Delft University of Technology

Rising lakes in the Kenyan Rift Valley

Lake Nakuru flooding: An analysis of the impact of spatial-temporal variability of precipitation and changes in moisture source and climate drivers

by

R.J.A. AMPTING

to obtain the degree of Master of Science at the Delft University of Technology,
Faculty of Civil Engineering and Geosciences Department of Water Management,
to be defended publicly on Tuesday June 20, 2023 at 3 PM.

Student number: 5431468

Project duration: September 1, 2022 – June 20, 2023

Thesis committee:

Prof. dr. ir. N.C. van de Giesen	TU Delft
Dr. ir. R. van der Ent	TU Delft
Dr. ir. M. Rutten	TU Delft
Dr. V. M. Kiluva	MMUST

An electronic version of this thesis is available at `http://repository.tudelft.nl/`.

TU Delft Delft University of Technology

Masinde Muliro University of Science and Technology — Technology for Development

Acknowledgements

My time as a student has come to an end. A time during which I have developed myself to the person I am today. I started with my bachelor at the University of Twente in Civil engineering and a minor in Melbourne, Australia. In three years I completed my bachelors and developed a strong passion for water, whereafter I moved to Delft for my Master in Water Management. My time in Delft started during the COVID-19 pandemic. I am grateful for the time I was able to spent at the study association, Dispuut Water Environment, which gave me a lot of purpose and interaction during the lock down and after. Further, the TU Delft Global Initiative is an inspiring organisation for which I had the privilege of joining the Delft Global Student Board. This community brings together students, researchers, and companies, to make an impact with research in low-and middle-income countries. Conducting my Master Thesis at Delft University of Technology has been a truly joyful and interesting experience that has shaped my academic journey, and I am immensely grateful for the unwavering support provided by the department, my friends and fellow students, and my family, without whom this accomplishment would not have been possible. Therefore I would like to specifically thank the following people.

First, I would like to thank my parents, Patrick and Rosemarie, and my sister Laura, for their unconditional support throughout this journey. I want to thank Ella for keeping me company through difficult times, her support in everything I do, and organising fun activities throughout my thesis. You kept me on my feet and I am truly grateful for that.
I had the privilege of living in the city of Kakamaga in Kenya for about three months were I made long lasting friendships and became familiar with another culture. I would like to mention the invaluable collaboration with Masinde Muliro University of Science and Technology (MMUST). I am deeply grateful for the support I received when I was going there on my own. My deepest gratitude is with my Kenyan friends, Nelly and Felix, for taking good care of my during my time in Kakamega. They are truly inspiring people and I wish them all the best in their future endeavours. I would also like to thank Dr. Kiluva, member of my assessment committee, for introducing me to MMUST and for our insightful discussions and exchange of experiences.
At our department Water Management Environmental Engineering at TU Delft I had the privilege of working at the 'Afstudeerhok'. This is an environment where Master graduates from our study come together to work on their thesis. In particular, a big thanks to; Helena, Guus, Valeria, Jan, David, Nienke, Annemieke, Josien, Charlie, Ron, Rogier, Diana, Sjoerd, Anja, Jorijn, and Mike. Most of the time I spent there writing my thesis and having valuable discussions with my friends and fellow students. It played a pivotal role in enabling me to openly exchange ideas, share my findings, and solidify my conclusions.
I extend my sincere appreciation to the members of my assessment committee. I am grateful for their time, commitment, and willingness to share their knowledge throughout the evaluation process. Nick supported me in my work throughout the thesis as chair of my committee. We have had critical discussions and he was always there to answer questions or to share ideas. He has been very open and approachable throughout my thesis which I really appreciate. Ruud introduced me to the moisture tracking model WAM2layers as part of one of my first courses during the Master's programme. The model intrigued me and for this reason I am grateful he was able to help me throughout the process of applying this model for moisture tracking purposes. Dr. Kiluva received me in Kenya with open arms and I am grateful for the guidance she gave me as I had thrown myself into this new culture and environment. Finally, Martine came with very insightful points regarding the content of my thesis and how to move things forward. It helped me to stay focused and make sure the content of this research was accurate and to the point. Additionally, I would like to thank Vincent, a fellow student from Wageningen University, for our discussions and his help with the moisture tracking model, as this was an important part in my research.
I would also like to express my gratitude for the financial support received from the Lamminga fund, TU Delft Global Initiative, and the Holland Scholarship. Their contributions have not only eased the financial burden but also enabled me to focus on my studies and research without distractions.

The time I have spent on my thesis has been challenging, but enriching. I am proud of this accomplishment and I am happy with the outcome. Enjoy reading!

Summary

Hydrological and meteorological research are of great value to comprehend the complex Eastern Africa system, and more specifically the Kenyan Rift Valley, as impacts of climate change and indisputable changes in the hydrology of the lakes have become more apparent in recent decades. Lakes in the Kenyan Rift Valley region, spanning from the south to the north of Kenya, have been experiencing an increase in their water levels resulting in severe flooding in some cases. The flooding has been a relatively gradual but steady process throughout the 2010s. However, since 2020 the lakes have been rising more rapidly. This research aims to understand the main drivers behind the overflowing of Lake Nakuru as one of the lakes in the Kenyan Rift Valley and specifically examines the hypothesis that increased precipitation in the catchment area is the primary cause of this phenomenon. The study utilises various methodologies to analyse spatio-temporal variability in precipitation, investigate the dynamics of Lake Nakuru, and identify the origin and driving forces behind the changing precipitation patterns.

A change point analysis is performed for weather station data throughout the catchment including lake precipitation and catchment precipitation. A water balance model is built to analyse the impact of precipitation on the lake through precipitation bin analysis and different catchment interactions such as catchment runoff, lake evaporation, and sub-surface catchment storage. Multiple linear regression has been applied to identify key catchment characteristics for the variability of the lake volume. Through moisture tracking as part of this research with WAM2layers it was possible to identify changes in moisture sources and climate drivers of precipitation in the catchment.

The findings of the research reveal significant changes in precipitation patterns, with a notable increase in precipitation since 2010. Change point analysis indicates that this increase coincides with the rise in lake volume, suggesting a strong correlation between precipitation and the overflowing of the lake. The water balance model employed in the study further emphasises the impact of precipitation on the lake, highlighting the complex interactions between precipitation and the catchment. Moisture tracking and climate driver analysis provide insights into the origins of the precipitation variability. The study identifies the western Indian Ocean region, particularly east of Madagascar, as an increasingly important moisture source for the catchment area. Moreover, the study highlights a negative correlation between El Niño and September precipitation, indicating the influence of climate oscillations on the changing precipitation patterns. These findings suggest that future climate changes, including increasing sea surface temperatures and intensifying La Niña and El Niño events, may further contribute to the likelihood of flooding in the area.

Overall, the research strongly supports the hypothesis that the overflowing of Lake Nakuru is primarily caused by changes in precipitation. The study emphasises the need for further research and monitoring to better understand the complex interactions between climate, precipitation, and the lake system.

Contents

1

Introduction

1.1. Problem Description

Kenya is no stranger to drought, as several parts in the north of the country have not experienced rain for a over two years. On the contrary, the southwestern region of Kenya experiences more severe rains and a phenomenon has been unfolding since 2010. Lakes in the Kenyan Rift Valley region, spanning from the south to the north of Kenya, have been experiencing an increase in their water levels resulting in severe flooding in some cases. The flooding has been a relatively gradual but steady process throughout the 2010s. However, since 2020 the lakes have been rising more rapidly (Baraka, 2022). Schools, farmland, wildlife reserves, and communities have flooded. Approximately 75,000 households were forced to leave their home, and as a results a population of nearly 400,000 people required ongoing humanitarian aid to this day (Kenya Ministry of Environment and UNDP, 2021). Inundated riparian land creates room for lake species to expand their territory, causing more human-wildlife interference, as for example schools are threatened by advancing crocodiles and hippos (Baraka, 2022). The photographs in Figure 1.1 demonstrate the significant impact of the rising lakes. The lakes are are not only vital to people and economy, but also wildlife. For example, three saltwater lakes located in the Kenyan Rift Valley, Lake Bogoria, Lake Nakuru, and Lake Elmentaita, sustain a great number of abundant marine and bird life. These ecosystems are threatened by extinction as the lakes' chemical and ecological compositions change.

Research performed by the Ministry of Environment and UDNP identified multiple causes for the rising lakes, such as increased precipitation, anthropogenic land degradation, and geological and tectonic activity. Furthermore, indications of cyclical events are proposed by the elders living near the lakes (Kenya Ministry of Environment and UNDP, 2021). Journalist Baraka (2022), who wrote an article in The Guardian called "A drowning world: Kenya's quiet slide underwater", has spoken to local people whom also suggest that the rise of the lakes is cyclical. A clear trend is observed looking at the Kenyan Rift Valley precipitation. As of 2010, yearly precipitation has increased by approximately 20% compared to the long term mean, increasing from 753 mm to 905 mm averaged over the entire Rift Valley (The World Bank Group, 2021). Therefore, for this study specific interest in changing precipitation patterns have been developed with its relation to the rising lakes. With understanding the changing patterns of precipitation and its impact on Lake Nakuru lies the focus of this research.

The overflowing of the lakes in the complete Rift Valley raises the question what the underlying cause is. Several researchers found a relation between precipitation and the rising lakes. However, their research does not eliminate other possibilities like tectonic plate movement and land use change, as they are not confident enough to be sure that precipitation is the only cause. Their findings showing a relation with overflowing lakes and increased precipitation, but they suggest this should be substantiated with more observational data and at a higher time resolution to establish a more sound theory about the rising lakes. A water balance model is suggested with bathymetry data, observational precipitation data, and more accurate evaporation data. Even though the increased precipitation is said to be evident from literature, the root causes for this variably have not been scrutinised. This research therefore implements moisture tracking and climate driver analysis to investigate the underlying causes and provide insights into possible future developments. The lake Nakuru catchment is used as a case study.

(a) (b) (c)

Figure 1.1: Pictures made from the lake shore and water representing; The impact of rising water levels at the shore of Lake Nakuru with dead trees and intruding terrestrial wildlife environment (a); Impact of rising water levels at the shore of Lake Naivasha with submerged hotel property (b); Impact of rising water levels on land separation on Lake Naivasha where an island has been created around 2019 to 2020 and properties became separated (c).

1.2. Research Objective

The objective of this research is to gain insights into the main drivers responsible for the overflowing of Lake Nakuru. Specifically, the aim is to test the premise that increased precipitation in the catchment and changing underlying precipitation drivers are the primary cause of the phenomenon. The research questions to answer in order to meet this objective are listed below:

1. What spatio-temporal variability in precipitation can be observed on different timescales in the Lake Nakuru catchment?
2. What are the dynamics of lake Nakuru and how do they relate to the precipitation variability?
3. What is the origin and the driving forces for the temporal variability in precipitation and lake Nakuru?

1.3. Report structure

This thesis report is structured as follows: Chapter 2 provides a literature study to shed light on the background of the topics relevant to this study. In Chapter 3, the study area is described in detail. Chapter 4 focuses on the data used in the research. Chapter 5 elaborates on the methodology used to answer the research questions. Chapter 6 presents an overview of the research findings, followed by a discussion of the results and related literature in Chapter 7. Chapter 8 concludes the research and presents the main findings. Finally, Chapter 9 offers recommendations for future research.

2

Literature study

The literature study provides and explains the main components of this research. This chapter elaborates on earlier findings on precipitation variability in Eastern Africa and Kenya combined with fluctuations of the Rift Valley lakes. The underlying climatic drivers and moisture source dynamics for precipitation are explained afterwards and is subdivided into; methods for moisture tracking; moisture sources and sea surface temperature (SST) dynamics; climatic drivers and large-scale circulations; and future changes in climate drivers. Finally, key findings from literature are extracted and listed.

2.1. Precipitation and Lake characteristics

Precipitation in Eastern Africa and Kenya is highly variable and seasonally dependent. Four seasons can be distinguished namely the dry season (JF), long rains season (MAM), continental rains season (JJAS), and the short rains season (OND) (Dyer and Washington, 2021; Yang et al., 2015). Inter-annual variations are most common with on average increasing precipitation since 2010 observed in Kenya (Dyer and Washington, 2021; Gichuru and Waithaka, 2016; Herrnegger et al., 2021; Kimaru et al., 2019). On a year to year basis 2009 to 2010 for example saw an increase in precipitation of 36.3% according to analysed TRMM data, whilst 2010-2011 precipitation reduced by 29.8% (Gichuru and Waithaka, 2016). Kimaru et al. (2019) investigated the temporal variability of precipitation and river flows at Lake Nakuru with Climate Hazards Group Infrared Precipitation with Station (CHIRPS) data on a yearly timescale. CHIRPS slightly overestimates the observational data from one single station that was included in this research. It is found with SPI indices that from 1981 to 1996 there was a high frequency of prolonged dry periods. Prolonged wet periods were experienced between 2010 and 2018 and is given as a possible explanation for the rising of the lake. The almost complete dry up of the lake in 1995 and 1996 due to dry periods whilst in 1997 an El Nino driven precipitation caused flooding. Recommendations are made suggesting the development of a water balance with the latest bathymetry survey to investigate whether the hydro-meteorological variability alone can account for the observed flooding.

Another case study by Langat et al. (2017) in the Tana river basin in southeastern Kenya, found that an increased streamflow is largely due to the increased precipitation in the highlands. They address the availability of water to the intra-seasonal and inter-annual variability in climate including El Nino. A positive annual monotonic trend was found for the higher elevated precipitation observation stations whilst a negative annual monotonic trend was found for lower elevated observations.

Herrnegger et al. (2021) explores the relationship between precipitation and lake characteristics on a yearly basis for all lakes in the Rift Valley and found a relationship between rainfall surplus and deficits with surface area fluctuations. This study suggested a lag of one year for lake Nakuru till precipitation is noticed in the surface area of the lake with a correlation of 0.515. Higher resolution investigations, like monthly and seasonal, might reveal more detailed relationships between precipitation and lake characteristics, where catchment characteristics cannot be omitted. A change point is found in precipitation at 2009 where the long term mean substantially differs from 1981 to 2009 and 2010 to 2020. The lake characteristics are defined by the Database for Hydrological Time Series of Inland Waters (DAHITI) (Schwatke et al., 2019). Furthermore, the actual evapotranspiration is calculated as the actual evaporation from ERA5 times the precipitation ratio between CHIRPS and ERA5, as ERA5 generally overestimates precipitation in Kenya. The article suggested that

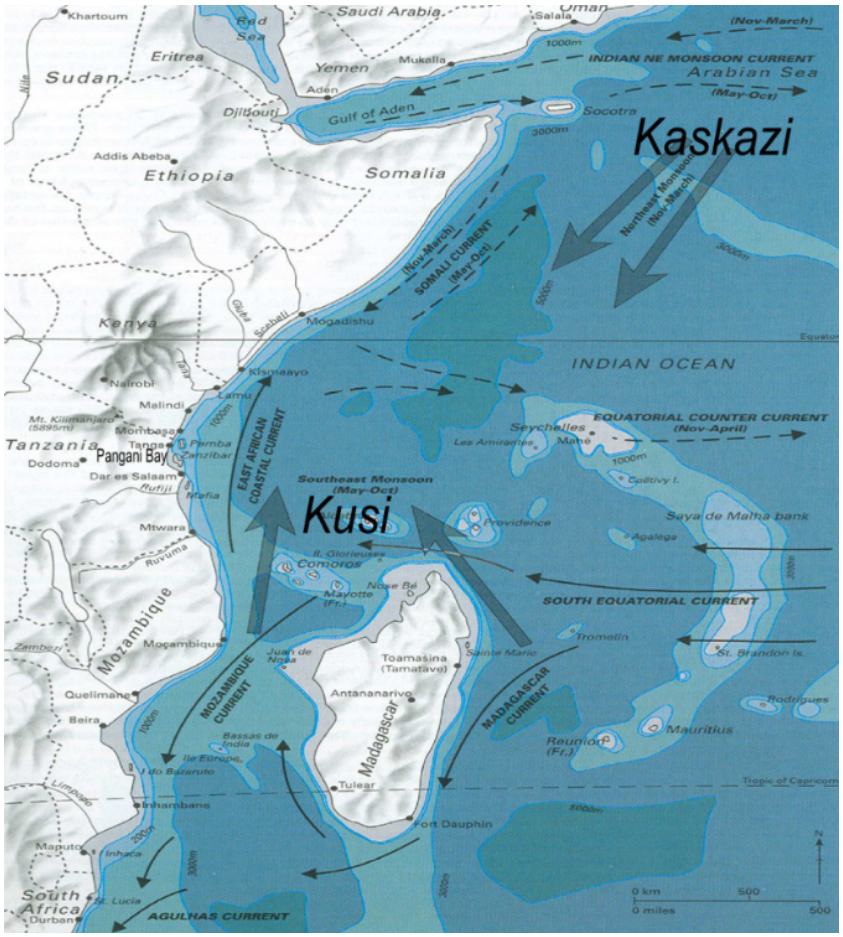


Figure 2.1: East African currents and monsoon winds in the west Indian Ocean (Pollard, 2007). Attention should be paid to the Monsoon systems from the south and the north over different periods. The Southeastern Monsoon manifests itself from May to October whereafter the ITCZ moves over the equator from North to South initiating the short rains season from October to December in Kenya. The Northeast Monsoon manifests from November to March followed by the ITCZ moving from south to north forcing the long rains season from March to May.

using a hydro-meteorological observation network on the ground would improve the accuracy of the analysis, but also more accurate evapotranspiration data is useful for the understanding of the lake and catchment dynamics (Herrnegger et al., 2021; Kimaru et al., 2019).

2.2. Moisture sources and climate drivers

The origin of East African precipitation is highly dynamic and dictated by numerous climatic influences. It receives moisture from three major air systems, the northeastern monsoon system and the southeastern monsoon system from the Indian Ocean, and South Westerly humid air from the Congo basin (Balagizi et al., 2018). The seasonality of precipitation is governed by the interchange between the two monsoon systems, controlled by the Indian Ocean and its temperatures visualised in Figure 2.1. The southeastern monsoon winds are present from May to October. The rain seasons in Eastern Africa are starting when these winds change with the ITCZ moving over Eastern Africa from North to South from October to December. The northeastern monsoon winds are occurring from around November to March which onsets the main rain season from March to May. Tierney et al. (2013) suggested that rainfall in East Africa is mainly dominated by the Indian Ocean by altering the local walker circulation, and the influence of the Pacific Ocean is minimal. The monsoon systems are mainly limited to East Africa because of the high plateaus and and Congo air boundary. To understand these dynamics it is important to understand where and how continental precipitation in Eastern Africa, and more specifically Kenya, is generated and how climate drivers affects this complex system.

Methods for moisture tracking

Several methods are used throughout previous studies for moisture tracking. In recent studies WAM2layers is used to model the moisture source and with that climatic drivers such as SST dynamics (Findell et al., 2019; Keys et al., 2022; Van der Ent et al., 2013; Van der Ent and Savenije, 2013). Van der Ent et al. (2013) compared different moisture tracking models. The tracking can be done online parallel to a climate model run, or posteriori i.e. offline with reanalysis data or output from climate models. It has been concluded that a one layer model would not be sufficient as it could not consider shear factors of wind. An effective way would be, if a moisture tracking model would be used, an offline 2D representation using a Eulerian grid as it reduces computation time, i.e. WAM2layers. The model first calculated the fluxes and states of the water balance within the domain of each grid cell. Second, the fluxes and states information is used to either forward or backtrack from or towards a specific location (Keys et al., 2022).

Moisture recycling in East Africa has also been researched using isotopic analysis (Balagizi et al., 2018; Levin et al., 2009; Otte et al., 2017; Soderberg et al., 2013). The isotope analysis found that for East African countries the main source is the Indian Ocean, but in specific cases of lacking data very difficult to assess for smaller areas. For a case study in Congo, Balagizi et al. (2018) found that East African Lakes are characterised by a certain d-excess indicator that is found in the Congo region. During dry season, evaporation from the lakes is more dominant, which signature is found back in the Congo region precipitation. This argument is supported by Van der Ent et al. (2010), whom studied moisture recycling ratios over the world and found that the main source of rainfall in the Congo basin is evaporated moisture over East Africa.

Moisture sources and SST dynamics

Yang et al. (2015) observed that during dry season (JF) the southern part of the ocean is warmer compared to the northern ocean region and Arabian sea and displays a north-south SST gradient. During the long rains season (MAM) the Indian Ocean gradually heats up, where the warmer temperatures are observed in front of the Kenyan Coast and more eastern in the Indian Ocean. During the continental rains season (JJAS) the ocean cools off with lower temperatures near the coast and higher lower temperatures deeper into the Indian Ocean. In the short rains season (OND) the same distribution of SST is seen as during the long rains season although lower. Comparing precipitation and SST in East Africa, both have their peak in April. Lowest values in SST are observed between July and September.

Van der Ent and Savenije (2013) identified oceanic sources for continental precipitation around the world from 1982 to 2009 using the WAM2LAYERS moisture tracking model. Furthermore, the SST of the source areas have been related to the precipitationshed, that is where evaporation from the ocean contributes to precipitation in a specific area. Not only SST adjacent to East Africa, but also the effects of El Nino 3.4 are assessed in several region. These regions show a variation of influences by local SST and El Nino. For instance, for West Africa the SST in the source region strongly correlates with precipitation in the rainy season and for Australia both local SST and El Nino have a big influence.

The areas relevant for this research are source area 10 delineated in the Southern Indian Ocean near Madagascar and source region 11 delineated in the Northern Indian Ocean and Arabian sea along the East African coast. SST is averaged over the surface area of the source region. The periods 1982 to 2009 for source region 10 show a slight positive correlation between precipitation and SST in some parts of Kenya for the months January, February, May, June, August, October, November, December. For source region 11 a positive correlation between precipitation and SST in some regions of Kenya is found for the months January, February, May, October, November, December. Furthermore, when comparing the precipitation to the SST in the El Nino region some positive correlations are observed for areas in Kenya in the months January, February, October, November, and December, but a negative correlation for the months March, July, August, and September (Van der Ent and Savenije, 2013).

Keys et al. (2022) present a 40-yr analysis of the source for Kenya's precipitation and the fate of evaporation that originates from Kenya. WAM2layers has been used to examine the annual and seasonal changes in moisture source and sinks. Kenya's precipitation consists for 85% of moisture originating from oceanic evaporation whilst 15% originates from continental evaporation, primarily from Kenyan highlands. This area remains an important source of moisture for Kenya. WAM2LAYERS was run at a global scale to calculated fluxes and states overall, after which tracking is used for Kenya specifically to backtrack and forward track moisture.

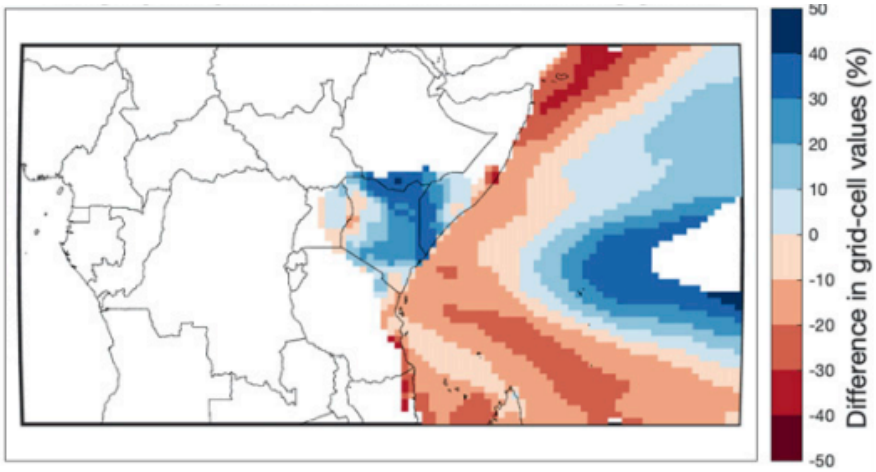


Figure 2.2: Percent difference in annual evaporation contribution to Kenya precipitation between wet and dry years (Keys et al., 2022). Blue indicates that for wet periods relative moisture production is higher for continental moisture and further Eastern in the Indian Ocean. Red periods indicate the regions where relative to wet periods moisture is primarily produced.

This study specifically focuses on the differentiation between wet and dry years in Kenya. During dry season the moisture tends to come from the Northeastern part of the Indian Ocean including the Somali coastline. The long rains season MAM moisture originates from more close to the Kenyan coastline in the Indian Ocean. The cool dry season displays moisture sources from the Southeast Indian ocean and the northern coastline of Madagascar and coastline of Tanzania. The short rains season the dominant moisture source is the coastline of Kenya, the western Indian Ocean and the coastline of Somalia. When comparing the wet and dry years it is found that during wet years the relative moisture contribution for precipitation is largest from continental evaporation from Kenya and moisture further east on the Indian Ocean. In dryer years, moisture originates relatively more to the coastline, and the Arabian Sea becomes relatively more important, as is visualised in Figure 2.2 (Keys et al., 2022).

Climatic drivers and large-scale circulations

Dyer and Washington (2021) concluded that during very wet seasons, the majority of rainfall variation occurs towards the end of March and beginning of April, resulting in an early start to the long rains season and a concentration of more rainfall earlier in the season. Various factors were identified as key drivers of rainfall variability, including local ascent above 700 hPa, mid-tropospheric moisture flux convergence, and moist static energy (latent heat) in the lower and mid-troposphere. Large-scale circulations were also found to play a significant role, with variations in the 200 hPa velocity potential and sub-seasonal correlations with sea surface temperatures in the western Indian Ocean identified as important factors. Moisture availability in East Africa was found to be linked to horizontal advection, with mid-tropospheric moisture strongly influenced by the strength of easterlies over central Africa (Cook and Vizy, 2013; Dyer and Washington, 2021). Black et al. (2003) suggested that during strong low-level easterlies in the North central Indian ocean there is enhanced precipitation in at the western Indian Ocean and drier conditions in the West. This phenomena is tied to the Indian Ocean Dipole (IOD) and corresponding Dipole Mode Index (DMI). It is suggested that a dipole index larger than 0.5°C for three consecutive months leads to enhanced rainfall over East Africa.

This phenomenon has been researched for the very wet short rains season starting from October to December 2019 leading into January 2020. The strong low-level easterlies weaken the westerly flows that usually transport the moisture away from Eastern Africa (Black et al., 2003). The event started early October through January and caused floods and landslides affecting an estimated 2.8 million people. The only precipitation event surpassing this event was in 1997 when large scale flooding occurred as well which is associated with the El Nino effect, whilst for the 2019 event there was no El Nino. There was a strong IOD event with very warm SST off the coast and anomalously cool SST in the eastern part of the Indian Ocean and is among the

strongest IOD events in 30 years. In 1997 the positive IOD led to extremely wet conditions over East Africa. Analysis of sea surface temperature data shows that strong East African rainfall is associated with warming in the Pacific and Western Indian Oceans and cooling in the Eastern Indian Ocean (Black, 2005; Black et al., 2003; Wainwright et al., 2021). Black (2005) suggested that earlier findings of the relationship between East African rainfall and the El Nino Southern Oscillation (ENSO) is manifested by the link between the ENSO and IOD.

Change in climate drivers

Most of the the East African climate is a bimodal precipitation annual cycle, although in some regions a unimodal distribution. East Africa has seen a decline of rainfall during the long rains season (Yang et al., 2015; Cook and Vizy, 2013). Several mechanisms have been proposed for this dying trend where related to the westward extension of the Indo-Pacific warm pool and associated Walker circulation, while other link it to SST over the Pacific basin to a La Nina-like pattern. (Yang et al., 2014) supports the observed drying trend of the long rains season, but assigns its cause to the natural decadal variability associated with the SST in the Pacific Ocean. Climate projections from CMIP5 show an increase in precipitation in the long rains season with global warming, which is opposite from the trend observed up until 2015.

Findell et al. (2019) focuses on historical and projected climate change and how it affects the oceanic and terrestrial moisture contribution to continental precipitation using WAM2layers (Van der Ent et al., 2014). They found that in a warming world, the contribution of oceanic evaporation to continental precipitation becomes larger compared to the contribution of continental evaporation to continental precipitation. Near surface temperatures increase so that the moisture capacity of air above the ocean increases. The theory and observations learn that with global temperatures warming the mean atmospheric moisture content increases and so potentially the total water column and storm intensities are enhanced.

Future projections from climate models indicate increasing rainfall during the short rains under future climate change, suggesting that events such as the increased rainfall during the 2019 short rains could become more frequent under climate change (Black et al., 2003). Cai et al. (2018) explored relationships between strong positive IOD events and global warming and found that with a global warming of 1.5°C the positive IOD events could happen twice as often, implicating more of these heavy rain events in the short rains season. Moreover, the rate at which the Indian Ocean SST is rising is among the fastest of any tropical region. It is found that the short rains season could be lengthened by two months in the twenty first century in the southern Kenya and Tanzania region (Cook and Vizy, 2013).

2.3. Key findings

Based on the studied literature, some key findings could be distilled to support the arguments of how this research is being conducted.

- Recommendations are made suggesting the development of a water balance with the latest bathymetry survey to investigate whether the hydro-meteorological variability alone can account for the observed flooding.
- WAM2LAYERS has been a widely used model for moisture tracking purposes in East Africa and other parts of the world.
- Catchment precipitation and oceanic evaporation in relation to SST and SST gradient in the Indian Ocean are interesting to consider for this study.
- SST in source regions show positive correlations with various regions in Kenya.
- During wet years the relative moisture contribution for precipitation is largest from continental evaporation from Kenya and moisture further east on the Indian Ocean. In dryer years, moisture originates relatively more to the coastline, and the Arabian Sea becomes relatively more important
- Drivers for precipitation in East Africa could be the IOD effect and El Nino Southern Oscillation, as they reveal both positive and negative correlations with precipitation in East Africa.

3

Study Area

For over a decade, the majority of lakes throughout the Kenyan Rift Valley have been steadily rising. One of these lakes is Lake Nakuru, the study area for this research. Despite being a small area of focus, it presents an intriguing case study, due to its complex geo-hydrological conditions, which may shed light on the overflowing of not only Lake Nakuru, but also other lakes in the Rift Valley. A better comprehension of the overall system is first established by detailing the Kenyan Rift Valley, before exploring the specific features of the Lake Nakuru catchment.

3.1. Kenyan Rift Valley

Kenya is situated in East Africa and is characterised by two rain seasons. The long rains season which lasts from March to May and the short rains season lasting from October to December. The temperature in Kenya is consistent over the year, with a mean annual temperature of 25°C, and has a mean annual precipitation of 703 mm (The World Bank Group, 2021). Precipitation in Kenya originates from the Indian Ocean, as low-level mean winds from southeastern direction move over Kenya between June and August, and shift direction to a more northeasterly wind between December and February, with the ITCZ moving over Kenya (Levin et al., 2009).

The ITCZ separates northerly and southerly air flows, whereas the Congo Air Boundary separates westerly flow from easterly low-level flows. The interplay between these two boundaries determines the meteorological situation in Kenya. This results in southeastern flows for the period between June and September called the continental rains season. During the long rains season between March and May and the short rains season between October and December, precipitation originates mainly from the southeastern region just above Madagascar. During dry season between January and February, precipitation originates from the Northeastern Indian Ocean and the Arabian sea.

In addition to these meteorological influences, the topography of Kenya is key for its complex meteorological system. High altitude plateaus between 1000 and 2000 m are interspersed by mountains of over 4000m, influencing weather patterns making them unique for different areas. The complex topography of the Rift Valley is visualised in Figure 3.1 with a range from 300 to 3700 m above mean sea level, and is of importance to the precipitation in the Kenyan Rift Valley. From the Kenyan highlands in the south of the Rift Valley it descents to the north. A channel is formed between two escarpments along the western and eastern boundary of the Rift Valley, as it leads towards the lake Turkana channel in the north, which separates the Kenyan and Ethiopian highlands.

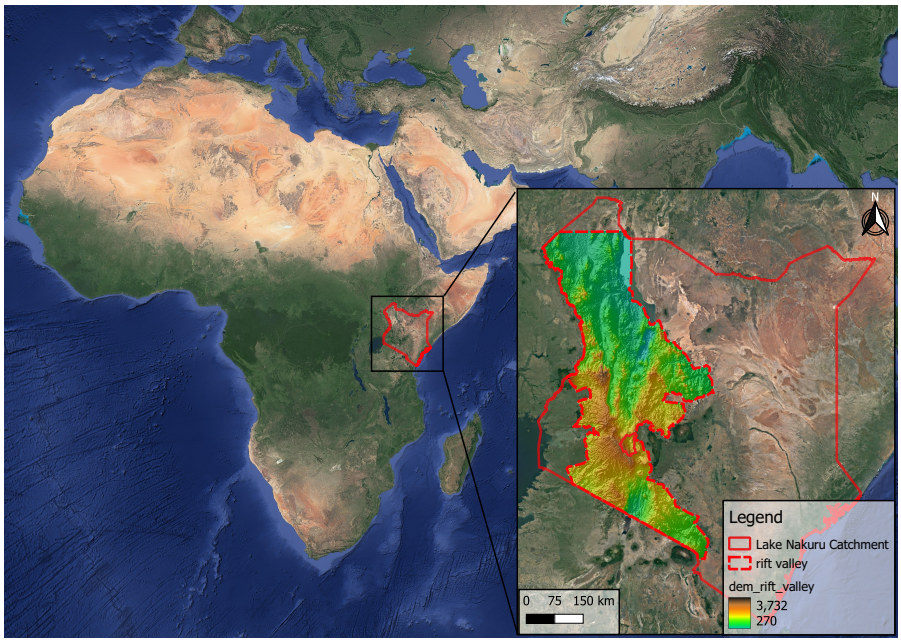


Figure 3.1: Study area of the Kenyan Rift Valley as digital elevation model indicating the wide variations in elevation. The catchment of Lake Nakuru is visualised in red in the middle along the eastern ridge of the Rift Valley on high elevation at about 2000 m.

3.2. Lake Nakuru catchment

Lake Nakuru is a salt water lake and situated at 1760 m above sea level with a catchment area of about 1600 km^2 laying in a closed hydrological basin along the southeastern escarpment of the Kenyan Rift Valley, geographically visualised in Figure 3.1. It experiences mean annual rainfall 876 mm and mean annual evaporation of 1800 mm. Lake Nakuru is a medium sized lake compared to other lakes with a surface area between 30 and 60 km^2 and varying depth between 0.5 and 8 m. Due to shallowness of the lake it is very sensitive to climatic variations as it has little buffer capacity to withstand inter- and intra-seasonal variability and experiences high evaporation rates. During the drought years of 1993 and 1996 this resulted in the lake almost drying out, whilst during wet years like 1997 and 2019 the lake has experiences major flooding (ilec, 2005; Kimaru et al., 2019; Odada, 2001).

Lake Nakuru has minimal underground inflow, and is therefore dependent on seasonal river inflow. Rivers drain into the lake, as well as treated wastewater from Nakuru city (Kanda and Suwai, 2013). Four rivers flow from the western escarpment towards the Lake. Most of the river water infiltrates to recharge groundwater before reaching the lake, and become influent as a result (Kanda and Suwai, 2013). Shallow groundwater flows directly to Lake Nakuru from the east and west rift edges. Deep groundwater is assumed to flow northwards, towards the lower laying parts of the Rift Valley, where Lake Bogoria is situated (Montcoudiol et al., 2019).

The lake is partially encircled by Nakuru city with a population of 400,000 people, that is rapidly expanding, putting stress on water supply and the environment. The lake is listed as a UNESCO heritage site and is known for hosting one of the largest populations of Flamingos. The national park surrounding the lake is among one of the highest biodiversity parks in the country. Over the last couple of decades, the catchment has transformed to a cultivated and highly populated and urbanised area. According to Raini (2009) the catchment has seen major deforestation practices from 1970 to 1986 with more than 80.000 ha deforested in the major forest of the catchment. In 1994, even more forest was assigned to other land-use practices by the government resulting in a 30% forest loss of the southwestern Mau forest. Investigations performed by Kenya Ministry of Environment and UNDP (2021) saw a downward trend in forest cover, while seeing an increase in cropland and wetland between 2000 and 2018. Forestland decreased from 8.5% to 7.7%, with cropland and wetland increasing from 35.4% and 2.6% to 46.6% and 3.3% respectively.

4

Data description

This study employs diverse datasets to portray catchment and atmospheric variables and analyse their long-term variability and trends. The data covers a 41-year period from 1981 to 2021 and are grouped based on the proposed research questions. Paragraph 4.1 clarifies the catchment characteristics, while paragraph 4.2 focuses on lake characteristics. Paragraph 4.3 details the data collection for moisture tracking and climate indicators.

4.1. Catchment characteristics

Precipitation

Accurate precipitation data are important for the purpose of this research. Kenya Meteorological Department (KMD) provided precipitation data of fourteen weather stations, from a period 1981 to 2021 (KMD, 2022). Table 4.1 provides details about latitude and longitude of the stations, as well as the elevation and mean precipitation. The data provided by KMD did not reveal any data gaps, indicating the KMD processed their data before sharing, as raw data over such a time period would likely have data gaps. The weather stations are quite evenly spread over the catchment area resulting in a dense dataset with a radial distance of no more than 10 km between two nearest weather stations.

The Thiessen polygon method was applied to determine the catchment average precipitation and lake precipitation, rather than taking the average over all stations. Lake precipitation is defined by the average Thiessen polygons for weather stations 6, 7, 10, and 11, taking into account contribution area of each weather station over the lake. According to the Thiessen method, the precipitation recorded at a particular station can be extrapolated halfway towards the next station in any direction (Schumann, 1998). These polygons are then used to estimate the relative area that is covered by the weather station's polygon, which is indicated by the area weight in Table 4.1. The KMD precipitation dataset is used in the remainder of this research for catchment wide precipitation analysis, water balance modelling, and moisture tracking analysis.

Evapo(transpi)ration

Lake evaporation and catchment evapotranspiration contribute to understanding the impact precipitation has on the lake. Data on catchment reference evaporation and actual evaporation from the lake are extracted from the Water Productivity Open-Access Portal (WaPOR) version 2. WaPOR is an online platform, developed by the Food and Agriculture Organisation of the United Nations (FAO), which provides access to satellite-based information on water resources in Africa and Southeastern Asia. WaPOR aims to improve water management and food security through providing high quality and open access data on water resources and hydrometeorological variables. The platform combines data from the Sentinel missions, MODIS and aqua satellites that is modelled to provide information on reference evaporation and land cover classifications to name a few. The time series are extracted from the portal using a shapefile for lake Nakuru and catchment boundaries (Food and of the United Nations, 2021).

Table 4.1: Fourteen weather stations from the Kenya Meteorological Department with corresponding location, elevation above mean sea level, mean precipitation, and the average weight according to the Thiessen polygon method (KMD, 2022). The elevations indicate differences of almost 1000 meters and precipitation mean deviates of 300 mm per year.

Station	Latitude [°]	Longitude [°]	Elevation [m]	Precipitation mean [mm/year]	Area weight [%]
Pts2	-0.4433	35.8903	2600	1123	5.2
Pts3	-0.3616	35.9595	2166	987	7.2
Pts4	-0.4521	35.9584	2380	979	6.1
Pts5	-0.5569	35.9432	2756	1050	8.3
Pts6	-0.3032	36.0432	1817	905	8.6
Pts7	-0.4002	36.0436	1951	915	7.3
Pts8	-0.5017	36.0482	2022	815	8.5
Pts9	-0.6143	36.0497	2564	810	11.3
Pts10	-0.2811	36.1231	1860	901	12.0
Pts11	-0.3991	36.1489	1845	1004	6.5
Pts12	-0.5172	36.1505	1838	719	7.9
Pts13	-0.6374	36.1596	2305	845	6.0
Pts14	-0.6253	36.2014	2352	912	2.7
Pts15	-0.4102	36.1805	1893	835	2.3

4.2. Lake characteristics

Lake characteristics are surface area, depth, and volume of lake Nakuru. A combination of two datasets has been used to approach the most accurate representation of the lake's behaviour. The datasets are the Database for Hydrological Time Series of Inland Waters (DAHITI) (Schwatke et al., 2019) and a bathymetry study with echo-sounding technology performed by Jomo Kenyatta University of Agriculture and Technology (JKUAT) (Iradukunda et al., 2020).

Database for Hydrological Time Series of Inland Waters

DAHITI is used to extract the time series for the surface area of the lake following the method introduced by (Schwatke et al., 2019). Landsat data level 2 product 'surface reflectance' and sentinel-2 have been used from January 1984 to June 2018. The Landsat data are processed in Google Earth Engine for cloud-processing whilst sentinel level 2 is already corrected for the atmosphere. As of 2015, Sentinel-2 data became available initiated by the Copernicus program of the European Space Agency (ESA). Two satellites were launched, sentinel-2A and sentinel-2B, providing a temporal resolution of five days. Sentinel Level-1C is used from 2018 for the extraction of the lake surface area after it has been processed. These satellite images are used as an input for the extraction of land-water masks.

Defining the land-water mask is performed by estimating five water indices; the New Water Index (NWI); Tasseled Cap for Wetness (TCwet); Automated Water Extraction Index for Shadow Areas (AWEIsh); Modified Normalised Difference Water Index (MNDWI); and Automated Water Extraction Index for Non-Shadow Areas (AWEInsh). This study combines these five to use the strength of each water index. The method calculates automated thresholds for separating water from land. If four out of five water indices agree whether a grid is water, the grid is assigned a water mask. Grids that are indicated by 0 or 1 water index as water, are set to land. Grids with 2 or 3 water indices are processed through a probability mask, where they get assigned water or land based on the long-term water probability of there being water or land.

The time series for surface area observations of the lake run from 1984 to 2020. Only 56 surface area extractions were made from 1984 to 2009, that is about two per year. Moreover, from 1989 to 2000, only two observations were made. After 2010, a total of 256 surface areas were extracted, that is 21 observations per year on average. Over 25 observations per year are provided for years 2013 to 2018, and 2020. The minimum surface area, depth, and volume, respectively 33.04 (+- 6.98) km^2, 0.73 m, and 0.014 km^3, were observed at 30 January, 2010. However, at 21 January, 1995, and 10 July, 1987, similar minimum values were observed. The lake characteristics reached its maximum of 63.49 (+- 0.6) km^2, 8.20 m, and 0.345 km^3 observed at 4 January, 2021. Within an 11-year time period the lake's surface area increased with 30.45 km^2 representing a 92% growth.

Table 4.2: Bathymetry survey results with depth from the surface of the lake, partial depth (delta_h),surface area at depth, partial volume, and total volume (Iradukunda et al., 2020). The maximum depth at the time of measuring was 6.2 m and a maximum surface area of 56.8 km^2.

Depth [m]	[m]	Surface area [m^2]	Partial volume [m^3]	Total volume [m^3]
6.2	0	0	0	0
6.0	0.2	1.23E+06	1.43E+04	1.43E+04
5.5	0.5	3.18E+07	1.04E+07	1.04E+07
5.0	0.5	3.73E+07	1.74E+07	2.79E+07
4.5	0.5	3.97E+07	1.93E+07	4.72E+07
4.0	0.5	4.18E+07	2.04E+07	6.76E+07
3.5	0.5	4.38E+07	2.14E+07	8.90E+07
3.0	0.5	4.54E+07	2.24E+07	1.11E+08
2.5	0.5	4.70E+07	2.32E+07	1.35E+08
2.0	0.5	4.83E+07	2.39E+07	1.59E+08
1.5	0.5	4.98E+07	2.47E+07	1.83E+08
1.0	0.5	5.20E+07	2.56E+07	2.09E+08
0.5	0.5	5.47E+07	2.70E+07	2.36E+08
0.0	0.5	5.68E+07	2.89E+07	2.65E+08

Bathymetry survey

The bathymetry study was performed by Iradukunda et al. (2020) in June 2019. The depth, surface area, partial volume, and total volume have been extracted with echo-sounding technology, from which the results are listed in Table 4.2. The lake variables are given for every 0.5 m layer in the lake. Measured from the surface, the lake has a maximum depth of 6.2 m. The partial volume of the lake gives the volume in that specific lake section, e.g. between 6 m and 5.5 m depth. The surface area is given for the top layer of this specific section, e.g. the surface area at 5.5 m depth is 3.18e07 m^2. The total volume is the accumulated sum of partial volumes. The bathymetry survey data are used to establish a relation between the measured surface area and volume. The bathymetry study describes the state of the lake at one specific point in time. In order to create an accurate time series for the volume of the lake, the surface area time series from DAHITI is used to fit the lake volume to lake surface area relation from bathymetry. DAHITI also provides data on the volume, which is estimated by combining the depth from altimetry and surface area from optical imagery. The volume is estimated from the minimum observed surface area, so it does not incorporate the actual bottom of the lake. It is for this reason that the relation between bathymetry observed surface area and volume is used to create a time series for the lake volume.

4.3. Moisture tracking and climate data

The moisture tracking model WAM2layers requires input data from surface variables and model level variables. ERA5 hourly data on single levels is used for this research (Hersbach, 2023). ERA5 replaces the ERA-Interim reanalysis data and provides data from 1940 to present. The data are acquired with a script that makes use of an CDS API key to download the data from the ERA5 database. The data are downloaded as NetCDF files for the domain 0 to 100 longitude and 40 to -40 latitude and time of interest from 1981 to 2021. The downloading script and WAM2layers code is provided on the GitHub repository (van der Ent et al., 2022).

The model level data is comprised of 22 model levels, a temporal resolution of one hour, and a grid size of 0.5°. Model variables are u and v component of wind and specific humidity. The u component of wind is the eastward component of the wind. The v component of wind is the northward component of wind. Both wind components are combined to define the horizontal wind speed and direction. The specific humidity gives the mass of water vapour per kilogram of moist air. The variables are defined for each grid per model level.

Surface variables used by the model are; total precipitation; evaporation; surface pressure; and total column of water. Total precipitation defines the accumulated liquid and frozen water falling to the earth's surface. The sum is taken for large-scale and convective precipitation. It omits fog, dew or precipitation that evaporates before it hits the Earth's surface. Evaporation represents the accumulated water that has evaporated from the surface of the Earth and includes a simplified version of transpiration. Surface pressure is the weight of all air in a vertical column above the surface of the Earth with unit Pascals. It is the pressure of the atmosphere on the land, sea surface, and in-land water. The total column of water is the accumulated amount of water in a vertical column from the Earth's surface to the top of the atmosphere. It comprises water vapour, liquid water, snow, cloud ice, and rain, in kg m^{-2}.

In various studies, sea surface temperature (SST) is associated with precipitation over land and is suggested to be an important driver for precipitation. SST is used in this research to assess the impact of it for different source regions on precipitation. The European Centre for Medium-Range Weather Forecasts (ECMWF) provides the Ocean Reanalysis System 5 (ORAS5). It is a combination of model data with observations globally covering 1979 to present at a resolution of 0.25°. Two global SST indices are used as well namely El Niño Southern Oscillation (ENSO) and the Indian Ocean Dipole acquired through NOAA (Huang et al., 2017).

5

Methodology

In this chapter, the research methods used in the study are outlined. The focus in Paragraph 5.1 is on the spatial and temporal analysis of precipitation with an emphasis on change point detection, which reflects on the first research question. Paragraph 5.2 examines how the variability of precipitation impacts the lake through the application of a water balance model. Paragraphs 5.3 and 5.4 describe the methodology used for moisture tracking, and moisture source analysis in relation to climate drivers respectively.

5.1. Spatial and temporal catchment analysis

The overflowing of lake Nakuru is a rather abrupt and sudden process initiated around the early 2010s, from which a steady rise is observed up until 2021 given historical data. It is therefore of considerable interest to look at sudden changes in precipitation patterns. To validate change points, an offline detection method is used, provided by the Ruptures package introduced by Truong et al. (2020). The algorithm takes a time series signal as input, where the goal is to find the optimal segmentation. It uses a built-in cost function including MSE, Binary Segmentation (binseg), and the Dynamic Time Warping (dtw). A recursive algorithm finds the optimal segmentation. The algorithm calculates the minimum cost for all possible segments up to that time. The algorithm backtracks to recover the indices of the change point once the optimal segmentation is found. Tuning parameters in the algorithm can be adjusted to improve the performance of the algorithm, including the penalty parameter that controls the complexity of the segmentation and the minimum size and jump parameter that control the minimum segment size and the maximum distance between two consecutive change points.

The significance of the change point is tested with the Mann-Whitney U test at a significance level of 5%, if a change point candidate is identified by the algorithm. The Mann-Whitney (MW) U test, also known as the Wilcoxon rank-sum test, does not demand a normal distribution and is suitable for smaller data sets. It also assumes two sub-samples were extracted from a larger sample of the same population. It assesses the significance of difference in median or mean over the two time series. If the segmentation is proven significant the first year after the year of change is called a change point. The MW test has been progressively used in recent decades for assessing changes and steps in hydrological time series like streamflow and precipitation. For instance, Liang et al. (2011) applied the method to find jumps in precipitation time series in the northeast of China and Keim and Muller (1992) investigated whether local heavy rainfall regimes had changed through the analysis of annual maximum storm series.

The KMD weather station data are assessed among each other to find statistical relations for elevation, distance apart, and rainfall amount. Change points in catchment precipitation are further analysed based on changes in climatology and daily, monthly and yearly precipitation. The SPI has been calculated for every year in the time series (WMO, 2012) to provide insight into the number of dry and wet years. Eventually periods of interest are distilled for further moisture tracking and climate driver analysis.

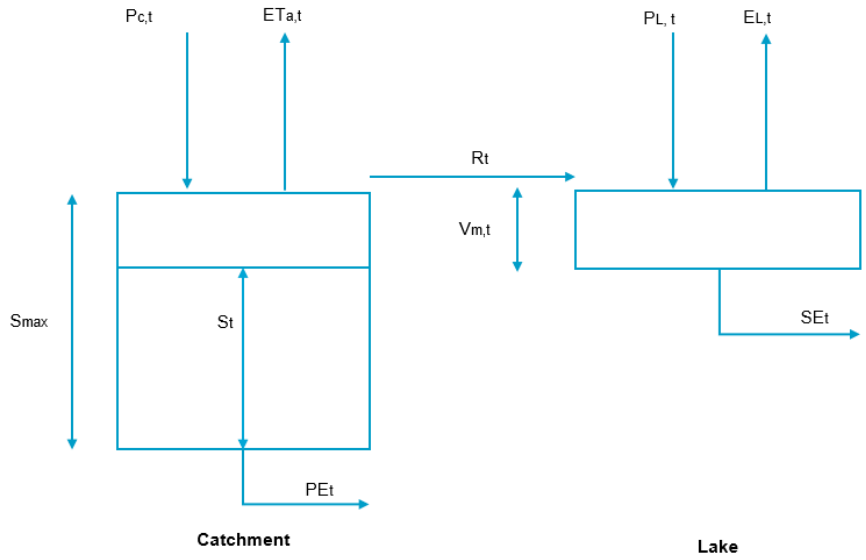


Figure 5.1: Water balance model for lake Nakuru presenting the catchment and lake model components; catchment precipitation ($P_{c,t}$), catchment actual evapotranspiration ($ET_{a,t}$) soil storage (S_t), groundwater percolation (PE_t), total catchment runoff (R_t), modelled lake volume ($V_{m,t}$), lake precipitation($P_{L,t}$), lake seepage (SE_t), and lake evaporation ($E_)L,t$) at time t and the maximum soil storage of the catchment (S_{max}).

5.2. Water balance model

This paragraph is divided into three subsections to detail the water balance modelling method. The first subsection explains the conceptual water balance model, including its corresponding components and relations. The second subsection describes the Monte Carlo model calibration, followed by an explanation of the methods used to interpret the model including a component analysis and a multiple linear regression model.

5.2.1. Conceptual model

To relate precipitation to the behaviour and fluctuations of the lake, a simple water balance model is introduced, conceptually visualised in Figure 5.1, with a daily time resolution (t=1 day). Model inputs consist of time series from 2009 to 2021 and are classified as; catchment precipitation; catchment potential evapotranspiration; lake precipitation; lake surface area; and lake actual evaporation. Lake Nakuru has no known surface water outlets (Kenya Ministry of Environment and UNDP, 2021). Shallow groundwater as a result of precipitation flows directly to Lake Nakuru from the east and west rift edges and from the south and from the north. Deep groundwater is assumed to flow northwards towards the lower laying parts where Lake Bogoria is situated (Montcoudiol et al., 2019). The outgoing flux for the lake would be actual evaporation ($E_{a,t}$) from the lake's surface and seepage (SE_t) from the bottom of the lake to deeper groundwater. Seepage from the lake is assumed to be low, for the reason that this is a salt water lake. The precipitation on the lake ($P_{L,t}$) and total catchment runoff (R_t) as a result of catchment response and runoff coefficient are the two incoming lake fluxes.

Catchment characteristics are also important to consider as it determines the response to precipitation. The maximum storage (S_{max}) is the average maximum amount of water that can be stored over the entire catchment area in the root zone. The incoming flux to the catchment is the catchment precipitation ($P_{c,t}$). The outgoing catchment fluxes would be groundwater percolation (PE_t) and actual evapotranspiration ($ET_{a,t}$). Percolation is assumed to not flow to the lake but rather recharge groundwater and subsequently flow downstream to lower laying areas in the Rift Valley. The catchment experiences a reference evapotranspiration used as an input for the model to calculate actual evapotranspiration based on the amount of water in the soil with respect to the maximum soil capacity. The output of the model comprises a time series of daily resolution of the catchment storage, actual catchment evapotranspiration, catchment runoff, total lake volume, and change in lake volume which is the volume that leaves or enters the lake at time t.

The equation for the actual evapotranspiration of the catchment at time t $ET_{a,t}$ is provided by equation 5.1, $ET_{ref,t}$ is the catchment reference evapotranspiration, S_t is the catchment storage at time t and S_{max} is the maximum storage of the catchment. Equation 5.2 and 5.3 represent the percolation PE_t and runoff coefficient $r_{c,\,t}$ at time t where PE_{max} and $r_{c,max}$ are their maximum defined parameters. Equation 5.4 describes lake seepage at time t SE_t which is a function of the lake area and maximum seepage for the lake SE_{max}.

$$ET_{a,t} = E_{ref,t} \times \frac{S_t}{S_{max}} \tag{5.1}$$

$$PE_t = PE_{max} \times \frac{S_t}{S_{max}} \tag{5.2}$$

$$r_{c,\,t} = r_{c,max} \times \frac{S_t}{S_{max}} \tag{5.3}$$

$$SE_t = SE_{max} \times \frac{A_{lake}}{A_{lake,max}} \tag{5.4}$$

Equation 5.5 describes lake influx from catchment runoff R_t, where P_t is the catchment precipitation at time t.

$$R_t = \begin{cases} r_{c,\,t} \times P_t & P_t \neq 0, S_t \leq S_{max} \\ r_{c,\,t} \times P_t + S_t - S_{max} & P_t \neq 0, S_i > S_{max} \\ 0 & P_t = 0, S_t \leq S_{max} \end{cases} \tag{5.5}$$

Equation 5.6 describes the catchment storage at time t S_t where the storage of previous time step S_{t-1} is considered with fluxes entering the catchment storage, namely catchment precipitation P_t and fluxes leaving the catchment storage i.e. catchment runoff at time t R_t, actual evapotranspiration at time t $ET_{a,t}$, and percolation at time t PE_t. If the catchment storage at time t exceeds the maximum catchment storage, the minimum value of both is taken for the catchment storage at time t+1 S_{t+1}.

$$S_t = \begin{cases} S_{t-1} + P_t - R_t - ET_{a,t} - PE_t & P \neq 0 \\ S_{t-1} - ET_{a,t} - PE_t & P_t = 0 \end{cases} \quad S_{t+1} = \begin{cases} min(S_t, S_{max}) & S_t > S_{max} \\ S_t & S_t \leq S_{max} \end{cases} \tag{5.6}$$

Eventually, the total volume contribution to the lake delta $V_{lake,mod,t}$ at time t is described by Equation 5.7, and totalled by the incoming fluxes at time t; runoff volume a time t $V_{runoff,t}$, lake precipitation volume $V_{P_L,t}$, and outgoing fluxes at time t; lake evaporation volume $V_{E_L\,t}$ and the lake seepage volume V_{SE_t}.

$$\triangle V_{lake,mod,t} = V_{runoff,t} + V_{P_L,t} - V_{E_L\,t} - V_{SE_t} \tag{5.7}$$

Equation 5.8 describes the total volume at time t with the lake volume at time t-1 $V_{lake,mod,t-1}$ and the total volume contribution at time t from catchment and lake interactions delta $V_{lake,mod,t}$.

$$V_{lake,mod,t} = V_{lake,mod,t-1} + \triangle V_{lake,mod,t} \tag{5.8}$$

5.2.2. Model calibration

The model is calibrated using a Monte Carlo simulation on the four parameters to represent the volume of the lake explained in Table 5.1. The ranges provided in the table are set to a certain range to reduce equifinality, as more parameter sets can perform equally well. Furthermore, there are four input parameters to consider, specifically runoff coefficient, maximum storage of the subsurface, percolation, and seepage. A wide but reasonable parameter range is chosen after literature and hand calibration. This means a total of 14256 parameter sets are made. A multiple objective function is used to select the right parameter set. The NSE and NSE_{log} are calculated for each parameter set and are described in the Equations 5.9 and 5.10. Finally

Table 5.1: Water balance model parameters with unites, range for calibration, and description.

Parameter	Unit	Range	Step	Description
S_{max}	[mm]	100 - 500	50	The maximum storage of the catchment sub-surface
$r_{c,max}$	[-]	0.05 - 0.15	0.01	Runoff coefficient of the catchment depending on catchment storage
SE_{max}	[mm/d]	0 - 4.8	0.4	Seepage from the lake to groundwater depending on lake volume
PE_{max}	[mm/d]	0 - 4.8	0.4	Percolation from the catchment sub-surface to groundwater depending on catchment storage

the Euclidean distance (D_e) is calculated for each parameter set to obtain the most balanced parameter set and is described in Equation 5.11. V_m is the modelled volume and V_o is the observed lake volume. As the lake volume is very sensitive to sudden changes and experiences high variability month to month, it is important to model both the high and low extremes. This multi-objective calibration therefore aims to model both the highs and lows using the NSE and NSE_{log} objective functions. The parameter set with the lowest D_e is often defined as the most balanced model and in this case taken as the parameter set for this model (Hrachowitz et al., 2014).

$$NSE = 1 - \frac{\sum_{i=1}^{n}(V_{m,i} - V_{o,i})^2}{\sum_{i=1}^{n}(V_{o,i} - \overline{V_o})^2} \tag{5.9}$$

$$NSE_{log} = 1 - \frac{\sum_{i=1}^{n}(log(V_{m,i}) - log(V_{o,i}))^2}{\sum_{i=1}^{n}(log(V_{o,i}) - log(\overline{V_o})}\Big)^2 \tag{5.10}$$

$$De = \sqrt{(1 - NSE)^2 + (1 - NSE_{log})^2} \tag{5.11}$$

5.2.3. Precipitation and lake behaviour

To investigate the difference in model behaviour to precipitation a monthly component analysis is conducted to evaluate the influence of precipitation on lake variability. The analysis is restricted to the period from 2009 to 2020, when sufficient measurements of lake characteristics are available. The analysis involves creating precipitation bins of approximately 25 mm, as depicted in Figure 5.2 according to the Freedman-Diaconis rule (Freedman and Diaconis, 1981), and assessing the monthly volume variations within each bin. To examine the impact of precipitation on the volume variation, all four fluxes contributing to lake volume change, namely lake precipitation, lake evaporation, lake seepage, and catchment runoff, are analysed. This analysis leads to the identification of three hypotheses, which aim to explain why precipitation events have different effects on the lake volume:

1. Examining the entirety of factors that influence lake volume, including lake precipitation, evaporation, seepage, and catchment runoff, instead of solely focusing on precipitation, can provide a more comprehensive explanation for the observed lake volume behaviour.

2. The fluctuation of the lake is closely linked to the variability of rainfall during the month.

3. The lake's variability is influenced by the precipitation from the previous month and the current month, impacting the catchment water storage capacity.

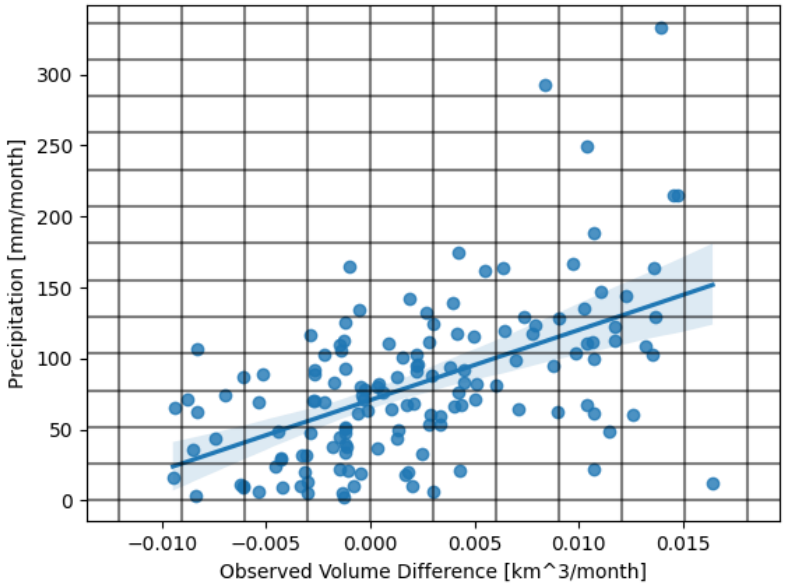


Figure 5.2: A scatter plot representing the precipitation bin distribution for component analysis and relation between monthly precipitation and monthly observed volume change. Take note of the different bins which should be viewed horizontally, where for the same range of precipitation, a wide range of volume differences are observed.

Subsequently, to understand the lake volume dynamics and contributing components, multiple linear regression (MLR) is performed (Holder, 1985). The general equation for MLR is:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \varepsilon \tag{5.12}$$

where:

- y is the dependent variable (the variable that is being predicted or explained)
- $x_1, x_2, \ldots, x_n$ are the independent variables (the independent variables that are used to predict the value of y)
- $\beta_0, \beta_1, \ldots, \beta_n$ are the regression coefficients (the parameters that determine the slope of the regression line)
- ε is the error term (the part of y that cannot be explained by the independent variables)

The independent variables are model outputs lake precipitation, lake evaporation, lake seepage, and catchment runoff. The dependent variable that is predicted is the lake volume. The coefficients of the independent variable give an impression of the importance of the variable towards predicting the volume of the lake. The goal of MLR is to estimate the values of the regression coefficients that minimise the sum of squared errors between the predicted values of y and the actual values of y. This is done using a method called ordinary least squares regression. It finds the values of the regression coefficients that minimise the residual sum of squares, which is the sum of the squared differences between the predicted values of the dependent variable and the actual values of the dependent variable. It is applied in various studies such as estimating the relationship between weather variables (rainfall, temperature, and other climate indices) and agricultural productivity (measured by crop yield) (Mediero and Kjeldsen, 2014).

5.3. Moisture tracking

The model WAM2layers is used for moisture tracking to track precipitation from the catchment area 10 days backward in time (Van der Ent et al., 2014). This is an offline tracking algorithm that makes use of Eulerian computation. The region of interest to perform backtracking is from -40 to 40 latitude and 0 to 100 longitude. This area is based on previous backtracking research for East Africa. Furthermore, the precipitation domain is given by the region from which backtracking is performed and is -0.7 to -0.2 latitude and 35.7 to 36.3 longitude to be representable for the lake Nakuru catchment area.

The next step is to process the data, so the model can read it and perform the backtracking analysis. It includes accumulating the data into two layers, deriving moisture fluxes from the available ERA5 data, and converting the data into the right units. In order to start the reprocessing, a configuration file is built with the correct date range and file locations.

After the processing step, the data is ready for the backtracking functionality of the model. The backtracking experiment tracks precipitation back in time. The model uses a file that defines the source region from where to track moisture. The source region file has latitude and longitude data as well as a grid containing 0 and 1 where 1 indicates in this case the lake Nakuru catchment and has the same domain as the downloaded data and thus the preprocessed files. Again the configuration file should contain the time period for the precipitation event that is backtracked, and the duration over which backtracking is performed. The data is used to solve the water balance for the tagged moisture for each grid and in two layers, called the upper and lower layer. The water mass balance that is solved is given by equation 5.4 (Findell et al., 2019; Van der Ent et al., 2013):

$$\frac{\partial S_{g,lower}}{\partial t} = \frac{(S_{g,lower}\, u)}{\partial x} + \frac{\partial (S_{g,lower}\, v)}{\partial y} + E_g - P_g \pm F_{v,g} \tag{5.13}$$

Where g describes the mass of water that has being tagged from the source area; t is time; u and v are the wind components; P is the surface precipitation, E is the surface evaporation, and F_v is the exchange between layers over the vertical. The upper and lower layer backward trajectory are somehow similar except that E_g is expected to only enter the lower layer. The fluxes are calculated over the boundaries of each grid cell over the u and v direction. Vertical exchange is a strong assumption as it tries to capture all exchanges of moisture due to rainfall that does not reach the surface, convection, and turbulence. However, results were reasonable physically in previous studies (Van der Ent et al., 2013).

5.4. Climate indices

To interpret the moisture tracking, this research makes use of climate indices to find relations with precipitation and lake overflowing. Moisture sources are of great importance when looking for reasons why precipitation patterns are changing. Therefore, this research investigates different moisture sources for catchment precipitation both spatially and temporally and further looks into larger scale global circulations that are at hand and could influence precipitation patterns for the time periods of interest, i.e. El Niño and the Indian Ocean Dipole (IOD).

5.4.1. Moisture sources

A distinction is made between land and sea evaporation contribution to the precipitation in the catchment. As Keys et al. (2022) suggests for Kenya as a whole, it is found that during wet years the relative moisture contribution to precipitation is more focused on continental evaporation and oceanic evaporation further east on the Indian Ocean. In dryer years, moisture originates relatively more from the coastline, and the Arabian Sea becomes relatively more important. It is also suggested that Kenya's precipitation consists for 85% of moisture originating from oceanic evaporation, whilst 15% originates from continental evaporation. Furthermore, Findell et al. (2019) suggested that in a warming world, the contribution of oceanic evaporation to continental precipitation becomes larger compared to the contribution of continental evaporation to continental precipitation. Near surface temperatures increase so that the moisture capacity of air above the ocean increases.

Not only the distinction between land and sea contribution is made. Moreover, sea contribution is divided into sea regions that are relevant to the time of year and moisture trajectory. Regarding the time series of interest, specific sea regions are identified based on literature and backtracking results. In defining the sea regions, distinctions are made between the coastal moisture source, off the coast evaporation in the western Indian ocean, and evaporation further east in the Indian Ocean contributing to precipitation in the catchment. In this way it is possible to identify sea regions that contribute to precipitation in the catchment for the tracking time period and account for possible changes. Precipitation throughout the year experiences different contribution regions of evaporation as wind patterns and SST shift with the ITCZ.

5.4.2. Sea surface temperature and climate indices

Changes in SST are key for investigating precipitation variability in the catchment. Analysis of SST data shows that strong East African rainfall is associated with warming in the Pacific and Western Indian Oceans and cooling in the Eastern Indian Ocean (Black, 2005; Black et al., 2003; Wainwright et al., 2021). Yang et al. (2015) observed that during dry season, from January to February, the southern part of the ocean is warmer compared to the northern ocean region and Arabian sea and displays a north-south SST gradient. During the long rains season, form March to May, the Indian Ocean gradually heats up, where the warmer temperatures are observed in front of the Kenyan Coast and more eastern in the Indian Ocean. During the continental rains season, from June to September, the ocean cools off with lower temperatures near the coast and higher lower temperatures deeper into the Indian Ocean. In the short rains season, from October to December, the same distribution of SST is seen as during the long rains season, although temperatures are lower.

Changing SST in the Indian Ocean can influence evaporation rates and moisture transport over East Africa, which in turn can impact precipitation patterns in the region. A warmer SST in the western Indian Ocean can increase the amount of water vapour in the atmosphere through evaporation, which can enhance rainfall over East Africa. Conversely, cooler SST in the eastern Indian Ocean can decrease evaporation rates and moisture transport, which can lead to drier conditions in the region (Gimeno et al., 2012). The key physical factors that affect evaporation rates on the ocean include SST, atmospheric moisture content, wind speed, and air-sea temperature and humidity differences. These factors can influence both the rate and spatial distribution of oceanic evaporation . Gimeno et al. (2013) presented results from a modelling study that analyses the relative importance of local and non-local controls on evaporation rates. The authors find that local controls, such as SST and atmospheric moisture content, are the dominant factors influencing evaporation rates on short timescales (hours to days). However, non-local controls, such as large-scale atmospheric circulation patterns and oceanic mixing, can also play a significant role in controlling evaporation rates on larger timescales (weeks to months). Regarding local controls of the evaporation rate, the sea regions identified for the time span of interest are analysed on their difference in SST. Black (2005) suggested that earlier findings of the relationship between East African rainfall and the ENSO is manifested by the link between the ENSO and IOD. Non-local controls include the ENSO and the IOD. These are compared with precipitation periods of interest for the catchment through Pearson correlation.

6

Results

The Results chapter of this thesis presents the findings of the study in a structured manner, by sequence of the research questions. The chapter begins with a spatial and temporal analysis of precipitation in the catchment of Lake Nakuru, followed by a detailed description of the water balance model that examines the interaction between precipitation and the lake. The chapter then investigates moisture tracking and sea surface temperature dynamics, and explores the impact of global circulations on the study area.

6.1. Spatial and temporal precipitation analysis

Change point analysis has been conducted on the individual and the Thiessen interpolated times series, lake precipitation, consisting of weather station 6, 7, 10 and 11, and catchment precipitation consisting of all weather stations. Figure 6.1 visualises the KMD weather station locations and whether the weather station experiences a change point in annual precipitation at 2010. Those that experience a change point are visualised in green, others in red. Nine out of fourteen weather stations experience a significant change point at 2010. This section is supported by supplementary material in Appendix A, where Table A.1 lists the change point and significance for every weather station, as well as lake precipitation and catchment precipitation. Weather stations around the lake, except for seven, experience a change point, as well as weather stations at the northwestern and eastern part of the catchment. No significant change points are observed in higher elevated areas in the southwestern and middle part of the catchment. Furthermore, higher laying weather stations seem less likely to experience a change point. In addition, both lake precipitation and catchment precipitation experience a significant change point at 2010.

The weather stations reveal diverse changes in precipitation when comparing the average precipitation prior and post change point. Station 11 and station 14 show a surge in precipitation of 41% and 48%, whereas station 8 and 9 remained roughly the same. Additionally, the other stations show increased precipitation of 13% for station 5, and 28% for station 15. Catchment precipitation shows a lower increase compared to lake precipitation, where they increase 18% and 25%, from 849.11 mm to 1005.05 mm, and 871.26 mm to 1090.65 mm, respectively. Comparing KMD stations among each other revealed the relationships between relative distance and precipitation. The correlation for stations within a radius of 15 kilometres suggests a high positive correlation above 0.75. As distance progresses the correlation becomes lower down to 0.63 on average, with extremes to a low correlation relation between certain weather stations. Furthermore, precipitation generally increases with elevation.

Figure 6.2 presents the baseline climatology of the catchment, established by averaging precipitation for each month from 1981 to 2009 and 2010 to 2021. Compared to the baseline, January shows a decrease in precipitation of -21% (-7.60 mm), while September exhibits the largest change of 64% (+37.81 mm). October and December also show notable changes, with increments of 30% (+21.14 mm and +15.95 mm, respectively). April (+33.66 mm), May (+23.07 mm), and July (+13.84 mm) represent changes of 20% to 30%, while March (+8.75 mm) and June (+6.31 mm) exhibit changes of around 10%. In contrast, February (-0.49 mm), August (+1.83 mm), and November (+1.68 mm) show only minor differences, with changes of 2% or less.

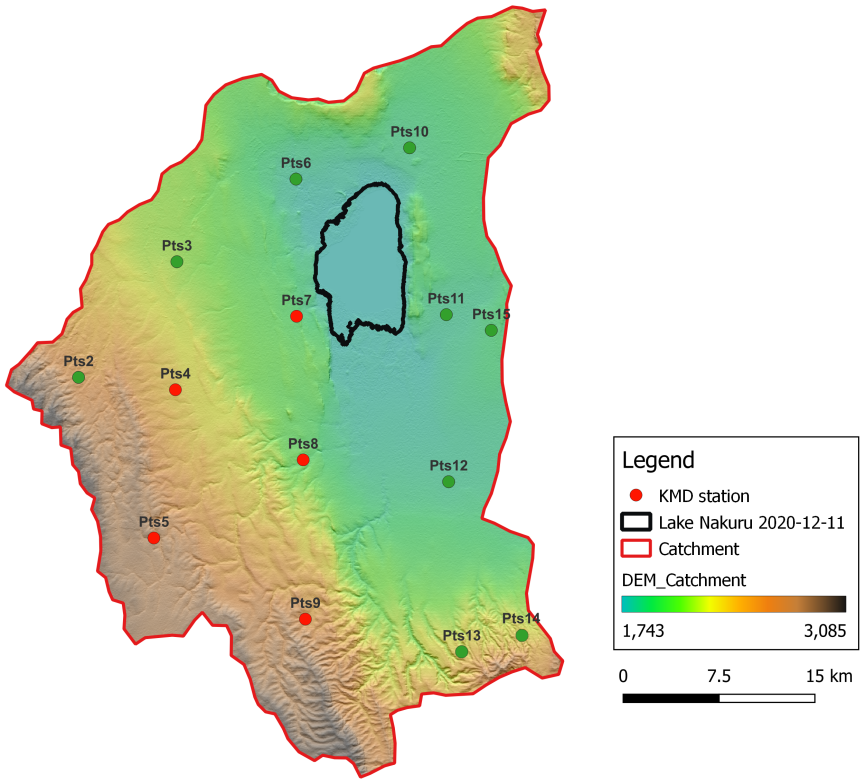


Figure 6.1: DEM map of lake Nakuru Catchment with delineation of lake Nakuru on December 11th, 2020. The KMD weather stations are visualised by their location and if there is a significant change point detected in 2010 (green) or not (red). Change points are observed in 9 out of 14 weather stations.

Before the change point, the catchment experienced one moderately wet year with 1 < SPI < 1.5 (1988) and one moderately dry year with -1 < SPI < -1.5 (2004). In addition, there were two very dry years with -1.5 < SPI < -1.75 (1999, 2000) and two extremely dry years with SPI < -2.0 (1984, 2009). From 2010 to 2021, the catchment experienced two moderately wet years (2010, 2014), two very wet years in a row with 1.5 < SPI < 1.75 (2018, 2019), and one extremely wet year with SPI > 2.0 (2020). It is noteworthy that no moderately dry or drier years were observed after the change point.

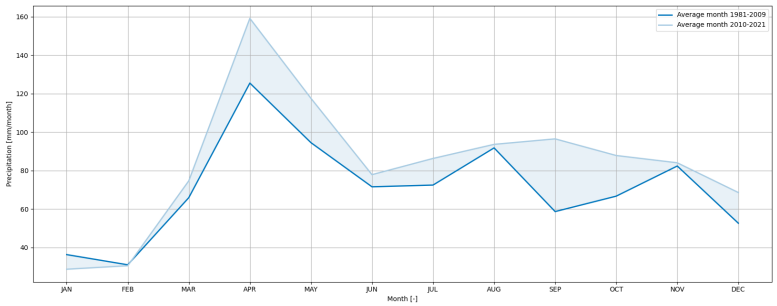
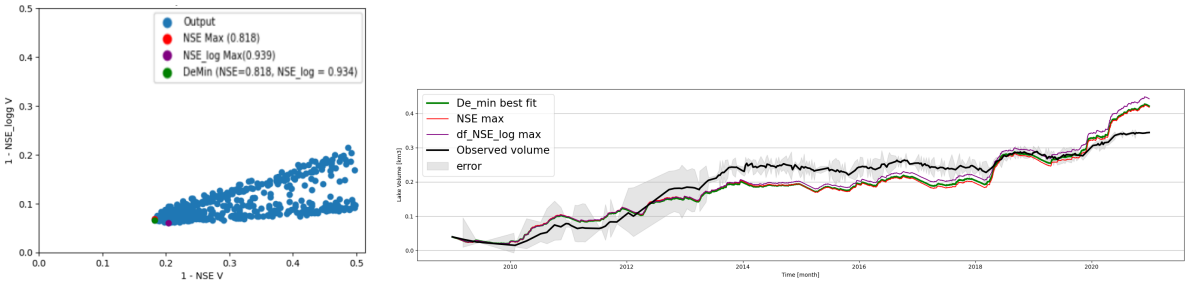


Figure 6.2: Climatology for lake Nakuru catchment displaying the difference between the average monthly precipitation prior to 2010 and post 2010. The shaded blue area displays the change in average monthly precipitation. Major relative and absolute increases are observed for April and September.



(a) Solution space model calibration

(b) Modelled volume for optimal parameter sets

Figure 6.3: Multi-objective calibration results with the solution space and optimal parameter sets for the water balance model where the objective functions equal 0.818, 0.939, and 0.194 for NSE (red), NSE_{log} (purple), and D_e (green) respectively (a) and time series of the observed volume in black and corresponding error, and the three optimal parameter sets for NSE, NSE_{log}, and D_e (b). Lower volumes and higher volumes are over-predicted whilst the mid range values are more under-predicted and parameter sets visually perform near equally well.

Table 6.1: Parameter sets for calibration results of the water balance model for the optimal objective functions NSE, NSE_{log}, and D_e

Variable	NSE	NSE_{log}	D_e
S_{max} [mm]	400	450	400
$r_{c,max}$ [-]	0.1	0.08	0.09
SE_{max} [mm/d]	2.0	1.6	1.6
PE_{max} [mm/d]	3.6	2.4	3.6

Table A.2 shows a change point in monthly precipitation over the catchment in September and October of 2010. Catchment precipitation has increased by 155.94 mm annually, with April and September contributing to 46% of this increase with 33.66 mm and 37.81 mm, respectively. The intensity of precipitation in these months has also increased. For instance, in April, there were 0.9 days per month with a 15 mm precipitation over the catchment before the change point, while after the change point, this increased to 2.7 days per month. Similarly, in September, there were 0.13 days per month with a daily precipitation intensity of 12 mm over the catchment area before the change point, whereas after the change point, this increased to 1.4 days per month. Figure A.1 and A.2 provide visualisations of the yearly and monthly precipitation intensities for these two significant months.

6.2. Water balance modelling

6.2.1. Precipitation and lake interaction

In order to identify the months of interest, it is important to understand how the lake responds to varying precipitation amounts in each month. Figure B.1 visualises the observed lake surface area and resulting lake volume over time, indicating an inclination from 2010 onwards. This section is supported by supplementary material in Appendix B, where Table B.1 provides the average monthly precipitation for each month during the model period (2009-2020), along with the accumulated volume contribution and the Pearson correlation of precipitation with observed monthly volume changes. It is found that three months, namely February, March, and October, contribute on average to a decline in lake volume with -0.024, -0.012, and -0.010 km^3, respectively. On the other hand, the months of April, May, and September contribute the most to the lake volume, with 0.0619, 0.0623, and 0.0532 km^3, respectively, and they also experience the highest average precipitation. Notably, prior to 2010, September had one of the lowest precipitation amounts, whereas it currently experiences the third-highest. The correlation between precipitation and observed volume change differs for each month. Specifically, April precipitation exhibits a high positive correlation with monthly volume change of 0.820, while May precipitation shows no to weak positive correlation with monthly volume change of 0.260. September precipitation demonstrates a moderate positive correlation with observed volume change of 0.527.

6.2.2. Model calibration

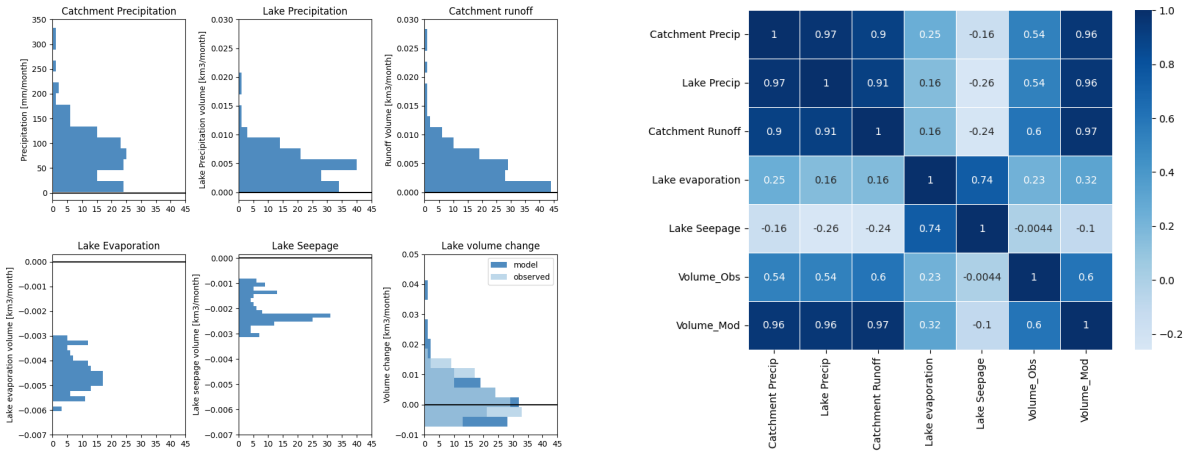
A total of 14256 parameter sets have been tested on the water balance model for which the solution space is visualised in Figure 6.3a. The parameter set with the lowest D_e values is taken as the most balanced model with a D_e value of 0.194 that include a NSE and NSE_{log} of 0.818 and 0.934 respectively. The parameter sets for the multi-objective function calibration are given in Table 6.1. The three parameter sets have maximum catchment storage and runoff coefficient roughly within the same range considering the total range for calibration. Figure 6.3b visualises the modelled volume for each optimum parameter set. What is chosen to be the most balanced model gives a maximum catchment storage of 400 mm, a runoff coefficient of 0.09, maximum lake seepage of 1.6 mm d^{-1} and maximum percolation from the catchment of 3.6 mm d^{-1}. The parameter set gives the maximum value as they are intertwined with catchment and lake variability. For all three models there is an overestimation for the lower volumes, however within the error range for the observed volume. Around an observed lake volume of 0.25 km^3 an under-prediction is noticed mostly outside of the error range. The dynamics are roughly the same, although lower. For lake volumes around 0.25 to 0.3 km^3 the model shows a good and mostly equal representation. Above 0.3 km^3 the model is over predicting the observed lake volume.

6.2.3. Model Output

To understand how the model behaves Figure 6.4a displays the histograms for the main model input, lake evaporation, catchment precipitation, and lake precipitation, and model outputs, catchment runoff, lake seepage, and modelled volume change per month. Catchment precipitation typically ranges from 0 to 125 mm per month, with maximum precipitation months extending to 300 mm. The catchment precipitation is translated to catchment runoff through the runoff coefficient and catchment storage. The rather uniform histogram for precipitation is translated to a exponentially decreasing catchment runoff histogram. The number of lower runoff volume is higher compared to low precipitation amount that could that the catchment has a dampening effect on precipitation. Lake precipitation follows roughly the same distribution as catchment precipitation as those are very well linked with a correlation of 0.97 where catchment precipitation and runoff experience a correlation of 0.9, visualised in Figure 6.4b. In general lake evaporation contributes about two times more to lake volume losses than lake seepage, as they range from -0.003 to -0.006, and -0.001 and -0.003 km^3 respectively. The volumes of lake evaporation and lake seepage are positively correlated with 0.74, most probably because they both depend on the surface area of the lake. Lake evaporation has a slight positive correlation with catchment precipitation of 0.25. Lake seepage has low to no correlation with other variables. Regarding Figure 6.4a, the observed volume change and the modelled volume change per month are visualised on the bottom right. They experience a moderately positive correlation of 0.6. In total, 57 months show a negative modelled volume change whilst for the observed volume change the is 63 months. The observed volume change displays more months in the second lowest bar of the histogram, whereas the modelled volume more in the lowest bar. Furthermore, the modelled volume change has a long tail with outliers up to 0.04 km^3, where the observed volume does not. The modelled volume therefore gives more extreme volume ranges as opposed to the observed volume.

6.2.4. Component analysis

To gain a comprehensive understanding of how precipitation affects the lake, it's crucial to not only examine the amount of precipitation, as many studies do, but also consider catchment characteristics. Predicting how a specific amount of precipitation of the same magnitude would impact the lake is challenging, but it's easier to predict for varying precipitation amounts. This relates to the first hypothesis, which suggests that the overall modelled change, encompassing lake precipitation, catchment runoff, lake evaporation, and lake seepage, is a superior predictor than catchment precipitation alone. To investigate this, the Pearson correlation of precipitation and modelled volume change is compared with the observed monthly volume change of the lake within the same precipitation bin. Considering all bins, precipitation and observed lake volume change are positively correlated with 0.537, whereas modelled volume change to observed volume change are stronger positively correlated with 0.604.



(a) Histogram of model entities

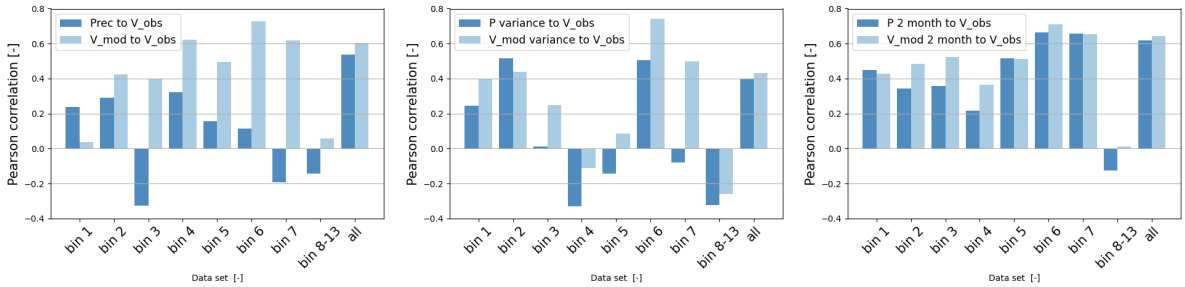
(b) Correlation matrix of model entities

Figure 6.4: Histograms for monthly totals for catchment precipitation and lake precipitation, catchment runoff, lake evaporation, lake seepage, and modelled lake volume change in km^3 per month (a) and Correlation matrix for the model variables contributing to the modelled monthly lake volume (b). Notably, the spread for modelled lake volume is higher with representing more extremes in the bottom right histogram. Correlation coefficients represent possible relations between mainly lake and catchment precipitation, but also moderate positive correlations for modelled and observed lake volume.

Figure 6.5a provides a bar chart of the correlation results for hypothesis one were it is argued that a more comprehensive explanation for the impact of precipitation on the lake is obtained by examining the impact of catchment characteristics rather than only precipitation. In the precipitation bins correlation values range from -0.326 for bin 3 (51.77 to 77.66 mm) to 0.322 for bin 4 (77.66 to 103.55 mm). Within the same range of precipitation it is therefore difficult to predict the influence on the lake volume. Hence, modelled volume change provides a better estimate of the impact on the observed volume, where the correlation ranges from no correlation in bin 1 and 8 to 13, to correlation values of 0.622, and 0.728 for bin 4 (77.66 to 103.55 mm) and bin 6 (129.44 to 155.32 mm) respectively.

Figure 6.5b provides a bar chart of the correlation results for hypothesis two where it is argued that the lake variability depends on the variability and intensity of rainfall, which is assessed by looking at the monthly variance with daily precipitation intensities. Lower precipitation ranges bin 1 (0 - 25.89 mm) and bin 2 (25.89 - 51.77 mm) experience a low positive correlation between observed lake volume change and variance in monthly rainfall. Also the variance in modelled volume change gives a moderate positive correlation. Bin 3 to 5 do not experience correlation, whereas for bin 6 a high positive correlation is found with variance in modelled lake volume and variance in monthly precipitation. Inter monthly variability could be a measure on lake impact for lower and higher precipitation bins with high correlations in the middle graph for bin 1, 2 and 6.

Figure 6.5c provides a bar chart of the correlation results for hypothesis three about the analysis that considers precipitation and catchment memory over two months. Overall, this approach enhances the correlation with observed lake volume to 0.618 and 0.644 for precipitation and modelled lake volume, respectively. Notably, incorporating a two-month precipitation accumulation results in significant improvements in the correlations with higher precipitation ranges bin 5,6, and 7, with regard to a one month period. Moreover, the modelled volume change over two months generally performs equally well or better than the accumulated two-month precipitation.



(a) Hypothesis 1

(b) Hypothesis 2

(c) Hypothesis 3

Figure 6.5: Component analysis for bin 1 to 13 and all months combined. Figure a presents hypothesis 1 where the correlation of precipitation and modelled volume change are compared with the observed volume change. Figure b presents hypothesis 2 where the effect of variance in precipitation and modelled volume is related to observed volume change. Figure c presents results for hypothesis 3 where a two month period is considered. Precipitation ranges above 25 mm show correlations above 0.4 up to 0.7 regarding modelled versus observed volume. Bin 1, 2, 6, and 7 show positive correlation with inter monthly variability. Including previous month in the analysis of lake impact generally increased the prediction of the model over all precipitation ranges.

In addition to analysing all months in the precipitation range it is also important to understand the extremes within each bin, as significant lake decrease or lake increase could be triggered by certain catchment or precipitation behaviour. The lower extremes and higher extremes are compared representing 15% of the sampled months at each end, or 2 samples at each end for bin 6 and 7 to 13, as their sample size is too small. Taking the minimum and maximum observed lake change over all months on the lower and upper 15% results in 21 months at each side. For the lower bound 86% received below average precipitation, whilst for the upper bound 71% received above average precipitation. With respect to modelled volume, 95% of the lower bound observations correspond with a below mean modelled volume change, whilst 76% of the upper bound correspond with an above mean modelled volume. The lower to negative volume changes are all caused by precipitation events lower than 100 mm per month, whilst for the high volume change the amount of precipitation is more ambiguous.

Hypothesis two refers to the monthly variance in precipitation and modelled volume change. In the lower bound, 71% receives below mean variance in precipitation, whilst 52% of the upper bound receives an above mean variance in precipitation. With respect to the variance in modelled volume, 81% of the lower bound observed volume change is modelled with a below average variance in modelled volume, whilst for the upper bound 71% is modelled with an above average variance in volume change. Regarding lower precipitation bins 1, 2 and 3, low variance in precipitation is associated with low to negative volume change. For higher precipitation bins this is not the case. Regarding variance in modelled volume, a low variance in general is associated with a lower volume change. In general, higher observed volume changes are less subject to variance in precipitation or modelled volume.

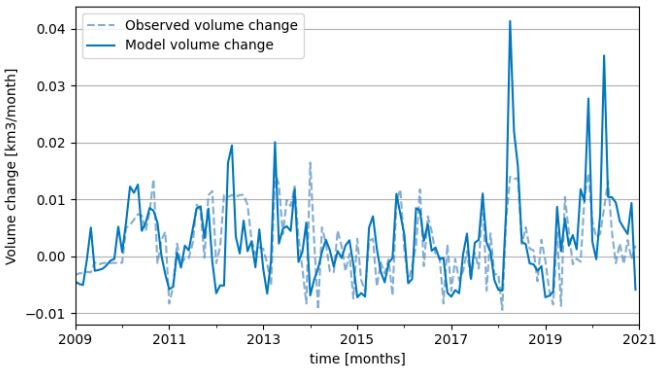
Hypothesis three refers to the memory of the catchment with a two month precipitation period. In 81.0% of the lower bound observed volume change, accumulated precipitation over two months is below mean precipitation. In 76.2% of the higher bound, accumulated precipitation over two months is above mean precipitation. Regarding modelled volume change, 81.0% of the lower bound observed volume change corresponds with a below mean modelled volume change over two months, whilst of the higher bound, 85.7% corresponds with above mean modelled volume change over two months.

6.2.5. Model Performance

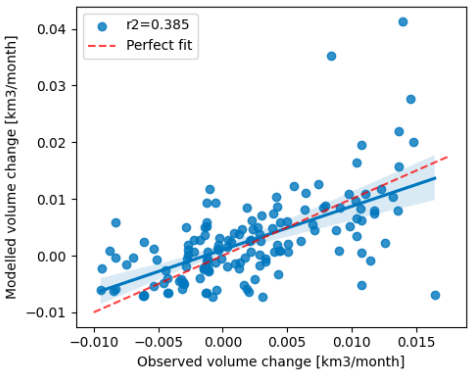
Figure 6.6 demonstrates the performance of the model by presenting the time series for one month, three months, and accumulated volume (a, c, e), along with their corresponding scatter plots (b, d, f) that visualise the observed volume against the modelled volume and corresponding r-squared value. Notable overestimation of three peaks occurring in 2018, 2019, and 2020 are identified. The model can account for 38.5% of the variability with an r-squared value of 0.385. These three peaks are also evident in the scatter plot as the three highest values. The linear regression is slightly less steep when compared to the perfect fit line. In greater detail, the time series visualisation exhibits the same order of magnitude, and hence significant positive and negative volume changes are usually well-matched. Based on Figure 6.6a, accurately representing near-zero volume changes on a monthly resolution can be a challenging task for the model. This is due to the potential for negative observed volume changes to be modelled as positive, and vice versa.

Figure 6.6c, considers a three-month time period. An improvement in the model performance is observed, with an increased ability to reproduce the variability in the lake volume, as reflected in the higher r-squared value of 0.625. The trend of the lake is well-matched, particularly for the period from 2010 to 2017. The extremes are still significantly overestimated in magnitude, with peak values that are roughly two times higher than the observed values, as observed in 2018, 2019, and 2020. Moreover, for the entire year of 2020, the model produces higher volume changes than what was observed. Additionally, the linear regression relation shows a similar trend to the perfect fit, with both lines being approximately parallel to each other. The line for the scatter plot is slightly above the perfect fit line, indicating a slight overestimation of the modelled volume changes. It goes without notice that in 2012 the observations are generally under-predicted by the model.

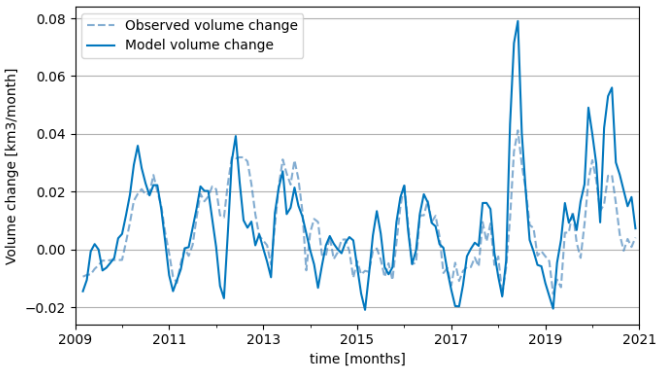
Figure 6.6e visualises the comparison of the accumulated modelled and observed volume. The model shows a r-squared value of 0.868. It has becomes clear that the model over-predicts the lower lake volumes in general from 0 to 0.1 km^3. Subsequently, the model generally under-predicts the lake volumes from 0.15 to 0.25 km^3. As the volume of the lake increases to about 0.3 to 0.35 km^3 the model over-predicts the volume again. We could argue that this over-prediction is caused by these three peaks described earlier in 2018, 2019, and 2020.



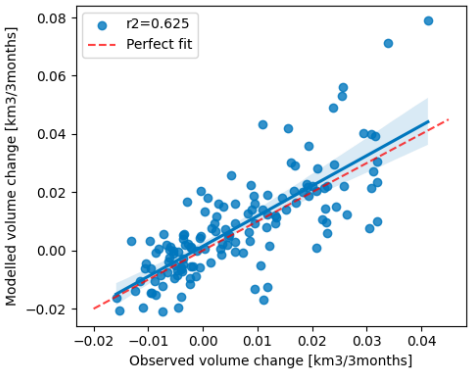
(a) 1-month lake volume change



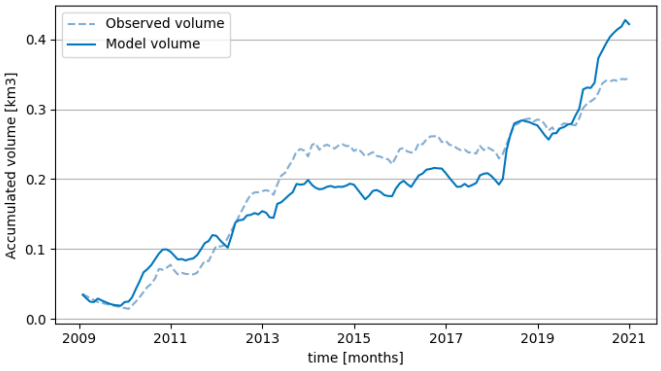
(b)



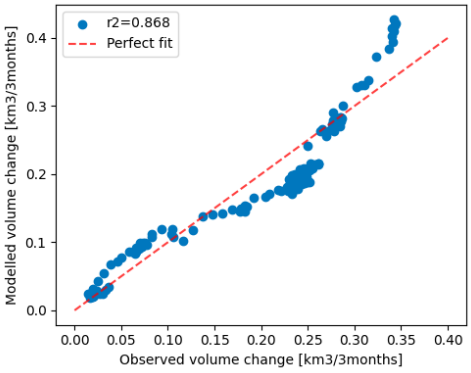
(c) 3-month lake volume change



(d)

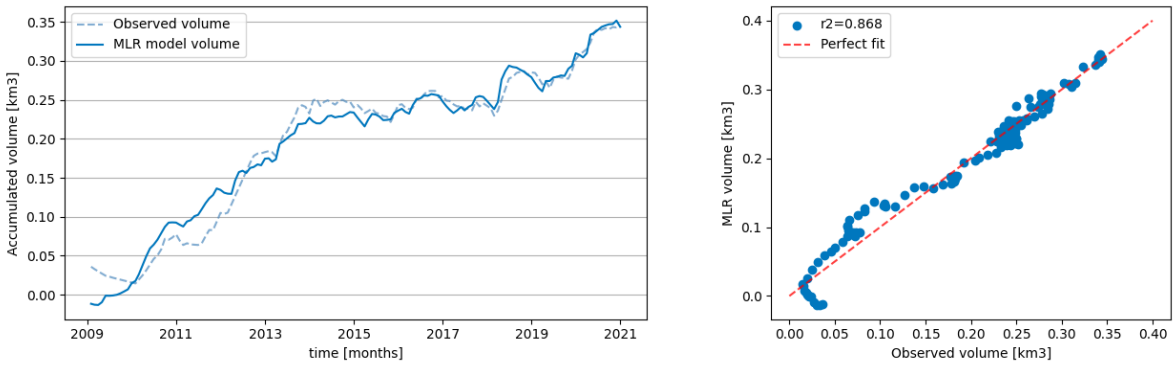


(e) Accumulated lake volume



(f)

Figure 6.6: Model performance with respect to modelled and observed volume where the time series for the one month volume change (a), the three month volume change (b), and the accumulated volume change (c) are given with on the x-axis the time in months, and on the y-axis the observed or modelled volume. Accompanied by their scatter plots in figure b, d, and f, respectively, where the x-axis represents the observed volume and y-axis the modelled volume. As the modelled volumes are taken into account for a longer time period the r^2-statistic improves. There is a good resemblance of lake variability with exception for high extremes in the years 2019, and 2020.



(a) MLR model and observed lake volume

(b) MLR model versus observed volume scatter plot

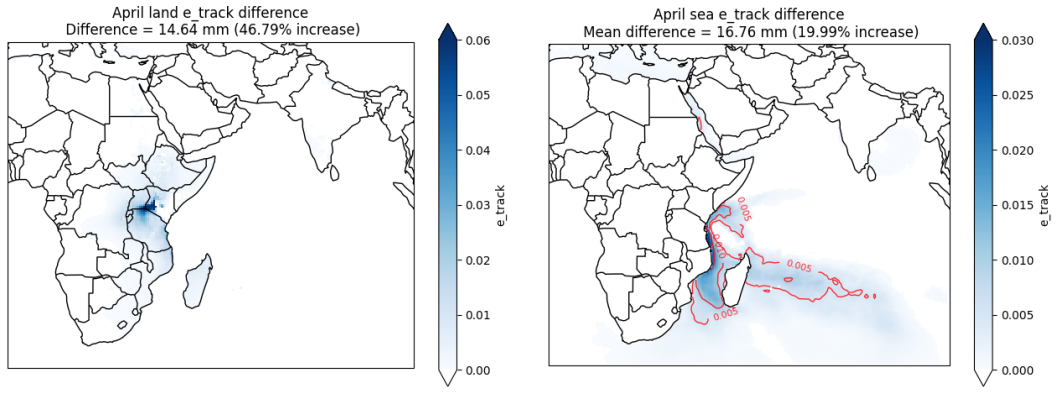
Figure 6.7: MLR model performance with respect to the observed lake volume with time series, with x-axis time in months, and y-axis accumulated lake volume (a) and accompanied scatter plot with x-axis modelled lake volume, and y-axis MLR model predicted volume are visualised (b). Compared to previous Figures there is a general improvement in observed lake volume resemblance with the scatter plot following the perfect fit, even though lower lake volumes are still mismatched in some parts.

Table 6.2: Correlation coefficients for MLR model constant and independent variables; lake precipitation; catchment runoff; lake evaporation; and lake seepage, with corresponding the coefficient, standard error, t-value, and p-value.

Independent variable [km^3]	Coefficient	Std err	t-value	P >| t |
Constant	-0.010	0.006	-1.784	0.077
Lake precipitation	2.152	0.468	4.601	0.000
Catchment runoff	-0.151	0.217	-0.694	0.489
Lake evaporation	-0.837	0.318	-2.635	0.009
Lake seepage	5.061	0.266	19.034	0.000

Regarding lake volume it is of interest to investigate the independent variables that make up the modelled lake volume. A MLR model is created of which the time series result and accompanied scatter plot are visualised in Figure 6.7. The coefficients for the independent variables and the constant are provided in Table 6.2. The 'Std err' column represents the standard error of the coefficient estimate, which measures the variability or uncertainty of the estimate. The 't-value' column shows the t-value, which is the ratio of the estimated coefficient to its standard error, and is used to test the null hypothesis that the true coefficient value is zero. The 'P>|t|' column shows the p-value, which is the probability of observing a t-value as extreme or more extreme than the observed value, assuming the null hypothesis is true. A small p-value indicates that the coefficient is statistically significant, meaning that it is unlikely to have occurred by chance, and that it provides evidence for a relationship between the corresponding independent variable and the dependent variable.

The p-values for lake precipitation, lake evaporation, and lake seepage are all less than 0.05, indicating that these variables are statistically significant predictors of the dependent variable. In contrast, the p-values for the constant and catchment runoff are greater than 0.05, indicating that these variables are not statistically significant predictors of the dependent variable in the presence of the other variables. The coefficient for lake precipitation and lake seepage are 2.152 and 5.061 respectively, which means they are significantly more important than catchment runoff and lake evaporation for predicting the observed lake volume. The MLR model results in a volume prediction that has a high r-squared value of 0.868. Thereby, the mid to high range lake volumes better predicted and match the observed lake volume as can be seen in both the time series as the scatter plot. It should be noted that looking at the time series the MLR model does not resemble the very lower range of lake volume with high accuracy. The same is true for the lake volume range from 0.08 to 0.12 km^3.



(a) (b)

Figure 6.8: April moisture tracking results of continental (a) and sea (b) evaporation representing the difference in evaporation prior and post the change point of 2010. Contour plots are indicating that the moisture contribution from coastal and east of Madagascar increases prior to post change point. Continental evaporation manifests a large absolute increase compared to oceanic evaporation.

6.3. Moisture tracking

The main aim of the moisture tracking is to find relations between changing precipitation patterns and the fluctuations in origin of this precipitation. The overall precipitation patterns are compared with origins of precipitation with regards to the change point observed. The months of April and September are investigated as they have proven to be most significant to precipitation and lake variability. They are assessed on their land versus sea moisture contribution, sea region contributions, and extremely low and high precipitation events.

6.3.1. April moisture tracking analysis

Continental and oceanic evaporation

Moisture tracking results for the month of April are visualised in Figure 6.8 which represents the modelled difference prior and post change point for continental and oceanic evaporation. This section is supported by supplementary material in Appendix C. The tracking visualisations prior and post change point for continental and sea contribution are visualised in Figure C.1 and C.2. In the month of April, continental evaporation contributes about 27.2% on average to the total precipitation in the catchment before 2010 whereas after 2010 this is 31.3%. In absolute terms this is 31.28 mm before the change point compared to 45.92 mm after 2010 an increase of 14.64 mm on average per year which is a surge of 46.8%, whilst the change in average April precipitation is 26.8%. During two wet years after the change point in 2013 and 2018 land contributions can reach above 50% of the total precipitation, whist low precipitation months like 2011 this can be as low as 13%. A moderately positive correlation is observed of 0.648 between the percentage of precipitation originating from the continent and April precipitation in the catchment. This relation intensified after the change point from 0.536 prior to a high positive correlation of 0.754 post change point. On the other hand the percentage of oceanic evaporation is then negatively correlated to April precipitation. Prior 2010 the mean sea moisture contribution is 83.86 mm or 72.83% and 100.62 mm or 68.66 % after 2010. This comes down to a mean increase of 16.76 mm or a 19.99%.

Notably, continental evaporation extends and intensifies more to the North, West and South of the catchment. Furthermore, there is a high concentration (dark blue) around and left of the catchment indicating intensified local moisture contribution. Regarding sea moisture, intensification is observed along the coast between Madagascar and Mozambique. The contour plot of 0.005 and 0.01 extends to the South of Madagascar indicating more evaporation from this region. It is thereby observed that in Figure C.2b the 0.02 contour extends further into the Western Indian Ocean. This can be clearly seen in the difference visualisation as well. Increases are observed before the coast of Tanzania as well as in the Western Indian Ocean just to the right of Madagascar.

Table 6.3: April contributions of six sea regions prior and post change point representing the percentage contribution of total oceanic evaporation and the absolute contribution to the moisture contributing to precipitation. Sea regions 1, 4 and 5, specifically mid coastal, equatorial oceanic, and south equatorial oceanic, have the largest moisture contribution, where sea region 5 increases in magnitude post 2010.

Region	Percent of sea contribution Prior 2010 [%]	Percent of sea contribution Post 2010 [%]	Absolute contribution to e_track to KMD precipitation Prior 2010[mm]	Absolute contribution to e_track to KMD precipitation Post 2010 [mm]
1	24.2	22.9	20.3	23.0
2	8.8	13.9	7.4	13.9
3	6.0	3.7	5.1	3.7
4	20.6	16.1	17.2	16.1
5	26.8	31.6	22.4	31.6
6	8.1	5.6	6.8	5.6

Sea regions analysis

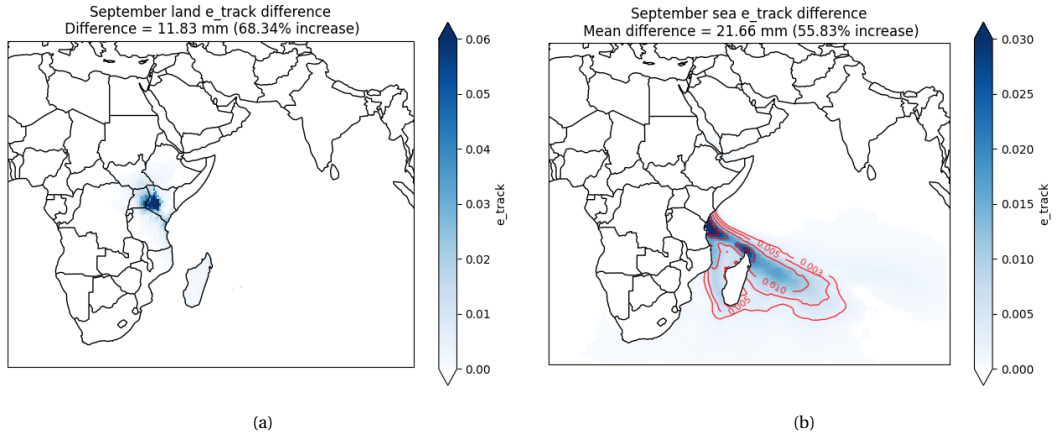
Appendix C displays visualisations of sea regions for the month of April, which include the distribution of six sea regions. Figure C.3 and Figure C.4 demonstrate the total contribution of sea regions before and after 2010, respectively. Figure C.5 illustrates the difference in tracked moisture from each region. As April precipitation is sometimes distributed over the full Indian Ocean extending till right of India. Six sea regions are defined representing; central coastal region (1); the southern coastal region (2); the northern coastal region (3); the Indian ocean off the coast along the equator (4), south of the equator and right of Madagascar (5), and north of the equator around India (6). Table 6.3 summarises the outcomes, presenting the percentage and absolute contribution of moisture to oceanic evaporation from each of the six regions before and after 2010.

Sea region 1, 4 and 5, are the most significant when it comes to moisture contribution. After 2010, the contribution of sea regions 1, 3, 4, and 6 to oceanic evaporation has decreased, while the contribution of regions 2 and 5 has increased from 8.8% to 13.9% and 26.8% to 31.6% respectively. This indicates that the importance of regions near and north of the equator has declined, while regions south of the equator have become more significant. Despite an overall increase in precipitation in April, regions 3, 4, and 5 have decreased in absolute terms, whereas region 1 has increased. Meanwhile, sea regions 2 and 5 have experienced a significant absolute increase, with their average evaporation levels rising from 7.4 mm to 13.9 mm and 22.4 mm to 31.6 mm, respectively, representing an absolute increase of 87.8% and 41.1%.

Table C.1 provides the relation expressed in correlation between April precipitation and sea region contribution as a percentage of oceanic evaporation, and between April precipitation and absolute evaporative contribution of the sea region. It has previously been mentioned that there is a moderate and high positive correlation for the percentage of evaporation originating from land compared to April precipitation, and thus negative for oceanic evaporation. Percentage wise there is no profound correlation between any of the sea regions and April precipitation over all years, however sea region 1 has the highest low correlation of -0.404. When we compare prior and post 2010, the negative correlation for sea region one changes from low to moderate negative correlation of -0.338 to -0.545. Comparing the absolute contributions from sea regions, region 1 has a high correlation with April precipitation of 0.863, and sea region 4 and 5 a moderate positive correlation of 0.654 and 0.664 respectively. Correlations for sea region 2, 3, and 6 are low positive. Interestingly, comparing prior and post 2010, sea regions 3 and 6 move from low positive correlation to moderate and high positive correlation respectively with 0.249 to 0.588 and 0.229 to 0.709. Other region correlation remain roughly equal. This could indicate that more moisture from the northern regions serves as an indicator for wet periods.

Table 6.4: Moisture source sea region contribution to three driest months April (1996, 2011, 2021) and three wettest months April (1988, 2018, 2020) including the percent contribution compared to the sea regions total and absolute contribution. Region 5 is most significant in its change to relative and absolute moisture contribution.

region	Contribution to dry extremes [%]	Contribution to dry extremes [mm]	Contribution to wet extremes [%]	Contribution to wet extremes [mm]
1	26.5	8.6	23.1	43.6
2	10.6	3.4	5.9	11.1
3	6.6	2.1	3.7	7.1
4	21.2	6.9	17.8	33.5
5	24.9	8.1	36.9	69.7
6	4.9	1.6	4.8	9.1



(a) (b)

Figure 6.9: September moisture tracking results of continental (a) and sea (b) evaporation representing the difference in evaporation prior and post the change point of 2010. Intensification of moisture contribution from the coastal areas as well as northeast of Madagascar with an increase also manifesting between the Mozambique and Madagascar. There is overall a major increase in continental evaporation extending more to the west of the catchment.

Low and high precipitation extremes

As part of understanding dynamics for wet and dry April months the three low and high extreme precipitation months are investigated for their moisture sources and sea region contribution of which the results are described in Table 6.4. The three driest years include 1996, 2011, and 2021. The three wettest April include 1988, 2018 and 2020. On average the three driest years have sea contribution of 32.45 mm opposed to 188.72 mm in the three wettest years. Relatively, during the driest months, moisture from sea represents 82.0% of the total tracked moisture compared to 66.6% in the wettest months. In the driest and wettest month the relative contribution of sea region 5 increased massively from 24.9% to 36.9%, whilst all other regions decreased in relative contribution most notably the mid coastal region 1 from 26.5% to 23.1% and the mid Indian Ocean region 4 from 21.2% to 17.8%. One can argue that the coastal region and Indian ocean around the equator become relatively less important during wet months and more important during dry months. The other coastal regions 2 and 3 have also decreased in relative importance from 10.6% to 5.9% and 6.6% to 3.7%.

6.3.2. September moisture tracking analysis

Continental and oceanic evaporation

Moisture tracking results for the month of September are visualised in Figure 6.9 which represents the modelled difference prior and post change point for continental and oceanic evaporation. This section is supported by supplementary material in Appendix D. Moisture tracking results for prior and post 2010 can be found in Figure D.1 and D.2. In the month of September 30.85% of the precipitation can be attributed to evaporation from land equalling 17.31 mm on average before 2010. After 2010 the contribution of continental evaporation increased to 32.52% or 29.14 mm on average per month September. This represents a 11.83 mm increase or 68.3% relatively. The largest difference is observed around and left of the catchment area.

Table 6.5: September contributions of five sea regions prior and post change point representing the relative contribution of total oceanic evaporation and the absolute contribution to the moisture contributing to precipitation. Region 4, to the east of Madagascar contributes the most to oceanic evaporation and increases significantly in absolute terms after 2010.

Region	Percent of sea contribution Prior 2010 [%]	Percent of sea contribution Post 2010 [%]	Absolute contribution to e_track to KMD precipitation pre 2010[mm]	Absolute contribution to e_track to KMD precipitation Post 2010 [mm]
1	14.0	14.3	5.4	8.7
2	6.8	8.1	2.6	4.9
3	16.9	17.0	6.5	10.3
4	44.1	45.7	17.1	27.6
5	18.0	14.6	7.0	8.8

The contribution of continental evaporation can be over 40% in the years of 2001 and 2011 and be as low as 15% in the years 1997 and 2002. The percent of land contribution and September precipitation represents a low positive correlation of 0.486 over all years. The correlation however is reduced after the change point from a moderate positive correlation of 0.622 to no correlation of 0.013. Alternatively, oceanic evaporation contributes to 69.15% or 38.8 mm prior 2010 and 67.5% or 60.47 mm post 2010, representing a 21.66 mm difference and 55.8% increase.

The tracking visualisations prior and post change point for continental and sea contribution are visualised in Figure D.1 and D.2. Continental evaporation became more intense around the left of the catchment region and there is a high concentration of moisture production around the catchment that contributes to the precipitation. Regarding sea moisture, intensification is observed along the coast between Madagascar and Mozambique. The contour plot of 0.005 and 0.01 extends to the South of Madagascar indicating more evaporation from this region. It is thereby observed that in comparing prior and post change point moisture source, the 0.005 and 0.010 contour stays reasonably in place whereas 0.020 and 0.030 contours extend further east of Madagascar. This is clearly observed in the difference visualisation. Large increases are also observed near the coast of the Kenyan-Tanzanian border extending to the north of Madagascar.

Sea regions analysis
Appendix D displays visualisations of sea regions for the month of September, which include the distribution of six sea regions. Figure D.3 and Figure D.4 demonstrate the total contribution of sea regions before and after 2010, respectively. Figure D.5 illustrates the difference in tracked moisture from each region. Given the distribution of September precipitation the sea is divided into five regions representing the central coastal region (1), the southern coastal region (2), northern Madagascar (3), eastern Madagascar and western Indian Ocean (4), and further east on the Indian ocean (5). Table 6.5 summarises the outcomes, presenting the percentage and absolute contribution of moisture to oceanic evaporation from each of the six regions before and after the change point.

It becomes evident that the moisture source from sea region 4 contributes most of the sea moisture to precipitation in the catchment with 44.1% of total sea moisture. Thereby, sea region 1, 3, and 5 contribute about the same percentage of sea moisture before 2010. After 2010 these distributions have slightly shifted with most notably the reduction in percentage contribution of region 5 from 18.0% to 14.6% and the increases of sea region 2 and 4 from 6.8% to 8.1% and 44.1% to 45.7% respectively. As September precipitation generally increases after the change point, the absolute contribution for all sea regions increases as well. The absolute increase is most profound in sea region 4 from 17.1 mm to 27.5 mm representing a 61.4% increase, where percentage wise region 1, 2, 3, and 5, increased with 61.1%, 88.5%, 58.5%, and 25.7%.

Table D.1 provides the relation expressed in correlation between September precipitation and sea region contribution as a percentage of oceanic evaporation, and between September precipitation and absolute evaporative contribution of the sea region. Percentage wise there is no profound correlation between any of the sea regions and September precipitation over all years. It is only after 2010 that sea region 3 exhibits a low negative correlation with precipitation. Furthermore, there are no drastic changes prior and post 2010 to the correlation of absolute moisture contribution and September precipitation. It is remarked that regions 1, 3 and 4 give a very high positive correlation whilst regions 2 and 5 give a moderately positive correlation.

Table 6.6: Moisture source sea region contribution to three driest months September (1996, 2011, 2021) and three wettest months September (1988, 2018, 2020) including the relative contribution compared to the sea regions total and absolute contribution

region	Contribution to dry extremes [%]	Contribution to dry extremes [mm]	Contribution to wet extremes [%]	Contribution to wet extremes [mm]
1	12.2	2.2	13.7	12.7
2	5.9	1.0	7.6	7.1
3	16.7	3.0	16.1	14.9
4	45.8	8.2	44.1	40.9
5	19.2	3.4	17.9	16.6

Low and high precipitation extremes
The three lowest and highest precipitation months for September are investigated as well for their moisture sources and sea region contributions of which the results are described in Table 6.6. The three driest years include 1997, 2002, and 2006. The three wettest April include 2013, 2017 and 2021. On average the three driest years have sea contribution of 17.82 mm opposed to 92.73 mm in the three wettest years. Relatively, during the driest months, moisture from sea represents 81.5% of the total tracked moisture compared to 69.6% in the wettest months. Sea region contribution as a percentage of total sea contribution remained relatively stable between wet and dry extremes.

6.4. Sea surface temperatures

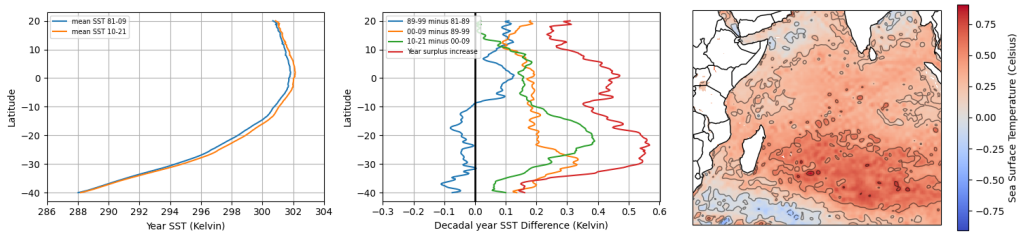
SST is an important metric for the rate of evaporation from the sea surface. Therefore the general changes in SST are observed first. Thereafter, the sea regions defined in moisture tracking are investigated for their variability in SST and what that means for catchment precipitation. Accordingly, the extremely low and high precipitation events, also defined from moisture tracking, are compared with each other regarding SST in the defined sea regions for April and September.

6.4.1. General SST variability

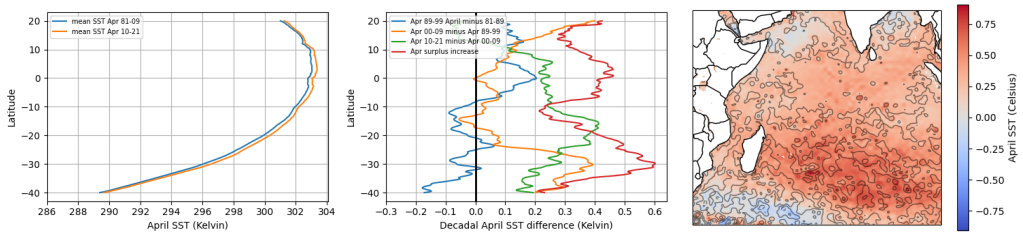
Therefore, the full year period, April, and September SST are investigated along with their relation to catchment precipitation. The results are visualised in Figure 6.10. Furthermore, the sea regions from April and September are assessed on their variability in SST. Looking at the left plots, the mean SST for the full year reveals the highest SST around the equator as well as for the month of September. Higher SST in April are observed around 5° latitude. Temperatures decreasing towards the -40° latitude to about 288 Kelvin for the yearly average, and to about 290 Kelvin and 286 Kelvin for April and September respectively. Temperatures between 20° and 0° latitude remain roughly the same.

Regarding decadal changes in the middle plots, we observe a slight increase around and above the equator for all the time periods regarding the change between the first two decades from 1981 to 1989 and 1990 to 1999 (blue). Below the equator temperatures remained the same on average except for September where a decrease in SST is observed of up to -0.25 Kelvin. The difference from the second to third decade (orange) gives us different patterns. For the full year period an increase in SST is observed over the full latitude range of around 0.2 Kelvin. Around -30° latitude an increase of 0.3 Kelvin is shown. Another pattern is observed for the month of April where mainly temperatures around 20° and -30° latitude are observed of about 0.3 Kelvin. Referring to September SST an increase of 0.3 Kelvin is observed over the full latitude range except for above 10° and below -30° latitude. Regarding the third to last decade (green) the main increase for the full year is observed between -10° to -30° latitude of 0.2 to 0.4 Kelvin increase.

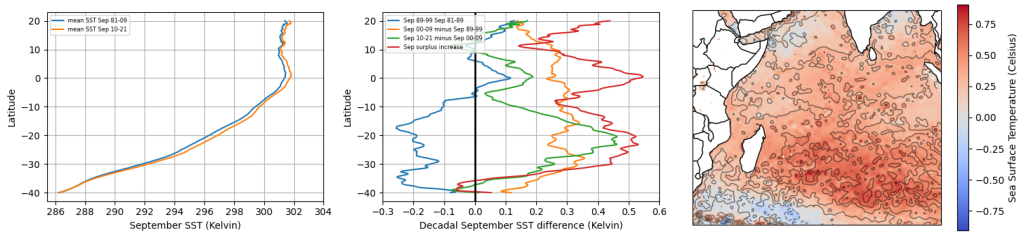
Regarding the month of September we see the increase in SST shift more towards the equator with up to 0.4 Kelvin increase around the -20° latitude ranges. Over the full latitude an increase of 0.2 Kelvin is observed. The month of September provides observations with a large increase between again -10° to -30° latitude of up to 0.45 Kelvin, whilst regions outside of this range remain the same increase about 0.1 Kelvin. Overall comparing the accumulated changes in SST the increased SST is most profound around the -20° to -30° latitude range all time periods. September SST however also reveals a surge around the equator. The right plots in Figure 6.10 provide a general overview of the changes in mean SST prior and post 2010.



(a) Full year SST



(b) April SST



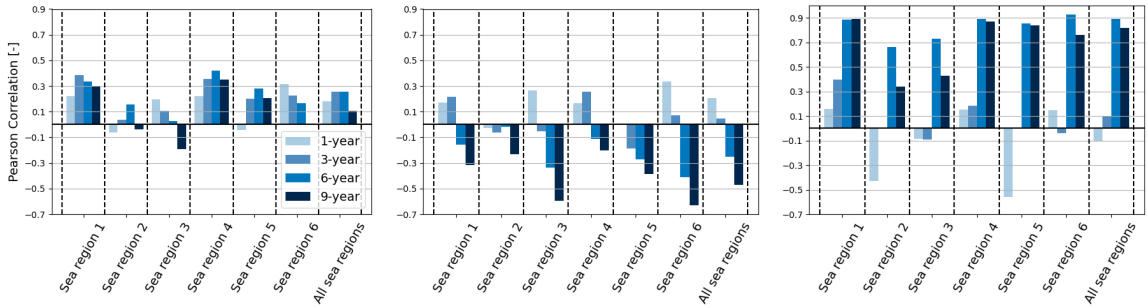
(c) September SST

Figure 6.10: Visualisations of the SST; as change over the latitude range for means from 1981 to 2009 and 2010 to 2021 (left); as a difference between 2 consecutive decades (middle); and as the difference between the means from 1981 to 2009 and 2010 to 2021 (right; with the full year SST (a), April SST (b), and September SST (c). April experiences the highest temperatures during the year whilst September for example experiences lower SST up to 4 Kelvin at 30° latitude (left). SST have been rising faster with the decade, and have most profoundly risen around -20 to -30° latitude for April and September. The same region where most of the moist increased production is found resulting from the moisture source analysis.

6.4.2. Regional SST variability

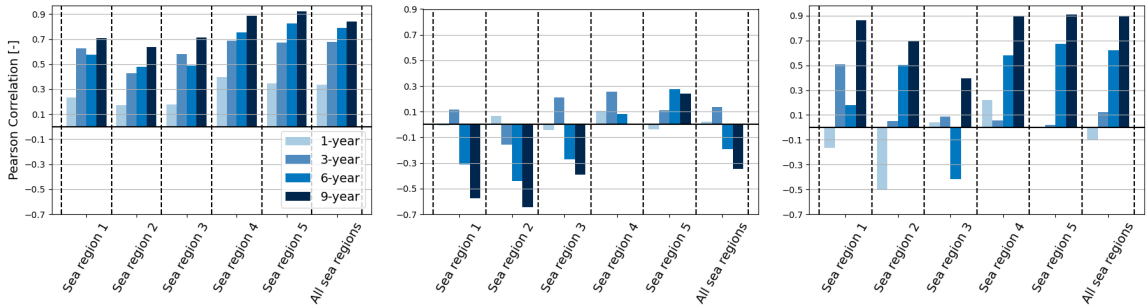
The correlation between April precipitation and sea region SST, as well as the mean SST across all regions, were analysed over one, three, six, and nine year periods to capture yearly and multi-yearly trends. This section is supported by supplementary material in Appendix E. The results are presented in Figure 6.11 and Table E.1, which shows that April precipitation did not exhibit strong correlations to any sea region across the entire 1981-2021 time period (left). However, when the time period was divided into prior- and post-2010, negative correlations were found between April precipitation and SST for the prior-2010 period (middle), while positive correlations were found for the post-2010 period (right).

Over time, a moderate trend correlation for the three-year mean was observed for region 1, particularly around 0.4. Positive correlations between SST and April precipitation were found for the six-year means, indicating that higher SST were associated with more precipitation on average in April for both individual and all sea regions. Although the correlation for the full time series was 0.255 and prior to 2010 was -0.253, a strong positive correlation of 0.819 was observed for the post-2010 period, suggesting a clear relationship between SST and April precipitation over a six-year time frame.



(a) 1981-2021 April SST (b) 1981-2009 April SST (c) 2009-2021 April SST

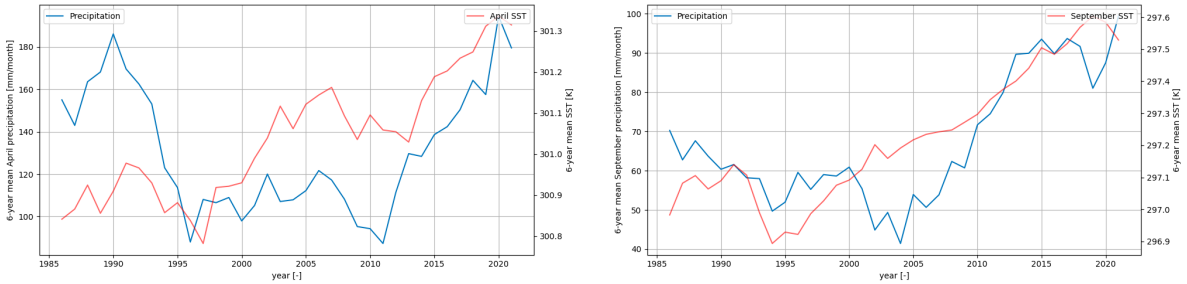
Figure 6.11: Correlation coefficients for April precipitation to sea region SST for all years together (a), only SST years prior 2010 (b), and SST years post 2010 (c). Regarding April precipitation there is positive correlation for sea regions 1 and 4, the coastal equatorial and oceanic equatorial region. The correlation between precipitation and SST becomes more profound when taking into account multiple year means. Precipitation after 2010 reveals very high correlation with SST over the 6 and 9 year mean. On a one year timeframe, low to no correlations are found.



(a) 1981-2021 September SST (b) 1981-2009 September SST (c) 2009-2021 September SST

Figure 6.12: Correlation coefficients for September precipitation to sea region SST for all years together (a), only SST years prior 2010 (b), and SST years post 2010 (c). September precipitation experiences low positive correlation for all sea regions over a one year mean. This relation becomes stronger already after taking into account the three year means for SST and precipitation. Negative correlations are found before 2010 over a multi year period, whereas with SST increasing means experience high correlations with post 2010 precipitation.

The relationship between September precipitation and SST is examined using the same approach as for April, and the findings are presented in Figure 6.12 and Table E.2. The results show a moderate positive correlation for sea region 4 in all years (left), while other regions display low or no correlation. However, for a three-year time period, the correlation becomes stronger, ranging from highly correlated for sea regions 1, 4, and 5 to moderately positive for sea regions 2 and 3. The six-year and nine-year means generally improve the correlation for all sea regions. In contrast, prior to 2010, negative to low positive correlations are observed across all regions and time periods. For the post-2010 period, high positive correlations are noted for the six-year and nine-year means, except for sea regions 1 and 3 on a six-year mean. Figure 6.13 provides the trend for all sea regions combined over a six-year period for the month of April and September.



(a) April 6-year mean

(b) September 6-year mean

Figure 6.13: The 6-year trend in April (a) and September (b) showing the relation between SST and precipitation. One should pay attention to the increasing mean temperatures of the sea regions trending with precipitation, where a rise is most profound from about 2010 for April and already from 2005 in September.

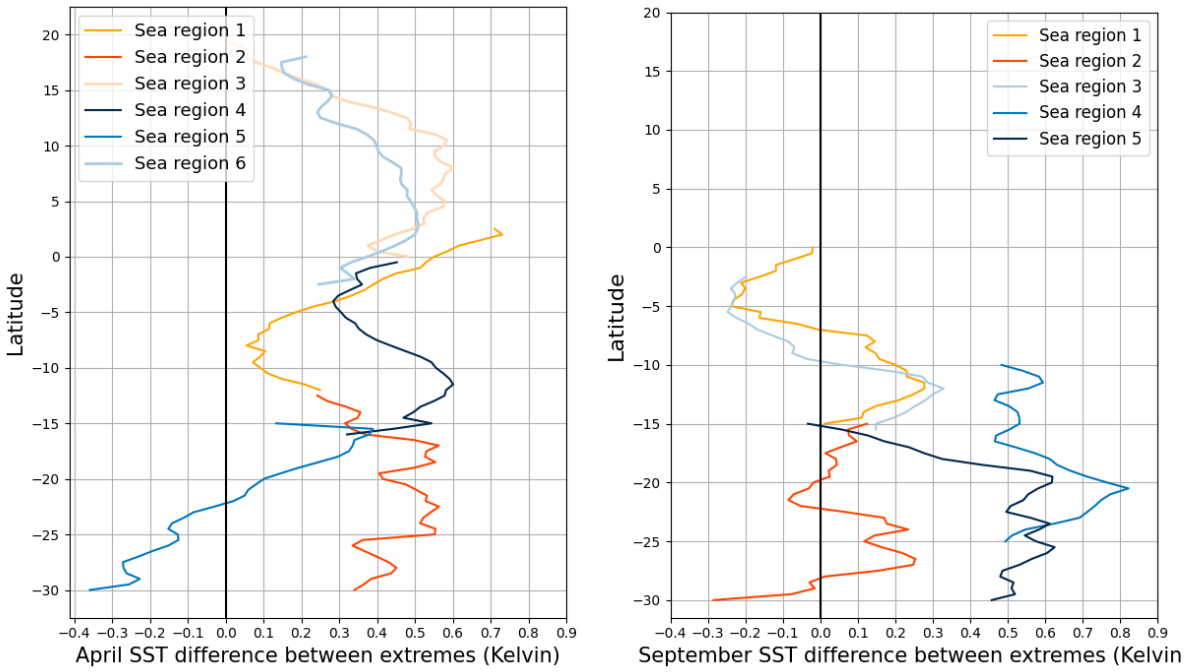
SST in sea regions during extreme precipitation

In order to strengthen the SST analysis, the three low and high extreme precipitation events from moisture tracking are also examined based on their corresponding sea region SST. The findings for the months of April and September are presented in Figure 6.14a and 6.14b, respectively. The SST differences are illustrated across the latitude range of six sea regions for the month of April. Regions 1, 2, and 3 represent the coastal regions, while regions 4, 5, and 6 are the oceanic regions located around, south, and north of the equator, respectively. The results indicate that higher temperatures are primarily observed for wet April months above the equator in regions 3 and 6, with differences ranging from 0 to 0.6 Kelvin. From the equator to the south, sea regions 2 and 4 display consistent differences ranging from 0.3 to 0.6 Kelvin. Sea region 1, which is the middle coastal region, exhibits a large variance, ranging from 0.1 at around -5° to -10° latitude to a maximum of 0.7° around the equator. Sea region 5, which is the oceanic region south of the equator, shows a difference of 0.4 Kelvin around -15° latitude to -0.3 Kelvin at about -30° latitude. Overall, the analysis shows that SST is higher for wet April months, with higher temperatures observed along the full coast and between 10° to -15° latitude at sea.

The month of September displays source regions mainly below the equator. Sea regions 1 and 2 are coastal regions, sea region 3 lies north of Madagascar, region 4 is the region east of Madagascar, and region 5 is further east on the Indian Ocean. Interestingly, SST differences are 0 to -0.2 Kelvin from about 0° to -5° latitude for the coastal and northern Madagascar regions, from which they experience larger differences to 0.3 Kelvin SST to about -10° latitude. The lower coastal region 2 shows very little difference in SST fluctuating between -0.1 and 0.2 Kelvin. Moisture sources in sea regions 4 and 5 were previously recognised to be significant to precipitation changes in September. Regarding SST they experience the largest difference as well. Peak differences are observed for sea region 4 constantly above 0.5 Kelvin with highs up to 0.8 Kelvin around the -20° latitude. Sea region 5 further up the ocean experiences differences above 0.5 Kelvin from a -20° latitude onwards.

6.5. Global Climate variability

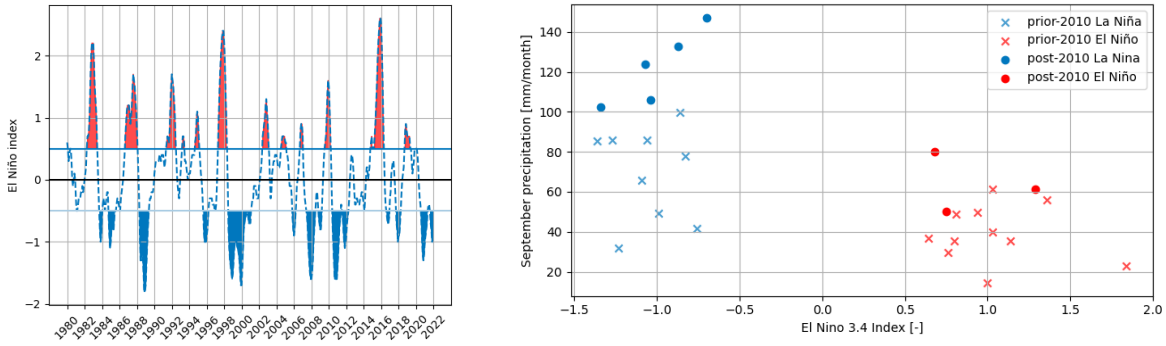
Apart from SST that are also directly influenced by global climate variability, the relation between these global climate variability and precipitation is investigated for the IOD and El Niño. Strong IOD events are related to period of high precipitation in the short rains season October, November, December experiencing a moderately high correlation of 0.688. This relation became stronger comparing prior- and post -2010 from 0.568 to 0.842. The flooding of 1997 and 2019 are tied to this phenomenon. Including September precipitation in the analysis gives an correlation of 0.539 overall, and an improvement from 0.388 to 0.732 comparing prior and post change point. The month September to IOD September experiences no to moderate negative correlation of -0.464 prior 2010 and -0.278 post 2010. No evidence of correlations is found for the months of April and IOD with 0.229 and -0.373 correlation.



(a) (b)

Figure 6.14: SST differences in sea regions defined in moisture tracking for extreme precipitation years for April (a) and September (b) with on the x-axis the difference in mean extreme SST per region and on the y-axis the latitude range for each sea region. For the month of April this regards the difference between SST for high extreme precipitation years 1988, 2018 and 2020 and low extreme precipitation years 1996, 2011, and 2021. For September this means SST difference in high extreme precipitation years 2013, 2017 and 2021, and low extreme years 1997, 2002, and 2006. Orange colours indicate coastal sea regions. Blue colours indicate oceanic regions. In September SST differences of up to 0.8 degrees over the latitude on average are observed. For September it mainly regars the full latitude range that experiences higher temperatures.

Figure 6.15a represents a time series of the second global climate indicator namely El Niño. Over the course of 42 years there have been 14 El Niño years and 14 La Niña years. Prior 2010 there have been 11 El Niño years and 9 La Niña years opposed to 3 and 5 post-2010. Figure 6.15b represents a visualisation of the El Niño and La Niña years against September precipitation. Overall there is a high negative correlation between the index and September precipitation of -0.620. This correlation became more profound comparing prior- and post-2010 where the correlation increased to -0.640 to a very high correlation of -0.800. La Niña events are associated with high September precipitation whilst El Niño events with low September precipitation. Prior 2010 the average precipitation of a La Niña year was 69.18 mm and post 2010 122.46 mm, whilst the average El Niño year prior 2010 was 39.16 mm compared to 63.82 mm post 2010, representing an increase of 77% and 63% respectively. September precipitation is highly negatively correlated with El Niño whilst the short rains season from October to December is strongly correlated with an El Niño.



(a) (b)

Figure 6.15: A time series visualisations of El Niño and La Niña between 1980 and 2021 (a) with the El Niño and La Niña years scattered against September precipitation prior and post 2010 (b). El Niño periods are defined in red and measure above 0.5 degrees above the mean for 3 consecutive months, whereas La Niña records three consecutive months with 0.5 degrees below the mean. Red points are indicated as El Niña years where blue points are indicated as La Niña. The crosses indicate prior-2010 and post-2010 are indicated by dots. There is a strong correlation between September precipitation and El Niño, resulting in wet September months during El Niño which have become more intense after 2010.

7

Discussion

Hydrological and meteorological research are of great value to comprehend the complex Eastern Africa system, and more specifically the Kenyan Rift Valley, as impacts of climate change (Black et al., 2003; Black, 2005) and indisputable changes in the hydrology of the lakes (Dyer and Washington, 2021; Gichuru and Waithaka, 2016) have become more apparent in recent decades. This study has been addressing variability in precipitation and lake dynamics for Lake Nakuru using a water balance model and moisture tracking. This was used to find patterns in changing climatology tied to the overflowing of the lake using observational precipitation data and bathymetry studies to obtain a more sophisticated understanding of the system. One must ascribe caution to the specific meaning of water balance model parameters as these are more conceptual than physical features, since they are not observed or studied on site. According to the literature study performed for this research, this is the first study to have employed moisture tracking with WAM2layers on a catchment scale, and the resulting findings are both adequate and insightful. The method used to explore moisture production and sea surface temperature (SST) variability in sea regions adds on our current understanding of the impact that land, coastal, and further oceanic evaporation has on precipitation in East Africa (Yang et al., 2015; Van der Ent and Savenije, 2013; Keys et al., 2022).

Precipitation variability

Years of low precipitation were interspersed with years of high precipitation over the lake Nakuru catchment. As of 2010 a change point has been identified in weather station precipitation data and catchment wide precipitation, in agreement with earlier findings for precipitation in the Rift Valley region. Earlier studies have identified as a breakpoint in 2009 (Herrnegger et al., 2021), whereas this study rather employs the terminology of a change point, which is the year after a point of change. Mean precipitation over lake and catchment area were investigated using the Thiessen method although this study does not take into account elevation difference. It is expected not to have great impact on this water balance model or the model outcomes as this model uses catchment wide precipitation. The Thiessen precipitation mean, and station precipitation mean data sets have been compared after which two important, but not very significant, differences are observed; first, the difference in the mean precipitation between 1981 and 2021 is about 20 mm; second, there is a dampening mainly seen in the higher precipitation years. Not sure is the impact on daily precipitation intensities on a catchment scale when implementing another method for determining the mean catchment precipitation. Furthermore, weather station areas as a contribution to lake precipitation is assumed to be stable over time. While lake precipitation is proven to be an important part of explaining the lake variability one might want to consider dynamic weather station contributions to lake precipitation for the model. Even so, a more detailed water balance model should consider modelling individual catchment regions contributing to the lake taking into account residence times, elevation, and land use changes. While this study's focus lies with the variability of precipitation and underlying drivers, taking into account previous catchment entities for the water balance model would benefit the understanding of how local changes, geography, and different catchment parts contribute to the dynamics of lake Nakuru.

The bimodal precipitation annual cycle for East Africa is disputable for the catchment itself. The long rains season from March to May is very profound whereas the short rains season from October to December does not show a major second jump in precipitation. On the other hand, lengthening of the short rains season (Cook and Vizy, 2013), as a result of climate change is an argument for the change point observed for the month of September for which precipitation increased from the third lowest to the third highest precipitation month surging 64% on average presumably extending the short rains season. Due to time and computation power constraints, with April being the second largest contributor to the absolute increase of precipitation, the two months April and September were identified as significant contributors the the change in precipitation and overflowing of the lake to be further analysed with moisture tracking. The prolonged dry periods and wet periods observed for the CHIRPS dataset (Kimaru et al., 2019) is supported by the observational data from the Kenyan Meteorological Department as SPI indicators before the change point mainly present dry to extremely dry periods whereas after the change point very wet and extremely wet periods were observed.

Implications water balance model decisions

It should be noted that because of limited data in lake observations and evaporation data the model employed in this research encompasses the time period from 2009 to 2020. Data prior 2009 revealed major data gaps of up to five years making it not possible to adequately model the dynamics of the lake. However a more detailed model was developed compared to previous studies (Herrnegger et al., 2021), specifically related to actual evaporation and precipitation data. The model uses observational data from fourteen weather stations for precipitation on a catchment and lake precipitation scale. Opposed to CHIRPS data which is primarily based on satellite data calibrated with observational data with a resolution of 0.5°. Future models could be made more sophisticated to model the variability of the lake with the 14 specific catchment compartments for which weather station data is available. The actual evaporation obtained through water balance modelling of the catchment is considered more reliable than the method employed by Herrnegger et al. (2021) where the actual evaporation was calculated through multiplying actual evaporation from ERA5 with the ratio between ERA5 and CHIRPS precipitation, under the assumption that the overestimation of precipitation by ERA5 compared to CHIRPS reflect directly to actual evapotranspiration under the same conditions. For this study, through the ratio of catchment storage compared to the maximum catchment storage, hence water availability, and reference evapotranspiration from WaPOR the actual evapotranspiration for the catchment has been obtained. This is important to accurately determine the impact precipitation has on the catchment and thus on the lake, because evaporation directly determines the net amount of precipitation available in the catchment.

Regarding the model parameters included in the model calibration caution should be taken with interpreting the values of the model parameters like runoff coefficient and catchment storage. The optimal values are obtained through calibration and represent conceptual rather than physical values and can differ heavily between different catchment areas and over time. Additionally, percolation from the catchment storage to groundwater is assumed to not end up in the lake as percolation is assumed to recharge deeper groundwater levels just as river flows that are mostly lost to recharge of groundwater before reaching the lake, for which most of the rivers become influent (Kanda and Suwai, 2013). Deep groundwater is assumed to flow northwards towards the lower laying parts of the Rift Valley where Lake Bogoria is situated (Montcoudiol et al., 2019). The catchment runoff is then assumed to take into consideration runoff flow and shallow subsurface flow towards the lake. The impact of groundwater flow is not known for lake Nakuru specifically. The assumptions above are mainly based on the literature study which suggests that underground inflow is minimum. The hydrology in the catchment is complex and future studies should therefore specifically focus on groundwater flows in the catchment and contribution to the lake. This could be realised with isotopic tracer hydrology which has been applied in several studies in East Africa (Levin et al., 2009; Kebede et al., 2009; Montcoudiol et al., 2019; Balagizi et al., 2018). Regarding model performance, the model overestimates parts of 2010 to 2012, most caused by a higher modelled volume for the period of 2010 only. As variability and magnitude change are quite stable between 2013 and 2018, the model on the one month and three month time frame indicates a significant underestimation for the year 2012 resulting in a continuing underestimation for the remaining time period. Large peaks in modelled volume in 2018, 2019, and 2020 cause the model to overestimate the lake volume by the end of 2020. Multiple Linear Regression has been applied to overcome the under-and overestimation of the model to see which model variables would be most significant for the variability in the model. The high performance indicator for the three month and accumulated modelled lake

volume indicates that the model is able to follow the dynamics of the lake accurately and is able to predict lake volumes in the same order of magnitude overall.

Moisture tracking and climate drivers

Researchers are increasingly interested in using moisture tracking to explore hydro-meteorological variability as it allows them to understand the underlying dynamics of climatological changes in a particular region. While moisture tracking studies have been implemented at the country scale, this study employs moisture tracking to examine precipitation on a lake catchment scale, making it one of the first studies to use WAM2layers for this purpose. Keys et al. (2022) found that 85% of Kenya's precipitation consists of moisture from oceanic evaporation and 15% from continental evaporation. Catchment specific moisture tracking for this study found that continental evaporation can have a significant larger contribution to precipitation of about 27% to 31% in the month of April and 31% to 33% for September precipitation. The relation between continental moisture and precipitation became more profound when the three driest and wettest years were compared. Extreme land contributions are found for the month of April with over 50% continental moisture contribution and over 40% for September precipitation, mostly for wet months. This study found that wet periods are tied to higher land evaporation for both April and September precipitation. On top of that, coastal moisture contribution is lower for wet periods compared to oceanic moisture contribution.

Overall, this study supports earlier findings related to land, coastal, and oceanic moisture production for dry and wet periods (Keys et al., 2022; Findell et al., 2019) now also on catchment scale. Complex geological and hydrological conditions of the Rift Valley, this study could result in different findings for other catchments. It is promising though that the relation find for Kenya is also found back in the lake Nakuru catchment regarding land, coastal, and oceanic moisture production. A positive correlation for April and September for continental contribution to precipitation is observed and with increasing relative contributions for precipitation comparing prior and post change point. It should be noted that this study specifically focuses on the months of April and September as these months are identified as the most significant contributors to the changing precipitation patterns in the catchment. It would be interesting to investigate all months and see if yearly patterns have changed. Furthermore, months that have seen no or negative change in precipitation are also interesting to investigate further so find similarities and differences between wetter months to be able to identify clearer causes.

Yearly SST over the Indian ocean and precipitation over the catchment both experience their peak in April. With a warming world Findell et al. (2019) argued that the ratio between oceanic and continental evaporation would become larger making oceanic evaporation more important. This study however found that continental evaporation became relatively more important compared to oceanic evaporation. Both April and September experience a gradient in SST with lower temperatures in the southern Indian Ocean. In April SST ranges from 304 Kelvin around the equator to 292 Kelvin south of Madagascar. SST in September range from 302 Kelvin around the equator to 290 Kelvin and lower south of Madagascar. In the area along the latitudes of Madagascar also the highest increments in SST are found. Comparing wet and dry periods SST in some areas to the right of Madagascar can differ almost 1 Kelvin. Moreover, these regions of the sea have also been identified as the most significant moisture contributors to precipitation in the catchment. Furthermore, these regions show the highest increase in produced moisture. Over time, SST has an effect on precipitation as mean SST for April and September both show moderate to high correlation with precipitation. Concerning El Niño and the Indian Ocean Dipole (IOD), this research supports the argument of the role El Niño and IOD have on precipitation. Where positive correlation are found with the short rains season (Black, 2005; Black et al., 2003). On the other hand, September precipitation experiences a high negative correlation with La Niña which is partly argued in previous research (Van der Ent and Savenije, 2013).

Future work

Although this work enhances the understanding of lake Nakuru variability and the impact of precipitation with underlying changing drivers, there are several important next steps. Whereas observation data for precipitation is available it is important to obtain observation data through fieldwork experiments regarding evaporation, catchment runoff, and possibly groundwater flows with isotopic tracers. This would enhance the understanding of the catchment. A more sophisticated water balance could be developed for the lake, dividing the model in different catchment parts and implementing retention times. These parts could be based on the locations of the KMD weather stations. Groundwater flow should be included in addition to using different runoff coefficients with respect to land classifications. Considering the added value of this study to our understanding of the lake Nakuru behaviour to precipitation variability, the aim should be to establish flood extent marks and mitigation plans for surrounding communities and wildlife habitat. The water balance model provides relatively accurate insights into the impact of precipitation on the lake, as it can predict the next month's impact if precipitation values, and the impact of preceding months are known. The under- and overestimation of the water balance model with respect to the observed lake volume should be compared to time-series of the surrounding lakes Naivasha and Elmenteita to see if there are similarities. The surrounding lakes could impact lake Nakuru through groundwater flow. It is furthermore interesting to elaborate on moisture tracking for all years which was not possible considering time constraints. This method of investigating precipitation variability and moisture sources produces interesting insights into the changing climatology of the catchment and could be applied to more catchment in the Rift Valley to investigate possible similarities and differences.

8

Conclusion

The objective of this research is to gain insights into the main drivers responsible for the overflowing of Lake Nakuru. Specifically, the aim is to test the premise that increased precipitation in the catchment is the primary cause of the phenomenon. This study has investigated spatial and temporal variability of precipitation in the lake Nakuru catchment by applying change point analysis and relating this variability to the lake through a water balance model. The origin and underlying causes for this variability in precipitation has been researched by applying moisture tracking with the WAM2layers model to identify moisture sources, local, and non-local climatic drivers to explain the changing hydro-meteorology of the catchment.

Change points in precipitation have been observed in catchment and lake precipitation including the majority of weather stations in 2010. Notably, this change point coincides with the rise in lake volume from 2010 onwards. Catchment precipitation has increased with about 19% while several weather stations indicate more precipitation up to 48%. The months of April and September together contribute 46% of the total increased precipitation over the catchment, which is further substantiated by the change point in September precipitation. Interestingly, the weather station change points are observed bordering the northern and eastern parts of the catchment along the mountain ridges. As the multiple linear regression analysis shows that lake precipitation is the most important independent variable as it comes to lake influx, it is interesting that the change points in precipitation are mainly focused to weather stations around the lake. Lake precipitation overall increased more than the catchment average precipitation, with 219 mm and 156 mm respectively, indicating a higher water availability in the lake's surroundings. This suggests that changes in lake precipitation is most likely the cause for the overflowing of the lake. The water balance model enhances our understanding of the effects of precipitation on the lake as it reveals different impacts for the same order of magnitude precipitation as a result of catchment interaction.

A combination of factors could be ascribed to the changing April and September months through moisture tracking an climate drivers. Over time the western Indian Ocean region east to Madagascar became an increasingly important moisture source for precipitation tied to a more than average sea surface temperature increase in this region specifically. This study reveals that moisture is produced more locally comparing prior and post change point with continental evaporation contribution increasing as well. Furthermore, a high negative correlation between El Niño and September precipitation indicate that La Niña periods have a strong positive effect on September precipitation. It is likely that with increasing sea surface temperatures, a La Niña and El Niño are increasing in magnitude over the coming years increasing the likelihood of flooding.

In conclusion, the results of this study strongly suggest that the flooding of Lake Nakuru is likely to be attributed to changes in precipitation. While there is limited research available on this subject, no significant evidence has been found to support the possibility of tectonic plate movement as a contributing factor. The sudden and significant shift in the lake's surface area and volume indicates that changes in land use alone are not a sufficient explanation for this phenomenon. Rather, it is possible that changes in land use and increased precipitation may have a mutually reinforcing impact on the overflowing of Lake Nakuru. Overall, these findings support the argument that a changing climate, in conjunction with increased precipitation, is the primary underlying cause for the overflowing of lake Nakuru.

Bibliography

Balagizi, C. M., Kasereka, M. M., Cuoco, E., and Liotta, M. (2018). Influence of moisture source dynamics and weather patterns on stable isotopes ratios of precipitation in central-eastern africa. Science of the Total Environment, 628:1058–1078.

Baraka, C. (2022). A drowning world: Kenya's quiet slide underwater. The Guadian.

Black, E. (2005). The relationship between indian ocean sea–surface temperature and east african rainfall. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 363(1826):43–47.

Black, E., Slingo, J., and Sperber, K. R. (2003). An observational study of the relationship between excessively strong short rains in coastal east africa and indian ocean sst. Monthly Weather Review, 131(1):74–94.

Cai, W., Wang, G., Gan, B., Wu, L., Santoso, A., Lin, X., Chen, Z., Jia, F., and Yamagata, T. (2018). Stabilised frequency of extreme positive indian ocean dipole under 1.5 c warming. Nature communications, 9(1):1419.

Cook, K. H. and Vizy, E. K. (2013). Projected changes in east african rainy seasons. Journal of Climate, 26(16):5931–5948.

Dyer, E. and Washington, R. (2021). Kenyan long rains: A subseasonal approach to process-based diagnostics. Journal of Climate, 34(9):3311–3326.

Findell, K. L., Keys, P. W., Van Der Ent, R. J., Lintner, B. R., Berg, A., and Krasting, J. P. (2019). Rising temperatures increase importance of oceanic evaporation as a source for continental precipitation. Journal of Climate, 32(22):7713–7726.

Food and of the United Nations, A. O. (2021). Improving water productivity in agriculture through remote sensing and modelling. WaPOR - Water Productivity Open-Access Portal.

Freedman, D. and Diaconis, P. (1981). On the histogram as a density estimator: L 2 theory. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, 57(4):453–476.

Gichuru, G. and Waithaka, H. (2016). Analysis of lake nakuru surface water area variations using geospatial technologies. In Scientific Conference Proceedings, number 1.

Gimeno, L., Nieto, R., Drumond, A., Castillo, R., and Trigo, R. (2013). Influence of the intensification of the major oceanic moisture sources on continental precipitation. Geophysical Research Letters, 40(7):1443–1450.

Gimeno, L., Stohl, A., Trigo, R. M., Dominguez, F., Yoshimura, K., Yu, L., Drumond, A., Durán-Quesada, A. M., and Nieto, R. (2012). Oceanic and terrestrial sources of continental precipitation. Reviews of Geophysics, 50(4).

Herrnegger, M., Stecher, G., Schwatke, C., and Olang, L. (2021). Hydroclimatic analysis of rising water levels in the great rift valley lakes of kenya. Journal of Hydrology: Regional Studies, 36:100857.

Hersbach, H., B. B. B. P. B. G. H. A. M. S. J. N. J. P. C. R. R. R. I. S. D. S. A. S. C. D. D. T. J.-N. (2023). Era5 hourly data on single levels from 1940 to present.

Holder, R. (1985). Multiple regression in hydrology. Institute of hydrology.

Hrachowitz, M., Fovet, O., Ruiz, L., Euser, T., Gharari, S., Nijzink, R., Freer, J., Savenije, H., and Gascuel-Odoux, C. (2014). Process consistency in models: The importance of system signatures, expert knowledge, and process complexity. Water resources research, 50(9):7445–7469.

Huang, B., Thorne, P. W., Banzon, V. F., Boyer, T., Chepurin, G., Lawrimore, J. H., Menne, M. J., Smith, T. M., Vose, R. S., and Zhang, H.-M. (2017). Extended reconstructed sea surface temperature, version 5 (ersstv5): upgrades, validations, and intercomparisons. Journal of Climate, 30(20):8179–8205.

ilec (2005). Managing lakes and their basins for sustainable use: A report for lake basin managers and stakeholders.

Iradukunda, P., Sang, J. K., Nyadawa, M. O., and Maina, C. W. (2020). Sedimentation effect on the storage capacity in lake nakuru, kenya. Journal of Sustainable Research in Engineering, 5(3):149–158.

Kanda, I. and Suwai, J. (2013). Hydrogeochemistry of shallow and deep water aquifers of menengai geothermal area, kenya rift valley. In Proceedings of the GRC Annual Meeting, Las Vegas, NV, USA, pages 403–409.

Kebede, S., Travi, Y., and Rozanski, K. (2009). The δ18o and δ2h enrichment of ethiopian lakes. Journal of Hydrology, 365(3-4):173–182.

Keim, B. D. and Muller, R. A. (1992). Temporal fluctuations of heavy rainfall magnitudes in new orleans, louisiana: 1871–1991 1. JAWRA Journal of the American Water Resources Association, 28(4):721–730.

Kenya Ministry of Environment and UNDP (2021). Rising water levels in kenya's rift valley lakes, turkwel gorge dam and lake victoria. a scoping report.

Keys, P. W., Warrier, R., Van Der Ent, R. J., Galvin, K. A., and Boone, R. B. (2022). Analysis of kenya's atmospheric moisture sources and sinks. Earth Interactions, 26(1):139–150.

Kimaru, A. N., Gathenya, J. M., and Cheruiyot, C. K. (2019). The temporal variability of rainfall and streamflow into lake nakuru, kenya, assessed using swat and hydrometeorological indices. Hydrology, 6(4):88.

KMD (2022). Precipitation daily data in lake nakuru catchment from kenya meteorological department.

Langat, P. K., Kumar, L., and Koech, R. (2017). Temporal variability and trends of rainfall and streamflow in tana river basin, kenya. Sustainability, 9(11):1963.

Levin, N. E., Zipser, E. J., and Cerling, T. E. (2009). Isotopic composition of waters from ethiopia and kenya: Insights into moisture sources for eastern africa. Journal of Geophysical Research: Atmospheres, 114(D23).

Liang, L., Li, L., and Liu, Q. (2011). Precipitation variability in northeast china from 1961 to 2008. Journal of Hydrology, 404(1-2):67–76.

Mediero, L. and Kjeldsen, T. R. (2014). Regional flood hydrology in a semi-arid catchment using a gls regression model. Journal of Hydrology, 514:158–171.

Montcoudiol, N., Burnside, N. M., Györe, D., Mariita, N., Mutia, T., and Boyce, A. (2019). Surface and groundwater hydrochemistry of the menengai caldera geothermal field and surrounding nakuru county, kenya. Energies, 12(16):3131.

Odada, E. (2001). Stable isotopic composition of east african lake waters.

Otte, I., Detsch, F., Gütlein, A., Scholl, M., Kiese, R., Appelhans, T., and Nauss, T. (2017). Seasonality of stable isotope composition of atmospheric water input at the southern slopes of mt. kilimanjaro, tanzania. Hydrological Processes, 31(22):3932–3947.

Raini, J. A. (2009). Impact of land use changes on water resources and biodiversity of lake nakuru catchment basin, kenya. African Journal of Ecology, 47:39–45.

Schumann, A. (1998). Encyclopedia of hydrology and lakes. encyclopedia of earth science.

Schwatke, C., Scherer, D., and Dettmering, D. (2019). Automated extraction of consistent time-variable water surfaces of lakes and reservoirs based on landsat and sentinel-2. Remote Sensing, 11(9):1010.

Soderberg, K., Good, S. P., O'Connor, M., Wang, L., Ryan, K., and Caylor, K. K. (2013). Using atmospheric trajectories to model the isotopic composition of rainfall in central kenya. Ecosphere, 4(3):1–18.

The World Bank Group (2021). Kenya current climate > climateology. observed average annual precipitation of kenya for 1901-2020.

Tierney, J. E., Smerdon, J. E., Anchukaitis, K. J., and Seager, R. (2013). Multidecadal variability in east african hydroclimate controlled by the indian ocean. Nature, 493(7432):389–392.

Truong, C., Oudre, L., and Vayatis, N. (2020). Selective review of offline change point detection methods. Signal Processing, 167:107299.

Van der Ent, R., Tuinenburg, O., Knoche, H.-R., Kunstmann, H., and Savenije, H. (2013). Should we use a simple or complex model for moisture recycling and atmospheric moisture tracking? Hydrology and Earth System Sciences, 17(12):4869–4884.

Van der Ent, R., Wang-Erlandsson, L., Keys, P. W., and Savenije, H. (2014). Contrasting roles of interception and transpiration in the hydrological cycle–part 2: Moisture recycling. Earth System Dynamics, 5(2):471–489.

van der Ent, R. J., Benedict, I. B., Weijenborg, C., Cömert, T., van de Koppel, N., Guo, L., and Kalverla, P. (2022). Wam2layers.

Van der Ent, R. J. and Savenije, H. H. (2013). Oceanic sources of continental precipitation and the correlation with sea surface temperature. Water Resources Research, 49(7):3993–4004.

Van der Ent, R. J., Savenije, H. H., Schaefli, B., and Steele-Dunne, S. C. (2010). Origin and fate of atmospheric moisture over continents. Water Resources Research, 46(9).

Wainwright, C. M., Finney, D. L., Kilavi, M., Black, E., and Marsham, J. H. (2021). Extreme rainfall in east africa, october 2019–january 2020 and context under future climate change. Weather, 76(1):26–31.

WMO (2012). Standardized precipitation index user guide. World Meteorological Organization, (1090).

Yang, W., Seager, R., Cane, M. A., and Lyon, B. (2014). The east african long rains in observations and models. Journal of Climate, 27(19):7185–7202.

Yang, W., Seager, R., Cane, M. A., and Lyon, B. (2015). The annual cycle of east african precipitation. Journal of Climate, 28(6):2385–2404.

Appendices

A

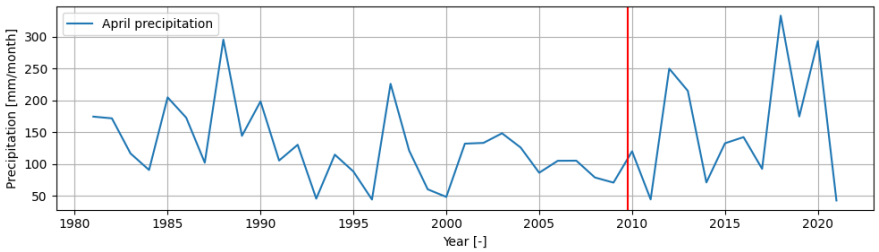
Change point detection

Table A.1: Change point detection results for weather stations 2 to 15, lake precipitation, and catchment precipitation. Eight out of fourteen weather stations experience a significant change point at 2010. Lake precipitation experiences a significant change point at 2010, however for the catchment precipitation average precipitation does not reach below 0.05 significance level.

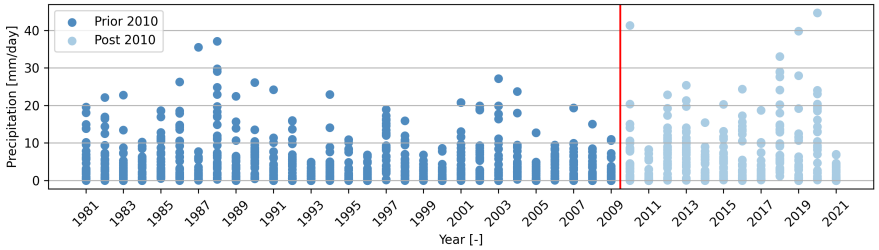
Time series	Elevation [m]	P_pre2010	P_post2010	Change [mm]	Change [%]	Change_Point	Significance
Pts2	2600	1040.21	1322.67	282.46	27.15%	2010	0.00068
Pts3	2166	916.82	1157.42	240.6	26.24%	2010	0.00398
Pts4	2380	925.79	1106.96	181.17	19.57%	2010	0.0882
Pts5	2756	1011.53	1142.85	131.32	12.98%	-	-
Pts6	1817	855.01	1024.64	169.63	19.84%	2010	0.01953
Pts7	1951	869.65	1024.82	155.17	17.84%	2010	0.06054
Pts8	2022	816.97	811.48	-5.49	-0.67%	-	-
Pts9	2564	813.22	802.75	-10.47	-1.29%	-	-
Pts10	1860	850.42	1021.58	171.16	20.13%	2010	0.02107
Pts11	1845	895.9	1266.31	370.41	41.35%	2010	0.0000
Pts12	1838	680.91	811.63	130.72	19.20%	2010	0.0122
Pts13	2305	786.5	985.07	198.57	25.25%	2010	0.0048
Pts14	2352	800.04	1182.03	381.99	47.75%	2010	0.0000
Pts15	1893	773.08	986.56	213.48	27.61%	2010	0.0008
Lake_Precipitation	1868	871.26	1090.65	219.39	25.18%	2010	0.0068
Catchment_Precipitation	2168	849.11	1005.05	145.51	18.37%	2010	0.0434

Table A.2: Change point detection results for monthly mean catchment precipitation. September and October experience a significant change point at 2010.

Time series	P_pre2010	P_post2010	Change [mm]	Change [%]	Percent of total change (%)	Change_Point	Significance
Jan	36.27	28.67	-7.60	-20.96%	-5.22%	-	0.4307
feb	30.97	30.48	-0.49	-1.59%	-0.34%	-	1
Mar	65.90	74.65	8.75	13.28%	6.02%	-	0.7635
Apr	125.47	159.13	33.66	26.83%	23.13%	-	0.3667
May	94.50	117.57	23.07	24.41%	15.85%	-	0.1400
Jun	71.53	77.85	6.31	8.82%	4.34%	-	0.6569
Jul	72.44	86.28	13.84	19.11%	9.51%	-	0.2343
Aug	91.77	93.60	1.83	2.00%	1.26%	-	0.57633
Sep	58.65	96.46	37.81	64.46%	25.98%	2010	0.0036
Oct	66.67	87.81	21.14	31.71%	14.53%	2010	0.0328
Nov	82.34	84.02	1.68	2.04%	1.15%	-	0.9657
Dec	52.61	68.56	15.96	30.33%	10.96%	-	0.4476

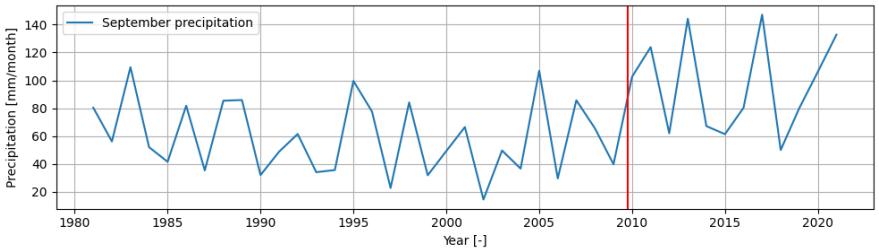


(a)

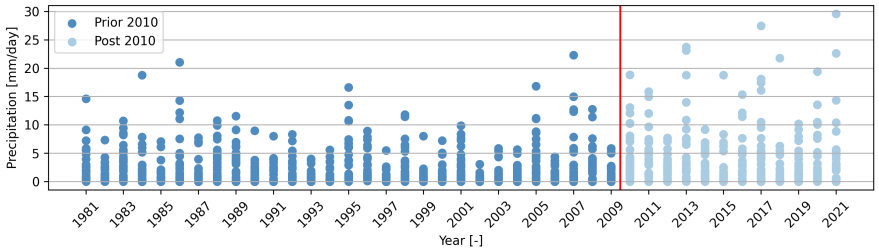


(b)

Figure A.1: A visualisation of the time series of yearly precipitation in the month of April, indicating more intense months observed as well as more intense daily precipitation, resulting in an increased precipitation of 26.83% on average per year comparing before and after 2010.



(a)



(b)

Figure A.2: A visualisation of the time series of yearly precipitation in the month of September, indicating more intense months observed as well as more intense daily precipitation, resulting in an increased precipitation of 64.46% on average per year comparing before and after 2010 for which the change point is significant in the month of September.

B

Water balance modelling

Table B.1: Monthly precipitation and total monthly contribution to volume increase accompanied by their Pearson correlation for the years 2009 to 2021

Month	Average monthly precipitation [mm/month]	Accumulated contribution to the total volume [^{3}]	Pearson correlation month observed volume change to precipitation [-]
JAN	27.3	0.00398	0.241
FEB	28.5	-0.02406	0.670
MAR	69.9	-0.01203	0.483
APR	152.3	0.06189	0.820
MAY	117.5	0.06231	0.260
JUN	73,5	0.03305	0.630
JUL	81.4	0.03347	0.130
AUG	89.0	0.04497	0.482
SEP	92.1	0.05315	0.527
OCT	85.0	-0.00980	-0.031
NOV	81.2	0.02217	0.645
DEC	72.9	0.02949	0.405

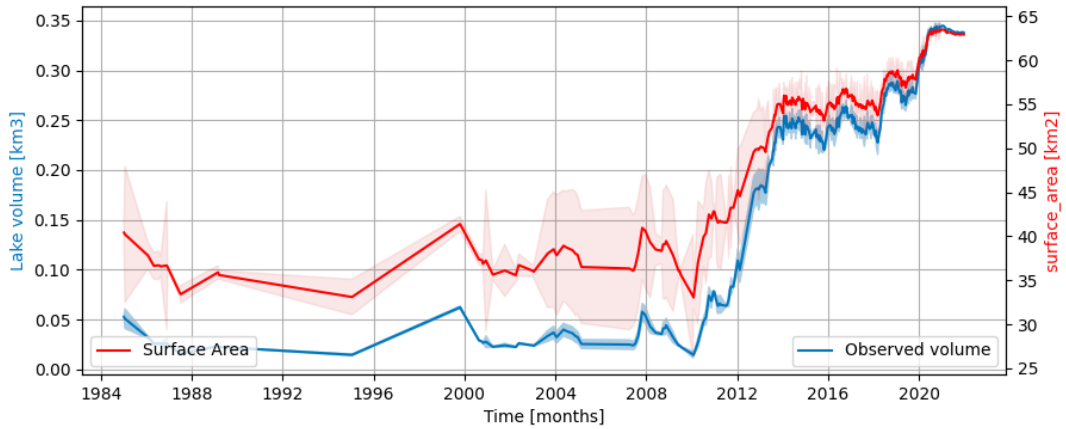


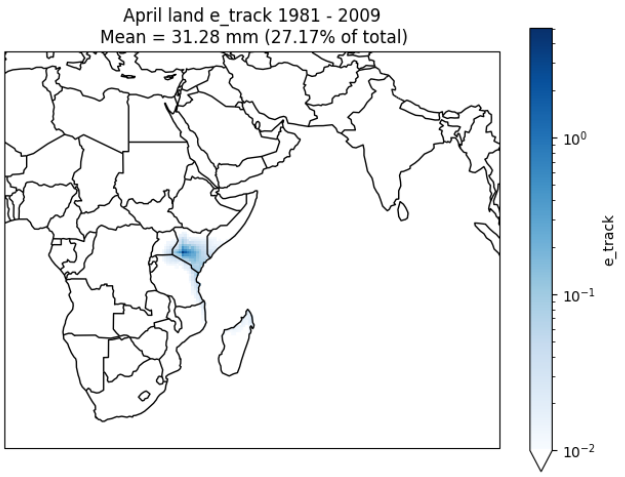
Figure B.1: Observed surface area (red) and lake volume (blue) from 1984 to 2021. An inclination is observed form 2010 onwards. Data from 1984 to 2009 is very scarse.

C

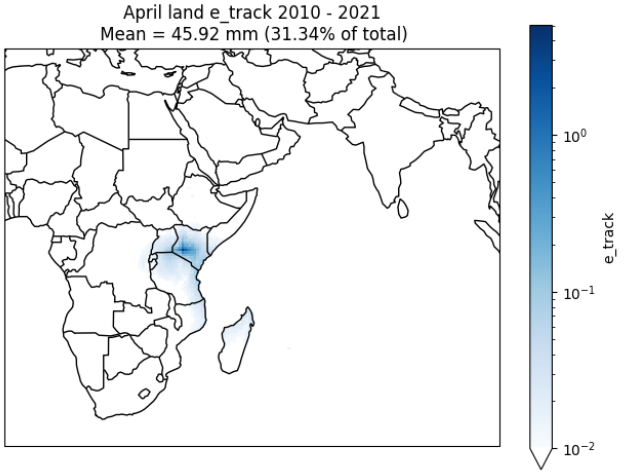
Moisture tracking April

Table C.1: Correlation results for moisture tracking sea regions April prior and post change point

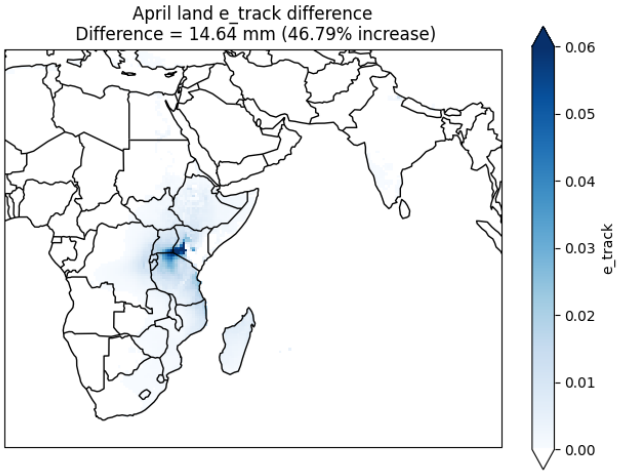
region	Correlation of KMD precipitation to % of sea evaporation from region	Correlation % before 2010	Correlation % after 2010	Correlation of KMD precipitation to absolute e_track from region	Correlation absolute before 2010	Correlation absolute after 2010
1	-0.404	-0.338	-0.545	0.863	0.860	0.873
2	0.177	0.062	0.234	0.457	0.338	0.488
3	-0.105	-0.111	0.125	0.291	0.249	0.588
4	-0.003	0.210	-0.292	0.654	0.704	0.730
5	0.0119	-0.042	-0.013	0.664	0.678	0.617
6	0.000	-0.012	0.240	0.370	0.229	0.709



(a)

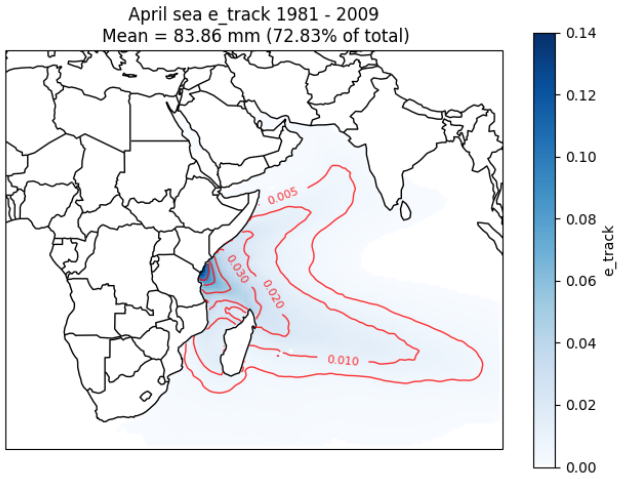


(b)

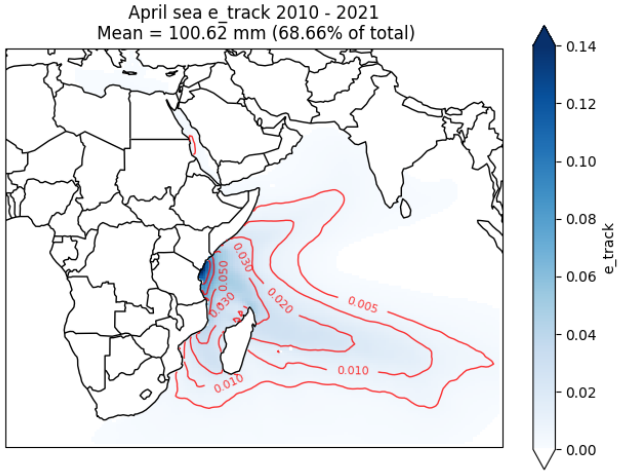


(c)

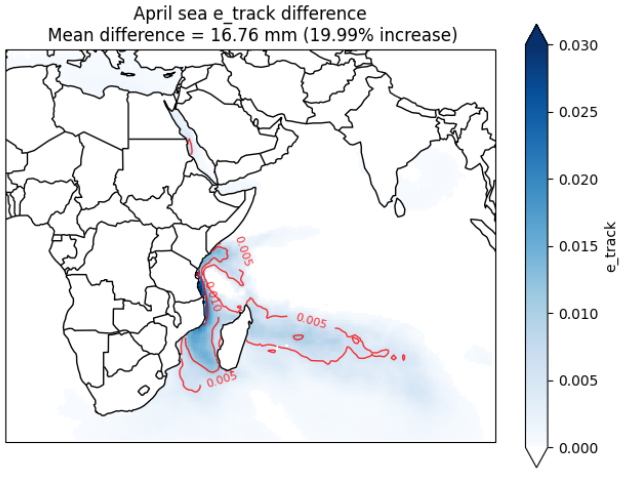
Figure C.1: April moisture tracking results of land evaporation before the change point (1981-2009) (a) and after change point (2010-2021) (b) including a visualisation of the difference in moisture source (c)



(a)

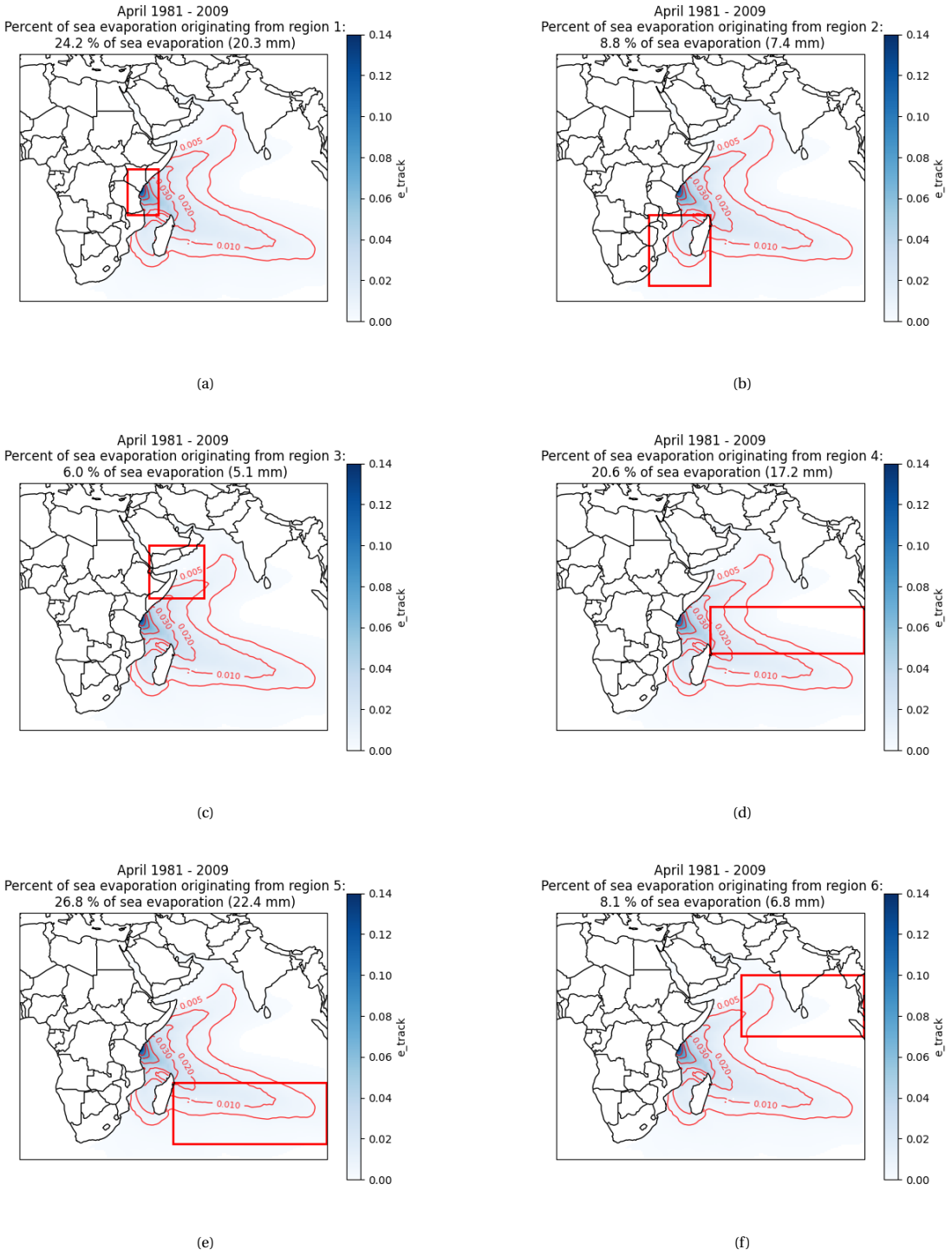


(b)



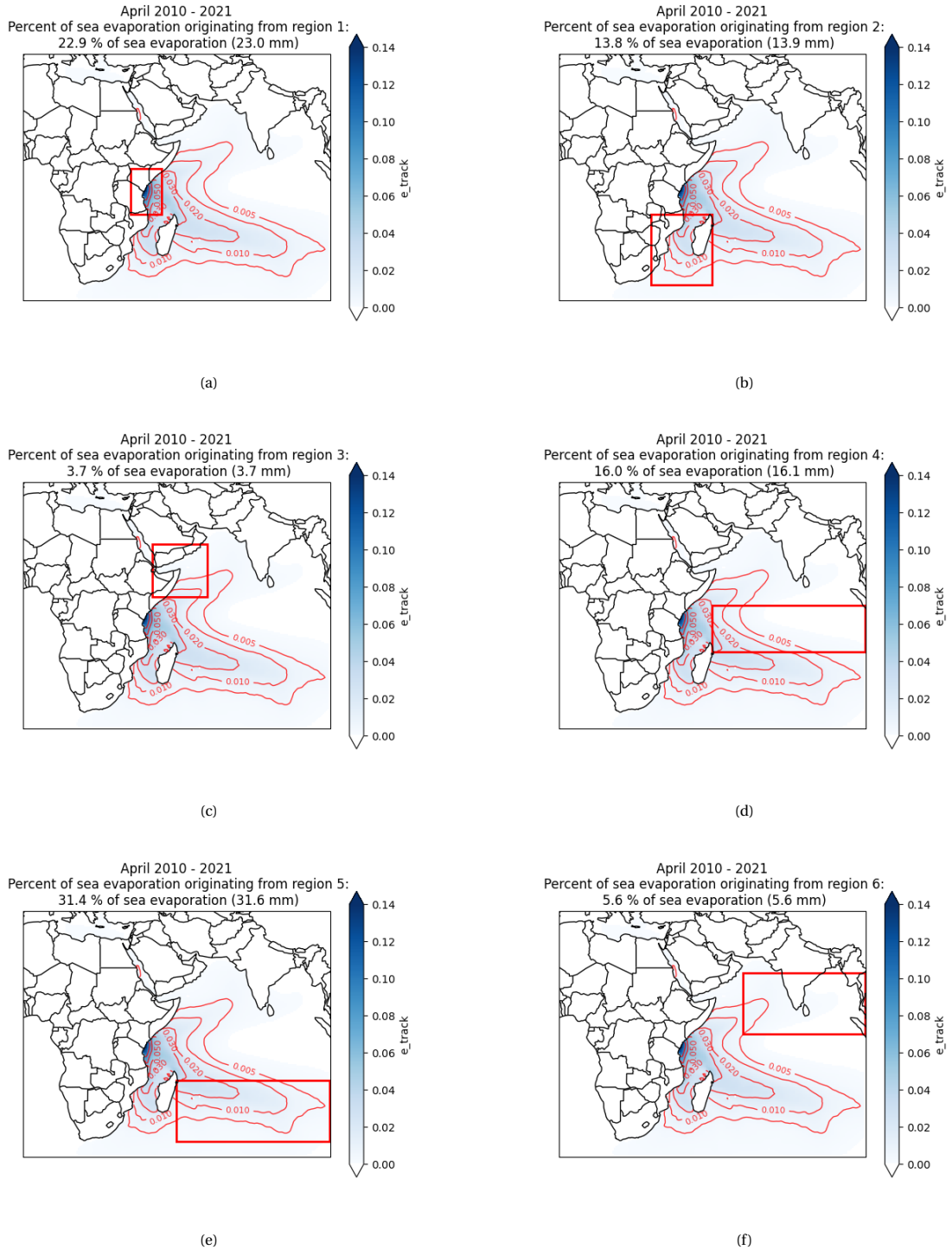
(c)

Figure C.2: April moisture tracking results of sea evaporation before the change point (1981-2009) (a) and after change point (2010-2021) (b) including a visualisation of the difference in moisture source (c)



(a) (b) (c) (d) (e) (f)

Figure C.3: April moisture tracking results of sea regions before the change point (1981-2009)



(a) (b) (c) (d) (e) (f)

Figure C.4: April moisture tracking results of sea regions after the change point (2010-2021)

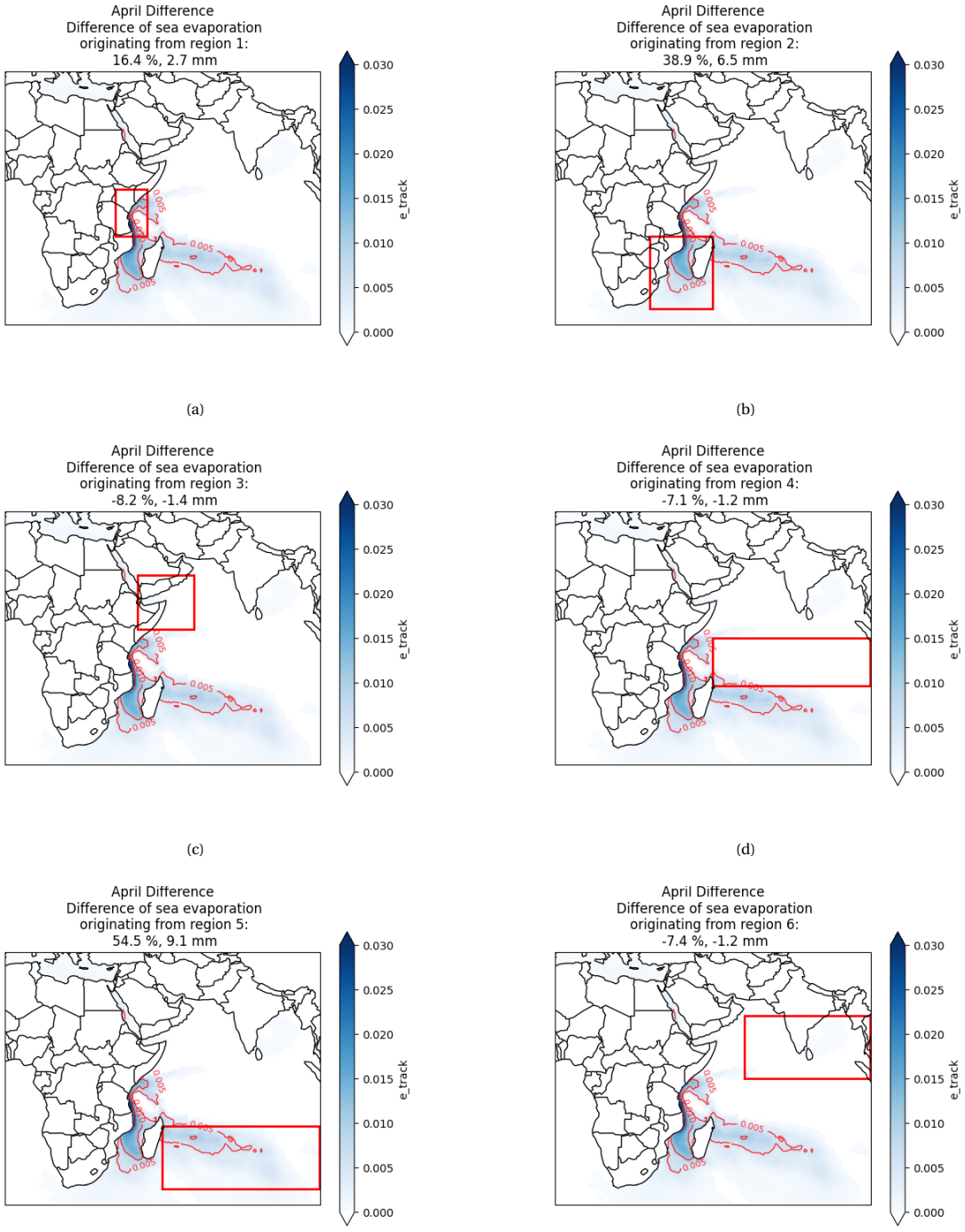


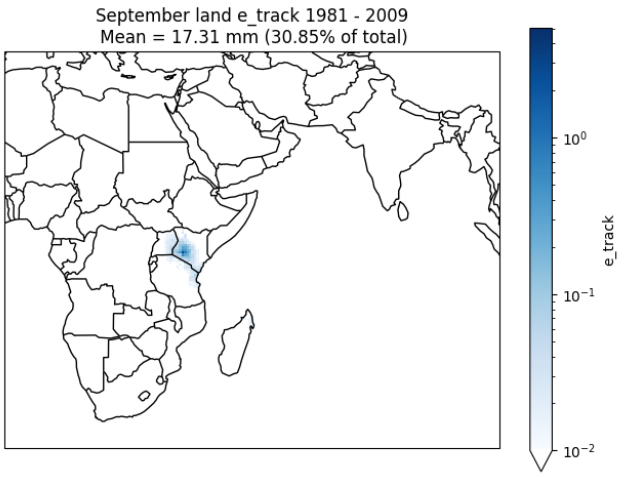
Figure C.5: April moisture tracking results of sea regions difference prior and post change point

D

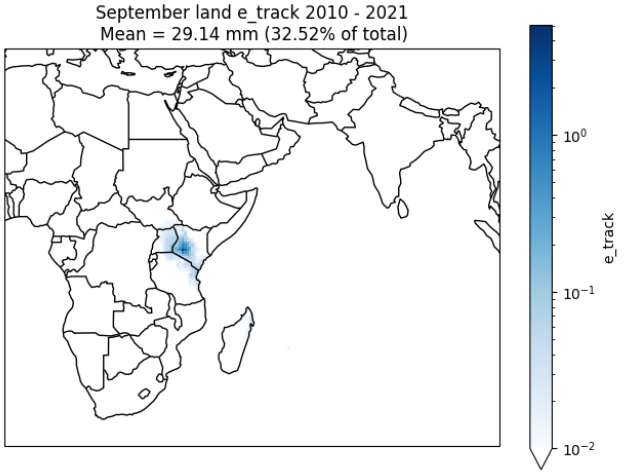
Moisture tracking September

Table D.1: Correlation results for moisture tracking sea regions September prior and post change point

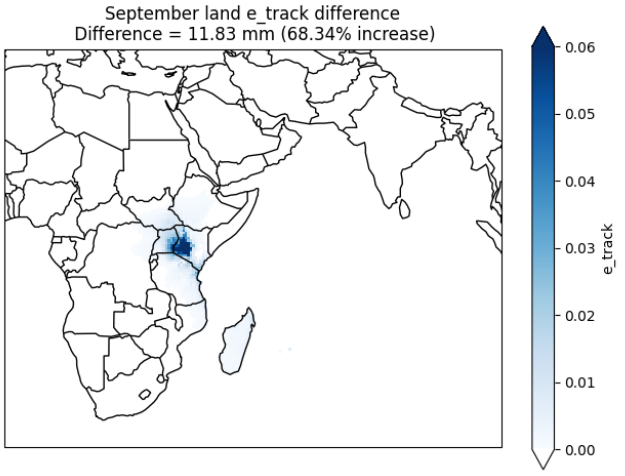
region	Correlation of KMD precipitation to % of sea evaporation from region	Correlation % before 2010	Correlation % after 2010	Correlation of KMD precipitation to absolute e_track from region	Correlation absolute before 2010	Correlation absolute after 2010
1	0.154	0.227	-0.191	0.944	0.924	0.931
2	0.29	0.216	0.254	0.682	0.589	0.671
3	-0.225	-0.165	-0.462	0.9	0.907	0.816
4	0.128	0.176	-0.139	0.931	0.928	0.883
5	-0.331	-0.387	0.133	0.632	0.622	0.677
6	0.000	-0.012	0.240	0.370	0.229	0.709



(a)

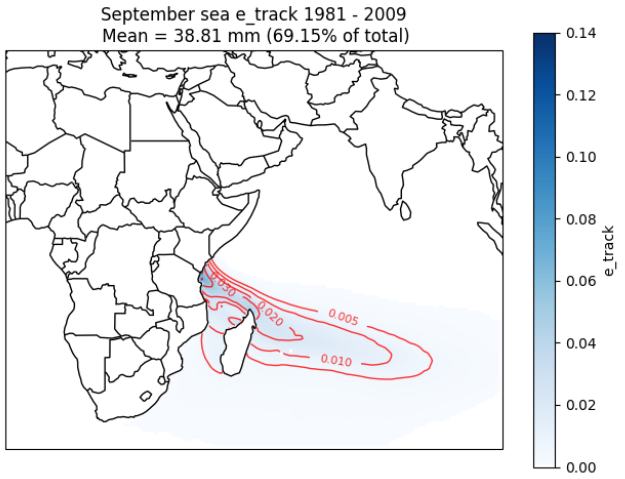


(b)

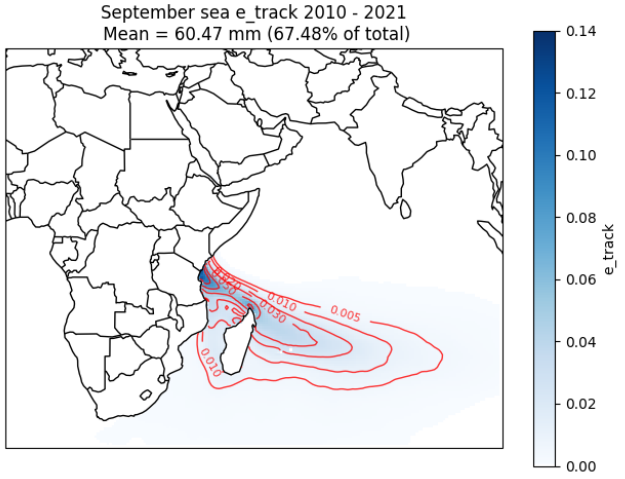


(c)

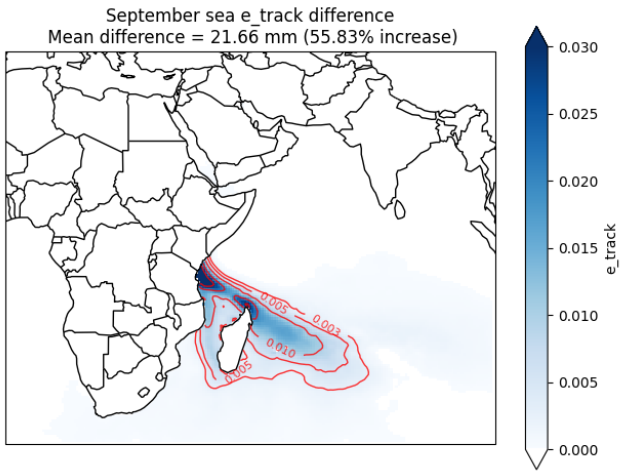
Figure D.1: September moisture tracking results of land evaporation before the change point (1981-2009) (a) and after change point (2010-2021) (b) including a visualisation of the difference in moisture source (c)



(a)

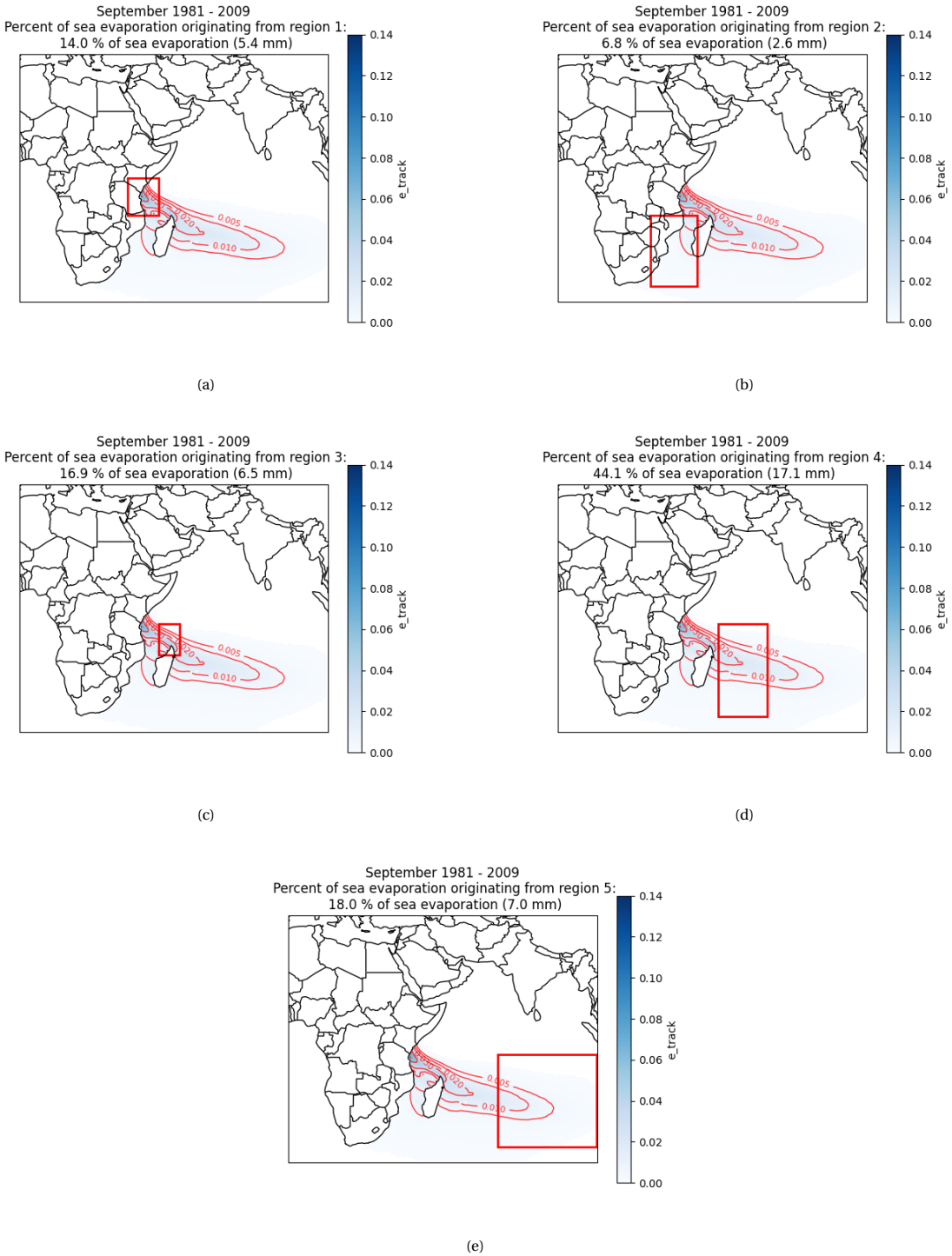


(b)



(c)

Figure D.2: September moisture tracking results of sea evaporation before the change point (1981-2009) (a) and after change point (2010-2021) (b) including a visualisation of the difference in moisture source (c)



(a)

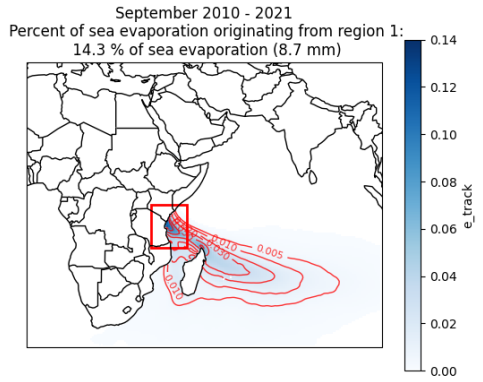
(b)

(c)

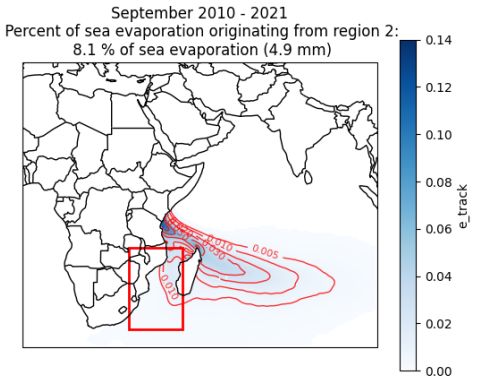
(d)

(e)

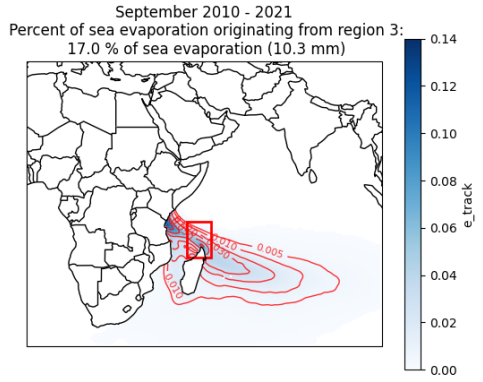
Figure D.3: September moisture tracking results of sea regions before the change point (1981-2009)



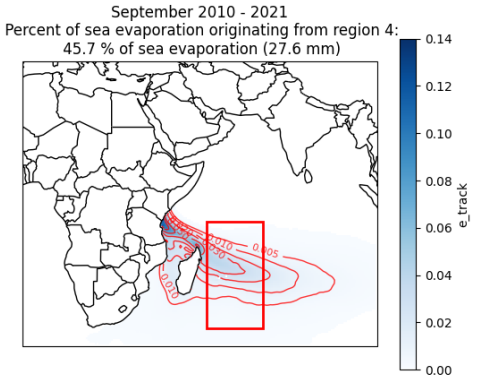
(a)



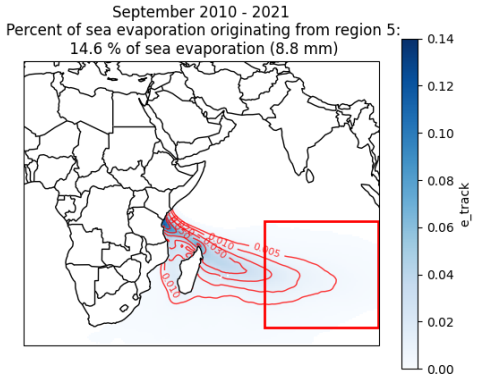
(b)



(c)



(d)



(e)

Figure D.4: September moisture tracking results of sea regions after the change point (2010-2021)

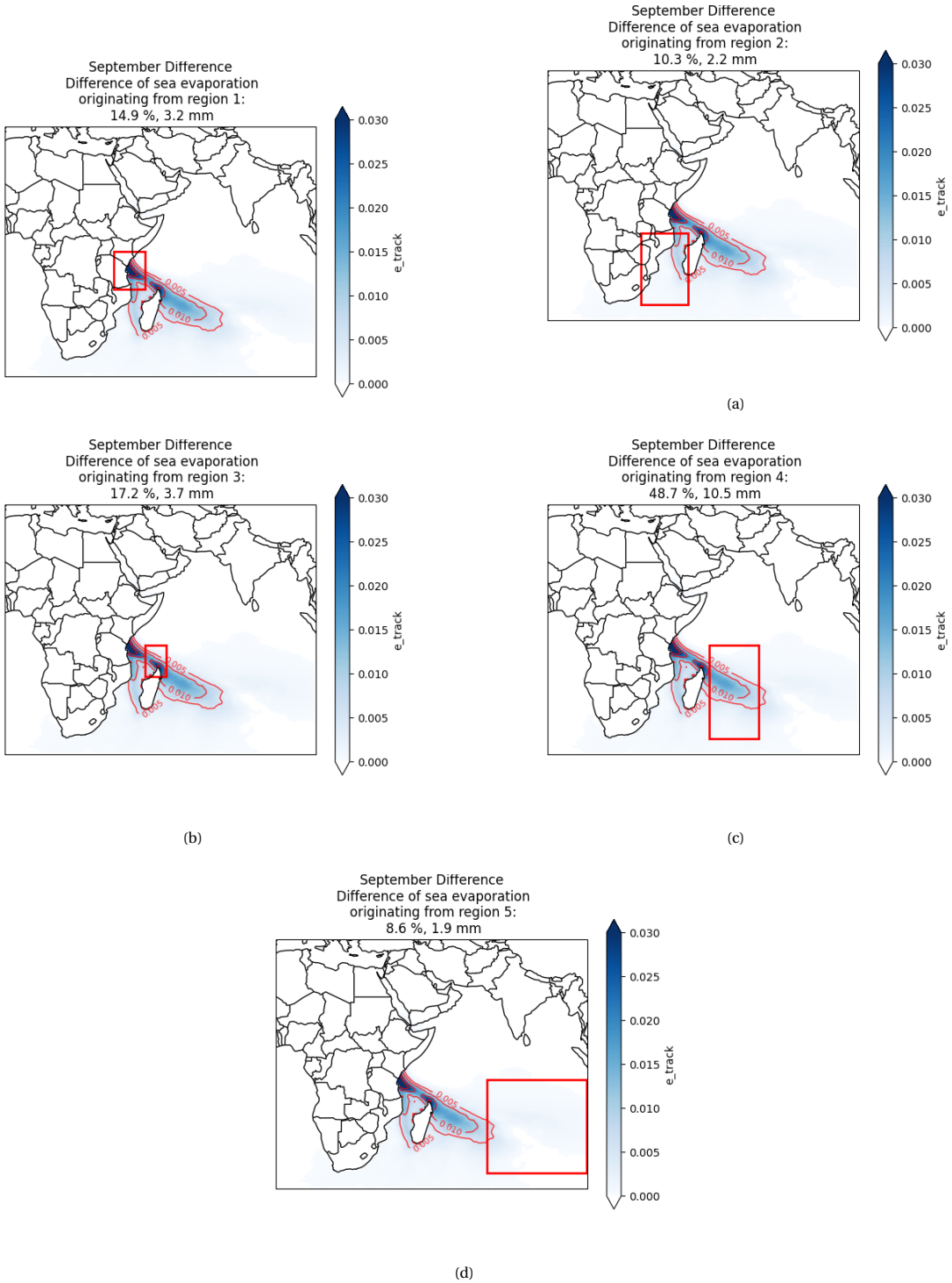


Figure D.5: September moisture tracking results of sea regions difference prior and post change point

E

Sea surface temperature analysis

Table E.1: Correlation of April SST from regions with April precipitation over a one year, three year, six year, and nine year mean

Region	Correlation 1 year	Prior 2010 1 year	Post 2010 1 year	Correlation 3 year	Prior 2010 3 year	Post 2010 3 year	Correlation 6 year	Prior 2010 6 year	Post 2010 6 year	Correlation 9 year	Prior 2010 9 year	Post 2010 9 year
1	0.219	0.171	0.157	0.387	0.215	0.398	0.333	-0.157	0.885	0.3	-0.316	0.889
2	-0.062	-0.027	-0.428	0.039	-0.064	-0.002	0.155	-0.018	0.664	-0.036	-0.233	0.338
3	0.195	0.265	-0.086	0.104	-0.055	-0.09	0.029	-0.335	0.727	-0.192	-0.594	0.429
4	0.223	0.164	0.153	0.353	0.254	0.183	0.42	-0.114	0.892	0.348	-0.2	0.871
5	-0.045	-0.001	-0.558	0.2	-0.188	-0.003	0.281	-0.274	0.856	0.207	-0.384	0.839
6	0.316	0.336	0.148	0.224	0.073	-0.039	0.167	-0.411	0.925	0.008	-0.628	0.76
All	0.182	0.205	-0.101	0.255	0.045	0.098	0.255	-0.253	0.889	0.105	-0.469	0.819

Table E.2: Correlation of September SST from regions with September precipitation over a one year, three year, six year, and nine year mean

Region	Correlation 1 year	Prior 2010 1 year	Post 2010 1 year	Correlation 3 year	Prior 2010 3 year	Post 2010 3 year	Correlation 6 year	Prior 2010 6 year	Post 2010 6 year	Correlation 9 year	Prior 2010 9 year	Post 2010 9 year
1	0.233	-0.008	-0.166	0.625	0.116	0.511	0.573	-0.31	0.18	0.71	-0.577	0.866
2	0.174	0.069	-0.504	0.425	-0.156	0.05	0.479	-0.443	0.501	0.637	-0.644	0.699
3	0.18	-0.043	0.041	0.583	0.21	0.087	0.487	-0.274	-0.416	0.715	-0.392	0.394
4	0.399	0.107	0.221	0.688	0.255	0.054	0.752	0.083	0.579	0.888	0	0.9
5	0.344	-0.039	-0.003	0.671	0.111	0.019	0.826	0.274	0.673	0.922	0.243	0.911
All	0.335	0.024	-0.104	0.676	0.136	0.121	0.788	-0.191	0.624	0.841	-0.344	0.898