EXAMINING OBSERVED AND SIMULATED EXTREME RAINFALL EVENTS OVER LAKE VICTORIA BASIN IN UGANDA

BY

OPIO RONALD
BSc. METEOROLOGY (Mak)

A THESIS SUBMITTED TO THE DIRECTORATE OF RESEARCH AND GRADUATE TRAINING IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF MASTER OF SCIENCE IN ENVIRONMENT AND NATURAL RESOURCES OF MAKERERE UNIVERSITY

JANUARY 2019
DECLARATION

I, Ronald Opio, do hereby declare, basing on my current knowledge and judgment, that this work is original and has never been published or presented to any institution for the award of a degree.

Signature and date.

This thesis has been submitted for examination with the approval of my supervisors:

Dr. Alex Nimusiima
Department of Geography, Geoinformatics and Climatic Sciences;
School of Forestry, Environmental and Geographical Sciences;
College of Agricultural and Environmental Sciences.

Dr. Geoffrey Sabiiti
Department of Geography, Geoinformatics and Climatic Sciences;
School of Forestry, Environmental and Geographical Sciences;
College of Agricultural and Environmental Sciences.
COPYRIGHT

© Ronald Opio, 2019

All rights reserved. No part of this thesis may be reproduced or distributed, in any form or by any means, without permission. Researchers, however, are free to use the content of this thesis on condition that they attribute it.
DEDICATION

For my mother, Sisily Achieng and kid brother, Edwin Paul Omalla.
ACKNOWLEDGEMENT

This study was supported by the WIMEA-ICT project using funding from the Norwegian Agency for Development Cooperation (NORAD). The project provided laboratory facilities for running the numerical experiments. Without this laboratory, this work would have been difficult to accomplish.

My academic supervisors, Dr. Geoffrey Sabiiti and Dr. Alex Nimusiima have been selflessly supportive on this journey to have this work completed. They rigorously critiqued the research idea to reveal the flaws and this shaped the work into what it is now. Furthermore, they gave me valuable input in writing the work and doing the scientific analyses. This work would not have reached this standard if it wasn’t for their guidance and support.

I also recognize my academic advisors, Associate Professor Charles Basalirwa and Dr. Daniel Waiswa for the mind opening suggestions that they gave me during consultations and ‘corridor’ discussions. Their input helped me prepare for my thesis defense. Many thanks go to my project supervisor, Dr. Isaac Mugume who strategically oversaw my research progress right from the idea level. He also mentored me on how to manage post graduate studies. I thank him very much for being there for me in this regard. Most importantly, sincere appreciation goes to the Almighty GOD for blessing me with the grace to do this work.
## TABLE OF CONTENTS

DECLARATION .................................................................................................................. Error! Bookmark not defined.
COPYRIGHT ...................................................................................................................... ii
DEDICATION ..................................................................................................................... iii
ACKNOWLEDGEMENT ...................................................................................................... iv
LIST OF TABLES ............................................................................................................... viii
LIST OF FIGURES ........................................................................................................... ix
ACRONYMS ..................................................................................................................... x
ABSTRACT ....................................................................................................................... xi

### CHAPTER ONE: INTRODUCTION

1.1 Background .................................................................................................................. 1
1.2 Research problem ........................................................................................................ 3
1.3 Objectives of the study ............................................................................................... 4
1.4 Specific objectives ....................................................................................................... 4
1.5 Research hypotheses ................................................................................................. 4
1.6 Justification ................................................................................................................ 4
1.7 Conceptual framework ............................................................................................. 5

### CHAPTER TWO: LITERATURE REVIEW

2.1 Rainfall patterns over LVB and the major drivers ..................................................... 7
2.2 WRF parameterization schemes and rainfall simulations ....................................... 8
2.3 Model grid resolution ............................................................................................... 10

### CHAPTER THREE: METHODOLOGY

3.1 Study area ................................................................................................................ 11
3.2 Data and its sources ............................................................................................... 12
  3.2.1 Boundary conditions ......................................................................................... 12
3.2.2 Observed rainfall data ........................................................................................................ 13

3.3 Data quality control .............................................................................................................. 13

3.3 Data analysis .......................................................................................................................... 13

3.3.1 Determining the spatial and temporal patterns of MAM season daily rainfall ...... 14

3.3.2 Assessing the skill of WRF model in simulating extreme rainfall under different parameterization settings ........................................................................................................ 15

3.3.3 Determining the effect of changing grid resolution on WRF model skill in simulating extreme rainfall .................................................................................................................. 21

CHAPTER FOUR: PRESENTATION OF RESULTS ........................................................................... 22

4.1 Observed spatial and temporal patterns of MAM season rainfall .................................... 22

4.1.1 Rainfall accumulation ........................................................................................................ 22

4.1.2 Rainfall anomaly ................................................................................................................ 23

4.1.3 Daily evolution of MAM season rainfall ......................................................................... 25

4.1.4 Trend of daily rainfall ........................................................................................................ 27

4.1.5 Extreme rainfall cases ...................................................................................................... 27

4.1.6 Case study period .............................................................................................................. 29

4.2 WRF simulations of extreme rainfall under different parameterization settings .......... 30

4.2.1 Observed and simulated spatial rainfall patterns ............................................................... 30

4.2.2 Comparison between observed and simulated rainfall amounts ..................................... 32

4.2.3 Observed and simulated temporal rainfall patterns ......................................................... 34

4.2.4 Accuracy and categorical skill scores ............................................................................. 36

4.2.5 Diurnal evolution of observed and simulated rainfall ..................................................... 38

4.3 Effect of grid resolution on WRF model skill in simulating extreme rainfall ............. 40

4.3.1 Effect of model resolution on the skill scores ................................................................. 40

4.3.2 Effect of model resolution on the spatial rainfall patterns ............................................ 42

4.3.3 Effect of model resolution on the temporal rainfall patterns ........................................ 42
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.3.4</td>
<td>Effect of domain size on the skill scores</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td><strong>CHAPTER FIVE: DISCUSSION OF RESULTS</strong></td>
<td>47</td>
</tr>
<tr>
<td>5.1</td>
<td>Spatial and temporal patterns of MAM season rainfall</td>
<td>47</td>
</tr>
<tr>
<td>5.2</td>
<td>WRF simulations of extreme rainfall under different parameterization settings</td>
<td>48</td>
</tr>
<tr>
<td>5.3</td>
<td>Effect of grid resolution on WRF model skill in simulating extreme rainfall</td>
<td>50</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Effect of domain size on WRF model skill</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td><strong>CHAPTER SIX: CONCLUSIONS AND RECOMMENDATIONS</strong></td>
<td>51</td>
</tr>
<tr>
<td>6.1</td>
<td>Conclusions</td>
<td>51</td>
</tr>
<tr>
<td>6.1.1</td>
<td>Spatial and temporal patterns of MAM season rainfall</td>
<td>51</td>
</tr>
<tr>
<td>6.1.2</td>
<td>WRF simulations of extreme rainfall under different parameterization settings</td>
<td>51</td>
</tr>
<tr>
<td>6.1.3</td>
<td>Effect of grid resolution on WRF model skill in simulating extreme rainfall</td>
<td>52</td>
</tr>
<tr>
<td>6.2</td>
<td>Recommendations</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td><strong>REFERENCES</strong></td>
<td>54</td>
</tr>
<tr>
<td></td>
<td><strong>APPENDICES</strong></td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>Appendix 1: Rainfall timeseries for the 2008 MAM season</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>Appendix 2: Demonstration of a trend in daily rainfall within individual MAM seasons</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>Appendix 3: Details of parameterization schemes considered for the study</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>Appendix 4: Atmospheric governing equations that are solved by the WRF ARW core</td>
<td>69</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1: Domain specifications used for investigating parameterization schemes .................. 16
Table 2: WRF model experimental setup .................................................................................. 18
Table 3: 2x2 contingency table ................................................................................................. 20
Table 4: Domain specifications used to study the effect of changing resolution ...................... 21
Table 5: Extreme rainfall events in the Lake Victoria Basin during the MAM season .......... 29
Table 6: Skill scores and ranks of the model runs. ................................................................. 37
Table 7: Skill scores generated by the 12 km and 4 km model runs ......................................... 41
Table 8: Ranks of the model skill scores for runs done at 4 km resolution ............................. 41
Table 9: Skill scores generated by model runs when different domain sizes are used .......... 46
LIST OF FIGURES

Figure 1: Conceptual framework for documenting WRF model skill ............................................. 6
Figure 2: Map of the Lake Victoria Basin of Uganda. ................................................................. 12
Figure 3: Domain configuration for WRF simulations ............................................................... 17
Figure 4: Total MAM season rainfall amount for each year ......................................................... 22
Figure 5: Normalized spatial rainfall anomaly ........................................................................... 24
Figure 6: Daily evolution of MAM season rainfall ................................................................... 26
Figure 7: Trend of MAM season daily rainfall ......................................................................... 28
Figure 8: Total rainfall amount for the 20-day period ............................................................... 31
Figure 9: Difference in total rainfall amount for the 20-day period ......................................... 33
Figure 10: Temporal rainfall accumulation ................................................................................ 35
Figure 11: 3-hourly rainfall accumulation .................................................................................. 39
Figure 12: Spatial rainfall accumulation for comparison of model resolution ......................... 44
Figure 13: Temporal rainfall accumulation for comparison of model resolution ....................... 45
Figure 14: Rainfall received during the MAM season of 2008 ..................................................... 66
Figure 15: Trend in daily rainfall for the 2008 and 2010 MAM seasons ....................................... 67
### ACRONYMS

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMET</td>
<td>Atmospheric Model Evaluation Tool</td>
</tr>
<tr>
<td>ARW</td>
<td>Advanced Research WRF</td>
</tr>
<tr>
<td>BMJ</td>
<td>Betts-Miller-Janjic scheme</td>
</tr>
<tr>
<td>ENSO</td>
<td>El Niño Southern Oscillation</td>
</tr>
<tr>
<td>GF</td>
<td>Grell-Freitas scheme</td>
</tr>
<tr>
<td>IOD</td>
<td>Indian Ocean Dipole</td>
</tr>
<tr>
<td>ITCZ</td>
<td>Inter-Tropical Convergence Zone</td>
</tr>
<tr>
<td>KF</td>
<td>Kain-Fritsch scheme</td>
</tr>
<tr>
<td>LVB</td>
<td>Lake Victoria Basin</td>
</tr>
<tr>
<td>MAM</td>
<td>March – April – May</td>
</tr>
<tr>
<td>MM5</td>
<td>Fifth-generation Mesoscale Model</td>
</tr>
<tr>
<td>MODIS</td>
<td>Moderate Resolution Imaging Spectroradiometer</td>
</tr>
<tr>
<td>NCAR</td>
<td>National Center for Atmospheric Research</td>
</tr>
<tr>
<td>NCEP</td>
<td>National Center for Environmental Prediction</td>
</tr>
<tr>
<td>NOAA</td>
<td>National Aeronautics and Space Administration</td>
</tr>
<tr>
<td>NWP</td>
<td>Numerical Weather Prediction</td>
</tr>
<tr>
<td>PBL</td>
<td>Planetary Boundary Layer</td>
</tr>
<tr>
<td>RAMS</td>
<td>Regional Atmospheric Modelling System</td>
</tr>
<tr>
<td>RRTM</td>
<td>Rapid Radiative Transfer Model</td>
</tr>
<tr>
<td>SBU_YLin</td>
<td>Stony-Brook University scheme</td>
</tr>
<tr>
<td>SON</td>
<td>September – October – November</td>
</tr>
<tr>
<td>TMPA</td>
<td>TRMM Multisatellite Precipitation Analysis</td>
</tr>
<tr>
<td>TRMM</td>
<td>Tropical Rainfall Measuring Mission</td>
</tr>
<tr>
<td>UBOS</td>
<td>Uganda Bureau of Statistics</td>
</tr>
<tr>
<td>UNMA</td>
<td>Uganda National Meteorological Authority</td>
</tr>
<tr>
<td>WRF</td>
<td>Weather Research and Forecasting Model</td>
</tr>
<tr>
<td>WSM3</td>
<td>WRF Single Moment 3 scheme</td>
</tr>
<tr>
<td>YSU</td>
<td>Yonsei University scheme</td>
</tr>
</tbody>
</table>
ABSTRACT

Rainfall extremes have strong connotations to socio-economic activities and human well-being particularly in the Lake Victoria Basin (LVB) of Uganda. Timely and reliable prediction and dissemination of extreme rainfall events is therefore of paramount importance to the region’s development agenda. The main objective of this study was to contribute to the prediction of rainfall extremes over this region using a numerical modelling approach. Finalized reanalyzes from the National Center for Environmental Prediction (NCEP) were downscaled using the Weather Research and Forecasting (WRF) model and the output verified against rainfall observations from the Tropical Rainfall Measuring Mission (TRMM). To start with, this study investigated a 10-year TRMM rainfall record and revealed a 20-day period in which both extremely heavy and low rainfall was received. The rainfall during this period was then simulated using the WRF model with a particular interest of investigating the performance of different combinations of cumulus and microphysical parameterization along with the model grid resolution and domain size. The results showed that the model was able to simulate the rainfall events and the most satisfactory skill was obtained with a model setup using the Grell 3D cumulus scheme in combination with the SBU_YLin microphysical scheme. This study concludes that WRF can be used for simulating extreme rainfall over western LVB. In the other 2 regions, central and eastern LVB, its performance is limited by its inability to simulate nocturnal rainfall. Furthermore, increasing the model grid resolution showed good potential for improving the model simulation especially when a large domain is used.
CHAPTER ONE: INTRODUCTION

1.1 Background

Rainfall is one of the important weather elements and its forecasting is crucial to society (Darji et al., 2015; Prakash et al., 2016; Valipour, 2016). Rainfall distribution in time and space influences among many things, food production and water availability. For example, rainfall seasonality determines crop yield in most of sub-Saharan Africa including Uganda where most agricultural activities are rainfed (Ray et al., 2015). Albeit these benefits, rainfall in its extreme cases has a rather undesirable impact on society and different socio-economic sectors. For example, extremely low rainfall is synonymous with drought and lack of clean domestic water while extremely heavy rainfall is a key contributor to hydro-meteorological disasters such as floods and landslides (Jayawardena, 2015; Thakur & Chandel, 2016). In addition, the large amounts of water associated with heavy rainfall extremes catalyze the spread of waterborne diseases (Cann et al., 2013; Olago et al., 2007). For this reason, accurate and timely rainfall forecasts are crucial to facilitate planning in rain-dependent sectors such as agriculture, public health, disaster management and transport.

End users of rainfall forecasts (such as farmers) are mainly interested in knowing rainfall quantities as well as the spatial and temporal distribution. In case of extreme rainfall, they would also like to get early warnings to enable them prepare for the event (Masinde et al., 2012). The challenge, however, for weather service providers especially in developing regions is the difficulty in prediction of these rainfall extremes, which is partly due to lack of advanced scientific tools and reliable data sets. In East Africa, this prediction problem is further compounded by the complex interaction of climate systems and the variation in local geography (Owiti & Zhu, 2012), ultimately resulting in less skillful forecasts.

One widely used technique for predicting rainfall extremes is numerical weather prediction (NWP) modelling and currently, the popular NWP models used for this purpose include the Weather Research and Forecasting (WRF) model (Skamarock et al., 2008), the Regional Atmospheric Modelling System (RAMS) (Cotton et al., 2003), and the fifth-generation Mesoscale Model (MM5) (Dudhia, 1993) among others.
In operational forecasting, NWP models are often used in tandem with other tools such as weather charts, weather observations and climatology. They have 3 unique strengths; objectiveness, ability to automate ‘routine’ tasks and flexible data outputs which can be integrated with observations to improve the forecast quality (Huang et al., 2012; Mugume et al., 2016a; Sills, 2009).

The use of models for making atmospheric predictions is increasing globally because of the increasing accuracy of model outputs and increasing computing capabilities in terms of processing speed and data storage coupled with the rapidly declining cost of computing hardware (Warner, 2011a). For numerical models, their good performance in predicting weather is partly because of advancements in the representation of the atmospheric governing equations (Palmer, 2017) and better definition of physical parameterization for representation of sub-grid scale processes (Bauer et al., 2015).

Parameterization is a means of understanding and representing weather and climate processes that either occur on very small scales (for example convection and radiation transfer), or are too complex and not adequately understood to be modelled explicitly (Stensrud, 2007). Through this, the model is able to include these processes in the simulation thus increasing the chance of deriving a more skillful forecast. Stensrud (2007) also explains a variety of physical parameterization schemes, including microphysical parameterization, land surface – atmosphere parameterization, planetary boundary layer and turbulence parameterization, convective (cumulus) parameterization, and radiation parameterization among others.

For rainfall simulations in particular, the most important are cumulus and microphysical parameterization because of the direct influence they have on convective and non-convective rainfall respectively (Klein et al., 2015). Cumulus parameterization represents moist convection, which directly determines atmospheric circulation and water vapor distribution (Stensrud, 2007), while microphysical parameterization describes atmospheric heat transfer (Gilmore et al., 2004a) and evolution of a cloud system in terms of its formation, growth and disintegration (Gilmore et al., 2004b) hence influencing the spatial distribution of simulated rainfall.

This study therefore explored the possibility of using an NWP model for simulating extreme rainfall over the Lake Victoria Basin (LVB) of Uganda.
1.2 Research problem

Occurrence of rainfall extremes has been associated with a variety of negative socio-economic impacts on human livelihood. For example, extremely heavy rainfall plays a principal role in soil erosion (Bamutaze et al., 2017), flooding (Lwasa, 2010), landslides (Mugagga et al., 2012) and transmission of waterborne diseases (Cann et al., 2013) while extremely low rainfall is a contributing factor to occurrence of drought (Dutra et al., 2013). Consequently, they cause substantial loss in the sectors of agriculture, public health, transport and disaster management among others especially in developing countries.

Prediction of rainfall extremes is still an eminent challenge most especially in developing countries like Uganda. This is partly due to lack of advanced forecast tools (Eza, 2013) and poor understanding of climate systems that control the frequency and intensity of these extremes especially under a changing climate (O’Gorman, 2015). To address the urgent need for predicting these extremes, numerical weather prediction (NWP) models are one of the tools that are widely applied.

The Uganda National Meteorological Authority (UNMA) has integrated NWP modelling in their weather forecasting service to support the generation of short and medium range forecasts (UNMA, 2016). However, over LVB there are limited studies exploring the strength of NWP models in simulating rainfall events. Mugume et al. (2017a, 2017b) evaluated the performance of cumulus parameterization schemes neglecting the influence of microphysical parameterization. Argent et al. (2015) evaluated the performance of several parameterization combinations on seasonal rainfall, which leaves the model performance in simulating extreme rainfall over shorter timescales unexplored. The present study was therefore targeted to fill this void.
1.3 Objectives of the study

The main objective of this study was to contribute to prediction of extreme rainfall over the Lake Victoria Basin in Uganda.

1.4 Specific objectives

The specific objectives of this study were:

1. To determine the observed spatial and temporal patterns of MAM season daily rainfall.

2. To assess the skill of WRF model in simulating extreme rainfall under different parameterization settings.

3. To determine the effect of changing grid resolution on WRF model skill in simulating extreme rainfall.

1.5 Research hypotheses

H₀: There is no significant trend in observed MAM season daily rainfall.

H₀: There is no significant variation in WRF model skill when different cumulus – microphysical parameterization combinations are used.

H₀: There is no significant variation in WRF model skill in simulating extreme rainfall at different grid resolutions.

1.6 Justification

This study sought to establish the skill of the WRF model in simulating extreme rainfall over LVB in Uganda and identify suitable combinations of cumulus and microphysical parameterization schemes for this purpose. Additionally, it explored the relevance of increasing the model’s grid resolution and changing the domain size.

These results contribute to the much-needed customization of the WRF model for operational rainfall forecasting over LVB and offer a basis for future research on the influence of physical parameterization on rainfall simulations in equatorial regions. Furthermore, this advancement in
numerical rainfall prediction is useful for provision of accurate rainfall forecasts to a variety of sectors: in agriculture, to facilitate planning of planting and harvesting periods, in public health to facilitate planning for response to waterborne diseases such as cholera, and in disaster management, to inform early warnings for heavy rain episodes which are vital to facilitate planning for disasters such as floods and landslides.

From an industry perspective, this study advances work in medium range weather forecasting over the LVB in Uganda, thus directly contributing to strategic objective 4 of the Uganda National Meteorological Authority (UNMA) which dwells on improving the accuracy and reliability of forecasts for the benefit of end-users (UNMA, 2013). Additionally, this study informs the Uganda National Climate Change Policy (Ministry of Water and Environment, 2015) which emboldens the need to predict rainfall extremes.

1.7 Conceptual framework

In this study, an attempt was made to simulate the observed extreme rainfall over LVB in Uganda using a numerical model, that is, the WRF model. Currently, the skill of WRF model for predicting short term extreme rainfall events is not well documented. This study makes an attempt to ascertain this skill using the framework presented in figure 1.

The model was first used in its default state (settings) to simulate extreme rainfall. If it exhibited satisfactory performance, in terms of reproducing the rainfall quantities and distribution in time and space, then the skill it yielded would be considered. However, if the model was unable to reproduce the observed rainfall, then modifications would be made in the model setup to improve its skill. This would be in two phases, the first phase was to change the combinations of cumulus and microphysical parameterization, and the second phase was to change the grid resolution.
No clear information on WRF model skill in simulating extreme rainfall over LVB

Examine WRF model skill using an extreme rainfall case study

Model does not detect rainfall events and simulates rainfall amounts and distribution incorrectly

Model detects rainfall events and correctly simulates rainfall amounts and distribution

Change model cumulus and microphysical parameterizations

Change model grid resolution

WRF Model Skill

Figure 1: Conceptual framework for documenting WRF model skill in simulating extreme rainfall.
CHAPTER TWO: LITERATURE REVIEW

This review focussed on 3 aspects. These were; understanding the rainfall patterns of LVB and its underlying drivers, exploring the basis for selection of the parameterization schemes investigated and exploring the current schools of thought with regard to using very fine model resolution.

2.1 Rainfall patterns over LVB and the major drivers

The LVB region was purposefully selected for the study because the climatological zones in which it is located have been reported to receive substantially high rainfall amounts during the March to May (MAM) season (Basalirwa, 1995; Majaliwa et al., 2015). Therefore, it seemed logical to explore the rainfall records of this region in the quest for extremely heavy rainfall.

The distribution of rainfall over LVB in time and space is a key determinant of socio-economic activities such as agriculture, transport and human settlement. Previous studies have revealed high spatial-temporal variation and an increase in rainfall amounts over this area. For example, Mugume et al. (2016b) investigated MAM season rainfall for LVB in the past 15 years (2000 to 2015) using decadal timesteps and concluded that there was an increase in the frequency of extreme rainfall, the overall rainfall amounts and rain days in the third and seventh dekads. Awange et al. (2013) found that there was slight increase in rainfall amounts in the period from 1998 to 2012, and Kizza et al. (2009) discovered that the long rains (March to May) had a positive trend during the past century (1901 to 2000).

This variation of rainfall over the LVB is driven by synoptic and mesoscale phenomena. The synoptic phenomena include; the Inter-Tropical Convergence Zone (ITCZ), perturbations in global sea surface temperatures, and subtropical anticyclones while the mesoscale phenomena include; topography and lake-induced mesoscale circulations (dominantly lake/land breeze). The influence of these systems is explained below.

Passage of the ITCZ over the equatorial part of Uganda and the LVB in particular creates the bimodal rainfall pattern, that is, the two observed rainfall seasons, March to May (‘long rains’) and October to December (‘short rains’) (Nicholson, 2017). Changes in global sea surface temperatures influence interannual and decadal rainfall variations over LVB. Particularly,
temperature changes over the equatorial Pacific and Indian Oceans account for two modes of variability, that is, El Niño Southern Oscillation (ENSO) and the Indian Ocean dipole (IOD) respectively (Nicholson, 2015; Omondi et al., 2013). Subtropical anticyclones generate synoptic wind flow into Uganda. For example, the St. Helena anticyclone drives the westerly flow of the Congo airmass which contributes to rainfall on the western side of the basin while the Mascarene anticyclone drives the easterly flow from the Indian Ocean which generates thunderstorm events over Uganda (Jury, 2017). Diurnal rainfall events over the basin are mainly driven by topography and lake-induced mesoscale circulations such as the lake/land breeze (Basalirwa, 1995). The influence of topography is substantial in the west most side of the basin, where the existence of highlands causes orographic lifting of moisture. On the other hand, the influence of the lake-induced circulations is most important in the watershed closet to the lake.

To aid the investigation of MAM season rainfall, techniques such as normalized anomaly, simple linear regression and the percentile method were used. Similar to Aragão et al. (2007), normalized spatial anomaly was used to compare MAM season rainfall in each year of interest to the long term rainfall records. The strength of this technique is that it reveals the spatial variations in the rainfall anomaly. To investigate the temporal patterns of daily rainfall, linear regression was used. This technique was also employed by Mugume et al. (2016b) to investigate dekadal rainfall over the LVB region. Lastly, to reveal the days that received extreme rainfall, the percentile method was used. Ngailo et al. (2016) used the same technique to study extreme rainfall over Tanzania. The above-mentioned methods are described in-depth in section 3.4.1.

2.2 WRF parameterization schemes and rainfall simulations

The WRF model was selected for the rainfall simulations because it is highly customizable owing to the fact that its source code is freely available for users to download and modify, and it also offers a wide variety of parameterization schemes for use in the simulations (Skamarock et al., 2008). The model also has a superior nesting capability which allows for simulations at resolutions as high as 0.5 km, and its outputs can be autonomously analyzed using existing state-of-the-science tools such as the Atmospheric Model Evaluation Tool (AMET) (Appel et al., 2011). This explains why the WRF model user community is on the rise and shows the confidence that scientists have in this model (Warner, 2011a).
Similar to the work of Mayor and Mesquita (2015) and basing on the earlier explanation, cumulus and microphysical parameterization were given high priority in this study and the experiments were designed to dwell on changing the combinations of the cumulus and microphysical schemes used while keeping all other parameterization schemes fixed.

For the investigation, the model was first tested with its default settings, employing the Kain-Fritsch (KF) cumulus scheme (Kain, 2004) and the WRF Single Moment 3 class (WSM3) microphysical scheme (Hong et al., 2004) because this configuration has shown satisfactory skill in 1-year long simulations over East Africa (Pohl et al., 2011). Additional experiments were set up using a set of 3 cumulus and 3 microphysical parameterization schemes. The cumulus schemes selected were; Grell 3D (Grell, 1993; Grell & Dévényi, 2002), Betts-Miller-Janjic (BMJ) (Janjić, 1994) and Grell-Freitas (GF) (Grell & Freitas, 2014) whereas the microphysical schemes selected were; a single moment scheme, Eta (Rogers et al., 1996, 2005), a double moment scheme, Stony-Brook University (SBU_YLin) (Lin & Colle, 2011), and a triple moment scheme, Milbrandt (Milbrandt & Yau, 2005b, 2005a).

Part of the motivation to select some of the schemes mentioned above was to benchmark the best performing runs recommended in earlier WRF rainfall studies over LVB. These were; Argent et al. (2015) who found that the BMJ cumulus scheme in combination with the Eta microphysical scheme gave most satisfactory performance for seasonal rainfall simulations over the entire LVB including parts of Democratic Republic of Congo. Also, Sun et al. (2014) acknowledged that the Grell 3D cumulus scheme in combination with the Eta microphysical scheme gave most satisfactory performance for a 5 day rainfall simulation over the entire LVB. This study thus set out to find out how these scheme combinations would perform in simulations of extreme rainfall.

To ascertain the model’s performance in simulating extreme rainfall, this study used the standard model evaluation techniques described by Warner (2011b). The continuous verification techniques used were the root mean square error and the mean error whereas the categorical techniques used were the probability of detection and the false alarm ratio.
2.3 Model grid resolution

Selection of optimal grid resolution for numerical experiments presents a challenge to scientists. While fine (high) resolutions are more computationally demanding, they offer advantages of capturing very small-scale processes and inhomogeneities in boundary conditions. Accordingly, the resultant model skill is often higher than when coarse (low) resolutions are used (Cardoso et al., 2013). On the other hand, as explained by Warner (2011b), as the grid resolutions become finer (higher), the model is able to resolve processes which have already been represented through parameterization. Consequently, this leads to duplication of processes and overall, an unrealistic representation. Also, Stensrud (2007) explained that at very high model resolutions, some of the parameterization assumptions become invalid. For these reasons, coarser resolutions could thus yield more realistic results than finer resolutions, moreover at a lower computational cost.

As an initial step to identifying optimal grid resolutions that yield the most realistic representation, scientists often compare model performance of larger coarse parent domains to smaller nest domains of the same experiment. For example, Cardoso et al. (2013) used the WRF model to compare the performance of a 27 km domain to that of a 9 km domain and found that the 27 km domain gave better agreement with the observations on a quantile basis. Similarly, Mass et al. (2002) compared the performance of the mesoscale model, MM5 at 36 km (parent domain), 12 km (nest domain) and 4 km (nest domain) resolution. The results showed that the 12 km resolution domain generated the best accuracy, outperforming the 4 km resolution domain. These 2 studies thus give evidence that when scores for domains of the same experiment are compared, the highest resolution domain might not necessarily exhibit the best performance.

This approach of comparing the performance of domains of the same experiment seems unrealistic because these domains not only have different grid resolution but also domain size. Therefore, disparities in the result could be argued as partly caused by the difference in domain size other than grid resolution alone. An experimental setup where only grid resolution varies could give better meaning. For this reason, to investigate the effect of changing model resolution, the present study developed a new approach of setting up 2 sets of independent WRF experiments with identical domain sizes but different grid resolutions. A comparison was then made between the performance of the 2 innermost model domains. To document the model’s performance at the different resolutions, the continuous and categorical scores highlighted in section 2.2 were used.
CHAPTER THREE: METHODOLOGY

3.1 Study area

This study focused on the LVB in Uganda (Figure 2). The entire basin, however, covers an area of 184,000 km\(^2\) (Onyutha et al., 2016) and is inhabited by over 30 million people (Awange et al., 2013). The basin is shared by five countries: Uganda, Kenya, Tanzania, Rwanda and Burundi and is an important watershed for Lake Victoria, the largest fresh water lake in Africa and second largest in the world.

About 43% (approximately 79,120 km\(^2\)) of the basin lies in Uganda (Onyutha et al., 2016), stretching from the east at the border with Kenya, to the south at the border with Tanzania. Specifically, it lies within longitudes 29° 40’00” east to 34° 20’00” east and latitudes 0° 40’00” north to 1° 40’00” south and its traversed by the equator at latitude 0°. The area is generally low lying near the lake shores (493 to 646 meters above sea level), and the altitude increases outward from the shores. The central part of the basin (region 2) has a relatively higher altitude (1193 to 1401 meters above sea level) whereas the south western part (region 1) has the highest altitude (about 1402 to 2514 meters above sea level).

The basin experiences equatorial climate (Kizza et al., 2009), receiving about 1200 mm of rainfall annually, spread over two rainfall seasons, March to May (MAM) and September to November (SON) (Mugume et al., 2016b). The rainfall patterns are mainly influenced by the seasonal shift of the Inter-Tropical Convergence Zone (ITCZ), changes in global sea surface temperatures, and lake-induced mesoscale circulations among others.

In the present study, the delineation of LVB into 3 regions of varying altitude was to aid the investigation of the spatial rainfall patterns. The motivation for this delineation based on altitude is because topography has been documented to influence rainfall regimes over the East African region where the LVB lies (Ogwang et al., 2014; Owiti & Zhu, 2012).
3.2 Data and its sources

Data consisted of reanalyzed data from a global data assimilation system which was used to provide boundary conditions for running the WRF model and observed rainfall data collected by satellites. The details of these data sets are described in the subsections below.

3.2.1 Boundary conditions

The WRF model simulations was initialized using National Center for Environmental Prediction (NCEP) Final Operational Model Global Tropospheric Reanalyses. This is 6–hourly data of 1° x 1° (110 km) resolution, collected by a global data assimilation system which collects data from atmospheric soundings, satellites, and global telecommunication systems. Parameters collected
include: convection, sea level pressure, surface winds, sea surface temperature, soil temperature, air temperature, geopotential height, humidity and evaporation. These are collected for atmospheric levels from 1000 hectopascals to 10 hectopascals (Kalnay et al., 1996). The boundary conditions used for the WRF experiments were for 21 days, corresponding to the case-study period selected with the addition of an extra day for model spin-up.

3.2.2 Observed rainfall data
Owing to the sparse density of ground observation stations over the study region, the present study used a satellite-based rainfall dataset, that is, Tropical Rainfall Measuring Mission (TRMM). This dataset has been found to closely mimic the rainfall climatology over East Africa (Dinku et al., 2007). Specifically, this study used a combined product, which is, TRMM Multisatellite Precipitation Analysis (TMPA). Observed estimates of daily and 3-hourly rainfall accumulations were obtained from TRMM 3B42 version 7. This product is a 0.25° x 0.25° (27.5 km) resolution precipitation estimate obtained by applying the Goddard Profiling Algorithm to extract rainfall estimates from a variety of passive microwave sources (Huffman et al., 2007). The dataset was downloaded from the GIOVANNI data handle (https://giovanni.gsfc.nasa.gov/giovanni/) of the National Aeronautics and Space Administration (NOAA). The data used was for the March to May seasons of 10 years, 2008 to 2017.

3.3 Data quality control
To aid comparison of the WRF model output and TRMM rainfall, it was essential to ensure that both data sets were of the same grid dimensions. For this purpose, the National Center for Atmospheric Research (NCAR) command language tool version 6.3 (NCAR, 2017) was used to remap the model output to 0.25° x 0.25° spatial resolution (to match the rainfall data) using the mass conservation remapping method (Erath et al., 2013), because of its ability to retain grid values as it allocates them onto a destination grid.

3.3 Data analysis
Rainfall extremes were identified by investigating rainfall observations in the March to May (MAM) season because it is the longest rainfall season in Uganda and thus the most important for crop growing (Nsubuga et al., 2014). A suitable case study period having several rainfall extremes
was then selected for simulation within the WRF model. The methods used to analyze the data are thus arranged following the objectives as shown in the subsections below.

3.3.1 Determining the spatial and temporal patterns of MAM season daily rainfall
This section presents the methods used to study the space-time patterns of the 10-year TRMM rainfall record (2008 to 2017).

3.3.1.1 Normalized anomaly
This method was used to compare the rainfall amount for each of the years of interest to the long-term historical rainfall records. Similar to Aragão et al. (2007), the rainfall anomaly for a year, \( x \) was calculated at each grid cell \((i, j)\) as a departure from the long term (20 year) mean of 1998 to 2017, normalized by the standard deviation (SD) (Equation 1).

\[
Rainfall_{anomaly}(i,j) = \frac{Rainfall_x(i,j)-Rainfall_{1998\ to\ 2017}(i,j)}{SD_{1998\ to\ 2017}(i,j)}
\]  

(1)

3.3.1.2 Multiple linear regression
This method was used to determine the magnitude, direction and significance of the temporal trend of rainfall. It was applied for all the 3 regions of the basin (Figure 2). Basically, it describes the linear relationship (Equation 2) between multiple variables (Wilks, 2006). For a predictand \( y \) with multiple predictors \( x_1, x_2, \ldots x_n \) the regression equation (line) is described as follows:

\[
\hat{y} = a_0 + a_1x_1 + a_2x_2 + \cdots + a_nx_n
\]

(2)

In equation 2 \( n \) is the number of predictor variables and \( a_0 \) is the regression constant. When \( n = 1 \), equation 2 reduces to a simple linear regression line which is described as, \( y = a_0 + bx \) where \( b \) is the slope.
3.3.1.3 Percentile method

The percentile method was used to define the rainfall extremes. Similar to Ngailo et al. (2016) a heavy rainfall extreme was defined as one where the rainfall amount exceeded the 99th percentile of the distribution of MAM seasonal rainfall. On the other hand, a low rainfall extreme was defined as one where the rainfall amount was below the 20th percentile of the same distribution.

3.3.2 Assessing the skill of WRF model in simulating extreme rainfall under different parameterization settings

This section presents the model setup, the technique used for perturbing the parameterization schemes and the statistical tests used to verify the model simulations.

3.3.2.1 Model setup and experimental design

Simulation of extreme rainfall events that occurred over the LVB was done using the WRF modelling system, Advanced Research WRF (ARW) core version 3.9 built by NCAR. It is a grid point model with a terrain – following vertical coordinate near the surface which transforms to a constant pressure at the highest level. WRF performs the numerical integration on a staggered Arakawa C-grid with 3rd order Runge-Kutta time integration (Skamarock et al., 2008). The model was setup with 27 vertical levels ending at the 50 hPa isobaric level. Three domains (Table 1 and Figure 3) of varying grid resolution were used with a superior 2-way domain interaction and a nesting ratio of 1:3 which has been found to yield realistic results (Liu et al., 2011). The model was initialized using NCEP finalized reanalyzes starting 24 hours before the period of interest to allow time for model spin-up.

The domain setup was as follows: a coarse domain centered at 0.73° N latitude, 22° E longitude covering Africa to encompass large scale synoptic systems such as the ITCZ and sub-tropical anticyclones; a nest domain d02, covering the Congo basin and the western region of the Indian Ocean to cater for moisture inflow from the Congo air mass and the Indian Ocean respectively. The later contributes to diurnal thunderstorms events over Uganda (Jury, 2017); and a nest domain d03, covering LVB and Lake Victoria to cater for the influence of lake surface temperatures on rainfall in the basin (Sun et al., 2015). The ARW solver was used with a combination of physical parameterization: microphysical, cumulus, planetary boundary layer, surface layer and atmospheric radiation parameterization, coupled with static land-use data based on Moderate
Resolution Imaging Spectroradiometer (MODIS) with 21 land-use categories (Friedl et al., 2002) and a special lake surface representation. This was setup with varying resolution, the coarse domain used a 10-minute resolution while d02 and d03 used a 5-minute resolution.

To determine the model skill associated with the use of different cumulus – microphysical scheme combinations, the model setup was identical in all experiments, except for the cumulus and microphysical schemes (Table 2). Similar to Argent et al. (2015), the default settings of the WRF model were used as the control run. The domain setup was such that the inner most domain was at 12 km resolution (Table 1).

Table 1: Domain specifications used for investigating parameterization schemes

<table>
<thead>
<tr>
<th>Domain</th>
<th>Specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Coarse (outer most domain)</strong></td>
<td>108 km spacing</td>
</tr>
<tr>
<td></td>
<td>(84 x 84 grids)</td>
</tr>
<tr>
<td>d02</td>
<td>36 km spacing</td>
</tr>
<tr>
<td></td>
<td>(163 x 73 grids)</td>
</tr>
<tr>
<td>d03</td>
<td>12 km spacing</td>
</tr>
<tr>
<td></td>
<td>(67 x 52 grids)</td>
</tr>
</tbody>
</table>
Figure 3: Domain configuration for WRF simulations.

To focus the discussion on cumulus and microphysical parameterization, popular options for accompanying physical parameterization were selected for all experiments (Table 2). For the planetary boundary layer physics, the Yonsei University (YSU) scheme (Hong et al., 2006) was used. For atmospheric radiation physics, the Dudhia scheme (Dudhia, 1989) and Rapid Radiative Transfer Model (RRTM) (Mlawer et al., 1997) were used for shortwave and longwave radiation transfer respectively. For land surface physics, the NOAH land surface model (Chen & Dudhia, 2001a, 2001b) was used. For surface physics, the MM5 soil temperature thermal diffusion model (Dudhia, 1996) was used.
<table>
<thead>
<tr>
<th>Experiment</th>
<th>Cumulus Scheme</th>
<th>Microphysics scheme</th>
<th>PBL scheme</th>
<th>Radiation (short-wave)</th>
<th>Radiation (long wave)</th>
<th>Land surface model</th>
<th>Surface scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control run</td>
<td>KF</td>
<td>WSM3</td>
<td>YSU</td>
<td>Dudhia</td>
<td>RRTM</td>
<td>NOAH</td>
<td>MM5</td>
</tr>
<tr>
<td>E1</td>
<td>GF</td>
<td>Eta</td>
<td>YSU</td>
<td>Dudhia</td>
<td>RRTM</td>
<td>NOAH</td>
<td>MM5</td>
</tr>
<tr>
<td>E2</td>
<td>GF</td>
<td>Milbrandt</td>
<td>YSU</td>
<td>Dudhia</td>
<td>RRTM</td>
<td>NOAH</td>
<td>MM5</td>
</tr>
<tr>
<td>E3</td>
<td>GF</td>
<td>SBU_YLin</td>
<td>YSU</td>
<td>Dudhia</td>
<td>RRTM</td>
<td>NOAH</td>
<td>MM5</td>
</tr>
<tr>
<td>E4</td>
<td>Grell 3D</td>
<td>Eta</td>
<td>YSU</td>
<td>Dudhia</td>
<td>RRTM</td>
<td>NOAH</td>
<td>MM5</td>
</tr>
<tr>
<td>E5</td>
<td>Grell 3D</td>
<td>Milbrandt</td>
<td>YSU</td>
<td>Dudhia</td>
<td>RRTM</td>
<td>NOAH</td>
<td>MM5</td>
</tr>
<tr>
<td>E6</td>
<td>Grell 3D</td>
<td>SBU_YLin</td>
<td>YSU</td>
<td>Dudhia</td>
<td>RRTM</td>
<td>NOAH</td>
<td>MM5</td>
</tr>
<tr>
<td>E7</td>
<td>BMJ</td>
<td>Eta</td>
<td>YSU</td>
<td>Dudhia</td>
<td>RRTM</td>
<td>NOAH</td>
<td>MM5</td>
</tr>
<tr>
<td>E8</td>
<td>BMJ</td>
<td>Milbrandt</td>
<td>YSU</td>
<td>Dudhia</td>
<td>RRTM</td>
<td>NOAH</td>
<td>MM5</td>
</tr>
<tr>
<td>E9</td>
<td>BMJ</td>
<td>SBU_YLin</td>
<td>YSU</td>
<td>Dudhia</td>
<td>RRTM</td>
<td>NOAH</td>
<td>MM5</td>
</tr>
</tbody>
</table>
3.3.2.2 Analysis of model output
The NCAR command language tool was used to visualize the observed and simulated rainfall data and to calculate the relevant statistics. The following methods were applied.

Difference variable
Maps were generated to facilitate a discussion, in spatial terms, of how simulated rainfall compares to observed rainfall. This was based on the difference variable (Willmott et al., 1985) which is described as follows:

\[
\text{Difference} = \text{Simulated} - \text{Observed}
\]  

A positive result is obtained when the model overestimates the observed rainfall and a negative result is obtained when the model underestimates the observed rainfall.

Accuracy scores
The Root Mean Square Error (RMSE) (Chai & Draxler, 2014) was used to determine model accuracy, and the Mean Error (ME) (Mugume et al., 2016a) was used to determine model bias. RMSE (Equation 4) compares how close simulated values are to observed values. A low RMSE value indicates high accuracy. ME (Equation 5) gives the direction and magnitude of model bias. A dispersion from zero ‘0’ in the negative direction indicates underestimation, while that in the positive direction indicates overestimation, and the zero ‘0’ value indicates absence of bias.

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n}(S_i - O_i)^2}
\]  

\[
ME = \frac{1}{n} \sum_{i=1}^{n}(S_i - O_i)
\]

In equations 4 and 5, \( S \) is the simulated value, \( O \) is the observed value, and \( n \) is the total number of comparisons.

Categorical scores
The scores used were the Probability of Detection (POD) and False Alarm Ratio (FAR). POD (Equation 6) indicates the ability of the WRF model to correctly simulate the occurrence of observed rainfall events. A high POD value indicates high detection ability. FAR (Equation 7)
indicates the possibility of the model to simulate rain events on days which, according to the observations, did not to have any rain. A low FAR value indicates high model accuracy. POD and FAR values range from 0 to 1 and they are derived from a 2x2 contingency table (Table 3) as described by Schaefer (1990).

**Table 3: 2x2 contingency table**

<table>
<thead>
<tr>
<th>Observations / WRF</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hit (A)</td>
<td>Miss (C)</td>
</tr>
<tr>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>False Alarm (B)</td>
<td>Correct Negative (D)</td>
</tr>
</tbody>
</table>

\[
POD = \frac{A}{A+C}
\]

\[
FAR = \frac{B}{A+B}
\]

**Rank Transformation**

To identify experiments that generated the overall best and worst skill, the present study used a simple rank transformation (Conover & Iman, 1981) to compare the skill scores of the experiments and thus show their relative performance. For RMSE, ME and FAR, ranks were assigned from the smallest value to the largest while for POD, the rank was assigned from the largest value to the smallest. Also, for ME values, only the magnitude was used for the ranking.

**Testing hypotheses**

The one-sample \( t \) test (Wilks, 2006) was used to test the hypothesis for this 2\(^{nd} \) objective (refer to page 4, sections 1.4 and 1.5). The \( t \) value (Equation 8) is used to compare the mean, \( M \) of a sample dataset, drawn from a population with a specified mean, \( X_0 \). This value is controlled by a parameter, \( \rho \) known as the degrees of freedom and its obtained as \( \rho = m - 1 \) where \( m \) is the number of independent observations in the sample dataset \( M \).

\[
t = \frac{M - X_0}{[Variance(M)]^{1/2}}
\]
3.3.3 Determining the effect of changing grid resolution on WRF model skill in simulating extreme rainfall

This section presents the model setup used to study the effect of model resolution and domain size. To study the effect of changing grid resolution, experiments done in objective 2 (Table 2, section 3.3.2.1) were repeated with the inner most domain set to 4 km resolution (Table 4). A comparison was then done between results of experiments at 12 km resolution (objective 2) and experiments at 4 km resolution (objective 3). Furthermore, to reveal the effect of changing domain size, a comparison was done between results of domain, d03 in objective 2 and domain, d02 in objective 3. Both domains were at 12 km resolution and only differed in size. The model setup used was similar to that described in section 3.3.2.1 except for the model resolution, which was changed, and, the methods used to analyze the model output were similar to those described in section 3.3.2.2 except for the technique used to test the hypotheses. Here, a paired sample $t$ test was used.

<table>
<thead>
<tr>
<th>Domain / Objective</th>
<th>Objective 2</th>
<th>Objective 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coarse</td>
<td>108 km spacing (84 x 84 grids)</td>
<td>36 km spacing (250 x 250 grids)</td>
</tr>
<tr>
<td>d02</td>
<td>36 km spacing (163 x 73 grids)</td>
<td>12 km spacing (490 x 220 grids)</td>
</tr>
<tr>
<td>d03</td>
<td>12 km spacing (67 x 52 grids)</td>
<td>4 km spacing (202 x 151 grids)</td>
</tr>
</tbody>
</table>

Testing hypotheses

The paired sample $t$ test (Wilks, 2006) was used to test the hypothesis for this 3$^{rd}$ objective (refer to page 4, sections 1.4 and 1.5). For two sample datasets $y_1$ and $y_2$ with equal number of observations, the $t$ value (Equation 9) is obtained as:

$$ t = \frac{\overline{D}}{\text{Standard Error}(D)} $$  \hspace{1cm} (9)

Where $\overline{D}$ is the mean of the difference ($D$) between the individual observations, $D = y_1 - y_2$. 

21
4.1 Observed spatial and temporal patterns of MAM season rainfall

This section presents results on the spatial and temporal patterns of observed rainfall across LVB during the MAM seasons of 10 years, that is, 2008 to 2017.

4.1.1 Rainfall accumulation

Overall, LVB received varying amounts of rainfall across all the 10 MAM seasons (Figure 4). The highest amount was received in 2008 whereas the lowest amount was received in 2010. For each season, there were notable differences in the amount of rainfall received in the 3 regions of the basin. Western LVB received rainfall between 190 and 380 mm with the highest amount received in 2013 and the lowest amount in 2014. Central LVB received rainfall between 210 and 480 mm with the highest amount received in 2017 and the lowest amount in 2010. Eastern LVB received rainfall between 320 and 600 mm with the highest amount received in 2008 and the lowest amount in 2012. Albeit these variations, eastern LVB received the highest overall amount of rainfall whereas western LVB received the lowest overall amount. Since the MAM season of 2008 received the highest amount of rainfall, the case study period investigated in the WRF model was selected from this season. Details of this period are presented in section 4.1.6.

![Figure 4: Total MAM season rainfall amount for each year.](image-url)
4.1.2 Rainfall anomaly

Figure 5 shows the spatial rainfall anomaly for each of the MAM seasons investigated. The anomalies ranged from -2.5 to just under 2.0 mm/day over the land. Central LVB had relatively large anomalies (both positive and negative) compared to eastern and western LVB.


Of the 2 years, 2008 and 2017, that were dominated by above normal rainfall, a keen interest was placed on 2008 because the above normal rainfall was mostly over the land area unlike in 2017 where it was mostly over the lake. Mainly, it is rainfall over the land area that affects human livelihood in the LVB. For this reason, the case study period (Section 4.1.6) investigated within the WRF model was selected in 2008.
Figure 5: Normalized spatial rainfall anomaly. Letters (a) to (j) correspond to the MAM seasons of 2008 to 2017 respectively.
4.1.3 Daily evolution of MAM season rainfall

Figure 6 shows the daily of evolution of observed rainfall over LVB based on time and area averaged values. Overall, rainfall received over LVB was ranging from 2.0 mm to just under 12.0 mm in a day and 3 rainfall peaks were observed for each of the regions. These were toward the end of March, in mid-April and early May. Most importantly, for each of the regions, the dekad toward the end of March received dominantly heavy rainfall that was unrivaled by any other dekad in the season. For this reason, the case study period (Section 4.1.6) selected for investigation within the WRF model partly covered this period.
Figure 6: Daily evolution of MAM season rainfall for (a) western LVB, (b) central LVB and (c) eastern LVB.
4.1.4 Trend of daily rainfall

Figure 7 shows the trend of MAM season daily rainfall from 2008 to 2017. Much as there were notable variations in the amount of daily rainfall received in the MAM season of each year, it did not have a significant trend ($p > 0.05$) in each of the regions of LVB.

4.1.5 Extreme rainfall cases

Figure 7 was also fitted with the percentile demarcations that were used to identify the rainfall extremes. For each region of LVB, the 99th and 20th percentile were calculated and considered as thresholds for identifying the extremely high and low values respectively. For western, central and eastern LVB, the threshold for extremely high values (99th percentile) was 24.0 mm, 26.8 mm, and 29.7 mm respectively while the threshold for extremely low values (20th percentile) was 0.2 mm, 0.3 mm and 0.2 mm respectively. The peak over and peak below threshold approach was then used to identify the extremely high and extremely low rainfall amounts respectively.

A total of 25 cases of heavy rainfall extremes and 583 cases of low rainfall extremes were identified in the MAM season for the period 2008 to 2017 (Figure 7 and Table 5). Generally, heavy rainfall extremes over LVB were ranging from 24.9 mm (the lowest amount) to 48.8 mm (the highest amount). On the other hand, the low rainfall extremes were ranging from 0.0 mm (the lowest amount) to 0.2 mm (the highest amount).

For heavy rainfall extremes, western and central LVB had 8 cases whereas eastern LVB had 9 cases, which was the highest count (Table 5). Eastern LVB had extremes with the highest rainfall amounts. This was followed by central LVB and then western LVB whose extremes had the least rainfall amounts. In each region of the basin, the extreme rainfall cases were spread out in all the months of the season, that is, March, April and May. However, unlike in eastern LVB, in western and central LVB 80% of the cases occurred in the month of March. On the other hand, for low rainfall extremes, western, central and eastern LVB had 209, 180 and 194 cases respectively (Table 5).
Figure 7: Trend of MAM season daily rainfall over (a) western LVB, (b) central LVB, and (c) eastern LVB. The black curvy segmented line is the regression line.
Table 5: Extreme rainfall events in the Lake Victoria Basin during the MAM season

<table>
<thead>
<tr>
<th>Region of LVB</th>
<th>Total number of heavy rainfall extremes</th>
<th>Total number of low rainfall extremes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Western</td>
<td>8</td>
<td>209</td>
</tr>
<tr>
<td>Central</td>
<td>8</td>
<td>180</td>
</tr>
<tr>
<td>Eastern</td>
<td>9</td>
<td>194</td>
</tr>
</tbody>
</table>

4.1.6 Case study period

The case study period for investigation within the WRF model was selected to be a 20-day period, 17th March 2008 to 5th April 2008. As earlier highlighted, the period was selected in the MAM season of 2008 because this season received the highest overall amount of rainfall (Section 4.1.1) with most of the rainfall received over the land area (Section 4.1.2). Furthermore, the period was selected to cover the end of March because the seasonal evolution (Section 4.1.3) indicated that this period received the highest total rainfall than any other period in the MAM season. It’s no wonder that several heavy rainfall extremes were observed in this period (Section 4.1.5). The period was also extended to cover early April because those days were observed to have received extremely low rainfall (Appendix 1).
4.2 WRF simulations of extreme rainfall under different parameterization settings

This section presents results on how WRF simulations compare to the observed rainfall.

4.2.1 Observed and simulated spatial rainfall patterns

Figure 8 shows the rainfall distribution from TRMM observations and the WRF model runs for the 20-day period. The TRMM observations (Figure 8a) show that varying amounts of rainfall were received in the different regions of LVB. Overall, the total rainfall amount received in LVB during this period was between 100 and 280 mm although rainfall amounts between 120 to 200 mm dominated the area. In western and central LVB, there were patches that received rainfall in the range of 200 to 280 mm. Also, eastern LVB had an area where rainfall in the range of 100 to 120 mm was received. Notably, central LVB received a higher rainfall amount compared to the other 2 regions.

Overall, the model runs (Figure 8, b to k) generated varying amounts of rainfall in the different regions of LVB. Majority of the model runs generated high rainfall amounts in the high-altitude areas, such as the southwestern corner of the basin and the Mt. Rwenzori and Elgon ranges. On the other hand, majority of the runs also gave low rainfall amounts in the low altitude areas of the basin, such as in the watershed neighboring the lake.

The control run (Figure 8b) generated a rainfall distribution in which the total rainfall ranged from 0 to over 320 mm. Eastern and central LVB were dominated by rainfall amounts in the range of 0 to 120 mm. Western LVB, however, had higher rainfall amounts, ranging from 60 to 280 mm. It is easily noticeable that the control run did not replicate the observed pattern satisfactorily. This was the motivation to explore other parameterization options available in the WRF model.
Figure 8: Total rainfall amount (mm) from TRMM data (a), the control run (b) and experimental runs E1 to E9 (c to k). The plots c to k are also arranged following the cumulus schemes (rows) and microphysical schemes (columns) used.
Overall, varying rainfall distribution patterns were generated when different cumulus and microphysical parameterization combinations were used (Figure 8, c to k). All model runs generated higher rainfall amounts in western LVB compared to the other two regions. This was contrary to the TRMM data which showed that the highest rainfall amounts were received in central LVB.

Model runs done with the SBU_YLin microphysical scheme (Figure 8 e, h, & k) generated higher rainfall amounts over LVB as compared to those done with other microphysical schemes. The rainfall amounts ranged from 30 to over 320 mm. Runs done with the Milbrandt microphysical scheme (Figure 8 d, g, & j) and those done with the Eta scheme (Figure 8 c, f, & i) generated rainfall in the range of 10 to over 320 mm.

In comparison to other model runs, those done with the Grell 3D cumulus scheme (Figure 8 f, g, & h) generated lower rainfall amounts in the southwestern corner of the basin. All the cumulus schemes, however, generated rainfall in the range of 10 to over 320 mm.

### 4.2.2 Comparison between observed and simulated rainfall amounts

Figure 9 shows the difference between simulated and observed rainfall for the 20-day period. Overall, majority of the model runs showed a disagreement with the observed rainfall in most parts of LVB. However, this disagreement was more pronounced in eastern and central LVB than it was in western LVB. There was dominant underestimation of rainfall in LVB, except for run E9 which overestimated rainfall most especially in western LVB. The best spatial representation was generated by model runs E6 and E2 whereas the control run, E5 and E9 generated the worst.

All model runs underestimated the rainfall in the watershed closet to the lake. Except for the runs done with the Grell 3D cumulus scheme, the other runs overestimated rainfall in the highland areas in the southwestern corner of the basin. Of the areas in which the model runs show close agreement, the error ranged from 25 to 50 mm and very few areas had an error below 25 mm.
Figure 9: Difference in total rainfall amount (mm) between WRF model output and TRMM data for the 20-day period.
4.2.3 Observed and simulated temporal rainfall patterns

Figure 10 shows the rainfall accumulation in each region of the basin during the 20-day period. Overall, the TRMM data had rainfall of about 200 mm, and the model runs simulated varying amounts of rainfall in each region of LVB.

In western LVB (Figure 10a), all model runs simulated the general pattern of TRMM rainfall, however, there were variations and some model runs were closer to the TRMM data than others in different days. For example, in the first 11 days, model run E3 generated rainfall that was closest to the TRMM data. From day 9 to day 10, model run E9 generated a rainfall amount similar to the TRMM data while for the last 5 days that were considerably dry, model runs E1, E5, and E8 simulated a pattern that was closest to the TRMM data.

In central LVB (Figure 10b), all model runs simulated the overall TRMM rainfall pattern for the first 5 days, however, model run E3 was closer to the TRMM data than the others. From day 6 to day 14, all the runs underestimated the TRMM rainfall. This underestimation agrees with that observed in the spatial patterns (Figure 9). Notably, in the last 5 dry days (day 15 to day 20) model runs E1, E5 and E8 simulated a pattern that was very close to that in the TRMM data.

In eastern LVB (Figure 10c), all model runs underestimated the TRMM rainfall from day 1 to day 13. This underestimation agrees with that observed in the spatial patterns (Figure 9). However, for the dry days (day 15 to day 20) the control run, E1, E5, E7 and E8 simulated a pattern that was similar to the TRMM data.
Figure 10: Temporal rainfall accumulation (mm) in (a) western LVB, (b) central LVB and (c) eastern LVB for the 20-day period.
4.2.4 Accuracy and categorical skill scores

Table 6 summarizes the skill scores generated by each model run, along with their respective ranks. The root mean square error (RMSE) of the model runs was between 109 and 226. Model run E6 gave the least RMSE of 109.26 and thus ranked 1st while run E9 gave the largest RMSE of 225.75 and thus ranked 10th. Model runs done using the Grell 3D cumulus scheme (E4, E5, & E6) generated the least RMSE compared to the other cumulus schemes.

Overall, the mean error (ME) generated by the model runs was between 15 and 100 mm in magnitude. Majority of the model runs generated a negative ME. Only model runs E3 and E9 generated a positive ME. Considering the magnitude of the ME, model run E6 generated the least value and thus ranked 1st while run E9 generated the largest value and thus ranked 10th.

The probability of detection (POD) generated by the models ranged from 0.86 to 1. Model runs E2 and E4 generated the highest POD thus ranked 1st while the control run generated the least POD and thus ranked 10th. The false alarm ratio (FAR) generated by the model runs was between 0.37 and 0.42. Model run E5 generated the least FAR whereas runs E3, E4, E6 and E9 generated the highest FAR.

At a 95% level of confidence, changing the microphysical and cumulus parameterization schemes (each experiment used different parameterization settings as shown in Table 2) significantly affected the ME (p < 0.05), POD (p < 0.05), and FAR (p < 0.05) whereas, it did not significantly affect the RMSE (p > 0.05).
Table 6: Skill scores and ranks of the model runs.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Accuracy scores</th>
<th>Categorical scores</th>
<th>Ranks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE</td>
<td>ME (mm)</td>
<td>POD</td>
</tr>
<tr>
<td>Control run</td>
<td>137.21</td>
<td>-60.18</td>
<td>0.86</td>
</tr>
<tr>
<td>E1</td>
<td>141.85</td>
<td>-32.53</td>
<td>0.94</td>
</tr>
<tr>
<td>E2</td>
<td>150.98</td>
<td>-16.81</td>
<td>1</td>
</tr>
<tr>
<td>E3</td>
<td>192.81</td>
<td>47.34</td>
<td>0.99</td>
</tr>
<tr>
<td>E4</td>
<td>117.77</td>
<td>-57.43</td>
<td>1</td>
</tr>
<tr>
<td>E5</td>
<td>125.60</td>
<td>-82.79</td>
<td>0.97</td>
</tr>
<tr>
<td>E6</td>
<td>109.26</td>
<td>-15.71</td>
<td>0.98</td>
</tr>
<tr>
<td>E7</td>
<td>144.12</td>
<td>-36.13</td>
<td>0.91</td>
</tr>
<tr>
<td>E8</td>
<td>147.36</td>
<td>-26.51</td>
<td>0.94</td>
</tr>
<tr>
<td>E9</td>
<td>225.75</td>
<td>99.09</td>
<td>0.92</td>
</tr>
</tbody>
</table>
4.2.5 Diurnal evolution of observed and simulated rainfall

Figure 11 shows the diurnal evolution of rainfall during the 20-day period based on 3-hourly rainfall accumulations. Generally, there was variation in the amount of both TRMM observed and WRF simulated rainfall during different times of the day, and this was evident across all the 3 regions of LVB. Furthermore, the WRF model runs generally generated majority of the rainfall during day time, that is, between 9 am (09:00 hours) and 6 pm (18:00 hours).

In western LVB (Figure 11a), the TRMM data shows that majority of the rainfall in this region was received in the evening and night time between 6 pm (18:00 hours) and 6 am (06:00 hours) of the following day. On the other hand, day time, that is, from 8 am (08:00 hours) to 5 pm (17:00 hours) received lower rainfall amounts. Contrary to the TRMM observations, all the WRF model runs generated most rainfall during day, that is, from 11 am (11:00 hours) to 6 pm (18:00 hours).

In central LVB (Figure 11b), none of the model runs simulated a diurnal rainfall pattern similar to that of the TRMM data. The TRMM observations show that majority of the rainfall was received between midnight (24:00 hours / 00:00 hours) and 6 am (06:00 hours) whereas the time between 9 am (09:00 hours) and 9 pm (21:00 hours) received lower amounts of rainfall. The WRF model runs generated most rainfall between 9 am (09:00 hours) in the morning and 3 pm (15:00 hours) in the afternoon. During the evening and night time, that is, from 6 pm (18:00 hours) to 6 am (06:00 hours), the model runs generally generated lower rainfall amounts and thus severely underestimated the TRMM data. This was with the exception of model run E3 which generated low rainfall amounts starting from 9 pm (21:00 hours).

In eastern LVB (Figure 11c), the TRMM observations show that majority of the rainfall was received in the evening and night time, that is, from 6 pm (18:00 hours) to 3 am (03:00 hours) the following day. Generally, except for model run E9, the rest of the runs simulated the observed diurnal rainfall pattern quite well for the first 15 hours of the day. Thereafter, all the runs severely underestimated the rainfall for the rest of the day.

In western LVB, WRF simulated rainfall compared quite well to the TRMM observations whereas in central and eastern LVB, there was a large disparity in the timing and amount of rainfall. The WRF model was incapable of simulating nocturnal rainfall.
Figure 11: 3-hourly rainfall accumulation (mm) of TRMM data and WRF simulations. (a), (b) and (c) correspond to western LVB, central LVB and eastern LVB respectively.
4.3 Effect of grid resolution on WRF model skill in simulating extreme rainfall

This section presents results on the effect of grid resolution on WRF model skill in simulating extreme rainfall.

4.3.1 Effect of model resolution on the skill scores

Each of the 10 model runs done at 12 km resolution were redone at a 4 km resolution, and the skill scores they generated are summarized in Table 7 alongside the scores they generated at the 12 km resolution. Changing model resolution significantly affected the RMSE (p < 0.05), the POD (p < 0.05), and FAR (p < 0.05) whereas the ME was not significantly affected (p > 0.05). Majority of the model runs generated lower RMSE values at the 4 km resolution compared to the 12 km resolution. Only model run E6 generated a higher RMSE at the 4 km resolution. For POD, majority of the model runs generated lower POD values at 4 km resolution than at 12 km resolution. Only model run E5 generated the same POD whereas model runs E7 and E9 generated higher POD values at 4 km resolution. For FAR, majority of the model runs generated higher FAR values at 4 km resolution compared to the 12 km resolution. Only model runs E1 and E6 generated the same FAR. None of the runs generated lower FAR at 4 km resolution.

It was also important to know the ranking of the scores of the 4 km model runs (Table 8) to enable a comparison with the ranking of the 12 km model runs (Table 6, section 4.2.4). For RMSE, at the 4 km resolution, model run E4 ranked 1st, contrary to the 12 km resolution where run E6 ranked 1st. Run E9 ranked 10th in RMSE at both resolutions. For ME, at the 4 km resolution, run E3 ranked 1st, contrary to the 12 km resolution where run E6 ranked 1st. Similarly, the control run ranked 10th, contrary to the 12 km resolution where run E9 ranked 10th. For POD, at the 4 km resolution, run E5 ranked 1st, contrary to the 12 km resolution where runs E2 and E4 ranked 1st. The Control run ranked 10th at both resolutions. For FAR, at the 4 km resolution, run E5 ranked 1st at both resolutions whereas runs E3, E4 and E9 ranked worst at both resolutions. This change in performance rankings of model runs in response to a change in model resolution triggered a further exploration into the spatial and temporal patterns of the rainfall generated in order to show how it relates to the change in the skill scores. For this purpose, 4 model runs were used; runs E4 and E6 to represent the best performing experiments and runs E9 and the control run to represent the worst performing experiments.
### Table 7: Skill scores generated by the 12 km and 4 km model runs

<table>
<thead>
<tr>
<th>Experiment</th>
<th>RMSE</th>
<th>ME</th>
<th>POD</th>
<th>FAR</th>
<th>RMSE</th>
<th>ME</th>
<th>POD</th>
<th>FAR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>12 km</td>
<td>4 km</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Model resolution</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control run</td>
<td>137.21</td>
<td>126.40</td>
<td>-60.18</td>
<td>0.86</td>
<td>0.39</td>
<td>-86.93</td>
<td>0.82</td>
<td>0.41</td>
</tr>
<tr>
<td>E1</td>
<td>141.85</td>
<td>120.46</td>
<td>-32.53</td>
<td>0.94</td>
<td>0.39</td>
<td>-86.43</td>
<td>0.87</td>
<td>0.39</td>
</tr>
<tr>
<td>E2</td>
<td>150.98</td>
<td>104.42</td>
<td>-16.81</td>
<td>1.00</td>
<td>0.40</td>
<td>-23.99</td>
<td>0.95</td>
<td>0.42</td>
</tr>
<tr>
<td>E3</td>
<td>192.81</td>
<td>129.57</td>
<td>47.34</td>
<td>0.99</td>
<td>0.41</td>
<td>-0.51</td>
<td>0.96</td>
<td>0.42</td>
</tr>
<tr>
<td>E4</td>
<td>117.77</td>
<td>100.64</td>
<td>-57.43</td>
<td>0.98</td>
<td>0.41</td>
<td>-29.88</td>
<td>0.94</td>
<td>0.42</td>
</tr>
<tr>
<td>E5</td>
<td>125.60</td>
<td>108.77</td>
<td>-82.79</td>
<td>0.97</td>
<td>0.38</td>
<td>-64.34</td>
<td>0.97</td>
<td>0.39</td>
</tr>
<tr>
<td>E6</td>
<td>109.26</td>
<td>113.39</td>
<td>-15.71</td>
<td>0.98</td>
<td>0.41</td>
<td>3.24</td>
<td>0.96</td>
<td>0.41</td>
</tr>
<tr>
<td>E7</td>
<td>144.12</td>
<td>114.19</td>
<td>-36.13</td>
<td>0.91</td>
<td>0.39</td>
<td>-42.47</td>
<td>0.93</td>
<td>0.41</td>
</tr>
<tr>
<td>E8</td>
<td>147.36</td>
<td>123.90</td>
<td>-26.51</td>
<td>0.94</td>
<td>0.39</td>
<td>-69.86</td>
<td>0.93</td>
<td>0.41</td>
</tr>
<tr>
<td>E9</td>
<td>225.75</td>
<td>162.53</td>
<td>99.09</td>
<td>0.92</td>
<td>0.41</td>
<td>69.08</td>
<td>0.94</td>
<td>0.42</td>
</tr>
</tbody>
</table>

### Table 8: Ranks of the model skill scores for runs done at 4 km resolution

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Accuracy scores</th>
<th>Categorical scores</th>
<th>Ranks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE</td>
<td>ME (mm)</td>
<td>POD</td>
</tr>
<tr>
<td><strong>Control run</strong></td>
<td>126.40</td>
<td>-86.93</td>
<td>0.82</td>
</tr>
<tr>
<td>E1</td>
<td>120.46</td>
<td>-86.43</td>
<td>0.87</td>
</tr>
<tr>
<td>E2</td>
<td>104.42</td>
<td>-23.99</td>
<td>0.95</td>
</tr>
<tr>
<td>E3</td>
<td>129.57</td>
<td>-0.51</td>
<td>0.96</td>
</tr>
<tr>
<td>E4</td>
<td>100.64</td>
<td>-29.88</td>
<td>0.94</td>
</tr>
<tr>
<td>E5</td>
<td>108.77</td>
<td>-64.34</td>
<td>0.97</td>
</tr>
<tr>
<td>E6</td>
<td>113.39</td>
<td>3.24</td>
<td>0.96</td>
</tr>
<tr>
<td>E7</td>
<td>114.19</td>
<td>-42.47</td>
<td>0.93</td>
</tr>
<tr>
<td>E8</td>
<td>123.90</td>
<td>-69.86</td>
<td>0.93</td>
</tr>
<tr>
<td>E9</td>
<td>162.53</td>
<td>69.08</td>
<td>0.94</td>
</tr>
</tbody>
</table>
4.3.2 Effect of model resolution on the spatial rainfall patterns

Overall, there was an improved level of detail in the simulated rainfall for the 4 km runs as compared to the 12 km runs (Figure 12). For the 2 best performing runs, that is runs E4 and E6, in western LVB, both runs generated slightly more rainfall at the 4 km resolution. Both had small areas in which the simulated rainfall exceeded 320 mm which was unlike that at 12 km resolution. In central LVB, run E4 generated considerably more rainfall at 4 km, a larger area had simulated rainfall between 120 and 200 mm whereas run E6 generated slightly less rainfall at 4 km.

In eastern LVB, run E4 generated considerably more rainfall at 4 km, the area was dominated by rainfall ranging from 60 to 280 mm contrary to the 12 km resolution where it was dominated by rainfall between 10 and 120 mm. Run E6 generated slightly less rainfall at 4 km than at 12 km, there were larger areas with rainfall less than 120 mm. Albeit these differences in the simulated rainfall (Figure 12), run E6 still generated a spatial rainfall pattern that was closer to the TRMM observations than that generated by run E4. However, run E4 made a considerable improvement in the amount and spatial distribution of rainfall at the 4 km resolution.

For the 2 worst performing runs, that is runs E9 and the control, in western LVB, run E9 generated slightly less rainfall at 4 km as compared to the 12 km resolution. However, in central LVB, it generated more rainfall at the 4 km resolution. The area next to the lake shore which received rainfall below 120 mm was considerably reduced in the 4 km run. In eastern LVB, run E9 generated less rainfall at the 4 km resolution. There were larger areas which received below 120 mm of rainfall. This was contrary to the pattern generated at 12 km resolution which showed that rainfall in this area was dominantly above 120 mm. The control run generated considerably less rainfall at 4 km resolution than at 12 km. This was the case in all the 3 regions of LVB.

4.3.3 Effect of model resolution on the temporal rainfall patterns

The temporal rainfall accumulations at the different resolutions (Figure 13) agree with the spatial patterns (Figure 12). Overall, the best performing runs, that is E4 and E6, generated similar temporal rainfall patterns at both resolutions, however, with notable deviations (which correspond to the spatial changes shown in Figure 12) and this was for all 3 regions of LVB.
On the other hand, for the worst performing runs, run E9 generated similar temporal patterns in western and eastern LVB at both resolutions, however, in central LVB, it generated more rainfall at the 4 km resolution. This was from about day 7 till the last day. The control run generated considerably less rainfall at the 4 km resolution compared to the 12 km resolution. This disparity started in the first week and became more evident toward the last day and this was the case in all 3 regions of LVB. This finding corresponds to the increase in the magnitude of the ME value at 4 km resolution. Other model runs which also exhibit an increase in the ME value at 4 km can be expected to have similar temporal accumulation. These runs are; E1, E2, E7 and E8.
Figure 12: Spatial rainfall accumulation (mm) for TRMM data and model runs E4, E6, E9 and the control at 12 km and 4 km resolutions.
Figure 13: Temporal rainfall accumulation for TRMM data and model runs E4, E6, E9 and the Control. Plots (a), (b) and (c) are for western LVB, central LVB and eastern LVB respectively.
### 4.3.4 Effect of domain size on the skill scores

The nature of the experimental setup allowed for a further exploration into the effect of changing the size of the domain. For this particular investigation, the skill scores generated by model runs at the 12 km resolution (done in objective 2) were compared to the skill scores generated by domain 2 of the high-resolution runs (done in objective 3). This domain was also at 12 km resolution, it only differed in size (it was a larger domain).

The results (Table 9) showed that changing the size of the domain significantly affected the RMSE ($p < 0.05$) and the FAR ($p < 0.05$), while the ME and POD were both not significantly affected ($p > 0.05$). Majority of the model runs generated lower RMSE values when a large domain was used. Only run E6 contradicted this and generated a higher RMSE value. For FAR, majority of the runs generated lower FAR values when a small domain was used. However, model runs E3, E6 and E9 generated the same FAR value for both domain sizes. None of the runs generated a lower FAR value when a large domain was used.

<table>
<thead>
<tr>
<th>Domain size</th>
<th>Small domain (12 km)</th>
<th>Large domain (12 km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment</td>
<td>RMSEs</td>
<td>MEs</td>
</tr>
<tr>
<td>Control run</td>
<td>137.21</td>
<td>-60.18</td>
</tr>
<tr>
<td>E1</td>
<td>141.85</td>
<td>-32.53</td>
</tr>
<tr>
<td>E2</td>
<td>150.98</td>
<td>-16.81</td>
</tr>
<tr>
<td>E3</td>
<td>192.81</td>
<td>47.34</td>
</tr>
<tr>
<td>E4</td>
<td>117.77</td>
<td>-57.43</td>
</tr>
<tr>
<td>E5</td>
<td>125.60</td>
<td>-82.79</td>
</tr>
<tr>
<td>E6</td>
<td>109.26</td>
<td>-15.71</td>
</tr>
<tr>
<td>E7</td>
<td>144.12</td>
<td>-36.13</td>
</tr>
<tr>
<td>E8</td>
<td>147.36</td>
<td>-26.51</td>
</tr>
<tr>
<td>E9</td>
<td>225.75</td>
<td>99.09</td>
</tr>
</tbody>
</table>
CHAPTER FIVE: DISCUSSION OF RESULTS

5.1 Spatial and temporal patterns of MAM season rainfall

Receipt of rainfall in LVB during the March to May season is a consequence of the presence of the ITCZ, a low-pressure region. During this season, the Mascarene anticyclone steers south-easterly moist winds from the Indian ocean to converge in this region, giving rise to rainfall (Awange et al., 2013). However, the variation in the amount of rainfall received in each of the 3 regions is explained by the local geography. For example, Basalirwa (1995) and Nicholson (2017) agree that the presence of a very large lake, that is, Lake Victoria, is a strong influencer of the climatology of the basin. This influence is experienced in form of the land-/lake-breeze circulation which is responsible for transporting vast amounts of moisture and thus drives convective activity in the watershed closest to the lake. This effect is most intense in eastern LVB, that is why it received the largest overall amount of rainfall.

The anomalously wet conditions experienced in 2008, could be attributed to the La Nina phase of ENSO which prevailed that year. As discussed by Camberlin and Philippon (2002), cool sea surface temperatures over the equatorial pacific are associated with low sea level pressure over the northern Indian Ocean, and a resultant northward shift of the ITCZ causing it to be positioned over Uganda earlier than it should. This accounts for the above normal rainfall receipts.

The cause for the anomalously dry conditions in the MAM seasons of 2010 and 2011 is not fully known, however, Lyon and Dewitt (2012) and Vigaud et al. (2016) associate it with sudden changes in sea surface temperatures most especially in the tropical Pacific ocean. Additionally, Camberlin and Okoola (2003) acknowledge that this suppressed rainfall can be related to a delayed onset of the MAM rains which happens when the meridional arm of the ITCZ is held further west, out of East Africa by strong equatorial easterly winds that are generated by sea level pressure anomalies in the south Atlantic and Indian oceans.

The lack of a trend in the daily rainfall events implies that the overall pattern of daily events in all the MAM seasons is the same. The season starts out with moderate rainfall amounts which peak in the middle of the season and gradually fall toward the end. This cycle is generally the same for all the MAM seasons, so when plotted side-by-side, there is no meaningful trend that can be drawn.
Daily events can however show a trend when individual seasons are considered. To demonstrate this, the wettest (2008) and driest (2010) MAM seasons were used (Appendix 2). Overall, there was a general decreasing trend of rainfall.

5.2 WRF simulations of extreme rainfall under different parameterization settings

The results show that changing the cumulus and microphysical parameterization schemes from the default affected the model’s bias, detection ability and the false alarm ratio. Overall, model runs done with the Grell 3D cumulus scheme showed closer proximity to the TRMM observations. Notably, model run E6 done with the Grell 3D cumulus scheme combined with the SBU_YLin microphysical scheme showed the most satisfactory skill and generated the best spatial representation of the rainfall. Furthermore, it outperformed the best ranking model runs suggested by Sun et al. (2014) and Argent et al. (2015) for WRF simulations of rainfall over LVB.

The variation in simulated rainfall distribution following a change in the microphysical parameterization scheme used shows that it is relevant to include these schemes in simulations of extreme rainfall. The graupel scheme, Milbrandt does not give the best representation of the rainfall field. This finding agrees with Mayor and Mesquita (2015) that including graupel prediction in a microphysical scheme does not necessarily give it superior performance in simulations of tropical rainfall.

The double-moment scheme, SBU_YLin, generated higher rainfall amounts compared to that generated by the single-moment scheme, Eta. As explained by Stensrud (2007) this is probably because it predicts both hydrometeor mixing ratio and concentration unlike single-moment scheme which only predicts the former. Also, the Milbrandt microphysical scheme generated lower rainfall amounts in comparison to the SBU_YLin scheme. This is the effect of added detail in the definition of the hydrometeor shapes within the gamma function of the triple-moment scheme, Milbrandt. It takes advantage of radar reflectivity data unlike the double-moment scheme, SBU_YLin (Warner, 2011b).

TRMM observations showed that rainfall in the basin is mostly received at night and early in the morning. Sun et al. (2014) demonstrated that during this time, the lake is warmer than the adjacent land, favoring the occurrence of a land breeze that causes the rainfall. The WRF model, however,
was unable to correctly simulate this nocturnal rainfall caused by the land breeze. As explained by Argent et al. (2015), there is insufficiency in the representation of the lake surface temperature within the model. The temperature of the lake is constantly colder than the adjacent land. This hinders the reversal of winds to blow from the land to the lake (that is, failure to generate the land breeze). The lack of nocturnal rainfall also explains why the model underestimated rainfall in central and eastern LVB. Mayor and Mesquita (2015) also acknowledged a similar failure of WRF over Cuba.

Notably, the higher rainfall amounts simulated in the high-altitude areas in comparison to the low altitude areas gives evidence of the terrain-sensitive nature of the WRF model which was also reported by Mugume et al. (2017b) and Ntwali et al. (2016).
5.3 Effect of grid resolution on WRF model skill in simulating extreme rainfall

Increasing the model’s resolution had both positive and negative effects on the model skill. In the positive sense, for majority of the runs, it yielded higher model accuracy and an improved level of detail in the spatial representation of rainfall. In the negative sense, for majority of the runs, it lowered the model’s detection ability. Majority of the runs showed a reduction of 2% or more in the number of events detected. Also, the false alarm ratio increased, with majority of the model runs falsely detecting 1 to 2% more rainfall events.

Deducing from the bias and corresponding temporal rainfall accumulation, there was no added value in increasing model grid resolution when using the Control run, E1, E2, E7 and E8 since all of them showed an increase in bias. For the Control run that uses the KF cumulus scheme, it was expected to show an improvement in skill at a higher resolution, however, the results show that it deteriorated. Perhaps, as acknowledged by Stensrud (2007), the scheme’s assumptions become invalid at high grid resolutions. Overall, the results show that increasing model resolution can be advantageous or disadvantageous depending on the parameterization scheme setup that is used.

The new approach of studying model resolution adopted in this study showed an increase in simulation accuracy at a higher resolution. This finding contradicts with that of the traditional approach adopted in studies such as Liu et al. (2011) and Argent et al. (2015), where the model accuracy is higher at lower resolutions (parent domains generate lower RMSE values than the high-resolution nest domains). This new approach gives an improved perspective on the effect of model resolution.

5.3.1 Effect of domain size on WRF model skill

Changing the size of the domain significantly affected the model’s accuracy and false alarm ratio. For majority of the model runs, the large domain was more accurate than the small domain. This is because synoptic features are better resolved in large domains (Heikkilä et al., 2011). The large domain used in this test covered an approximate area of more than 15,000,000 km², making it suitable to capture large scale features such as the ITCZ which is responsible for MAM season rainfall in the LVB. Notably, the large domain generated 1 to 2% more false alarms compared to the small domain.
CHAPTER SIX: CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusions

This study documented rainfall extremes in the Lake Victoria Basin of Uganda in the recent past (2008 to 2017) and demonstrated the possibility of predicting these extremes using a numerical weather prediction model, that is, the WRF model. The conclusions drawn under each objective are presented below.

6.1.1 Spatial and temporal patterns of MAM season rainfall

The investigation of rainfall patterns undertaken in the longest and most important rainfall season over LVB (March to May) for a period of 10 years (2008 to 2017) showed that the MAM seasons in the different years have similar daily rainfall patterns. Therefore, it can be expected that the regimes of daily MAM season rainfall in the near future might not differ much from what has been observed in the recent past. Notably, the investigation also exposed ENSO related periods of enhanced and suppressed rainfall with several heavy and low rainfall extremes. This affirms the importance of sea surface temperature changes on diurnal rainfall events over LVB during the MAM season. Also, the evidence of the occurrence of rainfall extremes in LVB in the recent past reechoes their importance and emboldens the need to study their possible occurrence in the future.

6.1.2 WRF simulations of extreme rainfall under different parameterization settings

For the numerical investigation, this study examined the skill of the WRF model in simulating extreme rainfall experienced in a 20-day period. In this section, uncertainty arising from the representation of sub-grid scale phenomena through parameterization was explored using a suite of 10 model experiments created by employing the model with its default settings (using the default parameterization schemes) and perturbing 3 cumulus and 3 microphysical parameterization schemes.

Overall, the model was able to simulate extreme rainfall over the LVB region, however, not with its default settings. This implies that it is useful to consider other parameterization schemes other than the default. For example, for the case study period, the Grell 3D cumulus scheme in combination with the SBU_YLin microphysical scheme gave more reliable simulations. These
schemes are thus worthy using in numerical simulations of extreme rainfall in equatorial regions. Additionally, model runs done with the Milbrandt microphysical scheme mostly generated the best simulations for the low extremes (dry days). It is therefore worth including this microphysical scheme in numerical simulations targeted at forecasting dry days.

6.1.3 Effect of grid resolution on WRF model skill in simulating extreme rainfall

In this section, uncertainty arising from the choice of model resolution and domain size was explored and the results showed that some model runs generated better simulations at a high resolution than at a low resolution while others generated better simulations at a low resolution than at a high resolution. This dualistic performance implies that the relevance of increasing the model grid resolution depends on the parameterization schemes used. In this study, it was mostly useful for experiments done with the Grell 3D cumulus scheme. Therefore, this scheme is suitable for high resolution numerical simulations of rainfall in equatorial regions. Furthermore, the higher accuracy generated by the large domain implies that there is added value in using a large domain in comparison to a small one.

6.2 Recommendations

The WRF model can be applied in operational extreme rainfall forecasting at national meteorological and hydrological services within the equatorial regions most especially in Uganda where the study area is located. The potential improvements that the model will bring about in extreme rainfall forecasting will trickle down and also be of benefit to rain-dependent sectors such as in agriculture for making cropping advisories, in disaster preparedness for making flood and drought forecasts and early warnings and in public health for making early warnings for water-borne disease outbreaks among others.

However, to achieve a holistic understanding of the WRF model’s ability to simulate extreme rainfall, further research is required to investigate the performance of the parameterization schemes at very high resolutions (between 3 km and 0.1 km) to test their stability and usefulness at such resolutions. It is also important to study the other possible sources of model uncertainty which could arise from the choice of boundary conditions, vertical resolution (the number of atmospheric levels used) and simulation lead time among others.
In order to achieve realistic in-land lake temperatures, there is need for modification of the land-use data, and land surface and surface layer parameterization. Also, there is need for improvement in the definition of all other model physics in order to achieve improvements in the overall rainfall simulations.
REFERENCES


Eza, J. (2013). *Challenges in forecasting severe weather events in uganda, issuance and delivery*


Ogwang, B. A., Chen, H., Li, X., Gao, C., Prof, T., & Author, C. (2014). The influence of Topography on East African October to December climate: Sensitivity experiments with RegCM4 Key Laboratory of Meteorological Disaster, Ministry of Education, Nanjing University of E-mail Addresses: bob_ogwang@yahoo.co.uk (Bob Alex Og. *Advances in Meteorology, 2014*, 1–27. https://doi.org/doi.org/10.1155/2014/143917


APPENDICES

Appendix 1: Rainfall timeseries for the 2008 MAM season

The 3 timeseries corresponding to the 3 regions of the basin (Figure 14) all agree that the highest amount of rainfall was received toward the end of March whereas the early days of April received barely any rainfall.

![Figure 14](image-url)  
**Figure 14:** Rainfall received during the MAM season of 2008.
Appendix 2: Demonstration of a trend in daily rainfall within individual MAM seasons

Using simple linear regression, the regression equations for 2008 were, \( y = -0.06299x + 6.751362 \) (p < 0.05), \( y = -0.04128x + 7.101023 \) (p > 0.05) and \( y = -0.03862x + 8.053839 \) (p > 0.05) for western, central and eastern LVB respectively. Similarly, for 2010, the equations were, \( y = -0.02418x + 3.375626 \) (p > 0.05), \( y = -0.025x + 3.522948 \) (p < 0.05) and \( y = -0.04695x + 6.159529 \) (p < 0.05) respectively.

![Figure 15](image.png)

**Figure 15**: Trend in daily rainfall for the (a) 2008 and (b) 2010 MAM seasons.
**Appendix 3: Details of parameterization schemes considered for the study**

<table>
<thead>
<tr>
<th>MICROPHYSICAL PARAMETERIZATION</th>
<th>CUMULUS PARAMETERIZATION</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>WRF Single Moment 3 class (WSM3)</strong> (Hong et al., 2004). This is a single moment microphysical scheme predicting three hydrometeors, that is, water vapor, cloud water or cloud ice, and rain or snow.</td>
<td><strong>Kain-Fritsch (KF) scheme</strong> (Kain, 2004). This is a mass flux convective scheme with a physically realistic representation of mass exchange between clouds and their surrounding environment.</td>
</tr>
<tr>
<td><strong>Eta scheme</strong> (Rogers et al., 1996, 2005). This is a single moment microphysical scheme predicting five hydrometeors, that is, water vapor, cloud water, cloud ice, rain, and precipitation ice.</td>
<td><strong>Grell 3D</strong> (Grell, 1993; Grell &amp; Dévényi, 2002). This is a mass flux convective scheme that may also be used on high resolution if subsidence spreading is turned on.</td>
</tr>
<tr>
<td><strong>Milbrandt scheme</strong> (Milbrandt &amp; Yau, 2005a, 2005b). This is a multi-moment microphysical scheme predicting six hydrometeors, that is, cloud water, cloud ice, rain, snow, hail and graupel.</td>
<td><strong>Betts-Miller-Janjic (BMJ)</strong> (Janjić, 1994). This is an adjustment type scheme relaxing toward a positive convective sounding. The instability is eliminated by nudging environmental profiles of temperature and humidity.</td>
</tr>
<tr>
<td><strong>Stony-Brook University scheme</strong> (SBU_YLIN) (Lin &amp; Colle, 2011). This is a double-moment microphysical scheme predicting five hydrometeors, that is, water vapor, cloud water, cloud ice, rain and precipitation ice.</td>
<td><strong>Grell-Freitas (GF)</strong> (Grell &amp; Freitas, 2014). This is a mass flux convective scheme based off a stochastic approach demonstrated by Grell and Dévényi (2002).</td>
</tr>
</tbody>
</table>
Appendix 4: Atmospheric governing equations that are solved by the WRF ARW core.

Warner (2011b) describes the following as the atmospheric governing equations, on which numerical weather models are based off: the momentum equations (i, ii, iii), the thermodynamic energy equation (iv), the continuity equation (v), water vapor equation (vi) and ideal gas law (vii).

\[
\begin{align*}
\frac{\partial u}{\partial t} &= -u \frac{\partial u}{\partial x} - v \frac{\partial u}{\partial y} - w \frac{\partial u}{\partial z} + \frac{uw \tan \varphi}{a} - \frac{uw}{a} - \frac{1}{\rho} \frac{\partial p}{\partial x} - 2\Omega(w \cos \varphi - v \sin \varphi) + F_{rx} \\
\frac{\partial v}{\partial t} &= -u \frac{\partial v}{\partial x} - v \frac{\partial v}{\partial y} - w \frac{\partial v}{\partial z} + \frac{u^2 \tan \varphi}{a} - \frac{uw}{a} - \frac{1}{\rho} \frac{\partial p}{\partial y} - 2\Omega u \sin \varphi + F_{ry} \\
\frac{\partial w}{\partial t} &= -u \frac{\partial w}{\partial x} - v \frac{\partial w}{\partial y} - w \frac{\partial w}{\partial z} + \frac{u^2 + v^2}{a} - \frac{1}{\rho} \frac{\partial p}{\partial z} - 2\Omega u \cos \varphi - g + F_{rz} \\
\frac{\partial T}{\partial t} &= -u \frac{\partial T}{\partial x} - v \frac{\partial T}{\partial y} + (\gamma - \gamma_d)w + \frac{1}{c_p} \frac{dH}{dt} \\
\frac{\partial \rho}{\partial t} &= -u \frac{\partial \rho}{\partial x} - v \frac{\partial \rho}{\partial y} - w \frac{\partial \rho}{\partial z} - \rho \left( \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} \right) \\
\frac{\partial q_v}{\partial t} &= -u \frac{\partial q_v}{\partial x} - v \frac{\partial q_v}{\partial y} - w \frac{\partial q_v}{\partial z} + Q_v \\
P &= \rho RT
\end{align*}
\]

From the above equations, \( u, v \) and \( w \) are velocity components, \( \rho \) is density, \( p \) is atmospheric pressure, \( T \) is temperature, \( q_v \) is specific humidity, \( \Omega \) is rotational frequency of the earth, \( \varphi \) is latitude, \( a \) is radius of the earth, \( \gamma \) is lapse rate of temperature, \( \gamma_d \) is dry adiabatic lapse rate, \( c_p \) is the specific heat capacity of air at constant pressure, \( g \) is acceleration due to gravity, \( H \) is heat loss or gain, \( Q_v \) is loss or gain of water vapor, and \( F_r \) is the frictional force.