Professor	Drought forecast research	Ethiopia
2	Field/International	ICCO	DRR Officer	DRR in Development projects	East Asia
3	HQ/International	WHH	Humanitarian Response Officer	Early Action	Madagascar
4	HQ/International	510/NLRC	Project Manager FBF	FBF	Southern Africa
5	HQ/International	FAO	EWEA Specialist	Early Warning Early Action	Mongolia, Kenya, Somalia, ...
6	HQ/International	IFRC	Senior Officer FbF	FBF	Global
7	Field/International	IFRC	DM Delegate Sahel Region	DRR	Sahel
8	HQ/International	WFP	Climate Scientist Consultant/ Meteorologist	DRR	Philippines, Nepal, Bangladesh, Haiti and Dominik Republic
9	HQ/International	Climate Centre	Senior Risk Adviser	FBF	Global,
10	Field/NS	MLRC	Country Representative/ Data team	-	Malawi
11	Government	DCCMS	Senior Meteorologist	-	Malawi
12	Government	MoAIWD	Hydrologist	-	Malawi
13	Field/International	WFP	Senior Project Manager FbF	FbF	Malawi

Table 4.1: Overview of the interviewed organisation. Department of Climate Change and Meteorological Services (DCCMS) and the Ministry of Agriculture, Irrigation and Water Development (MoAIWD).

4.1.2. Focus group discussions

Complementary to the semi-structured interviews were the FGD in Malawi. The protocol used in the FGD combines the developed protocol by the University of Leeds, as part of the IPACE-Malawi project and a protocol developed during this MSc thesis. Within the two protocols, there is a slightly different focus. The University of Leeds focuses on the understanding of the agricultural system and how weather information is received by smallholder farmers, whereas the protocol conducted as part of this research project focuses on the understanding of the livelihood of smallholder farmers during a drought period. In both protocols, a crop calendar is built to point out the most critical agricultural practises, to identify weather-related risks in the growing season and to identify severe drought years (appendix B). The groups for the FGD consisted of community members with a variety in agricultural experiences and gender to ensure broad knowledge of historical events, local traditions and experience with more innovative approaches in agricultural practices. The FGD took place in four villages, two of them followed the protocol of the University of Leeds, whereas the other areas used a merged protocol with adjustments from the MSc research protocol. The FGD pilot the protocol of the University of Leeds, to develop a standard protocol for further research in Malawi.

4.2. Study area selection

In the last two weeks of November 2018, a field visit was carried out with the aim to identify drought impact on smallholder farmers in the most drought-prone areas in the districts Chikwawa and Nsanje. This research aims to analyse the contribution of soil moisture monitoring at different spatial and temporal resolution, leading to research areas at different levels. The selected case-study areas follow the agricultural system in Malawi, which follows the governmental system (figure 4.2). The system exists of a national level, agricultural development division (ADD) and Extension Planning Areas (EPA). Malawi contains eight ADD's, twenty-eighth districts and hundred eighty-seven EPA's. Each EPA holds many Traditional Authorities (TA's) and villages (Kundhlande et al., 2015). Figure 4.2 illustrate the general governmental system and the study areas within this system.

The villages are selected based on both qualitative and quantitative criteria, the District Agricultural Development Officer (DADO) selected the most drought-prone areas within the district. In order to link the qualitative and quantitative re-

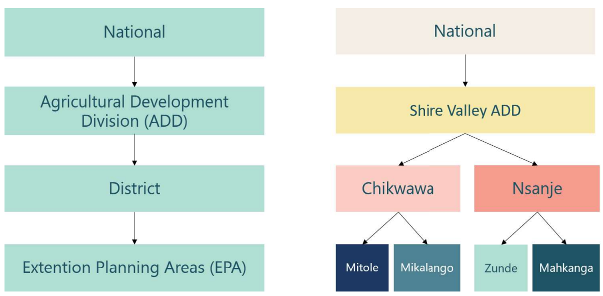


Figure 4.2: Four tier administrative hierarchical governmental sys-

search data the distance between the visited locations should at least exceed 25 km, due to the 25x25km resolution obtained from ESA. The quantitative research phase compares the model performance of using high and low spatial soil moisture resolution. This criterion in combination with the selection criteria of the FGD (variety in farmers gender, age and experience) resulted in the seven research areas presented in figure 4.3. These seven areas vary in spatial coverage from Mitole (11 km^2) to Nsanje + Chikwawa (2068 km^2). Please note that with Nsanje and Chikwawa is not refered to the complete two districts.

The locations experience a climatological and topographical difference. Mitole is located close to the high plateau at an altitude of (292 m). The other areas slightly lower at an altitude of 140 m (Mikalango), 85 m (Mahkanga) and 52 m (Zunde). Zunde is located close to the Lower Shire River with a distance of 2.5 km; also Mitole is located relatively close to the Lower Shire River with a distance of 3.5 km. The topography in this area is a bit different from the others, with a more rugged character. The areas experience various soil types varying between Cambisols (Mitole, Zunde), Vertisols (Mikalango), Lixisols (Zunde), Solonets (Zunde,Mahkanga) and Gleysols (Mahkanga) explained in chapter 3 (section 3.1.2).

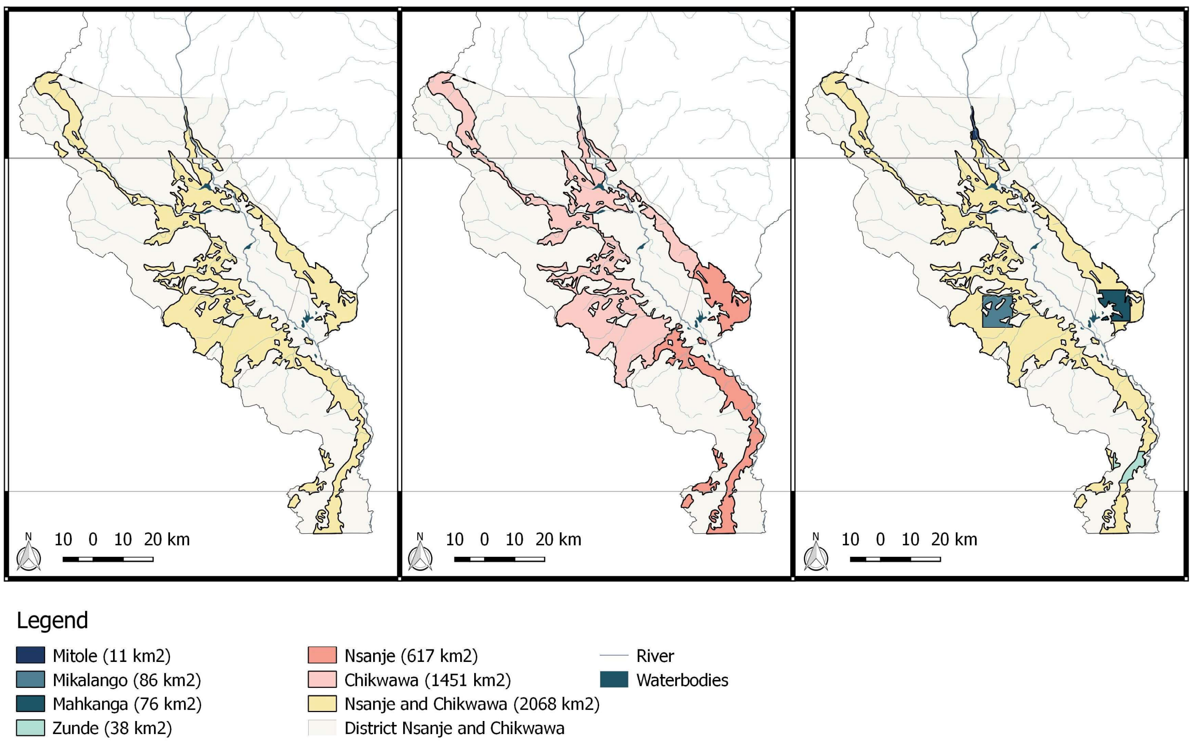


Figure 4.3: Field work locations within the Chikwawa and Nsanje district (Mitole, Mikalango, Zunde and Mahkanga)

4.3. Quantitative data collection and analysis

The quantitative research phase follows the qualitative data collection and analysis. This section explains the data collection and analysis used in the quantitative part. Subsection 4.3.1 elaborate on the candidate predictor and predictand selection criteria creating a shortlist of indicators mentioned in chapter 2 (section **??**-2.2.4). The drought forecast model developed in this research uses this shortlist of indicators to predict a drought at the end of the growing season. The outline of the drought forecast model is explained in more detail in section 4.4.

4.3.1. Candidate predictor and predictand selection criteria

The drought indicators presented in sections 2.2.1-2.2.4 uses multiple criteria in order to create a subset of variables in the drought forecast model. This selection follows from the data source characteristics: high temporal resolution, spatial resolution, geographic coverage and temporal coverage. Other selection criteria taken into account are the download accessibility, making use of other indicators, relevance in agricultural drought forecast, familiarity by the key stakeholders and implementation in current drought forecast/monitoring systems. This research aims to identify the contribution of using machine learning techniques with satellite-derived products into agricultural drought forecast. The satellite data company VdS provides high spatial resolution soil moisture data to implement in the model. The time frame of this data covers a period from June 2002 to December 2017, which is the guiding selection criteria for all other data sets. In this research the interest is in the growing season from November-March, which exclude season 2017/2018, furthermore, in 2012 there is a data gap caused by changing of the satellite. Table 4.2 present the drought indicators and their performance on the candidate predictor and predictand selection criteria. In this table, the accessibility of the open-source data guides the time frame of the datasets. Note that the rainfall decile, rainfall estimate, SPI, SPEI, VCI and VHI is calculated by using precipitation, evaporation and the NDVI dataset, instead of using a separate data source. The first exclusion of the shortlist is based on the required time frame of the data set from 2002-present. The drought indicators excluded from the shortlist, due to this reason are Ep, SPI, SPEI, Rainfall estimate, CMI, PHDI, SWI, VCI and LAI. The download accessibility is the second exclusion criteria, followed by the spatial coverage and indicators using drought indicators constructing their index (SWSI, RVI and VHI). After this exclusion, the shortlist exists of the following indicators: precipitation, MEI, ONI, soil moisture (SM), land surface temperature (LST), vegetation optical depth (VOD) and Normalised Difference Vegetation Index (NDVI). Additionally, to these indicators, one teleconnection criteria are included. Only one El Niño indicator is needed in the model preventing correlation between the variables in the drought forecast model. As El Niño indicator the ONI is selected, based on the recommendations from the senior meteorologist of the Malawi Meteorological Service.

		Download Accessibility	model in itself	Time frame	Time Interval	Spatial Coverage	Spatial Resolution	Relevant for agricultural drought	Familiarity with stakeholders	Implementation current drought forecast
Meteorological	Precipitation	easy	no	1981-present	Daily	Global	0.05°	yes	yes	FEWSNET, ODSS
	Evaporation	easy	no	2003-2019	Daily	Global	0.05°	yes	yes	FEWSNET, ODSS
	SPI	esay	no	01-16-1949 12-31-2012	Monthly	Global	1°		yes	USDM, ODSS
	SPEI	easy	no	01-01-1901 12-31-2011	Monthly	Global	0.5°			USDM
	Rainfall Deciles*	easy	no	See P	-	-	-			
	Rainfall estimate	easy	no	2012-present	Dekadal	Soutern Africa	0.08°			FEWSNET
	RAI*	easy	no	-	-	-	-			
	LST	easy	no	2010-present 2002-present	hourly	Global	0.05° 0.003°	yes	yes	ASIS, FEWS NET, ODSS
	MEI	easy	no	1950-present	Monthly	Global				
	ONI	easy	no	1949-present	Monthly	Global	0.005°		Yes	
	CMI	easy	yes	2015-present	Monthly	Global	0.05°			USDM
Hydrological	PHDI	hard	yes	2005-present	weekly	US	2.5°			
	SWSI*	hard	yes	-	-	-	-	no		
	SWI	easy	no	2007-present	10-daily	Global	0.1°	yes	yes	ODSS
Soil Moisture	PDSI	hard	yes	1998-2015	Weekly	US	2.5°	no		USDM
	SM	easy	no	1978-2015 2002-present	Daily	Global	0.25° 0.003°	yes	yes	USDM
	SWS	easy	no	-	-	-	-			
Vegetation	RVI	hard	no	-	-	-	-	-		
	NDVI	easy	no	1999-present	8-daily	Global		yes	yes	FEWS NET, ODSS
	VCI	easy	yes	2013-present	16-daily	Global, Continental	10° 0.01°	yes		
	VHI	easy	yes	1981-2012 2013-present	weekly	Global	0.04°	yes	yes	ASIS, USDM, ODSS
	LAI	easy	yes	1999-present 2014-present	dekadal	Global	0.01° 0.003°	yes		
	VOD	easy	no	1978-present 2002-present	Daily	Global	0.25° 0.003°	yes		
	NDWI	easy	no	2000-2017	Daily	Global				
	WRSI	easy	yes	1999-present	dekadal	Global	0.1°	yes		FEWS NET

Table 4.2: Candidate predictor and predictand selection criteria. Data sources: CHIRPS, ESA CCI, GIMMS, GEOGLAM, FEWS NET, NASA, NOAA, Copernicus Global Land Services and VanderSat

4.3.2. Data sources

The main data sources used in this research are open source satellite-derived products. Another data source used is the high-resolution soil moisture data obtained from VdS. The other datasets used are CHIRPS, ESA CCI, GIMMS, NOA and VdS. The dry spell data is developed from the CHIRPS dataset. Table 4.3 present the data collections used in the Machine Learning Model, followed by the description of the datasets.

Precipitation

The precipitation data is obtained from the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS version 2.0) Climate Hazard Groups (CHG) US Santa Barbara. The near-real-time daily dataset in-

Indicator	Unit	Spatial Resolution	Temporal resolution	Interval	Source
Precipitation	mm	0.05°	1981-present	Daily	CHIRPS
Dry spells	amount of dry days	0.05 °	1981-present	Daily	CHIRPS
Soil Moisture (high resolution)	m3/m3	0.003°	2002-06 to 2017-11	Daily	VdS
Soil Moisture (low resolution)	m3/m3	0.25 °	1978-2015	Daily	ESA CCI SM v04.4
Vegetation Optical Depth (VOD)	kg/m2	0.003°	2002-06/2017-11	Daily	VdS
Land Surface Temperture (LST)	Kelvin	0.003°	1999-present	Dekadal	VdS
Normalised Difference Vegetation Index (NDVI)	[0-250]	0.00225 °	1999-present	8-Daily	GIMMS MODIS Terra
Oceanic Niño Index (ONI)	std dev	0.005°	1949-present	monthly	NOAA

Table 4.3: Overview data collection

corporates 0.05° resolution satellite imagery over a 30+ year period from 1981-present. CHIRPS data is specifically designed for early warning agricultural drought monitoring supporting the United States Agency for International Development Famine Early Warning Systems Network (FEWS NET). Funk et al. (2015).

Dry Spells

One of the predictors is the amount, the length and frequency of dry days in the growing season. A dry day is defined with a day less than 0.85 mm/day rainfall (Barron et al., 2003). When talking about a sequence of dry days, it is called a dry spell. In this research dry spells are categorised into sequences of 0-2 dry days, 3-4 dry days, 5-9 dry day and a sequence of more than ten days. These sequences relate to the dry spell duration thresholds of three, five and ten days reducing the maize crop production ((Winsemius et al., 2014),(Barron et al., 2003)).

Soil Moisture, Vegetation Optical Depth and Land Surface Temperature

VdS provided the Soil Moisture (SM), Vegetation Optical Depth (VOD) and Land Surface Temperature (LST) data. The data has a spatial resolution of 300x300m and is obtained by the Advanced Microwave Scanning Radiometer (AMSR) from two different satellite missions, the first one is AMSR-E on NASA's EOS Aqua satellite and the second one is AMSR-2 on the JAXA-s GCOM-W1 satellite [1]. In the time series analysis, a gap between the two missions is visible excluding the 2012 data from the analysis. Soil Moisture is expressed as the volume of water in a volume of soil in [m3 m-3]. VOD measures the vegetation water content in the canopy with the dimension [kg m-2], which relates to the biomass, meaning high VOD values indicate high biomass. Land Surface Temperature is stated in Kelvin.

Soil Moisture ESA CCI SM v04.4

The second soil moisture dataset is downloaded from the Soil Moisture Climate Change Initiative (CCI) project of the ESA Programme. The project focusses on the multi-frequency radiometers and C-band scatterometers. ESA CCI SM 404.4 make use of both passive as active microwaves. The active data is provided by the Vienna University of Vienna (TU Wien) obtained from ERS-1, ERS-2 and MetOp-A and MetOp-B. VdS provides passive data. The data has a spatial resolution of 0.25 ∘ over the period 1978-2015 (ESA).

Normalised Difference Vegetation Index (NDVI)

The Normalized Difference Vegetation is retrieved from the *MCD43A4: MODIS/Terra and Aqua Nadir BRDF-Adjusted Reflectance Daily L3 Global 500 m SIN Grid V006*. The daily data is based on a 16-day interval period; the algorithm combines both the Terra and Aqua sensor products choosing the best representative pixel. As the name of the dataset already mentioned the spatial resolution is 500 m.

The Oceanic Niño Index(ONI)

One of the measures to express the El Niño-Southern Oscillation is the monthly Oceanic Niño Index (ONI). It is a 3-month running mean over a centred 30 year period. The index is updated every five years and calculating the ERSST.v5 Sea Surface Temperature (SST) anomaly (201, 2018), with a threshold value of +/- 0.5 C°. The ONI data is retrieved from the National Oceanic and Atmospheric Administration (NOAA).

[1] from May 2002 until October 2011 the data is retrieved from NASA's EOS Aqua mission. In May 2012 the JAXA's GCOM-W1 satellite was launched providing the AMSR-2 data nowadays

4.4. Outline Machine Learning Model

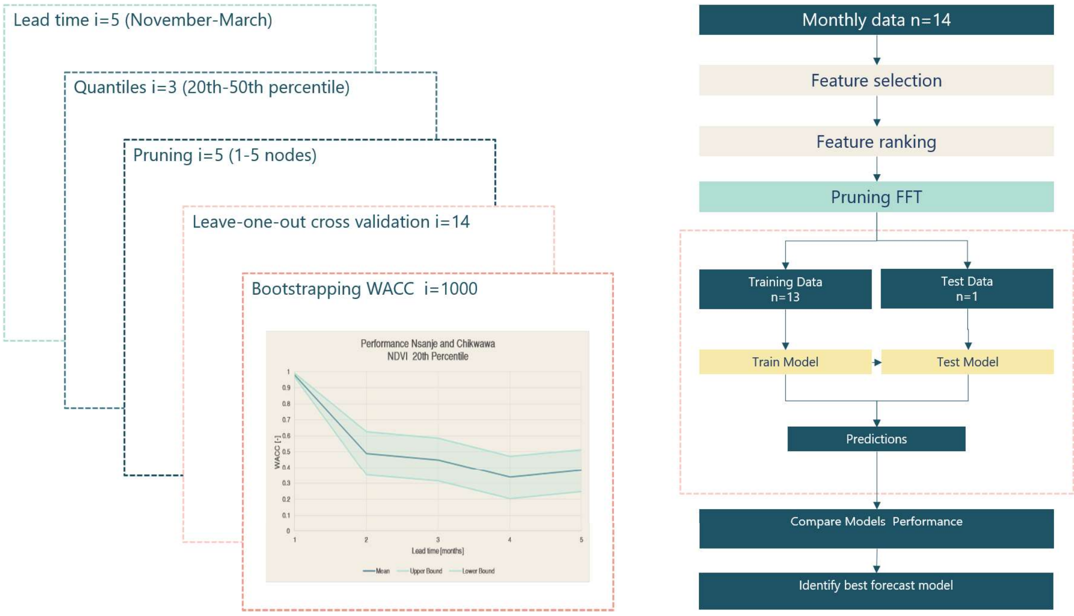


Figure 4.4: Outline drought forecast model. The model iterate over four levels, first over the months within the growing season, followed by the drought classification threshold values and the number of nodes in the decision tree (1 to 5) and last the cross-validation using Leave-one-out with replacement. The bootstrapping is down over the Leave-one-out cross-validation in order to calculate the confidence interval over the WACC. The steps are explained in more detail in section 4.4

This section describes the Machine Learning Model. In the study, the Fast-and Frugal decision tree is used to identify the best feature combination at different spatial and temporal resolution to predict drought in March (end of growing season). This process is repeated for all seven research areas to compare the effect of spatial distribution on the feature combination. The subsections explain the process steps in more detail. First, the project goal is defined followed by the data preparedness phase. Subsection 4.4.3 describes the algorithm and code used in the model. Section 4.4.4 specifies the validation and performance steps of the model; the last subsection elaborates on the output of the model.

4.4.1. Selection of the machine learning algorithm for agricultural drought modelling

The complex relationships between science, variability in data sources, types of models and understanding the potential impact of a hazard make it complicated for experts combining all this information in the development of an effective model. To provide disaster risk managers with a useful, accurate and efficient answer, Machine Learning can be of great value. A machine learning algorithm uses a variety of data input and rules to learn and perform a specific task. The output of the algorithm is adding new information by training from historical data, learning from consistent patterns. Deparday (2018) mentioned the following: "A model is only as good as the data used to train it". In Disaster Risk Management, machine learning is used to classify imagery and satellite-based data, but also different types of sensors and records of social media. Combining these different disciplines could solve more significant complex issues and support experts in their preparedness and mitigation activities to disasters (Deparday et al., 2018). In Forecast-based Financing, the combination of triggers with a specific threshold value lead a decision. As mentioned before the triggers and threshold values need to be agreed upon by local governmental agencies, NGO's and local communities. For decision makers to make an actionable decision to release funds and to start with the EAP the decisions make use of a binary classification of droughts. The outcome is either true (A drought is expected) or false (no drought is expected). In this binary classification system, an algorithm can support the decision process and automatise the fund release. One of the algorithms suitable for this purpose is the decision algorithm, which identifies a relation between cue values (indicators) and a decision. Within the decision algorithm, there is a distinction between compensatory (making use of all available data and cues in the data set, such as regression, random forest) and non-compensatory algorithms (making use of a partial subset of the cue information, such as a decision tree) (Phillips et al., 2017). Although the decision trees use partial information, the degree of simplicity varies a lot. Some say the 'more the better', however, in complex decision trees this can result in overfitting of the data. Dataheroes, Kastrun (2017) developed a machine Learning algorithm flowchart which links the best possible algorithms with the intention of developing a model (attached in appendix C). In the humanitarian sector, the decision should be simple, clear and accurate (Guimarães Nobre et al., 2019). One of the models meeting these requirements is the fast-and-frugal tree, explained in more detail in the next section.

4.4.2. Data preparedness

This section elaborates on the data preparedness and the data analysis done to identify the spatial distribution of the different locations. A land use map classified by GlobCover [2] mask the data over the rainfed agricultural land use illustrated in figure 4.5. The selected land use mask exist of three land use types: Maize, pulses, groundnuts & cassava, secondly, maize, sorghum, pulses & cotton and last maize, sorghum, sulrush millet, groundnuts & guarbeans presented in figure 4.6. This mask identifies the drought impact on smallholder farmers. The agricultural mask is derived in QGIS by selecting the districts Chikwawa and Nsanje as the boundary layer. Furthermore, all daily data are monthly averages.

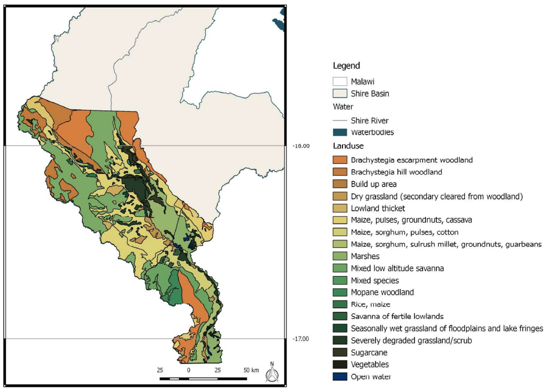


Figure 4.5: Landuse Nsanje and Chikwawa

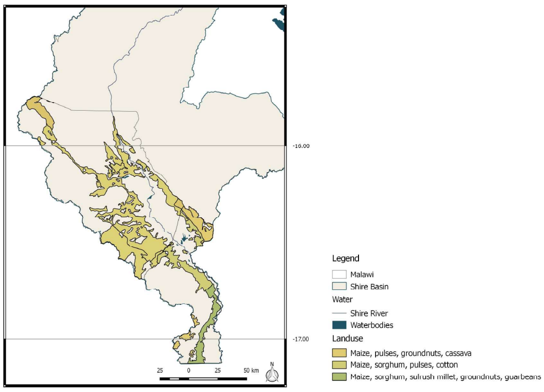


Figure 4.6: Land use mask rain-fed agriculture Chikwawa and Nsanje district

The FFT model does not use all data sets in the final model, based on the results of the Pearson correlations and the feature rank method in the *FFTree()*-package a top-five feature selection predicts droughts. The Pearson correlation defines the relation between the different features and excludes features creating noise. The values of the features are either a flux or a state. ONI, SM, VOD, NDVI and LST are implemented as state values in the model. ONI July is the most critical month predicting the impact of an El Niño on the growing season according to local experts in Malawi. Precipitation and dry spells are fluxes cumulative over the growing season. Cumulative values are used to include the history of the previous month in the next one. Secondly, the scale of the analysis varies between 11 km^2 (Mitole) and 2068 km ^{2} (Nsanje and Chikwawa). Over the period June 2002 - December 2017 there are 15 full growing seasons (not including season 2017-2018), in other words, the model can use 15 samples in total to predict a drought at the end of the growing season. In 2012 there is a data gap for which it is not possible to predict drought in the growing season 2011-2012. The total number of samples to work with is 14.

Time series analysis and anomalies

Over all seven regions, the time series provide information about historical climatic patterns for the selected agro-climatic indices. These time series show variation within the areas and their response to historical drought events. The selected indicators are plotted against each other to visualise their behaviour throughout 15 years. Besides the historical time series analysis the standardised anomalies (equation 4.1) of the indicators are calculated to define the deviation from the long-term mean divided by the standard deviation (equation 4.1 and 4.2). In this case, the long-term mean is over the 2002-2017 period.

$$z_{ij} = \frac{x_j - \overline{x_i}}{\sigma_i} \tag{4.1}$$

[2]GlobCover an ESA Initiative making use of a 300m MERIS sensor on board of the ENVISAT satellite mission

$$\overline{x_i} = \frac{1}{N}\Sigma r_{ij} \tag{4.2}$$

Pearson correlation

To determine the candidate agro-climatic predictors and predictands the correlation between the predictors is defined. To determine this relationship between variables the Pearson correlation method is used. This method is mostly used in diagnosing and selection of candidate predictors. If there is a large number of predictors, the mutual correlation could be leading to poor model performances (Hao et al., 2018). In this research, the Pearson correlation is calculated in python with the *scipy.stats*-package, which uses equation 4.3.

$$r = \frac{\Sigma(x-\overline{x})(y-\overline{y})}{\sqrt{\Sigma(x-\overline{x})^2\Sigma(y-\overline{y})^2}} \tag{4.3}$$

r is the correlation coefficient, x and y are the variables and $\overline{x}$ and $\overline{y}$ are the mean of the variables. To determine the candidate agro-climatic predictors and predictand a correlation coefficient higher than 0.70 is eliminated to reduce noise in the model.

4.4.3. Algorithm decision Fast and Frugal decision tree

Martignon et al (2008,) mentioned the Fast-and-Frugal tree (FFT) as one of the more simpler trees of the decision algorithms. A decision tree can be defined by a sequence of *nodes*, *branches* and *leaves*. In a FFT each node gives two branches with either an exit branch or a leaf and every node will give a decision-trigger (Phillips et al., 2017). The algorithm constructs a FFT according to four tasks: 1. Feature selection 2. Identify feature thresholds 3.Feature ranking and 4. Developing decision rule for each feature (positive or negative) shown in figure (4.8). One of the main reasons using the FFT is its ability to make a fast, frugal and accurate decision with limited information and without high-level knowledge of statistical training 2017, (Guimarães Nobre et al., 2019). Another reason mentioned by Phillips (2017) is the heuristic character of the FFT, meaning the model will perform well when the data contains uncertainties (Neth and Gigerenzer, 2015), (Keller et al., 2010). Other advantages using FFT in the humanitarian decision-making process is the fact it is transparent, it is easy to understand and when it does not perform well the error is easy to detect. This transparency is particularly useful when it is important to understand, implement, communicate and act quickly on a certain trigger.

Identify decision threshold for the predictand

The Fast-and Frugal model used in this research handles a binary classification system to identify a drought event yes or no. The exit branch uses a threshold value to classify a drought event. The thresholds used are the 20th, 40th and 50th percentile of the standardised anomalies for each predictand, represent the 20%, 40% and 50% largest negative standard deviations. The selected percentiles are based on the return period of a drought event and the number of available samples. Eliminating 'extreme' drought events with a five-year return period uses the 20th percentile over the 15 years. With these boundaries, the model will return three 'extreme' drought events used in the training and test set. The model performance will be more accurate with more data; there is no longer time series available, which decide to vary with the threshold value. The 40th and 50 percentile return six and seven drought events to train and test the model with.

Developing the code

The code is written in *R* using the *FFTree()* and *FFForest()* packages. First, the model calculates a classification threshold values for each predictor to maximise the weighted accuracy of the data set; this is done by *cuerank()*. The feature selection is based on the individual features and return the optimal threshold assuming a target value to be reached with that individual feature. The feature selection is followed by selecting the order (*order()*) based on their weighted accuracy. The model selects the top five features as input for the forecast model. The Fast-and Frugal Tree uses the *dfan()* algorithm in the growing process to calculate the threshold value and the exit structure. Instead of assuming cue independence as *ifan()* algorithm does, the *dfan()* iterative re-calculates the threshold values. This

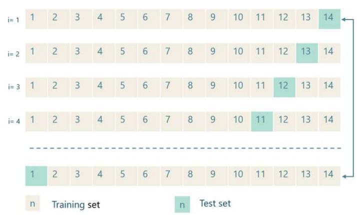


Figure 4.7: Principle leave-one-out cross-validation

iterative re-calculation detect specific predictive subsets. (Phillips et al., 2017) describe in detail the difference between theses algorithms. The tree is validated by pruning the tree from 1 to 5 leaves (figure 4.4). A pre-defined value in the model is the weighted factor *w.sens* , which is set to 0.75 to balance the sensitivity and specificity between the hits and misses (explained in the next subsection).

4.4.4. Performance Fast and Frugal trees

The Machine Learning algorithm needs a training and testing phase in order to test the performance of the model. Splitting the dataset create this training, validation and test set. The algorithm first needs to adjust the parameters to the task and to recognise patterns. Secondly, the validation set is used to determine the best performing model and last the test set is used to predict the criterion value with a new data set. A test set should be kept separately from the model in order to test the performance of the model to the accuracy (Deparday et al., 2018). Not in all models a validation set is used. At the end of the modelling process, the evaluation of a model should be based on the accuracy of the testing data, rather than on the accuracy of the training set. The data set used in the machine learning algorithm, need to be representative for the variability of the predictors and predictand (Deparday et al., 2018). Splitting the data can be either randomly by Monte Carlo cross-validation (10-90, 20-80, 30-70 per cent, commonly used in data mining), leave-one-out cross-validation or k-fold cross-validation. The importance of sample size splitting is dependent on the accurate measure needed: the level of accuracy and amount of variance. This research uses leave-one-out cross-validation as a splitting method. Keep in mind the training set is used to train the model, the more data, the better the model. If there is a large number of test samples, there is less variance expected in the results.

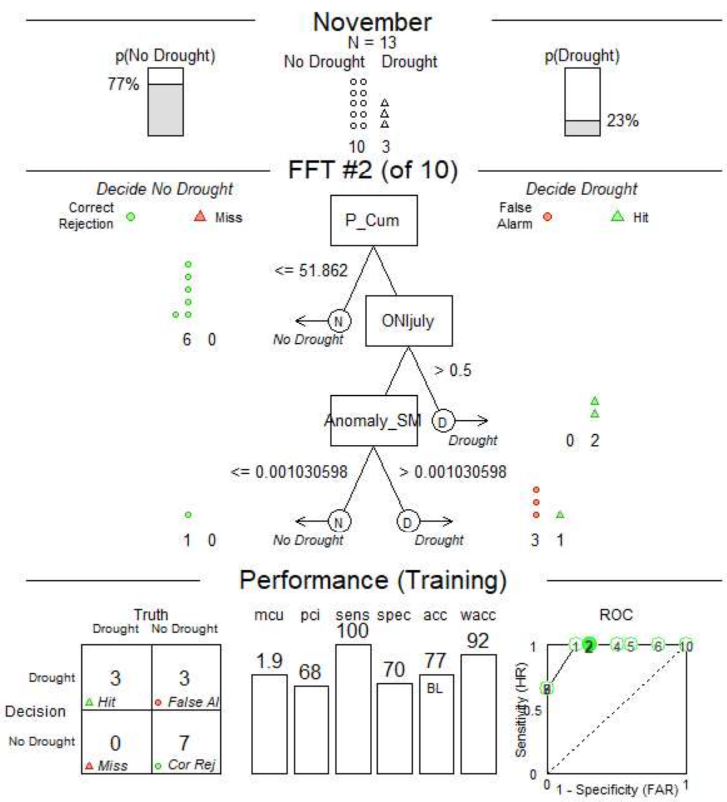


Figure 4.8: Concept Fast-and Frugal decision tree visualisation by *plot(dataset.fft, data='train')*. The illustration is split into three areas; the top area summarises the dataset including the number of samples and the ratio between negative and positive criterion values (no drought / drought). The bottom area represents the classification performance, including from left to right: confusion table, performance statistics summary of the plotted FFTree and the ROC curve of all FFTrees created plus competing for classification algorithms. The highlighted green dot is the is the currently plotted FFTree. The FFTree plotted is the one with the highest weighted accuracy (wacc) in the training phase, this model is the second best from the ten consttucted models for the month November and is pruned with 3 nodes. The performance statistics calculates; mcu: speed, pci: frugality, acc: accuracy, wacc: weighted accuracy, sens: sensitivity and spec: specificity.

The performance of a decision algorithm is expressed in a confusion table (bottom left corner in figure

4.8) evaluating its probability of predicting a criterion values. This confusion table classifies the following four outcomes: true positive (hits: hi), false positive (false alarm: fa), false negative (misses: mi) and true negative (correct rejection: cr). The confusion table combine measures of accuracy based on five areas: sensitivity, specificity, overall accuracy , weighted accuracy and balanced accuracy (equation 4.4, 4.5, 4.6, 4.7, 4.8). The weighted accuracy balance sensitivity and specificity representing the error trade-off, in the humanitarian sector, miss, have large consequences to the potential loss of life, which translates to a weighted factor above 0.5. (Phillips et al., 2017)

sensitivity:

$$sens = \frac{hi}{hi + mi} \tag{4.4}$$

specificity:

$$spec = \frac{cr}{cr + fa} \tag{4.5}$$

overall accuracy:

$$acc = \frac{hi + cr}{hi + fa + mi + cr} \tag{4.6}$$

weighted accuracy:

$$wacc = \frac{hi}{hi + mi} \times w + \frac{cr}{cr + fa} \times (1 - w) \tag{4.7}$$

balanced accuracy:

$$bacc : \frac{hi}{hi + mi} \times 0.5 + \frac{cr}{cr + fa} \times 0.5 \tag{4.8}$$

Another visual performance measure is the Receiver Operating Characteristic curve (ROC). The ROC-curve is used to describe a decision threshold and consider the cost/benefit analysis of the decision-making process, in terms of the sensitivity versus the specificity (**??**). This graph shows the trade-off between sensitivity and specificity (Phillips et al., 2017). The y-axis plot the hit rate and x-axis the false alarm. The most important parts of the ROC-curve are the (0,0) point, which represents a case with none positive classifications and the (1,1) point always experiencing a positive classification. Last point is (0,1) representing a perfect classification (ROC=1). The best performance can be selected by a point northwest (hit rate is high and false alarm is low) in the graph. The more a point is located in the upper right corner, the more humane, there is a high rate of hits but also a high rate of false alarm in decision making the ROC-curve can be of great value. The stakeholders should agree upon the hit-false alarm ratio and discuss the 'cost-benefit' of a decision followed by meeting this threshold. The ROC-curve shows a y=x line illustrating the randomly guessing of a class (ROC=0.5). When a point performs in the right lower triangle, the classifier performs worse than random guessing and can be neglected. When a classifier hits the y=x line nothing can be said about its performance; it does not give any useful information (Fawcett, 2005).

This research uses the weighted accuracy to specify the model performance with the confidence interval addressing the robustness. The confidence interval makes use of bootstrapping based on random sampling with replacement. With the Leave-one-out cross-validation for each of the 14 training and test iteration, the model returns either a False Alarm, Miss, Correct Rejection or a Hit, after these 14 iterations it sums the false alarm ratio (FAR), miss ratio (MR), correct rejection ratio (CR) and the hit ratio (HR) to calculate the WACC (equation 4.7). The model is bootstrapped over 14 times to calculate the WACC with a random order of leave-one-out samples. This process is iterated 1000 times so the confidence interval can be calculated over the mean ($\overline{x}$) and the standard deviation σ of the samples (equation 4.9). These values determine the lower and upper bound of the confidence interval. Figure 4.9 illustrate the leave-one-out and bootstrapping method.

The confidences interval for X is mean is calculated by:

$$\overline{X} \pm t\frac{\sigma}{\sqrt{n}} \tag{4.9}$$

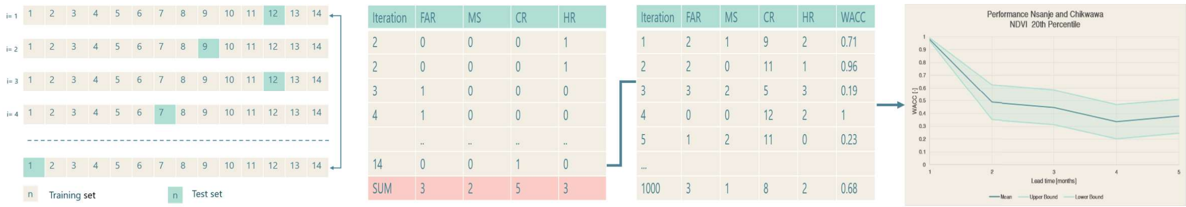


Figure 4.9: Bootstrapping over the WACC to calculate the confidence interval.

4.4.5. Final output

The output of the model returns the best combination of features in relation to the target value over the most critical months within the growing season to put EWEA in place. The outcome shows the best lead time for drought classification in March and which features are used for that lead time to perform a certain level of accuracy. In the end, this research compares the model performance of multiple model scenarios for each district. Table 4.4 shows the different scenarios for one of the districts, the y and x-values represent the selected candidate predictors and predictand based on qualitative and quantitative research.

Lead time	Predictor	Predictand	Lead time	Predictor	Predictand
1	$y_1, y_2, .., y_n$	x_1	1	$y_1, y_2, .., y_n$	x_2
2	$y_1, y_2, .., y_n$	x_1	2	$y_1, y_2, .., y_n$	x_2
3	$y_1, y_2, .., y_n$	x_1	3	$y_1, y_2, .., y_n$	x_2
4	$y_1, y_2, .., y_n$	x_1	4	$y_1, y_2, .., y_n$	x_2
5	$y_1, y_2, .., y_n$	x_1	5	$y_1, y_2, .., y_n$	x_2

Table 4.4: Model scenario's performed for one research area. Lead time March (1 month), February (2 months), January (3 months), December (4 months) and November (5 months). y_n and x_n represent the predictors and predictand used in the model.

5

Qualitative Results and Discussion

The qualitative data collection is used to explore the problem areas and to create an understanding of the information processing behaviour of both humanitarian organisations, agricultural officials and smallholder farmers. The semi-structured interviews and focus group discussions provide information about current drought forecast models and procedures used by the stakeholders, the weather information shared with smallholder farmers and understanding the agricultural practices. The questionnaire and protocols used in the semi-structured interviews and focus group discussions are presented in appendix B. The results of the qualitative research are used in the quantitative research phase to build a model designed to the needs of the end users. This chapter elaborates on the information processing behaviour at a national and local level. The question to be answered in this chapter is:

> *How do humanitarian actors and smallholders farmers currently process information on drought forecast?*

This question is clarified by first identifying the actors involved in current drought forecast procedures, followed by creating an understanding of the decision-making process of the humanitarian sector and smallholder farmers. The decision-making process of the smallholder farmers depends on their experiences of past droughts impacting their livelihood. These experiences are used to determines the spatial and temporal resolution necessary in the optimisation process of a drought forecast at national- and local level.

5.1. Information process behaviour in humanitarian organisations

In disaster preparedness and response, multiple actors work together, ranging from meteorological services to governmental institutes and from humanitarian organisation to local communities. In the development of an EAP different stakeholders are brought together to reach a joint agreement on who is responsible for which early action related to a specific hazard. All actors aim to act as soon as possible when a threshold is triggered. The different humanitarian organisation work alongside each other and unite with governmental organisations. In the development of an EAP or early warning system data is used at different scales, global, national and local level involving both international and national stakeholders. The MET Office provide meteorological data, whereas ideally farmer's organisations and local communities give officials indigenous knowledge in order to tailor the procedures according to the impact of a natural disaster on their livelihood.

This section elaborates on the stakeholders experienced and involved with developing a drought forecast for humanitarian purposes. The interview aims to collect information on the information processing behaviour in drought forecast over disaster preparedness and response phases. The nine humanitarian organisations interviewed are presented in table 4.1. These organisations are familiar with disaster risk reduction and/or forecast based financing. The International Federation of Red Cross and Red Crescent Societies (IFRC), Climate Center and 510 an Initiative of the Netherlands Red Cross has experience with Forecast-based Financing pilots in various countries, mainly focussed on floods, cold waves and typhoons. They gave information about current forecast procedures, the decision making factors and the time frame needed to act upon a specific trigger. Besides the international humanitarian organisation developing the Methodology of such systems experiences on the ground give an insight into project examples. WFP, FAO and the WHH more detailed information in specific areas.

Current forecast procedures

One of the first comments to develop a Forecast-based Financing procedure is to check if there is already an existing system available. The aim is to integrate existing procedures into the FbF protocols, rather than to develop a new parallel system (Climate Center, personal communication, 28th of August 2018). Current early warning systems primarily exist of meteorological forecast using among other trigger indicators, such as temperature and rainfall. With slow-onset hazards, contributing factors in the early warning systems are food security and livelihood indicators. Other commonly used indicators are NDVI, WRSI, ASI and market prices (figure (5.1)).

Increasing the reliability of a forecast model the use of different data sources is essential. Current drought forecast procedures make use of global, national and local data sources combining different resolutions. Another critical factor to take into account in a forecast model is the vulnerability on the ground. According to the FAO, without layering data information, it would be impossible to target the early action and understand whether it is meaningful to act. The combination of meteorological information from the MET Services and local data from the ground results in a two-way validation of a forecast. Historical time series analysis is used to test the reliability of local data sources by analysing if local sources manage to catch signs of past hazard. In these historical time series, there are certain points in time to revert to and validate the data. The quality of the data and the data availability strongly depend on the country. Every country uses different indicators in their forecast models, due to different circumstances. For example, in Mongolia the FAO hit the jackpot with available data sources, they had access to global, national and regional data. The local MET Office had a model available combining 15 different indicators (Personal contact, 13th of August 2018). When there is data available at different levels (global, regional and local) triangulate the data to understand what is happening at those three different levels. If not only one data source tells the story, but multiple sources at different levels (global, regional and local) the data is validated. Then there is enough information to set something in motion.

Decision making process

In the decision making process humanitarian organisation use different systems. Some organisations use an automatic trigger system and others a semi-automatic system. The design of an automatic trigger system release funds as soon as a danger level is reached (Climate Centre, personal communication, 28th of August 2018). In the semi-automatic system the trigger levels are checked by humans before funds are released ((WHH, personal communication, 9th of August 2018) and (FAO, personal communication, 15th of August 2018)). The experience from the different humanitarian organisation is that current systems do not always give an alert on time. Which makes one of the key aspects, how to improve the indicators of the already existing drought early warning systems and improve the results. National Societies mentioned the importance of combining scientific indicators, with the experiences on the ground. The indigenous indicators should align with modern scientific indicators. Once both thresholds meet, action can put in place to act on the predicted hazard.

> "You cannot intervene when the community says it is a drought. There is a discussion going on about who is the right person to declare a drought? Is it the government? Is it the NGO? Is it the International Red Cross and Red Crescent Movement or the Community? We argued that it must be the community; they should be the reference point if there is drought or not." - Ethiopian Red Cross Society

Within the International Red Cross and Red Crescent Movement, the decision-making process operates once there is a risk assessment, and the impact of the hazard is known. They target the most vulnerable areas and the areas most exposed and impacted in the past. After the risk assessment, the organisation designs an intervention map used as a decision-making tool, which tells them which areas need early actions when exceeding the agreed threshold. When a certain threshold is reached it automatically release funds to the national Red Cross Societies, who give support in whatever form is decided before. The critical element of this system is understanding which actions reduces what risk. In principle, the implemented action takes place within the window between the forecast and the crisis. Another moment activating actions is once a crisis hit, depending on the direct need and the time-frame to put preparedness for response in place. After each activation or on a multiyear basis the stakeholders should resubmit an update of early action protocol. For each hazard the stakeholder develops an early action protocol for a whole country; nevertheless, within the early action protocol, different actions can be distinguished per region.

In the decision-making process, it is essential to know the local capacity and operational procedures when a disaster hits. Both identifying the hit of an event and how to mitigate such is evenly important. A trigger should not be designed only on the certainty of an event but needs to include a vulnerability assessment, which shows what impact an event will have on a household. This reverse back to the disaster risk equation:

Risk = Exposure x Hazard x Vulnerability

Time frame acting on a trigger

In the previous paragraph, the operational side of the story is brought up quickly. This operational part is important to identify the time frame needed to put in place early actions. The standard time frame to act upon a drought trigger is about two or three months [1]. If the organisation has many things in place the time frame can be shorter, the actual problems are the operational, logistical, administrative aspect and availability of staff. Early signs can occur five months before the disaster hits, but that does not mean organisations can put in place the agreed early actions, In the forecast a certain level of scale is needed, the longer lead time of the forecast the larger the margin of error (IFRC, personal communication, 15th of August 2018). The time frame of a drought forecast is contextually driven by the geographical differences, different forecast systems and also the agricultural industry (FAO, personal communication, 13th of August 2018). Another factor influencing the identification of correct lead time is the fact the meteorological phenomenon drought will occur before the agricultural drought (WHH, personal contact, 9th of August 2018). Defining the type of drought in the development of a drought forecast is crucial. Besides the time frame of acting on trigger levels, the development of early action protocol takes about one year. After the different stakeholder accepts a protocol, stakeholders come together and evaluate the process, they refine the threshold levels if needed, and might consider different actions (WFP, Personal communication, 22nd of August 2018). During this refining and evaluation process, the input at the community level is evenly important as the experiences from humanitarian aid workers. All organisation mentions this combination of scientific knowledge, vulnerability assessments and indigenous knowledge as a key aspect in the development of early warning early action systems. Another aspect mentioned by the WFP is the anxiety in predictions when talking about preparing and responding to a trigger (WFP, Personal communication, 22nd of August 2018). Every day people make a decision based on certain probabilities; they have a plan B in mind.

> "If I tell you that tomorrow when you walk out the door you have a 50% chance of rain, all of a sudden people realise it does not have to be a 90% chance to take a jacket; there are probabilities related to the impact. The chance that you are dying by a disaster is a higher worry. However, within that uncertainty people can operate, they should prepare and think about a plan B. What happens when I lose my job or what happens when I get a flat tire, people are familiar with this kind of probabilities and prepare for such " -WFP

5.2. Information process behaviour national level

The EPA Officers are responsible for different villages varying between 367 and 50 villages spread over multiple Traditional Authorities (TA's). The Extension Planning Areas (EPA) Officers and Agriculture Extension Development Coordinator (AEDC) are working at the agricultural district offices within the district, from where they coordinate the agricultural activities. The EPA Officer is in direct contact with the farmers and fellow staff members. They provide information to the farmers about climate adaptation, afforestation, rearing of livestock and advise in using the best agricultural practices. Nowadays in Malawi, they are facing the effect of climate change, the EPA Officers advise farmers on new and different agricultural techniques, such as irrigation farming, conservation agriculture, pit planting, applying manure, making swales for water harvesting and share information about suitable crop types. In Mikalango they experience a short rainy season, so the EPA Officers advise farmers to plant early maturing crops. This section discusses the information gathered by interviewing the EPA Officers; first, the weather information sharing behaviour

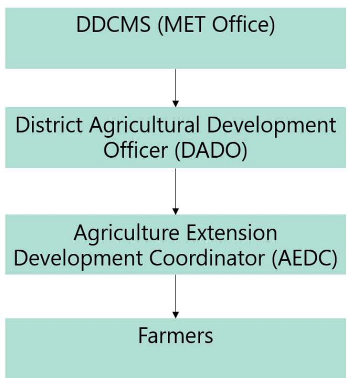


Figure 5.2: Information processing behaviour scheme

[1] Caution: It is difficult to put a number to the minimal lead time needed to put everything in place, it depends on different variables

Figure 5.1: Expressions commonly used in the semi-structured interviews when talking about drought forecast.

is mentioned, followed by defining the critical moment of providing farmers with weather information. The last section gives suggestions on how to improve the current weather forecast.

Weather information sharing behaviour

Two of the four EPA Officers mentioned the weather forecast information provided by the PRISCA project, a collaboration with FAO and the MET office (Department of Climate Change and Meteorological Services (DCCMS)). PRISCA is a participatory integrated climate service for agriculture, where the DDCMS provides weather-related information to farmers. This system involves farmers to understand weather predictions and support them to follow the procedures. (DCCMS and Farm Radio Trust provide the EPA Officers with weekly, monthly and seasonally forecast mainly communicated through WhatsApp messages and newspapers. At the end of August, the DCCMS share the seasonal forecast with the EPA Officers, providing the farmers time to decide on the best agricultural practices for the upcoming season. The weather forecast provided by DCCMS is simplified and translated from English to the local language (Chichewa) before shared with the farmers. The EPA Officers aim to give farmers exact but simple weather information. Sometimes the lead farmers and visitors from the districts come together; they discuss the weather information for a particular period. In general, the DCCMS give a weekly forecast, when they observe changes in the weather predictions they give an extra update in between. The information is used by farmers to take appropriate measures to cope with the effects of climate change using irrigation farming or starting small scale businesses. The farmers follow the advice of the EPA officers. The information is among others used to make plans for planting early maturing varieties. Saona Gama in Zunde remembers a season where a two-week dry spell was predicted; this information prepared the farmers on time to plant drought-resistant crops. Figure 5.2 illustrate the path weather information travels to the smallholder farmers.

Critical moment in the growing season

There are three moments within the growing season, where the weather forecast is crucial. At first, during and after land preparations, secondly, when the farmers receive the first rain and that farmers can start planting (Advice to start planting if 30 mm of rain is received). The last moment in time is when crops are about to mature. It is important for agricultural officials to inform farmers properly and before the planting season. In that period they provide the farmers with advice on which crops match the forecast. Another moment for important weather updates is in January/February when dry spells occur (mid-season).

Improvement current weather forecast

The priority in sharing weather forecast is sharing the message as quick as possible. Communication delay, due to lacking internet or access difficulties in reaching the more remote areas. All the EPA Officers mentioned the need of working phones and the distribution of radios. It would be beneficial for farmers to sit together, listen to the forecast and discuss the received information with each other at a set time of the week and support farmers adequately. Based on this discussion they can make better farming decisions and improving their harvest. Also, regional specific information is valued; the DCCMS made improvements in their weather forecast last season. The farmers valued this improved location-specific forecast. Other suggestions are the involvement of other ministries and sectors who can share information, such as the education and health department.

> "I am not all the time available to assist farmers with information; sometimes I lack WhatsApp data when this happens farmers will not get the required information in time." - Saona Gama

5.3. Focus Group Discussion

In the period from the 19th of November until the 23rd of November, four FGD were held. The District Agricultural Development Officer (DADO) selected the villages located in the most drought-prone areas. The four areas used in the quantitative data analysis cover the visited villages presented in figure 5.3. The FGD provided information about the agricultural practices, the climate-risks in the growing season and how the farmers receive weather information. The FGD facilitator summarised the information given by the farmers into crop calendars for a normal season and an unusual season. These crop calendars give an overview of the practises and risks in time, identifying the minimum lead time needed in early warning systems to prevent food security. All crop calenders constructed within the FGD are presented in appendix D. The crop calendars are developed by asking question about the main agricultural practices, what they expect to see on a field/corp when there are good weather conditions, what are the key climate-related risks, what they expect to see on the field/crop when those climate risks happen and last what other non-climatic factors affect the mentioned agricultural practices. This section discusses the results of these crop calendars.

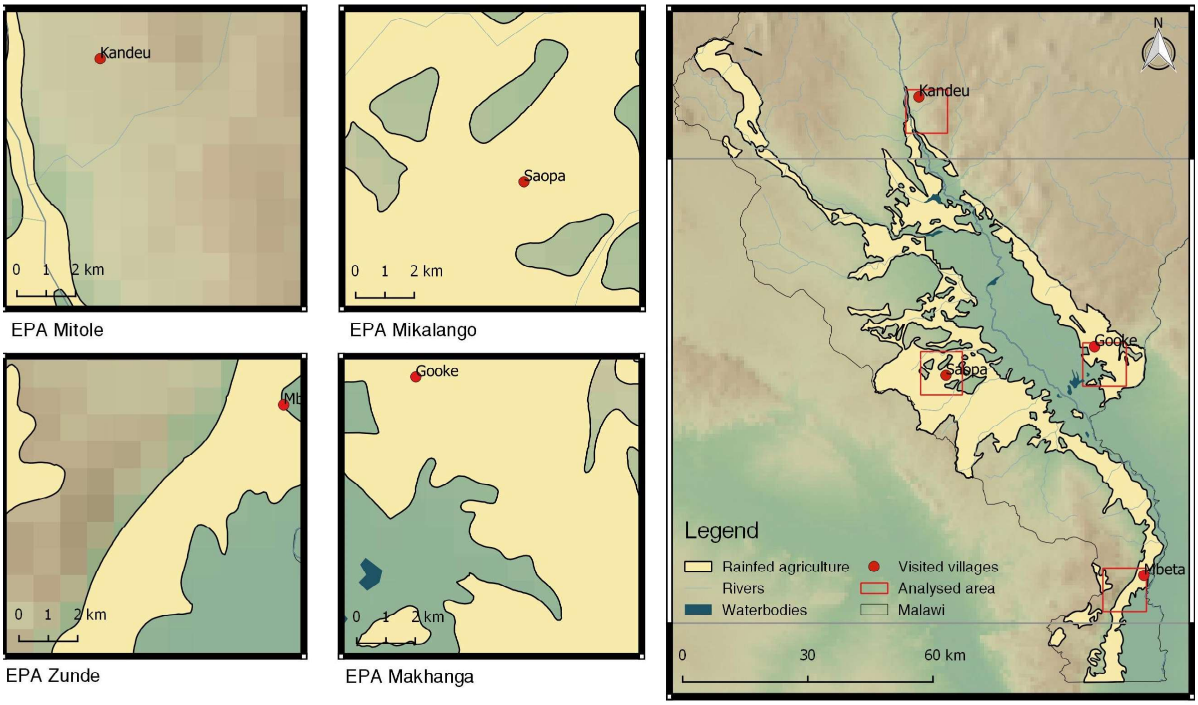


Figure 5.3: Location Field Visit Chikwawa and Nsanje district

Some background information about the visited areas; the visited locations show differences in climate, topography and geology presented in chapter 3 ((section 3.1.2), which influence the agricultural practices and the impact of a drought event. The main crops in the Lower Shire River Basin are maize, pigeon peas, sweet potatoes, beans, pumpkin, millet and sorghum. In Mikalango they also harvest cotton as a cash crop and in Zunde rice is grown along the Shire river. Harvesting takes place in March and April; March for cotton and millet and April mainly for maize. The farmers in the visited villages have an average farm holding between 0.5 and 2 acres.

Agricultural practices

The first agricultural practice mentioned is land preparation and manure making taking place in June/July or October. It takes about three months to apply the manure after making, and it should be ready when the rain comes to boost the fertility of the soil. In all areas, the farmers finish their land preparation before the first expected rains between November and December. Farmers experience the effect of a changing climate on their agricultural practices. All villages mentioned a shift in the start of the rain season. In Nsanje the land preparation was done between August and September, due to changing rainfall patterns this shifted to October. After land preparation farmers construct ridges in September/October in Mahkanga. In November they are digging pits for planting and waiting for rains to come in December and January. January exists of applying fertiliser and spraying pesticides to protect the crops from damages. In Nsanje the rain always starts in December the rain in November is unreliable. Village head in Zunde explained the agricultural practices in the past. They started with their land preparation in July, making ridges in August and September and the rain start in October. Now due to climate change, they receive planting rains in December. Nowadays, farmers plant already before the first rain increasing the probability the seeds receive the rain immediately. After planting, weeding takes place in the early growing stage in February. The general growing process is illustrated in figure 3.5

Figure D shows the start of the growing season, which used to be in October in the past, Nowadays this moved to November/December, shortening the growing season. There is a small difference in receiving rains between Chikwawa and Nsanje, in Nsanje the first rains are expected in December and planting takes place in January, while in Chikwawa this is a month earlier. The farmers experience the millet and sorghum being more drought resistant; they can withstand harsh climate conditions.

First rains on time characterise a good season, the first rains is an indicator for the rest of the season, which can give hope for a good harvest at the end of the season. People are encouraged when the weather conditions are healthy; they prepare for yield storage and work hard during the season. Signs of a good season are the observation by good soil moisture value and high germination rates in January. Healthy crops have large green leaves, sturdy stalks observed between January and February and large cobs and potatoes at the end of the season. For cotton, the amount of cotton balls is an indicator of a promising good harvest (figure 5.4).

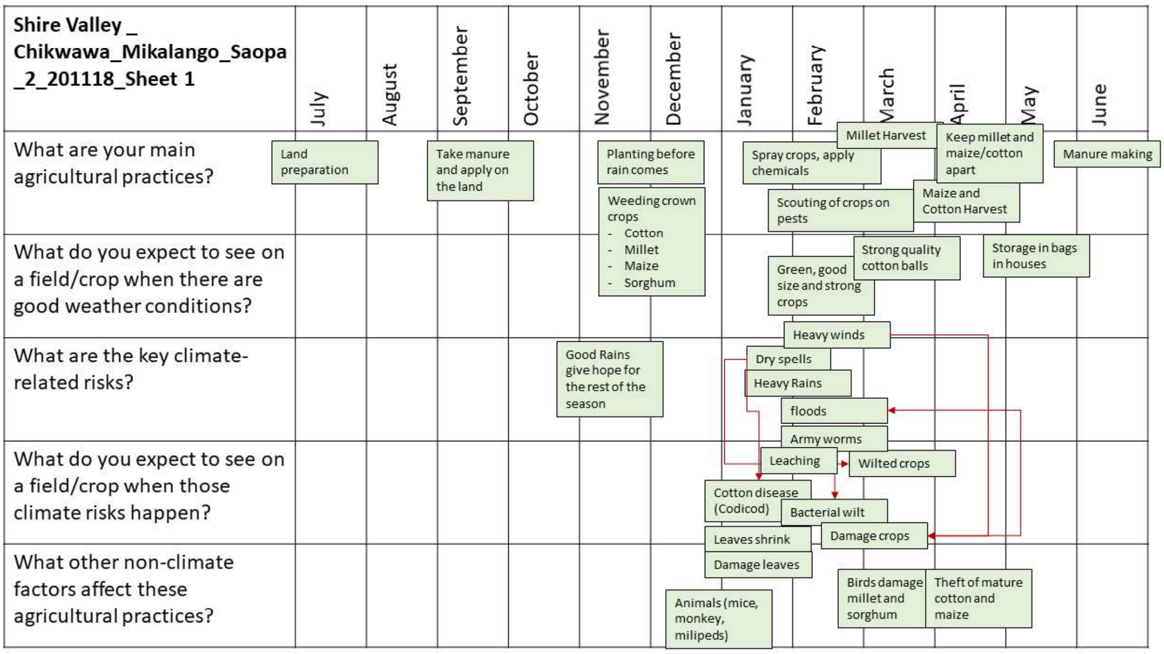


Figure 5.4: Example of one of the developed cropping calenders

Climate risks

The climate risks mentioned by the farmers are erratic rains, dry spells, strong winds and floods. The lack of rain mid-January, early February affects crop growth. Dry spells most often occur in the tender stage of the growing process in January and February. Dry spells of seven days damage crops like maize and cotton, which normally happen in February. In Nsanje farmers experience two or three weeks dry spells in January and February. Millet and sorghum can withstand harsh climate conditions and give yield at the end of the growing season when maize does not give any harvest. Besides dry spells in January and February heavy rains occur in that same period causing floods destroying the crops in February. Another known climate risk is destructive winds occurring in January and March when crops are maturing and ready to be harvested. Winds and temperature are indicators for rain prediction, strong winds disturb rainfall patterns, and high temperature is an indicator for rain.

> "Wind damages occur in March, when the crops have matured and ready to be harvested in the coming days. It pains to lose crops at the maturing level"- Farmers Mahkanga

During the FBD the farmers discuss two seasons; In the seasons 2016/2017 and 2017/2018, the crop production was affected by weather-related risks. In the season 2017/2018 in Mitole land preparation took place in October. The previous years they experienced a sequence of dry spells and poor rains, which demotivated them for this particular season. Dry spells starting at the 1st of February until the end of the growing season without rain, by March most crops, wilted. In Mikalango the season 2017/2018 rain did not come between November and January resulting in drying up and wilted crops. In February the rain came and destroyed the maize yield. Farmers who planted millet had some small harvest at the end of the season. Dry spells and attacks of fall armyworms resulted in low yield at the end of the season. The farmers in Zunde Nsanje faced a sequence of dry spells in the 2015/2016 season, forcing people to plant their crops close to the Shire River. In June/July they planted in the wetlands which provided them with some harvest in October. The farmers experienced the effect of the dry spells from the previous season and decided in the 2017/2018 season to plant millet and sorghum in the upper areas and in the wetlands of the Shire River maize. Also in Mahkanga the 2015/2016 was experienced as a difficult season, they faced erratic rains. On the 7th of December farmers planted their crops with the first rains, but soon after planting the crops were hit by three-week dry spells. After the rain came in January, the farmers planted again on the 10th of January, but a new dry spell occurred resulting in low yield.

Non-climate risks

Another type of risks farmers face is the non-climatic risks indirectly related to weather conditions. The most mentioned non-climatic risk affecting agricultural practices is the attack by fall armyworms in February. These fall armyworms caused many problems in both the 2015/2016 and 2017/2018 season damaging the leaves and maize cobs. The fall armyworms and other animals damaging the crops are indirectly related to weather conditions. Pesticides were used to fight against these worms, but this did not work. Eventually, by using Nimo tree leaves, they could resist the worms. Pesticide distribution by the EPA Officers did not solve the problems with the fall armyworms. The farmers did not have the capacity or knowledge to apply the pesticides as instructed, as well as the EAP Officers could not reach every farmer to support them with the distribution of pesticides. Another indirect effect of drought is that there was not enough drinking water for the livestock of farmers. Farmers used to have livestock which they sold when they had a bad harvest to buy food. Due to more dry conditions, there was no water available to feed the cattle. Other non-climatic risks are an outbreak of cholera, and other deceases, animals damaging crops and the accessibility of fertilisers. Another reason risk is that farmers do not have enough money to buy the drought resilient and hybrid seeds. It has been proven over the season's fertilisers help farmers in higher productions. There is a clear difference between the farmers who used fertilisers and the ones who did not. The fertiliser subsidy programs in Malawi are only beneficial for the selected few farmers receiving support. The government provide coupons to the poor, elderly and unproductive farmers. The strong farmers suggest giving these subsidised fertiliser coupons to them so that they can increase their productivity. Last non-climatic risk discussed is the storage problems, when there is a good harvest, farmers store their farm products, which is attacked by weevils and other animals, some also mentioned theft as a problem.

Receiving weather information

The farmer receives weather information from the lead farmers and agriculture officials. Some farmers said they receive this information once a week, others argued they had seen the lead farmers only once. The frequency of the weather information differs for each area. The weather information and updates are essential

Figure 5.5: Impression focus group discussions

for the farmers, this year (season 2018/2019) the extensions officers brief the farmers on the expected rains and advise the farmers to grow hybrid, early maturing and drought resistant varieties, instead of local maize varieties. Another communication organ is the Farm Radio Trust, farmers in Mikalango listen to the radio on Tuesday and Friday receiving weather information and utilise their farming. In all villages the extension officers organise community meetings, providing weather updates. The EPA Officers encourage the farmers planting early maturing varieties, drought-resistant crops and adopting new farming technologies like conservation smart agriculture. The farmers need direct weather information in order to prepare for the climate risks, making agricultural decisions, such as buying drought-resistant seeds when the forecast predicts dry spells. The information supports the farmers developing planting scheme during bad weather conditions to gain at least little harvest. Nowadays, the DCCMS is improving their weather forecast and providing farmers with more detailed information as mentioned in the previous section (section F.7).

> "The better part this year is that the weather information has been localised, meaning that people know which areas of Nsanje will receive more rains than other. I think such kind of information is vital to us." - The farmers in Mahkanga

Besides the weather information provided by the agriculture officials, they rely on indigenous knowledge by the elderly. They mentioned indicators, such as the number of fruits and green leaves on the Mbwemba tree, ants, direction of the wind for three to seven days associated with a good rainy season and a few days with high temperatures tells them when rain is about to fall.

> "As locals, we know how we can predict the weather. Sometimes when we see ants, high temperatures and the movement of the hippo from the water to where people live, that people will experience heavy rains" - Farmers in Mahkanga

5.4. Conclusion

This chapter helps to create understanding in the information processing behaviour and the decision-making process of both the humanitarian organisation and the smallholder farmers. The interviews expose the early warning process and the crucial timing of the forecast within the growing season. Answering the questions *What actors are involved in current drought forecast procedures in Malawi?*, *Which decision making factors define the information processing behaviour of humanitarian organisations and smallholder farmers?* and *What spatial and temporal resolution is necessary for the optimisation of drought forecast at national- and smallholders farmer scale?* ends with the answer of the first sub-question:

How do humanitarian actors and smallholders farmers currently process information on drought forecast?

What actors are involved in current drought forecast procedures in Malawi? There are a lot of different actors involved in disaster preparedness and response projects, varying between meteorological services and governmental institutes to the humanitarian organisations and local communities. Humanitarian organisations work alongside each other to develop EAP in collaboration with governmental organisations. The national organisations develop methodologies and forecast systems, such as ODSS explained in the previous chapter 2 (section 2.3). The information processing behaviour from the DCCMS to the smallholder farmers goes through the DADO, EPA Officers and lead farmers. The DCCMS provides the meteorological data to the DADO and EPA Officers, who share the information with the lead farmers. Other actors involved in Malawi are the Farmers Radio Trust who shares weather data through radio and local, national societies from an international humanitarian organisation, such as MLRC, who act on a developed EPA or forecast models. All international stakeholders agreed on the essence of combining scientific indicators with experiences on the ground. Ideally, the humanitarian aid workers, smallholder farmers, MET Office and governmental institutes refine the threshold values and trigger levels together after a disaster season.

Which decision making factors define the information processing behaviour of humanitarian organisations and smallholder farmers? In the development of an EAP or forecast model, it is essential to first check if there is already an existing system available. The organisation prefer to integrate the existing systems, rather than to develop a new system. The familiarity of forecast models and agro-climatic indicators by the involved stakeholders will be implemented in the development of the forecast model. Another given is the layering of data at different scales (global, national and local). The FAO mentioned a two-way validation between meteorological information from the local MET Offices and local data on the ground. These qualitative results support the decision in the quantitative data analysis and development of the model. Current forecast models primarily exist of meteorological trigger indicators, such as temperature and rainfall. The international humanitarian organisations suggest combining contributing factors, such as food security and livelihood indicators to improve agricultural drought forecast. A decision trigger should not be designed only on the probability of occurrence of an event but should include the vulnerability on the ground. Other factors to take into account in the decision-making process is the local capacity and the operational procedures when a disaster hits. At smallholder farmers level the information processing behaviour is following the governmental administrative system. The DDCCMS provide the EPA Officers with a seasonal and weekly weather update. This information is translated from English to the local language, before it is communicated to the farmers. There are certain thresholds guiding the EPA Officers to advice farmers to start planting. The first indicator is when 30 mm rainfall occurs between November and December. Farmers use this threshold as an indicator for the start of the planting season. Nowadays, the planting rains shifted from October to November/December. In Nsanje this first rains starts a little later than in Chikwawa. In Nsanje they start planting in December because the rains in November are unpredictable. Another climate risk influencing the decision-making process is the expected dry spells in January/February. Based on weather information farmers decide on planting early-maturing or drought resilient seed, use fertilisers or implement new agricultural techniques. Due to different climate conditions farmers start to plant different types and early maturing, drought resistant hybrid varieties. In the information processing behaviour farmers benefit from a forecast system communicated at a fixed time, this gives them the opportunity to discuss their planting decisions supported by agricultural officials.

What spatial and temporal resolution is necessary for the optimisation of drought forecast at national- and smallholders farmer scale? The humanitarian organisation recommends lead time between two or three months needed to put in place operational and logistical aspects. The lead time is depending on the local capacities and availability of staff. Another reference is the moment EPA Officers should inform farmers about the upcoming weather conditions to prepare for possible climate risks. There are three critical moments in the growing season. At first, during and after land preparation, secondly, when the farmers receive their first rains and last when crops are about to mature. Frequent weather updates are essential to make sure farmers are informed correctly over the period. The forecast should be designed combining the hazard information and the time needed to prepare to reduce the impact of such a hazard. Barriers agricultural officials experience in the information processing behaviour is the delay in communication. Some remote areas are difficult to reach and they need to inform farmers in person when farmers do not have access to a mobile phone, which slow down the weather information. The last factor to take into account is the increasing margin of error the longer the lead time of the forecast. Information implemented in the quantitative research phase is the time needed to put in place the operational and logical aspects of humanitarian assistance. Furthermore, the need of forecast information at the start of the season, just before planting and during the critical month January/February when there is a high risk of climate risks (dry spells, floods and destroying winds). Figure 5.6 combine the crucial moment within the growing season when weather information is needed and when drought forecast should be in place. The crop calendar of a normal year is presented in chapter 3 (section 3.2.1).

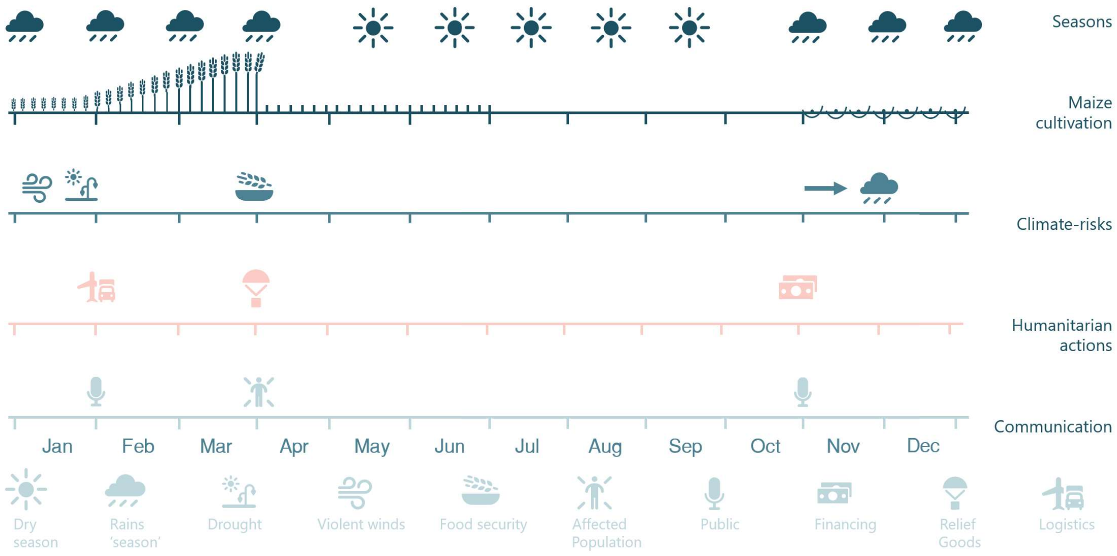


Figure 5.6: Critical moment of the year within the decision-making process

6

Quantitative Results and Discussion

The drought forecast model developed in this chapter makes use of the discussed qualitative data results from the previous chapter — these results support the analysis of the quantitative data results. This chapter first analyses the historical time series of the mentioned agro-climatic features; followed by addressing the correlation between these features. The historical time series analyses (section 6.1), together with the anomalies (section 6.2) and the correlation of the features (section 6.3) support the decisions made in the development of the drought forecast model, these sections together with the model performance answers the first sub-question *Which feature combination perform best in drought prediction at different spatial and temporal resolutions in the growing season?*. The last section 6.4 compares the model performance based on seven aspects of the forecast model for different model scenario's answering the question: *How do the different model scenarios perform in terms of accuracy?*. These results and discussion are used to answer the sub-question:

> *"What combination of agro-climatic features perform best in agricultural drought forecast in the Lower Shire River Basin?"*.

To answer this question the agro-climatic features are compared based on the influence of spatial resolution (ESA and VdS), threshold values predictand, type of predictand, top five feature ranking in the model, temporal resolution (lead time over the growing season) and spatial distribution. The confidence interval over the weighted accuracy of the forecast model distinguishes the robustness of the system. All graphs and tables supporting the discussion can be found in appendices E, F, G and H.

6.1. Historical time series analysis

A drought event can be characterised by the severity, spatial distribution and the duration, the return period, magnitude, predictability, timing and rate of onset (section 2.1) (Zargar et al.). These characteristics can be visualised in the historical time series of the selected indicators (figures E.1 - E.5). The climate seasonality in Malawi is visible for all features. The features show a clear difference between the wet- and dry season, with high peaks between November and March. This section compares the qualitative results with the historical time series analysis. In the previous chapters (3 and 5) the drought events in 2002, 2005, 2007, 2012 and 2015/2017 have been identified as 'extreme' events with high impact. In addition to these events, the farmers mentioned the drought events affecting their livelihood in 2010, 2015/2016 and 2017/2018. Another precondition discussed is the effect of the large-scale ocean-atmosphere system El Niño-Southern Oscillation (ENSO) on the precipitation pattern in Malawi. A strong El Niño in the seasons 2002/2003, 2009/2010 and 2015/2016 and a strong La Niña in the seasons 2007/2008, 2010/2011 and 2011/2012 influenced the climate patterns in Malawi(figure 3.6). The very strong El Niño visibly affected the precipitation pattern in 2015/2016, with exceptional low precipitation values varying between 90-150 mm/month. In a normal year, the precipitation rate reaches 250 mm/month. Another low peak occurred in the wet season of 2004 and 2009 influenced by the El Niño in 2002/2003 and 2009/2010. Contradicting to these low peaks are the high peaks occurring in 2012 related to the La Niña year 2011/2012. The maximum precipitation values between the seven areas vary between 258 (Mikalango) and 329mm (Zunde). Nsanje (319mm) give higher maximum values than Chikwawa (284mm). The plot for the area Nsanje and Chikwawa together show a more random graph, besides the high peak in the wet season it gives some smaller peaks in between the high peaks (figure F.7). The magnitude of the peaks is smaller compared to the peaks of the separate areas. Zunde follows this similar more random irregular pattern (figure F.7). Plotting the seven areas in one graph show the spatial variance within the different areas. In 2013 present three precipitation levels, the lowest peak is the combined Nsanje and Chiwkawa region, the middle level corresponds to Chikwawa, and highest precipitation rate takes place in Nsanje. The influence of averaging the precipitation data over the area is visible in the combined area of Nsanje and Chikwawa. The higher precipitation rate in Nsanje levels out when combining the two regions, resulting in lower precipitation values for the combined areas, meaning the intensity of a rain event is levelled out. Zunde show relatively low precipitation values in 2003 and 2007 compared to the other areas. The magnitude, duration and intensity within the areas do not return in clear differences between the precipitation pattern of Nsanje and Chikwawa (figure 6.10). These precipitation patterns indicate a difference in climatic conditions between the two areas.

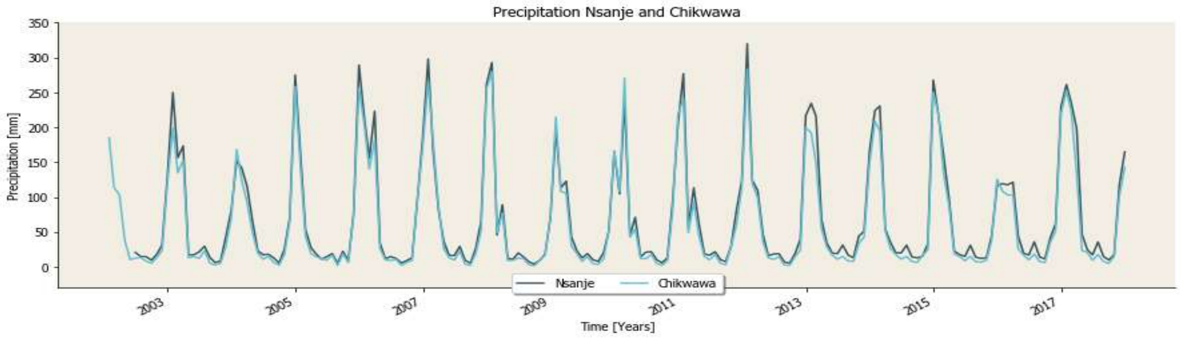


Figure 6.1: Time series precipitation Nsanje and Chikwawa

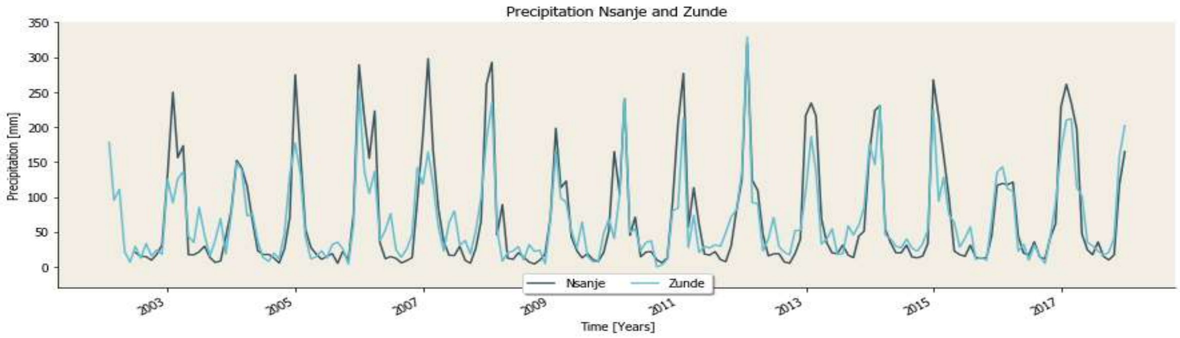


Figure 6.2: Time series precipitation Zunde

Soil Moisture

When looking at the next feature soil moisture in relation with precipitation a similar pattern in higher peaks during the wet season compared to the dry season is visible (figure 6.3 and 6.4 area Nsanje + Chikwawa and Mitole). There is a small delay reaching the highest value of the season for SM compared to the highest precipitation value. This time lag illustrates the soil response to the precipitation. Contradictory to this argument are the years 2004, 2009 and 2011 showing a different time lag, where SM peaks before precipitation. These values occur in a 'post-drought' year. The soil moisture plots give a second peak (lower than the first) starting at the end of the growing season until the dry season. This peak could indicate groundwater recharge after the harvest period. The lowest soil moisture values take place in the months February and March at the end of the growing season, just before harvesting when crops need water to mature (section 3.5). The soil moisture values are relatively steady for about three à four years; it seems like there is a certain degree of memory in the groundwater recharge reacting with a time lag on precipitation peaks. The years 2005,2006,2007 and 2008 show relative average or above average precipitation rates and, corresponding with the relative average soil moisture values in 2006, 2007 and 2008, the anomalies will be discussed more in section 6.2.

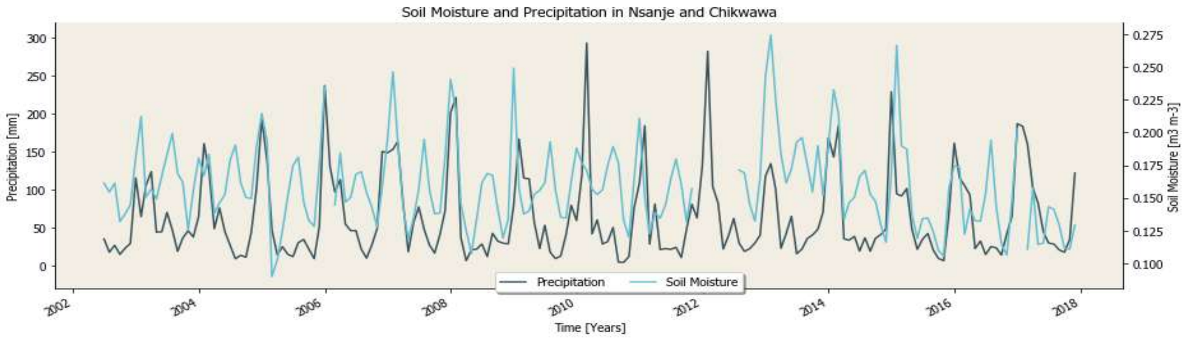


Figure 6.3: Time series SM and precipitation Nsanje and Chikwawa

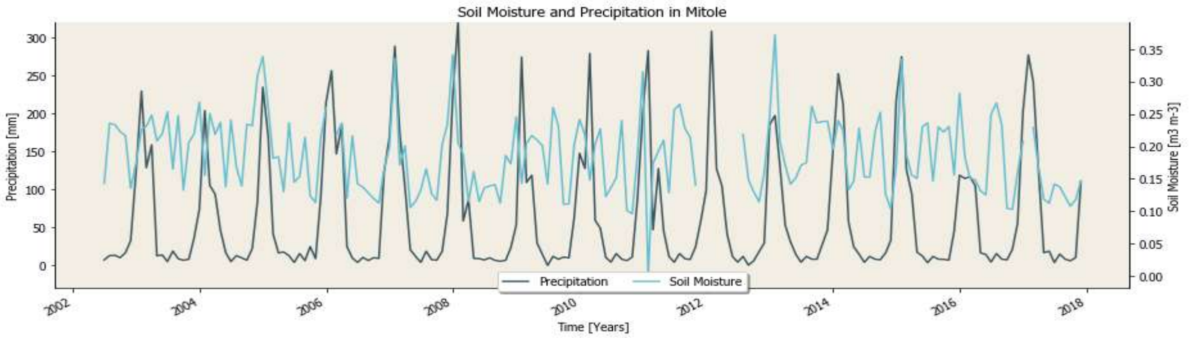


Figure 6.4: Time series SM and precipitation Mitole

Going back to the experience of the farmers and the natural hazard data from EMDAT and GDO the years 2002, 2005, 2007, 2010, 2012 and 2015/2016 experience impact from drought events. The soil moisture values show a sequence of relatively low standardised anomaly values over a period from 2015 to 2018, related to the strong El Niño and La Niña years (3.6). Such a sequence of low soil moisture values over a few years impact the livelihood of smallholder farmers even more than one year with low values. Comparing the different areas; Mitole reaches the highest soils moisture values up to 0.37 m^3/m^3, while the maximum value of the other areas varies between 0.27 m^3/m^3 and 0.30 m^3/m^3. Mitole shows heavier extremes compared to the other areas. The soil type, agricultural practises and altitude might affect these more substantial fluctuations. Mitole has a more hilly topography leaning against the plateau, the infiltration rate and groundwater tables influence these fluctuations. There is a small delay in reaching the soil moisture peak between Nsanje + Chikwawa area and Mitole. Mitole shows an earlier response to the start of the rainy season.

Land Surface Temperature

The third feature is the climate factor temperature. This factor together with solar radiation and precipitation influence the crop growth (Brouwer, 1986) (section 3.2.1). During the focus group discussions in the qualitative data collection phase, the farmers mentioned the temperature as an indicator to predict the start of the rainy season. They experience three very hot days just before the start of the season; this kind of daily information cannot be included in the historical time series. The highest temperatures occur during the dry season in July. Since 2012 the temperature gives a continuous higher value than the 2002-2011 period, possibly influenced by the change in the satellite in 2012 returning a systemic error. Although, the seasonality of Malawi climate with a dry and wet season explains the LST peak taking place before the precipitation peak. This argument does not hold for the second peak. This second peak takes place at the end of the rainy season between February and April. The increasing LST relate to the development of crops, in the harvesting stage the vegetation dries out and increase in temperature (Brouwer, 1986). The low precipitation rate in 2016 does not show higher or lower LST values compared to the other years, from 2012 onwards there is a low variation in LST between the years.

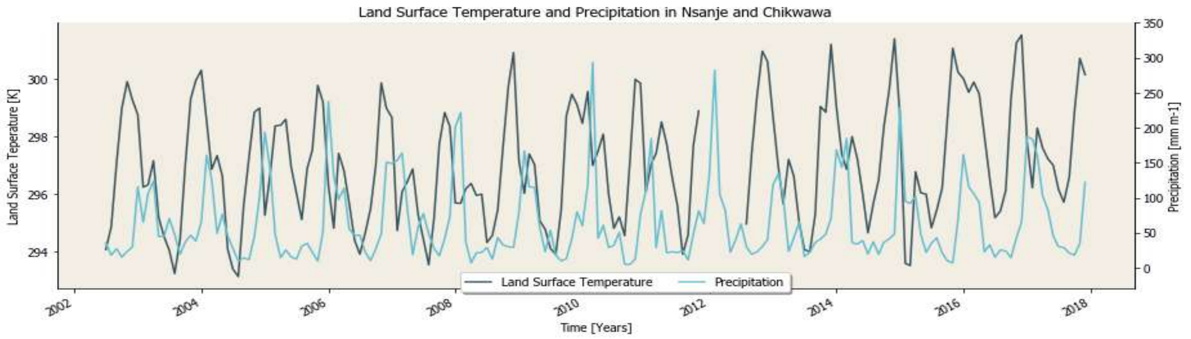


Figure 6.5: Time series LST and precipitation Nsanje and Chikwawa

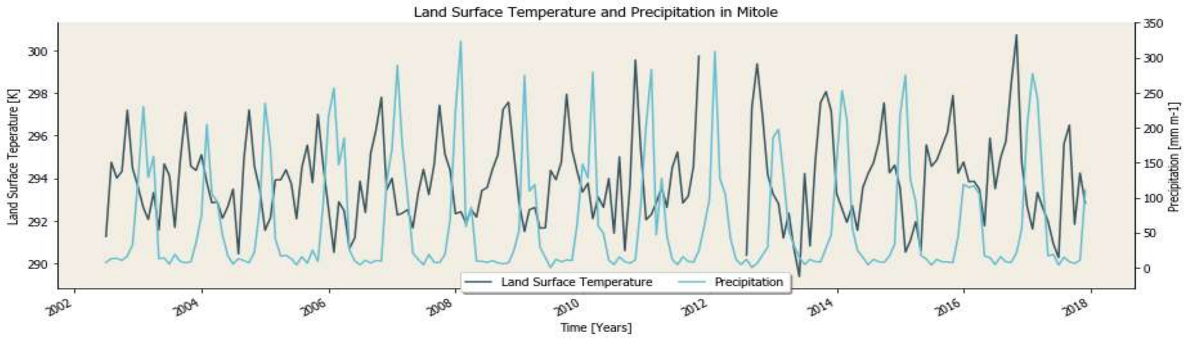


Figure 6.6: Time series LST and precipitation Mitole

Relatively low temperatures occur in 2005, 2008 and 2010, which are mentioned earlier as dry years. Chikwawa experience higher LST than Nsanje, within Nsanje, Zunde experiences slightly higher temperatures in the drought years 2005, 2010, 2016 compared to Mahkanga. In Chikwawa, the temperature difference between Mitole and Mikalango are relatively large and sometimes contradicting. There is a little delay in reaching the highest LST values between Mitole and Mikalango, Mitole reaches its peak earlier in the season. Although Mitole shows significant lower values compared to Mikalango, these low values do not translate to the whole Chikwawa area. The values of the whole Chikwawa area and Mikalango are comparable to the combined area Nsanje and Chikwawa. The combined areas return a little lower values, driven by the continually lower temperatures in the Nsanje areas. The LST values in relation with SM show a negative correlation; if the SM goes up the LST goes down, this should be seen in relation to the evaporation rates. Plants have a transpiration cooling effect, which could explain the negative relation between SM and LST (200, 2002). There is a slight difference between evaporation and transpiration in the growing process. Transpiration is the water escaping through the leaves of the plants to the atmosphere and evaporation is the water from the soil or on the leaves going to the atmosphere. At the initial stage of the crops, the evaporation rates are more critical, while at the end of the growing season the transpiration rates are more important. With dry harvesting crops, the transpiration rates decrease at the end of the season, explaining the increasing temperature at the end of the season (Brouwer, 1986) (section 3.2.1).

Vegetation indicators

The relation between temperature, solar radiation and precipitation on the growing process is visible in the historical time series. At first the influence of precipitation on the NDVI and VOD values. The seasonality of Malawi's climate and the growing season is visible in both the vegetation indices. The width of the peaks between the vegetation indicators NDVI and VOD is larger than the precipitation peak; there is a certain degree of crop growth response. The vegetation needs time to use the precipitation. NDVI peaks occur a little earlier than the VOD; they follow the same pattern and show the same amplitude. The observed greenness of the vegetation response quicker than the water content in the vegetation. Both indicators show low values in 2005, 2010, 2011, 2015, which correspond with the drought events mentioned by EMDAT, GDO and the farmers. All areas show a decreasing NDVI and VOD value over multiple years and a sudden jump from a low value to higher values starting over with a decreasing sequence. In almost all areas VOD experiences a lower value in 2015 compared to the NDVI value. Mikalango shows a significant difference in NDVI in the lower parts compared to the other areas.

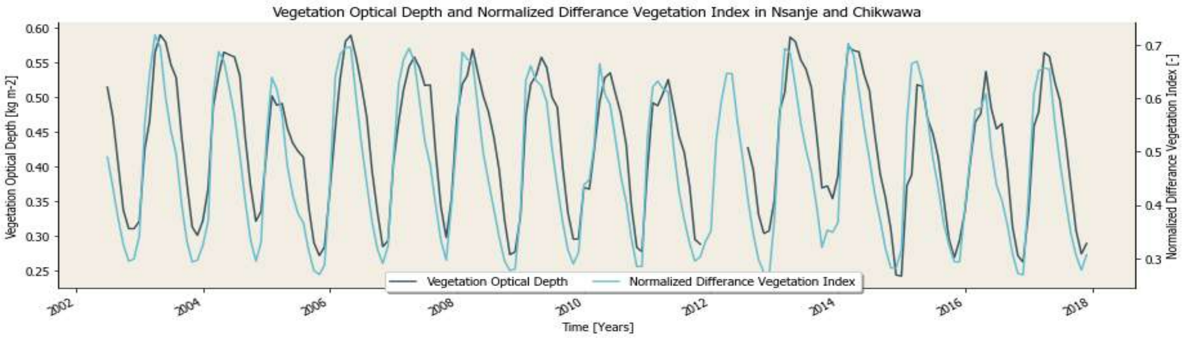


Figure 6.7: Time series VOD and NDVI Nsanje and Chikwawa

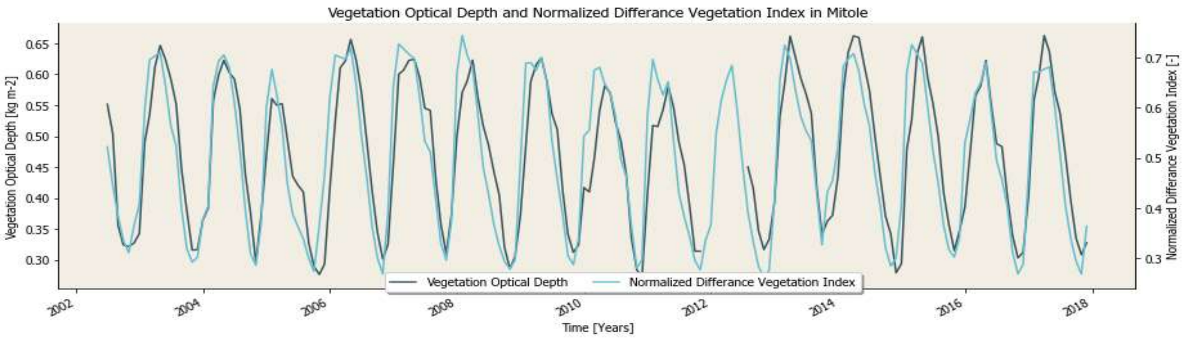


Figure 6.8: Time series VOD and NDVI Mitole

Comparing the time series of SM with the VOD and NDVI the vegetation response is visible. SM reaches its peak before VOD and NDVI reach their peaks, indicating a direct response of using the soil water content for vegetation growth. Another feature comparison is the relation between LST and the vegetation indices. There is a negative correlation between those features. When the LST increases the NDVI and VOD respond when LST reaches its peak. The vegetation uses the temperature for the growing process of the crops. The spatial difference between the NDVI and VOD values is visible for Mitole in the 2015/2016 season, Mitole does not show a substantial decrease in NDVI and VOD, although the farmers experienced low crop yields. The other areas return lower values of NDVI and VOD for the 2015/2016 season.

6.2. Anomalies

The standardised anomalies represent the standard deviation of the feature over 15 years from 2002-2018. Figure 6.9 presents the anomalies for the five main features P, SM, LST, VOD and NDVI in the area covering the rainfed agriculture in Nsanje and Chikwawa. From the seven areas, two areas Mikalango and Zunde have a larger y-axis limit varying between -3 and 3 standard deviation, mainly driven by the peaks of SM and LST in 2015. These peaks are also visible in the other areas, but with lower values reaching +/- 2 standard deviation. Within the different areas, Mitole shows large fluctuations over all five features (P, SM, LST, VOD and NDVI) compared to the others. The anomalies show a negative correlation between SM and LST, with high peaks in 2015. VOD and NDVI roughly follow the same standard deviation over the whole time series. In the years 2004, 2005, 2006, 2010 and 2015 most areas show large negative standardised deviations, meaning that those years have lower values than the average over that period over the whole time series.

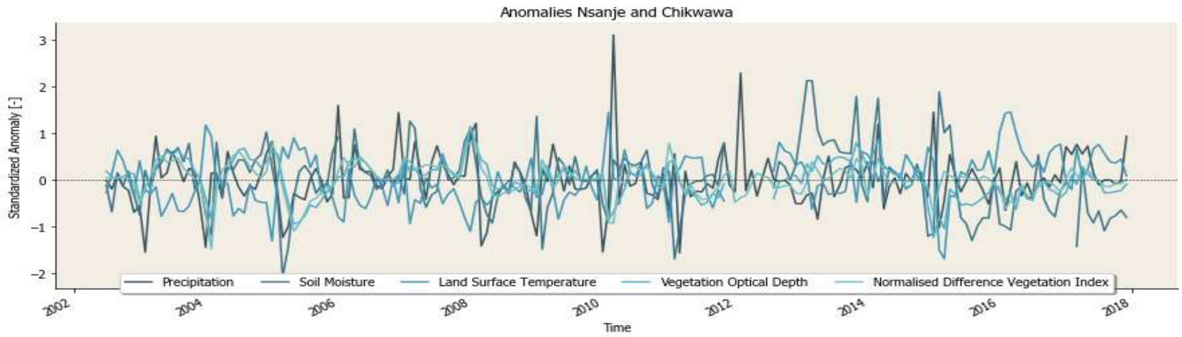


Figure 6.9: Anomalies Nsanje and Chikwawa area for all five candidate features

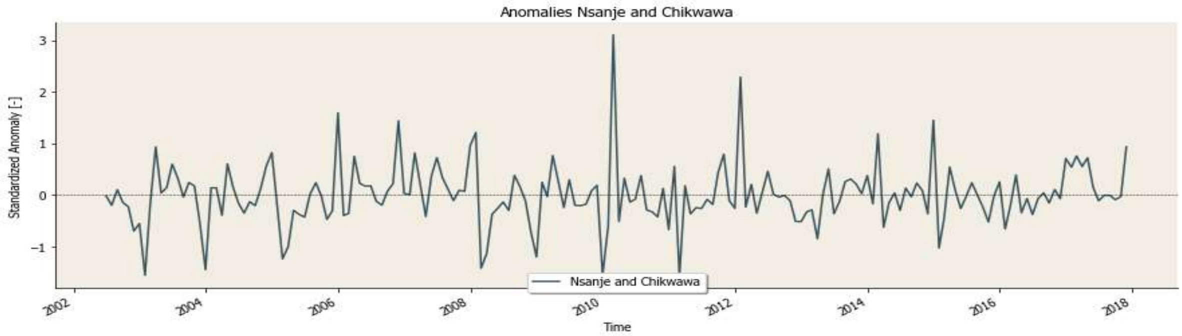


Figure 6.10: Time series anomalies precipitation Nsanje and Chikwawa

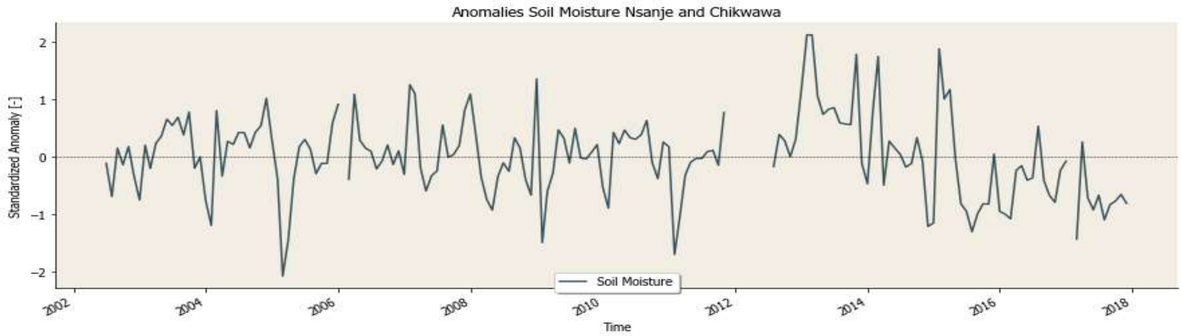


Figure 6.11: Time series anomalies SM Nsanje and Chikwawa

Another anomaly plotted are the predictands NDVI March and VOD March. The NDVI and VOD anomalies of March are used as predictand in the model to identify the drought classification at the end of the growing season. This drought classification is based on the threshold values representing the lowest 20%, 40% and 50% from the NDVI March and/or VOD March values illustrated figures 6.12 and 6.13 for the combined area Nsanje and Chikwawa. The graphs of the other areas are presented in appendix E. As mentioned in the methodology the percentiles are based on the drought extremes with a return period of five years (14 year period 20th percentile). Due to the low amount of data samples, the 40th and 50th percentile are calculated as well. These figures show the three percentiles derived from the NDVI and VOD March values at the end of the growing season. For most areas, the anomalies indicate 'extreme' drought events with a return period of 5 years in 2004, 2005/2006, 2010 and 2015.

The graphs plotting the anomalies show extremes in 2005 for all areas. With the largest extreme in NDVI Mitole giving a value of -3 standard deviations, while the others are between -1.5 and -2 standard deviation. The second extreme occurs in 2009 or 2010. The anomalies between NDVI and VOD in Zunde show significant differences compared to the other areas. The NDVI does not show vast extremes between 2002-2011, while for 2016 it returns an extreme value -3 standard deviations. The VOD show extremes with -1 standard deviation in 2005, 2007, 2010 and 2015 (figure 6.14), which is lower than the NDVI deviation. The NDVI pattern in

Mitole shows a significant negative peak in 2005 reaching -3 standard deviations, while until 2010 the value fluctuates close to the zero line. The VOD for Mitole plot a similar pattern except for the 2005 peak and the 2016 'drought' event visible for NDVI. In general, the percentiles for NDVI are a little bit higher than for VOD.

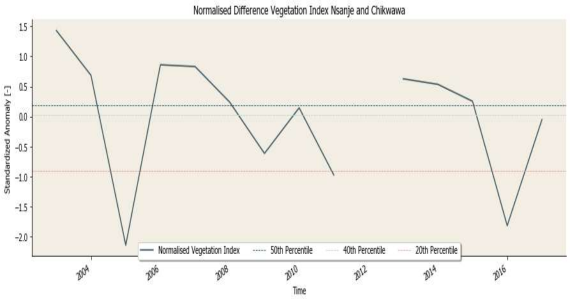


Figure 6.12: Time series predictand NDVI in March Nsanje and Chikwawa.

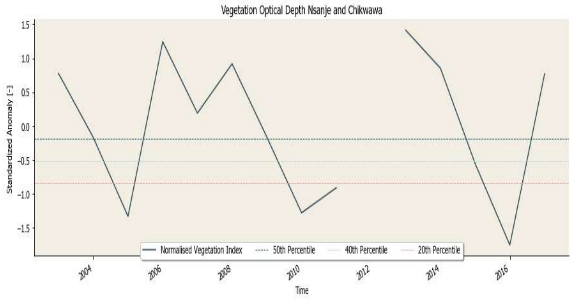


Figure 6.13: Time series predictand VOD in March Nsanje and Chikwawa.

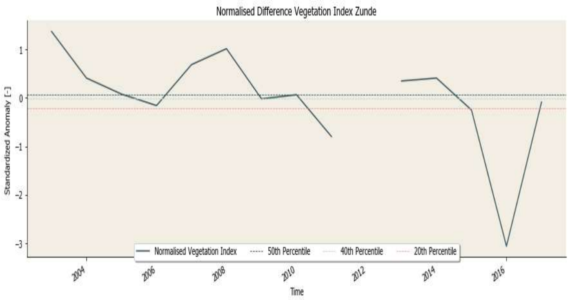


Figure 6.14: Time series predictand NDVI in March Zunde.

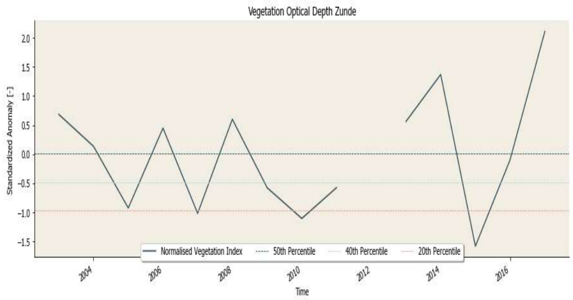


Figure 6.15: Time series predictand VOD in March Zunde.

6.3. Pearson correlation

The Pearson correlation is used to identify the linear relation between the selected features (section 4.3.2). The correlated features are: ONIjuly, P, SM, LST, Dry spells (Cat0, Cat1, Cat2 and Cat3 [1]), VOD and NDVI over the growing season and the state in March at the end of the growing season. All features correlated over the monthly mean, anomalies and cumulative values over the growing season. A heatmap presents the correlation coefficient (r) visualising the strongest and weakest correlation coefficient. Figure 6.16 presents the Pearson correlation matrix of Nsanje + Chikwawa. The correlation matrix for all the other areas is presented in appendix F.

The Pearson correlation analysis shows a few elements. First the correlation between the candidate predictands NDVI and VOD, between the monthly mean there is a strong correlation, varying between $r = 0.86$ and $r = 0.95$. The correlation between the anomaly of the NDVI/VOD with the NDVI/VOD of March is lower, returning values between $r = 0.77$ and $r = 0.17$. This low correlation factor for Zunde $r = 0.17$ might link to the low value of Nsanje $r = 0.33$. Zunde is located in the area Nsanje, averaging the extremes over the area; the low negative values in Zunde affect the average in Nsanje. The time series NDVI anomaly in March in Zunde show a different pattern than the other areas (discussed in section 6.2), this difference in standard deviation explains the r, there is a lower correlation between VOD and NDVI. VOD calculates the water content in the canopy while NDVI measures the greenness of the vegetation. Between 2002-2011 the anomalies are relatively stable, without extreme droughts, in 2016 an extreme 'drought' occur, the VOD March anomaly does show other extreme values. The r between LST and NDVI vary between the different areas. In Mitole there is a stronger negative correlation between LST and NDVI ($r = -0.73$) and Zunde ($r = -0.21$), meaning higher temperatures in Mitole, influence vegetation more negative.

The second feature explained is the correlation between the cumulative precipitation over the growing season and the predictands NDVI and VOD. The r between these features is similar for both and vary between

[1] Categories existing of a sequence of dry day (< 0.85 mm/day). Cat 0 = 0-2 dry days, Cat 1 = 3-4 dry days, Cat 2= 5-9 dry days and Cat 3 = more than 10 dry days

$r = 0.81$ and $r = 0.92$. When the precipitation rate increases the greenness and water content in the canopy increase as well.

Another discussion is the feature correlation between SM and the predictands NDVI and VOD showing a large difference when using the high-resolution SM data or the lower resolution SM data. The r between SM VdS and SM ESA range between $r = 0.48$ and $r = 0.63$, which does not show a very strong correlation. The correlation between NDVI and VdS SM over all areas return a r-value between $r = -0.09$ and $r = 0.19$, while the ESA SM data gives a r-value ranging between $r = 0.68$ and $r = 0.82$. The correlation between SM and VOD show similar values, for VdS ranging between $r = -0.10$ and $r = 0.05$ and ESA $r = 0.56$ and $r = 0.76$. Although the VOD data is also provided by VdS the difference between the r-value between SM correlated with NDVI or VOD stays the same. This low correlation coefficient for the high resolution data is not expected and therefore, these results should be handled with care, a more in depth data analysis is needed to understand whether these results are caused by some mistakes in the current analysis or that something else drive these results.

Third point: as mentioned in the time series analysis (section 6.1) LST and SM are negatively correlated with a value over the different areas between $r = -0.79$ and $r = -0.90$ for the SM data from VdS and $r = -0.41$ and $r = -0.75$ for the SM data from ESA. The correlation difference between the two datasets can be explained by the fact that the LST dataset is provided by VdS. They have the same resolution and use the same satellite data to construct the LST and SM data. The negative correlation means that when there are high LST values the SM values are low, with higher temperatures the soil moisture content is decreasing, with could indicate higher evaporation rates extracting water from the soil, decreasing the soil moisture content.

The last feature to discuss are the dry spells. The developing machine learning model uses one dry spell category as an input variable in order to reduce possible noise between the different dry spell categories. Cat2 and Cat3 show lower correlation with the other features, while Cat0 and Cat1 show similar r-values varying between $r = 0.57$ and $r = 0.82$. There is a slight difference between the Cat0 and Cat1 values. Over the different areas the r between Cat0 and P Cum, VOD Mean and NDVI Mean is in general more often higher, than the Cat1 r, but the mean r values over all areas is higher for Cat1 (Cat0 and Cat1 P Cum Mean $r = 0.69$, Cat0 VOD $r = 0.73$, Cat1 $r = 0.75$, Cat0 NDVI $r = 0.67$ and Cat1 NDVI $r = 0.68$). This higher mean value selects Cat1 as a feature in the machine learning model. The last comparison made, argued the spatial distribution in relation to the correlation coefficient. Within the r a wide variety between the different areas was visible for NDVI and VOD.

See Chat at conclusion

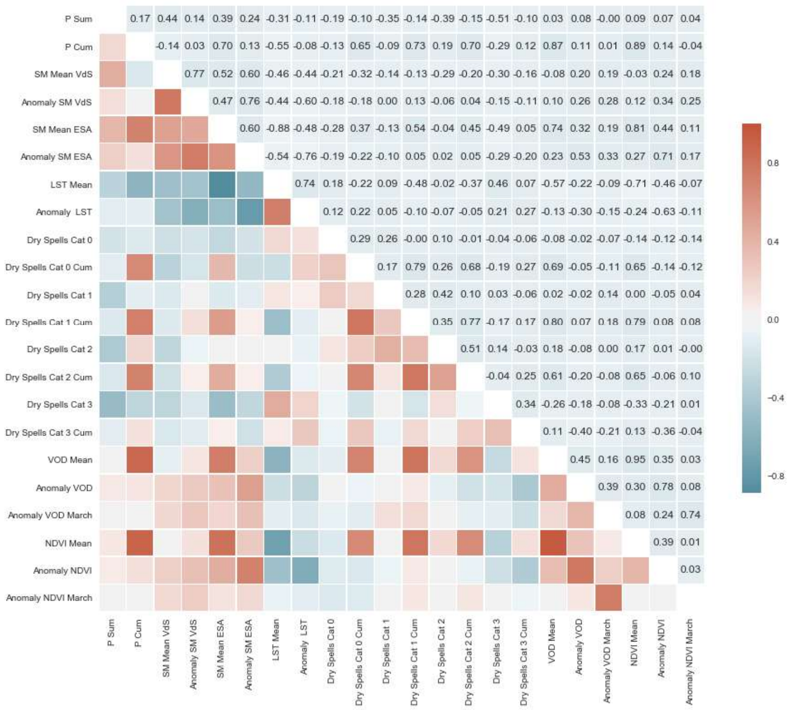


Figure 6.16: Pearson correlation matrix Nsanje and Chikwawa

6.4. Machine Learning forecast model

This section presents the result of the Fast-and Frugal decision tree (figure 6.17). The model performance is calculated with the weighted accuracy (WACC) over the growing season. The WACC exists of the False Alarm Ratio, Correct Rejections, Misses and Hits (equation 4.7 explained in Chapter 4 section 4.4.4). The weighted accuracy balance the sensitivity and specificity of the model, which is set to 0.75, giving more weight to the sensitivity of the model (chapter 4 section 4.4.4. Every area exists of 12 different models representing four models with three different threshold values. An overview of the main four model scenarios with the derived three percentiles is presented in table 6.1. In total this section analyses 76 different models over seven areas on various model aspects (ESA does not give sufficient soil moisture data to train and test the forecast model in Zunde). The model performance is divided into seven subsections comparing the influence of spatial resolution (ESA and VdS), threshold values, feature ranking, type of predictand, temporal resolution (lead time over the growing season), the confidence interval of the model, and the spatial distribution. The first subsection 6.4.1 compare high-resolution soil moisture data from the satellite company VdS and the lower soil moisture resolution from the ESA. Secondly, the model performance is measured by the effect of using different threshold values of drought classification (subsection 6.4.2), The third and fourth subsection (6.4.4 and 6.4.3) elaborate on the top five predictors, followed by the analysis of the model performance using different predictands. Subsection 6.4.5 compares the model performance in drought forecast with different lead time in the growing season. The confidence interval in section 6.4.6 discusses the robustness of the model. The spatial distribution of the different areas is described in the last (subsection 6.4.7).

The sections discuss the model performance according to the most critical times within the growing season to put in place early warning early action mentioned by the humanitarian organisations, EPA officers and smallholder farmers (chapter 5 section 5.1 and 5.3). On the one hand, humanitarian organisation need at least 2-3 months to put in place actions triggered by exceeding a specific threshold value predicting a drought. On the other hand, smallholder farmers need weather forecast at the start of the planting season (November) and mid-season when there is a high chance of dry spells, floods and destroying winds (January and February). For this reason, the model performance is analysed mainly focused on lead time 3 (LT3) and lead time 5 (LT5 (January and November)). Figure (5.6) gives an overview of the different time lines for each actor, showing the overlap of forecast needed in LT3 and LT5. Appendices H and G present all graphs and tables supporting the discussion.

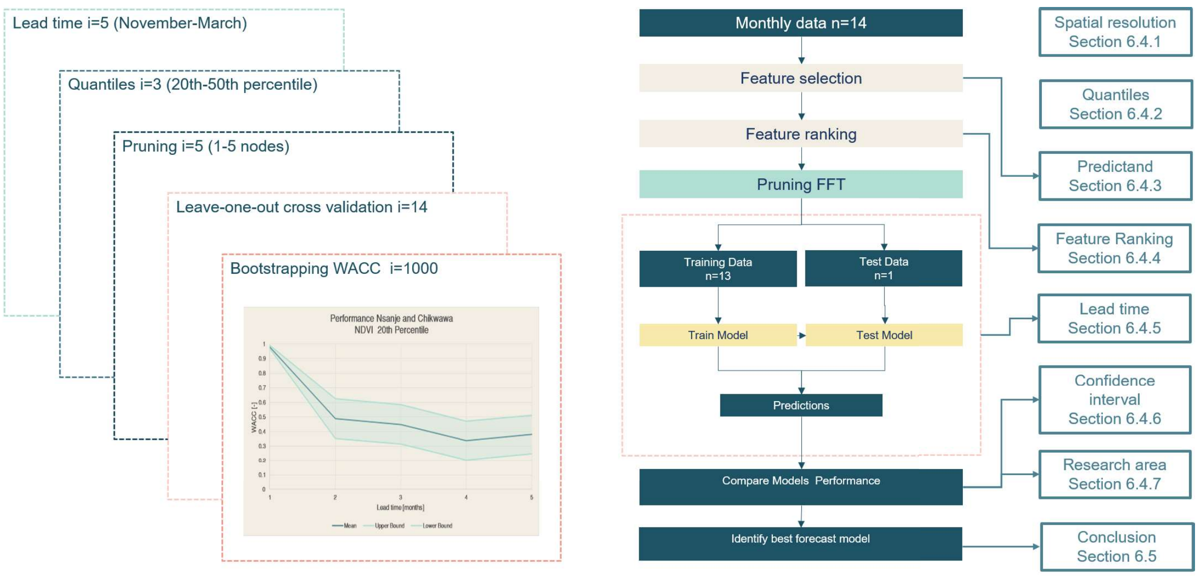


Figure 6.17: Outline drought forecast model. The model iterate over four levels, first over the months within the growing season, followed by the drought classification threshold values and the amount of nodes in the decision tree (1 to 5) and last the cross-validation using Leave-one-out with replacement. The bootstrapping is down over the Leave-one-out cross-validation in order to calculate the confidence interval over the WACC. The steps are explained in more detail in section 4.4

Lead time	Predictors	Predictand	Lead time	Predictors	Predictands
	VdS			ESA	
March	ONI_{July}, $SM_{March,VdS}$, P_{March}, LST_{March}, $NDVI_{March}$, $DSCat1_{March}$	$NDVI_{March,20}$ $NDVI_{March,40}$ $NDVI_{March,50}$	March	ONI_{July}, $SM_{March,ESA}$, P_{March}, LST_{March}, $NDVI_{March}$, $DSCat1_{March}$	$NDVI_{March,20}$ $NDVI_{March,40}$ $NDVI_{March,50}$
February	ONI_{July}, $SM_{Feb,VdS}$, P_{Feb}, LST_{Feb}, $NDVI_{Feb}$, $DSCat1_{Feb}$	$NDVI_{March,20}$ $NDVI_{March,40}$ $NDVI_{March,50}$	February	ONI_{July}, $SM_{Feb,ESA}$, P_{Feb}, LST_{Feb}, $NDVI_{Feb}$, $DSCat1_{Feb}$	$NDVI_{March,20}$ $NDVI_{March,40}$ $NDVI_{March,50}$
January	ONI_{July}, $SM_{Jan,VdS}$, P_{Jan}, LST_{Jan}, $NDVI_{Jan}$, $DSCat1_{Jan}$	$NDVI_{March,20}$ $NDVI_{March,40}$ $NDVI_{March,50}$	January	ONI_{July}, $SM_{Jan,ESA}$, P_{Jan}, LST_{Jan}, $NDVI_{Jan}$, $DSCat1_{Jan}$	$NDVI_{March,20}$ $NDVI_{March,40}$ $NDVI_{March,50}$
December	ONI_{July}, $SM_{Dec,VdS}$,P_{Dec}, LST_{Dec}, $NDVI_{Dec}$, $DSCat1_{Dec}$	$NDVI_{March,20}$ $NDVI_{March,40}$ $NDVI_{March,50}$	December	ONI_{July}, $SM_{Dec,ESA}$, P_{Dec}, LST_{Dec}, $NDVI_{Dec}$, $DSCat1_{Dec}$	$NDVI_{March,20}$ $NDVI_{March,40}$ $NDVI_{March,50}$
November	ONI_{July}, $SM_{Nov,VdS}$, P_{Nov}, LST_{Nov}, $NDVI_{Nov}$, $DSCat1_{Nov}$	$NDVI_{March,20}$ $NDVI_{March,40}$ $NDVI_{March,50}$	November	ONI_{July}, $SM_{Nov,ESA}$, P_{Nov}, LST_{Nov}, $NDVI_{Nov}$, $DSCat1_{Nov}$	$NDVI_{March,20}$ $NDVI_{March,40}$ $NDVI_{March,50}$
March	ONI_{July}, $SM_{March,VdS}$, P_{March}, LST_{March}, VOD_{March}, $DSCat1_{March}$	$VOD_{March,20}$ $VOD_{March,40}$ $VOD_{March,50}$	March	ONI_{July}, $SM_{March,ESA}$, P_{March}, LST_{March}, VOD_{March}, $DSCat1_{March}$	$VOD_{March,20}$ $VOD_{March,40}$ $VOD_{March,50}$
February	ONI_{July}, $SM_{Feb,VdS}$, P_{Feb}, LST_{Feb}, VOD_{Feb}, $DSCat1_{Feb}$	$VOD_{March,20}$ $VOD_{March,40}$ $VOD_{March,50}$	February	ONI_{July}, $SM_{Feb,ESA}$, P_{Feb}, LST_{Feb}, VOD_{Feb}, $DSCat1_{Feb}$	$VOD_{March,20}$ $VOD_{March,40}$ $VOD_{March,50}$
January	ONI_{July}, $SM_{Jan,VdS}$, P_{Jan}, LST_{Jan}, VOD_{Jan}, $DSCat1_{Jan}$	$VOD_{March,20}$ $VOD_{March,40}$ $VOD_{March,50}$	January	ONI_{July}, $SM_{Jan,ESA}$, P_{Jan}, LST_{Jan}, VOD_{Jan}, $DSCat1_{Jan}$	$VOD_{March,20}$ $VOD_{March,40}$ $VOD_{March,50}$
December	ONI_{July}, $SM_{Dec,VdS}$,P_{Dec}, LST_{Dec}, VOD_{Dec}, $DSCat1_{Dec}$	$VOD_{March,20}$ $VOD_{March,40}$ $VOD_{March,50}$	December	ONI_{July}, $SM_{Dec,ESA}$, P_{Dec}, LST_{Dec}, VOD_{Dec}, $DSCat1_{Dec}$	$VOD_{March,20}$ $VOD_{March,40}$ $VOD_{March,50}$
November	ONI_{July}, $SM_{Nov,VdS}$, P_{Nov}, LST_{Nov}, VOD_{Nov}, $DSCat1_{Nov}$	$VOD_{March,20}$ $VOD_{March,40}$ $VOD_{March,50}$	November	ONI_{July}, $SM_{Nov,ESA}$, P_{Nov}, LST_{Nov}, VOD_{Nov}, $DSCat1_{Nov}$	$VOD_{March,20}$ $VOD_{March,40}$ $VOD_{March,50}$

Table 6.1: Model scenarios per area existing of 12 different models representing four models with three different threshold values (20th percentile, 40th percentile and 50th percentile derived from the NDVI or VOD state in March). The four models differ in predictand (VOD or NDVI) and soil moisture data resolution (0.003 ° resolution from VanderSat and the lower 0.25 ° resolution from ESA). The input values of the models exist of the ONI value in July and the cumulative value per lead time for the other predictors.

6.4.1. Spatial resolution

The first aspect discusses the influence on the model performance when using high or low soil moisture resolution as a predictor in the forecast model. The resolution compared is the 0.003 ° spatial resolution from VdS and the lower 0.25 ° spatial resolution from ESA. The ESA data covers a period until 2015, while VdS contains data until 2017, resulting in a one year fewer data in the training and testing phase in the ESA model (VdS data contains a data gap over the growing season 2011-2012). Table 6.2 gives an overview of the mean WACC values for the different areas and different data sets. At first, the different graphs show a slightly better performance for the lower resolution data than the higher resolution, but this conclusion might be blunt, the differences are not that significant over the seven different areas and over the different lead times. The largest difference between lower and higher resolution occurs for $NDVI_{LT3,50}$ and $VOD_{LT3,40}$, where the lower resolution performs better for $NDVI_{LT3,50}$ and the higher resolution for $VOD_{LT3,40}$. The predictand $NDVI_{LT3,20}$ and $VOD_{LT5,50}$ ESA and VdS perform equally. When comparing the difference in model performance between the lower resolution data and the higher resolution over the different areas, there is an interesting given. Over the Nsanje + Chikwawa area the ESA and VdS data perform equally, for both the predictands; sometimes ESA performs better and sometimes VdS. Over the other percentiles, the lower resolution performs better for both lead times with predictands NDVI 40th percentile and VOD 20th percentile, while the higher resolution performs better for both lead times with NDVI 50th percentile and VOD 40th percentile. For the other predictand percentiles, either LT3 or LT5 performs better. ESA returns higher model performances for the areas Chikwawa for both the predictands, while in Nsanje the model performance is higher when using the VdS data (figure 6.18 - 6.21). Looking at the smaller areas VdS performance better with predictand VOD in Mikalango and Mahkanga, while ESA performs better with NDVI as predictand for Mikalango and Mahkanga. The model performance of VdS and ESA might be influenced by the datasets used for the other predictors. Although the correlation differences are small and the correlation coefficient is very low, somewhere around $r = 0.20$, the anomalies SM ESA is stronger correlated to the VOD March anomaly values. It should be taken into account that the LST and VOD values are obtained from VdS, which might result in a bias in the model performance. The area Mitole show contradicting results, the ESA is performing better with VOD as predictand, and VdS performs better with NDVI as predictand. The real advantage of using the high spatial resolution over the lower resolution is expected to be larger when focused on smaller areas; the areas in this research are probably too large to distinguish the effect of using the higher resolution (). In this research, the data average over the rain fed agricultural area, which removes the extremes of extreme dry- or extreme wet areas. Added to this averaging, is the importance of scale in the analysis. The scale is guiding in the decision-making which spatial resolution returns the best performance of the forecast model. Probably the ESA data is suitable in area analysis, while the high resolution of VdS shows its value at specified locations, the analysis on plot scale for example(). Another factor to keep in mind is the difference between the data sources used is not only limited to a difference in resolution but also covers different flagging of the area. Especially along the coastline, the values give different results ().

	NDVI				VOD				NDVI				VOD			
	ESA		VdS		ESA		VdS		ESA		VdS		ESA		VdS	
Areas	LT3	LT5	LT3	LT5	LT3	LT5	LT3	LT5	LT3	LT5	LT3	LT5	LT3	LT5	LT3	LT5
20th percentile																
Nsanje + Chikwawa	0.20	0.44	0.45	0.38	0.17	0.74	0.14	0.73	3	4	3	4	3	4	4	4
Chikwawa	0.17	0.18	0.17	0.05	0.19	0.74	0.17	0.74	4	4	4	4	5	5	5	4
Nsanje	0.40	0.12	0.65	0.71	0.67	0.36	0.14	0.37	-	4	-	4	4	2	4	2
Mitole	0.68	0.16	0.65	0.35	0.45	0.45	0.42	0.40	4	3	5	2	4	3	4	3
Mikalango	0.17	0.74	0.11	0.42	0.14	0.74	0.17	0.74	1	3	2	3	3	-	3	-
Zunde	0.95	0.09	-	-	0.08	0.71	-	-	x	x	2	4	x	x	-	3
Mahkanga	0.48	0.36	0.48	0.36	0.48	0.36	0.48	0.36	4	2	5	2	-	2	-	2
40th percentile																
Nsanje + Chikwawa	0.23	0.70	0.22	0.41	0.38	0.35	0.52	0.37	4	3	4	3	4	5	4	4
Chikwawa	0.41	0.47	0.26	0.37	0.75	0.51	0.34	0.38	2	5	1	5	5	2	5	2
Nsanje	0.78	0.53	0.93	0.57	0.78	0.27	0.78	0.27	5	2	5	3	2	4	2	4
Mitole	0.66	0.47	0.60	0.32	0.76	0.53	0.52	0.88	3	2	4	1	2	1	1	2
Mikalango	0.86	0.66	0.70	0.50	0.36	0.87	0.40	0.82	1	2	1	2	2	5	1	-
Zunde	0.23	0.47	-	-	0.52	0.57	-	-	x	x	2	2	x	x	-	4
Mahkanga	0.31	0.37	0.30	0.32	0.23	0.56	0.31	0.41	1	1	1	1	-	2	-	2
50th percentile																
Nsanje + Chikwawa	0.51	0.72	0.72	0.75	0.50	0.54	0.46	0.57	4	3	4	3	5	3	5	3
Chikwawa	0.54	0.64	0.46	0.57	0.76	0.43	0.71	0.43	2	5	1	5	5	2	5	2
Nsanje	0.63	0.64	0.55	0.64	0.92	0.37	0.92	0.47	3	2	2	3	3	2	3	2
Mitole	0.75	0.64	0.76	0.65	0.41	0.47	0.38	0.89	3	2	4	1	2	2	2	3
Mikalango	0.37	0.72	0.46	0.54	0.70	0.75	0.71	0.74	3	1	3	1	5	4	4	3
Zunde	0.34	0.44	-	-	0.58	0.61	-	-	x	x	2	3	x	x	-	2
Mahkanga	0.42	0.65	0.46	0.36	0.50	0.57	0.50	0.47	2	1	2	1	-	2	-	2

Table 6.2: Average WACC values LT3 and LT5 all model scenario's including the position of the NDVI and VOD in the feature ranking.

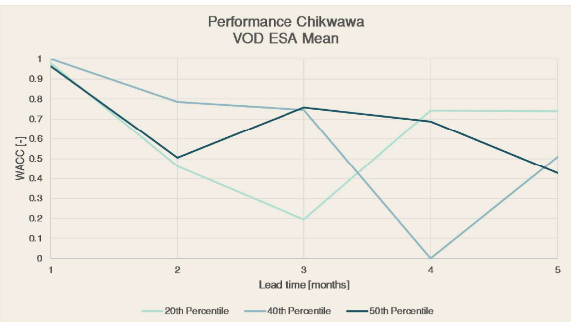


Figure 6.18: Model performance VOD ESA Mean WACC Chikwawa

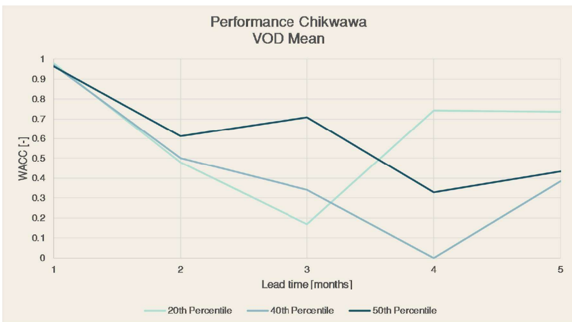


Figure 6.19: Model performance VOD VdS Mean WACC Chikwawa

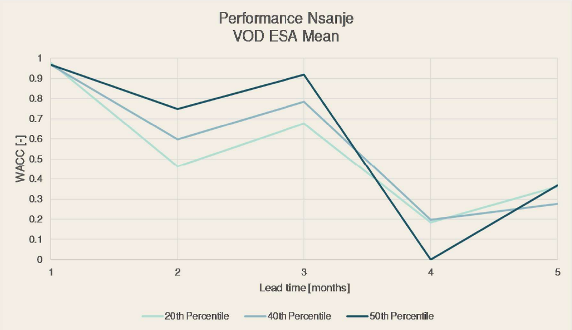


Figure 6.20: Model performance VOD ESA Mean WACC Nsanje

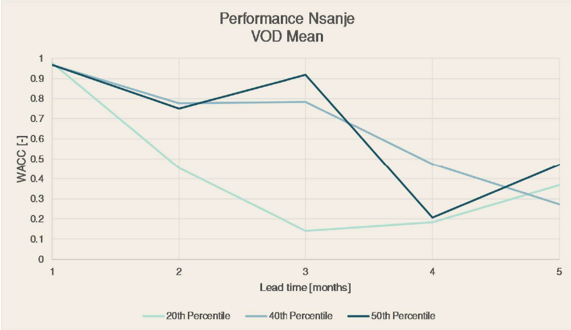


Figure 6.21: Model performance VOD VdS Mean WACC Nsanje

6.4.2. Threshold percentile

The second comparison made is based on the threshold values of the predictands.The percentiles represent with a 20th percentile an 'extreme' drougt event with a return period of five-years. While with a 50th percentile the return period is once in the two-years. Comparing the model performance of the 20th, 40th and 50th percentile for both predictands. Overall the 50th percentile show an increasing gradient over the growing season; there are fewer peaks and drops compared to the 20th and 40th percentile. Comparing the average values over all areas, NDVI 50th percentile shows the best performance at LT5 with an average around 0.65, while NDVI 40th percentile gives a value between 0.45 and 0.56 and NDVI 20th percentile a value ranging from 0.25 to 0.38. For VOD this is the opposite, where VOD 20th percentile give values around 0.62, VOD 40th percentile between 0.34 and 0.38 and VOD 50th percentile between 0.44 and 0.49. LT 3 shows over both predictand a better performance for the 50th percentile, with values between 0.63 and 0.73, the 40th percentile give lower values ranging between 0.57 and 0.64, where the predictand VOD performs slightly better than NDVI, last the 20th percentile vary between 0.15 and 0.42. The difference between VOD and NDVI at LT5 can be explained by how drought is defined. Over the whole growing season and region-specific Zunde NDVI VdS 20th percentile show differences compared to the other areas. It shows strong peaks over the season, while the 40th and the 50th percentile show smaller peaks. Chikwawa $NDVI_{20\%}$ and $NDVI_{40\%}$ from VdS and ESA show contradicting patterns, it seems like the 40th and 20th percentile alternates (figure 6.23). The different percentiles follow more or less the same peaks and drops, with exceptions for the contradicting peaks in Mahkanga (LT5 to LT4 and LT4 to LT3 20th percentile compared to the 40th and 50th percentile).

	NDVI ESA					NDVI VdS				
	Mean FAR	Mean MS	Mean CR	Mean Hit	Mean WACC	Mean FAR	Mean MS	Mean CR	Mean Hit	Mean WACC
	20th percentile					20th percentile				
1	0.06	0	0.94	1	0.99	0.09	0.00	0.91	1.00	0.98
2	0.3	0.22	0.7	0.78	0.76	0.24	0.78	0.76	0.22	0.37
3	0.33	0.89	0.67	0.11	0.25	0.33	0.67	0.67	0.33	0.42
4	0.39	0.56	0.61	0.44	0.49	0.42	0.56	0.58	0.44	0.48
5	0.36	0.89	0.64	0.11	0.25	0.48	0.67	0.52	0.33	0.38
	40th percentile					40th percentile				
1	0.13	0	0.88	1	0.97	0.13	0.00	0.88	1.00	0.97
2	0.46	0.17	0.54	0.83	0.76	0.46	0.17	0.54	0.83	0.76
3	0.52	0.53	0.48	0.47	0.47	0.52	0.53	0.48	0.47	0.47
4	0.58	0.39	0.42	0.61	0.57	0.63	0.33	0.38	0.67	0.60
5	0.58	0.39	0.42	0.61	0.57	0.88	0.44	0.13	0.56	0.45
	50th percentile					50th percentile				
1	0.14	0	0.86	1	0.97	0.14	0.00	0.86	1.00	0.97
2	0.81	0.43	0.19	0.57	0.48	0.52	0.33	0.48	0.67	0.62
3	0.61	0.39	0.39	0.61	0.56	0.72	0.33	0.28	0.67	0.58
4	0.71	0.33	0.29	0.67	0.58	0.76	0.33	0.24	0.67	0.56
5	0.62	0.24	0.38	0.76	0.66	0.81	0.19	0.19	0.81	0.65

Table 6.3: Model performance NDVI averaged over all areas

	VOD ESA					VOD VdS				
	Mean FAR	Mean MS	Mean CR	Mean Hit	Mean WACC	Mean FAR	Mean MS	Mean CR	Mean Hit	Mean WACC
	20th percentile					20th percentile				
1	0.09	0.00	0.91	1.00	0.98	0.09	0.00	0.91	1.00	0.98
2	0.21	0.67	0.79	0.33	0.45	0.21	0.67	0.79	0.33	0.45
3	0.30	0.78	0.70	0.22	0.35	0.41	1.00	0.59	0.00	0.15
4	0.15	0.56	0.85	0.44	0.56	0.15	0.56	0.85	0.44	0.55
5	0.24	0.44	0.76	0.56	0.61	0.24	0.44	0.76	0.56	0.61
	40th percentile					40th percentile				
1	0.08	0.00	0.92	1.00	0.98	0.13	0.00	0.88	1.00	0.97
2	0.46	0.17	0.54	0.83	0.76	0.42	0.22	0.58	0.78	0.73
3	0.48	0.33	0.52	0.67	0.64	0.62	0.40	0.38	0.60	0.55
4	0.79	0.83	0.21	0.17	0.18	0.79	0.67	0.21	0.33	0.31
5	0.67	0.61	0.33	0.39	0.38	0.96	0.56	0.04	0.44	0.34
	50th percentile					50th percentile				
1	0.14	0.00	0.86	1.00	0.97	0.14	0.00	0.86	1.00	0.97
2	0.62	0.33	0.38	0.67	0.60	0.62	0.29	0.38	0.71	0.63
3	0.44	0.22	0.56	0.78	0.73	0.56	0.22	0.44	0.78	0.70
4	0.90	0.57	0.10	0.43	0.35	0.76	0.67	0.24	0.33	0.31
5	0.95	0.43	0.05	0.57	0.44	0.90	0.38	0.10	0.62	0.49

Table 6.4: Model performance VOD averaged over all areas.

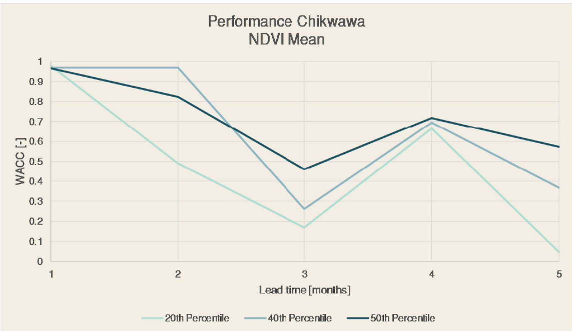


Figure 6.22: Model performance NDVI VdS Mean WACC Chikwawa

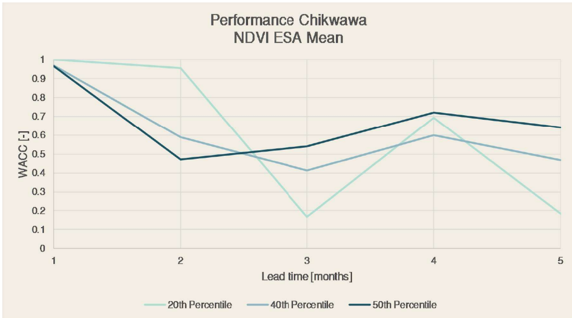


Figure 6.23: Model performance NDVI ESA Mean WACC Chikwawa

6.4.3. Predictand

The two predictands used in this model are the NDVI and VOD values in March. The model performance shows a slight difference in WACC. VOD show higher WACC values in general at LT4 compared to the NDVI values. The NDVI measure the health of the vegetation by the identification of the greenness (Bands NIR and IR). VOD measure the vegetation water content in the canopy, NDVI measures the reflectance photosynthesis by the NR and RED light scatter, shortly saying sunlight reflected by plants. During the growing season plants are absorbing visible sunlight — meaning less absorbing, results in less reflectance. The vegetation growth is limited to by the water availability.

This subsection compares the model performance difference by the use of different predictands, looking at LT3 and LT5. Like mentioned in the previous subsections the model performance fluctuates over the growing season 6.2. Comparing LT3 NDVI with LT3 VOD for ESA and VdS separately VOD turns out to be a better drought predicand. Within the areas were VOD performs better this happens more often for LT3, while the ones were NDVI performs better this is most of the time at LT5. When comparing the model performance per lead time over all the four different predictands (NDVI ESA, NDVI VdS, VOD ESA and VOD VdS) only at LT 3, 20th percentile, NDVI performs better. The other lead times and percentiles VOD returns higher WACC values. The difference is most significant for the 40th percentiles LT5. In general, VOD shows a better performance for LT3 and NDVI for LT5.

6.4.4. Ranking of features

There is a high variety over the spatial and temporal distribution in the top five feature ranking over the growing season. In some areas, the state of the predictand ranks in the top three, while in other areas it is excluded from the top five. The different hydrological, geographical and topographic characteristics might explain the variation in feature ranking. This section describes and discusses the variety in feature ranking.

The first feature to discuss is the state of the predictand. The model is trained using the NDVI and VOD March valueNDVI, VOD, SM, LST anomaly value. Precipitation and the number of dry spells cumulates over the growing season. to predict a drought at the end of the growing season. Within the season the model uses the NDVI and VOD monthly state values. The months returning the predictand state values as the first ranked feature is for all areas at LT1. Another month returning the NDVI state high in the top five is at LT4 in the combined area Nsanje and Chikwawa, Chikwawa and Mikalango. The model performance seems influenced by the ranking of the features, at the NDVI 20th percentile in Chikwawa the performance increases at LT4 and LT2, in these months the NDVI is high located in the top five. Over all areas, percentiles and months NDVI and VOD are included in the ranking.

The second feature analysed is SM, which together with NDVI occurs high in the feature ranking. SM is for most areas, percentiles and months ranked in the top three. When SM is low ranked it is mainly for the VdS data, the ESA SM data is steady ranked as the second predictor, while the SM VdS data vary between the first predictor or the last. Due to the higher resolution, the data might be more sensitive to changes, which translate to the model performance. In general, with NDVI as predictand, the ranking of SM is higher, compared to the VOD. Comparing the feature ranking of the location of SM within the top five between ESA and VdS, there is an inconsistency in which model performs better.

In some cases, the model performance does not show any different if ESA includes a feature and VdS not or the other way around. For example, Mitole VOD 20th percentile in February VdS ranks SM as the essential feature, while ESA does not take SM into account. The model performance of ESA is significantly higher than the VdS performance. With NDVI 40th percentile as predictand in Mitole, the model performance is not different if including SM or not. In another case the VdS performance is much higher, Nsane NDVI 20th percentile and 40th percentile in January, while ESA and VdS use almost the same feature ranking. The model performance for the 40th percentile is higher in January, using NDVI instead of DS in the 20th percentile.

The feature excludes most from the top five are ONI and DS. These predictors are either highly ranked or out of the ranking. Dry spells do not belong to the top five features in the March, which indicates that at the end of the growing season the impact of a dry spell is not essential. The crops are more resistant to a sequence of dry days, and the occurrence to this climate-risk might be lower. The impact of a dry spell is high in the maturing phase of the growing season when crops reach their maximum required water needs. The month DS occur in the top five is at LT 2, LT3 and for NDVI 40th percentile in LT4, which is mentioned by the farmers as the month with the climate-risks of erratic rains. This pattern of the climate-risks DS occurring at LT2 and LT3 covers all areas and percentiles. More in detail dry spells are higher ranked in Chikwawa and the areas within Chikwawa compared to Nsanje. DS does not occur in the top five in Mahkanga; this might relate to the soil type (section 3.2.1). Gleysols dominate Mahkanga, characterised with relatively high groundwater tables. ONI is a feature driving the drought predictions, especially in the first months of the season with VOD as predictand. In the forecast models with NDVI as predictand, the ONI returns to be more critical at the end of the growing season. Furthermore, there exists a spatial difference in the importance of ONI included in the drought forecast. ONI is more critical for the individual smaller areas with NDVI as predictand compared to the areas level. In most areas at LT4 ONI is excluded from the top five feature ranking with the predictand NDVI. At LT3, with VOD as predictand ONI is excluded from the top five in most areas. ONI is important for the 20th percentile, which indicate the influence of ENSO on the more 'extreme' drought events with a return period of five-years.

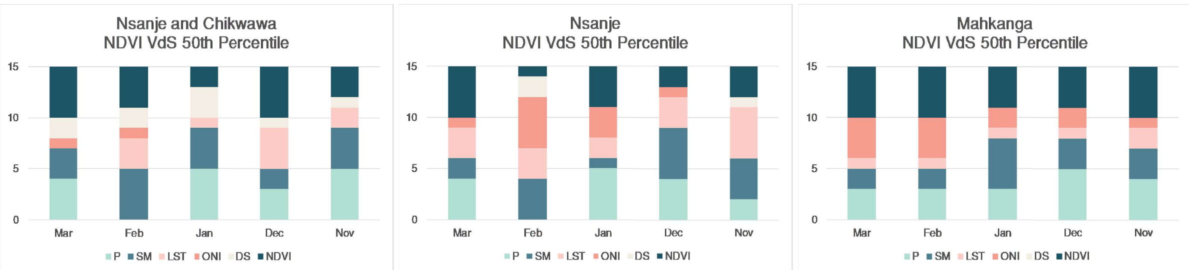


Figure 6.24: Feature ranking NDVI VdS 50th percentile Nsanje and Chikwawa area, Mahkanga and Mitole.

The last two features to discuss are P and LST. During the climate-risks prone months, at LT4, LT3 and LT2, P more often in the top three ranking. Toward the end of the growing season, the influence of P in drought forecast is decreasing. The vegetation state is becoming more critical as mentioned before. Like SM and NDVI, LST ranks within the top three over almost all areas and percentiles.

The feature ranking at LT3 exist of the following predictors; P, SM and NDVI with DS and ONI occurring for some areas with NDVI as predictand and SM, P and VOD/DS with VOD as predictand. For LT5 the predictors P, SM and LST are top three ranked, with VOD as predictand ONI occurs more often in the top five. The variety within the top five feature ranking and the top three ranking indicate that all features are essential in drought prediction. All features contain a certain degree in the ability to predict a drought. Especially at LT3 if there is a high change of climate risks, which makes dry spells and the vegetation status significant predictors for agricultural drought. At LT5 the initial conditions before the start of the season are crucial, such as soil moisture conditions, precipitation rate and land surface temperature. These features are drivers of crop growth (Brouwer, 1986).

6.4.5. Lead time growing season

The model performance is simplified by averaging the WACC over all areas. In figure **??** and **??** the average probability of a false alarm, correct rejection, misses and hits are presented together with the Mean WACC. This figure clearly shows the relation between the false alarm rate, misses and correct rejections and Hits; they balance each other. The WACC average shows a relatively straight line from LT5 to LT 3, with the lowest model performance at LT3, from where the WACC increases to 0.93 at the end of the growing season. The model performance increase over the growing season by reducing the false alarm. The longer the lead time, the harder it is to generate an accurate forecast. All areas show an already existing skill at LT5 with a WACC average of 0.51. The false alarm rate is higher with longer lead time, the probability of False Alarm reduce to the end of the growing season. The reduction in False Alarms drives the higher WACC values with an increasing probability of a correct rejection. The average probability of false alarm at LT5 is 59%, while for LT1 this probability decrease to 14% compensated by the correct rejection from LT5 41% to LT1 86%. Over the different predictands and percentiles can be seen the average WACC increases over the period, with a clear drop at LT3 in the 20th percentiles. It can be seen that the hit rate follow the line of the mean WACC value, guiding the model performance. With a hit rate close to the mean WACC the confidence interval is smaller (more about confidence interval (section 4.4.4). Another aspect to point out is the relationship between a high miss rate corresponding with a reducing in the mean WACC. It seems like the model performance reduces with an increasing miss rate and raise with a decreasing miss rate (figure 6.29 - 6.30.

Besides the observation of the relationship between an increasing probability of a miss with a decreasing WACC causing a drop. A drop in model performance could relate to a reaction on noise in the data, due to the low amount of samples. This drop could also indicate that something is happening when at LT3, which cause sudden changes in the data resulting in a decreasing performance; it can be a trigger signal indicating something is going on. At LT3 and LT2 there is a higher chance of climate-risks mentioned by the smallholder farmers and the agricultural officials in chapter 5 (section 5.3).

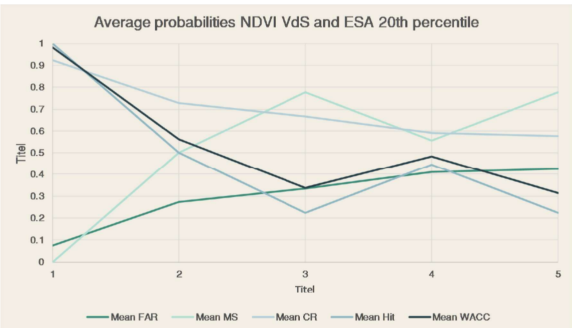


Figure 6.25: Average probabilities NDVI 20th percentile.

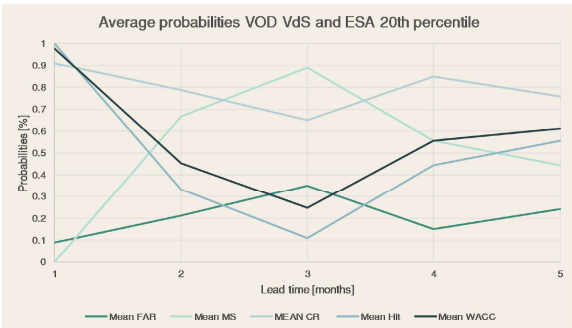


Figure 6.26: Average probabilities VOD 20th percentile.

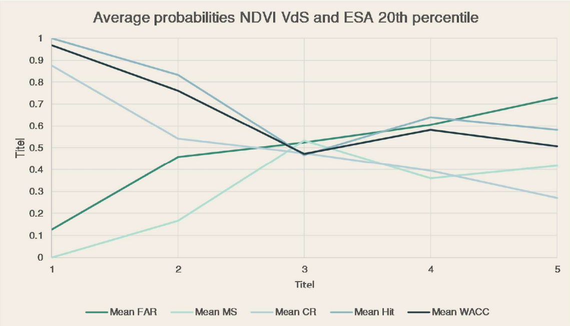


Figure 6.27: Average probabilities NDVI 40th percentile.

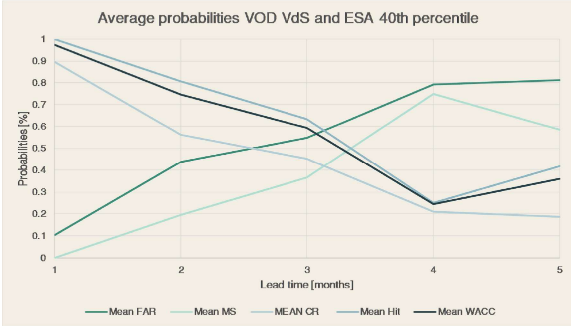


Figure 6.28: Average probabilities VOD 40th percentile.

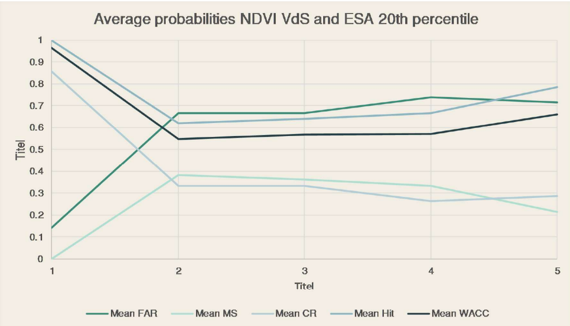


Figure 6.29: Average probabilities NDVI 50th percentile.

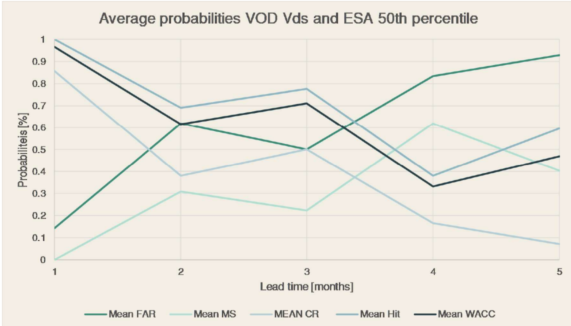


Figure 6.30: Average probabilities VOD 50th percentile.

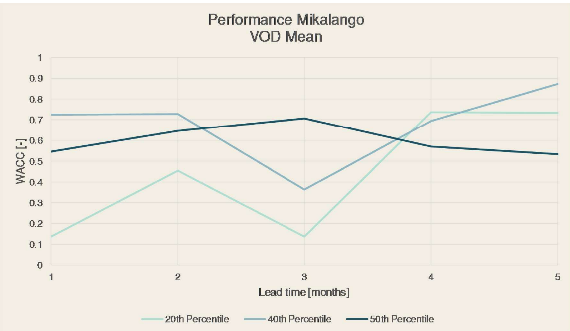


Figure 6.31: Model performance VOD VdS Mean WACC Mikalango

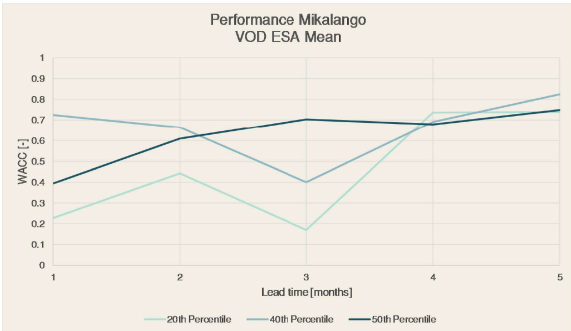


Figure 6.32: Model performance VOD ESA Mean WACC Mikalango

6.4.6. Confidence interval

The confidence interval is used to predict the robustness of the model performance. In general, it is expected that the confidence interval reduces over the growing season. In most cases, the confidence interval is relatively small, particular with LT5 a wider confidence interval is expected. The confidence interval is calculated over the mean WACC value. The low amount of sampling might cause the unexpected small confidence interval, the CI is calculated as mentioned in section 4.4.4 by bootstrapping the WACC values with n=14 with resampling for 1000 times. With a 20th percentile and n=14, the model predicts three droughts events over the datasets which are used to train and test the model. The confidence interval for LT1 is in almost all cases similar to the mean WACC; at the end of the growing season, the model shows a robust model performance. The combined area Nsanje and Chikwawa with the predicand NDVI and ESA data show an increasing confidence interval in LT3 and LT2 50th percentile compared to the 40th percentile and 40th percentile compared to the 20th percentile. With 50th percentile the model calculates with seven predicted drought events. The drop in WACC value for LT3 has a 0.31 difference. The drop at LT3 over the percentile also occur for the VOD VDS and ESA models over the area Nsanje and Chikwawa. Nsanje NDVI VDS LT2 shows a similar pattern with a small confidence interval for the 20th percentile with an increasing interval over the 40th and 50th percentile with a difference of WACC 0.39 and 0.28 (20-40 and 20-50). Confidence interval link to the feature rank at LT3 where the confidence interval reduces and increase again to LT2 (NsCH VOD 20 VdS, NsCh VOD ESA 20, NsCh NDVI ESA 20, Nsanje VOD VdS 20, Ch NDVI VdS 20, Ch NDVI ESA 20, Ch VOD VDS 20, Ch VOD ESA 20, Mikalango NDVI VSD and ESA 20, Mikalango VOD VDS and ESA 20, VOD ESA 20.

Predictor selection, remove one feature at the time and repeat the model to identify the weight of the feature on the model performance. It might be that there is noise in a feature, which case some sudden changes in model performance, as happens at LT3. More research is needed to identify the cause of the immediate increase of unexpected decrease in model performance. It can be a essential signal in agricultural drought forecast, a bias of the model or noise in the data.

6.4.7. Spatial Distribution

The last aspect of the model discussed is the model performance based on spatial distribution. The decreasing model performance towards LT3 is not shown for all research areas. The combined area Nsanje and Chikwawa return these lower values, while Nsanje shows an increase in performance and a decrease at LT2. Within the Chikwawa area, a stronger relationship is visible with Mikalango than with Mitole. Mikalango does show the mentioned drop at LT3, while the WACC in Mitole slightly decreases, but not as much as Chikwawa, Mikalango and the combined area show. The Nsanje area model performance pattern does not directly link to the averaged WACC in Zunde and Mahkanga. When comparing all areas, Mikalango VOD both VdS and ESA show unexpected results over all percentiles (figure 6.32. Although the confidence interval increases over the growing season, it seems like the performance is mirrored over the lead times, starting with a relatively high WACC, which falls over the growing season.

6.5. Conclusion

The qualitative research presented the candidate predictors and predictands, P, SM, DS, LST, NDVI and VOD. The developed Fast-and Frugal decision tree uses these features to predict a drought, through a binary drought classification applying the 20th, 40th and 50th percentile of the NDVI and VOD values in March at the end of the growing season. This chapter presents and discusses the model performance of the Fast-and Frugal decision tree, based on various model aspects. The aspects discussed are the influence of spatial resolution, threshold percentiles, temporal resolution, feature ranking, type of predictand, confidence interval and spatial distribution. Combining these aspects, with the provided experiences of humanitarian organisation and smallholder farmers gives a first indication of the best combination of agro-climatic indicators in agricultural drought forecast in the Lower Shire River basin in Malawi. The results answer the questions Which feature combination perform best in drought prediction at different spatial and temporal resolutions in the growing season? and How do the different model scenarios perform in terms of accuracy?. These answer to these questions will answer the second sub-question:

> **What combination of agro-climatic features perform best in agricultural drought forecast in the Lower Shire River Basin?**

Which feature combination perform best in drought prediction at different spatial and temporal resolutions in the growing season?

The different aspects of the model performance combined suggest the best feature combination in the growing season. As mentioned in the previous chapter the critical months in agricultural drought predictions are at LT5 and LT3. LT5 is at the start of the growing season when farmers start planting and at LT3 there is a high chance of climate-risks. Tabel 6.5 provides an overview of the spatial and temporal resolution with the corresponding best feature combination. In two-thirds of the models, the VOD is returning as best predictand compared with NDVI. With the 50th percentile VOD returns a highest value at LT3, compared to the 20th and 40th percentile, the opposite appears for LT5, where the 20th percentile show the highest mean WACC value. It is expected that the 50th percentile return better performances because the model contains more drought events to train and test the model with. With an increasing number of data samples the model can make a better prediction. With too much samples and variables there is a risk in overfitting, a balance need to be created. All areas show a certain degree of skill at the start of the season with a mean value of 0.51. The skill difference between the 20th, 40th and 50th percentile is lower at the start, than throughout the season. This difference in skill is driven by different the differences occurring within the growing season. For some areas the model performance show a sudden drop at LT3, this drops seems to correspond to an increasing number of misses predicted by the model. These drops occur more often with the 20th percentile compared to the 50th percentile.

At the start of the season, the initial conditions, such as soil moisture, precipitation rate and land surface temperature play an important role in agricultural drought predictions. These features drive the plant growth over the growing season. Over the season the vegetation indicators (VOD and NDVI) and DS get increasingly important in the drought prediction. At the 50th percentile, at LT5, SM rank as the first predictor with LST and P filling the top three. The 40th percentile shows similar features in the top five ranking, with an increasing importance of LST. Especially in the larger areas Nsanje and Chikwawa this feature appears as first ranked. This larger area average the LST, which might return a more stable temperature. The extremes are levelled out. The change of the feature order is also dependend on the predictand. It can be seen that the SM and LST sometimes alternate with the different predictands. At LT5 with a 20th percentile VOD, ONI wins a position in the feature ranking. The increasing importance of ONI could be explained by the fact that the 20th percentile identify the more extreme drought events (one in five-years), for which atmospheric teleconnections influence the occurrence of these events. The variety within the top five feature ranking and the top three ranking indicate that all features are essential in drought prediction, depending on the spatial characteristics and lead time within the growing season 6.24.

Nsanje experience higher precipitation rates compared to Chikwawa. The land surface temperature shows the opposite with higher values in Chikwawa. These opposite reactions of higher P rates, lower LST and higher SM, result in higher vegetation rates (VOD/NDVI) in Nsanje. The time series within the areas follow kind of the same patterns, except for Mitole which low-temperature values does not translate into the time series pattern of Chikwawa. In general, Mitole and Zunde are showing deviating values compared to the other areas, which might be driven by the spatial distribution of the areas. All extremes are levelled out by averaging the monthly values over rainfed agriculture. Due to the difference in the spatial area the extremes in the combined area of Nsanje and Chikwawa is levelled out more than the extreme in Mitole, which is the smallest area. This levelling out of the intensity of a rain event is visible in the historical time series. This spatial effect calculates through the drought forecast model and influences the feature ranking.

	LT3			LT5		
	Spatial resolution	Predictand	Predictor	Spatial resolution	Prectand	Predictor
	20th percentile			20th percentile		
Nsanje and Chikwawa	VdS	NDVI	P, LST, NDVI, DS, ONI	ESA	VOD	ONI, SM, P, VOD, LST
Chikwawa	ESA	VOD	DS, P, SM, LST, VOD	ESA/VdS	VOD	ONI, P, LST, VOD, SM
Nsanje	ESA	VOD	SM, ONI, LST, VOD, P	VdS	NDVI	SM, P, LST, NDVI, ONI
Mitole	ESA	NDVI	P, LST, DS, NDVI, SM	ESA	VOD	P, SM, VOD, DS, ONI
Mikalango	ESA	NDVI/VOD	NDVI, P, SM, ONI, LST	ESA/VdS	VOD	ONI, P, LST, DS, SM
Zunde	VdS	NDVI	P, NDVI, LST, DS, ONI	VdS	VOD	SM, ONI, VOD, P, LST
Mahkanga	all	all	ONI, P, SM, LST, NDVI	all	all	ONI, NDVI, P, LST, SM
	40th percentile			40th percentile		
Nsanje and Chikwawa	VdS	VOD	P, DS, SM, VOD, LST	ESA	NDVI	SM, LST, DS, NDVI, P
Chikwawa	ESA	VOD	DS, SM, P, LST, VOD	ESA	VOD	LST, VOD, SM, P, DS
Nsanje	VdS	NDVI	ONI, LST, P, SM, NDVI	VdS	NDVI	LST, SM, NDVI, P, DS
Mitole	ESA	VOD	SM, VOD, P, DS, ONI	VdS	VOD	SM, VOD, DS, ONI, LST
Mikalango	ESA	NDVI	NDVI, SM, ONI, P, LST	ESA	VOD	SM, P, DS, LST, VOD
Zunde	VdS	VOD	DS, ONI, SM, P, LST	VdS	VOD	SM, ONI, DS, VOD, LST
Mahkanga	ESA/VdS	NDVI/VOD	P, SM, LST, DS, ONI	ESA	VOD	LST, VOD, SM, P, ONI
	50th percentile			50th percentile		
Nsanje and Chikwawa	VdS	NDVI	P, SM, DS, NDVI, LST	VdS	NDVI	P, SM, NDVI, LST, DS
Chikwawa	ESA	VOD	DS, SM, P, LST, VOD	ESA	NDVI	SM, P, LST, DS, NDVI
Nsanje	ESA/VdS	VOD	ONI, LST, VOD, P, SM	ESA/VdS	VOD	SM, LST, VOD, ONI, DS
Mitole	VdS	NDVI	SM, LST, DS, NDVI, ONI	VdS	VOD	SM, DS, VOD, ONI, LST
Mikalango	VdS	VOD	ONI, LST, P, VOD, SM	ESA	VOD	SM, LST, P, VOD, DS
Zunde	VdS	VOD	DS, ONI, SM, P, LST	VdS	VOD	DS, VOD, LST, P, ONI
Mahkanga	ESA/VdS	VOD	P, SM, LST, DS, ONI	ESA	NDVI	NDVI, P, SM, LST, ONI

Table 6.5: Overview best feature combination at LT3 and LT5 over all areas.

How do the different model scenarios perform in terms of accuracy? The accuracy of the model fluctuate over the growing season; it is hard to say there is one model scenario with a highest accuracy. The model performance reacts strongly on the hit rate and miss rate of a model. The width of the confidence interval is sensitive to the misses in the model. Another factor influencing the accuracy of the model is the number of samples, the accuracy and robustness of a model increases with a larger amount of samples. Increasing the sample size could lead to a more robust system and could explain the reason for a sudden change within the model. Another method is to exclude one feature at the time and repeat the modelling. The weight of one feature can strongly influence the model performance. More research is needed to identify the cause of the immediate increase or an unexpected decrease in model performance. It can be an essential signal in agricultural drought forecast, a bias of the model or noise in the data. Other factors which influence the accuracy of the model are the low sample size; this experiment should be performed with a more extended period to identify the advantages using one dataset over the other. Another bias could be that the model uses VdS data as the data source for a various predictor and predictand features. The LST, SM and VOD are derived from VdS, which might give a bias in the model performance when comparing the ESA SM values with the VdS SM values. There is more thoroughly data analysis needed to understand the underlying causes of the model performance. Within this research, many variables are used to develop the Fast-and Frugal decision tree used as a tool for agricultural drought predictions. This model developed in this research provides a first indication of the influence of the different features in drought prediction — the model should be tested and clarified more into depth in further research.

7

Conclusion and Recommendations

This chapter answers the research questions presented in chapter 1 (section 1.4.1). The objective of this research is to assess the forecast skill of agricultural drought forecasting from satellite-derived products at different lead times before the end of the growing season using a variety of agro-climatic indices. Furthermore, this chapter provides recommendations for both further research and recommendations for 510 to develop and implement the results of this research.

7.1. Conclusion

The answer to the main research question follows from the qualitative and quantitative sub-questions. The qualitative sub-question illustrate the boundaries drought forecast should deal with; it gives information about the crucial months in the growing season to inform agricultural officials and smallholder farmers to reduce the impact of a drought. The quantitative part uses data analysis in the development of the fast-and-frugal decision tree in order to distinguish the best combination of agro-climatic indicators in the agricultural drought forecast in the case-study area of the Lower Shire River Basin in Malawi.

Sub-Question 1

> *How do humanitarian actors and smallholder farmers currently process information on drought forecast?*

In disaster risk management the focus of humanitarian aid shifted from the response and recovery to the mitigation and preparedness phase. The development of an early action protocol needs strong communication over the whole range of involved stakeholders, from governmental organisation to meteorological services and humanitarian organisation to smallholder farmers. Besides the different stakeholders and information sources implemented, the layering of information at different scales (global, national and local) and the integration of scientific knowledge and indigenous knowledge earns more attention. This two-way validation of ground data and scientific data supports the decision-making process, which actions are needed and who is responsible for these actions. Another aspect humanitarian organisation mentioned to take into account is the preference to develop an already existing system, rather than to develop a new parallel system. Although the drought impact does not only related to the meteorological drought, current forecast models use primarily meteorological trigger indicators, such as temperature and rainfall. The combination of these triggers with agricultural and vegetation information improve the agricultural drought forecast. A decision trigger should be designed not only on the probability of drought events but includes the vulnerability, local capacity and operational procedures on the ground. A natural risk is a function of the hazard itself, exposure and vulnerability.

At smallholder farmers level, the information processing behaviour is following the governmental administrative system. Smallholder farmers receive their information from agricultural officials, who receive information from the local meteorological office (DCCMS). The DDCCMS provide the EPA Officers with a seasonal and weekly weather update. This information is translated from English to the local language before communicated to the farmers. Barriers agricultural officials experience in the information processing behaviour is the delay in communication. Some remote areas are difficult to reach, some other need to inform farmers in person when farmers do not have access to a mobile phone, which slows down the communication of weather information. Farmers benefit from a forecast system communicated at a fixed time, where they can come together and discuss their planting decision supported by the agricultural officials. The most crucial times for farmers to receive weather information and planting advise is at the start of the season and mid-season, when the impact of climate-risks of dry spells, destroying winds and floods are high.

Summarised the factors influencing the decision making process in drought forecast is the first signs for planting rains (>30mm rainfall), expected climate risks mid-season, time to put in place operational procedures and local capacity. The time needed for operational and logistical procedures is about two or three months. These factors combined, results in the months November and January being the most critical months in agricultural drought forecast to inform smallholder farmers about the weather predictions.

Sub-Question 2

> *What combination of agro-climatic features perform best in agricultural drought forecast in the Lower Shire River Basin?*

There is a large variety within the best combination of agro-climatic indicators over the growing season and between the different areas. This variety indicates that all drought predictors used in this research contain a certain degree of essence in drought forecast. It is hard to say there is one best combination of agro-climatic indices. It is strongly dependent on the months within the growing season and the spatial characteristics. It can be said that at the start of the season initial climate conditions are slightly more critical, such as the SM, LST and precipitation rates. These are the features initiating a good start of the growing season. Over the growing season, the vegetation feature gains more importance, the development and the health of the crops indicate the impact of drought they can withstand. Another feature with increasing importance over the growing season is the dry spells. When comparing the feature ranking related to the threshold values of the 20th, 40th and 50th percentile, it shows their is a slight difference in feature ranking. The 20th percentile include ONI as a feature, especially with VOD as predictand. This ONI coming in with a 20th percentile is due to the fact this threshold value model the more extreme drought events with a five-year return period. This coming in of ONI might indicate that the ENSO influence these drought extremes. This relation between drought extremes and El Niño is visible in the historical time series and impact data from EM-DAT and GDO. Another feature difference is the model performance with different predictands, in two-thirds of the models, the VOD is returning as best predictand compared with NDVI. Despite the better performance of VOD in general, the NDVI appears to show higher WACC values at LT5. As mentioned earlier, the model performance and feature combination are influenced by the threshold values and the lead time over the growing season. An exciting outcome of the model performance is the drop at LT3. This drop is related to the number of misses the model predicts. This drop occurs more often with the 20th percentile than within the 50th percentile. Over the growing season, the drop might relate to the features included in the top five ranking. In most cases the drop happens, when the NDVI or VOD state is left out or lower ranked than the months with higher WACC. The last aspect to take into account is that probably the 50th percentile return better model performances because the model contains more drought events to train and test the model with, which increases the model performance. The reliability of the model increases with the more data used as a training and test set.

Research Question

> *To what extent can machine learning and satellite-derived products be used to obtain agricultural drought forecasts with high forecast skill in early warning early action?*

This research uses the machine learning algorithm fast-and-frugal decision tree. The advantage of this algorithm links to its transparency, simplicity and accuracy with limited data. This algorithm is, in particular, suitable when the decision-making process should be understood, implemented and communicated quickly after a trigger exceeds. In early warning early action the humanitarian organisation benefit from these model qualities. Machine learning and satellite-derived products can be in early warning early action systems when combined with indigenous knowledge and ground data. Especially the layering and validation of data are essential to increase the forecast skill of early warning early action. This layering and validation could be tested more in-depth in future research. In this research, the actual scale of the research areas was too coarse to identify the real effect and possible benefits from using high-resolution satellite-data, which could increase the forecast skill. By averaging the data over relatively large areas, the extremes level out. In drought forecast, the scale of the events is relatively large in comparison with the impact of floods, which makes that this research gives a first indication and estimate of which features are important at what time in the growing season, but does not directly increase the forecast skill. The model performance reacts strongly on the hit rate and miss rate of a model. The model performance needs to future specified. The interesting outcome of a drop occurring at LT3 mid-season could indicate an essential signal in agricultural drought forecast, a bias of the model or noise in the data. Within this research, many variables are used to develop the fast-and-frugal decision tree used as a tool for agricultural drought predictions; the different elements should be clarified more in detail to understand the effect of the independent variables on the forecast skill of the model.

7.2. Recommendations future research

This researched focus on various disciplines, covering both the analysis of the impact of drought and the forecast model itself. For further research, it is recommended to go more into depth in the different disciplines. The model itself should be validated and evaluated in more detail. As mentioned the model average the data over relatively large areas, which levels out the extremes and give more general information over a district. The high-resolution soil moisture data used in this research does not prove its quality over these large areas. The lower resolution data is suitable in agricultural drought forecast at the district level, while if the impact of drought on plot scale is required, the higher resolution data will probably be needed to increase the forecast skill. Another aspect of improving the model performance is the feature selection. The feature selection leave-one-out can be used to identify the influence of one feature on the model performance. The last data related comment is the implementation of weight factors to the features. Over the growing season, some features are more important than others; this could be translated into using certain weights per lead time. As an addition to the features used in this model, wind, start and end of the season could be included in the model. Moreover, this research discusses the size of the dataset and that a larger dataset will give a more robust model. It would be interesting to test the model over a longer time interval to test if the model performance would increase. Another validation of the model would be to test the model with an existing baseline, for example, the Global Food Security Index or the Integrated Phase Classification (IPC) used by FEWS NET (Glo), (FEWS NET, a). Furthermore, the combination of scientific thresholds and indigenous knowledge is suggested. The translation of indigenous knowledge to vulnerability threshold values would be valuable to include in the model. Besides the different data use the comparison of different machine learning algorithms will further develop agricultural drought forecast. Within the different machine learning models, an index can be used in the drought classification instead of a binary classification. The intensity of an event can be classified agricultural drought into different classes, such as extreme, severe, moderate and mild like is done in the US Drought Monitor (Abo). The last aspect of the model is to identify the cost-loss ratio when to act. Research on how to translate drought impact to damage and to identify the cost-loss ratio would of great value in this research area.

Besides the model aspects, the relation between other forecast models could improve the forecast skill of the agricultural drought model. In current forecast models, the relation between flood and droughts is not taken into account, while framers are more vulnerable for a sequence of natural hazards. A relatively low impact scientific drought could cause a high impact, due to the fact farmers are not prepared for a new natural hazard following a previous one. This layering of flood information with drought data could improve the agricultural drought forecast. Another model interesting to link to the agricultural drought prediction is the food security models.

All suggestions summarised:

- Increase, the time interval of the dataset
- Compare larger and smaller areas with each other to identify the influence of levelling out the extremes in the data.
- Include vulnerability threshold values translated from indigenous knowledge.
- Feature selection include; wind, the start of season and end of the season as predictors.
- Feature selection Leave-one-out to identify the influence of the individual features on the model performance.
- Add weight to the feature selection over the growing season.
- Validate the model with an existing baseline.

- Machine learning algorithms based on an index instead of binary classification.
- Translate drought impact to a cost-loss ratio by the humanitarian organisation when acting in vain.
- Model relation between flood- and drought events.
- Model impact of a sequence of natural hazards on agricultural drought predictions.
- Relate the agricultural drought forecast to a food security model.

7.3. Recommendations 510

This research gave 510 important insights for further development of impact-based forecasting for drought events. The machine learning model based on the fast and frugal decision tree allows for a seasonal forecast of whether that season will experience an agricultural drought. The prediction concentrates on two moments within the growing season, at the start of the season and mid-season. These lead times relate to the continuum weather and climate timespan relevant to the end-users requirements.

The Red Cross movement utilises or is planning to use forecasts at the seasonal, sub-seasonal to the short term timescales in the context of disaster risk reduction and (early) preparedness through FbF. The seasonal agricultural drought forecast model developed in this research can encourage humanitarians as well as smallholder farmers to take action already at the beginning of the dry season. The Early Action Protocol can incorporate the input for the triggers for a drought that will develop in due time. The Red Cross movement does not yet develop EAP for droughts as the focus has been first on developing the already very new and innovative early warning early action mechanism for sudden-onset disasters (such as floods, cold waves and typhoons).

The qualitative research performed in this research in the form of the focus group discussions with farmers resulted in several lines of possible early actions the farmers can take. The combination of these FGD with the results of the interviews with humanitarian organisations gave several possible early actions, which could be implemented by the professional responders. Possible interventions currently in place within the Red Cross Movement ranging from Crop, Cash, Livestock up to other livelihood related interventions. FAO has an even more extensive list of activities to take for reducing food insecurity, which also relates to drought. Overall, the seasonal forecast can be used to update contingency plans for droughts. More specifically in terms of crop and cash interventions, it can enable to take early actions at the start of the season (for example subsidies for drought resilient seeds). The ML model is now still at a too coarse scale in order to target specific communities. For this, one has to make use of, e.g. existing governmental social protection programs where vulnerable and poor households are pre-identified. It will also be still necessary to compare the ML model with existing baseline models, for example, the ones FEWS NET uses.

This research with the extensive and thorough analysis of possible data sets and features in agricultural drought forecast provide 510 with valuable insights for other ML models that 510 is developing. Furthermore, it gives 510 insight into if higher resolution data is valuable to invest in or not.

All in all, the ML forecasting model cannot be used directly to provide forecasts at different intermittent timescales throughout the dry season; however, the predictors used to provide pointers for the agro-climatic indices will be used in the NERC SHEAR IPACE project to develop (mobile) climate services for farmers.

A
Interview protocol

A.1. Formal request of an interview

05.08.2018

Formal request for an interview in the scope of our master theses

Dear Mrs Jones, dear Mr Lombardi,

As part of our master dissertations at the Delft University of Technology and the Maastricht University we are currently conducting expert interviews. We are both graduate interns at 510.global, the Netherlands Red Cross Humanitarian Data Team, with separate research topics and approaches. However, both our topics seek to get a broader knowledge about the current project landscape of the forecast-based financing pilots and especially the decision-making processes that lead to the introduction of an early action.

- Aim of the researches
  - Decision making in early action Dominik Semet

    The project is an explorative study to examine how the use of open data can strengthen the integration of food security aspects into the Early Warning Early Action/ Forecast Based Financing projects. The thesis shall provide an overview over possible early actions concerning food security as well as the data the respective decision making is based on. Therefore, food security interventions are mapped and assessed according to their usability in early action. In a second step the decision making process to implement these interventions is examined. Aim is thereby to develop approaches using open data to systematize and specify the humanitarian decision making concerning food insecurity.

  - Agricultural Drought Forecast Marijke Panis

    The focus of the research is to identify develop/optimize an agricultural drought forecast model using remote sensing data. Important in the research is first understanding the decision-making process in early warning early action as a reaction to a drought forecast based on a threshold value. Second the agro-climatic indicators which can be used in drought forecast and what threshold can be used to predict a drought with a certain probability. Last this threshold value indicators will be combined in a model to find the best performing indicators in drought forecast supporting the decision-making process. The objective of my study is to support humanitarian decision makers in their pre-disaster response by increasing the forecast skill of agricultural drought to reduce crop failure and food insecurity in Malawi.

- Interview goal

  This interview is used to get insight in the decision-making process of the humanitarian sector and other involved stakeholders in early warning early action and food severity caused by natural disasters. These insights will be used to identify the factors and actions responsible for the decision when to start early warning early action in forecast-based financing projects. to identify the current drought forecast situation, what stakeholders are involved, what the current decision-making process in drought

  forecast is, what indicators/indices are necessary to support humanitarian actors in

A.2. Information form interview

Natural Disasters in relation with agricultural drought and food severity

Marijke Panis	MSc Watermanagement	Delft Univeristy of Techology
Dominik Semet	MSc Public Policy	Maastricht University

Information for participants

Thank you for participating in these studies, which will take place from July 2018 to February 2019. With this document we would like to provide you with the purpose of the two different research studies and a description of the rights you have as a involved participant.

1. Research Topic and the aim of the research
   As mentioned above this interview will be used in two different research projects. The aim of the first research project is to identify the current situation of agricultural drought forecast, the decision-making process of local governments and other stakeholders in agricultural drought forecast and what indicators/indices are used in drought forecast models/procedures. The second research project focus on food security in the decision-making process in Early Warning Early Action approaches. The thesis shall provide an overview over used relief measures as well as over the data based used for the decision making in case of Food insecurity. Aim is thereby to develop approaches using open data to systematize and specify the decision making in cases of food insecurity. The data for both theses will be generated through semi-structured interviews and by the analysis of open source data from different platforms.
2. Taking part in the research
   Participation in these researches is voluntary, you can decide if you want to take part or not. If you want to take part of the research, we would like to ask you to sign the consent form, which can be send in return before the interview take place.
3. Involvement in research
   The interview will be semi-structured through either Skype/telephone or in person depending on the location of the interviewee. The interview will be audio-recorded in order to process the information after the meeting.
4. Content of the interview
   The interview is a way of collecting data and information about the humanitarian decision making process, forecast procedures and mechanisms to prevent food insecurity. This information will be used as part of the Master dissertation at Delft University of Technology (Marijke Panis) and Master dissertation at Maastricht University (Dominik Semet) in collaboration with 510 initiative of the Netherlands Red Cross. In the further the information can be used in research performed by 510.
5. Confidential/anonymized
   Because the interview will be audio-recorded these records will be kept as confidential as possible. Only Marijke Panis and Dominik Semet will have access to the audio files and will be deleted after finishing the Master Thesis. The data will be anonymized through data masking and it will be asked specifically when specific citations will be used in the report. Your name will be handled with care and not directly linked to the

content of the interview in the way that digital files, transcripts and summaries will be

A.3. Interview questions

Questionary Semi-Structured Expert Interviews **23.07.2018**

Marijke Panis	MSc Watermanagement	Delft Univeristy of Techology
Dominik Semet	MSc Public Policy	Maastricht University

Introduction – Information about research and goals

Overview – general

1) General information about the organization and respective position?
    a. What fields are you working in (Humanitarian, Development, DRR)?
    b. Current position and responsibilities within the organization?
2) Which early warning early action projects is the org. currently running?
    a. In which countries?
    b. Which disaster scenarios and which target groups are addressed?

Procedures and Instruments

1) What drivers of food Insecurity are you focussing on & what are specific instruments addressing these targets (general/EWEA)?
    a. Difference in addressing chronic and transitory food insecurity?
2) Can you give a short overview over most common instruments used by your organization to address food insecurity?
    a. Regarding floods & regarding droughts
    b. Before, during, after a disaster (DRR; Classic relief, etc.)
    c. For respective target groups (farmer groups, general population)
    d. In EWEA

Drought and flood forecasts

3) What forecast procedures or models are you using in respective EWEA projects?
    a. Indicators & Dates for the forecast (ndvi, spi, ground data gauge)?
        i. Climatological or agricultural indicators or a combination?
    b. What are first early warning sights for droughts?
    c. From whom is the information coming (met office, authorities, etc.)?
    d. On what time frame do you receive this information?
    e. Which other actors are involved?

f. Is additional local knowledge used in the forecast (Bottom up information, local

B

Protocal FGD

B.1. Protocol University of Leeds

Protocol for focus groups with farmers to identify key agro-climatic indices

This document sets out a protocol for conducting focus group discussions with groups of farmers. These can be composed of both male and female and should involve between 5 and 10 participants. The research team should be composed of one facilitator, one translator and the discussion should be audio recorded (if possible). The audio files need to be fully transcribed into English by a translator. Observations about the nature of responses and group dynamics should also be noted (for example, if responses are being dominated by one individual; that there doesn't appear to be consensus; that participants do not appear to understand the question etc.)

All outputs should be photographed and coded, and field notes taken to clearly document which photographs relate to which aspects of the discussion. The post-its added to the discussion sheets should be taped to it in order not to misplace any during travel!

0. Initiation
Seek and follow advice from key informant (District Officer or village head, for example) about appropriate ways of beginning the interview, e.g. with introductions from seniors, allowing everyone to introduce themselves, opening prayer etc.

1. Introduce the Research:
Thank you for accepting to participate in and contribute to this exercise which is being conducted as part of the IPACE-Malawi project, which is being undertaken by the Malawi Red Cross Society in partnership with organisations from the UK and the Netherlands. Through the project we hope to learn about the key climate challenges for agricultural systems in central and southern Malawi and contribute to improving the forecasting and delivery of agriculture-specific weather information to farmers' and humanitarian and disaster response organisations.

This exercise and discussion is designed to build an understanding of agricultural systems, and the weather related risks that they have experienced over recent years, as well as to understand the weather information that you currently receive and how it is used. You will be asked to collectively build a typical agricultural calendar and think about how practices/crops change in years of unusual weather conditions.

The exercise will take about two hours. You are reminded that you do not have to answer any questions that you do not feel conformable with, and that you are free to withdraw at any time.

Record the following:

Interview Information	
Enumerator's Name	
Date of Interview	________/_______ /________ *(DD/MM/YY)*
Start Time	
End Time	
Interview Code	*(ADD Name_District Name_ Interview No_DDMMYY)*

Participants' Information	
Participants' Names	
Province	

District Municipality	
Local Municipality	
Village Name	
GPS Coordinates	

2. Constructing a Crop/Farm Calendar

Using a large (A2) pre-prepared sheet (sheet 1) discuss the five pre-prepared rows and ask participants to describe and position (to indicate typical timing), agricultural practices, the status of crops and fields, and key climate-related risks that are encountered.

These can be anything that participants believe to be significant, but there may be a need for prompts or clarification questions. The following prompts are provided as examples:

Agricultural practices	Land preparation (specific techniques), planting (by crop), weeding, fertiliser/top-dressing application, drying (specific techniques), storage
Description of field/crop when there are good weather conditions	Weeds emerging (specific types), crop pests (specific types), tall/short stems, evidence of damage to crops, health of plants, moisture in grain
Key climate risks	Presence/absence of rain, rainfall quantities and durations, dry spells (absence of rain over specified duration), high temperatures, high winds, hail stones
Description of field/crop when identified climate risks happen	Weeds emerging (specific types), crop pests (specific types), tall/short stems, evidence of damage to crops, health of plants, moisture in grain
Other non-climatic conditions affecting their agricultural practices	Crop insurance, accessibility to seeds, availability of labour, availability of fertilizers

The enumerator should capture these points (in English or Chichewa) on post-it notes and where necessary recorded the span of time (in days/weeks/months) and dates/days of relevance of any description. Where a description is relevant across several months, post-it notes can be combined to indicate this spread.

This sheet should be discussed and iterated/added to until all of the group members are satisfied that it represents a complete picture of the typical crop/farm calendar. <u>In addition, arrows should be drawn in the sheet in order to identify the linkages between the agricultural practices and the other aspects discussed: description of field/crop in good conditions, climate related risks, description of field/crop when those climate risks occur and other non-climate factors influencing their agricultural practices.</u>

A post-it note should be added to top left of the sheet with the code: *ADD Name_District Name_ Interview No_DDMMYY_Sheet 1.* The sheet should be photographed such that it is clearly legible.

Sheet 1

ADD Name_District Name_ Focus Group No X_DDMMYY_Sheet 1	July	August	September	October	November	December	January	February	March	April	May	June
What are your main agricultural practices?												
What do you expect to see on a field/crop												

when there are good weather conditions?												
What are the key climate-related risks?												
What do you expect to see on a field/crop when those climate risks happen?												
What other non-climate factors affect these agricultural practices?												

2.1. Use of weather/climate information (for farmers)

Using the sheet developed above, ask participants to discuss:

- The most critical weather risks/events affecting their crops (probe to try and make these as specific as possible (i.e. 'consistent rainfall every day for the first three weeks after planting', rather than just 'rainfall')?

Note: If the participants identifiy more than 10 weather risks perhaps ask them to prioritize the key 6-8 weather risks (you can circle those in the sheet used above)

- Do they use/receive weather/climate information?

If <u>they receive</u> weather/climate information:	**If <u>they don't</u> receive weather/climate information:**
• What information do they use?	• Why don't they use it?
• How does that information help them with their crop-related decisions?	• If they had access, what information would be useful for them to have?
• From where do they get this information from (the source of the information)?	• How would that information help them with their crop-related decisions? What of these would be the critical decisions to them?
• How do they receive this information (through which mechanisms e.g. extension officer?)	• How would they like to receive this information (through which mechanisms e.g. extension officer)?
• In what format do they receive it (e.g. information on the radio, piece of paper, etc)?	• In what format would they receive to it (e.g. information on the radio, piece of paper, etc)?
• How often do they receive the information?	• How often would they like to receive it?
• Is the information they receive useful to inform your decisions? Why/why not?	• What conditions would need to be place for allowing them to use the information? e.g. access to the information, help

	understanding it, receiving it at the right time of the year. (Please ask them to describe these conditions in detail)
• Do you trust the information? Why/why not?	• Any other comments they would like to add to this?
• What other information would be useful to have? Why/why not?	
• How would you like to receive the information (e.g. extension officer, radio, etc) and in what format (e.g. verbal message, text message, etc)? Why?	
• Would it be more useful for you to receive this information at different times of the year? Why?	
• What else would you change about how this information so it would be more useful to help you manage your agricultural practices?	

3. Recalling and Describing Recent Seasons

Ask the participants to recall that season and think about how: (1) practices, (2) the description of the crop/field, and (3) the key challenges they found.

The facilitator/translator should use a different sheet (sheet 2 – see below) to add other post-its to it.

Note that this time, under the third row, participants are being asked to consider the key climate challenges that were experienced, as opposed to the risks that they are typically concerned with in an average season.

This sheet should be discussed and iterated/added to until all of the group members are satisfied that it represents a complete picture of the 2015/16 crop/farm calendar.

A post-it note should be added to top left of the sheet with the code: *ADD Name_District Name_ Interview No_DDMMYY_Sheet 2.* The sheet should be photographed such that it is clearly legible.

Sheet 2

ADD Name_District Name_ Focus Group No X_DDMMYY_Sheet 1	July	August	September	October	November	December	January	February	March	April	May	June
What were your main agricultural practices?												
What did you see on your field/crop during that season?												
What were the key climate												

challenges during this year?												

This process should be repeated for the most recent year of 2017/18 (sheet 3) season. If there's time ask participants about the latest season of 017/18 (if so, it may be easier to create a 3rd sheet to discus and include post-its

3.1. Use of weather/climate information during the El Niño year of 2015/16

Using **Sheet 2** developed above, ask participants to discuss:

- What was the most critical weather risk/event affecting their crops (probe to try and make these as specific as possible (i.e. 'consistent rainfall every day for the first three weeks after planting', rather than just 'rainfall')?
- Did they use weather/climate information? If yes, what information did they use?
- From where did you receive it? How did you receive it? When did you receive it?
- Did it help with their crop-related decision-making? If so, how?
- Would you like to improve that information (e.g. type of information, source, format for receiving, timing)? If so, how?

4. Close the Session

Thank participants for their time. Explain that this information will be combined with that collected in other locations and that we will use it as a basis of exploring how climate information and forecasts can be improved. Explain that this will be a long term process (to manage expectations about immediate benefits), but that we plan to share the findings of the research through the District Office/village head. Give seniors/district officer/village head opportunity to give final closing words, as appropriate.

…………………………………………………….

Questions for extension officers

These interviews should also be audio-recorded and afterwards fully transcribed (word by word) into English.

- How long have you been working as an extension officer in this section/TA?
- With what villages do you work with?
- What type of support/information do you provide to the villages you work with?
- Do you provide weather/climate information to the farmers?

If <u>they provide</u> weather/climate information:	**If they <u>don't provide</u> weather/climate information:**
• What type of information do you provide (e.g. forecasts for the next 5 days, seasonal forecast, etc)	• Can you explain why you don't provide weather/climate information to the farmers?
• From where do you receive this information (please describe the whole chain from whoever produces the information down to getting to you)?	• In your opinion, are the farmers with whom you work interested in receiving this type of information? Please explain why/why not?
• How do you receive this information? E.g. text message, radio, etc	• If the farmers are interested in receiving this type of information would you be able to provide them with that information? Please explain why/why not?
• In what format do you receive that information? (e.g. a text describing the weather/climate conditions for that period)?	• What would you require to be able to provide them with that information? e.g. having access to the weather/climate information, having training on how to understand the information, more resources to be able to reach the villages, etc.
• Do you work on that information before passing it on to the farmers? (e.g. do you change it in a way so that the farmers understand it better?). If so, what do you do?	• Would you be interested in providing that information to the farmers if you had the right conditions for doing it?
• How frequently do you receive this information?	
• How frequently do you give this information to the farmers?	

• Are there critical periods in the cropping calendar when this information is very relevant to the farmers? If so, why and when are those periods?	
• Do the farmers use this information to help them make decisions about how they manage their agricultural practices? If so, can you give us a few examples of how farmers have used it e.g. what type of decisions did they made based on this information and when?	
• In your opinion, how could this process of providing weather/climate information to the farmers be improved?	
• What else could be improved to help you better deliver this information to the farmers?	

B.2. Protocol data collection Lower Shire River Basin

AN INITIATIVE OF THE NETHERLANDS RED CROSS

Agricultural drought forecast in the Lower Shire River Basin (NERC SHEAR)

FIELD DATA COLLECTION PROGRAM PROPOSAL
LILONGWE AND LOWER SHIRE RIVER BASIN

WORKING DOCUMENT

Marijke Panis
mpanis@rodekruis.nl

Contents

Introduction

As part of my Thesis (MSc Watermanagement at Delft University of Technology), ECHO II and NERC SHEAR project of 510 an initiative of the Netherlands Red Cross, this document contains a summary of the research and how the field data on drought forecast will be used in this research.

The aimof the research project is to identify the current situation of agricultural drought forecast, the decision-making process of local stakeholders in agricultural drought forecast and what agro-climatic indicators/indices are used in current drought forecast models/procedures. This information will be used to meet the objective of my research to assess the forecast skill of agricultural drought from satellite-derived products at a different lead time before the end of the growing season using agro-climatic indices.

With the identification of different agro-climatic indices, a Machine Learning Model (Fast-and-Frugal Decision Tree) will be tested. The model is used to identify the optimal threshold values of the best skilled agro-climatic indices in agricultural drought forecast in the Lower Shire River Basin. Before testing the model the involved local stakeholders and the current drought forecast need to be identified, as well as the impact of agricultural drought on smallholder farmers.

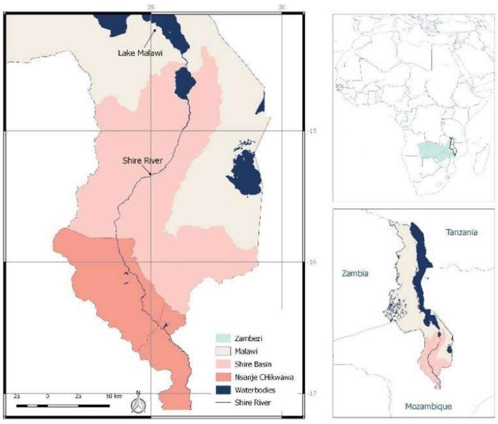
The field data collecting will exist in semi-structured interviews and expert meetings. The interviews will be used to get insight into the decision-making process of the involved stakeholders and smallholder farmers in early warning early action and drought forecast. These insights will be used to identify the factors responsible for the decision making and which agro-climatic indices are available and necessary to support (humanitarian) actors in there drought forecast. In the attachment, you can find the research proposal.

Research

The main research question relates to a broad objective of investigating the added value of implementing high spatial and temporal resolution soil moisture in agricultural drought forecast, while the sub-questions are specifically applied to the case-study area the Lower Shire River Basin in Malawi.

The main research question to be answered is:

To what extent can machine learning and satellite-derived products be used to obtain agricultural drought forecasts with high forecast skill in early warning early action?



This question will be answered through two focus areas, split into two sub-questions. First the current Drought risk procedure at both national and regional scale (Lower Shire River Basin) will be analysed by linking historical data and past drought events, followed by analysing current drought forecast models if used and the integration of agricultural and climatic drought indicator in these models. The prediction of drought is based on machine learning techniques to identify the value of high spatial and temporal resolution soil moisture monitoring and the resolution needed for accurate drought predictions at the regional level. Overarching is the practical implementation of a drought forecast mechanism to reduce drought impact and the risk of acting in vain. The research will be performed at two different levels, both at national scale looking at the humanitarian actors involved in Malawi and at the regional level.

1. **How do humanitarian actors and smallholders farmers currently process information on drought risks?**
(a) What actors are involved in the current drought forecast in Malawi?
(b) Which decision making factors can we distinguish in humanitarian sectors information processing behaviour?
(c) What spatial and temporal resolution is necessary for the optimisation of drought forecast at the national and regional level?

3. **What combination of agro-climatic features performs best in agricultural drought forecast in the Lower Shire River Basin?**
(a) Which feature combination perform best in drought prediction at different spatial and temporal resolutions in the growing season?
(b) How do the different model scenarios perform in terms of accuracy?

Model

The Machine Learning model exists of Input values (features) and a binary output value (Drought yes/no). The features are based on the agro-climatic indices SM, NDVI, LST, P, VOD, Dryspell days and the occurrence of drought in the growing season, while the output data is based on crop yield information to predict if there is a drought yes or no at the end of the season.

The outcome of the model is a decision tree, which identifies the weight of each indicator in the drought forecast. This Fast-and Frugal Decision tree also calculates the threshold value identifying the occurrence of drought.

At least eight models will be comparing forecast with different lead times and different composition of predictors. In table 1 the eight models are presented.

$$\begin{bmatrix} \text{F1} & \text{F2} & \text{F3} & \text{F4} \\ \text{P} & \text{SM} & \text{VOD} & \text{Dryspell} \\ '' & '' & '' & '' \\ '' & '' & '' & '' \\ '' & '' & '' & '' \end{bmatrix} \qquad \begin{bmatrix} \text{Y} \\ \text{Drought} \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

Input-Predictors		Output-Predictand	
MODEL	TRAINING SET	TEST SET	PREDICTOR
1	November-February	March	With SM
2	November-January	March	With SM
3	November-December	March	With SM
4	November	March	With SM
1	November-February	March	Without SM
2	November-January	March	Without SM
3	November-December	March	Without SM
4	November	March	Without SM

Table 1 Compare performance different models

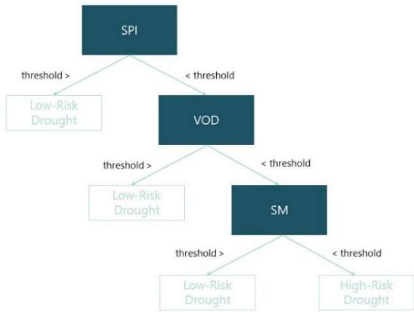
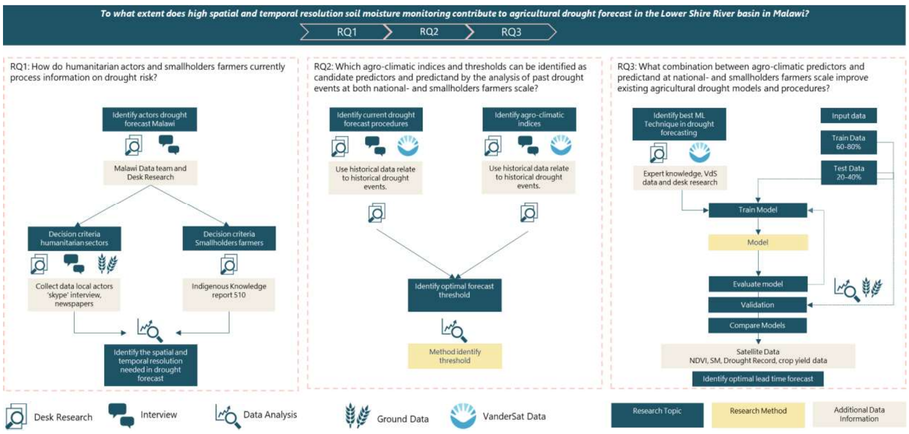


Figure 1: Concept of the FFT decision three

Aim Data Collection

The data collected during this field work will be used mainly in sub-question two. This question focusses on the identification of agro-climatic indices. This knowledge is implemented in the Fast-and Frugal Model to define the predictors and predictant.



1- Data collection Program overview:

In the period from the 10th of November until the 25th of November, a field mission of 2 weeks will take place in Lilongwe and the Lower Shire districts Nsanje and Chikwawa. This field mission exists of a program of semi-structured interviews and focuses group discussions.

Type of information to collect:

National Level:

- Information on existing drought forecast models and early warning early actions systems in Malawi
- Experiences with historical drought events in the Lower Shire River Basin
- Identification severity of drought (output value of ML-model, crop yield data)
- Validate and discuss the first results of the model and outcome of the research with local stakeholders
- Identification agro-climatic indices and ask access to crop yield data of Nsanje and Chikwawa
- Get feedback from partners

District Level or Community level ('Representative' smallholder farmers)

- Information on existing drought forecast models and early warning early actions systems
- Experiences with historical drought events
- Identification agro-climatic indices
- Identification severity of a drought
- Understanding the livelihood of the smallholder farmers during a drought event

Where:

The field mission starts in Lilongwe, where on Tuesday the 13th partners of the NERC SHEAR project come together. During this day, the preliminary results of the research will be presented and discussed. Together with the partners, the district focus group discussions are finalised. The two districts of the research are Chikwawa and Nsanje. Besides the FGD, 4 locations of 10x10 km will be selected for the timeseries analysis.

National Level	District Level	TA Level	Village Level (GVH)
Lilongwe	Chikwawa	Need to be confirmed	Need to be confirmed
	Nsanje	Need to be confirmed	Need to be confirmed

First week program:

Sun 11-11	Mon 12-11	Tue 13-11	Wed 14-11	Thu 15-11	Fri 16-11	Sat 17-11
Field collection planning	MRCR data team meeting	Kick- Off NERC SHEAR	Interviews National Level	Interviews National Level	Travel To Nsanje	TA Level - Nsanje
Field collection planning		Discuss planning FGD	Interviews National Level	Interviews National Level		

Overview NERC SHEAR IPACE

Revised planning:

Sun 18-11	Mon 19-11	Tue 20-11	Wed 21-11	Thu 22-11	Fri 23-11
Travelling to Chikwawa	Meeting Local MLRC, DADO (Chikwawa)	FGD 1 Chikwawa	Meeting Local MLRC, DADO (Nsanje)	FGD 3 Nsanje	Meet DDCMS (Blantyre)
	Meeting EPA Officer and FDG facilitator (Chikwawa)	FGD 2 Chikwawa	Meeting EPA Officer and FDG facilitator (Nsanje)	FGD 4 Nsanje	Travel back to Lilongwe

Daily Planning Monday-Friday

Time of day	# members	Mon 19-11
	Travel time within Chikwawa	
9:00-09:30	Gumbi Gumbi and Marijke	MLRC Office Chikwawa
9:30-12:00	DADO, FGD, Chris, Gumbi Gumbi and Marijke	Meet DADO
12:00-13:00	Lunch and Travel Time	
13:00-15:00	EPA Officer, Gumbi Gumbi and Marijke	EPA Officer 1
15:00-17:00	FGD, Gumbi and Marijke	FGD Facilitator Instructions
17:00	Travel Time Back To Stay in Chikwawa	

Time of day	# members	Tue 20-11
	Travel time Chikwawa to field	
9:00-12:00	Farmers 6-8 person, Translator, EPA Officer 1, Gumbi and Marijke	FGD 1
12:00-14:00	Travel Time and Lunch with EPA Officer 2	
14:00-17:00	Farmers 6-8 person, Translator, EPA Officer 2, Gumbi and Marijke	FGD 2
17:00	Travel Time Back To Nsanje	

Time of day	# members	Wed 21-11
	Travel time within Nsanje	
9:00-09:30	Gumbi Gumbi and Marijke	MLRC Office Nsanje
9:30-10:30	DADO, FGD, Chris, Gumbi Gumbi and Marijke	Meet DADO and EPA Officer 1
10:30-12:30		FGD Facilitator Instruction
12:30-14:00	Lunch and Travel Time	
14:00-17:00	EPA Officer 1, Gumbi Gumbi and Marijke	FDG 3
17:00	Travel Time Back To Stay Chikwawa	

Time of day	# members	Thu 22-11
	Travel time Nsanje to field	
9:00-12:00		Meet EPA Officer 2
12:00-14:00	Travel time and Lunch with EPA Officer 2	
14:00-17:00	Farmers 6-8 person, Translator, EPA Officer 2, Gumbi and Marijke	FGD 4
17:00	Travel Time Back To Chikwawa	

Time of day	# members	Fri 23-11
	Travel time Chikwawa to Blantyre	
9:00-10:30	Gumbi Gumbi and Marijke	Meeting DDCMS (Charles Vanya)
11:00-12:00	Gumbi Gumbi and Marijke	Meeting ... ?
12:00-17:00	Lunch and Travel Time back to Lilongwe	

2- The Strategy at National level :

On Wednesday and Thursday interviews can be scheduled in Lilongwe. Depending on the time it takes to talk with different stakeholders it would be ideal to speak to at least four stakeholders (two days, morning & afternoon). Besides the NERC SHEAR partners, the priority to talk to is DoDMA, DCCMS and MoAIWD. Secondly insurance companies and the WFP and FAO and last the Media agencies.

Who to interview :

- Specific objectives of National interview for the NERC SHEAR project on agricultural drought:
    - MRCS: Historical agricultural drought experiences, forecast, mitigation and response actions.
    - Partners NERC SHEAR (University of Leeds, UK MET Office, LUANAR)
    - DODMA: Access to crop yield data and associated impacts, for Chikwawa and Nsanje districts.
    - DCCMS: Understanding the current weather (drought) forecast systems and El Niño forecast in drought predictions.
    - MoAIWD: Access to crop yield data
    - LUANAR: Knowledge of specific drought impact on the growing season of rainfed agriculture. Agro-climatic indices in drought forecast.
    - NGO (including MRCS): knowledge on drought response

Questions at the National level and District Level:

1. Drought in general
    a. Are you familiar with the different types of drought?
    (meteorological, hydrological, agricultural and socio-economical)
    b. How do you define the severity of a drought?
        i. A dry spell and normal season
        ii. A dry spell and extended dry spell
        iii. Extended dry spell and drought
    c. How do you define the start and the end of a drought event?
    d. What are the main triggers of drought?
    e. Can you identify the first early warning signs of a drought?
    f. Can you describe the development of a drought (before, during and after)?

2. Historical (Agricultural)drought Events
    - Can you remember historical drought events in Malawi?
    - Which districts were affected by this drought?
    - Can you identify the impact of the drought?
    - Was there a drought forecast for those events and were there Early Warning Early Action measures taken?
        - What type of drought forecast?
        - What were EWEA measures taken to reduce the impact of the drought?
    - How were local smallholder farmers informed about the coming drought event?

3. Drought Forecast

- Who is calling out when Malawi is experiencing a drought?
- From whom is the information on drought forecast coming?
- What forecast procedures or models are you using?
  - Indicators & Data (NDVI, SPI, Rainfall patterns, etc.)
  - What agro-climatic indices do you use in drought forecast?
- What is the lead time of weather (drought) forecast?
- To whom and how do you communicate a drought forecast?
- What difficulties do you experience in drought forecast?
- Are the current early warning procedures reliable and accurate?
  - Possible improvements for reliability

4. Early Warning Early Action Measures
   - What mitigation actions can you take to reduce the impact of drought?
   - What is the minimum lead time needed to start with the mitigation actions?
   - What information is at least necessary to activate the mitigation actions?

5. How to get more information about drought occurrence and impact in the Lower Shire Basin?

Questions Extension Officer:

The information from the Extension Officers will be used to givean answer to the following questions:

- Which decision making factors can be distinguished on regional level information processing behaviour?
- What are the drought forecast procedures?
- What can agro-climatic indices be identified in agricultural drought forecast?

The focus during the interviews is to get insides in the information process behaviour at the regional level and how agricultural drought forecast can be implemented in the decision making of the Extension Officer. As an addition, when the Extension Officer has more knowledge about the drought forecast itself, information on the forecast procedures and agro-climatic indices used in the forecast will be discussed more in depth.

6. Drought in general
    a. Are you familiar with the different types of drought? (meteorological, hydrological, agricultural and socio-economical)
    b. How often does drought occur in your EPA/TA/Section?
    c. How do you define the severity of a drought? A dry spell and normal season
        i. A dry spell and extended dry spell
        ii. Extended dry spell and drought
    d. Can you identify the first early warning signs of a drought, if yes which?
    e. Can you describe the development of a drought (before, during and after)?

7. Historical (Agricultural)drought Events
    - Can you call back historical drought events in Malawi?
    - Which districts were affected by this drought?
    - Can you identify the impact of the drought?

8. Drought Forecast
    - Were there drought forecasts for the historical events?
    - Were there Early Warning Early Action measures taken?
    - What is the lead time of weather (drought) forecast?

9. Early Warning Early Action Measures
    - What mitigation actions can you take to reduce the impact of drought?
    - What is the minimum lead time needed to start with the mitigation actions?
    - What information is at least necessary to activate the mitigation actions?

3- Protocol Focus Group Discussions at Community level:

Introduction research

The Focus Group Discussions aim to understand the livelihood of smallholder farmers when a drought event occurs. By the construction of a crop calendar, the agricultural practices and impact of climatic-risks on the livelihood are discussed.

Questions at GVH Level as part of the FGD:

General Community level interview

Survey #:____________________ Date: ____________________
District:____________________ EPA: ____________________
Coordinates visited plot: ____________________
Size of agricultural plot: ____________________
Type of agriculture: ____________________

Name:____________________ Gender: ____________________
Age: ____________________ Time living in community: ____________________
Mean livelihood activities:____________________

Start the conversation with introducing yourself and asking them to introduce themselves.

1. Drought in general
   - Are you familiar with the different types of drought?
     - (meteorological, hydrological, agricultural and socio-economical)
   - How do you define the severity of a drought?
     - A dry spell and normal season
     - A dry spell and extended dry spell
     - Extended dry spell and drought
   - How do you define the start and the end of a drought event?
   - What are the main triggers of drought?
   - Can you identify the first early warning signs of a drought?
   - Can you describe the development of a drought (before, during and after)?

2. Drought specific
   - Can you remember historical drought events in your community?
     - When did the drought take place which period of the year?
     - What year and month can you remember?
   - How often does drought occur in your community?
   - Can you identify the impact of the drought?
   - Can you remember dry spell events in your community?
     - When did the drought take place which period of the year?
     - What year and month can you remember?
   - How often does dry spell occur in your community?
   - Can you identify the impact of a dry spell?

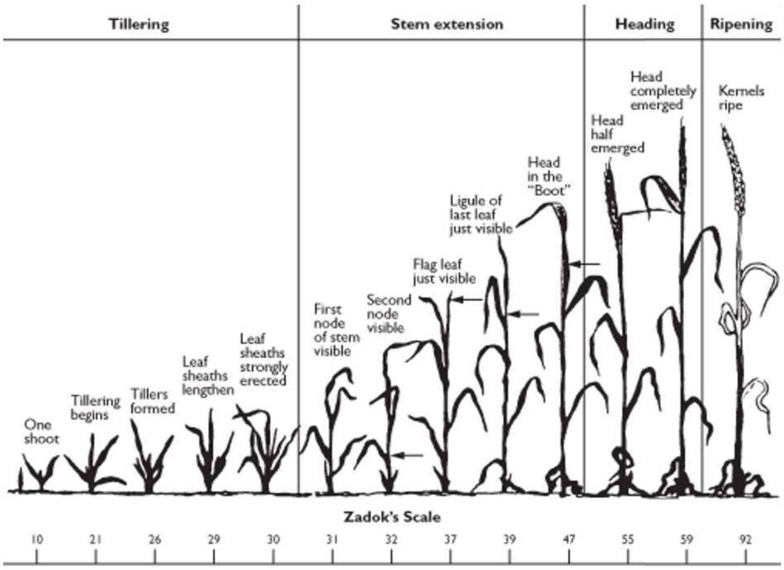
3. Drought forecast
   - Who is calling out when the community is experiencing a drought?
   - From whom is the information on drought forecast coming?
   - What forecast procedures or models are you using?
     - Indicators & Data (NDVI, SPI, Rainfall patterns, etc.)
     - What agro-climatic indices do you use in drought forecast?
   - What is the lead time of weather (drought) forecast?
   - What difficulties do you experience in drought forecast?

4. Before the Drought Early Warning Early Action
   - What actions do you take when a drought is coming?
     - Change of livelihood
     - Changing planting practice (crop type, planning period)
     - When will this happen?
     - Selling of livestock?
     - Did you move to other parts of the country?
   - What time do you need to prepare for a drought?
   - Who is giving you information about the weather forecast?
   - What information is at least necessary to start the preparations for a drought?
   - How do you want to receive information about an upcoming drought?
   - What do you need in case of an upcoming drought?

CUSTOMERS JOURNEY: GROWING PROCESS CROPS

2	Timeline	Months

1

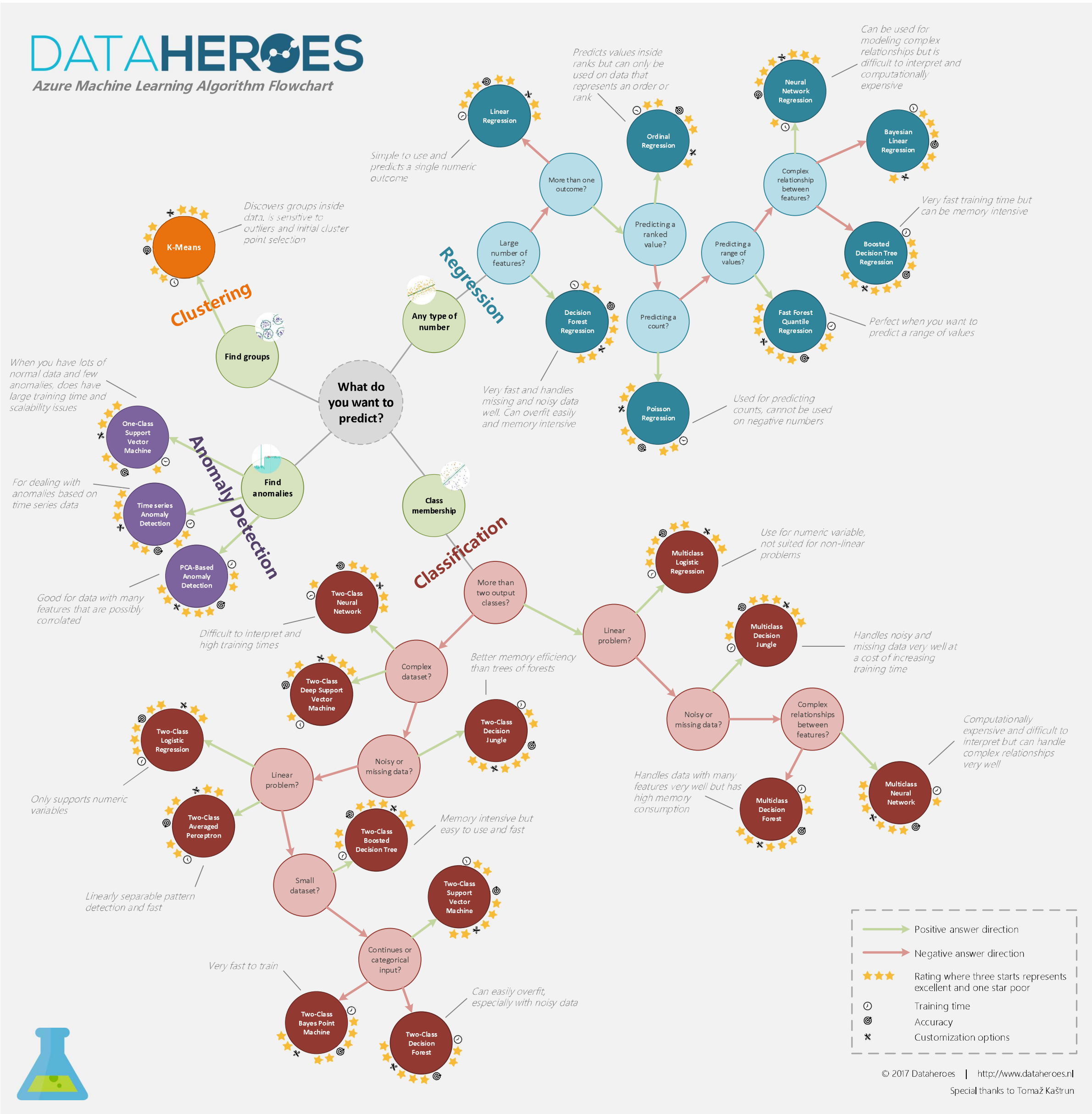


	Before	During	After
3			
4			
5			
6			

1: Start with drawing the growing process of maize or other crops (Warm-up)
2: Let them set their time line
3: Actions: before, during and after
4: Resources activities in normal situation
5: Risks before, during and after (unforeseen, drought, insect epidemy)
6: Resources to prevent risks

C

Cheat sheet Machine Learning Algorithms



D

Crop Calendars FGD

D.1. Crop calendar Mitole

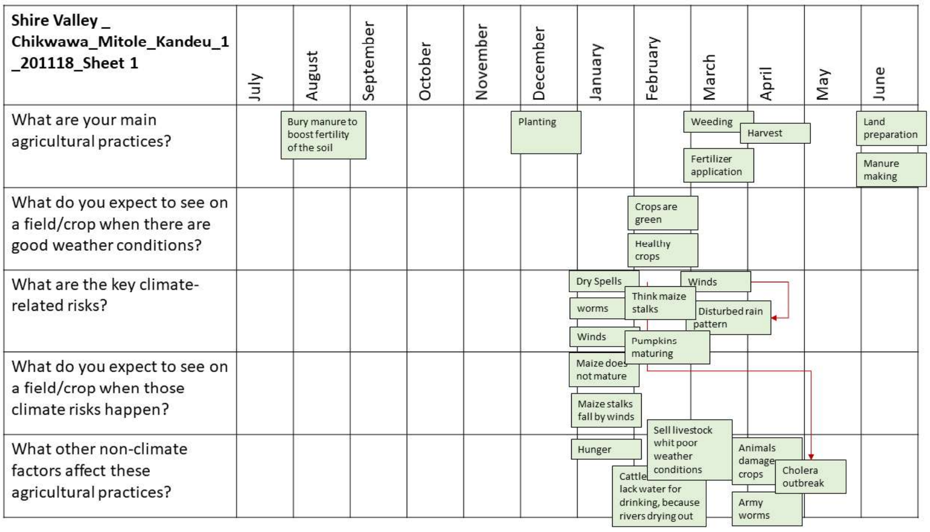


Figure D.1: Crop calendar Mitole normal growing season

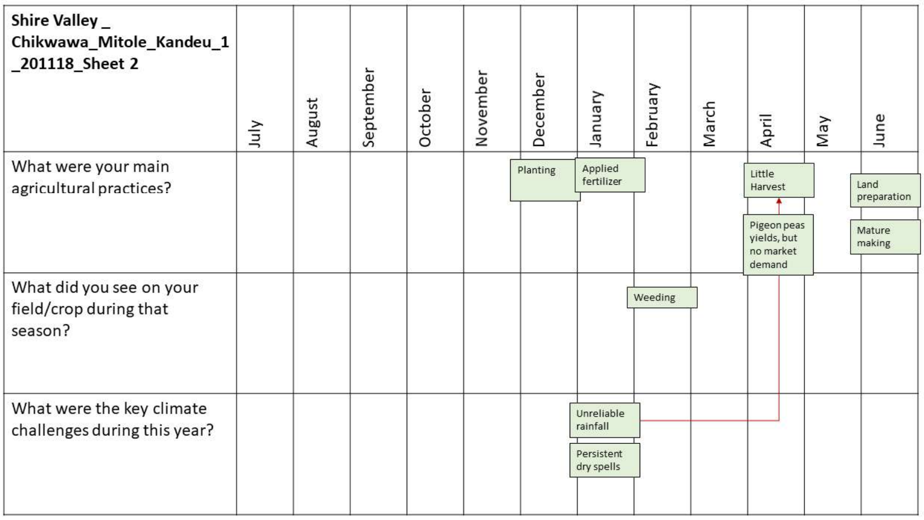


Figure D.2: Crop calendar Mitole season 2015 - 2016

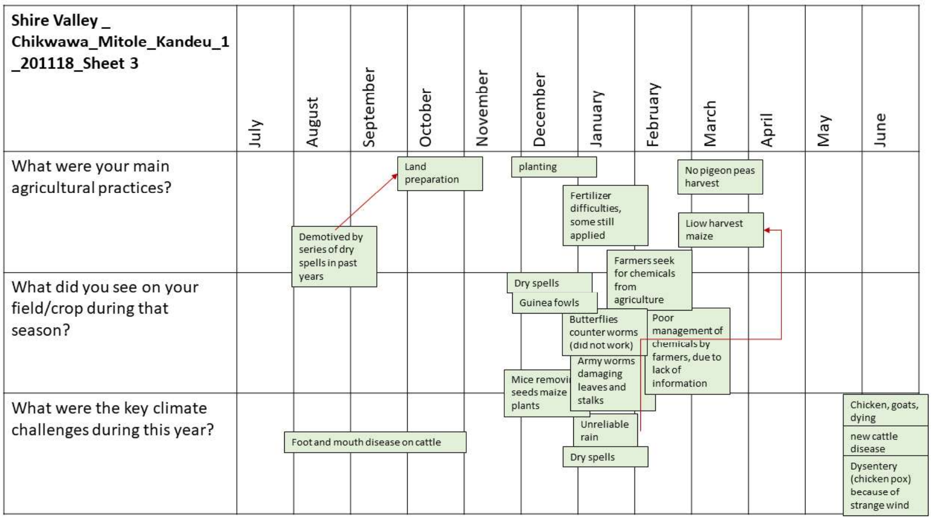


Figure D.3: Crop calendar Mitole season 2017 - 2017

D.2. Crop calendar Mikalango

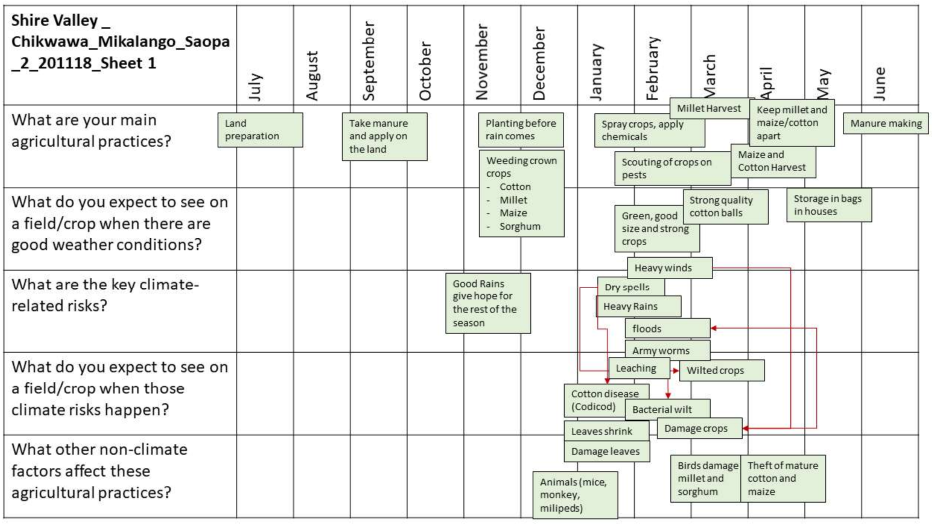


Figure D.4: Crop calendar Mahkanga season normal growing season

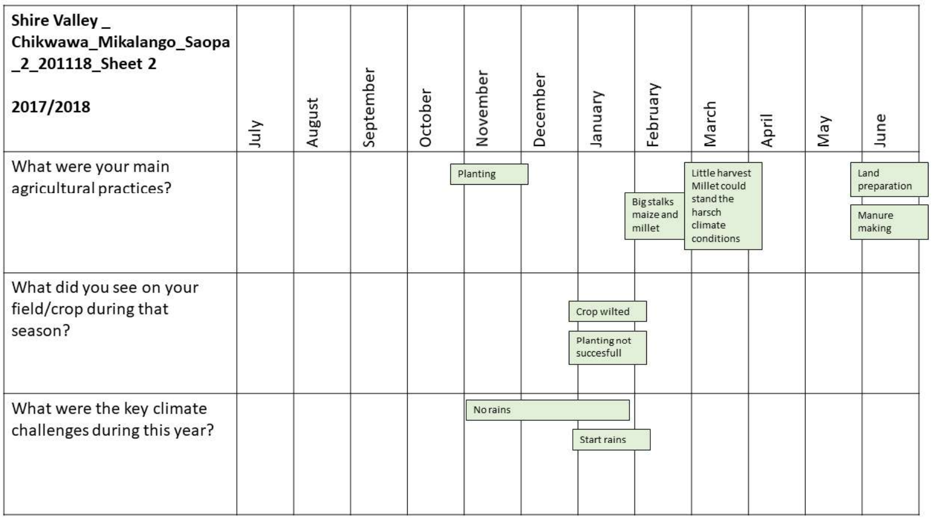


Figure D.5: Crop calendar Mikalango season 2017 - 2018

D.3. Crop calendar Zunde

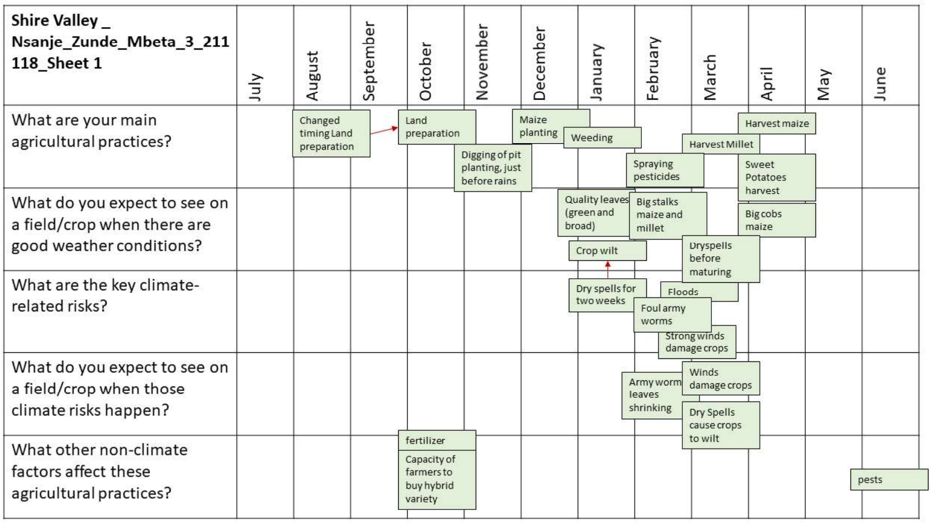


Figure D.6: Crop calendar Zunde season normal growing season

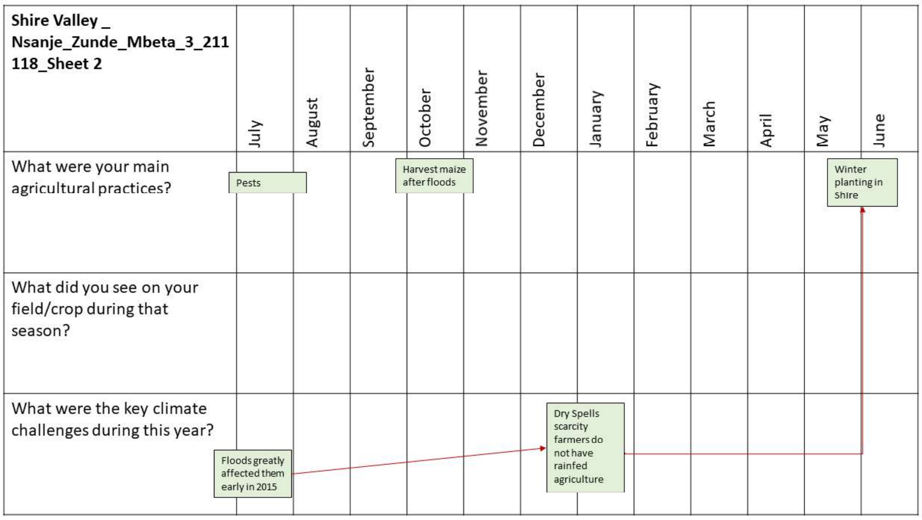


Figure D.7: Crop calendar Zunde season 2015 - 2016

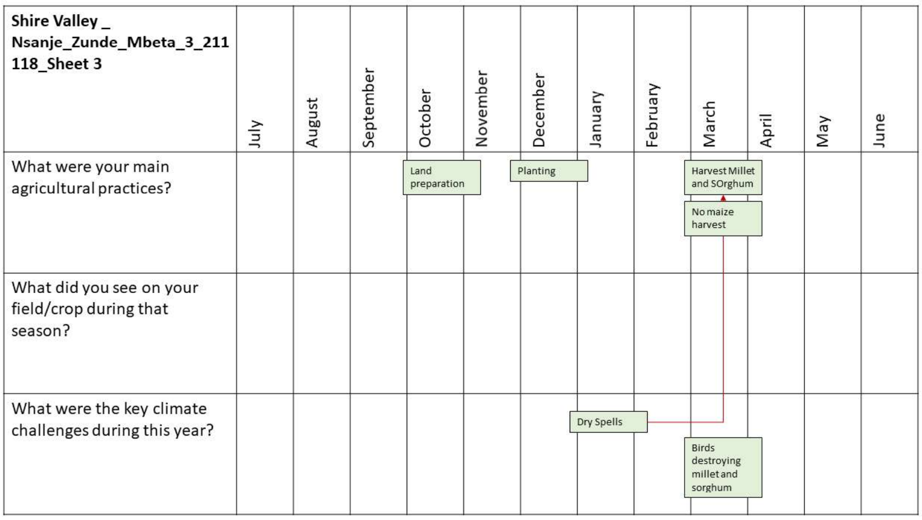


Figure D.8: Crop calendar Zunde season 2017 - 2018

D.4. Crop calendar Mahkanga

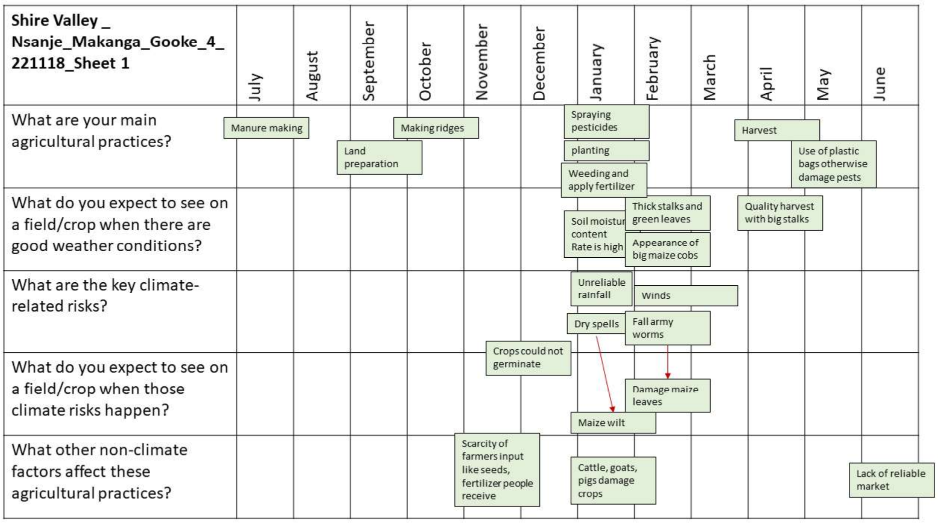


Figure D.9: Crop calendar Mahkanga season normal growing season

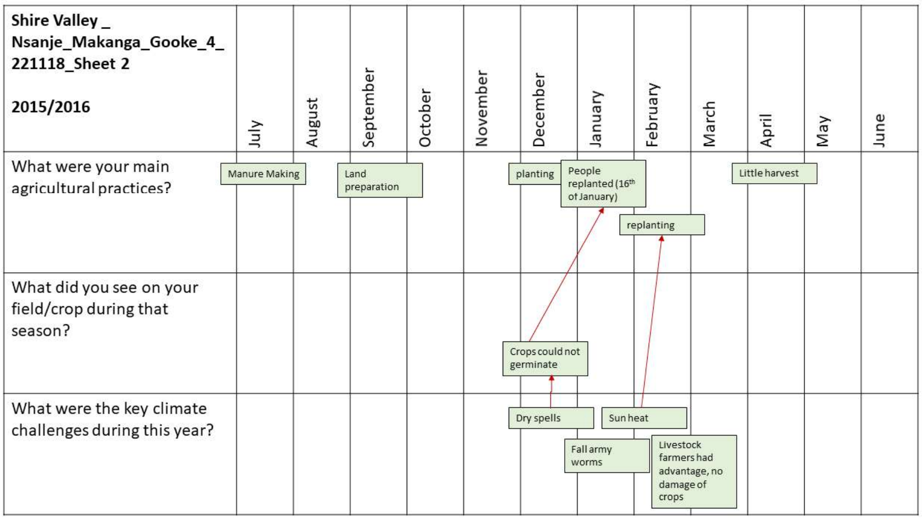


Figure D.10: Crop calendar Mahkanga season 2015 - 2016

E

Timeseries

E.1. Historical time series

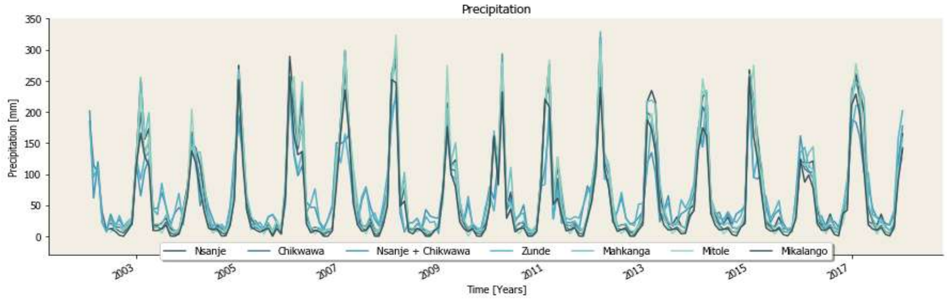


Figure E.1: Historical precipitation time serie

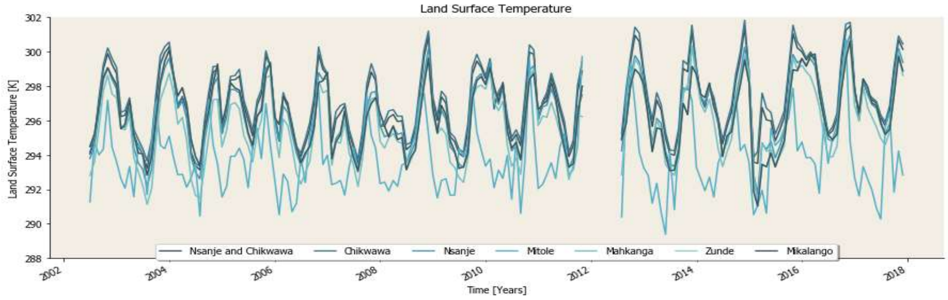


Figure E.2: Land surface temperature time serie

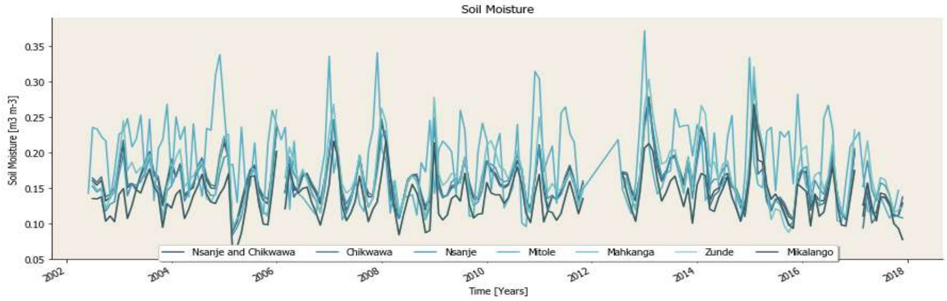


Figure E.3: Soil Moisture time serie

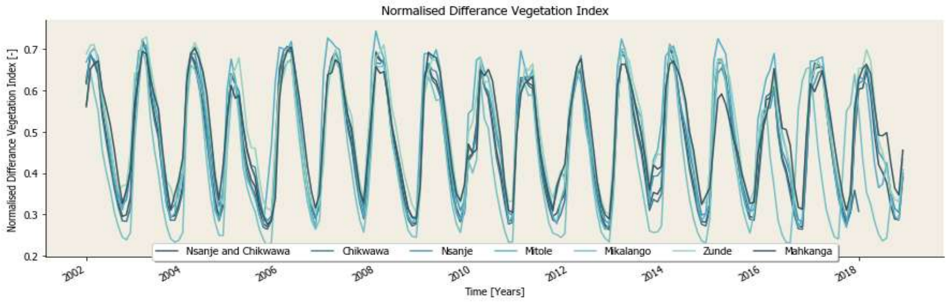


Figure E.4: Normalised Difference Vegetation Index time serie

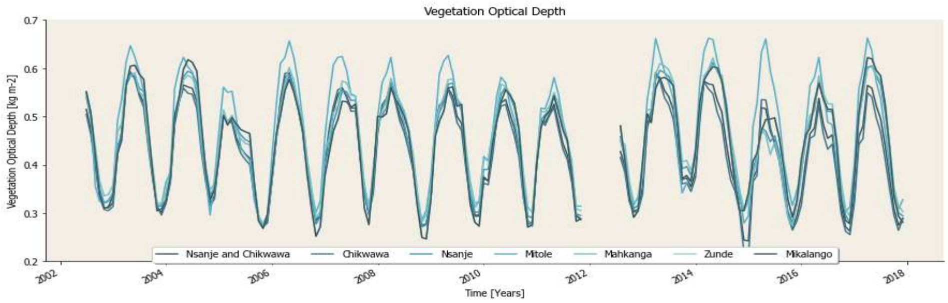


Figure E.5: Vegetation Optical Depth time serie

E.2. Historical time series standardised anomalies

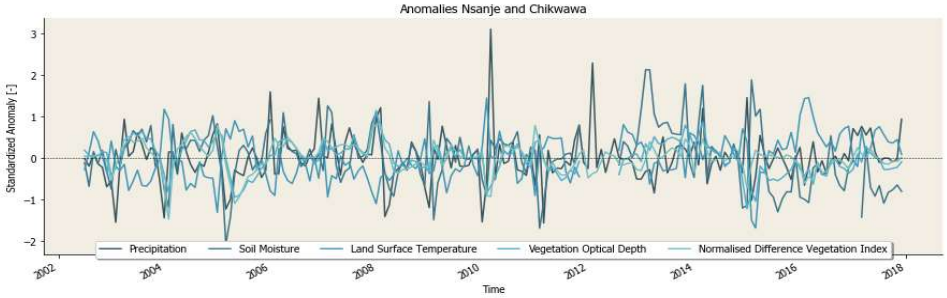


Figure E.6: Anomalies Nsanje and Chikwawa

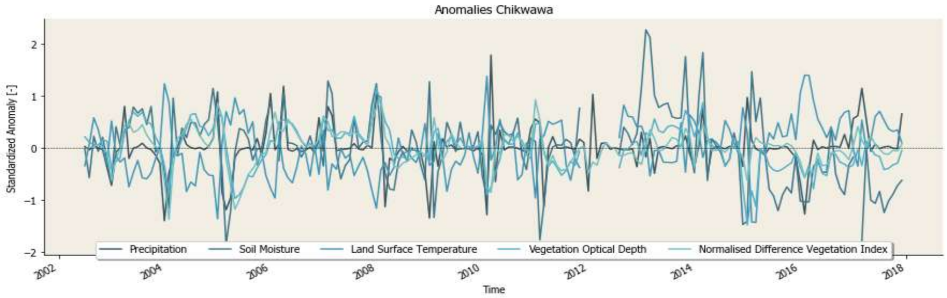


Figure E.7: Anomalies Chikwawa

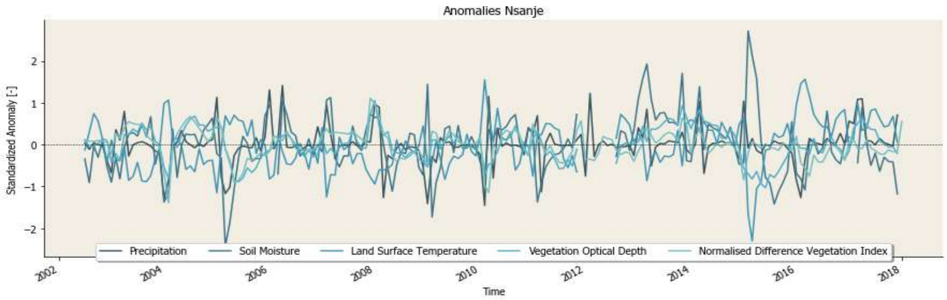


Figure E.8: Anomalies Nsanje

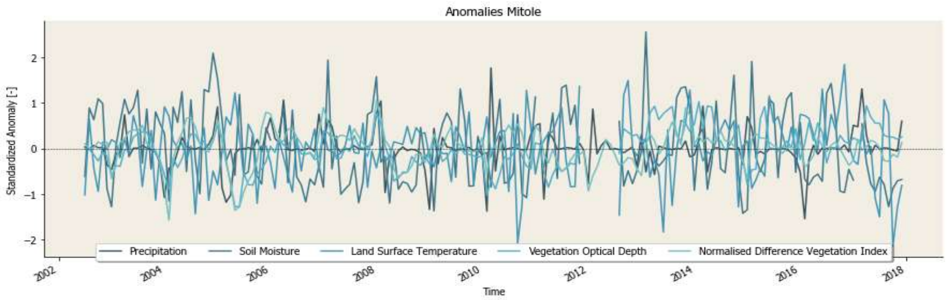


Figure E.9: Anomalies Mitole

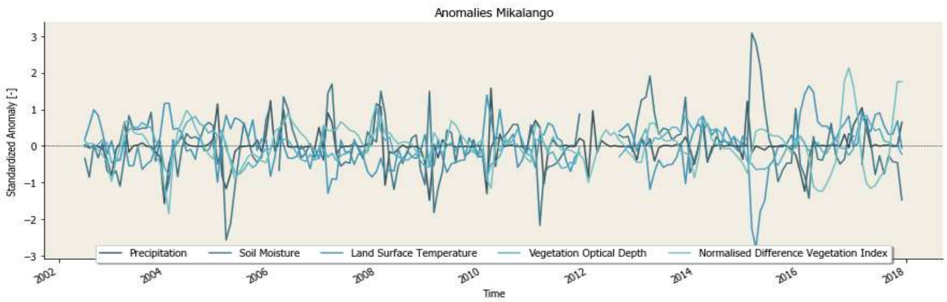


Figure E.10: Anomalies Mikalango

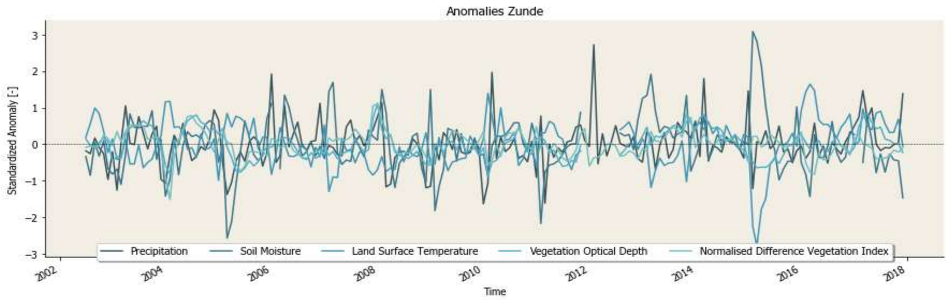


Figure E.11: Anomalies Zunde

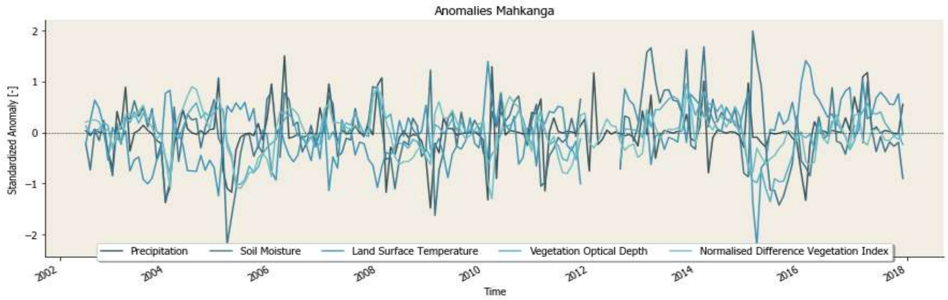
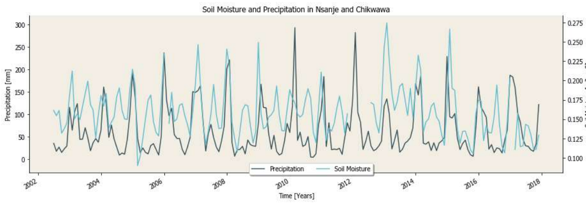
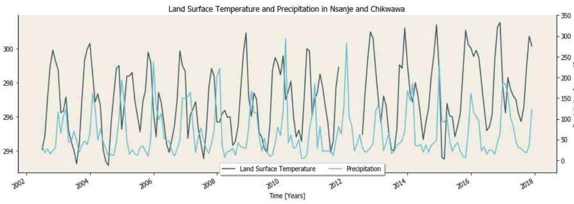
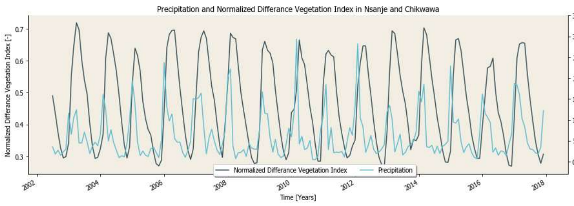


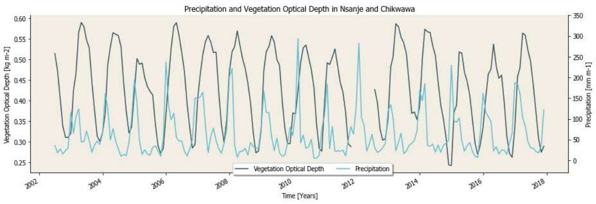
Figure E.12: Anomalies Mahkanga

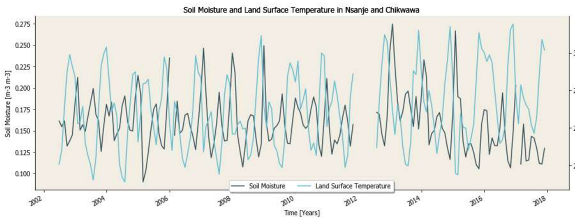
E.3. Nsanje and Chikwawa

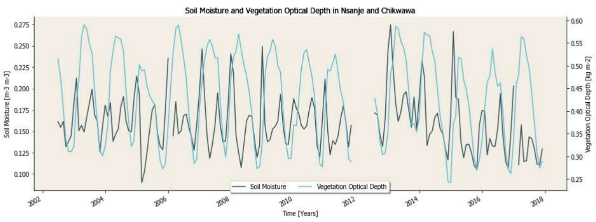


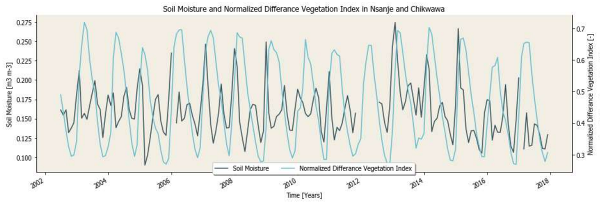


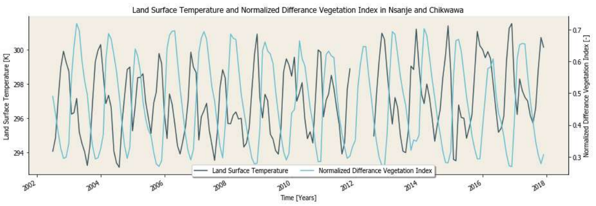


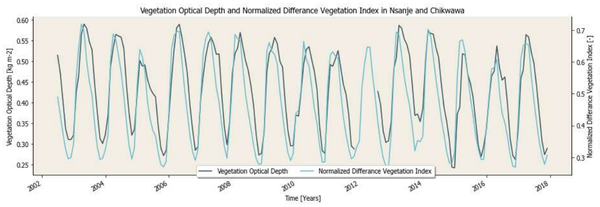


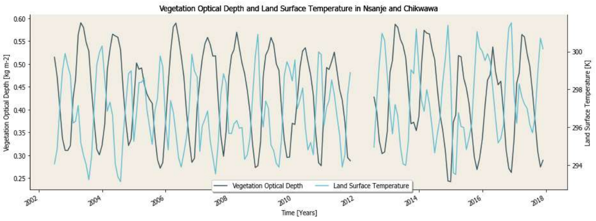




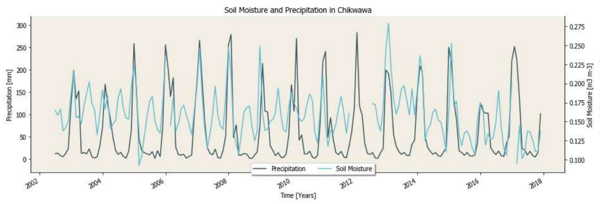

Soil Moisture and Normalized Differance Vegetation Index in Nsanje and Chikwawa
Soil Moisture [m3 m-3]
Normalized Differance Vegetation Index [-]
Soil Moisture
Normalized Differance Vegetation Index
Time [Years]

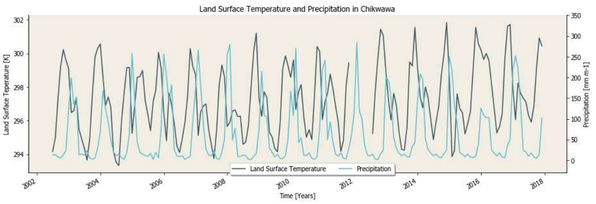

Land Surface Temperature and Normalized Differance Vegetation Index in Nsanje and Chikwawa
Land Surface Temperature [K]
Normalized Differance Vegetation Index [-]
Land Surface Temperature
Normalized Differance Vegetation Index
Time [Years]

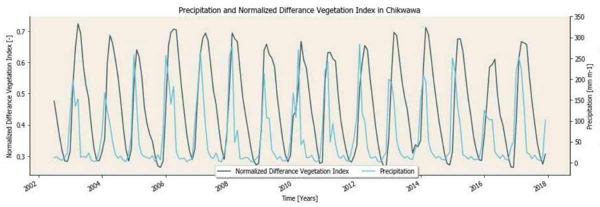

Vegetation Optical Depth and Normalized Differance Vegetation Index in Nsanje and Chikwawa
Vegetation Optical Depth [kg m-2]
Normalized Differance Vegetation Index [-]
Vegetation Optical Depth
Normalized Differance Vegetation Index
Time [Years]

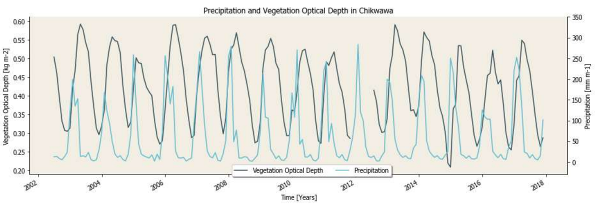

Vegetation Optical Depth and Land Surface Temperature in Nsanje and Chikwawa
Vegetation Optical Depth [kg m-2]
Land Surface Temperature [K]
Vegetation Optical Depth
Land Surface Temperature
Time [Years]

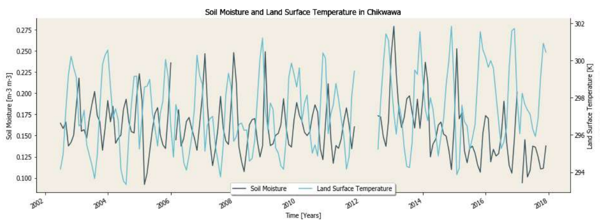
E.4. Chikwawa

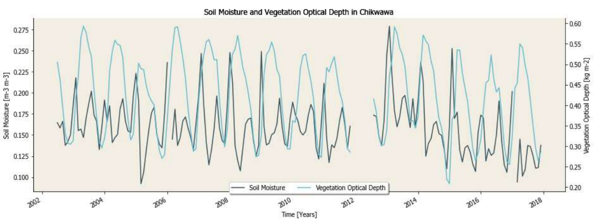

Soil Moisture and Precipitation in Chikwawa
Precipitation [mm]
Soil Moisture [m3 m-3]
Precipitation
Soil Moisture
Time [Years]

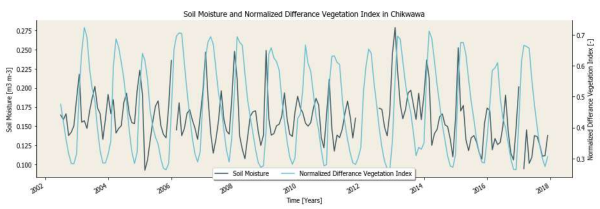

Land Surface Temperature and Precipitation in Chikwawa
Land Surface Temperature [K]
Precipitation [mm m-1]
Land Surface Temperature
Precipitation
Time [Years]

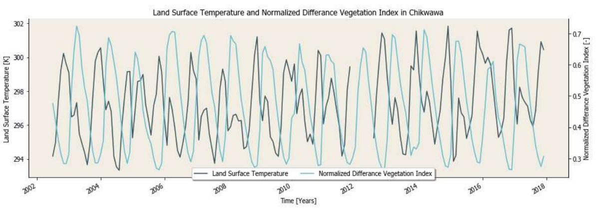

Precipitation and Normalized Differance Vegetation Index in Chikwawa
Normalized Differance Vegetation Index [-]
Precipitation [mm m-1]
Normalized Differance Vegetation Index
Precipitation
Time [Years]

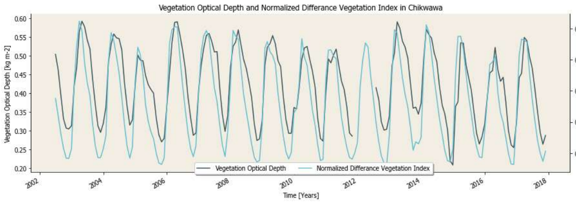

Precipitation and Vegetation Optical Depth in Chikwawa
Vegetation Optical Depth [kg m-2]
Precipitation [mm m-1]
Vegetation Optical Depth
Precipitation
Time [Years]

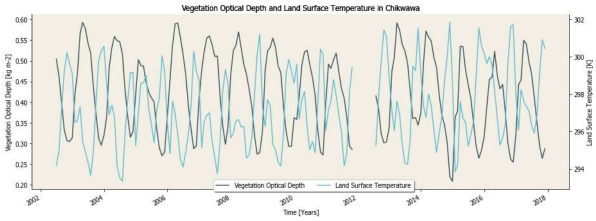

Soil Moisture and Land Surface Temperature in Chikwawa
Soil Moisture [m-3 m-3]
Land Surface Temperature [K]
Soil Moisture
Land Surface Temperature
Time [Years]


Soil Moisture and Vegetation Optical Depth in Chikwawa
Soil Moisture [m-3 m-3]
Vegetation Optical Depth [kg m-2]
Soil Moisture
Vegetation Optical Depth
Time [Years]

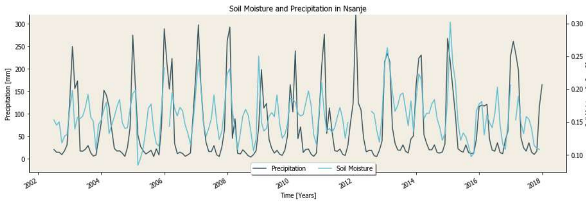

Soil Moisture and Normalized Differance Vegetation Index in Chikwawa
Soil Moisture [m3 m-3]
Normalized Differance Vegetation Index [-]
Soil Moisture
Normalized Differance Vegetation Index
Time [Years]

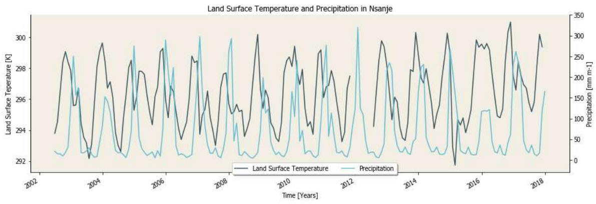

Land Surface Temperature and Normalized Differance Vegetation Index in Chikwawa
Land Surface Temperature [K]
Normalized Differance Vegetation Index [-]
Land Surface Temperature
Normalized Differance Vegetation Index
Time [Years]

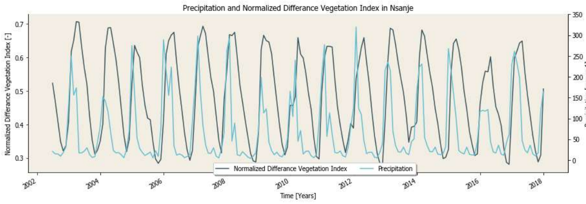

Vegetation Optical Depth and Normalized Differance Vegetation Index in Chikwawa
Vegetation Optical Depth [kg m-2]
Normalized Differance Vegetation Index [-]
Vegetation Optical Depth
Normalized Differance Vegetation Index
Time [Years]

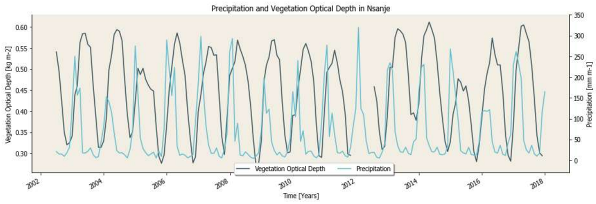

Vegetation Optical Depth and Land Surface Temperature in Chikwawa
Vegetation Optical Depth [kg m-2]
Land Surface Temperature [K]
Vegetation Optical Depth
Land Surface Temperature
Time [Years]

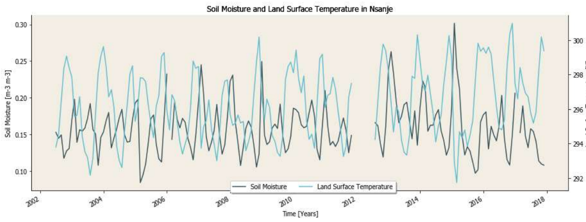
E.5. Nsanje

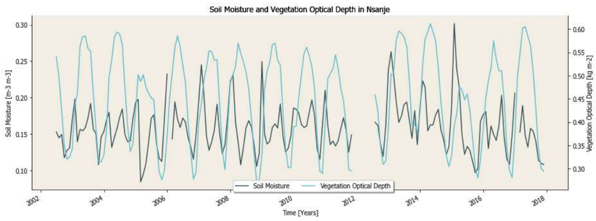

Soil Moisture and Precipitation in Nsanje
Precipitation [mm]
Soil Moisture [m3 m-3]
Precipitation
Soil Moisture
Time [Years]

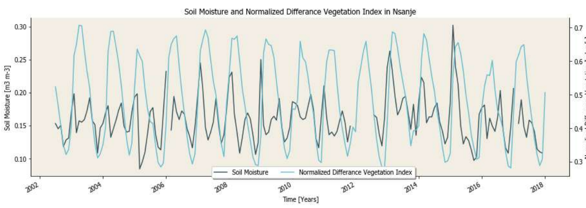

Land Surface Temperature and Precipitation in Nsanje
Land Surface Teperature [K]
Precipitation [mm m-1]
Land Surface Temperature
Precipitation
Time [Years]

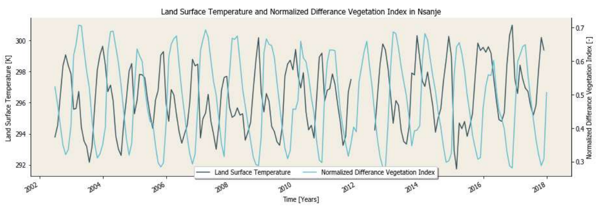

Precipitation and Normalized Differance Vegetation Index in Nsanje
Normalized Differance Vegetation Index [-]
Precipitation [mm m-1]
Normalized Differance Vegetation Index
Precipitation
Time [Years]

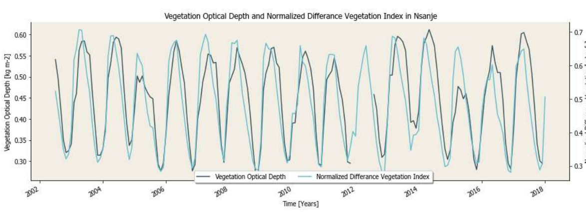

Precipitation and Vegetation Optical Depth in Nsanje
Vegetation Optical Depth [kg m-2]
Precipitation [mm m-1]
Vegetation Optical Depth
Precipitation
Time [Years]

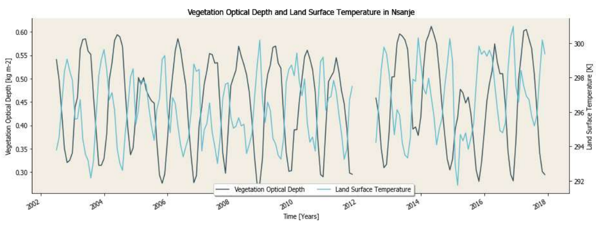

Soil Moisture and Land Surface Temperature in Nsanje
Soil Moisture [m-3 m-3]
Land Surface Temperature [K]
Soil Moisture
Land Surface Temperature
Time [Years]


Soil Moisture and Vegetation Optical Depth in Nsanje
Soil Moisture [m-3 m-3]
Vegetation Optical Depth [kg m-2]
Soil Moisture
Vegetation Optical Depth
Time [Years]

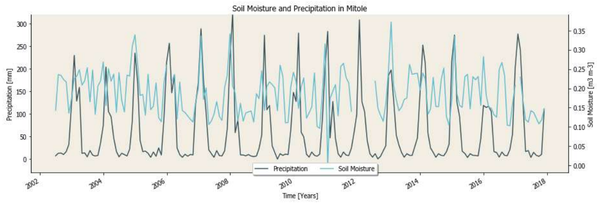

Soil Moisture and Normalized Differance Vegetation Index in Nsanje
Soil Moisture [m-3 m-3]
Normalized Differance Vegetation Index [-]
Soil Moisture
Normalized Differance Vegetation Index
Time [Years]

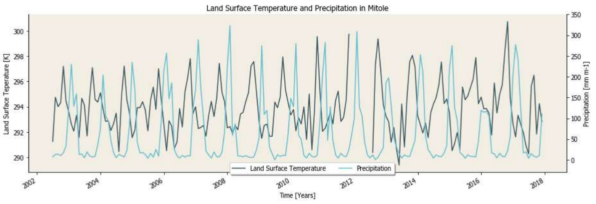

Land Surface Temperature and Normalized Differance Vegetation Index in Nsanje
Land Surface Temperature [K]
Normalized Differance Vegetation Index [-]
Land Surface Temperature
Normalized Differance Vegetation Index
Time [Years]

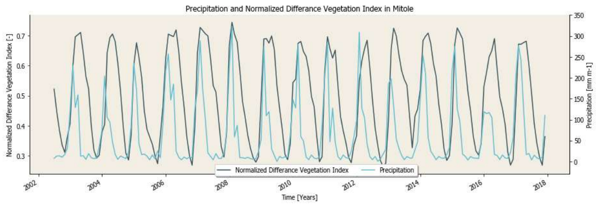

Vegetation Optical Depth and Normalized Differance Vegetation Index in Nsanje
Vegetation Optical Depth [kg m-2]
Normalized Differance Vegetation Index [-]
Vegetation Optical Depth
Normalized Differance Vegetation Index
Time [Years]

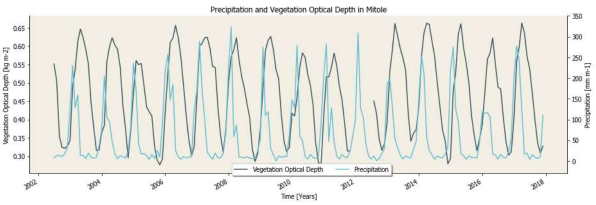

Vegetation Optical Depth and Land Surface Temperature in Nsanje
Vegetation Optical Depth [kg m-2]
Land Surface Temperature [K]
Vegetation Optical Depth
Land Surface Temperature
Time [Years]

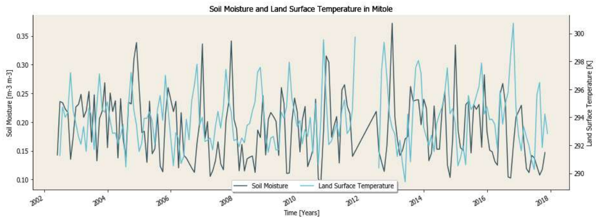
E.6. Mitole

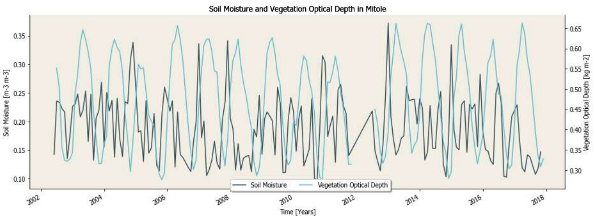

Soil Moisture and Precipitation in Mitole
Precipitation [mm]
Soil Moisture [m3 m-3]
Precipitation
Soil Moisture
Time [Years]

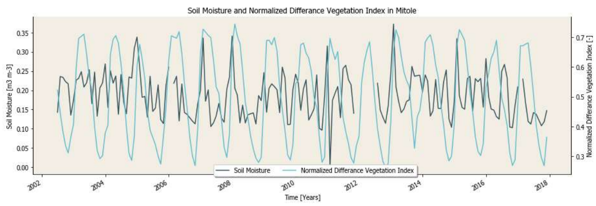

Land Surface Temperature and Precipitation in Mitole
Land Surface Teperature [K]
Precipitation [mm m-1]
Land Surface Temperature
Precipitation
Time [Years]

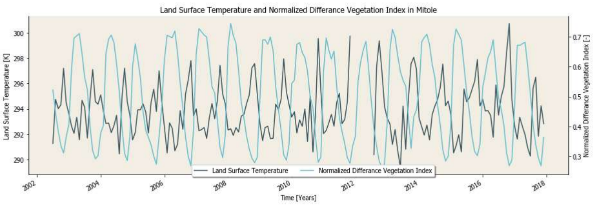

Precipitation and Normalized Differance Vegetation Index in Mitole
Normalized Differance Vegetation Index [-]
Precipitation [mm m-1]
Normalized Differance Vegetation Index
Precipitation
Time [Years]

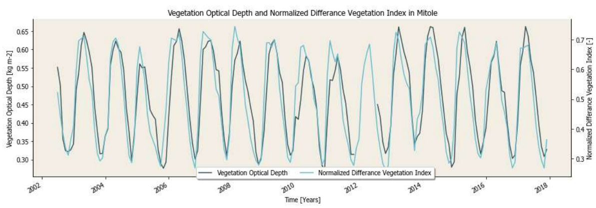

Precipitation and Vegetation Optical Depth in Mitole
Vegetation Optical Depth [kg m-2]
Precipitation [mm m-1]
Vegetation Optical Depth
Precipitation
Time [Years]

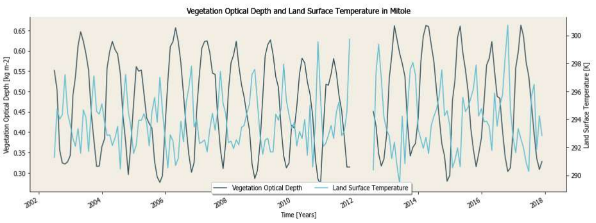

Soil Moisture and Land Surface Temperature in Mitole
Soil Moisture [m-3 m-3]
Land Surface Temperature [K]
Soil Moisture
Land Surface Temperature
Time [Years]


Soil Moisture and Vegetation Optical Depth in Mitole
Soil Moisture [m-3 m-3]
Vegetation Optical Depth [kg m-2]
Soil Moisture
Vegetation Optical Depth
Time [Years]

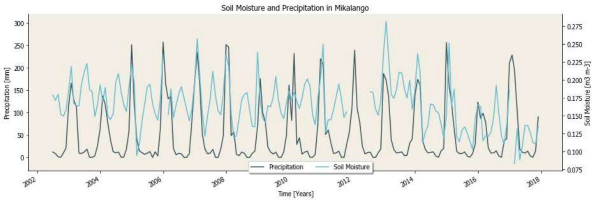

Soil Moisture and Normalized Differance Vegetation Index in Mitole
Soil Moisture [m3 m-3]
Normalized Differance Vegetation Index [-]
Soil Moisture
Normalized Differance Vegetation Index
Time [Years]

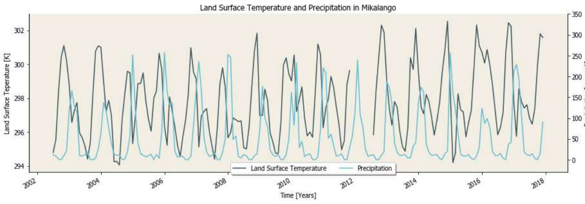

Land Surface Temperature and Normalized Differance Vegetation Index in Mitole
Land Surface Temperature [K]
Normalized Differance Vegetation Index [-]
Land Surface Temperature
Normalized Differance Vegetation Index
Time [Years]

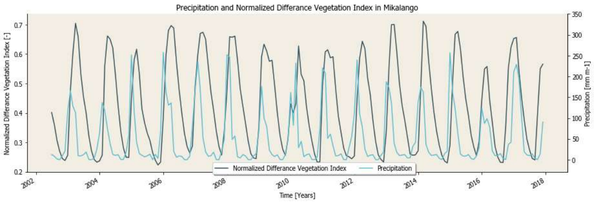

Vegetation Optical Depth and Normalized Differance Vegetation Index in Mitole
Vegetation Optical Depth [kg m-2]
Normalized Differance Vegetation Index [-]
Vegetation Optical Depth
Normalized Differance Vegetation Index
Time [Years]

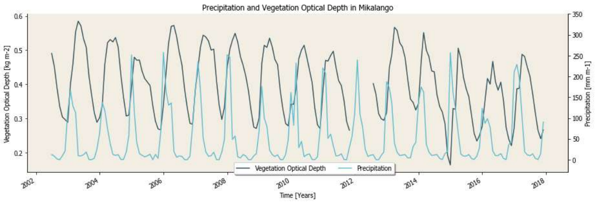

Vegetation Optical Depth and Land Surface Temperature in Mitole
Vegetation Optical Depth [kg m-2]
Land Surface Temperature [K]
Vegetation Optical Depth
Land Surface Temperature
Time [Years]

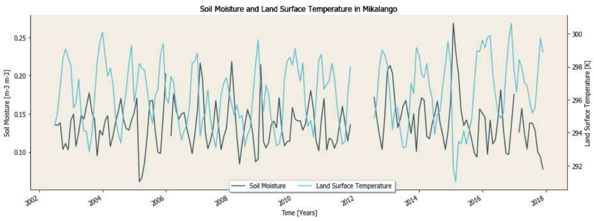
E.7. Mikalango

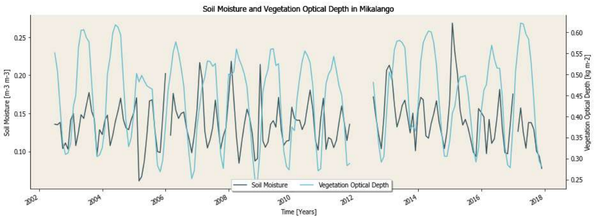


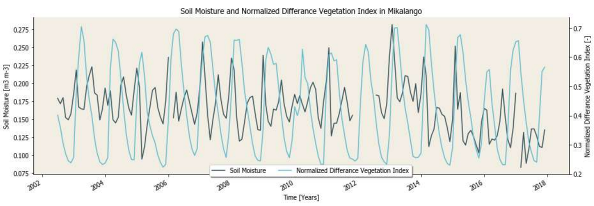


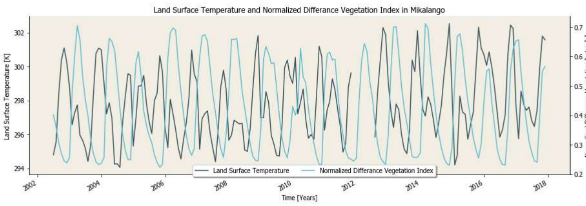


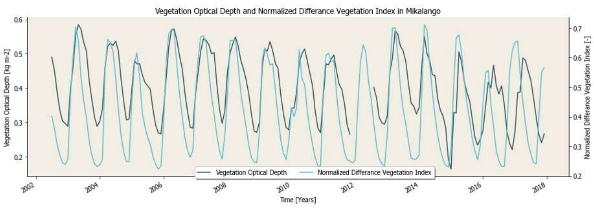


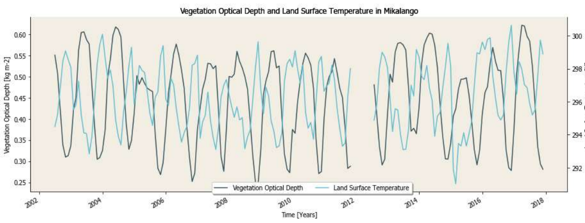




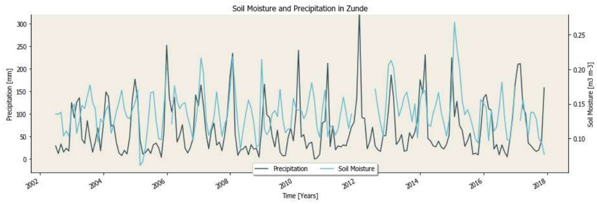




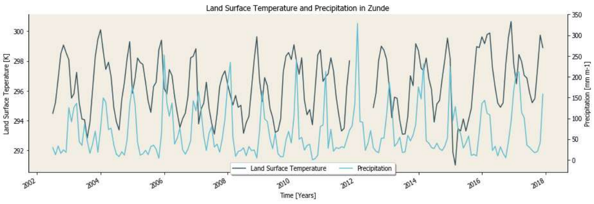




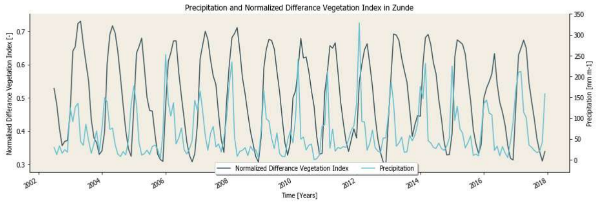
E.8. Zunde



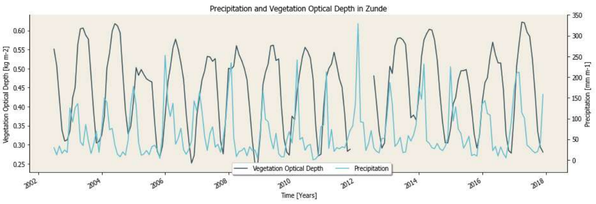
Soil Moisture and Precipitation in Zunde
Precipitation [mm]
Soil Moisture [m3 m-3]
Time [Years]
Precipitation
Soil Moisture



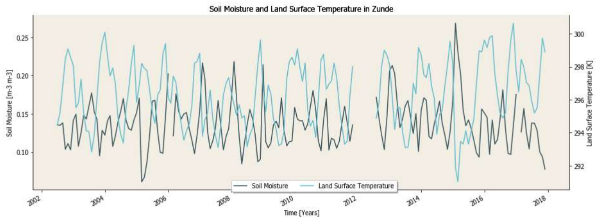
Land Surface Temperature and Precipitation in Zunde
Land Surface Teperature [K]
Precipitation [mm m-1]
Time [Years]
Land Surface Temperature
Precipitation



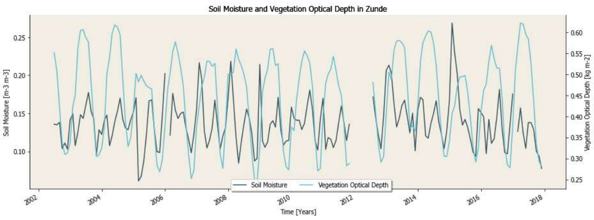
Precipitation and Normalized Differance Vegetation Index in Zunde
Normalized Differance Vegetation Index [-]
Precipitation [mm m-1]
Time [Years]
Normalized Differance Vegetation Index
Precipitation



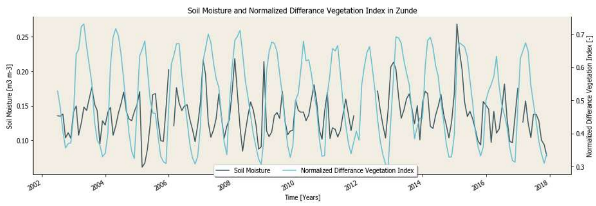
Precipitation and Vegetation Optical Depth in Zunde
Vegetation Optical Depth [kg m-2]
Precipitation [mm m-1]
Time [Years]
Vegetation Optical Depth
Precipitation



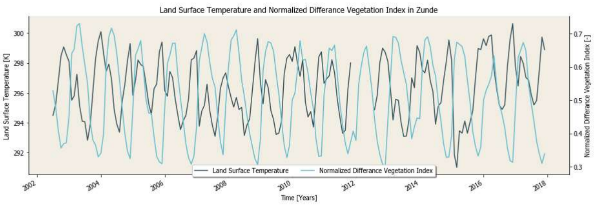
Soil Moisture and Land Surface Temperature in Zunde
Soil Moisture [m-3 m-3]
Land Surface Temperature [K]
Time [Years]
Soil Moisture
Land Surface Temperature



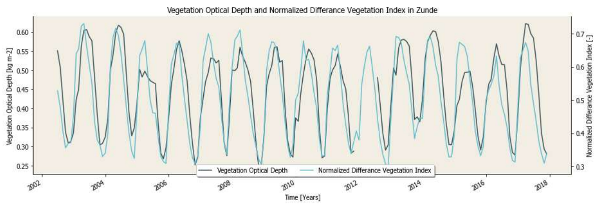
Soil Moisture and Vegetation Optical Depth in Zunde
Soil Moisture [m-3 m-3]
Vegetation Optical Depth [kg m-2]
Time [Years]
Soil Moisture
Vegetation Optical Depth



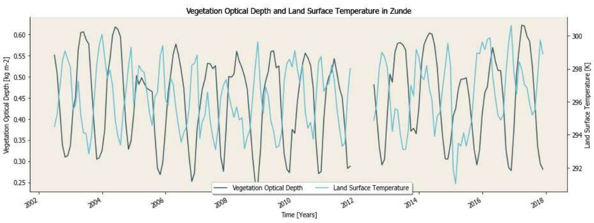
Soil Moisture and Normalized Differance Vegetation Index in Zunde
Soil Moisture [m3 m-3]
Normalized Differance Vegetation Index [-]
Time [Years]
Soil Moisture
Normalized Differance Vegetation Index



Land Surface Temperature and Normalized Differance Vegetation Index in Zunde
Land Surface Temperature [K]
Normalized Differance Vegetation Index [-]
Time [Years]
Land Surface Temperature
Normalized Differance Vegetation Index

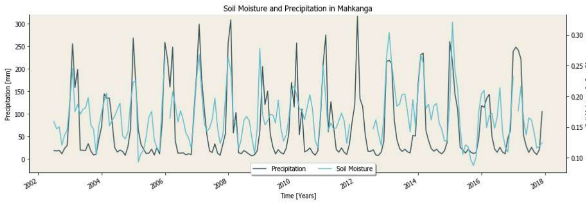


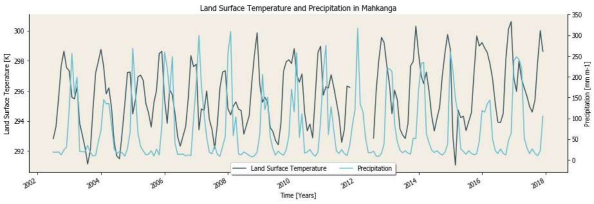
Vegetation Optical Depth and Normalized Differance Vegetation Index in Zunde
Vegetation Optical Depth [kg m-2]
Normalized Differance Vegetation Index [-]
Time [Years]
Vegetation Optical Depth
Normalized Differance Vegetation Index

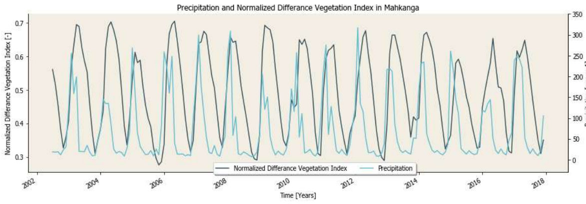


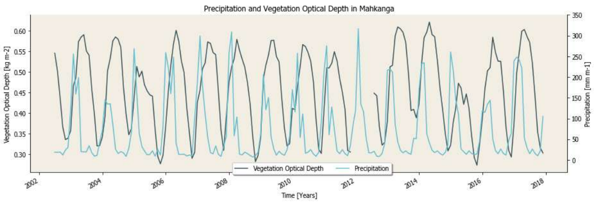
Vegetation Optical Depth and Land Surface Temperature in Zunde
Vegetation Optical Depth [kg m-2]
Land Surface Temperature [K]
Time [Years]
Vegetation Optical Depth
Land Surface Temperature

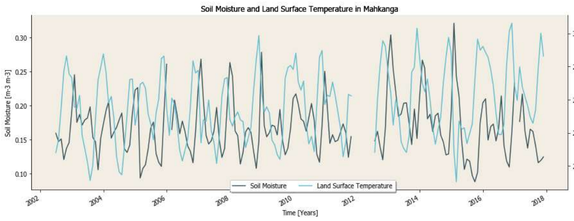
E.9. Mahkanga

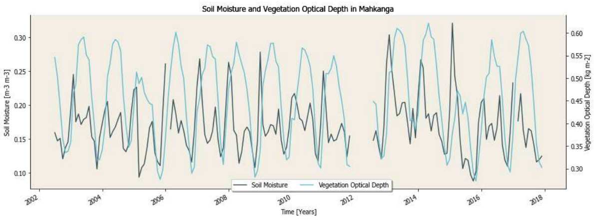

Soil Moisture and Precipitation in Mahkanga
Precipitation [mm]
Soil Moisture [m3 m-3]
Time [Years]
Precipitation
Soil Moisture

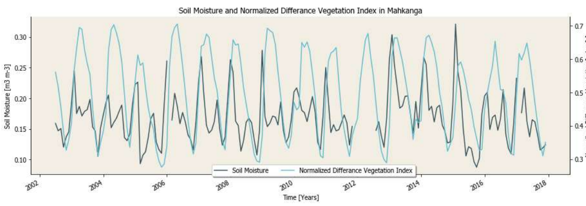

Land Surface Temperature and Precipitation in Mahkanga
Land Surface Teperature [K]
Precipitation [mm m-3]
Time [Years]
Land Surface Temperature
Precipitation

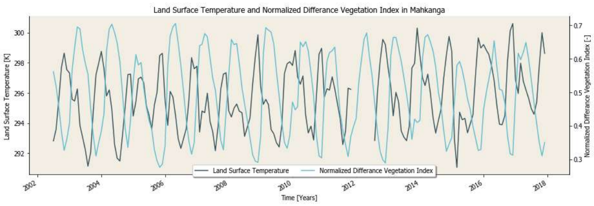

Precipitation and Normalized Differance Vegetation Index in Mahkanga
Normalized Differance Vegetation Index [-]
Precipitation [mm m-3]
Time [Years]
Normalized Differance Vegetation Index
Precipitation

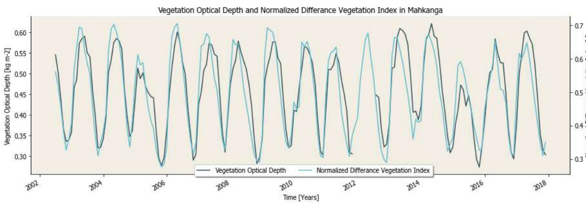

Precipitation and Vegetation Optical Depth in Mahkanga
Vegetation Optical Depth [kg m-2]
Precipitation [mm m-3]
Time [Years]
Vegetation Optical Depth
Precipitation

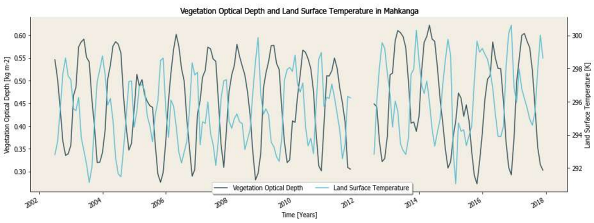

Soil Moisture and Land Surface Temperature in Mahkanga
Soil Moisture [m-3 m-3]
Land Surface Temperature [K]
Time [Years]
Soil Moisture
Land Surface Temperature


Soil Moisture and Vegetation Optical Depth in Mahkanga
Soil Moisture [m-3 m-3]
Vegetation Optical Depth [kg m-2]
Time [Years]
Soil Moisture
Vegetation Optical Depth


Soil Moisture and Normalized Differance Vegetation Index in Mahkanga
Soil Moisture [m3 m-3]
Normalized Differance Vegetation Index [-]
Time [Years]
Soil Moisture
Normalized Differance Vegetation Index


Land Surface Temperature and Normalized Differance Vegetation Index in Mahkanga
Land Surface Temperature [K]
Normalized Differance Vegetation Index [-]
Time [Years]
Land Surface Temperature
Normalized Differance Vegetation Index


Vegetation Optical Depth and Normalized Differance Vegetation Index in Mahkanga
Vegetation Optical Depth [kg m-2]
Normalized Differance Vegetation Index [-]
Time [Years]
Vegetation Optical Depth
Normalized Differance Vegetation Index


Vegetation Optical Depth and Land Surface Temperature in Mahkanga
Vegetation Optical Depth [kg m-2]
Land Surface Temperature [K]
Time [Years]
Vegetation Optical Depth
Land Surface Temperature

F

Pearson Correlation

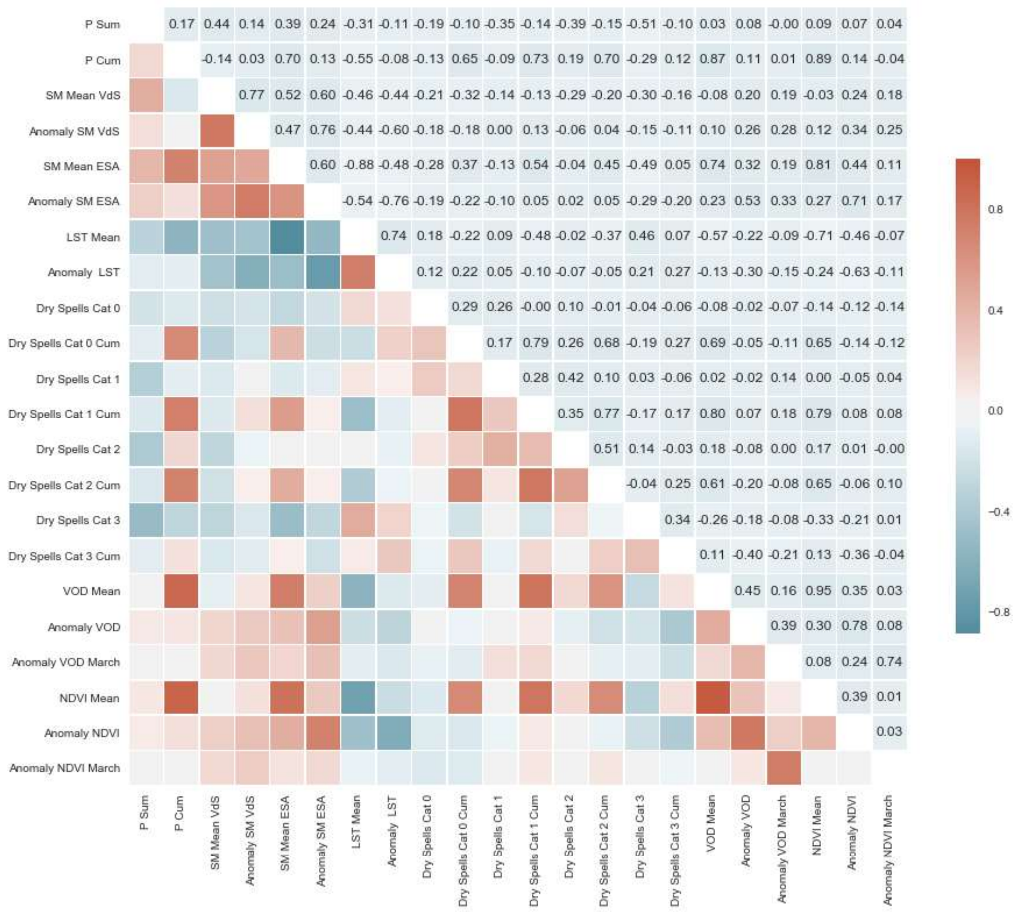


Figure F.1: Pearson correlation Matrix Nsanje and Chikwawa

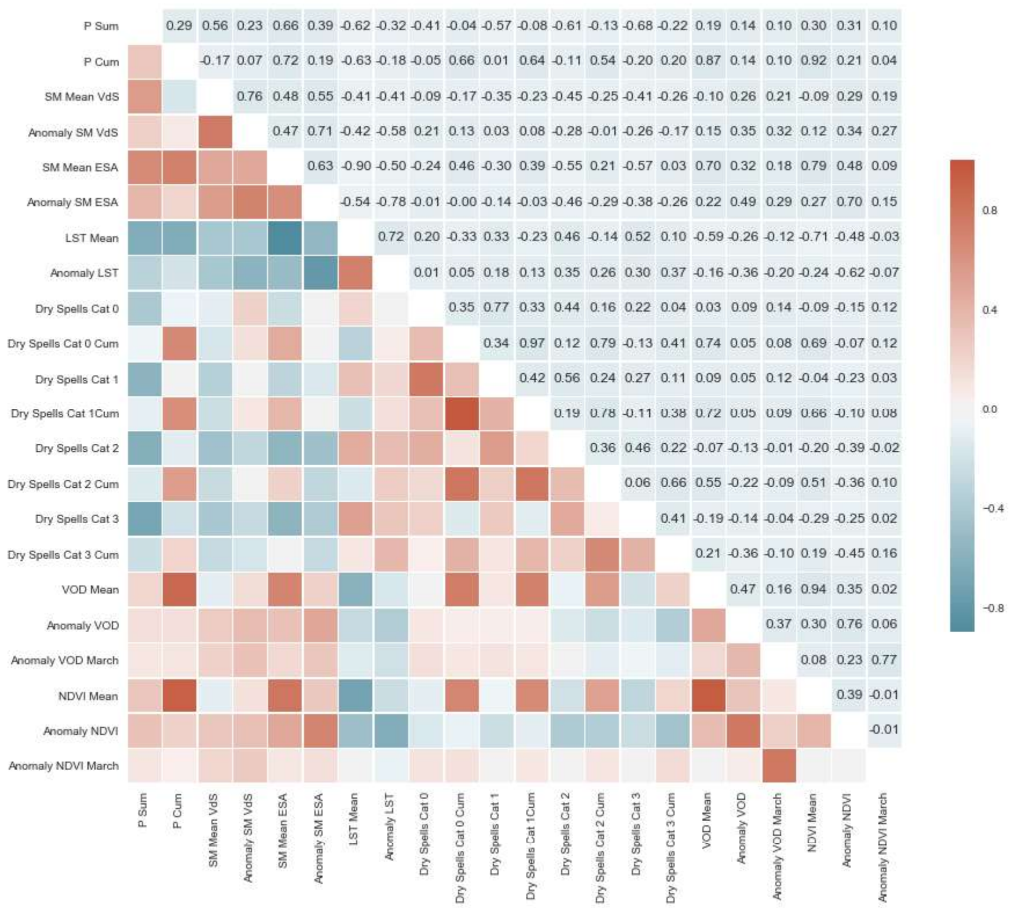


Figure F.2: Pearson correlation Matrix Chikwawa

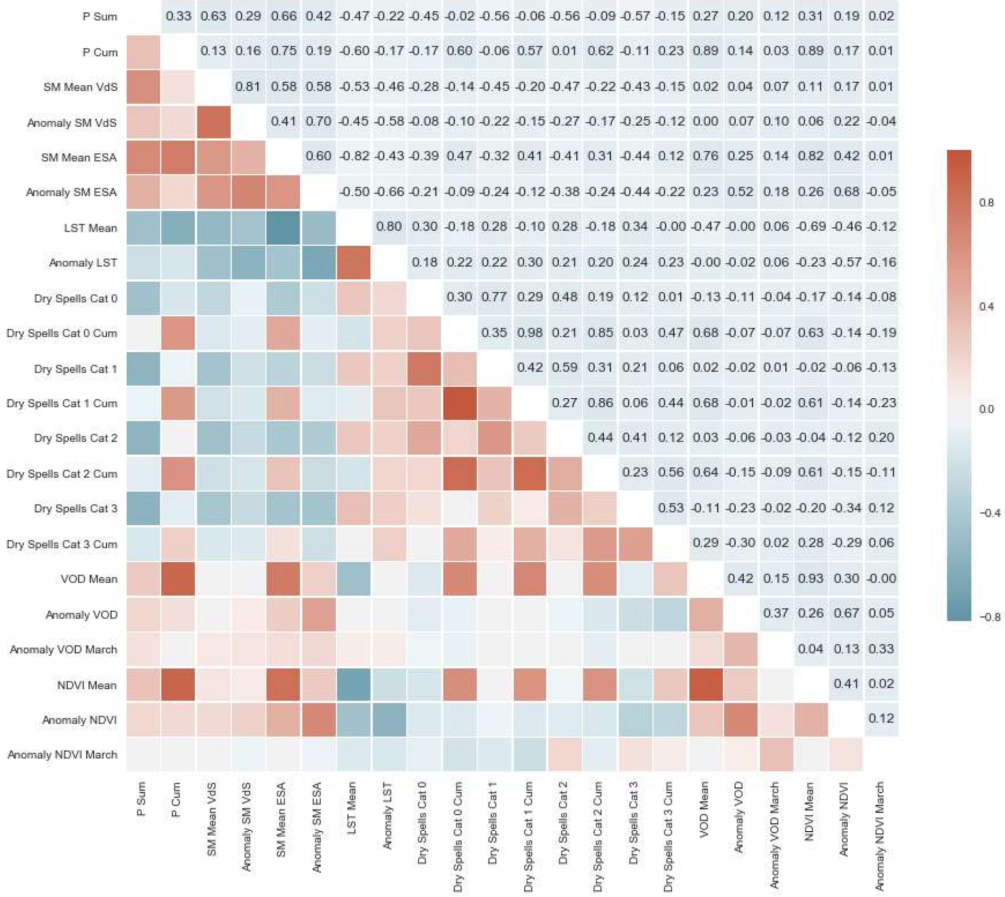


Figure F.3: Pearson correlation Matrix Nsanje

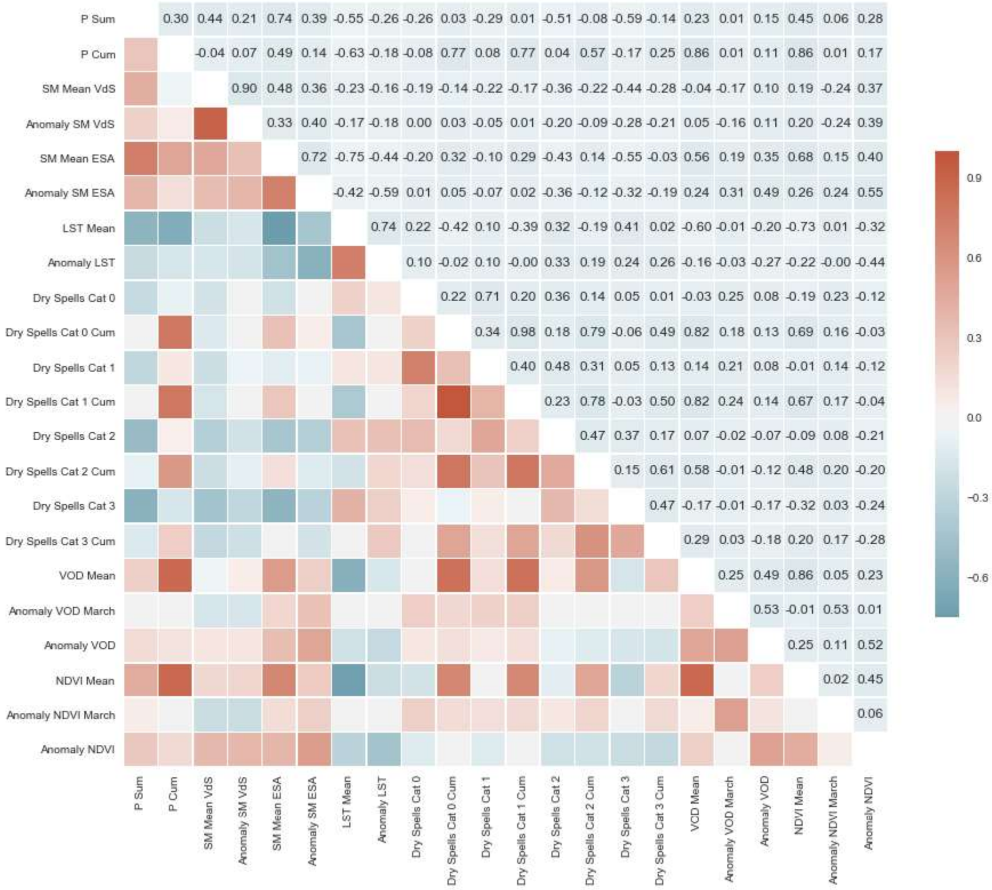


Figure F.4: Pearson correlation Matrix Mitole

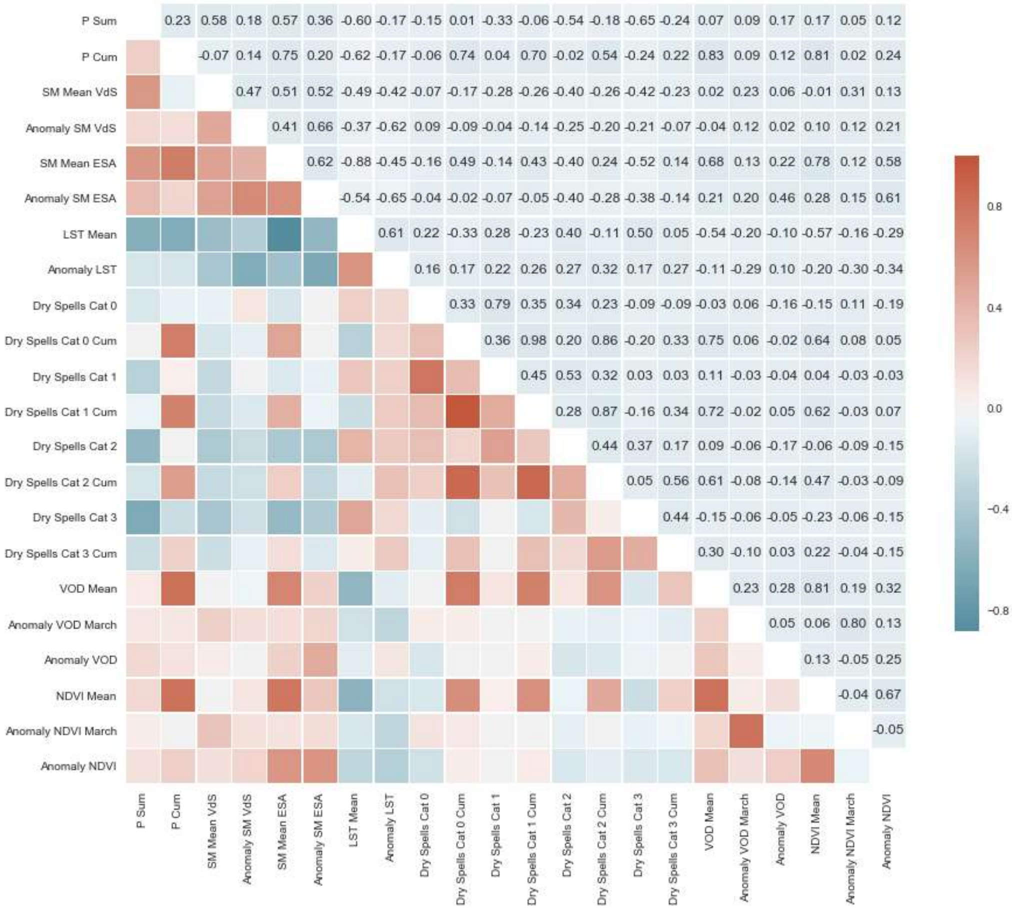


Figure F.5: Pearson correlation Matrix Mikalango

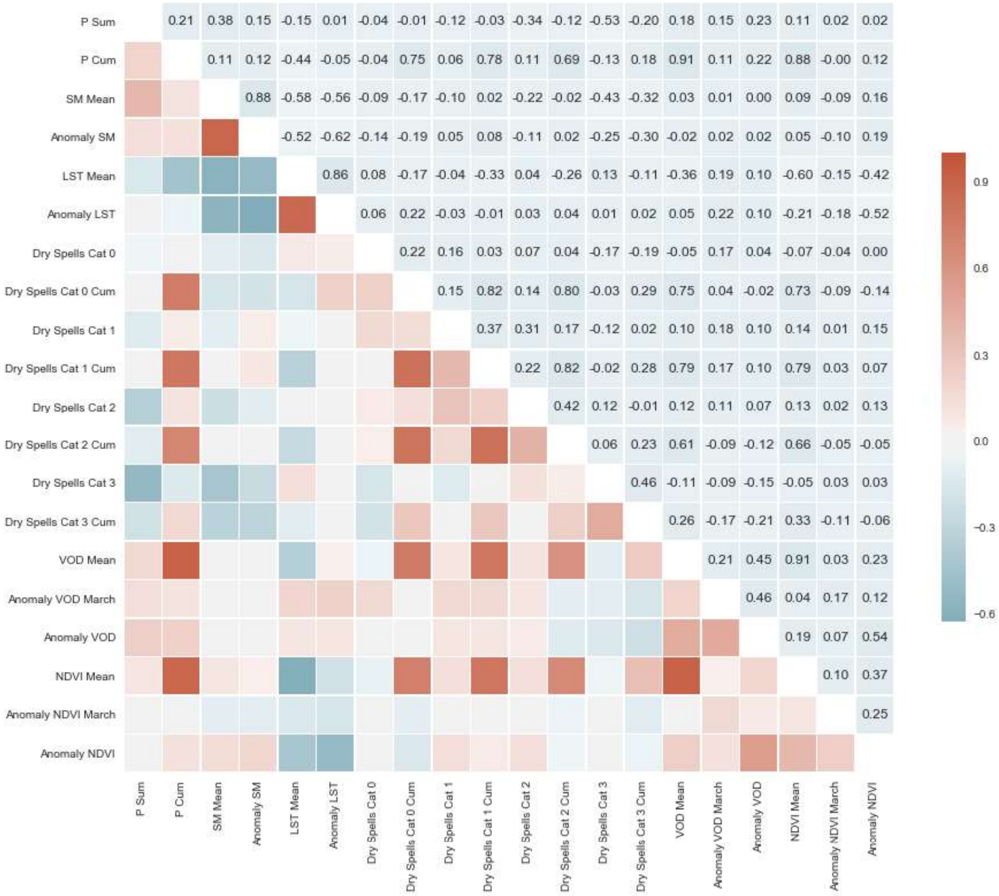


Figure F.6: Pearson correlation Matrix Zunde

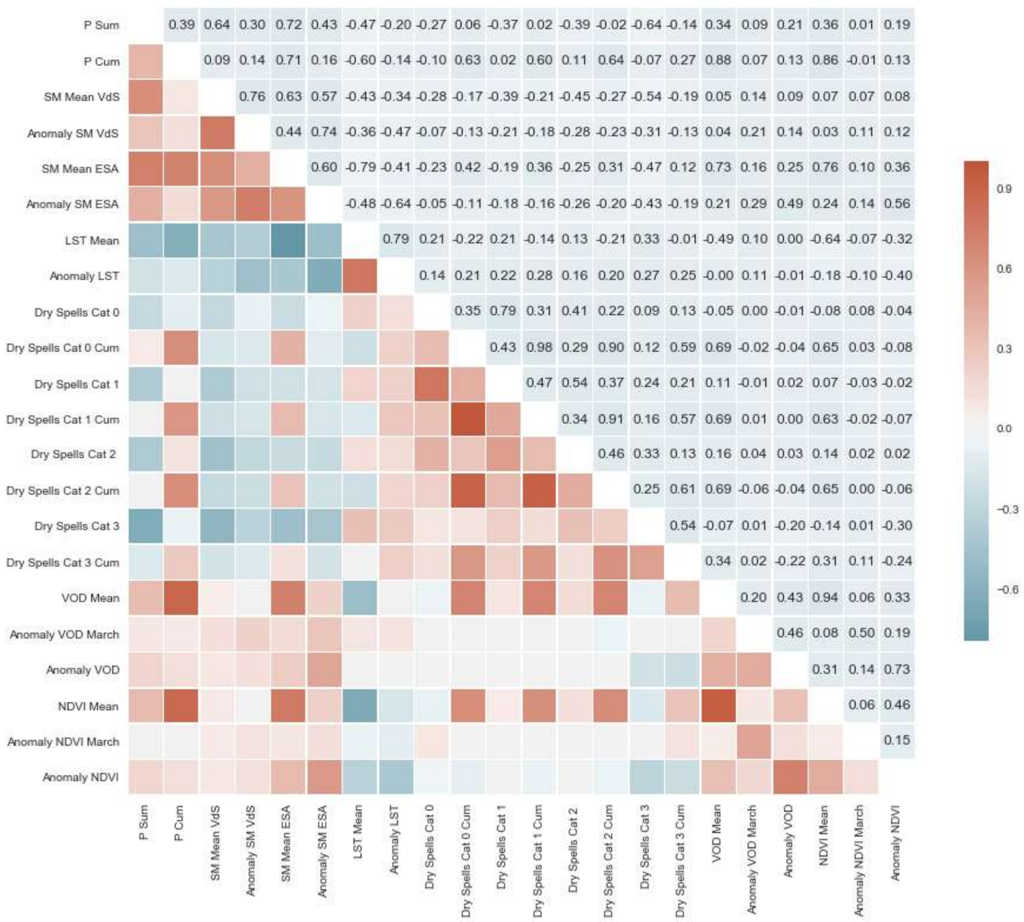


Figure F.7: Pearson correlation Matrix Mahkanga

G

Feature selection

G.1. Top five feature ranking

NDVI 20th Percentile

	NsCh		Nsanje		Chikwawa		Mitole		Mikalango		Zunde		Mahkanga	
	VdS	ESA	VdS	ESA	VdS	ESA	VdS	ESA	VdS	ESA	VdS	ESA	VdS	ESA
	LST	LST	NDVI	NDVI	LST	SM	LST	LST	NDVI	SM	NDVI		NDVI	NDVI
	NDVI	NDVI	LST	LST	NDVI	LST	NDVI	NDVI	LST	NDVI	P		P	SM
Mar	SM	SM	ONI	SM	SM	NDVI	ONI	SM	ONI	LST	SM	-	SM	P
	ONI	ONI	P	ONI	ONI	ONI	SM	ONI	DS	ONI	DS		ONI	ONI
	P	P	SM	P	P	P	P	P	SM	DS	ONI		LST	LST
	SM	SM	ONI	SM	SM	SM	NDVI	SM	LST	SM	P		NDVI	NDVI
	LST	LST	NDVI	ONI	NDVI	NDVI	SM	NDVI	SM	LST	NDVI		P	P
Feb	NDVI	NDVI	SM	NDVI	LST	LST	LST	LST	NDVI	NDVI	ONI	-	SM	SM
	ONI	ONI	DS	DS	P	P	ONI	ONI	ONI	ONI	SM		ONI	ONI
	P	P	LST	LST	ONI	ONI	P	P	DS	DS	DS		DS	DS
	P	P	ONI	ONI	P	P	P	P	SM	NDVI	P		ONI	ONI
	LST	LST	LST	LST	LST	LST	SM	LST	NDVI	P	NDVI		P	P
Jan	NDVI	NDVI	DS	DS	DS	DS	LST	DS	P	SM	LST	-	SM	LST
	DS	SM	P	SM	NDVI	NDVI	DS	NDVI	ONI	ONI	DS		LST	NDVI
	ONI	DS	SM	P	ONI	SM	NDVI	SM	LST	LST	ONI		NDVI	SM
	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI	SM	NDVI	NDVI	NDVI	P		SM	ONI
	P	P	SM	SM	SM	SM	NDVI	SM	DS	DS	SM		ONI	SM
Dec	SM	SM	P	P	P	P	P	P	P	SM	LST	-	P	P
	LST	LST	LST	LST	LST	LST	DS	DS	SM	P	NDVI		NDVI	NDVI
	DS	DS	ONI	ONI	DS	DS	LST	LST	ONI	ONI	DS		LST	LST
	P	SM	SM	SM	P	SM	P	SM	SM	SM	P		ONI	ONI
	SM	P	P	P	LST	P	NDVI	P	DS	DS	SM		NDVI	NDVI
Nov	LST	LST	LST	LST	SM	LST	DS	NDVI	NDVI	NDVI	LST	-	P	P
	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI	SM	DS	LST	LST	NDVI		LST	LST
	DS	DS	ONI	ONI	DS	DS	LST	LST	P	P	DS		SM	SM

NDVI 40th Percentile

	NsCh		Nsanje		Chikwawa		Mitole		Mikalango		Zunde		Mahkanga	
	VdS	ESA	VdS	ESA	VdS	ESA	VdS	ESA	VdS	ESA	VdS	ESA	VdS	ESA
	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI		NDVI	NDVI
	LST	SM	ONI	SM	SM	ONI	SM	LST	LST	SM	SM		ONI	SM
Mar	P	LST	P	ONI	ONI	SM	LST	ONI	SM	LST	DS	-	P	ONI
	SM	P	LST	P	DS	DS	ONI	SM	ONI	ONI	ONI		SM	P
	DS	DS	SM	LST	LST	LST	DS	DS	DS	DS	LST		LST	LST
	LST	SM	ONI	SM	SM	SM	NDVI	SM	SM	SM	NDVI		NDVI	NDVI
	NDVI	LST	SM	ONI	NDVI	NDVI	ONI	NDVI	NDVI	NDVI	SM		P	P
Feb	SM	NDVI	LST	LST	LST	LST	SM	ONI	LST	LST	ONI	-	SM	SM
	DS	DS	NDVI	NDVI	ONI	ONI	LST	LST	ONI	ONI	DS		ONI	ONI
	ONI	ONI	DS	DS	DS	DS	DS	DS	DS	DS	LST		DS	DS
	P	P	ONI	ONI	NDVI	SM	P	P	NDVI	NDVI	P		NDVI	NDVI
	DS	SM	LST	LST	ONI	NDVI	SM	DS	ONI	SM	NDVI		P	P
Jan	NDVI	DS	P	SM	P	ONI	DS	NDVI	P	ONI	ONI	-	SM	SM
	ONI	NDVI	SM	P	SM	P	NDVI	LST	SM	P	SM		ONI	ONI
	SM	ONI	NDVI	NDVI	LST	LST	LST	ONI	LST	LST	LST		LST	LST
	NDVI	NDVI	SM	P	NDVI	NDVI	SM	SM	NDVI	NDVI	ONI		P	P
	LST	SM	P	SM	LST	SM	NDVI	NDVI	SM	SM	SM		NDVI	NDVI
Dec	P	LST	LST	LST	SM	LST	LST	LST	P	P	LST	-	SM	SM
	SM	P	NDVI	NDVI	P	P	ONI	ONI	LST	LST	NDVI		DS	DS
	DS	DS	DS	DS	DS	DS	P	P	DS	DS	P		LST	LST
	LST	SM	LST	LST	P	SM	NDVI	SM	DS	DS	LST		NDVI	NDVI
	DS	LST	SM	NDVI	SM	P	P	NDVI	NDVI	NDVI	NDVI		P	P
Nov	NDVI	DS	NDVI	P	LST	LST	SM	P	LST	LST	P	-	SM	SM
	P	NDVI	P	DS	DS	DS	LST	LST	P	SM	ONI		LST	LST
	SM	P	DS	ONI	NDVI	NDVI	ONI	ONI	SM	P	DS		ONI	ONI

NDVI 50th Percentile

	NsCh		Nsanje		Chikwawa		Mitole		Mikalango		Zunde		Mahkanga	
	VdS	ESA	VdS	ESA	VdS	ESA	VdS	ESA	VdS	ESA	VdS	ESA	VdS	ESA
	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI		NDVI	NDVI
	P	P	P	SM	SM	ONI	SM	LST	SM	LST	SM		ONI	ONI
Mar	SM	DS	LST	P	ONI	DS	LST	ONI	LST	ONI	DS		P	SM
	DS	ONI	SM	LST	DS	SM	ONI	DS	ONI	SM	LST		SM	P
	ONI	SM	ONI	ONI	LST	LST	DS	SM	DS	DS	ONI		LST	LST
	SM	SM	ONI	SM	SM	SM	NDVI	SM	NDVI	NDVI	NDVI		NDVI	NDVI
	NDVI	NDVI	SM	ONI	NDVI	NDVI	DS	NDVI	SM	SM	SM		ONI	ONI
Feb	LST	LST	LST	LST	LST	LST	ONI	DS	LST	LST	DS		P	P
	DS	DS	DS	DS	ONI	ONI	SM	ONI	ONI	ONI	LST		SM	SM
	ONI	ONI	NDVI	NDVI	DS	DS	LST	LST	DS	DS	P		LST	LST
	P	P	P	P	NDVI	SM	SM	LST	SM	SM	P		SM	SM
	SM	DS	NDVI	SM	SM	NDVI	LST	DS	P	P	NDVI		NDVI	NDVI
Jan	DS	SM	ONI	NDVI	LST	LST	DS	NDVI	NDVI	NDVI	SM		P	P
	NDVI	NDVI	LST	ONI	ONI	ONI	NDVI	ONI	LST	LST	LST		ONI	ONI
	LST	LST	SM	LST	P	P	ONI	SM	ONI	ONI	DS		LST	LST
	NDVI	NDVI	SM	P	NDVI	NDVI	SM	SM	NDVI	NDVI	SM		P	P
	LST	SM	P	SM	LST	SM	NDVI	NDVI	SM	SM	LST		NDVI	NDVI
Dec	P	LST	LST	LST	P	LST	LST	LST	P	P	NDVI		SM	SM
	SM	P	NDVI	NDVI	SM	P	DS	DS	LST	LST	ONI		ONI	ONI
	DS	DS	ONI	ONI	DS	DS	P	P	DS	DS	P		LST	LST
	P	SM	LST	LST	P	SM	NDVI	SM	NDVI	NDVI	P		NDVI	NDVI
	SM	P	SM	NDVI	SM	P	P	NDVI	P	SM	LST		P	P
Nov	NDVI	NDVI	NDVI	P	LST	LST	SM	P	DS	P	NDVI		SM	SM
	LST	LST	P	DS	DS	DS	LST	LST	SM	DS	DS		LST	LST
	DS	DS	DS	ONI	NDVI	NDVI	DS	DS	LST	LST	ONI		ONI	ONI

VOD 20th Percentile

	NsCh		Nsanje		Chikwawa		Mitole		Mikalango		Zunde		Mahkanga	
	VdS	ESA	VdS	ESA	VdS	ESA	VdS	ESA	VdS	ESA	VdS	ESA	VdS	ESA
	VOD	VOD	VOD	VOD	VOD	VOD	VOD	VOD	P	SM	VOD		VOD	VOD
	LST	LST	P	SM	LST	LST	DS	DS	DS	P	SM		P	SM
Mar	ONI	SM	LST	P	P	SM	LST	LST	VOD	DS	LST	-	SM	P
	SM	ONI	DS	LST	ONI	P	ONI	SM	ONI	VOD	P		ONI	ONI
	P	P	SM	DS	SM	ONI	SM	ONI	LST	ONI	DS		LST	LST
	LST	SM	DS	DS	ONI	SM	SM	VOD	LST	LST	VOD		VOD	VOD
	ONI	LST	P	P	VOD	ONI	VOD	DS	VOD	SM	SM		P	P
Feb	VOD	ONI	LST	LST	P	VOD	DS	ONI	ONI	VOD	P	-	SM	SM
	P	VOD	VOD	VOD	SM	P	ONI	P	P	ONI	DS		ONI	ONI
	SM	P	SM	SM	LST	LST	P	LST	SM	P	LST		DS	DS
	DS	DS	SM	SM	SM	DS	ONI	SM	P	P	ONI		ONI	ONI
	SM	P	ONI	ONI	DS	P	DS	ONI	SM	SM	SM		P	P
Jan	P	VOD	LST	LST	P	SM	SM	DS	VOD	VOD	DS	-	SM	LST
	VOD	SM	VOD	VOD	LST	LST	VOD	VOD	LST	LST	P		LST	SM
	ONI	ONI	P	P	VOD	VOD	P	P	ONI	ONI	LST		DS	DS
	ONI	ONI	SM	ONI	ONI	ONI	DS	DS	ONI	ONI	LST		SM	ONI
	VOD	VOD	ONI	SM	SM	SM	P	SM	SM	DS	VOD		ONI	SM
Dec	SM	SM	P	P	DS	DS	SM	P	DS	SM	ONI	-	P	P
	P	P	LST	LST	P	P	LST	LST	VOD	VOD	SM		VOD	VOD
	DS	DS	DS	DS	VOD	VOD	VOD	VOD	P	P	P		LST	LST
	ONI	ONI	ONI	ONI	ONI	ONI	P	P	ONI	ONI	SM		ONI	ONI
	P	SM	VOD	VOD	P	SM	SM	SM	P	SM	ONI		VOD	VOD
Nov	SM	P	P	P	LST	P	VOD	VOD	LST	P	VOD	-	P	P
	VOD	VOD	LST	LST	VOD	LT	DS	DS	DS	LST	P		LST	LST
	LST	LST	DS	DS	SM	VOD	ONI	ONI	SM	DS	LST		SM	SM

VOD 40th Percentile

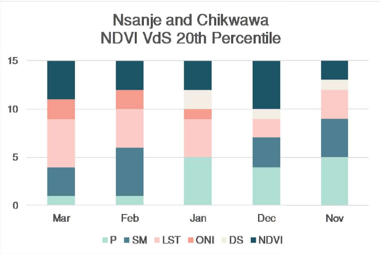
	NsCh		Nsanje		Chikwawa		Mitole		Mikalango		Zunde		Mahkanga	
	VdS	ESA	VdS	ESA	VdS	ESA	VdS	ESA	VdS	ESA	VdS	ESA	VdS	ESA
	VOD	VOD	VOD	VOD	VOD	SM	VOD	VOD	LST	LST	VOD		VOD	VOD
	P	SM	P	SM	P	VOD	SM	LST	ONI	ONI	DS		P	SM
Mar	DS	P	ONI	P	SM	P	LST	P	SM	SM	P	-	SM	P
	LST	DS	DS	ONI	LST	LST	P	SM	VOD	VOD	LST		ONI	ONI
	SM	LST	SM	DS	ONI	ONI	ONI	ONI	P	P	SM		LST	LST
	VOD	VOD	VOD	VOD	P	SM	LST	LST	LST	LST	VOD		VOD	VOD
	P	P	ONI	SM	LST	P	VOD	SM	ONI	ONI	P		P	SM
Feb	SM	DS	P	ONI	SM	LST	ONI	VOD	SM	SM	DS	-	SM	P
	DS	SM	SM	P	VOD	VOD	P	ONI	DS	DS	LST		ONI	ONI
	LST	LST	DS	DS	DS	DS	SM	P	VOD	VOD	ONI		LST	LST
	P	P	ONI	ONI	DS	DS	VOD	SM	VOD	SM	DS		P	P
	DS	DS	VOD	VOD	P	SM	SM	VOD	P	VOD	ONI		SM	SM
Jan	SM	SM	LST	SM	SM	P	P	P	LST	P	SM	-	LST	LST
	VOD	VOD	P	LST	LST	LST	DS	DS	ONI	LST	P		DS	DS
	LST	LST	SM	P	VOD	VOD	ONI	ONI	SM	ONI	LST		ONI	ONI
	SM	SM	SM	P	VOD	VOD	SM	DS	VOD	VOD	SM		SM	P
	ONI	ONI	P	SM	P	SM	DS	ONI	P	P	DS		P	VOD
Dec	P	P	ONI	ONI	SM	P	ONI	VOD	SM	SM	ONI	-	VOD	SM
	DS	DS	LST	LST	DS	DS	VOD	SM	DS	DS	LST		ONI	ONI
	VOD	VOD	VOD	VOD	ONI	ONI	P	P	LST	LST	VOD		LST	LST
	P	SM	P	SM	LST	LST	SM	VOD	P	SM	SM		LST	LST
	LST	P	SM	P	VOD	VOD	VOD	DS	DS	P	ONI		VOD	VOD
Nov	DS	LST	LST	LST	P	SM	DS	ONI	LST	DS	DS	-	P	SM
	VOD	DS	VOD	VOD	SM	P	ONI	SM	VOD	LST	VOD		SM	P
	ONI	VOD	ONI	ONI	DS	DS	LST	LST	SM	VOD	LST		ONI	ONI

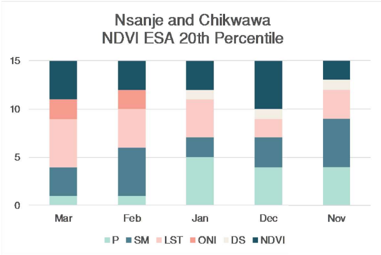
VOD 50th Percentile

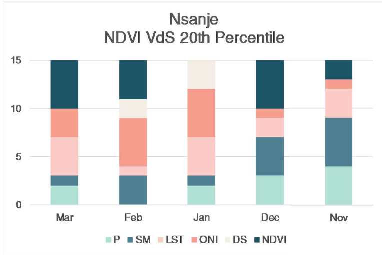
	NsCh		Nsanje		Chikwawa		Mitole		Mikalango		Zunde		Mahkanga	
	VdS	ESA	VdS	ESA	VdS	ESA	VdS	ESA	VdS	ESA	VdS	ESA	VdS	ESA
	VOD	VOD	VOD	VOD	VOD	VOD	VOD	VOD	LST	SM	VOD		VOD	VOD
	P	SM	P	SM	P	SM	LST	LST	ONI	LST	P		P	SM
Mar	DS	P	DS	P	LST	P	P	SM	P	ONI	DS		SM	P
	LST	DS	SM	DS	SM	LST	SM	P	VOD	P	SM		DS	DS
	SM	LST	ONI	ONI	DS	DS	DS	DS	SM	VOD	ONI		ONI	ONI
	P	P	VOD	SM	P	P	VOD	SM	LST	LST	VOD		VOD	VOD
	SM	DS	ONI	VOD	SM	VOD	SM	VOD	ONI	SM	P		P	SM
Feb	DS	VOD	P	ONI	VOD	SM	LST	LST	P	ONI	DS		SM	P
	VOD	SM	SM	P	DS	DS	ONI	ONI	SM	P	ONI		ONI	ONI
	LST	LST	LST	LST	LST	LST	P	P	DS	DS	SM		LST	LST
	P	P	ONI	ONI	DS	DS	DS	DS	ONI	SM	DS		P	P
	DS	DS	LST	LST	SM	SM	VOD	VOD	LST	ONI	ONI		SM	LST
Jan	LST	SM	VOD	VOD	P	P	ONI	SM	P	LST	SM		LST	DS
	SM	LST	P	P	LST	LST	P	ONI	VOD	P	P		DS	ONI
	VOD	VOD	SM	SM	VOD	VOD	SM	P	SM	VOD	LST		ONI	SM
	P	SM	SM	SM	P	SM	SM	DS	P	P	SM		SM	ONI
	DS	P	P	P	DS	P	DS	ONI	SM	VOD	VOD		ONI	SM
Dec	VOD	DS	LST	LST	VOD	DS	ONI	VOD	VOD	SM	P		P	P
	SM	VOD	VOD	VOD	SM	VOD	VOD	SM	LST	LST	LST		VOD	VOD
	ONI	ONI	ONI	ONI	ONI	ONI	P	P	DS	DS	ONI		LST	LST
	LST	LST	SM	LST	LST	LST	SM	DS	LST	SM	DS		LST	LST
	DS	DS	LST	VOD	VOD	VOD	DS	VOD	P	LST	VOD		VOD	VOD
Nov	VOD	VOD	VOD	ONI	P	SM	VOD	ONI	VOD	P	LST		ONI	SM
	P	P	ONI	DS	DS	P	ONI	LST	SM	VOD	P		P	ONI
	SM	ONI	DS	SM	ONI	DS	LST	P	DS	DS	ONI		SM	P

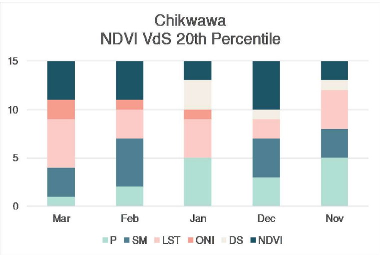
G.2. Overview feature ranking over the growing season

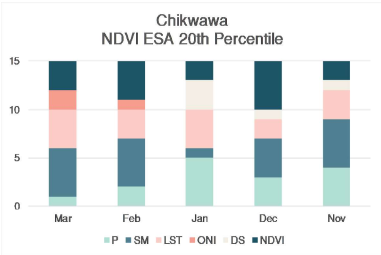
NDVI 20th Percentile

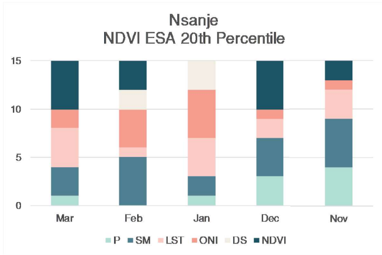


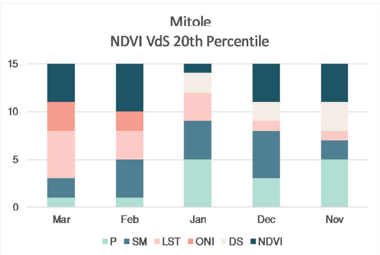


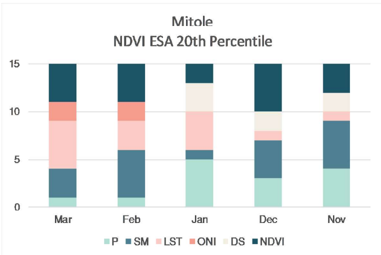


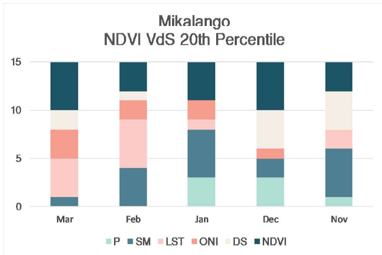


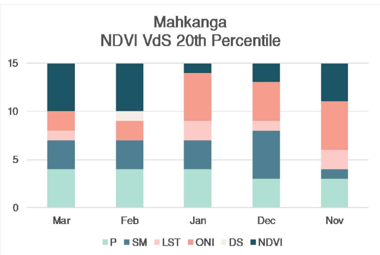


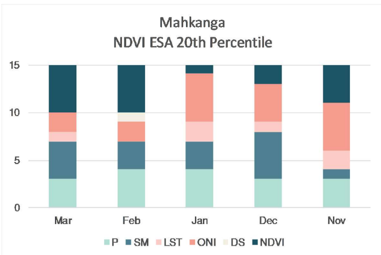


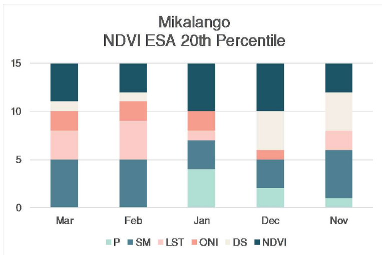


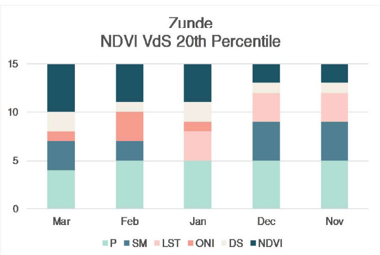




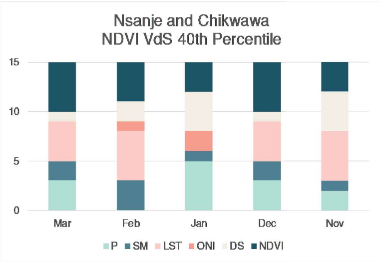


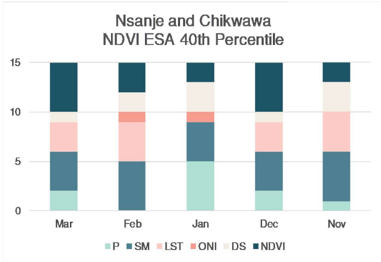


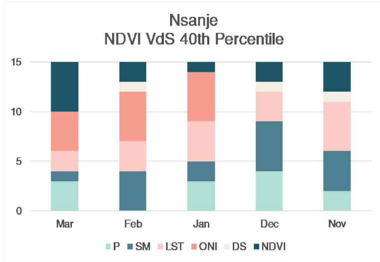


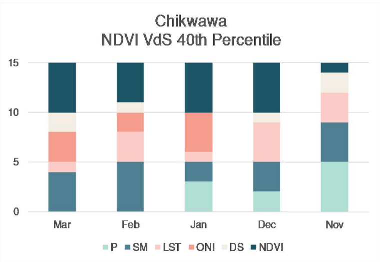


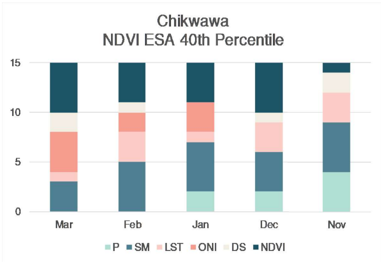
NDVI 40th Percentile

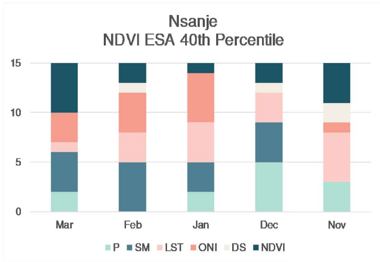


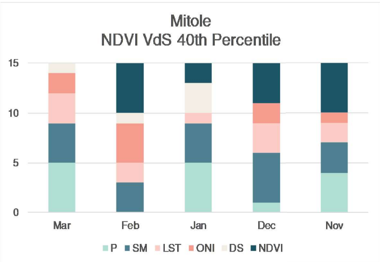


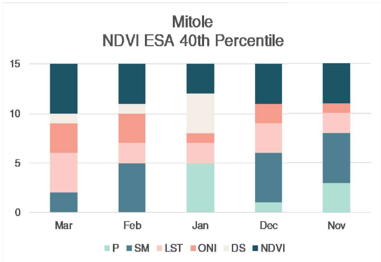


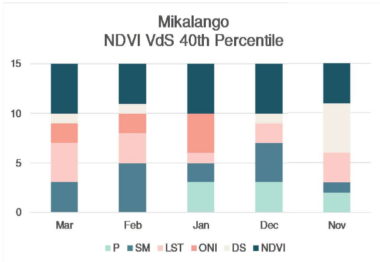


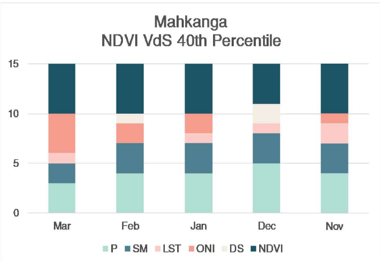


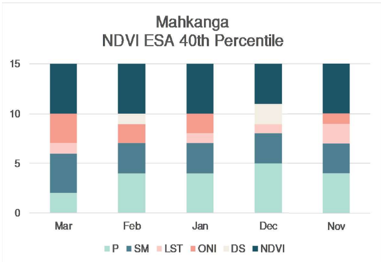


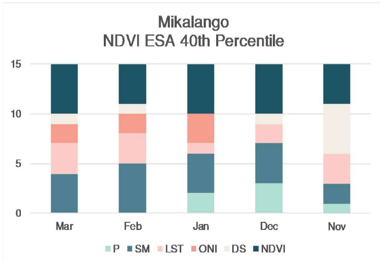


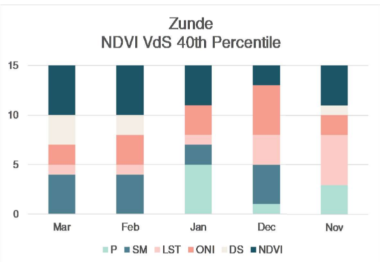




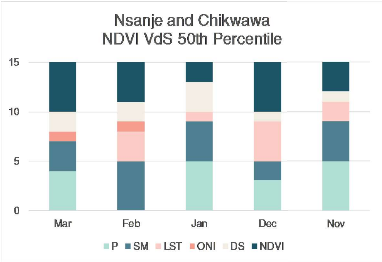


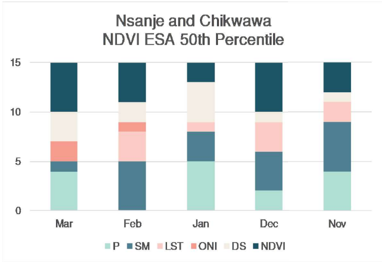


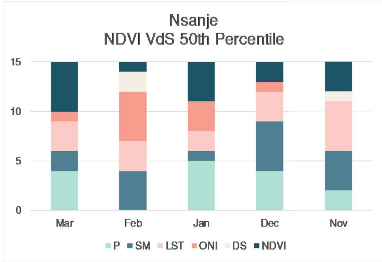


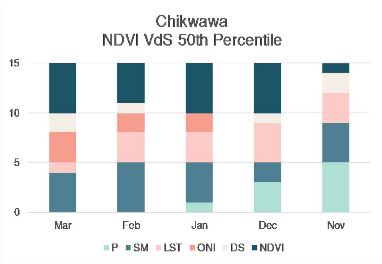


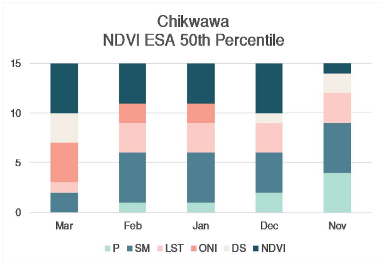
NDVI 50th Percentile

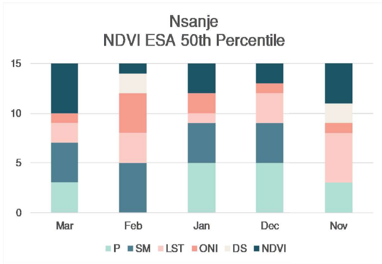


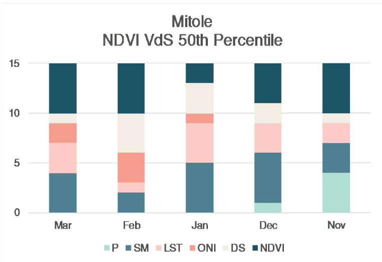


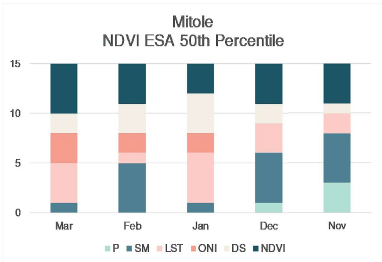


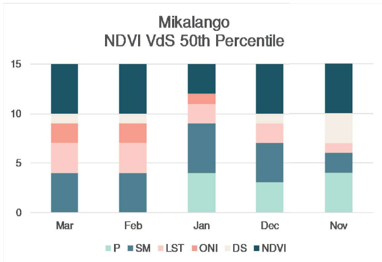


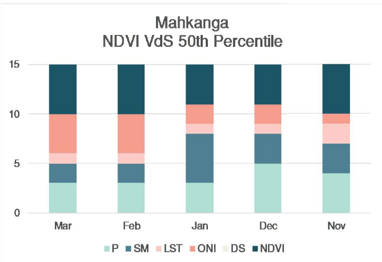


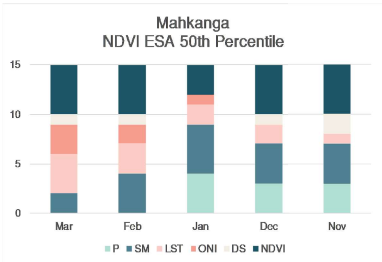


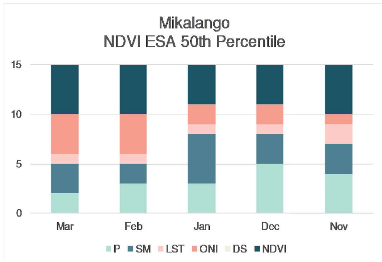


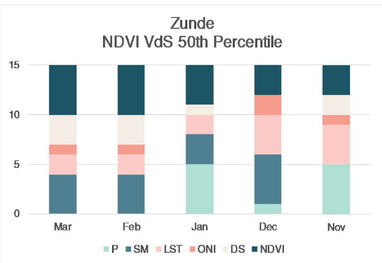




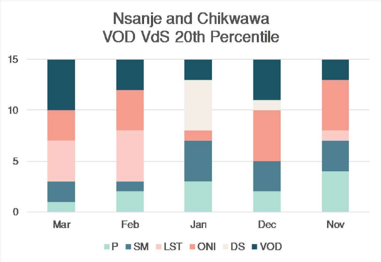


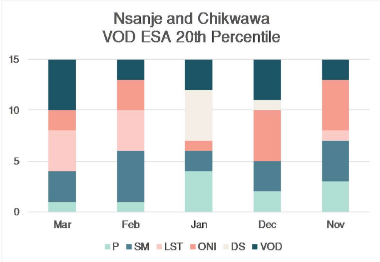


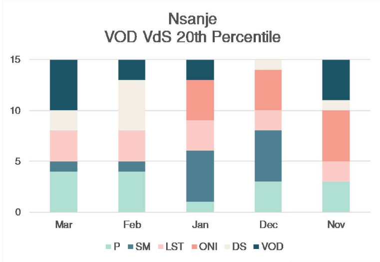


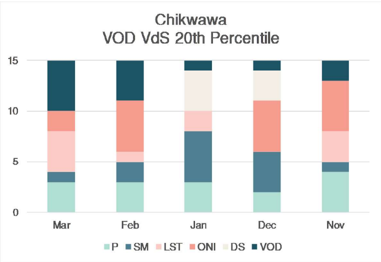


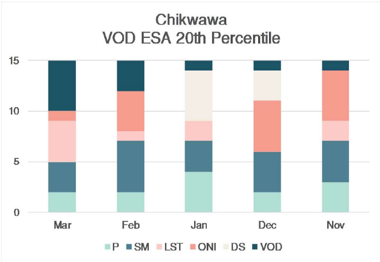
VOD 20th Percentile

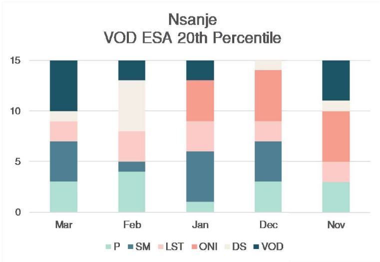


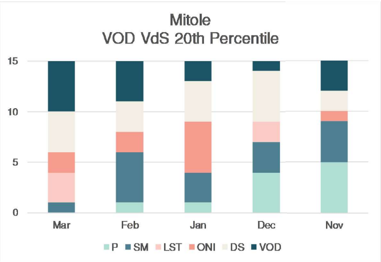


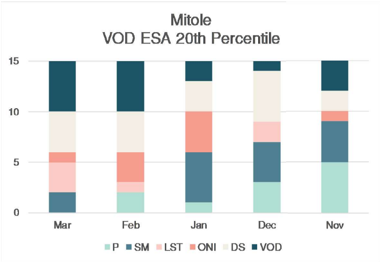


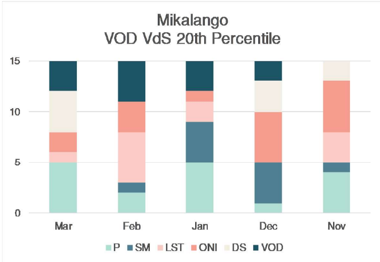


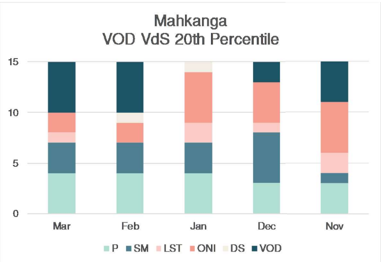


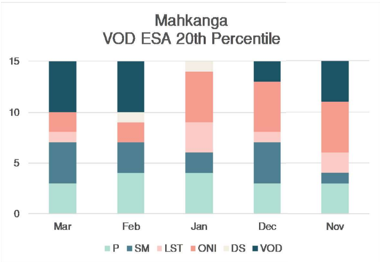


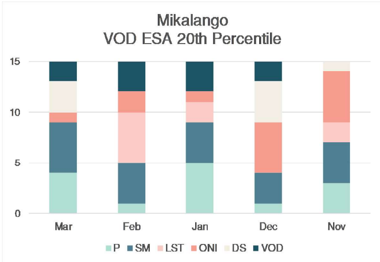


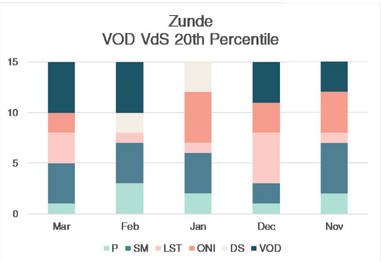




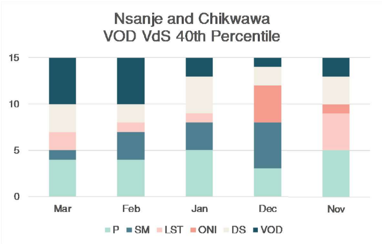


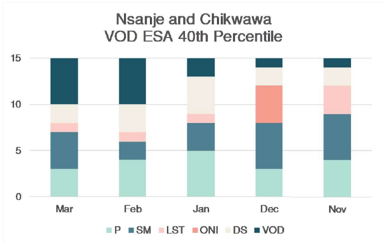


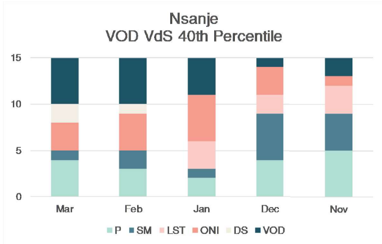


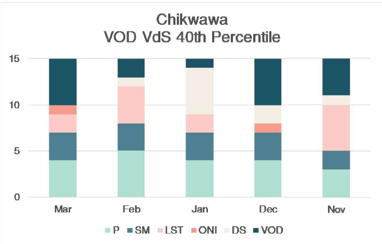


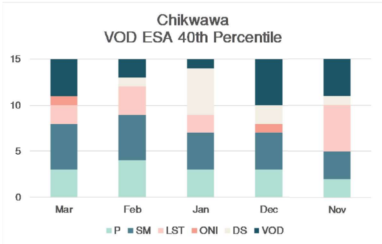
Top five features VOD 40th Percentile

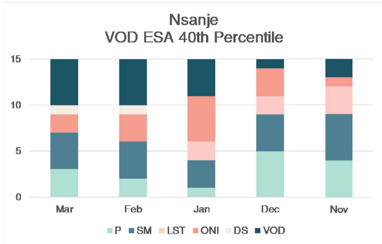


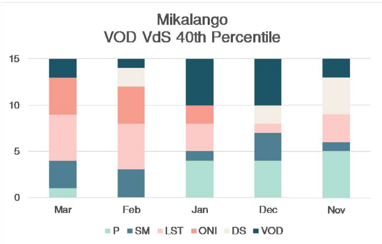


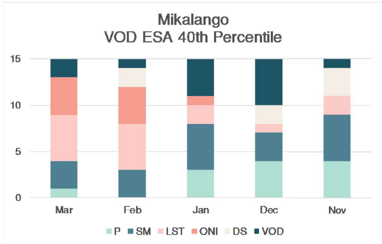


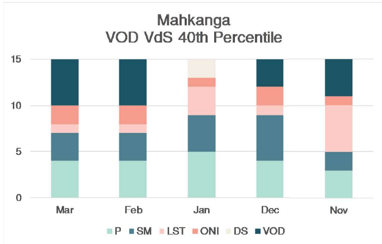


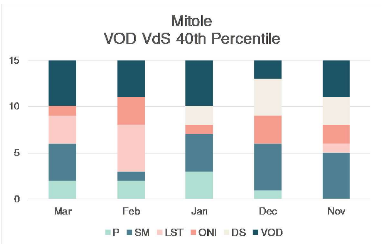


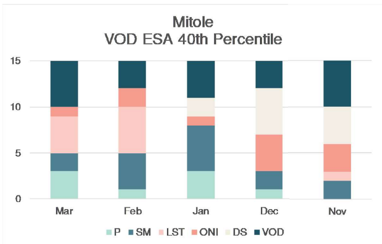


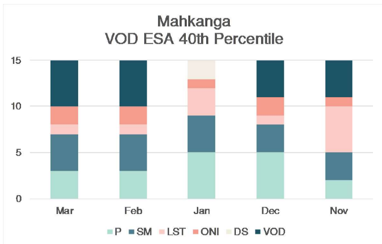


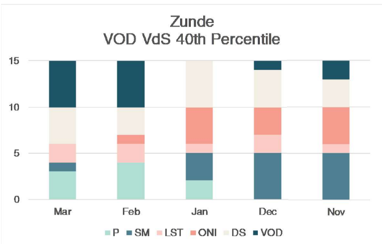




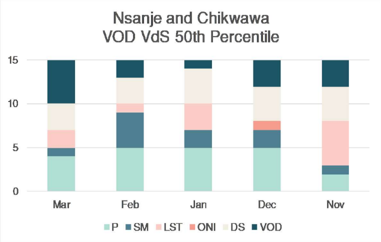


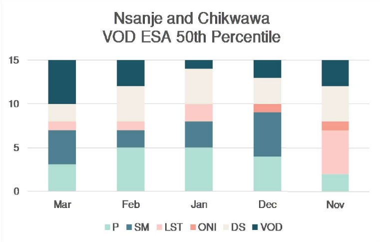


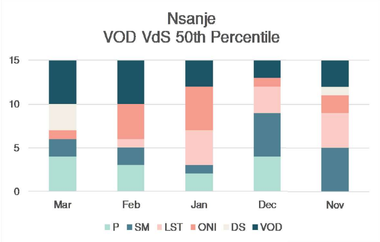


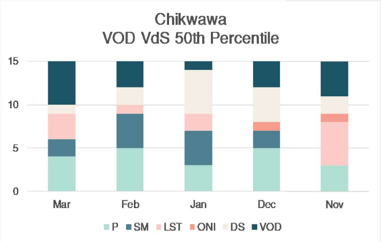


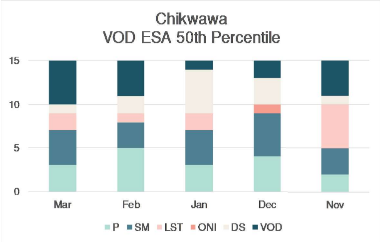
VOD 50th Percentile

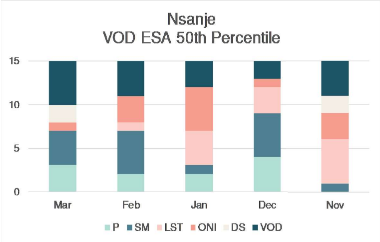


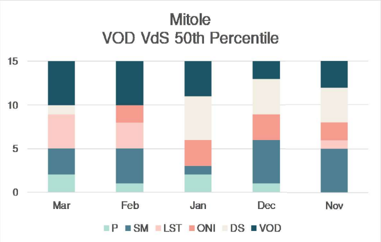


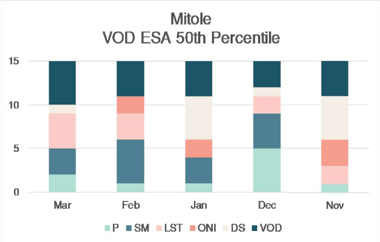


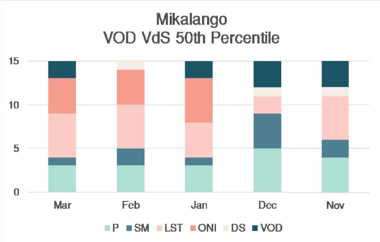


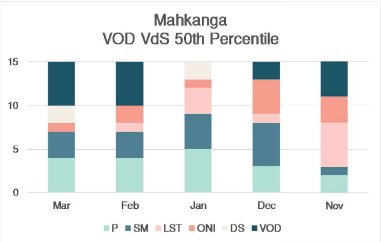


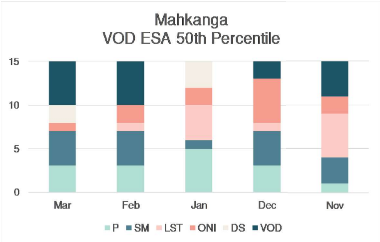


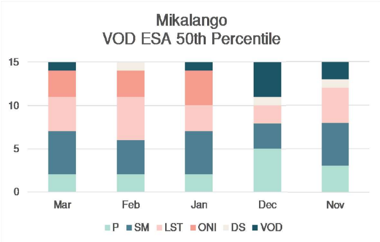


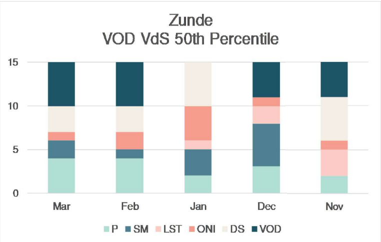






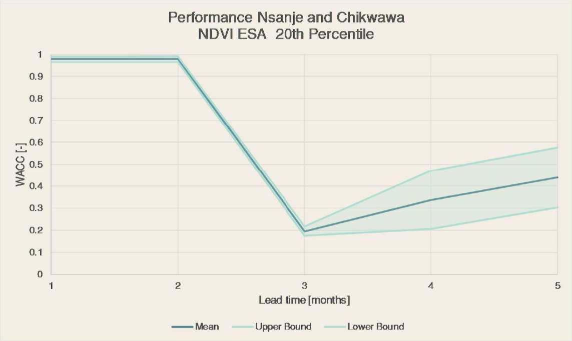






H
Model Performance

H.1. Nsanje and Chikwawa



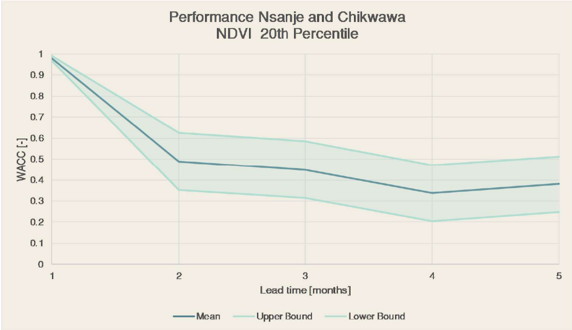
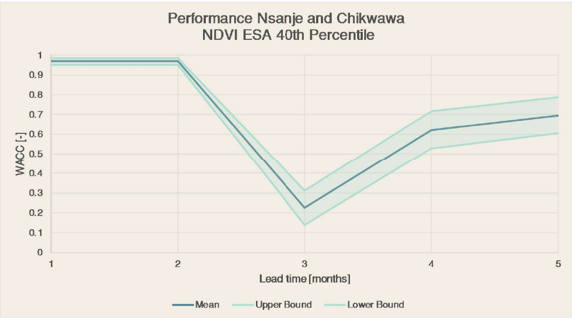
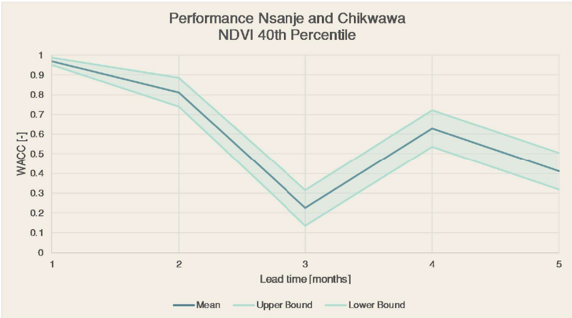
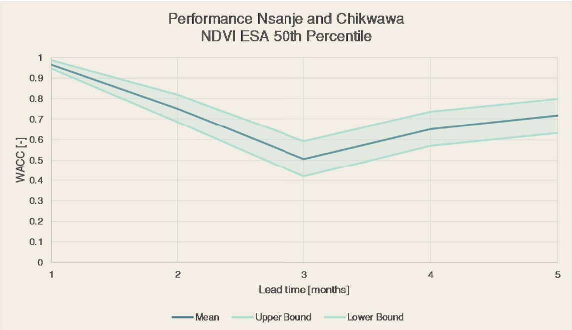
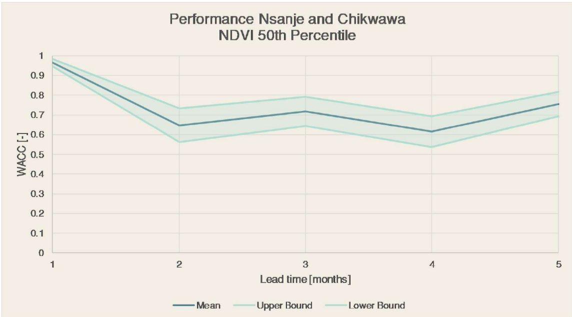


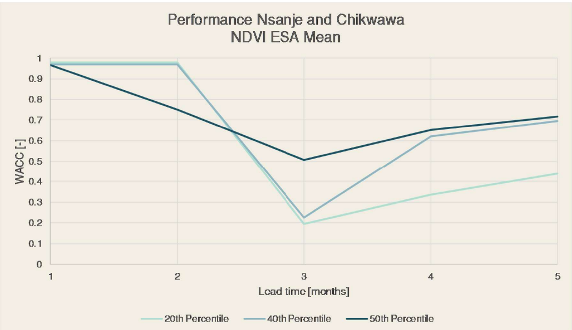
Figure H.1

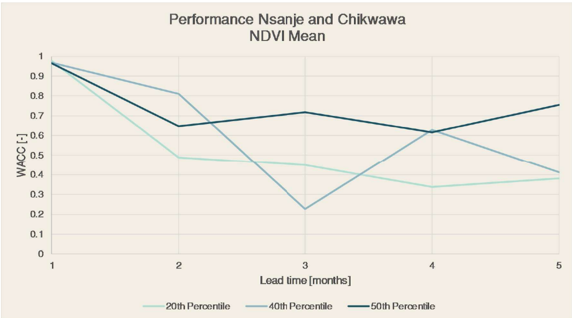


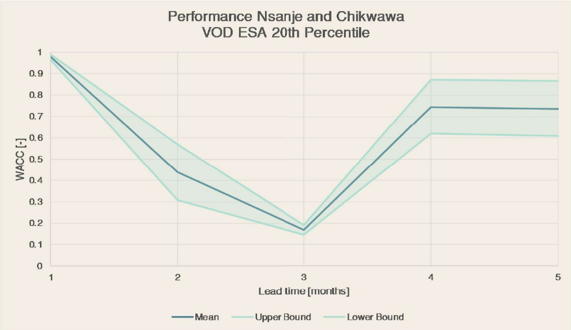


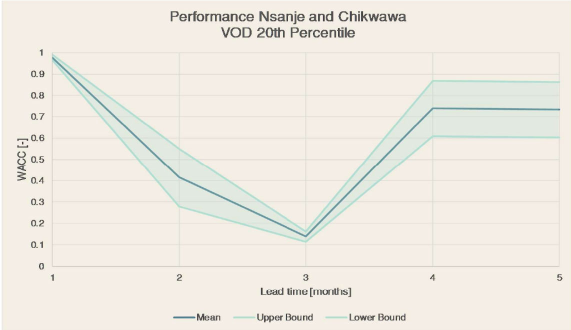


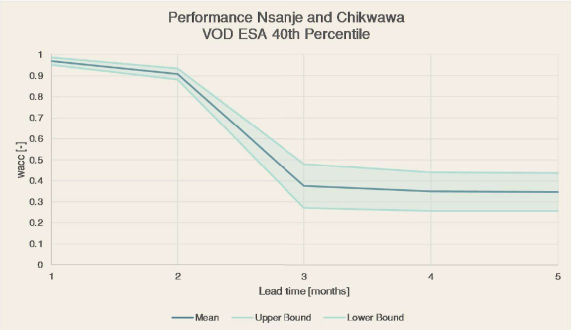


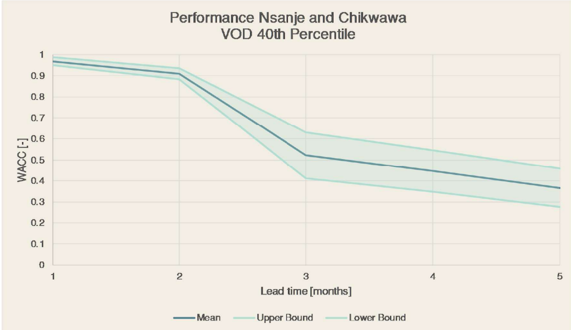


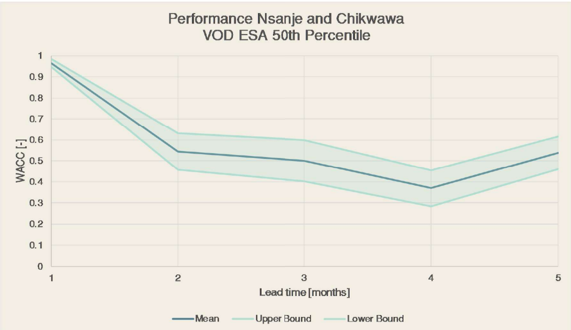


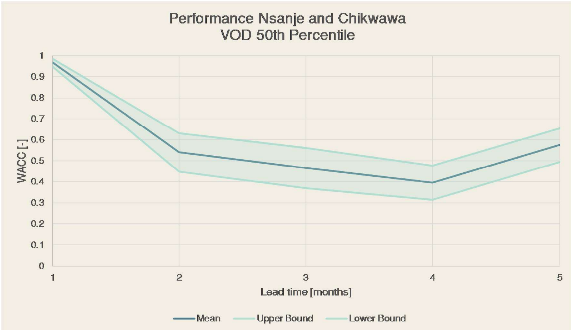

Performance Nsanje and Chikwawa
VOD ESA 20th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

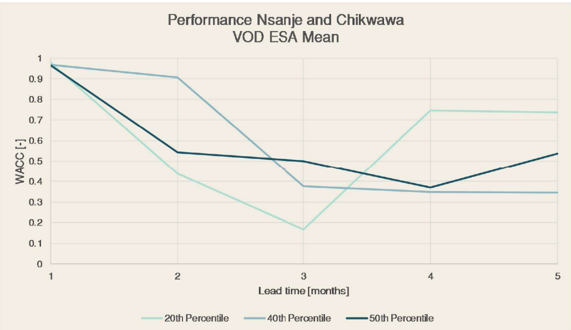

Performance Nsanje and Chikwawa
VOD 20th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

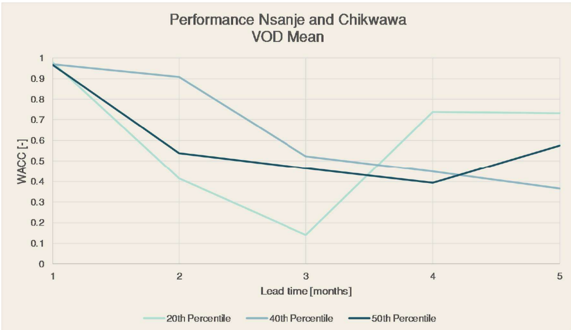

Performance Nsanje and Chikwawa
VOD ESA 40th Percentile
wacc [-]
Lead time [months]
Mean
Upper Bound
Lower Bound


Performance Nsanje and Chikwawa
VOD 40th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

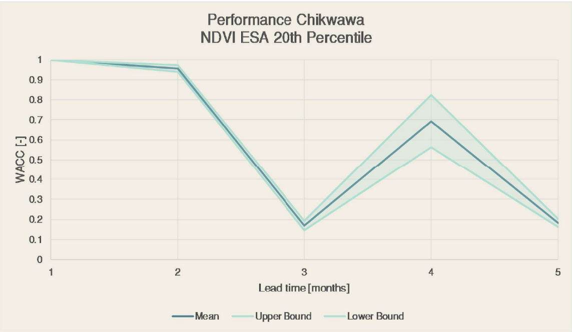

Performance Nsanje and Chikwawa
VOD ESA 50th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

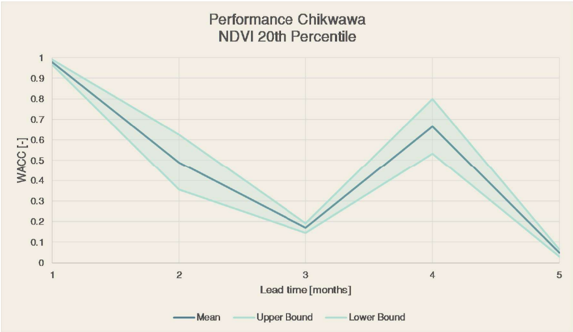

Performance Nsanje and Chikwawa
VOD 50th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

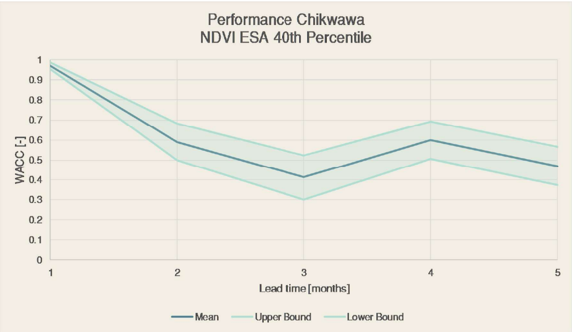

Performance Nsanje and Chikwawa
VOD ESA Mean
WACC [-]
Lead time [months]
20th Percentile
40th Percentile
50th Percentile

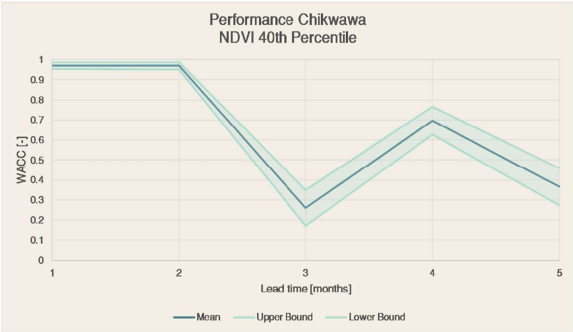

Performance Nsanje and Chikwawa
VOD Mean
WACC [-]
Lead time [months]
20th Percentile
40th Percentile
50th Percentile

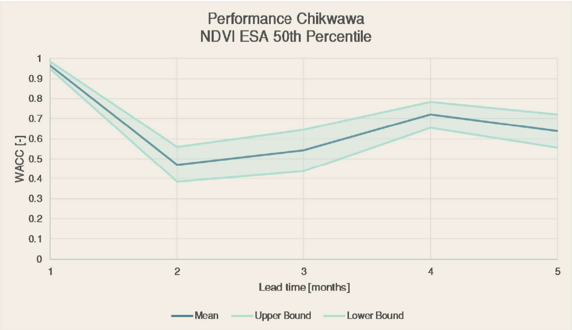
H.2. Chikwawa

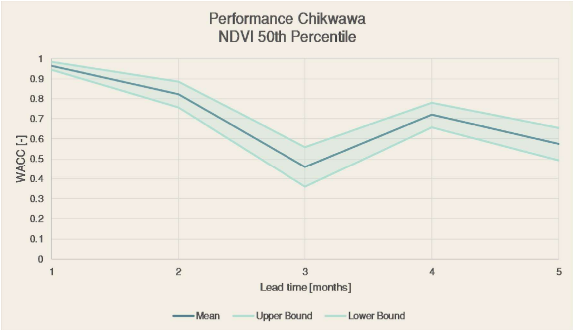

Performance Chikwawa
NDVI ESA 20th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

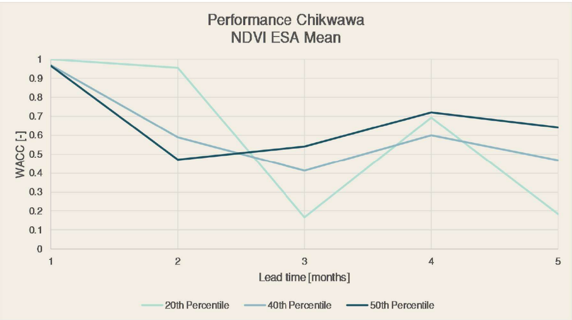

Performance Chikwawa
NDVI 20th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

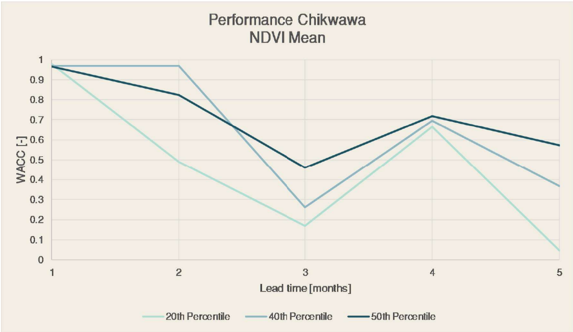

Performance Chikwawa
NDVI ESA 40th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

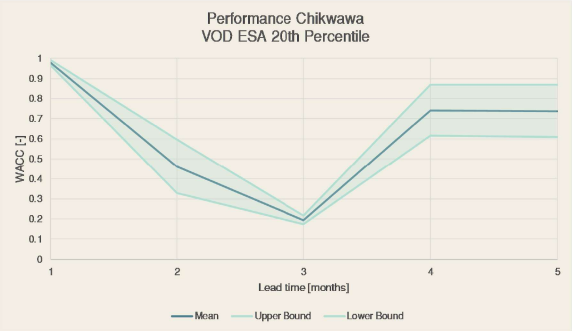

Performance Chikwawa
NDVI 40th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

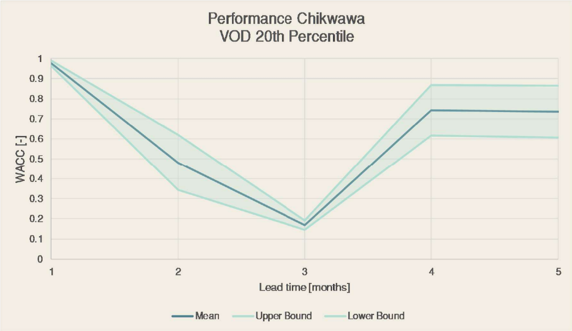

Performance Chikwawa
NDVI ESA 50th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

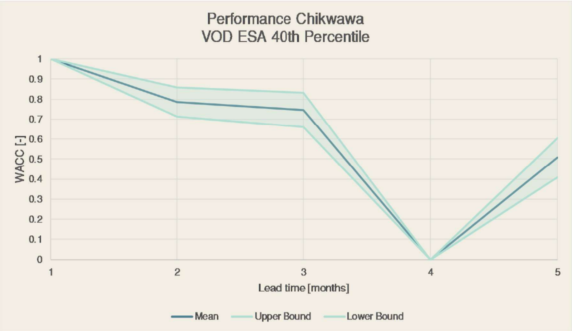

Performance Chikwawa
NDVI 50th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

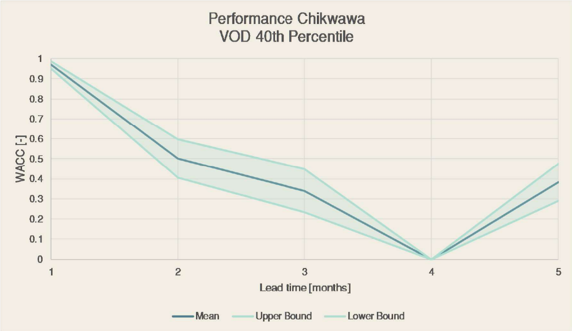

Performance Chikwawa
NDVI ESA Mean
WACC [-]
Lead time [months]
20th Percentile
40th Percentile
50th Percentile

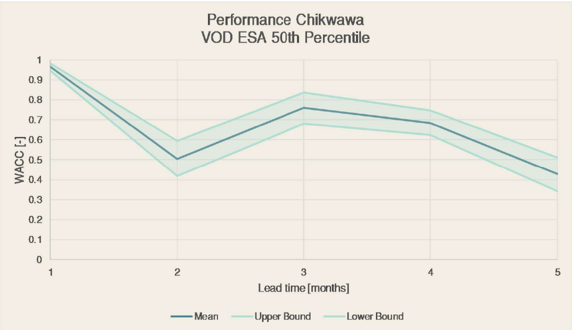

Performance Chikwawa
NDVI Mean
WACC [-]
Lead time [months]
20th Percentile
40th Percentile
50th Percentile

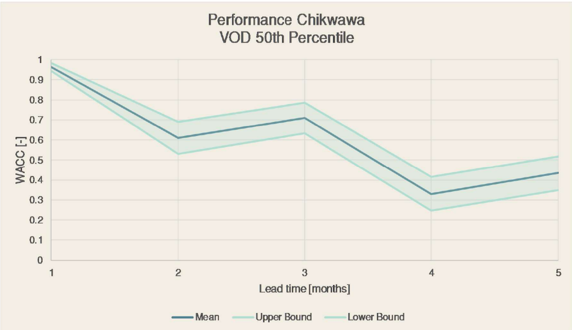

Performance Chikwawa
VOD ESA 20th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

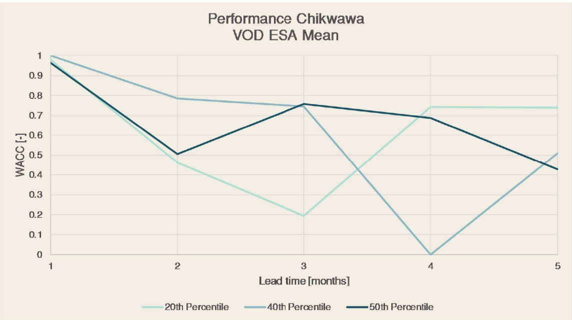

Performance Chikwawa
VOD 20th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

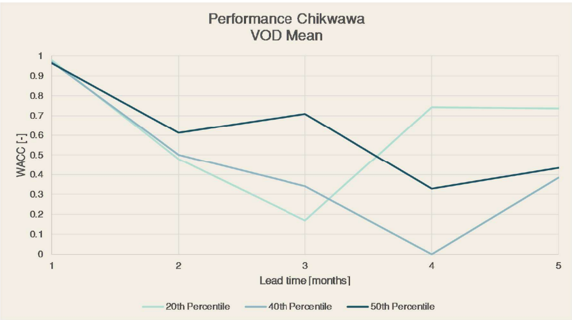

Performance Chikwawa
VOD ESA 40th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound


Performance Chikwawa
VOD 40th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

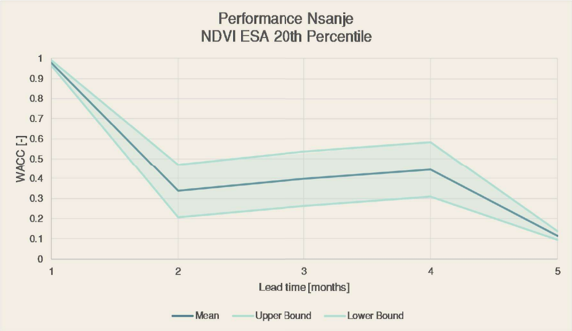

Performance Chikwawa
VOD ESA 50th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

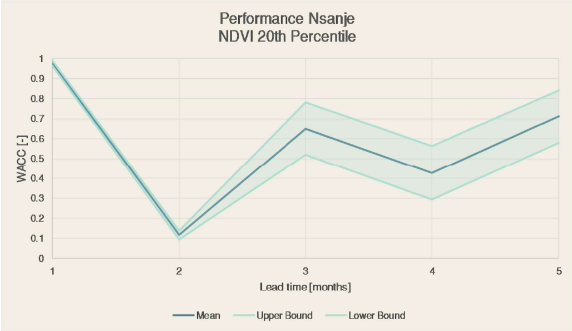

Performance Chikwawa
VOD 50th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

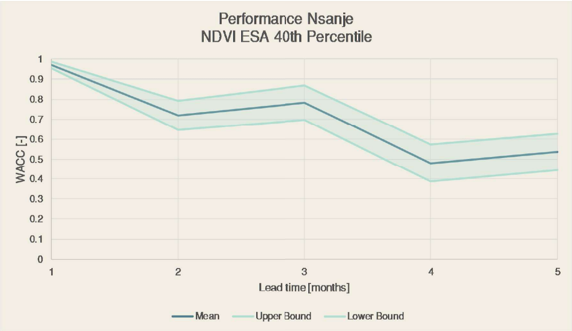

Performance Chikwawa
VOD ESA Mean
WACC [-]
Lead time [months]
20th Percentile
40th Percentile
50th Percentile

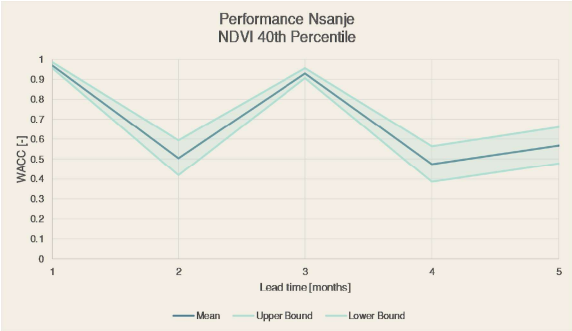

Performance Chikwawa
VOD Mean
WACC [-]
Lead time [months]
20th Percentile
40th Percentile
50th Percentile

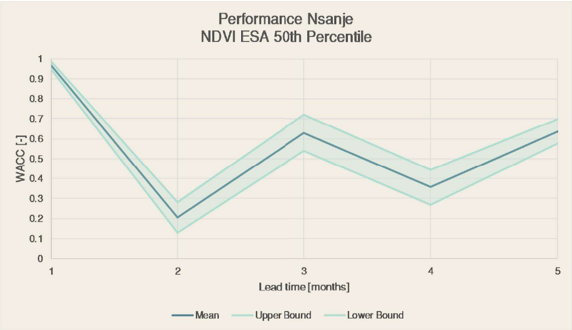
H.3. Nsanje

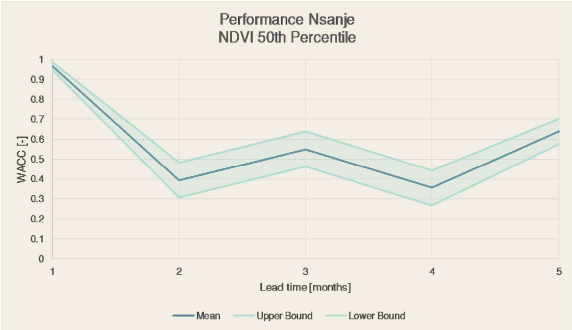

Performance Nsanje
NDVI ESA 20th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

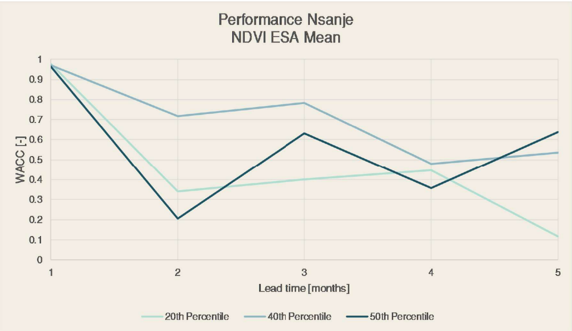

Performance Nsanje
NDVI 20th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

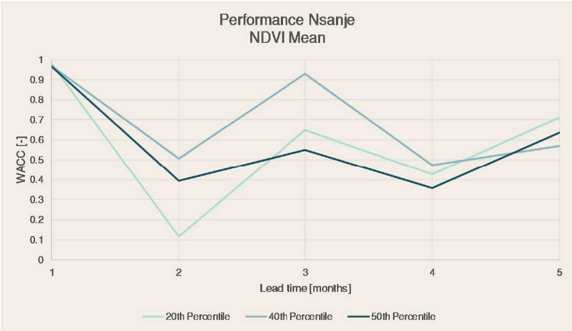

Performance Nsanje
NDVI ESA 40th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

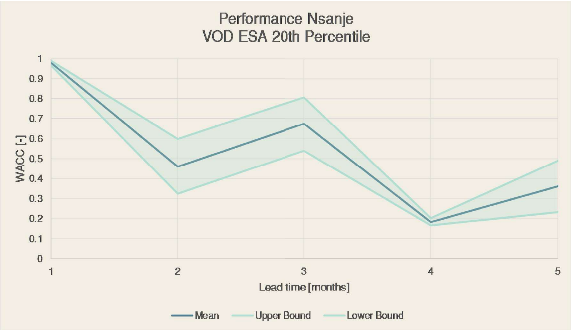

Performance Nsanje
NDVI 40th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

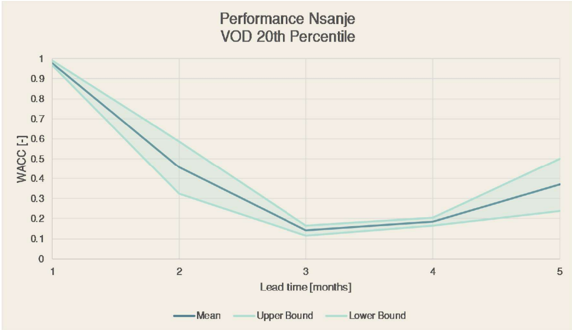

Performance Nsanje
NDVI ESA 50th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

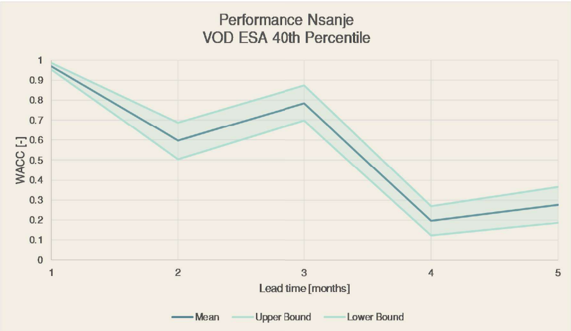

Performance Nsanje
NDVI 50th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

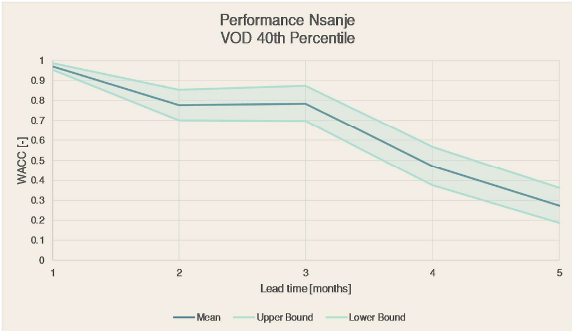

Performance Nsanje
NDVI ESA Mean
WACC [-]
Lead time [months]
20th Percentile
40th Percentile
50th Percentile

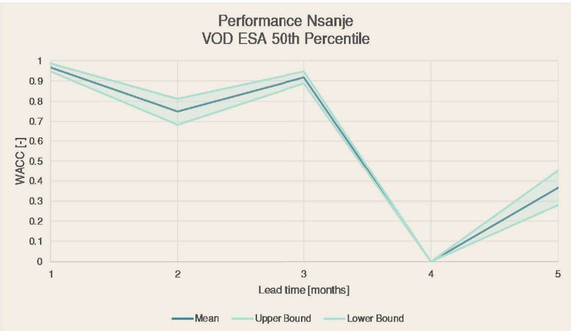

Performance Nsanje
NDVI Mean
WACC [-]
Lead time [months]
20th Percentile
40th Percentile
50th Percentile

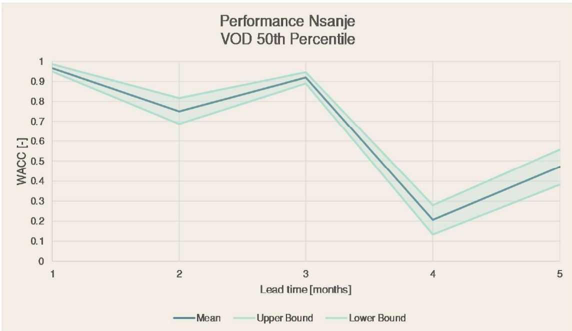

Performance Nsanje
VOD ESA 20th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

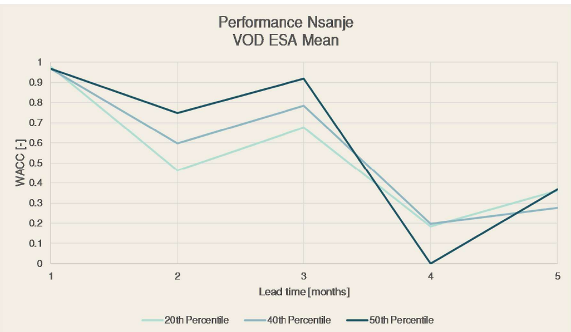

Performance Nsanje
VOD 20th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

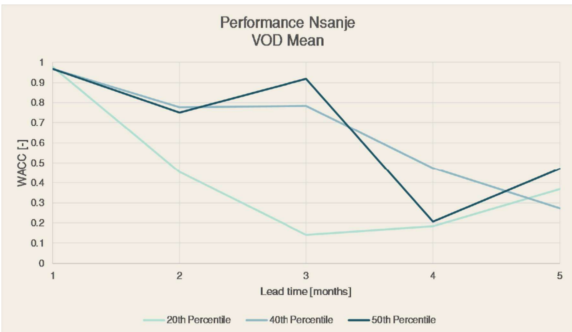

Performance Nsanje
VOD ESA 40th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound


Performance Nsanje
VOD 40th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

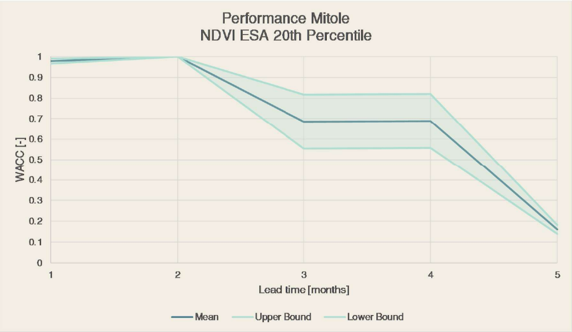

Performance Nsanje
VOD ESA 50th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

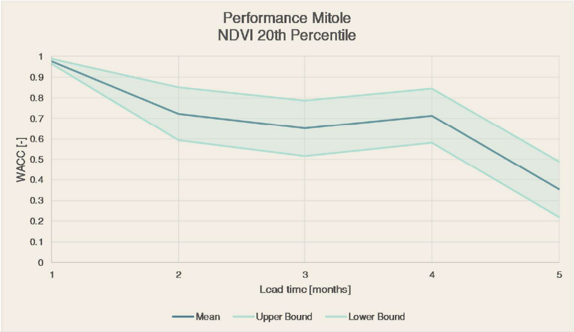

Performance Nsanje
VOD 50th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

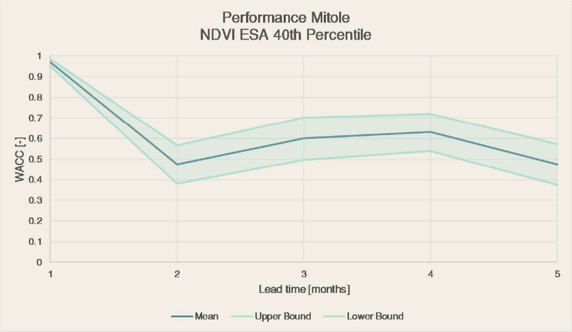

Performance Nsanje
VOD ESA Mean
WACC [-]
Lead time [months]
20th Percentile
40th Percentile
50th Percentile

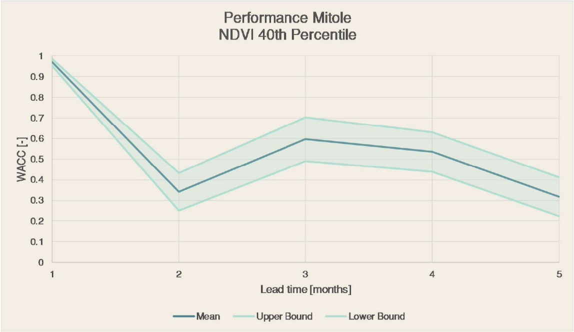

Performance Nsanje
VOD Mean
WACC [-]
Lead time [months]
20th Percentile
40th Percentile
50th Percentile

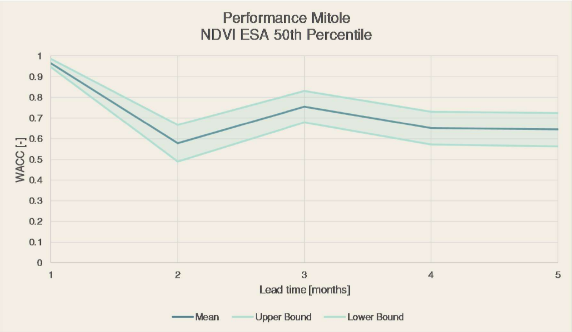
H.4. Mitole

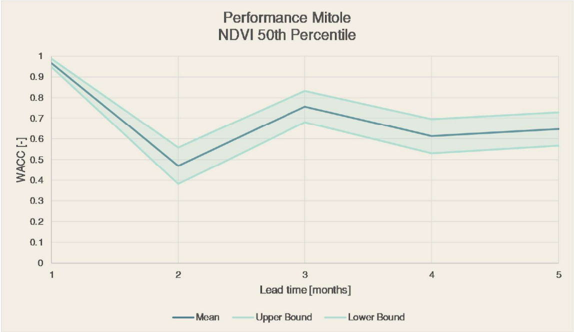

Performance Mitole
NDVI ESA 20th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

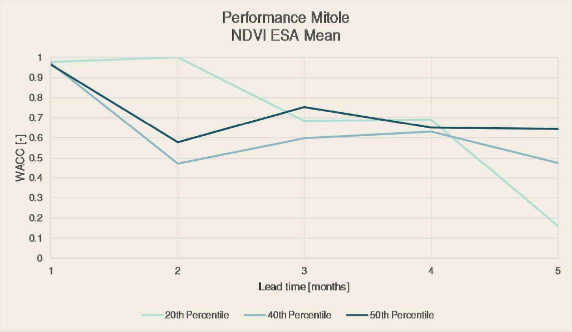

Performance Mitole
NDVI 20th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

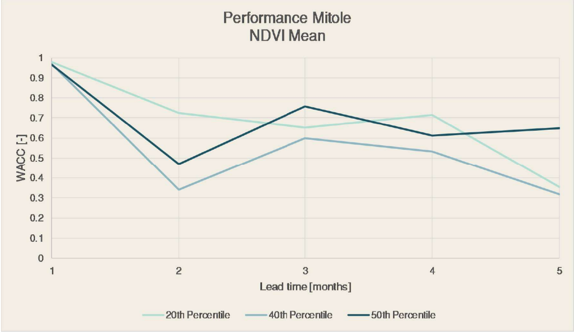

Performance Mitole
NDVI ESA 40th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

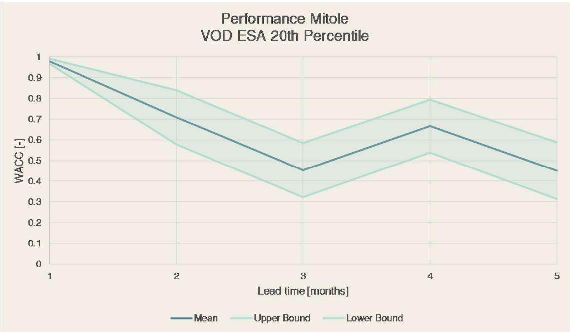

Performance Mitole
NDVI 40th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

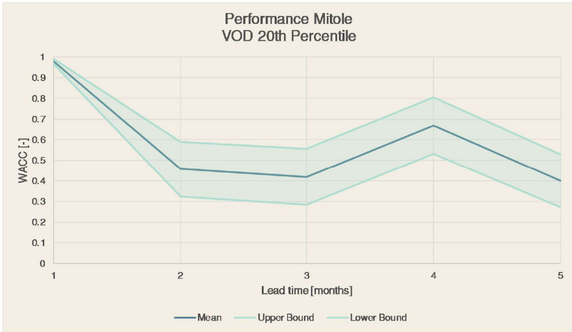

Performance Mitole
NDVI ESA 50th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

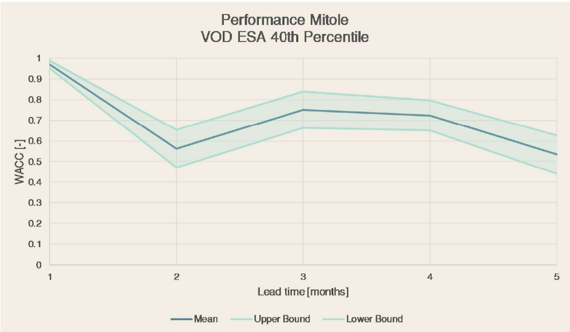

Performance Mitole
NDVI 50th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

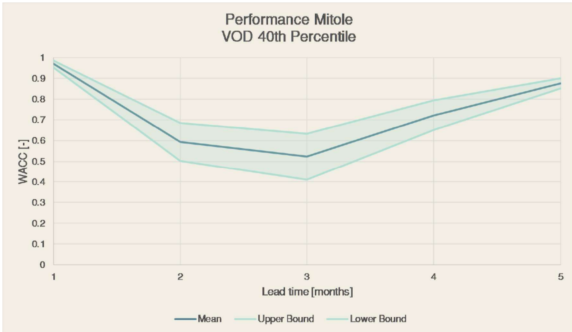

Performance Mitole
NDVI ESA Mean
WACC [-]
Lead time [months]
20th Percentile
40th Percentile
50th Percentile

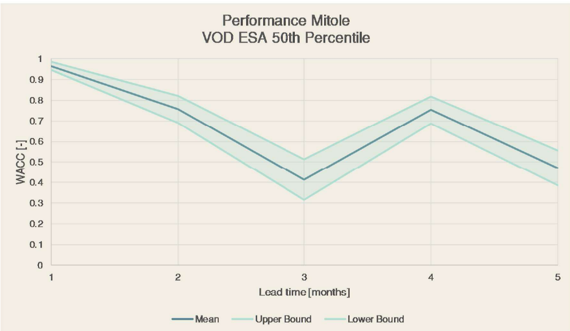

Performance Mitole
NDVI Mean
WACC [-]
Lead time [months]
20th Percentile
40th Percentile
50th Percentile

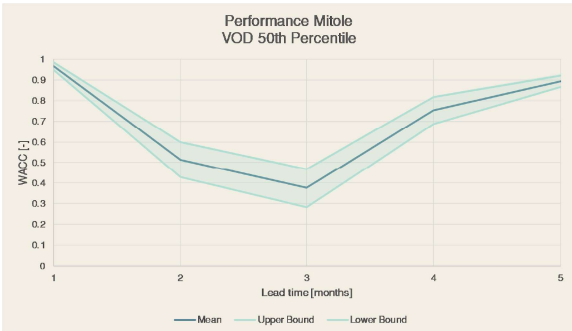

Performance Mitole
VOD ESA 20th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

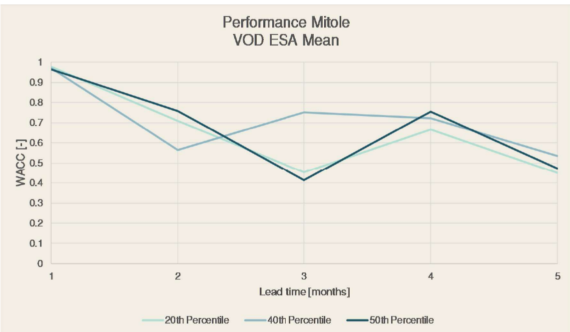

Performance Mitole
VOD 20th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

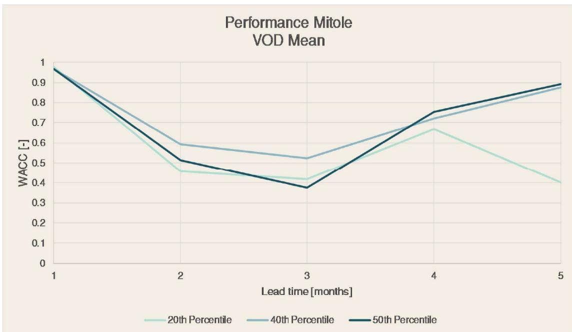

Performance Mitole
VOD ESA 40th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound


Performance Mitole
VOD 40th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

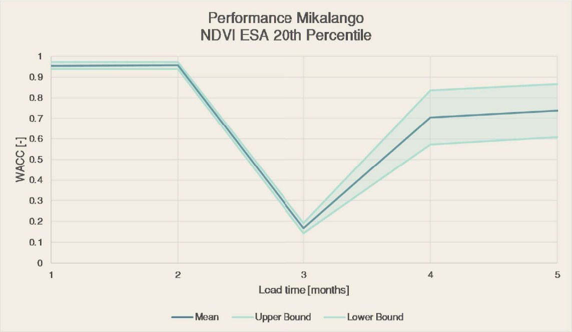

Performance Mitole
VOD ESA 50th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

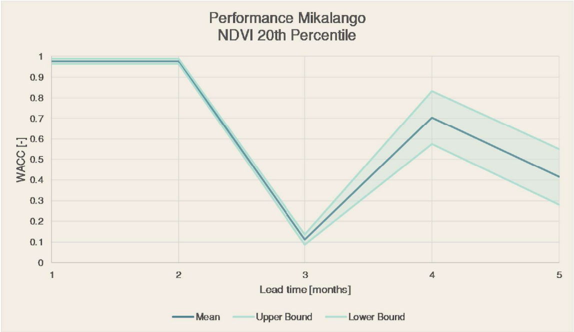

Performance Mitole
VOD 50th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

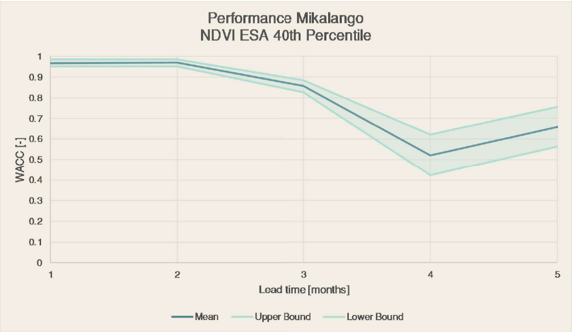

Performance Mitole
VOD ESA Mean
WACC [-]
Lead time [months]
20th Percentile
40th Percentile
50th Percentile

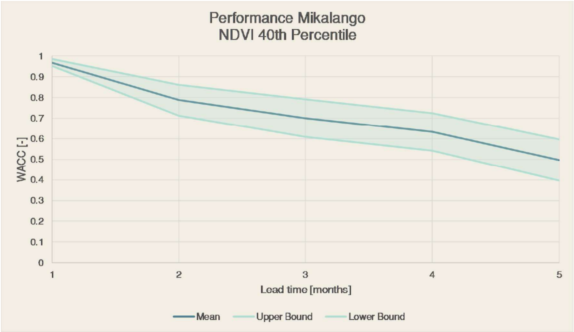

Performance Mitole
VOD Mean
WACC [-]
Lead time [months]
20th Percentile
40th Percentile
50th Percentile

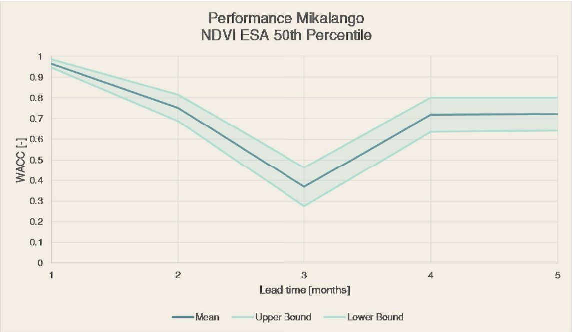
H.5. Mikalango

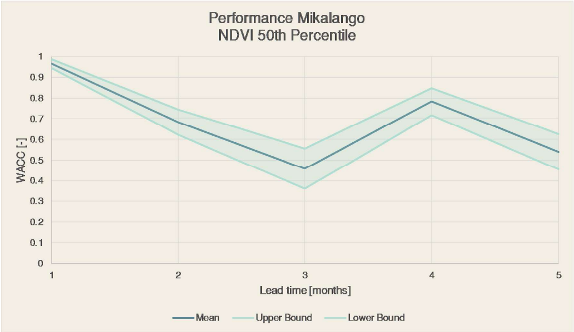


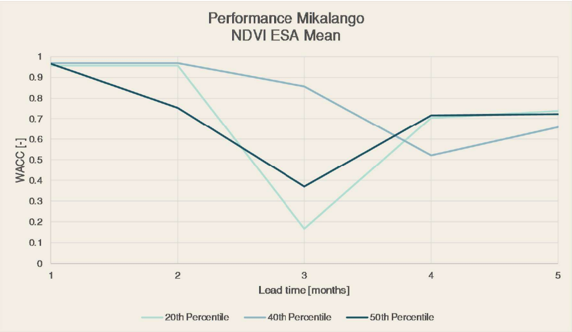


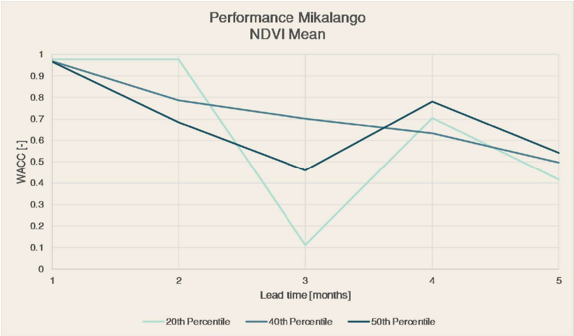


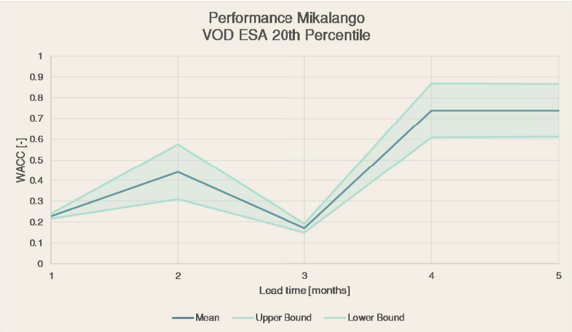




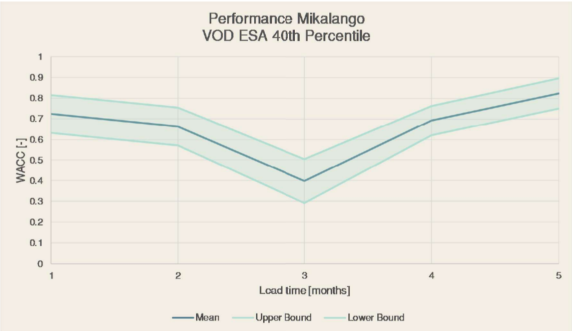


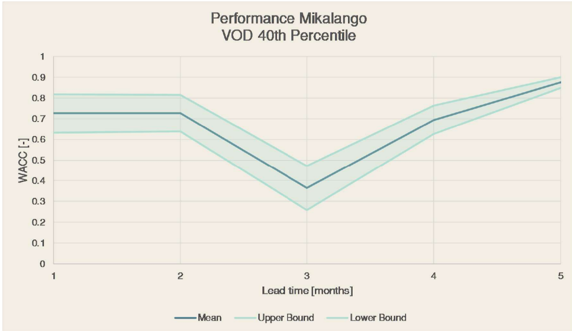


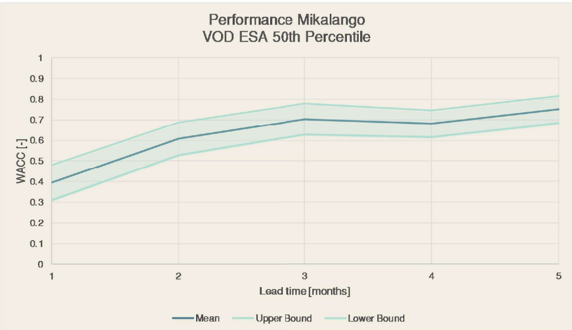


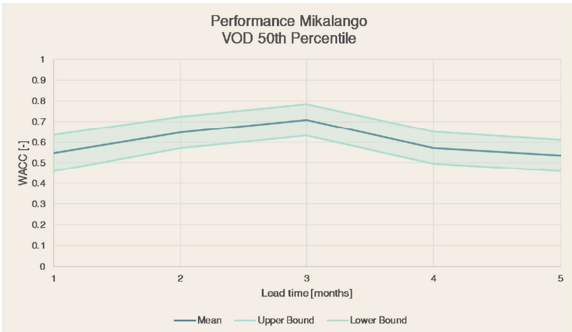

Performance Mikalango
VOD ESA 20th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

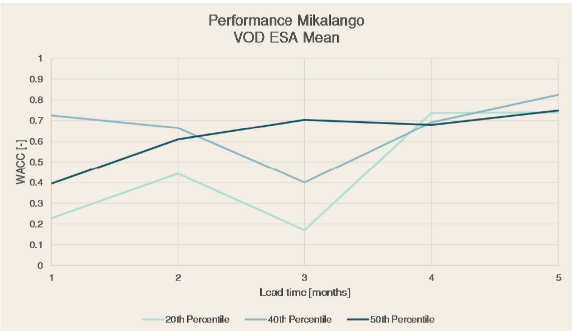

Performance Mikalango
VOD 20th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

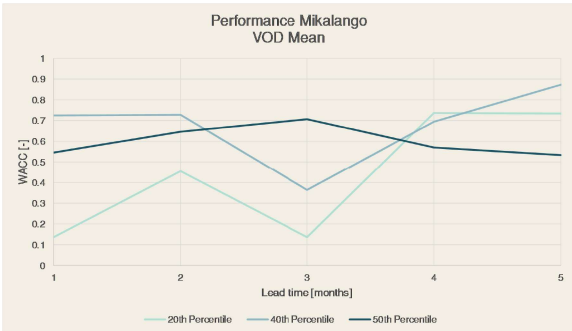

Performance Mikalango
VOD ESA 40th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound


Performance Mikalango
VOD 40th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound


Performance Mikalango
VOD ESA 50th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

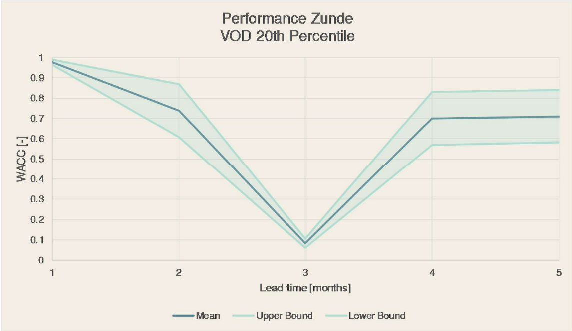

Performance Mikalango
VOD 50th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

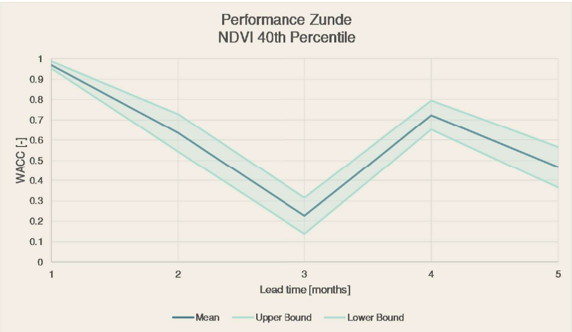

Performance Mikalango
VOD ESA Mean
WACC [-]
Lead time [months]
20th Percentile
40th Percentile
50th Percentile

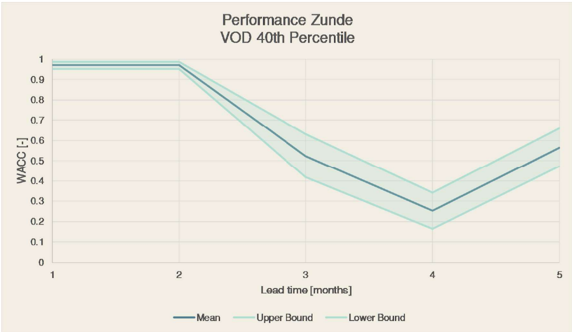

Performance Mikalango
VOD Mean
WACC [-]
Lead time [months]
20th Percentile
40th Percentile
50th Percentile

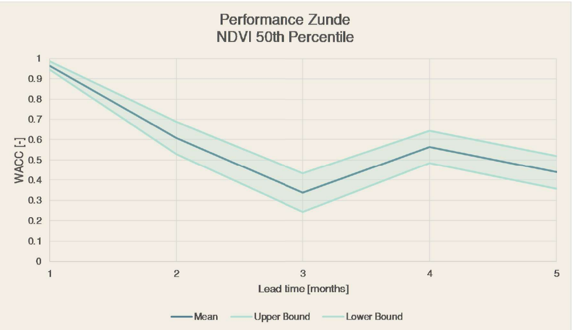
H.6. Zunde

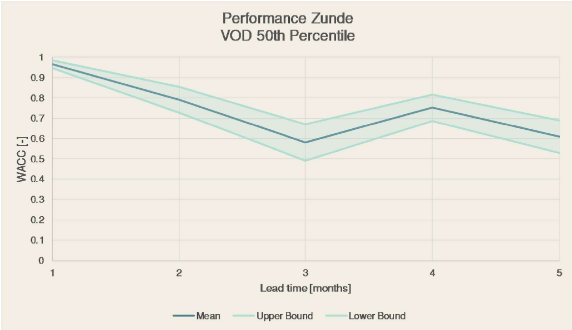

Performance Zunde
NDVI 20th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

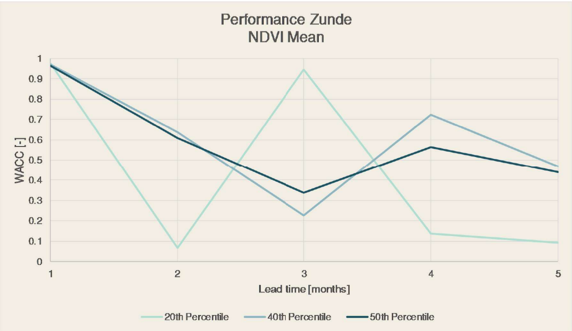

Performance Zunde
VOD 20th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

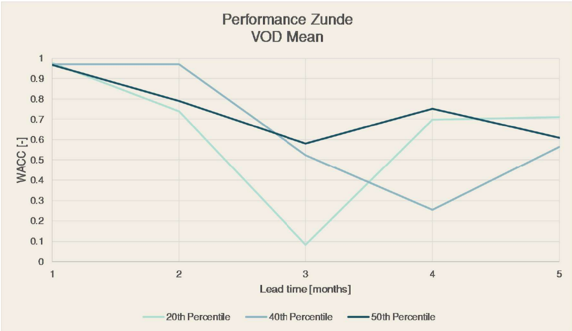

Performance Zunde
NDVI 40th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound


Performance Zunde
VOD 40th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound


Performance Zunde
NDVI 50th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

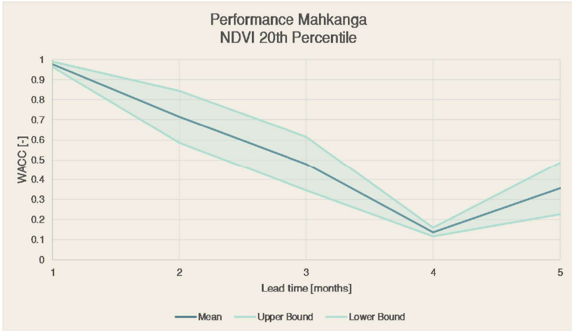

Performance Zunde
VOD 50th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

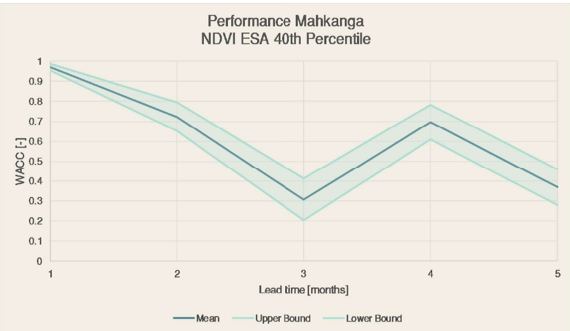

Performance Zunde
NDVI Mean
WACC [-]
Lead time [months]
20th Percentile
40th Percentile
50th Percentile

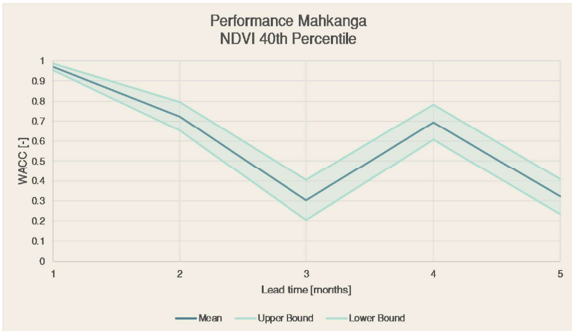

Performance Zunde
VOD Mean
WACC [-]
Lead time [months]
20th Percentile
40th Percentile
50th Percentile

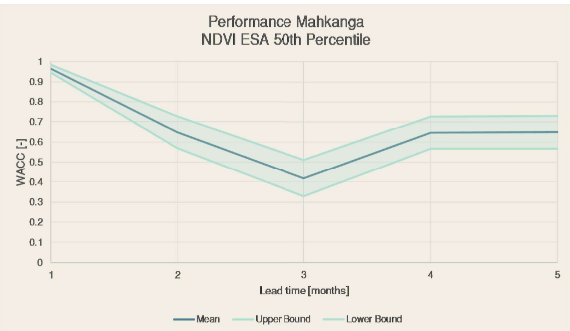
H.7. Mahkanga

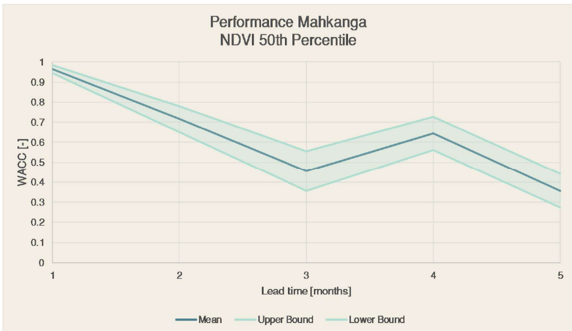

Performance Mahkanga
NDVI ESA 20th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

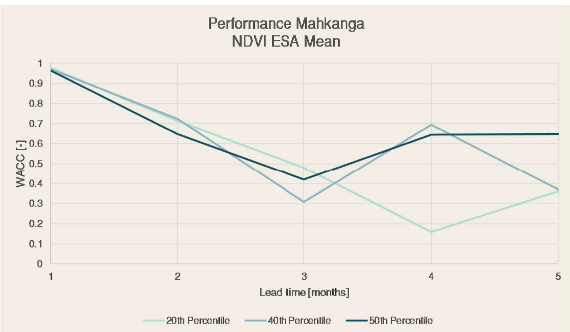

Performance Mahkanga
NDVI 20th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

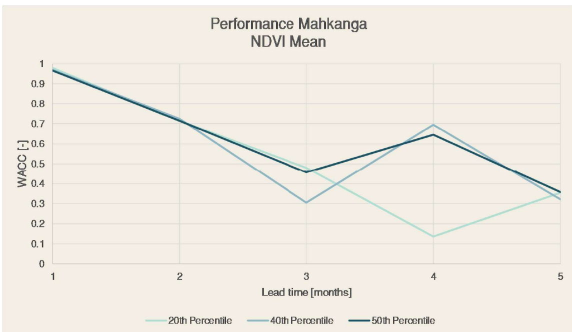

Performance Mahkanga
NDVI ESA 40th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

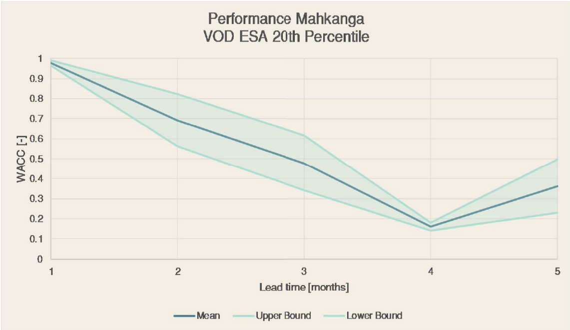

Performance Mahkanga
NDVI 40th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

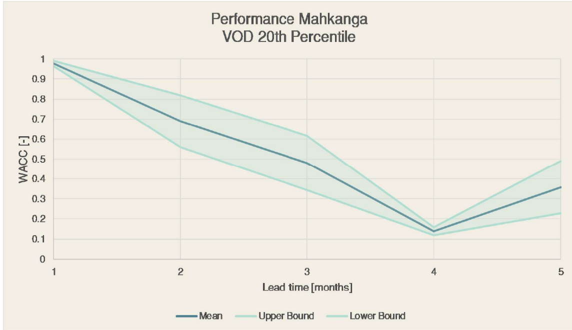

Performance Mahkanga
NDVI ESA 50th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

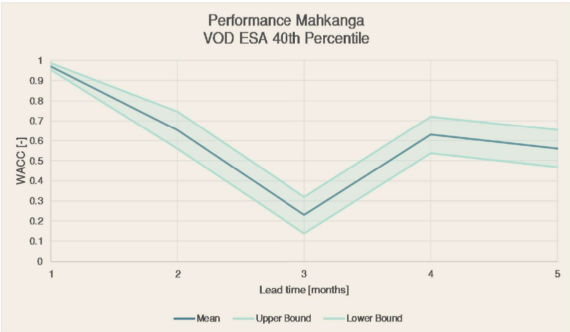

Performance Mahkanga
NDVI 50th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

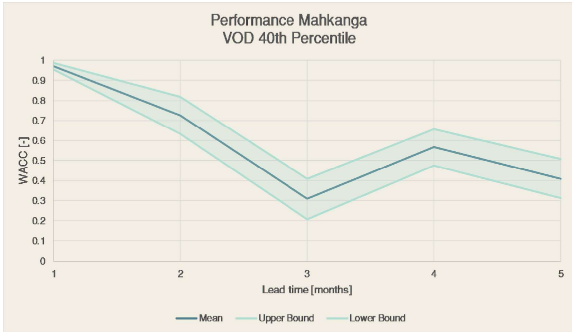

Performance Mahkanga
NDVI ESA Mean
WACC [-]
Lead time [months]
20th Percentile
40th Percentile
50th Percentile

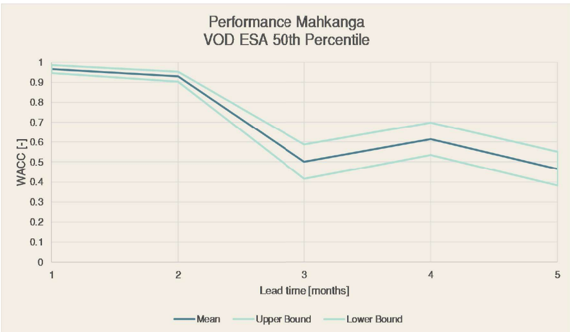

Performance Mahkanga
NDVI Mean
WACC [-]
Lead time [months]
20th Percentile
40th Percentile
50th Percentile

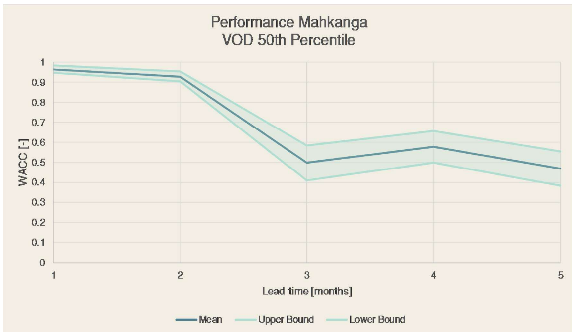

Performance Mahkanga
VOD ESA 20th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

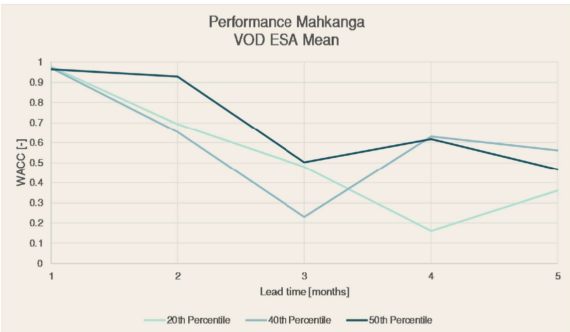

Performance Mahkanga
VOD 20th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound

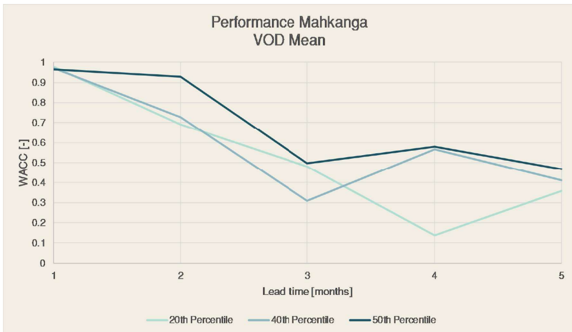

Performance Mahkanga
VOD ESA 40th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound


Performance Mahkanga
VOD 40th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound


Performance Mahkanga
VOD ESA 50th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound


Performance Mahkanga
VOD 50th Percentile
WACC [-]
Lead time [months]
Mean
Upper Bound
Lower Bound


Performance Mahkanga
VOD ESA Mean
WACC [-]
Lead time [months]
20th Percentile
40th Percentile
50th Percentile


Performance Mahkanga
VOD Mean
WACC [-]
Lead time [months]
20th Percentile
40th Percentile
50th Percentile

Bibliography

About the U.S. Drought Monitor, URL https://droughtmonitor.unl.edu/AboutUSDM.aspx.

EM-DAT, URL https://www.emdat.be/frequently-asked-questions.

ESA CCI Soil Moisture data, URL http://al-tc002.xtr.deltares.nl:8080/thredds/catalog/opendap/mirrored/pml/tuwien/ESACCI-SOILMOISTURE-L3S-SSMV-COMBINED_1978-2015-v03.2/1997/catalog.html?dataset=varopendap/mirrored/pml/tuwien/ESACCI-SOILMOISTURE-L3S-SSMV-COMBINED_1978-2015-v03.2/1997/ESAC.

Global Food Security Index, URL https://foodsecurityindex.eiu.com/.

Guidelines | EM-DAT, URL https://www.emdat.be/guidelines.

Types of Drought, URL http://drought.unl.edu/DroughtBasics/TypesofDrought.aspx.

White Paper: Agroclimatology FEWS NET, URL http://fews.net/sites/default/files/White%20Paper%20-%20Agroclimatology%20-%20Aug2017v2.pdf.

Seeing Leaves in a New Light, URL https://earthobservatory.nasa.gov/features/LAI/LAI2.php, 2002.

ENSO: Recent Evolution, Current Status and Predictions Outline Summary Recent Evolution and Current Conditions Oceanic Niño Index (ONI) Pacific SST Outlook U.S. Seasonal Precipitation and Temperature Outlooks, Tech. rep., URL http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/lanina/enso_evolution-status-fcsts-web.pdf, 2018.

Andrew Collings, Nick Maunder, M. M. A. M. and van Aalst, M.: World Disaster Report 2009 Focus on early warning, early action, Tech. rep., IFRC, Geneva, URL http://www.ifrc.org/Global/WDR2009-full.pdf, 2009.

Asa Giertz, Jorge Caballero, D. G. D. M. J. O. and German, G.: Agriculture Global Practice Technical Assistance Paper, Tech. Rep. 99941-MW, World Bank Group, Washington DC, 2015.

Bachmair, S., Stahl, K., Collins, K., Hannaford, J., Acreman, M., Svoboda, M., Knutson, C., Smith, K. H., Wall, N., Fuchs, B., Crossman, N. D., and Overton, I. C.: Drought indicators revisited: the need for a wider consideration of environment and society, Wiley Interdisciplinary Reviews: Water, 3, 516–536, https://doi.org/10.1002/wat2.1154, URL http://doi.wiley.com/10.1002/wat2.1154, 2016a.

Bachmair, S., Svensson, C., Hannaford, J., Barker, L. J., Stahl, K., and Bachmair, S.: A quantitative analysis to objectively appraise drought indicators and model drought impacts, Hydrol. Earth Syst. Sci, 20, 2589–2609, https://doi.org/10.5194/hess-20-2589-2016, URL www.hydrol-earth-syst-sci.net/20/2589/2016/, 2016b.

Barker, L. J., Hannaford, J., Chiverton, A., and Svensson, C.: From meteorological to hydrological drought using standardised indicators, Hydrol. Earth Syst. Sci, 20, 2483–2505, https://doi.org/10.5194/hess-20-2483-2016, URL www.hydrol-earth-syst-sci.net/20/2483/2016/, 2016.

Barron, J., Rockström, J., Gichuki, F., and Hatibu, N.: Dry spell analysis and maize yields for two semi-arid locations in east Africa, 117, 23–37, https://doi.org/10.1016/S0168-1923(03)00037-6, 2003.

Bhuiyan, C.: Various Drought Indices For Monitoring Drought Condition In Aravalli Terrain Of India, URL http://www.isprs.org/proceedings/xxxv/congress/comm7/papers/243.pdf, 2004.

Braman, L. M., van Aalst, M. K., Mason, S. J., Suarez, P., Ait-Chellouche, Y., and Tall, A.: Climate forecasts in disaster management: Red Cross flood operations in West Africa, 2008, Disasters, 37, 144–164, https://doi.org/10.1111/j.1467-7717.2012.01297.x, 2013.

Brocca, L., Hasenauer, S., Lacava, T., Melone, F., Moramarco, T., Wagner, W., Dorigo, W., Matgen, P., Martínez-Fernández, J., Llorens, P., Latron, J., Martin, C., and Bittelli, M.: Soil moisture estimation through ASCAT and AMSR-E sensors: An intercomparison and validation study across Europe, Remote Sensing of Environment, 115, 3390–3408, https://doi.org/10.1016/j.rse.2011.08.003, URL http://dx.doi.org/10.1016/j.rse.2011.08.003, 2011.

Brouwer, C.: Irrigation Water Management: Irrigation Water Needs, Training manual no. 3, FAO, Rome, URL http://www.fao.org/3/S2022E/s2022e00.htm#Contents, 1986.

Chen, J., Li, M., and Wang, W.: Statistical uncertainty estimation using random forests and its application to drought forecast, Mathematical Problems in Engineering, 2012, 1–13, https://doi.org/10.1155/2012/915053, 2012.

Coughlan De Perez, E., Van Den Hurk, B., Van Aalst, M. K., Jongman, B., Klose, T., and Suarez, P.: Forecast-based financing: an approach for catalyzing humanitarian action based on extreme weather and climate forecasts, Nat. Hazards Earth Syst. Sci, 15, 895–904, https://doi.org/10.5194/nhess-15-895-2015, URL www.nat-hazards-earth-syst-sci.net/15/895/2015/, 2015.

Coulibaly, J. Y., Mbow, C., Sileshi, G. W., Beedy, T., Kundhlande, G., Musau, J., and Coulibaly, Y. J.: Mapping Vulnerability to Climate Change in Malawi: Spatial and Social Differentiation in the Shire River Basin, American Journal of Climate Change, 4, 282–294, https://doi.org/10.4236/ajcc.2015.43023, URL http://www.scirp.org/journal/ajcchttp://dx.doi.org/10.4236/ajcc.2015.43023http://creativecommons.org/licenses/by/4.0/, 2015.

Creswell, J. W.: Research Design, Sage, California.

Critchley, W., Siegert, K., and Chapman, C.: A Manual for the Design and Construction of Water Harvesting Schemes for Plant Production, FAO, Rome, URL http://www.fao.org/3/u3160e/u3160e00.htm#Contents, 1991.

Deparday, V., Gevaert, C., Molinario Robert Soden, G., and Balog-Way, S.: M A C H I N E L E A R N I N G for DISASTER RISK MANAGEMENT A guidance note on how machine learning can be used for disaster risk management, including key definitions, case studies, and practical considerations for implementation ACKNOWLEDGMENTS PHOTO CREDITS MACHINE LEARNING FOR DISASTER RISK MANAGEMENT, Tech. rep., URL http://creativecommons.org/licenses/by/3.0/igo/www.worldbank.org, 2018.

DoDMA: Decentralized early warning systems in Malawi, Tech. rep., Government of Malawi Department of Disaster Management Affairs, URL http://meteorology.uonbi.ac.ke/sites/default/files/cbps/sps/meteorology/03June2015-RevisedEWSDraftReport(Piyasi).pdf, 2015.

Enenkel, M., Steiner, C., Mistelbauer, T., Dorigo, W., Wagner, W., See, L., Atzberger, C., Schneider, S., and Rogenhofer, E.: A combined satellite-derived drought indicator to support humanitarian aid organizations, Remote Sensing, 8, 1–7, https://doi.org/10.3390/rs8040340, 2016.

European Commission: Humanitarian aid, URL https://ec.europa.eu/echo/what/humanitarian-aid_en.

European Commission: Humanitarian Aid ECHO FACTSHEET, URL http://ec.europa.eu/echo/files/aid/countries/factsheets/thematic/humanitarian_aid_en.pdf, 2018.

FAO: Malawi Country fact sheet on food and agriculture policy trends Socio-economic context and role of agriculture, Tech. rep., Rome, URL http://www.fao.org/3/a-i4491e.pdf, 2015a.

FAO: The impact of disasters on agriculture and food security, Tech. rep., FAO, URL www.fao.org/publications, 2015b.

FAO: World reference base for soil resources 2014 International soil classification system for naming soils and creating legends for soil maps, Tech. rep., Rome, URL http://www.fao.org, 2015c.

Fawcett, T.: An introduction to ROC analysis, Pattern Recognition Letters, pp. 861–874, https://doi.org/10.1016/j.patrec.2005.10.010, URL www.elsevier.com/locate/patrec, 2005.

FEWS NET: Integrated Phase Classification | Famine Early Warning Systems Network, URL http://fews.net/nosso-trabalho/nosso-trabalho/classifica%C3%A7%C3%A3o-integrada-de-fases, a.

FEWS NET: Malawi | Famine Early Warning Systems Network, URL http://fews.net/southern-africa/malawi, b.

FEWS NET: FEWS NET Building Rainfall Assumptions for Scenario Development FEWS NET Guidance Document Series, Tech. rep., URL www.fews.net., 2018.

Field, C.B., V. B. T. S. D. D. D. K. M. M., K.J. Mach, G.K. Plattner, S. A. M. T., and Midgley, P.: Managing the risks of extreme events and disasters to advance climate change adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change, Tech. rep., IPCC, Cambridge, Uk and New York, Ny, USA, URL http://www.ipcc.ch/pdf/special-reports/srex/SREX_Full_Report.pdf, 2012.

Freitas, S., Trigo, I., Macedo, J., Barroso, C., Silva, R., and Perdigao, R.: Land Surface Temperature from multiple geostationary satellites, International Journal of Remote Sensing, Vol 34, 2013.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., and Michaelsen, J.: The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes, Scientific Data, 2, 150 066, https://doi.org/10.1038/sdata.2015.66, URL http://www.nature.com/articles/sdata201566, 2015.

Gao, B.-c.: NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space, Remote Sensing of Environment, 58, 257–266, https://doi.org/10.1016/S0034-4257(96)00067-3, URL https://www.sciencedirect.com/science/article/abs/pii/S0034425796000673, 1996.

German Red Cross: Closing The Gap: reconciling short-term disaster response with long-term risk reduction with 'forecast-based financing' (FBF), URL http://www.ifrc.org/PageFiles/212330/Closingthegap_Fbfpolicypaper_GRC.pdf, 2016.

GFDRR: Climate Risk and Adaptation Country Profile Climate Risk and Adaptation Country Profile Malawi, Tech. rep., The World Bank Group, Washington, URL http://sdwebx.worldbank.org/climateportal/countryprofile/doc/GFDRRCountryProfiles/wb_gfdrr_climate_change_country_profile_for_MWI.pdf, 2011.

Gibbs, W. and Maher, J.: Rainfall deciles as drought indicators, Bureau of Meteorology, Melbourne, URL https://www.worldcat.org/title/rainfall-deciles-as-drought-indicators/oclc/214962347, 1967.

Gobron, N.: LEAF AREA INDEX (LAI), URL www.fao.org/gtos/tems.

Golden Gate Weather Services: El Niño and La Niña Years and Intensities, URL https://www.ggweather.com/enso/oni.htm, 2019.

Government of Malawi: Annex II – Feasibility Study I Scaling-up Early Warning Systems and Use of Climate Information in Malawi – Feasibility Assessment Scaling-up Early Warning Systems and Use of Climate Information in Malawi – Feasibility Assessment, Tech. rep., URL http://www.adaptation-undp.org/sites/default/files/resources/feasibility_study_1.pdf, 2015.

Guimarães Nobre, G., Davenport, F., Bischiniotis, K., Veldkamp, T., Jongman, B., Funk, C. C., Husak, G., Ward, P. J., and Aerts, J. C.: Financing agricultural drought risk through ex-ante cash transfers, Science of The Total Environment, 653, 523–535, https://doi.org/10.1016/j.scitotenv.2018.10.406, URL https://linkinghub.elsevier.com/retrieve/pii/S0048969718343067, 2019.

Haak, E.: Towards a Governance Structure for the Data-Driven Prioritization of Humanitarian Aid A data ecosystem approach, Ph.D. thesis, Delft University of Technology, 2017.

Hao, Z., Xia, Y., Luo, L., Singh, V. P., Ouyang, W., and Hao, F.: Toward a categorical drought prediction system based on U.S. Drought Monitor (USDM) and climate forecast, Journal of Hydrology, 551, 300–305, https://doi.org/10.1016/J.JHYDROL.2017.06.005, URL https://www.sciencedirect.com/science/article/pii/S0022169417304043?via%3Dihub, 2017.

Hao, Z., Singh, V. P., and Xia, Y.: Seasonal Drought Prediction: Advances, Challenges, and Future Prospects, Reviews of Geophysics, 56, 108–141, https://doi.org/10.1002/2016RG000549, 2018.

IASC Working Group: Humanitarian System-Wide Emergency Activation: definition and procedures - IASC Transformative Agenda Reference Document - World | ReliefWeb, URL https://reliefweb.int/report/world/humanitarian-system-wide-emergency-activation-definition-and-procedures-iasc.

IFRC: Forecast-based Financing, URL http://www.forecast-based-financing.org/about/, a.

IFRC: Manual Forecast-based Financing: Phase I, URL http://fbf.drk.de/fileadmin/Content/Manual_FbF/01_Manual/01_01_Fbf_strategy_phase1.pdf, b.

IFRC: Manual Forecast-based Financing SDG, URL http://fbf.drk.de/fileadmin/Content/Manual_FbF/01_Manual/01_02_FbF_and_SFDRR_SDGs_PA.pdf, c.

IFRC: Early warning Early action, Tech. rep., IFRC, Geneva, URL http://www.ifrc.org/Global/Publications/disasters/ew-ea-2008.pdf, 2008.

Jordan, C. F.: Derivation of Leaf-Area Index from Quality of Light on the Forest Floor, Ecology, 50, 663–666, https://doi.org/10.2307/1936256, URL http://doi.wiley.com/10.2307/1936256, 1969.

Jury, M. R.: Malawi's Shire River Fluctuations and Climate, Journal of Hydrometeorology, 15, 2039–2049, https://doi.org/10.1175/JHM-D-13-0195.1, URL http://journals.ametsoc.org/doi/abs/10.1175/JHM-D-13-0195.1, 2014.

Kaštrun, T.: Dataheroes azure machine learning algorithm flowchart, URL http://www.dataheroes.nl/2017/07/21/azure-machine-learning-algoritme-flowchart/, 2017.

Keller, N., Cokely, E. T., Katsikopoulos, K. V., and Wegwarth, O.: Naturalistic Heuristics for Decision Making, Journal of Cognitive Engineering and Decision Making, 4, 256–274, https://doi.org/10.1518/155534310X12844000801168, URL http://journals.sagepub.com/doi/10.1518/155534310X12844000801168, 2010.

Khamis, M.: Climate change and smallholder farmers in Malawi, Action Aid, pp. 1–8, 2006.

Kita, S. M.: "Government Doesn't Have the Muscle": State, NGOs, Local Politics, and Disaster Risk Governance in Malawi, Risk, Hazards & Crisis in Public Policy, 8, 244–267, https://doi.org/10.1002/rhc3.12118, URL http://doi.wiley.com/10.1002/rhc3.12118, 2017.

Kogan, F.: Application of vegetation index and brightness temperature for drought detection, Advances in Space Research, 15, 91–100, https://doi.org/10.1016/0273-1177(95)00079-T, URL https://www.sciencedirect.com/science/article/pii/027311779500079T, 1995.

KOGAN, F. N.: Remote sensing of weather impacts on vegetation in non-homogeneous areas, International Journal of Remote Sensing, 11, 1405–1419, https://doi.org/10.1080/01431169008955102, URL https://www.tandfonline.com/doi/full/10.1080/01431169008955102, 1990.

Kundhlande, G., Franzel, S., Centre, W. A., and Simpson, B.: Farmer-to-farmer extension approaches in Malawi : a survey of Farmer-to-Farmer Extension Approach in Malawi : A Survey of Organizations Godfrey Kundhlande , Steven Franzel , Brent Simpson and Emma Gausi, 2015.

Liu, Y. Y., De Jeu, R. A. M., McCabe, M. F., Evans, J. P., and Van Dijk, A. I. J. M.: Global long-term passive microwave satellite-based retrievals of vegetation optical depth, Geophysical Research Letters, 38, 1–6, https://doi.org/10.1029/2011GL048684, 2011.

Lopez, A., Coughlan de Perez, E., Bazo, J., Suarez, P., van den Hurk, B., and van Aalst, M.: Bridging forecast verification and humanitarian decisions: A valuation approach for setting up action-oriented early warnings, Weather and Climate Extremes, pp. 1–8, https://doi.org/10.1016/j.wace.2018.03.006, URL https://doi.org/10.1016/j.wace.2018.03.006, 2018.

Lowcock, M.: Global Humanitarian Overview 2018, Tech. rep., UN OCHA, URL https://www.unocha.org/sites/unocha/files/GHO2018.PDF, 2018.

Martignon, L., Vitouch, O., Takezawa, M., and Forster, M. R.: Naive and yet Enlightened: From Natural Frequencies to Fast and Frugal Decision Trees, Tech. rep., URL http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.412.6594&rep=rep1&type=pdf.

Martignon, L., Katsikopoulos, K. V., and Woike, J. K.: Categorization with limited resources: A family of simple heuristics, Journal of Mathematical Psychology, 52, 352–361, https://doi.org/10.1016/J.JMP.2008.04.003, URL https://www.sciencedirect.com/science/article/pii/S0022249608000370, 2008.

Martínez-Fernández, J., González-Zamora, A., Sánchez, N., Gumuzzio, A., and Herrero-Jiménez, C.: 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, URL https://www.sciencedirect.com/science/article/pii/S0034425716300931, 2016.

Mckee, T. B., Doesken, N. J., and Kleist, J.: THE RELATIONSHIP OF DROUGHT FREQUENCY AND DURATION TO TIME SCALES, Tech. rep., URL http://www.droughtmanagement.info/literature/AMS_Relationship_Drought_Frequency_Duration_Time_Scales_1993.pdf, 1993.

Mcsweeney, C., New, M., and Lizcano, G.: UNDP Climate Change Country Profiles Malawi General Climate, pp. 1–27, URL http://country-profiles.geog.ox.ac.uk.

Mishra, A. K. and Singh, V. P.: A review of drought concepts, Journal of Hydrology, 391, 202–216, https://doi.org/10.1016/j.jhydrol.2010.07.012, URL http://dx.doi.org/10.1016/j.jhydrol.2010.07.012, 2010.

MoAIWD: Shire River Basin Management Program, Tech. Rep. June, MoAIWD, Lilongwe, 2016.

MOAIWD: Shire Basin Operational Decision Support System through enhanced Hydro-meteorological Services, Tech. rep., DHI, Horsholm, 2018.

Murphy, A. H.: What is a good forecast? an essay on the nature of foodness in weather forecasting, Weather and Forecasting, pp. 281–293, https://doi.org/https://doi.org/10.1175/1520-0434(1993)008<0281:WIAGFA>2.0.CO;2, 1993.

Murphy, A. H. and Winkler, R. L.: A General Framework for Forecast Verification, Monthly Weather Review, 115, 1330–1338, https://doi.org/10.1175/1520-0493(1987)115<1330:AGFFFV>2.0.CO;2, URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0493%281987%29115%3C1330%3AAGFFFV%3E2.0.CO%3B2, 1987.

Mwangi, E., Wetterhall, F., Dutra, E., Di Giuseppe, F., and Pappenberger, F.: Forecasting droughts in East Africa, Hydrology and Earth System Sciences, 18, 611–620, https://doi.org/10.5194/hess-18-611-2014, 2014.

Mylne, K. R.: Decision-making from probability forecasts based on forecast value, Meteorological Applications, 9, 307–315, https://doi.org/10.1017/S1350482702003043, 2002.

NDMC: Types of Drought, URL https://drought.unl.edu/Education/DroughtIn-depth/TypesofDrought.aspx.

Neth, H. and Gigerenzer, G.: Heuristics: Tools for an Uncertain World, in: Emerging Trends in the Social and Behavioral Sciences, pp. 1–18, John Wiley & Sons, Inc., Hoboken, NJ, USA, https://doi.org/10.1002/9781118900772.etrds0394, URL http://doi.wiley.com/10.1002/9781118900772.etrds0394, 2015.

Ngongondo, C., Xu, C.-Y., Gottschalk, L., and Alemaw, B.: Evaluation of spatial and temporal characteristics of rainfall in Malawi: a case of data scarce region, Theoretical and Applied Climatology, 106, 79–93, https://doi.org/10.1007/s00704-011-0413-0, URL http://link.springer.com/10.1007/s00704-011-0413-0, 2011.

NOAA: Glossary of Forecast Verification Metrics, Tech. rep., URL http://www.nws.noaa.gov/oh/rfcdev/docs/Glossary_Forecast_Verification_Metrics.pdf, 2012.

NOAA's Climate Prediction: NOAA's Climate Prediction Center, URL https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php.

Oreopoulos, L.: Water Cycle and Precipitation, URL https://atmospheres.gsfc.nasa.gov/climate/index.php?section=135.

Palmer, W. C.: Meteorological Drought, Tech. rep., US Weather Bureau, Washington, DC, 1965.

Palmer, W. C.: Keeping Track of Crop Moisture Conditions, Nationwide: The New Crop Moisture Index, Weatherwise, 21, 156–161, https://doi.org/10.1080/00431672.1968.9932814, URL http://www.tandfonline.com/doi/abs/10.1080/00431672.1968.9932814, 1968.

Phillips, N. D., Neth, H., Woike, J. K., and Gaissmaier, W.: FFTrees: A toolbox to create, visualize, and evaluate fast-and-frugal decision trees, Judgment and Decision Making, 12, 344–368, URL http://journal.sjdm.org/17/17217/jdm17217.pdf, 2017.

Potts, J. M.: Forecast Verification: A Practitioner's Guide in Atmospheric Science, Second Edition, Wiley-Blackwell, Sussex, second edn., https://doi.org/10.1002/9781119960003, URL http://danida.vnu.edu.vn/cpis/files/Books/Forecast%20Verification%20-%20A%20Practitioners%20Guide%20in%20Atmospheric%20Science.pdf, 2012.

Rojas, O.: Country-Level ASIS: an agricultural drought monitoring system, Tech. rep., FAO, URL http://www.fao.org/.

Rosso, C. C., Hochrainer-Stigler, S., Pflug, G., Condori, B., and Berndtsson, R.: Early warning and drought risk assessment for the Bolivian Altiplano agriculture using high resolution satellite imagery data, Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2018-133, URL https://www.nat-hazards-earth-syst-sci-discuss.net/nhess-2018-133/nhess-2018-133.pdf, 2018.

Rüth, A.: Framework: Action Plan of the Federal Foreign Office for Humanitarian Adaptation to Climate Change, URL http://www.drk.de/fileadmin/Presse/Downloads/Bild/Ausland/GRC__2015__FFO_Action_Plan_CCA.pdf, 2015.

Rüth, A., Devonna, K., Coughlan, E., Kelly, T., Jaime, C., Bailey, M., and Lux, S.: Forcast-based Financing: A policy overview, Tech. rep., DRK, 2017a.

Rüth, A., Lux, S., and Schölz, S.: Forecast-Based Financing An innovative approach, Tech. rep., German Red Cross, Berlin, 2017b.

Scaini, A., Sánchez, N., Vicente-Serrano, S. M., and Martínez-Fernández, J.: SMOS-derived soil moisture anomalies and drought indices: A comparative analysis using in situ measurements, Hydrological Processes, 29, 373–383, https://doi.org/10.1002/hyp.10150, 2015.

Seibert, M., Merz, B., and Apel, H.: Seasonal forecasting of hydrological drought in the Limpopo Basin: a comparison of statistical methods, Hydrol. Earth Syst. Sci, 21, 1611–1629, https://doi.org/10.5194/hess-21-1611-2017, URL www.hydrol-earth-syst-sci.net/21/1611/2017/, 2017.

Shafer, B. and Dezman, L.: Development of a surface water supply index (SWSI) to assess the severity of drought conditions in snowpack runoff areas | Western Snow Conference, in: Proceedings of the 50th Annual Western Snow Conference, pp. 164–175, Western Snow Conference, Colorado State University, URL https://westernsnowconference.org/node/932, 1982.

Stephens, E., Coughlan de Perez, Erin Kruczkiewicz, A., Boyd, E., and Suarez, P.: Forecast- based action, Red Cross, p. 37, 2015.

Stojic, M.: Disaster Management: Staying Ahead of the Storm « Earth Imaging Journal: Remote Sensing, Satellite Images, Satellite Imagery, URL http://eijournal.com/print/articles/disaster-management, 2013.

Svoboda, M. and Fuchs, B.: Handbook of Drought Indicators and Indices*, 1173, https://doi.org/10.1201/9781315265551-12, URL http://www.crcnetbase.com/doi/10.1201/9781315265551-12, 2017.

Tarpley, J. D., Schneider, S. R., and Money, R. L.: Global Vegetation Indices from the NOAA-7 Meteorological Satellite, Journal of Climate and Applied Meteorology, 23, 491–494, https://doi.org/10.1175/1520-0450(1984)023<0491:GVIFTN>2.0.CO;2, URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0450%281984%29023%3C0491%3AGVIFTN%3E2.0.CO%3B2, 1984.

The World Bank: Poverty & Equity Data Portal, URL http://povertydata.worldbank.org/poverty/country/MWI.

UNISDR: UNISDR Terminology on Disaster Risk Reduction, Tech. rep., UNISDR, Geneva, URL https://www.unisdr.org/files/7817_UNISDRTerminologyEnglish.pdf, 2009.

Verdin, J. and Klaver, R.: Grid-cell-based crop water accounting for the famine early warning system, Hydrological Processes, 16, 1617–1630, https://doi.org/10.1002/hyp.1025, URL http://doi.wiley.com/10.1002/hyp.1025, 2002.

Vicente-Serrano, S. M., Beguería, S., and López-Moreno, J. I.: A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index, Journal of Climate, 23, 1696–1718, https://doi.org/10.1175/2009JCLI2909.1, URL http://journals.ametsoc.org/doi/abs/10.1175/2009JCLI2909.1, 2010.

Vincent, K., Dougill, A. J., Mkwambisi, D. D., Cull, T., Stringer, L. C., and Chanika, D.: Analysis of Existing Weather and Climate Information for Malawi, URL http://www.futureclimateafrica.org/wp-content/uploads/2015/02/Malawi-existing-Analysis-of-existing-weather-and-climate-information-for-Malawi-final-with-disclaimer.pdf, 2014.

Wahlström, M.: Drought Risk Reduction Framework and Practices: Contributing to the Implementation of the Hyogo Framework for Action Photo cover: Drip system in Cauliflower in Churiya, Makawanpur, Nepal. Boosting Poor Households' Crop Production with Drip and Sprinkler, URL www.unisdr.orgwww.preventionweb.net, 2009.

Wanders, N., van Lanen, H. A. J., and Van Loon, A. F.: Indicators for drought characterization on a global scale, Tech. rep., Wageningen, URL http://edepot.wur.nl/160049, 2010.

Wetterhall, F., Winsemius, H. C., Dutra, E., Werner, M., and Pappenberger, E.: Seasonal predictions of agro-meteorological drought indicators for the Limpopo basin, Hydrology and Earth System Sciences, 19, 2577–2586, https://doi.org/10.5194/hess-19-2577-2015, 2015.

WFP: FoodSECuRE-Food Security Climate Resilience Facility Supporting community resilience-building before and after climatic shocks World Food Programme, Tech. rep., URL https://documents.wfp.org/stellent/groups/public/documents/communications/wfp279583.pdf?_ga=2.35076257.1096149526.1549441358-867894524.1549441358, 2016.

Winsemius, H. C., Dutra, E., Engelbrecht, F. A., Archer, E., Garderen, V., Wetterhall, F., Pappenberger, F., and Werner, M. G. F.: The potential value of seasonal forecasts in a changing climate in southern Africa, Hydrol. Earth Syst. Sci, 18, 1525–1538, https://doi.org/10.5194/hess-18-1525-2014, URL www.hydrol-earth-syst-sci.net/18/1525/2014/, 2014.

WMO: Drought monitoring and early warning: concepts, progress and future challenges Weather and climate information for sustainable agricultural development, Tech. rep., URL http://www.wamis.org/agm/pubs/brochures/WMO1006e.pdf, 2006.

Wolter, K. and Timlin, M. S.: El Niño/Southern Oscillation behaviour since 1871 as diagnosed in an extended multivariate ENSO index (MEI.ext), International Journal of Climatology, 31, 1074–1087, https://doi.org/10.1002/joc.2336, URL http://doi.wiley.com/10.1002/joc.2336, 2011.

Wong, G., Van Lanen, H. A. J., and Torfs, P. J. J. F.: Probabilistic analysis of hydrological drought characteristics using meteorological drought, Hydrological Sciences Journal – Journal des Sciences Hydrologiques Hydrological Sciences Journal, 58, 253–270, https://doi.org/10.1080/02626667.2012.753147, URL http://dx.doi.org/10.1080/02626667.2012.753147, 2013.

Xue, J. and Su, B.: Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications, Journal of Sensors, 2017, 1–17, https://doi.org/10.1155/2017/1353691, URL https://www.hindawi.com/journals/js/2017/1353691/, 2017.

Zargar, A., Sadiq, R., Naser, B., and Khan, F. I.: A review of drought indices, https://doi.org/10.1139/A11-013, URL www.nrcresearchpress.com.