ABSTRACT

GUDOSHAVA, MASILIN. An Investigation of the Sensitivity of East African Climate Variability to Urbanization. (Under the direction of Dr. Fredrick Semazzi and Dr. Lian Xie.)

In recent decades, the urbanization rate in Africa has been growing rapidly. It is projected that by 2030, the rate could increase by about 590% over its 2000 levels. By 2050 Africa is projected to have the second highest urban population. East Africa is among the regions that are expected to have a high increase in urbanization. The urbanization changes in East Africa will occur over Lake Victoria Basin. Lake Victoria Basin is the nerve centre of East Africa and supports over 35 million people. Rapid urbanization could affect the livelihoods of the lakeside community and cause further degradation of the already stressed ecosystem.

The aim of this research was to investigate the impact of urbanization on the climate of East Africa and the subsequent consequences on the energy and agricultural sectors. This was achieved by accomplishing the following objectives (i) Comprehensive customization and evaluation of the performance of the International Center of Theoretical Physics (ICTP) Regional climate model (ii) Investigation of the ability of the ICTP Regional climate model to reproduce the modes of climate variability from observational data. (iii) Investigate the impact of urbanization on the precipitation of East Africa (iv) Investigation of the dominant process between, moisture supply, thermal convergence and frictional convergence, and (v) Impact on agriculture and energy sector.

The study was conducted using two regional climate models, the ICTP Regional Climate Model (RegCM) and the Weather and Research Forecasting (WRF) model. The regional models were customized based on previous studies in the region, with more customization studies conducted for RegCM. Both models were able to capture the spatial variability of precipitation and temperature in the region however in some regions they overestimated the precipitation amount. We further analyzed the ability of the ICTP regional climate model to capture the inter-annual variability of precipitation during both the Long and Short Rains. Two domains were utilized, the East Africa, and Lake Victoria Basin domains. RegCM was able to capture the inter-annual variability of precipitation...
during the Short Rains and had less skill for the Long Rains. The Indian Ocean Dipole was found to be the leading source of inter-annual variability for both domains during the Short Rains. In this study we found the influence of the Indian Ocean Dipole on the Long Rains over the Lake Victoria Basin. For the East Africa domain the influence was rather weak suggesting that the region is inhomogeneous.

To investigate the impact of urbanization two urbanization scenarios were utilized, the low urbanization and the high urbanization scenario. The impact of urbanization on climate was mostly localized with the higher urbanization scenario having a higher impact. The changes in precipitation, temperature and partitioning of energy fluxes were found to be seasonal. The Long Rains season had higher changes in precipitation compared to the Short Rains. Sensitivity tests on identifying processes that are dominant in producing the precipitation pattern over urbanized regions were also conducted. Thermal convergence is found to be the dominant factor in reproducing the precipitation patterns. However if the sensitivity experiments are conducted with Lake Victoria converted to shrublands/grassland the precipitation over the region dramatically decreases. Thus when the lake is converted to another vegetation type, the moisture supply into the urbanized region is reduced. This might imply that if urbanization occurs in a region that is remote from a water body the total precipitation could reduce. The changes in precipitation were not uniform throughout the urbanized region. The western region had higher changes in precipitation in comparison with the eastern side. We proposed that the changes are mostly influenced by the interaction of the mesoscale flow, land-lake breezes and the mountain breezes.

The high frequency statistics for precipitation was calculated for agricultural purposes. In this case we calculated the number of consecutive dry days, periods with 5 continuous wet days and also the highest one day precipitation. Although both models reasonably capture the high precipitation frequency over the domain, the WRF model is able to capture the statistics more accurately than RegCM. With increase in urbanization in the future, an associated increase in the maximum daily precipitation in a season is observed. The Long Rains season had higher increases in maximum daily precipitation than the Short rains. The interaction of the Intertropical Convergence Zone and local urban circulation could enhance the precipitation during the Long Rains. The number
of continuous dry days during the Long Rains season is longer for the urbanized case in comparison with the control case. The changes in the high frequency precipitation imply that the season will have more intense rainfall events, but the number of days with rainfall will reduce.

In conclusion high urbanization over Lake Victoria Basin will increase the temperature and precipitation over the region. The minimum temperatures will increase thereby shortening the diurnal temperature range. The precipitation increases more on the western side of Lake Victoria Basin, than the eastern side. The frequency of the heavy precipitation increases in both models.

For wind application analysis, we use the WRF model because of its higher spatial resolution. The model is able to capture the spatial variability of the wind speed, when compared to ERA-Interim reanalysis data. The regional climate model is able to identify the Lake Turkana region that has been obtained in previous studies to have high wind power resources. This study has shown regions that have a potential for wind energy resources. These include a wider region over northern Tanzania other than the Singinda region, the northwestern part of Uganda, in addition there are more regions that were identified in Kenya. Changes in the land cover were not found to affect the viability of wind resources over the region.
An Investigation of the Sensitivity of East African Climate Variability to Urbanization

by
Masilin Gudoshava

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Marine, Earth and Atmospheric Sciences

Raleigh, North Carolina

2016

APPROVED BY:

Dr. Fredrick Semazzi
Co-chair of Advisory Committee

Dr. Lian Xie
Co-chair of Advisory Committee

Dr. Jared Bowden

Dr. Bin Liu

Dr. Alton Banks
DEDICATION

To my parents and siblings, you are my pillar of strength.
BIOGRAPHY

Masilin Gudoshava was born in Bulawayo, (aka The City of Queens and Kings) Zimbabwe. She obtained a BSc. Honors in Applied Mathematics at the National University of Science and Technology. After completing her BSc. Honors in Applied Mathematics she was awarded a Staff Development Fellowship by Great Zimbabwe University to study Industrial Mathematics at the National University of Science and Technology. In 2011 she was awarded a Fulbright Junior Staff Development fellowship to study a PhD in Atmospheric Sciences at North Carolina State University.
ACKNOWLEDGEMENTS

I would like to extend my profound gratitude to my academic advisors Drs. F. Semazzi, L. Xie, J. Bowden, B. Liu and my graduate school representative Dr. A. Banks for all the guidance that was rendered to me during the study. To my lab mates (Kara Smith, Rowan Agent, Michael Angus and Alexa Wood) would like to say for the help in improving my work. They say it takes the whole village to raise a child, I would like to say thank you to the following organizations and institutions for sponsoring my PhD studies in the past 5 years, National University of Science and Technology, the Fulbright Programme, Margret McNamara Memorial Fund, Schlumberger Faculty for the Future Fellowship and the International Peace Scholarship (IPS). To all the members of the IPS I say thank you for all the gifts, the companionship and encouragement. To my current Fulbright advisor, Megan thank you for your patience and support. I would also like thank all my friends for the support. Special mention goes to the late Professor C. P. Bhunu, thank you for all the encouragement and support you have shown me throughout my studies. I am immensely grateful to my family for their unwavering support and encouragement over the years. Last but not least I would like to thank the Lord God for giving me this opportunity. To God be the Glory.

All WRF computational work was done on the Yellowstone supercomputing system. The RegCM simulations were conducted on NCSU HPC. Analysis was done using the NCAR Command Language and also the Climate Data Operators.
TABLE OF CONTENTS

LIST OF TABLES .................................................. viii

LIST OF FIGURES .................................................. xi

Chapter 1 Introduction .............................................. 1
  1.1 Projected Urbanization Levels over East Africa ............... 1
  1.2 Impacts of Urbanization on Regional Climate ................. 3
  1.3 Research Hypothesis and Aim ................................ 7
    1.3.1 Limitations of the Study .............................. 10

Chapter 2 Evaluation of the Performance ICTP Regional Climate Model 11
  2.1 Introduction ............................................... 11
  2.2 Model Description and Experimental Setup .................... 13
    2.2.1 Model Validation ...................................... 18
  2.3 Results and Discussion .................................... 19
    2.3.1 Experiments with Higher Resolution .................... 23
      2.3.1.1 Cumulus Schemes .................................. 24
      2.3.1.2 Microphysics Schemes: ........................... 30
      2.3.1.3 Land Surface Schemes: .......................... 32
      2.3.1.4 Planetary Boundary Layer Scheme ................ 37
      2.3.1.5 Radiation Scheme ................................. 38
    2.3.2 Performance of the Optimal Conditions in Extreme years 42
    2.3.3 Evaluation of the Performance in Multiple Years ........ 49
  2.4 Conclusions ............................................... 55

Chapter 3 Analysis of East Africa Climatology using Satellite data and
  the ICTP Regional Climate Model ................................ 57
  3.1 Introduction ............................................... 57
  3.2 Data ...................................................... 60
  3.3 Methods of Analysis ....................................... 63
  3.4 Results and Discussion .................................... 64
    3.4.1 Rainfall Patterns ..................................... 64
    3.4.2 Empirical Orthogonal Function Analysis ................. 70
      3.4.2.1 Short Rains season ............................... 70
      3.4.2.2 Long Rains MAM season .......................... 78
  3.5 Conclusions ............................................... 102

Chapter 4 Impact of urbanization on the Climate of East Africa .... 104
  4.1 Introduction ............................................... 104
  4.2 Methods of Analysis ....................................... 109
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2.1</td>
<td>Precipitation Datasets</td>
<td>110</td>
</tr>
<tr>
<td>4.2.1.1</td>
<td>Satellite</td>
<td>110</td>
</tr>
<tr>
<td>4.2.1.2</td>
<td>Gridded</td>
<td>112</td>
</tr>
<tr>
<td>4.2.1.3</td>
<td>Satellite and Gridded</td>
<td>113</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Temperature Datasets</td>
<td>113</td>
</tr>
<tr>
<td>4.2.3</td>
<td>Urbanization Scenarios</td>
<td>113</td>
</tr>
<tr>
<td>4.3</td>
<td>Results and Discussion</td>
<td>116</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Model Validation</td>
<td>116</td>
</tr>
<tr>
<td>4.3.1.1</td>
<td>Evaluation for the Short Rains</td>
<td>116</td>
</tr>
<tr>
<td>4.3.1.2</td>
<td>Evaluation for the Long Rains</td>
<td>119</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Impact of Urbanization on the Climate of East Africa: High Urbanization Scenario</td>
<td>123</td>
</tr>
<tr>
<td>4.3.2.1</td>
<td>Land Surface Albedo</td>
<td>123</td>
</tr>
<tr>
<td>4.3.2.2</td>
<td>Surface Energy Fluxes</td>
<td>125</td>
</tr>
<tr>
<td>4.3.2.3</td>
<td>Surface Hydrology</td>
<td>129</td>
</tr>
<tr>
<td>4.3.2.4</td>
<td>Temperature</td>
<td>132</td>
</tr>
<tr>
<td>4.3.2.5</td>
<td>Precipitation</td>
<td>137</td>
</tr>
<tr>
<td>4.3.3</td>
<td>Impact of Urbanization on the Climate of East Africa: Low Urbanization Scenario</td>
<td>148</td>
</tr>
<tr>
<td>4.3.4</td>
<td>Surface Energy Fluxes</td>
<td>148</td>
</tr>
<tr>
<td>4.3.5</td>
<td>Temperature</td>
<td>150</td>
</tr>
<tr>
<td>4.3.6</td>
<td>Precipitation</td>
<td>150</td>
</tr>
<tr>
<td>4.3.7</td>
<td>Conclusions</td>
<td>151</td>
</tr>
</tbody>
</table>

**Chapter 5 Identification of the Dominant Processes** | 153
| 5.1    | Introduction | 153 |
| 5.2    | Methods of Analysis | 154 |
| 5.3    | Results and Discussion | 154 |
| 5.3.1  | Sensitivity Simulations with Lake Victoria Included | 155 |
| 5.3.2  | Sensitivity Simulation with Lake Victoria converted to a Vegetated Land Type | 157 |
| 5.3.3  | Conclusions | 157 |

**Chapter 6 Impact on the Agricultural Sector** | 165
| 6.1    | Introduction | 165 |
| 6.2    | Methods of Analysis | 167 |
| 6.3    | Results and Discussion | 168 |
| 6.3.1  | Control Simulation | 169 |
| 6.3.1.1| Statistics for Dry and Wet Days During the Long Rains | 169 |
| 6.3.1.2| Statistics for Dry and Wet Days During the Short Rains | 176 |
| 6.3.2  | Sensitivity Simulations | 179 |
| 6.4    | Conclusions | 183 |
LIST OF TABLES

Table 2.1 Summary of the customization experiments conducted for the RegCM model using a 36km resolution. ................................................................. 17
Table 2.2 Summary of evaluation statistics for RegCM model at a 36km resolution for the different simulations. RMSE is the root mean square error, SD is the standard deviation of the model, NMB is the normalized mean bias error, NME is the normalized mean error. .... 23
Table 2.3 Summary of RegCM simulations done to test the performance of the different cumulus schemes. The land/ocean indicate which schemes were used over land/ocean in the simulation. ........................................ 25
Table 2.4 Summary of evaluation statistics for RegCM simulations done to test the performance of the different large scale precipitation schemes. RMSE is the root mean square error, NME is the normalized mean error, NMB is the normalized mean bias, MB- mean bias, Cor is the pattern correlation, NRMSE is the normalized root mean square SD Mod.- standard deviation calculated from model output and SD Obs. - standard deviation for observations .......................... 33
Table 2.5 Summary of evaluation statistics for RegCM simulations done to test the performance of the Community Land Model. RMSE is the root mean square error, NME is the normalized mean error, NMB is the normalized mean bias, MB- mean bias, Cor is the pattern correlation, SD Mod.- standard deviation calculated from model output and SD Obs.- standard deviation for observations. The microphysics N/T is the Nogherroto/Tompkins scheme. .................. 35
Table 2.6 Summary of evaluation statistics for RegCM simulations done to test the performance of the different Planetary Boundary Layer schemes. RMSE is the root mean square error, NME is the normalized mean error, NMB is the normalized mean bias, MB- mean bias, Cor is the pattern correlation, SD Mod.- standard deviation calculated from model output and SD Obs.- standard deviation for observations. .................. 39
Table 2.7 Summary of evaluation statistics for RegCM simulations done to test the performance of the RRTM radiative scheme. RMSE is the root mean square error, NME is the normalized mean error, NMB is the normalized mean bias, MB- mean bias, Cor is the pattern correlation, SD Mod.- standard deviation calculated from model output and SD Obs.- standard deviation for observations. The microphysics N/T is the Nogherroto/Tompkins scheme. .................. 41
Table 3.1 Summary of the correlation of TRMM and model output EOF 1 with the different indices used in the study for the Short Rains Season. IOD-Indian Ocean Dipole, NINO3, QBO-Quasi-Biennial Oscillation, TNA-Tropical Northern Atlantic, TSA-Tropical Southern Atlantic, NAO-North Atlantic Oscillation, WP-West Pacific. The ** indicates correlations that are statistically significant at 95% and * at 90% significance of correlation was done using a t-test.

Table 3.2 Summary of the correlation of TRMM and model output EOF 1 with the different indices used in the study for the Long Rains Season. IOD-Indian Ocean Dipole, NINO3, QBO-Quasi-Biennial Oscillation, TNA-Tropical Northern Atlantic, TSA-Tropical Southern Atlantic, NAO-North Atlantic Oscillation, WP-West Pacific. The ** indicates correlations that are statistically significant at 95% and * at 90% significance of correlation was done using a t-test.

Table 4.1 Summary of WRF urbanization sensitivity simulations

Table 4.2 Model configuration used in the study for the REGCM and the WRF model.

Table 5.1 Summary of the experiments done to isolate the dominant processes.

Table 6.1 Intra-seasonal statistics normally calculated for agricultural purposes.

Table 7.1 Wind power classification.

Table B.1 Optimal physics combinations over East Africa from WRF simulations that have been utilized in past research.

Table B.2 Experiments Conducted to customize WRF model with NOAH land Surface model.

Table C.1 Summary of statistics for OND temperature for WRF and REGCM. OBS- observational datasets, SD Mod.- standard deviation calculated from model output and SD Obs. - standard deviation for observations, RMSE is the root mean square error, MB- mean bias, NME is the normalized mean error, NMB is the normalized mean bias.

Table C.2 Summary of Evaluation Statistics for WRF Simulations. SD Mod.- standard deviation calculated from model output and SD Obs. - standard deviation for observations, MB- mean bias, NME (%) is the normalized mean error, NMB (%) is the normalized mean bias, RMSE is the root mean square error.
Table C.3  Summary of Evaluation Statistics for RegCM Simulations. SD Mod.- standard deviation calculated from model output and SD Obs. - standard deviation for observations, MB- mean bias, NME (%) is the normalized mean error, NMB (%) is the normalized mean bias, RMSE is the root mean square error. 

Table C.4  Summary of statistics for MAM temperature for WRF and REGCM, OBS- observational datasets, SD Mod.- standard deviation calculated from model output and SD Obs. - standard deviation for observations, RMSE is the root mean square error, MB- mean bias, NME is the normalized mean error, NMB is the normalized mean bias. 

Table C.5  Summary of Evaluation Statistics for WRF Simulations. SD Mod.- standard deviation calculated from model output and SD Obs. - standard deviation for observations, MB- mean bias, NME (%) is the normalized mean error, NMB (%) is the normalized mean bias, RMSE is the root mean square error. 

Table C.6  Summary of Evaluation Statistics for RegCM Simulations during the Long Rains. SD Mod.- standard deviation calculated from model output and SD Obs. - standard deviation for observations, MB- mean bias, NME (%) is the normalized mean error, NMB (%) is the normalized mean bias, RMSE is the root mean square error. 

Table C.7  Summary of Evaluation Statistics for WRF D01 Simulations precipitation for the Short Rains. RMSE is the root mean square error, NME is the normalized mean error, NMB is the normalized mean bias, MB- mean bias, Cor is the pattern correlation, NRMSE is the normalized root mean square SD Mod.- standard deviation calculated from model output and SD Obs. - standard deviation for observations. 

Table C.8  Summary of Evaluation Statistics for WRF D01 temperature Simulations during the Long Rains. RMSE is the root mean square error, NME is the normalized mean error, NMB is the normalized mean bias, MB- mean bias, Cor is the pattern correlation, NRMSE is the normalized root mean square SD Mod.- standard deviation calculated from model output and SD Obs. - standard deviation for observations. 

Table C.9  Summary of Evaluation Statistics for WRF precipitation domain 1 during the Long Rains. RMSE is the root mean square error, NME is the normalized mean error, NMB is the normalized mean bias, MB- mean bias, Cor is the pattern correlation, NRMSE is the normalized root mean square SD Mod.- standard deviation calculated from model output and SD Obs. - standard deviation for observations. 

x
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Projected urbanization growth by 2030 over 2000 levels (Seto et al., 2012).</td>
<td>2</td>
</tr>
<tr>
<td>1.2</td>
<td>Illustration of the region of focus (a) Lake Victoria Basin (b) projected urbanization levels in 2030 over the 2000 levels over the Lake Victoria Basin (Seto et al., 2012).</td>
<td>3</td>
</tr>
<tr>
<td>1.3</td>
<td>Biogeophysical feedback processes in response to land cover change (Mölders, 2011).</td>
<td>5</td>
</tr>
<tr>
<td>1.4</td>
<td>Projected increases in Precipitation over East Africa: Low pass filtered rainfall March-May Average over Greater Horn of Africa (Rowell et al. (2015)).</td>
<td>7</td>
</tr>
<tr>
<td>1.5</td>
<td>Past (top, 1900-2010, Graph by Robert Simmon, based on data provided by the USDA) and Projected (2070-2100, Smith (2011)) (bottom) Lake Victoria levels trends.</td>
<td>8</td>
</tr>
<tr>
<td>2.1</td>
<td>Precipitation distribution using different physics combinations for the simulations run at a resolution of 36 km.</td>
<td>22</td>
</tr>
<tr>
<td>2.2</td>
<td>Precipitation evaluation statistics for the run at a resolution of 36 km. The red line represents the standard deviation of TRMM.</td>
<td>24</td>
</tr>
<tr>
<td>2.3</td>
<td>Precipitation distribution using different cumulus parameterization for a 25 km resolution simulation. (a) Grell AS (b) Grell AS/MIT-Emanuel (c) MIT-Emanuel/Grell AS (d) Grell FC (e) Grell FC/MIT-Emanuel (f) MIT-Emanuel/Grell FC (g) MIT-Emanuel and (h) the observations TRMM.</td>
<td>27</td>
</tr>
<tr>
<td>2.4</td>
<td>RMSE (blue bars) and Standard deviation (purple) over Congo Region using the different cumulus schemes. The red line is the standard deviation of the TRMM observational dataset.</td>
<td>28</td>
</tr>
<tr>
<td>2.5</td>
<td>RMSE (blue bars) and Standard deviation (purple) over East Africa Region using the different cumulus schemes. The red line is the standard deviation of the TRMM observational dataset.</td>
<td>29</td>
</tr>
<tr>
<td>2.6</td>
<td>RMSE (blue bars) and Standard deviation (purple) over Indian Ocean Region using the different cumulus schemes. The red line is the standard deviation of the TRMM observational dataset.</td>
<td>30</td>
</tr>
<tr>
<td>2.7</td>
<td>Spatial correlations of precipitation for the different regions using different cumulus schemes with the TRMM datasets.</td>
<td>31</td>
</tr>
<tr>
<td>2.8</td>
<td>Precipitation distribution using the Nogherroto/Tomkins microphysics scheme (a) Grell FC (b) MIT-Emanuel.</td>
<td>32</td>
</tr>
<tr>
<td>2.9</td>
<td>Precipitation distribution using the CLM land surface module.</td>
<td>34</td>
</tr>
<tr>
<td>2.10</td>
<td>Average evapotranspiration and 10 m winds for the different Land Surface Modules (a) BATS (b) CLM.</td>
<td>36</td>
</tr>
</tbody>
</table>
Figure 2.11 Precipitation distribution using the University of Washington planetary Boundary Layer Scheme at 25km resolution simulation.

Figure 2.12 Precipitation distribution using the RRTM radiation scheme for a 25km resolution simulation.

Figure 2.13 Precipitation distribution for 2000 using the Grell FC Scheme.

Figure 2.14 Statistical values for the evaluation of model during the Short rain season of 2000.

Figure 2.15 Precipitation distribution for 2006 using the Grell FC Scheme.

Figure 2.16 Clustered bar plot for the different statistics calculated to evaluate the performance of the model for the year 2006.

Figure 2.17 Precipitation distribution for 2010 using the Grell FC Scheme.

Figure 2.18 Clustered bar plot for the different statistics calculated to evaluate the performance of the model for the year 2010.

Figure 2.19 Taylor diagram for precipitation during the Short Rains season over East Africa.

Figure 2.20 Taylor diagram for temperature during the Short Rains season over East Africa.

Figure 2.21 Probability distribution function plot for temperature during the Short Rains over East Africa.

Figure 2.22 Taylor diagram for precipitation during the Long Rains over East Africa.

Figure 2.23 Taylor diagram for temperature during the Long Rains over East Africa.

Figure 2.24 Probability distribution frequency plot for temperature during the Long Rains over East Africa.

Figure 3.1 Elevation (m) and domain utilized in the study. The 2 subregions are that will be utilized in the study are highlighted. The inner domain is the Lake Victoria domain and the outer one is the East Africa domain.

Figure 3.2 Average precipitation over every 3 hour period for TRMM 3B42 datasets for the 16 year period. The plots are generated by calculating the average precipitation at each 3 hour time interval during the short rain season.

Figure 3.3 Diurnal variation for selected sites over the Basin for the TRMM 3B42, during the Short Rains season.

Figure 3.4 Monthly timeseries of rainfall variability over East Africa using TRMM. The Different lines were constructed using different latitudes in the region. The average line was calculated by considering the 16 year average precipitation over the region.
Figure 3.5 Monthly timeseries of rainfall variability over East Africa using TRMM data. The line plots are for the Extreme North and South parts of our East Arica Domain. The average line was calculated by considering the 16 year average precipitation over the region.

Figure 3.6 Spatial distribution for OND EOF1 over East Africa domain from models and observation (a) Grell FC (b) MIT Emanuel (c) TRMM.

Figure 3.7 Temporal Variation for the first mode for the Short Rains over the East Africa domain.

Figure 3.8 Correlation plot for OND EOF 1 timeseries with the different climate indices over East Africa.

Figure 3.9 Composites for EOF 1 during the Short Rains for the East Africa Domain (a) GRELL FC (b) MIT (c) TRMM.

Figure 3.10 Composite of the zonal wind using TRMM EOF1 timeseries over East Africa Region (a) above average precipitation (b) below average precipitation.

Figure 3.11 Spatial and temporal distribution for OND EOF1 Lake Victoria Basin Region from models and observation (a) Grell FC (b) MIT Emanuel (c) TRMM.

Figure 3.12 Timeseries plot for OND EOF 1 over Lake Victoria Basin.

Figure 3.13 Correlation plot for OND EOF 1 timeseries over Lake Victoria Basin with the different climate indices.

Figure 3.14 Short Rains Composites for EOF 1 for the East Africa Domain (a) GRELL FC (b) MIT (c) TRMM.

Figure 3.15 Vertical velocity (North South) components for the (a)LVB region (b) East Africa.

Figure 3.16 Spatial distribution for OND EOF2 over Lake Victoria Basin from models and observation (a) Grell FC (b) MIT Emanuel (c) TRMM.

Figure 3.17 Timeseries for OND EOF 2 timeseries over Lake Victoria Basin.

Figure 3.18 Cross section for Omega for OND EOF 2 timeseries over Lake Victoria Basin.

Figure 3.19 Spatial distribution for MAM EOF1 from models and observation (a) Grell FC (b) MIT Emanuel (c) TRMM.

Figure 3.20 MAM EOF1 timeseries for observational data and model over East Africa domain.

Figure 3.21 Correlation plots for MAM EOF1 timeseries and Climate indices over East Africa domain.

Figure 3.22 Long Rains Composites for EOF 1 for the East Africa Domain (a) GRELL FC (b) MIT (c) TRMM.

Figure 3.23 Spatial distribution for MAM EOF1 from models and observation over Lake Victoria Basin (a) Grell FC (b) MIT Emanuel (c) TRMM.

Figure 3.24 Timeseries for MAM EOF1 over Lake Victoria Basin.
Figure 3.25 Timeseries plot for the March-April and May season for the 2 regions utilized (a) March-April EA (b) March-April LVB (c) May EA(d) May LVB. .................................................. 92

Figure 3.26 Spatial distribution for March-April EOF1 from model and observation over East Africa (a) Grell FC (b) MIT-Emanuel (c) TRMM 93

Figure 3.27 March-April SST-circulation composites for EOF 1 East Africa Domain (a) Grell FC (b) MIT-Emanuel (c) TRMM. ................... 94

Figure 3.28 Spatial distribution for March-April EOF1 from models and observation over Lake Victoria Basin (a) Grell FC (b) MIT Emanuel (c) TRMM. .................................................. 95

Figure 3.29 March-April Composites for EOF 1 Lake Victoria Basin Region (a) Grell FC (b) MIT-Emanuel (c) TRMM. ............................... 96

Figure 3.30 Spatial distribution for May EOF1 from model simulations and observation over East Africa (a) Grell FC (b) MIT-Emanuel (c) TRMM. .................................................. 98

Figure 3.31 May SST-Circulation composites for EOF 1 East Africa Domain (a) Grell FC (b) MIT-Emanuel (c) TRMM. ............................... 99

Figure 3.32 Spatial distribution for MAY EOF1 from model and observation over Lake Victoria Basin (a) Grell FC (b) MIT Emanuel (c) TRMM. 100

Figure 3.33 May Composites for EOF 1 Lake Victoria Basin Region (a) GRELL FC (b) MIT (c) TRMM. .................................................. 101

Figure 3.34 Timeseries plot over the lake for the model and the observational data. The plots were obtained by averaging the precipitation over the lake with high loading. ................. 102

Figure 4.1 Proportion of urban dwellers in the different continents (United Nations, 2014). .................................................. 105

Figure 4.2 Countries with projected increase in urban population of greater than 25million by 2050 (United Nations, 2014). ...................... 106

Figure 4.3 Landuse Cover for the inner domain in the control simulation for WRF. .................................................. 111

Figure 4.4 Landuse cover for the high urbanization scenario. .............. 114

Figure 4.5 Landuse cover for the low urbanization scenario. ............... 115

Figure 4.6 Taylor diagram for temperature during the OND season (a) WRF (b) RegCM. .................................................. 117

Figure 4.7 Taylor diagram for temperature during the OND season (a) WRF (b) RegCM. .................................................. 119

Figure 4.8 Taylor diagram for temperature during the MAM season (a) WRF (b) RegCM. .................................................. 120

Figure 4.9 Taylor diagram for temperature during the Long Rains season (a) WRF (b) RegCM. .................................................. 122
Figure 4.10 Change in albedo between the control simulation and the high urbanization scenario for the 2 models (a) WRF MAM (b) RegCM MAM (c) WRF OND (d) RegCM OND.

Figure 4.11 Changes in energy fluxes during the Long Rains for the WRF model (a) Latent heat (b) Sensible heat flux.

Figure 4.12 Changes in energy fluxes during the Long Rains for the REGCM model (a) latent heat (b) sensible heat flux.

Figure 4.13 Changes in energy fluxes during the Short Rains season for the WRF Model (a) Latent heat flux (b) Sensible heat flux.

Figure 4.14 Changes in energy fluxes during the Short Rains season for the REGCM model (a) Latent heat flux (b) Sensible heat flux.

Figure 4.15 Diurnal variation of the Bowen ratio for the WRF model during the wet season (a) Long rains (b) Short rains.

Figure 4.16 Changes in evapotranspiration in the different seasons for the WRF model (a) JF (b) MAM (c) JJAS (d) OND.

Figure 4.17 Changes in surface evapotranspiration for the REGCM model (a) MAM (b) OND.

Figure 4.18 Changes in surface runoff in different seasons for the WRF model.

Figure 4.19 Changes in surface runoff for the REGCM model (a) MAM (b) OND.

Figure 4.20 Change in temperature between the control simulation and the high urbanization scenario during the (a) March-May (b) October to December season for the RegCM model.

Figure 4.21 Change in temperature between the control simulation and the high urbanization scenario in the different seasons (a) January-February (b) March-May (c) June-September (d) October-December for the WRF model.

Figure 4.22 Probability density function for the temperature for the during the 2 wet season for the 2 models (a) March-May WRF (b) March-May REGCM (c) October-December WRF (d) October-December REGCM.

Figure 4.23 Diurnal cycle of temperature for the WRF model during the 2 wet seasons (a) March-May (b) October-December.

Figure 4.24 Change in precipitation between the control simulation and the high urbanization scenario during the Long Rains.

Figure 4.25 Change in relative humidity between the control simulation and the high urbanization scenario during the Long Rains. Slice created at 1°S latitude. The lake and mountain ranges are located between the points 120-230.

Figure 4.26 Surface circulation and temperature for the control simulation (left column) and high urbanization scenario (right column) simulation and the high urbanization scenario during the MAM season.
Figure 4.27 Change in precipitation between the control simulation and the high urbanization scenario during the Short Rains. ........................ 141

Figure 4.28 Change in relative humidity between the control simulation and the high urbanization scenario for the Short Rains. Slice created at 1°S latitude. The lake and mountain ranges are located between the points 120-230. .......................... 142

Figure 4.29 Circulation and temperature the control simulation (left column) and high urbanization scenario (right column) simulation and the high urbanization scenario during the OND season. ....................... 144

Figure 4.30 Probability density function for the precipitation for the during the 2 wet seasons for WRF (a) MAM (<10mm) (b) MAM (10mm-30mm) (c) OND (<10mm) (d) OND (10mm-30mm). ........................ 145

Figure 4.31 Change in precipitation between the control simulation and the high urbanization scenario during (a)MAM (b) OND for RegCM. ........................ 146

Figure 4.32 Probability density function for the precipitation for the during the 2 wet seasons for RegCM (a) MAM (<10mm) (b) MAM (10mm-30mm) (c) OND (<10mm) (d) OND (10mm-30mm). ........................ 147

Figure 4.33 Change in surface energy fluxes for the low urbanization scenario during the long and short rains (a) MAM latent heat (b) MAM sensible heat (c) OND latent heat (d) OND sensible heat. ............ 149

Figure 4.34 Change in surface temperature for the low urbanization scenario during (a) MAM (b) OND. .................................................. 150

Figure 4.35 Change in precipitation for the low urbanization scenario during (a) MAM (b) OND. .................................................. 151

Figure 5.1 Precipitation difference between the urbanized control and the sensitivity simulations with Lake Victoria included (a) albedo (b) soil moisture (c) surface roughness. .............................. 156

Figure 5.2 Diurnal temperature variation of temperature in urbanized region for the urban control and the sensitivity where the albedo is increased. 157

Figure 5.3 Precipitation difference between the urbanized control and the sensitivity simulations with Lake Victoria is converted to vegetated land (a) albedo (b) soil moisture (c) surface roughness. ............ 158

Figure 5.4 Circulation over the region during at 15UTC over the region from the different sensitivity simulations. (a) urban (b) albedo (c) soil moisture (d) surface roughness. .............................. 159

Figure 5.5 Relative humidity over the region at 15UTC for (a) urban with lake (b) albedo with the lake removed. The mountain ranges and the lake are between 90-120. .............................. 160

Figure 5.6 Circulation over the region during at 00UTC over the region from the different sensitivity simulations. (a) urban (b) albedo (c) soil moisture (d) surface roughness. .............................. 161
Figure 5.7  Relative humidity over the region at 00UTC for (a) urban with lake
(b) albedo with the lake removed. The mountain ranges and the lake are between 90-120. .......................... 163

Figure 6.1  Africa agriculture exports and percentage of contribution to total
exports (Blein et al., 2013). .......................... 166

Figure 6.2  Onset and cessation plot for the Long Rains season and the Short
Rains season. .......................... 170

Figure 6.3  Consecutive dry days in the season from the observational and the
simulated datasets during the Long Rains. .......................... 171

Figure 6.4  Consecutive 5 wet day periods in the season from the observational
and the simulated datasets during the Long Rains. .......................... 171

Figure 6.5  Maximum daily precipitation in the season from the observational
and the simulated datasets during the Long Rains. .......................... 173

Figure 6.6  Number of dry days in the season from the observational and the
simulated datasets during the Short Rains. .......................... 175

Figure 6.7  Number of 5 consecutive wet days in the season from the observa-
tional and the simulated datasets during the Short Rains. .......................... 176

Figure 6.8  Maximum daily precipitation in the season from the observational
and the simulated datasets during the Short Rains. .......................... 177

Figure 6.9  Maximum daily precipitation in the season for the control and sen-
sitivity simulation during the Long Rains (a) Control (b) High urbanization scenario. .......................... 179

Figure 6.10 Maximum Daily Precipitation in the season for the control and sen-
sitivity simulation during the Long Rains (a) Control (b) High urbanization scenario. .......................... 180

Figure 6.11 Number of dry days in the season from the WRF control and the
high urbanization scenario during the Short Rains. .......................... 181

Figure 6.12 Number of dry days in the season from the control and the high
urbanization scenario during the Long Rains. .......................... 182

Figure 7.1 Sales lost and number of hours different countries have without
electricity. The red rectangle highlights the East African countries. .......................... 185

Figure 7.2 Wind turbine output with steady wind plot. The cut in speed is
the speed at which the turbine will start generating power, and the cut out speed occurs when the turbine shutdown due to excessively
high wind speed. .......................... 188

Figure 7.3 Shape parameters for the different years (a) 2006 (b) 2007 (c) 2008
(d) 2009 (e) 2010 from the WRF model. .......................... 190
Figure 7.18 JJAS distribution functions wind speed at 10m, 30m and 100m vertical level (a) Probability distribution function at 10m (b) Cumulative probability distribution function at 10m (c) Probability distribution function at 30m (d) Cumulative probability distribution function at 30m (e) Probability distribution probability distribution function at 100m (f) Cumulative probability distribution function at 100m.

Figure 7.19 October-December wind speed and wind power density spatial distribution at 10m (a) 00 UTC wind speed (b) 00 UTC power density (c) 06UTC wind speed, (d) 06UTC power density (e) 12 UTC wind speed (f) 12 UTC power density (g) 18UTC wind speed (h) 18UTC power density for WRF model.

Figure 7.20 October-December wind speed at 10m from ERA-Interim (a) 00 UTC (b) 06UTC (c) 12UTC (d) 18UTC.

Figure 7.21 Same as Figure 7.19 but for the 30m hub height level.

Figure 7.22 Same as Figure 7.19 but for the 100m hub height level.

Figure 7.23 October-December distribution functions wind speed at 10m, 30m and 100m vertical level (a) Probability distribution function at 10m (b) Cumulative probability distribution function at 10m (c) Probability distribution function at 30m (d) Cumulative probability distribution function at 30m (e) Probability distribution probability distribution function at 100m (f) Cumulative probability distribution function at 100m for WRF model.

Figure 7.24 January-February wind speed and wind power density difference plots (a) 00 UTC wind speed (b) 00 UTC power density (c) 06UTC wind speed, (d) 06UTC power density (e) 12 UTC wind speed (f) 12 UTC power density (g) 18UTC wind speed (h) 18UTC power density for WRF model using the high urbanization scenario.

Figure 7.25 March-May wind speed and wind power density spatial distribution at 30m wind land cover changes (a) 00 UTC wind speed (b) 00 UTC power density (c) 06UTC wind speed, (d) 06UTC power density (e) 12 UTC wind speed (f) 12 UTC power density (g) 18UTC wind speed (h) 18UTC power density for WRF model.

Figure 7.26 June-September wind speed and wind power density differences between the high urbanization scenario and the control at (a) 00 UTC wind speed (b) 00 UTC power density (c) 06UTC wind speed, (d) 06UTC power density (e) 12 UTC wind speed (f) 12 UTC power density (g) 18UTC wind speed (h) 18UTC power density for WRF model for the 30m height.
Figure 7.27 October-December difference plot for the wind speed and the wind density at 30m wind land cover changes (a) 00 UTC wind speed (b) 00 UTC power density (c) 06UTC wind speed, (d) 06UTC power density (e) 12 UTC wind speed (f) 12 UTC power density (g) 18UTC wind speed (h) 18UTC power density for WRF model for the control and high urbanization scenario.

Figure 7.28 Changes in the probability distribution function at 30m hub height for the different seasons in the WRF model (a)JF (b) MAM (c) JJAS (d) OND in the low urbanization scenario.

Figure 7.29 Changes in the probability distribution function at 30m hub height for the different seasons in the WRF model (a)JF (b) MAM (c) JJAS (d) OND in the high urbanization scenario.

Figure A.1 Spatial Differences in temperature for during the Short Rains season for some selected years for Grell FC scheme

Figure A.2 Spatial Differences in temperature for during the Long Rains season for each year for some selected years for the Grell FC scheme

Figure A.3 Spatial Differences in temperature for during the Short Rains season for some selected years for MIT-Emanuel scheme

Figure A.4 Spatial Differences in temperature for during the Long Rains season for selected years for MIT-Emanuel scheme

Figure B.1 Precipitation distribution obtained from the different physics combinations used in the simulation for Domain 1

Figure C.1 Average precipitation for the short rain season for WRF (left column), REGCM (middle column) and TRMM (right column)

Figure C.2 Average precipitation for the Long rains season for WRF (left column), REGCM (middle column) and TRMM (right column)

Figure C.3 Average precipitation for the Short Rains season for WRF (left column), and TRMM (right column) for Domain 1

Figure C.4 Average precipitation for the Long Rains season for WRF (left column), and TRMM (right column) for Domain 1
Chapter 1

Introduction

1.1 Projected Urbanization Levels over East Africa

In the past centuries humans have modified the land cover and the land use through activities such as agriculture, forestry, energy production, settlement and industry. Settlement in the form of urbanization is classified as one of the extreme changes on land cover and land use (Vitousek et al., 1997; Wang et al., 2012a), with population increase, migration, and industrialization being the prime activities that have significantly led to the expansion of urbanized regions. The rate of urbanization is expected to increase rapidly in the next decades in most developing regions, with Africa being expected to have the highest increase. It is projected that Africa will have a 590% increase in urbanization extent by 2030 compared to its 2000 levels (Seto et al., 2012; United Nations, 2012). The probability increases estimates in the urban extent were obtained by combing population and gross domestic product projections and current population densities. In Africa the increase in population in urbanized regions will mostly be due to the migration from rural areas (United Nations, 2014).

Increases in urbanization levels in Africa will be high in regions that are already urbanized and in regions that are in close proximity to water bodies. These regions include along the Nile river in Egypt, Coast of West Africa on the Gulf of Guinea, the Northern shores of Lake Victoria in Kenya, Uganda and extending into Rwanda and Burundi, the Kano region in Northern Nigeria and greater Addis Ababa, Ethiopia (Fig. 1.1). These regions have been highlighted in (Fig. 1.1). In this study we focus on Lake Victoria Basin (LVB).
Figure 1.1 Projected urbanization growth by 2030 over 2000 levels (Seto et al., 2012).

The LVB region is shown in Fig. 1.2a, it has a total area of 251 000km$^2$ (Lake Victoria Basin Commission et al., 2012) and is the nerve centre of the region. LVB supports livelihoods in five countries in East Africa through, fisheries, hydro-electric power, transportation and agriculture. The region has high probability of being urbanized by 2030 compared to the other regions in East Africa. Currently the population is over 35 million (Lake Victoria Basin Commission et al., 2012) and in 2030 the population will likely be over 105 million and hence exerting more pressure on the already stressed ecosystem around the lake (Swallow et al., 2009). Figure 1.2b illustrates the urban expansion over Lake Victoria Basin in 2030. The black dots on the map indicate the current levels of urbanization in the region, the yellow and red indicates the probability of the region being urbanized by the year 2030, with red showing a higher probability.

The urbanization rate over East Africa in the past decades has been growing steadily at a rate of approximately 3.5% per annum (African Union Commission et al., 2012), however with a projected economic growth, this rate is expected to surge. Despite the
steady growth in the economy, the urbanization levels and the economic growth over the past decades in Africa have not followed the trend that developed countries had, Africa has had a slower economic growth compared to the urbanization growths (Annez and Buckley, 2009).

1.2 Impacts of Urbanization on Regional Climate

Urbanization accounts for a small percentage of the total global land cover, approximately 3% in 2010 (Liu et al., 2014) but it has been shown that it has profound effects on the ecosystem and the climate of a region (Grimm et al., 2008). The impact of urbanization on regional and local climate has caught attention from researchers over the past three decades and two main methods have been used to investigate its impacts. These two main methods include using the Regional Climate Models (Bounoua et al., 2013; Comarazamy et al., 2013; Karaca et al., 1995; Shem and Shepherd, 2009; Wang et al., 2012a; Wenzheng et al., 2014; Zhang et al., 2014) and also observational datasets (Burian and Shepherd, 2005; Hand and Shepherd, 2009; Kishtawal et al., 2010; Oke, 1980). The
studies have shown coherent agreement on the impacts of urbanization on temperature, surface energy balance, the planetary boundary layer, however there is still uncertainty on how it will affect the precipitation. While Miao et al. (2011); Zhang et al. (2014) have found an increase in precipitation, Guo et al. (2006); Zhao et al. (2006) have found a decrease in the amount of precipitation. In addition to the impact on total precipitation studies in the midlatitudes (Argüeso et al., 2014; Daniels et al., 2014; Karaca et al., 1995; Mahmood et al., 2014; Miao et al., 2011; Yang et al., 2014a; Zhang et al., 2014) have shown that urbanized regions not only alter the amount of precipitation but also change the intensity, distribution and the frequency.

Urbanization influences the local and regional climate systems via biogeochemical and biogeophysical processes. Biogeochemical process alter the climate through changes in the concentration of the greenhouse gases. Human activities such as deforestation and agricultural activities have been linked with an increase in the amount of emitted greenhouse gases such as carbon dioxide, methane and nitrous oxide into the atmosphere. This study will focus on the biogeophysical processes which alter the physical features of the earth’s surface such as surface albedo, soil moisture, surface roughness, evapotranspiration (Fig. 1.3) hence altering the surface energy balance, momentum fluxes between the surface and overlaying atmosphere and moisture budget (Coutts et al., 2007; Lemonsu et al., 2004).

The surface energy balance plays a key role in influencing the dynamics and thermodynamics above the land surface (Guo et al., 2006; Shem and Shepherd, 2009; Shepherd et al., 2010; Van Den Heever and Cotton, 2007). Urbanized regions experience an increase in the sensible heat flux and ground flux while the latent heat flux decreases. The changes in the sensible heat and latent heat have been attributed to the increased short-wave absorption and reduced long wave emission in the urban areas. Zhao et al. (2013) attributed the increase in ground flux to the limited availability of surface water, which in turn reduces evaporative cooling in urban areas, and limits the heat released from the surface to the atmosphere.

Built up areas have impervious surfaces that increase surface runoff and hence reduces the infiltration rate which leads to a reduction in the soil moisture (Chen and Dudhia,
Figure 1.3 Biogeophysical feedback processes in response to land cover change (Mölders, 2011).

(2001) and changes to the water table. The decreased vegetation cover and the decreased soil moisture will reduce the evapotranspiration over the region and hence affect the temperature and precipitation.

The urban heat island (UHI) effect (Kalnay and Cai, 2003; Karaca et al., 1995; Mahmood et al., 2014; Zhang et al., 2014) is one of the most studied impacts of urbanization. The UHI develops because of the altered partitioning of the latent and sensible energy fluxes. Urbanized regions have higher sensible heat compared to vegetated regions. Studies have shown that the UHI is typically strongest during the nocturnal part of the diurnal cycle. Previous studies on the impact of urbanization on temperature have shown a reduced diurnal temperature variation (Kalnay and Cai, 2003; Wang et al., 2012b). Two
main reason have been postulated to explain this that is:

1. The urbanized regions have a lower albedo compared to the vegetated areas, at night the temperature cools slower than vegetated land. The urbanized region during the nocturnal hours will start loosing the heat absorbed during the day, thus warming up the atmosphere.

2. Anthropogenic heating at night could also contribute to elevated minimum temperatures.

Urbanization in a region that is close to a water body can alter the local circulation. The temperature increase in these regions lead to altered temperature gradients between the lake and the land. The difference in the temperature between the land and the lake surfaces results in temperature differences of the atmosphere above. Yang et al. (2014a) showed that the thermodynamic perturbations by urbanization on temperature and surface pressure fields enhances the intrusion of the land lake breeze and facilitate the formation of a convergence zone. The convergence zone will then create favorable conditions for deep convection over the urbanized region.

A limited number of studies have been conducted to investigate the impact of changing land cover over East Africa region. Semazzi and Song (2001) showed that deforestation over the Congo Basin will lead to a reduction in precipitation over the region. Deforestation led to changes in the trapped Rossby wave train activity, hence causing different responses in precipitation over Southern Africa with other regions receiving less precipitation while others had increased precipitation. Otieno and Anyah (2012) showed that the conversion of forestland to agricultural land led to a reduction in precipitation over Kenya. They attributed the decrease in precipitation to altered biogeophysical properties like albedo, surface roughness, and evapotranspiration. Moore et al. (2015) used projected land cover developed using the Landform Transformation Model. The projected land cover consists of mainly agriculture land and a few regions are urbanized. Their results indicate that land cover changes had more complex consequences than the impact of greenhouse gases. Past studies in the region have mainly focused on the conversion of vegetated land to another vegetated land, but none to our knowledge to date has utilized the extremely high change of landcover from vegetated to impervious land. In addition the spatial scales that were utilized in the simulations were relatively coarse, 36km was
the finest resolution by Moore et al. (2015).

1.3 Research Hypothesis and Aim

Over the past decades there has been a general decrease in mean precipitation over East Africa (Lyon et al., 2014; Lyon and DeWitt, 2012; Yang et al., 2014b). However most of the General Circulation Models (GCM’s) indicate that there will be an increase in precipitation in the next decades (Fig. 1.4), this phenomena is now known as the East Africa Climate Paradox (Rowell et al., 2015; Semazzi et al., 2014). Figure 1.4 shows the past and projected trends of the Long Rains (March-April-May). The Coupled Model Intercomparison Project Phase 5 (CMIP5) project a gradual increase in the Longs Rains in the future.

![Figure 1.4: Projected increases in Precipitation over East Africa: Low pass filtered rainfall March-May Average over Greater Horn of Africa (Rowell et al. (2015)).](image)
Figure 1.5 shows the trends in the lake levels. The lake levels have been decreasing over the past decades and in the future it is projected that the lake levels will increase (Smith, 2011). Although GCM’s provide fundamental information on the future trend of precipitation, uncertainty in projections arises from the inadequate treatment of the effects of urbanization. Most GCM’s do not consider changes in urbanization since it comprises of a very small percentage of the global land cover. In addition to this GCM’s are run at a coarse resolution (greater than 100km) and is unable to capture the dynamics in urbanized regions. East Africa is an agricultural based economy with 43% of the Gross Domestic Product (GDP) from agriculture (Salami et al., 2010; Waithaka et al., 2013). Any changes in precipitation can affect the agricultural production significantly, with droughts and water logging having devastating impacts on the socio-economic livelihoods of individuals in the region.

Figure 1.5 Past (top, 1900-2010, Graph by Robert Simmon, based on data provided by the USDA) and Projected (2070-2100, Smith (2011)) (bottom) Lake Victoria levels trends.
Based on preliminary numerical simulations and previous studies, we hypothesize that initially precipitation will decrease over the urbanized regions due to the changes in surface friction, albedo and evapotranspiration associated with vegetation degradation. However, in subsequent decades precipitation over the urbanized regions of the basin will begin to increase as the urban heat island effect in combination with the urban land supply of moisture, and supply of moisture from the lake dominate over the land cover response. Moreso these changes will alter the hydro-climatological balance over Lake Victoria Basin in a complex way that must be determined by numerical simulations. We further hypothesize that the additional moisture for this increase will result in reduced rainfall over some upstream moisture supply regions, specifically over Lake Victoria. The changes in the hydro-climatological balance over LVB could have dire implications for hydroelectric generation potential. Agricultural activities which make the backbone of the economy will also be greatly affect by changes in precipitation. A reduction in precipitation will lead to withering of crops, while an increase might lead to water logging. The cascade and knock on factors and processes could result in significant climate change.

The aim of this research is to investigate the impact of urbanization on the distribution, frequency and intensity of precipitation in the region, the most dominating physical processes that leads to the changes and the subsequent consequences on the energy and agricultural sectors over Lake Victoria Basin. This will be achieved by accomplishing the following objectives

1. Comprehensive customization and evaluation of the performance of the two Regional Climate models that will be used in the analysis( i.e. RegCM4 and WRFV3) over East Africa. We choose 2 models in order to investigate the response to changes in urban land cover from different models. These models will serve as sensitivity simulations. The different cumulus, microphysics, land surface schemes and radiation schemes will be fully tested for suitability over the region. Satellite and gridded datasets will then be used to evaluate the temperature and precipitation over the region.

2. Perform an Empirical Orthogonal Function analysis on the RegCM4 and compare the modes with TRMM datasets in order to ascertain that the model adequately captures the inter-annual variability of precipitation over East Africa.
3. Investigate how different urbanization scenarios affect the climate in the region, we pay particular attention to the precipitation intensity, frequency and distribution over East Africa. In this case we use two urbanization scenarios, the high urbanization and the low urbanization scenario based on Seto et al. (2012).

4. Investigation of the most dominant process between, moisture supply, thermal convergence and frictional convergence

5. Analysis of the impact of urbanization on the precipitation and the application sectors: For Agriculture the impact of urbanization on the onset and cessation of the short rains will be studied.

6. Explore the feasibility and viability of the wind energy over East Africa, in addition to this sensitivity to urbanization will also be investigated.

1.3.1 Limitations of the Study

- Since the future land use will be given by scenarios, there is uncertainty on how much urbanization will grow. Rate of urbanization growth a depends on economic, social and political stability.

- The lake model will be one dimensional and thus will not capture the full 3-D dynamics of the lake.

- East Africa does not have adequate insitu precipitation instruments, therefore we highly depend on the gridded datasets which have a coarse resolution and also Satellite datasets which are bias corrected from the sparse insitu measurements. Sylla et al. (2013) found that the observational datasets had biases in terms of total precipitation, however the correlation was very good. Thus in analyzing the results there is an associated error from the dataset used in the analysis.
Chapter 2

Evaluation of the Performance ICTP Regional Climate Model

2.1 Introduction

Regional climate models (RCM) have over the past decades been used for a number of different climate studies including but not limited to process studies, dynamical downscaling for climate change assessment, seasonal climate predictions and regional climate predictability. A number of these models have been developed for use in the midlatitudes and are not fully tested over the tropical regions, yet most of the countries in these regions depend mostly on rainfed activities. In this section we focus our analysis on tropical East Africa. East Africa has complex terrain that was formed due to tectonic movements. It is host to a number of lakes, commonly known as the East Africa Great lakes. The lakes in the region have a vast range of sizes, with Lake Victoria which is both large and shallow, Lake Tanganyika and Lake Malawi, which are deep, Lake Edward, Lake Albert and Lake Kyoga are relatively smaller. The population in the region is over 150 million people, and approximately 80% of the population depends on rainfed activities (EAC Secretariat, 2006). With the development of the RCM and rapid advances in computational power one would expect extensive use of regional climate models in this region, nevertheless a relatively small number of studies have been conducted over the region (Anyah and Semazzi, 2007; Anyah et al., 2006; Davis et al., 2009; Diro et al., 2012; Indeje et al., 2001; Ogwang et al., 2014; Segele et al., 2009; Sun et al., 1999; Sylla et al., 2010; Zaroug et al., 2014). Most studies in this region have been done using the International Centre for Theoretical Physics (ICTP) Regional Climate Model (RegCM) with only a limited number done using other RCMs for example the Weather Research and Forecasting model (WRF), Regional Atmospheric Modeling System (RAMS), Consortium for Small-scale
Modeling-Climate Limited-area Modeling (COSMO-CLM). Emphasis on this section will be on the performance of RegCM.

Sun et al. (1999) used RegCM2 to find the optimal physics combinations that can reproduce the precipitation pattern and distribution over the region. Their main focus was on the role that is played by cumulus convection schemes, non convective precipitation, radiative transfer formulation, surface processes, boundary layer formulation, and lateral boundary conditions. They found that adjusting the rainfall efficiency (default 0.3 to 0.7) in both Grell FC and Grell AS to 0.9 improved the precipitation over the region considerably. Spatially and statistically their study showed that Grell AS had a better representation of precipitation distribution and intensity in the region.

Anyah and Semazzi (2007) performed simulations with RegCM3, that included a more advanced large scale precipitation scheme (SUBEX) and a more advanced radiation scheme, they found that the model had a tendency to over predict precipitation with the Grell FC scheme. Davis et al. (2009) customized the model by tuning the microphysics parameters, enlarging the model domain towards the Indian Ocean and also testing the performance of a new convective scheme Massachusetts Institute of Technology-Emanuel (MIT-Emanuel). They found that the MIT-Emanuel scheme provided a more realistic spatial distribution of precipitation over the region. The MIT-Emanuel scheme also produced a more realistic partitioning of stratiform and convective rainfall. Overestimation of precipitation was ameliorated by reducing the relative humidity and the accretion rate in the microphysics scheme. Segele et al. (2009) also tested the performance of the MIT-Emanuel scheme over East Africa, however they tuned parameters in the cumulus scheme, to reduce the wet bias. They performed several sensitivity experiments by varying the key parameters that control precipitation, these include the rate of convective mass flux, the fraction of condensed water converted to precipitation, and the heating and moistening characteristics of the environment. The experiments showed that using an auto conversion rate of 0.01 in the cumulus parameterization significantly reduces the precipitation overestimation over the region.

Sylla et al. (2010) performed a simulation over the African domain, to examine the ability of RegCM3 to reproduce the seasonal mean climatologies, annual cycle and inter
annual variability. They found that although the model had positive biases over their Northern and Southern Equatorial domain, the model was able to capture the annual cycle of the precipitation in the region. Utilizing the tropical band in RegCM4, Zaroug et al. (2014) found that the simulations gave a general dry bias over the Congo Forest, and minimum bias errors of precipitation over Lake Victoria Basin.

Recently Ogwang et al. (2015) evaluated the ability of RegCM4 in reproducing the annual cycle and the June to August seasonal climatology over East Africa. Their findings showed that the model performs well in all the regions, however the Grell FC had a tendency to underestimate total precipitation, while the MIT-Emanuel overestimated the total precipitation. Their sensitivity experiments to model domain size showed that the MIT-Emanuel scheme performed better in their small domain which just included a small proportion of the Indian Ocean compared to the larger domain which included a bigger proportion of the Indian and Atlantic Ocean.

The latest version of RegCM model has been improved to include another resolved large scale precipitation scheme Nogherotto/Tompkins, provision, to use a mixed cumulus scheme, improved radiation scheme and land surface scheme. This study will start by customizing the model using the default configurations and then using the optimal combinations obtained in Davis et al. (2009). Further tests will be done on the land surface scheme and also the large scale precipitation schemes. The aim of this study is to find the optimal physics combinations available in RegCM 4.4 in reproducing the precipitation distribution over the region.

### 2.2 Model Description and Experimental Setup

RegCM4 is a hydrostatic, compressible, three-dimensional, regional climate model (Giorgi and Anyah, 2012; Giorgi et al., 2012) which comprises of a number of physics parameterization. The model has two planetary boundary layer (PBL) schemes; the Holtslag (Holtslag et al., 1990), and the University of Washington (UW) scheme (Bretherton et al., 2004). The Holtslag scheme has been tested in different latitudes, while the University of Washington scheme has only been mostly tested in the midlatitudes (Giorgi et al., 2012) hence care must be taken when using this scheme in tropical latitudes.
The Biosphere-Atmosphere Transfer Scheme (BATS) (Dickinson et al., 1993) and the Community Land model (CLM) (Steiner et al., 2005; Tawfik and Steiner, 2011) are the two land surface schemes available in the model. The BATS scheme is a 1-layer vegetation module, a 1-layer snow module, a force-restore model for soil temperatures, a 3-layer soil scheme, and a simple surface runoff parameterization. The scheme includes 22 land surface types and 12 soil color and soil texture types. The surface types include a suburban and urban regions. The parameters for the urbanized land use types were obtained from Table 1 of Kueppers et al. (2008). CLM contains five possible snow layers with an additional representation of trace snow and ten unevenly spaced soil layers with explicit solutions of temperature, liquid water and ice water in each layer. To account for land surface complexity within a climate model grid cell, CLM uses a mosaic approach to capture surface heterogeneity. The ability to account for land surface heterogeneity is essential over East Africa since the land cover in this region is highly inhomogeneous.

Six options are available for representing cumulus convection in the current version of RegCM and these are the Kuo-type, Grell, MIT-Emanuel, Kain Fritsch, Betts Miller and the Tiedtke. The Betts Miller currently is not working according to the source code and there are very limited studies if any that have used the Kain Fritsch scheme. The Kuo scheme (Anthes et al., 1987) activates convection when the column moisture convergence exceeds a threshold value. The scheme is known to produce the poorest rainfall distribution and intensity compared to all other schemes (Giorgi et al., 2012) hence this scheme will not be tested in this study. In the Grell scheme (Grell, 1993) clouds are considered as 2 steady state circulations including an updraft and a penetrative downdraft where no entrainment and detrainment occurs at the edges of the cloud and only allowed at the top or the bottom of the cloud. Two different closures for Grell scheme are available, an Arakawa-Schubert type closure in which all buoyant energy is immediately released at each time step and a Fritsch-Chappell type closure in which the available buoyant energy is released with a time scale typically of the order of 30 minutes. The Tiedtke scheme which was initially introduced for use in global simulations depends on the mass flux and moisture convergence (Tiedtke, 1989) and the convention is triggered when the lifted air parcel reaches the lifting condensation level. The most complex scheme in the model is the MIT-Emanuel (Emanuel and Živkovic-Rothman, 1999), the scheme simu-
lates more realistic atmosphere cloud dynamics compared to the other models. In the MIT-Emanuel parameterization, convection is triggered when the level of buoyancy is higher than the cloud base level. The scheme assumes that the cloud mixing is episodic and inhomogeneous, and convective fluxes modeling is based on sub-cloud-scale updrafts and downdrafts which are idealized (Giorgi et al., 2012). Precipitation is based on auto-conversion of cloud water into rain water and accounts for simplified ice processes.

Subgrid Explicit Moisture Scheme (SUBEX) and Nogherotto/Tompkins are the two large scale precipitation schemes in RegCM. The SUBEX scheme partitions each grid cell into a cloudy and non-cloudy fraction related to the average grid cell relative humidity (Pal et al., 2000). The model also includes simple formulations of raindrop accretion and evaporation. When a cloud water threshold is exceeded precipitation occurs. The specification of this threshold is based on empirical in-cloud observations of cloud liquid water amounts. Past studies have shown that the use of SUBEX as the microphysics scheme in RegCM3 (Davis et al., 2009) has a tendency to produce excessive precipitation. Tuning the minimum relative humidity, auto-conversion rate term and the rate of sub-cloud evaporation parameters have proven effective in reducing the excessive precipitation (Davis et al., 2009; Giorgi et al., 2012). Recently a new microphysics scheme Nogherotto/Tompkins (Nogherotto et al., 2012) has been implemented in RegCM4. The Nogherotto/Tompkins scheme is built upon the European Centre for Medium Weather Forecasts Integrated Forecast System. Major improvements in the model include the independent existence of liquid and ice water content, this allows for the existence of supercooled liquid water and mixed phase cloud. The rain and snow precipitate with a fixed, finite, terminal fall speed and can be then advected by the three dimensional wind. Currently the scheme has not yet been well tested in the model.

Other physics schemes in the model include the 1-dimensional Hostetler lake model, 2 pressure gradient schemes, chemistry model and 2 radiation schemes the NCAR CCM3, and the Rapid Radiative Transfer Model (RRTM). The RRTM uses the correlated-k method, and divides the longwave spectral region into 16 bands based on their homogeneity of contributing species and radiative transfer properties.

We conduct a series of simulations for the Short Rains in the year 2000 using a hor-
horizontal resolution of 36 km and 18 vertical layers. In some cases the number of vertical layers was increased to 23 as the model was becoming unstable with the use of 18 vertical layers. The spin-up time for all the simulations is 3 months thus the simulations were run from July to December. All simulations were done with the 1 dimensional lake model activated, and setup to be run in the tropics. The experiments were done with the initial and lateral boundary conditions obtained from the ERA-Interim gridded reanalysis data at 1.5° resolution and the NCEP reanalysis dataset. The sea surface temperatures are obtained from the National Oceanic and Atmospheric Administration. The Optimum Interpolated Sea Surface Temperature (OISST) is produced weekly on a 1° resolution. Both the SST and lateral boundary conditions are updated every 6 hours. The different experiments are mostly focused on the ability of the model to reproduce the rainfall distribution and intensity over the region. We tune different parameters in the SUBEX scheme as suggested in past research (Davis et al., 2009; Giorgi et al., 2012). Most experiments are performed using the MIT-Emanuel scheme and the SUBEX microphysics scheme. The following simulations were conducted for the 2000 (Table 2.1):
Table 2.1 Summary of the customization experiments conducted for the RegCM model using a 36 km resolution.

<table>
<thead>
<tr>
<th>Exp</th>
<th>IC</th>
<th>Microphysics</th>
<th>Cumulus</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>EIN15</td>
<td>SUBEX Default</td>
<td>MIT-Emanuel</td>
</tr>
<tr>
<td>b</td>
<td>EIN15</td>
<td>SUBEX, RH=0.6</td>
<td>MIT-Emanuel</td>
</tr>
<tr>
<td>c</td>
<td>EIN15</td>
<td>NT, Default</td>
<td>Grell FC over land and MIT-Emanuel over oceans</td>
</tr>
<tr>
<td>d</td>
<td>EIN15</td>
<td>SUBEX, RH=0.65</td>
<td>Grell FC over land and MIT-Emanuel over oceans</td>
</tr>
<tr>
<td>e</td>
<td>EIN15</td>
<td>NT, Default</td>
<td>Grell FC over ocean and MIT-Emanuel over land</td>
</tr>
<tr>
<td>f</td>
<td>NCEP</td>
<td>NT, Default</td>
<td>MIT-Emanuel</td>
</tr>
<tr>
<td>g</td>
<td>NCEP</td>
<td>NT, autoconversion threshold divided by 2</td>
<td>MIT-Emanuel</td>
</tr>
<tr>
<td>h</td>
<td>NCEP</td>
<td>NT, autoconversion threshold halved and rainfall speed reduced from 4 to 3 m/s</td>
<td>MIT-Emanuel</td>
</tr>
<tr>
<td>i</td>
<td>NCEP</td>
<td>SUBEX, RH=0.60 autoconversion threshold halved</td>
<td>MIT-Emanuel</td>
</tr>
<tr>
<td>j</td>
<td>NCEP</td>
<td>SUBEX, RH=0.50 autoconversion threshold halved, accretion rate over land reduced to 2 form 3</td>
<td>MIT-Emanuel over land and Grell FC over ocean</td>
</tr>
<tr>
<td>k</td>
<td>NCEP</td>
<td>SUBEX, RH=0.60 autoconversion threshold halved, accretion rate over land reduced to 2.5 form 3</td>
<td>MIT-Emanuel over oceans and Grell FC over land</td>
</tr>
<tr>
<td>l</td>
<td>NCEP</td>
<td>SUBEX, RH=0.80 autoconversion threshold halved, accretion rate over land reduced to 2.5 form 3</td>
<td>MIT-Emanuel over oceans and Grell FC over land</td>
</tr>
<tr>
<td>m</td>
<td>NCEP</td>
<td>SUBEX, RH=0.80 autoconversion threshold halved, accretion rate over land reduced to 2.5 form 3</td>
<td>MIT-Emanuel</td>
</tr>
<tr>
<td>n</td>
<td>NCEP</td>
<td>SUBEX, RH=0.80 autoconversion threshold halved, accretion rate over land reduced to 2.5 form 3</td>
<td>MIT-Emanuel over oceans and Grell FC over land 9 months spin up</td>
</tr>
</tbody>
</table>
2.2.1 Model Validation

A number of statistical metrics can be performed in order to validate a climate model. Pielke Sr (2013) proposed the use of the standard deviation, and the Root Mean Square Error (RMSE), which are calculated as follows:

\[ RMSE = \left( \sum_{i=1}^{N} (\phi_i - \phi_{iob})^2/N \right)^{1/2} \]  \hspace{1cm} (2.1)

where \( N \) is the number of observation, \( \phi_i \) and \( \phi_{iob} \) are the individual prediction and observations at the same grid point.

\[ \sigma_{iob} = \left( \sum_{i=1}^{N} (\phi_{iob} - \phi_{0ob})^2/N \right)^{1/2} \]  \hspace{1cm} (2.2)

\( \sigma_{iob} \) is the standard deviation of the observed data, \( \phi_{0ob} \) is the average value. In our analysis we use the Tropical Rainfall Measuring Mission (TRMM) as our observational data. The data is available at a resolution of 0.25° from 1998 to 2015. We use the monthly dataset in our analysis. The standard deviation for modeled data is calculated in the same way as that of the observed data. In order to assess the skill of the model the following criterion has to be satisfied:

1. Standard deviation of the observed data should be approximately equal to the standard deviation of the modeled/predicted data

2. The root mean square error should be less than the standard deviation of the observed datasets.

Other studies have also used the mean bias, normalized mean error (NMError) and the correlation in evaluating models. The mean bias is calculated as follows:

\[ Mean \ Bias = \sum_{i=1}^{N} (\phi_i - \phi_{iob})/N \]  \hspace{1cm} (2.3)

the NMError as

\[ NMError = \frac{\sum_{i=1}^{N} |(\phi_i - \phi_{iob})|}{\sum_{i=1}^{N} \phi_{iob}} \times 100\% \]  \hspace{1cm} (2.4)
In addition to these statistics we use the Taylor diagram (Taylor, 2001) to summarize the model performance. In constructing the Taylor diagram we calculate the normalized root mean square errors, standard deviations and the correlation. The construction of the Taylor diagram is based on the Law of Cosines:

\[ c^2 = a^2 + b^2 - 2ab \cos \phi \]  

(2.5)

The ratio of the variances and the correlation are then used to construct the diagram. The ratio of the variances are used to measure the ability of the model to capture the variability in observations, and the correlation indicates whether the model and the observations have the same patterns. The statistics are related by the following equation

\[ E^2 = 1 + \alpha^2 - 2\alpha R \]  

(2.6)

where E is the normalized root mean square error, \( \alpha \) is the ratio of the standard deviations

\[
\frac{\text{standard deviation of model}}{\text{standard deviation of the observations}}.
\]

(2.7)

The model will be most accurate when they are close to the reference value, that is when the ratio of the standard deviations and correlation are close to 1 and the normalized root mean square error approaches zero.

### 2.3 Results and Discussion

Considering that this version of RegCM has undergone a number of developments since the last comprehensive customization performed by Davis et al. (2009) using RegCM3 over the region of interest, the first simulation run was done with the default parameters in RegCM4. The default settings in RegCM4 tends to overestimate precipitation over land however there is underestimation of precipitation over the Indian Ocean (Fig. 2.1(a)) when the MIT- Emanuel cumulus scheme is utilized. Although the model overestimated the precipitation it does capture the essential distribution of precipitation over the region. The excess precipitation problem is consistent with other past studies (Davis et al., 2009; Elguindi et al., 2013; Segele et al., 2009). Thus since the model has a wet bias we tune parameters that will minimize the total biases, this is consistent with the studies that
have been conducted Segele et al. (2009) who tuned parameters in the MIT-Emanuel scheme and Davis et al. (2009) who had the microphysics parameters tuned.

A number of parameters that are sensitive to the total precipitation have been proposed and these include the relative humidity, evaporation, auto conversion rate, and the accretion rate. This study will proceed by using the parameters and optimal physics options used in Davis et al. (2009). First we test how sensitive precipitation is to reduced relative humidity over land. In the SUBEX scheme the fractional cloud cover is related to the relative humidity by the following equation

\[ FC = \sqrt{\frac{RH - RH_{\text{min}}}{RH_{\text{max}} - RH_{\text{min}}}} \] (2.8)

where \( RH_{\text{min}} \) is the cloud formation threshold, and \( RH_{\text{max}} \) is the maximum saturation. Clearly from this equation a reduction in the minimum relative humidity would lead to a reduction in the fractional cloud cover and hence a reduction in the total precipitation. Figure 2.1(j) uses the lowest RH and it has the weakest precipitation. Reducing just the accretion rate, and keeping the RH at 80% over land and 90% over the oceans, slightly reduces the rainfall bias over the region. Reducing both the accretion rate and the relative humidity weakens the intensity. In addition to the MIT-Emanuel scheme we also test the performance of the Grell FC scheme. Reducing the relative humidity when using the Grell FC scheme tends leads to further underestimation of total precipitation.

We also investigate the performance of the newly implemented large scale precipitation scheme (Nogherotto/Tompkins). The physics combination has a wet bias, as with the SUBEX scheme when the MIT-Emanuel cumulus scheme is used (Fig. 2.1(c-h)). However the bias is lower than in the SUBEX scheme. We proceed as before and tune parameters that are sensitive to moisture budgets. In this case we reduce the auto-conversion rate in the Nogherotto/Tompkins scheme. A reduction of the autoconversion scale factor leads to reduced precipitation. In addition to this we also reduce the speed of fall of rain. The lower the speed of fall of rain the higher the likely-hood for small rain droplets to evaporate before reaching the ground. Using this physics combination and tuned parameters there is a reduction in total precipitation bias over the region (Fig. 2.1(i)). In all the simulations the Ocean has weak precipitation, and the land has intense precipitation.
In calculating the statistics the boundaries were trimmed, thus the statistics does not include biases from the lateral boundaries. The statistics reveal the ability of the model to capture the spatial distribution pattern of total precipitation as all the correlations are relatively high (Table 2.2). The correlations between the different simulations and TRMM is greater than 0.65, with most simulations having a correlation greater than 0.70. The default settings for the SUBEX and MIT-Emanuel (experiment (a)) scheme have a correlation that is relatively high and is approximately 0.72 while that of Nogherotto/Tompkins and MIT-Emanuel (experiment e) is approximately 0.76. Thus the two large scale precipitation schemes perform well over the region in terms of reproducing the spatial distribution. The highest correlation is with experiment (c) where the Grell FC is used over land and the MIT-Emanuel over the oceans. The lowest correlation occurs with the same mixed cumulus schemes (experiment k), as mentioned before reducing the relative humidity for the Grell FC scheme leads to further dry bias. Hence over-tuning can be detrimental as it can lead to higher biases than the original biases that were in the default settings. The normalized mean bias error ranges from -15 to 60%, with the mixed convective scheme using Grell FC over land showing the driest bias and the MIT -Emanuel scheme with the Nogherotto/Tompkins having the highest normalized wet bias. The Grell FC scheme releases extra buoyant energy every 30 minutes and hence the simulations utilizing this cumulus scheme tends to underestimate the total precipitation. On the other hand the MIT-Emanuel scheme has a wet bias because once it is activated it produces very intense precipitation events.

Figure 2.2 shows the standard deviation of the observations and also the RMSE. The RMSE measures the spread in the errors between the model and the observations. The red line in the figure indicate the standard deviation of TRMM. In this case the seasonal RMSE is highest when the default setting of MIT-Emanuel and SUBEX schemes are utilized. Davis et al. (2009) proposed that the wet bias when the MIT-Emanuel scheme is utilized could be due to the enhanced low level moist westerlies from the Atlantic ocean that increases instability over the region. When tuning is done the seasonal RMSE is greatly reduced. Just like the RMSE the standard deviation of the model is higher in the case when the parameters are not tuned, with the default simulation having a standard deviation of approximately 500. This value is very high when compared with a
Figure 2.1 Precipitation distribution using different physics combinations for the simulations run at a resolution of 36km.
Table 2.2 Summary of evaluation statistics for RegCM model at a 36km resolution for the different simulations. RMSE is the root mean square error, SD is the standard deviation of the model, NMB is the normalized mean bias error, NME is the normalized mean error.

<table>
<thead>
<tr>
<th>Exp</th>
<th>Correlation</th>
<th>RMSE</th>
<th>SD</th>
<th>NMB (%)</th>
<th>NME</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>0.72</td>
<td>423.58</td>
<td>503.57</td>
<td>56</td>
<td>0.52</td>
</tr>
<tr>
<td>b</td>
<td>0.72</td>
<td>322.92</td>
<td>421.52</td>
<td>33</td>
<td>0.46</td>
</tr>
<tr>
<td>c</td>
<td>0.78</td>
<td>239</td>
<td>355.48</td>
<td>19</td>
<td>0.44</td>
</tr>
<tr>
<td>d</td>
<td>0.77</td>
<td>174.59</td>
<td>263.19</td>
<td>-15</td>
<td>0.45</td>
</tr>
<tr>
<td>e</td>
<td>0.76</td>
<td>369.19</td>
<td>439.50</td>
<td>60</td>
<td>0.47</td>
</tr>
<tr>
<td>f</td>
<td>0.75</td>
<td>276.97</td>
<td>372.44</td>
<td>33</td>
<td>0.44</td>
</tr>
<tr>
<td>g</td>
<td>0.76</td>
<td>266.09</td>
<td>363.14</td>
<td>31</td>
<td>0.43</td>
</tr>
<tr>
<td>h</td>
<td>0.76</td>
<td>195.36</td>
<td>290</td>
<td>4</td>
<td>0.39</td>
</tr>
<tr>
<td>i</td>
<td>0.74</td>
<td>183.53</td>
<td>271.34</td>
<td>-0.6</td>
<td>0.39</td>
</tr>
<tr>
<td>j</td>
<td>0.72</td>
<td>168.02</td>
<td>222.89</td>
<td>-11.8</td>
<td>0.411</td>
</tr>
<tr>
<td>k</td>
<td>0.65</td>
<td>212.21</td>
<td>191.86</td>
<td>-37</td>
<td>0.76</td>
</tr>
<tr>
<td>l</td>
<td>0.73</td>
<td>197</td>
<td>286.13</td>
<td>-6</td>
<td>0.46</td>
</tr>
<tr>
<td>m</td>
<td>0.74</td>
<td>258.49</td>
<td>353.97</td>
<td>26</td>
<td>0.43</td>
</tr>
<tr>
<td>n</td>
<td>0.71</td>
<td>209.26</td>
<td>294.50</td>
<td>-8</td>
<td>0.50</td>
</tr>
</tbody>
</table>

standard deviation of 264 for TRMM. Thus according to Pielke Sr (2013) skill criterion, the physics combination does not have skill when the parameters are not tuned. Reducing the accretion rate in the SUBEX scheme greatly reduces the statistical errors with the Grell scheme showing the best statistics. Tuning of the model improves skill in the model with experiments d, h, i, l, n showing the highest skill. However the models now have high dry biases. We proceed to customizing the model with a relatively higher resolution and hypothesize that the coarse resolution has a tendency to overestimate the total precipitation in the domain.

2.3.1 Experiments with Higher Resolution

Although the experiments with the 36km resolution produced precipitation that reasonably reproduced the distribution over the land, there are still a lot of discrepancies in the total precipitation over the region. In the oceans the model greatly underestimates precipitation and over land the model over predicts precipitation in most regions. Thus
another suit of experiments was carried out on the sensitivity of precipitation to the parameterization of the cumulus and microphysics schemes, enhanced radiation schemes, land surface schemes and the planetary boundary layer scheme. We use the same time period as stated in the previous section (Short Rains of 2000) and the same domain with a resolution of 25km.

2.3.1.1 Cumulus Schemes

While the coarse resolution (36km) experiments mainly utilized MIT-Emanuel scheme, the Grell FC and the Grell AS schemes that were previously used in Davis et al. (2009) and Anyah and Semazzi (2007) will also be tested in this suite of experiments. The customization of the model begins by testing 7 combinations of cumulus parameterization. The different combinations are shown in Table 2.3. The study will be done by using the
SUBEX microphysics, BATS1e land surface, scheme, Holtslag planetary boundary layer and a relatively higher resolution of 25 km. As has been shown before, the resolution of a simulation does affect the total precipitation (Gao et al., 2008, 2006; Rauscher et al., 2010) thus we expect that with higher resolution the precipitation biases should reduce.

Table 2.3 Summary of RegCM simulations done to test the performance of the different cumulus schemes. The land/ocean indicate which schemes were used over land/ocean in the simulation.

<table>
<thead>
<tr>
<th>Exp</th>
<th>Land</th>
<th>Ocean</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Grell AS</td>
<td>Grell AS</td>
</tr>
<tr>
<td>B</td>
<td>Grell AS</td>
<td>MIT-Emanuel</td>
</tr>
<tr>
<td>C</td>
<td>MIT-Emanuel</td>
<td>Grell AS</td>
</tr>
<tr>
<td>D</td>
<td>Grell FC</td>
<td>Grell FC</td>
</tr>
<tr>
<td>E</td>
<td>Grell FC</td>
<td>MIT-Emanuel</td>
</tr>
<tr>
<td>F</td>
<td>MIT-Emanuel</td>
<td>Grell FC</td>
</tr>
<tr>
<td>G</td>
<td>MIT-Emanuel</td>
<td>MIT-Emanuel</td>
</tr>
</tbody>
</table>

The higher resolution simulations show improved precipitation considerably for all the cumulus schemes considered (Fig. 2.3). The Grell AS scheme has a tendency to underestimate the precipitation over land and ocean (Fig. 2.3(a)-(b)), but the precipitation over oceans has greatly improved compared to the coarser resolution simulations. The Grell AS scheme removes the excess buoyant energy generated by large scale motions in a single time step and hence minimizes the convective precipitation (Giorgi and Shields, 1999). The Grell FC scheme produces reasonable precipitation totals, (Fig. 2.3(d)-(e)) however it also underestimated the total precipitation. The dry bias in precipitation in Grell FC is lower compared to Grell AS as the two closures dissipates the buoyant energy differently, while the Grell AS dissipates the energy at each time step, Grell FC dissipates the buoyant energy at time scales that are approximately half an hour. MIT-Emanuel scheme has a wet bias over the land however it has a dry bias over the oceans (Fig. 2.3(f)), this result is contrary to what the norm is over the ocean when the scheme is utilized.
Further customization was done with mixed convective schemes to access if the biases in precipitation can be reduced. With a mixed convective scheme, using MIT-Emanuel over oceans still under predicts the total amount of precipitation (Fig. 2.3(b),(f)). These results show us that MIT-Emanuel scheme does not perform well in the tropical oceans. The Grell options over ocean, although they also show weak precipitation, have improved total precipitation in comparison to MIT-Emanuel. When the MIT-Emanuel scheme is used over the land and a Grell scheme over oceans, there is a general overestimation of precipitation over land, consistent with the previous results when MIT-Emanuel is used as a standalone convective scheme. The MIT-Emanuel/Grell AS scheme produces the highest overestimation of precipitation over land, while MIT-Emanuel/Grell FC had lower wet bias over land. Both simulations show improved total precipitation over the oceans compared to those that utilize the MIT-Emanuel scheme over oceans.

These simulations show that using a higher resolution model greatly improves the precipitation over the region. In addition we have found the MIT-Emanuel scheme has a dry bias over ocean. Total bias could be reduced by using a mixed convective scheme. Statistical analysis of the performance of each scheme was done using the methods outlined in the previous section. To analyze how the schemes perform in specific regions over the domain, we subdivided the domain into smaller regions, the Congo Basin, the East Africa Region and the Oceans. The subdivision was done on the basis of the climate regimes and the land cover. The Congo region is a rainforest, which receives much higher precipitation compared to the East Africa Domain. The East Africa domain is made up of heterogeneous land cover.

In the Congo rainforest the minimum RMSE occur when using the Grell FC scheme, while the highest errors occur when the MIT-Emanuel scheme is used over land (Fig. 2.4). The red line in Fig. 2.4 indicates the standard deviation of the TRMM datasets. As indicated before a model has skill when the RMSE is less than the standard deviations of the observed data and the standard deviation of the model should almost be the same with the standard deviation of the observed data. Grell AS does not accurately capture the spatial distribution of precipitation in the region with the standard deviation being lower than from TRMM. From this subregion we can conclude that without tuning any
Figure 2.3 Precipitation distribution using different cumulus parameterization for a 25km resolution simulation. (a) Grell AS (b) Grell AS/MIT- Emanuel (c) MIT-Emanuel/Grell AS (d) Grell FC (e) Grell FC/MIT-Emanuel (f) MIT-Emanuel/Grell FC (g) MIT-Emanuel and (h) the observations TRMM.

of the microphysics and cumulus parameters Grell FC performs the best in terms of the statistical measures, however it does underestimate the total precipitation. These results
are consistent with Davis et al. (2009) findings that Grell cumulus schemes underestimate while MIT-Emanuel cumulus scheme overestimates the total precipitation. The mixed cumulus schemes, show a slightly reduced RMSE.

![Congo Region](image)

**Figure 2.4** RMSE (blue bars) and Standard deviation (purple) over Congo Region using the different cumulus schemes. The red line is the standard deviation of the TRMM observational dataset.

For the East Africa region, the Grell FC has the highest skill. The MIT-Emanuel scheme has a high variance, while the Grell AS has a low variance (Fig. 2.5). In comparison with the Congo region the differences in the modeled and observational data is less than that of their respective standard deviations. That is if we compare the standard deviation of the MIT-Emanuel scheme and the RMSE over the Congo region and over East Africa, the RMSE are relatively lower than the standard deviation over the East Africa region than the Congo region. This implies that the model is performing better over this region compared to the rainforest. We postulate that the data quality over the Congo region could be a major drawback in the statistics in this region. The region has sparse network of insitu datasets and thus it is difficult to calibrate the satellite observational datasets.
The model underperforms the most and shows little to no skill over the ocean in most simulations. Although the Grell schemes have been reported to have weaker precipitation over the ocean, they perform better compared to the MIT-Emanuel scheme (Fig. 2.6). The lack of model skill over oceans could be due to the input datasets, and also lack of an ocean model, hence the sea surface temperatures are not correctly captured in the model.

The correlations in all the simulations are relatively high (Fig. 2.7) except over the oceans where the correlations are as low as 0.49 in the mixed convective scheme simulations. In most of the simulations the East Africa region had higher correlations and thus performed better compared to the Congo region and the ocean. Zou and Zhou (2013) coupled an ocean model to RegCM3 and showed that without coupling the ocean model the atmosphere is driven by the prescribed SST forcing and hence the model does not perform well over oceans.

The normalized mean errors range from 33% to 52% with the MIT-Emanuel scheme having the highest normalized mean error. The normalized mean biases show that MIT-Emanuel has a wet bias over land and dry bias over the ocean. The Grell AS and Grell
FC scheme underestimates the precipitation over the domain with the Grell FC having the lowest under prediction. Hence from the initial customization of the higher resolution simulations we can conclude that the Grell FC cumulus scheme performs better compared to all the other schemes.

2.3.1.2 Microphysics Schemes:

These simulations use the same physics combinations as outlined in the previous section except for the microphysics scheme, in this case the Nogherroto/Tompkins scheme was utilized. The Grell FC (Fig. 2.8(a)) is able to capture the spatial pattern of precipitation over the domain, however the physics combination has a dry bias. The MIT-Emanuel scheme (Fig. 2.8(b)) is able to capture the spatial pattern of precipitation over the lake and the Congo region, the physics combination however has weak precipitation intensity over the Ocean. A comparison of the SUBEX and the Nogherroto/Tompkins scheme shows that the Nogherroto/Tompkins scheme has lower wet biases. Grell-AS became unstable when used in combination with the Nogherroto/Tompkins scheme after running for 2 months.
Figure 2.7 Spatial correlations of precipitation for the different regions using different cumulus schemes with the TRMM datasets.

Table 2.4 is a summary of the statistics carried out to evaluate the simulations. The Grell FC scheme has a dry bias in all the regions with precipitation biases ranging from 31mm to 39mm. Using the Pielke Sr (2013) skill criterion shows that the physics combination has skill in all the regions. The MIT-Emanuel scheme does not show skill in all the regions. In all the regions the standard deviation of the model is higher than that of the observations. Comparing with the SUBEX set of experiments the Nogherroto/Tompkins scheme in combination with the 2 cumulus schemes we tested produces less intense precipitation. In a study done by Nogherotto (2015) they showed that the SUBEX scheme has higher cloud fraction than in the Nogherroto/Tompkins scheme. The lower cloud fraction in the Nogherroto/Tompkins scheme leads to lower total precipitation when used in combination with any of the cumulus parameterization. The lower cloud fraction in the model leads to more reasonable precipitation when the MIT-Emanuel cumulus is utilized.
2.3.1.3 Land Surface Schemes:

The Community Land Model (CLM) produces less precipitation (Fig. 2.9) compared to the BATS1e model. The negative bias increases considerably when the scheme is used with the Nogherroto/Tompkins scheme. For Grell FC the use of CLM scheme and Nogherroto/ Tompkins scheme reduces precipitation over land. The spatial distribution of precipitation using the MIT-Emanuel scheme is less intense compared to the scheme that used BATS1e as the land surface scheme. Just like in the Grell FC scheme, when the CLM and Nogherroto/Tompkins scheme are used the total precipitation is reduced considerably. Spatially using the MIT-Emanuel scheme over land and the Grell FC scheme over oceans produces a pattern that closely resembles the precipitation distribution in observations.
Table 2.4 Summary of evaluation statistics for RegCM simulations done to test the performance of the different large scale precipitation schemes. RMSE is the root mean square error, NME is the normalized mean error, NMB is the normalized mean bias, MB- mean bias, Cor is the pattern correlation, NRMSE is the normalized root mean square SD Mod.- standard deviation calculated from model output and SD Obs. - standard deviation for observations.

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Region</th>
<th>RMSE</th>
<th>NME</th>
<th>NMB</th>
<th>MB</th>
<th>Cor</th>
<th>SD Mod.</th>
<th>SD Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grell FC</td>
<td>Congo</td>
<td>177</td>
<td>0.34</td>
<td>-0.08</td>
<td>-31</td>
<td>0.72</td>
<td>217</td>
<td>244</td>
</tr>
<tr>
<td></td>
<td>EA</td>
<td>147</td>
<td>0.34</td>
<td>-0.01</td>
<td>-39</td>
<td>0.78</td>
<td>192</td>
<td>221</td>
</tr>
<tr>
<td></td>
<td>Ocean</td>
<td>148</td>
<td>0.42</td>
<td>-0.14</td>
<td>-37</td>
<td>0.64</td>
<td>165</td>
<td>175</td>
</tr>
<tr>
<td>MIT</td>
<td>Congo</td>
<td>253</td>
<td>0.36</td>
<td>0.22</td>
<td>88</td>
<td>0.63</td>
<td>296</td>
<td>244</td>
</tr>
<tr>
<td></td>
<td>EA</td>
<td>204</td>
<td>0.42</td>
<td>0.18</td>
<td>59</td>
<td>0.75</td>
<td>295</td>
<td>221</td>
</tr>
<tr>
<td></td>
<td>Ocean</td>
<td>212</td>
<td>0.55</td>
<td>-0.12</td>
<td>-33</td>
<td>0.49</td>
<td>232</td>
<td>175</td>
</tr>
</tbody>
</table>

The CLM schemes performs better when used with the MIT-Emanuel scheme compared to the Grell scheme (Table 2.5) which tends to have a very dry bias in all the regions and all the physics combinations considered in this study. The highest dry bias occurs when the Nogherotto/Tompkins scheme, Grell FC and the CLM scheme are used, the precipitation biases are as high as 267mm. The normalized mean bias is high for the Grell scheme, over 49% over land and 38% over the ocean, while that for MIT-Emanuel has a maximum of 30% over land and about 37% over oceans. The correlation over East Africa for Grell FC scheme is even negative (−0.15) implying that the model has high/low precipitation in regions where the observed precipitation is low/high. The correlations in all the other simulations are reasonably high although they are lower when the BATS1e scheme. All the simulations have lower standard deviation and higher root mean square errors, thus they do not show skill. However the normalized root mean square are close to 1 and the correlations are relatively high.
Figure 2.9 Precipitation distribution using the CLM land surface module.
Table 2.5 Summary of evaluation statistics for RegCM simulations done to test the performance of the Community Land Model. RMSE is the root mean square error, NME is the normalized mean error, NMB is the normalized mean bias, MB- mean bias, Cor is the pattern correlation, SD Mod.- standard deviation calculated from model output and SD Obs.- standard deviation for observations. The microphysics N/T is the Nogherroto/Tompkins scheme.

<table>
<thead>
<tr>
<th>Cumulus</th>
<th>Microphysics</th>
<th>Region</th>
<th>RMSE</th>
<th>NME</th>
<th>NMB</th>
<th>MB</th>
<th>Cor</th>
<th>SD Mod.</th>
<th>SD Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grell FC</td>
<td>SUBEX</td>
<td>Congo</td>
<td>261</td>
<td>0.86</td>
<td>-0.49</td>
<td>-165</td>
<td>0.57</td>
<td>162</td>
<td>244</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EA</td>
<td>255</td>
<td>0.62</td>
<td>-0.49</td>
<td>-160</td>
<td>0.43</td>
<td>102</td>
<td>221</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ocean</td>
<td>203</td>
<td>0.58</td>
<td>-0.38</td>
<td>-105</td>
<td>0.36</td>
<td>127</td>
<td>175</td>
</tr>
<tr>
<td>Grell FC</td>
<td>N/T</td>
<td>Congo</td>
<td>358</td>
<td>2.14</td>
<td>-0.66</td>
<td>-267</td>
<td>0.27</td>
<td>104</td>
<td>244</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EA</td>
<td>323</td>
<td>0.77</td>
<td>-0.65</td>
<td>-216</td>
<td>-0.15</td>
<td>65</td>
<td>221</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ocean</td>
<td>220</td>
<td>0.61</td>
<td>-0.43</td>
<td>-119</td>
<td>0.26</td>
<td>124</td>
<td>175</td>
</tr>
<tr>
<td>MIT</td>
<td>SUBEX</td>
<td>Congo</td>
<td>240</td>
<td>0.51</td>
<td>-0.2</td>
<td>-81</td>
<td>0.49</td>
<td>199</td>
<td>244</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EA</td>
<td>189</td>
<td>0.39</td>
<td>-0.2</td>
<td>-65</td>
<td>0.49</td>
<td>153</td>
<td>221</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ocean</td>
<td>205</td>
<td>0.54</td>
<td>-0.34</td>
<td>-93</td>
<td>0.42</td>
<td>165</td>
<td>175</td>
</tr>
<tr>
<td>MIT</td>
<td>N/T</td>
<td>Congo</td>
<td>236</td>
<td>0.61</td>
<td>-0.29</td>
<td>-116</td>
<td>0.35</td>
<td>137</td>
<td>244</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EA</td>
<td>198</td>
<td>0.45</td>
<td>-0.29</td>
<td>-95</td>
<td>0.62</td>
<td>127</td>
<td>221</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ocean</td>
<td>188</td>
<td>0.53</td>
<td>-0.32</td>
<td>-87</td>
<td>0.44</td>
<td>135</td>
<td>175</td>
</tr>
<tr>
<td>MIT/Grell FC</td>
<td>N/T</td>
<td>Congo</td>
<td>227</td>
<td>0.55</td>
<td>-0.22</td>
<td>-90</td>
<td>0.56</td>
<td>188</td>
<td>244</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EA</td>
<td>194</td>
<td>0.45</td>
<td>-0.21</td>
<td>-68</td>
<td>0.58</td>
<td>154</td>
<td>221</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ocean</td>
<td>178</td>
<td>0.48</td>
<td>-0.17</td>
<td>-48</td>
<td>0.43</td>
<td>144</td>
<td>175</td>
</tr>
</tbody>
</table>

In general the CLM model has lower total precipitation compared to simulations that utilize the Grell scheme. The lower precipitation amounts in CLM compared to BATS1e are consistent with previous studies that compare the two land surface modules (Steiner et al., 2005; Wang et al., 2014b). In a study done by Steiner et al. (2005) over East Asia they showed that the surface water balance in the two models differ considerably. While CLM produces less rainfall it had a higher runoff rate, which was approximately 30% of the incident rain in some regions and most of the months, BATS had around 20 – 25% runoff despite its higher precipitation. Consequently this leads to more transpiration rates in CLM and less ground water evaporation. Thus the evapotranspiration rates in CLM are lower than in the BATS1e case. Figure 2.10 shows the average evapotranspiration from the different land surface schemes. The CLM scheme shows lower evapotranspiration fluxes compared to the BATS1e. This is more obvious in the Congo rainforest, where in BATS1e the fluxes are much higher than CLM. Over the lake BATS1e has higher...
fluxes than the CLM case. Although the circulation patterns are the same in the two simulations, the wind speed is a lot stronger in the BATS1e case compared to the CLM case, this reduces moisture influx into the Congo region from the Atlantic ocean, and also into the East Africa Coastal region form the Indian Ocean.

Figure 2.10 Average evapotranspiration and 10m winds for the different Land Surface Modules (a) BATS (b) CLM.
2.3.1.4 Planetary Boundary Layer Scheme

The default Planetary Boundary Layer (PBL) scheme is the Holtslag scheme. The scheme is known to produce relatively strong, and often excessive, turbulent vertical transfer. Although this problem has been reduced in the latest version, the problem was not totally removed (Elguindi et al., 2013). The second PBL scheme is the University of Washington (UW) PBL scheme. The major difference between the Holtslag PBL scheme and the UW scheme is that the UW model accounts for the production of turbulence by cloud-top radiative cooling. The simulations are done using the SUBEX microphysics scheme and the BATS1e scheme. When the UW PBL scheme is utilized total precipitation over the lake closely resembles that of observations. The western side is receiving more rain compared to the eastern side (Fig. 2.11). In the Grell FC case the model underestimates the total precipitation and the model wet bias is reduced for the MIT-Emanuel scheme.

Table 2.6 shows a summary of the statistics performed for the performance of each PBL scheme. Both schemes show relatively high correlations. The Grell FC scheme underestimates precipitation when used in combination with the UW scheme. In all the 3 subregions the scheme has low standard deviation and high RMSE. Thus the physics combination does not show skill. The MIT-Emanuel scheme generally overestimates the precipitation over the region. The standard deviation of the physics combination is generally higher than that of the model. Comparing the performance with the Hostlag, this scheme has a reduced wet bias when the MIT-Emanuel scheme is used.
2.3.1.5 Radiation Scheme

The NCAR Community Climate Model (CCM3), radiation component, follows the δ-Eddington approximation of Kiehl et al. (1996) and includes 18 spectral vertical intervals from 0.2 to 5μm. The optical properties of the cloud droplets are expressed in terms of the cloud liquid water content and an effective droplet radius. The Rapid Radiation Transfer Model (RRTM) is a spectral-band scheme with 16 bands using the correlated-k
Table 2.6 Summary of evaluation statistics for RegCM simulations done to test the performance of the different Planetary Boundary Layer schemes. RMSE is the root mean square error, NME is the normalized mean error, NMB is the normalized mean bias, MB- mean bias, Cor is the pattern correlation, SD Mod.- standard deviation calculated from model output and SD Obs. - standard deviation for observations.

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Region</th>
<th>RMSE</th>
<th>NME</th>
<th>NMB</th>
<th>MB</th>
<th>Cor</th>
<th>SD Mod.</th>
<th>SD Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grell</td>
<td>Congo</td>
<td>294</td>
<td>1.25</td>
<td>-0.53</td>
<td>-215</td>
<td>0.59</td>
<td>108</td>
<td>244</td>
</tr>
<tr>
<td>FC</td>
<td>EA</td>
<td>249</td>
<td>0.61</td>
<td>-0.53</td>
<td>-176</td>
<td>0.65</td>
<td>90</td>
<td>221</td>
</tr>
<tr>
<td></td>
<td>Ocean</td>
<td>185</td>
<td>0.52</td>
<td>-0.43</td>
<td>-120</td>
<td>0.6</td>
<td>105</td>
<td>175</td>
</tr>
<tr>
<td>MIT</td>
<td>Congo</td>
<td>286</td>
<td>0.41</td>
<td>0.22</td>
<td>310</td>
<td>0.53</td>
<td>310</td>
<td>244</td>
</tr>
<tr>
<td></td>
<td>EA</td>
<td>231</td>
<td>0.45</td>
<td>0.15</td>
<td>50</td>
<td>0.7</td>
<td>316</td>
<td>221</td>
</tr>
<tr>
<td></td>
<td>Ocean</td>
<td>220</td>
<td>0.54</td>
<td>-0.08</td>
<td>-21</td>
<td>0.51</td>
<td>249</td>
<td>175</td>
</tr>
</tbody>
</table>

We test the performance of the RRTM scheme using the MIT-Emanuel cumulus scheme since the Grell FC scheme has a dry bias tendency. In all the simulations using this scheme the MIT-Emanuel scheme underestimates precipitation. The land surface scheme that was used is the Community Land Model. Over the oceans we employ the Grell FC scheme and the model shows great improvement in terms of total precipitation over ocean.

Comparing the results from this scheme and the previous results using the CCM3 scheme we find that the scheme generally leads to underestimation of total precipitation. Major problems occur on the boundaries were there is an accumulation of precipitation in the southern boundary (Fig. 2.12). The accumulation in the boundaries could be preventing moisture from being advected into the inner domain. Compared to the CCM3 radiation scheme, this scheme physics combinations failed to capture the spatial distribution of precipitation over land. Hence more work needs to be done in the parameterization of the scheme over the tropical regions.
Figure 2.12 Precipitation distribution using the RRTM radiation scheme for a 25km resolution simulation.

The statistical evaluation of the performance of the radiation scheme shows that the model does not have skill when the RMSE and standard deviation is compared. The RMSE is greater than the standard deviation from TRMM, the standard deviation in the model is much lower than in observations. When MIT-Emanuel scheme is utilized over both land and ocean the correlation is low compared to the mixed convective scheme. The radiation scheme performs equally in simulations utilizing either the SUBEX or
the Nogherroto/Tompkins scheme. Hence we conclude that the scheme reduces bias in radiation effects that lead to cloud formation.

**Table 2.7** Summary of evaluation statistics for RegCM simulations done to test the performance of the RRTM radiative scheme. RMSE is the root mean square error, NME is the normalized mean error, NMB is the normalized mean bias, MB- mean bias, Cor is the pattern correlation, SD Mod.- standard deviation calculated from model output and SD Obs.- standard deviation for observations. The microphysics N/T is the Nogherroto/Tompkins scheme.

<table>
<thead>
<tr>
<th>Cumulus</th>
<th>Microphysics</th>
<th>Region</th>
<th>RMSE</th>
<th>NME</th>
<th>NMB</th>
<th>MB</th>
<th>Cor</th>
<th>SD Mod.</th>
<th>SD Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIT/Grell</td>
<td>SUBEX</td>
<td>Congo</td>
<td>273</td>
<td>0.7</td>
<td>-0.31</td>
<td>-126</td>
<td>0.41</td>
<td>195</td>
<td>244</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EA</td>
<td>212</td>
<td>0.44</td>
<td>-0.26</td>
<td>-85</td>
<td>0.55</td>
<td>185</td>
<td>221</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ocean</td>
<td>204</td>
<td>0.52</td>
<td>-0.28</td>
<td>-77</td>
<td>0.4</td>
<td>172</td>
<td>175</td>
</tr>
<tr>
<td>MIT</td>
<td>N/T</td>
<td>Congo</td>
<td>315</td>
<td>1.34</td>
<td>-0.5</td>
<td>-218</td>
<td>0.39</td>
<td>135</td>
<td>244</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EA</td>
<td>255</td>
<td>0.58</td>
<td>-0.43</td>
<td>-141</td>
<td>0.35</td>
<td>127</td>
<td>221</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ocean</td>
<td>191</td>
<td>0.53</td>
<td>-0.38</td>
<td>-120</td>
<td>0.35</td>
<td>127</td>
<td>175</td>
</tr>
<tr>
<td>MIT/Grell</td>
<td>N/T</td>
<td>Congo</td>
<td>264</td>
<td>0.7</td>
<td>-0.31</td>
<td>-126</td>
<td>0.42</td>
<td>171</td>
<td>244</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EA</td>
<td>208</td>
<td>0.45</td>
<td>-0.29</td>
<td>-96</td>
<td>0.55</td>
<td>135</td>
<td>221</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ocean</td>
<td>187</td>
<td>0.48</td>
<td>-0.24</td>
<td>-64</td>
<td>0.4</td>
<td>154</td>
<td>175</td>
</tr>
</tbody>
</table>
2.3.2 Performance of the Optimal Conditions in Extreme years

Often in customizing the model, a single normal year is chosen and an assumption is made that the optimal physics combination will hold in all years both normal and extreme for that particular region. However the best optimal conditions might not hold for all the other years, for example Sun et al. (1999), found that the optimal physics combinations over East Africa were obtained when the Grell AS scheme was used, but using the same combination Anyah et al. (2006) found that the combination greatly underestimated precipitation and opted to use the Grell FC. Hence in order to gain confidence that the optimal combination will produce reasonable precipitation in all the years we choose 2 years that have influence from large scale features. In this case ENSO and IOD are considered as they had the highest influence from previous studies. The years chosen in this case are the same as the ones used in Argent et al. (2014). That is we chose 2006 and 2010; 2006 had positive IOD and ENSO, with IOD being stronger than ENSO. 2010 had strong negative indices for both ENSO and IOD. In this case we expect 2006 to have more precipitation compared to 2010 and 2000, with 2010 being the driest year in the analysis.

From the customization experiments in the previous sections we concluded that the Grell FC cumulus scheme produced the most realistic results both spatially (Fig. 2.13) and statistically. Figure 2.14 shows that the chosen optimal conditions for the year 2000 show skill. Thus we hypothesize that the best optimal physics for a normal year will also be able to capture the spatio-temporal distribution of rainfall in an extreme year.

The model is able to capture increases in precipitation in a wet year. The RegCM simulation is able to capture the increases in precipitation over the Indian ocean (Fig. 2.15). Although in some cases it does underestimate the total precipitation. Over Congo the model is able to capture the high precipitation however it still does underestimate the total precipitation. Over East Africa the model underestimates precipitation in most cases, with the Lake, Kenya and Sudan having the highest underestimation. In these regions TRMM shows an average of approximately 300\text{mm} while RegCM simulates an average less than 100\text{mm}. The model does well in capturing the spatial pattern of precipitation over the lake but it however underestimates the total precipitation. Statistical evaluation
shows that the model performs badly over East Africa and has the highest underestimation in the region (Fig. 2.16). The correlation with the precipitation is 0.5 which is far much lower compared to the normal years. Thus in wet years the model generally underestimates the total precipitation.

In a dry year the model is able to capture the spatial pattern over the domain, although it generally underestimates the precipitation (Fig. 2.17). Over the Ocean the
model was able to capture the reduction in precipitation. The model was also able to capture the reduced precipitation over the lake. It however failed to capture the rainfall maxima over the Northern part of Lake Tanganyika. Statistical evaluation of the performance of the model shows that in dry years the model still exhibits skill as in a normal year (Fig. 2.18, Fig. 2.16). In all cases the RMSE is less that the standard deviation from the TRMM datasets. The standard deviations for the simulation also seem reasonable close to the standard deviation of TRMM. In the normal case East Africa had the best correlation and skill however for this case the correlation is high (0.77) but model underestimates more compared to the normal case.

We conclude that the model is able to capture the increase in an extreme wet year or a decrease in an extreme dry year. But the model is unable to capture the rainfall intensity. Often having weaker intensity in most regions for an extreme wet year.
Figure 2.15 Precipitation distribution for 2006 using the Grell FC Scheme.
Figure 2.16 Clustered bar plot for the different statistics calculated to evaluate the performance of the model for the year 2006.
Figure 2.17  Precipitation distribution for 2010 using the Grell FC Scheme.
Figure 2.18 Clustered bar plot for the different statistics calculated to evaluate the performance of the model for the year 2010.
2.3.3 Evaluation of the Performance in Multiple Years

We perform a 16 year integrated simulation for the period 1998 to 2013. The simulations are run from December of the previous year to December of the year we are interested in. Evaluation for these simulations closely follows the evaluations conducted for the shorter simulations. We evaluate precipitation for each year and for both the Short and Long Rains. In addition to the precipitation evaluations we evaluate the temperature using the Climate research Unit datasets.

The Taylor diagram for the Short Rains season (Fig. 2.19) shows that the model is able to capture the inter-annual variability of precipitation in the region. The bias in precipitation during the extreme years are not dramatically different from the normal years except for the year 2011, which on record was a very dry Long Rains season. The model standard deviations are different depending on the physics combination that is being utilized, the MIT-Emanuel scheme in general has higher normalized standardized deviations, while Grell FC has lower standardized deviations compared to the observational data. The correlations are at least 0.5 for most of the years. Hence we conclude that the model performance is good over East Africa in all the years.

The correlation for modeled temperature and CRU is very high ($r > 0.85$) in most cases, indicating that the model is able to reproduce the spatial distribution of temperature over the region (Fig. 2.20). During the short rains there are no major differences in the statistical performance of the different physics combinations. The normalized root mean square errors are very low indicating that the model has minimum temperature bias. The difference plots are shown in Appendix A. The difference plots for temperature show that the MIT-Emanuel scheme performs better as the errors are much lower compared to the Grell FC scheme.

We also investigate the performance of the model using the probability plots. The probability distribution plots (Fig. 2.21) show that during the Short Rains season the model is able to capture the temperature distribution. However the model are unable to capture correctly the frequency of the temperatures. More often the models capture a temperature of $24^\circ C$ while in general this high frequency occurs at a lower temperature.
In addition the frequency of lower temperatures is slightly higher in the simulation using the Grell FC cumulus physics. The MIT-Emanuel scheme has the same deficiencies like the Grell FC scheme. However the errors are much lower than in the Grell FC scheme. The extreme temperatures are captured more accurately in the MIT-Emanuel scheme.

During the Long Rains season the model simulations generally perform well in the different years for precipitation. The correlations are relatively high in all the years. The Taylor diagram (Fig. 2.22) for this season has more outliers compared to the OND season. The Long Rains season is more variable compared to the short rains season and the oceanic and atmospheric drivers for this season are no yet fully understood, this might mean that the model treatment for this season is inadequate and more work needs to be done. In addition all the customization was done for the Short Rains and the optimal physics combination for that season might not apply to the Long Rains season.

Figure 2.23 is the Taylor diagram for the temperature during the Long Rains season. The MIT-Emanuel scheme still outperforms the Grell scheme during the Long Rains
season just like during the Short Rains season. The MIT-Emanuel scheme has lower biases as shown in the difference plots in Appendix A. During the MAM season there are still higher biases along the Indian Ocean coastline. The errors could be due to lack of an ocean model, in our simulations, hence the air-sea interactions are not adequately treated. This leads to enhanced precipitation along the coastline, and hence a cold bias in temperature.

Although the model is able to capture the shape for the probability frequency of temperatures (Fig. 2.24), in both cases it overestimates the temperature were maximum frequency occurs. In addition the model is better suited to capture the lower temperatures frequency compared to the higher temperature frequency. Notably the frequency of the high temperatures is much lower in the Grell FC scheme. The MIT-Emanuel scheme also has a lower frequency for the high temperature ranges but it tails off at a higher temperature than Grell FC. Thus although the model is able to capture the inter-annual variability of the temperature in the region, there is need of more work that will focus on improving the temperature and precipitation biases.
Figure 2.21 Probability distribution function plot for temperature during the Short Rains over East Africa.

Figure 2.22 Taylor diagram for precipitation during the Long Rains over East Africa.
Figure 2.23 Taylor diagram for temperature during the Long Rains over East Africa.
Figure 2.24 Probability distribution frequency plot for temperature during the Long Rains over East Africa.
2.4 Conclusions

The model shows that it has skill and is able to reproduce the precipitation and temperature patterns over East Africa. In this Chapter we have examined the sensitivity of tuning parameters, in the microphysics scheme, the impact of using different physics combinations and the impact of resolution. Tuning of the minimum RH for SUBEX scheme alters the total rainfall over the region, with lower RH values having lower cloud fraction thus lower total precipitation. Choosing a very low RH value leads to dry bias. Tuning of the RH to lower values than the default in the source code is only advisable when using the MIT-Emanuel scheme. The Grell FC scheme has a dry bias with the default parameterization and hence increasing the minimum RH value will help in reducing the dry bias in the model. Thus for the coarse simulations we can conclude that RegCM simulated precipitation over East Africa can be improved by tuning the parameters in the microphysics schemes just as in the Davis et al. (2009) study. Further tuning of parameters in the coarse resolution did not improve the precipitation distribution and amount, thus higher resolution simulations were conducted.

Spatial comparison of the high and low resolutions simulation revealed the usefulness of using higher resolution, the total bias in the MIT-Emanuel scheme is lower. The higher resolution also performs better statistically compared to the lower resolution. Performance of the Grell AS, Grell FC and mixed cumulus schemes were also analyzed. All the cumulus schemes are able to reproduce the spatial patterns of precipitation over the region, but fails to capture the correct intensities. The Grell scheme using either closure have a dry bias while the MIT-Emanuel scheme has a wet bias. For the cumulus schemes we conclude that the Grell FC scheme had the least precipitation bias compared to other schemes when we use the BATS1e, SUBEX and Holtslag schemes.

When the MIT-Emanuel cumulus scheme is used with the Nogherroto/Tompkins microphysics, the total wet bias is reduced. Using this scheme with the Grell scheme further increases the dry bias. Thus for this region it is advisable that the Grell scheme is used in conjunction with the SUBEX scheme. We also tested the performance of the different land surface schemes, the BATS1e and the CLM. CLM is superior compared to BATS1e in terms of treatment of the heterogeneity and the energy balances. The CLM scheme
has lower evapotranspiration rates compared to the BATS1e scheme and hence it has a
tendency to have a dry bias. Although the use of both the SUBEX and the MIT-Emmanuel
scheme has been well known for producing the highest wet biases in the model, using
this physics combination with the CLM scheme reduces the total precipitation to a point
that the model underestimates the total precipitation. Hence when using this scheme one
will need to adjust the default RH value to a higher value in the SUBEX scheme. When
the UW PBL scheme and RRTM radiation schemes are used the Grell FC scheme does
not perform well, often with very high dry biases. The MIT-Emmanuel scheme however
performs a lot better and the wet biases are reduced.

We choose the MIT-Emmanuel scheme and the Grell FC scheme to analyze the ability
of the model to reproduce the inter-annual variability of precipitation over the region. We
evaluated both the temperature and the precipitation for both the Long and the Short
Rains. The models are able to reproduce the temperature over the region and have very
low temperature biases. In terms of precipitation the model was able to reproduce the
precipitation with minimum errors although there were years in which the model had
low spatial correlation, and overestimated the total precipitation.

This study has shown that the model is able to accurately reproduce the precipitation
patterns. Future simulations in the region could utilize the more advanced schemes,
especially the Community Land Model scheme which has a number of options that can
be turned on for use in a number of application sectors such as agriculture. In addition
the Nogherroto/Tompkins scheme can be instead of the SUBEX scheme, to reduce the
wet bias in the MIT-Emmanuel scheme. For longer simulations we chose to use the Grell
scheme with the default physics combination as in the source code, and also the optimum
conditions as identified by Davis et al. (2009). The simulations had very high correlations
for temperature and relatively high correlations for precipitation.
Chapter 3

Analysis of East Africa Climatology using Satellite data and the ICTP Regional Climate Model

3.1 Introduction

East Africa climate exhibits a bimodal rainfall pattern mainly controlled by the North-South migration of the Inter-tropical convergence zone (ITCZ). The ITCZ separates the northeast and southeast monsoons, and crosses East Africa twice every year, once during the Long Rains (March-April-May) and again during the Short Rains (October-November-December). The Long Rains tend to have higher total precipitation which is less variable on the inter-annual scale while the Short Rains are more variable and generally lower than the Long Rains (Mutai et al., 1998; Nicholson, 1996). The inter-annual variability of precipitation is also influenced by the presence of the East Africa Great Lakes, the complex terrain, sharp contracts in vegetation cover over the region and maritime effects mostly from the equatorial Indian Ocean and the Pacific Ocean.

The lake/land breezes mainly dominate the diurnal cycle over most lake basins. This is more evident over Lake Victoria because of its size (69,484 square km) and shape (Anyah and Semazzi, 2004). The reversal of the land-lake breeze is caused by the thermal gradient between the lake and land surface. During the day the land surface warms faster than the lake surface, this causes a drop in surface pressure over land and hence air flows from the lake towards the land. The mixing of the two different air masses forms a front where convergence will occur. Anyah et al. (2006) found that the upslope/downslope flow generated by the mountains enhanced the land/lake breeze. At night the pattern is
reversed, the land cools faster than the lake and the flow is from the land to the lake, leading to convergence over the lake. Just like during the day the nocturnal process is enhanced by the presence of the mountains through a katabatic wind. Convection over the western and northwestern parts of the lake is boosted by the humid Congo air mass (Anyamba, 1984), the background flow and also the bathymetry of the lake. Lake Victoria is shallower on the western side of the lake and hence the lake surface temperatures are a lot warmer compared to the eastern side.

Early studies on the inter-annual variability of precipitation over East Africa identified El Niño/Southern Oscillation (ENSO) as the leading source of variability (Mutai et al., 1998; Nicholson and Kim, 1997; Ogallo, 1988). Ogallo (1988) used rotated empirical orthogonal function analysis, and divided the region into smaller climatologically homogeneous regions. They found that increased precipitation was significantly associated with the El Niño events. Mutai et al. (1998) showed that ENSO was the dominant source of East Africa inter-annual variability, their studies used the rotated empirical orthogonal functions on sea surface temperatures from the tropical Pacific, Indian Ocean and Atlantic. During the 1997/1998 season the El Niño event was found to increase the total precipitation over Lake Victoria catchment by around 15 to 20% Birkett et al. (1999).

Later studies have shown that in addition to the teleconnection of East Africa precipitation with the Pacific Ocean, the sea surface temperature gradient changes between the east and the western parts of the Indian Ocean, also plays a major role in modulating the precipitation over East Africa. Saji et al. (1999) found evidence that the Indian Ocean SST anomalies have a significant impact on precipitation over East Africa; when the sea surface temperatures in the western Indian Ocean are warm and sea surface temperatures off the Sumatra are low they force anomalous southeasterly trade winds that increases the moisture influx into the region and hence enhances total precipitation. Birkett et al. (1999) also showed that while ENSO increased precipitation by 5 to 20%, warming of the Indian Ocean increased the precipitation by 20 to 160% which is a lot much higher than the impact of ENSO. Bowden and Semazzi (2007) used partial correlations on pentad CMAP rainfall data from 1979 to 2001 to investigate the inter-annual climate drivers of the Short Rains. Their studies revealed that both ENSO and IOD are important for the
inter-annual variability of the Short Rains. More studies indicating the importance of the Indian Ocean in modulating East Africa climatology have been carried out recently. However there is controversy on whether the Indian Ocean dipole and the ENSO are related, with different researchers proposing different explanations. Black et al. (2003) postulated that the warm phase of the El Niño Southern Oscillation can lead to an Indian Ocean Dipole mode, that will lead to extreme precipitation over the region. Pfeiffer and Dullo (2006) found high correlation of ENSO and IOD when they correlated the El Niño index with the Indian Ocean sea surface temperatures over a period of 150 years. On the other hand a number of studies argue that the Indian Ocean Dipole is independent of ENSO and is more influential in determining the inter annual variability of East Africa rainfall compared to ENSO (Behera et al., 2005; Manatsa et al., 2012; Saji et al., 1999; Saji and Yamagata, 2003).

Although the oceanic and atmospheric drivers for precipitation over East Africa have been established for the Short Rains season, the Long Rains season inter-annual drivers have been elusive (Camberlin and Philippon, 2002). A limited number of studies on the inter-annual variability of East Africa Long Rains have been done compared to the Short Rains. Past studies (Indeje et al., 2000; Ntale and Gan, 2004) have found that ENSO increases the precipitation over East Africa and this is likely to lead to wetter conditions during the Long Rains. Indeje et al. (2000) further showed that the increase in post El Niño events are not homogeneous throughout the region as some regions experience a decline in precipitation and others an increase in precipitation. However these links were not very strong often with low correlations. In addition Ogallo (1988) found that some of the extreme dry and wet years were not well correlated with ENSO implying that there could be other teleconnections that affect East Africa precipitation.

Camberlin and Philippon (2002) suggested that the season should not be analyzed in its entirety but rather group the onset and peak (March to April) and then analyze the cessation (May) separately, in this way the influence of ENSO on the March-April months was found to be significant at the 95% statistical level. In addition to the inter-annual variability a number of studies during this season have focused on the recent decadal decline of the Long Rains. Anthropogenic warming and natural variability of the SST have been identified as the major causes of the decadal decline in the Long Rains. Williams
and Funk (2011) proposed that anthropogenic forced rapid warming of the Indian ocean SST extends the warm pool and Walker circulation resulting is subsidence and hence drying over the region. In addition to the role of the Indian Ocean Lyon and DeWitt (2012) proposed that a shift to warmer SST over the western tropical Pacific and cooler SST over the central and eastern Pacific contribute to the decline of the Long Rains. Paleoclimate studies have shown that moist conditions in coastal East Africa are associated with cool SSTs in the eastern Indian Ocean and ascending circulation over East Africa, and this contributes more to the inter-annual variability of precipitation over the region (Tierney et al., 2013).

Whilst the region has experienced a decline in the long rains over the past decades, projections from the Coupled Model Inter-Comparison project Phase 5 (CMIP 5) models suggest that in the future, precipitation will increase over the East Africa region, a phenomena now known as the “East Africa Climate paradox”. Despite the consensus in the CMIP projections of an increase in precipitation of rainfall over the region in the future, Cook and Vizy (2013) projected a decrease in the precipitation using regional climate models. This discrepancy necessitates the investigation of the ability of a regional climate model in producing the spatial and temporal variability of precipitation.

This study utilizes the International Centre for Theoretical Physics (ICTP) Regional Climate Model version 4 (RegCM4) and satellite based datasets to evaluate the ability of the model to reproduce the observed inter-annual variability and investigate the drivers of climate in both the Long and the Short rains. It is instructive that if the regional climate models are able to capture the most dominant modes of variability that are produced by the satellite data then a historical simulation can be done in order to understand the physical mechanisms affecting the precipitation in the region and hence the decline in the Long Rains.

3.2 Data

The Tropical Rainfall Measuring Mission (TRMM) datasets are used for Empirical orthogonal Function analysis. The monthly TRMM datasets are produced by combining the estimates generated by the TRMM and other satellite products (3B42) and the CAMS
global gridded rain gauge data, produced by NOAA’s Climate Prediction Center and/or the global rain gauge product produced by the Global Precipitation Climatology Center (Huffman and Bolvin, 2011). To get the general variation of precipitation over the region monthly datasets (3B43) are used. The datasets spanned over 16 years that is from 1998 to 2013. While the TRMM dataset covers only a few years for EOF analysis its resolution is high compared to traditional gridded datasets that have been used in the region (Bowden and Semazzi, 2007; Mutai et al., 1998; Schreck and Semazzi, 2004; Smith, 2011). Nevertheless the dataset has biases over the region, especially because there is a sparse network of station data and a continuous decline in the number of station (Dinku et al., 2007). Adeyewa and Nakamura (2003) found that when the dataset is not bias corrected with observations it overestimates precipitation mostly during the Long Rains (MAM) season and the bias is reduced during part of the Short Rains (OND) season of our analysis. In contrast when they used the TRMM 3B43 monthly dataset the product closely matches rain gauge data. Thus we expect the bias to be high in regions where there are minimum number of station datasets. As mentioned before the region has complex terrain and therefore poses a challenge to satellite rainfall estimation algorithms, because of the warm orographic rainfall and ice over mountain tops (Dinku et al., 2007). Dinku et al. (2007) analyzed 10 satellite datasets over an Ethiopian highland with a relatively dense station network. Their findings showed that both the low temporal resolution (TRMM 3B43) and high temporal resolution (TRMM 3B42 3 hourly data) were in good agreement with the rain-gauge data, but over that region Tropical Applications of Meteorology using SATellite and other data (TAMSAT) and CPC Morphing Technique (CMORPH) performed better for high temporal resolution evaluation.

The National Centers for Environmental Prediction (NCEP) reanalysis wind, geopotential height, and sea surface temperature data was used to investigate rainfall air circulation relationships associated with the dominant TRMM EOF modes of regional climate variability. The reanalysis dataset is generated using a state of the science data assimilation scheme and numerical model (Kalnay et al., 1996). This Climate Data Assimilation System employs T62 horizontal resolution (about 210km, interpolated to a 2.58 latitude 32.58 longitude grid) and 28 vertical levels.

A number of indices are chosen for this study. Selection of the indices was done based
on past studies, the West Pacific index was chosen as a number of studies have shown that the gradient between the North Pacific and the West Pacific alters the Walker Circulation, causing subsidence over the region. The indices in the Atlantic ocean (TNA, NAO, TSA and AMSST) are motivated by the current work being done in the CLIMLAB on the role of the Atlantic ocean on the decadal variability of precipitation in the region. In addition to these indices we retain the traditional indices that are normally utilized in the region-Nino 3, IOD and the QBO. The following indices were used to determine the most dominant natural variability over Lake Victoria by correlating the indices with the time series of the EOF analysis.

- **Nino 3 index:** Several regions of the tropical Pacific Ocean have been chosen as being important for monitoring and identifying El Niño and La Niña. The Nino 3 index spans from 150°W to 90°W, and from 5°S to 5°N.

- **Dipole Mode Index:** The index is used to measure the strength of the gradient between the West and East Indian Ocean. It is obtained by calculating the difference between sea surface temperature (SST) anomalies in the western (50°E to 70°E and 10°S to 10°N) and eastern (90°E to 11°E and 10°S to 0°S) equatorial Indian Ocean.

- **The quasi-biennial oscillation (QBO):** The QBO is a quasi-periodic oscillation of the equatorial 30mb zonal wind between easterlies and westerlies in the tropical stratosphere. It propagates downwards at about 1km per month until they are dissipated at the tropical tropopause.

- **North Atlantic Oscillation (NAO):** The daily NAO index is constructed by projecting the daily (00Z) 500mb height anomalies over the Northern Hemisphere onto the loading pattern of the NAO.

- **Tropical North Atlantic Index:** The index is calculated using Sea Surface temperatures (55°W – 15°W, 5°N – 25°N) in eastern North Atlantic Ocean.

- **Tropical South Atlantic Index:** Calculated using the SST (30°W – 10°E, 20°S – 0°S)

- **West Pacific Index:** The West Pacific index is obtained by calculating the gradient between the North Pacific and the East Pacific. During winter and spring, the pattern consists of a north-south dipole of anomalies, with one center located over the
Kamchatka Peninsula and another broad center of opposite sign covering portions of southeastern Asia and the western subtropical North Pacific.

### 3.3 Methods of Analysis

Empirical Orthogonal Function (EOF) was first introduced in meteorology by Lorenz (1956) and is a fundamental tool that is now widely used for data reduction in meteorology and climate studies. A number of studies conducted over East Africa have used the tool to probe the physics underlying the variability of precipitation. In the EOF analysis we consider an $N$-dimensional vector time series $\mathbf{x}(t)$, sampled at $N$ discrete points in space. In this case we use the covariance matrix to conduct the EOF analysis. The covariance matrix of $\mathbf{x}$ is given by $C = \langle \mathbf{x} \mathbf{x}^T \rangle$ where the angle denote the probabilistic expectation.

In conducting an EOF analysis one of the most fundamental questions is “How many modes are not noise, that is how many modes should we retain?” Several statistical tests have been proposed for use in retaining the significant eigenmodes, these include, methods proposed by, North et al. (1982), Jolliffe (1972), Overland and Preisendorfer (1982), and Preisendorfer et al. (1981). The scree method is one of the subjective methods that has been proposed for use in retaining the most significant eigenmodes. In this case we use the method suggested by North et al. (1982).

Figure 3.1 shows the topography over our domain, we choose 2 subregions for further analysis that is the inner domain used in Schreck and Semazzi (2004) as our outer region for EOF analysis and the inner region is mostly focused on Lake Victoria Basin. The Lake Victoria region is chosen as it is the nerve centre of East Africa and supports approximately 40 million people through fisheries, agriculture, hydroelectric power, drinking water and transportation. In addition the climate in this region is largely influenced by large inland water bodies and the mountain ranges on either side of Lake Victoria.
3.4 Results and Discussion

3.4.1 Rainfall Patterns

Although TRMM datasets only span for a short period of time they have the advantage of having measurements over the lake that gridded insitu datasets do not have. Hence they offer a better analysis of the precipitations patterns over the lake. As mentioned in the data section East Africa region has limited data availability, and over the past years the number of stations used to produce gridded data has been declining. We construct 3-hourly plots for precipitation over East Africa in order to gain insight on the diurnal variation. Figure 3.2 is the average precipitation at each 3 hour interval. Lake Victoria receives the highest amount of precipitation in the region during the evening as we would
have expected. Maximum precipitation occurs in the South west of our domain. A num-
ber of theories have been postulated about the pattern over the lake, these include the
Lake Bathymetry, where the western side of the Lake is shallower than the Eastern side,
hence the lake is warmer in the Western side (Anyah and Semazzi, 2004; Anyah et al.,
2006). The background flow in the region is easterlies and this contributes to the asym-
metrical pattern. Over the land maximum precipitation occurs from 12UTC to 15UTC,
this is the time when the lake-breeze is most dominant. The rest of the region has less
precipitation compared to the lake and the coastal regions.

We further explore the diurnal timeseries at selected points over the lake. A diurnal
variation analysis of selected points over the lake shows when different maximum rainfall
occurs over the basin (Fig. 3.3). Precipitation at night seems to start from the North east
and this could be due to the steep topography to the east of Lake Victoria. The steep
topography generates very strong downslope (katabatic) winds at night since the air over
the mountain top is relatively colder than the air down the valley (Anyah et al., 2006).
This pattern leads to higher precipitation over the west coast compared to the east coast
of the lake, and is consistent with the Wet-Dry-Wet-Dry pattern in Asnani (1993). The
diurnal variation of precipitation over the region indicates that the rainfall maximum over
the lake do not occur at the same time. This is also supported by Burleyson (2013) plots
which showed the variation of percentage precipitation frequency in each 3 hour time step.
Figure 3.2 Average precipitation over every 3 hour period for TRMM 3B42 datasets for the 16 year period. The plots are generated by calculating the average precipitation at each 3 hour time interval during the short rain season.
Figure 3.3 Diurnal variation for selected sites over the Basin for the TRMM 3B42, during the Short Rains season.
We construct the annual cycle of precipitation by averaging the precipitation along between 28°E and 42°E, and then plot the average monthly precipitation over 16 years at different latitudes over the region (Fig. 3.4). The averages precipitation over the region is also obtained. In general the region exhibits a bimodal pattern, however rainfall pattern at each of the latitudes shows different intensities and to some extent different patterns. The Southern (4S – 8S) parts mostly have a unimodal rainfall pattern, and the rainy season is from December to April. Most of the Northern regions receive their peak rainfall in November and in April, with the least amount of precipitation from January to December. Although the other latitudes do not receive a fair amount of precipitation during the JJAS season, this region continues to receive precipitation. At the Equator the rainfall pattern is bimodal, and the region has distinctly high precipitation during November and March.

![Figure 3.4](image_url)

**Figure 3.4** Monthly timeseries of rainfall variability over East Africa using TRMM. The Different lines were constructed using different latitudes in the region. The average line was calculated by considering the 16 year average precipitation over the region.

Figure 3.5 is an extraction of the 4 line graphs from (Fig. 3.4). This plot clearly shows the antiphase rainfall over the northern parts and the southern parts of our domain. The
rainfall patterns are consistent with the North South Migration of the ITCZ during the boreal winter the ITCZ is located more in the South and during the boreal Spring it is moving out of the region toward the Northern parts of the region.

Figure 3.5 Monthly timeseries of rainfall variability over East Africa using TRMM data. The line plots are for the Extreme North and South parts of our East Arica Domain. The average line was calculated by considering the 16 year average precipitation over the region.
3.4.2 Empirical Orthogonal Function Analysis

To separate the most dominant modes of variability, empirical orthogonal function (EOF) decomposition is applied to the seasonal rainfall anomalies. The rainfall anomalies are constructed by calculating departures from the 1998 – 2013 mean in each grid square.

3.4.2.1 Short Rains season

In order to gain confidence in using TRMM datasets for analysis the EOF over the Greater Horn of Africa was calculated and compared with the EOF for Climate Research Unit (CRU) and CPC Merged Analysis of Precipitation (CMAP). The CRU and CMAP datasets have widely been used before over East Africa to study the rainfall variability (Bowden and Semazzi, 2007; Schreck and Semazzi, 2004; Smith, 2011). The results showed that the time series had the same amplitude and the same direction for most of the times.

Using the North et al. (1982) method, we retain only the first mode for the East Africa region. We employ 2 values as our cutoff for statistical significance, the 90% and the 95% significance levels. The cut off value for the 90% is 0.43, while that for the 95% is 0.50. The spatial distribution pattern for the large East Africa domain shows that EOF1 (Fig. 3.6) is dominated by one pattern and this has been attributed to Sea Surface Temperature anomaly forcing, of the El Niño/Southern Oscillation and Indian Ocean dipole (IOD) by past studies (Black et al., 2003; Bowden and Semazzi, 2007; Indeje et al., 2000; Manatsa et al., 2012; Ogallo, 1988; Saji et al., 1999; Schreck and Semazzi, 2004; Smith, 2011). The first EOF from TRMM explains approximately 58% of the total variance in the region of interest, while that using the Grell FC scheme and MIT-Emanuel scheme have a variance of 56% and 52.5% respectively.

The time series shows alternating dry and wet years, and a slight positive trend (Fig. 3.7). The correlation between the TRMM EOF 1 time series and those for the simulations using the Grell-FC scheme and the MIT-Emanuel scheme is 0.92 and 0.90 respectively (Table 3.1). This shows that the model is able to capture the inter-annual variability of precipitation over East Africa. The highest disagreement between the model and the observations is at the beginning of the time period (1998). Grell FC amplitude tends to be less than that of TRMM while the amplitude for MIT-Emanuel is greater.
Correlations of the EOF1 time series with the different climate indices shows that there is a statistically significant correlation with ENSO and IOD, however IOD has the higher correlation; $r > 80\%$ compared to $r > 48\%$ for ENSO (Table 3.1). All the other indices used in the study have low and statistically insignificant correlations.
Figure 3.7 Temporal Variation for the first mode for the Short Rains over the East Africa domain.

Figure 3.8 Correlation plot for OND EOF 1 timeseries with the different climate indices over East Africa.

Composites of the SST (Fig. 3.15) also supports the influence of the Pacific ocean and also the Indian Ocean. Above average rainfall is associated with the warming of the Pacific ocean, and warmer Indian Ocean temperatures along the coast of East Africa. In
Table 3.1 Summary of the correlation of TRMM and model output EOF 1 with the different indices used in the study for the Short Rains Season. IOD-Indian Ocean Dipole, NINO3, QBO-Quasi-Biennial Oscillation, TNA-Tropical Northern Atlantic, TSA-Tropical Southern Atlantic, NAO-North Atlantic Oscillation, WP-West Pacific. The ** indicates correlations that are statistically significant at 95% and * at 90% significance of correlation was done using a t-test.

<table>
<thead>
<tr>
<th>Region</th>
<th>Mode</th>
<th>IOD</th>
<th>NINO3</th>
<th>QBO</th>
<th>TNA</th>
<th>TSA</th>
<th>NAO</th>
<th>WP</th>
<th>TRMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>EA</td>
<td>Grell FC</td>
<td>0.84**</td>
<td>0.51*</td>
<td>0.28</td>
<td>-0.04</td>
<td>0.24</td>
<td>0.07</td>
<td>-0.12</td>
<td>0.92**</td>
</tr>
<tr>
<td></td>
<td>MIT</td>
<td>0.80**</td>
<td>0.57**</td>
<td>0.16</td>
<td>0.02</td>
<td>-0.18</td>
<td>-0.02</td>
<td>-0.16</td>
<td>0.90**</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>-0.84**</td>
<td>-0.46*</td>
<td>-0.30</td>
<td>-0.01</td>
<td>0.19</td>
<td>-0.20</td>
<td>0.01</td>
<td>1**</td>
</tr>
<tr>
<td>LVB</td>
<td>Grell FC</td>
<td>0.68**</td>
<td>0.43*</td>
<td>0.19</td>
<td>-0.02</td>
<td>-0.32</td>
<td>-0.22</td>
<td>-0.43*</td>
<td>-0.86**</td>
</tr>
<tr>
<td></td>
<td>MIT</td>
<td>0.69**</td>
<td>0.35</td>
<td>0.16</td>
<td>0.08</td>
<td>-0.33</td>
<td>-0.01</td>
<td>-0.26</td>
<td>-0.77**</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>-0.76**</td>
<td>-0.43*</td>
<td>0.19</td>
<td>-0.02</td>
<td>-0.18</td>
<td>-0.02</td>
<td>-0.16</td>
<td>1**</td>
</tr>
</tbody>
</table>

addition to the dipole SST over the Indian ocean coastal region the low level (850 hPa) wind composites show strengthening of the easterlies hence an increase of moisture influx from the Indian ocean. The Congo Rainforest and tropical Atlantic provide moisture from the western side of the region however the westerlies are weaker compared to the easterlies (Fig. 3.10). During the dry episodes, the sea surface temperatures are much cooler and the wind anomalies are mostly dominated by westerlies and hence limit moisture influx into the region. The cooler Indo-Pacific SST alters the Walker circulation during the dry episodes and hence divergence over East Africa. Although there are slight differences in the circulation maps from the model simulations and TRMM in general the patterns are the same, thus the model timeseries is able to capture the atmospheric and maritime forcing during the OND season as in the observations.
Figure 3.9 Composites for EOF 1 during the Short Rains for the East Africa Domain (a) GRELL FC (b) MIT (c) TRMM.
The loadings for LVB region (Fig. 3.11) shows that over Lake Victoria Basin the loading over the lake are higher compared to the other regions. Both model simulations reproduce this mode reasonably well. Correlations between the first loadings timeseries for the model and TRMM are statistically significant and are 0.72 and 0.78 for Grell FC and MIT Emanuel respectively. The timeseries for this mode has a 2 year cycle of below normal and above normal precipitation, however there is no trend in the timeseries (Fig. 3.12). Just like EOF1 of the larger domain EOF1 for LVB from TRMM explains approximately 58% of the total rainfall variability, while the EOF1 of Grell and MIT Emanuel explain 52.8% and 48.9%.

Figure 3.10 Composite of the zonal wind using TRMM EOF1 timeseries over East Africa Region (a) above average precipitation (b) below average precipitation.
Figure 3.11 Spatial and temporal distribution for OND EOF1 Lake Victoria Basin Region from models and observation (a) Grell FC (b) MIT Emanuel (c) TRMM.

The correlation of the LVB timeseries with the IOD is statistically significant with a value greater than 0.68 for the models and observational data (Fig. 3.13). Correlations with the NINO 3 index are statistically insignificant (Fig. 3.13) and are lower compared to the correlations with the bigger domain. This suggests that the rainfall drivers for the smaller domain and the larger domain are different. Analysis of the composites for SST and the circulation show that the SST anomalies for the wetter periods are lower over
the Pacific compared to the bigger region, and this could explain the lower correlation.

Figure 3.12 Timeseries plot for OND EOF 1 over Lake Victoria Basin.

Figure 3.13 Correlation plot for OND EOF 1 timeseries over Lake Victoria Basin with the different climate indices.
A comparison of the vertical velocity between Lake Victoria Basin and East Africa region shows that the East-West circulation are much stronger in the LVB domain compared to the East Africa region. While the east-west circulation is stronger over LVB, the composites for the North-South circulation (Hadley Cell) shows more vigorous upward motion for the East Africa region compared to the Lake Victoria Basin region.

The second mode of variability for the Lake Victoria Basin domain explains 9-23% of the variability from the observational data and the regional climate model. This loading has a dipole mode, which runs diagonally across the region, when the northern part is wet the southern part is dry (Fig. 3.16). The time series has a low frequency variability compared to the first mode. This mode is not significantly correlated with any of the chosen indices (Fig. 3.17). A study conducted by Schreck and Semazzi (2004) showed that over the recent decades there has been an increase in the global temperature and this was evident in their second EOF loading which had a dipole mode. Although our datasets shows the dipole mode we cannot conclude that this represents Global warming since our datasets is too short for such a conclusion. The vertical velocity plots over the region for the second mode have the region divided into 2. During the positive rainfall anomalies, the western region experiences strong subsidence thereby reducing total rainfall in that region in contrast with the eastern side which has strong rising motion (Fig. 3.18).

3.4.2.2 Long Rains MAM season

The Long Rains receive higher precipitation compared to the short rains season, however there is less inter-annual variability in this season compared to the Short Rains. A limited number of studies have been conducted to study the inter-annual variability of precipitation during the Long Rains (Camberlin and Philippon, 2002; Liebmann et al., 2014; Nicholson, 2015). Most of the recent studies have focused on the decadal variability. For the Long Rains, the association of the SST’s and seasonal rainfall variability have not yet been established for inter-annual rainfall variability. A few studies have proposed that a strong El Niño event in the Short Rains will have some reminiscent in the Long Rains. Nicholson (2015) suggested the drivers of the Long Rains are not homogeneous throughout the Long Rains season, and the researcher used individual months to build predictors for the Long Rains. A study done by Camberlin and Philippon (2002) showed
that the variability of the onset and peak is different from that of the cessation. Hence in order to get a comprehensive understanding of the season the Long Rains should not be analyzed as the whole season. The separation of the season allows for the identification of the underlying oceanic and atmospheric forcing (Camberlin and Philippon, 2002; Nicholson, 2015). We first analyze the whole Long Rains season and then split the season into two, the Onset and Peak (March-April) and the cessation (May).

East Africa domain has 2 significant modes of variability when the Long Rains are considered. Figure 3.19 is the loading pattern during this season, the loading is generally dominated by a negative loading. In Grell FC there is slightly higher loadings over the Lake and ocean, consistent with the loadings in TRMM. The MIT-Emanuel scheme has more areas with positive loadings that Grell FC and TRMM. The model is able to reproduce the year to year variation in precipitation (Fig. 3.20). Correlations of the time series (Fig. 3.21) with the indices does not have any statistically significant correlations at the 95% level, however some correlations are significant at the 90% level. Thus since there are significant correlations at the 90% level we construct the SST and wind composites for this season. The SST and wind composites for MAM EOF 1 (Fig. 3.22) shows some hot spots that that there is warming in the Pacific Ocean, a dipole mode in the Indian Ocean and some warming in the Atlantic ocean. In the Pacific Ocean there is an ENSO like pattern however the pattern is weak and does not extend to the West Pacific. The gradient between the North Pacific and the Central Pacific is very weak, however correlation of the West pacific index and the timeseries at the 90% confidence level is significant. The correlations with the IOD are not significant at the 95% significance level but are significant for TRMM at the 90% level. Both model simulations do not have significant correlations with the Indian Ocean at the 90% level. Analyzing the SST-wind composites for the Indian Ocean shows that there is moisture influx into the region from the Indian Ocean, and a weak IOD signature.

The Lake Victoria Basin loading for TRMM has a high loading over the lake, however both MIT-Emanuel and Grell FC schemes are unable to capture this (Fig. 3.19). The time series shows alternating above and below average rainfall anomalies, with the last 5 years experiencing a decline in total precipitation. The Indian ocean modulates the precipitation variability over Lake Victoria Basin and the correlation between the Grell
FC timeseries and the IOD index is 0.67 and 0.61 for TRMM. With the MIT-Emanuel scheme a correlation of 0.64 is for the second EOF loading. During the wet years the SST temperatures are warmer along the Indian Ocean Coastal region and this enhance the low level easterlies that bring in moisture into the region. Thus the drivers of precipitation over this region is different from the larger domain. This implies the need to analyze the domain different, since we might be mixing regions with different climatic regimes. The southern part of the larger domain is in a region that experiences unimodal precipitation, while the Basin is mostly bimodal.

Since a number of studies have already alluded to the fact that drivers of precipitation in this season are temporarily distinct we adopt the temporal separation of the season suggested in Camberlin and Philippon (2002). Figure 3.25 is the timeseries for the intra-variability of precipitation. The EOF timeseries for TRMM is significantly correlated with that of the model during the March-April time period for the Lake Victoria Basin and May for the East Africa domain. When the season is split only one EOF is retained for the March-April time period. The analysis of the larger domain for the March-April time period has a statistically significant correlation with the West Pacific Index. Composites for SST show an ENSO like pattern over east Pacific, but this pattern does not extend to the West Pacific. A contrast of north Pacific and west Pacific show much warmer temperatures in north Pacific. This temperature gradient in the in the Pacific has been shown to alter the Walker circulation and hence the total precipitation in the region. The correlation of the EOF1 timeseries for TRMM with the West Pacific index has a statistically significant value of 0.5. Both the model simulation timeseries do not have statistically significant correlations with the West Pacific index.

For the Lake Victoria Basin EOF1 the model simulations are able to reproduce the spatial pattern of the loading as in the TRMM datasets (Fig. 3.28). Correlations with the IOD are statistically significant and all the other correlations are not. In addition to the correlation with the IOD the rainfall during this season is modulated by the 30mb stratospheric zonal winds. The QBO is significantly correlated with the TRMM timeseries. In a study done by Indeje and Semazzi (2000) they found that there were significantly high correlations with the QBO concentrated over the western regions of eastern Africa with values over 0.8. They reported that these high correlation occur more with lagged time-
series rather than month to month. Both model simulations are not correlated with and of the indices that correlate with the TRMM datasets. When the coastal temperatures are high along the East Africa domain the precipitation over the domain increases. The lower winds composited show strengthening of the easterlies during this time period and hence an increase in the moisture transport into the region.
Figure 3.14 Short Rains Composites for EOF 1 for the East Africa Domain (a) GRELL FC (b) MIT (c) TRMM.
Figure 3.15 Vertical velocity (North South) components for the (a)LVB region (b) East Africa.
Figure 3.16 Spatial distribution for OND EOF2 over Lake Victoria Basin from models and observation (a) Grell FC (b) MIT Emmanuel (c) TRMM.
Figure 3.17 Timeseries for OND EOF 2 timeseries over Lake Victoria Basin.

Figure 3.18 Cross section for Omega for OND EOF 2 timeseries over Lake Victoria Basin.
Figure 3.19 Spatial distribution for MAM EOF1 from models and observation (a) Grell FC (b) MIT Emanuel (c) TRMM.
Figure 3.20 MAM EOF1 timeseries for observational data and model over East Africa domain.

Figure 3.21 Correlation plots for MAM EOF1 timeseries and Climate indices over East Africa domain.
Figure 3.22 Long Rains Composites for EOF 1 for the East Africa Domain (a) GRELL FC (b) MIT (c) TRMM.
Table 3.2 Summary of the correlation of TRMM and model output EOF 1 with the different indices used in the study for the Long Rains Season. IOD-Indian Ocean Dipole, NINO3, QBO-Quassi-Biennial Oscillation, TNA-Tropical Northern Atlantic, TSA-Tropical Southern Atlantic, NAO-North Atlantic Oscillation, WP-West Pacific. The ** indicates correlations that are statistically significant at 95% and * at 90% significance of correlation was done using a t-test.

<table>
<thead>
<tr>
<th>Region</th>
<th>Season</th>
<th>Mode</th>
<th>IOD</th>
<th>NINO3</th>
<th>QBO</th>
<th>TNA</th>
<th>TSA</th>
<th>NAO</th>
<th>WP</th>
<th>TRMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAM</td>
<td>MAM</td>
<td>Grell</td>
<td>0.35</td>
<td>0.18</td>
<td>-0.47*</td>
<td>0.04</td>
<td>-0.17</td>
<td>-0.23</td>
<td>0.10</td>
<td>-0.43</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FC</td>
<td>-0.13</td>
<td>-0.07</td>
<td>0.46*</td>
<td>-0.13</td>
<td>0.10</td>
<td>-0.11</td>
<td>-0.16</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MIT</td>
<td>0.46</td>
<td>-0.39</td>
<td>-0.03</td>
<td>-0.14</td>
<td>-0.06</td>
<td>-0.31</td>
<td>-0.41</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TRMM</td>
<td>-0.35</td>
<td>-0.08</td>
<td>-0.06</td>
<td>-0.15</td>
<td>0.04</td>
<td>-0.12</td>
<td>-0.31</td>
<td>-0.51**</td>
</tr>
<tr>
<td>EA</td>
<td>Mar-Apr</td>
<td>Grell</td>
<td>-0.23</td>
<td>-0.25</td>
<td>0.27</td>
<td>-0.21</td>
<td>-0.30</td>
<td>-0.05</td>
<td>-0.07</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>Mar-Apr</td>
<td>FC</td>
<td>0.07</td>
<td>0.04</td>
<td>-0.60**</td>
<td>0.16</td>
<td>-0.33</td>
<td>-0.07</td>
<td>-0.15</td>
<td>-0.67**</td>
</tr>
<tr>
<td></td>
<td>Mar-Apr</td>
<td>MIT</td>
<td>-0.32</td>
<td>-0.21</td>
<td>-0.34</td>
<td>0.01</td>
<td>0.20</td>
<td>-0.24</td>
<td>0.0</td>
<td>-0.6**</td>
</tr>
<tr>
<td></td>
<td>Mar-Apr</td>
<td>TRMM</td>
<td>0.23</td>
<td>0.05</td>
<td>-0.43*</td>
<td>0.26</td>
<td>-0.22</td>
<td>0.02</td>
<td>0.06</td>
<td>1</td>
</tr>
<tr>
<td>MAM</td>
<td>MAY</td>
<td>Grell</td>
<td>0.67**</td>
<td>0.04</td>
<td>0.09</td>
<td>0.24</td>
<td>0.26</td>
<td>0.31</td>
<td>-0.20</td>
<td>0.63**</td>
</tr>
<tr>
<td></td>
<td>MAY</td>
<td>FC</td>
<td>-0.32</td>
<td>-0.11</td>
<td>-0.11</td>
<td>-0.01</td>
<td>-0.46</td>
<td>0.00</td>
<td>-0.21</td>
<td>-0.61**</td>
</tr>
<tr>
<td></td>
<td>MAY</td>
<td>MIT</td>
<td>0.64**</td>
<td>0.06</td>
<td>-0.30</td>
<td>0.05</td>
<td>-0.08</td>
<td>0.36</td>
<td>-0.49*</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>MAY</td>
<td>TRMM</td>
<td>0.61**</td>
<td>0.06</td>
<td>0.34</td>
<td>0.00</td>
<td>0.41</td>
<td>-0.02</td>
<td>-0.71**</td>
<td></td>
</tr>
<tr>
<td>LVB</td>
<td>Mar-Apr</td>
<td>Grell</td>
<td>0.39</td>
<td>0.30</td>
<td>-0.12</td>
<td>0.23</td>
<td>0.00</td>
<td>0.41</td>
<td>-0.02</td>
<td>-0.71**</td>
</tr>
<tr>
<td></td>
<td>Mar-Apr</td>
<td>FC</td>
<td>0.28</td>
<td>0.26</td>
<td>-0.11</td>
<td>0.07</td>
<td>-0.13</td>
<td>0.34</td>
<td>-0.00</td>
<td>-0.72**</td>
</tr>
<tr>
<td></td>
<td>Mar-Apr</td>
<td>MIT</td>
<td>0.53**</td>
<td>0.16</td>
<td>-0.54</td>
<td>0.00</td>
<td>0.40</td>
<td>-0.31</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mar-Apr</td>
<td>TRMM</td>
<td>0.23</td>
<td>0.31</td>
<td>0.57**</td>
<td>-0.09</td>
<td>0.48</td>
<td>0.05</td>
<td>0.14</td>
<td>-0.42</td>
</tr>
<tr>
<td></td>
<td>MAY</td>
<td>Grell</td>
<td>0.32</td>
<td>-0.19</td>
<td>-0.35</td>
<td>-0.01</td>
<td>-0.22</td>
<td>0.25</td>
<td>0.02</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>MAY</td>
<td>FC</td>
<td>0.13</td>
<td>0.40</td>
<td>-0.48**</td>
<td>0.10</td>
<td>-0.29</td>
<td>0.14</td>
<td>0.33</td>
<td>1</td>
</tr>
</tbody>
</table>

89
Figure 3.23 Spatial distribution for MAM EOF1 from models and observation over Lake Victoria Basin (a) Grell FC (b) MIT Emanuel (c) TRMM.
Figure 3.24 Timeseries for MAM EOF1 over Lake Victoria Basin
Figure 3.25 Timeseries plot for the March-April and May season for the 2 regions utilized (a) March-April EA (b) March-April LVB (c) May EA (d) May LVB.
Figure 3.26 Spatial distribution for March-April EOF1 from model and observation over East Africa (a) Grell FC (b) MIT-Emanuel (c) TRMM
Figure 3.27 March-April SST-circulation composites for EOF 1 East Africa Domain (a) Grell FC (b) MIT-Emanuel (c) TRMM.
Figure 3.28 Spatial distribution for March-April EOF1 from models and observation over Lake Victoria Basin (a) Grell FC (b) MIT Emanuel (c) TRMM.
Figure 3.29 March-April Composites for EOF 1 Lake Victoria Basin Region (a) Grell FC (b) MIT-Emanuel (c) TRMM.
Correlations for the May rainfall from the model and observational data is statistically significant at the 90% level (Table 3.2). The timeseries for TRMM and Grell FC is correlated with the QBO index with the Grell FC simulation having the higher value. Indeje and Semazzi (2000) created 8 different homogeneous regions and correlated them with QBO. Their findings revealed the heterogeneity of the region, some regions had high positive correlations while for some the correlations were low. Thus in our case since the domain is bigger we expect that the signal will be diluted and weaker as we are including other regions that are not highly influenced by the QBO. The composites plots for the May season has above normal SST over both the Indian and the Pacific ocean. We propose that the Long Rains are modulated in a complex manner by the interaction of the different ocean basins and atmospheric circulation.
Figure 3.30 Spatial distribution for May EOF1 from model simulations and observation over East Africa (a) Grell FC (b) MIT-Emanuel (c) TRMM.
Figure 3.31 May SST-Circulation composites for EOF 1 East Africa Domain (a) Grell FC (b) MIT-Emanuel (c) TRMM.
Figure 3.32 Spatial distribution for MAY EOF1 from model and observation over Lake Victoria Basin (a) Grell FC (b) MIT Emanuel (c) TRMM.
Figure 3.33 May Composites for EOF 1 Lake Victoria Basin Region (a) GRELL FC (b) MIT (c) TRMM.
We investigate why the model was unable to capture the high loading over the lake. This was done by taking the average precipitation over the lake region where precipitation is high for the 16 years. Both simulations are able to capture the variability of precipitation over the region, however they underestimate the amounts. Over this region the MIT-Emanuel scheme is very dry during the wet season and during the wet seasons it under-predicts the total precipitation. The Grell FC is able to capture the precipitation over the dry season, but however it underestimates the precipitation. While the TRMM datasets shows a negative trend over the region both simulations do not show a trend. Hence the model might not be able to capture the physics and dynamics in the region.

![Figure 3.34](image)

**Figure 3.34** Timeseries plot over the lake for the model and the observational data. The plots were obtained by averaging the precipitation over the lake with high loading.

### 3.5 Conclusions

In this study we investigated the ability of the ICTP RegCM model to reproduce the dominant modes of variability from TRMM. We use 2 simulations, one using the Grell FC scheme and the other one using the MIT-Emanuel scheme. When the EOF analysis is
applied to the 2 different simulations we conclude that the model is able to reproduce the most dominant modes of variability. This is especially so during the Short Rains season where the model time series correlation and that for TRMM is approximately 0.9. When the Lake Victoria domain is considered the correlation between the first modes of TRMM and from the model is reduced. This might imply that the model is unable to capture the mesoscale features over the lake and more work is required to accurately capture the dynamics over the lake. Nonetheless the correlations are statistically significant. For the short rains this study reveals that in the recent years the inter-annual variability of precipitation over the East Africa is mostly dominated by the variability of the Indian Ocean, rather than the ENSO events.

Past studies have shown a weak link between the East Africa rainfall and the SST in the Pacific and Indian ocean. Although the model skill during the Long Rains is not as high as the one during the Short Rains, the model timeseries was able to reproduce the influence of the Indian ocean over the Lake Victoria Basin Domain. The drivers for the larger domain were however elusive. To further this season the study utilized the recommendation from Camberlin and Philippon (2002) to split the season to onset-peak and cessation. The March-April (onset-peak) period is positively correlated to the West Pacific with the TRMM datasets, however both model simulations fail to capture the contribution of the West Pacific. The West pacific has been linked to the East Africa through modification of the Walker Circulation. During the cessation period the QBO indices are significantly correlated to the EOF1 timeseries. Thus for the cessation period more work needs to be done to understand if the percentage contribution of the QBO and the other indices. Recent studies on the cessation have identified different sources of inter-annual variability. With some studies suggesting that the West Pacific plays a major and role, while others have identified the role of the Atlantic Ocean.

In conclusion although the model can simulate the dominant modes of variability during the Short Rains as in the TRMM dataset, the model has less skill during the Long Rains especially when the East Africa domain is considered. More work needs to be done to understand the physical mechanisms the drive the Long Rains season.
Chapter 4

Impact of urbanization on the Climate of East Africa

4.1 Introduction

In the recent decades Africa has been experiencing rapid urbanization, which is expected to continue to increase in the near future (United Nations, 2012, 2014). Figure 4.1 shows the percentage proportion of the total urban dwellers. In the past the percentage of total urban population in Africa has been relatively low compared to the other continents. Nonetheless, in the future the region is projected to have a higher percentage of urban dwellers (approximately 21% in 2050) coming second after Asia (United Nations, 2014). The increase in urbanization could be mostly caused by the migration from rural regions to urbanized regions. The alarming rate of increase in urbanization has caught the attention of the current African Union Chairperson that she dedicated 2 Head of States Summits to the discussion of urbanization in the continent in 2015 (UN Habit, 2015).

A number of countries are projected to have an increase in total population of greater than 25 million by 2050, this includes 9 countries in Africa (Fig. 4.2). Of these nine countries, four of the countries are in East Africa and these are Ethiopia, Kenya, Uganda and United Republic of Tanzania (United Nations, 2014). Currently the total population in Kenya, Uganda and Tanzania is estimated at approximately 150 million. Between 2000 and 2030 the projected population is expected to treble (Seto et al., 2012). Seto et al. (2012) projected that the high increase in population growth in Kenya and Uganda could mostly likely occur over Lake Victoria Basin (LVB). As mentioned before this region is the nerve centre of East Africa and most activities depend on rainfed activities.
The rapid change in urbanization will alter the surface properties of land, and lead to land-atmosphere feedbacks. Conversion of natural vegetated land to paved surfaces alters the moisture, momentum and energy fluxes (Guo et al., 2006; Shem and Shepherd, 2009; Shepherd et al., 2010; Van Den Heever and Cotton, 2007).

A number of studies have been done globally to assess the impact of land cover on the climate in the future (Feddema et al., 2005; McCarthy et al., 2010; Oleson, 2012; Oleson et al., 2011). Tropical regions have been shown to have different responses to changes in land cover, Feddema et al. (2005) showed that conversion of broadleaf forests over the Amazon Basin does not have the same impact as over Indonesia, with the Amazon Basin having an increase in minimum temperature but over Indonesia there are just slight changes. These responses reinforce the importance of using a limited area climate model to access the impact of urbanization in a particular region.

The Weather and Research Forecasting (WRF) model is one of the most extensively tested models in urbanization simulations. The model includes an urban canopy model that has the properties of a city. The model has been applied in a number of studies in midlatitudes (Argüeso et al., 2014; Daniels et al., 2014; Karaca et al., 1995; Mahmood et al., 2014; Miao et al., 2011; Yang et al., 2014a; Zhang et al., 2014), however very limited studies have been conducted in the low latitudes. Table 4.1 summarizes some of the
studies that has been done on the impact of urbanization on precipitation. Most of the experiments have used one way nesting, with the inner most domain running at a high resolution without the cumulus scheme. Their findings indicate that there is a general increase in the amount of precipitation over the urbanized region. Miao et al. (2011) reports that the magnitude depends on the level of urbanization, high urbanized regions...
are likely to receive an increase in precipitation, while low or moderate urbanization lead to a decrease in precipitation. Recently Li et al. (2016) studied the impacts of urbanization in Singapore, they found that increased urbanization leads to higher precipitation and the sea breezes were more influential to precipitation increase than the city design.

While WRF model has been extensively used in the study of the impacts of urbanization on climate, the studies using RegCM are limited. The first attempt to model the impact of urbanization on the climate in RegCM was done by Kueppers et al. (2008), and they found that the diurnal temperature range was reduced. Huszar et al. (2014) coupled the Single-layer Urban Canopy Model to RegCM4 and investigated the impact of urbanization on the climate in central Europe. The study showed that precipitation increased in some cities, but in other cities the increase was not statistically significant. In this section we investigate the impact of changing land cover from vegetated land to urbanized land on the climate in the region. In particular we pay attention to the impacts of the different levels of urbanization to the precipitation and temperature. This study however does not consider the impact of projected climate change but at the impact of urbanization on the climate of the region. Thus when we use future urbanization scenarios we keep the lateral boundary conditions the same with the control simulation. Climate change and urbanization interacts in a complex manner and leads to different intensities in the urban heat island and also total precipitation (McCarthy et al., 2010; Moore et al., 2015). The interaction of the land cover change and the greenhouse gases emission impacts could be investigated in another project in the future.

Here we offer a brief summary of the interactions and expected change in the urban heat island (UHI) and hence precipitation. Oleson et al. (2011) showed that in East Africa the June to August (JJA) season has higher urban heat island effect compared to the December to February (DJF) season. The JJA temperatures are approximately 1°C higher than the DJF. The differences in the UHI in the different seasons are due to the urban-rural surface moisture characteristics. McCarthy et al. (2010) showed that the increase in urban heat island is heterogeneous, with the Middle East having a temperature increase of approximately 3.2°C while East Africa has a temperature increase of approximately 1°C when future urban growth is considered. While McCarthy et al. (2010), shows that the urban heat island effect will increase in the future, Oleson et al. (2011) shows that
<table>
<thead>
<tr>
<th>Authors and Domain</th>
<th>Microphysics</th>
<th>Cumulus</th>
<th>SW and LW</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yang et al. (2014a). Milwaukee-Lake Michigan, Midwest of the US, 22-23 July 2010. Three one way nesting, 9, 3, and 1km</td>
<td>WRF single-moment 6 class</td>
<td>None</td>
<td>Dudhia, RRTM</td>
<td>Deep convection over the urbanized region. Surface skin temperature 5K higher. Large scale rainfall pattern is not sensitive to local scale urban forcing</td>
</tr>
<tr>
<td>Miao et al. (2010) 5 one way domain nesting 40.5km being the most coarse and 0.5km being the finest domain</td>
<td>WRF single-moment 6 class</td>
<td>Grell Devenyi for the 40.5 and 13.5km</td>
<td>Dudhia, RRTM</td>
<td>Urbanized regions break the squall line to convective cells. Lower urbanization leads to surface drying and weak confluence and uplift. Further urbanization led to stronger UHI and vertical mixing, raising PBL height and weakening the capped inversion intensity, lead to more accumulated precipitation.</td>
</tr>
<tr>
<td>Kusaka et al. (2014) 4km grid resolution. 27Jul-1Sep 2001-08</td>
<td>WRF single-moment 6 class</td>
<td>None</td>
<td>Dudhia, RRTM</td>
<td>Positive anomaly of approximately in precipitation over urbanized regions</td>
</tr>
<tr>
<td>Shem and Shepherd (2009) Southeast US. 30km, 10, and 3.3km nested domains</td>
<td>Ferrier Kain Fritsch for two outer domains</td>
<td>None for finer nest</td>
<td>Dudhia, RRTM</td>
<td>Higher accumulated total rainfall over the urbanized case study compared to non urbanized.</td>
</tr>
<tr>
<td>Zhang et al. (2010) Yangtze delta 2 one way nesting 20km and 5km</td>
<td>Lin et al. (1983)</td>
<td>None for finer nest</td>
<td>Simple Short-wave, RRTM</td>
<td>Change in precipitation is low during winter. However it is high during summer. The percentage is approximately 15% higher PBL is 50m higher than the no urban case</td>
</tr>
<tr>
<td>Zhan et al. (2013) Beijing-Tianjin-Tangshan Metropolitan</td>
<td>Lin et al. (1983) Grell Devenyi</td>
<td>None for finer nest</td>
<td>CAM, CAM</td>
<td>Increase in some parts of the urbanized region and a decrease in precipitation over other urbanized regions due to extended urban boundary and increased secondary outflow activity. Buildings of different heights can not only cause mechanical turbulence but also hinder the moving slow precipitation system, thus leading to the increase of precipitation.</td>
</tr>
<tr>
<td>Shimadera et al. (2015) One way nesting Japan, 3 and 1km domains</td>
<td>WRF single moment 6 class</td>
<td>None</td>
<td>Dudhia, RRTM</td>
<td>Increase in precipitation of up to 20mm/day over the urbanized regions</td>
</tr>
</tbody>
</table>
the rural areas will warm up fast and hence reduce the contrast between urban areas and rural areas, thereby reducing the UHI effect. However their simulations considered static land cover and hence there was no meaningful increase in the urbanization footprint. Olesen (2012) considered 3 different Representative Concentration Pathways (RCPs), and found that the average UHI stays constant for RCP2.6 and RCP4.5, however the daytime UHI decreases while the nocturnal stays constant. In the RCP8.5 scenario both the daytime and nocturnal UHI decrease globally. Over LVB the Northern part experiences an increase in UHI intensity (0.2 – 0.6°C) in all scenarios and seasons while the western part experiences a decline in the UHI effect (−0.1 – 0.3°C).

4.2 Methods of Analysis

The study will utilize both the Weather and Research Forecasting model (WRFv3.7) and the fourth generation model for the ICTP Regional Climate Model version 4.4 (RegCM4.4). The simulations are conducted for a continuous 5 years and have a spin-up of 7 months, thus they run from June 2005 to December 2010. The models are run at different resolutions, RegCM4 is a hydrostatic model (Elguindi et al., 2013) with the finest resolution being 20km, subgridding can be done for finer resolution, however when subgridding is done it only calculates other variables at the finer resolution but does not calculate the precipitation at the finer resolution. WRFv3.7 is a non hydrostatic model (Wang et al., 2014a) that can be run at very fine resolution, with some researchers using resolutions that are lower than 1km.

Table 4.2 summarizes the model physics and resolution used in this study. We maintain the same resolution as that was used in Chapters 2 and 3 of 25km for the RegCM simulations. For the WRF simulations we use 2 domains, the parent domain has a resolution of 36km and the nested domain at 12km. We adopt a one way nesting, meaning that there is no two way feedback between the outer domain and the inner domain, the inner domain uses the lateral boundary conditions from the outer domain, however it does not overwrite the variables in the first domain. A one way nesting allows for the investigation of the impact of urbanization on the climate when different resolutions are utilized.

Figure 4.3 shows some of the land cover types that are available in WRF and over our
region of interest. The regions around the lake are mostly non-irrigated croplands and pasture. Croplands and pasture have different vegetation properties depending on the season. During the dry seasons we expect the soil moisture and the vegetation fraction to decrease compared to the wet season. Thus in this region we expect the surface albedo to change between the wet and dry seasons. The two models use land cover datasets from the Global Land Cover Characterization (GLCC) however the land use classes are defined differently. RegCM uses 22 land surface categories while WRF uses 24 land surface categories for the United States Geological Survey classification.

### 4.2.1 Precipitation Datasets

Evaluation techniques, that were applied in Chapter 2 will also be applied to the analysis in this section. Sylla et al. (2013) evaluated the performance of RegCM3 using different observational datasets and found that the models performs differently according to the dataset that is utilized, this mostly occurred when the higher frequency precipitation was evaluated. Hence in addition to using the Tropical Rainfall Measuring Mission (TRMM) for precipitation evaluation, other gridded and satellite datasets are utilized, so as to obtain the spread of uncertainty from observations.

#### 4.2.1.1 Satellite

Recently the use of satellite data to evaluate regional climate models over the African region has been on the rise (Argent et al., 2015; Davis et al., 2009; Sun et al., 2014; Sylla

---

**Table 4.2** Model configuration used in the study for the REGCM and the WRF model.

<table>
<thead>
<tr>
<th>Content</th>
<th>RegCM4.0</th>
<th>WRFv3.7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution</td>
<td>25km</td>
<td>36km and 12km</td>
</tr>
<tr>
<td>Map projection</td>
<td>Normal Mercator</td>
<td>Normal Mercator</td>
</tr>
<tr>
<td>Cumulus</td>
<td>MIT-EMAN(Emanuel, 1991)</td>
<td>BMJ(Janjic, 1994)</td>
</tr>
<tr>
<td>Planetary boundary layer</td>
<td>Holtslag (Holtslag and Boville, 1993)</td>
<td>YSU Hong et al. (2006)</td>
</tr>
<tr>
<td>Microphysics</td>
<td>SUBEX(Pal et al., 2000)</td>
<td>WSM6(Hong and Lim, 2006)</td>
</tr>
<tr>
<td>Land Surface Processes</td>
<td>BATS (Dickinson et al., 1993)</td>
<td>NOAH (Chen and Dudhia, 2001)</td>
</tr>
</tbody>
</table>
et al., 2013) since the datasets are readily available and has high spatial and temporal variability. In this section we consider 2 datasets that solely rely on satellite data, the Climate Prediction Center(CPC) MORPHing technique (CMORPH) and the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN). These datasets have been evaluated before in the region and have been shown to perform reasonably well. Biases were mostly high in regions that had complex topography. We will discuss the performance of each of the datasets in the region in the following paragraphs.

CMORPH is available at high temporal and spatial resolutions, it includes temporal resolutions of 30 minutes, 3 hourly and daily and a spatial resolution of 8km and 0.25° (Joyce et al., 2004). The data is obtained from the low frequency passive microwave derived precipitation estimates that are generated from the following satellites; NOAA

Figure 4.3 Landuse Cover for the inner domain in the control simulation for WRF.
polar orbiting, US Defense Meteorological Satellite Program satellites, and TRMM. The CMORPH datasets have been found to perform well over East Africa, often having lower biases compared to other precipitation datasets. Compared to the TMPA-3B42 dataset, the CMORPH dataset has a lower precipitation underestimation (Habib et al., 2012).

The PERSIANN dataset uses neural network approach applied to geostationary infrared brightness temperatures to obtain the precipitation rates (Hsu and Sorooshian, 2009). The spatial resolution for the data is 0.25° and is available every 6 hours. The network parameters are updated through an adaptive procedure using the instantaneous rain rate estimates from TRMM. Sorooshian et al. (2000) showed that when the data is combined with TRMM products the bias in the dataset is reduced. The errors are further reduced for accumulated daily data. The PERSIANN datasets was evaluated over the region performed well in detecting regions that have high precipitation, however they did not capture the correct amount (Dinku et al., 2008). Hirpa et al. (2010) evaluated the dataset over Ethiopia and found that the dataset considerably underestimates rainfall in high-elevation areas.

4.2.1.2 Gridded

Gridded rainfall datasets have been widely used over East Africa (Kim et al., 2014; Ogwang et al., 2015; Sylla et al., 2013) for different climate studies, including diagnostic evaluation of regional climate models. One drawback of using these datasets is that in the past decades the number of stations has been on the decline in East Africa (Harris et al., 2014; Rowell, 2013). In this study we use 2 gridded datasets over the region, the East Anglia Climate Research Unit (CRU) and the Global Precipitation Climatology Centre (GPCC). Both these datasets are available at monthly timescales and are available over land. The GPCCv7 dataset is available in different resolutions (0.5°, 1°, 2.5°), this study utilizes the 0.5° resolution (Schneider et al., 2008). The CRU TS 3.23 (Harris et al., 2014) datasets is also available at various spatial resolutions just like the GPCC datasets, in this study we use the 0.5° degrees resolution from both datasets.
4.2.1.3 Satellite and Gridded

The Global Precipitation Climatology Project (GPCP) 1DD is a daily dataset with a resolution of 1° produced by combing data from the TRMM satellite and ground based data.

4.2.2 Temperature Datasets

Datasets for evaluating temperature are more limited. The analysis utilizes 2 gridded global datasets, the CRU(TS3.23) and the University of Delaware (UDELAWAREv4.0.1) monthly gridded datasets. The datasets have a spatial resolution of 0.5°, and are only available over land.

4.2.3 Urbanization Scenarios

The urbanization scenarios used in this section closely follow the projected urbanization rates from Seto et al. (2012). We employ 2 urbanization scenarios, the low urbanization and the high urbanization scenario. The low urbanization scenario is obtained by considering the urbanization values that had higher probabilities of being urbanized. In this case we define higher probabilities of urbanization to be 90% and above. Thus we are considering most of the regions that have red shading in the Seto et al. (2012) urbanization projections. The higher urbanization scenario uses regions that will be expected to be urbanized regardless of how low the projected probability of urbanization in the region. Thus we are taking all values that have a range of probability of 1% to 100% for urbanization. The metadata for the urbanization is obtained from the following website http://urban.yale.edu/data and has a grid spacing of 5km. Figure 4.4, shows the high urbanization scenario used in the study and Fig. 4.5 is for the low urbanization scenario. The urbanization mostly occurred around Lake Victoria shores.
Figure 4.4 Landuse cover for the high urbanization scenario.
Figure 4.5 Landuse cover for the low urbanization scenario.
4.3 Results and Discussion

The results are presented in two sections. The first section discusses the performance of the models. We evaluate the temperature and precipitation. The second section discusses how urbanization will impact the climate at a seasonal scale. We also inter-compare the sensitivity of the two models to land cover.

4.3.1 Model Validation

In this section we start by evaluating the performance of both models during the Short and the Long Rains. Although statistics for both resolutions of WRF are calculated, results presented in this chapter refer to the inner domain of the simulation. For the performance and statistics of the parent domain refer to Appendix C. Thus from hereafter WRF simulations refers to Domain 2 of WRF. The regions evaluated for RegCM and WRF are the same. In RegCM we extract the same region as in WRF that is we analyze the following region (8°S – 8.5°N, 15°E – 48°E).

4.3.1.1 Evaluation for the Short Rains

RegCMv4.4 and WRFv3.7 capture the spatial distribution of seasonal temperature in the region (Fig. 4.6), with both models having high correlations ($r > 0.84$). The ratio of the variances from using either observational datasets is close to 1. The normalized root mean square errors are relatively low ($nrmse < 0.25$), thus the model is able to capture the variability of temperature in the region, in a case of perfect agreement the normalized root mean square error should approach zero and the correlation 1. The observational datasets used have different variability. When WRF is evaluated against CRU we can conclude that in most of the years the model spatial variability is high, however when the UDELAWARE dataset is used the spatial variability of WRF will be less than the observation. For the RegCM model, the temperature variability is generally less than that of the observations, but the values are close to 1.

In most cases both models have a cold bias, however the bias is low with temperatures having a bias of less than 1°C (Table C.1). The model biases are low when using the
UDELAWARE dataset and relatively higher when the CRU dataset is used. The normalized mean bias errors and the normalized mean error is less than 5% for all the years and all the simulations thus our model does very well in reproducing the temperature over the region. Past studies have shown that improper presentation of the land surface characteristics can lead to cold biases in addition the lateral boundary conditions can also cause a cold bias (Branković et al., 2012; Caldwell et al., 2009). Both models use the 2000 (GLCC) land cover datasets, which might not be a true representative of the current land cover, thus the models could capture a slightly high albedo than observed (Rauscher et al., 2010). When evaluated using the Pielke Sr (2013) evaluation measure the models show skill in simulation the mean temperature. In all cases the root mean square error was far much less than the standard deviation of the observations.

Both models are able to capture spatial variability of precipitation in the region (Fig. C.1). Figure C.1 shows the total precipitation from year to year during the short rains season, the year 2006 has the highest precipitation compared to the other years, while 2010 has the lowest total precipitation.
Figure 4.7 summarizes the spatial variability of precipitation in the region from the 2 models. The model and observations have a high correlation when the WRF model is utilized, the least value is $r > 0.69$ and the maximum is $r < 0.86$. The RegCM correlations are relatively lower compared to the WRF correlations with a range of $r > 0.4$ to $r < 0.77$. Both the models are able to reproduce the spatial pattern over the region. Although the models are able to capture the spatial pattern over the domain they are unable to capture the intensity of the precipitation, often over predicting the values. In most of the years the standard deviation of the models is higher than that of the observations. The observational datasets have different standard deviations. GPCC and CRU have much lower standard deviation, this could be because they use limited station data. In addition the number of stations have been on the decline in the region and hence the interpolation is using only a limited number of datasets.

Table C.2 is a summary of the precipitation statistics that were calculated for WRF during the OND season. The normalized mean bias errors are very low in all the years. The normalized mean errors are relatively higher indicating that for normalized mean bias there was cancelation of errors. Thus our model is over predicting precipitation in some regions while under-predicting precipitation in others. Evaluation of model performance using the standard deviation and the root mean square shows that the model has skill in some years and with respect to some datasets. The years 2007, 2008, and 2010 have the highest skill. These years had moderate precipitation while 2006 had very high precipitation. Thus from these results, WRF model is able to skillfully model the precipitation during moderate rain season, and overestimates the total precipitation in an extreme wet year. Precipitation differences in model and observational datasets has been attributed to the forcing data and also the physics combinations being utilized (Crétat and Pohl, 2012; Pohl et al., 2011).

For RegCM (Table C.3) the mean bias are less than 65mm for the OND season and the normalized mean bias errors are mostly less than 20%. However just as in the WRF model the NME are relatively high (40 – 60%). This implies that our models underestimates and overestimate precipitation in the different regions.
4.3.1.2 Evaluation for the Long Rains

During the Long Rains season the temperature for both models have high correlations with the observations (Fig. 4.8). The correlation in both models are greater than 0.88 in all the years and when using the different observational datasets. The normalized
root mean square errors are relatively low indicating that the differences in the observations and the modeled values are low. The standard deviations of RegCM model is slightly lower than that of the observations. For the WRF model the standard deviations are close to the standard deviation of the observation when evaluated against the CRU datasets and slightly lower for the UDELAWARE dataset. In this case the WRF model performs slightly better than the RegCM model.

The temperature mean biases in this season are low with all the values being less than 1°C, and most of them close to zero. In most of the years and against the different observational datasets the standard deviation of the observations is almost equal to the standard deviation of model output, and the root mean square errors are much less than the standard deviations. Thus we can conclude that the model shows skill in reproducing the temperature amplitude and variance. The normalized mean bias and the normalized mean errors are all low implying that the model errors are small throughout the domain, and no regions have high underestimation or overestimation.
For precipitation, there is more spread in the performance of the model from year to year and within the different observational datasets for (Fig. 4.9). The performance of the model in terms of precipitation is lower for this season compared to the Short Rains season. This could be because the Short Rains season was used as the testbed, and the optimal conditions during this season were used for the Long Rains season too. In addition to the customization procedure the season physical mechanisms are not yet well understood and might contribute to incorrect representation of the season in models. The spatial correlations from both models are relatively high, with all the correlations being greater that 0.45 and some being as high as 0.78. Hence in terms of spatial correlation both models are able to reproduce the precipitation patterns over the region. However the models are failing to capture the amplitude of precipitation during the season.

Table C.5 and Table C.6 are the summary statistics for the WRF and RegCM precipitation evaluation. The NMB has a range of between 1% to about 30%, while the NME has errors that range from 40% to 50%. Thus the models less accurately capture the amount of precipitation in the season, but are able to capture the spatial correlations.

From the evaluation we can conclude that both models are able to capture the temperature and precipitation variability. Both models show higher skill in simulating the temperature, with the WRF model showing higher skill than the RegCM model. In both models the Short Rains season has superior statistics compared to the Long Rains.
Figure 4.9 Taylor diagram for temperature during the Long Rains season (a) WRF (b) RegCM.
4.3.2 Impact of Urbanization on the Climate of East Africa: High Urbanization Scenario

4.3.2.1 Land Surface Albedo

Urbanized regions have lower surface albedo than the vegetated land that it replaces (Fig 4.10). These regions have impervious, dark layers that trap shortwave radiation more efficiently than vegetated regions. The original land cover detects by how much the albedo changes. In regions that were initially croplands we expect a larger change compared to a region that initially was forested. Surface albedo for urbanized regions are approximately 0.15 (Sailor, 1995), while those for croplands vary seasonally and with soil type (Betts, 2000) with surface albedo as high as 0.21.

The surface albedo is insensitive to the different wet seasons. The changes in surface albedo are insignificantly different for the Long and the Short Rains. The WRF model is more sensitive to changes in land cover. The changes in albedo in the WRF model are slightly higher than the RegCM model. The differences are due to the different formulation and parameterization for the different vegetation types and also the methods of albedo calculation.
Figure 4.10 Change in albedo between the control simulation and the high urbanization scenario for the 2 models (a) WRF MAM (b) RegCM MAM (c) WRF OND (d) RegCM OND.
4.3.2.2 Surface Energy Fluxes

Urbanized regions have been found to alter the partitioning of the latent and sensible heat fluxes (Wang et al., 2012a). Our study shows consistent results with the changes in energy partitioning over urbanized regions (Fig. 4.11 -Fig. 4.14). In these regions the sensible heat flux increases, while the latent heat flux decreases. The urbanized regions have lower latent heat fluxes due to a decrease in moisture availability through reduction in soil moisture and evapotranspiration. During the Long Rains the WRF models has a lower change in the sensible heat fluxes over the north-eastern part of the basin, while for RegCM the minimum changes occur over the western side of the domain. The uneven changes in the sensible heat could be due to the differences in the properties of the initial land cover. The regions that were originally savannas have the least changes in the sensible heat and latent heat fluxes, while the croplands have a higher change.

![Figure 4.11](image-url)

**Figure 4.11** Changes in energy fluxes during the Long Rains for the WRF model (a) Latent heat (b) Sensible heat flux.

The latent heat fluxes reduce more during the OND (Fig. 4.13,) season compared to during the MAM (Fig. 4.11) season. Since the water cycle affects partitioning of energy fluxes, we propose that this could affect the changes in fluxes in the different seasons. The short rains are preceded by the longer dry season (June-September (JJAS)) and hence during this time the soil is much drier. In addition the MAM season receives more
Figure 4.12 Changes in energy fluxes during the Long Rains for the REGCM model (a) latent heat (b) sensible heat flux.

precipitation in comparison to the OND season, thus soil moisture is higher.

Figure 4.13 Changes in energy fluxes during the Short Rains season for the WRF Model (a) Latent heat flux (b) Sensible heat flux.
Figure 4.14 Changes in energy fluxes during the Short Rains season for the REGCM model (a) Latent heat flux (b) Sensible heat flux.

Figure 4.15 is the Bowen ratio for the WRF model during the OND and MAM seasons. The Bowen ratio is higher for the urbanized regions in comparison to the non urbanized regions. This implies that the sensible heat fluxes increased over the region and/or the latent heat fluxes reduced. The highest changes in magnitude occur during the nocturnal hours, where the Bowen ratio of urbanized regions is substantially higher. Although the partition changes in terms of magnitude, urbanization does not change the diurnal cycle pattern of Bowen ratio.
Figure 4.15 Diurnal variation of the Bowen ratio for the WRF model during the wet season (a) Long rains (b) Short rains.
4.3.2.3 Surface Hydrology

Urbanization leads to a reduction in the vegetation fractions, the reduced vegetation fraction leads to a decrease in the total evapotranspiration over the urbanized regions (Fig. 4.16). Surface evapotranspiration changes are seasonally dependent. During the MAM season there is minimum reduction of the evapotranspiration over the Northern part of the domain for WRF. While for RegCM the minimum changes occur during the OND season. These differences could be due to the original soil moisture properties and the variable land cover during the seasons. In WRF we also investigate changes in the drier seasons. The January-February (JF) season experiences lower changes in evapotranspiration than the JJAS season. We propose that the changes are mostly due to changes in soil moisture and also the vegetation type in the season.

Figure 4.16 Changes in evapotranspiration in the different seasons for the WRF model (a) JF (b) MAM (c) JJAS (d) OND.
The surface runoff in the urbanized regions is higher than in the non-urbanized region. Urbanized regions are made up of mostly impervious layers and this reduces the amount of water that seeps into the soil. The reduction of seepage water leads to an increase in the amount of surface runoff (Fig. 4.18, Fig. 4.19). The two wet seasons show comparable changes in the surface runoff for both models. Higher surface runoff in the two seasons is high in the mountain regions. The REGCM model has modest changes in comparison to WRF.
Figure 4.18 Changes in surface runoff in different seasons for the WRF model.

Figure 4.19 Changes in surface runoff for the REGCM model (a) MAM (b) OND.
4.3.2.4 Temperature

The conversion of the land cover from mostly croplands to urbanized regions results in an increase in the temperature over the urbanized regions. The urbanized region has a low albedo, thus they absorb more shortwave radiation at the surface than what vegetated land such as croplands, grasslands. The average increase in temperature is approximately 3°C, with other regions having an increase of greater than 4°C. The change in temperature in the RegCM model is lower (Fig 4.20) than that of the WRF model (Fig 4.21). This is because the albedo and energy fluxes changes in RegCM are lower in comparison to WRF.

\[ \text{Temperature} \]

\[ 3^\circ C \]

\[ 4^\circ C \]

This is because the albedo and energy fluxes changes in RegCM are lower in comparison to WRF.

**Figure 4.20** Change in temperature between the control simulation and the high urbanization scenario during the (a) March-May (b) October to December season for the RegCM model.
Figure 4.21 Change in temperature between the control simulation and the high urbanization scenario in the different seasons (a) January-February (b) March-May (c) June-September (d) October-December for the WRF model.

We also explore the differences in the temperature for WRF model during the dry season. The seasons show different changes in temperatures, JJAS has the lowest increases in temperature while JF has the highest increases in temperature (Fig 4.21). The soil moisture has been found to play a fundamental role in modulating the surface temperature. During the wet season the soil moisture has been found to dampen the temperature variations (Jiang et al., 2015). The lower changes in temperature in the JJAS season could be due to the higher wind speed. During the JJAS season the Somali Jet strengthens. Thus we propose that both the soil moisture and the wind speed play a major role in the seasonal variations of temperatures over LVB.

In this study we define our Urban Heat Island (UHI) as the difference between the 2m temperature of the sensitivity simulation and the 2m temperature of the control sim-
ulation. From Fig. 4.21 the UHI is least during the JJAS season (approximately 2 – 3°C) and highest during the JF season (approximately 3 – 4°C). The two wet seasons have an increase in temperature that is comparable. The highest changes in temperature occur over the North Western parts of LVB. The Urbanized regions on the eastern side of the domain experience the lowest temperature change.

The probability distribution function (Fig. 4.22) is constructed by extracting temperatures at each grid point and each time series over the following region (6°S – 6°N, 25°E – 40°E). There is a slight shift in the mean temperature however there is no major change in the variance. There is a general increase in the frequency of warmer minimum temperatures for both seasons. The maximum temperatures however do not shift much in comparison with the minimum temperatures. These trends in the surface temperature can be explained by considering the properties of the urban land cover. During the day the surfaces absorb more heat compared to the vegetated areas, at night the heat is released more slowly and hence elevating the minimum temperatures. This leads to a reduction in the diurnal temperature range over the urbanized regions.
We analyze the changes in the diurnal cycle using the region identified in the PDF analysis ($6^\circ S - 6^\circ N, 25^\circ E - 40^\circ E$). For the diurnal cycle we only show the plots for the WRF model. During most of the day the average temperature of the simulation conducted with urbanized regions is slightly higher than the control simulation (Fig. 4.23). At night however the temperatures over the urbanized scenario simulation does not cool down.
compared to the control simulation. The decrease in the diurnal temperature range is approximately $1^\circ C$ during the Long Rains season and $0.7^\circ C$ during the Short Rains season.

![Temperature-Time Graph](image)

(a) MAM

![Temperature-Time Graph](image)

(b) OND

**Figure 4.23** Diurnal cycle of temperature for the WRF model during the 2 wet seasons (a) March-May (b) October-December.
4.3.2.5 Precipitation

A number of possible mechanisms have been proposed for the increase in the precipitation over urbanized regions. These include (i) increased convergence due to the high surface roughness of the urbanized regions compared to the other types of land cover. (ii) urban heat island generated convective clouds (iii) increased aerosols in the atmosphere which then increases the cloud condensation nuclei.

We first discuss changes in precipitation for the WRF model. Figure 4.24 is the difference plot for precipitation in the year 2006 and 2009. Over the urbanized regions there is generally an increase in the daily precipitation. Precipitation increases are as high as 3mm/day in some regions. The amount of precipitation changes from year to year, with the year 2009 having the highest increase in precipitation compared to the other years. In addition to the year to year variability, the precipitation change is not uniform over the urbanized regions. In 2009 the northern part of Lake Victoria receives increased precipitation, while the southern part has a slight decrease. For the year 2006 Lake Victoria does not have significant changes in precipitation. The western side of LVB along the mountain ranges have high increases in the total precipitation, while the changes on the eastern side are relatively weaker. We propose that urbanization interacts with the lake breezes and mountain breezes in a way that leads to high increases in the precipitation.

![Change in precipitation between the control simulation and the high urbanization scenario during the Long Rains.](image)

**Figure 4.24** Change in precipitation between the control simulation and the high urbanization scenario during the Long Rains.
The difference plots for the moisture over the mountain ranges and the lake at 1°S latitude are depicted in Fig. 4.25 for the Long Rains. The difference plots show elevated relative humidity over the mountain ranges on the western side of the lake. Increases in moisture are more evident during the nocturnal hours (Fig. 4.25). The year 2009 has a higher increase in the moisture content than the year 2006.

**Figure 4.25** Change in relative humidity between the control simulation and the high urbanization scenario during the Long Rains. Slice created at 1°S latitude. The lake and mountain ranges are located between the points 120-230.
Figure 4.26 is an illustration of the changing land lake breeze in 6 hour intervals and also the temperature changes for the Long Rains. During the nocturnal hours the region has elevated temperatures, that are up to 4°C higher than over the non-urbanized regions. The temperature difference are non existent at 09UTC hours. In general as stated before the urbanized regions have elevated temperatures at night compared to during the day. The temperature differences lead to altering the land lake breeze in the region. In the late evening when the lake breeze is dominant, the land is much warmer than the lake, this leads to stronger lake breezes in the region, and hence an increase in total precipitation. At night the warmer temperatures over the land, weaken the land breeze. This is most evident at 00UTC when the surface wind are calm in the northern part of the urbanized regions.

The two seasons are also different, with MAM season having higher precipitation compared to the OND season (Fig. 4.27). The changes in the absolute precipitation could also depend on the magnitude of precipitation in a particular season. The moisture changes during the Short Rains (Fig. 4.28) are lower in comparison to the Long Rains.
Figure 4.26 Surface circulation and temperature for the control simulation (left column) and high urbanization scenario (right column) simulation and the high urbanization scenario during the MAM season.
Figure 4.27 Change in precipitation between the control simulation and the high urbanization scenario during the Short Rains.
Figure 4.28 Change in relative humidity between the control simulation and the high urbanization scenario for the Short Rains. Slice created at 1°S latitude. The lake and mountain ranges are located between the points 120-230.
We plot the probability density functions for daily precipitation to show the shifts in precipitation (Fig. 4.30). During the Long Rains season the light rain events in urbanized regions are slightly higher than the control simulation. For the heavier rain events the frequency of obtaining heavier rainfall is higher for the urbanized case. A number of studies have already alluded that urbanized regions have intense precipitation because of the UHI induced convection. During the Short Rains there are instances in which over urbanized regions the probability of light rain is lower than in the control scenario. However the frequency of the region receiving heavy rain is higher in comparison with the Long Rains. The differences in the precipitation in the two seasons could be caused by the interaction of the large scale drivers and the local urbanization effects.

The changes in the RegCM model (Fig. 4.31) did not have major changes from year to year. We show the results for the year 2006. The high precipitation changes occur over the western side of the domain, however some changes occur in the eastern side. The 2 models have consistent results in terms of where the highest increase in precipitation will occur. Just as in the WRF case the changes in total precipitation is higher during the Long Rains than the Short Rains. The RegCM model does not have changes in the light precipitation events. For the 10\(mm\) – 30\(mm\) the urbanized regions had higher precipitation compared to the control simulation. The changes are higher in the MAM season.

For the high urbanization scenario, we conclude the changes in the urbanization will increase the total precipitation over the region. Both models show a higher increase on the western side of Lake Victoria compared to the eastern side. The greater increase in precipitation on the western side could be due to the interaction of the background flow, the land-lake breeze and the mountain breezes. The differences in the two seasons could be due to the interaction of the local circulation with the Intertropical Convergence Zone.
Figure 4.29 Circulation and temperature the control simulation (left column) and high urbanization scenario (right column) simulation and the high urbanization scenario during the OND season.
Figure 4.30 Probability density function for the precipitation for the during the 2 wet seasons for WRF (a) MAM (<10mm) (b) MAM (10mm-30mm) (c) OND (<10mm) (d) OND (10mm-30mm).
Figure 4.31 Change in precipitation between the control simulation and the high urbanization scenario during (a)MAM (b) OND for RegCM.
Figure 4.32 Probability density function for the precipitation for the during the 2 wet seas-
sons for RegCM (a) MAM (<10mm) (b) MAM (10mm-30mm) (c) OND (<10mm) (d) OND (10mm-30mm).
4.3.3 Impact of Urbanization on the Climate of East Africa: Low Urbanization Scenario

Some studies have highlighted that the intensity of the precipitation changes (Miao et al., 2011) and the increase in the temperatures is mainly proportional to the size of the urbanized regions. Thus in addition to the high urbanization scenario we investigate the impact of moderate changes in urbanization on the climate. We hypothesize that the changes in the climate variables for the low urbanization scenario will be less than that of the higher urbanization scenario. We will analyze the WRF model in this case.

4.3.4 Surface Energy Fluxes

The surface energy fluxes change for this scenario are mostly localized just like the high urbanization scenario (Fig. 4.33). The changes in the energy fluxes are lower during the Long Rains season. Our results are consistent with other studies that have shown that the changes in the energy fluxes are proportional to the size of urbanization (Miao et al., 2011). The differences in the changes are due to differences in the moisture content between the two seasons, the high moisture content during the Long Rains minimizes changes in energy partition in comparison with the Short Rains.
Figure 4.33 Change in surface energy fluxes for the low urbanization scenario during the long and short rains (a) MAM latent heat (b) MAM sensible heat (c) OND latent heat (d) OND sensible heat.
4.3.5 Temperature

Figure 4.34 is the difference in temperature between the low urbanization scenario and the control simulation. Since temperature changes in urbanized regions are localized, this scenario has lower temperature increases. The UHI in this case is relatively weaker than in the high urbanization scenario. Temperature changes are less than 3°C for both seasons. The Long Rains season has slightly warmer temperatures than the Short rains season. Although the changes in sensible heat are higher during the Short Rains seasons, cloud cover during the Long Rain season is higher and could possibly increase the surface temperatures.

![Figure 4.34 Change in surface temperature for the low urbanization scenario during (a) MAM (b) OND.](image)

4.3.6 Precipitation

The low urbanization scenario shows minimum changes in the precipitation difference (Fig. 4.35). During the MAM season increases in precipitation are only over the mountain ranges. Over the northern part of Lake Victoria there are no changes in the precipitation. The changes in the OND season are rather very mixed. The eastern side of Lake Victoria has precipitation decline. The mountain region generally experiences an increase in precipitation however the increases are very small compared to the high urbanization
scenario.

![Figure 4.35](image)

**Figure 4.35** Change in precipitation for the low urbanization scenario during (a) MAM (b) OND.

### 4.3.7 Conclusions

This study investigated the impact of urbanization on East Africa Climate. Projected urban land cover datasets are obtained from Seto et al. (2012). We focused our analysis on its impact on the energy fluxes, temperature and precipitation. Urbanized regions have higher sensible heat fluxes, and reduced latent heat fluxes due to the modified urban surface. The diurnal variation of the Bowen ratio shows that the ratio does not change much between 4 UTC and 8 UTC however more pronounced changes are obtained between 16 UTC and 22 UTC. In general the Bowen ratio over urbanized regions increases.

There is a shift in the average temperatures in urbanized regions towards warmer temperatures. The shifts in minimum temperatures are higher than the maximum temperatures for both models. The diurnal cycle range of temperature is reduced by approximately 1°C during the Long Rains, while during the Short Rains the reduction is approximately 0.7°C. The differences in the hourly temperatures are more evident during the nocturnal hours. The changes in temperature for the high urbanization scenario are approximately 2°C higher than the low urbanization scenario. Increase in temperature
around the lake modified the lake-land breezes. During the day the land is much warmer than the lake and hence this strengthens the lake breeze, at night the land breeze is weakened in some cases due to the warm land surface.

Both models show increases in the total precipitation over the urbanized regions when the high urbanization scenario is utilized. The highest increases in precipitation occur on the western side of Lake Victoria Basin. We propose that the increase on this side are a combination of the prevailing easterlies, interaction of the urban circulation, lake breezes and the mountain breezes. These results also show that urbanizing on the eastern side of the Basin has a lesser impact on the total precipitation. The PDF plots for both models show an increase in the heavy precipitation during the Long Rains, but minimum changes in the light precipitation. During the OND season the urbanized regions has less frequency of the light rain, while the RegCM model does not have changes. The heavier precipitation events are also high in the Short Rains season just like the Long Rains season.

In our hypothesis we proposed that the high levels of urbanization have a greater impact on the climate compared to low levels of urbanization. This study has shown that over LVB the increase in the urbanized land will lead to an increase in total precipitation, temperature and sensible heat. The different regions that will be urbanized will not have a uniform increase in the total precipitation. The eastern part of LVB will have the least impact of increased urbanization, and the western part of LVB will have the highest impact. These results are very important for future urban planning. The western side of LVB is likely to have higher frequencies of flooding in comparison with the eastern side.
Chapter 5

Identification of the Dominant Processes

5.1 Introduction

The objective of this study is to identify some of the processes that are dominant in producing the climate characteristics of an urbanized region. In our hypothesis we identify 3 processes that could lead to a change in climate in the region, these are thermally driven convergence, frictionally driven convergence and moisture supply, from both the lake and the soil. These processes are directly affected by the land cover properties. In this study we test the sensitivity of the albedo, surface roughness, soil moisture and moisture from Lake Victoria in producing the precipitation patterns over the region.

Charney (1975) showed that an increase in the surface albedo will lead to surface cooling, increased divergence and hence a reduction in the total precipitation. The surface albedo over urbanized regions is generally lower than that of vegetated land. A number of studies have shown that over urbanized regions the surface temperature is higher than the surrounding rural temperatures (Arnfield, 2003; Gallo and Xian, 2014; Rizwan et al., 2008). The changes in the temperature have been attributed to the changes in the land cover from vegetated land cover datasets to impervious layers. The impervious dark layers absorb more shortwave radiation compared to the vegetated regions and this leads to the urban heat island. Sailor (1995) investigated the impact of increasing the surface albedo over urbanized regions in Los Angeles and their study mainly focused on temperature changes. They found that increasing albedo over the urbanized regions, decreased peak summertime temperatures by as much as 1.5°C.
Sud and Smith (1985) studied the effect of altered surface roughness in deserts on the circulation. They showed that the reduced low surface roughness over deserts reduces the exchange of momentum and virtual temperature fluxes between the ground and the planetary boundary layer. Dirmeyer and Shukla (1994) argued that although it is expected that a reduction in the surface roughness will lead to a decrease in the frictional convergence and hence total precipitation, it is possible that the increased thermally driven convergence can offset the frictional convergence. In this study we will investigate how reducing the surface roughness over urbanized regions will affect the precipitation distribution.

5.2 Methods of Analysis

The simulations in this section are done by altering one of the land surface characteristics that we propose to dominate the changes in the precipitation patterns and keep all other characteristics the same with the urban control simulation. Table 5.1 outlines the experiments that were carried out. The surface roughness and the albedo values are altered in the VEGPAM table in WRFv3.7. The soil moisture variables are changed in the land surface scheme. The lake is switched off by converting the land surface to a more dry land cover. In this case we convert the inland water body to shrubland/grasslands. This land surface cover was used as it has lower soil moisture, lower leaf area index and deep rooting zone and hence the evapotranspiration is lower in this case. We conduct the sensitivity tests for the OND season in 2007 since this is the year with the least errors for both models.

After conducting the experiments we rank the process with the most dominant impact by calculating the differences in the precipitation amount over the urbanized regions. We calculate the differences of the sensitivity simulation from the urbanized case.

5.3 Results and Discussion

We present our results in 2 sections. The first section discusses the differences in precipitation of the sensitivity simulation with the lake included and the second section discusses on the sensitivity simulations, when the lake is converted to vegetated regions.
Table 5.1 Summary of the experiments done to isolate the dominant processes.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Control Simulation</td>
</tr>
<tr>
<td>B</td>
<td>Urbanized all parameters are kept the same</td>
</tr>
<tr>
<td>C</td>
<td>Urbanization albedo is converted to Vegetation albedo (croplands albedo 0.23)</td>
</tr>
<tr>
<td>D</td>
<td>Urbanization roughness is converted to the Vegetated roughness (cropland albedo 0.05)</td>
</tr>
<tr>
<td>E</td>
<td>Urbanized soil moisture kept to vegetated land moisture by converting the soil properties</td>
</tr>
<tr>
<td>F</td>
<td>Same as C with the lake converted to vegetated land</td>
</tr>
<tr>
<td>G</td>
<td>Same as D with the lake converted to vegetated land</td>
</tr>
<tr>
<td>H</td>
<td>Same as E with the lake converted to vegetated land</td>
</tr>
</tbody>
</table>

5.3.1 Sensitivity Simulations with Lake Victoria Included

Figure 5.1 is the spatial distribution of the precipitation differences between the urban control simulation and the sensitivity simulations. In this case we investigate the changes in precipitation by altering either the albedo, surface roughness and the soil moisture and not altering the lake. The changes in the different surface properties lead to different changes in the precipitation. Changes in albedo have the highest changes in the total precipitation. The surface roughness produces the least changes in precipitation. For this case the ranking in the changes in precipitation is (i) surface albedo (ii) soil moisture (iii) surface roughness. We discuss how each of the sensitivity could lead to the changes in the precipitation in the following paragraphs.

Changing of minimum albedo and maximum albedo leads to reduced precipitation (Fig. 5.1a). This is consistent with past studies that have shown that when surface albedo is increased total precipitation reduces (Charney, 1975; Doughty et al., 2012). Figure 5.2 shows the average diurnal cycle of temperature over the urbanized region for the control urbanized simulation and the simulation with altered albedo. During the day the temperature differences are insignificant, but at night the temperatures over regions that have higher albedo are lower. This is because the surface does not absorb as much shortwave radiation as the surfaces with lower albedo, and loose the heat faster at night.
The lower temperatures leads to a weaker urban heat island at night, reduced convergence and hence a reduction in the total precipitation on the western side of Lake Victoria.

Compared to the albedo, surface roughness and soil moisture responses are lower. The soil moisture was changed to that of the grasslands/shrublands, which is slightly higher than that of the urban regions. The minimal increases in soil moisture lead to minimal changes in the total precipitation over the urbanized regions. Only a few regions experience an increase in the total precipitation, the increase could be due to increased evaporation.

Figure 5.1 Precipitation difference between the urbanized control and the sensitivity simulations with Lake Victoria included (a) albedo (b) soil moisture (c) surface roughness.
Doughty et al. (2012) studied the impact of reducing the surface roughness over the Amazon basin and found that, reducing the surface roughness can lead to increased precipitation. They propose that the reduced surface roughness can lead to an increase in the surface temperature and hence the total precipitation. However a study done by Sud and Smith (1985) showed that decreasing surface roughness reduces convergence and hence the total precipitation in a region. In our simulation decreasing the surface roughness over the urbanized regions leads to slight changes in the total precipitation over the region. In this sensitivity simulation changing the surface roughness while keeping all other properties constant does not lead to a decrease in precipitation.

5.3.2 Sensitivity Simulation with Lake Victoria converted to a Vegetated Land Type

Figure 5.3 shows the spatial distribution of the 3 different experiments conducted when the lake is converted to grassland/shrubland. Around the lake the rainfall decreases in all the simulations. Thus converting the lake to a different land cover does not affect the precipitation just over the lake but also on the Lake Basin. A comparison on the east and
the west side of the lake shows that the western side of the basin experiences a higher decrease in the total precipitation. We propose that the reduction is mainly due to the interaction of the background flow, lake breezes and mountain breezes. In the case of the control simulation the land lake breezes are enhanced by the easterlies that dominate in the region. When the lake is convert to another land cover the moisture supply for orographic lifting is reduced. This leads to drier conditions over the mountain region.

Figure 5.3 Precipitation difference between the urbanized control and the sensitivity simulations with Lake Victoria is converted to vegetated land (a) albedo (b) soil moisture (c) surface roughness.
The circulation over the region is modified (Fig. 5.4) when the lake is converted to a vegetated land cover. Without the Lake over the region there is no wind reversal due to the land-lake breeze. The southeasterlies dominate the flow during the day over the lake and over the urbanized regions. For the case where the lake is included there is reversal in the flow. The circulation over the mountain regions do not change much, however the moisture content is significantly reduced over this region (Fig. 5.5). At night the surface flow over the lake region is weaker (Fig. 5.6) for the experiments with the increase and altered soil moisture, while the one with modified soil moisture has almost the same magnitude. In addition to the weaker flow the moisture content over the lake region and mountain ranges is low. During the night the relative humidity (Fig. 5.7) is lower than during the day. Thus the changes in the land properties leads to modified circulation, moisture availability and hence a reduction in the total precipitation over the region. The highest impact is experienced during the nocturnal hours were there is a significant reduction in moisture availability.

In terms of the changes in total precipitation, the increased albedo case has the highest changes in the precipitation patterns, and the case with increase land soil moisture has the lowest reduced precipitation. The high soil moisture case has higher evaporation, compared to the case where the soil moisture is for urbanized regions. Thus the reduction in the total precipitation is lower. As discussed in the previous section, increased albedo leads to a reduction in the total precipitation. In this case the reduced moisture supply and high albedo dramatically reduces the total precipitation over the urbanized region and beyond. When the surface roughness is reduced and there is no moisture supply the total precipitation is reduced. In these experiments we can conclude that without the moisture supply from the lake the urbanized region experiences a decrease in the total precipitation. The ranking for this set of experiments is (i) albedo (ii) surface roughness (iii) soil moisture.
Figure 5.4 Circulation over the region during at 15UTC over the region from the different sensitivity simulations. (a) urban (b) albedo (c) soil moisture (d) surface roughness.
Figure 5.5 Relative humidity over the region at 15UTC for (a) urban with lake (b) albedo with the lake removed. The mountain ranges and the lake are between 90-120.
Figure 5.6 Circulation over the region during at 00UTC over the region from the different sensitivity simulations. (a) urban (b) albedo (c) soil moisture (d) surface roughness.
Figure 5.7 Relative humidity over the region at 00UTC for (a) urban with lake (b) albedo with the lake removed. The mountain ranges and the lake are between 90-120.
5.3.3 Conclusions

The studies conducted are idealized studies that can help in future urban planning. In this study we have shown that the most dominant process in producing the precipitation distribution over the region is the moisture supply from the lake. All the simulations that where conducted without the moisture supply from the lake had a reduction in precipitation of up to 5 mm/day. Thus the presence of Lake Victoria in close proximity to the urbanized region increases the thermally, and orographic driven convergence over the lake basin by moisture supply. The relative humidity plots showed a substantial lower moisture content when the lake is removed. The reduction in the moisture content is more evident during the nocturnal hours.

For the simulations conducted with the lake included but land surface characteristics changes the change in surface albedo gives the highest changes in the precipitation. When the surface albedo is increased the thermal contrasts between the lake and the land is altered. This leads to reduced convection in the region and hence a reduction in the total precipitation. When surface roughness is increased it leads to minimum changes in the total precipitation. The roughness effect could be diminished because of the higher thermally induced convergence. Changing the soil moisture leads to a bifurcation of precipitation changes. The precipitation is some regions is reduced while in others it increases. When the moisture supply from Lake Victoria is reduced the region with increased soil moisture has the least changes in the reduced precipitation. The region still has high moisture content and hence evaporation is high.
Chapter 6

Impact on the Agricultural Sector

6.1 Introduction

As the population and urbanization extent increases over East Africa, the economic sectors that highly depend on rainfall will be greatly impacted, these sectors include agriculture, energy (hydroelectric power) and fisheries. Not only will agriculture be affected by the conversion of the mostly croplands to urbanized region, but the precipitation patterns will also be affected. Figure 6.1 shows the agricultural output for Africa and the total exports. East Africa has the highest number of countries that have high exports, with some countries having a total of over 50% of exports, thus it is the backbone of the total exports. Disruptions to the rainfall pattern would derail the economic developments in the region.

In the previous sections we were only interested in the impact of precipitation in terms of the seasonal/monthly rainfall for the region. In this section we focus more on high resolution temporal analysis of precipitation data. It is essential to analyze precipitation data at high frequency since crops need an amount of water at more regular intervals. Agricultural activities in East Africa are mostly rainfed with a very low percentage of agricultural activities depending on irrigation. The different crops thrive in different climatic conditions, thus what might be considered a maximum climatic threshold for one crop might not be the same threshold for another crop.

In determining the onset and cessation of that rainfall period a number of methods have been proposed, these range from exceedance of accumulated rainfall threshold over a certain period to the ones that consider accumulated anomalous rainfall. These methods
require that there are no consecutive long dry spells after the onset, otherwise it will be a false onset and crops might not be able to germinate. Camberlin and Okoola (2003) proposed a method based on the Principal Component Analysis (PCA) where the daily data is initially square root transformed before PCA is conducted. After obtaining the timeseries from the PCA they then calculated the cumulative series from the 1st of February. The minimum points are thus the onset and the maximum points are the cessation. A method similar to Camberlin and Okoola (2003), was proposed by Liebmann et al. (2007), however they did not use the PCA analysis. Their method uses the anomalous calculation, where the onset is the anomalous accumulation is positive relative to the previous day and the cessation is when the anomalous accumulation is maximum.
A number of studies have proposed different values as a rainy day. Segele and Lamb (2005) proposed using 3 thresholds based on their knowledge of the different regions they analyzed over Ethiopia, these values were 0.1\(mm\), 5\(mm\) and 10\(mm\). Other studies (Camberlin et al., 2009; Camberlin and Okoola, 2003) have used 1\(mm\) as a threshold for a rainy day. In our study we adopt the 1\(mm\) threshold as defined in Camberlin et al. (2009); Camberlin and Okoola (2003).

### 6.2 Methods of Analysis

Analysis will initially be done on the control simulation for both WRFv3.7 and RegCM4.4. In addition to analysis of the model output, we also use 2 observational satellite data, and these are TRMM and CMORPH to validate the performance of the model in reproducing the high frequency statistics for precipitation. Both these datasets have a resolution of 0.25 degrees. CMORPH data is provided at 3 hourly intervals, to get the daily data we accumulate the precipitation for each day. The following table (Table 6.1) is a summary of the statistics that are normally calculated for agricultural purposes. We use some of the statistics in our analysis.

In order to calculate the onset and cessation over the region we subdivide the temporal time period into 2, the short rains (ASONDJF) and the long rains (JFMAMJ). This temporal division closely follows the study done by Camberlin et al. (2009) over East Africa to determine the variability of the onset, cessation and intensity of precipitation in the region. We also use these two time periods in general to calculate the statistics in Table 6.1, however we will not display all the plots. For the onset and the cessation we use the method proposed by Liebmann et al. (2007). The method uses the anomalous calculation, where the onset occurs when the anomalous accumulation is positive relative to the previous day and the cessation is when the anomalous accumulation is maximum. The anomalous accumulation quantity is defined by the following

\[
A(day) = \sum_{n=1}^{day} [R(n) - \bar{R}]
\]  

where \(R(n)\) is the daily precipitation and \(\bar{R}\) is the climatological annual daily average.
Table 6.1 Intra-seasonal statistics normally calculated for agricultural purposes.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Definition</th>
<th>units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dry Days (DD)</td>
<td>Number of dry days in the season</td>
<td>days</td>
</tr>
<tr>
<td>5 or more days (5DD)</td>
<td>mean frequency of dry spell that lasts 5 or more days (overlap is allowed in calculation)</td>
<td>days</td>
</tr>
<tr>
<td>Mean length of dry day</td>
<td>average duration of dry days</td>
<td>days</td>
</tr>
<tr>
<td>Longest Dry</td>
<td>longest consecutive dry days</td>
<td>days</td>
</tr>
<tr>
<td>Wet Days (DD)</td>
<td>Number of wet days in the season</td>
<td>days</td>
</tr>
<tr>
<td>3 or more days (3WW)</td>
<td>mean frequency of wet spells that lasts 3 or more days (overlap is allowed in calculation)</td>
<td>days</td>
</tr>
<tr>
<td>Mean length of wet days</td>
<td>average duration of wet days</td>
<td>days</td>
</tr>
<tr>
<td>Longest Wet</td>
<td>longest consecutive wet days</td>
<td>days</td>
</tr>
<tr>
<td>90%threshold</td>
<td>percentage that a wet day exceeds the 90% threshold</td>
<td></td>
</tr>
<tr>
<td>Frequency of rainy days</td>
<td>percentage number of days, within the season when precipitation exceeds 1mm</td>
<td></td>
</tr>
</tbody>
</table>

It is advised that the calculation be started 10 days prior to the driest month and is summed over the year. The onset will be the minimum value. While the beginning of the cessation is the maximum value.

The 2 models output precipitation differently. RegCM outputs average precipitation every 3 hours in the surface files and every 6 hours in the atmospheric files. We first convert these values to daily totals by adding the precipitation every 3 hours and convert the units to mm/day. For the WRF model, precipitation is accumulated, and to get daily data we de-accumulate the precipitation. We investigate the changes in the agricultural metrics in the year 2007-2008 seasons.

6.3 Results and Discussion

We divide our results section into 2 sections, the first section discusses the different statistics for the control simulation and the second analyzes how the conversion of land cover will affect the daily precipitation locally and regionally. Thus we use the precipitation
output from the urbanization sensitivity experiments.

6.3.1 Control Simulation

We use the daily precipitation datasets described in the previous chapter, to evaluate the ability of the model to capture the onset and the cessation of both the Long Rains and the Short Rains. Figure 6.2 shows a sample plot for the onset and cessation over Uganda. The 2 models are able to reproduce the temporal anomalies of precipitation. The RegCM model has accumulated anomalies that are higher than that of TRMM and CMORPH. For most of the time the WRF simulation has the same magnitude of accumulated anomalies as the observational datasets. Both models are able to capture the onset of the rainfall season, which is around the 5\textsuperscript{th} of August. This marks the beginning of the short rain season, and the cessation occurs at the beginning of November. While the spread of the onset is low, the spread in the cessation is a lot higher with the WRF model having an earlier cessation during the short rains. Thus the length of the short rains season is a lot shorter in the WRF model than the observations. RegCM onset occurs much earlier than in the observations and the WRF model. The onset occurs in late December, but for the observations this occurs at the beginning of February. The cessation of the long rains occurs during the peak of the long rains, in April. The models are able to capture the onset of the following season.

6.3.1.1 Statistics for Dry and Wet Days During the Long Rains

Figure 6.3 is the number of longest dry days period over the region from the models and the observations. The semi-deserts in the Somalia and Kenya have the highest number of dry days (greater than 60 days) compared to all other regions. Thus the regions do not have precipitation for approximately a third of the time during the Long Rains season. The Congo region and Lake Victoria have the lowest number of days without precipitation. The regions receive rains more than 90\% of the time.

Spatially the observational datasets are similar except over the southern part of the domain. In TRMM the number of dry days are lower as compared to the CMORPH datasets. Both models are able to capture the number of dry days in the model. The WRF model is able to capture dry patterns in most of the regions however it overestimates the
Figure 6.2 Onset and cessation plot for the Long Rains season and the Short Rains season.

dry days in the Congo region, and underestimates along the Indian Ocean coastal region. RegCM produces less days without precipitation. The MIT-Emanuel scheme produces excessive precipitation once it is activated and hence all the light events will be intense.
Figure 6.3 Consecutive dry days in the season from the observational and the simulated datasets during the Long Rains.
We also plot the number of wet days during the season. In this case instead of showing the total number of wet days we show the number of consecutive 5 wet days in the seasons. The calculation allows for an overlap from the previous five days to the next. The pattern is not surprising and follows what we expected that the Congo rainforest will have the highest number of 5 day periods that receive precipitation in a row.

The observational datasets have almost the same pattern, in most of the regions (Fig. 6.4). The TRMM dataset has lower number of 5 days periods with continuous precipitation compared to the CMORPH datasets. Discrepancies over the Congo region can be due to the unavailability of station datasets to bias correct the satellite datasets. Both models overestimate the frequency of 5 day rainfall accumulations over the Congo Region and also the southern part of the domain. The RegCM model over-predicts the 5 day accumulation statistic, with the highest over-prediction over the Indian Ocean and also the coastal regions. In a study done by Gianotti et al. (2012) over the maritime continent they found that RegCM model has the following deficiencies: underestimates the frequency of dry periods and frequency of high intensity rainfall, in addition it over-estimates frequency of low-intensity rainfall.
Figure 6.4 Consecutive 5 wet day periods in the season from the observational and the simulated datasets during the Long Rains.
In addition to the dry and wet days frequency we also plot the maximum daily precipitation during the Long Rains. A number of regions receive extremely high precipitation in some days, with some areas receiving over 100\textit{mm} precipitation per day (Fig. 6.5). The 2 observational datasets have almost the same pattern, with the minor differences over Lake Victoria, where the maximum precipitation over the Western part of the lake differ. This region has high cloud cover and hence the retrieval algorithms could be inadequate for the region. WRF model is able to capture the spatial pattern of the highest rainfall days in most of the regions, but in general it underestimates the values. The underestimation occurs over the Kenya, Lake Victoria and also the coastal regions. The RegCM model on the other hand has a number of regions that have light precipitation compared to the observations. However comparing the model with the observations shows the same spatial pattern.

In general during the Long rains season the model is able to reproduce the statistic that we investigated. With the WRF model producing the statistics that are close to the observations as compared to the RegCM. The RegCM model especially had over the Kenya/Ethiopian region, where it underestimated the number of dry days.
Figure 6.5 Maximum daily precipitation in the season from the observational and the simulated datasets during the Long Rains.
6.3.1.2 Statistics for Dry and Wet Days During the Short Rains

During the Short Rains season there are major differences in the observational datasets for the number of consecutive dry days (Fig. 6.6). The CMORPH dataset has longer drier periods compared to the TRMM datasets. WRF has the same pattern with TRMM dataset with almost the same number of dry days over land. In a few regions the model underestimated the length of the dry days. The RegCM model just like in the Long Rains, has shorter dry days over the ocean and also the eastern part of the domain. Over the Congo region the models and observations have better agreement.

![Figure 6.6](image)

**Figure 6.6** Number of dry days in the season from the observational and the simulated datasets during the Short Rains.
Just like the Long Rains season the models overestimates the number of consecutive wet days during the short rains (Fig. 6.7). The Short Rains has less rains that have consecutive 5 days with rainfall.

![Image of rainfall patterns](image_url)

*Figure 6.7* Number of 5 consecutive wet days in the season from the observational and the simulated datasets during the Short Rains.

The spatial patterns for the maximum 1 day precipitation are similar for the models and the observations (Fig 6.8). However the intensities differ, the RegCM model has the lowest intensity. On the western side of the domain the models can capture the intensity of precipitation, better compared with the eastern side. Over Lake Victoria, RegCM model underestimates the intensity.
Figure 6.8 Maximum daily precipitation in the season from the observational and the simulated datasets during the Short Rains.
6.3.2 Sensitivity Simulations

The WRF simulation was able to reproduce the high frequency precipitation than RegCM. We proceed by calculating the sensitivity of urbanization using the WRF model. We also just show the statistics for the high urbanization scenario as the changes with the low urbanization scenario were low. We zoom into the region were changes in land cover occur. Figure 6.9 is a comparison of the precipitation for the 2 urbanization scenarios and the control simulation for WRF during the Short Rains season. The high urbanization scenario has more intense precipitation over the urbanized regions. Over Lake Victoria the precipitation decreases in the western side of the domain. Past studies have shown that urbanizing a region increases the intensity of the precipitation in a region. The low urbanization scenario has lower changes in the intensity of precipitation (not shown).

![Figure 6.9](image)

Figure 6.9 Maximum daily precipitation in the season for the control and sensitivity simulation during the Long Rains (a) Control (b) High urbanization scenario.

During the Long Rains the increase in the maximum precipitation is higher compared to the Short Rains season. The higher precipitation is more evident over the northern part of Lake Victoria Basin. Changes in the intensity of the rainfall can be attributed
to the interaction of the urbanized regions with the northward moving Inter-tropical convergence zone.

*Figure 6.10* Maximum Daily Precipitation in the season for the control and sensitivity simulation during the Long Rains (a) Control (b) High urbanization scenario.
During the Short Rains the consecutive number of dry days does not change due to an increase in urbanization. Thus this season will not be affected in terms of the changes in number of dry days. However the Long Rains season has different patterns. During this season increased dryness occurs over Kenya and the Indian Ocean. While over the Northern parts of Tanzania, the number of consecutive dry days reduces by about 20 days.

**Figure 6.11** Number of dry days in the season from the WRF control and the high urbanization scenario during the Short Rains.
Figure 6.12 Number of dry days in the season from the control and the high urbanization scenario during the Long Rains.
6.4 Conclusions

In this section we investigated the ability of the two Regional climate models to reproduce the high frequency statistics. We calculated the longest consecutive dry day, number of 5 consecutive wet days in a time period and also the highest 1 day precipitation in a season. WRF model was able to capture the number of consecutive dry days, however the model over estimated the number of 5 day wet periods. The model performs equally well in both seasons. RegCM on the other hand is unable to capture the number of consecutive dry days over the Indian Ocean coastline, often over-predicting the number. The over prediction does not mean the model has higher precipitation than the observation or WRF but however the model produces too many light rain days. This is also evident in the highest 1 day amount in the season. The RegCM has lower maximum precipitation values in both seasons.

We proceed to use the WRF model for the sensitivity analysis. During the Long Rains the maximum precipitation increases for the urbanized scenario. In addition there is a decrease in the number of dry days over the northern part of Tanzania and also an increase in dryness over the coastal regions. The Long rains have a higher increase in the total one day precipitation. The Short rain season however does not have changes in the total number of dry days.

In conclusion the increased urbanization leads to an increase in the wettest day for both the Long and the Short Rains due to the enhanced thermal convergence over the region. The Long Rains have higher changes in the wettest day in a season, and this could be due to the interaction of the Intertropical convergence zone with the local circulation. The number of dry days are longer during the Long Rains season, which might imply that the season becomes shorter. Thus during the MAM season there will be more intense precipitation, but for a shorter period of time.
Chapter 7

Impact on the Energy sector

7.1 Introduction

In East Africa there are three main sources of energy and these are hydroelectric power, methane and biomass burning and oil (International Energy Agency, 2014; Othieno and Awange, 2016). At the moment the current energy production levels do not meet the demand in the region, with some of the regions experiencing long hours of power shortages (International Energy Agency, 2014; WindForce, 2013). Figure 7.1 shows the average number of hours some countries in Sub-Sahara Africa spend without power, and East African countries are amongst the countries that have substantial hours without power (International Energy Agency, 2014). As the lake side population increases there will be a competition in use of the available water resources, that is, drinking water, hydroelectric power and irrigation. This poses a great threat to the viability of the hydroelectric power energy in the region. Hydroelectric power is the main source of electricity for the region, especially in Uganda (Othieno and Awange, 2016). Currently most industries depend on hydroelectric power for their day to day production (Othieno and Awange, 2016). In the past years the region has been exploring ways of taping their natural resources of energy, which includes wind and solar energy, with the focus being mostly on the viability of wind power energy (Deichmann et al., 2011; International Energy Agency, 2014; Karekezi and Kithyoma, 2002; WindForce, 2013).

A few areas in the region have already setup wind turbines, these include Ngong near Nairobi, Marsabit in Kenya and the 400 kW machine at Chunya Catholic Mission in Tanzania (Othieno and Awange, 2016; WindForce, 2013). In addition to the wind turbines a 310MW wind farm project is underway in Kenya, Lake Turkana region. After com-
Figure 7.1 Sales lost and number of hours different countries have without electricity. The red rectangle highlights the East African countries.

pletion it is expected that this will provide approximately 20% of the current installed electricity generating capacity (Kenya Vision 2030, 2015; WindForce, 2013). The project consists of building 365 wind turbines Vestas V52 of hub height 45m and nominal power 850kW. In Tanzania there is a project underway to construct a 100MW wind farm in Singinda (World Bank, 2015). Pallabazzer and Sebbit (1998) studied the possibility of setting up wind turbines in Uganda and recommended that the country did not have very good wind resources, thus wind resources in the region could be used for domestic and agricultural water pumping. However they also suggested that it is possible to setup wind farms in regions where there is complex terrain and could have high wind speeds. In addition to the lake Turkana region and the Singinda region this study seeks to explore if there are other regions that have high potential for wind resources in East Africa, and how urbanization might affect the viability of the wind resources in the future. Increased urbanization extent could affect the viability of wind resources in the region by either encroaching in regions that have high wind potential. In addition, urbanized regions have higher surface roughness and hence there will be a reduction in the wind speed. No observational datasets are currently available to validate the results, thus we use reanalysis data to evaluate, hence this study should be taken as a attempt to explore regions of
high wind resources in the region. The Atlas map constructed by the National Renewable Energy Laboratory (NREL) (http://www.nrel.gov/gis/data_wind.html) will be used to inter-compare if regional climate models can capture the main wind features in the region.

Kamau et al. (2010) used wind data from the Kenya Meteorological department for the period 2001 to 2006 to study the diurnal, monthly and inter-annual variability of wind using empirical methods. They found that the region had an average wind speed of greater than 11\(ms^{-1}\) at a height of 10\(m\). The power law was then used to find the wind speed at 50\(m\) and 100\(m\). At 100\(m\) the power density was found to be between 1776\(Wm^{-2}\) and 2202\(Wm^{-2}\). The Weibull scale parameter was found to be greater than 2, implying that the wind speed in the area is not very variable. In addition to the wind speed they also analyzed the wind direction and found that the wind direction throughout the year does not to change much. Kainkwa (2000) studied the wind resources at Basotu, Tanzania and found that the windy season, is from June to November, with October being the windiest month. The mean wind power during this season was at least 114\(Wm^{-2}\).

In order to identify regions that have good wind resources a wind power classification table has been constructed. We will use this classification to draw conclusion on the region that might have good wind power resources based on the model simulation.

**Table 7.1** Wind power classification.

<table>
<thead>
<tr>
<th>Wind Power Class</th>
<th>Wind Power Density</th>
<th>Resource Potential</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0-200</td>
<td>Poor</td>
</tr>
<tr>
<td>2</td>
<td>200-300</td>
<td>Marginal</td>
</tr>
<tr>
<td>3</td>
<td>300-400</td>
<td>Fair</td>
</tr>
<tr>
<td>4</td>
<td>400-500</td>
<td>Good</td>
</tr>
<tr>
<td>5</td>
<td>500-600</td>
<td>Excellent</td>
</tr>
<tr>
<td>6</td>
<td>600-800</td>
<td>Outstanding</td>
</tr>
<tr>
<td>7</td>
<td>&gt;800</td>
<td>Superb</td>
</tr>
</tbody>
</table>
7.2 Methods of Analysis and Data

To capture the variability of wind more accurately, simulations with finer grid resolution are essential. This section will use the WRF output since it has the finest resolution. In addition to analyzing the wet seasons we also analyze the dry seasons that is we analyze all the four seasons in East Africa, the January-February (JF) season, March-May (MAM), June-September (JJAS) and the October to December (OND) season at 6 hourly levels. Several statistics have been proposed and are used in analyzing the wind potential in a region, these include calculating the Weibull shape and scale parameters, average wind speed, and also the power density.

The wind variation is mostly summarized using the Weibull distribution.

\[
f(v) = \frac{k}{X} \left( \frac{v}{X} \right)^{k-1} \exp \left( - \left( \frac{v}{X} \right)^k \right)
\]  

(7.1)

\( v \) is the wind speed in \( ms^{-1} \), \( X \) is the Weibull scale parameter in \( ms^{-1} \), measure for the characteristic wind of the distribution, it is proportional to the mean wind speed. \( k \) is the Weibull form parameter and specifies the shape of the distribution. A small value for \( k \) signifies very variable winds, while constant winds are characterized by larger \( k \). The Weibull parameters will be calculated and displayed for each year to show the wind variability in the region. The wind power density (\( P \)) is defined as

\[
P = 0.5 \rho A v^3
\]

(7.2)

where \( \rho \) is the density of air and \( A \) is the area swept by the rotors. Albert Betz a German Physicist found that approximately 40% of energy is lost when energy is converted from kinetic to mechanical. That is, no wind turbine is able to convert all the kinetic energy to mechanical energy, this law is commonly referred to as the Betz Limit, and the theoretical maximum efficiency of a wind turbine is approximately 59%. Thus the wind power density is calculated as follows

\[
P = 0.5 C_p \rho A v^3
\]

(7.3)

where \( C_p \) is the power coefficient.
Figure 7.2 Wind turbine output with steady wind plot. The cut in speed is the speed at which the turbine will start generating power, and the cut out speed occurs when the turbine shutdown due to excessively high wind speed.

Figure 7.2 is a typical power curve for a wind turbine. The graph shows that turbines need a minimum amount of wind for generation of wind to start, this speed is what is commonly referred to as the cut in speed. The power between the cut in speed and the rated output power is calculated by equation 7.3.

We use the ERA Interim reanalysis (Dee et al., 2011) data produced by the European Centre for Medium-Range Weather Forecasts (ECMWF) to show the 10m wind patterns in the different seasons, at every 6 hourly intervals. The data has a spatial resolution of about 80km and is available every 6 hours.

7.3 Results

7.3.1 Inter-annual Variability of Wind Resources

The shape parameters are used to analyze the wind variability in the region. The variations of wind is essential to model as this will determine if a region will have a viable wind farm. The less the parameter is the more variable the wind speed is at the region,
the higher values indicate more stable winds in a region. Large scale phenomenon like the
ENSO or IOD can affect the wind resources in a particular year. In our 5 year analysis
there are different large scale phenomenon, the year 2010 was a negative IOD year and
also an La Niña year, while 2006 is a strong IOD and El Niño year the other years are
neutral years. Figure 7.3 is the shape parameters of the Weibull distribution at 10m for
yearly data. A larger shape parameter indicates regions with less variable wind speeds.
Like wise a small shape parameter indicates more variable wind speed and the wind are
mostly in the lower ranges. Over our domain we can conclude that in general the Congo
region has the least speed and also the wind are more variable. In Kenya and Tanzania
the shape parameters in some regions are relatively high, this implies that there are more
steady winds in those regions. There were no major difference in the Weibull parameters
from year to year, thus we conclude that the large scale phenomenon did not have a huge
impact on the shape parameters.
Figure 7.3 Shape parameters for the different years (a) 2006 (b) 2007 (c) 2008 (d) 2009 (e) 2010 from the WRF model.
7.3.2 Seasonal Changes in the Wind Resources

Wind resources in a region can vary from season to season due to large scale circulation features, in addition to the large scale flow the local atmospheric and geographic conditions influence the wind speed. We show the resulting wind speeds and associated wind power density in each season at 6 hourly time intervals for the lower levels and the upper levels in each season. In this study we consider the 10m, 30m, the 100m heights. These 3 levels are chosen so as to minimize interpolation errors. As mentioned before East Africa comprises of a number of large inland water bodies, has highly variable terrain and some countries are on the coastlines, these different land features affect the mesoscale circulation in the region, and hence the wind patterns.

7.3.3 January-February Season

This season is dominated by the northeast monsoons that originates from a semi-permanent high pressure system (Okoola, 1999). The average 10m wind speed for the entire domain during the season is approximately $2.6 ms^{-1}$ (Fig. 7.4). The daytime hours and the night hours show significant differences in the wind speeds, with the daytime having more regions with high wind speed compared to night time. Along the Indian Ocean and Lake Victoria coastlines the wind speed is generally high during the day compared to the nocturnal hours. The high wind speeds are due to the sea breezes, which are normally stronger than the land breezes. The Congo region has the lowest average 10m wind speed due to the vegetation cover in the region. The forests have high surface roughness and therefore reduce the wind speed in the region. The region with high wind speeds over Kenya from our model simulation coincides with the region with the largest wind farm in Sub-Saharan Africa. Thus this gives us confidence that WRF model simulates the wind speed in the region reasonably well. The high wind speeds over Kenya have been attributed to the interaction of the large scale flow with the local topography to create convergence zones (Indeje et al., 2001). The Turkana-Marsabit corridor occurs in between the Ethiopian and Kenyan highlands and it is known to be the one of the windiest regions. In Uganda the wind speed is high in the northern part of the country and low over the southern parts, with wind averages of less than $2.5 ms^{-1}$. 

191
Figure 7.4 January-February wind speed and wind power density spatial distribution at 10m (a) 00 UTC wind speed (b) 00 UTC power density (c) 06UTC wind speed, (d) 06UTC power density (e) 12 UTC wind speed (f) 12 UTC power density (g) 18UTC wind speed (h) 18UTC power density.
Figure 7.5 is the ERA-Interim average wind speed at 6-hourly intervals at a 10\(m\) height for the January to February season. Spatially the wind speed for WRF (Fig. 7.4) has the same pattern with the ERA-Interim (Fig. 7.5). However the wind speed from the WRF model is higher than the ERA interim. The differences in the wind speed could be due to the differences in model spatial resolution. The lower wind speeds in the reanalysis data could be due to the low resolution in the model. In addition East Africa has limited observational datasets, this could lead to underestimation of the wind speed. It has also been shown that low resolution simulations might be unable to capture mesoscale features, like low level jets and sea breezes.

![Figure 7.5 January-February wind speed at 10m from ERA-Interim](a) 00 UTC (b) 06 UTC (c) 12 UTC (d) 18 UTC.

According to the wind profile power law we expect that the wind speed at higher
levels is faster than at 10m. Clearly the 30m (Fig. 7.6) and the 100m (Fig. 7.7) hub heights have higher wind speeds, with the 100m level having the highest wind speed. At these heights regions in Uganda that had marginal potential for a wind farm now have high steady winds throughout the day. The diurnal variation of wind speed in the upper levels is consistent with the low levels.

The wind power density is proportional to the cube of the wind speed, thus we expect regions that have high wind speed to also have high wind power density. The right columns of Fig. 7.4, Fig. 7.6 to Fig. 7.7 represent the corresponding power density to the wind speed. At 10ms$^{-1}$ the Congo region and some parts of East Africa have wind power density of less than 20Wm$^{-2}$. Thus at this height most places in the Congo region and East Africa are not suitable for a wind power farm. At 30m there is improved power density values, with some of the Eastern parts of the domain having marginal wind power to excellent wind resources. At 100m hub height more areas have wind power density greater than 500Wm$^{-2}$ implying that the wind resources are excellent in the region.
Figure 7.6 Same as Figure 7.4 but for the 30m hub height level.
Figure 7.7 Same as Figure 7.4 but for the 100m hub height level.
Figure 7.8 is the probability distribution function of the average wind speed at three hourly intervals for the 3 levels. During most of the times, the wind speed at 10m with the highest frequency is between 1\( m/s \) to 1.5\( m/s \). From the graph we can deduce that 18UTC has relatively lower wind speed and 09 UTC had the highest frequency of obtaining winds speed greater than 2.5\( m/s \). At 30m hub height the frequency of getting wind speeds with higher than 2.5\( m/s \) increases. For the 100m hub height the frequency of higher wind speeds increases, and there is a higher percentage of frequency of non calm wind. Likewise the cumulative probability distribution function shows that the higher hub height has a high frequency of high wind speeds. For the 10m winds the cumulative probability distribution function show that over the region the probability of getting wind speeds of greater than 2.5\( m/s \) the cut in speed for some wind turbines is approximated 50%. This number might seem low but this is because the region is mostly dominated by the Congo region which has low average wind speeds. At 30m hub height the percentage of winds that have a wind speed 2.5\( m/s \) or higher increases to over 75%. That it is most ideal to set up wind turbines at this height as compared to the 10m level if a small turbine is desired. The 100m hub height shows approximately 85% of obtaining wind speeds greater than the cut in speed. Thus during the JF season the region has good mean average wind speed especially when higher hub heights are considered.
Figure 7.8 January-February distribution functions wind speed at 10m, 30m and 100m vertical level (a) Probability distribution function at 10m (b) Cumulative probability distribution function at 10m (c) Probability distribution function at 30m (d) Cumulative probability distribution function at 30m (e) Probability distribution function at 100m (f) Cumulative probability distribution function at 100m.
7.3.4 March-May Season

The season has the same pattern as the JF season, however the wind speed are lower during this season over the coastal regions, and the southern part of the domain has high wind speed (Fig. 7.9) as compared to the JF season (Fig. 7.4). During this season we expect the Inter-Tropical Convergence Zone (ITCZ), to be moving northwards and hence influence the wind patterns in the Northern part of Tanzania. In addition to the ITCZ the moisture content in the atmosphere increases and hence wind speed reduces, upon reaching the land surface, they dump out moisture and hence wind speed increases over land. For the coastal regions the low wind speed, could be due to a combination of the moist winds and also the high cloud cover. High cloud cover will reduce the amount of insolation that reaches the surface. This reduction reduces the thermal contrasts between the coastal regions and hence reduces the intensity of the sea breezes. Just like the JF season the Congo region has the least wind speeds due to the vegetation cover. In Uganda the region with high wind speed is reduced as compared to the JF season. The high wind speeds are predominantly confined to the Northwestern side of the country. In Kenya the wind speed is high along the Turkana-Marsabit channel the wind speed is high.
Figure 7.9 March-May wind speed and wind power density spatial distribution at 10m (a) 00 UTC wind speed (b) 00 UTC power density (c) 06 UTC wind speed, (d) 06 UTC power density (e) 12 UTC wind speed (f) 12 UTC power density (g) 18 UTC wind speed (h) 18 UTC power density.
Just as in the JF season the MAM WRF average wind speed have the same spatial
distribution, as the ERA-Interim. The model is able to reproduce the diurnal variation
of the wind speeds in the region. As mentioned before the low resolution in the ERA-
Interim might be missing some mesoscale features hence the wind speed is lower when
compared to the WRF model. Thus during the MAM season the WRF model is also able
to reproduce the average wind patterns over the region.

Figure 7.10 March-May wind speed at 10m from ERA-Interim (a) 00 UTC (b) 06UTC (c)
12UTC (d) 18UTC.

Figure 7.11 and Fig. 7.17) are the 30m and 100m height plots. At the 30m hub height
(Fig. 7.11) the average wind speed is greater than 2.7\text{ms}^{-1}, and most of the regions have
good wind resources including some parts of the Congo rainforest which have average
wind speeds greater than $4ms^{-1}$. At the 100$m$ level all the 3 countries have regions that have a fairly high probability of wind power harnessing. The wind speed at this level are mostly over $5ms^{-1}$ for most times of the day. The maximum wind speed at this level is about $19ms^{-1}$ with an average wind speed of over $3ms^{-1}$ at all time levels. The wind power density plots correspond to the wind speed patterns, with maximum power occurring over some parts of Kenya, Tanzania, and northern part of Uganda. During this season there are a number of regions in all 3 East Africa countries, that have fair to excellent wind resources.
Figure 7.11 Same as Figure 7.9 but for the 30m hub height level.
Figure 7.12 Same as Figure 7.9 but for the 100m hub height level.
The probability distribution function of wind at 10m show that the frequency of high wind speed occurs at 09UTC and at 21UTC. There is a high probability of getting more calm conditions (Fig. 7.13(a)). During this season there is a probability of about 40 to 50% of getting wind speeds greater than 2.5ms\(^{-1}\) (Fig. 7.13(b)). This probability is a lower compared to the JF season. Thus at the 10m level the wind speed during the MAM season is lower compared to the JF season.

At 30m the probability of getting higher wind speeds increases (Fig. 7.13(c)), with high wind speeds at most of the times. At this level the 21UTC has the lowest frequency of low wind speeds. The cumulative probability distribution function, shows that there is a high probability of getting higher wind speeds (Fig. 7.13(d)). The probabilities in this case, are all high, with probabilities of getting high wind speeds being greater than 70%.

The 100m hub height shows higher wind frequencies compared to all the other levels (Fig. 7.13(e)). The peaks of non calm winds is higher than all the other levels. The cumulative distribution function (Fig. 7.13(f)) has over 80% of the period of time when the winds are greater than 2.5ms\(^{-1}\). At this level there is a high probability that there will be generation of power most of the times.
Figure 7.13 March-May distribution functions wind speed at 10m, 30m and 100m vertical level (a) Probability distribution function at 10m (b) Cumulative probability distribution function at 10m (c) Probability distribution function at 30m (d) Cumulative probability distribution function at 30m (e) Probability distribution probability distribution function at 100m (f) Cumulative probability distribution function at 100m for the WRF model.
7.3.5 June-September Season

This season has the highest wind speeds over the eastern side of East Africa. During this season the southeast monsoonal flow dominates over the eastern part of our domain (Okoola, 1999). The JJAS season is characterized by the strengthening of the Somali Jet, which increases the wind speed in the region. In addition to the strengthening of the Somali Jet the sea breezes are also strong during this period. The region generally receives less rainfall during this period, thus less cloud cover and more intense insolation reaching the surface. The western side as expected has low wind speeds due to the forests that have high surface roughness.

The 10m wind speeds (Fig.7.14) are generally high in the Eastern side of the domain compared to the western side of the domain, with the wind speeds over the Congo region being a lot much lower compared to the MAM season. Although during this season most regions in the eastern side of the domain have good wind resources, the wind speed over Uganda is low. This could be due to the effect of the position of the ITCZ.
Figure 7.14 June-September wind speed and wind power density spatial distribution at 10m (a) 00 UTC wind speed (b) 00 UTC power density (c) 06UTC wind speed, (d) 06UTC power density (e) 12 UTC wind speed (f) 12 UTC power density (g) 18UTC wind speed (h) 18UTC power density.
The WRF model (Fig. 7.14) has high wind speeds over the ocean, Somalia, parts of Kenya, Tanzania and Ethiopia. This is consistent with the reanalysis data spatial distribution (Fig. 7.15). Just like the other seasons mentioned before the model can capture the diurnal cycle of the wind speed.

Figure 7.15 June-September wind speed at 10m from ERA-Interim (a) 00 UTC (b) 06UTC (c) 12UTC (d) 18UTC.

The wind speed over Uganda are surprisingly low for all the levels and during most of the times. Even at a 100m hub height (Fig.7.17) the only time when the wind speed is relatively high over the Southern part of the country is at approximately 18UTC. Only the north western part of the domain has high wind speed. The power density for the southern part of Uganda is less than 20Wm$^{-2}$, which is very low for any viable wind
power farms, but the North western part has wind power density that is over $200 W m^{-2}$. In addition to this region, most of Kenya and Tanzania has wind power density that is greater than $200 W m^{-2}$ with a maximum wind power density of $3210 W m^{-2}$. 
Figure 7.16 Same as Figure 7.14 but for the 30m hub height level.
Figure 7.17 Same as Figure 7.14 but for the 100m hub height level.
The probability distribution function for this season has 2 peaks, one occurring at approximately $1.5ms^{-1}$ and the other one occurring at $8ms^{-1}$. This possibly shows that the east side and the west side of the domain have different wind regimes, with the western side of the domain having relatively low wind speed compared to the eastern side. The cumulative probability distribution function shows that during this season there is a probability of over 60% that wind turbines will generate power.

The wind speed increase with high altitude, with the $10m$ height having the lowest wind speed and the $100m$ having the highest wind speed. At $100m$ there are some regions in the north eastern side of the domain that shows high wind speeds.
Figure 7.18 JJAS distribution functions wind speed at 10m, 30m and 100m vertical level
(a) Probability distribution function at 10m (b) Cumulative probability distribution function at 10m (c) Probability distribution function at 30m (d) Cumulative probability distribution function at 30m (e) Probability distribution function at 100m (f) Cumulative probability distribution function at 100m.
7.3.6 October-December Season

During this season the southeast monsoonal flow reverses and flow is northeasterly. As stated before the northeasterly flow is weaker compared to the southeasterly flow. Thus this season generally has low wind speeds in comparison to the preceding JJAS season. During the OND season there is high cloud cover compared to the dry JJAS season. The high cloud cover weakens all thermally induced mesoscale circulations and hence a reduction in the overall wind speed just like the MAM season. For this season the wind speed are generally lower over the coastal regions however they are relatively high in the northern part of Tanzania, Lake Turkana region and also the northwest part of Uganda (Fig. 7.19).
Figure 7.19 October-December wind speed and wind power density spatial distribution at
10m (a) 00 UTC wind speed (b) 00 UTC power density (c) 06UTC wind speed, (d) 06UTC
power density (e) 12 UTC wind speed (f) 12 UTC power density (g) 18UTC wind speed (h)
18UTC power density for WRF model.
The 10\text{m} OND WRF (Fig. 7.19) speed has the same spatial distribution with that of the ERA-Interim (Fig. 7.20). Just like in the ERA-Interim, the wind speed in WRF is also high between 06UTC to 12UTC. The major differences in the model and reanalysis data is in the magnitude of the wind speed as stated for the other seasons. Thus for all the seasons we can conclude that the model is able to reproduce the spatial patterns of the wind speed in the region.

Figure 7.20 October-December wind speed at 10m from ERA-Interim (a) 00 UTC (b) 06UTC (c) 12UTC (d) 18UTC.
Figure 7.21 Same as Figure 7.19 but for the 30m hub height level.
Figure 7.22 Same as Figure 7.19 but for the 100m hub height level.
Figure 7.23 October-December distribution functions wind speed at 10m, 30m and 100m vertical level (a) Probability distribution function at 10m (b) Cumulative probability distribution function at 10m (c) Probability distribution function at 30m (d) Cumulative probability distribution function at 30m (e) Probability distribution function at 100m (f) Cumulative probability distribution function at 100m for WRF model.
7.3.7 Sensitivity to Changes in Land Cover

Land cover changes alters the mesoscale and regional circulation and thus this might pose as a challenge to the viability of wind power. We analyze the changes in the impact on wind speed change at both the high low urbanization scenarios. We inter-compare the 30m height winds in the control simulation with those in the urbanized scenario.

7.3.7.1 JF High Urbanization Scenario

Changes in the wind speed are normally localized (Fig. 7.24). The lowest changes in the wind speed occurs in the afternoon where the region experiences generally lower wind speeds. At about 06UTC the wind speed over the southern part of the urbanized regions increases, while over the urbanized region it decreases. Over the urbanized regions the wind speed slightly reduces at night in the Northern parts and the western part of Lake Victoria Basin. A reduction of wind speed over these regions could be attributed to the changes in surface roughness, which will be dramatically increased from a value of 0.05 to a value of 0.5. In addition to the regions that are urbanized the Northern parts of Tanzania experience an increase in the average wind speed. The increases mostly occur during the nocturnal hours. At about 18UTC there is an average increase in the wind speed over the region with just a few regions having reduced wind speeds.
Figure 7.24 January-February wind speed and wind power density difference plots (a) 00 UTC wind speed (b) 00 UTC power density (c) 06UTC wind speed, (d) 06UTC power density (e) 12 UTC wind speed (f) 12 UTC power density (g) 18UTC wind speed (h) 18UTC power density for WRF model using the high urbanization scenario.
7.3.7.2 MAM High Urbanization Scenario

During the MAM season the changes in winds speed are mostly localized (Fig. 7.25). The wind power density for the MAM season does not change much. Over Lake Victoria only slight changes in the wind speed are experienced, during the nocturnal hours there is a dipole in changes in the wind speed. The Northern parts have reduced wind speeds, while the southern parts have higher wind speeds. The Power density only changes slightly and these changes should not affect the viability of wind power during this season.
Figure 7.25 March-May wind speed and wind power density spatial distribution at 30m wind land cover changes (a) 00 UTC wind speed (b) 00 UTC power density (c) 06UTC wind speed, (d) 06UTC power density (e) 12 UTC wind speed (f) 12 UTC power density (g) 18UTC wind speed (h) 18UTC power density for WRF model.
7.3.7.3 JJAS High Urbanization Scenario

During the JJAS season urbanization has a more intriguing effect on the wind pattern (Fig. 7.26) in comparison to the MAM season (Fig. 7.25). While the southern region of Uganda experiences a reduction in the wind speed at about 00UTC and 18UTC there is an increase in the average wind speed at about 06UTC. This could be due to land breezes shifting time when they are most dominant in the region. During the day the wind speeds in the urbanized are generally higher than in the non-urbanized regions. This is contrary of the MAM season where there is a reduction in the wind speed throughout the day. We propose that the south-easterly monsoonal winds are enhanced by the sea breezes during this season and hence lead to higher wind speeds. Over the Congo region the wind speeds are slightly elevated during the day and lower in the late evening. The changes in the power density are not statistically significant and will not alter the wind resources in the region.
Figure 7.26 June-September wind speed and wind power density differences between the high urbanization scenario and the control at (a) 00 UTC wind speed (b) 00 UTC power density (c) 06 UTC wind speed, (d) 06 UTC power density (e) 12 UTC wind speed (f) 12 UTC power density (g) 18 UTC wind speed (h) 18 UTC power density for WRF model for the 30m height.
7.3.7.4 OND High Urbanization Scenario

During the OND season there are slight changes in the wind speed over the urbanized regions and the regions close to these areas (Fig. 7.27). The northern part of lake Victoria has a reduction in the wind speed during the nocturnal hours. With a gradual increase in the wind speed during the day, which reaches a maximum in the late afternoon hours. The changes in the wind speed however do not affect the changes in the wind power density much with the highest changes being less than 20$W m^{-2}$. Thus during this season we do not expect urbanization to have an impact on the wind power viability.
Figure 7.27 October-December difference plot for the wind speed and the wind density at 30m wind land cover changes (a) 00 UTC wind speed (b) 00 UTC power density (c) 06UTC wind speed, (d) 06UTC power density (e) 12 UTC wind speed (f) 12 UTC power density (g) 18UTC wind speed (h) 18UTC power density for WRF model for the control and high urbanization scenario.
7.3.8 Changes in the Probability Distributions

Figure 7.28 shows the changes in the probability distribution of wind at the 30\textit{m} height in the different seasons. In all the other season the changes in the wind speed is minimum, only a few regions have either an increase or a decrease in the wind speed. The impact does not seem to change with the time period, with all the 3 hourly time periods having the same shape.

We do not show the difference plots for the low urbanization scenario as the changes are very minimum, however we show the changes in the PDF (Fig. 7.29). The low urbanization scenario shows lower changes in the wind speed as compared to the high urbanization scenario. The probability of obtaining no change in the wind speed is higher for this case. Thus the lower the region that is urbanized the lower the changes in the wind speed.
Figure 7.28 Changes in the probability distribution function at 30m hub height for the different seasons in the WRF model (a) JF (b) MAM (c) JJAS (d) OND in the low urbanization scenario.
Figure 7.29 Changes in the probability distribution function at 30m hub height for the different seasons in the WRF model (a) JF (b) MAM (c) JJAS (d) OND in the high urbanization scenario.
7.4 Conclusions

In this Chapter we have assessed the wind potential in the region and how future urbanization scenarios might impact the wind power energy potential. We conclude that the regional climate model was able to identify the regions that have been obtained in previous studies to have high wind power resources, and WRF model is able to capture the spatial variability of the wind speed, when compared to ERA-Interim reanalysis data. In addition to these areas the north eastern part of Uganda also has a potential for good wind resources. In addition to the coastal regions identified by Mukasa et al. (2013) we find that there is a wider area in Kenya, Somalia and Tanzania that can be used for wind resources. The movement of the ITCZ has a great impact on the northern part of Tanzania and the Southern parts of Uganda. During the JJAS the ITCZ is northwards and leads to lower wind speeds over Uganda. During the JF season, the ITCZ has moved southwards and wind speeds over northern parts of Tanzania are low.

It must be noted however that due to the lack of observational datasets the patterns are not validated with observational datasets, but rather with the reanalysis data which also highly depend on the density of the insitu datasets. This study has shown that wind speed in the region is seasonal, higher wind speeds occur during the dry season when the South easterly monsoon is more dominant, and weaker when the North easterly monsoon is dominant. The Turkana channel has the highest wind speed due to the complex topography that leads to channeling. In the future it will be essential to collect observational datasets at a number of regions that have been identified to have high wind potential.

Urbanized regions do not have a high impact in wind speed, with the average change in wind speed being $\pm 1ms^{-1}$. The changes in the high urbanization scenario experiences slightly higher changes in wind speed in comparison with the low urbanization scenario. We propose that the changes in wind speed around the lake are mostly affected by the changes in the land lake breeze. The intense surface heating causes the modification of the land-lake breeze. The lowest changes in the wind speed occur at 12UTC. During the nocturnal hours the land breeze is reduced. This is because the urbanized regions loose the heat slower at night, thus the thermal contrasts between the land and the lake are reduced. and hence there is a reduction in the wind speeds over the region. At 18UTC
the wind speeds increases slightly along the edges of urbanization.

In conclusion, WRF model is able to reproduce the wind climatology over the region. Increased urbanization around the lake does not seem to impact urbanization as the increases will be remote from regions of high wind potential.
Chapter 8

Conclusions and Recommendations

8.1 Conclusions

The aim of this study was to investigate the impact of urbanization on the climate of East Africa. The study was divided into 4 major sections that is the (i) customization stage, (ii) evaluation of the ability of a regional climate model to reproduce the climate variability over the region, (iii) the impacts of urbanization and the processes that are dominant in reproducing the patterns (iv) applications of how rapid urbanization will affect the agricultural and energy sectors.

It is essential to first find the optimal physics combination over a region that will minimize the model bias before using the model for sensitivity studies. The customization stage was done for both RegCM4.4 and WRFv3.7 model, during the Short Rains season, with more customization experiments done for the RegCM model. In both models customization was done based on previous customization studies over the region. The last comprehensive customization for RegCM over the region was done with the previous version of the model, and a number of new physics schemes have been incorporated in the model. The RegCM model is customized using 2 resolutions the 36 km and the 25 km. Spatial comparison of the 36 km and 25 km resolutions simulations revealed the usefulness of using high resolution. Total precipitation bias is reduced when using the higher resolution. Tuning of parameters in the SUBEX scheme reduces total wet bias when the MIT Emanuel scheme is utilized, and reasonably reproduces the total precipitation in the region. This study has also revealed that in addition to the cumulus scheme the microphysics and the land surface scheme play a major role in reproducing total precipitation over the region. When the Grell scheme and the Nogherroto/Tompkins physics schemes
are used, the model has a dry bias. However when the MIT-Emanuel scheme is used with the Nogherroto/Tompkins scheme the wet bias in the scheme is reduced. The reduction occurs without tuning the parameters in the microphysics scheme. Thus in the future the MIT-Emanuel scheme can be used with the Nogherroto/Tompkins scheme. The use of the Community Land Model surface scheme also reduced the wet bias in the MIT-Emanuel scheme without tuning the SUBEX scheme. The CLM scheme is a more advanced scheme in comparison with the BATS1e and it has the ability to model different land cover in one grid cell. Thus it is suitable in East Africa which has heterogeneous land cover. A lot of work still needs to be done to find the optimal parameters over the region, but the current results show an improvement in total precipitation with the MIT-Emanuel, Nogherroto/Tompkins and Community Land Model physics scheme combination.

In order to ascertain that the customization stage are just not adequate for a single year and can reproduce the inter-annual variability of precipitation, we perform the Empirical Orthogonal Function (EOF) analysis for the RegCM model. In this case we choose the MIT-Emanuel scheme which can partition the large scale and convective precipitation more accurately than all the other schemes in the model, and the simplistic Grell FC scheme which had the least precipitation bias. We adjust the SUBEX parameters as suggested in Davis et al. (2009) for the MIT-Emanuel scheme. When the EOF analysis is applied to the 2 different simulations we conclude that the model is able to reproduce the most dominant modes of variability as the observational Tropical Rainfall Measuring Mission data. This is especially so during the Short Rains season where the correlation of the model time series and that for TRMM is approximately 0.9. The Lake Victoria Basin domain timeseries has a lower correlation with observations than the East Africa domain. We propose that the model might not be able to capture the mesoscale circulations in the region correctly, since the land lake breezes mostly dominate this region. In our simulations we employ a 1-dimensional lake model which is unable to capture the full lake dynamics and hence increases biases over the region. In addition to insufficient treatment of the lake, the EOF method could be a source of reduced accuracy since it depends on a covariance matrix, which might not be able to split the temporal and spatial modes more accurately as in the larger East Africa domain. During the MAM season the model has less skill, this could be due to inadequate customization for the season and also inadequate understanding of the climate drivers during this season. For the Lake
The Victoria Basin region the model and TRMM first EOF timeseries was significantly correlated with the Indian Ocean Dipole index. Composite analysis of the circulation, zonal flow, vertical velocity and the sea surface temperatures was also done. All the model composites showed that the model temporal evolution is consistent with the observed conditions. We conclude that the RegCM model is able to reproduce the inter-annual variability of precipitation in the region.

To investigate the impact of urbanization on the climate, we use the urbanization land cover projections based on Seto et al. (2012) studies. The rapid urbanization in the region will impact the climate over Lake Victoria Basin. The high urbanization scenario had higher increases in the temperature, energy fluxes and also precipitation than the low urbanization scenario. In all the urbanized regions and for all seasons the evapotranspiration reduces, and the surface runoff increases. These changes are mainly due to the replacement of vegetated land with more impervious land cover. Changes in the water cycle and the surface properties modified the energy partitioning over the urbanized regions, sensible heat fluxes increased and latent heat fluxes reduced. The partition of the sensible and latent heat fluxes is less evident during the early hours of the day. In the late evening the Bowen ratio increases over the urbanized regions.

The 2 models have differences in the changes in temperature, the RegCM model has lower temperature changes in comparison with the WRF model. Urbanization shifts the mean of the temperature to warmer temperatures, however there were minimum changes in the variance of temperatures. There is a general increase in the minimum temperatures over the urbanized region. Diurnal temperature ranges are reduced in both seasons. Seasonally the changes in diurnal temperature range are lower for the Long Rains during the day in comparison with the Short Rains. These differences can be attributed to the differences in total precipitation for the 2 seasons. During the Long Rains the soil moisture is high and hence evaporative cooling minimizes the effect of the UHI. The size of the urbanized region was found to have an impact on the temperature increase. The changes in temperature for the high urbanization scenario are approximately 2°C higher than the low urbanization scenario. The increase in temperature around the lake modifies the lake-land breezes. During the day the land is much warmer than the lake and hence this strengthens the lake breeze. Modification of the land-lake breeze leads to
modification of total precipitation over the urbanized regions. The urbanized regions have higher precipitation when compared to the vegetated land that it replaces. During both seasons the changes in the lighter precipitation is not high, however there is a difference for precipitation greater than 10mm/day. Urbanized regions have a higher probability of getting heavier precipitation. The precipitation changes are not uniform throughout the urbanized regions. The western side of LVB had increased precipitation that the eastern side. These changes are due to an interaction of the mesoscale flow, lake breezes and mountain breezes.

Sensitivity simulations were done to investigate which processes were dominant in producing the precipitation characteristics of an urbanized region. We identify 3 processes that could lead to a change in precipitation in the region, these are thermally driven convergence, frictionally driven convergence and moisture supply, from both the lake and the soil. These processes are directly affected by the land cover properties. In this study we test the sensitivity of the albedo, surface roughness, soil moisture and the moisture from Lake Victoria in producing the precipitation patterns over the region. For the simulations conducted with the lake included but land surface characteristics altered, the change in surface albedo gives the highest changes in the precipitation. When the surface albedo is increased the thermal contrasts between the lake and the land is altered. At night a region with high albedo looses heat faster than when the albedo is low. Increasing the surface albedo over urbanized regions will lead to a reduction in the total precipitation. When surface roughness is increased it leads to minimum changes in the total precipitation. The roughness effect could be diminished because the thermal convergence due to the low albedo over urbanized regions dominates over the frictional convergence. Increasing the soil moisture leads to a bifurcation in precipitation. Some regions have an increase in precipitation and other regions have a decrease. The experiments that were conducted without the moisture supply from the Lake Victoria had a reduction in precipitation of up to 5mm/day. This shows that if urbanized regions were in regions that were remote from close bodies, there is a possibility that the total rainfall could decrease. The decrease could lead to a positive feedback that will lead to more decreases in precipitation.

The applications section investigated the impact of urbanization on the agricultural and energy sectors. We analyzed the impact of urbanization on the dry and the wet
In the urbanized regions there is an increase in the wettest day in both seasons. Increases could be due to increases in the thermal convergence over the urbanized region. The increases in the wettest 1 day precipitation are higher in the Long Rains season as compared to the Short Rains season. These changes could be due to the interaction of the ITCZ and the local circulation. The number of consecutive dry days does not change during the Short Rains, however the Long Rains have an increase in the number of dry days. During the long rains there will be an increase in total precipitation, though more intense rainfall events, however the season will have more dry days. Thus in urbanized regions there will be an increase in the number of dry days.

The last section investigated the wind resources in the region and the sensitivity to changes in land cover. WRFV3.7 is able to capture the spatial variability of the wind speed, when compared to ERA-Interim reanalysis data. This study has shown that wind speed in the region is seasonal. Higher wind speeds occur during the dry season when the south easterly monsoon is more dominant. The Turkana channel has the highest wind speed due to the complex topography that leads to channeling. Other regions that have high wind speed is the Northern region of Tanzania along mountain ranges and also the North Eastern region in Uganda which is very close to the Turkana region. In these regions we propose that hub height of turbines be greater than 50m. At this height the wind speed are relatively high and the power density will be in regions that can allow for power production throughout the day. Urbanized regions do not have a high impact in wind speed, with the average change in wind speed being ±1ms⁻¹. The changes in the high urbanization scenario experiences slightly higher changes in wind speed in comparison with the low urbanization scenario.

8.2 Recommendations

This study has highlighted the ability of the Regional Climate models to reproduce the temporal and spatial variability of precipitation; the impact of rapid urbanization on the climate and how changes in urbanization are likely to impact the agricultural and energy sectors in East Africa.

The first section of our study mostly focused on the customization of the RegCM and
the ability of the model to reproduce the inter-annual variability of precipitation. In our analysis we chose 2 domains, East Africa and Lake Victoria Basin. Although the correlations between the TRMM timeseries and Lake Victoria Basin timeseries are statistically significant they are slightly lower than that for East Africa domain and TRMM. This might imply that over the Basin the model does not adequately capture the dynamics. Also a timeseries plot that was constructed using the grid points on the western side of the lake showed that the model is unable to capture the amplitude of precipitation, often underestimating. In this study we used a 1-Dimensional lake model which assumes static conditions and does not capture the lake circulation. Past studies done over the region have shown that inclusion of the 3-Dimensional lake model improves the representation of the precipitation. Thus future studies should incorporate the 3-Dimensional lake model to investigate the inter-annual variability of precipitation. During the Long Rains season, the study was unable to identify the teleconnections for the East Africa Domain. However for the Lake Victoria Basin domain the Indian Ocean Dipole was found to influence the precipitation. This shows that the region is inhomogeneous and in order to understand the climate drivers during the Long Rains, smaller homogeneous regions need to be analyzed.

Increased urbanization led to an increase in the surface temperature and the total precipitation. One fascinating result is that the eastern region does not have the same response as the western side. This sensitivity could be fundamental in future urban planning. On the western side urban planners will need to develop more efficient drainage system that will reduce the likelihood of flooding in the region. On the eastern side there might be no need in investing in more efficient drainage systems.

One limitation of this study is that for the projected urbanization experiments we used lateral boundary conditions for the current years. Thus the study did not account for increases in greenhouse gases in the atmosphere. The Intergovernmental Project on Climate Changes proposed a 2°C threshold as the maximum increase in temperature that will not be destructive to livelihoods. Above this temperature increase, human livelihoods and the ecosystem will be negatively impacted. Not much emphasis has been placed on how the UHI could contribute to the increases in temperature. We propose that in the future a study can be conducted that investigates the impacts of urbanization on temperature in a warmer planet. Thus in the future we propose a suit of experiments that
will investigate the impact of both increases in greenhouse gases and urbanization on the climate. Important information that can be extracted from this simulation will include by how much the urban heat island will change in a warming environment. East Africa is projected to have an increase in the precipitation in the future, the question is “Will the interaction of the urban heat island induced convection and the warmer world produce more intense precipitation that will lead to flooding in the lakeside communities?”. This study also highlighted differences in the response to urbanizations of the eastern, western and northern part of the Basin. The western part of the domain received the highest precipitation increases. It will be interesting in the future to model different urbanization designs and how they could impact the total precipitation in the region.

For the wind energy sector, we identified regions that could have high potential for wind energy. However these results are not validated. East Africa has limited number of station data. Thus we propose that they could setup more observational data. Different statistical techniques can be used to identify the optimal number of stations and also the optimal positions of the stations.
REFERENCES


Annez P C and Buckley R M, (2009); Urbanization and growth: Setting the context. Urbanization and growth, 1, 1.


Anyah R and Semazzi F, (2004); Simulation of the sensitivity of Lake Victoria basin climate to lake surface temperatures. Theoretical and Applied Climatology, 79, 55.


Asnani G, (1993); Climatology of the tropics. Tropical Meteorology, 1, 100.

241


Birkett C, Murtugudde R and Allan T, (1999); Indian Ocean climate event brings floods to East Africa’s lakes and the Sudd Marsh. Geophysical Research Letters, 26, 1031.


Camberlin P and Okoola R, (2003); The onset and cessation of the long rains in eastern Africa and their interannual variability. Theoretical and Applied Climatology, 75, 43.


Dickinson R E, Kennedy P and Henderson-Sellers A, Biosphere-atmosphere transfer scheme (BATS) version 1e as coupled to the NCAR community climate model, National Center for Atmospheric Research, Climate and Global Dynamics Division, 1993.

243


Doughty C E, Loarie S R and Field C B, (2012); Theoretical impact of changing albedo on precipitation at the southernmost boundary of the ITCZ in South America. Earth Interactions, 16, 1.

EAC Secretariat E, (2006); Agriculture and Rural Development Policy for the East African Community.


Gallo K and Xian G, (2014); Application of spatially gridded temperature and land cover data sets for urban heat island analysis. Urban Climate, 8, 1.

Gao X, Shi Y, Song R, Giorgi F, Wang Y and Zhang D, (2008); Reduction of future monsoon precipitation over China: Comparison between a high resolution RCM simulation and the driving GCM. Meteorology and Atmospheric Physics, 100, 73.


Guo X, Fu D and Wang J, (2006); Mesoscale convective precipitation system modified by urbanization in Beijing City. Atmospheric Research, 82, 112.


Holtslag A and Boville B, (1993); Local versus nonlocal boundary-layer diffusion in a global climate model. Journal of Climate, 6, 1825.


Huffman G and Bolvin D, (2011); TRMM and other data precipitation data set documentation. Laboratory for Atmospheres, NASA Goddard Space Flight Center and Science Systems and Applications, Inc.


Kainkwa R, (2000); Wind speed pattern and the available wind power at Basotu, Tanzania. Renewable energy, 21, 289.


Kamau J, Kinyua R and Gathua J, (2010); 6 years of wind data for Marsabit, Kenya average over 14m/s at 100m hub height; An analysis of the wind energy potential. Renewable Energy, 35, 1298.


Karekezi S and Kithyoma W, (2002); Renewable energy strategies for rural Africa: is a PV-led renewable energy strategy the right approach for providing modern energy to the rural poor of sub-Saharan Africa? Energy Policy, 30, 1071.


Lake Victoria Basin Commission L et al., (2012); Regional Transboundary Diagnostic Analysis of the Lake Victoria Basin.


Lorenz E N, (1956); Empirical orthogonal functions and statistical weather prediction.


Nogherotto R, (2015); A numerical framework for multiple phase cloud microphysics in regional and global atmospheric models.


Okoola R E, (1999); A diagnostic study of the eastern Africa monsoon circulation during the Northern Hemisphere spring season. International Journal of Climatology, 19, 143.

Oleson K, (2012); Contrasts between urban and rural climate in CCSM4 CMIP5 climate change scenarios. Journal of Climate, 25, 1390.


Otieno V O and Anyah R, (2012); Effects of land use changes on climate in the Greater Horn of Africa.


Pfeiffer M and Dullo W C, (2006); Monsoon-induced cooling of the western equatorial Indian Ocean as recorded in coral oxygen isotope records from the Seychelles covering the period of 1840–1994AD. Quaternary Science Reviews, 25, 993.

Pielke R A, (2001); Mesoscale Meteorological Modeling.


Rauscher S A, Coppola E, Piani C and Giorgi F, (2010); Resolution effects on regional climate model simulations of seasonal precipitation over Europe. Climate dynamics, 35, 685.


Rowell D P, (2013); Simulating SST teleconnections to Africa: What is the state of the art? Journal of Climate, 26, 5397.


Sailor D J, (1995); Simulated urban climate response to modifications in surface albedo and vegetative cover. Journal of applied meteorology, 34, 1694.


Saji N and Yamagata T, (2003); Possible impacts of Indian Ocean dipole mode events on global climate. Climate Research, 25, 151.

Schneider U, Fuchs T, Meyer-Christoffer A and Rudolf B, (2008); Global precipitation analysis products of the GPCC. Global Precipitation Climatology Centre (GPCC), DWD, Internet Publikation, 112.


Segele Z T and Lamb P J, (2005); Characterization and variability of Kiremt rainy season over Ethiopia. Meteorology and Atmospheric Physics, 89, 153.


Semazzi F and Song Y, (2001); A GCM study of climate change induced by deforestation in Africa. Climate Research, 17, 169.


Smith K, (2011); The Role of the Dominant Modes of Climate Variability over Eastern Africa in Modulating the Hydrology of Lake Victoria.

Steiner A L, Pal J, Giorgi F, Dickinson R and Chameides W, (2005); The coupling of the Common Land Model (CLM0) to a regional climate model (RegCM). Theoretical and Applied Climatology, 82, 225.


Taylor Z B, (2011); The impact of cumulus parameterization schemes on the convective transport of biomass burning emissions using WRF-CHEM.


UN Habit U, (2015); Africa Heads of State summits to discuss urbanisation, says AU Chair.


Williams A P and Funk C, (2011); A westward extension of the warm pool leads to a westward extension of the Walker circulation, drying eastern Africa. Climate Dynamics, 37, 2417.


Zaroug M, Giorgi F, Coppola E, Abdo G and Eltahir E, (2014); Simulating the connections of ENSO and the rainfall regime of East Africa and the upper Blue Nile region using a climate model of the Tropics.


Zou L and Zhou T, (2013); Can a regional ocean–atmosphere coupled model improve the simulation of the interannual variability of the western North Pacific summer monsoon? Journal of Climate, 26, 2353.
Appendix A

Temperature Difference Plots

This section presents the difference plots for the RegCM Evaluation. We start by showing the results using Grell FC and then using MIT-Emanuel.
Figure A.1 Spatial Differences in temperature for during the Short Rains season for some selected years for Grell FC scheme
Figure A.2 Spatial Differences in temperature for during the Long Rains season for each year for some selected years for the Grell FC scheme
Figure A.3 Spatial Differences in temperature for during the Short Rains season for some selected years for MIT-Emanuel scheme
Figure A.4 Spatial Differences in temperature for during the Long Rains season for selected years for MIT-Emanuel scheme
Appendix B

Customization of the Weather Research and Forecasting Model

B.1 Introduction

A limited number of studies have been done on the customization of Weather Research and Forecasting model over East Africa. Pohl et al. (2011) conducted 58 one-year simulations over East Africa. They found that the Shortwave (SW) radiation, convective scheme, land use category and domain size have the highest impact on the precipitation. Recently Argent et al. (2014) and Sun et al. (2014) conducted customization of WRF over East Africa with more emphasis on Lake Victoria Basin. They both found that without the correct representation of the Lake surface temperatures the model is unable to correctly capture the precipitation over the region. Sun et al. (2014) coupled WRF with Princeton Ocean Model (POM) and found that the precipitation over the lake is greatly improved by capturing the correct hydrodynamics over the Lake.

Argent et al. (2014) found that by just changing the cumulus parameterization the simulated rainfall over the region can dramatically change. This supports the conclusions from Pohl et al. (2011) that precipitation is highly sensitive to the cumulus parameterization. The WRF model used by Argent et al. (2014); Pohl et al. (2011) did not contain any lake model, but recently a lake model has been incorporated in WRF model.

Table B.1 summarizes the optimal physics combinations that were obtained by Argent et al. (2014); Pohl et al. (2011); Sun et al. (2014). Their studies agree on the best longwave and shortwave radiation schemes and also the planetary boundary layer scheme. The simulations use different cumulus and microphysics schemes. The research will consider the different physics combinations for the radiation schemes, cumulus, microphysics and
The objective of this study is to find the optimum physics combination that reduces the precipitation bias over the East Africa Domain.

**Table B.1** Optimal physics combinations over East Africa from WRF simulations that have been utilized in past research

<table>
<thead>
<tr>
<th></th>
<th>LW Radiation</th>
<th>SW Radiation</th>
<th>Cumulus</th>
<th>Microphysics</th>
<th>PBL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argent et al. (2014)</td>
<td>RRTM</td>
<td>Dudhia</td>
<td>BMJ</td>
<td>ETA</td>
<td>ACM2</td>
</tr>
<tr>
<td>Sun et al. (2014)</td>
<td>RRTM</td>
<td>Dudhia</td>
<td>Grell-3D</td>
<td>ETA</td>
<td>ACM2</td>
</tr>
<tr>
<td>Pohl et al. (2011)</td>
<td>RRTM</td>
<td>Dudhia</td>
<td>KF</td>
<td>WSM6</td>
<td>ACM2</td>
</tr>
</tbody>
</table>

### B.1.1 Model Description and Experimental Design

WRF model is a mesoscale model with advanced dynamics, physics and numerical schemes. It is a compressible, non-hydrostatic, Euler equation, mesoscale meteorological model. WRF features multiple dynamical cores, a 3-dimension variational (3DVAR) data assimilation and a software architecture accommodating computational parallelism and system flexibility. It is viable across scales ranging from meters to thousands of kilometers. The ARW (Advanced Research WRF) core version used in this study is a state-of-the-art modeling system that is utilized by investigators throughout the scientific community (Shamarock et al., 2008). Table B.2 gives a summary of simulations done to customize the WRF model. The WRF model offers a wide range of cumulus schemes. In this case 3 cumulus schemes were chosen based on past studies. These schemes the Kain Fritsch identified by Pohl et al. (2011), the GRELL 3D identified by Sun et al. (2014) and the Betts-Miller-Janjic (BMJ) used by Argent et al. (2014). The KF scheme is an entraining-detraining model, with parcel buoyancy calculated as a function of parcels mixed laterally between the environment and the updrafts (Kain, 2004; Kain and Fritsch, 1990). Updrafts in the KF scheme are represented using a steady-state entrainment/detrainment plume model in which greater entrainment is favored by large parcel buoyancy and moist environments. It was mainly designed for relatively high resolution of about 20-25km. The KF closure assumption is based on a Convective Available Potential Energy (CAPE) removal process. The updraft, downdraft, and environmental mass fluxes are rearranged
Table B.2 Experiments Conducted to customize WRF model with NOAH land Surface model

<table>
<thead>
<tr>
<th>LW Radiation</th>
<th>SW Radiation</th>
<th>Cumulus</th>
<th>Microphysics</th>
<th>PBL</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAM</td>
<td>CAM</td>
<td>Grell-3D</td>
<td>WSM6</td>
<td>YSU</td>
</tr>
<tr>
<td>RRTMG</td>
<td>RRTMG</td>
<td>Grell-3D</td>
<td>WSM6</td>
<td>YSU</td>
</tr>
<tr>
<td>RRTMG</td>
<td>RRTMG</td>
<td>Grell-3D</td>
<td>LIN</td>
<td>YSU</td>
</tr>
<tr>
<td>RRTM</td>
<td>Dudhia</td>
<td>KF</td>
<td>WSM6</td>
<td>ACM2</td>
</tr>
<tr>
<td>RRTM</td>
<td>Dudhia</td>
<td>Grell-3D</td>
<td>ETA</td>
<td>ACM2</td>
</tr>
<tr>
<td>RRTM</td>
<td>Dudhia</td>
<td>BMJ</td>
<td>ETA</td>
<td>ACM2</td>
</tr>
<tr>
<td>RRTM</td>
<td>Dudhia</td>
<td>BMJ</td>
<td>WSM6</td>
<td>ACM2</td>
</tr>
<tr>
<td>RRTM</td>
<td>Dudhia</td>
<td>BMJ</td>
<td>WSM6</td>
<td>YSU</td>
</tr>
</tbody>
</table>

in the mass column until 90% of the CAPE is removed. The scheme closure assumption is not well suited to tropical environments and can result in overly vigorous convection (Pielke, 2001). The BMJ cumulus parameterization is a moisture adjustment scheme that primarily relaxes the sounding to a referenced well-mixed profile (Betts and Miller, 1993; Janjic, 1994). This adjustment is completed by ensuring that the vertical temperature and moisture profiles are realistic. Unlike other cumulus parameterization schemes, BMJ is based solely on thermodynamics, not momentum. The Grell-Devenyi 3D cumulus parameterization scheme (Grell and Dévényi, 2002) was developed from the Grell scheme (Grell, 1993). It employs a large ensemble of closure assumptions and parameters that are commonly used in numerical models and uses statistical techniques to determine the optimal value for feedback to the entire model. The scheme employs a variety of mass-flux and closure assumptions, feedback assumptions, and trigger functions at each grid point. The various closure assumptions include vertical advection of moisture, vertical velocity, and CAPE removal, similar to that in KF (Taylor, 2011). The feedback assumptions include perturbations based on cloud size and entrainment strength and detrainment from updrafts. The differences in static control are combined with differences in dynamic control (the modulation of the convection by the environment) which are based on CAPE, vertical velocity, and moisture convergence.

Three microphysics schemes are also chosen for comparison and these are the WSM6, which was developed from the WSM5 scheme by adding graupel processes (Hong and
Lim, 2006), the Lin scheme (Lin et al., 1983), which utilizes the bulk water microphysical parameterization technique to represent the precipitation fields and the Eta scheme (Rogers et al., 2001).

Two PBL schemes are tested the Yonsei University scheme (YSU), which is a non-local-$K$ scheme with explicit entrainment layer and parabolic $K$ profile in unstable mixed layer, and the Asymmetric Convective Model (ACM) with non-local upward mixing and local downward mixing. Three pairs of radiation scheme were used, CAM for both short-wave and long wave, RRTMG for both the LW and SW, and the RRTM and Dudhia for LW and SW respectively. Three domains were designed for the simulation, that is the 36km, 12km and 4km. In order to save on the computational allocation simulations were run with one way nesting of the 36km and 12 km resolution. For a few simulations only the 36km resolution was run. Once the best combination was obtained the simulation incorporated the 4km resolution. The lateral boundary forcing data was from the National Centers for Environmental Predictions (NCEP) operational Global Final (FNL) Analyses (NCEP/FNL)

**B.1.2 Results and Discussion**

**B.1.2.1 Precipitation**

The precipitation plots (Fig. B.1) show that WRF model generally over predicts precipitation over the Indian ocean, and over land, however it underestimates precipitation over the lake as compared to the TRMM satellite datasets. The results are consistent with Argent et al. (2014) findings who also found a overestimation of precipitation over oceans and under precipitation over the lake. In order to improve the precipitation over the lake, Sun et al. (2014) coupled an ocean model to the WRF model to correctly simulate the hydrodynamics of the lake. The results showed great improvement on the precipitation over the lake as compared to the WRF model without any lake model. Using the 1- Dimensional lake model, the precipitation over Lake Victoria Basin, increased considerably as compared to the simulations without the lake model. The different physics combinations lead to different patterns in precipitation over the region. Simulations (d) and (h) (Fig. B.1 (d), Fig. B.1(h)) are able to capture the asymmetrical pattern over the lake however they still underestimate the total precipitation over the lake. The rest of the other simulations are unable to capture this pattern and greatly underestimate precipi-
As mentioned before the oceans perform badly in almost all the simulations with simulations (a) and (d) having the worst simulations over the ocean. The input data and the lack of an ocean model are the major factors that lead to biases over the oceans. The simulations utilizing the BMJ scheme show more realistic precipitation amounts over the ocean. With the use of YSU PBL scheme producing minimum biases as compared to the ones utilizing the ACM PBL scheme.

The model tends to over predict the precipitation amount in the Congo Basin. This occurs especially over North of Lake Tanganyika in all the simulations except (Fig. B.1 e), which greatly underestimate precipitation. Although this region also has high precipitation in TRMM the precipitation from the model is much higher than in the observational datasets. The simulation using Pohl et al. (2011) optimal physics combination has the highest overestimation of precipitation over the region. This is consistent with the drawbacks of the KF cumulus scheme, which is known to produce a lot of rainfall over the tropics. The simulations with the BMJ cumulus scheme produces the most realistic precipitation over the Congo region. This region is difficult to evaluate because of lack of station data to bias correct the TRMM satellite datasets. This overestimation seems to appear in all simulation runs. Run 1 which includes GRELL cumulus scheme produced the poorest distribution of rainfall over the region (Fig. B.1 a). This result is consistent with the results obtained by Pohl et al. (2011). Using the LIN microphysics scheme with the KF cumulus scheme over predicts the total precipitation both on land and the ocean. However using WSM6 microphysics scheme seem to substantially lower the precipitation over the land and give a reasonable precipitation distribution.

Over East Africa region the model greatly underestimates the total precipitation. With run (Fig. B.1 (e)) showing the worst precipitation distribution over this region. This simulation utilized Grell 3D cumulus physics and the Eta microphysics. Exponential and linear boundary conditions were used in the simulations however there were no major changes on the response of the lateral boundaries. Considerable improvements in the boundaries occurred when the model used larger buffer zone. In this case 11 relaxation points where used (Fig. B.1).
Figure B.1 Precipitation distribution obtained from the different physics combinations used in the simulation for Domain 1

B.2 Conclusion

This section of the study customized the WRF model. Different cumulus, microphysics and planetary boundary layer schemes were considered. The Grell scheme had the worst spatial pattern in comparison with the other cumulus schemes. The BMJ cumulus scheme performs better than all the other schemes when spatial comparison is done. However the scheme does have some deficiencies over East Africa domain where the model underestimates the total precipitation. When the YSU scheme is utilized the spatial distribution
of the precipitation improves over the Region. Our optimal physics combination is using the BMJ cumulus scheme, WSM6 microphysics scheme, the YSU PBL, RRTM longwave radiation and the Dudhia shortwave radiation.
Appendix C

WRF and RegCM evaluation

Table C.1 Summary of statistics for OND temperature for WRF and REGCM, OBS- observational datasets, SD Mod.- standard deviation calculated from model output and SD Obs. - standard deviation for observations, RMSE is the root mean square error, MB- mean bias, NME is the normalized mean error, NMB is the normalized mean bias

<table>
<thead>
<tr>
<th>Year</th>
<th>OBS.</th>
<th>SD Mod.</th>
<th>SD Obs.</th>
<th>MB</th>
<th>NMB</th>
<th>NME</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>WRF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>CRU</td>
<td>2.75</td>
<td>2.47</td>
<td>-0.49</td>
<td>-2.6</td>
<td>4.24</td>
<td>1.44</td>
</tr>
<tr>
<td>2006</td>
<td>UDELAWARE</td>
<td>2.75</td>
<td>2.58</td>
<td>-0.03</td>
<td>-0.85</td>
<td>3.99</td>
<td>1.38</td>
</tr>
<tr>
<td>2007</td>
<td>CRU</td>
<td>2.79</td>
<td>2.61</td>
<td>-0.19</td>
<td>-1.32</td>
<td>3.85</td>
<td>1.34</td>
</tr>
<tr>
<td>2007</td>
<td>UDELAWARE</td>
<td>2.79</td>
<td>2.72</td>
<td>0.05</td>
<td>-0.43</td>
<td>4.72</td>
<td>1.54</td>
</tr>
<tr>
<td>2008</td>
<td>CRU</td>
<td>2.75</td>
<td>2.7</td>
<td>0.33</td>
<td>0.88</td>
<td>3.88</td>
<td>1.28</td>
</tr>
<tr>
<td>2008</td>
<td>UDELAWARE</td>
<td>2.75</td>
<td>2.97</td>
<td>-0.08</td>
<td>-0.84</td>
<td>4.87</td>
<td>1.6</td>
</tr>
<tr>
<td>2009</td>
<td>CRU</td>
<td>2.71</td>
<td>2.55</td>
<td>-0.49</td>
<td>-2.3</td>
<td>4</td>
<td>1.4</td>
</tr>
<tr>
<td>2009</td>
<td>UDELAWARE</td>
<td>2.71</td>
<td>2.82</td>
<td>0.22</td>
<td>0.44</td>
<td>4.45</td>
<td>1.47</td>
</tr>
<tr>
<td>2010</td>
<td>CRU</td>
<td>2.72</td>
<td>2.59</td>
<td>-0.46</td>
<td>-2.25</td>
<td>4.52</td>
<td>1.53</td>
</tr>
<tr>
<td>2010</td>
<td>UDELAWARE</td>
<td>2.72</td>
<td>2.86</td>
<td>-0.42</td>
<td>2.13</td>
<td>5.08</td>
<td>1.68</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>OBS.</th>
<th>SD Mod.</th>
<th>SD Obs.</th>
<th>MB</th>
<th>NMB</th>
<th>NME</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>REGCM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>CRU</td>
<td>2.39</td>
<td>2.39</td>
<td>-0.51</td>
<td>-2.3</td>
<td>4.72</td>
<td>1.42</td>
</tr>
<tr>
<td>2006</td>
<td>UDELAWARE</td>
<td>2.39</td>
<td>2.53</td>
<td>0.01</td>
<td>0.05</td>
<td>3.78</td>
<td>1.23</td>
</tr>
<tr>
<td>2007</td>
<td>CRU</td>
<td>2.42</td>
<td>2.51</td>
<td>-0.21</td>
<td>-1.05</td>
<td>3.77</td>
<td>1.24</td>
</tr>
<tr>
<td>2007</td>
<td>UDELAWARE</td>
<td>2.42</td>
<td>2.65</td>
<td>0.05</td>
<td>-1.5</td>
<td>4.36</td>
<td>1.43</td>
</tr>
<tr>
<td>2008</td>
<td>CRU</td>
<td>2.41</td>
<td>2.6</td>
<td>0.52</td>
<td>2.12</td>
<td>4.14</td>
<td>1.31</td>
</tr>
<tr>
<td>2008</td>
<td>UDELAWARE</td>
<td>2.41</td>
<td>2.88</td>
<td>0.12</td>
<td>0.48</td>
<td>4.55</td>
<td>1.5</td>
</tr>
<tr>
<td>2009</td>
<td>CRU</td>
<td>2.5</td>
<td>2.45</td>
<td>-0.54</td>
<td>-2.11</td>
<td>4.78</td>
<td>1.46</td>
</tr>
<tr>
<td>2009</td>
<td>UDELAWARE</td>
<td>2.5</td>
<td>2.76</td>
<td>0.25</td>
<td>1</td>
<td>4.17</td>
<td>1.36</td>
</tr>
<tr>
<td>2010</td>
<td>CRU</td>
<td>2.4</td>
<td>2.48</td>
<td>-0.46</td>
<td>-1.8</td>
<td>4.2</td>
<td>1.35</td>
</tr>
<tr>
<td>2010</td>
<td>UDELAWARE</td>
<td>2.4</td>
<td>2.78</td>
<td>-0.41</td>
<td>-1.74</td>
<td>5.02</td>
<td>1.65</td>
</tr>
</tbody>
</table>
Figure C.1 Average precipitation for the short rain season for WRF (left column), REGCM (middle column) and TRMM (right column)
Figure C.2 Average precipitation for the Long rains season for WRF (left column), REGCM (middle column) and TRMM (right column)
Table C.2 Summary of Evaluation Statistics for WRF Simulations. SD Mod.- standard deviation calculated from model output and SD Obs. - standard deviation for observations, MB- mean bias, NME (%) is the normalized mean error, NMB (%) is the normalized mean bias, RMSE is the root mean square error

<table>
<thead>
<tr>
<th>Year</th>
<th>OBS.</th>
<th>SD Mod.</th>
<th>SD Obs.</th>
<th>MB</th>
<th>NMB</th>
<th>NME</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>CMORPH</td>
<td>356</td>
<td>307</td>
<td>9.2</td>
<td>2.1</td>
<td>42.8</td>
<td>271</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>356</td>
<td>189</td>
<td>105</td>
<td>32</td>
<td>50</td>
<td>271</td>
</tr>
<tr>
<td></td>
<td>GPCP</td>
<td>356</td>
<td>201</td>
<td>61.32</td>
<td>15</td>
<td>41</td>
<td>248</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>356</td>
<td>195</td>
<td>59.7</td>
<td>17</td>
<td>46</td>
<td>272</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>356</td>
<td>215</td>
<td>55</td>
<td>13</td>
<td>41</td>
<td>252</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>356</td>
<td>214</td>
<td>65</td>
<td>16.5</td>
<td>47</td>
<td>272</td>
</tr>
<tr>
<td>2007</td>
<td>CMORPH</td>
<td>283</td>
<td>364</td>
<td>-97.7</td>
<td>-25</td>
<td>46.6</td>
<td>246</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>283</td>
<td>195</td>
<td>-18</td>
<td>-5.8</td>
<td>35</td>
<td>164</td>
</tr>
<tr>
<td></td>
<td>GPCP</td>
<td>284</td>
<td>241</td>
<td>-20.05</td>
<td>-6.4</td>
<td>35</td>
<td>152</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>284</td>
<td>218</td>
<td>-35</td>
<td>-4.4</td>
<td>34</td>
<td>159</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>284</td>
<td>223</td>
<td>6.47</td>
<td>2.25</td>
<td>38.4</td>
<td>156</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>284</td>
<td>239</td>
<td>-20</td>
<td>-6</td>
<td>37</td>
<td>163</td>
</tr>
<tr>
<td>2008</td>
<td>CMORPH</td>
<td>277</td>
<td>379</td>
<td>-103.19</td>
<td>-25</td>
<td>42.9</td>
<td>238</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>277</td>
<td>173</td>
<td>-10</td>
<td>-3.4</td>
<td>39.3</td>
<td>179</td>
</tr>
<tr>
<td></td>
<td>GPCP</td>
<td>284</td>
<td>167</td>
<td>35.53</td>
<td>13.72</td>
<td>43.4</td>
<td>175</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>284</td>
<td>163</td>
<td>18.09</td>
<td>6.5</td>
<td>45.8</td>
<td>198.9</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>284</td>
<td>174</td>
<td>33</td>
<td>12.7</td>
<td>44</td>
<td>177</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>284</td>
<td>175</td>
<td>34</td>
<td>13</td>
<td>47</td>
<td>177</td>
</tr>
<tr>
<td>2009</td>
<td>CMORPH</td>
<td>271</td>
<td>387</td>
<td>-123.6</td>
<td>-29</td>
<td>45</td>
<td>263</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>277</td>
<td>171</td>
<td>-20</td>
<td>-6</td>
<td>40</td>
<td>183</td>
</tr>
<tr>
<td></td>
<td>GPCP</td>
<td>277</td>
<td>180</td>
<td>-2.71</td>
<td>-0.87</td>
<td>39</td>
<td>170</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>277</td>
<td>145</td>
<td>4.64</td>
<td>1.52</td>
<td>52</td>
<td>222</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>277</td>
<td>191</td>
<td>-10</td>
<td>-3.1</td>
<td>40</td>
<td>178</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>277</td>
<td>169</td>
<td>14.7</td>
<td>5</td>
<td>50</td>
<td>204</td>
</tr>
<tr>
<td>2010</td>
<td>CMORPH</td>
<td>271</td>
<td>387</td>
<td>-136.93</td>
<td>-35.8</td>
<td>47</td>
<td>271.9</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>271</td>
<td>161</td>
<td>-52</td>
<td>-17</td>
<td>42</td>
<td>174</td>
</tr>
<tr>
<td></td>
<td>GPCP</td>
<td>271</td>
<td>212</td>
<td>-23</td>
<td>-8.5</td>
<td>34.7</td>
<td>143</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>271</td>
<td>192</td>
<td>-40</td>
<td>-14</td>
<td>39</td>
<td>166</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>271</td>
<td>217</td>
<td>-32</td>
<td>-11</td>
<td>35</td>
<td>146</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>271</td>
<td>205</td>
<td>-15</td>
<td>-6</td>
<td>39</td>
<td>156</td>
</tr>
</tbody>
</table>
Table C.3 Summary of Evaluation Statistics for RegCM Simulations. SD Mod.- standard deviation calculated from model output and SD Obs. - standard deviation for observations, MB- mean bias, NME (%) is the normalized mean error, NMB (%) is the normalized mean bias, RMSE is the root mean square error

<table>
<thead>
<tr>
<th>Year</th>
<th>OBS.</th>
<th>SD Mod.</th>
<th>SD Obs.</th>
<th>MB</th>
<th>NMB</th>
<th>NME</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>CMORPH</td>
<td>306</td>
<td>291</td>
<td>-48.02</td>
<td>-10.3</td>
<td>38.6</td>
<td>262</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>306</td>
<td>200</td>
<td>62.52</td>
<td>18.4</td>
<td>46.3</td>
<td>268.6</td>
</tr>
<tr>
<td></td>
<td>GPCP</td>
<td>306</td>
<td>228</td>
<td>-10.21</td>
<td>-2.4</td>
<td>38</td>
<td>243.8</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>306</td>
<td>234</td>
<td>10.4</td>
<td>2.82</td>
<td>45</td>
<td>275</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>306</td>
<td>244</td>
<td>-15.6</td>
<td>-3.6</td>
<td>39.6</td>
<td>252</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>306</td>
<td>258</td>
<td>-23.07</td>
<td>-5.2</td>
<td>42</td>
<td>272</td>
</tr>
<tr>
<td>2006</td>
<td>CMORPH</td>
<td>248</td>
<td>309</td>
<td>-41</td>
<td>-11.9</td>
<td>46.8</td>
<td>239</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>248</td>
<td>214</td>
<td>-26.3</td>
<td>9.6</td>
<td>44</td>
<td>217</td>
</tr>
<tr>
<td></td>
<td>GPCP</td>
<td>248</td>
<td>245</td>
<td>-18.2</td>
<td>-5.7</td>
<td>42</td>
<td>196.8</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>248</td>
<td>258</td>
<td>-65</td>
<td>-1.8</td>
<td>43.4</td>
<td>214</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>248</td>
<td>233</td>
<td>6.77</td>
<td>2.28</td>
<td>44.3</td>
<td>196</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>248</td>
<td>257</td>
<td>-18</td>
<td>-5.6</td>
<td>47</td>
<td>227</td>
</tr>
<tr>
<td>2008</td>
<td>CMORPH</td>
<td>272</td>
<td>301</td>
<td>-34</td>
<td>-9.8</td>
<td>45</td>
<td>247</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>272</td>
<td>208</td>
<td>-20</td>
<td>12</td>
<td>49</td>
<td>258</td>
</tr>
<tr>
<td></td>
<td>GPCP</td>
<td>272</td>
<td>164</td>
<td>50.5</td>
<td>19</td>
<td>50.2</td>
<td>223</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>272</td>
<td>174</td>
<td>25.3</td>
<td>30</td>
<td>60</td>
<td>273</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>272</td>
<td>174</td>
<td>49.14</td>
<td>18</td>
<td>52</td>
<td>228</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>272</td>
<td>187</td>
<td>36.9</td>
<td>13</td>
<td>56.6</td>
<td>247</td>
</tr>
<tr>
<td>2009</td>
<td>CMORPH</td>
<td>285</td>
<td>322</td>
<td>-34</td>
<td>-9</td>
<td>47.5</td>
<td>270</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>285</td>
<td>186</td>
<td>6.7</td>
<td>19.2</td>
<td>51</td>
<td>269</td>
</tr>
<tr>
<td></td>
<td>GPCP</td>
<td>285</td>
<td>195</td>
<td>27.04</td>
<td>8.69</td>
<td>46</td>
<td>232</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>285</td>
<td>217</td>
<td>11</td>
<td>21</td>
<td>58</td>
<td>303</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>285</td>
<td>209</td>
<td>23.46</td>
<td>7.5</td>
<td>48</td>
<td>237</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>285</td>
<td>207</td>
<td>28.7</td>
<td>9.28</td>
<td>52.8</td>
<td>253</td>
</tr>
<tr>
<td>2010</td>
<td>CMORPH</td>
<td>269</td>
<td>333.5</td>
<td>-25</td>
<td>-7.7</td>
<td>46.3</td>
<td>233</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>269</td>
<td>187</td>
<td>-21.67</td>
<td>18.9</td>
<td>47</td>
<td>236</td>
</tr>
<tr>
<td></td>
<td>GPCP</td>
<td>269</td>
<td>270</td>
<td>41</td>
<td>16</td>
<td>45.6</td>
<td>192</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>269</td>
<td>224</td>
<td>-14.8</td>
<td>22</td>
<td>48</td>
<td>242</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>269</td>
<td>220</td>
<td>35</td>
<td>13.3</td>
<td>46</td>
<td>196</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>269</td>
<td>230</td>
<td>34.7</td>
<td>13.29</td>
<td>51</td>
<td>215</td>
</tr>
</tbody>
</table>
Table C.4 Summary of statistics for MAM temperature for WRF and REGCM, OBS- observational datasets, SD Mod.- standard deviation calculated from model output and SD Obs. - standard deviation for observations, RMSE is the root mean square error, MB- mean bias, NME is the normalized mean error, NMB is the normalized mean bias

<table>
<thead>
<tr>
<th>Year</th>
<th>OBS.</th>
<th>SD Mod.</th>
<th>SD Obs.</th>
<th>MB</th>
<th>NMB</th>
<th>NME</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WRF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>UDELAWARE</td>
<td>2.94</td>
<td>2.96</td>
<td>0</td>
<td>-0.08</td>
<td>3.7</td>
<td>1.32</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>2.94</td>
<td>3.16</td>
<td>0.16</td>
<td>0.52</td>
<td>4.6</td>
<td>1.53</td>
</tr>
<tr>
<td>2007</td>
<td>UDELAWARE</td>
<td>2.92</td>
<td>3.08</td>
<td>0.08</td>
<td>-0.22</td>
<td>-1.44</td>
<td>3.48</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>2.96</td>
<td>3.07</td>
<td>-0.08</td>
<td>-0.64</td>
<td>3.83</td>
<td>1.38</td>
</tr>
<tr>
<td>2008</td>
<td>UDELAWARE</td>
<td>2.88</td>
<td>3.17</td>
<td>0.04</td>
<td>-0.18</td>
<td>-0.81</td>
<td>3.9</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>2.88</td>
<td>3.02</td>
<td>0.07</td>
<td>0</td>
<td>3.1</td>
<td>1.37</td>
</tr>
<tr>
<td>2009</td>
<td>UDELAWARE</td>
<td>2.95</td>
<td>3.11</td>
<td>-0.14</td>
<td>-0.54</td>
<td>3.86</td>
<td>1.48</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>2.95</td>
<td>3.02</td>
<td>-0.07</td>
<td>-0.3</td>
<td>3.1</td>
<td>1.37</td>
</tr>
<tr>
<td>2010</td>
<td>UDELAWARE</td>
<td>2.95</td>
<td>3.11</td>
<td>-0.14</td>
<td>-0.54</td>
<td>3.86</td>
<td>1.48</td>
</tr>
<tr>
<td></td>
<td>REGCM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>UDELAWARE</td>
<td>2.6</td>
<td>2.9</td>
<td>-0.16</td>
<td>-0.6</td>
<td>3.5</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>2.7</td>
<td>3</td>
<td>-0.06</td>
<td>0.12</td>
<td>4</td>
<td>1.36</td>
</tr>
<tr>
<td>2007</td>
<td>UDELAWARE</td>
<td>2.67</td>
<td>3.3</td>
<td>-0.06</td>
<td>-0.24</td>
<td>4.2</td>
<td>1.45</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>2.6</td>
<td>2.8</td>
<td>-0.18</td>
<td>-0.7</td>
<td>3.6</td>
<td>1.2</td>
</tr>
<tr>
<td>2008</td>
<td>UDELAWARE</td>
<td>2.6</td>
<td>3.1</td>
<td>-0.01</td>
<td>-0.6</td>
<td>4.2</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>2.7</td>
<td>3</td>
<td>-0.08</td>
<td>-0.3</td>
<td>3.7</td>
<td>1.3</td>
</tr>
<tr>
<td>2009</td>
<td>UDELAWARE</td>
<td>2.7</td>
<td>3.2</td>
<td>0.59</td>
<td>0.46</td>
<td>4.7</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>2.61</td>
<td>2.89</td>
<td>0.02</td>
<td>0.06</td>
<td>3.57</td>
<td>1.22</td>
</tr>
<tr>
<td>2010</td>
<td>UDELAWARE</td>
<td>2.61</td>
<td>3</td>
<td>-0.07</td>
<td>-0.28</td>
<td>3.87</td>
<td>1.36</td>
</tr>
</tbody>
</table>
Table C.5 Summary of Evaluation Statistics for WRF Simulations. SD Mod.- standard deviation calculated from model output and SD Obs. - standard deviation for observations, MB- mean bias, NME (%) is the normalized mean error, NMB (%) is the normalized mean bias, RMSE is the root mean square error.

<table>
<thead>
<tr>
<th>Year</th>
<th>OBS.</th>
<th>SD Mod.</th>
<th>SD Obs.</th>
<th>MB</th>
<th>NMB</th>
<th>NME</th>
<th>RMSE</th>
<th>CORREL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CMORPH</td>
<td>240</td>
<td>149</td>
<td>-198</td>
<td>-42</td>
<td>51</td>
<td>304</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>240</td>
<td>133</td>
<td>-90</td>
<td>-25</td>
<td>46</td>
<td>210</td>
<td>0.52</td>
</tr>
<tr>
<td>2006</td>
<td>GPCP</td>
<td>240</td>
<td>147</td>
<td>-95</td>
<td>-22</td>
<td>45</td>
<td>200</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>240</td>
<td>147</td>
<td>-95</td>
<td>-22</td>
<td>45.2</td>
<td>201</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>240</td>
<td>172</td>
<td>-101</td>
<td>-27</td>
<td>45.6</td>
<td>220</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>240</td>
<td>142</td>
<td>-96</td>
<td>-26</td>
<td>48</td>
<td>229</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>CMORPH</td>
<td>208</td>
<td>180</td>
<td>-200</td>
<td>-45</td>
<td>50</td>
<td>277</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>208</td>
<td>135</td>
<td>-90</td>
<td>-27</td>
<td>43</td>
<td>179</td>
<td>0.65</td>
</tr>
<tr>
<td>2007</td>
<td>GPCP</td>
<td>208</td>
<td>176</td>
<td>-99</td>
<td>-27</td>
<td>40.4</td>
<td>171</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>208</td>
<td>176</td>
<td>-99</td>
<td>-27</td>
<td>40.8</td>
<td>180</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>208</td>
<td>171</td>
<td>-100</td>
<td>-29</td>
<td>40.8</td>
<td>182</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>208</td>
<td>177</td>
<td>-98</td>
<td>-28</td>
<td>44</td>
<td>196</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>CMORPH</td>
<td>251</td>
<td>180</td>
<td>-178</td>
<td>-37</td>
<td>48</td>
<td>297</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>251</td>
<td>134</td>
<td>-45</td>
<td>-13</td>
<td>38</td>
<td>174</td>
<td>0.75</td>
</tr>
<tr>
<td>2008</td>
<td>GPCP</td>
<td>251</td>
<td>150</td>
<td>-31.7</td>
<td>-7.6</td>
<td>42.8</td>
<td>179</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>251</td>
<td>150</td>
<td>-32</td>
<td>-7.6</td>
<td>43</td>
<td>189</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>251</td>
<td>138</td>
<td>-17</td>
<td>-5.3</td>
<td>43</td>
<td>187</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>251</td>
<td>159</td>
<td>-45</td>
<td>-12</td>
<td>41</td>
<td>190</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>CMORPH</td>
<td>216</td>
<td>176</td>
<td>-168</td>
<td>-38</td>
<td>45</td>
<td>262</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>216</td>
<td>150</td>
<td>-37</td>
<td>-11.7</td>
<td>33</td>
<td>139</td>
<td>0.76</td>
</tr>
<tr>
<td>2009</td>
<td>GPCP</td>
<td>216</td>
<td>175</td>
<td>-33</td>
<td>-7.6</td>
<td>43</td>
<td>160</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>216</td>
<td>174</td>
<td>-33</td>
<td>-7.7</td>
<td>43</td>
<td>169</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>216</td>
<td>169</td>
<td>-35</td>
<td>-11</td>
<td>39</td>
<td>161</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>216</td>
<td>187</td>
<td>-36</td>
<td>-11</td>
<td>45</td>
<td>185</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>CMORPH</td>
<td>248</td>
<td>139</td>
<td>-138</td>
<td>-29</td>
<td>45</td>
<td>272</td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>248</td>
<td>143</td>
<td>0.02</td>
<td>0.01</td>
<td>37</td>
<td>180</td>
<td>0.68</td>
</tr>
<tr>
<td>2010</td>
<td>GPCP</td>
<td>248</td>
<td>142</td>
<td>43.05</td>
<td>19</td>
<td>53</td>
<td>208</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>248</td>
<td>142</td>
<td>43</td>
<td>19</td>
<td>53</td>
<td>208</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>248</td>
<td>130</td>
<td>24.39</td>
<td>7.6</td>
<td>44</td>
<td>195</td>
<td>0.63</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>248</td>
<td>147</td>
<td>26.09</td>
<td>8.13</td>
<td>50</td>
<td>216</td>
<td>0.51</td>
</tr>
</tbody>
</table>
Table C.6 Summary of Evaluation Statistics for RegCM Simulations during the Long Rains

<table>
<thead>
<tr>
<th>Year</th>
<th>OBS.</th>
<th>SD Mod.</th>
<th>SD Obs.</th>
<th>MB</th>
<th>NMB</th>
<th>NME</th>
<th>RMSE</th>
<th>CORREL</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>CMORPH</td>
<td>226</td>
<td>190</td>
<td>-71</td>
<td>-19.4</td>
<td>44</td>
<td>245</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>226</td>
<td>163</td>
<td>-57</td>
<td>1.32</td>
<td>45</td>
<td>248</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td>GPCP</td>
<td>226</td>
<td>166</td>
<td>-17</td>
<td>-5.5</td>
<td>42.2</td>
<td>202.6</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>226</td>
<td>155</td>
<td>-62</td>
<td>-17</td>
<td>46</td>
<td>250</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>226</td>
<td>180</td>
<td>-22</td>
<td>-7</td>
<td>43</td>
<td>208</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>226</td>
<td>187</td>
<td>-34</td>
<td>-10</td>
<td>47</td>
<td>230</td>
<td>0.41</td>
</tr>
<tr>
<td>2007</td>
<td>CMORPH</td>
<td>202</td>
<td>198</td>
<td>-59.8</td>
<td>-16.4</td>
<td>48</td>
<td>258</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>202</td>
<td>148</td>
<td>-11.8</td>
<td>5.9</td>
<td>39.4</td>
<td>211.9</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>GPCP</td>
<td>202</td>
<td>172</td>
<td>7.81</td>
<td>2.6</td>
<td>40.6</td>
<td>191</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>202</td>
<td>189</td>
<td>-24</td>
<td>-7.2</td>
<td>42</td>
<td>225</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>202</td>
<td>183</td>
<td>2.8</td>
<td>0.91</td>
<td>41.5</td>
<td>196</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>202</td>
<td>201</td>
<td>-11</td>
<td>-3.4</td>
<td>45</td>
<td>216</td>
<td>0.43</td>
</tr>
<tr>
<td>2008</td>
<td>CMORPH</td>
<td>212</td>
<td>183</td>
<td>-71.6</td>
<td>-19.1</td>
<td>45.5</td>
<td>251.6</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>212.7</td>
<td>146.6</td>
<td>-22.6</td>
<td>4.05</td>
<td>43.3</td>
<td>218</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>GPCP</td>
<td>212</td>
<td>138</td>
<td>25.3</td>
<td>9.1</td>
<td>41.2</td>
<td>180</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>212</td>
<td>164</td>
<td>-12.2</td>
<td>-3.8</td>
<td>50</td>
<td>234</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>212</td>
<td>143</td>
<td>31</td>
<td>11.4</td>
<td>44</td>
<td>183</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>212</td>
<td>175</td>
<td>34</td>
<td>13</td>
<td>47</td>
<td>177</td>
<td>0.48</td>
</tr>
<tr>
<td>2009</td>
<td>CMORPH</td>
<td>210.9</td>
<td>174</td>
<td>-9</td>
<td>-2.8</td>
<td>45</td>
<td>211</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>210</td>
<td>182</td>
<td>-37</td>
<td>2.1</td>
<td>38</td>
<td>202</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>GPCP</td>
<td>210</td>
<td>149</td>
<td>16</td>
<td>5.9</td>
<td>42.7</td>
<td>181</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>210</td>
<td>182</td>
<td>-26</td>
<td>-8.2</td>
<td>45.8</td>
<td>221</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>210</td>
<td>161</td>
<td>11.2</td>
<td>4</td>
<td>44</td>
<td>186</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>210</td>
<td>193</td>
<td>-20</td>
<td>-6.4</td>
<td>48</td>
<td>220</td>
<td>0.41</td>
</tr>
<tr>
<td>2010</td>
<td>CMORPH</td>
<td>229</td>
<td>177.7</td>
<td>-77</td>
<td>-19.5</td>
<td>47.3</td>
<td>279</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>229</td>
<td>156</td>
<td>-13</td>
<td>-3.8</td>
<td>44</td>
<td>253</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>GPCP</td>
<td>229</td>
<td>128</td>
<td>23</td>
<td>-8.5</td>
<td>7.8</td>
<td>212</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>229</td>
<td>150</td>
<td>19.1</td>
<td>15.1</td>
<td>52</td>
<td>259</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>229</td>
<td>141</td>
<td>19</td>
<td>6.2</td>
<td>42.5</td>
<td>215</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>229</td>
<td>172</td>
<td>-5</td>
<td>-1.6</td>
<td>46</td>
<td>236</td>
<td>0.33</td>
</tr>
</tbody>
</table>
Table C.7 Summary of Evaluation Statistics for WRF D01 Simulations precipitation for the Short Rains. RMSE is the root mean square error, NME is the normalized mean error, NMB is the normalized mean bias, MB- mean bias, Cor is the pattern correlation, NRMSE is the normalized root mean square SD Mod.- standard deviation calculated from model output and SD Obs. - standard deviation for observations

<table>
<thead>
<tr>
<th>Year</th>
<th>OBS.</th>
<th>RMSE</th>
<th>NME</th>
<th>NMB</th>
<th>MB</th>
<th>Cor</th>
<th>SD Mod.</th>
<th>SD Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>CMORPH</td>
<td>502</td>
<td>314</td>
<td>136</td>
<td>39.5</td>
<td>70.7</td>
<td>406</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>502</td>
<td>222</td>
<td>226</td>
<td>47</td>
<td>63.4</td>
<td>272.4</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>GPCP</td>
<td>502</td>
<td>222</td>
<td>160</td>
<td>52</td>
<td>72.5</td>
<td>403</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>502</td>
<td>259</td>
<td>188</td>
<td>27</td>
<td>64</td>
<td>252</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>502</td>
<td>276</td>
<td>159</td>
<td>50</td>
<td>72</td>
<td>402</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>502</td>
<td>283</td>
<td>160</td>
<td>50</td>
<td>72</td>
<td>402</td>
<td>0.69</td>
</tr>
<tr>
<td>2007</td>
<td>CMORPH</td>
<td>326</td>
<td>293</td>
<td>35</td>
<td>13.5</td>
<td>19.3</td>
<td>234</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>326</td>
<td>226</td>
<td>60</td>
<td>25</td>
<td>44</td>
<td>188</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td>GPCP</td>
<td>326</td>
<td>224</td>
<td>51.5</td>
<td>21</td>
<td>50</td>
<td>195</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>326</td>
<td>265</td>
<td>33.76</td>
<td>7</td>
<td>40.9</td>
<td>188</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>326</td>
<td>231</td>
<td>72</td>
<td>32</td>
<td>56</td>
<td>205</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>326</td>
<td>253</td>
<td>56</td>
<td>23</td>
<td>56</td>
<td>226</td>
<td>0.74</td>
</tr>
<tr>
<td>2008</td>
<td>CMORPH</td>
<td>335</td>
<td>283</td>
<td>51.54</td>
<td>19.3</td>
<td>61.1</td>
<td>253.9</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>335</td>
<td>225</td>
<td>72.7</td>
<td>29.6</td>
<td>49.6</td>
<td>212</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>GPCP</td>
<td>335</td>
<td>174</td>
<td>109</td>
<td>52.5</td>
<td>69</td>
<td>243</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>335</td>
<td>187</td>
<td>105</td>
<td>50</td>
<td>63</td>
<td>229</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>335</td>
<td>182</td>
<td>109.84</td>
<td>52.69</td>
<td>69</td>
<td>244</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>335</td>
<td>196</td>
<td>102</td>
<td>48</td>
<td>70</td>
<td>257</td>
<td>0.73</td>
</tr>
<tr>
<td>2009</td>
<td>CMORPH</td>
<td>347</td>
<td>303</td>
<td>50.5</td>
<td>18.6</td>
<td>65</td>
<td>272</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>347</td>
<td>208</td>
<td>81.9</td>
<td>34.1</td>
<td>50.9</td>
<td>211</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>GPCP</td>
<td>347</td>
<td>205</td>
<td>90</td>
<td>39</td>
<td>63.5</td>
<td>246</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>347</td>
<td>224</td>
<td>86</td>
<td>36</td>
<td>58</td>
<td>238</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>347</td>
<td>216</td>
<td>90</td>
<td>39</td>
<td>63</td>
<td>244</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>347</td>
<td>212</td>
<td>97</td>
<td>43</td>
<td>68</td>
<td>251</td>
<td>0.76</td>
</tr>
<tr>
<td>2010</td>
<td>CMORPH</td>
<td>262</td>
<td>303</td>
<td>-24</td>
<td>-10</td>
<td>55.9</td>
<td>212.8</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>262</td>
<td>202</td>
<td>-14</td>
<td>-5.9</td>
<td>48</td>
<td>189</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>GPCP</td>
<td>262</td>
<td>202</td>
<td>22.31</td>
<td>11</td>
<td>47.6</td>
<td>156</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>262</td>
<td>224</td>
<td>-8</td>
<td>-3.5</td>
<td>47</td>
<td>191</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>262</td>
<td>211</td>
<td>20</td>
<td>9.7</td>
<td>48</td>
<td>158</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>262</td>
<td>220</td>
<td>21.08</td>
<td>11</td>
<td>52</td>
<td>179</td>
<td>0.74</td>
</tr>
</tbody>
</table>
Table C.8 Summary of Evaluation Statistics for WRF D01 temperature Simulations during the Long Rains. RMSE is the root mean square error, NME is the normalized mean error, NMB is the normalized mean bias, MB- mean bias, Cor is the pattern correlation, NRMSE is the normalized root mean square SD Mod.- standard deviation calculated from model output and SD Obs. - standard deviation for observations.

<table>
<thead>
<tr>
<th>Year</th>
<th>OBS.</th>
<th>RMSE</th>
<th>NME</th>
<th>NMB</th>
<th>MB</th>
<th>Cor</th>
<th>SD Mod.</th>
<th>SD Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>CRU</td>
<td>2.67</td>
<td>2.37</td>
<td>0.28</td>
<td>-1.26</td>
<td>4.29</td>
<td>1.42</td>
<td>0.87</td>
</tr>
<tr>
<td>2006</td>
<td>UDELAWARE2.67</td>
<td>2.49</td>
<td>0.35</td>
<td>-1.42</td>
<td>4.89</td>
<td>1.66</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>CRU</td>
<td>2.7</td>
<td>2.66</td>
<td>-0.04</td>
<td>-1.85</td>
<td>4.25</td>
<td>1.52</td>
<td>0.87</td>
</tr>
<tr>
<td>2007</td>
<td>UDELAWARE2.7</td>
<td>2.85</td>
<td>0.15</td>
<td>-1.5</td>
<td>5.54</td>
<td>1.9</td>
<td>0.79</td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td>CRU</td>
<td>2.66</td>
<td>2.69</td>
<td>0.23</td>
<td>-0.73</td>
<td>4.1</td>
<td>1.45</td>
<td>0.87</td>
</tr>
<tr>
<td>2008</td>
<td>UDELAWARE2.66</td>
<td>2.83</td>
<td>0.02</td>
<td>-1.9</td>
<td>5.5</td>
<td>1.9</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td>CRU</td>
<td>2.65</td>
<td>2.48</td>
<td>0</td>
<td>-1.69</td>
<td>4.1</td>
<td>1.38</td>
<td>0.89</td>
</tr>
<tr>
<td>2009</td>
<td>UDELAWARE2.65</td>
<td>2.72</td>
<td>0.27</td>
<td>-1.02</td>
<td>5</td>
<td>1.7</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>CRU</td>
<td>2.65</td>
<td>2.61</td>
<td>-0.31</td>
<td>-2.3</td>
<td>4.3</td>
<td>1.47</td>
<td>0.88</td>
</tr>
<tr>
<td>2010</td>
<td>UDELAWARE2.65</td>
<td>2.76</td>
<td>-0.2</td>
<td>-2.23</td>
<td>5.07</td>
<td>1.76</td>
<td>0.83</td>
<td></td>
</tr>
</tbody>
</table>
Figure C.3 Average precipitation for the Short Rains season for WRF (left column), and TRMM (right column) for Domain 1.
Table C.9 Summary of Evaluation Statistics for WRF precipitation domain 1 during the Long Rains. RMSE is the root mean square error, NME is the normalized mean error, NMB is the normalized mean bias, MB- mean bias, Cor is the pattern correlation, NRMSE is the normalized root mean square SD Mod.- standard deviation calculated from model output and SD Obs. - standard deviation for observations

<table>
<thead>
<tr>
<th>Year</th>
<th>OBS.</th>
<th>RMSE</th>
<th>NME</th>
<th>NMB</th>
<th>MB</th>
<th>Cor</th>
<th>SD Mod.</th>
<th>SD Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>CMORPH</td>
<td>339</td>
<td>267</td>
<td>37</td>
<td>13</td>
<td>62</td>
<td>269</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>339</td>
<td>200</td>
<td>64</td>
<td>25</td>
<td>44</td>
<td>190</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>GPCP</td>
<td>339</td>
<td>199</td>
<td>59</td>
<td>5.7</td>
<td>43.5</td>
<td>192</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>339</td>
<td>199</td>
<td>59</td>
<td>5.7</td>
<td>43.5</td>
<td>152</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>339</td>
<td>195</td>
<td>76</td>
<td>31</td>
<td>59</td>
<td>239</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>339</td>
<td>208</td>
<td>71</td>
<td>27</td>
<td>60</td>
<td>266</td>
<td>0.66</td>
</tr>
<tr>
<td>2007</td>
<td>CMORPH</td>
<td>363</td>
<td>282</td>
<td>26</td>
<td>8.7</td>
<td>57.7</td>
<td>251</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>363</td>
<td>177</td>
<td>85</td>
<td>36</td>
<td>44</td>
<td>176</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>GPCP</td>
<td>363</td>
<td>207</td>
<td>79</td>
<td>2.7</td>
<td>41</td>
<td>132</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>363</td>
<td>207</td>
<td>78.53</td>
<td>2.71</td>
<td>42</td>
<td>132</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>363</td>
<td>200</td>
<td>76</td>
<td>31</td>
<td>55</td>
<td>218</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>363</td>
<td>221</td>
<td>68</td>
<td>23</td>
<td>56</td>
<td>279</td>
<td>0.67</td>
</tr>
<tr>
<td>2008</td>
<td>CMORPH</td>
<td>361</td>
<td>291</td>
<td>61</td>
<td>20</td>
<td>70</td>
<td>316</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>361</td>
<td>181</td>
<td>118</td>
<td>48</td>
<td>45</td>
<td>192</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>GPCP</td>
<td>361</td>
<td>189</td>
<td>127</td>
<td>17.6</td>
<td>49</td>
<td>160</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>361</td>
<td>190</td>
<td>127.26</td>
<td>18</td>
<td>49</td>
<td>203</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>361</td>
<td>164</td>
<td>142</td>
<td>65</td>
<td>82</td>
<td>293</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>361</td>
<td>203</td>
<td>111</td>
<td>44</td>
<td>69</td>
<td>300</td>
<td>0.64</td>
</tr>
<tr>
<td>2009</td>
<td>CMORPH</td>
<td>378</td>
<td>285</td>
<td>66</td>
<td>23</td>
<td>73</td>
<td>333</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>378</td>
<td>205</td>
<td>112</td>
<td>46</td>
<td>39.6</td>
<td>173</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>GPCP</td>
<td>378</td>
<td>207</td>
<td>116</td>
<td>13</td>
<td>50.8</td>
<td>157</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>378</td>
<td>207</td>
<td>116</td>
<td>13</td>
<td>50</td>
<td>156</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>378</td>
<td>184</td>
<td>127</td>
<td>57</td>
<td>76</td>
<td>302</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>378</td>
<td>219</td>
<td>104</td>
<td>43</td>
<td>71</td>
<td>315</td>
<td>0.62</td>
</tr>
<tr>
<td>2010</td>
<td>CMORPH</td>
<td>360</td>
<td>267</td>
<td>67</td>
<td>22</td>
<td>69</td>
<td>329</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>CRU</td>
<td>360</td>
<td>188</td>
<td>134</td>
<td>55</td>
<td>49</td>
<td>205</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>GPCP</td>
<td>360</td>
<td>178</td>
<td>85</td>
<td>38</td>
<td>60</td>
<td>177</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>GPCC</td>
<td>360</td>
<td>178</td>
<td>154</td>
<td>38</td>
<td>60</td>
<td>176</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>360</td>
<td>169</td>
<td>140</td>
<td>59</td>
<td>77</td>
<td>304</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>TRMM</td>
<td>360</td>
<td>199</td>
<td>126.65</td>
<td>50.82</td>
<td>52</td>
<td>297</td>
<td>0.68</td>
</tr>
</tbody>
</table>
Figure C.4 Average precipitation for the Long Rains season for WRF (left column), and TRMM (right column) for Domain 1