

Student thesis series INES nr 528

Behind the early warning: Improving impact-based forecasting of riverine floods in Malawi using passive microwave remote sensing

Lone C. Mokkenstorm

2020

Department of

Physical Geography and Ecosystem Science

Lund University

Sölvegatan 12



Lone C. Mokkenstorm (2020)

Behind the early warning: Improving impact-based forecasting of riverine floods in Malawi using passive microwave remote sensing

Master degree thesis, 30 credits in *Physical Geography and Ecosystem Science*
Department of Physical Geography and Ecosystem Science, Lund University

Level: Master of Science (MSc)

Course duration: *January 2020* until *June 2020*

Disclaimer:

This document describes work undertaken as part of a program of study at the University of Lund. All views and opinions expressed herein remain the sole responsibility of the author, and do not necessarily represent those of the institute.

Behind the early warning: Improving impact-based forecasting of riverine floods in Malawi using passive microwave remote sensing

Lone C. Mokkenstorm

Master thesis, 30 credits, in *Physical Geography and Ecosystem Science*

Supervisors:

Andreas Persson

Department of Physical Geography and Ecosystem Science

Marc van den Homberg

510 – an initiative of The Netherlands Red Cross

Exam committee:

Zheng Duan, Department of Physical Geography and Ecosystem Science

David E. Tenenbaum, Department of Physical Geography and Ecosystem Science

Acknowledgements

This thesis was carried out at 510, an initiative of The Netherlands Red Cross, in the context of the Hazard Monitoring and Warning work package of the ECHO V project by the Danish Red Cross (lead), the British Red Cross, the Malawi Red Cross and 510. The document was submitted to Lund University in completion of an MSc in Physical Geography and Ecosystem Science.

The project would not have been possible without my supervisors, Andreas Persson (LU) and Marc van den Homberg (510). Both of your dedication, enthusiasm and endless support provided the best guidance possible to make this project a success. My sincere thanks to the Malawi Data Team as well for arranging additional data from the Malawi Department of Water Resources, and to the Global Runoff Data Centre for providing the Karonga discharge data. I would also like to express my gratitude to Hessel Winsemius, who brought my attention to this interesting topic, and Oscar Keunen, for always being there to exchange and discuss ideas.

On a personal note, I would like to thank my parents, Bianca and Jan Mokkenstorm, my brothers Mats and Niels Mokkenstorm -special thanks for the occasional Python-troubleshooting- and my fiancée Robert Hoekman for their unconditional support. Lastly, I would like to thank my fellow students from both Lund University and 510: It was a challenging time to be graduating, but the 'Fika Fridays' in Lund and socially-distanced study sessions in The Hague certainly made the road seem a lot easier.

Abstract

This thesis investigates whether freely available, coarse-resolution, Passive Microwave Remote Sensing (PMRS) data (37 GHz) can be effectively used for early warning systems for floods in Malawi. The Shire River Basin in Chikwawa and the smaller-scale North Rukuru River Basin in Karonga were studied using two alternative, ratio-based satellite indices that make use of the signal difference between wet and dry pixel cells: The m index and the r_{cmc} index. The m index is directly related to the r_{cm} index, introduced by Brakenridge et al. (2007), and divides the brightness temperature in a relatively stable, dry calibration cell by the brightness temperature measured in the cell with the river. The r_{cmc} index is an adaptation of this, and uses an additional wet calibration cell. It was investigated whether these indices could aid in the detection and forecasting of flood events and their magnitude. The findings pertaining to detection skill showed that at both study sites, r_{cmc} and m detected a similar seasonality to the observed discharge hydrographs, as long as the downstream virtual gauging station was located at a sufficient distance from a large water body. A regression analysis showed that the indices' relationship with observed discharge had a moderately strong, positive correlation in Chikwawa, but not in Karonga. Flood occurrence detection skill was assessed using an impact database. A flood threshold corresponding to a return period of 5 years was determined to see if the indices could simulate historical flood events. Both indices did not detect the majority of registered floods, which is likely a consequence of the method used to determine the trigger threshold. There were no upstream virtual gauging stations present that had a sufficient lag time with the downstream satellite signal. A possible forecasting system using merely the downstream satellite signals was shown to have a sufficient accuracy at a lead time of up to nearly 3 days, although in an operational setting, the forecasts would not reach the calculated trigger threshold at this lead time. Overall, the PMRS-model showed a better performance in Chikwawa when compared to the global runoff model GloFAS. As it also does not require extensive input data when used as an Early Warning System (EWS), as many smaller-scale EWS do, we suggest that when perfected, the PMRS-method is implemented in a coupled EWS solution, including a PMRS-model, a global forecasting model and a more detailed national model. This would offer early warnings in data-scarce regions and at a variety of lead times. In order for this to be effective, we suggest that more research be done on correctly setting the trigger threshold, and into the potential spatial interpretation of r_{cmc} .

Keywords: Physical Geography, Ecosystem Analysis, Early Warning, Riverine Floods, Humanitarian Work, Impact-based Forecasting, Malawi

Popular Summary

Early warning systems can support humanitarian operations by forecasting hazard impacts and aid to release funds before disasters take place. One challenge humanitarian organizations currently face, however, is a lack of historical and real-time data to set up such a system. Whereas optical satellite remote sensing data has been used in the past to address these issues, cloud cover and infrequent satellite overpasses often yield this data suboptimal for this purpose. This study therefore investigates whether openly available, coarse-resolution, passive microwave satellite data –which is less impeded by cloud-cover and is measured on a daily basis– can be effectively used for early warning systems for floods in the Shire River and the North Rukuru River in Malawi. Two alternative indices were calculated from the raw data and further studied.

Flood Detection

Firstly, the potential for detecting both river discharge and individual flood events was assessed. At both study sites, the indices detected a seasonality similar to that of the observed discharge, as long as the satellite data pixel studied was sufficiently far from a large water body. The indices were moderately correlated to discharge in the Shire River, but not in the North Rukuru River. A comparison of the satellite data to an impact database with flood events showed that the indices did not detect the majority of registered floods, although this is likely a consequence of the method used to determine the threshold of what ‘counts’ as a flood.

Flood forecasting

Satellite data from the upstream part of the river catchment could unfortunately not be used to forecast the satellite signal and hence the flood events downstream, as the correlation between the two was strongest without a time shift. If it would have been stronger with a time shift of at least one day, this would add time to the window available to give out early warnings. Therefore, a forecasting model was set up using *just* the satellite data from the downstream point of interest. Statistically, this model showed to have sufficient accuracy with a lead time up to nearly three days, but a test conducted with a historical flood event showed that this model would, in practice, not have triggered an early warning before the flood event happened. Overall, the passive microwave system presented in this thesis had a stronger relationship with discharge in the study area than the existing global runoff model GloFAS does. As it also does not require extensive input data when used as an early warning system, we suggest that when perfected, the PMRS-method is implemented in a coupled solution, including a remote sensing-model, a global forecasting model and a more detailed national model. The use of these systems could offer early warnings in data-scarce regions and at a variety of lead times, which has the potential to make humanitarian aid in response to floods faster and more efficient.

Table of Contents

1. Introduction	1
1.1 Flood detection and remote sensing	1
1.2 Relevance, thesis aim & research questions	2
2. Background	3
2.1 Riverine floods	3
2.1 Forecast-based financing and impact-based forecasting	3
2.2 Early warning systems	4
2.3 Remote sensing	6
2.4 Flood detection using passive microwave remote sensing	8
3. Methods	11
3.1 Case study: Malawi	11
3.2 Virtual gauging stations	14
3.3 Data acquisition and treatment	16
3.4 Relationship with discharge	19
3.5 Relationship with flood occurrence and timing	20
3.6 Assessment of ability of prediction	22
4. Results	25
4.1 Observed satellite signals	25
4.2 Detection of discharge	27
4.3 Detection of flood occurrence	30
4.4 Relationship between upstream and downstream satellite signals	34
4.5 Forecasting model	35
5. Discussion	39
5.1 Raw T_b signals	39
5.2 Flood detection	40
5.3 Flood forecasting	43
5.4 Comparison with existing early warning systems	45
5.5 Assumptions and limitations of research	46
6. Conclusion	49
6.1 Detection potential	49
6.2 Forecasting potential	49
6.3 Implications for humanitarian work	50
7. References	51

List of Figures

Figure 1: Phases of the disaster management cycle. Source: 510 an initiative of the Netherlands Red Cross (2020b), used with permission 3

Figure 2: Schematic overview of the relationship between early warning systems, triggers and early action protocols in the context of Forecast-based Financing (FbF) / Impact-based Forecasting 5

Figure 3: Simplified electromagnetic spectrum, including identified regions and atmospheric windows. Source: NASA (2008), public domain 7

Figure 4: Schematic overview of the respective locations of the cells/pixels used in this research to calculate the indices, including wet calibration cell (C_w), the dry calibration cell (C_d) and the measurement cell containing the river (M) 9

Figure 5: Overview map of Malawi. Administrative units and districts studied in this thesis, namely Chikwawa in the southwest and Karonga in the north..... 12

Figure 6: Overview map of Malawi: Topography and Hydrology 13

Figure 7: Upstream and downstream Virtual Gauging Stations (VGSs) along the Shire and North Rukuru River. One grid cell is approximately 25 x 25 kilometres. The labels on the cells are used for referencing them throughout this thesis, with the letter corresponding to the district Chikwawa (C) or Karonga (K)..... 15

Figure 8: Illustration of the effect of the mean-filter method in Chikwawa in 2015 compared to a 15-day, 3rd order Savitzky-Golay on the brightness temperature (T_b)-signal from the measurement cell with the river (M), the dry calibration cell (C_d) and the wet calibration cell (C_w). Data source: Brodzik (2016)..... 17

Figure 9: Schematic representation of the use and processing of raw brightness temperature (T_b) data to T_b -series in the wet calibration (T_{bCw}), dry calibration (T_{bCd}) and measurement cell (T_{bM}), eventually calculated into the CM- (r_{cm}) and CMC-ratio (r_{cmc}) and the m index (m)..... 19

Figure 10: Conceptual overview of the model including processing steps. sm = submodels, f_s = stationary forecast, f = final forecast, t = point in time that represents ‘present’ when forecasting..... 25

Figure 11: Locations of the measurement cells (M , green), dry calibration cells (C_d , red) and wet calibration cells (C_w , black) in the North Rukuru and Shire watersheds. The dots present the upper left corner of the T_b -cell in question. There are more C_d -cells than M -cells, because the method used to determine C_d 's location allows it to change each year. Data source: Brodzik (2016)..... 26

Figure 12: Brightness temperature (T_b) values in Kelvin (K) of the wet calibration (C_w , blue line), dry calibration (C_d , green line) and measurement (M , orange line) cells in the downstream virtual gauging station in Chikwawa (cell C0). Data source: Brodzik (2016)..... 27

Figure 13: Brightness temperature (T_b) values in Kelvin (K) of the wet calibration (C_w , blue line), dry calibration (C_d , green line) and measurement (M , orange line) cells in the downstream virtual gauging station in Karonga (cell K0). Data source: Brodzik (2016)..... 27

Figure 14: Observed discharge and downstream (cell C0) satellite signal at Chikwawa. r_{cmc} = CMC-ratio, m = m index. Data sources: Brodzik (2016); National Hydrological Services of Malawi (2019) 28

Figure 15: Mean monthly discharge and downstream (cell C0) satellite signals for Chikwawa. r_{cmc} = CMC-ratio, m = m index. Data sources: Brodzik (2016); National Hydrological Services of Malawi (2019)..... 28

Figure 16: Observed discharge and downstream (cell K0) satellite signals at Karonga. r_{cmc} = CMC-ratio, m = m index. Data sources: Brodzik (2016); Global Runoff Data Centre (2020)..... 28

Figure 17: Mean monthly discharge and downstream (cell K0) satellite signal for Karonga, including data from cell K2 as this will be used as a proxy for K0 due to the proximity of Lake Malawi to K0. r_{cmc} = CMC-ratio, m = m index. Data sources: Brodzik (2016); Global Runoff Data Centre (2020)..... 28

Figure 18: Spearman’s correlation (ρ) and regression results of satellite signals and discharge for Chikwawa (C0-cell). r_{cmc} = CMC-ratio, m = m index. Data sources: Brodzik (2016); National Hydrological Services of Malawi (2019) 30

Figure 19: Spearman’s correlation (ρ) and regression results of satellite signals and discharge for Karonga (K2-cell). r_{cmc} = CMC-ratio, m = m index. Data sources: Brodzik (2016); Global Runoff Data Centre (2020)..... 30

Figure 20: Peaks of the CMC-ratio (r_{cmc}), m index (m) and discharge per hydrological year vs. their respective return periods in years for Chikwawa and Karonga. The curve in red is a 10th degree polynomial curve through the data points. Data source: Brodzik (2016). 31

Figure 21: Discharge in wet season measured at the Chikwawa gauging station and registered floods in Chikwawa, green lines with * denote floods with a return period (rp) of 5 years or higher. Data sources: National Hydrological Services of Malawi (2019); Dartmouth Flood Observatory (2020); GLIDE (2020); Guha-Sapir (2020)..... 31

Figure 22: Satellite signals in wet season and registered floods with a return period (rp) of 5 years or more in Chikwawa (cell C0). r_{cmc} = CMC-ratio, m = m index. Data sources: Brodzik (2016); Dartmouth Flood Observatory (2020); GLIDE (2020); Guha-Sapir (2020)..... 32

Figure 23: Satellite signals in wet season and registered floods. r_{cmc} = CMC-ratio, m = m index. Registered flood return periods could not be calculated due to the absence of discharge data. Data sources: Brodzik (2016); Dartmouth Flood Observatory (2020); GLIDE (2020); Guha-Sapir (2020) 32

Figure 24: Discharge as simulated by the Global Flood Awareness System (GloFAS) and peaks per hydrological year vs. return periods for Chikwawa (cell C0). Data source: Harrigan et al. (2020)..... 33

Figure 25: Non-parametric time-lagged cross correlation of upstream (cells K3-K4; cells C1-C12) and downstream (cell K2; cell C0) m along the North Rukuru (left) and Shire River (right). Correlations have been performed on stationary data..... 34

Figure 26: Non-parametric time-lagged cross correlation of upstream (cells K3-K4; cells C1-C12) and downstream (cell K2; cell C0) CMC-ratio (r_{cmc}) along the North Rukuru (left) and Shire River (right). Correlations have been performed on stationary data..... 34

Figure 27: Performance of the submodels (predictions vs. test observations) for the m index (m , bottom) and the CMC-ratio (r_{cmc} , top) in Chikwawa (A) and Karonga (B). RMSE = root mean squared error.. 36

Figure 28: Main model forecast performance for Chikwawa in the 2015 wet season. A: Satellite signal forecasts vs. observed signals. B: R^2 vs. lead time (days). C: RMSE versus lead time (days). l_t = lead time, r_{cmc} = CMC-ratio, RMSE = Root Mean Squared Error 36

Figure 29: Main model forecast performance for Karonga in the 2015 wet season. A: Satellite signal forecasts vs. observed signals. B: R^2 vs. lead time (days). C: RMSE versus lead time (days). l_t = lead time, r_{cmc} = CMC-ratio, RMSE = Root Mean Squared Error 37

Figure 30: Performance of model in forecasting exceedance of the trigger threshold during the 2015 floods. rp = return period, r_{cmc} = CMC-ratio 38

List of Tables

Table 1: Overview of regions on the electromagnetic spectrum used for remote sensing observations, including their wavelengths and abbreviation. Data source: Chuvieco (2016)..... 6

Table 2: Gauge data availability for Chikwawa and Karonga (% of total days with data per year). Red means no data was available, dark green means all data was available (linear coloring scale) 18

Table 3: Confusion matrix example. In the context of this study, ‘Observed’ refers to events reported in the impact database, and ‘Modelled’ refers to events simulated by exceedances of the trigger threshold. 21

Table 4: Confusion matrix for the satellite signals at cells K2 and C0. 1 entry corresponds to 1 flood season between 1978 and 2017. ‘Observed’ refers to events reported in the impact database, and ‘Modelled’ refers to events simulated by exceedances of the trigger threshold. r_{cmc} = CMC-ratio, m = m index, rp = return period. 33

Table 5: Success metrics for the satellite signals at cells C0 and K2. r_{cmc} = CMC-ratio, m = m index, rp = return period, POD = Probability of Detection, FAR = False Alarm Ratio, CSI = Critical Success Index..... 33

Table 6: Confusion matrix for the Global Flood Awareness System (GloFAS) at cell C0 in Chikwawa. 1 entry corresponds to 1 flood season between 2000 and 2017. ‘Observed’ refers to events reported in the impact database, and ‘Modelled’ refers to events simulated by exceedances of the trigger threshold. rp = return period. 34

Table 7: Results of the non-parametric time-lagged cross correlation between satellite and discharge data for the Shire (C) (1979-2009) and North Rukuru (K*) (1979-1990). r_{cmc} = CMC-ratio, m = m index, ρ = Spearman’s correlation coefficient. The optimal lag time is the lag time at which ρ is highest 35*

Table 8: Summary of relationships between passive microwave remote sensing indices and observed discharge found in existing literature. The upper section shows regressions, the lower correlations. (obs) Q = (observed) discharge, sim Q = simulated discharge, r_{cm} = CM-ratio, r = Pearson’s correlation coefficient, ρ = Spearman’s correlation coefficient..... 42

List of abbreviations and symbols

C_d	Dry calibration cell
C_w	Wet calibration cell
EM	Electromagnetic
EWS	Early Warning System
FbA	Forecast-based Action
FbF	Forecast-based Financing
GloFAS	Global Flood Awareness System
IbF	Impact-based Forecasting
IFRC	International Federation of Red Cross Red Crescent Societies
M	Measurement cell
m	Flood magnitude, relative r_{cm}
MW	Microwave
ODSS	Operational Decision Support System
PMRS	Passive Microwave Remote Sensing
r_{cm}	CM-ratio
r_{cmc}	CMC-ratio
rp	Return Period
SSA	Sub-Saharan Africa
T_b	Brightness temperature
VGS	Virtual Gauging Station

Image sources and permissions

All figures in this report are own work, unless stated otherwise. If raw data have been used to create the graphs, the source of these data is stated in the image caption. Figure 1 was created by 510, an initiative of The Netherlands Red Cross, and permission has been obtained to use it. Figure 3 is made by (NASA 2008), which is part of the public domain and has been used in this thesis in compliance with the NASA media guidelines: news outlets, schools, and text-book authors may use NASA content without needing explicit permission.

1. Introduction

Natural hazards have a tremendous impact on communities across the world. They often disrupt the functioning of a society and cause human, material, economic or environmental losses that the community or society cannot cope with using its own resources. When this is the case, the hazard in question is considered a disaster (IFRC 2020b). Whether or not a hazard results into a disaster depends on the severity of the hazard itself, the vulnerability of the society or community at risk, and their capacity, i.e. the ability or inability to reduce potential negative effects (IFRC 2020b).

Of all hazards, floods are among the most frequent and globally widespread (Wallemacq et al. 2015; Vinet 2018). Between 1995 and 2015, flooding accounted for 47% of all weather-related disasters, affecting 2.3 billion people globally (Wallemacq et al. 2015). The number of floods occurring annually has also been rising in recent years (Wallemacq et al. 2015), a trend which is likely to persist in the light of climate change (Chidanti-Malunga 2011; IPCC 2014). Simultaneously, floods have been causing more economic, material, and human losses due to the rapid growth of the global population and economy in recent years. Consequentially, they are regarded as disasters more frequently as well (Hussain et al. 2005).

One region that is increasingly affected by weather-related disasters is Sub-Saharan Africa (SSA). Economies here are largely dependent on agriculture, leading to a substantial economic dependency on weather and hydrology. This is amplified by the fact that only a fraction of cultivated land in SSA is irrigated (Svendsen et al. 2009). Rain-fed agriculture is very vulnerable to extreme weather events and related natural hazards, such as floods and droughts, which cause severe damage to crops and hence the livelihoods of local communities (Chidanti-Malunga 2011). An example of a rain-fed and flood-prone economy in SSA is Malawi (Chidanti-Malunga 2011). A series of devastating floods in January 2015 affected more than one million people in the country, with 230,000 people displaced, 172 reported missing, and 170 reported fatalities (Guha-Sapir 2020). As extreme events like this are expected to happen more frequently in the near future (Mijoni and Izadkhah 2009; Macharia et al. 2015), flood risk will increase as well, unless drastic disaster management measures are taken (Šakić Trogrlić et al. 2018).

1.1 Flood detection and remote sensing

Continuous monitoring and detailed forecast systems are needed to increase preparedness for floods and take humanitarian action before disasters strike. Unfortunately, there are many data-sparse areas in Africa (Hasan et al. 2019) and Malawi specifically (Ngongondo et al. 2011; Mwale et al. 2012), where gauge or meteorological data, which often constitute important components of these systems, are unavailable or are not of satisfactory quality. In order to account for this, alternative strategies for flood forecasting are being explored. One of these is the use of satellite remote sensing for monitoring and forecasting (De Groeve 2010; Palmer et al. 2015). The potential of optical remote sensing data is widely covered in literature (e.g. Kwak et al. (2017); Hoang et al. (2019)). However, these data are less suitable for this purpose for two reasons (D'Addabbo et al. 2018; Weintrit et al. 2018):

- Cloudiness often occurs in times of flood surges, whereas optical sensors cannot penetrate cloud cover.
- Most satellites carrying optical sensors have a relatively long return period, ranging from 1-2 days (SPOT) to several days (Sentinel, Landsat). This resolution is not ideal for preventative monitoring of floods.

Both are addressed by using passive microwave imagery rather than optical satellite imagery. The sensors that use this part of the spectrum are less limited by cloud cover, atmospheric haze or sun illumination and have a shorter return period (D'Addabbo et al. 2018), often 12 hours (SSM/I, SMMIS, AMSR-E). However,

low radiation intensity in the microwave spectrum causes the spatial resolution of the data to be relatively low, meaning the data is most suitable for analyses over larger-scale water features (Smith 1997).

One of the ways in which floods can be assessed using passive microwave remote sensing is the CM-ratio. This satellite-derived signal uses the Brightness Temperatures obtained from Ka-band passive microwave radiometry, comparing the values from a measurement cell with a dry calibration cell, where a higher value depicts higher water content in the measurement cell. This concept was first developed by Brakenridge et al. (2007). In areas with a well delineated flood plain, that becomes inundated during high water flows, increases of the ratio over time can be synchronous with discharge increases, as in-pixel water area expands (Hirpa et al. 2013). In addition to the use of a measurement and dry calibration cell, passive microwave radiometry can also forecast floods with an additional wet calibration cell using a method called the CMC-ratio (Neisingh 2018a). This helps to distinguish inundation extent from discharge and interpret the CM-ratio spatially. In contrast to the CM-ratio, the CMC-ratio has the potential to calculate the fraction of surface water in the cell that is caused by inundation. A more extensive description of the physical basis of both signals is given in Section 2.4.

1.2 Relevance, thesis aim & research questions

The applicability of the CM-ratio for flood detection and forecasting has been investigated in, for example, Namibia (De Groeve 2010), and the Brahmaputra and Ganges Watershed (Hirpa et al. 2013). However, the applicability in the context of Malawi has only been studied with commercially available, high spatial resolution Passive Microwave Remote Sensing (PMRS) data (Kramer 2018). Furthermore, the potential of introducing a wet calibration cell, hereby enhancing inundation detection, has been proposed in the MSc thesis by Neisingh (2018a) but was not further investigated in other studies. This thesis therefore aims to investigate whether openly available PMRS data of 25 km-resolution can be used for forecasting floods in Malawi. In order to achieve this aim, this thesis will focus on two calibrated variants of the PMRS-derived index r_{cm} : m , which calibrates using historical signal data, and r_{cmc} , which uses an additional wet calibration target and can be interpreted spatially. The two main research questions and sub-questions are outlined below.

1. What is the potential of PMRS-data for the detection of the occurrence and magnitude of riverine floods in downstream areas in Malawi?

- 1.1 What is the relationship between the m - and r_{cmc} -signal and observed discharge?
- 1.2 What is the relationship between m - and CMC-signal and flood timing and occurrence?

2. What is the potential of PMRS-data for forecasting downstream flood events in Malawi?

- 2.1 What is the relationship between upstream and downstream m - and r_{cmc} -signals in the different areas of interest?
- 2.2 What is the predictive accuracy of a riverine flood forecast model, possibly based on these relationships?

In order to create a flood forecasting model, the temporal and quantitative relationship between the CM-ratio at different upstream Virtual Gauging Stations (VGSs) and downstream areas of interest can be established. In order to assess predictive accuracy, this thesis will mainly focus on forecasted flood return period and exceedance of trigger levels as derived from discharge forecasts. The study will be conducted with the support of 510, an initiative of The Netherlands Red Cross, and the scope of the work is shaped to meet the goals and areas of intervention as stipulated in the the Netherlands Red Cross ECHO-III and ECHO-V projects funded by the European Civil Protection and Humanitarian Aid Operations.

2. Background

2.1 Riverine floods

Floods have several different definitions, based on either geophysical factors, economic impact or the number of people affected. In the context of this thesis, a flood is defined as an event where high water flows cause water to inundate floodplains/terrains outside of the major river channel, causing economic damage and/or affecting people's safety.

Riverine or fluvial floods, the focus of this research, are flood events that are most often caused by rainfall. Apart from precipitation, their occurrence and propagation time is influenced by watershed characteristics, such as slope and surface condition (Vinet 2018). This means floods generally have natural causes, being influenced by factors such as geomorphology, geology, vegetation, soil conditions, and weather patterns (Parker 2000). However, their magnitude and frequency is also influenced by human actions, such as the removal of forest cover, soil erosion, and the substitution of permeable surfaces with impermeable ones (Parker 2000). In Africa for example, floods mainly occur within the monsoon season, or wet season, as extreme precipitation in the upstream catchment propagates downstream (Parker 2000).

Riverine floods, in contrast to, for example, flash floods, can be forecasted in FbF due to their relatively slow onset. Flash floods, which are fast-onset events, are therefore beyond the scope of this thesis.

2.1 Forecast-based financing and impact-based forecasting

The rapid growth of the world population and global economy, new developments in hydrological modelling and geotechnology, and the previously mentioned increasing risk on extreme weather events call for new innovations and strategies across all phases of the disaster management cycle for floods and other natural hazards (Figure 1). This is the mission of the data-driven organization 510, an initiative of The Netherlands Red Cross that aims to "shape the future of humanitarian aid by converting data into understanding, and putting it in the hands of humanitarian relief workers, decision makers and people affected, so that they can better prepare for and cope with disasters and crises."

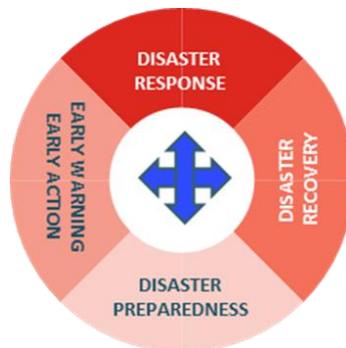


Figure 1: Phases of the disaster management cycle. Source: 510 an initiative of the Netherlands Red Cross (2020b), used with permission

One of the phases as defined by 510 is 'Early Warning Early Action' (Figure 1). Whereas this step has been widely – and in many cases, successfully - adapted in the context of, for example, cyclones (UNEP and GEAS 2014; Roy and Kovordanyi 2015; Fakhruddin and Schick 2019; Sheel et al. 2019) and heatwaves (Lowe et al. 2011), early action for riverine floods is still facing some barriers. For example, humanitarian funding is generally lacking in situations where a flood is *likely*, albeit not *certain* to strike (Coughlan de Perez et al. 2016). Furthermore, there is a high probability that a forecasted flood is not followed by

humanitarian action, due to the fact that flood forecasts, contrary to, for example, cyclone forecasts, often contain high uncertainties (Coughlan de Perez et al. 2016; Gros et al. 2019). The Forecast-Based Financing (FbF) programme is being developed by the International Federation of Red Cross and Red Crescent Societies (IFRC) to address these issues. The aim of the program is to “anticipate disasters, prevent their impact if possible, and reduce human suffering and losses” (IFRC 2020b). It is therefore a mechanism to enable access to funding for early action and preparedness for response to different types of natural hazards (Climate Centre ND). One part of FbF is Impact-based Forecasting (IbF): A process including data collection and integration, where the data is used to predict the impact of impending disasters on people in areas prone to these disasters.

In order to reach these goals, FbF, and therefore IbF, make use of pre-agreed upon automated triggers, using in-depth forecast information and risk analysis to activate humanitarian funding before disasters take place (IFRC 2020b). The triggers (also called trigger mechanisms) often take the shape of certain measurements or forecasts and automatically trigger an allocated budget when exceeded. The triggers, actions and targeting are agreed upon in advance by the humanitarian organization(s) involved, meteorological services and local communities at risk, and summarized in Early Action Protocols (EAPs), also known as Standard Operating Procedures (Wilkinson et al. 2018). All Red Crescent and Red Cross Societies that have successfully developed an EAP, will gain access to the central fund that automatically allocates resources: Forecast-based Action (FbA). EAPs are the result of collaborations between local stakeholders, the local Red Cross / Red Crescent Society in the country, and, where applicable, partner national societies.

2.1.1 Forecast-based Financing for floods

FbF-programmes exist for different kinds of hazards. For floods, only the EAPs of Peru and Bangladesh have been approved so far. However, flood-EAPs are being developed by Red Cross/Red Crescent societies in many other countries, including but not limited to Malawi, Uganda, Kenya, Zambia, Nepal, Indonesia and Eswatini (IFRC 2020a). Other organizations implementing FbF-initiatives and testing FbF-pilots include the Food and Agricultural Organization of the United Nations (FAO), the World Food Programme (WFP) and the Start Network.

There are large differences in the proposed approaches in these EAPs: Some make use of simple upstream discharge detection and local trigger modelling (e.g. Mali), whereas others make use of more complex or global hydrological models, such as the Global Flood Awareness System (GloFAS) (e.g. Peru). Bangladesh was one of the first countries to adapt an FbF-system for floods, where the triggers are based on a hydrological model that considers upstream river conditions rather than rainfall. During the monsoonal floods in June to August 2017, an activation took place in the district of Bogra, which led to a cash distribution to 1039 beneficiaries in July, which had positive effects regarding food intake, financial health and psychosocial stress at a household-level (Gros et al. 2019). A second pilot in Bangladesh was activated during the monsoon season of 2019, in the Char Islands (IFRC 2019). In Peru, FbF for floods is in place in the northern Pacific region Piura, where El Nino events often lead to flooding. The project here uses a combination of forecast models, including GloFAS (Lumbroso et al. 2017). In Uganda, two activations in 2015 and 2016 took place for floods that used GloFAS, which led to the digging of trenches and distribution of food storage bags prior to flooding (Jjemba et al. 2018).

2.2 Early warning systems

Within the framework of Early Warning Early Action and FbF/IbF, the IFRC is increasingly investing into human-centered Early Warning Systems (EWS). These are the systems that provide hazard forecasts to which the trigger thresholds are applied, and they are developed with the local network in mind. According to UNISDR (2009), an EWS is defined as the set of capacities that is necessary to generate and share

meaningful warning information to communities and organizations at risk of a hazard, in a timely manner. A human-centered EWS generally consists of four components:

1. Knowledge of risks
2. Monitoring, analysis, and forecasting of hazards
3. Communication or dissemination of alerts and warnings
4. Local capabilities to respond to the warnings received

This thesis will mainly focus on the second component, as areas of interest have already been identified by 510 and the ECHO project based on existing risks (component 1). The use of an effective EWS can be very valuable in decreasing hazard risk and supporting FbF: So-called trigger mechanisms, which are used to activate the FbF-funding, can be connected to the forecasts from an EWS (Figure 2). In the context of this thesis, the focus will mostly be on the second aspect: The monitoring, analysis, and forecasting of hazards. EWSs can be managed on different levels:

- Local EWSs can take many different shapes but are often based on a bottom-up approach involving local communities and personal networks. An example is warning dissemination from upstream river gauging stations when a certain threshold is exceeded. Sometimes, indigenous or traditional knowledge are included in these warning systems as well. Bottom-up approaches tend to better address the vulnerabilities of the communities at risk and harness participation of the vulnerable when compared to national, top-down approaches (Šakić Trogrlić et al. 2018).
- National EWSs are managed by the government and often make use of more complex hydrological models. This is a top-down approach, involving different stakeholders. An example is the Operational Decision Support System (ODSS), an EWS managed by the government that is in place in the Southern Region of Malawi (Ammentorp and Richaud 2016).
- Non-governmental organizations may be involved in the development of National or local EWSs or may set up their own EWS. These can take shape as different combinations of bottom-up and top-down approaches.

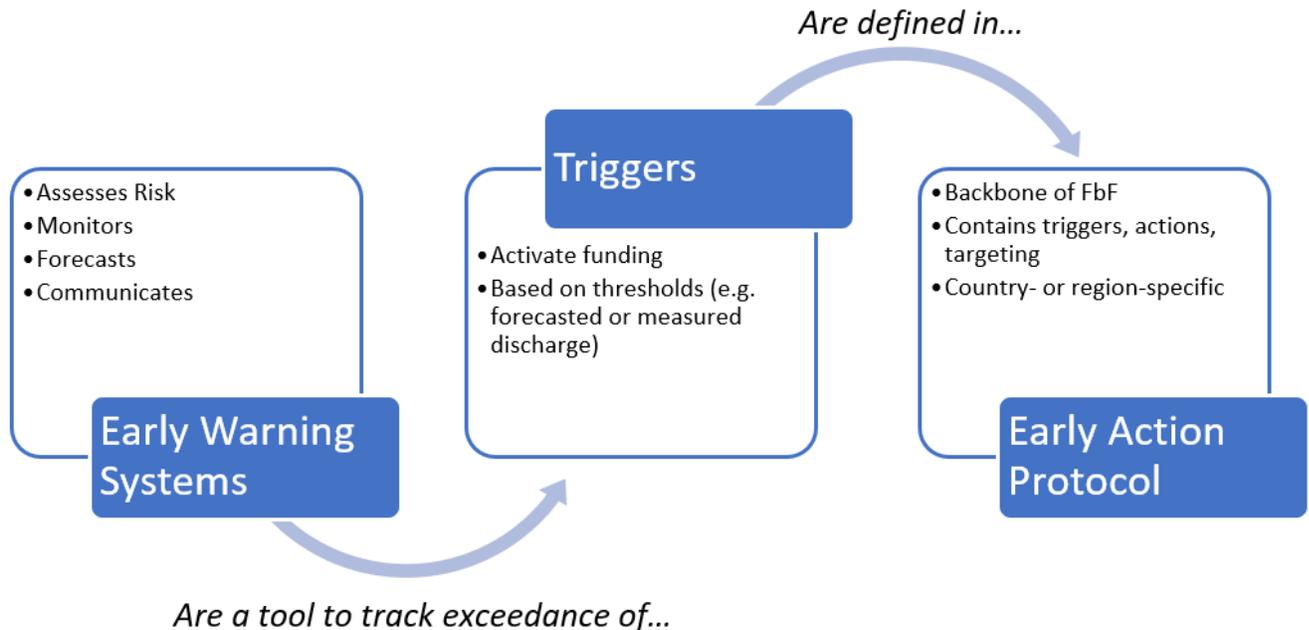


Figure 2: Schematic overview of the relationship between early warning systems, triggers and early action protocols in the context of Forecast-based Financing (FbF) / Impact-based Forecasting

2.3 Remote sensing

The EWS proposed in this thesis is based on the concept of Remote Sensing (RS). This tool uses measured electromagnetic (EM) radiation originating from features or objects, either reflected or emitted, to identify and categorize these features/objects from a distance. RS makes use of either passive sensors, that detect the reflected or emitted EM-radiation originating directly from the surface, or active sensors, that detect the response from surfaces irradiated by the sensor's energy source.

2.3.1 Physical basis of RS

The type of EM radiation (E), λ , in terms of where it can be found on the EM spectrum (see below), can be expressed in terms of frequency (f) or wavelength (λ), using the following two equations:

$$E = hcf \quad \text{Eq. 1}$$

$$E = h\left(\frac{c}{\lambda}\right) \quad \text{Eq. 2}$$

, where E stands for EM radiation, c denotes the Planck constant (6.626×10^{-34} J s) and h stands for the speed of light (3×10^8 m/s), the speed at which EM radiation is transmitted. Based on these equations, it is important to note that shorter wavelengths have higher energy content and longer wavelengths have lower energy content. The continuum of radiant energy, plotted as a function of λ or f, is called the EM spectrum. This spectrum is divided into different regions, defined with intervals of λ or f (Table 1).

Table 1: Overview of regions on the electromagnetic spectrum used for remote sensing observations, including their wavelengths and abbreviation. Data source: Chuvieco (2016)

Region		Wavelength
Visible	VIS	0.4 μm – 0.7 μm
Near/Photographic Infrared	NIR	0.7 μm – 0.9 μm
Mid-Infrared	MIR	1.2 μm – 8 μm
Thermal Infrared	TIR	8 μm – 14 μm
Microwave	MW	> 1 cm

In some regions of the EM spectrum, most of the energy is transmitted through the atmosphere. These regions are referred to as atmospheric windows. In other regions, certain molecules present in the atmosphere (e.g. H₂O, CO₂ and O₂) absorb or backscatter radiation (Figure 3).

Radiant energy is the most common energy unit and is expressed in Joules (J). It describes the total energy that is radiated in all directions, away or towards a surface. The Stefan-Boltzmann law (Eq. 3) describes the radiant energy that is emitted by a perfect blackbody over all wavelengths, per unit area and unit time. This is also called the exitance (M) and is calculated as

$$M = \sigma T^4 \quad \text{Eq. 3}$$

, where σ is the Stefan-Boltzmann constant, and T is temperature. Following this law, total radiant energy emitted by an object is dependent on the temperature of this object. Note that this law describes a perfect blackbody, an object that absorbs and re-emits rather than reflects all energy it receives.

As perfect blackbodies do not exist in nature, another term is added to this equation by Kirchhoff's Law when naturally emitting surfaces are studied: This is emissivity (ϵ), the fraction of incident radiation that is absorbed and emitted by an object. Almost all objects and materials have an emissivity that varies with wavelength or frequency and are therefore called selective radiators. Emissivity (ϵ) is mainly a function of

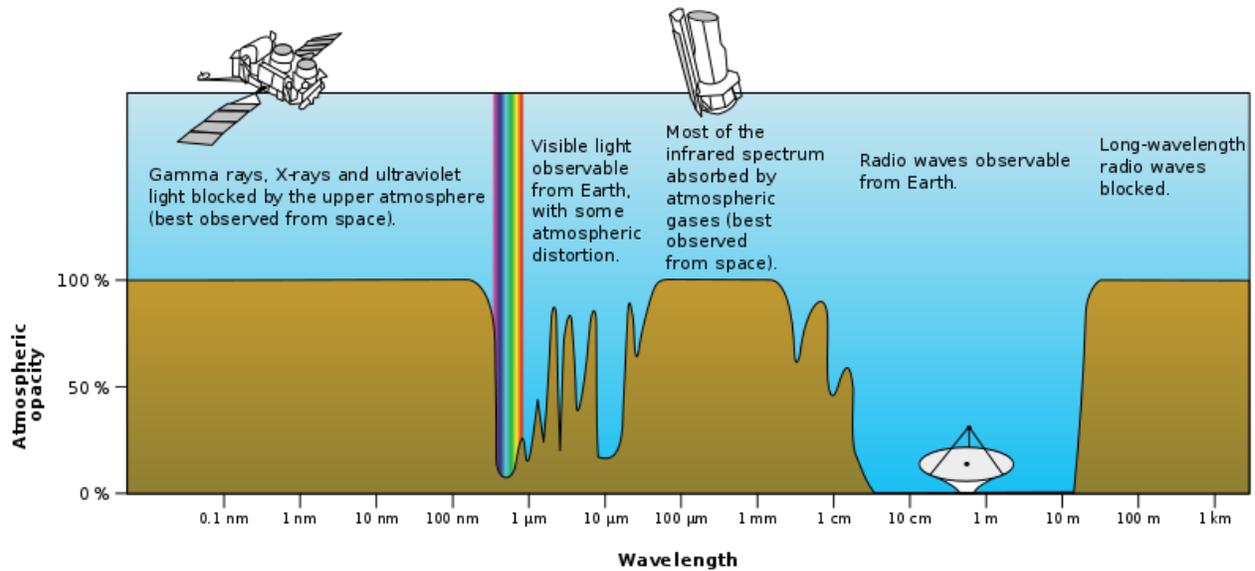


Figure 3: Simplified electromagnetic spectrum, including identified regions and atmospheric windows. Source: NASA (2008), public domain

the dielectric constant of the material, also known as relative permittivity. This describes the ability of a certain material to store electrical energy in an electrical field. Solid materials are non-conducting and have a relatively low dielectric constant (~ 3), whereas liquid materials such as water are good conductors and have a higher dielectric constant (~ 80). This means that, when water is added to a grid cell with dry soil, the dielectric constant of this grid cell rises. As a result, the reflectance of this wet soil increases, and ϵ decreases, leading to a lower M . Apart from the dielectric constant, surface roughness, incidence angle and polarization also have an influence on emissivity.

2.3.2 Passive microwave sensors and brightness temperature

Many satellites are equipped with sensors that are used for RS purposes. Utilizing the previously mentioned atmospheric windows, these sensors measure EM radiation that is emitted or reflected from the Earth's surface, transmitted through the atmosphere in a particular wavelength. RS makes use of both passive and active sensors: Active sensors send out a signal using their own source of energy, and therefore rely on the reflection of this 'artificial' signal back from the Earth surface. In contrast, passive sensors rely on incoming radiation that is either directly emitted by the Earth surface or emitted by the Sun and subsequently reflected by the Earth. Passive sensors therefore do not have their own source of energy (Chuvieco 2016).

As shown in Figure 3, the microwave (MW) region of the EM spectrum contains a large atmospheric window, meaning it is almost entirely not inhibited by the atmosphere. An advantage of this is that MW radiation is almost unaffected by clouds. However, its long wavelength and low frequency also mean that its energy content is relatively low (see Eq. 1 and 2), which leads to the fact that it is also relatively difficult to detect. A large area therefore needs to be observed for a good signal-to-noise ratio, meaning the spatial resolution of passive MW sensors is generally low. The MW part of the EM spectrum can be divided into several subregions that are used for different monitoring purposes. The part of the Ka-band subregion that will be used in this study follows the approach taken by Brakenridge et al. (2007) and has a frequency of 36.5-37 GHz. Passive MW imagery at this frequency can, for example, be used to observe land surface temperature (Holmes et al. 2009) at the vertical polarization, snow cover on mountain slopes (Pilant 1994), patterns of snowmelt (Arndt et al. 2016) or monitor rapid intensification of cyclones (Kieper and Jiang 2012). An overview of existing satellites carrying sensors in this wavelength can be found in Appendix A1.

Passive microwave sensors (or radiometers) measure raw antenna counts, which can be translated to the exitance emitted from the top of the atmosphere. This exitance (M), expressed as the temperature that a blackbody would have if it would emit the same amount of radiation in this particular wavelength, is called the brightness temperature (T_b). T_b is influenced by factors such as soil moisture, vegetation water content and surface temperature, but also surface roughness, vegetation type and soil texture (Njoku et al. 2003; Prigent et al. 2006). Properties of the emitting surface can therefore be derived from T_b using an inverse microwave radiative transfer model.

Prigent et al. (2006) applied such a model to examine the emissivity of objects derived from different SSM/I-observations, one of the instruments used in this study. The authors found that at a frequency of 37 GHz, open-water surfaces showcase a relatively low emissivity in both horizontal and vertical polarizations when compared to areas with less standing water. This in turn leads to a lower measured T_b over areas with large river systems and surface water. This can be observed in Figure 11, where the outline of Lake Malawi is clearly visible through the presence of lower T_b values.

2.4 Flood detection using passive microwave remote sensing

2.4.1 A brief review of microwave remote sensing flood monitoring

RS was introduced to the realm of flood risk monitoring and management over 40 years ago. The use of these data for the monitoring of floods, floodplain inundation and river hydrodynamics has increased substantially in the last two decades (Petropoulos and Islam 2017). Images in the visible and infrared ranges are available over long historical time periods, making them useful for long-term flood studies (Petropoulos and Islam 2017). However, sensors that operate within the MW region of RS are frequently used for flood monitoring as well, primarily due to their ability to operate in situations of adverse weather conditions.

Studies have shown the utility of active MW sensors, such as altimeters, to estimate surface water elevation and measure water level change (e.g. Alsdorf (2003)) or estimate downstream discharge, using downstream altimetry data in combination with upstream discharge data (Birkinshaw et al. 2010). Furthermore, images from SAR sensors can be used to extract maximum channel width and water surface width, which can be combined with river slopes to estimate discharge (Bjerklie et al. 2005).

Starting a discussion on the utility of passive MW sensors as well, Brakenridge et al. (2007) demonstrated that PMRS-data can be used to estimate discharge and watershed runoff, making use of a metric that will be referred to as the CM-ratio (r_{cm}). An advantage of his method is the use of wide-swath sensors, which provide daily or near-daily measurements over a near-global coverage. Using the same approach, De Groeve et al. (2007) demonstrated how this method could be used to detect floods in near real-time on a global scale. An experiment in Namibia executed by De Groeve (2010) showed that the signal had a good correspondence with in situ observed discharge values, although variable local conditions on the ground have a large impact on the signal-to-noise ratio. Hirpa et al. (2013) used upstream rather than downstream satellite signals to forecast floods in the Brahmaputra and Ganges watershed, and found that PMRS-data is a useful source of flood information in data-scarce areas.

2.4.2 CM-ratio (r_{cm})

The satellite-derived signal r_{cm} is calculated as

$$r_{CM} = \frac{T_{bCd}}{T_{bM}} \tag{Eq. 4}$$

, where T_{bM} is the T_b of a (wet) measurement cell containing a substantial fraction of surface water, and T_{bCd} is the T_b of a closely located, dry calibration target (Figure 4).

In order to derive wetness from differences in T_b , one needs to correct for the other factors influencing T_b , including variable vegetation and surface temperatures (Njoku et al. 2003). Assuming these factors have a similar influence in the measurement cell (M) and the dry calibration cell (C_d), the resulting r_{cm} should simulate the presence of surface water. If a river stage rises, the ratio rises along with it, and if overbank flow occurs in a large part of cell M, the r_{cm} rises sharply (Brakenridge et al. 2007). The r_{cm} can therefore function as a discharge estimator.

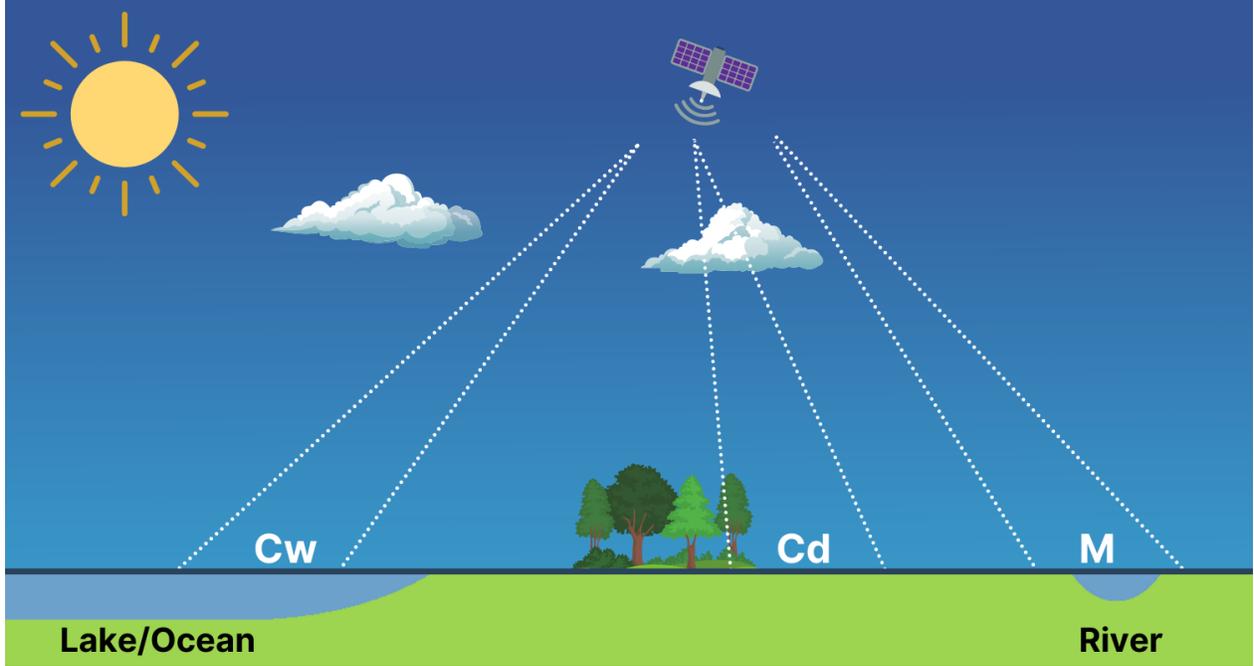


Figure 4: Schematic overview of the respective locations of the cells/pixels used in this research to calculate the indices, including wet calibration cell (C_w), the dry calibration cell (C_d) and the measurement cell containing the river (M)

Surface temperature, vegetation and other environmental variables have a considerable effect on raw brightness temperature values. The CM-methodology is therefore based on the assumption that cells that are located within a reasonable distance from one another are similarly affected by these variables (Van Dijk et al. 2016). Brakenridge et al. (2007) therefore located C_d within the temperature-correlation length of M, removing this factor from the ratio. Choosing two cells that are located close to one another also means that the T_b measurements are approximately taken at the same time, and are unaffected by uncorrelated changes in vegetation, soil characteristics or rainfall distribution. At the same time, however, C_d must be located outside of the river reach of M, yielding it a relatively stable calibration target.

2.4.3 CMC-ratio (r_{cmc}) and m

The first standardized signal that will be studied is m : It is directly related to r_{cm} and uses historical r_{cm} -signals to express the observed r_{cm} as the number of standard deviations from the base value. It is calculated as

$$m = \frac{r_{CM} - \text{avg}(r_{CM})}{SD(r_{CM})} \quad \text{Eq. 5}$$

, where r_{CM} is the r_{cm} as calculated in Eq. 4, $\text{avg}(r_{CM})$ is the average r_{cm} over the complete historical period available, and $\text{SD}(r_{CM})$ is the standard deviation of r_{cm} over the same period. The standardized index was introduced by De Groeve et al. (2007) and further elaborated upon in by De Groeve (2010).

The second standardized signal is r_{cmc} . It is similar to r_{cm} , but introduces an additional calibration cell to the equation. The concept was proposed by Neisingh (2018a) as part of an MSc thesis. The author had identified two shortcomings of the r_{cm} : The signal could not easily be interpreted in a spatial way, and it does not distinguish what fraction of surface water is caused by actual flooding. He calculated the r_{cmc} as

$$r_{CMC} = \frac{T_{b_{C_d}} - T_{b_M}}{T_{b_{C_d}} - T_{b_{C_w}}} \quad \text{Eq. 6}$$

, where $T_{b_{C_d}}$ is the T_b in the dry calibration cell (C_d), T_{b_M} the T_b in the measurement cell containing the river, and $T_{b_{C_w}}$ the T_b in a second calibration cell, one that is permanently wet (C_w). In contrast to r_{cm} , which was developed to mimic discharge patterns, r_{cmc} is an estimation of the surface water fraction present in the cell and is hence related to flood extent. C_w should ideally be located close to C_d and M . Note that the identification of C_w can be difficult in areas lacking large water bodies, considering the coarse spatial resolution of most passive microwave RS-data.

3. Methods

3.1 Case study: Malawi

3.1.1 *Topography, hydrology and climate*

The Republic of Malawi is a landlocked country in the southeast of Africa. It borders Zambia, Tanzania and Mozambique and has a surface area of approximately 118,484 km², one fifth of which is covered by water (Macharia et al. 2015). The country is divided into three large administrative regions, each of which is divided into districts (Figure 5). Its topography is dominated by the Great Rift Valley, which crosses the country from north to south (Figure 6).

In the Great Rift Valley lies Lake Malawi (Figure 6), that borders 75% of the eastern boundary of Malawi's land mass (Macharia et al. 2015). The lake forms the source of the Shire River, that flows out of the lake on the southern end to eventually join the Zambezi River in Mozambique.

Malawi has a subtropical climate with a strong bimodal seasonal pattern. The climate is governed by the oscillations of the Inter-Tropical Convergence Zone (ITCZ), interactions with the air mass above Congo, as well as the meridian south-eastern trade winds and monsoonal north-eastern winds (Macharia et al. 2015). The main growing season is the wet and warm season, taking place from November to April and accounting for 95% of all annual precipitation (Macharia et al. 2015; Malawi Meteorological Services 2020). From May to August, cool and dry winter weather dominates (17-27°C), with frost occurring in some areas between June and July (Malawi Meteorological Services 2020). September and October are the hottest and driest months, with average temperatures ranging from 25°C to 37°C (Malawi Meteorological Services 2020).

3.1.2 *Population and economy*

In 2018, Malawi had a population of 18.1 million people (World Bank 2018a), of which 83.1% based in rural areas (World Bank 2018b). The country is the third poorest country in the world (Troglić et al. 2019), with approximately half of the population living below the poverty line (World Bank 2016). Its economy is mostly agro-based, with 35% of the country's Gross Domestic Product (GDP) originating from the agricultural sector (Macharia et al. 2015). In 2011, 97% of all rural households in Malawi were engaged in agricultural activities (Davis et al. 2017), indicating the importance of agriculture for the Malawians livelihoods. However, only 2.3% of all cultivated land, and 66.0% of all arable land suitable for irrigation was irrigated in this year (FAO 2011), yielding the country very dependent on the weather. It is therefore estimated that the floods and droughts together reduce the country's GDP by 1.7% (Pauw et al. 2011).

3.1.3 *Floods in Malawi & early warning systems*

Malawi experiences annually fluctuating weather conditions, especially in the southern region of the country (Hendriks and Boersma 2019), leading to recurring flood and drought events. Flooding accounts for approximately 40% of all recorded disasters in the country (Mijoni and Izadkhah 2009), affecting millions of lives and causing displacement, economic damage and casualties. One of the earliest written record of a flood in the country dates back to 1942, when the Lower Shire Valley in the southern region experienced severe flooding (Mijoni and Izadkhah 2009). Despite the floods being a reoccurring phenomenon, the drought and floods that occurred in the growing season of 2014-2015 were far more destructive than most recent disasters (Hendriks and Boersma 2019). As recently as 2019, floods occurring in the wake of Cyclone Idai had a destructive impact on the country as well (Hendriks and Boersma 2019), claiming 60 lives and affecting more than 975,000 people (Guha-Sapir 2020).

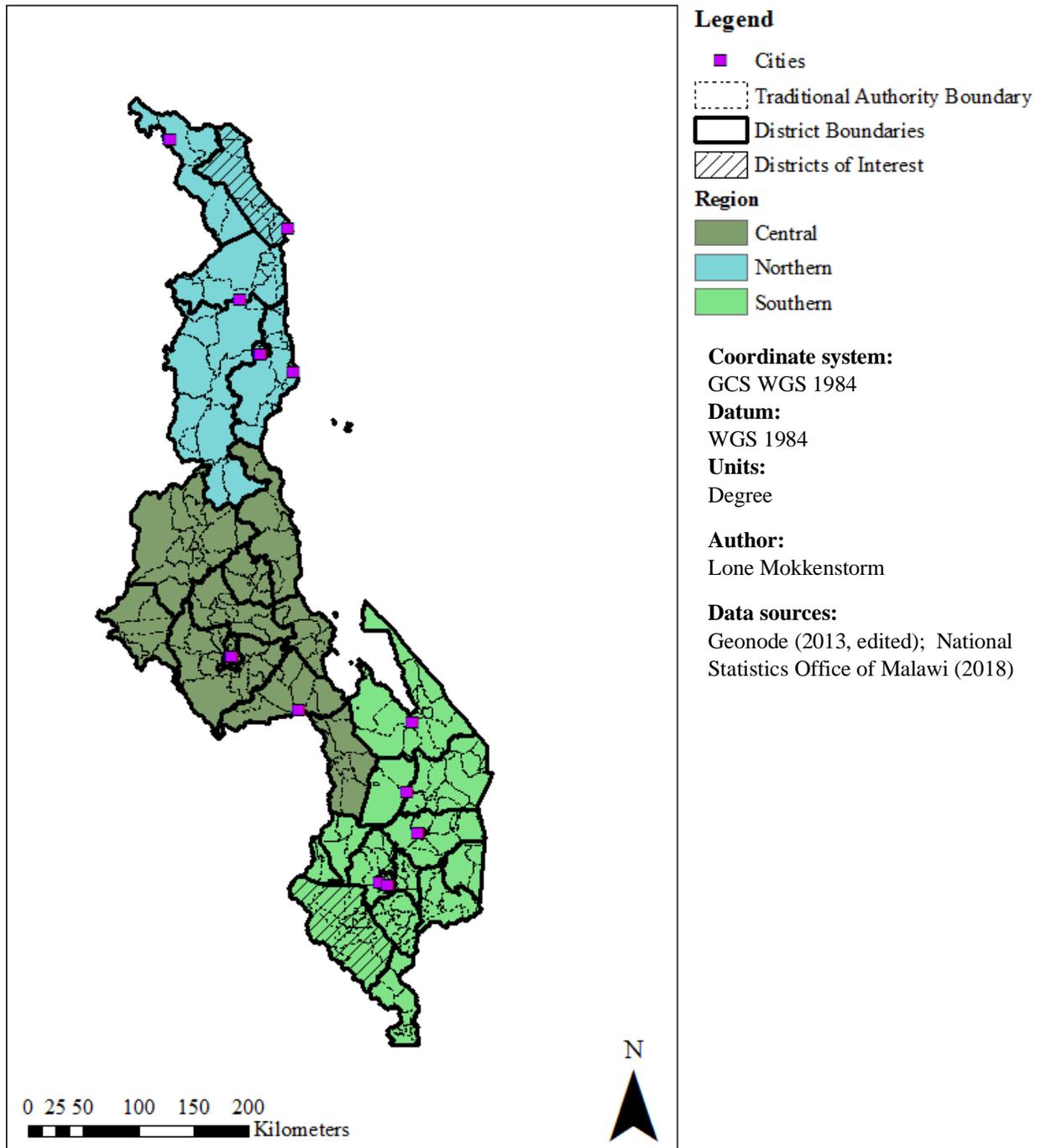


Figure 5: Overview map of Malawi. Administrative units and districts studied in this thesis, namely Chikwawa in the southwest and Karonga in the north.

Many of the current flood risk management practices in Malawi are the result of community-based systems, funded by international donors. An example of community-based flood risk management is the set-up of community-based Early Warning Systems (CB-EWS), that monitor rainfall and water level gauges and disseminate messages downstream when deemed necessary (Šakić Trogrlić et al. 2018). These practices are the product of local government structures called Civil Protection Committees (CPCs). CPCs operate on village- and district level. Village CPCs create Village Contingency Plans and Action Plans, in which

hazards and risks are identified and mitigation strategies are proposed, while district CPCs coordinate and oversee mitigation work at lower administrative levels and are in contact with the government (Šakić Trogrlić et al. 2018).

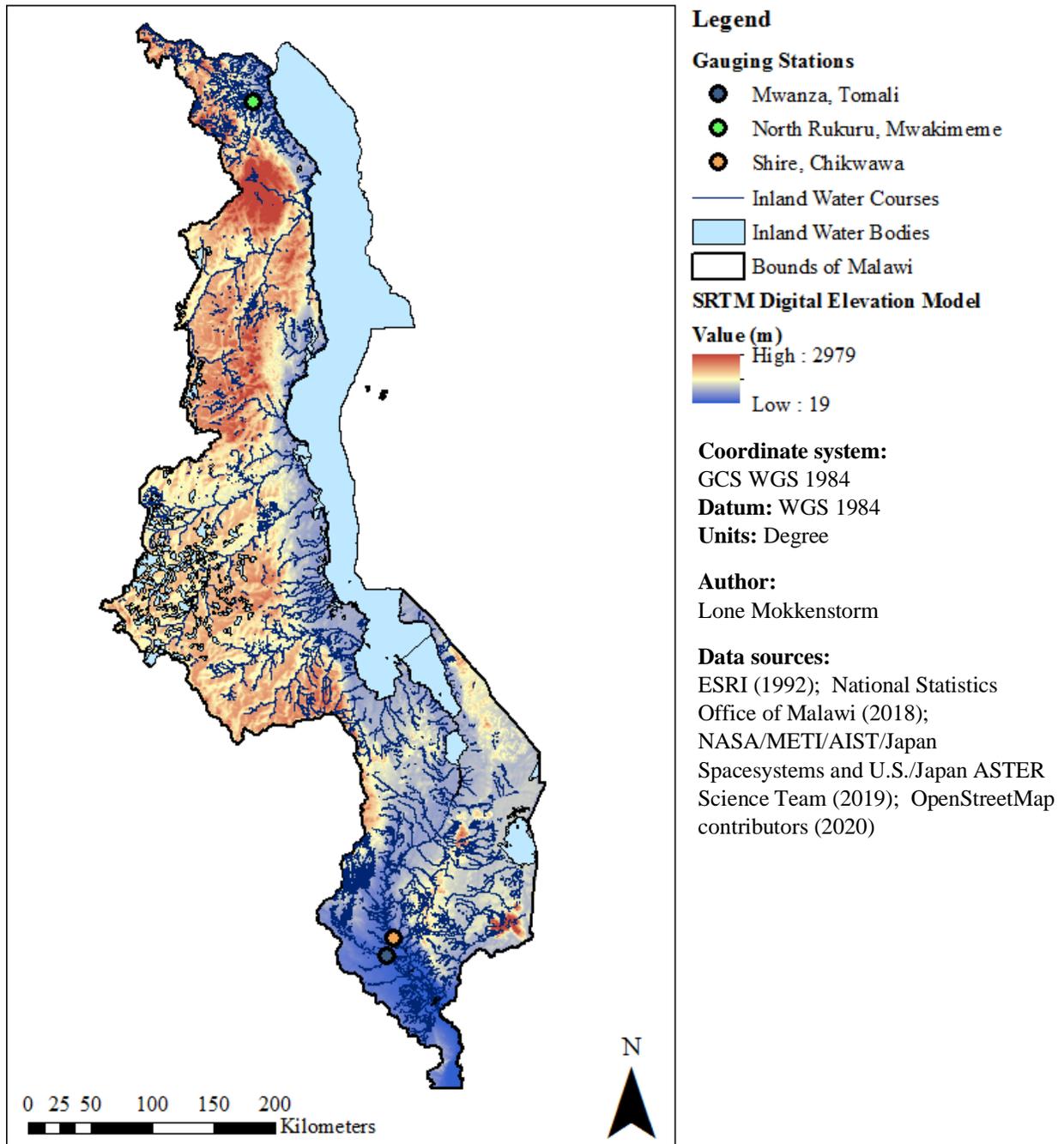


Figure 6: Overview map of Malawi: Topography and Hydrology

The official national EWS consists of the Operational Decision Support System (ODSS), a meteorological, hydrological and hydraulic flood forecasting and warning system that predicts riverine floods on the short-term, i.e. approximately 3 days. It is currently only operational in the Lower Shire Valley. This area, spanning across the southern districts of Nsanje and Chikwawa, is considered one of the areas in Malawi

most severely affected by floods (Mijoni and Izadkhah 2009). ODSS forecasts floods with a lead time of approximately 3 days. It also has components that predict weather, droughts, seasonal water levels and crop phenology (Ammentorp and Richaud 2016). The ODSS component that predicts short-term river flow and water level variations is based on the MIKE 11 modelling software developed by the Danish Hydraulic Institute, which uses the NAM hydrological model for rainfall-runoff modeling: A lumped, conceptual rainfall-runoff model that also takes soil moisture into account (Nielsen and Hansen 1973). The potential of applying ODSS in the rest of the country is currently being investigated.

As mentioned in Section 2.1.1, many existing FbF systems are based upon the global model GloFAS (Lumbroso et al. 2017; Jjemba et al. 2018), which is based on rainfall data as well as a hydrological and hydraulic model. The model is freely available and provides medium-range forecasts. Whereas the maximum lead time of GloFAS is 30 days, the daily forecast data is provided with a maximum lead time of 15 days. For the case of Malawi, the applicability of GloFAS for FbF has been investigated by 510 (the data initiative of The Netherlands Red Cross) for the Lower Shire Valley (Teule 2019). It was found that GloFAS cannot accurately predict absolute discharge values in the area but could be used in forecasting systems if trigger levels would be set correctly.

3.2 Virtual gauging stations

3.2.1 Selection of downstream points of interest

The study areas were selected using both a humanitarian perspective and a technical perspective: Districts of interest were selected based on their vulnerability, exposure to riverine floods and lack of coping capacity as reported on the Community Risk Dashboard by 510 (510 an initiative of the Netherlands Red Cross 2020a). From the districts with the highest scores on these levels, Karonga and Chikwawa were selected as the most relevant districts, as they were respectively mentioned in the ECHO III and ECHO V projects of the Red Cross as ‘areas of intervention’ and are therefore relevant for the work of the Malawi Red Cross Society and its partner National Societies. Within these districts, downstream Virtual Gauging Stations (VGSs) were selected (Figure 7). These sites were selected from a technical perspective: They had to be situated relatively downstream in the watershed, have a known record of riverine flooding, and have a visible flood plain as identified from optical satellite imagery. The latter factor was visually assessed, making use of the time slider function in Google Earth Pro 7.3.2.5776. The areas of interest also had to be relatively densely populated to yield them important for humanitarian action.

Karonga – North Rukuru

For the district of Karonga, the downstream point of interest (9°56′15″S, 33°50′44″E, K0 on Figure 7) is situated in the grid cell covering the capital of the district, Karonga Town. The area is prone to several natural hazards, but floods are locally regarded as the most serious (Manda and Wanda 2017). Riverine flooding of the North Rukuru River has occurred here every wet season between 2009 and 2016, where the floods in 2010 and 2016 were reported especially severe (Manda and Wanda 2017). The floodplains on the river (up to 125 metres wide from river to end) are home to many informal settlements, which are especially vulnerable to flooding, even frequent low-impact events (Manda and Wanda 2017). The North Rukuru River is the main river in this low-lying region, meandering through the grid cell, entering Karonga Town, and eventually draining into Lake Malawi. It originates in the Nyika Plateau (Figure 7). The North Rukuru River shows a strong seasonal pattern in streamflow, with a wet season starting in November / December, reaching a high peak before gradually reclining from May onwards to a discharge of nearly 0 (see Section 4.2.1).

Chikwawa - Shire

For the district of Chikwawa, the focus is on the Shire River, the largest river in Malawi that originates in Lake Malawi and flows into the Zambezi River in Mozambique. The downstream point of interest is located

along the Lower Shire River, covering the city of Chikwawa (16°04'07"S, 34°49'50"E, C0 on Figure 7), which is frequently hit by riverine floods, including the recent severe floods in 2015 and 2019 (Guha-Sapir 2020). At Chikwawa, the Shire River showcases a unimodal seasonal pattern in streamflow, with a peak occurring in February or March, corresponding to the peak precipitation season. Discharge ranges from an average of approximately $400 \text{ m}^3 \text{ s}^{-1}$ in the dry season to an average of $600 \text{ m}^3 \text{ s}^{-1}$ in the wet season (see Section 4.2.1). The floodplains along the river are up to approximately 800 metres wide from river to end, and seem to flood less frequently than those along the North Rukuru.

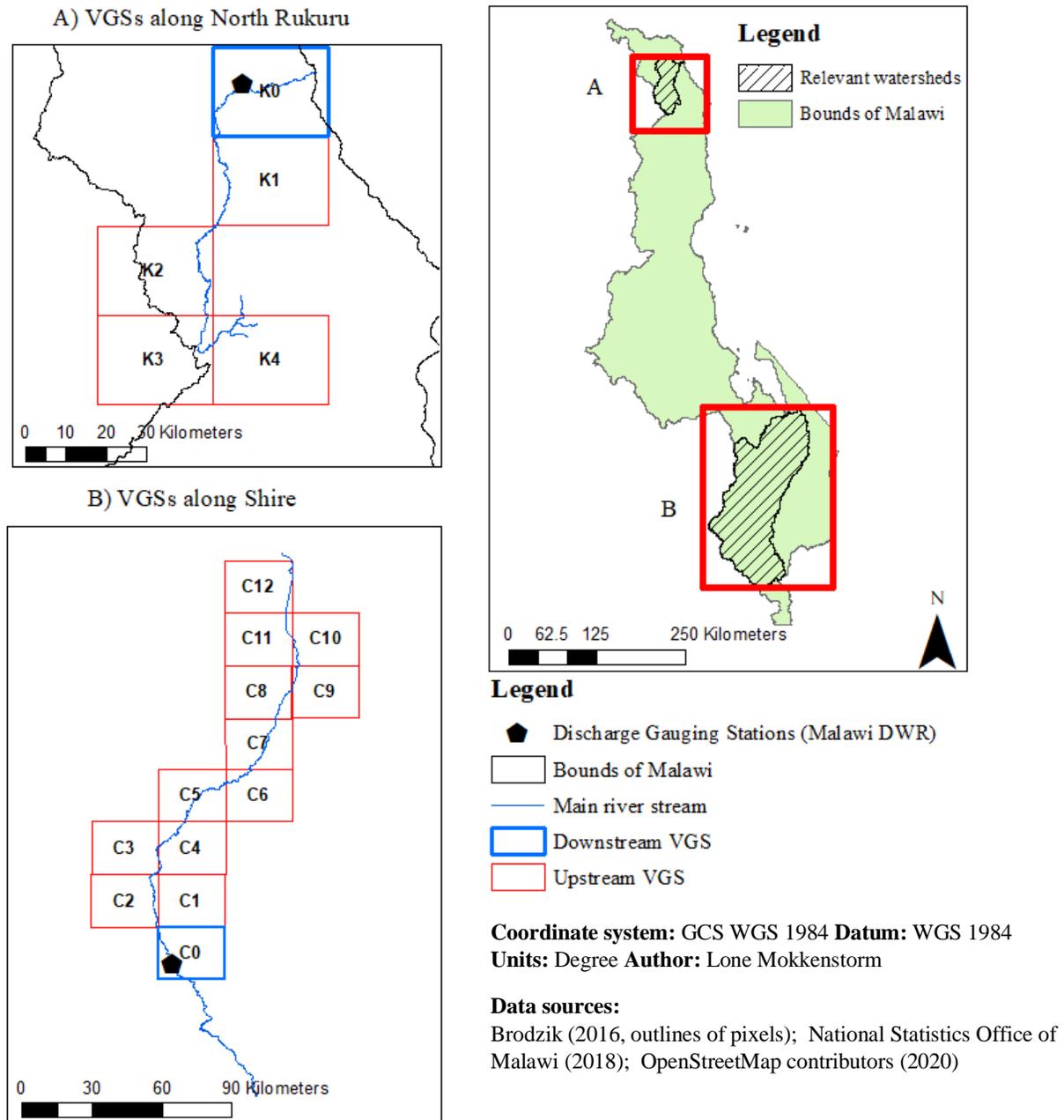


Figure 7: Upstream and downstream Virtual Gauging Stations (VGSs) along the Shire and North Rukuru River. One grid cell is approximately 25 x 25 kilometres. The labels on the cells are used for referencing them throughout this thesis, with the letter corresponding to the district Chikwawa (C) or Karonga (K)

3.2.2 Selection of upstream points of interest

For both the points of interest in Karonga and Chikwawa, a series of upstream points of interest were selected, following respectively the flow of the Shire River and North Rukuru River (Figure 7). Due to the lack of a comprehensive shapefile containing the headstreams of these rivers, these had to be identified first in ArcMap 10.5.1. Making use of waterway lines from OpenStreetMap (OpenStreetMap contributors 2020), supplemented with a basin map and flow accumulation map (created with the ASTER GDEM (NASA/METI/AIST/Japan Spacesystems and U.S./Japan ASTER Science Team 2019) the main flow of both rivers was identified. Subsequently, the river flow was traced upstream, where each EASE-grid cell upstream from the downstream point of interest (Section 3.2.1) was marked as upstream VGS. The centroids of the cells were used as input for the CM- and CMC-calculation scripts (Appendix A3).

Early on in the research process, the satellite signals extracted from the K0-cell - that was considered the appropriate downstream VGS for Karonga - showcased a flipped seasonality compared to the observed discharge data (see Section 4.2.1 for more information). As the first VGS with a realistic seasonal oscillation was cell K2, this cell was treated as a proxy for the original downstream VGS.

3.3 Data acquisition and treatment

Data analysis was conducted in Spyder (Anaconda 3) with Python 3.7, and graphs were created with Spyder and Microsoft 365 Excel. The scripts can be accessed through the link in Appendix A3.

3.3.1 Brightness temperature data

Brightness temperature data were obtained from NASA's Making Earth System Data Records for Use in Research Environments (MEaSUREs) Calibrated Enhanced-Resolution Passive Microwave Daily EASE-Grid 2.0 Brightness Temperature ESDR, Version 1 (Brodzik 2016), a freely available dataset including passive microwave T_b data from different platforms and sensors. Whereas the MEaSUREs dataset itself spans from October 25th, 1978 to July 1st, 2017, some platforms, including the DMSP-F17 used in this study, are still operational and therefore still record data. All acquired data was Level 3-processed (NASA standards) prior to acquisition, meaning the raw data were processed to sensor units (T_b), calibrated, and mapped onto a grid, in this case into a resampled grid resolution of 25 kilometers. Due to the gridding, the data may be temporally averaged or ignore overlapping satellite swaths altogether. Gridded, calibrated data allows for long-term studies between sensors, although some differences between sensors do remain (Brodzik 2016). The data was downloaded utilizing an adapted version of a script by Neisingh (2018b) and the software WGet (Appendix A3).

Brakenridge et al. (2007) used the brightness temperature measured in the 36.5 GHz channel of the Advanced Microwave Scanning Radiometer (AMSR-E). This frequency was chosen over others because because it suffered from less interference than other options: The 6.92 and 10.65 GHz channels had a low spatial resolution and were interfered by from radio frequencies, while the 89.0-GHz channel was too close to the oxygen spectral line, and the 8.7- and 23.8-GHz channels were too close to the water vapor spectral line (Brakenridge et al. 2007). This means that the oxygen and water molecules in the atmosphere would absorb radiations at these frequencies. The authors used imagery in the horizontal polarization and the descending (evening) node. These settings are in accordance with the optimal settings for r_{cmc} as defined by Neisingh (2018a), and those used by other previous research on the topic of CM-signals (De Groeve et al. 2007; De Groeve 2010; Hirpa et al. 2013). Therefore, horizontally polarized imagery from the descending node was used in this research as well. As a frequency, 37 GHz was chosen over 36.5 GHz since data at this frequency in MEaSUREs covered a much longer period.

A long-term timeseries of T_b was created by downloading MEaSUREs data from different platforms and sensors, and combining them in one dataset, spanning from 1978 to 2017 (Appendix A2).

Generally, EM-radiation in the MW region is unaffected by cloudy conditions. However, some cloud-induced noise could remain in the T_b signal in cases of thunderstorms or heavy rain events. In order to eliminate this, and account for the fact that the satellites do not achieve full swath coverage near the equator, a filter method should be applied to the data. For this research, the widely used Savitzky Golay low-pass filter (that uses a local least-squares polynomial approximation (Savitzky and Golay 1964)) was considered. However, this method makes use of a centered moving window, which is not ideal for forecasting purposes. As we are interested in forecast skill, the filter method should be able to be applied to the data in real-time and cannot be looking to neighbouring values in the future, as this would remove valuable response time in an FbF-context. Therefore the T_b data for each cell was filtered using the same approach as De Groeve (2010) and Van Dijk et al. (2016). A filter was applied to the values that takes the average of the preceding 4 values and the current value. This window removes some data gaps, most notably in the period 1978-1987, when the satellite return period is two days rather than one, and eliminates some noise, whilst also retaining the flood peaks in the data. Furthermore, it allows for comparison with the research results by De Groeve (2010) and Van Dijk et al. (2016). Figure 8 shows the effect of this filtering method on the measured T_b signal of the wet calibration cell (C_w), the dry calibration cell (C_d) and the measurement cell M in a wet season in 2015, and illustrates the effect a 15-day Savitzky-Golay filter of the 3rd order (as applied by Neisingh (2018a)), in combination with a backwards fill, would have had on the same data. A backwards fill is a fill method where missing values are filled with the next available data point.

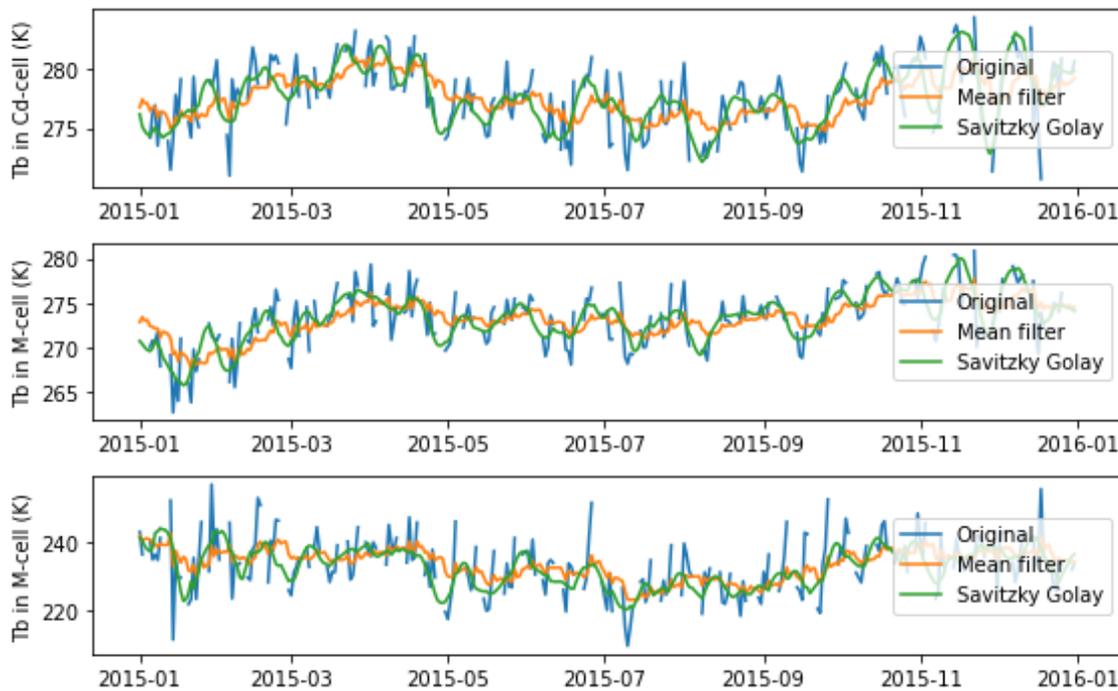


Figure 8: Illustration of the effect of the mean-filter method in Chikwawa in 2015 compared to a 15-day, 3rd order Savitzky-Golay on the brightness temperature (T_b)-signal from the measurement cell with the river (M), the dry calibration cell (C_d) and the wet calibration cell (C_w). Data source: Brodzik (2016).

3.3.2 Gauge data

Average daily discharge at the gauging station at Chikwawa (1L12), measured with an automated gauging system (Malawi Disaster Management Portal 2020), was obtained from the National Hydrological Services of Malawi. The original dataset spanned May 1977 to November 2009 (Table 2). Data from the Mwakimeme station in Karonga (8A5), a manual gauging station, was obtained from the Global Runoff Data Centre (GRDC) (56068 Koblenz, Germany). The data spanned from November 1968 to March 1991. For both stations, discharge entries older than 1979 (the first full year of the MEaSURES dataset) were dropped.

Table 2: Gauge data availability for Chikwawa and Karonga (% of total days with data per year). Red means no data was available, dark green means all data was available (linear coloring scale)

	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993
Chikwawa	91	100	91	85	96	87	60	100	100	100	100	95	91	89	100
Karonga	100	99	98	97	99	99	100	100	100	79	100	100	25	0	0
	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Chikwawa	93	100	56	95	93	87	85	83	84	73	70	42	47	68	83
Karonga	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

3.3.3 Historical flood impact and hazard data

A flood impact database was created by combining impact data from EM-DAT, the international disasters database by the Centre for Research on the Epidemiology of Disasters (Guha-Sapir 2020), with flood hazard data from the Dartmouth Flood Observatory (Dartmouth Flood Observatory 2020) and the GLIDE-database (GLIDE 2020). All floods that were tagged to have taken place in the districts of Chikwawa and Karonga were selected. Floods tagged as flash floods or floods having a duration of one day were dropped, as the distinction between natural, short-term signal fluctuations and flash floods cannot be made by just looking at the satellite signal, and because many flash floods tend to go unreported. Floods that only had a starting or ending month and no precise dates recorded were assumed to start on the first day of the month, or end on the last day of the month, respectively. After this pre-treatment, the database counted a total of 15 floods for Chikwawa and 12 for Karonga.

As many of the flood entries contained information on the affected districts only, and not the precise location or river where the flood took place, the resulting flood database was quality checked with an analysis of annual discharge extremes at our downstream points of interest. The maximum discharge observed within the timespan of these floods was extracted in order to 1: Check whether the reported flood occurred during annual peak discharge, and the flood therefore took place in our downstream area of interest and 2: Calculate the return period of each flood to check which floods exceeded a determined 5-year threshold.

3.3.4 Calculation of m and r_{cm}

Satellite signals were calculated from the observed T_b in the cells C_d , C_w and M , after the aforementioned backward-looking filter method had been applied to it. The two signals that were compared were m and r_{cm} . Whereas these signals are related to the r_{cm} , they are more standardized with respect to the average conditions at that point in time and at that location (Figure 9). In order to calculate m , r_{cm} was calculated first (Eq. 4). From r_{cm} , a standardized, unitless satellite signal magnitude m was calculated following De Groeve (2010) (Eq. 5). The mean and standard deviation were calculated per year to detect anomalies compared to the base level of that particular year. For the calculation of r_{cm} , the r_{cm} was not needed first (Figure 9). It was calculated using Eq. 6.

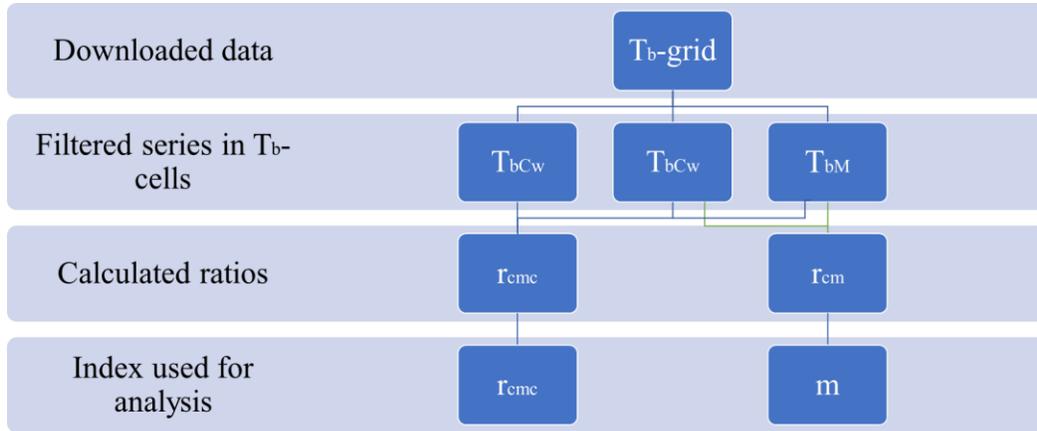


Figure 9: Schematic representation of the use and processing of raw brightness temperature (T_b) data to T_b -series in the wet calibration (T_{bCw}), dry calibration (T_{bCd}) and measurement cell (T_{bM}), eventually calculated into the CM- (r_{cm}) and CMC-ratio (r_{cmc}) and the m index (m)

For the calculation of both signals, the location of calibration cells has to be defined first. C_d was chosen by applying a kernel of 5×5 cells, i.e. a squared buffer of 50 km (2 cells of 25 km) in each direction around M. Within this kernel, the cell with a yearly average M-signal closest to the 95th percentile of all cells was chosen as the C_d , a method similar to that of Van Dijk et al. (2016). This process was repeated each year. While this would mean that C_d would likely not be at the same location each year, it would account for potential changes in river reach throughout the years.

The wet calibration target C_w , which is needed for the calculation of r_{cmc} , was manually chosen as the nearest cell to M that was fully covered by a large water body. The C_w -cell for Chikwawa was therefore located in Lake Chilwa, and the C_w -cell for Karonga in Lake Malawi, the closest large and permanent bodies of water for the two downstream points of interest. By selecting these locations, it was assumed that surface temperatures, emissivity and atmospheric moisture in these locations were similar as in the M-cell.

3.4 Relationship with discharge

3.4.1 Treating interannual variability and seasonality

The relationship between the satellite signals and observed discharge was established using a regression analysis. Chikwawa's discharge showed a climate-driven (Jury 2014), significant interannual baseflow trend (Kelly et al. 2019), whereas the standardized nature of the satellite signals meant they did not show this trend. As this would mean that the relationship with discharge would change throughout time, and as the aim of this research is to set up a forecasting system – yielding the recent values the most relevant –, it was decided to conduct this part of the research looking at the 5 most recent years of discharge record, which were also relatively stable in baseflow. As the seasonal oscillations were assumed to have an equal impact on both datasets in this relatively short period, the r_{cmc} , m and discharge data were not made stationary prior to the regression analysis. This approach was also applied to Karonga for the sake of consistency. Furthermore, only the months of December up to and including April were used in this step, as we are primarily interested in the wet season. Frequently occurring low discharge values in the dry season would potentially impact the relationship, and have also shown to not be well simulated by r_{cm} (Brakenridge et al. 2007).

3.4.2 Regression analysis between observed discharge and modelled satellite data

The 5 years of data from Chikwawa (cell C0) and Karonga (cell K2) was divided into two random subsets, both comprising half of the data. The relationships between the discharge data and both m and r_{cmc} were tested for normality of residuals, homoscedasticity, and linearity to assess whether a parametric or non-parametric correlation would be appropriate. Normality was tested with a Shapiro-Wilk test at 5% significance level and a Quantile-Quantile plot (QQ-plot). The QQ-plot is a scatterplot where two sets of quantiles are plotted against one another; if the sets came from the same distribution (in our case, a normal distribution), the points would form a linear line. Homoscedasticity was tested with a median-based Levene's test. Linearity was assessed visually with a scatterplot. The Shapiro-Wilk test and the Levene's test yielded a p-value of less than 0.05 for the discharge data and both r_{cmc} and m , meaning the null hypotheses of normality and homoscedasticity were rejected. Furthermore, the relationship between discharge and satellite signals did not seem linear. A non-parametric test using Spearman's correlation coefficient (ρ) was therefore deemed most appropriate. ρ is calculated as

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2-1)} \quad \text{Eq. 7}$$

, where the data is ranked from 1 to n (the number of data points), and $d_i = x_i - y_i$, the difference in ranking i and n . Spearman's ρ is a non-parametric approach that can be used when the assumptions of parametric tests do not hold, as is the case with this data. The relationship was simulated with a second-degree polynomial regression for Chikwawa and Karonga, applied to the training data. This degree was chosen because it was the lowest order at which monotonicity could be achieved in Chikwawa. In Karonga, monotonicity could not be achieved with any polynomial order. Spearman's ρ was calculated using the testing data and the polynomial predictions. The significance of the coefficient was assessed by calculating a two-sided p-value to test the null hypothesis that the predictions and test-values are uncorrelated ($\alpha = 0.05$).

3.5 Relationship with flood occurrence and timing

3.5.1 Setting the threshold with an extreme value analysis

To assess how well floods are detected by our VGSs, a threshold of what a flood is needs to be defined first. As FbF-programs are designed to be activated once every five years, this study looks at floods with a return period (or return interval) of five years or more. The return period (rp) represents the average estimated time it takes for a flood of a certain severity to repeat itself, where a flood with a rp of five years has a probability each year of 0.05 (5%) to return.

An extreme value analysis was done on the original discharge data (i.e. the complete dataset including the trend). For each hydrological year (May to April), the annual maximum was selected. The signal values at these peaks were sorted, after which a rank was attributed to them. The probability of each occurring value was calculated as

$$p = \frac{n-r+1}{n+1} \quad \text{Eq. 8}$$

, where n is the total number of peaks and r is the rank of the value. This number was then converted into a rp in years by taking the inverse of it:

$$rp = \frac{1}{p} \quad \text{Eq. 9}$$

A 10th degree polynomial curve was fitted to the existing data points to find the discharge value corresponding to 5 years. Note that this model cannot provide meaningful conclusions on floods with long rp values, as the length of our observations is relatively short. The r_{cmc} and m -thresholds were identified by

using one of two alternative approaches, depending on the outcome of the discharge-signal regression relationship identified previously:

1. If the discharge-signal correlation proved strong ($\rho > 0.7$) and significant, the m and r_{cnc} values corresponding to $rp=5$ would be identified by solving the polynomial equation for the 5-year discharge threshold.
2. If the discharge-signal correlation was not strong and/or not significant, the values corresponding to $rp=5$ would be calculated in a similar manner as the discharge threshold, applying the abovementioned flood frequency analysis on the original r_{cnc} – and m -signals between 1978 and 2017.

The resulting discharge, m and r_{cnc} -thresholds with $rp=5$ were used in the detection performance assessment.

3.5.2 Assessing performance

A confusion matrix was constructed for the different satellite indices (Table 3) to assess their performance in detecting floods. For each event in each hydrological year in which the 5-year threshold was exceeded, the presence of a flood in the database was also assessed. ‘Observed’ here refers to the presence of a flood in the database, and ‘Modelled’ to the satellite signal’s exceedance of the 5-year threshold. If the flood in the database occurred within 14 days from the flood detected by the signal, it was considered a ‘hit’ (H), i.e. a correctly detected flood. This margin was chosen because observed flood peaks in discharge data may precede a maximum in floodplain inundation, and this time difference can be up to a couple of days (Brakenridge et al. 2007).

Table 3: Confusion matrix. In the context of this study, ‘Observed’ refers to events reported in the impact database, and ‘Modelled’ refers to events simulated by exceedances of the trigger threshold.

		Observed	
		Yes	No
Modelled	Yes	Hits (H)	False Alarms (FA)
	No	Misses (M)	Correct Negatives (CN)

Different metrics can be calculated from the confusion matrix, which in turn can be used to evaluate the forecasting or detection skills of a model. However, the large number of days in the year where no flood occurred would lead to a high number of Correct Negatives (CNs) and hence impact any success score involving the CNs in its calculation. Therefore, the Critical Success Index (CSI), False Alarm Ratio (FAR) and the POD were calculated (Eq. 10, 11 and 12), skill scores commonly used for model/forecast evaluation that do not make use of the number of CNs. They are calculated as:

$$POD = \frac{Hits}{Hits+Misses} \quad \text{Eq. 10}$$

$$FAR = \frac{False\ Alarms}{Hits+False\ Alarms} \quad \text{Eq. 11}$$

$$CSI = \frac{Hits}{Hits+False\ Alarms+Misses} \quad \text{Eq. 12}$$

The POD represents the fraction of the reported floods that was successfully forecasted, whereas the FAR represents the fraction of all alarms that were false according to the records, and therefore is a measure of failure to exclude non-event cases. The CSI, also called the Threat Index, stands for the fraction of modelled floods that was also reported. The metric is often used by organizations such as the US National Weather Service to indicate the value of a warning. It is independent of the CNs, but varies directly with H, and varies inversely with the number of FAs and Ms (Schaefer 1990). As no single statistic can describe the

skill of a model by itself, all three metrics were considered in evaluating the skill of the satellite signals in detecting discharge.

3.5.3 Comparison with GloFAS

In order to compare the detection skill of the presented PMRS-model with existing EWSs that are used in a humanitarian context, a comparison of the Chikwawa confusion matrix was done with hindcasting data (simulated real-time discharge, also called reanalysis) from the Global Flood Awareness System (GloFAS). The global model had daily average discharge estimates available (2000-2018) in a global netcdf file with a grid of 0.1° by 0.1° cells, obtained through the Copernicus Climate Data Store and distributed by the Copernicus Climate Change Service by ECMWF (Harrigan et al. 2020). A 5-year rp threshold was generated from the discharge estimates (2000-2019) using the same method as was used for the satellite data (see Sections 3.5.1-3.5.2). This threshold, in combination with the discharge estimates and the compiled flood database, was used to generate a confusion matrix and compare the performance of this global model when similar data treatment is applied. The confusion matrix was based on all recorded floods, not just the floods that had an rp of 5 years or more.

3.6 Assessment of ability of prediction

3.6.1 Choosing the predictor for the forecasting model

An ex-ante forecasting model was created to assess the ability of r_{cmc} and m to forecast floods. A time-lagged cross correlation (TLCC) was first conducted to identify which VGS (see Figure 7), if any, could be used for forecasting downstream satellite signals, in the same manner as Hirpa et al. (2013) found upstream VGSs to be useful in forecasting downstream floods. If a VGS showed a strong correlation (correlation coefficient >0.7) and also had a lag time of at least one day, this VGS would be used for the forecasting model. If such an upstream VGS did not exist, the forecast model would be based on the satellite signals at the downstream VGSs (i.e. C0 and K2, since K2 was selected as a proxy for K0 (Section 4.2.1)).

3.6.2 Stationarity and lag times

As the TLCC assumed stationarity, the datasets were made stationary first. Whereas the Augmented Dickey-Fuller Test at a 95% confidence level rejected the null-hypothesis that the downstream (C0, K2) and upstream discharge data and satellite signals in the wet season (i.e. December to April) were non-stationary, this test does not account for seasonality but merely for the presence of a unit root. A visual examination of the datasets confirmed that the data indeed has a long-term trend. This is especially true for the discharge at Chikwawa, where the discharge timeseries shows a clear and significant change in baseflow over the years (Kelly et al. 2019), driven by long-term changes in precipitation (Jury 2014). Interannual trends were removed from all datasets in linear segments, where locations where the annual mean value changes shift from positive to negative or vice versa were used as break points. The data also showed strong seasonal oscillations, which is why the seasonal component was subsequently removed from the data by fitting a second-degree polynomial oscillating curve and subtracting this from the detrended values. After detrending, another Augmented Dickey-Fuller Test at 95% confidence level was conducted, which again rejected the null-hypothesis that the datasets were non-stationary. The effect of the detrending can be seen in Appendix C1.

The TLCC was conducted by artificially shifting the stationary data with a lag time ranging from -50 to 50 days. At each lag time, a pairwise correlation was conducted and Spearman's ρ was calculated to see how the correlation changed with different time steps, and at which lag time the correlation would be strongest (i.e. what would be the 'optimal lag time'). Depending on the outcome of the TLCC, either the upstream or the downstream satellite signals were selected as predictor dataset for the model.

3.6.3 Submodels for forecasting: Linear regression between predicted and observed satellite signals at one downstream location

The most suitable predictor dataset was identified in the TLCC, which turned out to be downstream VGS signals (see Section 4.4) rather than upstream VGS signals. The following steps were therefore applied to both the r_{cmc} and the m-data. The stationary dataset, covering the wet season, was first split in half into a training set and a testing set. Regression models that would come together in the overall forecasting model, hereafter referred to as ‘submodels’, were created using shifted versions of the training data, simulating forecasting lead times (i.e. the time between the ‘present day’ observation and the forecast in days).

The training set was tested for normality of residuals, homoscedasticity and linearity of the relationship with the shifted versions up to 30 days lead time. A visual assessment of the scatterplots showed that the assumption of linearity holds in both locations, especially for lower forecasting lead times. A median-based Levene’s test showed that m showcases homoscedasticity in Karonga up to a lead time of 18 days, and r_{cmc} up to 10 days. In Chikwawa, the test showed that homoscedasticity holds up to a lead time of 18 days for m and up until 14 days for r_{cmc} . Normality of the residuals was tested with a QQ-plot as well as a Shapiro-Wilk test at 5% significance level. Both methods showed that the residuals of the fit are not normally distributed at neither of Chikwawa nor Karonga. However, as the QQ-plot did not show extreme deviations from normality, and considering normality tests are sensitive to datasets with large sample sizes, it was decided to still proceed with a linear, parametric regression. After creating the linear (sub)model with the training data, its performance was assessed by comparing the predicted values from the test data with the observed values from the test data. This was done using two parameters, including the coefficient of determination (R^2). It is calculated as

$$R^2 = 1 - \frac{\sum(y - \tilde{y})^2}{\sum(y - \bar{y})^2} \quad \text{Eq. 13}$$

, where \bar{y} is the mean of response variable y , and \tilde{y} stands for the corresponding predicted values. R^2 describes what fraction of variance in the target values is accounted for by the model. It is a common method to assess the overall quality of model fit, the analysis of association. The submodels were created for a forecasting lead time of only 10 days to ensure the assumption of homoscedasticity would still hold. After the linear regressions were run on the training data for each lead time, the observed stationary satellite signals from the test dataset were used to assess the skill of each submodel, using the R^2 as well as the Root Mean Squared Error (RMSE). The RMSE is calculated as

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\tilde{y}_i - y_i)^2}{n}} \quad \text{Eq. 14}$$

, where \tilde{y} again stands for the predicted values corresponding to the observed y . The RMSE analyzes coincidence and explains how well the magnitude of the observed signal is simulated by the submodel.

As a rule of thumb, the performance of the sub-model was regarded sufficient if the R^2 was higher than 0.7 and if the RMSE was below 10% of the maximum value present in the test data.

3.6.4 Model set-up and assessing model performance

For forecasting satellite signals up to 10 days at a point in time t , the workflow of the model is as follows:

1. Model makes satellite data at point t stationary, resulting in stationary signal s_{st}
2. Model runs submodels sm_1 - sm_{10st} on s_{st} , resulting in 10 stationary forecast values f_{s1} - f_{s10} .
3. Model adds the trend that was removed in step 1 back to the data, resulting in 10 forecast values f_{1-10} .

Each timestep within a multi-day forecast is based upon three key elements:

- The submodel describing the relationship of that timestep with the r_{cmc} observed at t
 - Different for all three forecasts (individual submodels)
 - Submodels are based on all historical data up until and including t
- The trend observed at time t (T_t)
- The difference between the trend at t and $t-1$ (T_{Δ})
 - Used as an expected increase/decrease in trend per time step
 - For a 3-day forecast, this T_{Δ} is thus multiplied by 3 before being added to T_t
 - As T_{Δ} is a linear extrapolation, it is of the same magnitude between the time interval $T_t - T_{t+1}$ and $T_{t+1} - T_{t+2}$

Hence, in an operational setting, the different submodels would be merged into one final model: Each lead time has its own submodel, so a multi-day forecast would apply a different submodel to each timestep, creating a forecasted time series by merging the submodel forecasts together and adding the trend to these forecasts (Figure 10). The overall skill of the model was assessed for each lead time individually by looping through the satellite values observed in the wet season in 2015, making the most recent day in each iteration the t for that loop (Figure 10) and performing a forecast (at the lead time studied). The year 2015 was chosen, because it was a hydrologically relevant year in which no discharge data was available. A single year was chosen rather than a series of years to reduce the computational load of the experiment: each single data point of each timeseries required extracting the data up until the day before the data point one wants to compute, calculating the trend, applying the appropriate submodel and adding the computed, extrapolated trend. The performance was visually assessed, and by calculating the R^2 and RMSE for each lead time (see Section 3.6.3 for the equations).

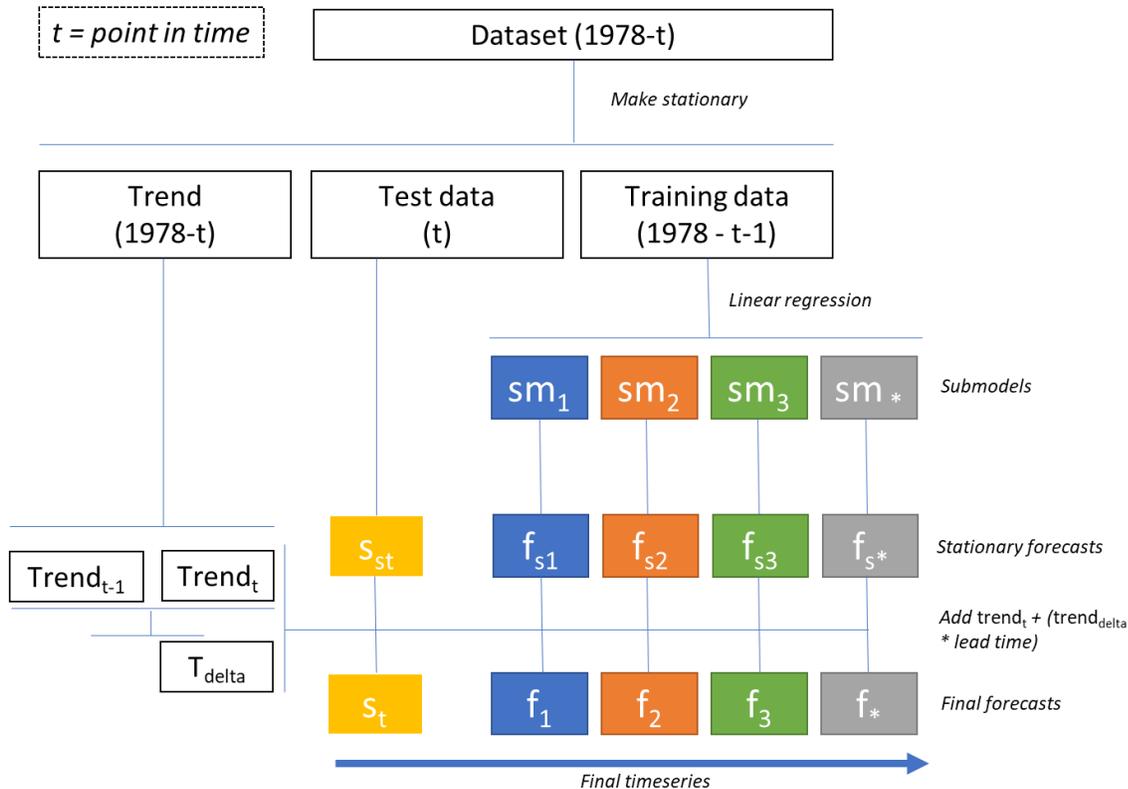


Figure 10: Conceptual overview of the model including processing steps. sm = submodels, f_s = stationary forecast, f = final forecast, t = point in time that represents 'present' when forecasting.

4. Results

4.1 Observed satellite signals

4.1.1 Locations of selected cells

The locations of the cells that were used to extract C_d , C_w , and M can be found in Figure 11. The location of C_w (black dots in Figure 11) was manually chosen and did not differ depending on which M -cell was studied. Using the automated method to select C_d , different grid cells within a 50 kilometer radius from the upstream and downstream VGS-cells (green dots in Figure 11). were selected. The algorithm allows the location of C_d (red dots in Figure 11) to differ from year to year. This is why C_d is present at different locations surrounding M in Figure 11, and why there are more C_d -cells than M -cells.

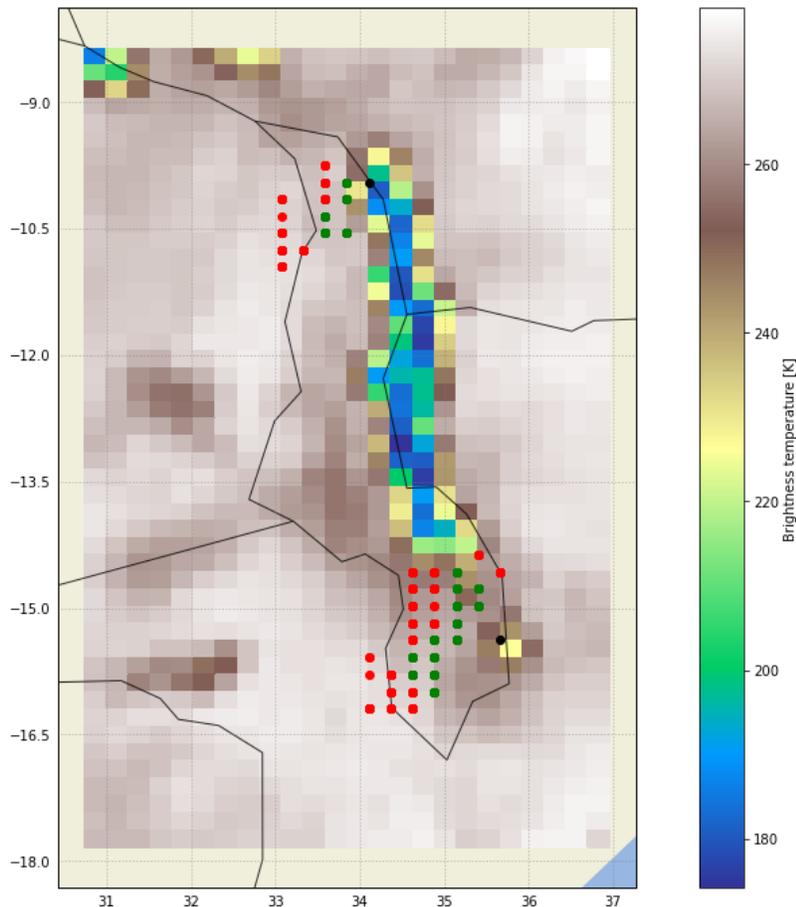


Figure 11: Locations of the measurement cells (M , green), dry calibration cells (C_d , red) and wet calibration cells (C_w , black) in the North Rukuru and Shire watersheds. The dots present the upper left corner of the T_b -cell in question. There are more C_d -cells than M -cells, because the method used to determine C_d 's location allows it to change each year. Data source: Brodzik (2016).

4.1.2 T_b per cell

The observed timeseries of C_d , M and r_{cm} for the downstream VGSs in Chikwawa (C0) and Karonga (K0) can be found in Figures 12 and 13, respectively. In Chikwawa, C_d and M show a similar trend over time

throughout the timeseries (Figure 12). C_w shows a clear peak in 1996, the reason for which will be discussed in Section 5.1.2.

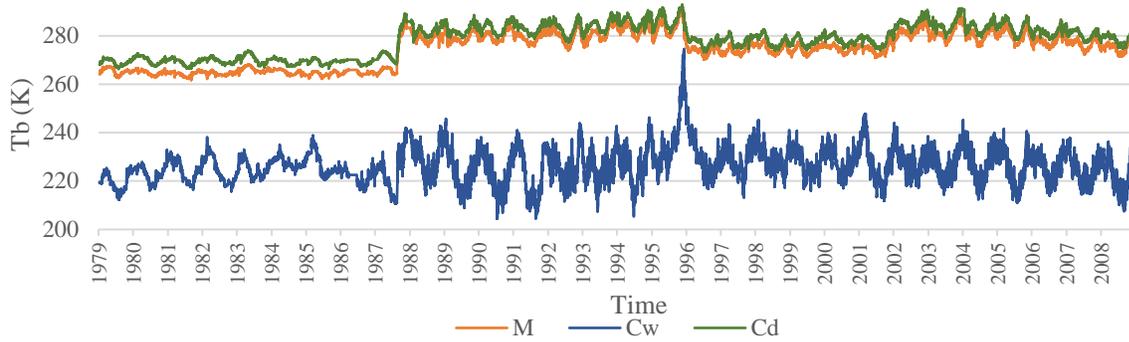


Figure 12: Brightness temperature (T_b) values in Kelvin (K) of the wet calibration (C_w , blue line), dry calibration (C_d , green line) and measurement (M , orange line) cells in the downstream virtual gauging station in Chikwawa (cell C0). Data source: Brodzik (2016).

In the downstream Karonga cell (K0), C_d and M do not follow a similar trend throughout the timeseries, and the difference in T_b between C_d and M is much larger than was the case in Chikwawa (Figure 13). In contrast, VGSs situated more upstream did show a C_d - M pattern more similar to Chikwawa. Potential reasons for this will be discussed further in Section 5.1.2.

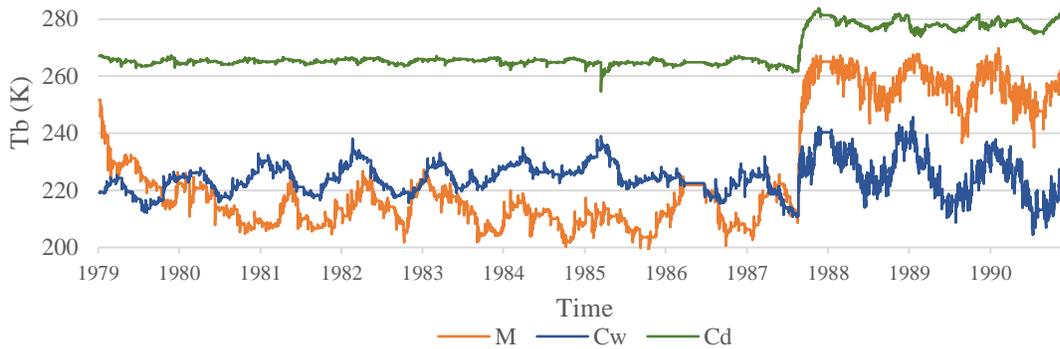


Figure 13: Brightness temperature (T_b) values in Kelvin (K) of the wet calibration (C_w , blue line), dry calibration (C_d , green line) and measurement (M , orange line) cells in the downstream virtual gauging station in Karonga (cell K0). Data source: Brodzik (2016).

4.1.3 CM-ratio (r_{cm}), CMC-ratio (r_{cmc}) and m

From the observed T_b in the cells, different satellite indices can be calculated. In this study, two standardized alternatives to the r_{cm} are used: m and r_{cmc} . Appendix B2 illustrates the difference between these signals and the ‘traditionally used’ r_{cm} in K0 and C0. In the remainder of this study, the r_{cm} itself is not touched upon, but it must be noted that m has a perfect linear correlation with r_{cm} ($r = 1.0$) and findings based on m can therefore be connected to r_{cm} . r_{cmc} does show a slight signal difference from r_{cm} (Appendix B2).

4.2 Detection of discharge

4.2.1 Seasonality of discharge and signals

In Chikwawa, the signals observed in the downstream VGS follow a similar seasonal oscillation as the discharge. Note the previously mentioned interannual trend present in the discharge series (Figures 14-15).

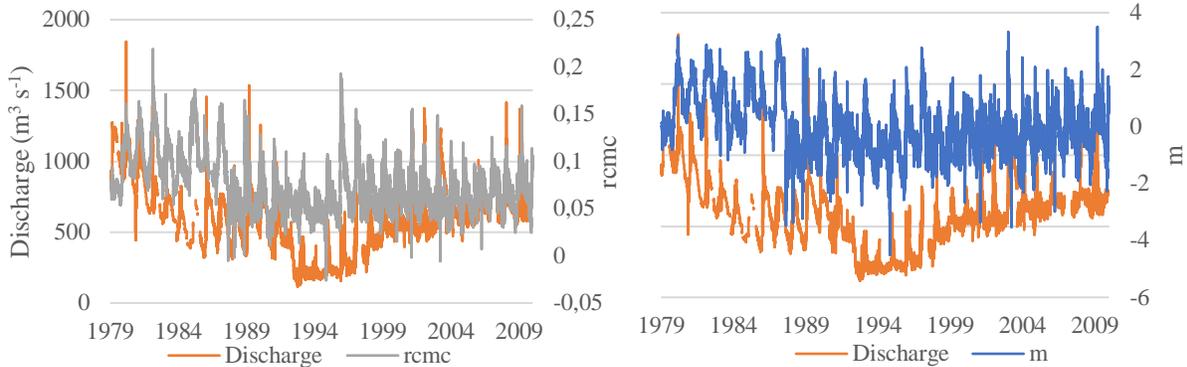


Figure 14: Observed discharge and downstream (cell C0) satellite signal at Chikwawa. $rcmc$ = CMC-ratio, m = m index. Data sources: Brodzik (2016); National Hydrological Services of Malawi (2019)

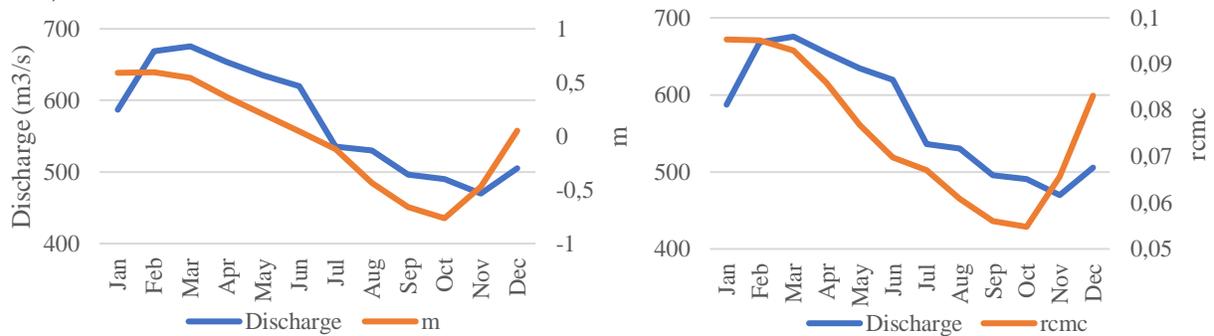


Figure 15: Mean monthly discharge and downstream (cell C0) satellite signals for Chikwawa. $rcmc$ = CMC-ratio, m = m index. Data sources: Brodzik (2016); National Hydrological Services of Malawi (2019)

In Karonga, the seasonal oscillations of satellite signals and discharge data do not coincide (Figures 16-17), as the peak of the satellite signals occurs later in the year than the discharge peaks, and the $rcmc$ values are unusually high. The reasons behind this phenomenon will be discussed in Section 5.2.1. K2 showed a better correspondence of seasonality (Figure 17, Appendix B1) and more realistic $rcmc$ -values than K0 and K1. It was therefore decided to use K2 as a proxy VGS for the Karonga downstream area of interest in this research.

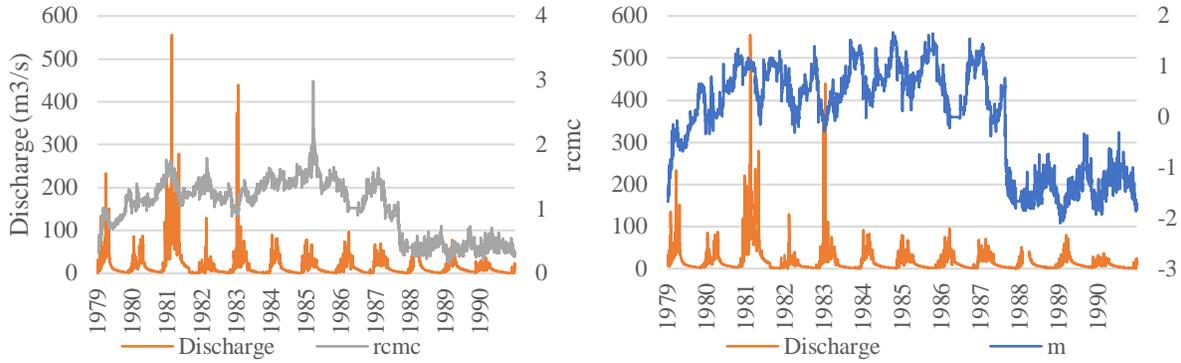


Figure 16: Observed discharge and downstream (cell K0) satellite signals at Karonga. $rcmc = CMC$ -ratio, $m = m$ index. Data sources: Brodzik (2016); Global Runoff Data Centre (2020).

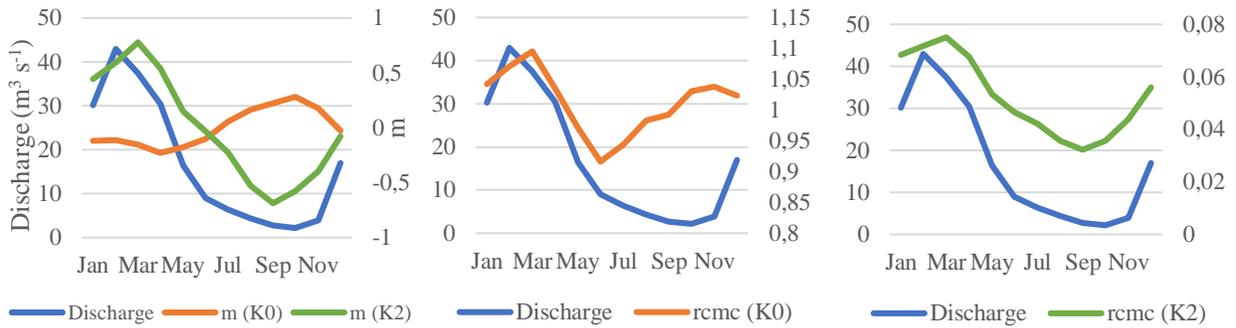


Figure 17: Mean monthly discharge and downstream (cell K0) satellite signal for Karonga, including data from cell K2 as this will be used as a proxy for K0 due to the proximity of Lake Malawi to K0. $rcmc = CMC$ -ratio, $m = m$ index. Data sources: Brodzik (2016); Global Runoff Data Centre (2020).

4.2.2 Relationship downstream signals and discharge

Figures 18 and 19 show the 2nd-degree polynomial regression applied to the discharge data at both locations. For Chikwawa, both m and r_{cmc} depicted moderately strong correlations between discharge and satellite signals (Figure 18). For Karonga, the correlations were very weak (Figure 19). All correlations yielded a p-value of smaller than the α of 0.05 and were therefore considered significant. The results of the regression based on the test data show that the model severely underpredicts higher discharge values in Karonga (Figure 19), which is likely the product of the insufficient fit of the regression line to the training data.

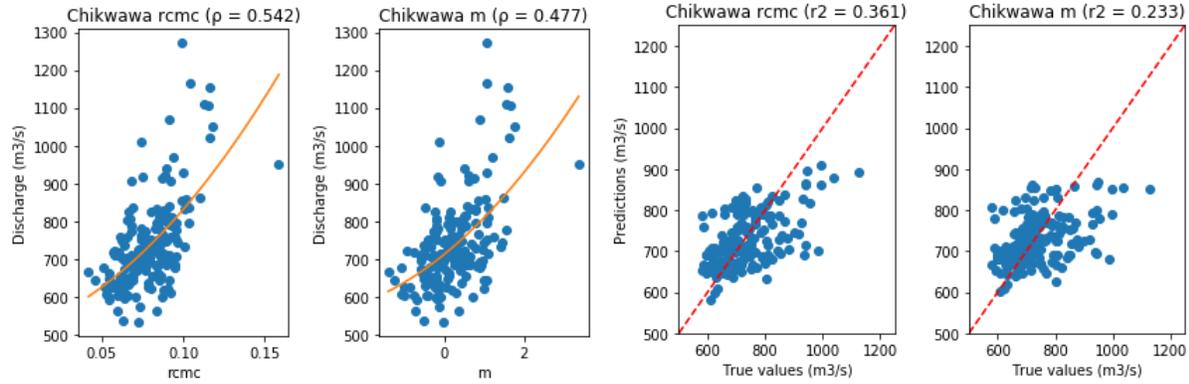


Figure 18: Spearman’s correlation (ρ) and regression results of satellite signals and discharge for Chikwawa (C0-cell). rcmc = CMC-ratio, m = m index. Data sources: Brodzik (2016); National Hydrological Services of Malawi (2019)

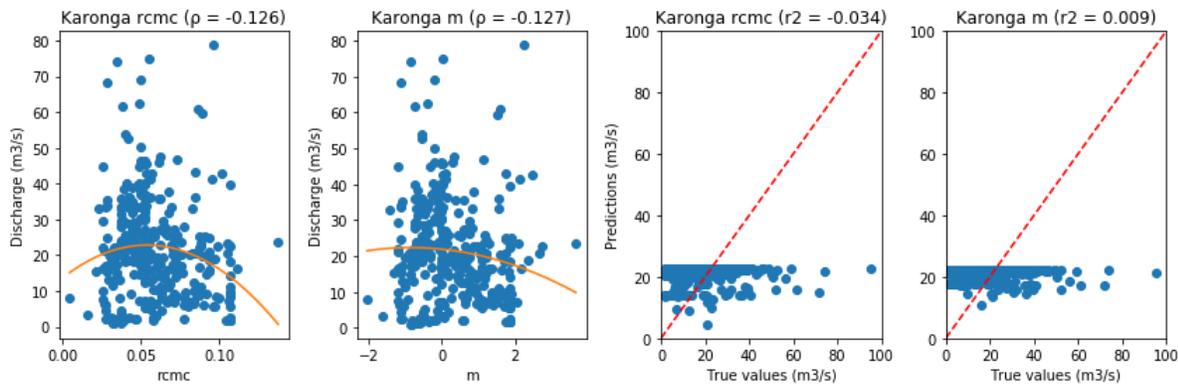


Figure 19: Spearman’s correlation (ρ) and regression results of satellite signals and discharge for Karonga (K2-cell). rcmc = CMC-ratio, m = m index. Data sources: Brodzik (2016); Global Runoff Data Centre (2020).

4.3 Detection of flood occurrence

4.3.1 Flood database quality

In Chikwawa, a total of 15 floods were registered, of which 12 occurred during the period of the available discharge data (1978-2009), and three outside of this period. Of the data within the discharge period, one flood took place during a data gap, and three took place in a period where the annual discharge peak did not take place (indicating that either no flood took place, or that it was not the most severe event of the year). The resulting 8 reported floods in the database (i.e. 67% of the reported floods within the discharge period) took place during annual discharge peaks at the downstream area of interest. In Karonga, the quality of the flood records could not be validated, as the floods occurred outside the timespan of the discharge dataset.

4.3.2 Flood threshold

Due to the limited strength of the correlation between high discharge values and satellite signals, the satellite thresholds were defined with an extreme values analysis and not with the regression relationship. The plots used to identify the satellite and discharge values with a 5-year rp are found in Figure 20. The curves show that in Chikwawa, the flood threshold was defined as a discharge of $1349 \text{ m}^3 \text{ s}^{-1}$, which corresponds to an r_{cmc} of 0.169 and an m of 2.677. In Karonga, the flood threshold was defined as a discharge of $233 \text{ m}^3 \text{ s}^{-1}$, with satellite thresholds 0.142 for r_{cmc} and 2.983 for m .

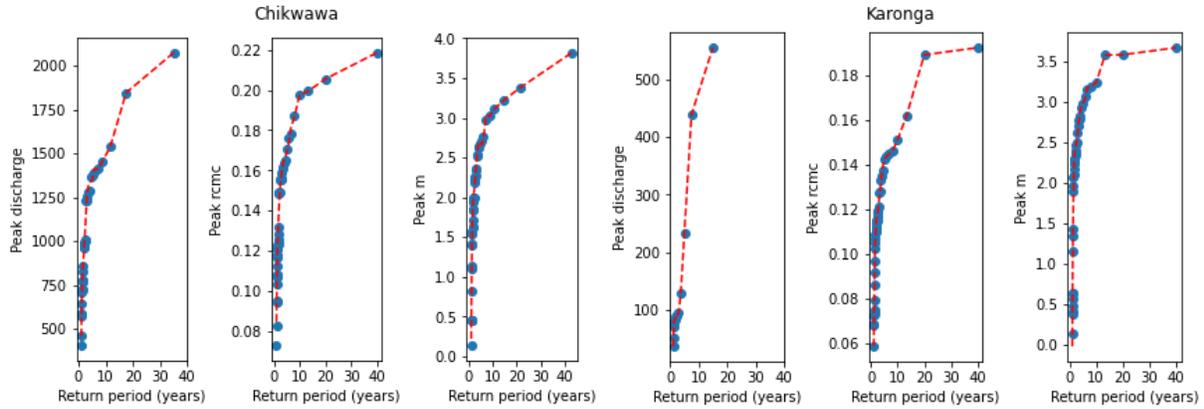


Figure 20: Peaks of the CMC-ratio (r_{cmc}), m index (m) and discharge per hydrological year vs. their respective return periods in years for Chikwawa and Karonga. The curve in red is a 10th degree polynomial curve through the data points. Data source: Brodzik (2016).

4.3.3 Contingency

For 3 of the 13 floods in Chikwawa, including the severe flooding in 2015, the rp could not be calculated due to the absence of discharge data. A rp of 5 years or higher was observed in 4 of the 13 registered floods, based on their maximum discharge value (events with * in Figure 21).

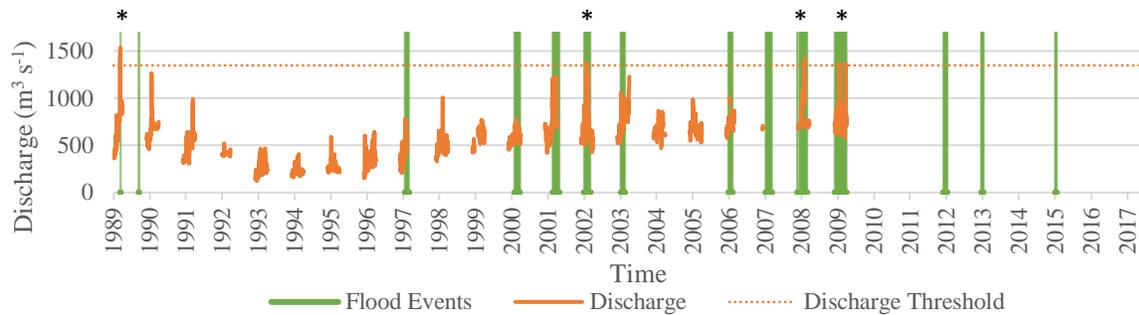


Figure 21: Discharge in wet season measured at the Chikwawa gauging station and registered floods in Chikwawa, green lines with * denote floods with a return period (rp) of 5 years or higher. Data sources: National Hydrological Services of Malawi (2019); Dartmouth Flood Observatory (2020); GLIDE (2020); Guha-Sapir (2020).

Flash floods that were not registered in the flood database could have impacted the average daily discharge and increased the number of ‘Misses’ if the flash flood would not be visible the satellite signal. However, Figure 21 shows that this is not the case: All exceedances of the threshold are present during registered flood events. In the C0-cell in Chikwawa, both m and r_{cmc} exceeded the 5-year threshold in 5 flood seasons between 1989 and 2017 (Figure 22). m scored better than r_{cmc} in the detection of flood events with a rp larger than 5 years (Tables 4-5). When looking at all registered floods rather than the floods where $rp \geq 5$, the detection skill went up for both r_{cmc} and m .

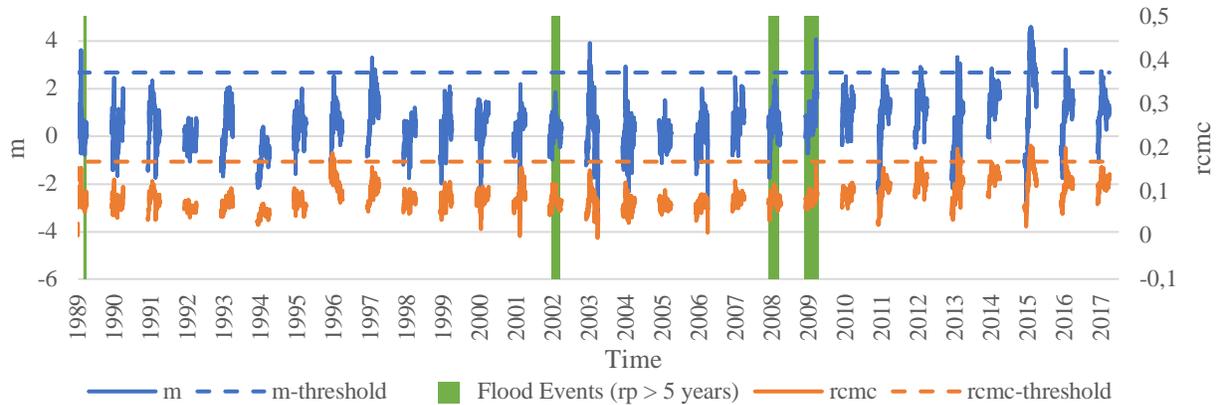


Figure 22: Satellite signals in wet season and registered floods with a return period (rp) of 5 years or more in Chikwawa (cell C0). $rcmc$ = CMC-ratio, m = m index. Data sources: Brodzik (2016); Dartmouth Flood Observatory (2020); GLIDE (2020); Guha-Sapir (2020)

In the K2-cell in Karonga, $rcmc$ exceeded the 5-year threshold in 2 flood seasons between 1978 and 2017, and m exceeded the threshold in 7 seasons (Figure 23). Most flood events in this period occurred at or near peaks in the satellite signal, although only 2 and 0 registered flood events from the database were simulated by the signal m and $rcmc$, respectively (Table 4). The rp values of these floods could not be calculated.

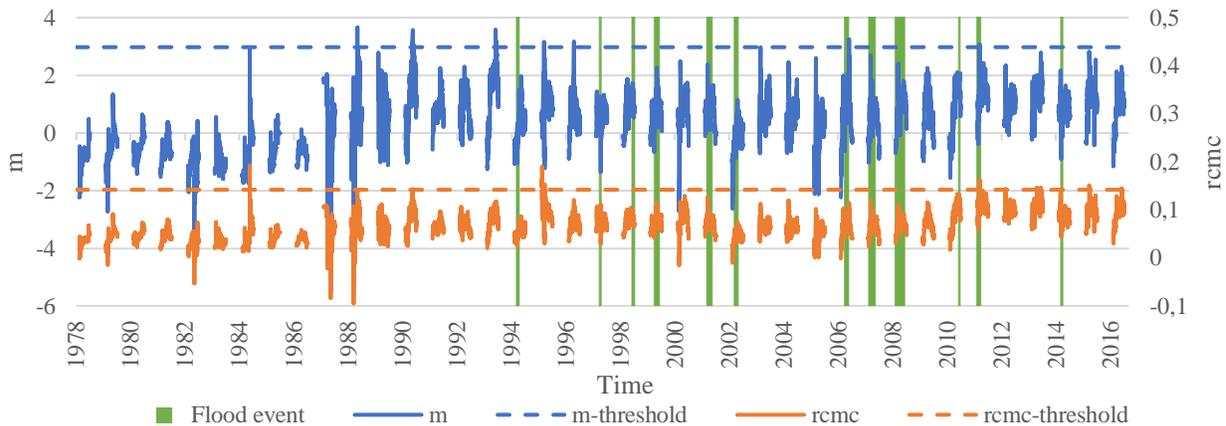


Figure 23: Satellite signals in wet season and registered floods. $rcmc$ = CMC-ratio, m = m index. Registered flood return periods could not be calculated due to the absence of discharge data. Data sources: Brodzik (2016); Dartmouth Flood Observatory (2020); GLIDE (2020); Guha-Sapir (2020)

The POD, FAR and CSI calculated from the confusion matrix can be found in Table 5. For both locations, the lowest FARs, highest PODs and highest CSIs were found in the simulations done with m .

Table 4: Confusion matrix for the satellite signals at cells K2 and C0. 1 entry corresponds to 1 flood season between 1978 and 2017. ‘Observed’ refers to events reported in the impact database, and ‘Modelled’ refers to events simulated by exceedances of the trigger threshold. r_{cmc} = CMC-ratio, m = m index, rp = return period.

Flood Events	A: Observed K2 (North Rukuru)		B: Observed C0 (Shire)		C: Observed C0 ($rp \geq 5$) (Shire)		
	Yes	No	Yes	No	Yes	No	
Modelled (m)	Yes	2	5	4	1	1	7
	No	10	23	9	26	3	29
Modelled (r_{cmc})	Yes	0	2	2	3*	0	5
	No	12	26	11*	25	4	31

* One year counted double because floods were observed and modelled, but not within 14 days of one another

Table 5: Success metrics for the satellite signals at cells C0 and K2. r_{cmc} = CMC-ratio, m = m index, rp = return period, POD = Probability of Detection, FAR = False Alarm Ratio, CSI = Critical Success Index.

	A: Karonga (North Rukuru)		B: Chikwawa (Shire)		C: Chikwawa ($rp \geq 5$) (Shire)	
	m	r_{cmc}	m	r_{cmc}	m	r_{cmc}
POD	0.167	0.000	0.308	0.154	0.250	0.000
FAR	0.714	1.000	0.200	0.600	0.875	1.000
CSI	0.118	0.000	0.286	0.125	0.091	0.000

4.3.4 Comparison with GloFAS

The simulated GloFAS-values (2000-2018), extracted from the hindcasting NetCDF-file (Harrigan et al. 2020) at the coordinates of Chikwawa show very high discharge-values compared to the observed values (Figure 24). Because of this, the 5 year rp -threshold of GloFAS was also relatively high: Approximately $3432 \text{ m}^3 \text{ s}^{-1}$, based on discharge peaks from the wet seasons in these years. The corresponding confusion matrix can be found below (Table 6). If only the flood events with an rp of at least 5 years were counted, the POD of GloFAS was 0, the FAR was 1 and the CSI was 0 (Table 6B). If all floods within 2000-2018 were included, the POD was 0.182, the FAR 0 and the CSI 0.182 (Table 6A).

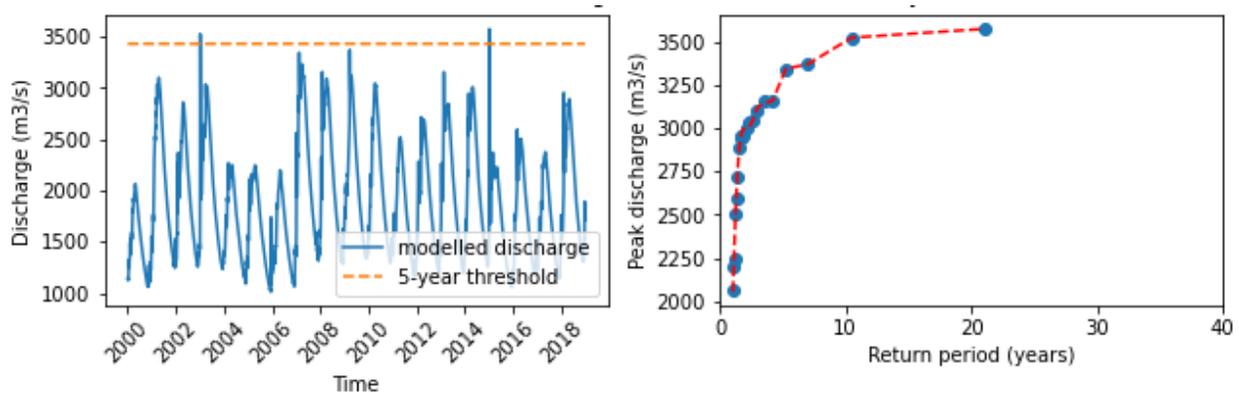


Figure 24: Discharge as simulated by the Global Flood Awareness System (GloFAS) and peaks per hydrological year vs. return periods for Chikwawa (cell C0). Data source: Harrigan et al. (2020)

Table 6: Confusion matrix for the Global Flood Awareness System (GloFAS) at cell C0 in Chikwawa. 1 entry corresponds to 1 flood season between 2000 and 2017. ‘Observed’ refers to events reported in the impact database, and ‘Modelled’ refers to events simulated by exceedances of the trigger threshold. rp = return period.

Flood Events	A: Observed Chikwawa		B: Observed Chikwawa ($rp \geq 5$)	
	Yes	No	Yes	No
Modelled (GloFAS)	Yes	2	0	2
	No	9	7	12

4.4 Relationship between upstream and downstream satellite signals

Not all VGSs showed a positive, Spearman's correlation coefficient (ρ) curve, that reached a peak near or after the 0-day lag time, as would be the case for VGSs that simulate downstream values well (Figures 25-26). The lines in the graphs correspond to the VGS cells depicted in Figure 7.

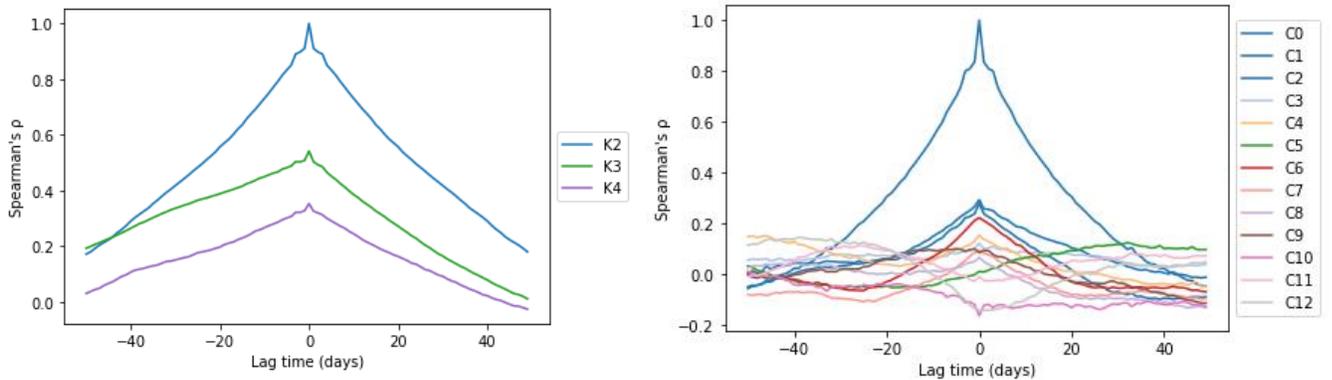


Figure 25: Non-parametric time-lagged cross correlation of upstream (cells K3-K4; cells C1-C12) and downstream (cell K2; cell C0) m along the North Rukuru (left) and Shire River (right). Correlations have been performed on stationary data.

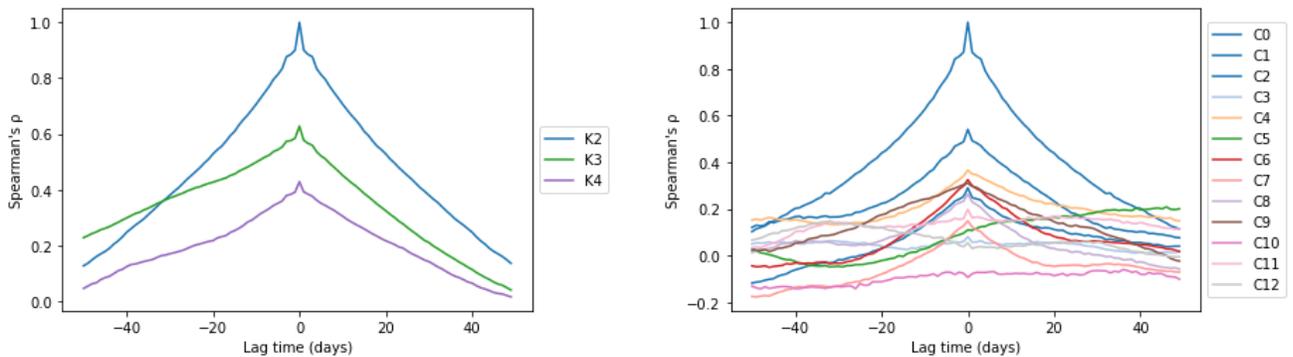


Figure 26: Non-parametric time-lagged cross correlation of upstream (cells K3-K4; cells C1-C12) and downstream (cell K2; cell C0) CMC-ratio (r_{cmc}) along the North Rukuru (left) and Shire River (right). Correlations have been performed on stationary data.

The lag times where the maximum ρ occurs (i.e. the optimal lag times) per VGS cell are found in Table 7. The highest upstream correlations were found at cells C1 for the Shire, and K3 for the North Rukuru, both with a lag time of 0 days. None of the VGSs in this study showcased a positive lag time of more than 0 days,

whereas that would have been realistic based on their distance from the gauging station. Instead, lag times either tended towards negative or positive extremes that are not realistic for their location upstream or remained close to 0. Therefore, it was decided to proceed with the downstream VGSs K2 and C0 themselves for the forecasting component of this research, and not with upstream gauges. This way, it was investigated whether real-time satellite signals could be used to predict future signals for flood forecasting.

Table 7: Results of the non-parametric time-lagged cross correlation between satellite and discharge data for the Shire (C*) (1979-2009) and North Rukuru (K*) (1979-1990). r_{cmc} = CMC-ratio, $m = m$ index, ρ = Spearman’s correlation coefficient. The optimal lag time is the lag time at which ρ is highest

VGS	Correlation with downstream r_{cmc}		Correlation with downstream m	
	Optimal lag time (days)	max ρ	Optimal lag time (days)	max ρ
C0	0	1.000	0	1.000
C1	0	0.541	0	0.279
C2	0	0.290	0	0.293
C3	0	0.081	0	0.121
C4	0	0.366	0	0.152
C5	46	0.209	32	0.124
C6	0	0.325	0	0.221
C7	0	0.149	0	0.095
C8	0	0.264	0	0.065
C9	0	0.310	-5	0.099
C10	36	-0.060	-50	0.010
C11	0	0.198	-26	0.113
C12	-32	0.148	-39	0.144
K2	0	1.000	0	1.000
K3	0	0.628	0	0.542
K4	0	0.429	0	0.353

4.5 Forecasting model

4.5.1 Submodel performance

Figure 27 shows the RMSE and the R^2 of the different submodels (over the period 1998-2017) as a proxy for model performance. When a long lead time-submodel was used, the RMSE of the predictions of the satellite signals vs. the observed (test) satellite signals was higher and the R^2 of the model was lower than when a short lead time-submodel was used. The light blue lines in Figure 27 indicate the thresholds for the acceptable performance (see Section 3.6.3). A well-performing submodel should have an R^2 exceeding this limit, and an RMSE staying under this limit. As the RMSE of r_{cmc} was far below the acceptable limit at both locations, no blue line is visible in the upper right plots of Figures 27A and 27B, that depict the RMSE of r_{cmc} . At Chikwawa, R^2 remained under the thresholds at all lead times for both r_{cmc} and m (Figure 27A), indicating the performance of the submodels is not satisfactory. r_{cmc} generally achieved higher R^2 -values and lower RMSE-values than m . In Karonga, the R^2 -values were almost identical for r_{cmc} and m , and the submodels performed with an R^2 above 0.7 up until a lead time of 1 day (Figure 27B).

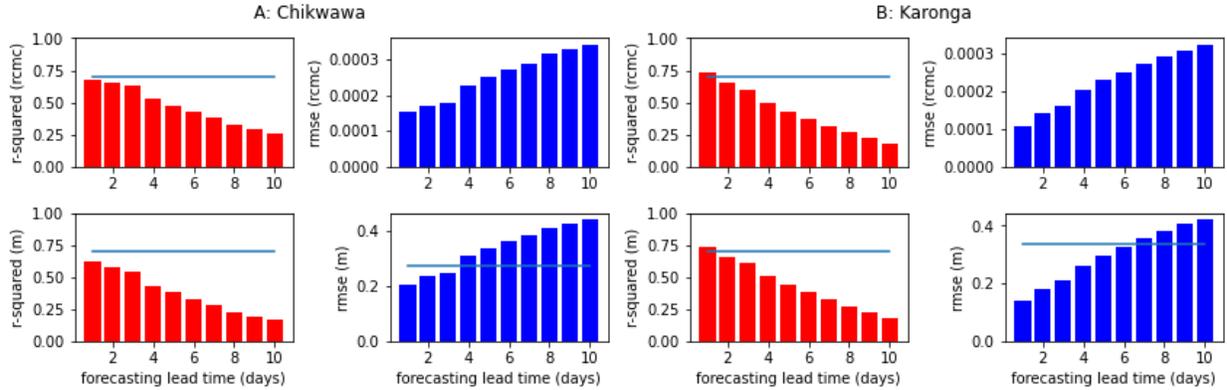


Figure 27: Performance of the submodels (predictions vs. test observations) for the m index (m , bottom) and the CMC-ratio (r_{cmc} , top) in Chikwawa (A) and Karonga (B). RMSE = root mean squared error.

4.5.2 Main model forecast performance

The performance of the main model in the wet season of 2015 is presented in Figures 28 and 29. The values are forecasted at a lead time up to 4 days, combining the forecasted values in timeseries. As could be observed in Figure 27, r_{cmc} proved a better predictor than m , which is why the forecasts presented here are based on r_{cmc} . Forecast skill at both locations decreases with an increasing lead time, as can be observed both visually (Figures 28A, 29A), in the graph showing the R^2 (Figures 28B, 29B) and the RMSE (Figures 28C, 29C). In Chikwawa, the overall R^2 only exceeds the ‘acceptable limit’ of 0.7 up until a lead time of 2 days. In Karonga, the overall R^2 does not exceed the threshold at all. The RMSE stayed under the ‘acceptable limit’ of 10% of the maximum occurring value within the validation dataset at each lead time at both locations. The RMSE in Karonga is lower, however, likely due to the fact that the observed r_{cmc} values were lower as well.

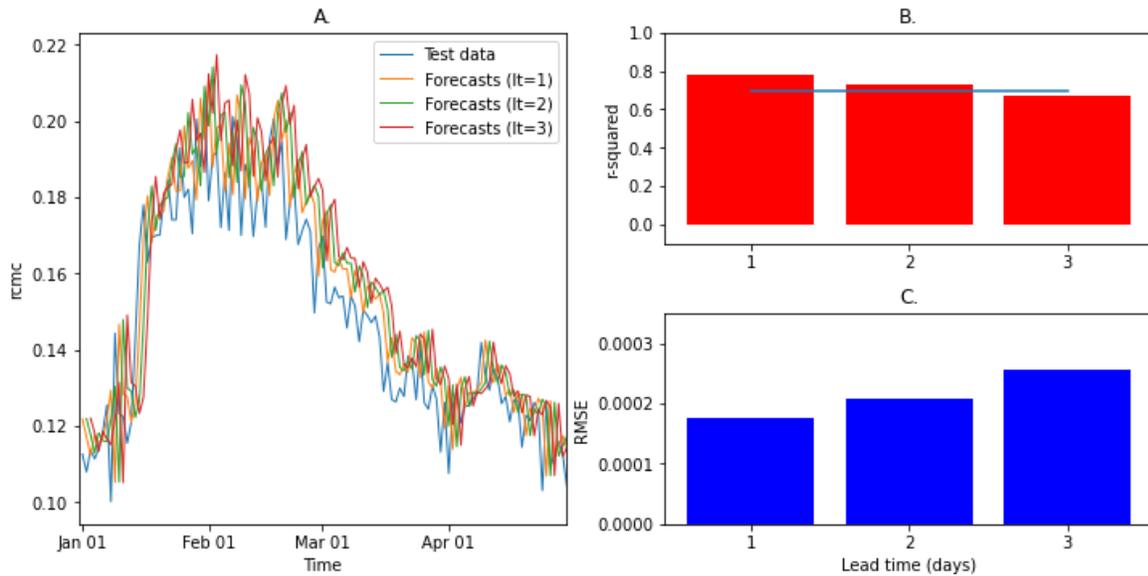


Figure 28: Main model forecast performance for Chikwawa in the 2015 wet season. A: Satellite signal forecasts vs. observed signals. B: R^2 vs. lead time (days). C: RMSE versus lead time (days). l_t = lead time, $rcmc$ = CMC-ratio, RMSE = Root Mean Squared Error

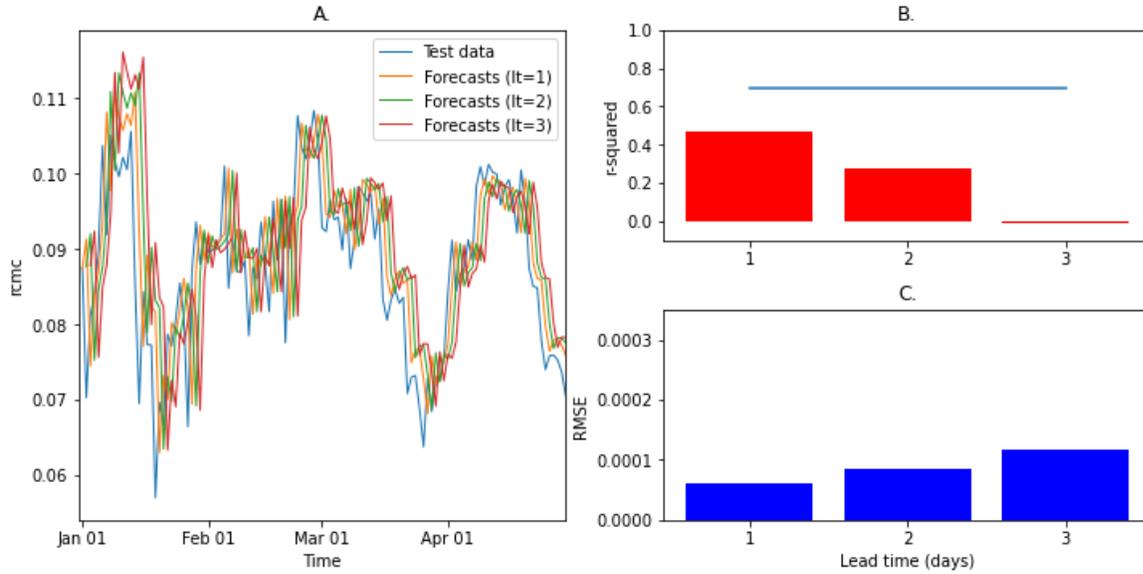


Figure 29: Main model forecast performance for Karonga in the 2015 wet season. A: Satellite signal forecasts vs. observed signals. B: R^2 vs. lead time (days). C: RMSE versus lead time (days). lt = lead time, $rcmc$ = CMC-ratio, $RMSE$ = Root Mean Squared Error

In practice, the model presented in this project will be used at a certain time (t), combining the different lead times in one forecast, and will therefore not generate timeseries as presented in Figures 28 and 29. The model's ability to simulate a flood event in an operational setting is illustrated with an example of the 2015 floods (Figure 30), an event for which no discharge data was available, but an event that has been documented across databases and literature (e.g. Guha-Sapir (2020)).

The future values can be forecasted with the PMRS-model up to nearly three days ahead, due to the sufficient skill at these lead times (Figure 28B). Figure 30 therefore shows a 3-day forecast for r_{cmc} in Chikwawa from the time t , which is January 13. Unfortunately, the model does not forecast an exceedance of the trigger threshold on January, whereas this is indeed visible in the detection system (blue line). Instead, the forecasted values do increase from the values observed at time t , but not sharply enough. The first date on which the confidence interval of the model did exceed the threshold was January 15, one day before the actual exceedance took place. Hence, in this particular case study, the model does not seem to foresee the sharp increase observed between January 13 and January 16, and the system would therefore have not given out an early warning for this flood.

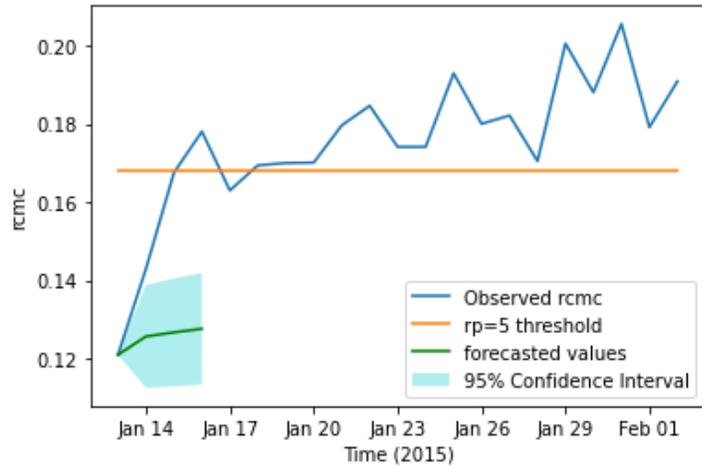


Figure 30: Performance of model in forecasting exceedance of the trigger threshold during the 2015 floods. *rp* = return period, *rcmc* = CMC-ratio

5. Discussion

The purpose of this study was to assess whether PMRS could be used for FbF-purposes in Malawi. The ability of the PMRS-derived ratios to both detect and forecast flood events was assessed for the potential use in Early Warning Systems in an FbF-context.

5.1 Raw T_b signals

5.1.1 Break points in T_b -timeseries

The raw T_b data form the basis of this research, measured in a wet calibration cell (C_w), a dry calibration cell (C_d) and a measurement cell containing a river (M), which in our case represents our VGSs. In the timeseries containing the extracted T_b -signals from both locations, points where a transition to a different satellite sensor/platform was made are clearly visible in the form of amplitude shifts, for example in 1987 (Figure 12). This can be attributed to the fact that the MEaSUREs-dataset is resampled from different satellite overpasses and platforms, and the resampled cells correspond to slightly different ‘real’ ground footprints for each satellite. The aforementioned phenomenon is especially apparent in the VGS in Karonga (K0), which is bordered by a large water body, making the slight different footprints show very different signals depending on whether the true footprint was taken over the lake or not. The upward and downward shifts in signal are mostly cancelled out by converting C_d and M into a ratio such as m or r_{cmc} as visible in Figures 14 and 16, since the ratios represent the relative difference between cells from the same platform in the same period. However, in Karonga especially, a shift remains visible in 1987 after conversion to a ratio, where the Nimbus-satellite was merged with the first satellite of the DMSP-series (Figure 16). The reason behind this is unknown, although it might have to do with deterioration of the Nimbus satellite sensor, causing less sensitivity, or the amount of water in the real ground footprint of the data prior to resampling. The amplitude shift could have a negative impact on the PMRS-method, because it might change the satellite signal’s relationship with discharge through time, or cause more threshold exceedances in earlier years.

5.1.2 T_b -signal assessment

In an ideal case study, the C_d -signal and M-signal have a similar signal when no flood event is taking place. The r_{cmc} and m -signal would ideally only show a spike if the C_d -signal and M-signal differ, indicating a flood event. The proximity of the C_d - and M-lines in Figure 12 indicate that this is roughly the case for cell C0 in Chikwawa. Whereas the lines are not perfectly aligned, they follow a similar variability and a distinctly different oscillating pattern and seasonal amplitude when compared to the C_w -cell. This latter cell shows a sharp peak in T_b in 1996 (Figure 12), which showed to impact the r_{cmc} -signal in this year substantially (Figure 14). This can be connected to severe receding of Lake Chilwa in 1995-1996 (Njaya et al. 2011), causing the T_b of C_w to rise and even approximate the T_b -value visible in M, which is supposed to be a much drier cell than C_w . In K0, which was initially intended as the downstream VGS for Karonga, the signal of M showed to be more similar in its oscillations as C_w (covering Lake Malawi) than to the oscillations of C_d (Figure 13). This was also the case for K1, the first best proxy alternative to K0 (Appendix B1). The VGSs K0 and K1 are situated so close to Lake Malawi, that its T_b value does not follow the seasonal cycle of the river, but of Lake Malawi (which also shows much larger fluctuations). As a result, the signal-to-noise ratio of the r_{cmc} and m -signals is relatively high, and the detection skill of the ratios calculated from these VGSs is low. This was the reason that K2 was eventually selected as a proxy downstream VGS for Karonga.

5.2 Flood detection

5.2.1 Relationship between m , r_{cmc} and observed discharge

The satellite signals and discharge data follow similar seasonal oscillations in the downstream VGS in Chikwawa (C0) (Figure 15). The interannual trend present in the discharge series is not visible in the satellite signals, however, due to the fact that the satellite indices are standardized relative to the prevailing hydrometeorological conditions at that point in time (Figure 14): m is relative to the annual mean and standard deviation, and r_{cmc} is relative to the T_b of the nearest driest and wettest cells (i.e. in a wetter year, C_d might show a 'wetter' signal' due to precipitation). A potential problem with this is that the relationship between discharge values and satellite signals will be different in dry and in wet years. This proved to be one of the main challenges encountered while performing a regression on the data in Chikwawa. When the last 5 years – which showcased relatively stable discharge values – were used in the regression, a positive relationship could be quantified between discharge and both r_{cmc} and m (Figure 18). The correlation coefficients were significant, yet only moderately strong ($\rho=0.542$ for r_{cmc} and $\rho=0.477$ for m). The relatively low R^2 shows that the model discharge predictions do not fit the test data very well and, judging from Figure 18, also underestimate high values, which are especially important for flood prediction.

In the original downstream VGS in Karonga (K0), the seasonal oscillations did not coincide at all (Figure 17), which can be attributed to the previously mentioned problems with the original T_b -signal in the M-cell. The result is an r_{cmc} -series that contains extremely high values compared to what was observed in Chikwawa (Figure 16), and that follows a seasonality that is radically different from the discharge-series in Karonga (Figure 17). When moving further upstream to the proxy VGS K2, the values and seasonal variation seems more realistic (Figure 17). The rest of the research project therefore focused on K2 as a proxy for the downstream area of interest in Karonga, and did not take into account K0 anymore. The regression analysis done on K2 shows that the second polynomial fails to identify the positive correlation that was expected between m/r_{cmc} and discharge. The regression line shows a downward slope ($\rho=-0.126$ for r_{cmc} and $\rho=-0.127$ for m), not covering the high discharge and satellite values in the training set. The scatterplots suggest this is likely not merely a statistical problem, as they show that the relationship between discharge and satellite signals in K2 is not as clear-cut as in C0; the highest discharge-values in Karonga were not found where the highest satellite values occurred, whereas this was roughly the case in Chikwawa. As a result of this, the regression on K2 has an R^2 of nearly 0, and the model severely underestimates discharge values (Figure 19). One possible explanation for the weak correlation at Karonga could be that the North Rukuru River and its flood plains are of a much smaller scale than the Shire River in both width, length (Figure 7) and discharge (Figures 14,16), meaning its contribution to signal changes in the relatively large PMRS-cell (25x25 km) is smaller than is the case for the Shire River.

When comparing average daily discharge data to T_b -related satellite indices, it is important to keep in mind that the measurements are temporally different. The satellite data are measured once a day, during the 'snapshot' of a single satellite overpass, whereas the discharge data are averaged values calculated from multiple sub-daily values. This may explain why the higher-end discharge values are not well captured in the regression relationship: Whereas a short-term discharge peak occurring on a given day will affect the average daily discharge dataset, the satellite overpass on that day might have occurred before or after the peak, meaning this short-term peak might not be represented in the satellite data in the same manner.

These findings can be compared to similar studies looking into PMRS flood forecasting in different or similar study areas (Table 8). Kramer (2018), who conducted a study in Chikwawa as well, using high-resolution (100x100m), commercially available PMRS data, found a Pearson's r of -0.784 when correlating an r_{cm} from this satellite data to discharge data. Whereas this is a stronger correlation than was found in this

study, it was calculated using discharge data simulated from a global model rather than observed discharge, which complicates comparisons of results.

Brakenridge et al. (2007) introduced the r_{cm} with his research on a number of rivers, using a similar spatial resolution as was used in this study (Table 8), comparing it to average daily discharge for three locations, and monthly values for three others. The R^2 in our study cannot be directly compared to the ones in Brakenridge et al. (2007), because in our study, R^2 was calculated based on the relationship between predicted and observed discharge, rather than the relationship between observed discharge and the satellite signal as in Brakenridge et al. The scatterplots in Brakenridge et al. (2007) that also included daily rather than monthly discharge data do indicate a slightly better fit between satellite data and discharge than could be observed in Figure 18. A quick survey through the historical satellite imagery available on Google Earth Pro 7.3.3 shows that the rivers where the experiment of Brakenridge et al. (2007) was done range in width from 150 to 350 metres (Wabash River), and 100 to 150 metres (Red River). This is a similar size to the Shire River, although smaller floodplains are visible on the satellite imagery near Chikwawa compared to the study areas in the US. As flooding of floodplains increases the surface water fraction and therefore the value of the satellite indices considerably, this could explain why the data from the United States is slightly more sensitive than the Malawian data to higher discharge values and therefore shows stronger correlations. In another study on PMRS flood detection, De Groeve (2010) reports “excellent correlation” between discharge values and m . However, no correlation coefficient is reported in their study, complicating a comparison with the results from our study. Van Dijk et al. (2016) conducted a PMRS-experiment on a global scale, correlating simulated water extent to monthly discharge. Both the parametric- and non-parametric correlation coefficient of the correlation between observed and simulated discharge were 0.2 on average, although this varied from region to region, and Africa was underrepresented in this study compared to Europe and North America. The stations that were in East-Africa showcased particularly high coefficients in this study, larger than 0.9. Keunen (2020), who researched the potential of PMRS in the Zambezi River basin using similar methods as in this study, observed a ρ of 0.86 between r_{cm} and discharge. Thus, whereas the Chikwawa PMRS-data shows a moderate positive correlation with discharge in Chikwawa, this correlation is relatively weak when compared to the results from other studies (Table 8).

Keeping in mind the ‘snapshot’ trait of the satellite data, this could explain why the R^2 is much higher at locations where average monthly discharge data was used rather than daily values (most notably at three locations studied by Brakenridge et al. (2007), see Table 8). When discharge data is averaged over a longer period, short-term peaks affect the final value less than when they are averaged over a day, reducing the occurrence of extremely high discharge values. And, as Figure 18 shows, the extreme values are captured less well by the regression. Therefore, it can be expected that the R^2 would be higher when the regression would look at monthly hydrograph data as was done for three locations in Brakenridge et al. (2007).

Table 8: Summary of relationships between passive microwave remote sensing indices and observed discharge found in existing literature. The upper section shows regressions, the lower correlations. (obs) Q = (observed) discharge, sim Q = simulated discharge, r_{cm} = CM-ratio, r = Pearson's correlation coefficient, ρ = Spearman's correlation coefficient

River	Country	Corr./reg. parameter	Value	Polynomial order	Sources
Shire	Malawi	R^2 (obs Q vs. sim Q) (r_{cm})	0.48	2	This study
Shire	Malawi	R^2 (obs Q vs. sim Q) (r_{cmc})	0.54	2	This study
Multiple	Global	r and ρ (obs Q vs. sim Q)	0.20 ¹	Various	Van Dijk et al. (2016)
Shire	Malawi	ρ (r_{cm} vs. Q)	0.23	2	This study
Shire	Malawi	ρ (r_{cmc} vs. Q)	0.36	2	This study
Wabash	United States	R^2 (r_{cm} vs. Q)	0.63	4	Brakenridge et al. (2007)
Red	United States	R^2 (r_{cm} vs. Q)	0.74	4	Brakenridge et al. (2007)
Siret	Romania	R^2 (r_{cm} vs. Q)	0.64	3	Brakenridge et al. (2007)
Niger	Mali	R^2 (r_{cm} vs. Q)	0.91 ¹	2	Brakenridge et al. (2007)
Mekong	Cambodia	R^2 (r_{cm} vs. Q)	0.97 ¹	3	Brakenridge et al. (2007)
Lena	Russia	R^2 (r_{cm} vs. Q)	0.96 ¹	3	Brakenridge et al. (2007)
Brahmaputra	Bangladesh	r (r_{cm} vs. Q)	$\sim 0.68^2$	1	Hirpa et al. (2013)
Ganges	Bangladesh	r (r_{cm} vs. Q)	$\sim 0.59^2$	1	Hirpa et al. (2013)
Zambezi	Zambia	ρ (r_{cm} vs. Q)	0.86	2	Keunen (2020)

¹ Monthly rather than daily discharge/satellite data used

² Maximum correlation found when using upstream VGSs to forecast downstream VGSs

5.2.2 Relationship between m , r_{cmc} and flood events

The metrics and graphs (Figures 22-23, Tables 4-5) show that the majority of registered floods still go undetected with the current threshold configuration. Whereas the registered floods occur at or near annual peaks in m and r_{cmc} , these peaks in satellite data often do not reach high enough to cross the 5-year rp threshold and be classified as a flood. This explains the relatively low POD and CSI, and the high FAR found for both locations and satellite indices (Table 5). In all cases where a peak did not occur directly within the period that was registered as a flood in the impact database (but within 14 days), the satellite peak occurred *after* the discharge peak.

When comparing the performance per satellite signal (r_{cmc} or m), the metrics show that m has a lower chance than r_{cmc} of under-reporting actual flood events (relatively high POD), a lower chance of misclassifying an event as a flood event (relatively low FAR), and a higher chance that a forecasted flood event was an actual flood event (relatively high CSI). This was the case for both locations. As r_{cmc} performed better in detecting absolute discharge values (see Section 5.2.1), this high flood event detection skill of m is interesting.

A comparison between the confusion matrices (Tables 4A-B) and success metrics (Tables 5A-B) of the two rivers, when taking into account all floods, showed that the POD and CSI were higher at C0 (Shire) than at K2 (North Rukuru), and the FAR was lower; Hence, the presented PMRS-method showed greater flood event detection skill and a lower rate of false alarms in the downstream cell of the Shire River than in the downstream cell of the North Rukuru.

For the Shire River, a confusion matrix was also constructed using just the floods that had an $rp \geq 5$ years (Table 4C, Table 5C). When comparing this table with the one for the Shire that included all registered floods (Table 4B, 5B), the POD is higher and the FAR is much lower for the table including all floods. As

the calculated thresholds were based on an rp of 5 years, these relatively low scores for $rp \geq 5$ (Table 5C) can indicate two things:

1. The calculated 5-year threshold for discharge is not accurate. With a lower threshold, Figure 21 would have more floods marked as having an rp of at least 5 years, leading to more observed flood entries into Tables 4C and 5C and possibly a higher number of ‘Hits’.
2. And/or: The threshold configuration for the satellite signals is not accurate. A lower threshold would include more events as a ‘Hit’ and less as a ‘Miss’

One thing one has to keep in mind when comparing Tables 4B to 4C and Tables 5B to 5C is that the analysis in Tables 4B and 5B included many more observed flood events than in Tables 4C and 5C. The probability of achieving a ‘Hit’ or ‘Miss’ was therefore lower for the latter confusion matrix. Comparisons between the two Shire confusion matrices are therefore biased. Whereas the use of success metrics (POD, CSI, FAR) should standardize some of this bias, research has shown that the CSI is still highly dependent on event frequency (Schaefer 1990), which means the comparison should still be made with caution.

An adjustment of the threshold using more advanced rp calculations could potentially improve the metrics for both the North Rukuru and Shire. As described in Section 3.5.1, the flood threshold for a 5-year rp was set using a relatively simple extreme value analysis, while this threshold is one of the primary factors influencing the contingency and success metrics presented in Table 5. The accuracy of the discharge thresholds can be assessed by comparing it to those used in existing literature. Teule (2019), for example, who assessed the hydrological skill of the global model GloFAS in the Shire River Basin, used percentiles rather than an extreme value analysis as applied in this study to define a flood threshold, and the author also studied different return periods than $rp=5$. This approach is also applied by the Red Cross in Zambia, where the GloFAS model is embedded in the EAP. On the other hand, Jury (2014) defines floods in Chikwawa as events with a discharge of at least $1200 \text{ m}^3 \text{ s}^{-1}$. Whereas no rp was mentioned with this threshold, it still is substantially lower than the $1349 \text{ m}^3 \text{ s}^{-1}$ found in this study. Based on these findings, we suggest more research is done on correctly setting the flood threshold of $rp=5$.

In Chikwawa, the gauge records and flood database only overlapped partially, meaning some registered floods were omitted from the $rp \geq 5$ analysis as no discharge records were available. One of these events is the severe 2015 flooding, which is in fact detected by m and r_{cmc} , but the event is not included in the confusion matrix with $rp \geq 5$. Lack of discharge data meant that its rp could not be calculated in this study. However, it is estimated by the government of Malawi that the flood event had a rp of 500 years (Government of Malawi 2015), meaning this is an example of a ‘Hit’ that went unreported in Table 4C.

5.3 Flood forecasting

5.3.1 TLCC

The TLCC was conducted to assess which VGS, if any, could be used as predictor in the forecasting model. The assessment showed that only a few upstream VGSs showed a strong positive correlation with the satellite signal observed at the downstream VGS (Figures 25-26). The strongest correlations with C0 for the Shire River were found in VGS C1 (r_{cmc}) and C2 (m). Along the North Rukuru, the correlations with K2 were strongest for K3 for both indices. In the VGS cells along the North Rukuru, only two upstream VGSs (K3, K4) were available due to the use of K2 as a proxy for K0 as downstream VGS.

Thus, VGSs situated close to the downstream VGS generally showed the highest correlations, although this does not mean that correlation strength constantly decreased with distance. An example is C9, that showed a relatively high ρ and an optimal lag time of 0 for r_{cmc} , but is situated further upstream than C7, which had a lower ρ . This suggests that more factors than the main river discharge influence the m and r_{cmc} signals,

such as surface characteristics (flood plain width, geomorphology) or the presence of other rivers in cells. When comparing the two satellite signals, the r_{cmc} exhibited slightly stronger upstream-downstream correlations than m in the majority of VGS cells along the Shire River and the North Rukuru River.

A secondary aim of the TLCC was to assess whether time lags were present between changes in the upstream and downstream satellite signals. The lines in Figures 25 and 26 that show a maximum ρ at a lag time of 0 days, correspond to VGSs that show the highest correlation when used in a detection setting (i.e. their optimal lag time is real-time). This means that the delay between upstream and downstream satellite signals is on average not longer than 24 hours. The majority of the downstream *and* upstream VGSs showed this pattern. This is in contrast with the findings of Hirpa et al. (2013), who used upstream VGSs to forecast the discharge at a downstream point. The authors of this study did find positive lag times when using VGSs further upstream.

That the downstream VGSs C0 and K2 showed a lag time of 0 days, is in line with the findings of Brakenridge et al. (2007). The authors found a lag time at some of the downstream VGSs studied, but no lag time at others. They suggest that this has to do with the geography of the river floodplains; some floodplains are situated in flat terrain, and flooding is induced by local precipitation or snowmelt. Here, discharge and the satellite signal rise synchronously. However, floodplains with a steeper gradient situated in drier conditions have their inundation governed more by small-scale topography, hydraulic connectivity and local resistance to flow, which may lead to positive lag times. The immediate response of the satellite signals in C0 and K2 indicate that the North Rukuru River and the Shire River belong to the first group described.

An optimal upstream VGS as predictor for the forecasting system should have a positive lag time (>0) and a relatively high and positive correlation coefficient. Unfortunately, none of the VGSs complied with these criteria. This means we could only use the downstream VGSs K2 and C0 as predictors, and not the upstream VGSs. The results of these forecasts are discussed in the next section.

Of course, this study looks at a long-term timeseries of data, wherein surface characteristics and climate have changed as well, and the relationships found in the TLCC are dependent on the datapoints included in the analysis. The TLCC was implemented using the data from the complete study period 1978-2017, and data points from only the wet season. When only a few recent years are studied, or data over the whole year are included, the ρ coefficients, lag times and therefore the suitability of the upstream VGSs may be different as well.

5.3.2 Submodel forecast performance

From Figure 27, it can be derived that as the lead time of a forecast increases, the R^2 decreases. This means that submodels with a longer lead time have increasingly less variability of their forecasts around their mean. Furthermore, the RMSE of the model increases with longer lead times, meaning submodels with longer lead times describe the magnitude of the observed values increasingly less accurately as well. This outcome was as expected, as long-term forecasts are generally associated with more uncertainty than short-term forecasts. r_{cmc} showed a better performance at longer lead times than m in Chikwawa, but at Karonga, both satellite indices showed a similar score. If a critical threshold of an R^2 of 0.7 is maintained, as visible in Figure 27, the submodels themselves can only perform forecasts up to 1 day in advance for Karonga, and 0 days in advance in Chikwawa. However, the overall performance at different lead times of the main model will be different, as this model brings the trend back into the data and extrapolates the trend for future timesteps.

5.3.3 Main model forecast performance

Figures 28 and 29 show the performance of the model at each lead time individually, when the model follows the workflow depicted in Figure 10. Figures 28 and 29 show that the gap between forecasted and observed satellite values increases as the applied lead time increases as well. Based on the submodel results, this was expected. Whereas the submodels performed well ($R^2 > 0.7$) up to a lead time of 1 day in Karonga and not at all in Chikwawa, the main model paints a different picture. At Chikwawa, the maximum forecasting time has been increased to nearly 3 days (the R^2 at a lead time of 3 days is 0.67, which could still be considered sufficient if rounded) when the main model is used, whereas at Karonga, the main model does not show a sufficient R^2 at any lead time anymore. This suggests that the retrending step applied when creating the main model had a positive impact on forecasting skill at Chikwawa, and a negative impact in Karonga. It is important to note that a relatively simple detrending method was used to remove the trend from the data and extrapolate the trend at the day to be forecasted. Using more advanced methods may change the effect this step has on the forecasts. This skill assessment was conducted using the wet season of 2015 as training data due to the computational intensity of the experiment. A multi-year skill assessment might provide a more comprehensive picture of overall model skill.

The model's ability to simulate a flood event was illustrated with an example of the 2015 floods (Figure 30). The model does not foresee the sharp increase observed between January 13 and January 16 and therefore does not give out an early warning. The fact that the forecasted signal does not increase sufficiently likely has to do with the way the trend was estimated. Because the curve describing the trend does not perfectly fit the original data, the linear regressions described by the submodels is influenced by upwards or downward trends remaining in the 'stationary' timeseries. Further research could be done on how to extract the seasonality more accurately from the original satellite series.

5.4 Comparison with existing early warning systems

GloFAS is used in many national EWSs and is also currently used by the Red Cross in the EAP for Zambia. As it is a coupled ocean-atmosphere general circulation model, it has its strength in representing large-scale modes of variability such as the North Atlantic Oscillation (NAO) or the well-known El Niño–Southern Oscillation (ENSO), among others (Emerton et al. 2018). In many countries in SSA, including Malawi, teleconnections are less prevalent to absent and the forecasting skill of models such as GloFAS is therefore also lower. If no robust national EWS is present in those countries, PMRS provides a promising alternative.

5.4.1 Detecting discharge

The performance of our PMRS-model cannot be directly compared to discharges as simulated by the global model GloFAS: The relationship between r_{cmc} and discharge as found in our study is not strong enough to make a direct comparison. However, the statistical indices of the discharge estimation between GloFAS and the PMRS-estimation presented here can be compared. Teule (2019) did a regression analysis on the discharges as simulated by the global GloFAS model in Chikwawa with observed discharge data. The global model showed to have a correlation with discharge in Chikwawa, quantified by the Pearson's correlation coefficient (which can be interpreted similarly to Spearman's ρ) of 0.25. The R^2 was 0.06, indicating a lower skill than the statistics found for both r_{cmc} and m in this study. This suggests that in Chikwawa, the PMRS-method described in this study can more accurately estimate precise discharge values compared to GloFAS.

This previous point was confirmed while doing the extreme value analysis. The 5-year r_p threshold is more than double the threshold found for the discharge data. This is in line with the findings by Teule (2019), who also indicated that GloFAS tends to systematically overestimate discharge at this location.

5.4.2 Detecting events

Exceedance of trigger thresholds can also be compared directly, as GloFAS-data with a 0-day lead time is available for a cell surrounding Chikwawa (Figure 24). If the trigger thresholds for GloFAS are calculated the same way as the thresholds in this model were calculated (see Section 3.5.1), the GloFAS success metrics are slightly better than the PMRS-method when using r_{cmc} (Table 5B) if all floods are taken into account. If only the floods with a $rp \geq 5$ years are taken into account, the success metrics are the same as for r_{cmc} (Table 5C).

However, when compared to the PMRS-method when using m , the POD is much lower, and the FAR much higher. This shows that with the current methodology and in the current setting of Chikwawa and the Shire River, m performs better than GloFAS in detecting flood events and it generates fewer false alarms when used as an EWS. Caution should be taken when comparing the metrics, as the total number of floods within the GloFAS study period is lower than the period used to construct the confusion matrices in Table 4.

5.4.3 Practical implications

Unfortunately, no historical data on ODSS, the EWS currently in place in the Lower Shire River Basin, was available. Therefore the discharge detection-skill of our model cannot be compared to that of ODSS, which would be a more relevant comparison applied to the current situation in the study area. What is interesting to point out, however, is that the maximum lead time of our model was established to be 2 to nearly 3 days, which is similar to the forecasting lead time of the ODSS system, but lower than the forecasting lead time of GloFAS.

Thus, when used in a detection setting, the PMRS-method presented in this study performs better than GloFAS in both detecting absolute discharge and flood events in Chikwawa. The forecasting skill could not be compared due to the lack of GloFAS forecasting data. However, even if the PMRS-method would have a higher forecasting skill, the use of GloFAS might still be preferred if the forecast is done with a long lead time. Furthermore, the applicability of the PMRS-method is very location-dependent: Whereas the detection skill for discharge is relatively high in Chikwawa (C0), this is not the case in Karonga (K2), and GloFAS might still present a better alternative in the K2-cell.

Despite having a shorter or equal lead time to ODSS, it could still be interesting to compare the forecasting skill of the proposed PMRS-method with that of ODSS. This is because ODSS is a relatively complex EWS: For the flow and flood forecasting, the system makes use of high-frequency (max. hourly) real-time rainfall and water level data and precipitation forecasts (Ammentorp and Richaud 2016). As many data-scarce regions do not have access to high-quality, high-frequency hydrological data, the realization of ODSS (or a comparable system) is not realistic. The presented PMRS-method only makes use of one source of PMRS-data to be operational (and possibly temporarily discharge data for validation), meaning the method could be an asset for these regions and in some cases (such as our study area in Chikwawa, along the Shire) a better alternative than GloFAS.

5.5 Assumptions and limitations of research

5.5.1 Suitability and representativeness of C_w and C_d

The way in which the calibration cells were chosen and the way this data was treated impacted the outcome of this study substantially. This is something that needs to be considered whilst interpreting the results presented in this work.

Firstly, the r_{cmc} method is based upon the use of a wet calibration target (C_w) covering a complete grid cell. Such a target of a sufficient size is not always present in the vicinity of the study area. For the case study of

Chikwawa, for example, the closest suitable target was located over Lake Chilwa, which is located approximately 125 kilometers from Chikwawa. Whether or not the T_b signals from C_w at Lake Chilwa and M at Chikwawa were affected by uncorrelated changes in vegetation, soil characteristics or rainfall distribution was not studied in detail, which is a shortcoming of this study, and a shortcoming of the use of the r_{cmc} method in this case study. Whereas the drying of Lake Chilwa in 1996 is recognizable in C_w and the resulting r_{cmc} for Chikwawa for that particular year, the discharge records do not show this pattern at all (Figures 12, 14), implying the Shire was likely unaffected by this change. Hence, this is an example of a situation where the large distance between M and C_w likely affected the accuracy of the analysis negatively. In the case of Karonga, C_w was located much closer to the VGSs. This, however, led to a whole different set of complexities: The similarity of the signal originating from the M - and C_w -cells indicates that M should also not be in too close of a proximity to a large water body, as the signal of the water body may influence what is observed in the M -cell (Figure 13).

Thus, whereas C_w should not be located too far from M , which could lead to geophysical, hydrological or biological influences impacting the r_{cmc} -signal such as in Chikwawa in 1996, C_w should also not be located too close to M , as the presence of a large water body in or near the M -cell impacts the signal-to-noise ratio of the T_b -signal in M . The choice to use C_w (and thus the r_{cmc} -signal rather than the m -signal) in an analysis should therefore be taken cautiously, especially in the case of coastal regions and should depend on the geography of the study area in question.

Secondly, the use of a dry calibration cell in the calculations of m and r_{cm} is based upon the simplifying assumption that there is no water present in the dry calibration cell C_d . As C_d will never completely be without surface water due to the presence of small streams and puddles even in the driest regions, it was therefore assumed in this research that the influence of small streams would be averaged out when observing the T_b of a large grid cell.

5.5.2 *Effect of filter method*

As mentioned in Section 3.3.1, a backwards-looking mean filter was applied to the raw T_b data prior to the calculation of the satellite indices. This filled gaps in the data and filtered out some signal noise. This filtering method was chosen, despite the fact that a Savitzky-Golay filter with the same window size did seem to conserve the day-to-day variability of the raw data more accurately than the mean-filter method (Figure 8) whilst filtering out high-frequency noise as well. However, the centered nature of the Savitzky-Golay window is not desirable in a forecasting setting, whereas the applied mean-filter made use of a backward-looking window.

Some tests that were done with the Savitzky-Golay-filtered r_{cmc} and m signal showed a much better correlation with discharge compared to the filter method applied in this thesis; Applying a different filter method to the raw data could therefore potentially significantly alter the outcomes of this study. However, whereas applying stronger filters to data can bend results more positively towards significance, the question remains when an advanced filtering method such as the Savitzky-Golay filter is justified, whether it can be used under operational conditions (such as triggering early action of humanitarian actors), and how it impacts the relevance of the satellite data in connection to the real hydrological relationships. More research should therefore be done on which filtering method is most optimal in the context of the use of PMRS for flood FbF.

5.5.3 *Hydrological complexities*

This research is built upon the assumptions that generally, a widening of a river upstream will lead to a widening of the river and/or bank overflow downstream, and that an increase in discharge leads to a widening of the river. Of course, the propagation of a flood in reality depends on many other factors,

including antecedent soil moisture and river diversions caused by bank overflow. A model that takes into account additional variables could therefore potentially achieve more accurate forecasts. Recognition of the impact of geomorphic factors on the signal response was also discussed by Brakenridge et al. (2007), who observed that the shape and slope of the river impact whether or not there is a lag between peak discharge records and peak signals.

Furthermore, presence of man-made barriers such as hydroelectric dams has changed the hydrological relationships between upstream and downstream points of interest. For example, Jury (2014) showed that Mangochi (upstream on the Shire, just south of Lake Malawi) and Chikwawa had similar discharge cycles up until the 1990s, after which hydroelectric yield manipulated the observed baseflow. This shows that certain hydrological patterns, cycles and extremes upstream can have a clear relationship with observations downstream, but that these relationships can change over time when new barriers are constructed or changes are made to the water retainment of existing dams. Whereas this factor has not been taken into account in this study, the Kapichira Hydroelectric Power Station is situated just upstream from Chikwawa, which may have impacted the results of this study.

Lastly, the relatively large size of the VGS cells means that one cannot be certain no other rivers or water bodies are present in the VGS. This became apparent when studying K0, where not only Lake Malawi is situated within the cell, but also some other, mainly non-perennial rivers in the delta. This factor may impact the signals in Chikwawa as well: Whereas no medium-to-large rivers other than the Mwamphanzi – a tributary of the Shire – are visible within the bounds of the VGS, smaller streams and ponds can still influence the observed signal. The relatively strong relationships found in a study for Chikwawa on fine-resolution PMRS data (see Section 5.2.2) suggest more research could be done on the impact of the spatial resolution of the PMRS-data on the detection skill. Finer resolution grids (up to 3.125 km x 3.125 km) are indeed available in the MEaSURES dataset, but were not used in this study due to computational limitations and because longer timeseries were available of the coarser resolution grid.

5.5.4 Bias due to Correct Negatives

Whereas it was assumed that the flood database was relatively complete, floods may have taken place and gone unreported, especially if floods took place in a location with little to no settlement (low exposure) or where high protection led to a low impact (low vulnerability) . This affects the relevance of the FAR- and CSI-metrics: They make use of FAs, and their trustworthiness depends on the assumption of a comprehensive database. Based on the conducted quality check of the database, however, it is assumed that floods that are described in the database did indeed take place and that the number of Hits and Misses in the confusion matrix and metrics derived from it (i.e. the POD) can be trusted. The POD should therefore be leading in the interpretation of the success metrics in this study. In the context of this study, this previous point showed to be less relevant, however. Simulations done with m showcased not only the highest POD, but also the lowest FAR and the highest CSI.

6. Conclusion

The aim of this thesis was to investigate whether openly available, coarse-resolution PMRS data could be used for EWS for floods in Malawi, and the Shire River Basin in Chikwawa, the North Rukuru River Basin in Karonga in particular. Two different PMRS-indices were used to achieve this aim: r_{cmc} , which is relatively new and was proposed by Neisingh (2018a), and m , which can be seen as the relative r_{cm} . The r_{cm} has been studied more widely at different locations (e.g. Brakenridge et al. (2007); De Groeve (2010)).

6.1 Detection potential

The first research question pertained to whether PMRS could be used for detecting riverine floods and estimate their magnitude. The aim of this question was to quantify the satellite indices' relationships with discharge and flood events and potentially use these relationships as a basis for a forecasting system. The findings showed that at both locations, r_{cmc} and m showed a similar seasonality to the observed discharge hydrographs, as long as the downstream VGS was located at a sufficient distance from a large water body. However, the regression showed that in Chikwawa, whereas the indices' relationship with observed discharge had a moderately strong, positive correlation, this was not the case for Karonga. We suggest this has to do with the size of the North Rukuru River: The river has a much lower discharge than that of the Shire River (Figures 14,16), and whereas fluctuations are large relative to the baseflow, they might not lead to a river widening that will be distinctly visible in a PMRS grid of 25x25 km: The R^2 in Chikwawa was only 0.361 and 0.233 for r_{cmc} and m , respectively.

Apart from absolute discharge values, flood occurrence detection skill was assessed as well. The trigger thresholds were calculated from the discharge and satellite data as the peak value with a rp of 5 years. However, at this threshold configuration, both r_{cmc} and m did not detect the majority of registered floods. In fact, when only the registered floods with a 5 year rp or more were used, the detection skill went down even more. Judging from the peaks of the satellite data that occur near the registered floods, however, we suggest more research is done in setting a correct trigger threshold, as this could provide a substantial improvement from the success metrics presented in this research.

When comparing the detection potential of m and r_{cmc} , the latter performed better than m for absolute discharge estimation. However, when looking at flood occurrence, the success metrics proved to be better for m . If more research is done into the trigger thresholds, the flood event detection comparison should be done again to assess whether at a different threshold configuration, m still performs better. If this is the case, an EWS could be set up where both m and r_{cmc} are used: m for estimating flood timing, and r_{cmc} for estimating precise discharge if needed.

6.2 Forecasting potential

For the second research question, we investigated the potential of PMRS for forecasting flood events. A TLCC was conducted, but no VGS was identified that had 1) a satellite signal that was strongly correlated to the downstream satellite signal and 2) showed a positive lag time > 0 at the point of maximum correlation. Therefore, the forecasting model was based on the downstream satellite signals only, using r_{cmc} rather than m due to the higher R^2 and lower RMSE of r_{cmc} . Model forecasting skill was tested on the wet season in 2015, and proved to be sufficient up to a lead time of nearly 3 days in Chikwawa if r_{cmc} was used. In Karonga, the model unfortunately failed to forecast the future signal accurately.

When used in an operational early warning setting, using the 2015 flood in Chikwawa as an example, the model unfortunately fails to forecast exceedance of the trigger threshold 3 days in advance, but only

recognizes this 1 day ahead. This experiment could be conducted again once the trigger threshold configuration has been perfected, to see if the proposed EWS would give the trigger in a timely manner.

6.3 Implications for humanitarian work

6.3.1 r_{cmc} vs. m

m is directly related to r_{cm} . r_{cm} has been researched more extensively and is therefore considered more ‘traditional/mainstream’ in the academic literature on flood forecasting using PMRS. However, our research also has shown the potential of r_{cmc} , an index that was introduced by (Neisingh 2018a), but which has not yet been further researched. The results in this study indicated that the r_{cmc} has a stronger positive correlation with discharge than m in our study area. We therefore suggest more research is done in the potential applicability of r_{cmc} as a predictor in EWSs for floods. An interesting follow-up research project could be done on linking r_{cmc} data to high-resolution Digital Elevation Models (DEMs) of the study area. Since r_{cmc} can be interpreted spatially (see Neisingh (2018a)), flood extents could be simulated on maps just after (or even during) flood events, before the sky has cleared up and therefore before optical satellite imagery of the affected region is available. This could facilitate humanitarian missions by being able to identify affected areas early on and reach the people who are in need of humanitarian aid the most.

6.3.2 *Real-time data availability*

This research was done by using historical PMRS-data from the MEaSURES-dataset that spanned a period until 2017. Of course, should PMRS be included in an EWS, it is important to gain access to openly available recent data in real-time. The World Meteorological Organisation (WMO) has a list of satellites and sensors that operate at or near the 36.5-37.0 GHz frequency, of which multiple are currently operational and providing data in (near) real-time. An example is RosHydromet’s Meteor-M N2 satellite series, that have the MTCZA-GY sensor, operating at 36.7 GHz and providing full global coverage twice a day in real-time. The DMSP-F17 satellite, which was used in this study, is also still operational at the 37.0 GHz bandwidth (sensor: SSMIS) at the time of writing. More research could be done in which agencies provide open data access, or which partnerships could be made to achieve this.

6.3.3 *Alternative EWSs*

A comparison with the detection skill of GloFAS in Chikwawa showed that the presented PMRS-method would be preferred over GloFAS when it comes to both flood event detection and discharge estimation. For forecasting, no GloFAS-data was available to make a quantitative comparison with the PMRS-forecasts. No comparison could be made at all with the EWS currently in place in Chikwawa (ODSS) due to the absence of detecting or forecasting data from this system. The lead time of the PMRS-method in Chikwawa presents no significant time win compared to the EWS currently in place (ODSS) or GloFAS. However, it might still be interesting to compare the forecasting skill with that of ODSS and/or GloFAS when data becomes available. The relatively low data demand of the presented PMRS-method means that it has a potential to be used for EWSs in data-scarce, ungauged regions, whereas this is more difficult to achieve for ODSS.

Overall, a coupled EWS solution where a global forecasting model is combined with a PMRS-model, and a national model such as ODSS seems optimal. It would offer the end user forecasts at different lead times with increasing performance for the shorter lead times. The combination will also increase the overall spatial coverage of the EWS as the global models and the PMRS-model can cover the areas where no national model is available and/or where input data is scarce. Further research will be necessary also in terms of how to well communicate the uncertainties associated with each of the systems and how to practically implement such a coupled system.

7. References

- 510 an initiative of the Netherlands Red Cross. 2020a. Community Risk Dashboard. Rode Kruis Dashboards. Retrieved 2-3 2020, from https://profiles.510.global#!/community_risk.
- 510 an initiative of the Netherlands Red Cross. 2020b. Phases of the disaster management cycle. Presentation template: 510, an initiative of The Netherlands Red Cross.
- Alsdorf, D. 2003. Water Storage of the Central Amazon Floodplain Measured with GIS and Remote Sensing Imagery. *Annals of the Association of American Geographers*, 93: 55-66. DOI: 10.1111/1467-8306.93105
- Ammentorp, H. C., and B. Richaud, 2016. Sub-components A3(II): Operational Decision Support System through Enhanced Hydro-Meteorological Services., Ministry of Agriculture, Irrigation and Water Development, Report, Lilongwe.
- Arndt, S., S. Willmes, W. Dierking, and M. Nicolaus. 2016. Timing and regional patterns of snowmelt on Antarctic sea ice from passive microwave satellite observations. *Journal of Geophysical Research: Oceans*, 121: 5916-5930. DOI: 10.1002/2015JC011504
- Birkinshaw, S. J., G. M. O'Donnell, P. Moore, C. G. Kilsby, H. J. Fowler, and P. A. M. Berry. 2010. Using satellite altimetry data to augment flow estimation techniques on the Mekong River. *Hydrological Processes*, 24: 3811-3825. DOI: 10.1002/hyp.7811
- Bjerklie, D. M., D. Moller, L. C. Smith, and S. L. Dingman. 2005. Estimating discharge in rivers using remotely sensed hydraulic information. *Journal of Hydrology*, 309: 191-209. DOI: <https://doi.org/10.1016/j.jhydrol.2004.11.022>
- Brakenridge, R., S. Nghiem, E. Anderson, and R.-P. Mic. 2007. Orbital microwave measurement of river discharge and ice status. *Water Resources Research*, 43. DOI: 10.1029/2006WR005238
- Brodzik, M. J., D. G. Long, M. A. Hardman, A. Paget, and R. Armstrong. 2016. MEASUREs Calibrated Enhanced-Resolution Passive Microwave Daily EASE-Grid 2.0 Brightness Temperature ESDR, Version 1. ed. NASA National Snow and Ice Data Center Distributed Active Archive Center. Boulder, Colorado USA: NSIDC.
- Chidanti-Malunga, J. 2011. Adaptive strategies to climate change in Southern Malawi. *Physics and Chemistry of the Earth, Parts A/B/C*, 36: 1043-1046. DOI: <https://doi.org/10.1016/j.pce.2011.08.012>
- Chuvieco, E. 2016. *Fundamentals of Satellite Remote Sensing: An Environmental Approach, Second Edition*. CRC Press, Inc.
- Climate Centre, ND. Glossary of terms for forecast-based financing. Red Cross / Red Crescent Climate Centre, Report.
- Coughlan de Perez, E., B. van den Hurk, M. K. van Aalst, I. Amuron, D. Bamanya, T. Hauser, B. Jongma, A. Lopez, et al. 2016. Action-based flood forecasting for triggering humanitarian action. *Hydrol. Earth Syst. Sci.*, 20: 3549-3560. DOI: 10.5194/hess-20-3549-2016
- D'Addabbo, A., SpringerLink, D. Capolongo, and A. Refice. 2018. *Flood Monitoring through Remote Sensing*. Springer International Publishing.
- Dartmouth Flood Observatory. 2020. Global Active Archive of Large Flood Events.
- Davis, B., S. Di Giuseppe, and A. Zezza. 2017. Are African households (not) leaving agriculture? Patterns of households' income sources in rural Sub-Saharan Africa. *Food Policy*, 67: 153-174. DOI: 10.1016/j.foodpol.2016.09.018
- De Groeve, T. 2010. Flood monitoring and mapping using passive microwave remote sensing in Namibia. *Geomatics, Natural Hazards and Risk*, 1: 19-35. DOI: 10.1080/19475701003648085
- De Groeve, T., Z. Kugler, and R. Brakenridge. 2007. Near Real Time Flood Alerting for the Global Disaster Alert and Coordination System. *Proceedings ISCRAM2007*.
- Emerton, R., E. Zsoter, L. Arnal, H. Cloke, D. Muraro, C. Prudhomme, E. Stephens, P. Salamon, et al. 2018. Developing a global operational seasonal hydro-meteorological forecasting system: GloFAS-Seasonal v1.0. *Geoscientific Model Development*, 11: 3327-3346. DOI: 10.5194/gmd-11-3327-2018
- ESRI. 1992. Inland Water. ed. Digital Chart of the World. DIVA-GIS.
- Fakhrudin, B. S. H. M., and L. Schick. 2019. Benefits of economic assessment of cyclone early warning systems - A case study on Cyclone Evan in Samoa. *Progress in Disaster Science*, 2: 100034. DOI: <https://doi.org/10.1016/j.pdisas.2019.100034>
- FAO. 2011. Irrigation and drainage development. ed. Food and Agriculture Organization of the United Nations (FAO).
- Geonode. 2013. Main Cities. MASDAP.
- GLIDE. 2020. GLIDE Search. ed. Asian Disaster Reduction Centre.

- Global Runoff Data Centre. 2020. MWAKIMEME (65312401) daily data. ed. GRDC. GRDC Data Download portal.
- Government of Malawi, 2015. Malawi 2015 Floods Post Disaster Needs Assessment Report, Report.
- Gros, C., M. Bailey, S. Schwager, A. Hassan, R. Zingg, M. M. Uddin, M. Shahjahan, H. Islam, et al. 2019. Household-level effects of providing forecast-based cash in anticipation of extreme weather events: Quasi-experimental evidence from humanitarian interventions in the 2017 floods in Bangladesh. *International Journal of Disaster Risk Reduction*, 41: 101275. DOI: <https://doi.org/10.1016/j.ijdr.2019.101275>
- Guha-Sapir, D. 2020. EM-DAT. ed. CRED / UCLouvain. Brussels, Belgium.
- Harrigan, S., E. Zsoter, L. Alfieri, C. Prudhomme, P. Salamon, F. Wetterhall, C. Barnard, H. Cloke, et al. 2020. River discharge and related historical data from the Global Flood Awareness System v2.1. ed. Copernicus Emergency Management Service (CEMS). Copernicus Climate Data Store: Copernicus Climate Change Service by ECMWF.
- Hasan, E., A. Tarhule, Y. Hong, and B. Moore. 2019. Assessment of Physical Water Scarcity in Africa Using GRACE and TRMM Satellite Data. *Remote Sensing*, 11. DOI: 10.3390/rs11080904
- Hendriks, T. D., and F. K. Boersma. 2019. Bringing the state back in to humanitarian crises response: Disaster governance and challenging collaborations in the 2015 Malawi flood response. *International Journal of Disaster Risk Reduction*, 40. DOI: 10.1016/j.ijdr.2019.101262
- Hirpa, F. A., T. M. Hopson, T. De Groeve, G. R. Brakenridge, M. Gebremichael, and P. J. Restrepo. 2013. Upstream satellite remote sensing for river discharge forecasting: Application to major rivers in South Asia. *Remote Sensing of Environment*, 131: 140-151. DOI: <https://doi.org/10.1016/j.rse.2012.11.013>
- Hoang, T., N. Phu, M. Ombadi, H. S. U. Kuolin, S. Sorooshian, and K. Andreadis. 2019. Improving Hydrologic Modeling Using Cloud-Free MODIS Flood Maps. *Journal of Hydrometeorology*, 20: 2203-2214. DOI: 10.1175/JHM-D-19-0021.1
- Holmes, T., R. de Jeu, M. Owe, and H. Dolman. 2009. Land surface temperature from Ka band (37 GHz) passive microwave observations. *Journal of Geophysical Research*, 114. DOI: 10.1029/2008JD010257
- Hussain, M., M. H. Arsalan, K. Siddiqi, B. Naseem, and U. Rabab. 2005. Emerging geo-information technologies (GIT) for natural disaster management in Pakistan: an overview. In *Proceedings of 2nd International Conference on Recent Advances in Space Technologies, 2005. RAST 2005.*, 487-493.
- IFRC. 2019. Monsoon Floods in Bangladesh 2019: FbF activated in the Jamuna/Brahmaputra river basin. Retrieved 28-2 2020, from <https://www.forecast-based-financing.org/2019/07/25/4946/>.
- IFRC. 2020a. Forecast-Based Action (FbA) by the DREF. Retrieved 21-2 2020, from <https://media.ifrc.org/ifrc/fba/>.
- IFRC. 2020b. What is Forecast-based Financing? Retrieved 21-2 2020, from <https://www.forecast-based-financing.org/about/>.
- IPCC, 2014. Climate Change 2014: Synthesis Report. IPCC, Report.
- Jjemba, E. W., B. K. Mwebaze, J. Arrighi, E. C. de Perez, and M. Bailey. 2018. Chapter 19 - Forecast-Based Financing and Climate Change Adaptation: Uganda Makes History Using Science to Prepare for Floods. In *Resilience*, eds. Z. Zommers, and K. Alverson, 237-242. Elsevier.
- Jury, M. R. 2014. Malawi's Shire River Fluctuations and Climate. *Journal of Hydrometeorology*, 15: 2039-2049. DOI: 10.1175/JHM-D-13-0195.1
- Kelly, L., D. Bertram, C. Ngongondo, and R. Kalin. 2019. *Characterization of Groundwater Discharge to Rivers in the Shire River Basin*.
- Keunen, O. 2020. Flood Forecasting Using Passive Microwave Radiometry in the Zambezi River, Zambia. MSc Thesis Delft University of Technology
- Kieper, M. E., and H. Jiang. 2012. Predicting tropical cyclone rapid intensification using the 37 GHz ring pattern identified from passive microwave measurements.
- Kramer, T., 2018. Relating VanderSat microwave brightness temperature to discharge of the Shire River in Southern Malawi. 510, Report.
- Kwak, Y., J. Park, and Y. Iwami. 2017. Large flood mapping using syncro water index coupling with hydro data and time series modis images. 610-611. IEEE.
- Lowe, D., K. L. Ebi, and B. Forsberg. 2011. Heatwave Early Warning Systems and Adaptation Advice to Reduce Human Health Consequences of Heatwaves. *International Journal of Environmental Research and Public Health*, 8. DOI: 10.3390/ijerph8124623
- Lumbroso, D., L. Boelee, E. Zsoter, H. Cloke, P. Samuels, and J. Bazo. 2017. *Analysis of the uncertainty in flood predictions of GLoFAS for Piura in the Pacific Region of Peru*.
- Macharia, D., M. Kinglsey, M. Mwale, D. Ongo, and G. Koech. 2015. *Malawi Hazards and Vulnerability Atlas*. Malawi Disaster Management Portal.
2020. Malawi - Environment - River gauge stations. Malawi Atlas.

- Malawi Meteorological Services. 2020. Climate of Malawi. Retrieved, from <https://www.metmalawi.com/climate/climate.php>.
- Manda, M., and E. Wanda. 2017. Understanding the nature and scale of risks in Karonga, Malawi. *Environment and Urbanization*, 29: 15-32. DOI: 10.1177/0956247817692200
- Mijoni, P. L., and Y. O. Izadkhah. 2009. Management of floods in Malawi; case study of the Lower Shire River Valley. *Disaster Prevention and Management: An International Journal*, 18: 490-503. DOI: 10.1108/09653560911003688
- Mwale, F. D., A. J. Adeloje, and R. Rustum. 2012. Infilling of missing rainfall and streamflow data in the Shire River basin, Malawi - A self organizing map approach. *Physics and Chemistry of the Earth*, 50-52: 34-43. DOI: 10.1016/j.pce.2012.09.006
- NASA. 2008. Atmospheric electromagnetic opacity, Vectorized by User:Mysid. In *Inkscape*, Atmospheric electromagnetic opacity.svg. Wikimedia commons: NASA.
- NASA/METI/AIST/Japan Spacesystems, and U.S./Japan ASTER Science Team. 2019. ASTER Global Digital Elevation Model V003. NASA EOSDIS Land Processes DAAC,.
- National Hydrological Services of Malawi. 2019. Shire at Chikwawa (11212) discharge. ed. National Hydrological Services of Malawi. Lilongwe, Malawi.
- National Statistics Office of Malawi. 2018. Malawi administrative level 0-3 boundaries. Humanitarian Data Exchange.
- Neisingh, W. 2018a. Inundation mapping in a dynamic, data-scarce environment using Ka-band passive microwave radiometry. Ouémé Delta, Benin. MSc Thesis. Delft: TU Delft
- Neisingh, W. 2018b. JAXA_AMSR2. Satellite Cookbook. Retrieved, from https://github.com/hcwinsemius/satellite-cookbook/blob/master/NSIDC-AMSRE/JAXA_AMSR2.ipynb.
- Ngongondo, C., C.-Y. Xu, L. Gottschalk, and B. Alemaw. 2011. Evaluation of spatial and temporal characteristics of rainfall in Malawi: a case of data scarce region. *Theoretical and Applied Climatology*, 106: 79-93. DOI: 10.1007/s00704-011-0413-0
- Nielsen, S. A., and E. Hansen. 1973. Numerical Simulation of the rainfall-runoff process on a daily basis. *Hydrology Research*, 4: 171-190. DOI: 10.2166/nh.1973.0013
- Njaya, F., K. Snyder, D. Jamu, J. Wilson, C. Howard-Williams, E. Allison, and N. Andrew. 2011. The natural history and fisheries ecology of Lake Chilwa, southern Malawi. *Lancet*, 37: 15-25. DOI: 10.1016/j.jglr.2010.09.008
- Njoku, E. G., T. J. Jackson, V. Lakshmi, T. K. Chan, and S. V. Nghiem. 2003. Soil moisture retrieval from AMSR-E. 215-229. USA: IEEE.
- OpenStreetMap contributors. 2020. HOTOSM Malawi Waterways (OpenStreetMap Export). ed. Humanitarian OpenStreetMap Team (HOT). Humanitarian Data Exchange.
- Palmer, S. C. J., T. Kutser, and P. D. Hunter. 2015. Remote sensing of inland waters: Challenges, progress and future directions. 1-8.
- Parker, D. J. 2000. Floods in Africa. In *Floods*, 188-202. London: Routledge.
- Pauw, K., J. Thurlow, M. Bachu, and D. E. Van Seventer. 2011. The economic costs of extreme weather events: a hydrometeorological CGE analysis for Malawi. *Environment and Development Economics*, 16: 177-198. DOI: 10.1017/S1355770X10000471
- Petropoulos, G. P., and T. Islam. 2017. *Remote Sensing of Hydrometeorological Hazards*. CRC Press.
- Pilant, D. 1994. Close range 37 GHz passive microwave observations of snow on north and south facing slopes. 2071.
- Prigent, C., F. Aires, and B. Rossow William. 2006. Land surface microwave emissivities over the globe for a decade. *Bulletin of the American Meteorological Society*, 87: 1573.
- Roy, C., and R. Kovordanyi. 2015. The current cyclone early warning system in Bangladesh : Providers' and receivers' views. *International Journal of Disaster Risk Reduction*, 12: 285-299. DOI: 10.1016/j.ijdrr.2015.02.004
- Šakić Trogrlić, R., G. B. Wright, A. J. Adeloje, M. J. Duncan, and F. Mwale. 2018. Taking stock of community-based flood risk management in Malawi: different stakeholders, different perspectives. *Environmental Hazards*, 17: 107-127. DOI: 10.1080/17477891.2017.1381582
- Savitzky, A., and M. J. E. Golay. 1964. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, 36: 1627-1639. DOI: 10.1021/ac60214a047
- Schaefer, J. T. 1990. The Critical Success Index as an Indicator of Warning Skill. *Weather and Forecasting*, 5: 570-575. DOI: 10.1175/1520-0434(1990)005<0570:TCSIAA>2.0.CO;2

- Sheel, M., J. Collins, M. Kama, D. Nand, D. Faktaufon, J. Samuela, V. Biaukula, C. Haskew, et al. 2019. Evaluation of the early warning, alert and response system after Cyclone Winston, Fiji, 2016. *Bulletin of the World Health Organization*, 97: 178-189C. DOI: 10.2471/BLT.18.211409
- Smith, L. C. 1997. Satellite remote sensing of river inundation area, stage, and discharge: a review. *Hydrological Processes*, 11: 1427-1439. DOI: 10.1002/(SICI)1099-1085(199708)11:10<1427::AID-HYP473>3.0.CO;2-S
- Svendsen, M., M. Ewing, and S. Msangi. 2009. Measuring Irrigation Performance in Africa. *International Food Policy Research Institute (IFPRI), IFPRI discussion papers*.
- Teule, T. 2019. Assessing Two Methods to Potentially Improve the Flood Early Warning System in the Lower Shire Valley in Malawi. MSc Thesis. Amsterdam: VU
- Trogrlić, R. Š., G. B. Wright, A. J. Adeloye, M. J. Duncan, M. J. C. van den Homberg, F. D. Mwale, and J. Mwafalirwa. 2019. Characterising local knowledge across the flood risk management cycle: A case study of Southern Malawi. *Sustainability (Switzerland)*, 11. DOI: 10.3390/su11061681
- UNEP, and GEAS. 2014. Cyclone Phailin in India: Early warning and timely actions saved lives. *Environmental Development*, 9: 93-100. DOI: <https://doi.org/10.1016/j.envdev.2013.12.001>
- UNISDR, 2009. UNISDR Terminology on Disaster Risk Reduction. United Nations International Strategy for Disaster Risk Reduction, Report, Geneva, Switzerland.
- Van Dijk, A. I. J. M., G. R. Brakenridge, A. J. Kettner, H. E. Beck, T. De Groeve, and J. Schellekens. 2016. River gauging at global scale using optical and passive microwave remote sensing. *Water Resources Research*, 52: 6404-6418. DOI: 10.1002/2015WR018545
- Vinet, F. 2018. *Floods*. Elsevier.
- Wallemacq, P., D. Guha-Sapir, D. McClean, Cred, and Unisdr. 2015. *The Human Cost of Weather Related Disasters - 1995 - 2015*.
- Weintrit, B., K. Osińska-Skotak, and M. Pilarska. 2018. Feasibility study of flood risk monitoring based on optical satellite data. *Miscellanea Geographica: Regional Studies on Development*: 172. DOI: 10.2478/mgrsd-2018-0011
- Wilkinson, E., L. Weingärtner, R. Choularton, M. Bailey, M. Todd, D. Kniveton, and C. C. Venton, 2018. Forecasting hazards, averting disasters. Implementing forecast-based early action at scale. Overseas Development Institute, Report, London.
- World Bank. 2016. Malawi - Poverty headcount ratio at national poverty lines. The World Bank Group.
- World Bank. 2018a. Malawi - Population, total. ed. World Development Indicators. The World Bank Group.
- World Bank. 2018b. Malawi - Rural Population. The World Bank Group.
- World Meteorological Organization. 2020. List of all Satellites. Retrieved 22-3 2020, from <https://www.wmo-sat.info/oscar/satellites>.

Appendices

A: Data treatment and downloads

A1: Overview of satellites carrying passive sensors in the Ka band frequencies of interest (36.5 to 37 GHz) in chronological order of launch. Data: World Meteorological Organization (2020)

Satellite	Agency	Launch	Sensor	Frequency (GHz)	IFOV (km)	Temporal coverage of sensor	Present in MEaSUREs?
SeaSat	NASA	1978	SMMR	37.0	17 x 29	1978 – 1978	No
Nimbus 7	NASA	1978	SMMR	37.0	17 x 29	1978 – 1987	Yes
DMSP – F08	US DoD, NOAA	1987	SSM/I	37.0	24 x 36	1987 – 2006	Yes
DMSP – F10	US DoD, NOAA	1990	SSM/I	37.0	24 x 36	1990 – 1997	Yes
DMSP – F11	US DoD, NOAA	1991	SSM/I	37.0	24 x 36	1991 – 1995	Yes
ERS-1	ESA	1991	ATSR	36.5		1991– 2000	No
DMSP – F12	US DoD, NOAA	1994	SSM/I	37.0	24 x 36	1994 – 2008	No
DMSP – F13	US DoD, NOAA	1995	SSM/I	37.0	24 x 36	1995 – 2008	Yes
ERS-2	ESA	1995	MWR	36.5	20	1995 – 2011	No
DMSP – F14	US DoD, NOAA	1997	SSM/I	37.0	24 x 36	1997 – 2007	Yes
TRIMM	NASA, JAXA	1997	TMI	37.0	9 x 16	1998 – 2015	No
DMSP – F15	US DoD, NOAA	1999	SSM/I	37.0	24 x 36	1999 – 2018	Yes
Aqua	NASA	2002	AMS-R-E	36.5	9 x 14	2002 – 2011	Yes
Envisat	ESA	2002	MWR	36.5	20	2002 – 2012	No
ADEOS-II	JAXA	2002	AMS-R	36.5	7.6 x 13	2003 – 2003	No
DMSP – F16	US DoD, NOAA	2003	SSMIS	37.0	24 x 36	2003 – 2019	Yes
FY-3A	CMA, NRSCC	2008	MWRI	36.5	18 x 30	2008 – 2015	No
Meteor-M N1	RosHydroMet, Roscosmos	2009	MTVZA-GY	36.7	30 x 67	2010 – 2014	No
HY-2A	NSOAS, CAST	2011	MWRI	37.0	15 x 22	2012 – 2016	No
Megha-Tropiques	ISRO, SNES	2011	MADRAS	36.5	20 x 31	2013 – 2013	No
DMSP – F19	US DoD, NOAA	2014	SSMIS	37.0	24 x 36	2014 – 2016	Yes
Meteor-M N2	RosHydroMet, Roscosmos	2015	MTVZA-GY	36.7	30 x 67	2015 – 2017	No
Coriolis	US DoD, NASA	2003	WindSat	37.0	8 x 13	2003 – present	No
DMSP – F17	US DoD, NOAA	2006	SSMIS	37.0	24 x 36	2006 – present	Yes
DMSP – F18	US DoD, NOAA	2009	SSMIS	37.0	24 x 36	2009 – present	Yes
FY-3B	CMA, NRSCC	2010	MWRI	36.5	18 x 30	2010 – present	No
GCOM-W	JAXA	2012	AMS-R-2	36.5	7 x 12	2012 – present	No
SARAL	CNES, ISRO	2013	AltiKa	36.5		2013 – present	No
FY-3C	CMA, NRSCC	2013	MWRI	36.5	18 x 30	2013 – present	No
GPM Core Observatory	NASA, JAXA	2014	GMI	36.5	8.6 x 14	2014 – present	No
Sentinel-3A	ESA	2016	MWR	36.5	20	2016 – present	No
FY-3D	CMA, NRSCC	2017	MWRI	36.5	18 x 30	2017 – present	No
Sentinel-3B	ESA	2018	MWR	36.5	20	2018 – present	No
HY-2B	NSOAS, CAST	2018	MWRI	37	15 x 22	2019 – present	No
Meteor-M N2-2	RosHydroMet, Roscosmos	2019	MTVZA-GY	36.7	18 x 30	2019 – present	No

A2: Subsets downloaded from MEaSURES

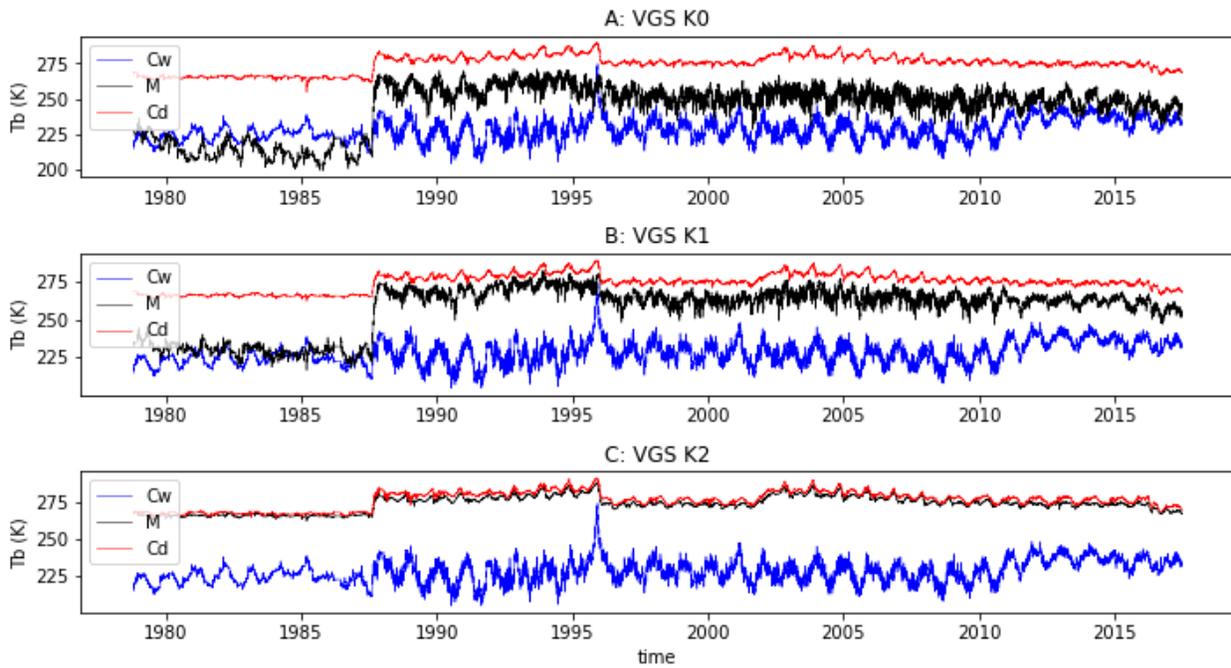
Satellite	Sensor	Subset downloaded (dd-mm-yyyy)
Nimbus	SMMR	01-01-1979 – 31-12-1987
DMSP – F08	SSM/I	01-01-1988 – 31-12-1991
DMSP – F10	SSM/I	01-01-1992 – 31-12-1995
DMSP – F13	SSM/I	01-01-1996 – 31-12-2001
DMSP – F15	SSM/I	01-01-2002 – 31-12-2010
DMSP – F17	SMMIS	01-01-2011 – 01-07-2017

A3: Scripts for downloading and analyzing data

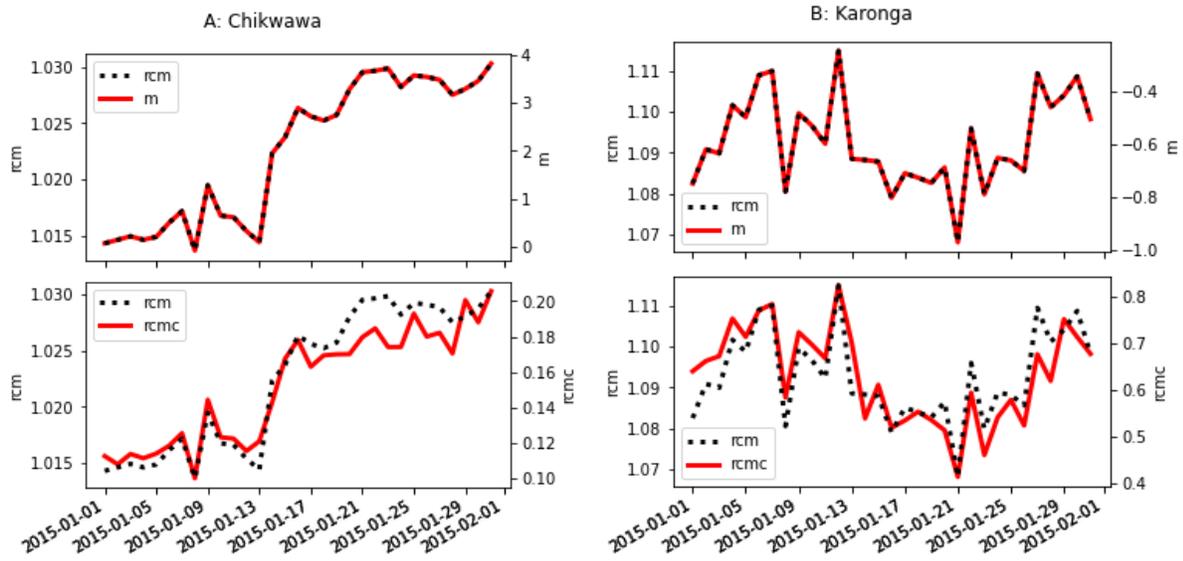
All scripts used in this thesis can be found at https://github.com/lcmok/LU_THESIS

B: Detection skill

B1: Visualization of raw Tb signals in K0, K1 and K2

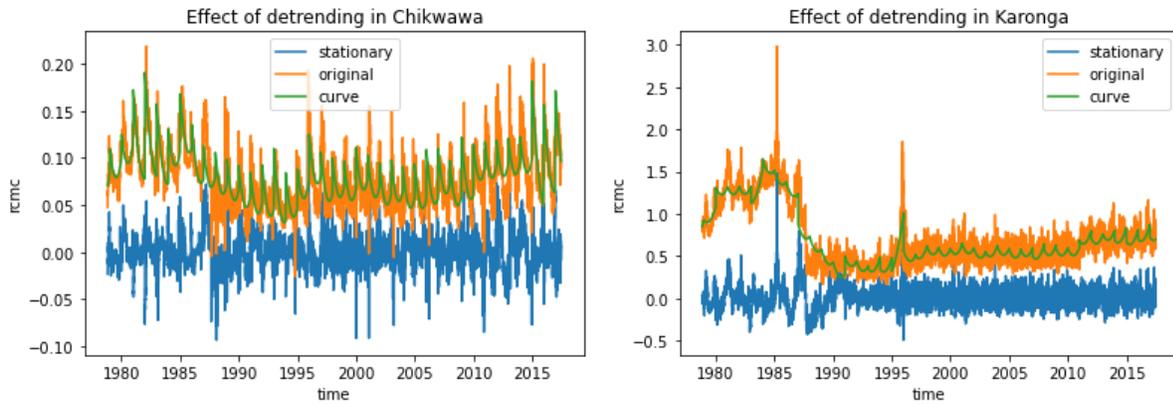


B2: Illustration of the difference between m , rcm and $rcmc$ in Chikwawa (C0) and Karonga (K0) in January 2015

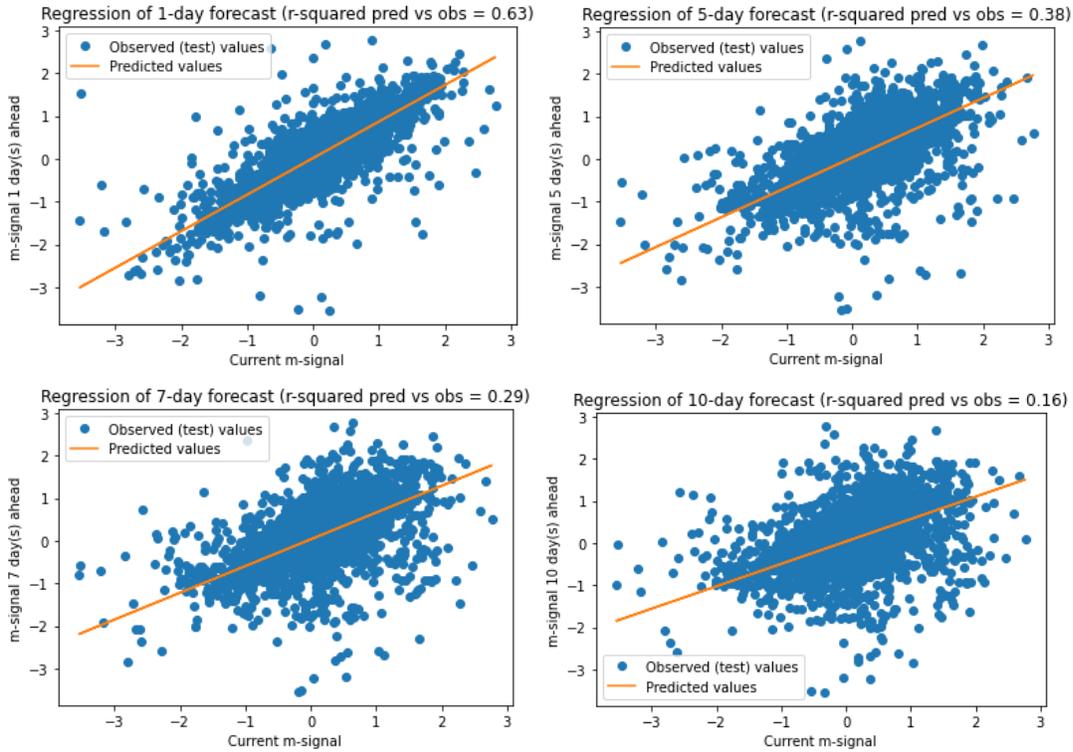


C: Model set-up and prediction

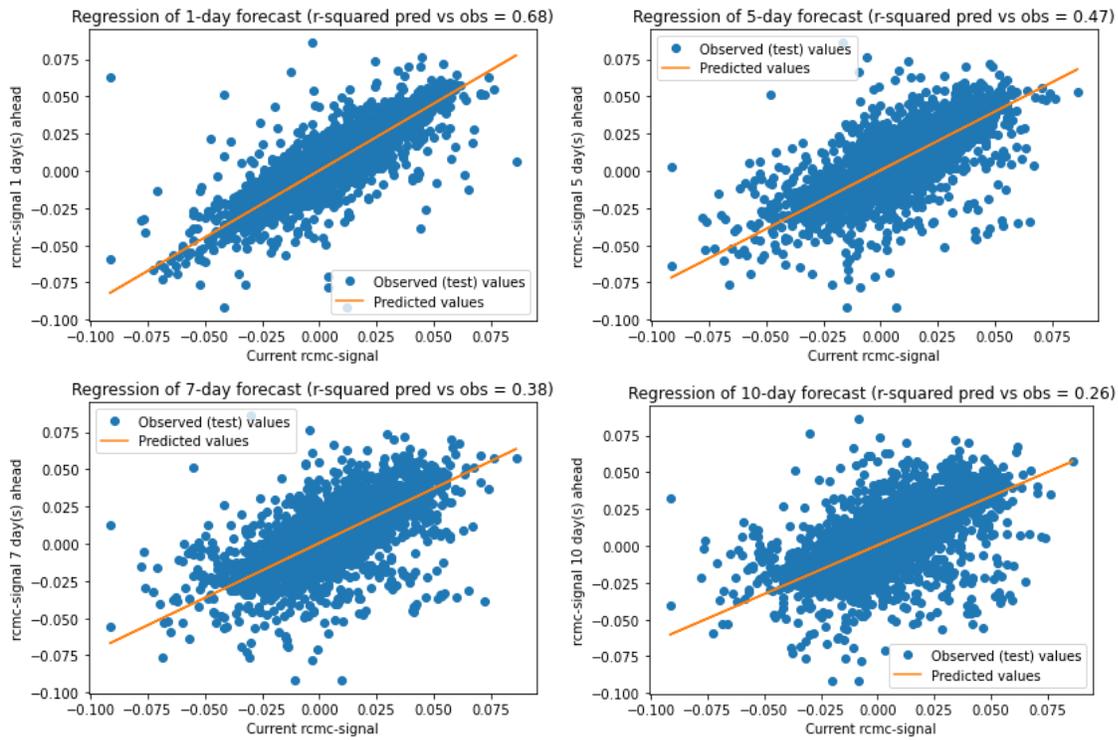
C1: Trend estimation at Chikwawa (C0) and Karonga (K2)



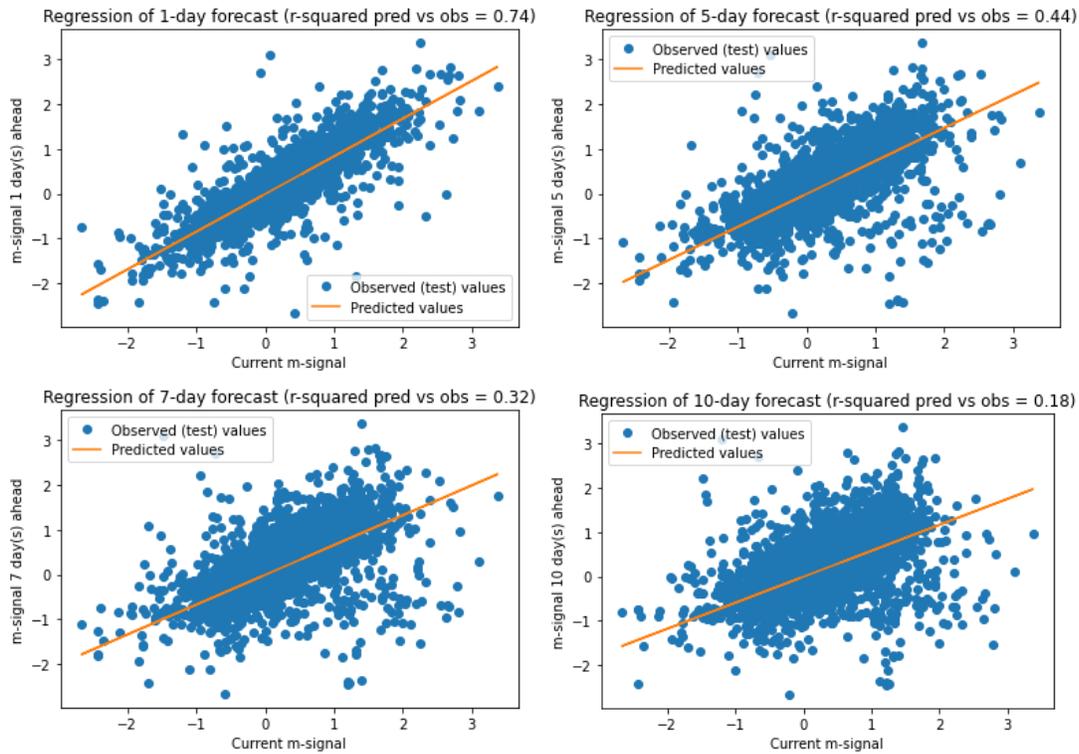
C2: Regression plots of submodels for m in Chikwawa



C3: Regression plots of submodels for r_{cmc} in Chikwawa



C4: Regression plots of submodels for m in Karonga



C5: Regression plots of training data used for forecasting r_{cmc} in Karonga at different lead times

