

RAINFALL VARIABILITY AND ESTIMATION FOR HYDROLOGIC MODELING

A REMOTE SENSING BASED STUDY AT THE SOURCE
BASIN OF THE UPPER BLUE NILE RIVER



ITC dissertation number 166
ITC, P.Box 6, 7500 AA Enschede, the Netherlands

The research described in this thesis was funded by the Netherlands organization for international cooperation in higher education. The research was undertaken at the faculty of Geo-information science and earth observation (ITC), University of Twente, Enschede, The Netherlands.

Cover: The pictures are taken in the Upper Blue Nile basin
Copyright © 2010 by Alemseged Tamiru Haile, Enschede, the Netherlands
Printed by: ITC printing department, Enschede, the Netherlands
ISBN 978-90-6164-286-8

RAINFALL VARIABILITY AND ESTIMATION FOR HYDROLOGIC MODELING

A REMOTE SENSING BASED STUDY AT THE SOURCE BASIN OF
THE UPPER BLUE NILE RIVER

DISSERTATION

to obtain
the degree of doctor at the University of Twente,
on the authority of the rector magnificus,
prof.dr. H. Brinksma,
on account of the decision of the graduation committee,
to be publicly defended
on Thursday 25 February 2010 at 15:00

by

Alemseged Tamiru Haile
born on 28 February 1978
in Addis Ababa, Ethiopia

This dissertation has been approved by:

Prof. dr. V.G. (Victor) Jetten

Dr.ing. T.H.M. (Tom) Rientjes

promotor

assistant-promotor

ABSTRACT

Rainfall is one of the meteorological forcing terms in hydrologic modelling and therefore its spatial variability in coverage, frequency and intensity affects simulation results. Rainfall variability in particular under the effect of orography adjacent to a large water body is not fully explored. Such study is done for the Gilgel Abbay watershed of the Lake Tana basin (Ethiopia). The study area is the source basin of the Upper Blue Nile River which is one of the major contributors to the River Nile. The livelihood in the Lake Tana basin largely depends on rainfed agriculture and therefore understanding rainfall variability in the basin is required. As part of the study, a set of recording rain gauges have been installed to observe rainfall at high resolution.

First, rainfall variability in the Lake Tana basin is evaluated by statistical analysis of rain gauge observations. Furthermore, a convective index is derived from remote sensing observations to infer the pattern of rainfall variability in the basin. Results suggest that orography and the presence of Lake Tana largely affect the diurnal cycle, frequency and intra- and inter-event properties of the rainfall. The rainfall varies significantly at scales much smaller than inter-station distances suggesting that the existing rain gauge network may be inadequate to fully capture the space-time pattern of the rainfall. Such affects the accuracy of spatial rainfall estimation that serves to specify the input to hydrologic models.

Second, two remote sensing based approaches have been developed to estimate spatial rainfall: (i) a multi-spectral remote sensing approach, and (ii) a conceptual cloud model approach with inputs from remote sensing and typical ground based observations (pressure and temperature). Results show the potential of remote sensing observations for rainfall estimation although the ground based data still provided some limitations at this point.

Third, the effect of the rainfall variability on the accuracy of the simulated stream flows by a physically based rainfall-runoff model is evaluated. The effect of rain gauge density and configuration on rainfall representation and consequently on stream flow simulation is evaluated through a set of performance measures. The large rainfall variability in the study area caused the accuracy of the simulated flow to be significantly affected by both the density and the configuration of the network. The use of rainfall from a single rain gauge resulted in a relative difference of up to 100 % between the simulated and observed stream flows. It is also shown that simulated stream flow largely differs if uniform rainfall input is compared to non-uniform rainfall input. This study is relevant to

hydrologic modeling since much research has focused on model development and assessing parameter uncertainty while less attention is given to aspects that relate to effects of rainfall representation.

ACKNOWLEDGEMENTS

This research would have not been a lot less easy without the help and support of many people. Therefore, it is a great pleasure for me to acknowledge those who contributed to the success of my Ph.D. study.

First, I would like to thank my assistant-promotor Tom Rientjes for not only critically reading my thesis but also for the interesting discussion we always had on both academic as well as non-academic matters. I am indebted to him for useful suggestions and encouragements. His comments were very instrumental to improve my scientific writing skills. Tom gave me a lot of freedom to set the objectives of my thesis. I hope that our relation will continue in the future.

I would like to thank my Promotor Prof. Victor Jetten for his willingness to take me as his Ph.D. student during a late stage and a difficult period of my Ph.D. He encouraged me a lot to complete the thesis in time. I always enjoyed our discussion and appreciated the way he interacts with students.

My acknowledgement with high appreciation goes to Ambro Gieski, Paolo Regianni and Mekonnen Gebremichael who provided myriad comments and useful suggestions. I highly appreciate their contribution. Special thanks to Paolo for providing me access to use the REW model and for his helpful hints about the use of the model.

Thanks to the staffs of the Water Resources department of ITC. Ir. Arno van Lieshout, Ms. Anke Walet, Ms. Tina Butt-Castro and Ms. Loes Colenbrander have always been very willing to provide me support letters and administrative information. I highly appreciate their support. Arno was also supportive for the success of my field work. The ITC librarians were very helpful in many ways and particularly in helping me to access several references that were not available in the ITC library.

Thanks to my office mates and the Ethiopian community in Enschede for the wonderful time we have together.

My acknowledgement goes to the Netherlands organization for international cooperation in higher education (Nuffic) that funded this research. I would like to thank those who facilitated the deployment and installation of the rain gages that provided the data for this research: the Ethiopian embassy in Brussels, the Bahir Dar head office of the Ethiopian meteorological agency and the primary schools in the Lake Tana basin. I thank the Ethiopian Meteorological Agency and the hydrological and GIS departments of the Ministry of Water Resources of Ethiopia for providing me the meteorological, hydrologic and GIS data of the study area. I also

thank the European organization for the exploitation of meteorological satellites (EUMETSAT) for making available the meteosat images free of charge.

I am grateful to my parents, brothers and sister for their unwavering support and prayers. Time at ITC would have been a lot slower had it not been for my fiancé Elsi who showed me unwavering support and accompanied me in Enschede for 18 months.

CONTENTS

List of figures	ix
List of tables.....	xiv
1 INTRODUCTION.....	1
1.1. Background	2
1.2. Research objectives.....	5
1.3. Structure of the Thesis	6
2 THE LAKE TANA BASIN.....	9
2.1. Description of the basin.....	10
2.1.1. Temperature	11
2.1.2. Rainfall variation	14
2.2. Summary	20
3 RAINFALL VARIABILITY USING GROUND BASED AND REMOTE SENSING OBSERVATIONS.....	23
ABSTRACT	24
3.1. Introduction	25
3.2. Data Sets	26
3.2.1. Rainfall	26
3.2.2. Cloud infrared temperature	28
3.3. Method of Analysis.....	29
3.3.1. Spatial variability	30
3.3.2. Spatial correlation structure	30
3.3.3. Diurnal variability	31
3.3.4. Fractals of rainfall intermittence.....	32
3.4. Results.....	33
3.4.1. Spatial variability.....	33
3.4.2. Spatial correlation structure	37
3.4.3. Diurnal cycle.....	40
3.4.4. Fractals of rainfall intermittence.....	54
3.5. Discussion	55
3.6. Conclusion.....	58
4 RAIN EVENT PROPERTIES.....	61

ABSTRACT	62
4.1. Introduction	63
4.2. Method of analysis	66
4.2.1. Properties of rain events	66
4.2.2. Dimensionless event hyetographs.....	67
4.2.3. Conditional probability of rainfall occurrences.....	69
4.3. Results.....	70
4.3.1. Rain event properties	70
4.3.2. Relation between rain event properties.....	72
4.3.3. Temporal variation of rain event properties.....	72
4.3.4. Spatial variation of rain event properties.....	74
4.3.5. Dimensionless event hyetographs.....	82
4.3.6. Conditional probability of rainfall occurrences.....	88
4.4. Discussion and conclusion.....	89
5 REMOTE SENSING BASED RAINFALL DETECTION AND ESTIMATION	95
ABSTRACT	96
5.1. Introduction	97
5.2. Data Sets	98
5.2.1. Remote sensing observations	98
5.2.2. Ground based observations.....	101
5.3. Method of Analysis	101
5.3.1. Rainfall detection	101
5.3.2. Rainfall estimation.....	103
5.3.3. Performance measures	103
5.4. Results.....	105
5.4.1. Rainfall detection	106
5.4.2. Rainfall estimation.....	109
5.4.3. Comparison against rain gauge observations	110
5.5. Conclusion.....	111
6 REMOTE SENSING BASED CONCEPTUAL CLOUD MODELLING FOR RAINFALL SIMULATION	115
ABSTRACT	116
6.1. Introduction	117
6.2. Model Approach	119

6.2.1. Model equations	119
6.2.2. Model inputs	124
6.2.3. Model parameterization	125
6.3. Regional sensitivity analysis.....	132
6.4. Results.....	134
6.4.1. Data analysis.....	134
6.4.2. Model sensitivity.....	135
6.4.3. Model calibration.....	137
6.5. Discussion and conclusion.....	142
APPENDIX: LIST OF SYMBOLS.....	147
7 SENSITIVITY OF THE REPRESENTATIVE ELEMENTARY WATERSHED MODEL TO RAINFALL REPRESENTATION.....	151
ABSTRACT	152
7.1. Introduction	153
7.2. Study Area.....	155
7.3. Methods	157
7.3.1. The REW model	157
7.3.2. Model calibration.....	158
7.3.3. Rainfall representation.....	159
7.3.4. Sensitivity to network density and configuration.....	162
7.4. Results.....	163
7.4.1. REW model calibration	163
7.4.2. Effects of the gauge network configuration and density on rainfall and runoff estimation	165
7.4.3. REW model sensitivity to network configuration and density	168
7.4.4. Effect of model resolution	172
7.4.5. Model sensitivity to rainfall variability	173
7.4.6. Effects of rainfall interpolation	175
7.5. Discussion and Conclusion.....	176
8 SUMMARY, CONCLUSIONS AND RECOMMENDATIONS.....	181
8.1. Summary and conclusions	182
8.1.1. Rainfall variability	182
8.1.2. Rain event properties	185
8.1.3. Remote sensing for rainfall detection and estimation	187

8.1.4. Remote sensing based conceptual cloud modeling for rainfall estimation	188
8.1.5. Sensitivity of the REW model to rainfall	189
8.2. Recommendations.....	191
REFERENCES	193

List of Figures

Figure 2.1: <i>The topography of the Lake Tana basin and its major watersheds. The boundaries are represented by white lines.</i>	10
Figure 2.2: <i>Inter-annual variability of the daily minimum temperature at four stations in the Lake Tana basin (time period 1994 – 2003). The 95 % confidence intervals of the mean values are also shown. Note: the first month corresponds to January.</i>	12
Figure 2.3: <i>Inter-annual variability of the daily maximum temperature at four selected stations in the Lake Tana basin (time period 1994-2003). The 95 % confidence intervals of the means are shown.</i>	13
Figure 2.4: <i>The diurnal cycle of temperature at Durbet weather station in Gilgel Abbay watershed.</i>	14
Figure 2.5: <i>Estimated and fitted Spherical semivarioagram for the annual and JJA rainfall of the Lake Tana basin.</i>	15
Figure 2.6: <i>The long-term mean annual and seasonal (JJA) rainfall of the Lake Tana basin. Fraction of rainfall represents the ratio of the JJA rainfall over the annual rainfall. Note: figure 2.6d shows the stations that are shown by the numbers while the name of the stations is presented in table 2.1.</i>	16
Figure 2.7: <i>Box plots of the daily rainfall at four stations in the Lake Tana basin.</i>	19
Figure 3.1: <i>Digital Elevation Model of Lake Tana basin with watershed boundaries of the major rivers and location and names of rain gauges as indicated by numbers. Transect A-A' is described in section 3.5.</i>	27
Figure 3.2: <i>Rainfall observations at Jema for the time period June 1 to August 31, 2007. Note the Date is in month/day format.</i>	29
Figure 3.3: <i>Statistics of hourly rainfall in June-August, 2007. In the diagrams, a line is added that shows the main river that separates the east and west sides of Gilgel Abbay watershed. Note: for the mean rainfall, the conditioning was done on non-zero rainfall values.</i>	36
Figure 3.4: <i>Exceedance probability of the hourly rainfall (a) and (b), and mean rainfall conditioned on specific rainfall depth (c).</i>	37
Figure 3.5: <i>Spatial correlation structure of the rainfall observations (a) and (c), and values of correlation distance (d_0) for several assumed</i>	

values of correlation at zero distance (c_0) (b) and (d). Note: For the daily rainfall, the correlation coefficient and its averaged values are shown in small hollow and large solid circles, respectively.....	39
Figure 3.6: Cross-correlation of the hourly rainfall.....	40
Figure 3.7: Hourly rainfall variation (a) and comparison between the maximum hourly and 3-month accumulated hourly rainfall (b) in JJA at Jema station.	42
Figure 3.8: Rainfall diurnal cycle at selected four stations.	43
Figure 3.9: Nocturnal (2100-0900 LST) rainfall distribution as a percentage of the seasonal rainfall.	44
Figure 3.10: Scatter plot of catchment-average mean (a) and standard deviation (b) of the CI and the frequency of rainfall occurrence averaged over each hour (expressed in LST) in JJA, 2007.	46
Figure 3.11: Diurnal variation of the convective index (CI).....	47
Figure 3.12: Variations of elevation and CI along transect A–A' which is shown in figure 3.1. Distance is measured from the Lake near Bahir Dar.	48
Figure 3.13: Convective index over Lake Tana, its basin area and the shores of the lake.	49
Figure 3.14: The maximum (peak) CI and Local time to this peak CI for various temperature thresholds. The maximum CI values are divided by the maximum values to obtain the normalized values.....	50
Figure 3.15: Results after applying a harmonic analysis to the rainfall frequency and the CI. Results are only for 5 rain gauges, the complete results are given in table 3.6.....	52
Figure 3.16: Log-log plot of Number of boxes with rain ($N(\lambda)$) against the size of boxes, i.e. the scale factor (λ). Note: The slope of the second scaling regime is 1.0 for all the stations.....	56
Figure 4.1: Box plot of the rain event depths in JJA of 2007 as observed at eight rain gauge stations. The lower and the upper bars indicate the 25 % quartile minus 1.5 IQR and the 75 % quartile plus 1.5 IQR, respectively. IQR refers to the interquartile range which is defined by the height of the box. The bar inside the boxes shows the median while the upper and lower edges of the boxes show the 75 % and the 25 % quartiles.	75
Figure 4.2: Box plot of the rain event durations in JJA of 2007 as observed at eight rain gauge stations.....	77

Figure 4.3: Box plot of average rain event intensity in JJA of 2007 as observed at eight rain gauge stations.....	78
Figure 4.4: Cumulative probability of inter-event time in JJA of 2007 as observed at three rain gauge stations.....	79
Figure 4.5: Relation between event properties and terrain elevation. The event properties are (a) Event depth, (b) Event duration, (c) Event intensity and (d) Inter-event time (IET). The median of each of the event properties is selected for the regression and the regression equations are shown in Table 4.4.....	82
Figure 4.6: The observed dimensionless hyetograph of the rain events at two stations for three quantiles.....	83
Figure 4.7: The observed and the modelled dimensionless hyetograph of the rain events at Bahir Dar station for three quantiles.	85
Figure 4.8: The observed and the modelled dimensionless hyetograph of the rain events at Jema station for three quantiles.....	86
Figure 4.9: The observed and the modelled dimensionless hyetograph of the rain events at Sekela station for three quantiles.	87
Figure 4.10: Conditional probability of rainfall occurrences against inter-station distance. The exponential model is fitted with an R2 value of 0.86 and 0.63 for the 1 hour and the 6 hour rainfall, respectively.....	90
Figure 5.1: Comparison of $T_{10.8}$ and TRMM PR rain rates for July 12, 2007 at 10:45. (a) and (b) show the original images while (c) shows the collocated (matched) image and (d) shows relation between $T_{10.8}$ and rainfall rate after collocation by the principle that the highest rainfall rate corresponds to the lowest $T_{10.8}$. The hollow circles in (a) and (b) indicate the space shift between the highest convective rainfall locations from the two images. The map scale is 1:500,000.....	100
Figure 5.2: Categorical statistics for rainfall detection using brightness temperature ($T_{10.8}$), rate of temperature change ($\Delta T_{10.8}$ (here Delta $T_{10.8}$)), gradient of ($\nabla T_{10.8}$ (here Grad. $T_{10.8}$)) and brightness temperature difference ($T_{10.8} - T_{6.2}$). NB: The rainfall rate threshold is 1.0 mm h^{-1} . The scales of the axis of the categorical statistics are not equal for the selected indices.....	107
Figure 5.3: (a) Relation between estimated mean rainfall rate and brightness temperature ($T_{10.8}$). The 95 % confidence interval of the mean rainfall	

rate is indicated by the vertical bars. Note: T is defined as $T_{10.8}$ minus the threshold T_t which is set to 180 K. (b) Relation between standard error of mean rainfall rate for each 1 K temperature interval and the number of data pairs of TRMM PR and MSG-2 data used to estimate the mean.... 110

Figure 5.4: (a) POD and (b) bias of the Thermal Infrared (TIR) based exponential model and inverse distance weighting (IDW) rainfall estimates; (c) shows the plot of the estimated rainfall by both methods against the reference rainfall as obtained from the rain gauges..... 112

Figure 6.1: Schematization of the two layer model. T and p are temperature and pressure while subscripts 0, b, i and t indicate the temperature and the pressure at the ground surface, the cloud bottom, the interface of the layers and the cloud top surfaces, respectively. 123

Figure 6.2: Rain generation time at different pressure levels of a cloud layer. 131

Figure 6.3: Cumulative normalized (Cum. Norm.) distribution of the parameters for abias. Note: Black lines represent parameter values with high model performance (bin 1) while brighter lines represent parameter values that give lower model performance (bin 2 – 10). 136

Figure 6.4: Cumulative normalized (Cum. Norm.) distribution of the parameters for M4E. Note: Dark lines represent parameter values with high model performance (bin 1) while brighter lines represent parameter values that give low model performance (bin 2 – 10). 137

Figure 6.5: Observed and simulated rainfall intensity of four rain events. ... 140

Figure 6.6: Observed rainfall and estimated updraft velocity for July 22, 2007 event. Note that at first, the updraft increases rapidly but levels off after some time period. 141

Figure 7.1: The rain gauge network and the delineated representative elementary watersheds (REWs) of the Upper Gilgel Abbay watershed that drains to Lake Tana. 156

Figure 7.2: Observed and simulated flow at the watershed outlet. The watershed average rainfall is shown at the top. Note: the calibration period range between 732 – 1431, the first evaluation period ranges between 1 – 731 and the second evaluation period ranges between 1432 – 2192. 164

Figure 7.3: Performance of spatial rainfall estimation using specific number

of rain gauge stations with respect to the reference area-average rainfall. The lower and the upper bars indicate the 25 % quartile minus 1.5 IQR and the 75 % quartile plus 1.5 IQR, respectively where IQR is the interquartile range which is the height of the box. The middle bar of the boxes shows the median while the upper and lower bars of the boxes show the 75 % and the 25 % quartiles.167

Figure 7.4: *Relative difference (see equation 12) between the simulated stream flow and the reference stream flow for a specific number of rain gauge stations. Note: The rain gauge configuration for respective rain gauge numbers that produced the maximum GORE index is selected....170*

Figure 7.5: *Effect of removing Durbet station (a) and Sekela station (b) on the simulated flow as compared to the reference flow. Durbet is located in a valley while Sekela is located on a mountain.172*

Figure 7.6: *Simulated flow for (a) rainfall with maximum depth near the watershed outlet (at Durbet) and (b) on the mountain (at Sekela). Note: The time is measured since the start of simulation of the event.176*

List of Tables

Table 2.1: <i>Spatial features of the rainfall stations. Note: The Coordinate System projection is Universal Transverse Mercator, Datum Adindan, Clark Ellipsoid (1880)</i>	17
Table 2.2: <i>The median value of the daily rainfall at four stations in the Lake Tana basin</i>	20
Table 3.1: <i>Characteristics of the location of the rain gauge stations: Note: The Coordinate System projection is Universal Transverse Mercator, Datum Adindan, Clark Ellipsoid (1880)</i>	28
Table 3.2: <i>Statistics of the hourly rainfall observations</i>	34
Table 3.3: <i>Correlation between number of rain hours and elevation. Note: Distance to the stations is measured from the centre of Lake Tana</i>	38
Table 3.4: <i>Rainfall depth (% of the seasonal rainfall) in JJA, 2007</i>	43
Table 3.5: <i>Results from the harmonic analysis</i>	53
Table 3.6: <i>Fitted parameters of the harmonic analysis. Note: phase angle is interpreted and presented as time to the amplitude of the diurnal and semidiurnal cycles</i>	54
Table 4.1: <i>Statistics of rain events at Jema in two wet seasons that are JJA of the years 2007 and 2008</i>	71
Table 4.2: <i>Lower triangle of correlation matrix between the properties of the selected rainfall events at Jema. The events are observed in two wet seasons that are JJA 2007 and 2008</i>	72
Table 4.3: <i>The temporal variation of the median of rain event properties at Jema for MIT = 30 min. The events are observed in two wet seasons that are JJA 2007 and 2008</i>	73
Table 4.4: <i>Relation between rain event properties and terrain elevation (Elev.)</i>	81
Table 4.5: <i>Absolute difference between two dimensionless hyetographs (Dmax) for Kolmogorov-Smirnov test. Note $D_{0.05,10} = 0.4092$</i>	84
Table 4.6: <i>Absolute difference between observed and simulated dimensionless hyetographs (Dmax) for Kolmogorov-Smirnov test. Note $D_{0.05,10} = 0.4092$</i>	87

Table 4.7: Parameters of the beta distribution model that is fitted to the dimensionless hyetographs.	88
Table 5.1: Contingency table	104
Table 5.2: Evaluation of rainfall detection using indices from MSG-2 images	108
Table 6.1: Autocorrelation of the model inputs for the July 22, 2007 event. The data are observed between 12:00–18:00 Local Standard Time. Note: Ground–station data is available every 30 minutes while remote sensing data is available every 15 minutes.	134
Table 6.2: The 1-hour lag cross-correlation of the model inputs for the calibration period (July, 22, 2007) and the validation period (August 16, 2007). The data are observed between 12:00–18:00 Local Standard Time.	135
Table 6.3: Model parameter values of the two layer model.	139
Table 6.4: Observed and simulated characteristics of the four rain events. A negative time shift indicates a delay in model response as compared to observations.	141
Table 6.5: The relation between rainfall initiation time and the time to maximum updraft and cloud formation is observed for the four events. .	142
Table 6.6: Best 10 performing parameter sets in terms of M4E	144
Table 6.7: Best 10 performing parameter sets in terms of abias	144
Table 7.1: Model efficiency for the calibration and validation data sets.....	165
Table 7.2: Statistics of the GORE index for estimation of the spatial rainfall for events with intensity of higher than 5 mm h^{-1}	168
Table 7.3: Sensitivity for spatially averaged rainfall that is estimated by a specific number of stations. Note: the stations with maximum GORE were selected.....	169
Table 7.4: Effect of removing a specific station on the performances of the REW model: The name of the removed station and its elevation are shown on the top of each column.	171
Table 7.5: Comparison of sensitivity to rainfall of the REW model using a second and third Strahler order discretization for the time period June 1 to June 9, 2007.	173
Table 7.6: Correlation matrix for rainfall characteristics and model	

*performance measures for simulated stream flow with a uniform rainfall
input with the non-uniform rainfall input for a reference flow.....174*

**Table 7.7: Effects of interpolation on REW model outputs. Note: the volume
of the simulated stream flow is divided by the size of the
watershed area.....176**

1

INTRODUCTION



1.1. BACKGROUND

SCIENTIFIC RELEVANCE

Rainfall is one of the meteorological forcing terms of hydrological systems and understanding its variability in coverage, magnitude and frequency is of key importance. Rainfall variability is caused by various factors that, for instance, relate to atmospheric processes, effects of orography and large water bodies such as lakes (Buytaert et al., 2006; Barros et al., 2004; Ba and Nicholson, 1998; Johnson and Hanson, 1995). The effect that such factors have on rainfall variability differs from region to region which makes rainfall variability difficult to predict.

Rainfall variability is commonly affected by orography that influences the arrival directions of wet air masses. The variation in daily, monthly and annual rainfall as caused by factors such as terrain elevation, slope and aspect is relatively well explored (e.g. Buytaert et al., 2006; Johnson and Hanson, 1995; Basist et al., 1994) and some generic understanding has been gained. Studies show that rainfall amounts that are accumulated at daily to annual time scales generally increase with terrain elevation but also rainfall is affected by terrain slope and aspect with wind and lee side effects. However, variability of sub-daily rainfall as caused by these factors is not well explored (e.g. Allamano et al., 2009; Loukas and Quick, 1996). For instance, the effect of orography on the properties of rain events such as duration, depth, and intensity, and the length of the time period between two consecutive events, i.e. inter-event time, is not well explored. This is particular for the tropical areas of the African continent that generally have poor observation networks.

Rainfall studies are commonly restricted by availability of time series observations. Although radar provides spatial coverage of rainfall, such data is commonly not available since it is considered too expensive. Also, the spatial coverage of radar is constrained by high mountain ranges. Rain gauges serve as the main source of rainfall data in many regions. Rain gauges record rainfall data at discrete points in space and time while rain gauges are often sparsely and unevenly distributed in space. For rainfall monitoring, World Meteorological Organization (WMO) recommends an average rain gauge inter-station distance of 25 – 30 km in flat areas and

approximately half that distance in mountainous areas (see Gandin, 1970). Such requirement, however, is rarely met in practice.

In practice, the density and configuration of a rain gauge network is determined based on the availability of funds, accessibility of site and the purpose of the network. Rain gauges in developing countries are commonly installed in towns that are located along main roads that provide accessibility. As a result, relatively inaccessible areas such as mountainous areas may remain uncovered by the observation network.

As part of an effort to overcome rainfall data unavailability, various model approaches are proposed in literature to synthetically generate rainfall data. Examples include models that are based on dimensionless event hyetographs (e.g. Garcia-Gazman and Aranda-Olivier, 1993; Huff, 1967) and models that are based on the scaling properties of rainfall (e.g. Olsson and Berndtsson, 1998; Over and Gupta, 1994). The calibration and validation of these approaches commonly rely on rainfall data at high, i.e. sub-daily, resolutions.

A second alternative to overcome unavailability of rainfall data is the use of remote sensing. Observations by remote sensing provide spatial coverage in contrast to the point observations by single rain gauge. Additional advantage of remote sensing in rainfall studies is that the observations are commonly available free of charge. Such advantages have triggered researchers to develop satellite remote sensing based methods of rainfall estimation.

Since the late 1960s, a plethora of satellite remote sensing based rainfall estimation methods has been developed to meet the demands by various applications for rainfall data. Reviews of the methods are presented by Stephens and Kummerow (2007); Levizzani et al. (2002); Barrett and Martin (1991); Kidder and Vonder Haar (1995); Petty (1995). Reviews revealed that the methods in general have poor performance. As stated by Barrett and Beaumont (1994), satellite sensors view from the top of the atmosphere downwards to the land surface but not upwards from the land surface that would be a more logical approach for rainfall observation. As such, through satellite remote sensing only “proxy” variables of rainfall are observed. Those variables which include cloud reflectance and cloud top temperature only have a weak and indirect relation to surface rainfall rate.

A conceptual cloud model that integrates remote sensing and ground based observations to simulate rainfall is proposed by Georgakakos

and Bras (1984a,b). That model is evaluated and extended using radar observations by Andrieu et al. (2003); Bell and Moore (2000b); French and Krajewski (1994) and French et al. (1994). Advantages of the approach are the model structure that may be considered simple and the use of readily available observations as model inputs. In the cloud model, the vertical profile of a cloud system was represented by a single model layer and the approach showed to have some limitations. Increasing the number of the model layers is considered necessary to account for the differences in rain generation time along the vertical profile of a cloud system. Georgakakos and Bras (1984a,b) proposed the conceptual model at first to overcome restrictions by rain gauge data availability. It is noted the model has not been evaluated for regions where radar data is unavailable and where the ground based observation network is poor.

Despite the poor performance of remote sensing based methods of rainfall estimation, remote sensing has found applications in rainfall studies mainly since it provides a spatial coverage of rain producing clouds. Some applications of remote sensing include: rainfall detection to study the diurnal cycle of rainfall (Dai 2001; Imaoka and Spencer, 2000), to monitor and characterize clouds that produce heavy rainfall (Curtis et al. 2007; Feidas, 2003; Laing et al., 1999) and to analyze the scaling behaviour of rainfall (Gebremichael et al., 2008; Nykanen, 2008).

Studies of rainfall variability and estimation commonly are restricted by data availability. Rainfall estimation becomes challenging for areas where the rainfall largely varies over small time and space domain while the observation network is sparse. In such cases, the accuracy of the estimated rainfall may deteriorate with consequences for its further use.

Studies by Segond et al., (2007); Bell and Moore (2000a) revealed that rainfall variability can largely affect hydrological model outputs while Chaubey et al., (1999); Younger et al., (2009) showed that rainfall variability affects estimation of model parameters. Beven and Hornberger (1982) noted that spatial variability of rainfall mainly affects the timing of peak runoff. However, these studies were carried out in regions where either the rainfall is relatively uniform or the rain gauge network is dense. Where data availability is poor, studying the effect of rainfall variability on simulation results of hydrological models can be overcome by using synthetically generated rainfall and runoff data (e.g. Arnaud et al., 2002; Obled et al., 1994).

SOCIETAL RELEVANCE

Lake Tana is the largest lake in Ethiopia and is considered the source of the Upper Blue Nile River which contributes approximately 50 % to the stream flow of the Nile River. According to Conway (1997), the outflow from Lake Tana contributes 8 % to the Upper Blue Nile flow in Ethiopia.

The outflow from Lake Tana contributes to the Tis Abbay I and Tis Abbay II hydropower plants. There is an ongoing construction of a 12 km tunnel to transfer water from Lake Tana to the Beles basin of the Upper Blue Nile River. In terms of irrigation, 6 dams are proposed in the Lake Tana basin (see, BCOEM, 1999). The Koga dam is nearly completed while the construction of Megech and Ribb dams is about to be commenced but the remaining dams are only at a pre-feasibility study stage. In addition to irrigation and hydropower purposes, Lake Tana serves for fishing, navigation and tourism. The reader is referred to SMEC (2007) for some information about the possible impacts of the various water resources projects in the Lake Tana basin.

Despite its major socio-economic importance, the hydrology of the Lake Tana basin is not well documented in literature. The lake has a large surface area, i.e. approximately 3100 km² and therefore receives large volumes of rainfall that may significantly affect the water balance of the lake. However, the lake rainfall has not been monitored and as a result hydrological studies extrapolate rainfall from inland stations (e.g. Wale et al., 2009; Kebede et al. 2006). Such studies can benefit from understanding the spatio-temporal variability of the rainfall in the basin.

1.2. RESEARCH OBJECTIVES

The main objective in this study is to evaluate the spatio-temporal pattern of rainfall variability and its effect on runoff. Sub-objectives of the study are:

- To evaluate the spatio-temporal variability of rainfall as affected by orography and the presence of a lake.
- To develop and evaluate a remote sensing based approach for rainfall detection and estimation.

- To develop and evaluate a conceptual cloud model for rainfall simulation with remote sensing and ground based observations as model inputs.
- To evaluate the performance of a physically based rainfall-runoff model and its sensitivity to a rain gauge network density and configuration.

As part of this study, a network of 10 recording rain gauges and 1 automatic weather station has been installed in May 2007. The location and terrain attributes of the network are shown in figure 3.1 and table 3.1 in Chapter 3 of this thesis.

By the objectives of this study, the data from the rain gauges is used to study the rainfall diurnal cycle, the spatio-temporal structure of rainfall and rain event properties. The rain gauge observations are also used to evaluate the remote sensing based rainfall estimates and to assess the sensitivity of a rainfall-runoff model to rain gauge network density and configuration.

Remote sensing observations are obtained from the Meteosat Second Generation (MSG-2) and the Tropical Rainfall Measuring Mission (TRMM) satellites. Thermal infrared (TIR) observations from MSG-2 are used to analyze patterns in rainfall diurnal cycle. In this thesis, two parsimonious rainfall model approaches are developed and evaluated to estimate rainfall from remote sensing observations. Surface rainfall rates from Precipitation Radar (PR) of TRMM satellite served as the ground truth to calibrate the remote sensing indices for rainfall detection and estimation.

1.3. STRUCTURE OF THE THESIS

This thesis is structured in eight chapters. In **Chapter 1**, the scientific and the social relevance of the study are presented. The research objectives in this study are stated in this Chapter.

In **Chapter 2**, the study area which is the Lake Tana basin is described and includes a description of its geographic and climatic settings. The long-term rainfall and stream flow statistics of the Lake Tana basin are analyzed and presented in Chapter 2.

In **Chapter 3**, the spatio-temporal patterns of the rainfall in the Lake Tana basin and the Gilgel Abbay watershed in particular are analyzed. The rainfall data that are obtained from 8 recording rain gauges of the network

are analyzed to study the space-time structure of the rainfall, the diurnal cycle of the rainfall and its temporal intermittence. In Chapter 3, a convective index was developed using thermal infrared (TIR) observations of cloud top surface to analyze the spatial pattern of the rainfall diurnal cycle.

In **Chapter 4**, the properties of the rain events of the Lake Tana basin and their temporal variability are analyzed. The rain event properties include frequency of occurrences, intensity and duration of rain events, and inter-event time. Dimensionless event-hyetographs are derived in this Chapter and a simple model that only has two parameters is fitted to the hyetographs.

In **Chapter 5**, the potential of remote sensing observations for rainfall detection and estimation is evaluated. First, a set of indices was derived using multi-spectral observations from MSG-2. The performance of the indices for rainfall detection is evaluated in terms of selected performance measures. In Chapter 5, an exponential model was developed to estimate rainfall using TIR brightness temperature of cloud top surface. The PR of TRMM satellite provided the rainfall rates that served as the ground-truth for calibration and validation of the rainfall detection and estimation approach.

In **Chapter 6**, a conceptual cloud model is developed to simulate convective rainfall using readily available model inputs. The cloud model operates based on a parsimonious formulation of conservation equations of mass for a cloud layer. The model inputs are ground-surface temperature, pressure and dew-point temperature, and TIR brightness temperature of cloud top surface. Model sensitivity is evaluated through Regional Sensitivity Analysis (RSA).

In **Chapter 7**, the sensitivity of a physically based rainfall-runoff model to rain gauge network density and configuration is evaluated. The accuracy of the estimated rainfall is evaluated through a set of performance measures by comparing it against a reference rainfall. The main objective is to determine how rain gauge network density and configuration affects the stream flow simulated by the Representative Elementary Watershed (REW) rainfall-runoff model. In Chapter 7, model sensitivity to model resolution and rainfall variability is evaluated.

Introduction

In **Chapter 8**, the main results of this thesis are summarized. Some concluding remarks and recommendations for future studies are presented in Chapter 8.

2

THE LAKE TANA BASIN



2.1. DESCRIPTION OF THE BASIN

The centre of Lake Tana is located approximately at 12° 00' N and 37° 15' E in Ethiopia, East Africa. The lake has a surface area of approximately 3100 km² with a mean and a maximum depth of 9 m and 14 m, respectively and an altitude of 1786 m above mean sea level.

The Lake Tana basin has a surface area of approximately 15,000 km² with a north-south length of 200 km and a west-east length of 165 km. The geology of the basin is dominated by basalts while the soil is dominated by luvisols which have a soil texture of clay to silty clay. The basin is mostly covered by agricultural land.

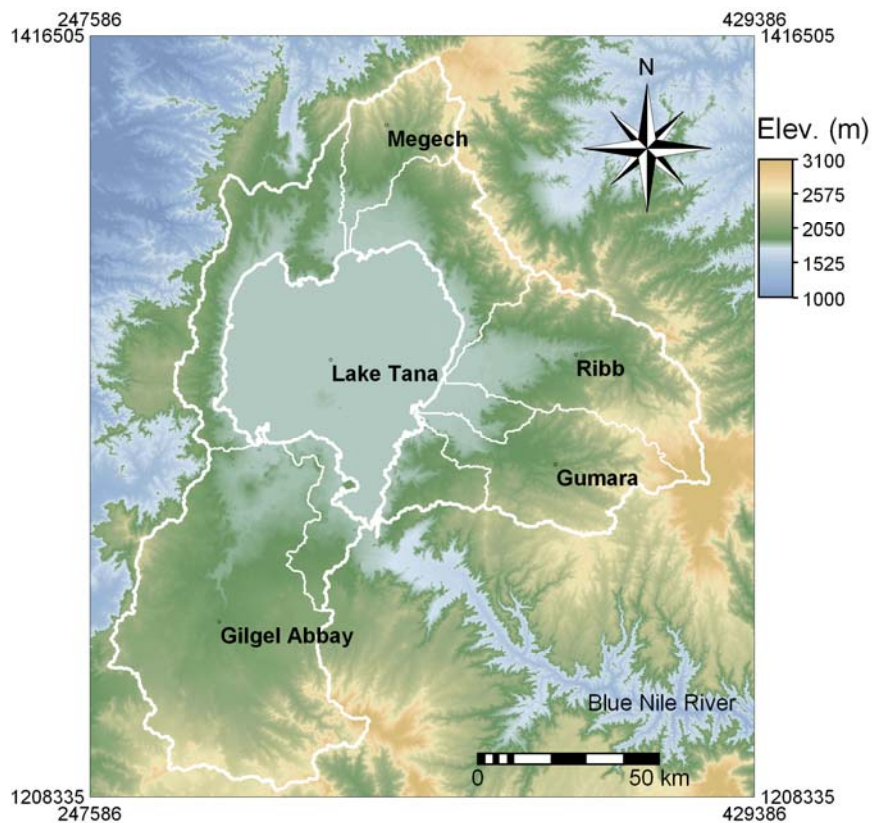


Figure 0.1: *The topography of the Lake Tana basin and its major watersheds. The boundaries are represented by white lines.*

More than 40 watersheds contribute to the inflow of Lake Tana. According to Wale et al. (2009), the gauged river inflow to the lake is 1280 mm yr⁻¹ while the ungauged river flow is 880 mm yr⁻¹. Most of the recording rain gauges that provided the rainfall data for the analysis in Chapter 3 to 7 of this thesis are installed in Gilgel Abbay watershed. Gilgel Abbay is the largest watershed of the Lake Tana basin with a surface area of about 5000 km² and an altitude that ranges between 1790 – 3500 m above mean sea level. The Gilgel Abbay watershed has mountain ranges in the southern part, extensive flat plains near its main river mouth and a large water body (Lake Tana) north of its outlet (see figure 2.1). The other major watersheds of Lake Tana basin are Gummara, Ribb and Megech that have a surface area of approximately 1400 km², 1900 km², and 850 km², respectively. The surface areas of the watersheds are extracted through watershed area delineation based on a Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM) which has 90 m resolution.

2.1.1. Temperature

INTER-ANNUAL VARIABILITY

The inter-annual variability of surface temperature in the Lake Tana basin is analysed using data from four ground based weather stations. These stations are Dangila, Bahir Dar, Debre Tabor and Gondar which are located in the southern part of the basin, the south shore of the lake, the eastern part and the northern part of the basin, see figure 2.6. The daily temperature data that is used in this analysis is recorded for the time period 1994 – 2003.

The inter-annual variability of the daily minimum temperature is shown in figure 2.2 and represents monthly averaged values. The figure shows that the patterns of the minimum temperature of the four stations are somewhat similar. The minimum temperature is highest in April or May and reaches its lowest value in January. The mean of the minimum temperature ranges between 8.7 °C at Dangila to 13.8 °C at Debre Tabor.

Figure 2.3 shows the inter-annual variability of the maximum daily temperature which is averaged for each month. The temporal pattern of the maximum temperature differs from that of the minimum temperature with their lowest value occurring in different months. Figure 2.3 shows that the maximum temperature of the four stations has a similar pattern with higher

value in April and lowest values in July and August. Bahir Dar and Gondar have higher temperature values while Debre Tabor has lowest values. The mean of the maximum temperature ranges between 22 °C at Debre Tabor and 27.3 °C at Bahir Dar.

DIURNAL VARIABILITY

The diurnal cycle of the surface temperature is analysed using the data that was recorded at Durbet station. The data was recorded in June – August (JJA) of 2007 by the weather station that has been installed as part of the present study.

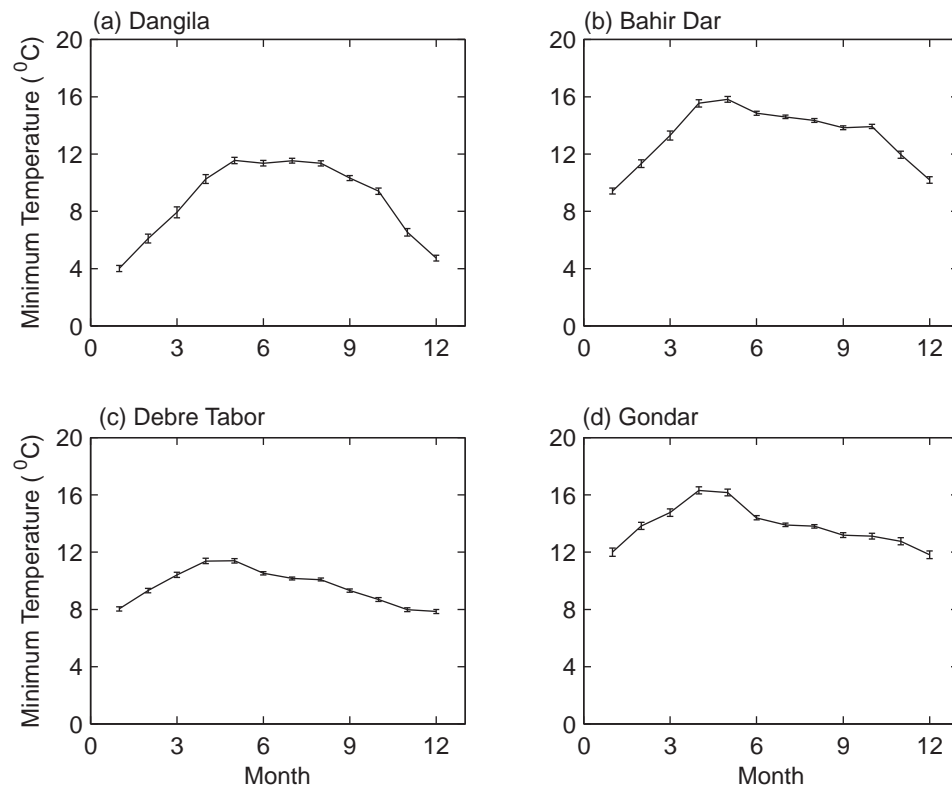


Figure 0.2: Inter-annual variability of the daily minimum temperature at four stations in the Lake Tana basin (time period 1994 – 2003). The 95 % confidence intervals of the mean values are also shown. Note: the first month corresponds to January.

Figure 2.4 shows the diurnal cycle of temperature at Durbet. The figure shows that the diurnal cycle has maximum temperature at 14:00 Local Standard Time (LST) and minimum temperature at 7:00 LST. The temperature ranges between 14 °C – 22 °C showing large diurnal variability.

The Lake Tana basin

The figure also shows the 95 % confidence interval of the estimated mean of the temperature on each LST. The confidence interval is larger in the afternoon than in the evening or in the morning and is due to large variation in the afternoon temperatures.

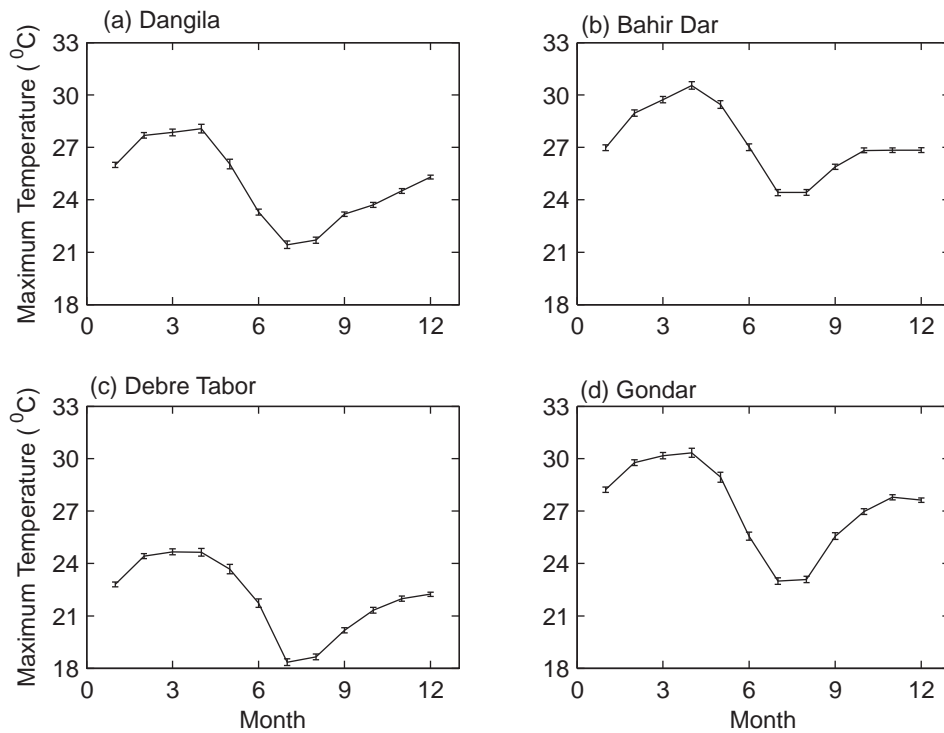


Figure 0.3: Inter-annual variability of the daily maximum temperature at four selected stations in the Lake Tana basin (time period 1994-2003). The 95 % confidence intervals of the means are shown.

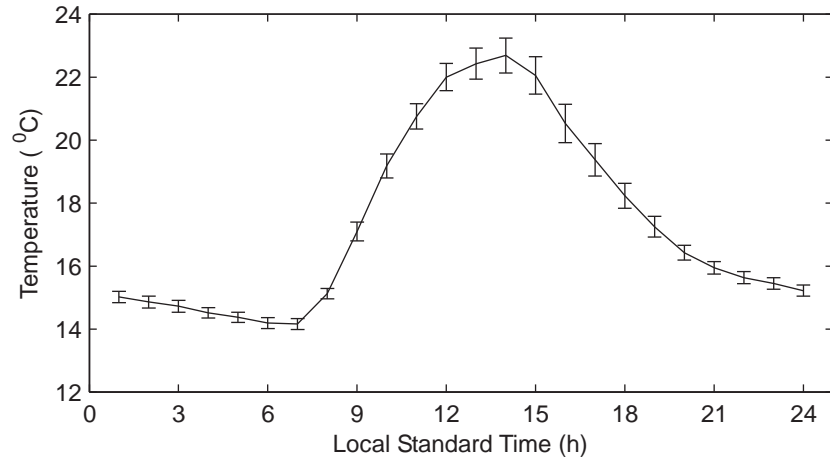


Figure 0.4: The diurnal cycle of temperature at Durbet weather station in Gilgel Abbay watershed.

2.1.2. Rainfall variation

THE ANNUAL AND SEASONAL RAINFALL

The annual and the wet season rainfall of the Lake Tana basin are estimated using rainfall records of 15 stations (see figure 2.6d) in the time period 1987 – 2006. First, the mean annual rainfall of each station is estimated and then the annual rainfall is interpolated over the basin at grid elements of 200 m. The rainfall is interpolated using the ordinary kriging technique with a spherical semi-variogram model (see, figure 2.5a). The estimated parameters of the model are: nugget = 0, sill = 2×10^5 , and range = 80 km. The annual rainfall of the grid elements is averaged to estimate the basin rainfall. The mean annual rainfall of the Lake Tana basin is estimated to be 1424 mm while the mean annual rainfall of Lake Tana is estimated to be 1534 mm. As such, there is a difference of 110 mm with the lake receiving larger rainfall depth than the land surface of the basin.

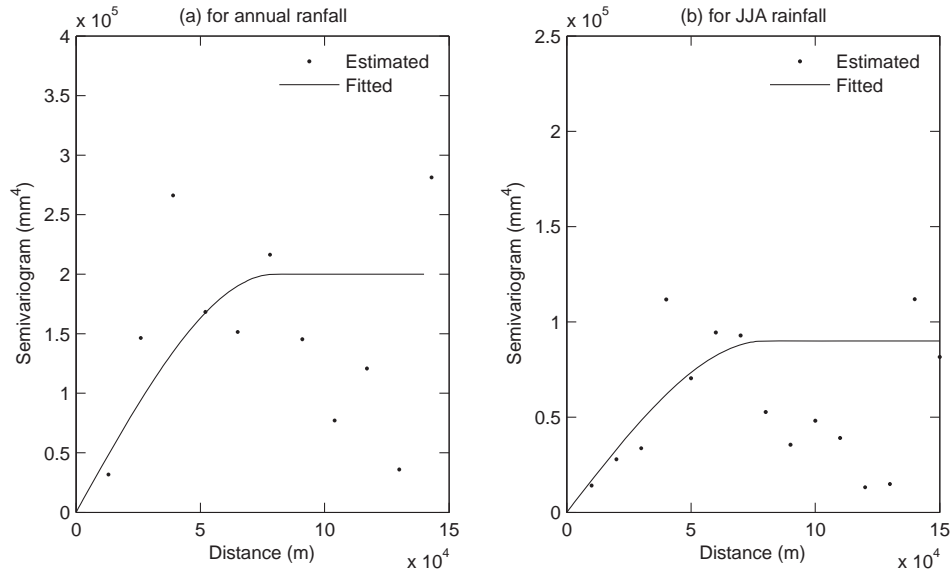


Figure 0.5: *Estimated and fitted Spherical semivariogram for the annual and JJA rainfall of the Lake Tana basin.*

The wet season rainfall is estimated following the same procedure that is applied to estimate the annual rainfall. The time period between June – August (JJA) is considered as the main wet season of the basin. The parameters of the semi-variogram model for the wet season rainfall are (see figure 2.5b): nugget = 0, sill = 9×10^4 and range = 80 km. In JJA, the mean rainfall of the basin is 948 mm which is 67 % of the annual rainfall of the basin while the mean rainfall of the Lake is 1030 mm which is also 67 % of the annual rainfall of the Lake.

Figures 2.6a-d show some features of the rainfall of the Lake Tana basin. The stations' code is shown in figure 2.6d while the name of the stations is presented in Table 2.1. Sekela station has missing values in the dry season and therefore is not used to estimate the annual rainfall but it is used to estimate the JJA rainfall since the wet season rainfall record at this station covers the full time period.

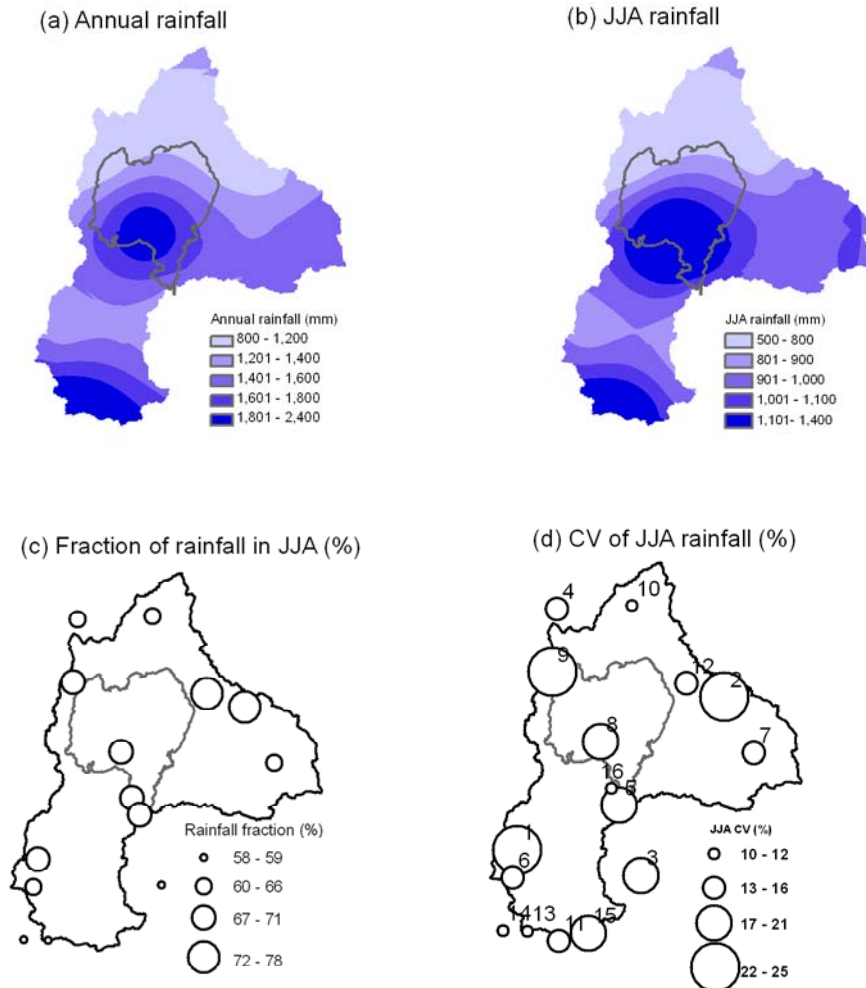


Figure 0.6: The long-term mean annual and seasonal (JJA) rainfall of the Lake Tana basin. Fraction of rainfall represents the ratio of the JJA rainfall over the annual rainfall. Note: figure 2.6d shows the stations that are shown by the numbers while the name of the stations is presented in table 2.1.

The annual rainfall in the Lake Tana basin ranges between 830 – 2368 mm which shows large spatial variability with a maximum rainfall as large as 2.8 times the minimum rainfall. Figure 2.6a shows that the south part of the basin receives the largest amount of annual rainfall while the north part of the basin receives the smallest amount of rainfall. The rainfall of the area that is situated between the mountains in the south part of the basin and Lake Tana shows mixed properties. In the south part, the rainfall

The Lake Tana basin

depth decreases as the distance from the mountain increases until some intermediate location and then the rainfall increases with a decrease in distance to the lake.

Figure 2.6b shows the JJA rainfall of the basin. Overall, the JJA rainfall has a similar spatial pattern as that of the annual rainfall. However, figure 2.6c shows that there is some spatial variation in terms of the fraction of the annual rainfall that falls in the wet season. The areas near the lake receive about 70 % of their annual rainfall in JJA while the mountain areas in the south part of the basin receive 60 % of their annual rainfall in JJA. In terms of coefficient of variation (CV), figure 2.6d shows that there is no clear spatial pattern in the temporal variation of the seasonal rainfall.

Table 0.1: *Spatial features of the rainfall stations. Note: The Coordinate System projection is Universal Transverse Mercator, Datum Adindan, Clark Ellipsoid (1880).*

Station Code	Station Name	Easting (m)	Northing (m)	Altitude (m)
1	Dangila	264717	1244326	2127
2	Bahir Dar	321101	1282608	1798
3	Sekela	304733	1215046	2715
4	Adet	332531	1245493	2218
5	Injibara	272548	1216055	2592
6	Gundil	289084	1211018	2549
7	Abbay Sheleko	267202	1259116	2075
8	Zege	316786	1291520	1791
9	Kidamaja	259514	1216449	2462
10	Gondar	327882	1387682	2123
11	Addis Zemen	376577	1339504	2111
12	Aykel	288018	1386105	2153
13	Debre Tabor	392163	1310040	2744
14	Deke Istifanos	311120	1315878	1799
15	Delgi	285688	1352646	1865
16	Infranz	356388	1346686	1889

Figure 2.7 shows some features of the daily rainfall at four stations in the Lake Tana Basin. Injibara is located on a mountain in the southern part of the basin; Bahir Dar is located on the south shore of Lake Tana; Gondar is located in the northern part of the basin and Debre Tabor is located in the eastern part of the basin. The data of Bahir Dar was recorded in the time period 1981 – 2006 while the data of Injibara, Gondar and Debre

Tabor was recorded in the time period 1987 – 2006. The days with missing data are excluded during the analysis and the daily rainfall is arranged in the respective months.

The top and the bottom horizontal line of the box plot indicate the 75 % quartile plus 1.5IQR and the 25 % quartile minus 1.5IQR, respectively where IQR is the interquartile range. The IQR is defined by the difference between the 75 % quartile which is represented by the top edge of the box and the 25 % quartile which is represented by the bottom edge of the box. Therefore, the IQR is represented by the size of the box.

Figure 2.7 shows that January, February, March, November and December are the driest months of the basin. The IQR suggests that Injibara station recorded high daily rainfall in April which shows that the wet season starts first at Injibara, i.e. at the southern part of the Lake Tana basin. In May, all of the four stations recorded relatively large rainfall depth but a large range of daily rainfall is recorded at Injibara. Also, Injibara is characterized by a wider range of daily rainfall in October as compared to the remaining stations.

In terms of skewness, a median value that is observed closer to the 25 % or the 75 % quartiles than to the median indicates a skewed distribution. For instance, the daily rainfall in May and April is largely skewed since the median is closer to the 25 % quartile than the 75 % quartile. However, the skewness is less pronounced for the July and the August rainfall since the median is closer to the middle of the box which indicates that 50 % of the rainfall records that are above and below the median value are distributed over a similar range of values.

Table 2.2 shows the median value of the daily rainfall at the selected stations in the Lake Tana basin. Most of the days in June and September are rainy while Injibara recorded significant rainfall in the months of May and October. It is shown that the median value of the September rainfall at Gondar is close to zero that indicates Gondar has much more non-rainy days than rainy days in this month. Overall, Injibara has the longest wet season and the largest daily rainfall depth as compared to the other stations.

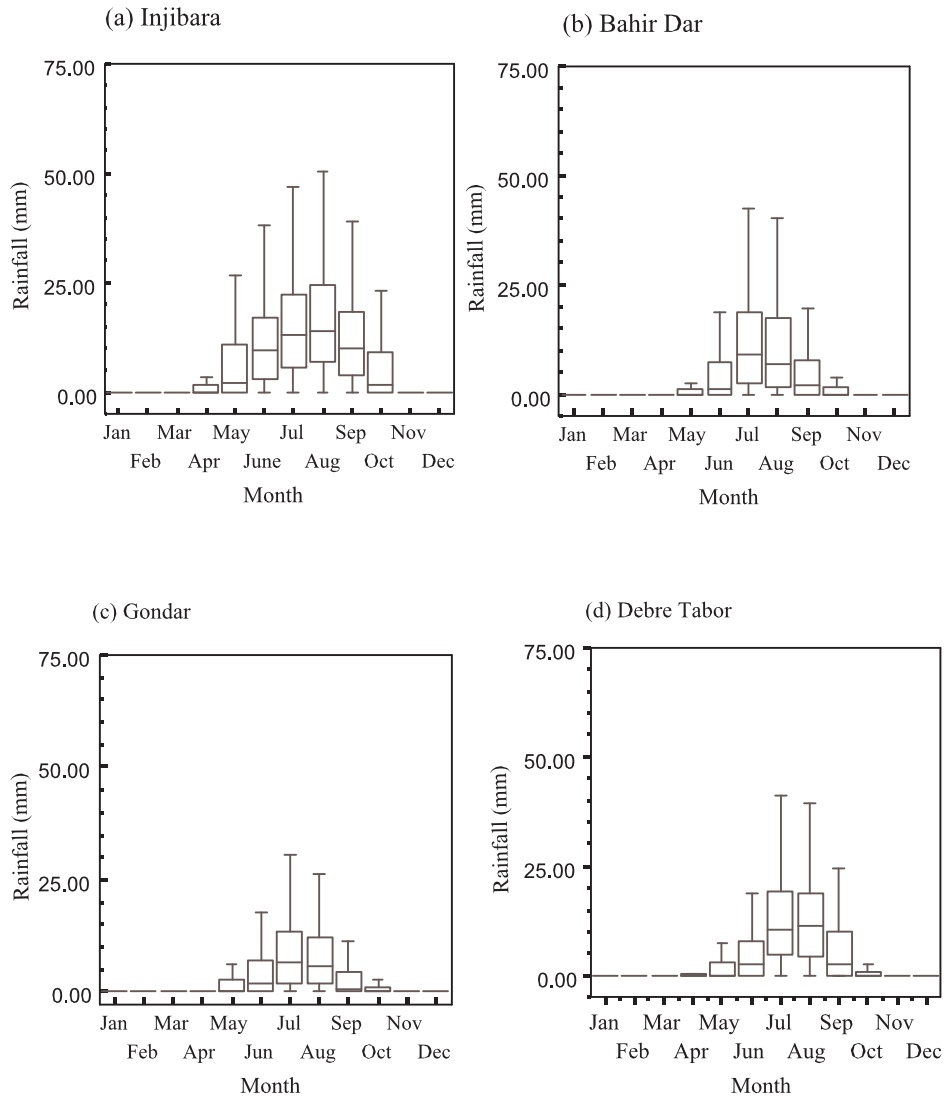


Figure 0.7: Box plots of the daily rainfall at four stations in the Lake Tana basin.

In all the wet months, Injibara receives the largest daily rainfall while Gondar receives the lowest daily rainfall, see Table 2.2. In terms of median, the daily rainfall of Injibara is as large as 2.1 times that of Bahir Dar and 2.5 times that of Gondar. The daily rainfall at Debre Tabor is approximately 3 mm smaller than that of Injibara in July and August while it is 4 times smaller than that of Injibara in June and September.

The Lake Tana basin

Table 0.2: *The median value of the daily rainfall at four stations in the Lake Tana basin.*

Station	Apr.	May	Jun.	Jul.	Aug.	Sept.	Oct.
Injibara	0.0	2.1	9.5	13.0	14.1	9.8	1.8
Bahir Dar	0.0	0.0	1.2	8.9	6.9	2.1	0.0
Gondar	0.0	0.0	1.8	6.6	5.7	0.4	0.0
Debre Tabor	0.0	0.0	2.5	10.4	11.3	2.8	0.0

2.2. SUMMARY

The inter-annual variability of temperature of the Lake Tana basin is analyzed using data from four meteorological stations. The daily maximum and minimum temperature in the Lake Tana basin have a somewhat similar inter-annual pattern. The maximum temperature peaks in April and reaches its lowest value in July. The annual-mean of the minimum temperature ranges between 8.7 °C at Dangila to 13.8 °C at Debre Tabor while the annual-mean maximum temperature ranges between 22 °C at Debre Tabor and 27.3 °C at Bahir Dar. A distinct pattern is observed in the diurnal cycle of the temperature in the basin with the minimum and peak temperature occurring on 7:00 LST and 14:00 LST, respectively. The peak of the diurnal cycle is 1.6 times its minimum.

The mean-annual rainfall of the Lake Tana basin and Lake Tana is estimated to be 1424 mm and 1534 mm, respectively showing that the lake has larger rainfall depth than the land surface of the basin. In JJA, the mean-seasonal rainfall of the basin and the lake is 948 mm and 1030 mm, respectively which is 67 % of the annual rainfall of the respective areas.

The annual rainfall in the Lake Tana basin ranges between 830 mm in the northern part of the basin and 2368 mm in the southern part with the maximum rainfall as large as 2.8 times the minimum rainfall. The areas near the lake and the mountain areas in the southern part of the basin receive approximately 70 % and 60 %, respectively of their annual rainfall in JJA.

The rainy period in the mountain areas south of Lake Tana starts in May and ends in October. However, the rainfall in the other parts of the basin starts in June and ends in September. In most parts of the basin, large amounts of daily rainfall occur in July and August but the mountains in the south also receive significant daily amounts of rainfall in June and September. Therefore, the large spatial variation that is evident in the

The Lake Tana basin

annual and the season rainfall of the basin is caused not only by large differences in daily rainfall amounts but also by differences in the length of the wet periods in the basin.

3

RAINFALL VARIABILITY USING GROUND BASED AND REMOTE SENSING OBSERVATIONS



ABSTRACT¹

Water resources of the Blue Nile River are of key regional importance to the north-eastern African countries. However, little is known about the characteristics of the rainfall in the basin. In this study, the space-time variability of rainfall is evaluated in the vicinity of Lake Tana which is the source of the Blue Nile River. The analysis was based on hourly rainfall data from a network of newly installed rain gauges, and cloud temperature indices from the Spinning Enhanced Visible and Infrared Imager (SEVIRI) sensor of the Meteosat Second Generation (MSG-2) satellite. The spatial and temporal patterns of rainfall were examined using not only statistical techniques such as exceedance probabilities, spatial correlation structure, harmonic analysis and fractal analysis but also marginal statistics such as mean and standard deviation. In addition, a convective index was established from remote sensing images to infer the spatial and temporal patterns of rainfall. Heavy rainfall is frequent at stations that are relatively close to the lake. The correlation distances for the hourly and the daily rainfall are found at about 8 and 18 km, respectively. The rainfall shows a strong spatially varying diurnal cycle. The nocturnal rainfall was found higher over the southern shore of Lake Tana than the mountainous area further to the south. Maximum convection occurs between the 1600 – 1700 Local Standard Time (LST) over the Gilgel Abbay, Ribb and Gumara catchments, and between 2200 – 2300 LST over Lake Tana and the Megech catchment. In addition, the hourly rainfall of the station with highest elevation is relatively closely clustered as compared to those stations at lower elevation. The study provides relevant information to understand rainfall variation with elevation and distance from a lake. This understanding benefits climate and hydrological studies, water resources management and energy development in the region.

Key words: Blue Nile, Lake Tana, rainfall variability, diurnal cycle, MSG, SEVIRI

¹ This chapter is based on: Haile, A. T., Rientjes, T., Gieske, A., Gebremichael, M., 2009. **Rainfall variability over mountainous and adjacent lake areas: the case of Lake Tana basin at the source of the Blue Nile River**, *Journal of Applied Meteorology and Climatology*, 48(8), 1696 – 1717.

3.1. INTRODUCTION

Rainfall is of major importance to water resources assessment and management and therefore considerable research has been devoted towards characterizing its spatial and temporal variability. Information about space-time variability of rainfall is important for a range of applications in climate and hydrology. Examples of applications include evaluation of the rainfall predictions from global climate models (Lee et al., 2007), evaluation of satellite-based rainfall products (Gebremichael et al., 2003; Bowman et al., 2005), rain gauge network design (Rodriguez-Iturbe and Bras, 1976), rainfall modeling (Arnaud et al., 2007) and rainfall-runoff modeling (Reggiani and Rientjes, 2005).

Accurate estimation of the space-time variability of rainfall is one of the major challenges in hydrometeorology. Rainfall variability is often influenced by several factors that may interact. For instance, in mountainous areas, rainfall is influenced by the spatial orientation of terrain in terms of terrain altitude, slope and aspect. The terrain orientation may affect wind patterns and cause rain shading which often results in variations in rainfall distributions (e.g., Buytaert, 2006; Gebremichael et al., 2007). Also, a large water body commonly affects rainfall distributions since the water body can influence local meteorological conditions (e.g., Ba and Nicholson, 1998).

The inter-annual variability of the Ethiopian rainfall and the prevailing large scale circulation that influence its distributions have been studied by many authors (for instance, Eklundh and Pilesjö, 1990; Conway and Hulme, 1993; Camberlin, 1997; Bewket and Conway, 2007). However, the role of terrain in affecting the diurnal and inter-seasonal variability of the Ethiopian rainfall is not clearly known. This is, in part, due to the lack of sub-daily rainfall observations.

In this Chapter, the spatial and temporal pattern of rainfall variability is analyzed at the source region of the Blue Nile River basin. This region is mountainous and is located in a tropical humid area adjacent to Lake Tana, a large inland lake in the Ethiopian highlands. Therefore, this region can be used to study rainfall variability that is affected by orography and the presence of a large water body. In the present study, several statistical techniques have been applied to assess various aspects of spatial

and temporal patterns. These are marginal statistics, exceedance probabilities, spatial correlation structure of the observation network, the rainfall diurnal cycle and fractal analysis. The data consisted of in-situ rainfall observations from a network of eight tipping-bucket rain gauges, and cloud-top temperature observations from the Meteosat Second Generation (MSG-2) geostationary satellite with the SEVIRI sensor.

3.2. DATA SETS

3.2.1 Rainfall

As part of the present study, ten rain gauges (Hobo S-RGA-M002) were installed in May 2007 in the valley and mountain area of the Gilgel Abbay catchment as shown in figure 3.1. The rain gauges have a funnel diameter of 15.4 cm, a tipping bucket mechanism with a resolution of 0.2 mm and a measurement range of 0–10 cm per hour. One of the rain gauges was installed at Sekela (Gish Abbay) which is considered to be the legendary source of Blue Nile River (see Hurst et al., 1959). To the knowledge of the author, the rainfall in the area has never been monitored at high temporal and spatial resolution. Table 3.1 shows the characteristics of the rain gauge stations. The information in table 3.1 is based on a Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM) of 90 m resolution.

Out of the ten rain gauges, eight rain gauges provided complete observations in the wet season. Two of the rain gauges, which were placed at Wotet Abbay and Gult, did not provide observations for the time periods June 1 to August 24 and June 23 to August 31, 2007, probably because of human interference. These stations are excluded from the analysis since the data covers a time period much smaller than the time period of the study which extends from June 1 to August 31 of 2007, i.e. June, July August (JJA). As a sample case, the hourly rainfall observation at Jema station is shown in figure 3.2. The figure shows three peak events with hourly rainfall higher than 30 mm and some regularity is observed in the time of occurrence of rainfall events higher than 5 mm.

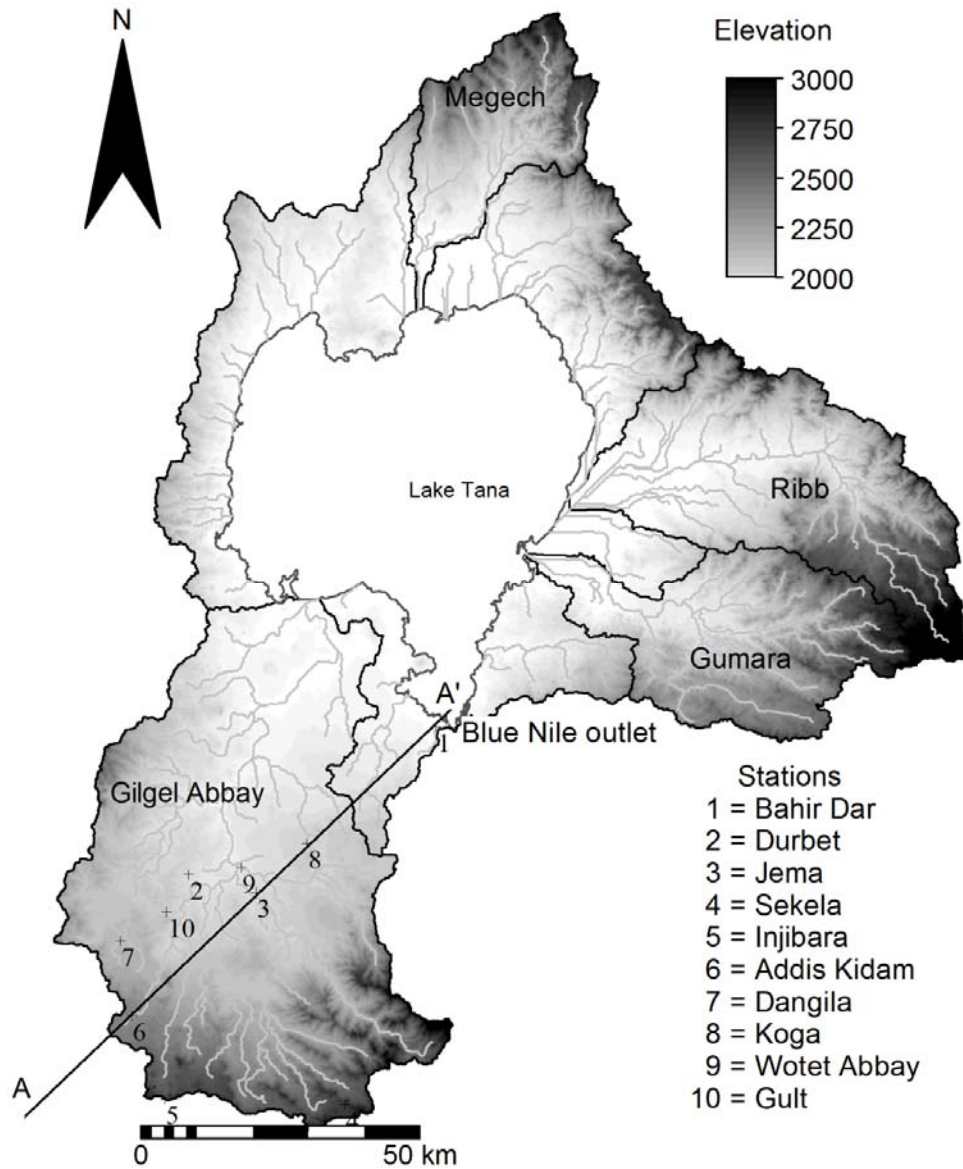


Figure 3.1: Digital Elevation Model of Lake Tana basin with watershed boundaries of the major rivers and location and names of rain gauges as indicated by numbers. Transect A-A' is described in section 3.5.

3.2.2. Cloud infrared temperature

Surface observations of rainfall can be used to evaluate temporal variability at discrete point locations. Evaluation of rainfall variability over space domains requires the use of remote sensing data that provides spatial coverage. Remote sensing data for this study is thermal infrared temperature data from the SEVIRI sensor on the Meteosat Second Generation (MSG-2) satellite and images are received via the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) Multicast Distribution System (EUMETCast). The International Institute for Geo-Information Science and Earth Observation (ITC) receives the images under the license from EUMETSAT. Information about the MSG image processing that is undertaken at ITC can be found in Gieske et al. (2005) and Maathuis et al. (2006).

Table 3.1: *Characteristics of the location of the rain gauge stations: Note: The Coordinate System projection is Universal Transverse Mercator, Datum Adindan, Clark Ellipsoid (1880).*

Stations	Geographic coordinates				
	Easting (m)	Northing (m)	Elevation (m)	Slope (%)	Aspect (-)
Bahir Dar	321101	1282608	1798	1.11	West
Durbet	276868	1256270	1984	2.07	East
Jema	288938	1252964	1970	2.48	West
Sekela	304733	1215046	2715	4.60	South-west
Injibara	272548	1216055	2592	5.15	South
Addis Kidam	266687	1231060	2370	1.61	South
Dangila	264717	1244326	2127	1.94	North-west
Koga	298104	1261710	2011	0.95	North-west
Wotet	286263	1257411	1917	3.23	North-east
Abbay Gult	272969	1249501	2063	2.92	North-east

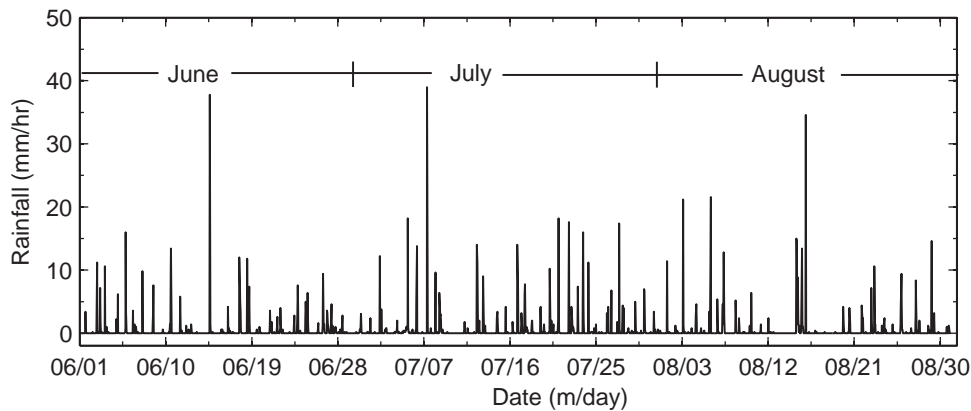


Figure 3.2: Rainfall observations at Jema for the time period June 1 to August 31, 2007. Note the Date is in month/day format.

The thermal infrared (TIR) brightness temperatures in the range of 10–12.5 μm have been used in studies of rainfall estimation since the late 1960s. In TIR based rainfall estimation, cloud top brightness temperatures are commonly used as proxy variables to infer rainfall rates (see Adler et al., 1993; Todd et al., 1995). The underlying physical assumption for this method is that relatively cold clouds can be associated with thick and high clouds that often produce relatively high rainfall rates. Some applications of TIR temperatures for diurnal cycle assessments are also presented for instance by Ba and Nicholson (1998); Ohsawa et al. (2001).

In this study, the 10.8 μm channel is selected and the brightness temperature images are analysed at hourly time intervals for the time period of JJA in 2007. Images in this study are of 3 km spatial resolution. For the period from August 26 to 31, images were not available by ITC receiver failure. As such, the 10.8 μm TIR data was analysed for the time period June 1 – August 25 for which high resolution data was available.

3.3. METHOD OF ANALYSIS

Statistical methods and indices are applied to analyse the different aspects of spatial and temporal patterns of the observed rainfall. Statistical methods have been applied to the network data to assess the effect of elevation and the effect of distance to the lake on rainfall variability. Also,

procedures have been applied to the remote sensing data to analyze cloud behaviour over selected time and space domains.

3.3.1. Spatial variability

First, the spatial variability of the statistics of the data that is recorded at the rain gauges is considered. The statistics include: mean, coefficient of variation (standard deviation divided by mean), probability of rainfall occurrence and exceedance. The exceedance probability values were estimated using the Weibull plotting position method that provides unbiased estimates (Stedinger et al., 1993). Examples of the Weibull method in rainfall studies can be found in Rappold (2005).

3.3.2. Spatial correlation structure

For analysis of spatial rainfall distributions, correlation functions are defined to measure the spatial arrangement at the order level of the second moments. Correlation functions can be used for both descriptive and for predictive purposes. In a descriptive mode, the correlation functions are used to infer the dominant meteorological process (see Gebremichael and Krajewski, 2004) while in a predictive mode, the correlation functions are used in spatial estimation. The correlation functions are usually expressed as a function of Euclidean distance that also is in this study. In the correlation analysis, the Pearson product-moment correlation, $\rho_g(d)$, is first estimated for each pair of gauges at a specific separation distance that was followed by fitting a correlation model to the correlation values. Studies such as by Ciach and Krajewski (2006) used the following equation that is also adapted here:

$$\rho_g(d) = c_0 \exp \left[- \left(\frac{d}{d_0} \right)^{s_0} \right] \quad [3.1]$$

where: c_0 , d_0 and s_0 are fitting parameters to the estimated correlation values. Here, d is the separation distance between two gauge locations, c_0 is the correlation value for the near-zero distances, d_0 is the correlation distance (the spatial decorrelation scale) which is commonly referred to as the 'scale parameter' while s_0 is the 'shape parameter'.

The partial correlation coefficient is applied to analyze the relation between seasonal rainfall and elevation, after removing the effect of the Euclidean separation distance. The formula for the partial correlation, $r_{YX|W}$, between two random variables Y and X with effects of a controlling, third, variable W reads:

$$r_{YX|W} = \frac{r_{XY} - r_{XW}r_{YW}}{\sqrt{1-r_{XW}^2}\sqrt{1-r_{YW}^2}} \quad [3.2]$$

where: r_{YX} , r_{XW} , and r_{YW} are correlation coefficients between rainfall (here Y variable) and elevation (here X variable), elevation and the controlling variable W (here distance with respect to a reference point), and rainfall and the controlling variable, respectively.

3.3.3. Diurnal variability

To understand the diurnal cycle of rainfall over the Lake Tana basin, both rain gauge and remote sensing data are analysed. The accumulated rainfall and the frequency of rainfall occurrences at each hour of the day is also analysed and the effects of extreme events on the diurnal cycle characteristics is examined. The 10.8 μ m images are applied to characterize the temporal and spatial patterns of convective activity. The analysis was following a commonly applied procedure to establish the diurnal cycle from TIR images similar to that applied in Ohsawa et al. (2001) and Barros et al. (2004) among others. The seasonal cloudiness for each hour of the day is inferred from the frequency index that also is referred to as the convective index (see Ba and Nicholson, 1998). In this work, the convective index (CI) over a given pixel for a time period T at a specific hour h is defined as

$$CI_h = 100 \left[\sum_{i=1}^N t_i \right] N^{-1} \quad [3.3]$$

where: t is equal to one when the observed brightness temperature is in the range of specified temperature values, otherwise it is equal to zero; N is the number of observations at local time h over the time period T . In this study, T corresponds to the time period that spans from June 1 to August 25, 2007. The established indices correspond to brightness temperature values lower than 210 K, between 210 and 225 K, and between 225 and 240 K to account for high, mid and low level clouds, respectively.

A harmonic analysis is applied to the frequency of rainfall occurrences and the convective index data. The results provide information on the spatial consistency and strength of the diurnal and the semidiurnal cycles. The following form of Fourier decomposition is applied:

$$F(t) = m + S_1 + S_2 + \dots + S_n \quad [3.4]$$

where: $F(t)$ is the fitted series, m is the daily mean value of the diurnal cycle and S_n is the n^{th} harmonic of the series. In rainfall studies, the main interest is on the first and second harmonics that are often applied to determine the diurnal and the semidiurnal cycles, respectively, see also Dai (2001). The harmonics of the series are expressed by:

$$S_1 = a_1 \cos(\omega t - \phi_1) \quad [3.5]$$

$$S_2 = a_2 \cos(2\omega t - \phi_2) \quad [3.6]$$

where: a and ϕ are the amplitude and phase angle of the harmonics, ω equals $2\pi/24$ since the number of hours in a day is 24. The parameters a and ϕ are determined using the method of least squares. The fraction of variance (σ_i^2) of each harmonic can be determined from $a_i^2/2$ and as such the variance explained by each harmonic is σ_i^2/σ^2 , where σ is the standard deviation of the observations.

3.3.4. Fractals of rainfall intermittence

The fractal dimension of the rainfall intermittency (rain – no-rain) of the time series is estimated to assess rainfall scaling behaviour. By the box-counting method, the fractal dimension of the rain – no-rain observations is estimated at hourly time intervals. For each station, 2048 hourly rainfall records, i.e. 85.3 days which include both the rain and no-rain cases for the time period from June 1 to August 25 are analysed.

In the box-counting method, the time domain is divided into non-overlapping boxes of size λ . For each box size, which is consecutively decreased by a factor 2, the number of boxes of size λ that contain rain, $N(\lambda)$, are counted. The fractal dimension, D , is applied to characterise the behaviour of the process under consideration. In the present study, if the rainfall intermittency is scale-invariant, then for any change in the box size, $N(\lambda)$ will linearly be changed by a factor λ^{-D} :

$$N(\lambda) \propto \lambda^{-D} \quad [3.7]$$

The box size $N(\lambda)$ is plotted against λ on a log-log scale to evaluate the scaling behaviour of the rainfall intermittency where the process is termed scale-invariant if the plot shows a straight line. Further information on the box-counting method is available in the work of De Lima and Grasman (1999).

3.4. RESULTS

3.4.1. Spatial variability

Marginal statistics of the hourly rainfall observations are shown in table 3.2. The statistics of the rainfall changed when no-rain hours are excluded from the analysis. For instance, without excluding the no-rain hours, the median and interquartile range (IQR), which was calculated as the difference between the 75 % and 25 % quartiles, of the hourly rainfall are zero. However, after excluding the no-rain hours, the median and IQR are between 0.6 – 0.8 mm and 1.6 – 2.6 mm, respectively. Considering rain hours only, the mountain stations have relatively small IQR suggesting relatively small variability. As shown in table 3.2, the relatively high Kurtosis that was observed at Koga and Addis Kidam suggests that most of the variance at these stations is due to relatively high variability in the observations.

Most of the stations on the mountain areas of Gilgel Abbay have higher probability of rainfall occurrences than those on low elevation areas, see figure 3.3b. The lowest probability of rainfall occurrences was found on the southern part of the lake shore, i.e. Bahir Dar. The seasonal rainfall that was observed in JJA 2007 is in the range of 810 mm (at Durbet) and 1185 mm (at Jema), see figure 3.3c. The conditional mean rainfall was estimated as the ratio of the sum of the hourly rainfall values over the corresponding number of rainy hours. Figure 3.3d shows that the highest conditional mean rainfall was observed at Jema while the lowest conditional mean rainfall was observed at Sekela that is located on the mountain.

Table 3.2: *Statistics of the hourly rainfall observations.*

	B. Da	Dur.	Jema	Sek	Inji.	A.Ki	Dan.	Kog.
(a) Unconditional statistics								
Data (hours)	2208	2208	2208	2208	2208	2208	2208	2208
Mean (mm)	0.41	0.37	0.54	0.48	0.53	0.38	0.47	0.41
Median (mm)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
St. Dev.(mm)	2.06	1.52	2.35	1.69	2.11	1.80	1.95	2.15
IQR (mm)	0.0	0.0	0.0	0.2	0.2	0.0	0.0	0.0
Skewness (-)	8.7	7.2	8.5	7.1	8.9	13.3	7.7	12.3
Kurtosis (-)	95.0	65.2	98.6	70.4	102.3	278.8	74.2	231.6
(b) Statistics conditioned on rainfall occurrence								
Data (hours)	369	433	464	593	625	443	490	390
Mean (mm)	2.47	1.87	2.56	1.77	1.86	1.89	2.13	2.32
Median (mm)	0.8	0.6	0.8	0.6	0.8	0.8	0.8	0.6
St. Dev.(mm)	4.50	2.99	4.61	2.89	3.64	3.64	3.69	4.67
IQR (mm)	2.2	2.1	2.6	1.8	1.6	1.8	2.0	2.0
Skewness (-)	3.6	3.2	4.0	3.9	5.0	6.9	3.7	5.6
Kurtosis (-)	15.0	12.3	22.0	20.9	31.1	69.2	16.4	47.4

As shown in figure 3.3e, the values of the coefficient of variation (CV) are greater than one indicating a standard deviation that exceeds the mean rainfall. The largest CV was observed at Koga while the smallest CV was observed over the mountain at Sekela. The maximum hourly rainfall and the Julian day of the maximum hourly rainfall are shown in figure 3.3f–g. The highest maximum rainfall which is 55 mm hr⁻¹ was observed at Koga station on the same Julian day on which the maximum hourly rainfall at Durbet was observed. However, the maximum rainfall at the remaining stations was observed on different Julian days, which suggests a convective nature of the rainfall formation that also applies to most other days.

The exceedance probabilities of the hourly rainfall are shown in figure 3.4a and 3.4b. To clearly show the difference between the stations in terms of the probabilities the result is presented in two plots that are for hourly rainfall in the ranges 0.2 – 2.5 mm and 5 – 20 mm. These ranges are selected based on visual judgment of the rainfall distribution. Figures 3.4a–b show that the highest exceedance probability of low hourly rainfall is observed at Sekela while the highest exceedance probability of high hourly rainfall is observed at Jema. As shown in figure 3.4c, Durbet has the lowest

mean rainfall conditioned on any specific threshold while Jema has the highest conditional mean rainfall.

Although the exceedance probabilities of the hourly rainfall show spatial variation, there is no unique relation between the variation of exceedance probabilities and elevation or distance to the lake. Therefore, the data is first arbitrarily grouped into five classes of hourly rainfall and then the number of rain hours within a specified range is counted, see table 3.3. For hourly rainfall lower than 10 mm, a direct relation is found between the number of rainfall occurrences and terrain elevation. However, for hourly rainfall greater than 10 mm, an inverse relation is found between the number of rainfall occurrences and elevation. Table 3.3 also shows the partial correlation between the number of rainfall hours and elevation after removing the effect of distance to the center of Lake Tana, which is a potential source of moisture for rain formation in the area. Equation [3.2] was applied for this purpose. In general, the correlation between the number of rainfall occurrences and elevation decreased when the effect of distance to Lake Tana was removed.

A correlation coefficient of 0.340 is found between the accumulated JJA rainfall and elevation. The partial correlation between the two variables became rather small (-0.001) when distance from the centre of Lake Tana was considered as a controlling factor. To evaluate the effect of the choice of reference points from which distance is measured, the partial correlation is estimated by changing the reference points since finding best reference location is far from trivial. Using nine alternative reference locations that is the location of the rain gauges, the estimated partial correlation ranges between 0.134 and 0.403 with a mean of 0.268 and a CV of 0.500. Therefore, the seasonal rainfall distribution of the 2007 rainy season in the study area may not simply be explained by elevation and inter-station distance only. The effect of rainfall variation on runoff is evaluated in Chapter 7.

Rainfall variability over mountainous and adjacent lake area

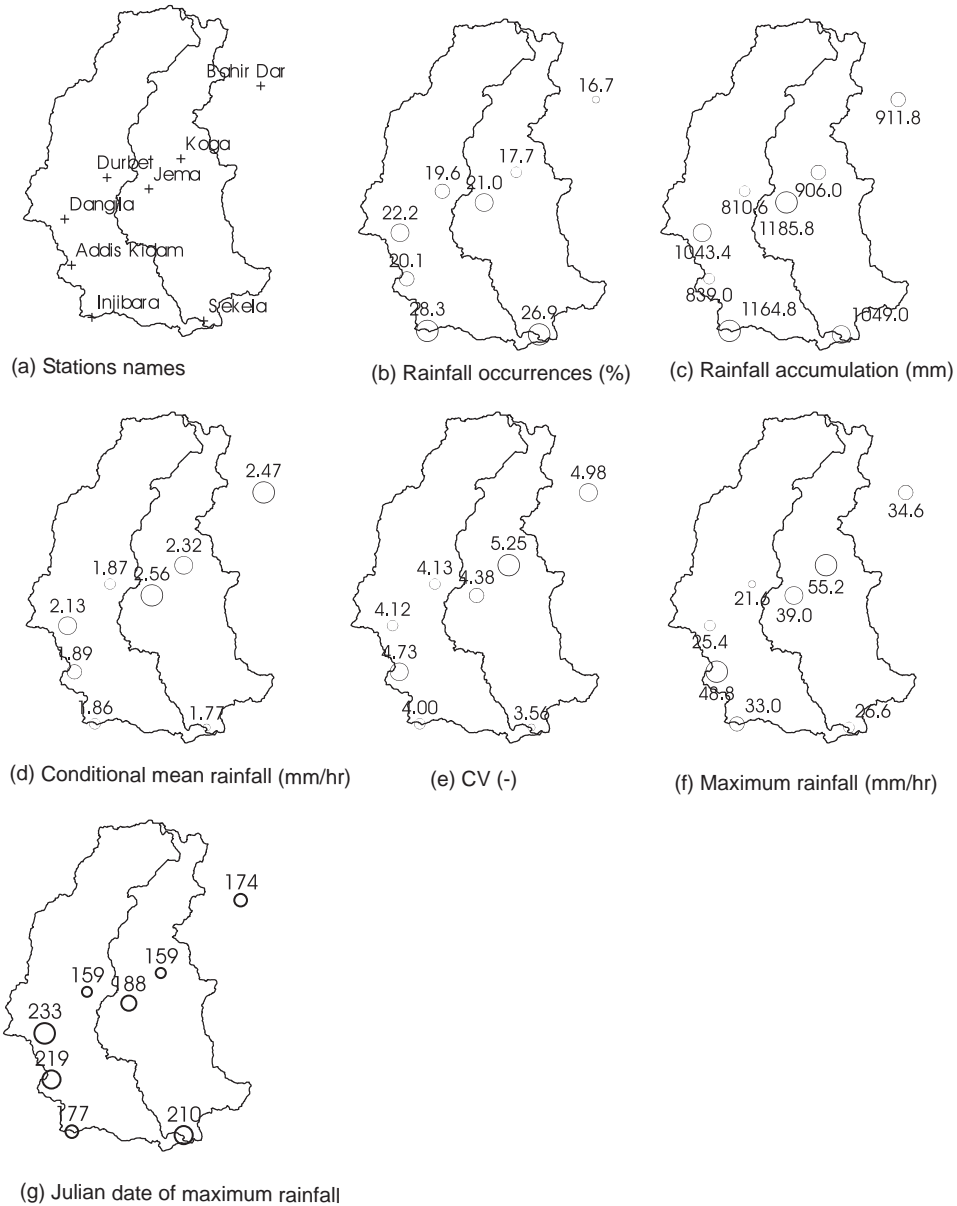


Figure 3.3: Statistics of hourly rainfall in June-August, 2007. In the diagrams, a line is added that shows the main river that separates the east and west sides of Gilgel Abbay watershed. Note: for the mean rainfall, the conditioning was done on non-zero rainfall values.

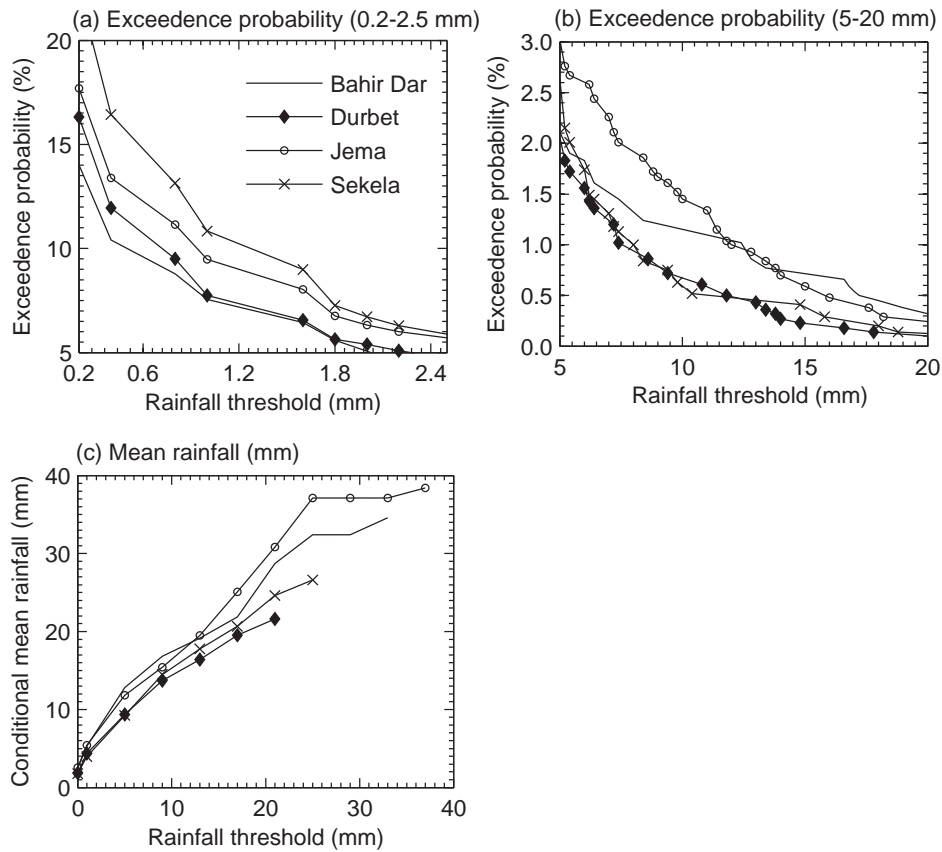


Figure 3.4: Exceedance probability of the hourly rainfall (a) and (b), and mean rainfall conditioned on specific rainfall depth (c).

3.4.2. Spatial correlation structure

A model is established between the inter-station distance and the corresponding correlation values for the hourly rainfall. At inter-station distance smaller than 12 km, the data set was too small to accurately establish the parameters. This is similar to the results shown in Young et al. (2000). Figure 3.5b and 3.5d show the changes in the value of the correlation distance (d_0) with changes in the value of the correlation at zero inter-station distance (c_0). The results in these figures are obtained by optimizing the two parameters (d_0 and s_0) of equation [3.1] while constraining c_0 . For

instance, the correlation distance of the hourly rainfall varies between 8 – 12 km for c_0 values between 0.8 and 1. The result suggested that accurate estimation of the correlation distance of the hourly rainfall requires rain gauges with inter-station distance smaller than 10 km. However, to obtain the approximate values of the correlation parameters for the hourly rainfall, at first the exponential model is fitted such that the correlation at zero inter-station distance (c_0) is equal to 1.0. Second, the same procedure is repeated without forcing the model.

Table 3.3: *Correlation between number of rain hours and elevation. Note: Distance to the stations is measured from the centre of Lake Tana.*

Hourly rainfall	Average number of rain hours	Correlation	
		Elevation	Elevation (Effect of distance removed)
0.2 mm to 1.0 mm	265	0.90	0.52
1.0 mm to 5.0 mm	160	0.86	0.36
5.0 mm to 10.0 mm	32	0.45	-0.27
10.0 mm to 15.0 mm	9	-0.90	-0.90
≥ 15.0 mm	9	-0.34	0.22

In the first case, the shape parameter is 0.51 with a correlation distance (d_0) of 8.1 km, see figure 3.5a. In the second case, it is found that $c_0 = 0.9$, the shape parameter is 0.54 and the correlation distance is 9.8 km for which the correlation became 0.37. Therefore, less than 15% of the total variance at a specified station is explained by the variance of a station located beyond some 8 – 10 km. The value of the shape parameter revealed a sharp decline in the correlation structure that is a characteristic of a rainfall structure which largely varies over a relatively small domain. Also, the value is smaller than the commonly reported values in the literature, see Ciach and Krajewski (2006) and Gebremichael et al. (2007).

As shown in figure 3.5c, there is a decreasing trend in the daily correlation structure of the rainfall. It was not possible to fit an exponential model of the form in equation [3.1] due to large scattering. Instead, the correlation coefficient is averaged at equal distance intervals following the same procedure that was applied by Şen and Habib (2001). A decreasing trend of an exponential form was obtained when the correlation coefficient was averaged at 12 km interval. After fitting the modified-exponential

model for the averaged daily correlation structure with $c_0=1$, the parameter values are a shape parameter (s_0) of 0.46 and a correlation distance (d_0) of 18.4 km.

Figure 3.6 shows the cross-correlation of the hourly rainfall for five pairs of stations. The pattern in the figure suggests that in some cases, the ± 1 -hr lag cross-correlation is higher than that for the 0-hr lag. Examples are the cross-correlation of the hourly rainfall at the station pairs Durbet-Koga, Durbet-Bahir Dar and Koga-Dangila. The fact that the 1-hr lag cross-correlation is higher than the 0-hr lag cross-correlation can be caused by the movement of convective clouds. It also indicates that the movement of clouds between these stations takes place within 1-hr.

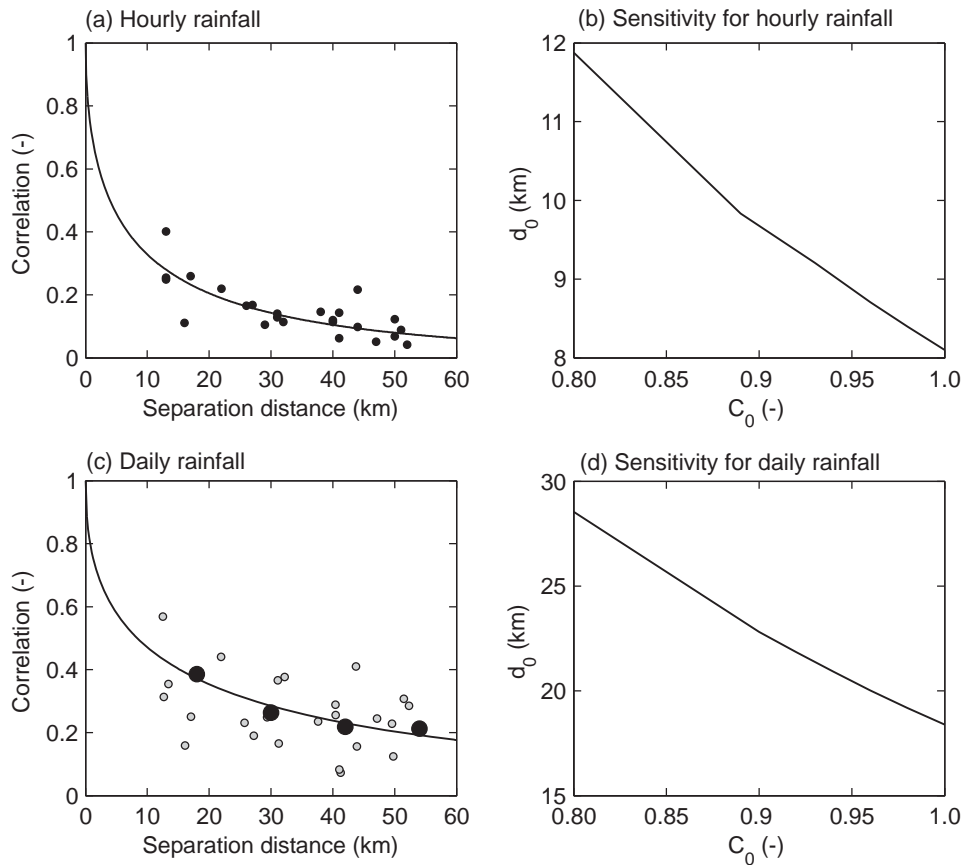


Figure 3.5: Spatial correlation structure of the rainfall observations (a) and (c), and values of correlation distance (d_0) for several assumed values of correlation at zero distance (c_0) (b) and (d). Note: For the daily rainfall, the correlation coefficient

and its averaged values are shown in small hollow and large solid circles, respectively.

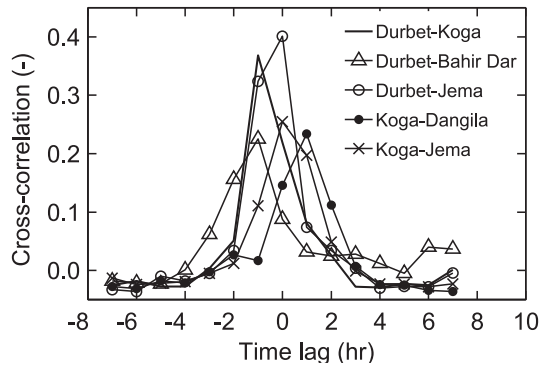


Figure 3.6: Cross-correlation of the hourly rainfall.

The 1-hr lag cross-correlation for Koga-Dangila is higher than the 0-hr lag cross-correlation suggesting movement of clouds from Koga to Dangila stations. The -1-hr lag cross-correlation for Durbet-Bahir Dar is higher than the 0-hr lag cross-correlation suggesting movement of clouds from Bahir Dar to Durbet stations.

3.4.3. Diurnal cycle

RAINFALL OCCURRENCES

In this section, results of the analysis on the diurnal variability of the observed rainfall are present. Figure 3.7a shows hourly rainfall observations at Jema station for the time period of June 1 – August 31, 2007. The figure shows that in the specified time period, three events with an hourly rainfall of higher than 30 mm were observed between 1700 and 1800 LST. For the same station, a comparison is made between the diurnal distributions of the maximum hourly rainfall and the total rainfall that was accumulated for each hour over the season. The objective is to evaluate to what extent the pattern of the diurnal cycle is affected by the contribution from the maximum hourly rainfall.

Figure 3.7b shows that the afternoon rainfall at Jema station is higher than the rainfall in the remaining periods of a day. In the afternoon, the maximum hourly rainfall at Jema was only about 25 % of the seasonal rainfall in each hour. However, in the morning (0900 – 1200 LST) the

maximum hourly rainfall was about 50 - 70 % of the seasonal rainfall in each hour. Compared to the late morning rainfall, the afternoon rainfall distribution at Jema was less affected by the contribution from the maximum rainfall. Such observations are of relevance for real-time flood forecasting studies but also in soil erosion studies where information on timing and magnitude of the maximum rainfall is required at high temporal resolution.

Figure 3.8 shows the diurnal distribution of the hourly rainfall and the frequency of rainfall occurrence in JJA at four stations namely Sekela (on the mountain), Bahir Dar (on the shore of Lake Tana), and Jema and Durbet (on opposite sides of the main valley). The threshold for rainfall occurrence is assumed to be 0.2 mm which is the lowest measurement range of the rain gauges. The frequency of rainfall occurrence is commonly used to assess rainfall diurnal variation and it shows the percentage of rain hours at each time during the day. After analysing the data from eight rain gauges, the maximum rainfall frequency was observed between 1600 and 1700 LST over land and on 2300 LST over the lake shore. The maximum frequency of rainfall occurrence varied between 23 and 61 %, with a mean of 40 % and a standard deviation of 10.8 %.

Over the southern part of the lake shore, the total rainfall depth and rainfall frequency in 0900 – 1700 LST were found low with a total rainfall of less than 10 mm and rainfall frequency of less than 10 % on each LST. However, over the mountains and inland areas, the smallest depths and frequencies were observed during a shorter period of time than over the lake shore at Bahir Dar (i.e. 0900 – 1200 LST).

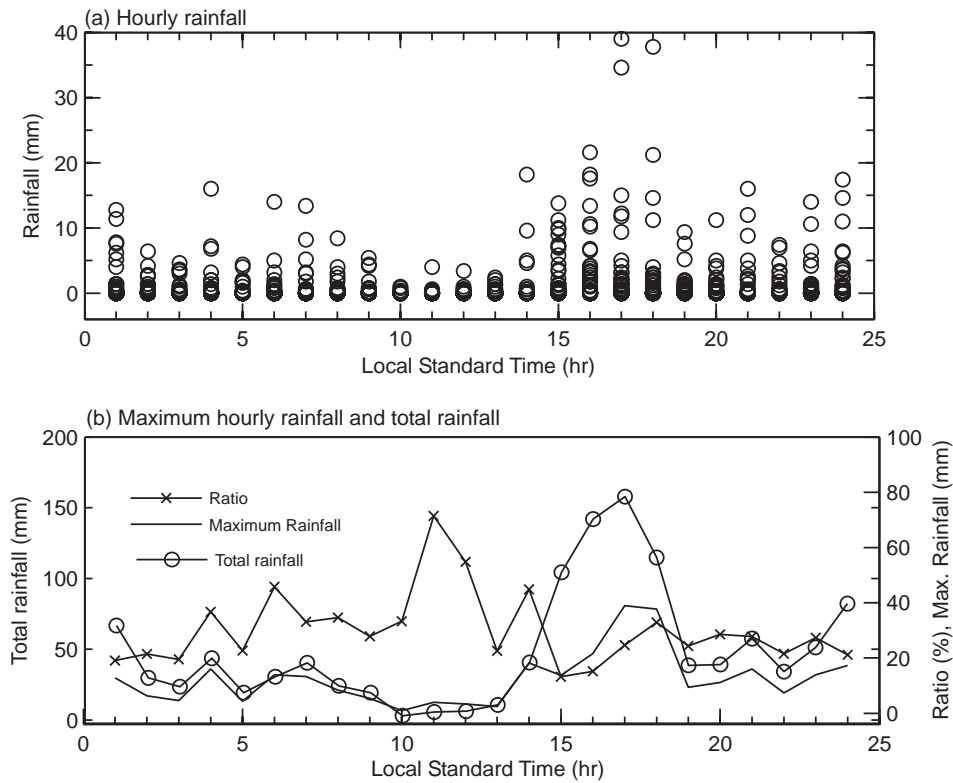


Figure 3.7: Hourly rainfall variation (a) and comparison between the maximum hourly and 3-month accumulated hourly rainfall (b) in JJA at Jema station.

Table 3.4 also shows that the percentage of nocturnal (in this case 2100 – 0900 LST) and afternoon-evening rainfall varies spatially. The selection of this time period for the comparison is because daily rainfall is recorded on 0900 LST. Also the selected time period of a day helps to make a comparison between nocturnal and afternoon-evening rainfall distribution. For instance, the nocturnal rainfall on the lake shore (Bahir Dar) was 73 % of the seasonal rainfall. However, the nocturnal rainfall on the mountain (Injibara) was about 24 % only.

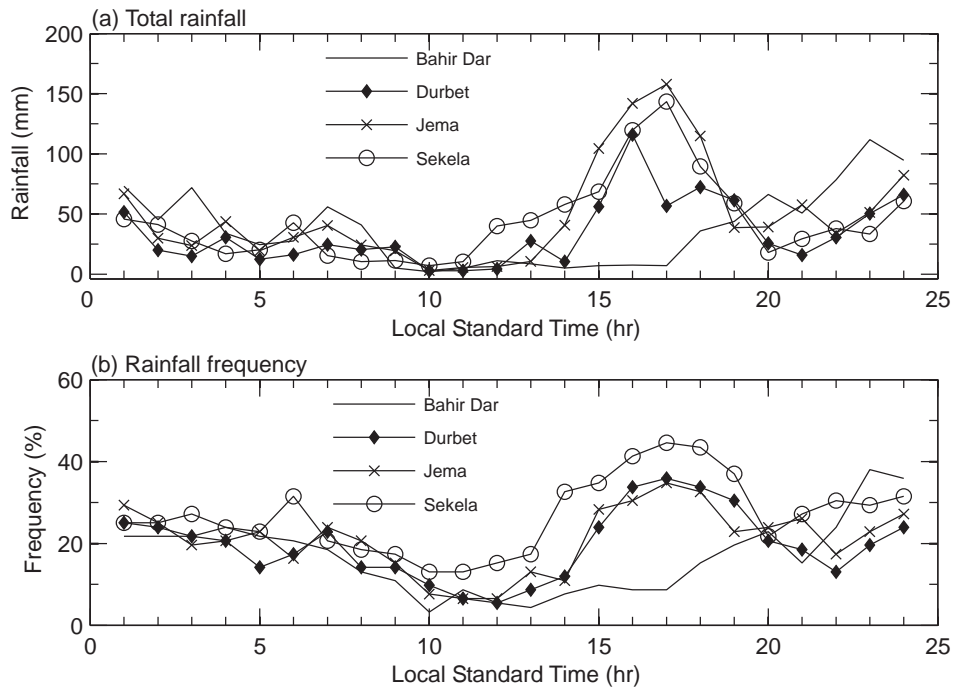


Figure 3.8: Rainfall diurnal cycle at selected four stations.

Figure 3.9 shows that the nocturnal rainfall distribution is inversely related to distance to the centre of Lake Tana. About 70 % of the variance in the spatial distribution of the nocturnal rainfall in the season is explained by distance to the centre of the lake. The relation of the seasonal rainfall with elevation and distance is discussed in section 3.5.1.

Table 3.4: Rainfall depth (% of the seasonal rainfall) in JJA, 2007.

Stations	B.Dar	Dur.	Jem	Sek	Inj	A.Ki	Dan	Kog
09-21LST	27.4	44.4	60.8	65.4	75.9	65.9	61.4	61.3
21-09LST	72.6	55.6	39.2	34.6	24.1	34.1	38.6	38.7

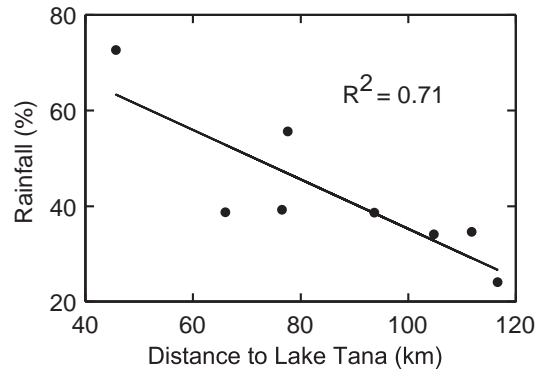


Figure 3.9: Nocturnal (2100-0900 LST) rainfall distribution as a percentage of the seasonal rainfall.

COMPARING RAINFALL OCCURRENCES AND CONVECTIVE INDEX

Although the rain gauge network covers Gilgel Abbay, rain gauges are not available that provide sub-daily observations over Lake Tana and over the remaining catchments. An alternative data source for ungauged areas is obtained from remote sensing images that provide spatial coverage. First, it is evaluated whether the convective index (*CI*) derived from the TIR images can be used to infer the area-average diurnal variations over Lake Tana and its catchments.

The relation between rainfall magnitude and *CI* values over a lake and its surrounding land would be similar when the precipitation efficiency over both surfaces is the same. Ba and Nicholson (1998) argue that this assumption can be reasonably applied for humid areas in tropical conditions. Since Lake Tana has a similar setting and since the interest is more interested on rainfall occurrence than magnitude, it can be assumed that the analysis is not much affected by any difference in effective precipitation.

The area-average convective index (*CI*) is compared against the frequency of rainfall occurrence over Gilgel Abbay. To obtain area-averaged values, the frequency of rainfall occurrence was interpolated using the inverse distance weighting (IDW) method with power 2. It is assumed that the results of the comparison will not be much affected by the bias due to the interpolation method. The interpolation was done for a grid element size of 3x3 km² which is equivalent to the resolution of the MSG-2 images for Ethiopia.

Scatter plots of the statistics of the CI and the frequency of rainfall occurrence over Gilgel Abbay are shown in figure 3.10. The CI values were calculated for temperature values less than 240 K. The regression equations and the coefficient of determination, R^2 , are also shown in figure 3.10. Results indicate that about 60 % and 80 % of the variance of the standard deviation and the mean of the frequency of rainfall over Gilgel Abbay is explained by the convective index. This suggests that the CI values can be applied to study the diurnal variation of convective activity over the lake and over its ungauged catchments.

CONVECTIVE INDEX

The spatial variation of the normalized convective index for specified hours of a day is shown in figure 3.11. Somewhat different conditions prevail in the distribution of CI values over Lake Tana and its catchments. First, in the morning, maximum CI values were observed in the form of cells that are somewhat scattered over the basin. Then these isolated cells of maximum CI merged on mid afternoon to form large areas of maximum CI. On 1600 LST the maximum CI was observed over the ridges of Gilgel Abbay which are located on the southern part of Lake Tana basin. Between 1600 and 2000 LST, the location of maximum CI moved from the mountains of Gilgel Abbay towards Lake Tana. On 2200 LST, a somewhat opposite trend to that of the 1600 LST trend was observed with the minimum CI observed over the ridges of Gilgel Abbay while the maximum CI was mostly found over Lake Tana and over Megech that is located on north of Lake Tana.

Figure 3.12 shows the convective index variation with terrain elevation along a transect that passes through Bahir Dar and Addis Kidam. The orientation of the transect is shown in figure 3.1. The CI values were estimated for temperature values less than 240 K. There is a direct relation between CI and terrain elevation on 1500 and 1800 LST. However, there is an inverse relation in the early night hours. For the remaining LST, a trend is also observed in the variation of CI with respect to elevation although this trend is much less pronounced.

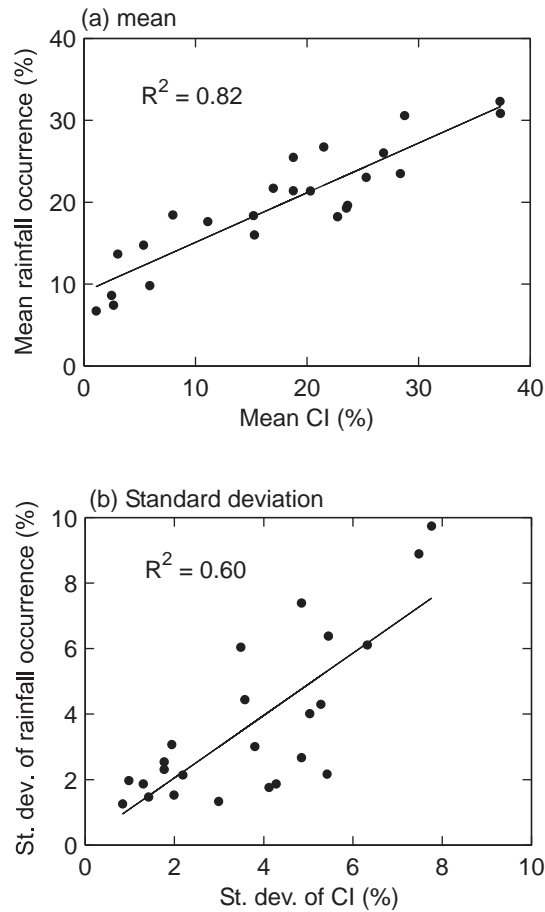


Figure 3.10: Scatter plot of catchment-average mean (a) and standard deviation (b) of the CI and the frequency of rainfall occurrence averaged over each hour (expressed in LST) in JJA, 2007.

A comparison is also made between the area-average CI over Lake Tana and over four of its major tributaries, i.e. the Gilgel Abbay, the Gumara, the Ribb and the Megech catchments. The Gilgel Abbay catchment is located south of Lake Tana, the Gumara and Ribb catchments are located east of the lake and finally the Megech catchment is in the north. The mean CI in the time period June – August was 17.5, 15.9, 14.1, 13.5 and 17.4 % over the Gilgel Abbay, Gumara, Ribb, Megech catchments and Lake Tana. As shown in figures 3.13a and 3.13b, somewhat different conditions prevail over Lake Tana and its catchments.

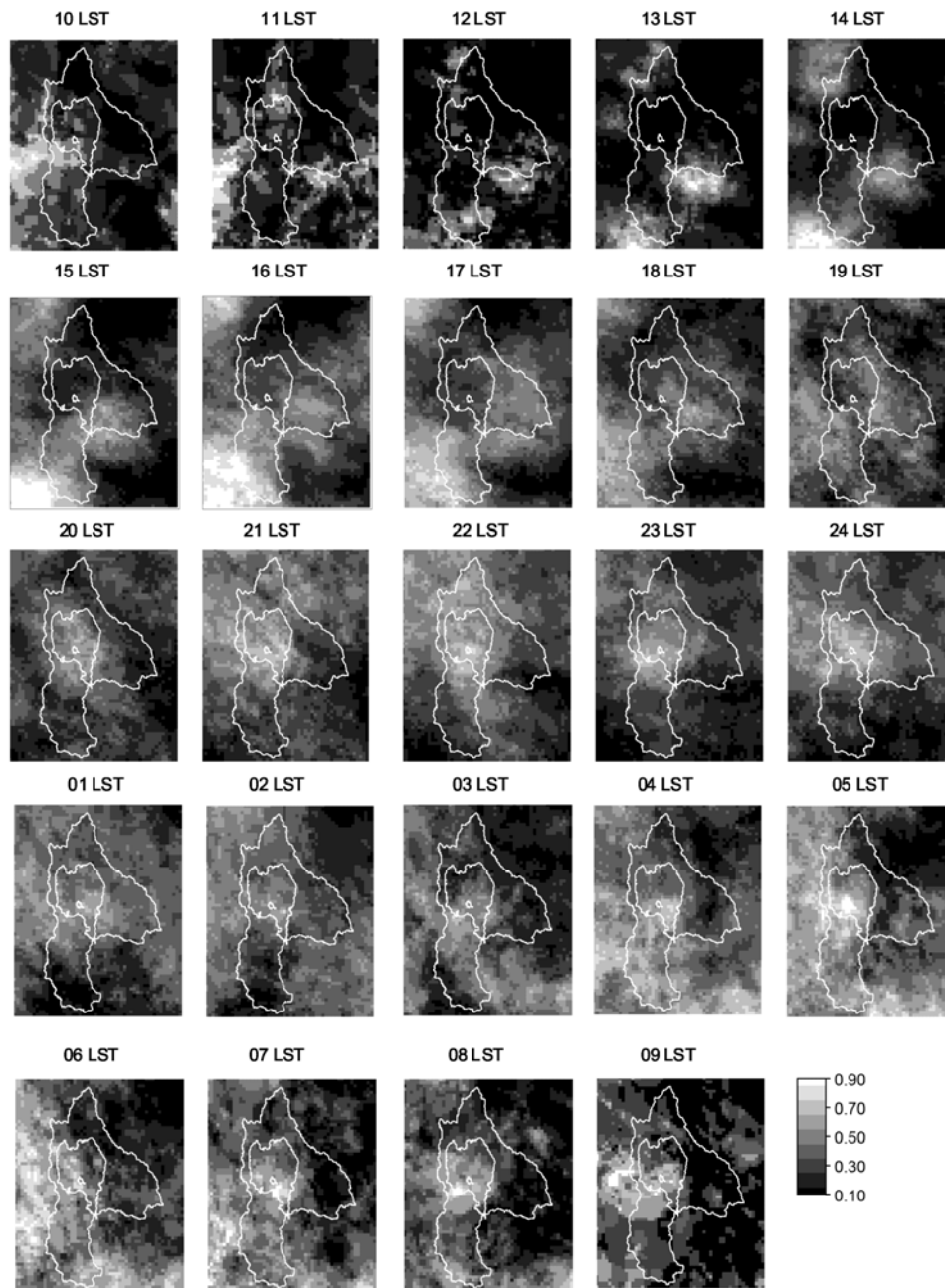


Figure 3.11: *Diurnal variation of the convective index (CI).*

A CI maximum first occurs over the Gilgel Abbay catchment between 1600 and 1700 LST (see figure 3.13a). This is in agreement with

observations from the rain gauges that were placed in the same catchment. The *CI* maxima over the Gumara and Ribb catchments were observed on about 1700 LST. However, in the afternoon, a relatively active convection was observed for a larger time interval over the Gumara than over the Ribb catchment. As such, several cloudy hours prevailed over Gumara than Ribb. The pattern of the diurnal cycle over the Megech catchment appears to be similar to that over Lake Tana, with a *CI* maximum observed between 2200 and 2300 LST.

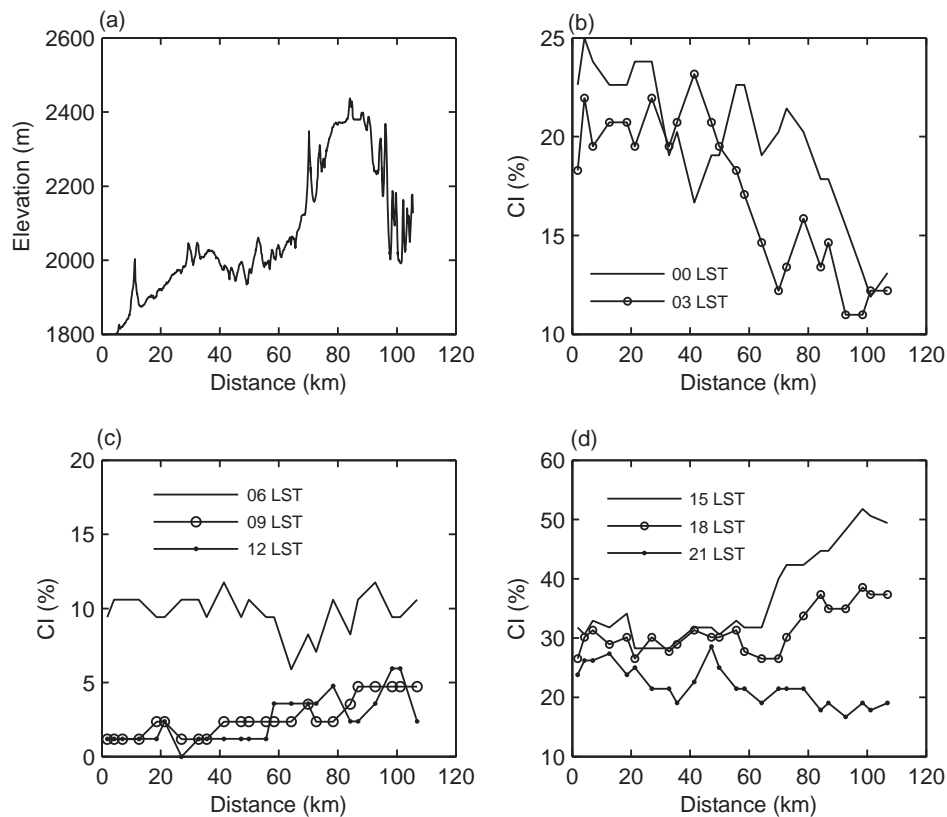


Figure 3.12: Variations of elevation and *CI* along transect A–A' which is shown in figure 3.1. Distance is measured from the Lake near Bahir Dar.

Figure 3.13b shows the coefficient of variation (CV) of the *CI* values at each specific LST. The maximum CV over the four catchments, which is in the range of 1.0–4.0, was observed in the late morning. However, over Lake Tana, a relatively high CV, which is about 1.0, was observed in the early afternoon. Figure 3.13c shows the *CI* distribution covering the shore of

Lake Tana. The highest convective activity can be observed along the southern parts of the lake shore in the afternoon.

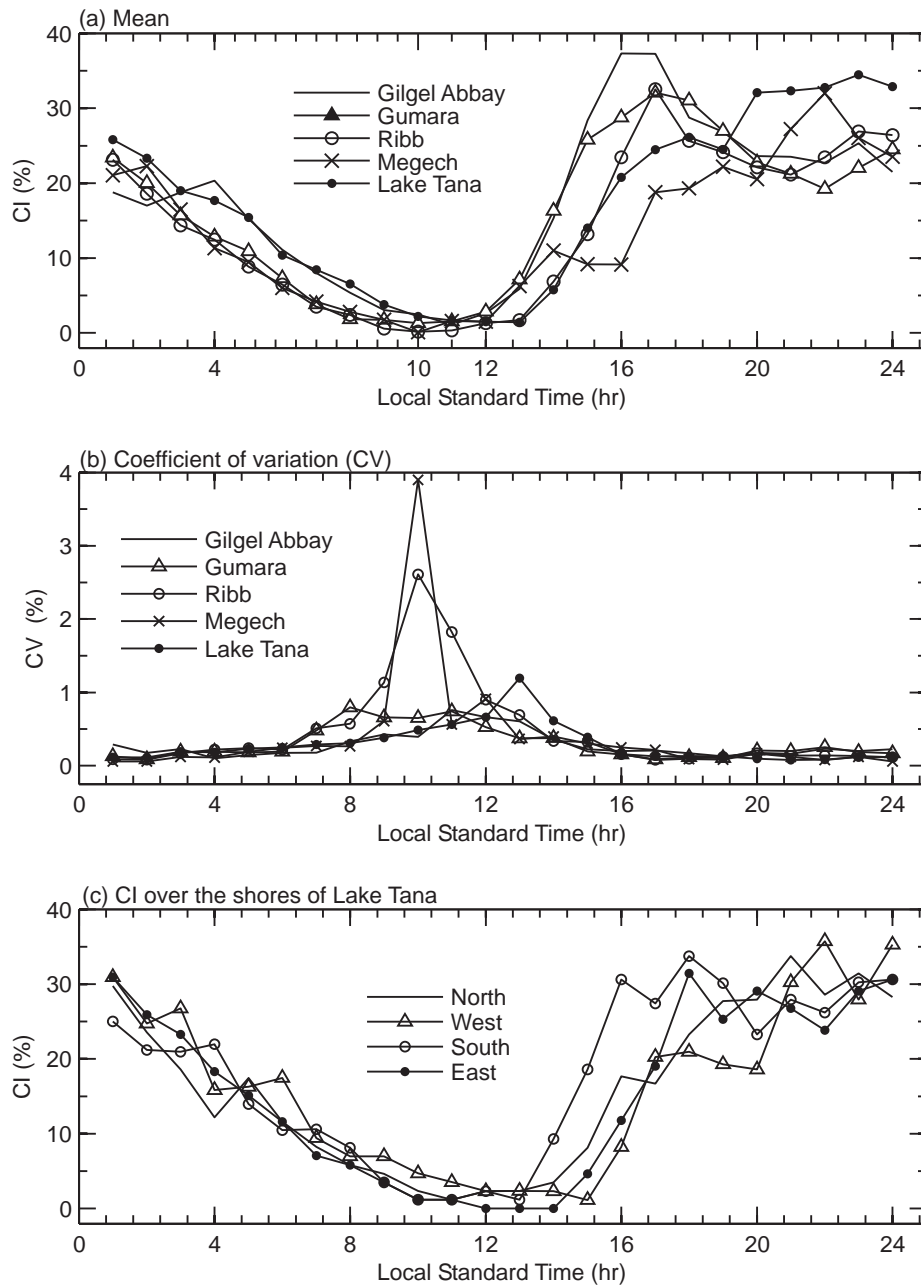


Figure 3.13: Convective index over Lake Tana, its basin area and the shores of the lake.

The temporal and spatial distribution of the maximum *CI* values is also evaluated. The selected *CI* values correspond to brightness temperature values less than 210 K, 210–225 K and 225–240 K that are assumed to represent high, mid and low level clouds, respectively. The *CI* values were divided by the maximum values over the study area to obtain the normalized values. The maximum *CI* is relatively small for high, mid and low level clouds over the eastern ridges of Gilgel Abbay catchment, see figure 3.14. The same figure shows a trend in the occurrence time of the maximum *CI*. Commonly, high and thick clouds first seem to develop over the southern part of the Gilgel Abbay catchment, next over its northern part, and finally over Lake Tana.

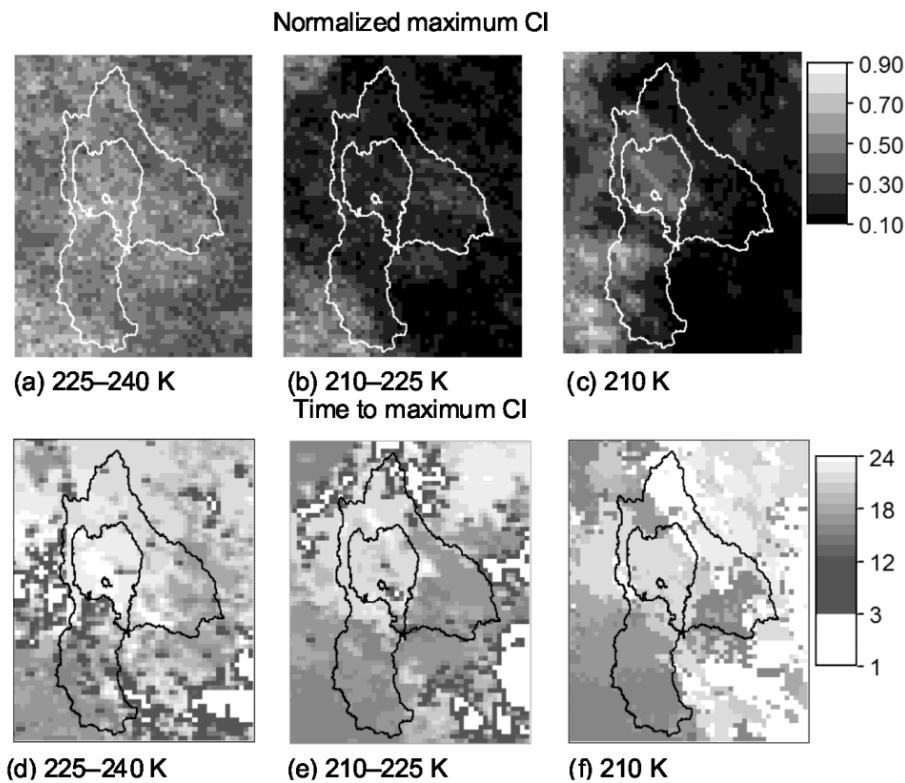


Figure 3.14: *The maximum (peak) CI and Local time to this peak CI for various temperature thresholds. The maximum CI values are divided by the maximum values to obtain the normalized values.*

HARMONIC ANALYSIS

A trigonometric series of the form in equations [3.4 – 3.6] was fitted to the frequency of rainfall occurrences and convective index (CI) data. This procedure was used to determine the amount of the diurnal and semidiurnal rainfall variations and the degree to which the first two harmonics represent the observed variations. The CI values are estimated for temperature values less than 240 K.

The fitted diurnal and semidiurnal cycles of the area-averaged CI over Lake Tana and its major tributaries, and the frequency of rainfall occurrences at five stations in the basin are shown in figure 3.15. The variance of the observed CI cycle that is explained by both the first and second harmonics is in the range of 77–99 % with the first harmonic explaining most of the variance. The amplitude of the semidiurnal cycle is found to be relatively small over Lake Tana and Megech catchment, north of the lake.

The variance in the observed variation of the rainfall frequency that is explained by both the first and the second harmonics is in the range of 63–76 % (see also table 3.5). The variance that is explained by the first harmonic is in the range of 28 % (at Dangila) and 76 % (at Bahir Dar). However, the variance explained by the second harmonic and the amplitude of the semidiurnal cycle at Bahir Dar which is on the lake shore was found very small. For this station, the second harmonic explains 0.4 % of the variance in the observed diurnal cycle.

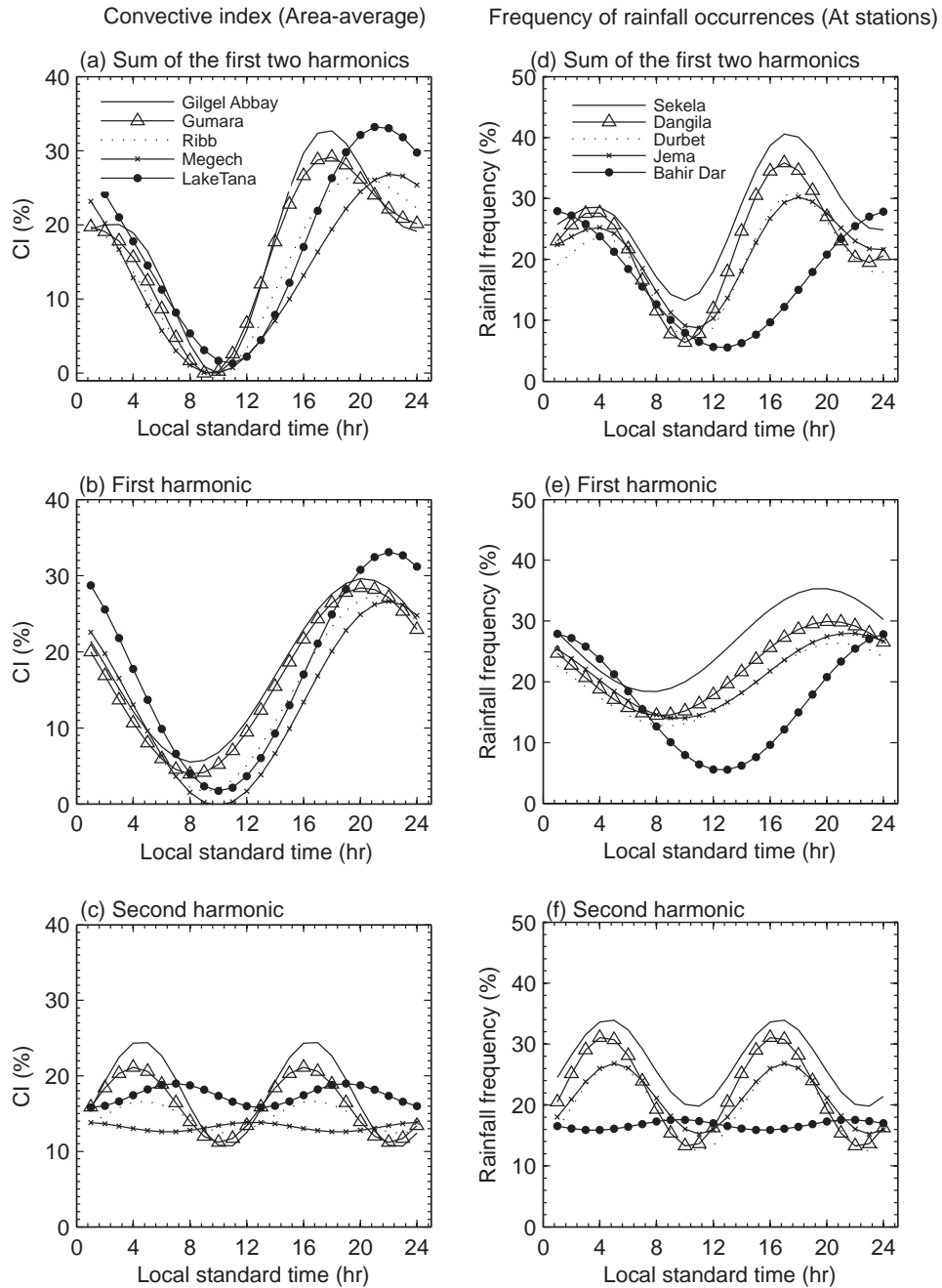


Figure 3.15: Results after applying a harmonic analysis to the rainfall frequency and the CI. Results are only for 5 rain gauges, the complete results are given in table 3.6.

Table 3.5: *Results from the harmonic analysis.*

Stations	Variance explained (%)		Sum of the two harmonics	
	First harmonic	Second harmonic	Amplitude (%)	LST to the amplitude (h)
Rainfall occurrence				
Bahir Dar	76	0.4	28	00
Durbet	32	36	31	18
Jema	38	26	30	18
Sekela	43	31	41	17
Injibara	48	26	49	17
A. Kidam	48	20	33	18
Dangila	28	39	36	17
Koga	36	27	24	17
Convective index				
Gilgel Abb	63	22	33	18
Gumara	70	11	29	18
Ribb	75	2	26	20
Megech	99	0	27	22
LakeTana	92	1	33	21

Table 3.6 shows the parameters of the harmonic analysis that are fitted to the observations. The phase angle is interpreted as the time to the amplitudes, which can be obtained from the phase angle (φ) divided by the frequency (ω). The mean of the diurnal cycle varies between 16 and 17 % with large mean values found at the mountain stations and low values found at stations that are relatively near Lake Tana. However, the maximum amplitude of the diurnal cycle is found at the mountain station (Injibara) and the lake shore (Bahir Dar) while low values are found at intermediate stations between the mountain and the Lake shore. The diurnal cycle peaks between 18:00 hour at Injibara and 00:00 hour at Bahir Dar. The amplitude of the semi diurnal cycle at the lake shore, i.e. Bahir Dar, is 0.85 % only while that at the mountain station, i.e., at Injibara is about 9 %. The time to the amplitudes of the semi diurnal cycles at each station is also presented in table 3.6.

Table 3.6: *Fitted parameters of the harmonic analysis. Note: phase angle is interpreted and presented as time to the amplitude of the diurnal and semidiurnal cycles.*

Stations	Mean of the diurnal cycle (%)	Diurnal cycle		Semi diurnal cycle	
		Amplitude (%)	Time to the amplitude (hr)	Amplitude (%)	Time to the amplitude (hr)
Bahir					
Dar	16.72	11.25	0.60	0.85	3.6 and 15.6
Durbet	19.56	6.82	20.80	7.20	5.0 and 17.0
Jema	21.01	7.00	21.60	5.79	5.0 and 17.0
Sekela	26.86	8.51	19.60	7.17	4.6 and 22.6
Injibara	28.30	12.40	18.10	9.08	4.8 and 16.8
Addis					
Kidam	20.06	8.85	19.50	5.72	5.1 and 17.1
Dangila	22.19	7.75	20.20	9.04	4.4 and 16.4
Koga	17.35	5.61	21.80	4.89	4.5 and 16.5

3.4.4. Fractals of rainfall intermittence

The fractal dimension of the rainfall intermittency is estimated to analyse the scaling behaviour of the rainfall observations. Figure 3.16 shows a log-log plot of the number of boxes that contain rain, $N(\lambda)$, against the size, λ , of the boxes. The same figure shows that the rain-no rain distribution is scale-invariant. Two regimes that fit equation [3.7] are identified. The first scaling regime extends from 2 hour to 8 hour while the second regime extends from 2.7 days onwards.

For the first scaling regime, the values of the slopes are determined with a maximum standard error of about 0.02. The slopes also define the fractal dimension that ranges from 0.42 to 0.50 with a standard deviation of 0.03. High values of the fractal indicate temporal clustering of rainfall. The correlation between the fractal dimension and elevation is 0.69 when all stations are considered and it is 0.74 for the stations on the west. Thus, the rainfall at higher elevation areas are mostly characterised by closely clustered rainfall events as explained by larger fractal dimensions. For the second scaling regime, the fractal dimension for all stations is 1 which indicates that for the time period between June 1 and August 25, rainfall was observed at least once in 2.7 days.

3.5. DISCUSSION

The results of this study revealed the variation of rainfall and its relation to elevation and distance to a lake at the source of the Blue Nile River. A high difference in seasonal rainfall was observed between Durbet and Jema stations that are 12.5 km apart. The Jema station has a higher seasonal (JJA) rainfall, more frequent rainfall, a higher hourly rainfall and a larger coefficient of variation than the Durbet station. The rainfall depths recorded in JJA 2007 show a difference of about 375 mm between the rainfall of the two stations. These stations are at about the same distance from the main valley axis and the center of Lake Tana while Durbet is situated about 15 m higher than Jema. The terrain slope of Durbet and Jema is 2.07 % and 2.48 % with an east and west facing terrain slope, respectively. Therefore, the difference in rainfall may be partly caused by the differences in terrain aspect. However, multi-annual observations are required before definite conclusions can be drawn.

The shape parameter of the correlation structure of the hourly rainfall as determined here is much less than the commonly reported values, e.g., Ciach and Krajewski (2006) and Young et al. (2000). This suggests that the rainfall distribution varies substantially over a relatively small domain. The correlation distances (d_0) of the hourly rainfall and daily rainfall are determined here at about 9 km and 18 km, respectively. However, d_0 was found to be sensitive to the correlation at zero-distance. Thus, care must be exercised in interpreting the correlation structure for sparse rain gauges. In some cases, the lag-1 hour cross-correlation was found higher than the lag-0 hour cross-correlation suggesting a possible movement of clouds with a speed less than the inter-station distance per hour. These results reveal information for the estimation of area-averaged rainfall and possible improvements to the existing rain gauge network.

A north-south stretching trend in the spatial variation of the rainfall diurnal cycle was observed. The maximum frequency of rainfall events occurs between 1600 and 1700 LST over the mountain areas while the maximum frequency occurs around 2300 LST over the southern part of the lake shore at Bahir Dar. Using *in situ* rainfall data and thermal infrared images, Ba and Nicholson (1998) showed that the maximum rainfall over Lake Victoria is observed between 0500 – 0800 LST. As such, the rainfall over Lake Tana peaks earlier than the rainfall over Lake Victoria.

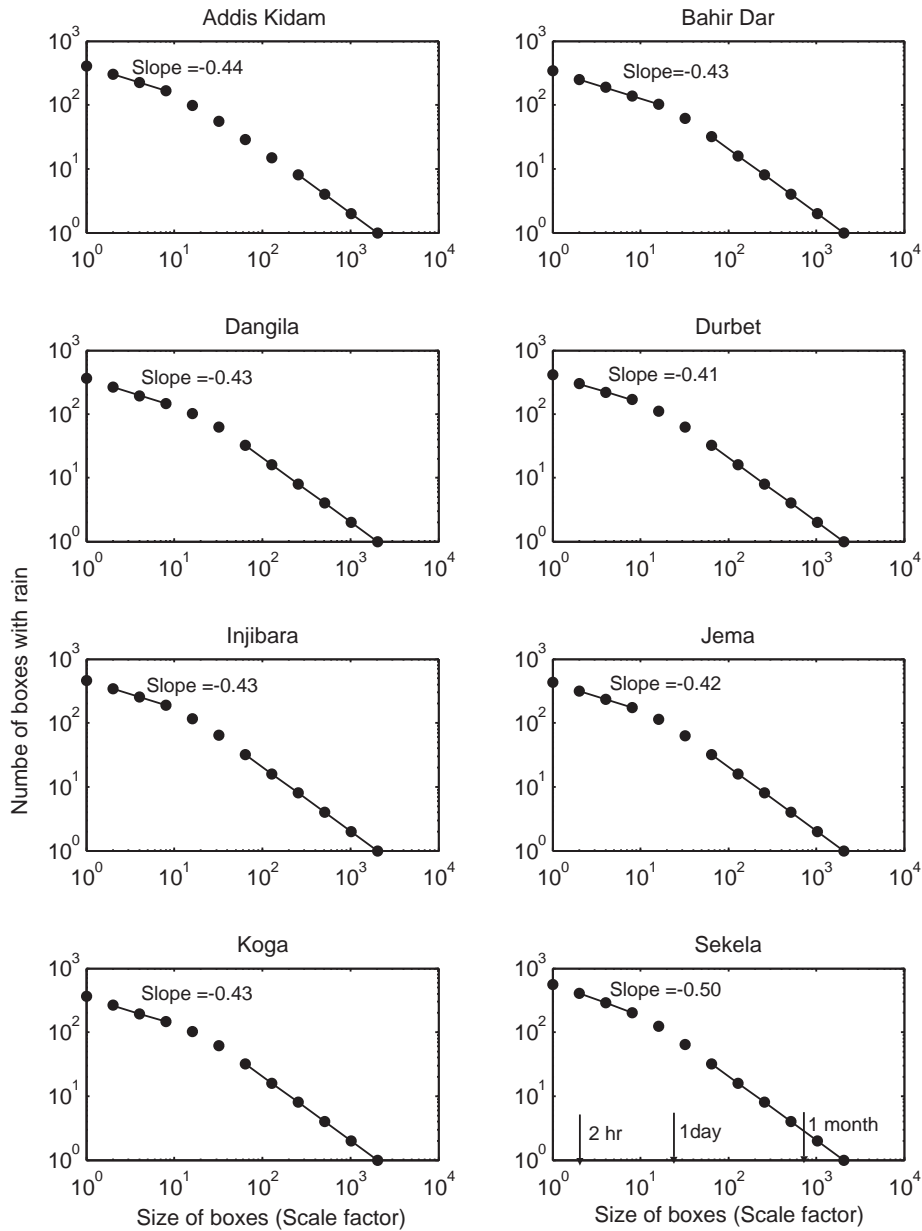


Figure 3.16: Log-log plot of Number of boxes with rain ($N(\lambda)$) against the size of boxes, i.e. the scale factor (λ). Note: The slope of the second scaling regime is 1.0 for all the stations.

It is also shown that convective index (CI) values that are computed from remote sensing data can be used as proxy variables to estimate convective activity over the lake and its ungauged catchments. In general, it was observed that in the afternoon, the CI value increases from the area near Lake Tana to the mountain areas of Gilgel Abbay catchment. The reverse was observed in the evening and early morning hours. The CI revealed that the afternoon conditions in the north-eastern and south-western parts of the lake shore are less favourable for convective development as compared to the shores on the other sides of the lake.

The harmonic analysis of the lake CI results show a similar result with that of the rainfall at Bahir Dar that suggests the second harmonics does not have much contribution to the diurnal variation near and over the lake. However, there is some difference in the time to the amplitude of the rainfall over the lake and its shore. The harmonic analysis of the Gilgel Abbay CI results show a similar result with that of the rainfall at most of the stations in the catchment that suggest the LST to the amplitude is 1700 or 1800 LST. The amplitude of the Gilgel Abbay CI is within the range of the observed amplitude at the stations.

The fractal analysis revealed that the rainfall intermittence is scale-invariant and that, based on a Pearson's correlation of 0.69, a direct relation could be established between fractal dimension and terrain elevation. Except for the highest elevation area, the difference in the fractal dimension values is within the standard error of the estimates. The relation suggests that the rainfall at higher elevation area is mostly characterised by closely clustered rainfall events. However, the smallest fractal dimension does not correspond to the lowest elevation that is Bahir Dar which is also close to the lake. Instead the smallest value of the fractal dimension is found at Durbet which recorded the lowest seasonal rainfall. The result is not in agreement with Gebremichael et al. (2007) who reported closely clustered rainfall events at low elevations for the rainfall in a mountainous area of Sonora, Mexico.

This study provides relevant information for hydrologic and climatic studies in the basin. Information on rainfall variability at high temporal resolution is required for the simulation of runoff production as well as real-time runoff forecasting for the flood prone area. Information on rainfall intensities and its temporal distribution is highly relevant to study gully erosion that largely has affected the basin. Information on the diurnal and spatial variability of rainfall is commonly required by climatic studies

and this study provides such information for a mountainous area that is adjacent to a lake. In the future, it is recommended to install hydro-meteorological sensors, for sub-daily observations, on the ungauged catchments of the Lake Tana basin and on the islands of the lake. This will provide data that is necessary to study the rainfall characteristics over a longer time period.

3.6. CONCLUSION

Several statistical techniques are applied to determine the pattern in rainfall variation with respect to terrain features and distance to lake. The results suggested that variation of rainfall at the source of Blue Nile River is affected by terrain elevation and distance to the centre of the lake. The lowest conditional mean rainfall of the hourly observations was observed over the mountains of Sekela where also the hourly rainfall with the smallest coefficient of variation was observed. There appears to be less rainfall variation over high elevation areas than low elevation areas of the Gilgel Abbay catchment close to Lake Tana. The lowest probability of rainfall occurrences was found on the southern part of the lake shore, i.e. Bahir Dar. Heavy rainfall events of higher than 10 mm hr^{-1} were frequent at stations relatively close to Lake Tana. Light to moderate hourly rainfall of less than 10 mm was frequent over mountain areas. In addition, frequent rainfall and convective activity were observed in the afternoon over the southern mountains and in the night over the southern part of the lake shore. The amount of nocturnal rainfall over the lake shore was about 75 % of the total and is much higher than the nocturnal rainfall over the mountain areas. The results of the fractal analysis revealed that the rainfall at the station with the highest elevation is mostly characterised by closely clustered rainfall events.

The CI analysis revealed that, commonly, the time at which high and thick clouds were observed showed a spatial trend stretching from mountain areas towards Lake Tana. High and thick clouds first occurred over mountain areas of Gilgel Abbay (in the afternoon), next over valleys and lowlands of Gilgel Abbay (in the late afternoon), and finally over Lake Tana and Megech catchment (in the night). The amplitude of the semidiurnal cycle of convective activity was found to be rather small over Lake Tana and the Megech catchment, north of the lake. Also, the

Rainfall variability over mountainous and adjacent lake area

amplitude of the semidiurnal cycle of the rainfall frequency at Bahir Dar which is on the lake shore was found very small.

Rainfall variability over mountainous and adjacent lake area

4

RAIN EVENT PROPERTIES



ABSTRACT¹

An understanding of rain event properties over space and time is of great importance for hydrology as well as for climate studies. Rainfall data for the present study has been obtained from a field campaign in two wet seasons of June – August (JJA) of 2007 and 2008 in the Gilgel Abbay watershed that is situated at the source basin of the Upper Blue Nile River in Ethiopia. The rainfall data was recorded at eight stations. In the present study, the spatial and temporal patterns of the rain event properties are analysed. The event properties are rain event depth, event duration, mean event intensity, peak intensity and the time span between two consecutive rain events which is referred to as inter-event time (IET). Dimensionless event hyetographs are established by relating fractions of event intensities to the corresponding fractions of event durations. The spatial variation of the characteristics of hyetographs is also evaluated. A model in the form of the beta distribution function is applied to reproduce the dimensionless hyetographs. The results reveal that rain event depth is more related to peak intensity than to event duration. At the start and towards the end of the wet season, the rain events have larger depth with longer duration and longer IET than the rain events in the mid-season. Mean event intensity and IET are strongly related to terrain elevation. Overall, Sekela which is on a mountain area has the shortest IET while Bahir Dar which is at the south shore of the lake has the longest IET.

Keywords: Blue Nile River, Lake Tana, Rainfall, rain event

¹ This chapter is based on: Haile, A. T., Rientjes, T., Jetten, V., 2009: **Rain event properties and dimensionless rain event hyetographs at the source of the Blue Nile River.** *Hydrological Processes*. Prepared for submission.

4.1. INTRODUCTION

An understanding of rain event properties at specific spatio-temporal scales is of great importance for both hydrology as well as for climate studies. According to Brown et al. (1985), rain events are a convenient way of summarizing a time series of rainfall amounts into entities that are defined so that they are meaningful in terms of a particular application. Such applications include studies of runoff generation (e.g. Kusumastuti et al., 2007; Milly and Egelson, 1987a), soil erosion (e.g. Angel et al., 2005), interception losses (e.g. Zeng et al., 2000) and rainfall modeling (e.g. Wooliser and Osborn, 1985).

In rainfall-runoff modeling, daily or coarser resolution observations of rainfall commonly serve as model inputs. By use of such observations, the assumption of spatial uniformity of rainfall is commonly applied and therefore spatially averaged rainfall estimates can serve as inputs to rainfall-runoff models, e.g. Haile et al. (2009d). However, Milly and Egelson (1987b) stated that the constraining assumption of spatial uniformity is not especially helpful in the analysis of single events. Properties of single rain events often vary at scales much smaller than a watershed scale and event properties consequently affect runoff generation.

Lack of high resolution observations of rainfall often restricts the applicability of rainfall-runoff models at high temporal and spatial resolutions. In literature, it is shown that this restriction can be overcome by the use of stochastic rainfall models. For instance, Kusumastuti et al. (2007) evaluated the effects of catchment storage on runoff generation by coupling a stochastic rainfall model with deterministic rainfall-runoff models. Their stochastic rainfall model considers variability of rain event duration and inter-event time (IET) and the occurrences of infrequent tropical cyclones. For a synthetic watershed, Kusumastuti et al. (2007) showed that variability within a rain event which is referred to as intra-event variability can largely affect quick runoff generation mechanisms and consequently affects flood generation. Through rainfall-runoff modeling with the objective to study the scaling behaviour of flood frequency, Robinson and Sivapalan (1997); Menabde and Sivapalan (2001) showed that the runoff discharge is proportional to the watershed area for relatively small watersheds with a response time less than the mean duration of rain events. However, the

authors noted that the stream flow at the watershed outlet will be affected by the branching structure of the stream network in case the response time is greater than the mean duration of rain events.

Rain event properties also serve as main inputs to soil erosion studies. For instance, the maximum 30-minute rainfall intensity is one of the main inputs to the Revised Universal Soil Loss Equation (RUSLE), e.g. Diodato and Bellocchi (2007). However, lack of long-term rainfall records at the required spatial and temporal resolution often restricts the applicability of erosion models in many regions. van Dijk et al. (2005) state that this problem has been overcome in literature in two ways: (i) by using synthesised idealised rain events, or (ii) by using empirical relations between hydrological model variables and more readily available rainfall characteristics. An example of the second one is the empirical relation between the rainfall erosivity factor within RUSLE and the monthly or annual rainfall amount, see Renard and Freimund (1994); Diodato and Bellocchi (2007). The use of synthesised or empirical relations to derive rain event properties can benefit from information about these properties on a short term base through observation and analysis.

To overcome restrictions by lack of high resolution data, sub-daily rainfall data can be generated from daily or coarser-resolution rainfall data by rainfall disaggregation models, see Hingray and Haha (2005); Segond et al. (2006); Arnaud et al. (2007); Koutsoyiannis and Onof (2001). Rainfall disaggregation models extend the short-term records of sub-daily rainfall to long-term data of sub-daily rainfall by disaggregating the long-term daily records which are commonly available. However, the calibration of disaggregation models requires availability of short-term sub-daily records at one or more rain gauge stations, e.g. Arnaud et al. (2007); Koutsoyiannis and Onof (2001).

Huff (1967) established a relation between the fraction of rain event depth and the corresponding fraction of rain event duration through dimensionless hyetographs. These hyetographs are also called "Huff curves", see Tsubo et al. (2005). Huff (1967) showed that the dimensionless hyetograph for the area in Illinois was affected little by event depth and event duration. Woolhiser and Osborn (1985) observed that the dimensionless hyetograph of an experimental watershed in South-eastern Arizona was affected by event duration. Garcia-Guzman and Aranda-Oliver (1993) advocated evaluating the assumption that the hyetographs are invariant with respect to event depth, duration and area averaging.

However, such requires large data sets covering several years to obtain sufficient data for statistical analysis.

Rainfall models that are based on dimensionless hyetographs are presented in Woolhiser and Osborn, 1985; Garcia-Guzman and Aranda-Oliver (1993). Other examples are Bras (1979) who applied the dimensionless hyetograph concept to study the sampling of the rainfall-runoff process and Tsubo et al. (2005) who developed dimensionless hyetographs for the Highveld of South Africa. Similar studies, however, are noticeably absent for the Upper Blue Nile Basin in Ethiopia.

The use of only one or two rain gauge stations may not be adequate to evaluate the spatial pattern of rain event properties (see Brown et al., 1985; Tsubo et al., 2005). To analyze the effect of terrain attributes such as elevation on rain event properties, see Loukas and Quick (1996); Palecki et al. (2005), records from spatially distributed rain gauge stations are required. However, studies that report on the spatial and temporal patterns of rain event properties are absent for many geographic locations and such also applies for the Upper Blue Nile basin.

As part of the present study, a network of 10 recording rain gauges was setup in May 2007 at the source basin of the Upper Blue Nile River. From the network, eight stations recorded the rainfall data of the full wet season of the year 2007, i.e. June – August (JJA) while only 2 rain gauges recorded the full wet season rainfall of 2008. The remaining gauges did not record the full season rainfall probably due to human interference.

In Haile et al. (2009a), the orientation of the rain gauges is presented and the spatial patterns of the diurnal cycle of the basin rainfall are analyzed which showed that both orography and the presence of Lake Tana affects the spatial pattern of the diurnal cycle, see Chapter 3 of this thesis. Using the same rain gauge network as in the present study, Haile et al. (2009d) showed that runoff simulations in the basin are largely sensitive to rainfall representation and reported runoff volume errors as large as 15 – 40 % when the model input is only from 3 – 5 rain gauges instead of 8 rain gauges, see Chapter 3 of this thesis. However, these studies did not explicitly explore the rain event properties in the basin.

In the present study, the temporal patterns of rain event properties are analyzed using rainfall records of Jema station which recorded the rainfall data for the wet season of the years 2007 and 2008. This station is situated at an intermediate location between the mountains of the Gilgel

Abbay watershed and Lake Tana. Haile et al. (2009d) suggested that Jema station should be part of any rain gauge network that is applied to estimate rainfall input to a runoff model, see Chapter 7 of this thesis.

In this Chapter, the spatial patterns of rain event properties are evaluated using data from eight stations that recorded the JJA 2007 recorded. Rain event properties that are of general interest to the study of hydrological processes are rain event depth, intensity, duration and inter-event time (IET). Also, dimensionless event hyetographs are developed for selected stations while the spatial variation of the characteristics of the hyetographs is evaluated. Finally, a beta distribution type model is fitted to the observed dimensionless hyetographs.

4.2. METHOD OF ANALYSIS

4.2.1. Properties of rain events

Selected criteria to identify rain events have influence on the properties of rain events that are derived. For instance, the number, duration and intensity of rain events can change subject to the criterion to identify events (see Dunkerley, 2008b). Commonly, the minimum inter-event time (MIT) and/or the minimum event depth serve to identify rain events or to classify a series of rain observations as a single event. Dunkerley (2008a, b) reviewed a large number of criteria and reported that the applied MIT varies from 3 min to 24 h while the minimum event depth varies from the measurable amount that is the resolution of the rain gauge to 13.0 mm. The consequence of a change in criterion is for instance that the mean event intensity may decrease as the MIT increases since the events may include several rainless periods.

By absence of a commonly applied criterion in literature, a set of criteria has been chosen to identify rain events, i.e. to mark the start and end time of the events. The applied criteria are a minimum event depth of 1.0 mm and a minimum inter-event time (MIT) of 30 minute. According to Brown et al. (1985), rain events identified by such criteria can be termed “primitive” events and are based on the simplest and most logical approach without the use of additional information regarding the synoptic weather condition such as cloudiness.

The rain event properties that are analyzed in the present study include event depth, event duration, mean event intensity, peak intensity of an event and inter-event time (IET) which is defined as the length of the dry period between two consecutive events. These event properties are selected since they are considered to be relevant to hydrologic studies. The reader is referred to Dunkerley (2008a,b) for further discussion on the importance of these event properties in hydrologic studies.

The event depth (d_e) is defined as:

$$d_e = \sum_{i=1}^{E_d} I_i \quad [4.1]$$

where: E_d is the event duration while I_i is the one-minute rainfall depth at the i^{th} minute since the start time of the rain event.

The mean event intensity (R_e) is:

$$R_e = \frac{d_e}{E_d} \quad [4.2]$$

where: d_e and E_d are the depth and the duration of the rain event, respectively.

The peak intensity (I_{max}) is:

$$I_{max} = \max_{1 \leq i \leq E_d} (60I_i) \quad [4.3]$$

where: max indicates the maximum while the constant 60 is applied to convert the one-minute rainfall depth to hourly rainfall intensity. The remaining terms are as defined previously.

Relations between the various event properties are evaluated through the Pearson's product-moment correlation. The effect of terrain elevation on the event properties is also evaluated.

4.2.2. Dimensionless event hyetographs

A dimensionless hyetograph relates the fractions of rain event depth to the respective fractions of rain event duration. Following Garcia-Guzman and Aranda-Oliver (1993); Tsubo et al. (2005), the 10 %, 50 % and 90 % fractions of event depth are selected to develop dimensionless hyetographs

since these correspond to small, normal and large event depths. The spatial variability of the properties of the hyetographs is evaluated using rainfall records at three stations. These stations are Bahir Dar which is at the south shore of Lake Tana, Sekela which is on the mountain area and Jema which is at an intermediate location between Bahir Dar and Sekela. Haile et al. (2009a) showed that the diurnal cycle of the rainfall at Bahir Dar is largely different from the other stations in the basin while in Haile et al. (2009d) it is shown that the rainfall of Jema and Sekela stations largely affect the accuracy of runoff simulations in the Gilgel Abbay watershed of Lake Tana basin.

The dimensionless hyetographs at the selected stations are evaluated graphically and statistically. Following Tsubo et al. (2005), the Kolmogorov-Smirnov goodness of fit test is applied to evaluate whether differences between the hyetographs are statistically significant. The absolute value of the difference (D) in the fractional event depths of two dimensionless hyetographs at the increment j of the fraction of the event duration reads:

$$D_j = \left| H_{j,1} - H_{j,2} \right| \quad [4.4]$$

where: H_j is the fraction of the accumulated rain event depth over the time interval $(0,t)$ where t is time instant which is expressed as the fraction of the total event duration, for instance $t = j/10$ for process with 10 increments where $j = 1, 2, 3, \dots, 9, 10$. The subscripts 1 and 2 indicate the two dimensionless hyetographs that are compared.

The test statistic (D) for the Kolmogorov-Smirnov test is defined as:

$$D = \max(D_j) \quad [4.5]$$

where: \max indicates that the maximum D_j in equation [4.4] is used as the test statistic. The critical values $D_{\alpha,n}$ for a selected significance level α and number of increments n are presented in text books on statistics such as Zar (1996) that is used in this study. The null hypothesis is that the two hyetographs are from the same statistical distribution and therefore will be rejected if D in equation [4.5] exceeds $D_{\alpha,n}$.

A model in the form of a beta distribution function is applied to reproduce the observed dimensionless hyetographs by assuming that the increment process of the rain events can be represented by a beta

distribution function. The increment process Z_j for $0 < Z_j < 1$ at a fixed time is, see Woolhiser and Osborn (1985); Garcia-Guzman and Aranda-Oliver (1993):

$$Z_1 = H_1 \quad [4.6]$$

$$Z_j = \frac{H_j - H_{j-1}}{1 - H_{j-1}} \quad [4.7]$$

where: H_j is as defined previously and $j = 2, 3, \dots, 9$.

Following Garcia-Guzman and Aranda-Oliver (1993); Tsubo et al. (2005), a model in the form of the beta distribution function is fitted to the distribution of the increment process. The beta distribution function for the increment process Z is:

$$f(Z) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot Z^{\alpha-1}(1-Z)^{\beta-1} \quad [4.8]$$

where: Γ is the gamma function, $\alpha > 0$. Both are shape parameters which are fitted using techniques of optimization. In this study, the maximum likelihood method is applied using a built in function in Matlab®. Garcia-Guzman and Aranda-Oliver (1993) described that when $\alpha < \beta$, it is more probable to have all observed H_j close to zero than close to one while for $\alpha = \beta > 1$, the beta density is symmetrical and has a maximum at $t = 0$. For $\alpha = \beta < 1$, the beta density has a minimum at $t = 0$. The data base in the present study is considered too limited to evaluate the sensitivity of the model parameters to factors such as rain event duration.

4.2.3. Conditional probability of rainfall occurrences

The estimation of the conditional probability of rainfall occurrence at any two stations is important to understand the spatial structure of rainfall intermittence in particular the size of rain events in terms of geographic location. Examples of such studies are found in Gebremichael et al. (2007) who studied the spatial behavior of hourly rainfall intermittence for an area in Sonora, Mexico and in Robinson (1994) who evaluated the intermittence of daily rainfall for Colorado and North Carolina. Haile et al. (2009a), see Chapter 3 of this thesis, evaluated the temporal structure of the

hourly rainfall intermittence in the Gilgel Abbay watershed but ignored the spatial structure of the rainfall intermittence that is evaluated in the present study.

The probability (P_{ij}) that it is raining at station j while it is raining at station i as well can be estimated through the conditional probability that is defined as:

$$P_{ij} = \frac{r_b}{r_b + r_i + r_j} \quad [4.9]$$

where: r_b is the number of time steps when there was rain at both stations; r_i is the number of rainfall occurrences at station i while it is not raining at station j ; r_j is the number of rainfall occurrences at station j while it is not raining at station i . $P_{ij} = 0$ shows that it does not rain at all at station j when it is raining at station i while P_{ij} shows that it always rains at station j when it is raining at station i . Note that P_{ij} is not necessarily equal to P_{ji} . The conditional probability of rainfall occurrences has been estimated for an arbitrary chosen time interval of 1 hour, 6 hour and 1 day to evaluate the change in probability with changing time scales. The use of 6 hour and 1 day intervals allows considering effect of event durations longer than 1 hour. An arbitrarily chosen rainfall depth of 1 mm is defined to differentiate rainy from non-rainy time periods.

4.3. RESULTS

4.3.1. Rain event properties

In this section, the rain event properties at Jema station are analysed. Jema station is selected since this rain gauge recorded rainfall data for two consecutive wet seasons that are in JJA of 2007 and 2008. Also, the station is located at somewhat an intermediate distance between the mountain areas of Gilgel Abbay watershed and Lake Tana and therefore, following the work in Haile et al. (2009a), which is presented in Chapter 3 of this thesis, the rain event properties are, presumably, affected by orography and the presence of the lake. A total of 236 rain events have been observed at Jema with 122 events in JJA 2007 and 114 events in JJA 2008.

Rain event properties

Table 4.1 shows the statistics of the rain events at Jema station. A maximum event depth of 67.6 mm has been recorded in the two seasons. This event depth is 12 times higher than the median of the observed event depths which is only 5.6 mm indicating that the event can be considered an extreme event. In terms of event duration, a single event has lasted up to 344 minutes, i.e. about 5.7 hours. However, the median of the duration of the rain events is only about 1 hour which is 6 times shorter than the maximum duration.

Table 0.1: *Statistics of rain events at Jema in two wet seasons that are JJA of the years 2007 and 2008.*

Statistics	Depth (mm)	Duration (min.)	Peak intensity (mm h ⁻¹)	Event intensity (mm h ⁻¹)	Inter-event time (h)
Minimum	1.20	3.00	12.00	0.93	0.52
Maximum	67.6	344.00	180.00	54.67	94.62
Mean	9.11	76.55	42.56	10.01	17.12
Median	5.60	60.00	36.00	6.18	13.42
Std. Dev.	9.84	66.97	35.45	10.45	16.67

The maximum peak intensity at Jema is 180 mm h⁻¹ which is equivalent to 12 bucket tips per minute. In terms of median, the peak intensity is 36 mm h⁻¹ which is equivalent to 3 bucket tips per minute. The median of the peak intensity is 6 times the median of the mean event intensity indicating large differences which is partly explained by the presence of several rainless periods in a single rain event that causes low mean event intensity.

Mean event intensities range between 0.93 mm h⁻¹ and 54.67 mm h⁻¹ with a median of 6.18 mm h⁻¹. The median of the IET is 27 times the specified MIT which is 30 minutes. Such shows that changing the MIT value, for instance, to 1 hour will not largely change the statistics of the estimated event properties. In terms of median, the IET at Jema is about 13.42 hours that is about 13 times the median of the rain event duration indicating longer dry periods than event duration. The median value of the IET also shows that the daily rainfall depths at Jema are mostly caused by one or two rain events. It is noted that the presence of multiple events in a day restricts the applicability of daily rainfall records for event based analysis. As such, the rainfall needs to be recorded at sub-daily time scales for such analysis.

4.3.2. Relation between rain event properties

The relation between the rain event properties has been evaluated for the rainfall records of Jema station during the two consecutive wet seasons. Table 4.2 shows the lower triangle of the correlation matrix between the rain event properties. The depth of the rain events is positively correlated to all of the remaining event properties. Overall, event depth increases with an increase in event duration or mean event intensity. Event depth has the strongest relation with peak intensity which reveals that event depth at Jema is more related to peak intensity than event duration.

Table 0.2: Lower triangle of correlation matrix between the properties of the selected rainfall events at Jema. The events are observed in two wet seasons that are JJA 2007 and 2008.

	Depth	Duration	Mean intensity	Peak I_1	Inter-event time
Depth	1.000				
Duration	0.446(*)	1.000			
Mean intensity	0.481(*)	-0.313(*)	1.000		
Peak I_1	0.739(*)	0.028	0.800(*)	1.000	
Inter-event time	0.059	0.046	0.087	0.067	1.000

* Correlation is significant at the 0.01 level (2-tailed)

Although the relation is not strong, rain event duration is negatively correlated to mean event intensity and therefore events with short duration have relatively high mean event intensity. Such can be partly caused by high rainfall intermittency that is due to the presence of several rainless periods in a single event as the event duration increases which consequently results in relatively low mean event intensities. The correlation value also shows that events with high peak intensity also have high mean event intensity. Overall, high event intensity at Jema is more related to peak intensity than to event duration. The IET at Jema has shown a weak relation with all of the other rain event properties.

4.3.3. Temporal variation of rain event properties

To evaluate the intra-season variability of rain event properties, the rain event properties at Jema have been analysed for each of the three

Rain event properties

months of the wet seasons of 2007 and 2008. Table 4.3 shows the median of the rain event properties for each of the three months during the wet season. It is noted that June is at the beginning of the wet season while August is towards the end of the wet season. In terms of the median of the rain event depth, there are some differences between the events in June, July and August. June has the highest rain event depth while July has the lowest event depth with a 2 mm difference between the median of the event depth of the two months.

Table 0.3: *The temporal variation of the median of rain event properties at Jema for MIT = 30 min. The events are observed in two wet seasons that are JJA 2007 and 2008.*

Statistics	Depth (mm)	Duration (min.)	Peak intensity, I_1 (mm h ⁻¹)	Event intensity (mm h ⁻¹)	Inter- event time (h)
Median					
June	6.6	64.0	36	6.52	17.6
July	4.7	55.5	24	6.11	11.33
August	5.4	61.0	36	6	13.56
CV					
June	0.998	0.773	0.824	1.002	0.882
July	1.161	0.905	0.860	1.030	0.928
August	1.020	0.931	0.797	1.093	1.024

In terms of the median value, the rain events in July have the shortest duration while the events in June have the longest duration with a 10.5 minute difference between the median of the event durations in the two months. In terms of peak intensity, the events in July have the lowest intensity. In terms of mean event intensity, the rain events in August have the lowest intensity while the events in June have the highest intensity. Also, the July events have the shortest IET which shows that the rain events are likely to be clustered while the June rain events have the longest IET. There is a 6.3 hour difference between the median values of the IETs of the events of June and July.

Overall, the event properties at the start and towards the end of the wet season have somewhat similar characteristics while the events in the mid-season differ in this respect. At the start and towards the end of the wet season, the rain event depth is large with long event duration and long IET as compared to events in the mid-season.

The coefficient of variation (CV) in table 4.3 shows that the rain event depth in July is more variable as compared to that of the other months. Also, there is some difference in terms of the variability of the rain event duration in each of the months. The event duration in June is less variable than for the other months of the wet season.

Table 4.3 also shows that the peak intensity of the events in August has the smallest variability while the difference in CV of mean event intensity in the three months is less than 10 %. In terms of IET, the largest variability has been observed in August. Overall, the mid-season events have depth, duration and peak intensity that are more variable than for events at the start and end of the wet season.

4.3.4. Spatial variation of rain event properties

In this section, the spatial variation of rain event properties is analysed based on rainfall records from 8 stations. The number of rain events is 122, 100, 140, 112, 177, 122, 105, and 153 for Addis Kidam, Bahir Dar, Dangila, Durbet, Injibara, Jema, Koga and Sekela, respectively. The largest number of events is observed at Sekela station that is located on a mountain area while the smallest number of events is observed at Bahir Dar which is at the south shore of Lake Tana. Haile et al. (2009d) showed that considering the rainfall of the Sekela station is largely important for runoff modelling of the Gilgel Abbay watershed. The result in the present study suggests that one of the reasons for the importance of the station could be due to the relatively large number of rain events that are observed as compared to the other stations.

RAIN EVENT DEPTH

Statistical properties of the rain event depths of the eight stations are presented in figure 4.1. The top and the bottom horizontal bars represent the 25 % quartile minus 1.5IQR and the 75 % quartile plus 1.5IQR, respectively. IQR is defined as the interquartile range which is the 75 % quartile minus the 25% quartile and is as such defined by the size of the box. The 25 % and the 75 % quartiles are shown by the bottom edge and the top edge of the box, respectively while the median is shown by the line inside the box.

Rain event properties

By the box plot in figure 4.1 some suggestions can be made with respect to the rain event depth in the basin. Overall, the rain event depths have skewed distribution where event depths higher than the median are distributed over a larger range than the event depths that are lower than the median. In terms of median, Addis Kidam, Dangila, Injibara and Sekela stations which are on the mountain areas have events with lowest rainfall depths as compared to the stations on the lowland areas. In terms of median, the lowest event depth is 3.6 mm at Injibara which is located on a mountain area while the highest depth is 6.0 mm at Jema which is situated at an intermediate location between the mountain areas and the lowland areas near Lake Tana. On average, Jema receives an event depth of 1.67 times that of Injibara indicating significant spatial variation of event depth in the study area. Bahir Dar which is on the south shore of the lake has a median value of 4.8 mm event depth.

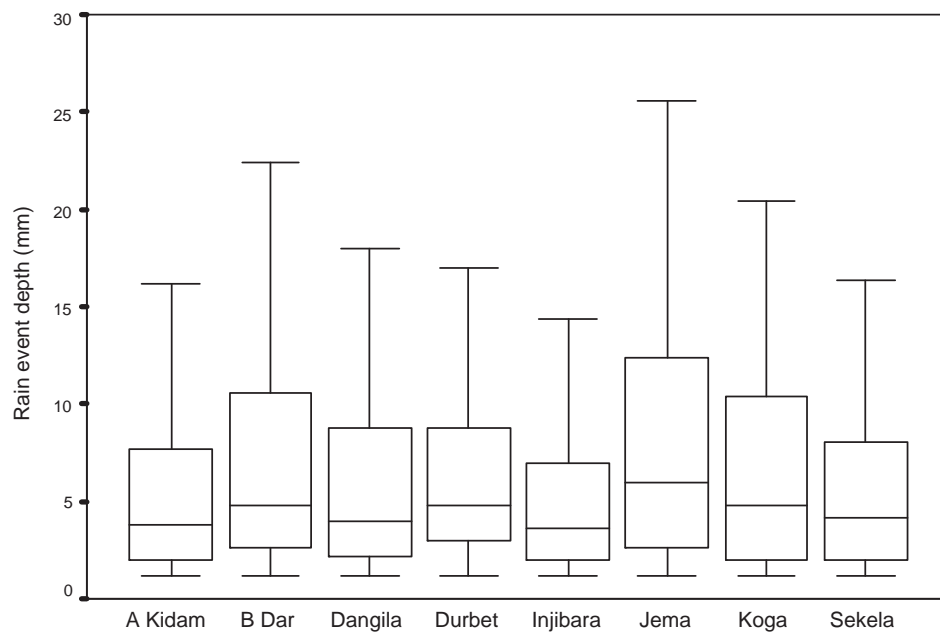


Figure 0.1: Box plot of the rain event depths in JJA of 2007 as observed at eight rain gauge stations. The lower and the upper bars indicate the 25 % quartile minus 1.5 IQR and the 75 % quartile plus 1.5 IQR, respectively. IQR refers to the interquartile range which is defined by the height of the box. The bar inside the boxes shows the median while the upper and lower edges of the boxes show the 75 % and the 25 % quartiles.

The stations that are relatively close to the lake which are Bahir Dar, Koga and Jema station have lowest terrain elevation and have a larger IQR that shows these stations have event depth which is more variable than for the stations near the mountain areas. Such implies that the variability of event depth decreases by the combined effect of an increase in distance from the lake and an increase in terrain elevation.

The box plot also shows that the 75 % quartiles of the event depths of the stations near the lake are larger than the event depths in the mountain areas. Injibara station which is situated on a mountain area has the smallest 25 % quartile of event depth.

RAIN EVENT DURATION

Figure 4.2 shows a box plot of the rain event duration at the eight stations. Similar to the rain event depth, some suggestions can be made regarding the spatial variation of rain event duration. Overall, the distribution of the event duration is skewed since events that have longer durations than the median are distributed over a larger range than event durations that are shorter than the median. However, this skewness is less pronounced compared to the skewed distribution of the rain event depth. As such, the mean event duration is less affected by extreme values as compared to the event depth in the study area.

In terms of median, Addis Kidam, Dangila, Injibara and Sekela stations which are higher in the mountain areas have events with longest duration compared to events of the remaining stations that are in the lowland. An exception is Jema which has relatively low elevation but on average has event duration which is comparable to that of the events on the mountain areas. Koga and Bahir Dar which are relatively close to the lake have events that have the shortest duration on average.

Durbet, Jema and Sekela have a larger IQR of event duration than the remaining stations and thus event duration of these stations is relatively more variable. Although the median of the event duration at Sekela and Injibara are comparable, the IQR does not show that the event duration is equally variable at these stations. The rain event duration at Sekela is more variable as compared to the event duration at Injibara.

MEAN EVENT INTENSITY

Figure 4.3 shows a box plot of the mean event intensities at the eight stations. Similar to the rain event depth, the distribution of the mean event intensity is skewed. In terms of median, Addis Kidam, Dangila, Injibara and Sekela station which are all located on mountain areas have events with lower intensity than the stations that are relatively close to the lake. The lowest median value of event intensity is 3.5 mm h^{-1} at Injibara while the highest median value of event intensity is 6 mm h^{-1} at Jema and Koga that are relatively close to Bahir Dar which has a median value of 5.5 mm h^{-1} . As such, on average the maximum intensity in the study area (i.e. at Jema) is 1.7 times that of the minimum intensity. Bahir Dar, Koga and Jema which are relatively close to the lake and at low elevation in the basin have a larger IQR than the stations near the mountain areas indicating that the event intensity at the lowland areas is more variable.

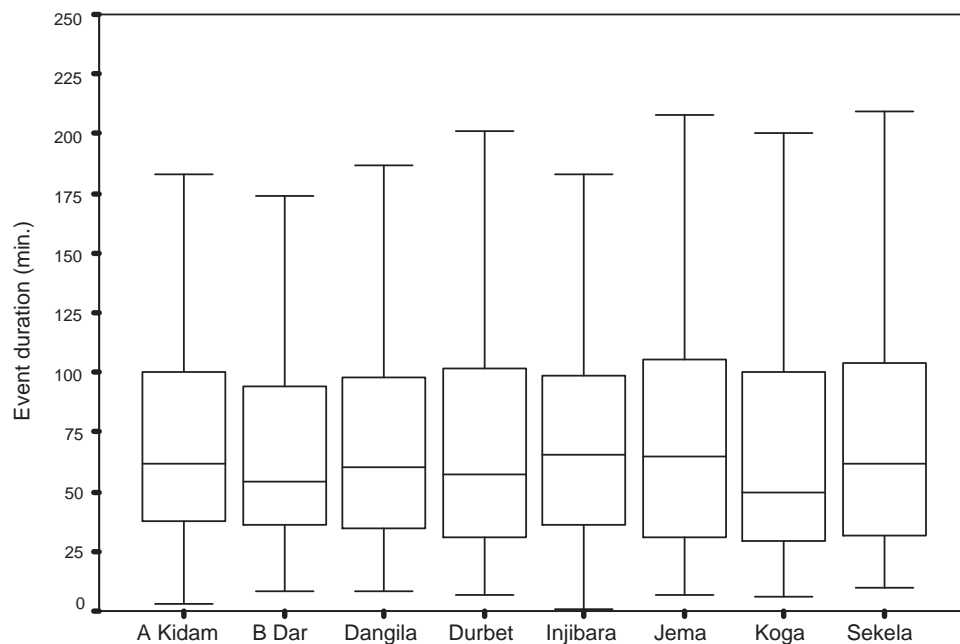


Figure 0.2: Box plot of the rain event durations in JJA of 2007 as observed at eight rain gauge stations.

Rain event properties

The 75 % quartiles of the rain event intensities at the stations which are near the lake are higher than that of the intensities on the mountain areas. Injibara station which is located on the mountain area has the lowest 25 % quartile of event intensities from the network. Over all, the event intensities in the study area have similar spatial pattern to that of the rain event depths.

INTER-EVENT TIME

Figure 4.4 shows the cumulative distribution of the IET at three stations. The cumulative distribution shows some difference which is consistent throughout the observed ranges of IET. For the same IET, the number of occurrences of dry periods shows a spatial variation. The 10 % quantile is 1.4 hours, 1.5 hours and 1.0 hour at Sekela, Jema and Bahir Dar, respectively. The median is 9.0 hours, 13.5 hours and 16.5 hours at Sekela, Jema and Bahir Dar, respectively while the 90 % quantile is 29.5 hours, 31 hours and 42.5 hours at Sekela, Jema and Bahir Dar, respectively.

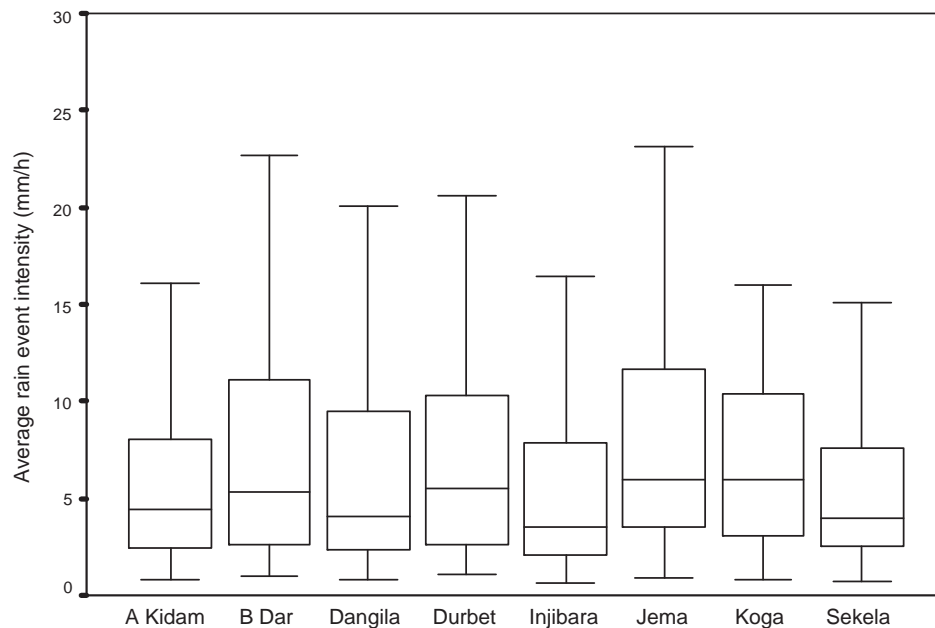


Figure 0.3: Box plot of average rain event intensity in JJA of 2007 as observed at eight rain gauge stations.

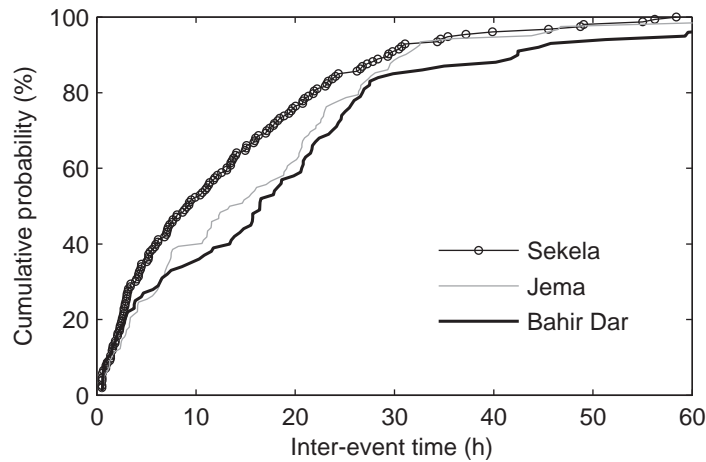


Figure 0.4: Cumulative probability of inter-event time in JJA of 2007 as observed at three rain gauge stations.

At a selected cumulative probability, Sekela has the shortest IET while Bahir Dar has the longest IET. The cumulative distribution reveals that, on average, the lowland areas have a dry period length of 1.8 times that of the mountain areas. Jema which is situated at an intermediate distance between the mountain areas and the lake has an IET which is shorter than at Bahir Dar but longer than at Sekela.

RELATION BETWEEN ELEVATION AND EVENT PROPERTIES

In the previous paragraphs, it is shown that event properties of the study area have some relation with terrain elevation. Therefore, a regression analysis is performed to evaluate such relation. Dairaku et al. (2004) performed a linear regression to relate event properties and terrain elevation. In the present study, a non-linear regression is found to be suitable since it showed a small improvement in the coefficient of determination (R^2) value as compared to a linear regression.

Table 4.4 shows the regression equation that relates event properties to terrain elevation. The 10 %, 50 % and 90 % quantiles are selected to consider low, normal and high event properties, respectively. The R^2 value in table 4.4 shows that terrain elevation weakly explains the variation in the 10 % quantile event depth. However, the 50 % and the 90% quantiles of event depth are strongly related to terrain elevation. The results suggest that normal and large event depths in the study area are largely affected by

orography. In terms of event duration, the 10 % quantile is weakly related to terrain elevation while the 90 % quantile is more strongly related to terrain elevation as compared to the relation between other quantiles and terrain elevation. The spatial distribution of long event duration is affected by orography but that of short event duration is not affected by orography.

The R^2 value shows that the variations in the three quantiles of mean event intensity are largely related to the variation in terrain elevation. The lowland areas in the study area receive rainfall with high intensity as compared to the mountain areas. However, in terms of peak event intensity, a relation could not be established between the 10 % and the 50 % quantiles with terrain elevation since quantiles of peak intensity are found to be equal at all stations. However, the 90 % peak intensity is related to terrain elevation and the slope of the regression equation shows that the events in the lowland areas have high peak intensities. Although the 10 % quantile of IET is weakly related to terrain elevation, the 50 % and the 90 % quantiles of IET have strong relation to terrain elevation. Overall, the high quantiles of event properties are largely affected by orography while low quantiles of event properties except that of event intensity are weakly affected by orography.

Figure 4.5 shows the relation between event properties and terrain elevation. The median of the event properties is selected to evaluate the relation while the regression equations are presented in table 4.4. The plot shows that rain event depth, mean event intensity and IET decrease with an increase in terrain elevation while event duration increases with an increase in terrain elevation. As such, the mountain areas receive rainfall of relatively small depth, low intensity, long duration and short IET. Haile et al. (2009a) showed that the mountain areas of the Lake Tana basin receive relatively large seasonal rainfall depth. The result of the present study showed that the large seasonal rainfall of the mountain areas is mainly due to frequent and long duration events. However, the seasonal rainfall of the lowland areas is largely due to events with high intensity but short duration.

Rain event properties

Table 0.4: *Relation between rain event properties and terrain elevation (Elev.)*

	Equation	R ²
Event depth		
10 % Quantile	-0.0002Elev.+1.8730	0.1527
Median	2007.1Elev. ^{-0.7954}	0.4915
90 % Quantile	14311Elev. ^{-0.8796}	0.5161
Event duration		
10 % Quantile	11.629Elev. ^{0.0546}	0.0042
Median	3.4611Elev. ^{0.3694}	0.3312
90 % Quantile	4986.8Elev. ^{-0.4521}	0.5248
Mean event intensity		
10 % Quantile	648.84Elev. ^{-0.7763}	0.5628
Median	34407Elev. ^{-1.1559}	0.6819
90 % Quantile	459710Elev. ^{-1.3301}	0.6755
Peak event intensity		
10 % Quantile	-	-
Median	-	-
90 % Quantile	-0.0327Elev.+154.41	0.4884
Inter-event time		
10 % Quantile	0.0169Elev. ^{0.5509}	0.0732
Median	3611447Elev. ^{-1.64}	0.9100
90 % Quantile	259888Elev. ^{-1.1659}	0.5666

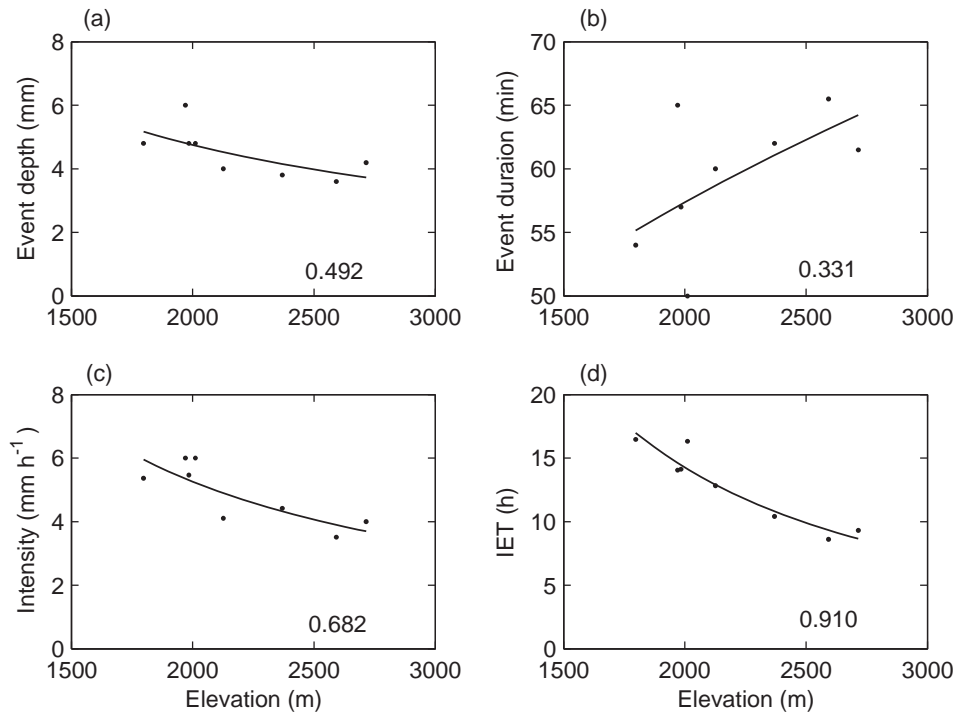


Figure 0.5: Relation between event properties and terrain elevation. The event properties are (a) Event depth, (b) Event duration, (c) Event intensity and (d) Inter-event time (IET). The median of each of the event properties is selected for the regression and the regression equations are shown in Table 4.4.

4.3.5. Dimensionless event hyetographs

Figure 4.6 shows the observed dimensionless hyetographs of Bahir Dar at the south shore of Lake Tana and that of Sekela on a mountain area. Some differences are observed between the hyetographs of the two stations. In terms of the 10 %, 50 % and 90 % quantiles, Bahir Dar has higher fractions of event depths than Sekela for all ranges of fractions of event duration. This shows that Bahir Dar receives larger rainfall depth within an event than Sekela. The largest difference between the dimensionless hyetographs of the two stations is observed for the 10 % quantile.

The difference in the dimensionless hyetographs of Bahir Dar, Jema and Sekela is also statistically evaluated using the Kolmogorov-Smirnov test, see equation [4.4 – 4.5]. Table 4.5 shows the observed test statistic (D)

Rain event properties

which is here defined as the maximum difference in the fractions of event depths of two stations, see equations [4.4 – 4.5]. The critical value of the test statistic for the Kolmogorov-Smirnov goodness of fit test for an arbitrary selected significance level of 0.05 % and number of increments $n = 10$, (i.e. $D_{0.05,10}$) is 0.4092. The observed D as shown in table 4.5 is less than the critical value in all cases indicating there is no statistically significant difference between the dimensionless hyetographs of any of the two selected stations. However, the outcome of the test depends on the n value that is the number of increments. The Kolmogorov-Smirnov test shows that the dimensionless hyetographs do not have statistical difference for $n < 160$. The median of the rain event durations at Bahir Dar, Jema and Sekela is 54, 65 and 62 minutes, respectively and therefore applying $n = 160$ is equivalent to using an increment that is much smaller than 0.5 minutes.

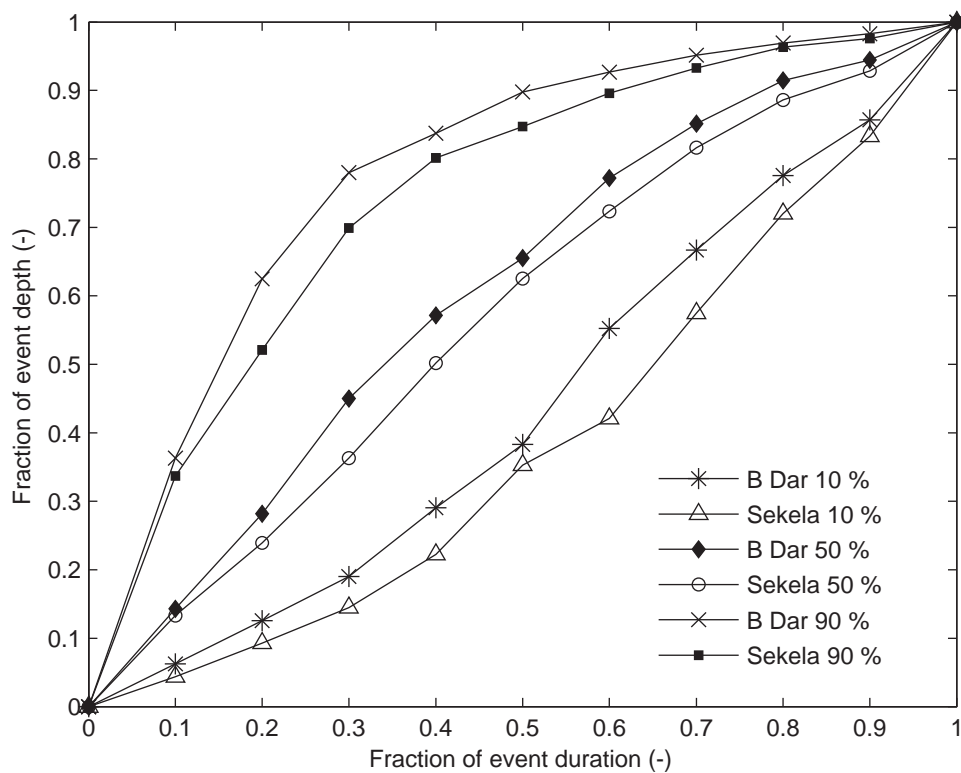


Figure 0.6: The observed dimensionless hyetograph of the rain events at two stations for three quantiles.

Rain event properties

Figure 4.7 – 4.9 show the observed and simulated dimensionless hyetographs of the rain events at Bahir Dar, Jema and Sekela stations, respectively. The hyetographs are simulated based on the beta probability distribution model in equation [4.8].

Table 0.5: Absolute difference between two dimensionless hyetographs (D_{max}) for Kolmogorov-Smirnov test. Note $D_{0.05,10}=0.4092$.

Stations	Quantiles	Bahir Dar	Jema
Jema	10 %	0.0601	-
	50 %	0.0864	-
	90 %	0.1007	-
Sekela	10 %	0.1313	0.0716
	50 %	0.0868	0.0267
	90 %	0.1040	0.0340

Figure 4.7 shows that overall the beta model satisfactorily reproduced the shape of the dimensionless hyetographs. For the 10 % quantile, the model overestimated the fraction of event depth at the middle fractions of event duration but satisfactorily reproduced the depth towards the end of the fractions of event duration. This suggests that the model developed in this study has difficulties to reproduce low rainfall intensities of Bahir Dar.

The model satisfactorily reproduced the 50 % quantile hyetograph over all fractions of event duration showing that normal rainfall intensities could be reproduced using the model. However, the model underestimated the fractions of event depth at all fractions of event duration for the 90 % quantile but the difference is too small as compared to that of the 10 % quantile.

Figure 4.8 shows that the beta model largely overestimates the fractions of the 10 % quantile event depth at Jema but relatively small difference are observed for the 50 % and the 90 % quantiles. Figure 4.9 show that the model reproduced the 10 % quantile hyetograph of Sekela better than that of Jema. At the start and middle fractions of event depth, the modelled hyetograph of Sekela deviates from the observed hyetograph but the modelled and simulated hyetographs show very good agreement at large fractions of event duration. For low intensities, the dimensionless hyetograph of Sekela follows the beta cumulative distribution more than that of Jema. The simulated and the observed 10 % quantile hyetographs show a better agreement at Sekela than that at Jema station. The beta model

Rain event properties

shows some bias in simulating small and large dimensionless event depths. Small event depths are consistently overestimated while large event depths are underestimated at the selected stations.

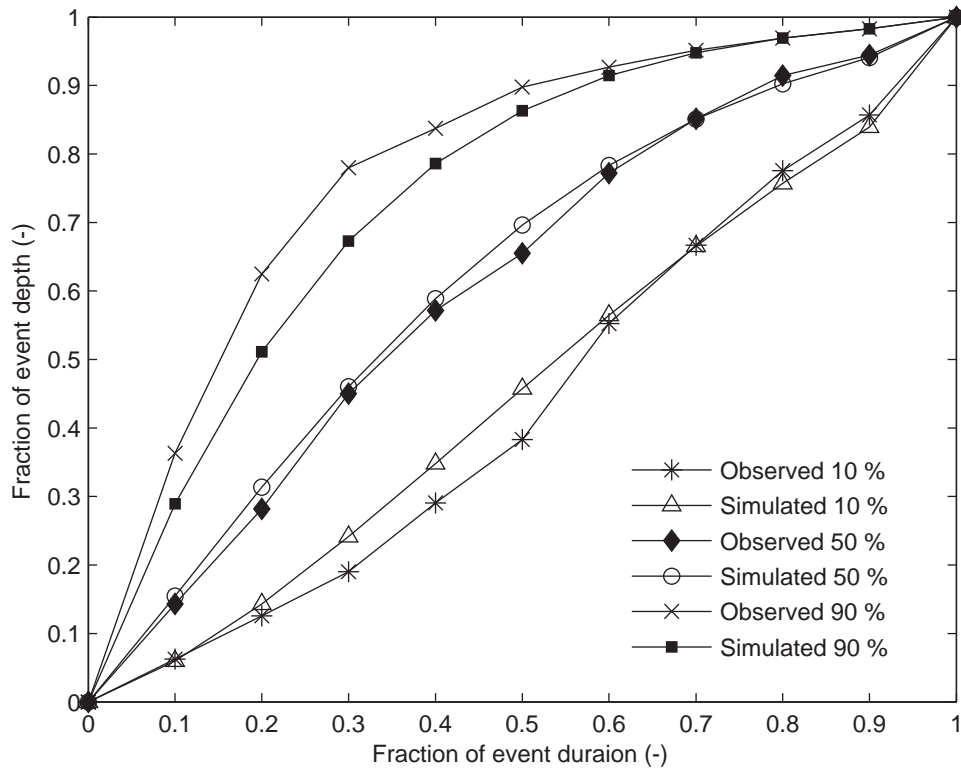


Figure 0.7: The observed and the modelled dimensionless hyetograph of the rain events at Bahir Dar station for three quantiles.

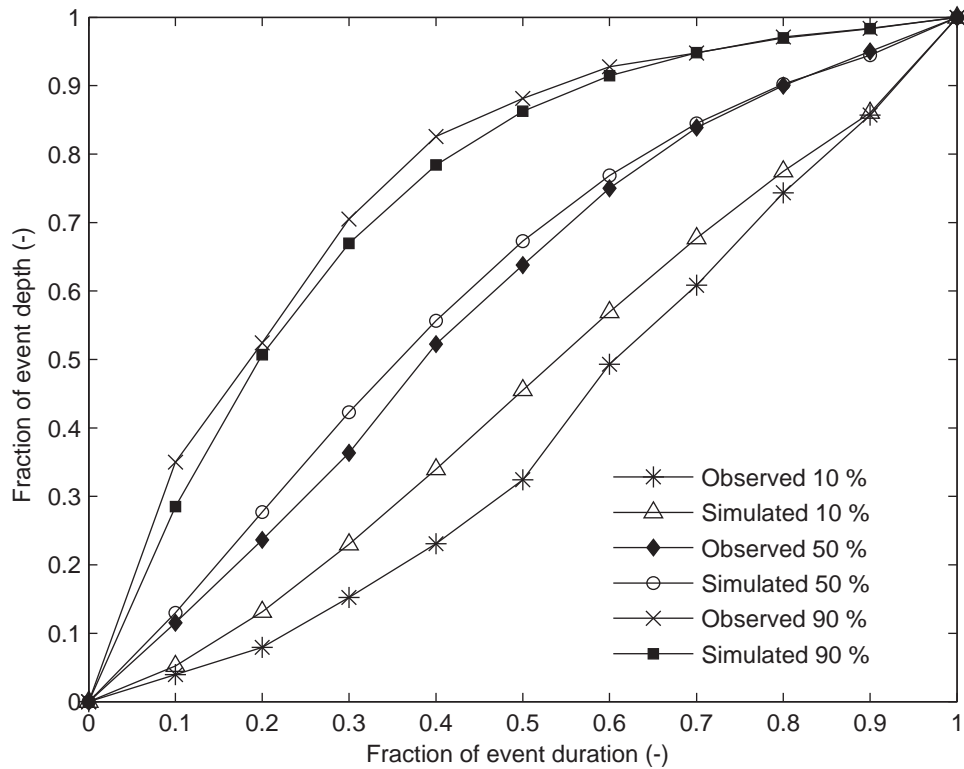


Figure 0.8: The observed and the modelled dimensionless hyetograph of the rain events at Jema station for three quantiles.

The difference between the simulated and the observed dimensionless hyetographs is also statistically evaluated using the Kolmogorov-Smirnov test. The test statistic D which indicates the maximum difference between the simulated and the observed hyetographs is shown in table 4.6 and is less than the critical value in all cases. The result shows that there is no statistically significant difference between the modeled and the observed hyetographs of the three stations. The largest deviation between the simulated and the observed hyetographs is observed for the 10 % quantile at Jema station.

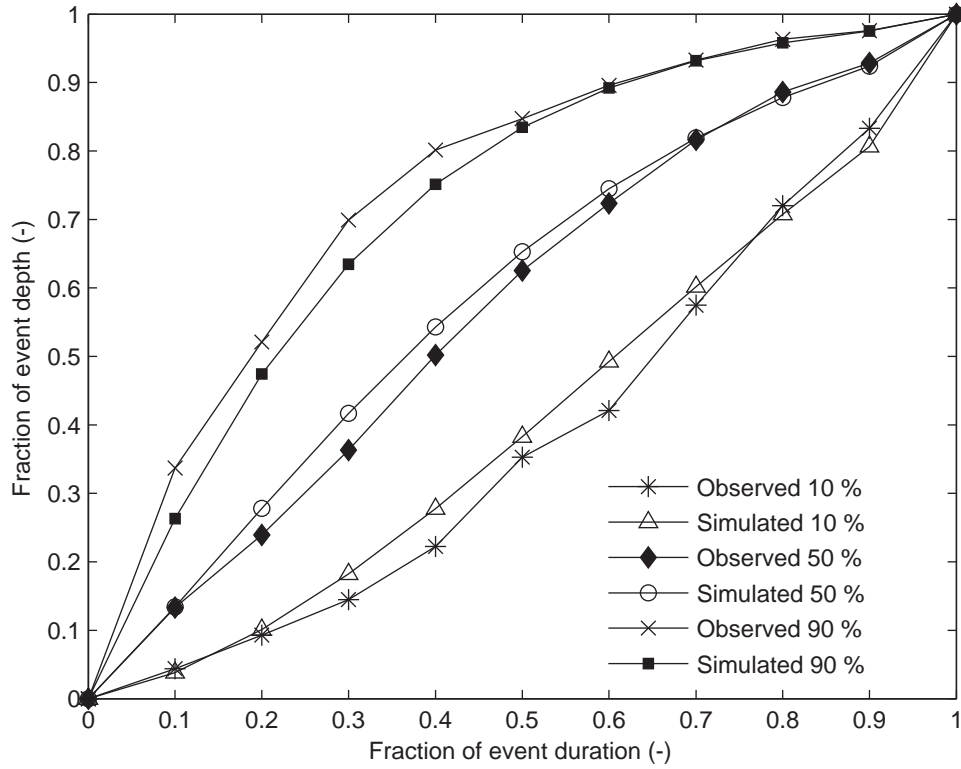


Figure 0.9: The observed and the modelled dimensionless hyetograph of the rain events at Sekela station for three quantiles.

Table 0.6: Absolute difference between observed and simulated dimensionless hyetographs (D_{max}) for Kolmogorov-Smirnov test. Note $D_{0.05,10}=0.4092$.

Station	10 % quantile	50 % quantile	90 % quantile
Bahir Dar	0.0744	0.0412	0.1136
Jema	0.1307	0.0591	0.0647
Sekela	0.0712	0.0538	0.0737

The parameter values for the beta model are shown in table 4.7. Both parameter values increase with an increase in terrain elevation.

Table 0.7: Parameters of the beta distribution model that is fitted to the dimensionless hyetographs.

Parameter	10 % quantile	50 % quantile	90 % quantile
Bahir Dar			
α	2.146	5.956	25.005
β	9.385	16.221	44.087
Jema			
α	1.705	3.680	21.333
β	6.884	9.778	37.237
Sekela			
α	1.510	5.077	22.088
β	7.344	15.363	43.573

4.3.6. Conditional probability of rainfall occurrences

Figure 4.10 shows the plot of the conditional probability (P_{ij}) of rainfall occurrences (equation [4.9]) against inter-station distance. The conditional probability at a selected inter-station distance increased when the rainfall aggregation time changed from 1 hour to 6 hour. Maximum P_{ij} values of 0.35 and 0.55 occurred at the smallest inter-station distance are estimated for the 1 hour and the 6 hour rainfall, respectively.

It is shown that the conditional probability decreases exponentially with an increase in inter-station distance for both aggregation times. As such an exponential model of the following form is fitted to the data points using least square fitting:

$$P_{ij} = a \times \exp(-b \times d) \quad [3.10]$$

where: P_{ij} is the conditional probability of rainfall occurrences which indicates that rains at station j while it also rains at station i , a and b are empirical fitting constants and d is the inter-station (separation) distance between two selected stations. The exponential model is fitted with an R^2 value of 0.86 and 0.63 for the 1 hour and the 6 hour rainfall, respectively. The values of the constants for the 1 hour rainfall are $a = 0.4057$ and $b = 0.0231$ while the values of the constants for the 6 hour rainfall are $a = 0.5927$ and $b = 0.0103$. Although it is not shown in figure 4.10, the conditional probability of daily rainfall occurrences ranges between 0.61– 0.82 with R^2 value of 0.0005 for the fitted exponential model showing a poor model

performance. This is caused by weak relation between inter-station distance and daily rainfall occurrence.

4.4. DISCUSSION AND CONCLUSION

Studies show that changes in event properties affect the response of hydrologic models (e.g. Kusumastuti et al., 2007) but lack of information on event properties restricts the applicability of hydrologic models (e.g. Renard and Freimun, 1994; Diodato and Bellocchi, 2007). For locations without sub-daily rainfall records, generation of synthetic rainfall data is required to allow application of hydrologic models at high temporal resolutions. Such rainfall generation requires knowledge on the spatial and temporal patterns of rain event properties on a short-term base. Few studies address the spatial pattern of rain event properties (e.g. Bidin and Chappel, 2006; Dairaku et al., 2004). The objective in the present study is to understand the rain event properties in the source basin of the Upper Blue Nile River and to evaluate the relation between rain event properties and orography of the basin.

The temporal characteristics of rain event properties have been analyzed to evaluate whether event properties change with time. Also, relations between the various rain event properties are evaluated to enhance our understanding of rainfall and provide information that may be useful for synthetic generation of rainfall data. The considered event properties are rain event depth, duration, mean intensity, peak intensity and inter-event time (IET).

The relation between rain event properties is evaluated through correlation analysis which showed that some relation exists between the rain event properties. Overall, rain event depth increases with an increase in event duration, mean event intensity, and peak intensity. In terms of correlation, the rain event depth in the study area is more related to peak intensity than to event duration. As such, large event depth in the study area is caused not only by relatively a long duration event but also by high intensity. The analyses show that the IET has a weak relation with the other rain event properties.

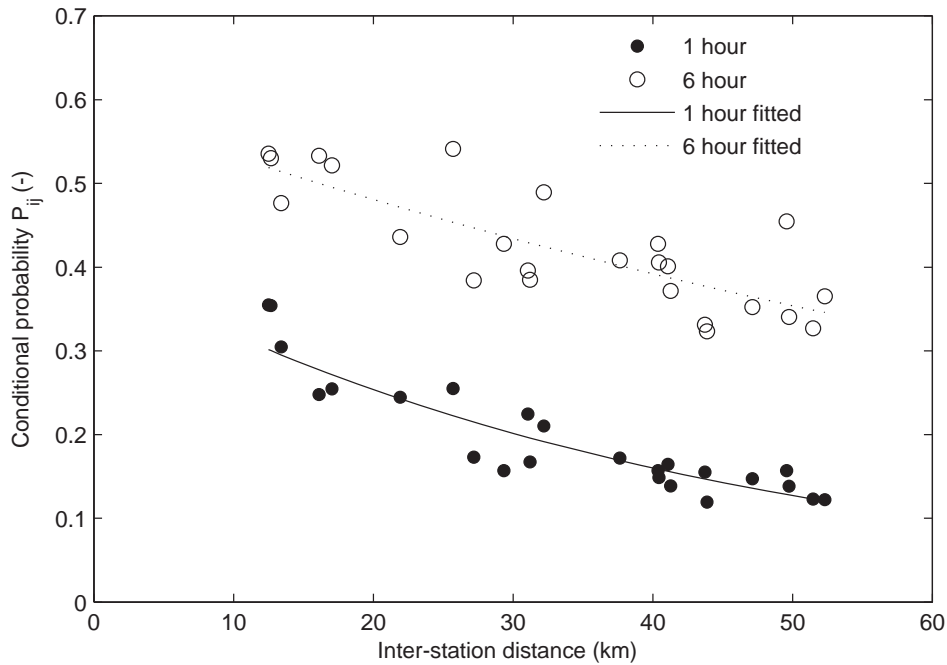


Figure 0.10: Conditional probability of rainfall occurrences against inter-station distance. The exponential model is fitted with an R^2 value of 0.86 and 0.63 for the 1 hour and the 6 hour rainfall, respectively.

The results also show that at the start and towards the end of the wet season, the rain event depth is relatively large with long event duration and long IET. Such indicates that two consecutive events as observed in the early and the late season are separated by longer dry periods than the events in the mid-season. In terms of median, the June events have event depth that is 2 mm higher and event duration that is 10.5 minutes longer than the July events. The IET in June is 6.3 hours longer than that in July. Also, mid-season events have depth, duration and peak intensity that are more variable as compared to the events at the start and at the end of the wet season. Such differences may affect hydrologic processes. Long IET may result in large evapotranspiration between two consecutive rain events and therefore decreases the effective rainfall which is available for runoff.

Haile et al. (2009a) showed that there is a weak relation between seasonal rainfall and terrain elevation or distance to the lake but the diurnal cycle of rainfall in the study area is strongly affected by orography. The results of the present study revealed that terrain elevation of the Lake Tana

basin affects rain event properties in particular high quantiles. The interquartile range shows that the variability of event depth and mean event intensity decreases with an increase in distance from the lake and with an increase in terrain elevation. As such, the low land areas have largely variable event depth and intensity as compared to the mountain areas of the basin.

The coefficient of determination (R^2) value indicates that terrain elevation weakly relates to the 10 % quantile event depth, event duration, peak intensity, and IET. Such suggests that events with small values of properties such as small rainfall depth are not affected by terrain elevation. In terms of the median values, the maximum spatial difference in the watershed is 2.4 mm, 2.5 mm h⁻¹, 15.5 minutes and 7.85 hour for event depth, mean event intensity, event duration, IET, respectively. Also, the median of the event durations in the watershed is in the range of 1 hour ±10 minutes. Such differences reveal that the rain events in the study area are characterized by a short duration which is the case for convective events that are commonly observed in the study area.

Rainfall amount is expected to increase with an increase in terrain elevation due to orographic effects. Although some studies evaluated the effect of orography on annual, seasonal or daily rainfall, there are few studies in literature that evaluate the relation between orography and rainfall at the scale of rain events. In the present study such relation is evaluated for the rain events of the Lake Tana basin. The results revealed that (i) the low land areas have events with relatively more variable depth and intensity than the mountain areas, (ii) on average, the low land areas receive rain events of larger depth, higher intensity, shorter duration and longer IET than the mountain areas; (iii) relatively high quantiles of event properties, such as large event depth, are affected by terrain elevation; (iv) relatively low quantiles of event properties, such as small event depth, except mean event intensity are weakly related to terrain elevation; (v) both low and high event intensity are strongly related to terrain elevation with low lands.

The results are similar to that by Barros et al. (2000) who, using 4 month data in Nepal, showed that the lower-elevation stations received higher intensity rainfall over short durations. However, the results are somewhat in contrast to that of Dairaku et al. (2004) who showed that relatively low intensity events in the mountains of northwestern Thailand are weakly affected by terrain elevation. Such difference can be caused not

only by climatic factors but also due to the presence of Lake Tana in the present study area that affect rain event properties.

A dimensionless rain event hyetograph serves to generate rainfall intensity for data scarce areas. A dimensionless rain event hyetograph is developed for the Gilgel Abbay watershed which is situated in the study area. The 10 %, 50 % and 90 % quantiles of fractions of rain event depth at the corresponding fractions of event duration are used to develop the dimensionless event hyetographs. For the same event duration, the areas near Lake Tana that are at lower elevation receive larger rainfall depth than the mountain areas for all ranges of rain events, i.e. light, moderate and heavy. Although such suggests that temporal patterns of rainfall intensity within rain events of the study area show some spatial difference, the difference is found to be statistically insignificant.

The beta probability distribution model is fitted to the observed dimensionless hyetographs. The model consistently underestimated large dimensionless depths while it overestimated small dimensionless depths of selected stations. However, the Kolmogorov-Smirnov test showed that the deviation between the modelled and the observed hyetograph is not significant. This shows that the model satisfactorily reproduced the shape of the observed dimensionless hyetographs in particular normal dimensionless depths of the selected stations. As pointed out by Garcia-Guzman and Aranda-Oliver (1993), the beta model is simple to implement since it requires estimation of only two parameters. The model can be used to generate synthetic rainfall data, for instance, for runoff and soil erosion studies in the absence of rainfall observations at the required resolution.

Haile et al. (2009a) evaluated the temporal intermittence of the hourly rainfall of the Lake Tana basin using fractal analysis and found that rainfall occurs at least once in 2.7 days and that rainfall of the mountain stations is closely clustered. In the present study, the spatial intermittence of the hourly rainfall is evaluated. The conditional probability of rainfall occurrences at any two stations is evaluated and an exponential model is fitted to relate such probability to inter-station distance. The conditional probability of 1 hour and 6 hour rainfall occurrences is approximately 0.35 and 0.55, respectively for inter-station distances of 10 km. Such small probability of occurrences at scales smaller than 10 km shows the high variability of the rainfall in the basin. The increase in the conditional probability with an increase in the aggregation time of the rainfall data can be caused by: (i) the presence of rain events with a duration that is longer

than 1 hour; (ii) the movement of rain producing clouds at a speed that is lower than the inter-station distance per hour, or (iii) the occurrence of multiple rain events in the 6 hour or 1 day time period.

The fitted exponential model between the conditional probability of rainfall occurrence and inter-station distance performed satisfactorily with an R^2 value of 0.86 and 0.63 for the 1 hour and the 6 hour rainfall, respectively. However, the model performed poor when it is fitted to the conditional probability of daily rainfall occurrences since the probabilities only vary over a very small range that does not show relation with inter-station distance and as such a trend could not be established.

Haile et al. (2009d) have shown that the accuracy of the simulated stream flows of the Gilgel Abbay watershed in Lake Tana basin depends on whether a low land or a mountain rain gauge is considered. In the same work, a relative error between observed and simulated stream flow of up to 100 % is reported when rainfall data from a single station served as model input. Such indicates large variability in the rainfall over the study area that largely affects the accuracy of runoff simulations. In particular the rainfall of Sekela and Jema is very important and should be considered for runoff modelling. The present study showed that Jema has high intensity events while Sekela has the largest number of events with relatively short IET. Also, the two stations have longest event duration as compared to the stations in the low land areas.

Based on these result of the present study, we speculate that the rain events of mountain areas have less soil erosive power than the rain events in the low lands. We recommend that future studies should evaluate the effect of soil erosion in the watershed since severe erosion has been observed in the watershed during the field campaign of this study.

5

REMOTE SENSING BASED RAINFALL DETECTION AND ESTIMATION



ABSTRACT¹

Many remote sensing based rainfall products have spatial resolutions $\geq 0.25^\circ$ and temporal resolutions ≥ 1 day which are coarser than what is typically needed in hydrology. In this study, the rainfall data that served as the ground-truth for rainfall is obtained from the Precipitation Radar of the Tropical Rainfall Measuring Mission (TRMM) which acquires data at 5 km resolution once or twice a day. The remote sensing data for rainfall detection and estimation is obtained from the Spinning Enhanced Visible and Infrared Imager (SEVIRI) of the Meteosat Second Generation (MSG-2) which acquires data at 3 km resolution at 15 minutes interval. Three MSG-2 channels are evaluated for rain detection in the Upper Blue Nile area in Ethiopia by the following indices: (1) the 10.8 μm brightness temperature, (2) the rate of change of the 10.8 μm brightness temperature, (3) the space gradient of the 10.8 μm brightness temperature, (4) the brightness temperature difference (BTD) at the 10.8 μm and the 6.2 μm and (5) the BTD at the 10.8 μm and the 12.0 μm channels. The evaluation was made through categorical statistics that are *bias*, probability of detection, false alarm ratio and Heidke skill score. In this work also, an exponential model was developed for thermal infrared based rainfall estimation. The model was evaluated using observations from a rain gauge network that was installed at the source basin of the Upper Blue Nile River.

Keywords: Rainfall detection, Rainfall estimation, MSG-2, TRMM PR, Blue Nile

¹ This chapter is based on: Haile, A. T., Rientjes, T., Gieske, A., Gebremichael, M., 2009. **Multispectral remote sensing for rainfall detection and estimation at the source of the Blue Nile River**, *International Journal of Applied Earth Observation and Geoinformation*, <http://dx.doi.org/10.1016/j.jag.2009.09.001> (In press).

5.1. INTRODUCTION

Accurate rainfall estimation requires frequent observations with continuous spatial coverage. Such observations are available from geostationary satellites which carry sensors that measure radiances at a range of wavelengths. The most commonly used is the thermal infrared (TIR) radiance that is recorded in the wavelength range of 10.0 – 12.5 μm (e.g., Griffith et al., 1978; Negri et al., 1984; Adler and Negri, 1988). However, TIR data do not provide direct observations of either surface rainfall rates or cloud profiles of, for instance, cloud water content that affect the rainfall process.

In TIR based rainfall estimation, rainfall rates are inferred from cloud top temperatures. The underlying physical assumption is that relatively cold clouds are associated with thick and high clouds that tend to produce high rainfall rates. The limitation of this approach is that different vertical profiles of clouds, which result in different rainfall rates, can have the same cloud top temperatures. A second limitation is the difficulty to detect clouds that produce rainfall. TIR based approaches assume that rain occurs when the cloud top temperature is less than a selected threshold. For instance, a threshold of 253 K was applied by Griffith et al. (1978) while a threshold of 235 K was applied by Arkin (1979) and Arkin and Meisner (1987) for the GOES precipitation index (GPI). However, as shown by Todd et al. (1995), the effectiveness of a constant temperature threshold is questionable as it ignores several factors such as the rain water profile in a cloud system that govern rain generation.

The limitation associated with the inability of TIR sensors to detect signals from the cloud profile that affect rainfall generation can be overcome by the use of microwave (MW) sensors which respond primarily to precipitation-size hydrometeors in the cloud profile. Currently, MW sensors are mounted only on low-altitude orbiting satellites and therefore the observations are snapshots that are available once or twice a day. Such observation frequency makes it difficult to capture the temporal dynamics of cloud systems.

The complementary strengths of the TIR and the MW sensors have led to the development of algorithms that use multispectral channels for rainfall estimation. Some examples of such algorithms are PERSIANN (Hsu

et al., 1997; Sorooshian, et al., 2000; Hong et al., 2005), NRLgeo (Turk and Miller, 2005), TRMM 3B42 (Huffman et al., 2007). Most of these products are available at spatial resolutions $\geq 0.25^\circ$ and temporal resolutions ≥ 1 day. However, the spatial resolutions in particular are coarser than what is typically needed in hydrology and water resources.

Motivated by the need to develop rainfall products of high temporal resolution, in the present study, a relationship is established between MW based rainfall rates and TIR brightness temperatures at a resolution of 5 km. The MW based rainfall rates are obtained from the orbiting TRMM (Tropical Rainfall Measuring Mission) precipitation radar (PR) sensor. The TIR brightness temperatures are obtained from the geostationary MSG-2 (Meteosat Second Generation) channel of $10.8 \mu\text{m}$ which is probably the most commonly used channel for rainfall detection and estimation. Various methods are evaluated and tested for rain detection using TIR based indices from MSG-2. The TRMM PR rainfall estimates served as a reference data to evaluate the rain detection methods. Also, rainfall is estimated by a simple exponential rainfall model using MSG-2 TIR brightness temperature for which the accuracy of the estimates was evaluated using rainfall observations from a network of rain gauges.

In addition to the $10.8 \mu\text{m}$ brightness temperature ($T_{10.8}$), various TIR indices are evaluated for rain detection. The TIR indices are the rate of change of $T_{10.8}$ ($\Delta T_{10.8}$), the space gradient of $T_{10.8}$ ($\nabla T_{10.8}$), and the brightness temperature differences (BTD) between the $10.8 \mu\text{m}$ and the $6.2 \mu\text{m}$ channels ($T_{10.8} - T_{6.2}$), and the $10.8 \mu\text{m}$ and the $12.0 \mu\text{m}$ channels ($T_{10.8} - T_{12.0}$).

5.2. DATA SETS

5.2.1. Remote Sensing observations

Observations from remote sensing which are used in this study are rainfall rate estimates from TRMM PR and brightness temperatures from three MSG-2 channels. The TRMM PR rainfall rates served as the reference data set for calibration and evaluation of the brightness temperature indices for rainfall detection and estimation. Among the variables in the TRMM PR product 2A25, we used only the near-surface rainfall rate to calibrate and

evaluate the MSG-2 brightness temperatures for rainfall detection and evaluation. The 2A25 products represent snapshots of rainfall rates at a horizontal resolution of 5 km and a temporal frequency of 1 or 2 observations per day depending on the latitude. The 2A25 products are processed by the TRMM science team using the algorithms described in Iguchi et al. (2000) and Meneghini et al. (2000). The 2A25 products and other TRMM products can be downloaded free-of-charge following the links at <http://daac.gsfc.nasa.gov/precipitation/>.

Observations from MSG-2 sensors have a spatial resolution of 3 km at nadir and a temporal resolution of 15 minutes. The following three channels of MSG-2 are used in the present study: water vapor (WV) channel at 6.2 μm , TIR channel at 10.8 μm , and TIR channel at 12.0 μm . Information about MSG-2 observations is available from EUMETSAT (<http://www.eumetsat.int>). Software to convert MSG-2 radiance observations to brightness temperatures and to re-project the images to Universal Transverse Mercator (UTM) projection at selected grid element sizes is developed by Gieske et al., 2005; Maathuis et al., 2006.

The following processing steps are followed here. First, the region of interest is defined as a region covering 5° latitude by 5° longitude with the Lake Tana basin located at the centre. Second, 17 convective events are selected for this region based on TRMM PR observations for the time period June – August of the years 2006 and 2007. Convective events were selected if the lowest TIR temperature in an image was less than 210 K which indicates heavy rainfall. It was also checked whether cloud systems were not a result of merging of multiple clouds that affect cloud top temperatures. Third, the MSG-2 images that were acquired close to the TRMM PR overpasses are identified for collocation of the two data sets in time. This procedure is necessary since the TRMM PR and MSG-2 observations are seldom acquired at the same time. Fourth, the MSG-2 images are re-sampled at the resolution of the TRMM PR rainfall estimates to match up the respective resolutions at 5 km. Fifth, the PR surface rainfall rates and the TIR brightness temperatures are collocated in space to coincide observations.

Similar to the work by Vicente et al. (1998), the surface rainfall rates and the brightness temperatures are collocated in space for each of the 17 convective cloud fields. The two data sets are joined based on the principle that the lowest TIR temperature matches with the highest PR rainfall. Figure 5.1 shows that there is a shift of one or two pixels between the highest

rainfall rate and the lowest brightness temperature. The rainfall rate decreases exponentially with an increase in brightness temperature after the two images are collocated. Such a procedure assures that the highest rainfall rate is assigned to the lowest cloud top brightness temperature. This procedure is selected, as also pointed out by Vicente et al. (1998), to minimize collocation errors for instance due to the small differences in time at which images are acquired. The procedure is applied for each of the 17 events. Collocation is manually checked to rule out errors that could occur when multiple clouds are within the same image. In total, about 4500 pixels of cloud top temperature and rainfall rates are collocated.

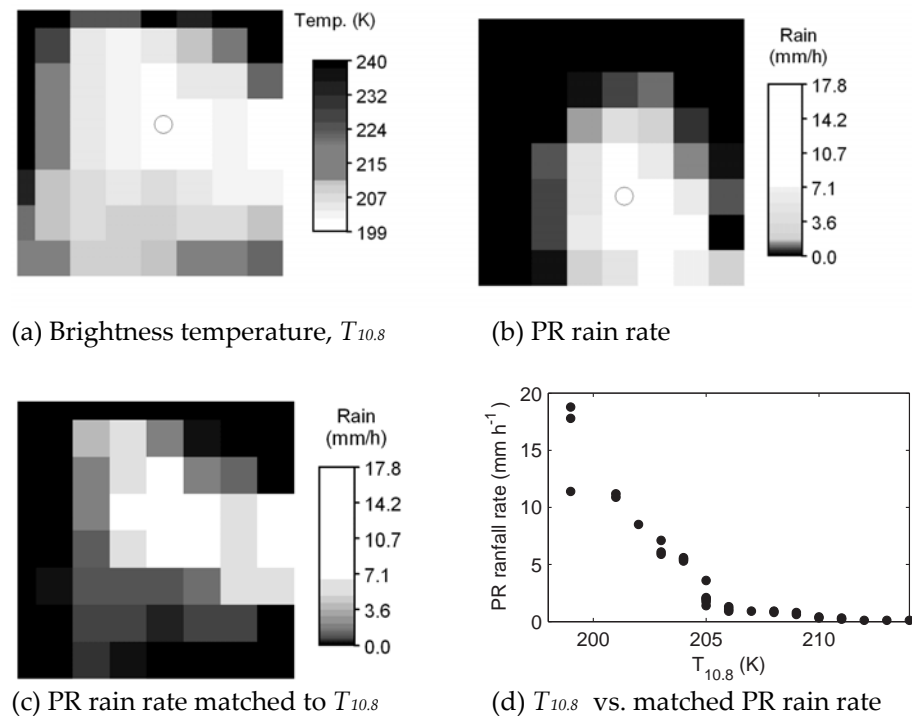


Figure 5.1: Comparison of $T_{10.8}$ and TRMM PR rain rates for July 12, 2007 at 10:45. (a) and (b) show the original images while (c) shows the collocated (matched) image and (d) shows relation between $T_{10.8}$ and rainfall rate after collocation by the principle that the highest rainfall rate corresponds to the lowest $T_{10.8}$. The hollow circles in (a) and (b) indicate the space shift between the highest convective rainfall locations from the two images. The map scale is 1:500,000.

5.2.2. Ground based observations

The ground based observations are obtained from 10 tipping-bucket rain gauges that were installed in the Lake Tana basin in May 2007 and rain data has been collected since then. The inter-station distance of the rain gauges ranges between 5 km to 82 km with an average of 35 km. The terrain elevation of the stations is between 1798 m and 2715 m above mean sea level. Further information on the rain gauge network layout and the statistics of the observed rainfall is given in the work of Haile et al. (2009) and in chapter 3 of this thesis. On the basis of the rain gauge data, 9 convective rainfall events are identified for the time period June – August 2007 that are used in the present study.

5.3. METHOD OF ANALYSIS

5.3.1. Rainfall detection

To evaluate the effectiveness of TIR based rainfall detection, the various TIR indices that yield information on spatial and temporal cloud dynamics are evaluated. The indices are developed based on the brightness temperatures at three of the MSG-2 channels. The selected indices are (1) the brightness temperature at the 10.8 μm channel ($T_{10.8}$); (2) the rate of change of brightness temperature over two consecutive images ($\Delta T_{10.8}$); (3) the space gradient $\nabla T_{10.8}$ over connected pixels in an image; (4) the brightness temperature differences (BTD) at the 10.8 μm channel and the 6.2 μm channels ($T_{10.8} - T_{6.2}$) and (5) BTD at the 10.8 μm channel and the 12.0 μm channels ($T_{10.8} - T_{12.0}$). The indices are discussed in the following paragraphs.

The rate of change in $T_{10.8}$ over time ($\Delta T_{10.8}$) provides information on cloud stage development (e.g., Griffith et al., 1978; Vicente et al., 1998) and is defined as:

$$\Delta T_{10.8} = \frac{T_{10.8,i} - T_{10.8,i-1}}{\Delta t} \quad [5.1]$$

where: Δt is the time difference between two consecutive images (i.e. 15 minute for MSG-2), and i is a time index.

The space gradient of $T_{10.8}$ ($\nabla T_{10.8}$) serves to yield information on rain producing clouds (e.g., Weiss and Smith, 1987) and is defined as:

$$\nabla T_{10.8} = \sqrt{\frac{d^2 T_{10.8}}{dx^2} + \frac{d^2 T_{10.8}}{dy^2}} \quad [5.2]$$

where: x and y are the Cartesian coordinates. The space gradient of $T_{10.8}$ in equation [5.2] is estimated using a central difference approximation for the second derivatives.

One of the challenges in TIR-based rainfall detection is to differentiate convective cloud systems that produce rainfall from other high-level (i.e. cold) cloud systems such as cirrus clouds that do not produce rainfall. The top surface of both cloud systems can have equally low $T_{10.8}$. However, the cirrus cloud systems only are thin that as such do not produce rainfall. Such cloud systems may be incorrectly classified as rain producing clouds if the brightness temperature at the 10.8 μm channel is used as an indicator. Improving the accuracy of cirrus clouds detection could therefore lead to a better rainfall detection capability. Cirrus clouds could be detected using information from three MSG-2 channels as further described below.

Fritz and Laszlo (1993) described that for cloud tops in the troposphere where air temperature decreases with height, the water vapor above the cloud top usually is colder ($T_{10.8} > T_{6.2}$) than the cloud top temperature at 10.8 μm that as such could provide information that is useful for rainfall detection. On the other hand, if the cloud is thick and the top is above the tropopause, then $T_{10.8} < T_{6.2}$. Therefore, following Kuligowski (2002), the index $T_{10.8} - T_{6.2}$ is used for rainfall detection by identifying clouds with tops that are in the troposphere or above the tropopause.

A second approach to assess the presence of cirrus clouds is based on the 10.8 μm and the 12.0 μm channels. At the two channels, the emissivity values for cirrus clouds may differ substantially while the emissivity of thick clouds is close to 1. Therefore, thick clouds have small values of $T_{10.8} - T_{12.0}$ while thin clouds have large values of $T_{10.8} - T_{12.0}$ (Inoue, 1987) which indicate cirrus clouds.

5.3.2. Rainfall estimation

Following the approach by Vicente et al. (1998); Boi et al. (2004) and Wardah et al. (2008), we fitted an exponential model to relate the TRMM PR rainfall rates to the MSG-2 brightness temperature $T_{10.8}$ that is most commonly used in rainfall estimation:

$$R = a \times \exp(-b(T_{10.8} - T_t)) \quad [5.3]$$

where: R the rainfall rate averaged at each 1 K temperature interval; T_t is a threshold temperature; a , b and c are constants, i.e. empirical parameters that are estimated by least squares fitting; R is in mm hr^{-1} and $T_{10.8}$ is in K. A threshold temperature (T_t) of 180 K was used to limit the constant values to practical values. It is noted that $T_{10.8}$ is always higher than 180 K at the source of the Blue Nile River.

5.3.3. Performance measures

Categorical statistics are used to evaluate the effectiveness of the selected TIR indices to detect rainfall. In this evaluation, the TRMM PR surface rainfall rates are considered as the ground truth and in a sense serve as a reference. The categorical statistics in the present study are: *bias*, probability of detection (*POD*), false alarm ratio (*FAR*) and Heidke skill score (*HSS*).

The categorical statistics can be defined by using the standard 2×2 contingency table where reference is made to Kuligowski (2002); Doswell et al. (1990); Todd et al. (1995) for detailed information. Table 5.1 shows the contingency table of rain detection by a selected TIR index in reference to the rain detection by TRMM PR. In Table 5.1, h is the number of pixels for which both the TIR index and TRMM PR detected rainfall and therefore indicates the number of pixels correctly classified as rainy by the TIR index; f is the number of pixels for which rainfall is detected by only the TIR index but not by TRMM PR and as such indicates the number of pixels incorrectly classified as rain by the TIR index; m is the number of pixels for which rainfall is not detected by the TIR index while it is detected by TRMM PR and as such it indicates the number of pixels incorrectly classified as non-rainy by the TIR index; z is the number of pixels for which rainfall is not detected by both TIR index and the TRMM PR which indicates the number

of pixels correctly classified as non-rainy by the TIR index. The categorical statistics are defined as follows:

$$bias = \frac{f + h}{m + h} \quad [5.4]$$

$$POD = \frac{h}{m + h} \quad [5.5]$$

$$FAR = \frac{f}{f + h} \quad [5.6]$$

$$HSS = \frac{C - E}{N - E} \quad [5.7]$$

$$C = h + z \quad [5.8]$$

$$E = \frac{(h + f)(h + m) + (f + z)(m + z)}{(h + m + z + f)} \quad [5.9]$$

After substituting for C and E , equation [5.7] reads:

$$HSS = \frac{2(zh - fm)}{(z + f)(f + h) + (m + h)(z + m)} \quad [5.10]$$

where: C is the number of pixels for which both the TIR index and the TRMM PR report the same result. E is the expected number of rainy pixels correctly classified by the TIR index due purely to chance, see Doswell et al. (1990). N is the total number of pixels that is $(h+f+m+z)$.

Table 5.1: Contingency table

		TIR based rainfall detection	
		Rain	No-rain
TRMM PR rainfall detection	Rain	h	f
	No-rain	m	z

Bias is the ratio of the number of incorrect rain detection by a TIR index to the total number of rain detection by TRMM PR and indicates whether the selected TIR index underestimates or overestimates the number of rainy pixels that are detected by TRMM PR. A *bias* greater than 1 implies that the TIR index overestimates the number of rainy pixels while a *bias* of less than 1 implies the TIR index underestimates. The *POD* is the ratio of the number of correct rain detection by a TIR index to the total number of rain detection by TRMM PR and it indicates the fraction of times the reference rainy pixels are correctly detected by the selected index. A *POD* = 1 indicates that the selected index correctly detects all rainy pixels. The *FAR* is the ratio of the number of incorrect rain detection by the TIR index to the total number of rain detection by the TIR index and it indicates the fraction the selected index detects rainy pixels that were not detected by the TRMM PR.

The *HSS* values range from -1 to +1. Some cases of *HSS* are (i) *HSS* = 0 for ($h = 0$) and ($m = 0$) which indicates that the reference data set (i.e. TRMM PR) did not detect any rainy pixel. (ii) *HSS* = 0 for ($h = 0$) and ($f = 0$) which indicates the TIR index did not detect any rainy pixel. (iii) *HSS* = -1 for ($h = 0$) and ($z = 0$) and ($f = m$) that indicates the TIR index did not correctly detect any rainy and non-rainy pixels, and the number of pixels that are incorrectly classified as rainy and the number of pixels that are incorrectly classified as non-rainy by the TIR index are equal. (iv) *HSS* = 1 for ($m = 0$) and ($f = 0$) which indicates the TIR index did not incorrectly classify any pixel.

These categorical statistics are estimated for various ranges of threshold rainfall rates to evaluate the range of rainfall rates over which the selected indices perform satisfactorily. For instance, if a threshold of 1 mm h⁻¹ is specified then a pixel is classified as rainy when the TRMM PR rain rate exceeds 1 mm h⁻¹.

5.4. RESULTS

The statistics of the rainfall rates, which correspond to the 17 cloud fields, are summarized. The frequency of occurrences of TRMM PR rainfall rates that exceeded 1, 5, 10, 15 and 20 mm h⁻¹ thresholds are 25.1, 8.9, 4.6, 2.4 and 0.6 percent of the total 4500 pixels considered. The mean rainfall for rates higher than 1 mm hr⁻¹ is 5.7 mm hr⁻¹ and has a coefficient of variation

of 1.0 indicating that the magnitude of the standard deviation is equal to the mean.

5.4.1. Rainfall detection

Five TIR indices are applied to differentiate rainy from non-rainy pixels. The indices are $T_{10.8}$, $\Delta T_{10.8}$, $\nabla T_{10.8}$, $T_{10.8} - T_{6.2}$ and $T_{10.8} - T_{12.0}$. The estimated categorical statistics are shown in Fig. 2 for a rainfall rate detection threshold of 1.0 mm hr^{-1} which indicates a pixel is considered rainy when the TRMM PR rainfall rate exceeds 1.0 mm hr^{-1} . For the TIR indices, for instance, if the $T_{10.8}$ threshold is 210 K then a pixel is considered rainy when $T_{10.8}$ is lower than 210 K. Figure 5.2 shows only the results for the four best performing indices.

Figure 5.2 shows the categorical statistics are not equally sensitive to the TIR indices. In terms of *HSS*, the performance of the rain detection shows higher sensitivity to the $T_{10.8}$ index than the other indices. The highest *HSS* was obtained when using $T_{10.8}$ or $\Delta T_{10.8}$ while the lowest *HSS* was obtained when using $T_{10.8} - T_{12.0}$ (not shown in Figure 5.2) which was applied to differentiate cirrus from thick clouds. The result suggests the importance of considering both cloud top temperature and temporal evolution of clouds in rainfall detection algorithms.

In terms of *bias*, the performance of the rain detection is sensitive to all the indices except for the $\Delta T_{10.8}$ index. The *bias* indicates $\Delta T_{10.8}$ overestimated the number of rainfall occurrences while the other indices overestimated or underestimated rainfall occurrences depending on their threshold values.

In terms of *POD*, it is possible to obtain a *POD* value of 1.0 at some value of $T_{10.8}$ or $\Delta T_{10.8}$ indices. However, this is at the expense of large *FAR*. For instance, the *POD* when using $\Delta T_{10.8}$ is always 1.0 but the *FAR* is still large which suggests that the two categorical statistics should be interpreted jointly since TIR indices that are selected based on high *POD* can result in larger number of rainy pixels than detected by TRMM PR.

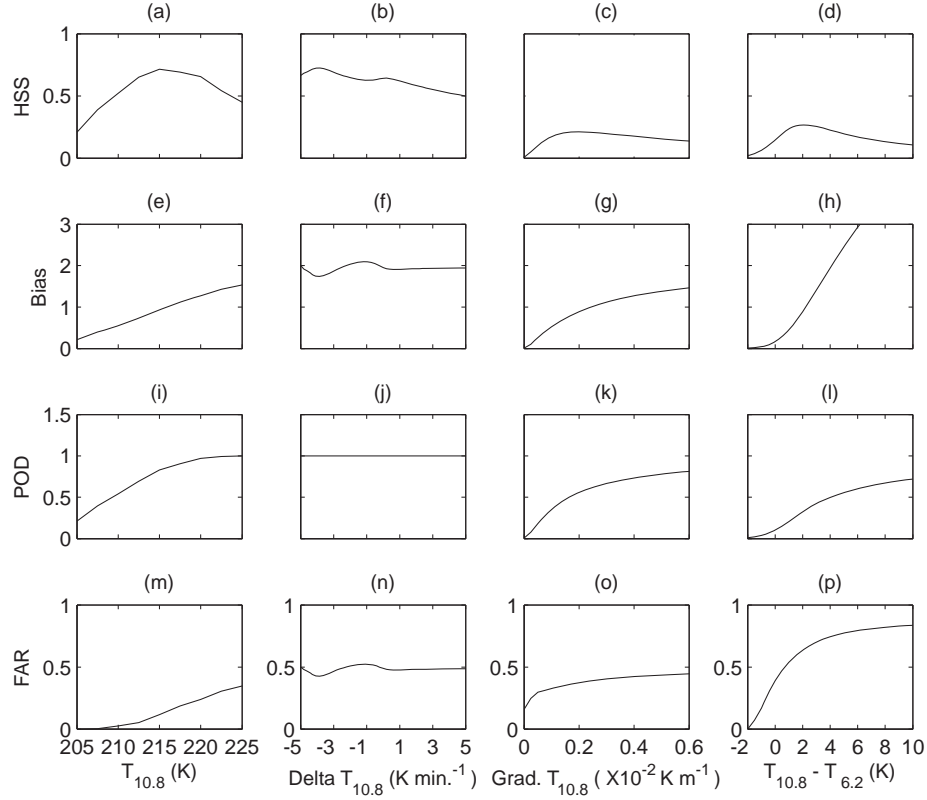


Figure 5.2: Categorical statistics for rainfall detection using brightness temperature ($T_{10.8}$), rate of temperature change ($\Delta T_{10.8}$ (here Delta $T_{10.8}$)), gradient of $T_{10.8}$ ($\nabla T_{10.8}$ (here Grad. $T_{10.8}$)) and brightness temperature difference ($T_{10.8} - T_{6.2}$). NB: The rainfall rate threshold is 1.0 mm h^{-1} . The scales of the axis of the categorical statistics are not equal for the selected indices.

Next, the optimized threshold values are identified for each of the TIR indices by maximizing the HSS and the resulting statistics are shown in Table 5.2 for threshold rainfall rates of 1.0 mm hr^{-1} and 10 mm hr^{-1} . Considering rainfall rates higher than 10 mm hr^{-1} resulted in very few data samples for the statistics and hence is ignored here.

Table 5.2: *Evaluation of rainfall detection using indices from MSG-2 images*

	$T_{10.8}$ (K)	$\Delta T_{10.8}$ (K min. ⁻¹)	$\nabla T_{10.8}$ ($\times 10^{-2}$ K m ⁻¹)	$T_{10.8} - T_{6.2}$ (K)	$T_{10.8} - T_{12.0}$ (K)
Rainfall threshold = 1.0 mm h ⁻¹					
Maximum <i>HSS</i>	0.716	0.724 and 0.642	0.210	0.334	0.092
Maximum <i>HSS</i> at	215.0	-3.8 and 0.2	0.2	6.0	2.5
<i>POD</i>	0.829	1.000	0.558	0.489	0.468
<i>BIAS</i>	0.937	1.928	0.890	0.635	0.929
<i>FAR</i>	0.115	0.476	0.374	0.230	0.435
Rainfall threshold = 10.0 mm h ⁻¹					
Maximum <i>HSS</i>	0.535	0.210 and 0.191	0.120	0.269	0.100
Maximum <i>HSS</i> at	205.0	-3.4 and 0.2	0.1	2.0	0.5
<i>POD</i>	0.580	1.000	0.344	0.324	0.200
<i>BIAS</i>	0.979	8.022	1.805	0.890	1.079
<i>FAR</i>	0.407	0.876	0.810	0.636	0.804

Considering all the indices, larger values of *HSS* and *POD* and smaller values of optimized *bias* and *FAR* were mostly found for the 1.0 mm hr⁻¹ threshold as compared to the 10 mm hr⁻¹ threshold. For both thresholds, the overestimation of rainfall occurrences was larger for $\Delta T_{10.8}$ than for the other indices. As such, the large *POD* value of 1.0 for $\Delta T_{10.8}$ is at the expense of many false alarms of rainfall.

For both rainfall rate thresholds, the *HSS* values showed slight sensitivity to $\Delta T_{10.8}$. The *HSS* results also suggest that the rainfall was produced during the growing, the mature and the dissipation cloud stages. Many studies assume that clouds produce rainfall during the growing stage only (e.g. Griffith et al. 1978; Vicente et al. 1998). However, our results suggest that such an assumption may affect the accuracy of the rainfall detection approach since results indicate that rainfall occurs during the three cloud stages.

In terms of *HSS*, $T_{10.8}$ and $(T_{10.8} - T_{6.2})$ performed better for detection rainfall rates higher than 10 mm hr⁻¹ than the other indices. However, $(T_{10.8} - T_{12.0})$ showed the lowest performance with an optimized *HSS* value that is

close to zero which according to Kuligowski (2002) suggests that the number of correctly detected rainy pixels is not much higher than expected to be estimated by random guessing.

For $T_{10.8}$, the optimum HSS value was found at 215 K for the 1.0 mm hr^{-1} rainfall threshold and at 205 K for the 10 mm hr^{-1} rainfall threshold. The result suggests that the accuracy of the traditionally used relatively high temperature thresholds as in Griffith et al. (1978); Arkin (1979); Arkin and Meisner (1987) is questionable for convective events.

In terms of *bias*, a temperature threshold that is higher than 215 K leads to overestimation of rainfall occurrences. At the optimum HSS , $T_{10.8}$ largely outperforms the other indices in terms of *bias*, *POD* and *FAR* although the HSS values for $T_{10.8}$ and $\Delta T_{10.8}$ indices are comparable.

In this study, the TIR indices are evaluated separately where two of the indices are a combination of the brightness temperature from two channels. The use of combined indices is left for future work.

5.4.2. Rainfall estimation

The mean of the TRMM PR rainfall rates at each brightness temperature $T_{10.8}$ and the 95 % confidence interval of the mean rainfall rate are shown in figure 5.3a. The confidence interval (CI) is estimated using:

$$CI = \bar{R} \pm 1.96 \frac{\sigma}{\sqrt{n}} \quad [5.11]$$

where: \bar{R} and σ are the mean and the standard deviation of the TRMM PR rainfall rates at each 1 K temperature interval, respectively and n is the number of data points at each temperature interval.

Figure 5.3a shows an exponential decrease of rainfall rate with an increase in $T_{10.8}$. In the same figure it is shown that the calibrated model has a standard error of 0.84 mm hr^{-1} with constant values $a = 45.851$ mm hr^{-1} , $b = 2.419 \times 10^{-3}$ and $c = 2.012$. It is also shown that the confidence interval of the mean rainfall rate is relatively larger at low temperature values than at high temperature values.

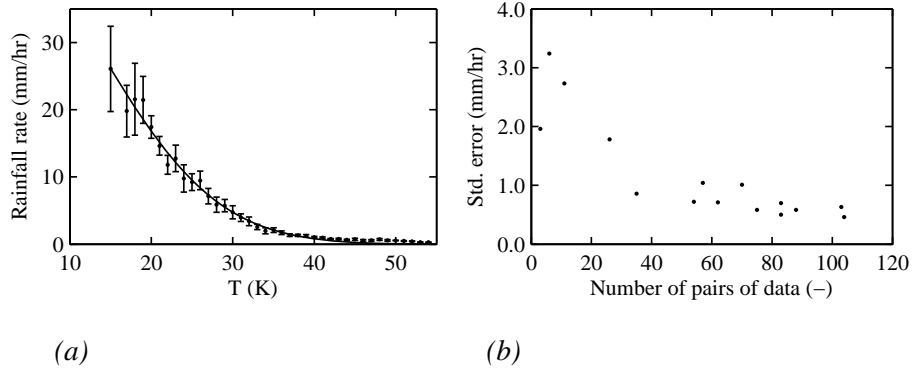


Figure 5.3: (a) Relation between estimated mean rainfall rate and brightness temperature ($T_{10.8}$). The 95 % confidence interval of the mean rainfall rate is indicated by the vertical bars. Note: T is defined as $T_{10.8}$ minus the threshold T_t which is set to 180 K. (b) Relation between standard error of mean rainfall rate for each 1 K temperature interval and the number of data pairs of TRMM PR and MSG-2 data used to estimate the mean.

A larger confidence interval at low temperatures can be caused by difficulties in estimating rainfall from high level clouds using TIR temperature as proxy variable and since high clouds are rarely observed in the data set. The latter is also indicated by the results in figure 5.3b in which the standard error of the mean rainfall rate is shown to decrease with an increase in the number of data points. In figure 5.3b, the largest standard error is associated to the lowest brightness temperature while the smallest standard error is associated to $T_{10.8} = 210$ K. Although it is not shown in the figure, the standard errors at temperature values higher than 210 K are close to zero.

5.4.3. Comparison against rain gauge observations

The rainfall estimates by the $T_{10.8}$ based exponential rainfall model (i.e. Eq 5.3) are compared against the observations from the rain gauge network. Figure 5.4 shows that *POD* for the rainfall model abruptly decreased with an increase in rainfall depth threshold while in terms of *bias*, the model underestimated the number of rainfall occurrences. However, it cannot be conclusive on the cause of such low performance for high rainfall depths partly since the ‘true’ spatial rainfall is not known.

The performance of the rainfall model is also compared against that of an inverse distance weighting (IDW) interpolation of the rain gauge observations. A power of 2.5 was applied for the IDW scheme since smaller weights were found more biased. The IDW estimates were defined through cross-validation, i.e. by removing the station for which the IDW estimates are made. In total, 73 data points have been used for comparison against the reference rainfall that is obtained from the rain gauges.

Figure 5.4a shows that in terms of *POD*, the IDW method have performed better than the TIR based rainfall model. In terms of *bias*, unlike the rainfall model, the IDW overestimates the number of rainfall occurrences, see figure 5.4b. The result suggests that even the interpolation technique can result in poor performance for relatively high rainfall rates. This is also shown by the scatter plot in figure 5.4c which indicates that both methods perform poor when they are compared against the reference data. An assessment of the errors in the rainfall estimates at hydrological relevant spatial and temporal scales such as the catchment scale is necessary. Such is a challenging topic in remote sensing based rainfall studies. For instance, Hossain and Huffman (2008) used a dense rain gauge network to investigate the errors in satellite based rainfall estimates, the spatial variation of the errors and the temporal variation of the errors for hydrological relevant scales that range between 0.04^0 to 1.0^0 . However, the data base in the present study is too limited and the rain gauge network is too sparse to carry out such studies and therefore future work in the region should focus on quantifying the errors at scales that are effective in hydrology.

5.5. CONCLUSION

In this study, observations from MSG-2 are used to detect rainfall and to estimate rainfall intensities while rainfall rates from TRMM PR served as ground truth. A network of 10 rain gauges was installed in the Lake Tana Basin to evaluate the remote sensing based rainfall estimates. The analysis revealed that by optimizing selected categorical statistics, optimum threshold values can be defined for image based indices that can be applied to detect rainfall.

For detecting rainfall rates higher than 1.0 mm hr^{-1} , the $10.8 \text{ }\mu\text{m}$ brightness temperature ($T_{10.8}$) and the rate of $T_{10.8}$ change ($\Delta T_{10.8}$) indices

performed better than the other indices that are the space gradient ($\nabla T_{10.8}$) and the brightness temperature differences (BTD) that are $(T_{10.8} - T_{6.2})$ and $(T_{10.8} - T_{12.0})$. For detecting rainfall rates higher than 10.0 mm hr^{-1} , $T_{10.8}$ and $(T_{10.8} - T_{6.2})$ indices performed better than the other indices. However, for both rainfall thresholds, the BTD $(T_{10.8} - T_{12.0})$ which is commonly applied to differentiate between cirrus clouds and convective clouds performed poor.

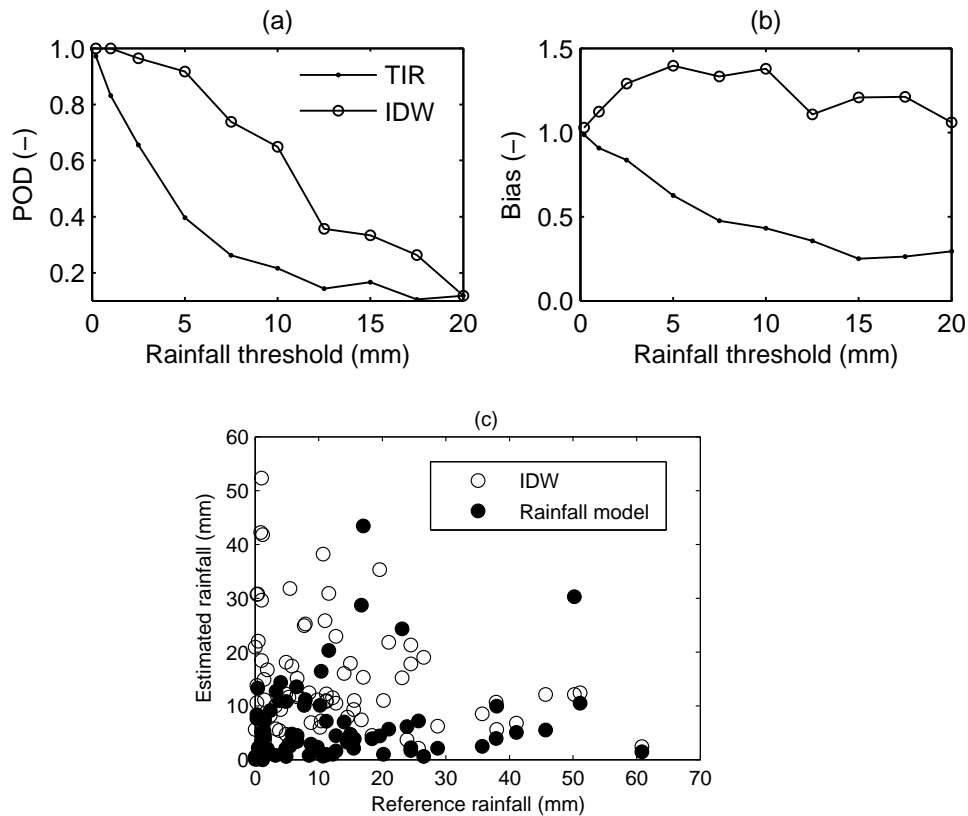


Figure 5.4: (a) POD and (b) bias of the Thermal Infrared (TIR) based exponential model and inverse distance weighting (IDW) rainfall estimates; (c) shows the plot of the estimated rainfall by both methods against the reference rainfall as obtained from the rain gauges.

Thresholds of 215 K and 205 K for $T_{10.8}$ can be used to detect convective rainfall rates higher than 1.0 mm hr^{-1} and 10 mm hr^{-1} , respectively. These temperature thresholds are much lower than the commonly applied thresholds in TIR based rainfall estimation such as 235 K

or 253 K. The two commonly applied constant temperature thresholds overestimated the number of rainfall occurrences in the study area.

For the TIR based rainfall model, the *POD* decreased rapidly with an increase in the selected threshold of rainfall depth. In terms of *POD*, the performance of the rainfall model was poor compared to that of the IDW interpolation method. In terms of *bias*, unlike the IDW approach, the rainfall model underestimated the number of rainfall occurrences.

Validation of the remote sensing estimates in areas like the Lake Tana basin is still a research challenge partly due to lack of knowledge on the 'true' spatial rainfall rate as affected by many local factors such as the presence of Lake Tana and the mountain areas. It is noted that presumably the rain gauge network in the basin is too sparse to capture the time-space distribution of rainfall. As such, research is required to study the rainfall structure for scales smaller than 5 km and to study the effect of variability of rainfall at scales smaller than the pixel scale on the accuracy of the rainfall detection and estimation.

6

REMOTE SENSING BASED CONCEPTUAL CLOUD MODELLING FOR RAINFALL SIMULATION



ABSTRACT¹

Motivated by the need for rainfall prediction models in data scarce areas, a simple storage based cloud model is adapted to routinely available thermal infrared (TIR) data. The data is obtained from the Spinning Enhanced Visible and InfraRed Imager (SEVIRI) of the Meteosat Second Generation (MSG-2) satellite. Model inputs are at 15 minute intervals of TIR temperature of cloud top and at 30 minute interval observations of pressure, temperature and dew point temperature from ground based stations. The sensitivity of the parsimonious cloud model to its parameters is evaluated by a regional sensitivity analysis (RSA) which suggested that model performance is sensitive to few parameters. The model was calibrated and tested for four convective events that were observed during the wet season in the source basin of the Upper Blue Nile River. Simulation results are satisfactory and indicate the effectiveness of this approach as evaluated by selected performance measures. The various characteristics of the rainfall events are satisfactorily simulated when ground based and remote sensing observations are combined.

Key words: Blue Nile, Lake Tana, rainfall simulation, conceptual model, MSG-2, Remote sensing

¹ This chapter is based on: Haile, A. T., Rientjes, T., Gieske, A., Jetten, V. G., Gebremichael, M., 2009. **Satellite remote sensing and conceptual cloud modelling for convective rainfall simulation**, *Advances in Water Resources*. Conditionally accepted.

6.1. INTRODUCTION

Prediction of convective rain events in data scarce areas is often hampered by unavailability of meteorological data at the ground surface as well as unavailability of atmospheric data that affect rain production. Rainfall prediction models in general can be grouped in statistical and numerical weather prediction models. Characteristic to statistical approaches is that the models do not simulate cloud dynamics to predict rainfall but much more rely on statistical analysis of historic rainfall observations. Statistical models commonly apply a set of multiple linear regression equations and as such are relatively simple. As opposed to these approaches, Numerical Weather Prediction (NWP) models are complex and data-demanding by the use of gridded and layered model domains for which sets of partial differential equations are solved. These approaches however have difficulty to predict highly dynamic convective rainfall systems since they commonly operate at much larger scales as compared to those at which convective cloud systems develop. Also, NWP models require a large amount of input data that commonly restricts their applicability.

An alternative to these approaches is to use models that have a physical base with a simple structure. Such models should only require input data which is routinely available but also should be parsimonious to allow application in areas that are data scarce. Such approaches are advocated, for instance, by Georgakakos and Bras (1984) and Dolciné et al. (1998) and the reader is also referred to the works by Wardah et al. (2008) and Haile et al. (2009b), see Chapter 5 of this thesis, where remote sensing observations serve as model inputs to estimate convective rainfall based on simple exponential models.

In the present study, the objective is to develop a parsimonious conceptual cloud model to predict convective rainfall in a data scarce area while input data must be through remote sensing and simple routinely recorded ground observations. The rationale for the study is to further develop and evaluate the simple cloud model by Georgakakos and Bras (1984a,b) through the availability of the observations of the Meteosat Second Generation (MSG-2) satellite at 15 minutes interval. In this study, it

is hypothesized that such high frequency of observation reflects on cloud properties that affect rainfall generation in convective cloud systems.

Georgakakos and Bras (1984a,b) developed a simple one layer, storage based cloud model with routinely available observations from meteorological stations as inputs. The approach did not explicitly model the details of micro-physical cloud processes but, by storage based accounting principles, the model simulated rainfall by reserving the mass balance. In the same work, the modeling of the cloud water phase was ignored although this phase acts as an intermediate phase between the water vapor and the rain water phases.

French and Krajewski (1994) and French et al. (1994) extended the model by Georgakakos and Bras (1984a,b) by using ground-based radar observations at temporal resolution of 10 to 15 minutes and satellite based thermal infrared (TIR) observations at temporal resolution of 1 hour. Results of this work were satisfying when satellite observations were combined with radar observations but the relatively low temporal resolution of 1 hour must be considered too low to capture the dominant dynamic properties of convective cloud systems. The model by Georgakakos and Bras (1984a,b) was also considered by Bell and Moore (2000b) and Andrieu et al. (2003) who used ground-based radar observations as model input. Both works indicated that model inputs at (higher) temporal resolution of 10 – 15 minutes could improve the performance of the model but results also indicated some deficiencies of the applied cloud model. Deficiencies appear to be primarily related to mismatches between rainfall intensities and rainfall volumes for the various events simulated. Also, Andrieu et al. (2003) speculate that ignoring the time required for condensation of water vapour is the main reason that rainfall is produced too quickly in comparison to field observations of rainfall.

A major challenge to further develop the conceptual model by Georgakakos and Bras (1984a,b) is to improve model performance while retaining the parsimonious nature to allow application in data scarce areas. With respect to model performance, the model should accurately simulate duration and the onset point when rainfall starts to fall from a cloud and it should accurately simulate the rainfall intensity distribution during the various stages of cloud development. Also, the cloud model must produce good estimates of the observed rainfall volume. Although much knowledge on model complexity and effective performance evaluation is gained in

hydrology, only few efforts are reported in the previous works to evaluate effects of model complexity and related model parameterisation.

In this study, the model by Georgakakos and Bras (1984a,b) is adapted and its performance is evaluated for the three objectives described above when two coupled reservoirs are defined to simulate convective rain events. Such allows considering the differences in rainfall response time along the vertical cloud profiles. This study differs from the work by French et al. (1994) since the model inputs are remote sensing observations at 15 minutes interval and ground based observations at 30 minutes interval. These relatively high observation frequencies are considered in this study to be essential for capturing the dominant properties of convective rain producing cloud systems. The model performance is evaluated for a data scarce region at the source basin of the Blue Nile River that has poor spatial and temporal correlation structure (Haile et al., 2009a) and where radar observations are absent.

The sensitivity of the model to its parameters is also evaluated. Model sensitivity is assessed through Regional Sensitivity Analysis (RSA) as available in the Monte Carlo Simulation Toolbox (MCST) (Freer et al., 1996; Wagener et al., 2001). In rainfall-runoff modeling such analysis has proven to be very effective in the work by Demaria et al. (2007) for instance. In this study, the RSA approach is applied to evaluate the sensitivity and performance of the conceptual cloud model to its parameters.

6.2. MODEL APPROACH

6.2.1. Model equations

Continuity equations for saturated water vapour, cloud water, rain water and air density are given by French and Krajewski (1994) and Bell and Moore (2000b) and here are presented as follows:

$$\frac{\partial Q_v}{\partial t} + \nabla \cdot Q_v \vec{u} = -\Delta Q_c \quad [6.1]$$

$$\frac{\partial Q_c}{\partial t} + \nabla \cdot Q_c \vec{u} = \Delta Q_c - \Delta Q_R \quad [6.2]$$

$$\frac{\partial Q_R}{\partial t} + \nabla \cdot Q_R \vec{u} - \frac{\partial(W_T Q_R)}{\partial z} = \Delta Q_R \quad [6.3]$$

$$\frac{\partial \rho_a}{\partial t} + \nabla \cdot \rho_a \vec{u} = 0 \quad [6.4]$$

where: \vec{u} is a vector of wind velocities in x , y and z directions of a Cartesian coordinate system, respectively; t is the simulation time step and W_T is the terminal velocity of rain water in still air. The quantities Q_v , Q_c and Q_R are saturation water vapor, cloud water and rain water concentrations, respectively; ρ_a is the density of dry air; $\partial / \partial t$ indicates the local time derivative; ∇ is the divergence operator; Δ indicates the rate of change; ΔQ_C and ΔQ_R are the rate of conversion of saturated water vapour to cloud water and the rate of conversion of cloud water to rain water, respectively.

ONE LAYER MODEL

The set of continuity equations (equations [6.1 – 6.4]) commonly is at the core of the algorithms of the complex NWP models but require simplifying assumptions when applied to the simple storage based cloud model as developed in this study. Assumptions however must be plausible while the principle of mass conservation must be respected. A first assumption is that the rate of conversion of water vapour to cloud water (ΔQ_c) is equal to the rate of cloud water conversion to rain water (ΔQ_R) with a transfer efficiency of 100 % (French and Krajewski, 1994). This assumption suggests that the water vapour input is directly transferred into liquid water and makes equation [6.2] redundant. A second assumption is that air is incompressible and thus air density becomes a fixed property that does not require updating during a simulation and makes equation [6.4] redundant.

By the first assumption, equation [6.1] and equation [6.3] can be combined to:

$$\frac{\partial Q_v}{\partial t} + \nabla \cdot Q_v \vec{u} + \frac{\partial Q_R}{\partial t} + \nabla \cdot Q_R \vec{u} - \frac{\partial(Q_R W_T)}{\partial z} = 0 \quad [6.5]$$

When assuming $\nabla \cdot \vec{u} = 0$, the second and fourth terms of equation [6.5] can be expressed as follows:

$$\nabla \cdot Q_v \vec{u} = u \frac{\partial Q_v}{\partial x} + v \frac{\partial Q_v}{\partial y} + w \frac{\partial Q_v}{\partial z} \quad [6.6]$$

$$\nabla \cdot Q_R \vec{u} = u \frac{\partial Q_R}{\partial x} + v \frac{\partial Q_R}{\partial y} + w \frac{\partial Q_R}{\partial z} \quad [6.7]$$

where: u , v , and w are wind velocities in x , y and z directions of a Cartesian coordinate system, respectively. The remaining terms are as defined previously.

Air is assumed to be incompressible and its saturation vapour density is assumed to be locally steady and horizontally uniform (Bell and Moore, 2000b) and thus only vertical variation of Q_v has to be considered. Thus the first term of equation [6.5] and the first and second term of the right-hand side of equation [6.6] are eliminated. By the assumption that rain water content is horizontally uniform, the first and second term of the right-hand side of equation [6.7] are eliminated. After substituting the non-zero terms of equations 6.6 – 6.7 in equation 6.5, after rearranging for the local derivative of the rain water concentration and when grouping similar terms, equation [6.5] becomes:

$$\frac{\partial Q_R}{\partial t} = -w \frac{\partial Q_v}{\partial z} + \frac{\partial}{\partial z} [Q_R (W_T - w)] \quad [6.8]$$

where: w is updraft velocity which is the wind velocity in the vertical direction. The remaining terms are defined previously.

By integration between the bottom level (z_b) and the top level (z_t) of a unit area of cloud layer equation [6.8] becomes:

$$\frac{dS}{dt} = w_0 (Q_{vb} - Q_{vt}) - (W_T - w_0) S_b \quad [6.9]$$

$$S = \int_{z_b}^{z_t} Q_R dz \quad [6.10]$$

where: S is the liquid water amount that is vertically integrated over the cloud layer; w_0 is the vertically averaged updraft velocity; S_b is liquid water (i.e. rain water) concentration at the bottom of the cloud layer; Q_{vb} and Q_{vt} are the saturation water vapour concentrations at the cloud bottom and cloud top surfaces, respectively. The remaining terms are

defined in the previous paragraphs. Following Georgakakos and Bras (1984a,b), a vertically-averaged updraft velocity w_0 is assumed in the cloud layer. The left hand side of equation [6.9] indicates that S is defined as a function of time.

TWO LAYER MODEL

In the work by Georgakakos and Bras (1984b) and Bell and Moore (2000b), it is shown that it is rather difficult to simulate rainfall events accurately with a one layer model. Here, it is hypothesized that these mismatches are caused by incorrect estimates of the response time that applies to the entire cloud layer. Such limitation restricts the models applicability when several peak intensities characterise a rain event. To overcome such limitations, Sinclair (1994) used a vertically varying rain formation time in one dimensional cloud modelling by arguing that the rain formation time in the upper part of a cloud system is longer than the rain formation time in the lower part of a cloud system. Such difference could result in several peak intensities in the hyetograph of a rain event. It is noted that rainfall intensity observations in the study area, which is the Gilgel Abbay watershed in Lake Tana basin, show that the rain events usually have more than one peak intensity.

Following the line of argument that is presented in the previous paragraph, the assumption that a cloud system can be represented by a single layer of reservoir (see Georgakakos and Bras, 1984a,b) has been relaxed. Instead, an alternative approach that represents the cloud system by two layers of reservoir is proposed, see figure 6.1. It is noted that the set of equation 6.5 – 6.10 easily can be applied for more than two layers but such can be against the premise for the development of a parsimonious conceptual model.

In this work a two layer model is proposed and equation [6.9] is adapted as follows:

$$\frac{dS_1}{dt} = w_0(Q_{vb} - Q_{vi}) - (W_{T1} - w_0)S_{b1} \quad [6.11]$$

$$\frac{dS_2}{dt} = w_0(Q_{vi} - Q_{vr}) - (W_{T2} - w_0)S_{b2} \quad [6.12]$$

where: the subscripts 1 and 2 indicate the lower and the upper part or, in modelling terms, layers of a cloud model. S is the liquid water

amount that is vertically integrated over the cloud layer; S_b is the rain water concentration at the bottom of the cloud layer; Q_{vb} and Q_{vt} are saturation water vapour concentrations at the bottom of the lower layer and at the top of the upper layer, respectively; Q_{vi} is saturation water vapour concentration at the interface of the two layers; W_T is the terminal velocity of rain water in the layers. The vertically-averaged updraft velocity (w_0) in the two layers is assumed equal while the updraft velocity is maximum at the interface of the two layers that is assumed where the pressure gradient in the cloud column is highest. Figure 6.1 shows the model schematization for the two layer model and reference is made to Georgakakos and Bras (1984b) and French and Krajewski (1994) for further description.

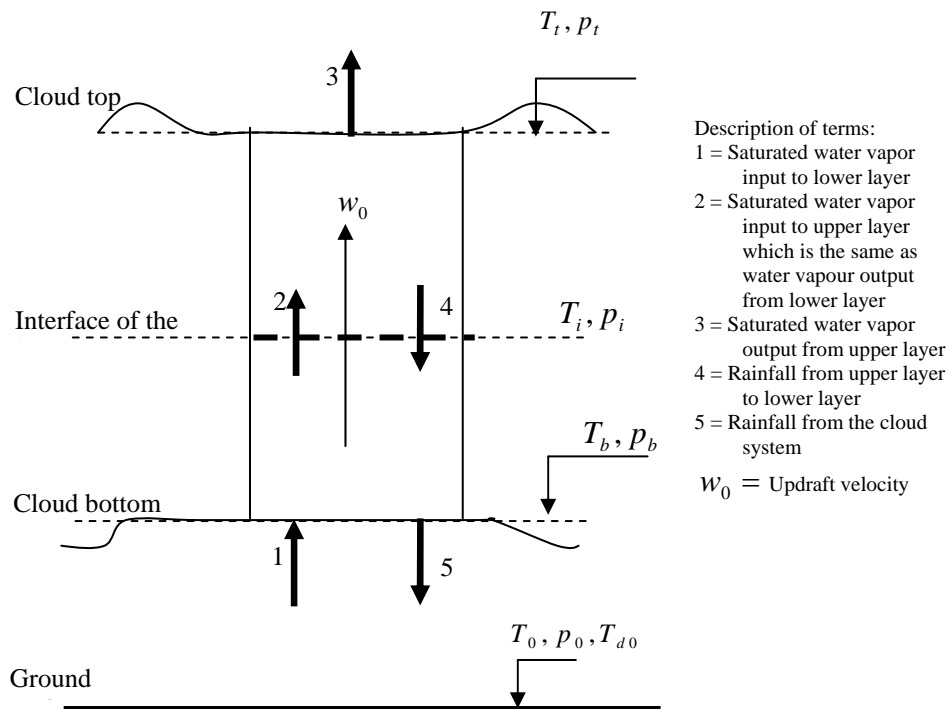


Figure 6.1: Schematization of the two layer model. T and p are temperature and pressure while subscripts 0, b, i and t indicate the temperature and the pressure at the ground surface, the cloud bottom, the interface of the layers and the cloud top surfaces, respectively.

Equations [6.11, 6.12] describe the mass balance of the cloud layers and their solution requires defining the terms which govern the rate of mass exchange. These terms are pressure and temperature at the bottom and top surfaces of the layers. In the following sections, the model inputs are presented first which is followed by the estimation of pressure and temperature.

6.2.2. Model inputs

The model inputs are ground-station pressure (p_0), temperature (T_0) and dew point temperature (Td_0), and cloud top brightness temperature (T_t) from the sensor of the MSG-2 satellite. The ground-station observations have a temporal resolution of 30 minutes while the MSG-2 images have a resolution of 15 minutes. Following French et al. (1994), the 10.8 μm brightness temperature is used as a proxy variable of cloud top temperature based on the assumption that the 10.8 μm emissivity of thick clouds is 1.0. Such assumption is plausible when convective clouds are thick and is commonly applied in TIR-based rainfall estimation (Adler et al., 1993; Boi et al., 2004; Todd et al., 1995). It is noted that the brightness temperature images have a spatial resolution of 3 km at nadir.

Ten tipping-bucket raingauges and a weather station were installed in the Gilgel Abbay watershed in May 2007. Information on the raingauge network and the statistics of the collected rainfall data are presented in the work of Haile et al. (2009a). Data from Durbet station have been used to calibrate the conceptual model and to evaluate model simulations. The station is selected since both rainfall and weather variables that are pressure, temperature and dew-point temperature are observed at the station. Rainfall at this station is dominated by light to moderate rainfall and has few convective rain events during the three month for which high resolution observations are available. It is noted that other network stations in the Gilgel Abbay area have recorded (many) more convective events but pressure, temperature and dew-point temperature are not available.

6.2.3. Model parameterization

PARAMETERIZATION OF THE ONE LAYER MODEL

First, the model parameterization for a one layer model will be presented that will be followed by a discussion on the difference between the parameterization of the one layer and the two layer models.

Cloud bottom pressure and temperature

The governing equations to estimate the lifting condensation level (cloud bottom) pressure (p_b) and temperature (T_b) are, see Georgakakos and Bras (1984a):

$$p_b = \frac{1}{\left(\frac{T_0 - Td_0}{223.15} + 1\right)^{3.5}} \cdot P_0 \quad [6.13]$$

$$T_b = \frac{1}{\left(\frac{T_0 - Td_0}{223.15} + 1\right)} \cdot T_0 \quad [6.14]$$

where: the terms are as defined previously. Equations 6.13 – 6.14 are applicable for the dry adiabatic condition, i.e. for $T_{d0} < T_0$. At initiation stage of cloud formation in the model, the cloud bottom temperature just exceeds the cloud top temperature, which is obtained from MSG-2 observations. As such, the time instant at which the cloud bottom temperature exceeds the cloud top temperature is considered as the onset point for the simulation period, i.e. $t = 0$.

Estimation of cloud top pressure from remote sensing observation

The cloud top temperature obtained from the TIR images has been used to estimate the cloud top pressure. First, the cloud layer is sliced in to several thin layers that have equal temperature differences between their top and bottom levels. Next, the pressure of the top of the first layer is estimated using the pressure and the temperature of the cloud bottom. The estimated pressure and the temperature of the top of the first layer serve as the start values to estimate the pressure at the top of the second layer. This

is repeated until the pressure of the cloud top is estimated while the entire procedure is repeated for all simulation time steps.

The equation to estimate the pressure can be expressed in finite difference form as follows (Stull, 2000):

$$p_{j+1} = p_j + \left(\frac{\partial p}{\partial T} \right)_j \cdot (T_{j+1} - T_j) \quad [6.15]$$

And,

$$\left(\frac{\partial p}{\partial T} \right)_j = \frac{p_j \left[1 + \left(\frac{b \cdot \omega_s}{T_j^2} \right) \right]}{a \cdot T_j + c \cdot \omega_s} \quad [6.16]$$

$$j = 0, 1, \dots, \text{int} \left(\frac{(T_b - T_t)}{(T_{j+1} - T_j)} \right) \quad [6.17]$$

where: T_j is the temperature at pressure level p_j while T_{j+1} is the temperature at pressure level p_{j+1} ; ω_s is saturation mixing ratio. The index j is a space index that indicates the pressure level while “int” indicates that an integer value is considered as a result of rounding-off. The temperature and pressure at the cloud bottom serve as start values, i.e. initial values. The solution for equation (15 – 16) is obtained by estimating the pressure from an initial temperature T_b at the cloud bottom and an initial pressure p_b at the cloud bottom up to temperature T_t at the cloud top by introducing a constant temperature step of $(\Delta T = T_{j+1} - T_j)$ which should be small enough (here 0.1 K) to avoid instability of the solution procedure and to ensure sufficiently accurate pressure values.

In equations 6.15 – 6.16, ω_s is the saturation mixing ratio at pressure p_j and temperature T_j . The equation to estimate the saturation mixing ratio is given in Georgakakos and Bras (1984a) and in meteorological text books such as by Stull (2000) and will be presented in the subsequent paragraphs. Following Stull (2000), the remaining terms of equation [6.16] are: $a = 0.28571$, $b = 1.35 \times 10^7 \text{ K}^2$, $c = 2488.4 \text{ K}$.

Rate of liquid water input

The rate of liquid water input to a unit area of cloud layer which results from the ascent of a moist air is given by Georgakakos and Bras (1984a):

$$I = \rho_m \cdot \Delta\omega \cdot w_0 \quad [6.18]$$

where: I is the rate of liquid water input to a cloud layer; ρ_m is the vertically averaged density of moist air for the cloud layer; $\Delta\omega$ is the change in mixing ratio between the cloud bottom and the cloud top surface and w_0 is the vertically averaged updraft velocity in the cloud layer.

The approximation introduced by Georgakakos and Bras (1984a) to estimate the saturation mixing ratio at a cloud top surface results in non-numeric values for cloud top temperatures lower than 223 K by a negative number to the power 3.5. However, MSG-2 observations indicate that temperatures of lower than 223 K are common for the top of the cloud systems at the source of the Blue Nile River and equations 6.20–6.21 are applied in this study, (see Stull, 2000).

The change in mixing ratio in equation [6.18] is defined as follows:

$$\Delta\omega = \omega_{s0} - \omega_{st} \quad [6.19]$$

The saturation mixing ratio at ground surface is defined as follows, see Stull (2000):

$$\omega_{s0} = \frac{e_{s0}}{p_0 - e_{s0}} \quad [6.20]$$

With,

$$e_{s0} = 611 \cdot \exp\left[\frac{L}{\mathfrak{R}_v} \cdot \left(\frac{1}{273} - \frac{1}{T_{d0}}\right)\right] \quad [6.21]$$

where: ω_{s0} is the saturation mixing ratio at the dew point temperature and the pressure at the ground surface; ω_{st} is the saturation mixing ratio at temperature and pressure at the cloud top surface; e_{s0} is saturation vapour pressure at ground surface dew point temperature T_{d0} ; exp is the base of the natural logarithm; L is latent heat of vaporization and

\mathfrak{R}_v is the gas constant for water vapour. The remaining terms are as defined previously. The saturation mixing ratio at the cloud top ω_{st} is estimated using the same form of equations 6.21 – 6.22 by replacing p_0 and T_{d0} by p_t and T_t .

The updraft velocity in equation [6.18] is estimated following Georgakakos and Bras (1984b), French and Krajewski (1994); Bell and Moore (2000b):

$$w_0 = e_1 (c_p \Delta T)^{0.5} \quad [6.22]$$

With,

$$\Delta T = T_m - T' \quad [6.23]$$

where: e_1 is a parameter for updraft velocity; c_p is the specific heat of dry air under constant pressure; T_m is the cloud temperature at pressure p' in a cloud layer assuming pseudo-adiabatic ascent and T' is the ambient temperature at pressure p' in a cloud layer. Following Georgakakos and Bras (1984b), the pressure p' and the temperature T' are:

$$p' = p_b - 0.25(p_b - p_t) \quad [6.24]$$

$$T' = \frac{T_o}{p_o^{0.286}} (p')^{0.286} \quad [6.25]$$

where: p_b and p_t are the pressure at the cloud bottom and the cloud top.

At each simulation time step, the temperature T_m at pressure level p' in a cloud layer is estimated by re-arranging equations 6.15 – 6.16 to solve for T_{j+1} instead of p_{j+1} until p_{j+1} equals p' . In the procedure, a constant pressure step of ($\Delta p = p_{j+1} - p_j$) and start values of temperature and pressure at the cloud bottom are used.

Rainfall rate

The rainfall rate from a cloud layer is estimated by the assumption that the rate is governed by the liquid water amount that is vertically integrated and by the response time of the cloud layer. The equation reads:

$$R = \frac{S}{\tau} \quad [6.26]$$

where: R is the rainfall rate from a cloud layer; S is the liquid water amount that is vertically integrated over the layer and τ is the response time of the layer, i.e. the time required to drain a unit depth of the cloud layer. The response time is defined as:

$$\tau = \frac{1}{W_T - w_0} \quad [6.27]$$

where: the terms are as defined previously. As discussed in the work by Bell and Moore (2000b), the response time of a cloud layer is considered a calibration parameter.

By introducing the parameters, the equation for the one layer model (equation [6.9]) now reads:

$$\frac{dS}{dt} = I - \frac{S}{\tau} \quad [6.28]$$

All terms are as defined previously.

PARAMETERIZATION OF THE TWO LAYER MODEL

The parameters of the two layer model are similar to the one layer model with the extension that the respective terms are substituted. The final form of the governing equations for a two layer model can be written as:

$$\frac{dS_1}{dt} = I_1 + \frac{S_2}{\tau_2} - \frac{S_1}{\tau_1} \quad [6.29]$$

$$\frac{dS_2}{dt} = I_2 - \frac{S_2}{\tau_2} \quad [6.30]$$

where: the subscripts 1 and 2 indicate the lower and the upper layer of the cloud model, respectively; S is the liquid water amount that is vertically integrated over a layer; I is the rate of liquid water input to a layer; τ is the response time of the cloud layer that is the time required to drain a unit depth of liquid water of a cloud layer and t is the simulation time instant.

The rates of liquid water that serves as input to the two layers are defined as:

$$I_1 = \rho_m \cdot (\omega_{s0} - \omega_{si}) \cdot w_0 \quad [6.31]$$

$$I_2 = \rho_m \cdot (\omega_{si} - \omega_{st}) \cdot w_0 \quad [6.32]$$

where: ω_{s0} , ω_{si} and ω_{st} are the saturation mixing ratio at ground surface, at the interface of the two layers and at the cloud top, respectively. These terms are estimated using a similar form to equations 6.18 – 6.19 but using the temperature and the pressure at the respective cloud surface levels. The remaining terms are defined as previously.

Rain generation time

In this study, the rain generation time that is the time difference between the cloud initiation and the rain start time is parameterized by adapting the equation by Sinclair (1994) as follows:

$$t_r = \alpha \left(0.5 + \frac{1}{\pi} \arctan^{-1} \left(\frac{p_m - p}{50} \right) \right) + t_{event} \quad [6.33]$$

where: t_r is the rain generation time in a cloud layer, p_m (hPa) is the pressure level below which the saturation water vapour is assumed to condense immediately after entering the cloud system and α is a constant. The correction t_{event} is introduced to account for mismatches in rain generation time due to effects of unknown initial conditions and the deficiency of the model to simulate microphysical processes such as water vapour condensation. Both effects may introduce time shifts between the start of the simulated and the observed rainfall events. It is noted that equation [6.33] also allows to extend the model equations to a larger number of cloud layers since the rain generation time is estimated based on pressure levels within a layer.

Figure 6.2 shows the formation time for three p_m values that are 500 hPa, 600 hPa and 700 hPa at different pressure levels p and $\alpha = 7000$ and $t_{event} = 0$. The figure shows that the water vapour at high pressure levels condenses faster than the water vapour at low pressure levels. Such a difference in formation time cannot be well simulated by a single layer model and constitutes the rationale for the proposed two layer model approach in this work.

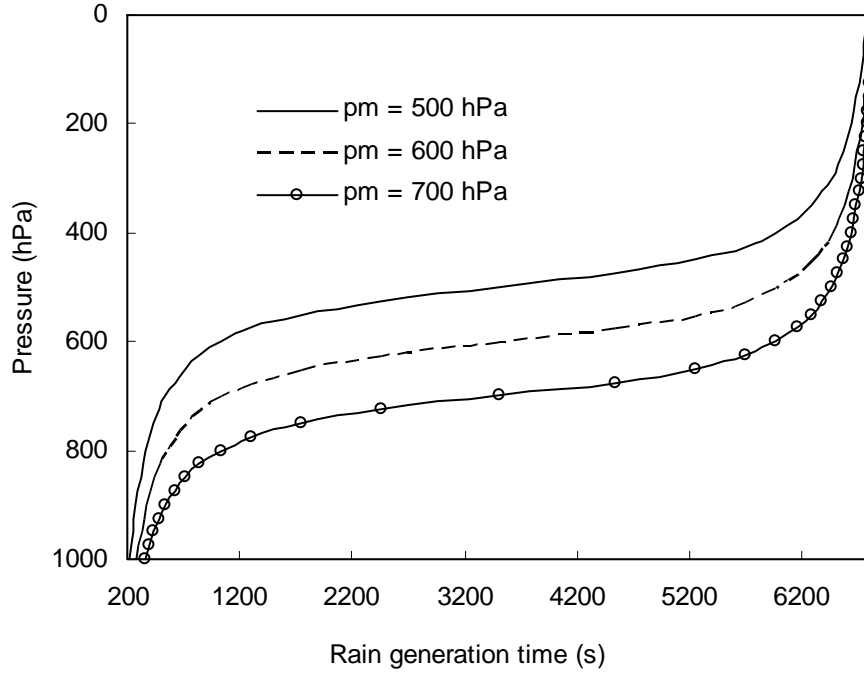


Figure 6.2: Rain generation time at different pressure levels of a cloud layer.

In this study, the interest is on the time instant at which rain starts to fall off the lower layer. Therefore to arrive at a single response time, the rain generation time is estimated for a single pressure level and therefore $p_m - p$ is replaced by a single parameter dp .

The rain generation time for the two layers is estimated using equation [6.33] by replacing $(p_m - p)$ by dp_1 and dp_2 for the lower and upper layer, respectively.

The set of equations 6.29 – 6.30 has been solved by implicit time integration based on the backward Euler method. A computer code was developed in Matlab® for both the one layer and the two layer cloud models.

6.3. REGIONAL SENSITIVITY ANALYSIS

The July 22, 2007 rain event was selected for model sensitivity analysis and calibration. The model inputs were observations from the Durbet weather station and observations from MSG-2. The distribution of the rainfall intensities of the selected event shows two peaks that suggests a time lag between the responses from the lower and the upper part (i.e. layer) of the cloud. The cloud developed in the vicinity of the weather station and provided the opportunity to test the model.

The Regional Sensitivity Analysis (RSA) is applied to identify sensitive model parameters and to define optimal ranges of parameter values. RSA is a non-parametric method that evaluates large number of parameter sets in terms of model performance. RSA has many applications in hydrological modelling (Demaria et al., 2007; Freer et al., 1996; Wagener et al., 2001) and supports model evaluation by revealing information on how well parameters are identified but also allows for identifying deficiencies of the selected model structure (see, Lamb et al., 1997).

The RSA module that is available in the Monte Carlo Analysis Toolbox (MCAT) is used (see Wagener et al., 2001]). The RSA in MCAT is based on a procedure provided by Freer et al. (1996) and, in brief, has the following steps. First, for each parameter, a prior value range is specified that is commonly based on a uniform distribution. Then a large number of parameter sets (for instance, 10,000 – 100,000) are randomly generated by a Monte Carlo sampling procedure. For each parameter set the model is run and parameter sets are ranked with respect to selected objective functions that served as a measure of model performance. Commonly, the objective functions are transformed through normalization where higher values indicate better performance. In the procedure, it is common to select the best performing parameter sets for further analysis. In the work by Spear and Hornberger (1980), these parameter sets are termed behavioural and usually only the best 10 % to 20 % of all parameter sets are selected. In the procedure, results are rearranged and spread over ten bins of an equal number of parameter sets where the first bin includes sets that indicate best model performance while the tenth bin includes the parameter sets that perform poorest. Model sensitivity to single parameters is evaluated based on the cumulative distribution of the objective functions values versus the range of the parameter values in each bin. A steep gradient in the

distribution function indicates high sensitivity and suggest that parameter are well identifiable.

In the present study, the following objective functions are selected:

$$abias = \frac{\left| \sum_{i=1}^N (S_i(\theta) - O_i) \right|}{\sum_{i=1}^N O_i} \quad [6.34]$$

$$M4E = \frac{1}{N} \sum_{i=1}^N (S_i(\theta) - O_i)^4 \quad [6.35]$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (S_i(\theta) - O_i)^2} \quad [6.36]$$

$$R^2 = \left(\frac{\sum_{i=1}^N (O_i - \overline{O_i})(S_i(\theta) - \overline{S_i(\theta)})}{\sqrt{\sum_{i=1}^N (O_i - \overline{O_i})^2 \sum_{i=1}^N (S_i(\theta) - \overline{S_i(\theta)})^2}} \right)^2 \quad [6.37]$$

where: S_i is the simulated rainfall at the i^{th} time step; θ is a specific parameter set; O_i is the observed rainfall at the i^{th} time step, and N is the total number of observations. The objective function *abias* is the absolute bias and measures the relative volumetric error by weighing the errors equally while *RMSE* is the root mean square error and indicates overall performance by weighing the larger errors more than the smaller errors. The coefficient of determination (R^2) is the square of the correlation coefficient between the observed and the simulated rainfall and also indicates overall performance. A fourth objective function used is the *M4E* that is an indicator of goodness-of-fit for peak values since large errors are weighted to the power of 4 (see De Vos and Rientjes, 2008).

6.4. RESULTS

6.4.1. Data analysis

First, statistics of the model inputs which are obtained from the Durbet ground station and satellite remote sensing are shown. Table 6.1 shows the autocorrelations of the inputs for July 22, 2007 for lag times up to 1 hour. Both the ground stations and the satellite remote sensing observations were recorded between 12:00 – 18:00 Local Standard Time (LST). Table 6.1 shows that the model inputs have high autocorrelation. For instance, the 30 minute lag autocorrelations range from 0.56 – 0.8 which suggests that current values of the inputs carry high information content about their values after 30 minutes. The result indicates that the model inputs can be forecasted with simple techniques. Table 6.1 also shows that the dew point temperature and TIR show the most rapid decrease in autocorrelation with lag time.

Table 6.1: *Autocorrelation of the model inputs for the July 22, 2007 event. The data are observed between 12:00–18:00 Local Standard Time. Note: Ground–station data is available every 30 minutes while remote sensing data is available every 15 minutes.*

Time lag	Ground-station			Remote sensing
	Pressure	Temperature	Dew point temperature	TIR brightness temperature
15 minute	–	–	–	0.907
30 minute	0.696	0.798	0.56	0.717
45 minute	–	–	–	0.479
1 hour	0.428	0.53	0.099	0.226

Table 6.2 shows the cross-correlations for the inputs of the July 22 event. In the table, the cross-correlation value of 0.268 can be interpreted as the correlation between the current pressure value and the expected temperature value after 1-hour. In the same manner, the cross-correlation value of 0.774 can be interpreted as the correlation between the current temperature value and the expected pressure value after 1 hour. Table 6.2 also shows that the current value of the ground-station temperature in general carries high information content about the expected (after 1 hour) values of the other model inputs. The cross-correlations in table 6.2 are larger than the autocorrelations in table 6.1 which suggests that any of the

conceptual cloud model inputs can be predicted based on the current values of the other model inputs.

Table 6.2: The 1-hour lag cross-correlation of the model inputs for the calibration period (July, 22, 2007) and the validation period (August 16, 2007). The data are observed between 12:00–18:00 Local Standard Time.

	Pressure	Temperature	Dew point temperature	TIR brightness temperature
July 22, 2007				
Pressure	0.428	0.268 ¹	0.719	0.268
Temperature	0.774	0.53	0.370	0.719
Dew point temp.	0.541	0.495	0.099	0.696
TIR temperature	0.687	-0.189	-0.339	0.226

¹Note: the cross-correlation value of 0.268 is interpreted as the correlation between the current pressure and the expected temperature after 1-hour. The other values are interpreted in a similar manner.

6.4.2. Model sensitivity

In this study, six parameters of the model are evaluated by the RSA. These parameters are the updraft velocity parameter (e_1); the rain generation parameters of the upper and the lower layers dp_1 and dp_2 , respectively; the response times of the upper and the lower layers that are τ_1 and τ_2 , respectively and the parameter t_{event} that accounts for the change in rain generation time by changing the onset time of rain formation by updraft of moisture. A constant value of $\alpha = 7000$ has been specified for both layers after a preliminary sensitivity analysis of the rain generation time in equation [6.33].

For the RSA, at first the number of Monte-Carlo model runs was set to 10, 000 but was increased to 30,000 until the cumulative distribution of the normalized values of the objective function for each parameter in the 10 bins showed consistency and stability. Following Freer et al. (1996), the parameter sets with a coefficient of determination (R^2) value greater than 0.5 are considered behavioural. The R^2 value of 0.5 indicates a moderate performance and in this study it is noted that the value was arbitrarily chosen for performance evaluation.

Figure 6.3 shows the distribution of the normalized cumulative frequency of the *abias* objective function for 6 model parameters. Each

curve in a box of the figure represents the cumulative distribution in a certain bin that were ranked from 1 to 10 based on the likelihood values. Model sensitivity is indicated by the location of the cumulative distribution functions in objective function space and by the shape and the gradient of the respective distributions where a steep gradient indicates high sensitivity while a flat gradient indicates low sensitivity. The model is said to be completely insensitive to a certain parameter if the cumulative distributions of the objective functions for all 10 bins have uniform distribution that as such then can be overlain. Such could be for a single objective function but, possibly, also for multiple objective functions (see figures 3 and 4).

From figure 6.3 it can be concluded that the model is most sensitive to the updraft velocity parameter (e_1) since the plot shows steep gradients for the cumulative distribution curves for all 10 bins. The parameter e_1 controls the rate of liquid water input in the cloud layers. The RSA indicates the importance to accurately estimate e_1 to obtain small *abias* objective function values which is the absolute value of relative volume error. In terms of *abias*, the model is least sensitive to those parameters that govern the rain generation time and the peak rainfall rates.

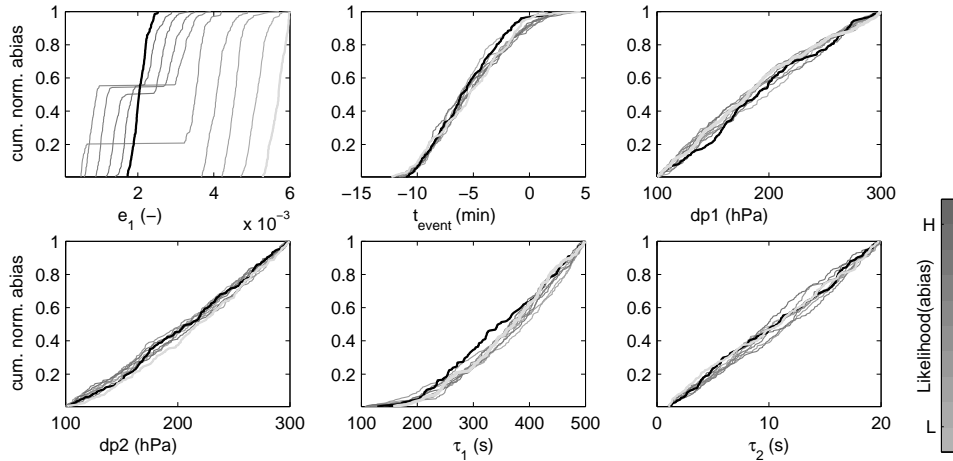


Figure 6.3: Cumulative normalized (Cum. Norm.) distribution of the parameters for *abias*. Note: Black lines represent parameter values with high model performance (bin 1) while brighter lines represent parameter values that give lower model performance (bin 2 – 10).

Figure 6.4 shows the cumulative distribution for the selected six parameters in each of the 10 bins that were ranked with respect to $M4E$ which mainly measures the model ability to reproduce peak rainfall intensities. Compared to the cumulative distribution functions in terms of $abias$, there was no major change in the shape of the distribution function except some slight changes for the parameters e_1 and τ_1 . In terms of $M4E$, the model is most sensitive to the updraft velocity parameter e_1 since the distribution functions show large differences. The model is also somewhat sensitive to the event marker (t_{event}) and the response time of the lower layer (τ_1) while it is least sensitive to the parameters of the rain generation time in the upper layer (dp_2) and the response time of the upper layer (τ_2). The RSA also suggests that the model performance is more sensitive to parameter (τ_1) than to parameter (τ_2). This indicates the importance to accurately estimate the lower layer response time for rainfall simulation.

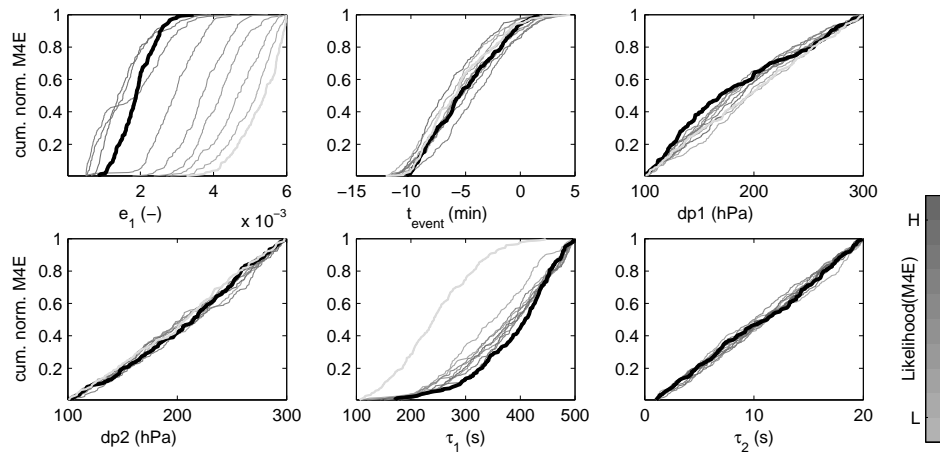


Figure 6.4: Cumulative normalized (Cum. Norm.) distribution of the parameters for $M4E$. Note: Dark lines represent parameter values with high model performance (bin 1) while brighter lines represent parameter values that give low model performance (bin 2 – 10).

6.4.3. Model calibration

The July 22, 2007 rain event has served for calibration of the cloud model. During calibration, a constant value of $\alpha = 7000$ has been specified for both layers after a preliminary sensitivity analysis. The parameter

governs the time at which the rain starts to fall from the respective layers of the cloud model. The conceptual model has been manually calibrated. First, the values of the updraft velocity parameter (e_1) have been optimised since this parameter controls the volume of rainfall that is considered the most important. Second, the start time of the rainfall is calibrated by changing the values of the rain generation parameter of the upper and the lower layers dp_2 and dp_1 , respectively which is followed by the response times of the upper and the lower layers τ_2 and τ_1 , respectively. Also, the correction factor t_{event} is considered to correct for the change in the rain generation time as caused by effects of uncertain initial conditions for the four events. As shown by the RSA, the model performance is most sensitive to e_1 and τ_1 . It is noted that the calibrated parameters values of the July 22 event have served as start values to the remaining rain events that were observed on June 1, June 8 and August 16 of the year 2007, see table 6.3.

The updraft velocity parameter (e_1) affects the rate of saturation water vapour input to the cloud layers. Overall, the calibrated value of e_1 is similar to that reported by Georgakakos and Bras (1984b) who used a value of 0.002. In the present study, the same initial conditions are specified for all events. As such, it can be speculated that the small difference in the value of e_1 for the events is partly caused by this assumption. The four events have equal values of dp_1 but have somewhat different values of dp_2 .

Table 6.3 also shows that all events have equal response time (i.e. τ_1) for the lower layer except for the June 1 event which has smallest response time. However, τ_2 shows some changes with a change in the event property. The calibrated response times τ_1 and τ_2 suggest that rain falls faster from the upper layer than from the lower layer. In terms of rain generation time, the values for t_{event} range between -44 minutes for the June 1 event to +32 minutes for the June 8 event.

Table 6.3: *Model parameter values of the two layer model.*

Symbol	Parameter	Jun. 1	Jun. 8	Jul. 22	Aug. 16	Unit
e_1	For updraft velocity	0.041	0.0020	0.0020	0.0040	–
t_{event}	Correction for rain generation time	-44	32	0	- 33	min.
dp_1	Rain generation time of lower layer	105	105	105	105	hPa
dp_2	Rain generation time of upper layer	150	150	180	180	hPa
τ_1	Lower layer response time	210	310	310	310	s
τ_2	Upper layer response time	3	3	3	3	s
Values of objective functions						
	<i>abias</i>	-0.110	-0.041	0.172	0.049	–
	<i>M4E</i>	0.192	0.321	0.024	0.274	mm ⁴
	<i>RMSE</i>	0.014	0.232	0.172	0.422	mm

The simulated and the observed rainfall hyetographs of the July 22 event are shown in figure 6.5a. The plot indicates that the cloud model satisfactorily reproduced both the shape and the volume of the observed hyetograph. The start and the end time of the simulated event agree to those of the observed event. Peak intensities are also reproduced satisfactorily although not all peaks could be simulated with exactness. The simulated and the observed hyetographs of the rainfall intensities of the June 1, June 8 and August 16 events are shown in figure 6.5b – 6.5d.

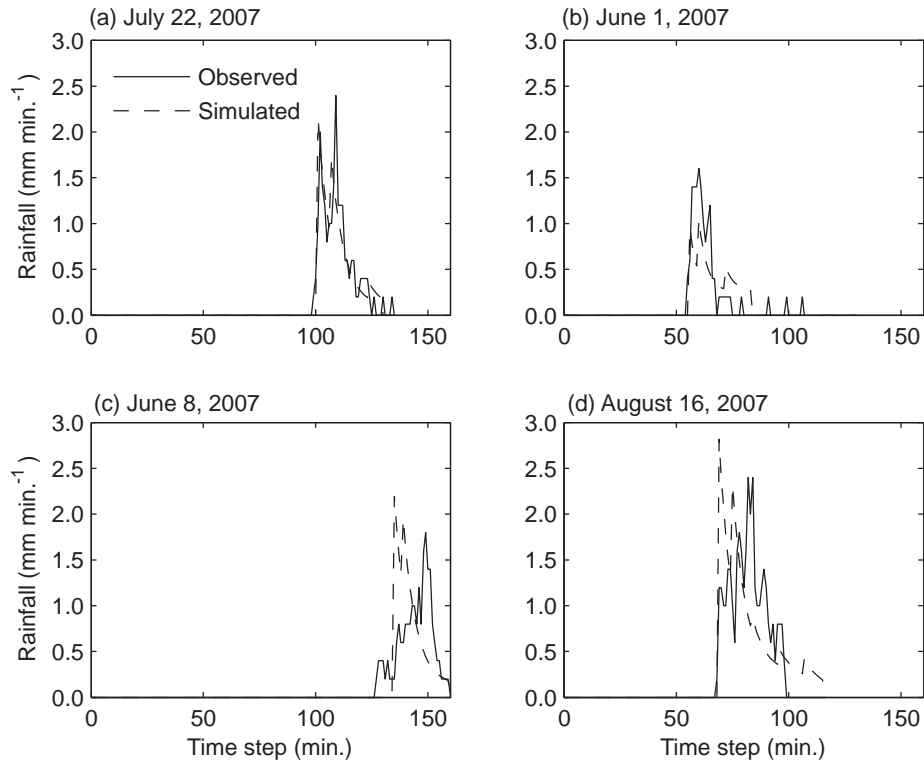


Figure 6.5: Observed and simulated rainfall intensity of four rain events.

Table 6.4 shows accumulated values of the observed and the simulated event depth for the four events. After calibrating each rain event independently, the simulated and the observed rainfall depth of the four events showed a good agreement. Calibrating the rain events using the parameter values of the July 22 event resulted in some deterioration but overall in reasonable agreement between the simulated and the observed rainfall depths for the June 1 and June 8 events. However, the model underestimated the observed rainfall depth of the August 16 event. Results of table 6.4 also reveal that the main limitation of using the July 22 parameter values is that the rain start time of the other three events is not estimated satisfactorily. The calibrated parameters for the July 22 event cause a time shift between the start time of the simulated and observed rainfall. The time shift ranges between -46 minutes for the June 1 event which indicates a delay in model response and 26 minutes for the June 8 event which indicates a model response that is too early.

It is observed that the maximum updraft velocity increases steeply during the initial phase of cloud formation but then levels off, see figure 6.6. This characteristic of the updraft velocity is observed for three events out of four. In this study, it is noted, however, that June 1 event is exceptional since rain starts to fall while the updraft is still increasing before it levels off. As such, for the three events a relation can be established between the time at which the steep increase of the updraft levels off and the time at which rain starts to fall.

Table 6.4: Observed and simulated characteristics of the four rain events. A negative time shift indicates a delay in model response as compared to observations.

	Calibrated for each event			Using July 22 parameters		
	Observed depth (mm)	Simulate d depth (mm)	Time shift (min.)	Observed depth (mm)	Simulate d depth (mm)	Time shift (min.)
June 1	14.4	13.4	-1	14.4	11.3	-46
June 8	21.6	21.8	-8	21.6	14.3	26
July 22	21.6	21.8	-2	21.6	21.8	-2
Aug. 16	36.0	37.8	-1	36.0	19.0	-33

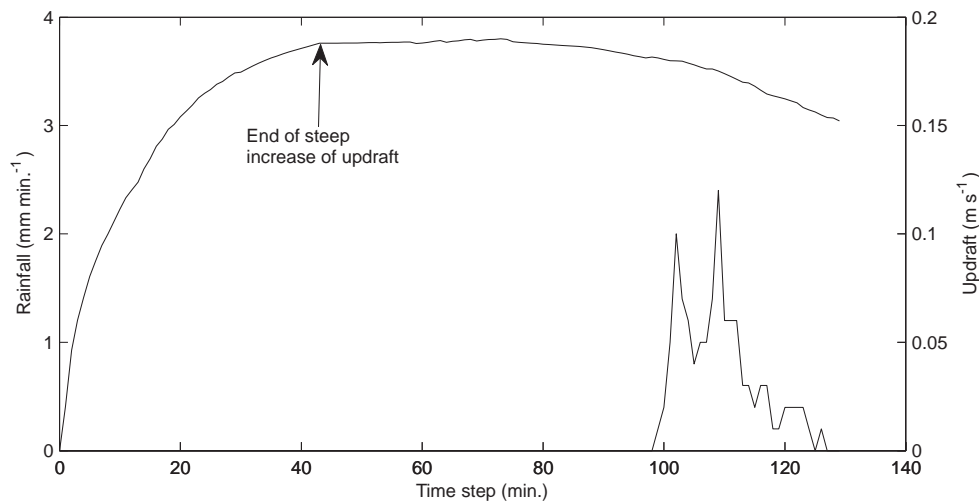


Figure 6.6: Observed rainfall and estimated updraft velocity for July 22, 2007 event. Note that at first, the updraft increases rapidly but levels off after some time period.

Table 6.5 shows the difference in the start time of the observed rain event and the time at which the maximum updraft velocity (w_{max}) occurs. In the present study, the time to w_{max} is considered to be the time at which

the steep increase in the updraft levels off since the time at which the cloud is formed. For three out of four events, this time difference ranges between 50 – 59 minutes. As such, if the time to reach w_{\max} is used as a rain event marker, then the rain events start 54 ± 5 minutes after w_{\max} occurred. The validity of the characteristics of the updraft velocity for the three events should, however, be evaluated for other events as well since there are some limited observations of convective rainfall intensity at Durbet station in June – August of 2007. Table 6.5 shows that there is a very large variation for the time difference between the start time of the observed rain event and the time at which the cloud is formed. The time difference for the four events ranges between 68 – 127 minutes which indicates larger variation than that indicated by using the w_{\max} time.

Table 6.5: *The relation between rainfall initiation time and the time to maximum updraft and cloud formation is observed for the four events.*

	Jun. 1	Jun. 8	Jul. 22	Aug.16	Unit
1. Time to max. updraft	15:45	17:20	14:42	15:44	h:min.
2. Cloud formation time	14:20	16:05	14:02	15:26	h:min.
3. Start of observed rain	15:15	18:12	15:41	16:34	h:min.
4. Row 3 – row 1	-30	52	59	50	min.
5. Row 3 – row 2	85	127	99	68	min.

Note: Row 3 – row 1 is the time difference between the time at which the observed rainfall started and the time at which the maximum updraft is observed. Row 3 – row 2 is the time difference between the time instants at which rainfall started and the time at which the cloud is formed.

6.5. DISCUSSION AND CONCLUSION

The main impetus for developing the conceptual model is to have a parsimonious model approach that simulates the various rainfall characteristics accurately but with relatively low requirement for model input. The number of inputs should be small but also simply be obtained. This is following the comments in Georgakakos and Bras (1984a), where it is stated that model inputs should be readily available but, in case not available, predictable through simple techniques such as regression. In

Georgakakos and Bras (1984a), it is described that such is considered crucial for applications in near real time rain prediction. The conceptual model proposed in this study meets these requirements since inputs are weather observations at ground surface and satellite infrared observations that are readily available from the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT). The model is an extension of that proposed by Georgakakos and Bras (1984a,b). Bell and Moore (2000b) described that the model by Georgakakos and Bras (1984a,b) is a simple mass-balance model of the lower atmosphere while it is discussed that the model structure has much in common with conceptual water-balance models as are applied in hydrology.

The structure of the proposed conceptual cloud model in this study is not comparable to the structure of physically based numerical weather prediction (NWP) models which generally simulate the complex interactions and processes of all three water phases in a cloud system. Also, NWP models have their own limitations and according to Pedder et al. (2000) an obvious limitation is the model's inability to represent small-scale precipitation events explicitly.

As discussed in the Sections on the model equations and the model parameterization, many assumptions are introduced for model development. For instance, the horizontal movement of clouds which can cause moist air advection is not explicitly considered although Pedder et al. (2000) noted that it is not clear whether it is in fact realistic to neglect horizontal advection terms. With respect to the model development, French and Krajewski (1994) suggested that future work in rainfall simulation will indicate that either remote sensing based formulation is the best way to proceed or that another, quite different path, leads to a solution of the rainfall estimation and forecasting problem. This work shows effectiveness of a remote sensing based approach.

In the present study, (i) a two layer model of rainfall simulation is developed and evaluated, (ii) weather observations at ground station and satellite thermal infrared observations at 30 minute and 15 minute resolution served as model inputs, (iii) a RSA is performed to evaluate the sensitivity of the model to its parameters for various objective functions and (iv) model performance is evaluated for selected rain events and a list of calibrated parameter values is presented.

A regional sensitivity analysis that considers single evaluation criteria has been applied to evaluate the sensitivity of the model to six parameters. The analysis is performed for 4 objective functions and revealed that the model is sensitive to only few parameters. Model sensitivity changed from one objective function to another (see table 6.6 and 6.7) suggesting that multiple sets of optimal parameter values can be established as subject to the selected objectives. Changes in sensitivity also suggest that calibration of the cloud model inherently involves a number of criteria but such complex, multi-criteria, evaluation has been ignored here for reasons of brevity.

Table 6.6: *Best 10 performing parameter sets in terms of M4E*

Rank	e_1	t_{event}	pd_1	pd_2	τ_1	τ_2	M4E
1	0.0025	-0.4	124	249	444	1	0.1084
2	0.0020	-2.7	146	284	342	18	0.1152
3	0.0021	1.6	105	164	434	9	0.1298
4	0.0020	-2.5	134	255	346	17	0.1301
5	0.0017	-0.1	104	204	431	15	0.1421
6	0.0016	-0.3	111	216	300	4	0.1481
7	0.0018	1.2	101	142	360	13	0.1492
8	0.0017	-0.2	108	180	359	19	0.1493
9	0.0022	-4.1	156	276	477	6	0.1493
10	0.0026	-1.8	136	218	386	2	0.1548

Table 6.7: *Best 10 performing parameter sets in terms of abias*

Rank	e_1	t_{event}	pd_1	pd_2	τ_1	τ_2	abias
1	0.0021	-6.7	178	247	248	6	0.0000
2	0.0022	-8.8	270	159	282	16	0.0000
3	0.0020	-5.5	133	219	416	11	0.0001
4	0.0020	-9.1	268	296	498	4	0.0002
5	0.0021	-1.9	143	129	322	8	0.0002
6	0.0023	-10.0	298	228	212	19	0.0003
7	0.0020	-6.8	293	233	460	12	0.0003
8	0.0022	-1.3	108	261	212	12	0.0003
9	0.0021	-1.6	111	122	376	19	0.0005
10	0.0021	-8.8	270	267	300	7	0.0006

In terms of the *abias* objective function the model is highly sensitive to the updraft velocity parameter (e_1) and the response time of the lower

layer (τ_1) while the model showed low sensitivity to the other four parameters. The former determines the rate of moist air input to the cloud layer while the latter determines the rainfall rate. In terms of the *M4E* objective function, the model is highly sensitive to the updraft velocity parameter (e_1) but also it is somewhat sensitive to the event marker (t_{event}) and the response time of the lower layer (τ_1). The result reveals that the updraft velocity and the response time of the lower layer affect the volume as well as the peak of the rainfall hyetograph but the rain generation time largely affects the peak rain rate while it has less effect on rainfall volume.

The two layer model has been successfully calibrated for the July 22 rain event. The model reproduced the shape of the hyetograph well that is characterized by timing, event duration and peak intensities but also the volume of the observed rainfall hyetograph is simulated well.

In the calibration, first an agreement between the observed and simulated rainfall volume is found by changing the value of the updraft velocity parameter (e_1). This is followed by changing the values of the rain generation time to satisfactorily simulate the onset point rain starts to fall. Finally, the response time parameters are calibrated to find an agreement between the observed and simulated peak rainfall intensities. Most of the parameters did not change significantly when calibration was performed independently for the four events. The peak intensities and the durations of all four events have satisfactorily been reproduced by the model. When using the calibrated parameter values of the July 22 event for the other three events, the model satisfactorily simulated the accumulated rainfall depth. The main limitation has been with the rain initiation time where the model produced rainfall too early for June 8 and where the model showed a delay for June 1 and August 16 events as compared to the observed hyetograph. The difference between the simulated and the observed rain initiation time is in the order of 30 minutes which, in hydrology, is considered small. Except for the June 1 event the start time of the events can be estimated with a maximum timing error of 5 minutes that is the difference in the simulated and observed rain initiation time. However, the procedure to arrive at the start time of rainfall through the updraft velocity should be further evaluated for other events.

Results are encouraging considering the many sources of uncertainty in the modelling that relate to the model assumptions, the movement of clouds that is not explicitly considered, the observations that

are only for the cloud top and the ground surface and the fact that observations for the cloud properties are absent along its profile.

It is noted that in the work by French et al. (1994), the timing of the simulated rainfall has been adjusted by assimilating radar observations that indicate the rain water content of the cloud layer. Such observations are absent for the present study area and a new procedure has been developed. The present study is a first attempt to simulate convective cloud systems by two layers of reservoir. It is concluded that results of this study are encouraging, particularly considering the (very) low model input data requirement. To be able to generalise on applicability of the cloud model, tests should be performed on a larger number of watersheds but also on a larger number of events. Some critical issue is that in this work the focus was on simulating convective cloud systems that are characterised by two rainfall peaks as commonly observed in the Gilgel Abbay catchment. The application of the cloud model to soil erosion and runoff studies that commonly require accurate estimations of event duration, accumulated rainfall depth and peak intensity should be tested.

APPENDIX: LIST OF SYMBOLS

Q_v	Saturation water vapour concentration	[M L ⁻³]
Q_c	Cloud water concentration	[M L ⁻³]
Q_R	Rain water concentration	[M L ⁻³]
Q_{vb}	Saturation water vapour concentration at a cloud bottom	[M L ⁻³]
Q_{vi}	Saturation water vapour concentration at interface of two layers	[M L ⁻³]
Q_{vt}	Saturation water vapour concentration at a cloud top	[M L ⁻³]
\vec{u}	Vector of wind velocities	[L T ⁻¹]
u, v, w	Wind velocities in x , y and z directions	[L T ⁻¹]
w_{\max}	Maximum updraft velocity	[L T ⁻¹]
t	Time instant	[T]
x, y	Distances along a horizontal direction of a Cartesian coordinate	[L]
z	Distance along a vertical direction of a Cartesian coordinate	[L]
z_b	Cloud bottom level	[L]
z_t	Cloud top level	[L]
W_T	Terminal velocity of rain water in still air	[L T ⁻¹]
ρ_a	Density of dry air	[M L ⁻³]
ρ_m	Density of moist air	[M L ⁻³]
w	Updraft velocity	[L T ⁻¹]
w_0	Vertically-averaged updraft velocity	[L T ⁻¹]
S	Vertically integrated liquid water amount in the cloud layer	[M]
S_1	Vertically integrated liquid water amount in the lower cloud layer	

		[M]
S_2	Vertically integrated liquid water amount in the upper cloud layer [M]	
S_b	Rain water concentration at a cloud bottom	[M L ⁻³]
S_{b1}	Rain water concentration at the bottom of the lower cloud layer	
		[M L ⁻³]
S_{b2}	Rain water concentration at the bottom of the upper cloud layer	
		[M L ⁻³]
I	Input mass rate to a cloud layer	[M T ⁻¹]
I_1	Input mass rate to the lower cloud layer	[M T ⁻¹]
I_2	Input mass rate to the upper cloud layer	[M T ⁻¹]
p	Pressure	[M L ⁻¹ T ⁻¹]
p_m	Pressure at a freezing level in a cloud	[M L ⁻¹ T ⁻¹]
p_0	Pressure at ground surface	[M L ⁻¹ T ⁻¹]
p_b	Pressure at cloud bottom	[M L ⁻¹ T ⁻¹]
p_i	Pressure at an interface of two layers of a cloud	[M L ⁻¹ T ⁻¹]
p'	Pressure for updraft velocity estimation	[M L ⁻¹ T ⁻¹]
p_t	Pressure at cloud top	[M L ⁻¹ T ⁻¹]
T	Temperature	[Θ]
T_0	Temperature at ground surface	[Θ]
T_{d0}	Dew point temperature at ground surface	[Θ]
T_b	Temperature at cloud bottom	[Θ]
T_i	Temperature at an interface of two layers of a cloud	[Θ]
T_m	Cloud temperature for updraft velocity estimation	[Θ]
T'	Ambient temperature for updraft velocity estimation	[Θ]

ω_s	Saturation mixing ratio	[M M ⁻¹]
$\Delta\omega$	Change in saturation mixing ratio in a cloud layer	[M M ⁻¹]
ω_{s0}	Saturation mixing ratio at ground surface dew point temperature	[M M ⁻¹]
ω_{si}	Saturation mixing ratio at the interface the two cloud layers	[M M ⁻¹]
ω_{st}	Saturation mixing ratio at a cloud top	[M M ⁻¹]
a	A constant parameter	[-]
b	A constant parameter	[Θ^2]
c	A constant parameter	[Θ]
e_{s0}	Saturation vapour pressure at Td_0	[M L ⁻¹ T ⁻¹]
\mathfrak{R}_v	The gas constant for water vapour	[M L ² T ⁻² M ⁻¹ Θ^{-1}]
e_1	Parameter for updraft velocity	[-]
c_p	Specific heat of dry air under constant pressure	[M L ² T ⁻² M ⁻¹ Θ^{-1}]
R	Rainfall rate	[M T ⁻¹]
t_r	Rain generation time	[T]
t_{event}	Correction for rain generation time	[T]
τ	Response time	[T]
τ_1	Response time of the lower cloud layer	[T]
τ_2	Response time of the upper cloud layer	[T]
α	A constant parameter	[-]
dp	Rain generation parameter	[M L ⁻¹ T ⁻¹]
dp_1	Rain generation parameter for the lower cloud layer	[M L ⁻¹ T ⁻¹]
dp_2	Rain generation parameter for the upper cloud layer	[M L ⁻¹ T ⁻¹]
i, j	Indices	[-]

Remote sensing based conceptual cloud modelling for rainfall simulation

$S_i(\theta)$	Simulated rainfall at i^{th} time step	[M T ⁻¹]
O_i	Observed rainfall at i^{th} time step	[M T ⁻¹]
R^2	Coefficient of determination	[-]

7

SENSITIVITY OF THE REPRESENTATIVE ELEMENTARY WATERSHED MODEL TO RAINFALL REPRESENTATION



ABSTRACT¹

The Representative Elementary Watershed (REW) model is evaluated for the Upper Gilgel Abbay watershed, Upper Blue Nile basin in Ethiopia. The watershed is characterised by mountainous terrain which receives large amounts of highly variable rainfall in space and time. The REW model has been calibrated and validated based on 6 years of daily stream flow and rainfall records from 10 non-recording rain gauges. A network of 8 recording rain gauges which were installed to supplement the existing network served as a reference network to evaluate the effect of hourly rainfall representation in REW modelling. By progressively decreasing the reference network density and by changing the configuration, 256 networks are generated. Network performance is evaluated in terms of the Goodness of Rainfall Estimation (GORE) index and 72 % of the configurations produced non-representative rainfall estimates. Results suggest that increasing network density does not necessarily improve the accuracy of the estimated rainfall, unless the network configuration is representative. The density and configuration of the selected rain gauge network is shown to affect flows simulated by the model. Modified Nash-Sutcliffe model efficiency (NSE_m) of 0.343 for 2 rain gauges and 0.943 for 6 rain gauges with optimal network configuration are shown. Rainfall input from a single rain gauge has caused extremely poor performance of the REW model. Understanding the spatio-temporal variability of the observed rainfall is found essential to decide whether a rainfall input should be represented by a uniform or non-uniform distribution. Results of this study also revealed that REW model performance is affected by the model resolution.

Key words: REW model, rainfall variability, Blue Nile, Gilgel Abbay

¹ This chapter is based on: Haile, A. T., Rientjes, T., Reggiani, P., 2009. **Model sensitivity to rainfall representation: the representative elementary watershed model**, *Water Resources Research*, Revised for re-submission.

7.1. INTRODUCTION

The accuracy of rainfall-runoff model outputs is affected by uncertainty in model parameters and meteorological forcing as well as deficiencies in model structure. Model calibration efforts (e.g. Vrugt et al., 2006; Wagener and Wheeler 2006; de Vos and Rientjes, 2008) commonly address parameter uncertainty while uncertainty by meteorological forcing received relatively little research attention. A number of studies showed that runoff hydrographs can be sensitive to rainfall representation (see, Dawdy and Bergman, 1969; Wilson et al., 1979; Beven and Hornberger, 1982) and consequently affect model parameter estimation (e.g. Chaubey et al., 1999; Bárdossy and Das, 2008; Younger et al., 2009).

Arnaud et al. (2002) discussed that the sensitivity of rainfall-runoff models to rainfall is very specific depending on the scale of the basin, the rainfall variability in the area and the mechanisms involved in runoff generation. A brief summary of how rainfall affects properties of simulated stream flow hydrographs is provided by Segond et al. (2007). In general, the response of a rainfall-runoff model to rainfall is affected by rainfall type (Bell and Moore, 2000a; Ajami et al., 2004; Segond et al. 2007), spatial location of a rainfall event with respect to a watershed outlet (Bras and Rodriguez-Iturbe, 1976; Woods and Sivapalan, 1999), watershed scale (Smith et al., 2004; Andréassian et al., 2001; Obled et al., 1994), geology (Naden 1992; Dodov and Foufoula-Georgiou, 2005) and soil type (Obled et al., 1994).

Other studies showed that runoff generation processes are affected by duration, coverage and spatial location of rainfall. Robinson and Sivapalan (1997); Menabde and Sivapalan (2001) showed that the stream flow of watersheds can be affected by the mean duration of rainfall if the watershed response time is less than the mean rainfall duration. Syed et al. (2003) found that the area coverage of a storm core, which is defined as an area with a 10-minute intensity of greater than 25 mm h⁻¹, was a better predictor of runoff volume and peak flow than the area coverage of the storm. In the same work, it was reported that the location of the storms became more important as the watershed size increased with storms located at the centre of the watershed producing more runoff than those located near the outlet or near the head of the watershed. Also, Morin et al. (2006) showed that simulated peak flows is affected by rainfall location and may

increase by a factor of 2 when the storm location is shifted by less than 4 km towards the watershed outlet.

Analyzing aspects of spatial representation of rainfall to runoff modelling has been a topic of ongoing research. Rainfall may be represented by uniform and non-uniform distribution while models may rely on lumped, semi-distributed and fully distributed model domain. For a semi-distributed model of 71 km² watershed, Obled et al. (1994) found that non-uniform rainfall representation did not improve simulation results. Smith et al. (2004) concluded that for a spatially more uniform rainfall observations, the use of a distributed model does not always lead to improved results as compared to the use of a lumped model. In the work of Brath et al. (2004), a distributed model of a 1050 km² watershed was not noticeably sensitive to rainfall representation through uniform or non-uniform distributions. In the work of Bell and Moore (2000a), the performance of a distributed model over a lumped model was only marginally better for stratiform rainfall but significantly better for convective rainfall. It is noted that stratiform rainfall is characterised by lower spatio-temporal variability as compared to a convective rainfall and suggests more uniformity in the observed rainfall.

In a number of studies, it is shown that the density of a rain gauge network affects how well rainfall can be represented and thus how network density may affect the stream flow response of a rainfall-runoff model. Obled et al. (1994) suggested that, when comparing a network of 5 or 21 gauges, 5 rain gauges are sufficient for a semi-distributed TOPMODEL of a 71 km² experimental watershed in Southeast France. Andréassian et al. (2001) suggested an optimally configured network of 8 rain gauges as a reference network for lumped GR3, TOPMODEL and IHAC rainfall-runoff model approaches of the 1120 km² Serein watershed in France. For the event based, semi-distributed RORB model of the rural watersheds of Lee basin of size 1400 km² in UK, Segond et al. (2007) suggested 16 rain gauges for a watershed of size 1000 km² and 7 to 8 rain gauges for watersheds of size 80 – 280 km². The authors pointed out that the suggested network density may not be appropriate for representing summer convective events.

Most of the studies described above are undertaken by strictly following network design criteria. However, results for data scarce regions in the tropics are barely presented. In addition, it is still not fully clear whether model sensitivity to rainfall significantly changes with model resolution. The main objectives in the present study are (i) to evaluate the

performance of the semi-distributed REW model approach to the Upper Gilgel Abbay watershed in the Blue Nile basin, (ii) to evaluate the sensitivity of the REW model to rain gauge density and configuration, (iii) to evaluate the effect of spatial model resolution on stream flow simulation, (iv) to assess model sensitivity to spatially uniform vs. non-uniform rainfall representations. For this study, the existing network of the Gilgel Abbay watershed is used. Network stations are located close to villages as constrained by site accessibility. Such is common in developing countries where the infrastructure is poor and where design criteria are largely ignored. In Haile et al. (2009a), it is shown that the Gilgel Abbay watershed has large spatio-temporal variability of rainfall at short inter-station distances. Results of the study indicated clear differences in the pattern and magnitude of the rainfall that was recorded in the uplands and in the lowlands of the watershed. The effect of the rainfall variability and its representation on runoff is, however, not known and is the motivation for the present study.

7.2. STUDY AREA

The watershed of the Upper Gilgel Abbay is situated between latitudes of 10°56'N – 11°22' N and longitudes of 36°49' E – 37°24' E. The watershed covers a surface area of about 1655 km² and is characterized by mountainous terrain and agricultural land that is dominated by clay to clay-loam soils. The annual rainfall in the years 2000 – 2005 ranged between 1600 – 2100 mm indicating a large variability between relatively dry and wet years.

The rainfall records that are used for the present study are obtained from 2 complementary networks as shown in figure 7.1. The base network has 10 non-recording rain gauges that provided daily rainfall data for 6 years. The reference network has 8 recording rain gauges that were installed in May 2007. For model calibration, the REW rainfall input was estimated using daily observations from the base network. The stations are Abbay Sheleko, Adet, Bahir Dar, Dangila, Gundil, Injibara, Kidamaja, Sekela, Wotet Abbay and Zege. The reference network provided the data for evaluating the effect of rainfall representation on runoff modelling. The network includes Addis Kidam, Bahir Dar, Dangila, Durbet, Injibara, Jema, Kega and Sekela. The meteorological data that served as input to the estimation of

Sensitivity of the Representative Elementary Watershed model to rainfall

potential evapotranspiration (PET) at daily base is obtained from Dangila, Adet and Bahir Dar stations which are operated by the National Meteorological Agency of Ethiopia.

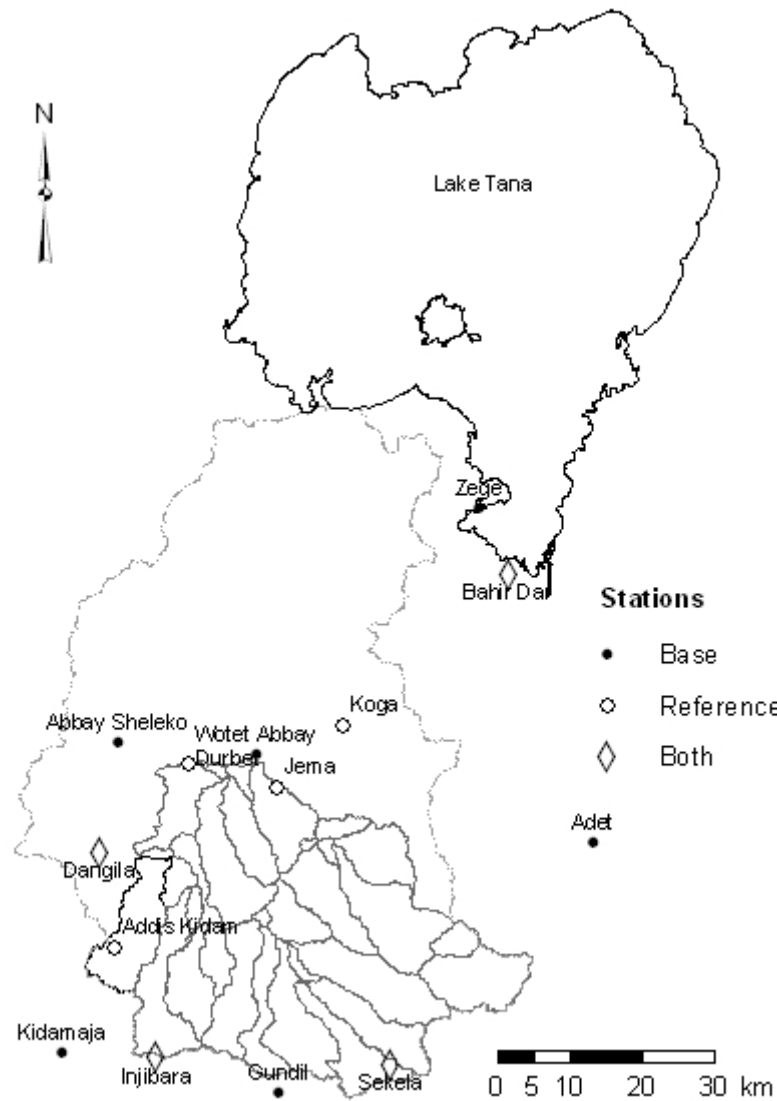


Figure 7.1: *The rain gauge network and the delineated representative elementary watersheds (REWs) of the Upper Gilgel Abbay watershed that drains to Lake Tana.*

7.3. METHODS

7.3.1. The REW model

In the REW approach, a watershed is divided into a series of irregularly shaped modeling entities that are the REWs. The subdivision is based on a topographic analysis following a selected Strahler order topology. The modeling volumes are called “representative” for the following reasons (Reggiani et al., 1998):

1. The REW is defined in such a way as to encompass all basic functional components of a watershed (channels, hill slopes) and constitutes a single functional unit, which is representative of other sub-entities of the entire watershed due to its repetitive character.
2. The REW is the smallest and therefore the most elementary unit into which the watershed is discretized for a given scale of interest.

Each REW is divided into six regions or simply flow zones based on the nature of the flow and the flow domain. The approach includes zones for simulating i) infiltration excess overland flow, ii) saturation excess overland flow, iii) channel flow, iv) unsaturated subsurface flow, v) saturated sub-surface flow (i.e. ground water flow), (vi) shallow subsurface flow to simulate interflow at shallow depth.

Interaction between neighboring REWs and between flow zones takes place in terms of mass and momentum exchanges across the boundary interfaces. The REWs are confined by the atmosphere on the land surface and by an impermeable layer at the bottom.

The governing equations expressed in terms of ordinary differential equations (ODE's) are complemented by a set of constitutive relationships that are at the core of the REW model formulation. The governing equations which are solved in the approach have the following general form:

$$\frac{d\psi}{dt} = \sum e_i^\psi + R + G \quad [7.1]$$

where: ψ refers to state variables such as soil water saturation and piezometric heads; e_i^ψ refers to exchange of mass or force among various

phases and sub regions; R refers to external supply of, for instance, mass while G refers to internal generation. For a complete description of the REW approach, the reader is referred to Reggiani et al. (1999); Reggiani and Rientjes (2005).

7.3.2. Model calibration

Model discretization based on a second Strahler order topology resulted in 33 REWs as shown in figure 7.1. Actual evapotranspiration is estimated based on the PET and the simulated soil moisture content of the REWs at each simulation time step. Spatially uniform PET and rainfall were specified as the model inputs.

After warming, the REW model of the Upper Gilgel Abbay watershed has been calibrated and validated for the time period covering 2000 – 2005. The model is calibrated against daily stream flows for the period 2002 – 2003 as recorded at Wotet Abbay that is at the outlet of the watershed. The performance of the model is validated for the periods 2000 – 2001 and 2004 – 2005.

The REW model has few parameters that require estimation. Rientjes et al. (2003); Zhang et al. (2005); Varado et al. (2006) showed that acceptable flow predictions can be attained by optimising porosity, hydraulic conductivity of the unsaturated zone, and hydraulic conductivity of the saturated zone. After a preliminary sensitivity analysis these parameters proved to be the most sensitive for the Upper Gilgel Abbay as well. In addition, the shallow subsurface interflow zone requires specifying the saturated conductivity of a shallow subsurface flow layer. As such, 4 parameters were manually calibrated first by changing the porosity and hydraulic conductivity of the saturated zone to simulate the base flow which is followed by changing the values of hydraulic conductivity of the unsaturated zone and the shallow subsurface interflow zone to simulate the high flows. A list of the remaining parameters and their values can be found in Reggiani and Rientjes (2005) that is adapted here. Similar to Reggiani and Rientjes (2005), all model parameters are assumed time and space invariant to exclusively focus on assessing the effects of rainfall representation on stream flow simulation.

Aim of the model calibration was to assess the capability of the model to reproduce the integrated response of the watershed in terms of

Sensitivity of the Representative Elementary Watershed model to rainfall

stream flow at the outlet of the watershed. It is noted that other state variables as simulated by the model were not validated since observations of these variables at the required time and space scales were not available in the Upper Gilgel Abbay.

The following objective functions are selected:

$$abias = \frac{\left| \sum_{j=1}^N (Q_j^s - Q_j^o) \right|}{\sum_{j=1}^N Q_j^o} \quad [7.2]$$

$$NSE = 1 - \frac{\sum_{j=1}^N (Q_j^s - \overline{Q_j^o})^2}{\sum_{j=1}^N (Q_j^o - \overline{Q_j^o})^2} \quad [7.3]$$

where: Q^s is the simulated flow for a specific parameter set, Q^o is the observed flow, and N is the total number of data elements. The over bar indicates time-averaged quantities, while j is a time index. The parameter *abias* is the absolute bias which measures the volumetric error while *NSE* is the Nash-Sutcliffe model efficiency that measures overall model performance. An *abias* value of zero indicates that there is no volumetric error while a value larger than zero indicates that there is under- or overestimation of the observed volume. $NSE = 1.0$ indicates a perfect match between the simulated and the observed flows while $NSE = 0.0$ indicates that model simulations are as accurate as the mean of the observed flows.

7.3.3. Rainfall representation

The sensitivity of the REW model to rainfall representation was assessed by using subsets of the reference network for different network configurations. Subset networks of k gauges from the reference network resulted in 256 configurations. Here, it is noted that the rainfall as estimated using subset networks is referred to as the estimated rainfall while the rainfall that is estimated using the reference network is referred to as the reference rainfall.

Sensitivity of the Representative Elementary Watershed model to rainfall

The relatively sparse network and the highly convective nature of rainfall in the study area restrict the use of interpolation methods such as Kriging since it is difficult to construct the semivariograms for convective events at hourly time interval. As such, the uniform rainfall at hourly time steps is interpolated using the Thiessen polygon method which is also applied by Segond et al. (2007):

$$\overline{R}_j = \sum_{i=1}^N a_i R_{ij} \quad [7.4]$$

where: \overline{R} is the spatially uniform rainfall estimate, a_i is the Thiessen weight for station i , R is the rainfall depth recorded at a specific station, j is a time index, and N is the total number of rain gauge stations. A critical note to the procedure is that some bias is introduced to the estimation since stations are not well distributed over the area as caused by site accessibility.

The spatial coefficient of variation (CV_s) is used to assess the spatial variability of the estimated rainfall with respect to the reference rainfall. The equation for CV_s is defined for each time step (j) using a relation that is given by Arnaud et al. (2002):

$$CV_s j = \frac{\sqrt{\sum_{i=1}^N a_i (R_{ij} - \overline{R}_j)^2}}{\overline{R}_j} \quad [7.5]$$

All terms are defined in the previous paragraph. Equation [7.5] applies for each time step but is averaged over the simulation period to arrive at a single CV_s value for the simulation period.

In this study, a temporal coefficient of variation (CV_t) is proposed to compare the temporal variation over the wet period of June – August (JJA) of the estimated rainfall and that of the reference rainfall. The index reads:

$$CV_t = \frac{CV}{CV^{ref}} \quad [7.6]$$

where: CV is the coefficient of variation of the estimated rainfall for a selected network density and configuration while CV^{ref} is the coefficient of variation of the reference rainfall. is defined as the standard deviation of

Sensitivity of the Representative Elementary Watershed model to rainfall

the JJA hourly rainfall divided by the mean of the hourly rainfall. $CVt=1$ indicates that both the estimated rainfall of a selected network and that of the reference network have equal temporal variability. $CVt > 1$ indicates that the estimated rainfall has a larger temporal variability than the reference rainfall while $CVt < 1$ indicates that the estimated rainfall has a smaller temporal variability than the reference rainfall.

Andréassian et al. (2001) proposed Goodness of Rainfall Estimation (*GORE*) and *Balance* indices. The *GORE* index evaluates the goodness of rainfall estimates while *Balance* indicates the volumetric error of the estimated rainfall.

The *GORE* index reads:

$$GORE = 1 - \frac{\sum_{j=1}^N \left(\sqrt{\overline{R_j}} - \sqrt{\overline{R_j^{ref}}} \right)^2}{\sum_{j=1}^N \left(\sqrt{\overline{R_j^{ref}}} - \sqrt{\overline{R_j^{ref}}} \right)^2} \quad [7.7]$$

where: R is the estimated rainfall, R^{ref} is the reference rainfall, the over bar indicates averaged values over time in this case JJA, j is a time index, and N is the total number of time steps. A *GORE* value of $-\infty$ indicates very poor performance of the configured network while a value of 1.0 indicates the estimated rainfall is as good as the reference rainfall.

The *Balance* index reads:

$$Balance = \frac{\sum_{j=1}^N R_j}{\sum_{j=1}^N R_j^{ref}} \quad [7.8]$$

All terms are as defined in the previous paragraph. A *Balance* value that is larger than 1.0 indicates that the estimated rainfall overestimates the reference rainfall while a value smaller than 1.0 indicates that the estimated rainfall underestimates the reference rainfall.

The root mean square error (*RMSE*) is used to evaluate the deviation of the estimated rainfall from the reference rainfall and reads:

$$RMSE = \sqrt{\frac{\sum_{j=1}^N (R_j - R_j^{ref})^2}{N}} \quad [7.9]$$

All the terms are as defined previously.

7.3.4. Sensitivity to network density and configuration

The model is run for the calibrated parameter set of the watershed to evaluate the REW model sensitivity to rainfall input. Therefore, in the following, stream flow as simulated using the rainfall from any of the sub-set networks is referred to as simulated flow while flow as simulated using the reference network is referred to as reference flow. To evaluate the differences in simulated flows by sub-set networks, a modified Nash-Sutcliffe Efficiency (NSE_m) coefficient that is similar to that by Segond et al. (2007) is applied which reads:

$$NSE_m = 1 - \frac{\sum_{j=1}^N (Q_j^s - Q_j^{ref})^2}{\sum_{j=1}^N (Q_j^{ref} - Q_j^{ref})^2} \quad [7.10]$$

where: Q^s is the simulated flow for a specific rainfall input from a subset of the reference network, Q^{ref} is the reference flow, j is an index of time step, and N is the total number of time steps.

The differences in the quantiles of the simulated and the reference flows of the same rank is also compared to evaluate if some flow ranges are under- or overestimated. Similar to Arnaud et al. (2002), the Hazen plotting position formula is applied to estimate the cumulative frequencies. The relative difference (E_q) between the quantiles of the same rank that indicates equal probabilities of exceedance can be computed as follows:

$$E_q = \frac{Q_q - Q_q^{ref}}{Q_q^{ref}} \quad [7.11]$$

where: Q_q is the q^{th} quantile of the simulated flow, Q_q^{ref} is the q^{th} quantile of the reference flow.

Assessments on the effect of rainfall representation on the simulated stream flows are through the use of spatially uniform and spatially non-uniform rainfall inputs. The analysis is performed for the model that is based on the second Strahler order discretization (33 REWs). The spatially non-uniform rainfall is estimated for the centroid of the REWs by the Inverse Distance Weighting (IDW) method of interpolation using distance weighting power of 2. For objective comparison, the volume of the uniform rainfall (UR) is constrained by conserving the volume of the non-uniform rainfall (NUR) at hourly time steps. By this conservation principle, the effects of only spatial rainfall variability can be studied (see, Younger et al., 2009).

The effect of rainfall interpolation is assessed for three interpolation methods. These are (i) the Thiessen polygon method, (ii) inverse distance method and (iii) nearest rain gauge to the REWs. The methods are applied to estimate the rainfall of each REW. In this study, the effect of model resolution is also assessed for discretizations that are based on a second and third Strahler order topology.

7.4. RESULTS

7.4.1. REW model calibration

The REW model is calibrated against daily stream flows that were observed for the period 2002 – 2003. After a preliminary sensitivity analysis, the optimised parameters are soil porosity ($\epsilon = 0.44$), saturated hydraulic conductivities of the unsaturated zone ($K_{su} = 10^{-5} \text{ m s}^{-1}$), the saturated zone ($K_{ss} = 10^{-6} \text{ m s}^{-1}$), and the shallow subsurface flow zone ($K_{si} = 10^{-4} \text{ m s}^{-1}$).

Figure 7.2 shows the observed and the simulated flows at the watershed outlet and the spatially averaged rainfall for the calibration period for days that range between 732 – 1431, for the first validation period for days that range between 1 – 731 and for the second validation period for days that range between 1432 – 2192. In general, the REW model satisfactorily reproduced the observed flows. Some small peaks are simulated prior to the wet season that are absent in the stream flow records although heavy rainfall events were reported in the rainfall records. In terms of peak flows, the model reproduced most of the peaks except for

Sensitivity of the Representative Elementary Watershed model to rainfall

some overestimation on days 1684, 2034 and 2035 for instance. The rainfall records showed that heavy events were observed on these days and it is speculate that the overestimation is caused by poor observation of the spatial distribution of high intensity rainfall by the rain gauge network.

Table 7.1 shows the values of the performance measures for the calibration and the two validation periods. The Nash-Sutcliffe efficiency for the calibration period is 0.75 with an *abias* of 6.65 %. The performance measures for the validation periods are a Nash-Sutcliffe efficiency of 0.60 – 0.64 and an *abias* of 5.60 – 9.05 % which are comparable to those reported for REW modelling results of other watersheds (see Zhang et al., 2005; Varado et al. 2006).

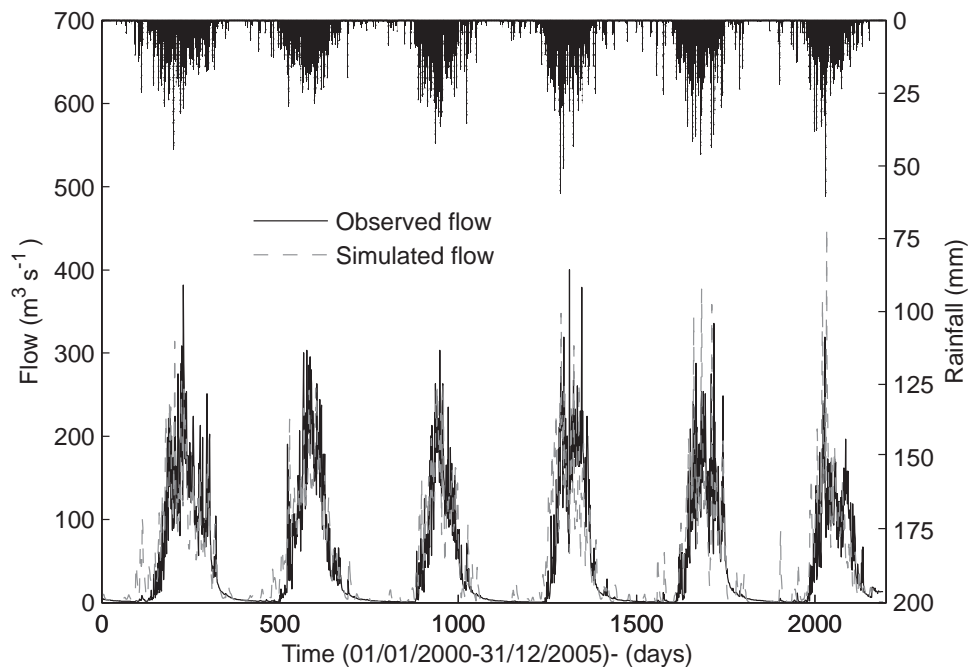


Figure 7.2: Observed and simulated flow at the watershed outlet. The watershed average rainfall is shown at the top. Note: the calibration period range between 732 –1431, the first evaluation period ranges between 1 – 731 and the second evaluation period ranges between 1432 – 2192.

Sensitivity of the Representative Elementary Watershed model to rainfall

Table 7.1: *Model efficiency for the calibration and validation data sets.*

Data set	Period	Nash-Sutcliffe	
		efficiency	<i>abias</i> (%)
Calibration	2002-2003	0.75	6.65
Evaluation 1	2000-2001	0.60	5.60
Evaluation 2	2004-2005	0.64	9.05

7.4.2. Effects of the gauge network configuration and density on rainfall and runoff estimation

First, the effect of network density on the estimated rainfall is evaluated. Rainfall estimates are obtained for fixed densities of subsets of the reference network by changing the network configuration. Figure 7.3 shows the differences in the estimated rainfall by selected indices.

Figure 7.3a shows the *GORE* index which indicates that the goodness of the rainfall estimation as compared to the reference rainfall decreases with decreasing network density. Maximum value of the *GORE* can reach up to 1.0 which indicates perfect performance for network sizes larger than 6 rain gauges. The maximum values for networks of 1 or 2 rain gauges are in the order of 0.3 and 0.7, respectively. The maximum value of the *GORE* index indicates the best configuration. A minimum value of -0.98 was obtained for network configurations of less than 3 rain gauges indicating very poor performance. The result suggests that the overall performance of rainfall estimation using any subset of the reference network can range between extremely poor and perfect with respect to the reference rainfall. For instance, if rainfall estimates with *GORE* > 0.8 are considered to be representative, i.e. with acceptable accuracy, then about 72 % of all the possible ensemble networks produced non-representative rainfall estimates.

Figure 7.3b shows that decreasing the network density leads to a large deviation of the *Balance* index with reference to the optimal value of 1.0. The middle-fifty percent of the *Balance* index (the 75 % quartile minus the 25 % quartile) increases with a decrease in network density indicating that variability increases when the density of networks decreases. Overall, any two configurations of networks with equal but few rain gauges can have large difference in terms of the estimated rainfall. For instance, one of the configurations can overestimate while the other configuration can

Sensitivity of the Representative Elementary Watershed model to rainfall

underestimate the reference rainfall. The result suggests that careful selection of network configuration is more important for low density networks than it is for high density networks. This analysis is undertaken for the ensembles of sub-set networks and thus the site specific characteristics of the stations are not considered.

Figure 7.3c shows that the *RMSE* increased with a decrease in network density. The *RMSE* varies between 0.0 mm h⁻¹ for network size of 7 rain gauges and 2.3 mm h⁻¹ for network size of 1 rain gauge. This shows that a decrease in network density leads to a large deviation between the estimated and the reference rainfall. Figure 7.3d shows that the temporal coefficient of variation (*CVt*) increases rapidly with a decrease in network density. For all cases, the *CVt* values are larger than 1.0 indicating that the rainfall estimates have larger temporal variability than the reference rainfall.

Values of the selected performance measures suggest that the accuracy of the estimated rainfall deteriorates with a decrease in network density. A reason for this could be the convective nature of the rainfall. Haile et al. (2009a) showed that localized events are very frequent and that rainfall in the watershed is highly variable with maximum rainfall intensity at one station while the remaining stations recorded very low to zero rainfall intensity.

A multivariate regression analysis is performed to relate the *GORE* index as estimated for networks of only 4 specified stations which are considered categorical variables. The regression equation, which is fitted with an *R*² value of 0.87 and a standard error of 0.07, reads:

$$GORE_{4s} = 1.075 - 0.126A - 0.238B - 0.181D - 0.196DU - 0.050I - 0.044J - 0.212K + 0.158S \quad [7.12]$$

where: The subscript 4s shows that only 4 gauges are considered; *A* is for Addis Kidam, *B* is for Bahir Dar, *D* is for Dangila, *DU* is for Durbet; *I* is for injibara; *J* is for Jemma; *K* is for Koga and *S* is for Sekela. In equation [7.12], *A* = 0 when Addis Kidam is 'not in network' and *A* = 1 when Addis Kidam is 'in network'. This applies for the remaining stations as well. As such, the *GORE* index can be estimated for any other combination of 4 'in-network' stations.

Sensitivity of the Representative Elementary Watershed model to rainfall

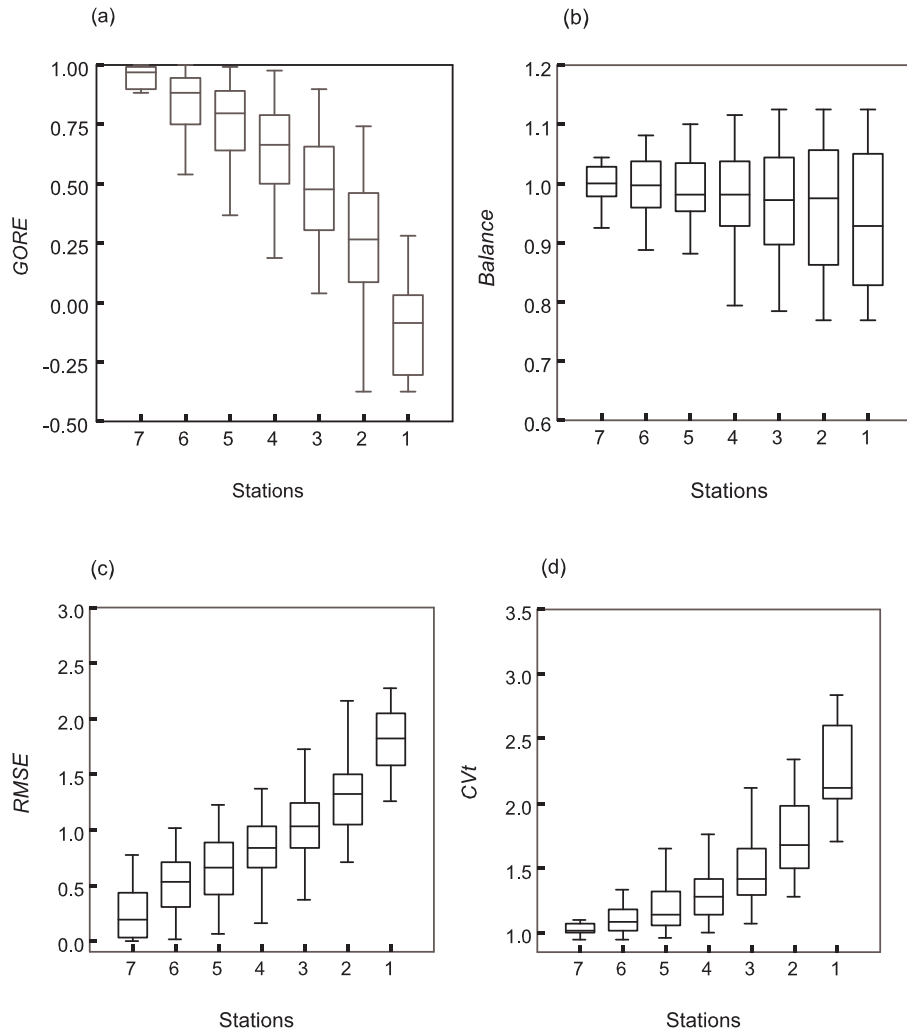


Figure 7.3: Performance of spatial rainfall estimation using specific number of rain gauge stations with respect to the reference area-average rainfall. The lower and the upper bars indicate the 25 % quartile minus 1.5 IQR and the 75 % quartile plus 1.5 IQR, respectively where IQR is the interquartile range which is the height of the box. The middle bar of the boxes shows the median while the upper and lower bars of the boxes show the 75 % and the 25 % quartiles.

In the previous paragraphs, the variation of the estimated rainfall which is estimated using different network densities is discussed without considering the intensities of the rainfall events. To differentiate between

Sensitivity of the Representative Elementary Watershed model to rainfall

light to medium and high intensity events, an arbitrarily selected threshold of 5 mm h^{-1} is applied for which the performance of the network for rainfall estimation is assessed. The records of the reference rainfall contain 23 events with intensities higher than the selected threshold.

In table 7.2, the statistics of the *GORE* index for all possible network configurations are shown for events with intensities higher than 5 mm h^{-1} . The same table shows that the value of the minimum and 25 % quartile of the *GORE* index for all of the selected network densities are negative. This suggests that certain configurations of the network may poorly estimate the spatial rainfall. Table 7.2 also shows that in terms of the 75 % quartile, the configurations of network densities of smaller than 5 rain gauges resulted in a negative value of *GORE* index. In terms of the median, only a network size of 7 rain gauges resulted in a satisfactorily estimates while results of the rainfall estimation deteriorated when smaller than 7 rain gauges are used. In terms of the maximum values, the *GORE* index is larger than 0.7 when more than 3 gauges of the best performing network configurations are used to estimate the spatial rainfall. This suggests that the best performing configurations of networks of 4 rain gauges are adequate to satisfactorily estimate rainfall events of intensity higher than 5 mm h^{-1} . Overall, most of the configurations of the networks of specific densities are not adequate to accurately estimate rainfall of such intensities.

Table 7.2: *Statistics of the GORE index for estimation of the spatial rainfall for events with intensity of higher than 5 mm h^{-1} .*

Statistics	Number of stations						
	7	6	5	4	3	2	1
Minimum	-8.5	-12.9	-17.6	-21.4	-24.9	-28.1	-32.1
25 % quartile	-0.8	-0.4	-9.3	-12.0	-14.8	-19.4	-26.4
Median	0.9	-0.4	-3.7	-7.7	-11.9	-15.2	-19.8
75 % quartile	1.0	0.8	-0.1	-3.4	-6.5	-12.4	-18.7
Maximum	1.0	1.0	0.9	0.8	0.6	-1.9	-14.5

7.4.3. REW model sensitivity to network configuration and density

The sensitivity of the REW model for specific configurations of the rain gauge network is shown in table 7.3 with network performance for selected indices. For each network density, the configuration that produced

Sensitivity of the Representative Elementary Watershed model to rainfall

the maximum *GORE* index is selected since this shows the best configuration for representing rainfall with reference to the entire network.

Table 7.3 shows that the rain gauges at Sekela and Jema are part of the best performing configurations of any of the network densities. Sekela is located high on a mountain which is the head source of Gilgel Abbay river while Jema is located at relatively low elevation and east of the main stream of Gilgel Abbay river. Haile et al. (2009a) showed that the rainfall at Sekela is light to medium while that at Jema is heavy with intensities as high as 39 mm h⁻¹.

Table 7.3 shows that similar to the value of the *GORE* index, the *NSE_m* for stream flow simulation decreases with decreasing network density. The coefficient of determination *R*² between the *NSE_m* and the *GORE* index is 98 % and indicates that *NSE_m* is strongly related to the *GORE* index. The *NSE_m* changed from 0.943 when 6 rain gauges were used to 0.344 when 2 rain gauges were used while extremely poor performance is indicated when a single rain gauge was used.

Table 7.3: Sensitivity for spatially averaged rainfall that is estimated by a specific number of stations. Note: the stations with maximum *GORE* were selected.

	Number of stations					
	6	5	4	3	2	1
Rain gauges for best configuration	Dur, Sek, A. Ki, Jema, Dan, Inji	Dur, Sek, A.Kim, Jema, Inji	Sek, A. Ki, Jema, Inji	Sek, Jema, Inji	Sek, Jema	Sekela
Indices for estimated rainfall						
<i>GORE</i> index	1.000	0.992	0.977	0.899	0.740	0.284
<i>Balance</i>	1.002	0.994	1.017	1.073	1.063	0.996
<i>CVt</i>	1.004	1.019	1.080	1.165	1.407	1.699
Performance measures for simulated flow						
<i>NSE_m</i>	0.943	0.813	0.810	0.597	0.344	-0.365
bias (%)	2.100	-2.468	-1.109	5.481	-1.140	6.771

The *bias* is estimated using equation [7.2] by ignoring the absolute value. It must be noted that the *bias* values of the simulated flow are not uniquely related to the errors in the rainfall input since underestimation of

Sensitivity of the Representative Elementary Watershed model to rainfall

the rainfall does not necessarily cause an underestimation of the simulated flows with respect to the reference flow.

Figure 7.4 shows the relative difference between the simulated flow quantiles and the reference flow quantiles of the same rank (see, equation [7.11]). The simulated flows are based on rainfall input which is estimated using the best performing configuration of networks with 1, 3 and 5 rain gauges, see table 7.3 for selected stations. The relative difference is largest for high and low flows. The relative difference for high flows is up to 40 % for a configuration of 3 rain gauges while the difference for high as well as for low flows is up to 15 % for 5 rain gauges. The relative difference for 3 and 5 stations is smaller than 5 % for middle ranged flows indicating that rainfall representation is particularly poor for these flows. Rainfall input as estimated from only the mountain rain gauge Sekela resulted in extremely large relative differences of up to 100 % for almost all ranges of the simulated flow.

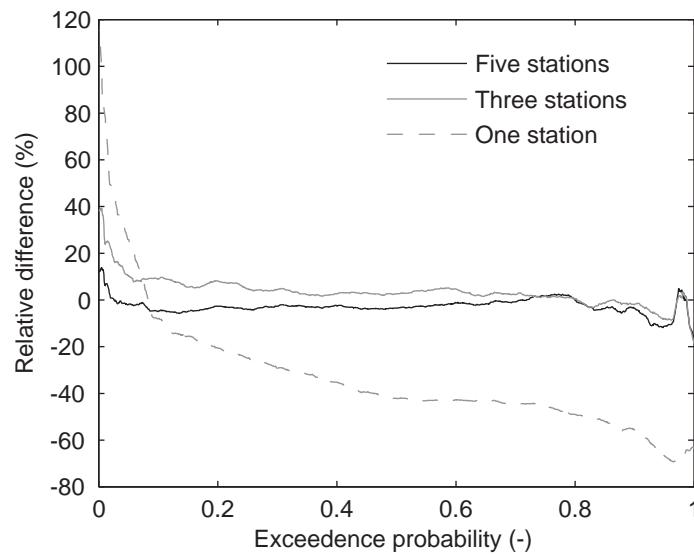


Figure 7.4: Relative difference (see equation 12) between the simulated stream flow and the reference stream flow for a specific number of rain gauge stations. Note: The rain gauge configuration for respective rain gauge numbers that produced the maximum GORE index is selected.

The effect of the configuration of the rain gauge network on the accuracy of the simulated flows is evaluated. Rainfall is estimated using data from 7 rain gauges by removing 1 rain gauge at a time from the reference network of 8. Removing Bahir Dar station did not affect the

Sensitivity of the Representative Elementary Watershed model to rainfall

performance of the network in estimating rainfall, presumably, since the station is located somewhat far from the watershed. As such this station was not considered for evaluating the sensitivity of the model and the REW model was run for the 7 subsets of the reference network.

Table 7.4: Effect of removing a specific station on the performances of the REW model: The name of the removed station and its elevation are shown on the top of each column.

	A. Ki	Inji	Sek	Dan	Jema	Dur	Koga
Elev.(m)	2370	2592	2715	2127	1970	1984	2011
Indices for estimated rainfall							
<i>GORE</i>	0.945	0.916	0.712	0.993	0.885	0.990	1.000
<i>Balance</i>	1.038	0.965	1.043	0.993	0.924	1.017	1.002
<i>CVt</i>	1.020	1.099	1.299	1.016	0.950	1.031	1.004
<i>CVs</i>	1.114	1.091	1.091	1.113	1.141	1.116	1.136
Performance measures for simulated flow							
<i>NSE_m</i>	0.825	0.641	0.504	0.780	0.853	0.928	0.816
<i>bias (%)</i>	-6.244	-14.656	-7.441	-3.201	-8.013	-3.758	-6.433

Table 7.4 shows that the *GORE* index of the rainfall and the *NSE_m* of the flow are related where higher *GORE* values result in higher *NSE_m* values. The *NSE_m* ranged between 0.504 and 0.928 when the mountain rain gauge Sekela and the lowland rain gauge Durbet were removed, respectively. In terms of *bias*, removing any of the rain gauges causes underestimation of the flow. The *bias* is largest when the mountain station Injibara was removed. In Haile et al., 2009a, it is shown that the mountains of Gilgel Abbay watershed have large seasonal rainfall and are dominated by light to medium but frequent rainfall events when compared to the rainfall in the lowlands which are relatively heavy but infrequent. As such, removing a mountain station may result in a large rainfall error which consequently may cause underestimation of simulated flows.

Figure 7.5 shows the sensitivity of the model when Sekela and Durbet stations are removed from the reference network. Figure 7.5a shows that the difference is very small when Durbet is removed but that the difference becomes large when Sekela is removed. Figure 7.5b shows that removing Sekela results in a systematic underestimation of low flows at the beginning of the wet season (e.g. on hours 500, 750, 1000 since start of simulation) while it also results in underestimation of the peaks in the beginning of the wet season (i.e. until simulation hour 850) and

overestimation of some of the peak flows towards the end of the wet season (i.e. on hours 1853 and 2032).

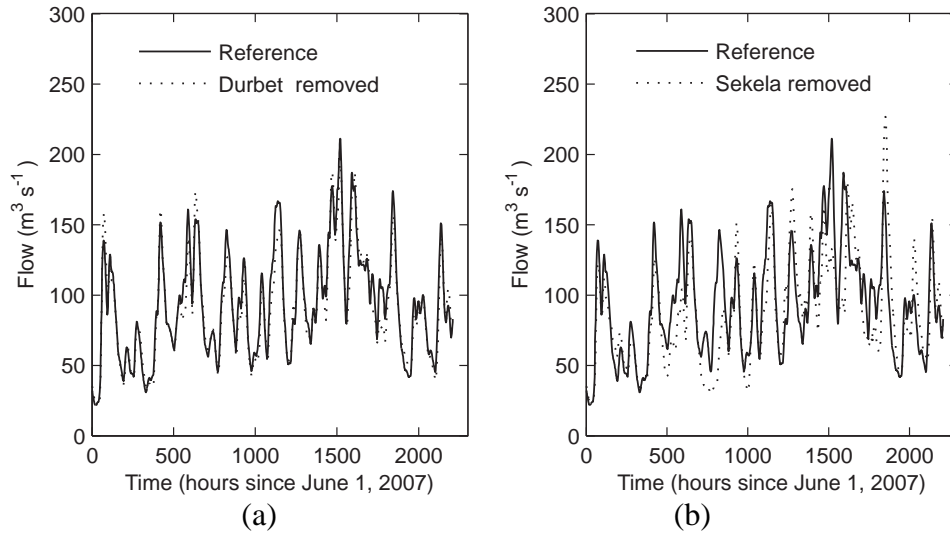


Figure 7.5: Effect of removing Durbet station (a) and Sekela station (b) on the simulated flow as compared to the reference flow. Durbet is located in a valley while Sekela is located on a mountain.

7.4.4. Effect of model resolution

The effect of model resolution on the sensitivity of the REW model to rainfall is evaluated by applying a second and a third Strahler order discretization that resulted in 33 REWs and 9 REWs, respectively. The comparison is made for the time period that spans from June 1, 2007 to June 9, 2007 at the beginning of the wet season while NSE_m , *bias* and peak flow indices are used.

Table 7.5 reveals that model performance for the third order discretization is less sensitive to rain gauge density as compared to the second order discretization since the former resulted in higher values of NSE_m with relatively smaller range of values. In terms of *bias*, there is no clear pattern in the sensitivity by the two model resolutions. The differences in the simulated peak flows are generally much smaller for the third order than the second order discretization which again suggests that the coarser resolution model is less sensitive to rainfall.

7.4.5. Model sensitivity to rainfall variability

Stream flow volumes and the peak flows as simulated for uniform rainfall (UR) and non-uniform rainfall (NUR) representations are compared. For the comparison, 11 events in JJA, 2007 are selected and performance is evaluated through the NSE_m coefficient, see table 7. 5. For the events, NSE_m ranged between 0.66 and 0.99 which shows that the REW model response to rainfall representations largely varies subject to the rain events. The *bias* ranged between -13.8 % and 18.7 % while the difference in the simulated peak flow ranged between -16.9 % and 25.0 %. This again suggests that the sensitivity of the model to rainfall representation varies from event to event. As such, the spatial coefficient of variation (CVs), the temporal coefficient of variation (CVt) and the antecedent rainfall are estimated for the selected events to reveal if these can be related to the model sensitivity to rainfall representation.

Table 7.5: Comparison of sensitivity to rainfall of the REW model using a second and third Strahler order descritization for the time period June 1 to June 9, 2007.

	Six stations	Five stations	Four station	Three stations
Second Strahler order (33 REWs)				
NSE_m	0.990	0.962	0.877	0.858
<i>bias</i>	0.119	-1.073	6.383	6.103
Error in peak flow (%)	3.5	13.2	22.2	21.4
Third Strahler order (9 REWs)				
NSE_m	0.994	0.995	0.988	0.986
<i>bias</i>	2.3	-2.7	-7.7	-4.6
Error in peak flow (%)	9.5	4.0	1.6	7.1

Table 7.6 shows the correlation between these rainfall indices and the performance measures as estimated for stream flow using UR and NUR representation. The correlation values are estimated using the data for 11 events in JJA of 2007. The CVs is estimated for the event rainfall while the antecedent rainfall is estimated as the accumulated rainfall that is observed during a 5 days period preceding the specific events.

Sensitivity of the Representative Elementary Watershed model to rainfall

Table 7.6: Correlation matrix for rainfall characteristics and model performance measures for simulated stream flow with a uniform rainfall input with the non-uniform rainfall input for a reference flow.

Rainfall characteristics	Performance measured		Difference in peak flow (%)
	NSE_m (-)	<i>bias</i> (%)	
CVs (-)	-0.533	-0.3	-0.1
CVt (-)	-0.318	-0.1	0.7
Antecedent rainfall (mm)	0.299	-0.5	-0.4

The correlation value in table 7.6 shows that there is an inverse relation between CVs and NSE_m which indicates that a larger value of CVs results in a smaller value of NSE_m . Therefore, the simulated stream flow for the UR representation becomes significantly different from that is simulated for the NUR representation when the rainfall largely varies in space, i.e. for high CVs values. In terms of correlation, rainfall temporal variability (CVt) and antecedent rainfall have moderate effect on the NSE_m .

The correlation values indicate that the *bias* is moderately related to CVs and the antecedent rainfall while it is weakly related to CVt . The result indicates that a smaller CVs value results in a larger *bias* value. For small antecedent rainfall, UR resulted in a larger volume in the simulated stream flows compared to NUR. Such suggests that differences in model output by UR and NUR become larger for dry antecedent conditions as compared to wet conditions.

The effect of rainfall characteristics on the difference in peak flow as a result of UR as compared to NUR is also analyzed. Correlation values in table 7.6 are estimated using the absolute value of the difference in peak flow. The correlation values indicate that the difference in the peak flow by UR and NUR is significantly affected by CVt and moderately affected by the antecedent rainfall while it is weakly affected by the CVs . This suggests that the difference in peak flow as a result of the spatial rainfall representations becomes larger for rainfall with larger temporal variability.

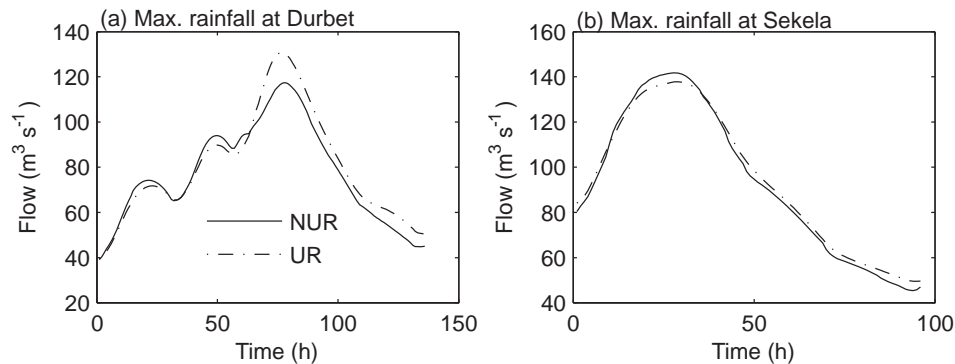


Figure 7.6: Simulated flow for (a) rainfall with maximum depth near the watershed outlet (at Durbet) and (b) on the mountain (at Sekela). Note: The time is measured since the start of simulation of the event.

In this study, the effect of the location of the maximum rainfall depth on the response of the runoff model is evaluated. Figure 7.6a shows the simulated stream flows using UR and NUR representations for maximum rainfall depth that is observed at Durbet which is located in a lowland area near the watershed outlet. The figure shows that the UR resulted in a peak flow which is 11.4 % higher than that by the NUR. Such presumably can be explained since the maximum rainfall depth that falls near the watershed outlet has a short flow path that is ignored by specifying a UR. Figure 7.6b shows the simulated flow for the event for which the maximum rainfall was recorded by the mountain gauges at Sekela. The figure shows that the UR and the NUR resulted in somewhat similar hydrographs with a peak flow of UR that is 2.8 % lower than by the NUR. Overall, the results show the effect of maximum rainfall location in the watershed in assessing effects of rainfall representation.

7.4.6. Effects of rainfall interpolation

In this study, 3 interpolation methods are compared which are the Thiessen polygon, the inverse distance Weighting (DW) methods and a nearest neighbour approach where rainfall of the nearest gauge is assigned to the REWs. The comparison is performed for the third order discretization. Results for 3 events are summarized in table 7.7 in terms of the volume and the peak of the simulated stream flow.

Table 7.7 shows that the IDW method results in smallest volume error of simulated stream flows. In terms of percentage, the volume of the

Sensitivity of the Representative Elementary Watershed model to rainfall

simulated flow using the IDW method is 12.8 %, 1.1 % and 0.6 % smaller than the Thiessen method. The comparison of the peak flows also shows similar results where the IDW method resulted in lower peaks than the other methods. In terms of percentage, the volume of the simulated peak flow using the IDW method is 13.0 %, 11.4 % and 3.3 % smaller than the simulated peak by the Thiessen method. Overall, the use of the different interpolation methods resulted in larger percentage differences in the peak than in the volume of the simulated flow. Such large difference in the characteristics of the stream flows suggests that the selected interpolation method has large effect on the result of model simulations and may contribute to uncertainty of model results.

Table 7.7: *Effects of interpolation on REW model outputs. Note: the volume of the simulated stream flow is divided by the size of the watershed area.*

		Interpolation method	Event 1	Event 2	Event 3
Volume of simulated stream flow (mm)		IDW	18.3	14.7	30.9
		Thiessen	20.9	14.9	31.1
		Nearest gauge	22.2	14.1	33.5
Peak stream flow ($\text{m}^3 \text{s}^{-1}$)		IDW	96.0	109.9	189.0
		Thiessen	110.4	124.1	195.5
		Nearest gauge	107.3	115.2	220.2

7.5. DISCUSSION AND CONCLUSION

In this study, the REW model is evaluated for the 1655 km² watershed of the Upper Gilgel Abbay with specific focus on evaluating effects rainfall representation. All model parameters are assumed spatially uniform and time invariant to exclusively focus on rainfall representation. Previous studies that applied the REW model also assumed uniformity of parameters, e.g. Reggiani and Rientjes (2005); Zhang et al. (2005); Varado et al. (2006). and focussed on assessing the models ability to simulate the integrated watershed response in terms of stream flow.

The base network of 10 rain gauges provided the rainfall data for model calibration. The REW model is calibrated for daily stream flow observations with a Nash-Sutcliffe coefficient of 0.75 and an absolute bias

Sensitivity of the Representative Elementary Watershed model to rainfall

(*abias*) of 6.6 %. Such values indicate good model performance and did not change much when the model was applied to validation periods.

Effect of rain gauge network on rainfall estimation: The variation of the estimated spatially uniform rainfall with a change in the density and configuration of a rain gauge network is evaluated. In the work of Haile et al. (2009a), the observations by the reference network revealed the presence of large rainfall variability in the study area. Stations, however, are installed close to small towns which may result in some bias in the analysis since stations are very weakly distributed. The uneven distribution leads, for instance, to unknown extrapolation errors of rainfall to some parts of the watershed that as such may affect UR and NU representation. To be able to quantify on these effects, rain gauges should also be available on the central and northeastern parts of the watershed as well.

In the present study, the Thiessen polygon method is applied to estimate spatially uniform rainfall for each of the sub-set networks. The interpolation method is simple and does not require optimizations of any parameters. To analyse how network density and configuration affects rainfall estimation, 256 sub-networks are established from a network of 8 recording rain gauges. Based on the GORE value of lower than 0.8, about 72 % of the generated networks produced non-representative rainfall estimates. The values of both *RMSE* and temporal coefficient of variation (CV_t) rapidly decrease with a decrease in network density. Although the GORE index for denser networks is mostly high (i.e. up to 1.0) it is shown that some configurations of denser networks can result in a low value. Such indicates poor representation and the analysis indicates that is particular for heavy rainfall events. The result suggests that an increase in network density does not necessarily improve the accuracy of the estimated rainfall unless the network configuration is representative.

Model sensitivity to rain gauge network: The REW model of the Upper Gilgel Abay is found to be sensitive to both the density as well as to the configuration in terms of location of the rain gauges. The relative difference between the simulated high and low stream flows of the same rank as explained by the exceedance probability becomes as large as 15 – 40 % when 3 and 5 rain gauges are used while it becomes as large as 100 % when 1 rain gauge is used. The use of 1 rain gauge resulted in larger differences over the whole range of the simulated flows.

Sensitivity of the Representative Elementary Watershed model to rainfall

It is shown that model sensitivity to rainfall depends on whether a lowland or a mountain rain gauge is considered. Removing the lowland rain gauge Durbet from the reference network resulted in a NSE_m value of 0.93 while removing the mountain rain gauge Sekela resulted in a NSE_m value of 0.50. Removing Sekela station caused underestimation of the simulated low flows at the beginning of the wet season while it caused overestimation of some of the peak flows during the wet season.

For a satisfactory performance of the REW model, the rain gauge at Sekela located at the head source of the watershed and the rain gauge located at Jema that is east of the main river at low elevation should always be part of any selected rain gauge network. These stations become important since: (i) there are no other stations east of Jema and north of Sekela and therefore removing the two stations affects the interpolation weights to the stations that are south of the main river of the watershed, (ii) Sekela and Jema have different rainfall properties as compared to the remaining stations in the study area.

Effects by (i) can be evaluated by installing additional rain gauges on the areas north of the main river of the watershed. With regard to (ii), Haile et al. (2009a) showed that both Sekela and Jema have large seasonal rainfall but Jema receives more frequent and high intensity rain events while Sekela receives clustered hourly rainfall with mostly light to moderate rainfall intensities. Haile et al. (2009c) showed that rainfall in the lowland areas of Gilgel Abbay can be characterized by large depth, high intensity, short duration and long inter-event time as compared to the mountain areas. As a result of the differences in rainfall properties, removing Jema and Sekela from the network can result in larger differences in the estimated spatial rainfall as compared to removing any of the other stations. In the present study, such differences are found to significantly affect the simulations of the REW model.

Effect of model resolution: The sensitivity of the REW model to spatial rainfall representation is analyzed by using model discretization scheme that is based on second and third Strahler order topology which resulted in 33 and 9 REWs, respectively. The second order discretization showed a higher sensitivity to rainfall representation suggesting as compared to the third order discretization. Zhang et al. (2005) stated that higher order discretization results in a lower number of REWs but a larger subsurface volume for each REW which causes delayed responses from the subsurface

Sensitivity of the Representative Elementary Watershed model to rainfall

storage while peak flow at the outlet are suppressed. This was also observed by comparison of results of third and second order discretization with the latter resulting in larger overestimation of peak flows. Based on the results, it may be concluded that the increase in runoff dynamics by the finer resolution model may increase the sensitivity of the model to the errors in the rainfall input.

Effect of rainfall variability: Arnaud et al. (2009) showed for fictitious basins that UR results in a significantly different runoff volume and peak flow as compared to a NUR representation. A similar conclusion can be drawn from the results of the present study. It is shown that the response of the REW model to rainfall varies subject to the rain events. In terms of NSE_m , the model becomes more sensitive to rainfall representation when the rainfall is more variable in space. The difference in peak flow for UR and NUR becomes larger when rainfall is characterized by larger variability over time and for small volumes of antecedent rainfall. For a maximum rainfall depth near the watershed outlet, the UR resulted in a peak flow which is 11.4 % higher than the peak flow by NUR. However, for a maximum rainfall depth on the mountain, the UR resulted in a peak flow which is 2.8 % lower than by the NUR. It is noted that the observation time series in this study may not cover the entire ranges of rainfall variability in the watershed. Such requires increasing the density of the rain gauge network.

The effect of rainfall interpolation on simulated stream flows is evaluated in this study. The IDW method resulted in a smaller stream flow volume as well as peak flow than by the Thiessen polygon method and the nearest neighbour approach where the rainfall of nearest rain gauge is assigned to the REWs. Based on the results of the analysis, it is concluded that the use of different interpolation methods results in relatively larger difference in the peak flow than the stream flow volume.

Sensitivity of the Representative Elementary Watershed model to rainfall

8

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS



8.1. SUMMARY AND CONCLUSIONS

The analysis of ground based observations in the Lake Tana basin shows that the rainfall is largely variable and is affected by orographic factors and the presence of Lake Tana. The inter-station distance between the rain gauges is larger than the scales of the rainfall and as such the network does not satisfactorily capture the rainfall variability that has a strong effect on the hydrology of the basin. The spatial variability of the rainfall is so large that a few rain gauges are not sufficient to simulate discharges adequately. Simulated discharges are found to be affected by not only network density but also network configuration.

A network of 8 recording rain gauges is used to capture the rainfall variability of the basin in space and time. The temporal analysis shows a diurnal pattern that is location specific suggesting different mechanisms of rainfall generation in the lowlands, around the lake and in the uplands. The network of the recording rain gauges does not cover the entire basin. Remote sensing images can help as an additional source of data, but not directly, since the remote sensing signals come from cloud top surfaces not from the surface rainfall. Conclusions that are drawn from the work in this thesis and recommendations for future work are presented in this Chapter.

8.1.1. Rainfall variability

The spatial pattern of the rainfall variability in the Lake Tana basin is evaluated using a set of statistical methods applied to a network of 8 recording rain gauges and supporting satellite data. The rain gauges recorded the rainfall of the full wet season from June – August of 2007. For the areas that are not covered by the rain gauge network, the diurnal cycle of the rainfall was inferred using cloud-top observation at the 10.8 μm channel of the Meteosat Second Generation (MSG-2) geostationary satellite, see Chapter 3.

SPATIAL PATTERN OF HOURLY RAINFALL

Statistics of the rainfall observations suggest that the upland areas of Gilgel Abbay watershed in the Lake Tana basin receive rainfall more

Summary, conclusions and recommendations

frequently as compared to the lowland areas. For hourly rainfall lower than 10 mm, a direct relation has been found between the number of rainfall occurrences and terrain elevation. However, for hourly rainfall higher than 10 mm, an inverse relation has been found between the number of rainfall occurrences and terrain elevation. As such, it can be concluded that relatively smaller hourly rainfall mostly occurs on mountainous areas of Gilgel Abbay while relatively larger hourly rainfall mostly occurs on the lowland areas of the watershed.

The spatial structure of the rainfall is studied through the Pearson's product-moment correlation that is first estimated for the rainfall observations at a pair of stations. In the procedure, this is followed by fitting a correlation model, which is in the form of an exponential equation, to the correlation values and the inter-station distances. To summarize the results, the correlation distance of the hourly and the daily rainfall is 8 – 10 km and 18 km, respectively which indicates that stations that are separated by distances that are larger than the correlation distance have statistically insignificant relation. As such, the inter-station distance in the watershed should not exceed 8 – 10 km or 18 km if the spatial structure of the hourly and daily rainfall needs to be satisfactorily observed by the network. In some cases, the ± 1 hour lag cross-correlation has been found higher than the 0 hour lag correlation which can be caused by the movement of convective clouds. If two consecutive hours are considered, the rainfall depth of station 1 in the first hour is consistently larger or smaller than the rainfall depth of station 2 in the second hour. Two conclusions can be drawn from the correlation analysis in this study: (i) on average, the inter-station distance of the rain gauges in the study area is much smaller than the spatial scale at which there is significant correlation of the rainfall and (ii) the hourly and daily rainfall of the study area is largely variable and erratic.

TEMPORAL PATTERNS: DIURNAL CYCLE

In this thesis, the analysis of the diurnal cycle of the rainfall observations indicated the following: (i) in the Lake Tana basin, rainfall rarely occurs in the morning, i.e. between 0900 – 1200 Local Standard Time (LST), (ii) in the upland areas of Gilgel Abbay, rainfall occurs mostly between 1600 and 1700 LST while the rainfall of Bahir Dar which is situated at the south shore of Lake Tana occurs mostly around 2300 LST, (iii) only 24 % of the seasonal rainfall of Injibara which is situated on a mountain area is

Summary, conclusions and recommendations

nocturnal rainfall that occurs between 2100 – 0900 LST while 73 % of the seasonal rainfall of the south shore is nocturnal rainfall, see figure 3.1 for a map of the area and locations of the station. The result suggests a difference in the mechanisms that cause the rainfall in the mountain areas and that in the lowland areas. In the mountain areas, rainfall occurrence is affected mostly by orographic factors while in the lowland areas, rainfall is affected mostly by the presence of Lake Tana.

The spatial pattern in the rainfall diurnal cycle of the Lake Tana basin is inferred using a convective index (*CI*) that is derived from observations at the 10.8 μm channel of MSG-2. A *CI* value of 1 at a particular LST indicates that a rain producing cloud occurred every day at that LST while a *CI* value of 0 at a particular LST indicates that a rain producing cloud did not occur at that LST throughout the time period of the study. The *CI* values of Gilgel Abbay show a diurnal cycle which is similar to the diurnal cycle that is estimated using the rain gauge observations. Therefore, it is concluded in this thesis that remote sensing observations can be used to study the rainfall diurnal cycle of the areas which are not covered by the recording rain gauges.

The *CI* values suggest that the LST at which intense rainfall occurs show a spatial trend that stretches from the upland areas of the watersheds of the basin towards the lowland areas near the lake. At 1500 and 1800 LST, the frequency of rainfall occurrence could directly be related to terrain elevation while the two are inversely related in the early night hours. In this study, it is noted that such difference in the direction of the relation between rainfall occurrence and terrain elevation can be caused by the relatively strong effects of orography on the afternoon rainfall and the presence of Lake Tana on the nocturnal rainfall. The *CI* values suggest that in Gilgel Abbay, Gumara, and Ribb watershed, rainfall occurs mostly between 1600 LST and 1700 LST while in Megech watershed and Lake Tana, rainfall occurs mostly between 2200 and 2300 LST. As such, it can be concluded that the rainfall of Lake Tana and Megech mostly occur in the same LST of a day on average while that of the rainfall of the remaining three watersheds also occur on the same LST.

A harmonic analysis showed that the characteristics of the rainfall diurnal cycle in the basin have a significant spatial variation which suggests spatial differences in the mechanisms that affect rainfall occurrence. The analysis also revealed that the first and second harmonics explain most of the variance in the rainfall diurnal cycle in Gilgel Abbay, Ribb and Gumara.

However, only the first harmonic explains most of the variance in the diurnal cycle at the south shore of the lake, Megech watershed and Lake Tana. Therefore, it is concluded in this thesis that the semidiurnal cycle of the rainfall in Gilgel Abbay, Ribb and Gumara are significant and should be considered in the harmonic analysis while the semidiurnal cycle of the rainfall in Lake Tana and Megech is too small to be considered.

SCALING ISSUES

The box-counting method has been applied to study the temporal rainfall intermittency in the study area. The results suggest that the rainfall intermittency of Gilgel Abbay watershed has two scaling regimes. The first scaling regime extends from 2 hour to 8 hour while the second scaling regime extends from 2.7 days onwards. The fractal dimension of the second scaling regime is 1.0 which indicates saturation and as such rainfall has been observed at least once in 2.7 days. As explained by the Pearson's correlation coefficient of 0.69, there is a direct relation between the fractal dimension of the first scaling regime and terrain elevation in the watershed. Therefore, it is concluded that the hourly rainfall of the upland areas of Gilgel Abbay watershed is closely clustered than that of the lowland areas.

8.1.2. Rain event properties

To characterize the rainfall better, the spatial and temporal patterns of the rain event properties in the Lake Tana basin are analysed (see Chapter 4). These event properties are rain event depth, event duration, mean event intensity, peak intensity and the length of dry time period between two consecutive rain events which is also termed as inter-event time (IET).

Results indicate that an increase in rain event depth can be related to an increase in event duration, mean event intensity, and peak intensity. Based on correlation analysis, it is concluded in this thesis that rain event depth in the study area is more related to peak intensity than event duration. However, IET has a weak relation with all of the remaining rain event properties. For instance, an event that is preceded by a short dry period does not necessarily have large event depth or high event intensity. Results also showed that the rain events at the start and towards the end of the wet season have relatively large depth, long duration and are separated by long dry time periods as compared to the events in the mid-season. It can

Summary, conclusions and recommendations

be concluded that the rain events at the start and towards the end of the wet season of the study area are mostly convective events.

The spatial variation of rain event properties shows that there is a weak relation between terrain elevation and small values, i.e. the 10 % quantile, of event depth, event duration, peak intensity, and IET. In terms of the median value, rain event depth, mean event intensity and IET in the study area decrease with an increase in terrain elevation while event duration increases with an increase in terrain elevation. It is concluded that the large seasonal rainfall depth in the mountain areas of the study area is mainly due to relatively frequent events that have long duration.

A dimensionless event hyetograph is developed for the rain events of the study area since there is a relatively simple model approach that is based on such hyetographs. The beta probability distribution model is fitted to the dimensionless hyetographs of the observed rain events. Overall, the model satisfactorily reproduced the shape of the dimensionless hyetographs with small errors. This suggests that such models can be used to synthetically generate rainfall intensity but rainfall records of several years are required to become conclusive.

Subsequently, the conditional probability of rainfall occurrences at any two stations is estimated. Such analysis provides information on the spatial structure of rainfall intermittence. The conditional probability of rainfall occurrences at two stations is shown to increase when the rainfall aggregation time is increased from 1 hour to 6 hour and then to 1 day. In this thesis, it is suggested that this is caused either by the presence of events with duration longer than 1 hour or by the presence of multiple events in 6 hour or a day.

An exponential model is fitted to define the conditional probability in terms of inter-station distance. The model performed satisfactorily for the 1 hour and the 6 hour rainfall, respectively. However, the model performed poor when it is fitted to the conditional probability of the daily rainfall occurrences. It can be concluded that for short time scales, the exponential model can be used to estimate conditional probability of rainfall occurrence in terms of inter-station distance but for larger than daily time scales, the probabilities have approximately equal values with some scattering and as such the exponential model cannot be used.

Based on the results, it can be speculated that (i) the difference in rain event properties in the study area is caused by a combination of orographic

factors and the presence of Lake Tana, and (ii) the use of terrain elevation as a covariate of the interpolation of large rain event depths may improve the accuracy of the estimated rainfall but this requires further study.

8.1.3. Remote sensing for rainfall detection and estimation

The next logical step is to evaluate the effectiveness of remote sensing based indices for rainfall detection and estimation. The indices are established based on the observations at selected channels that are the 6.2 μm water vapor (WV) channel and the 10.8 μm and the 12.0 μm thermal infrared (TIR) channels of MSG-2. Rainfall rates from the Precipitation Radar (PR) of Tropical Rainfall Measuring Mission (TRMM) satellite served as the ground-truth in the evaluation of the indices, see Chapter 5.

The following indices are applied to detect rainfall: (i) the brightness temperature that is recorded at the 10.8 μm channel ($T_{10.8}$) which provides information cloud thickness, (ii) the rate of change of $T_{10.8}$ ($\Delta T_{10.8}$) which is used to infer the rate of vertical development of a cloud top surface, (iii) the space gradient of $T_{10.8}$ ($\nabla T_{10.8}$) that provides information about the texture of a cloud top surface, (iv) the brightness temperature difference (BTD) between the 10.8 μm and the 6.2 μm channel ($T_{10.8} - T_{6.2}$) which is used to infer cloud thickness and top height with respect to the tropopause, and (v) the BTD between the 10.8 μm and the 12.0 μm channels ($T_{10.8} - T_{12.0}$) which is applied to differentiate between thick clouds and thin clouds with high top surface. The performances of these indices for rainfall detection are evaluated through a set of categorical statistics that are: *bias*, Probability of Detection (*POD*), False Alarm Ratio (*FAR*) and Heidke Skill Score (*HSS*).

The result using the $\Delta T_{10.8}$ index revealed that rainfall is produced not only during the cloud growing and mature stages but also during the dissipation stage. This is in contrary to the assumption by some TIR based rainfall estimation algorithms where rainfall is assumed to occur during the growth stage only.

For detecting low rainfall rates, the $T_{10.8}$ and the $\Delta T_{10.8}$ indices performed better than the other indices while for detecting high rainfall rates, the $T_{10.8}$ and the BTD ($T_{10.8} - T_{6.2}$) indices performed better than the other indices. It has been shown that the commonly applied $T_{10.8}$ thresholds such as 235 or 253 K overestimated the number of rainfall occurrences. As such, it is concluded in this thesis that the rainfall of the Lake Tana basin is

caused mostly by thick clouds that have high top surfaces. Also, it is concluded that $T_{10.8}$ and $\Delta T_{10.8}$ are the preferred indices for rainfall detection while the BTM ($T_{10.8} - T_{6.2}$) index can provide additional information to detect high rainfall rates. However, it is concluded that BTM ($T_{10.8} - T_{12.0}$) is the least preferred index for rainfall detection.

Results in Chapter 5 suggest that rainfall rate in the Lake Tana basin decreases exponentially with an increase in brightness temperature of cloud top surface. Therefore an exponential model has been developed for TIR based rainfall estimation. The model parameters have been fitted by the method of least squares with a standard error of 0.84 mm h^{-1} . It is concluded that accurate evaluation of the model requires increasing the existing rain gauge network density.

8.1.4. Remote sensing based conceptual cloud modeling for rainfall estimation

One of the major sources of uncertainty in remote sensing based rainfall estimation is that the observations are indirect, see Chapter 1 of this thesis. For instance, thermal infrared (TIR) based rainfall estimation methods rely on cloud top surface temperature to estimate surface rainfall rates. However, the TIR observations do not provide any information on the dynamics of the physical properties below the cloud top surface. As such, TIR based methods are indirect and do not simulate the processes that cause rainfall and such limitation introduces uncertainties in the estimated rainfall. To overcome such limitations, a conceptual cloud-rainfall model for rainfall simulation has been developed in Chapter 5 of this thesis. The model inputs are weather observations at ground surface and satellite infrared observations which are readily available. In this thesis, these inputs are TIR observations from MSG-2 at 15 minute interval and ground surface pressure (P_0), temperature (T_0) and dew-point temperature (T_{d0}) that are recorded at 30 minute interval.

For the cloud model development, the mass balance equations for saturated water vapour, cloud water, rain water and air density have been simplified based on a set of plausible assumptions. The simplified equation defines the rate of change in the liquid water amount of a cloud layer in terms of the difference between the net saturation water vapour input and a rainfall rate as model output. In this thesis, it is concluded that the cloud

model is a parsimonious model that has a similar structure as that of hydrological conceptual models and is not data demanding.

Model sensitivity has been evaluated through Regional Sensitivity Analysis (RSA) which is implemented in the Monte Carlo Simulation Toolbox (MCST). Based on the RSA, it is concluded that the cloud model is most sensitive to the updraft velocity and the response time parameter. The model has been calibrated by mainly changing the values of the sensitive parameters. The model performed satisfactorily when evaluated for four events except that there is some time shift between the start of the simulated and the observed rainfall which is in the order of 30 minute that is considered small in hydrology.

It is also shown that the start time of the rainfall is related to the steep increase and then leveling-off the updraft velocity. It is concluded in this study that the use of such characteristics of the updraft velocity for simulating the rain start time is more promising than the use of the time elapsed since cloud initiation time. The latter is characterized by uncertain initial conditions.

Overall, simulations by the conceptual cloud model are satisfactory, and integrating ground based and remote sensing observations through cloud modeling is a feasible option for rainfall estimation.

8.1.5. Sensitivity of the REW model to rainfall

In hydrological modelling, much emphasis has been given to the study of parameter uncertainty. However, input uncertainty such as rainfall received less attention although such uncertainty can largely propagate to runoff model simulations. In Chapter 7 of this thesis, the sensitivity of the Representative Elementary Watershed (REW) model to rain gauge network density and configuration and to rainfall variability is evaluated and the results are summarized in this section.

First, 256 subset networks are generated from a reference network of 8 rain gauges by changing the network density and configuration. The accuracy of the estimated rainfall using the subset networks is evaluated through a set of performance measures which reflect to what extent the characteristics of the simulated rainfall match to those of the reference rainfall. Only 72 % of the networks produced a representative rainfall estimate as indicated by a Goodness of Rainfall Estimation (*GORE*) index of

larger than 0.8. The accuracy of the estimated rainfall decreased with a decrease in network density. It is also observed that different accuracy levels are obtained for the same network density but different rain gauge configurations. Sekela and Jema are part of the best performing configurations of any network density. Therefore it is concluded that these stations should always be considered for accurate rainfall estimation.

In this thesis, the performance of the REW model is evaluated for the Upper Gilgel Abbay watershed. The performance of the model was found satisfactorily in reproducing the stream flows of the watershed with the values of the performance measures that are similar to those reported for watersheds in other geographic regions. To evaluate model sensitivity to rain gauge network density and configuration, the REW model is run for the rainfall input that is estimated using the best configured networks of 1 – 8 rain gauges. A network of 1 rain gauge resulted in a modified Nash-sutcliffe model Efficiency (NSE_m) of -0.368 indicating a very poor performance while the best configured network of 6 rain gauges resulted in NSE_m of 0.943. As such, it is concluded in this thesis that REW model simulations are significantly affected by the density and configuration of the rain gauge network.

The difference between the simulated flow quantiles and the observed flow quantiles of the same rank are compared for the best performing configurations of 1, 3 and 5 rain gauges. The relative difference is up to 15 – 40 % for high and low flows while it is within 5 % for middle ranged flows. The results indicate that the effect of rainfall representation is large when high and low flows are simulated. It is found that a coarse resolution model, which is discretized based on large Strahler order topology, of the study area is less sensitive to rain gauge network density than fine resolution model.

In Chapter 5 of this thesis, the difference in simulated stream flows as a result of uniform rainfall (UR) input as compared to a non-uniform rainfall (NUR) input is analysed. It is found that the extent of the difference is affected by the spatial and temporal variability of rainfall as well as the antecedent rainfall amount. It is concluded that large temporal and spatial variation and low antecedent rainfall could lead to large model sensitivity to rainfall variability. Such sensitivity is found to be affected by the location of the maximum rainfall with respect to the watershed outlet.

8.2. RECOMMENDATIONS

It is shown that the rainfall of the Gilgel Abbay watershed in the Lake Tana basin varies largely across space and over time. The diurnal cycle of the rainfall in the remaining watersheds of the basin and Lake Tana were inferred from an index that is derived using remote sensing observations. However, the rain event properties in these watersheds and the lake must be evaluated in the future by establishing a network of recording rain gauges similar to that established for the Gilgel Abbay as part of this study.

Rain event properties are important in land degradation studies, notably in soil erosion modelling where rainfall intensity and energy are driving forces. Based on the result of the analysis of the rain event properties, it can be speculated that the rain events of mountain areas have less soil erosive power than the rain events of the lowland areas. However, further work is recommended to evaluate soil erosion in the watershed since severe erosion has been observed in the watershed during the field campaign of this study.

The results of this thesis showed that the existing rain gauge network of the study area is too sparse to reasonably capture the space-time pattern of the rainfall field. The results suggest that the basin rainfall vary at scales that are much smaller than the scale of most remote sensing based rainfall estimates. Such can affect the performance of remote sensing based rainfall estimation method.

In remote sensing based rainfall estimation, one of the major problems is the 'beam-filling' effect. By this effect, the radiance recorded for an image pixel element comes not only from a rain producing cloud but also from other surfaces such as land and water surface. The beam-filling effect can be pronounced in regions where the rainfall has a very small correlation distance which is the case in the Lake Tana basin. Based on the results in this thesis, it is recommended that future work should analyze the pattern of the basin rainfall at scales that are compatible to the spatial resolution of remote sensing observations that is commonly in the order of 5 km. Such analysis is important in the calibration and evaluation of remote sensing based rainfall estimates.

In this thesis, two approaches are evaluated for rainfall estimation using remote sensing observations: (i) a multi-spectral remote sensing

Summary, conclusions and recommendations

approach to detect and estimate rainfall, (ii) a conceptual cloud modeling to simulate rainfall. In particular, the results revealed that the second approach is promising but it must be further evaluated for larger number of rain events. Emphasis must be given to install additional weather stations to provide sufficient spatial coverage. Also, the simulated rainfall must be evaluated for its effectiveness to serve as an input to rainfall-runoff and soil erosion modeling.

In hydrological modeling, much research attention has been given to the issue of model parameter uncertainty while less attention is given to input uncertainty. However, the effect of rainfall representation in rainfall-runoff modeling is still not fully known because conclusions of studies are usually model and watershed specific. In this thesis, such analysis is performed for the Gilgel Abbay watershed but should be repeated to the remaining watersheds of the Lake Tana basin to capture the full temporal and spatial rainfall variability in the area. It is recommended that such studies should also analyze the effect of rainfall input uncertainty on model parameter uncertainty.

REFERENCES

- Adler, F. R., Negri, A. J., 1988: A satellite infrared technique to estimate tropical convective and stratiform rainfall. *J. Appl. Meteorol.*, **27**, 30-51.
- Adler, R. F., Negri, A. J., Keehn, P. R., Hakkarinen, I. M., 1993: Estimation of monthly rainfall over Japan and surrounding waters from a combination of low-orbit MW and geosynchronous IR data. *J. Appl. Meteor.*, **32**, 335-356.
- Ajami N. K., Gupta, H., Wagener, T., Sorooshian, S., 2004: Calibration of a semi-distributed model for streamflow estimation along a river system, *J. Hydrol.*, **298**, 112 – 135.
- Allamano, P., Claps, P., Laio, F., Thea, C., 2009: A data-based assessment of the dependence of short-duration precipitation on elevation. *Phys. Chem. Earth*, **34**, 635 -641.
- Andréassian V., Perrin, C., Michel, C., Usart-Sanchez, I., Lavabre, J., 2001: Impact of imperfect rainfall knowledge on the efficiency and the parameters of watershed models. *J. Hydrol.*, **250**, 206 – 223.
- Andrieu, H., French, M. N., Krajewski, W. F., Georgakakos, K. P., 2003: Stochastic-dynamical rainfall simulation based on weather radar volume scan data. *Adv. Water Resour.*, **26**(5), 681-693.
- Angel, J. R., Palecki, M. A., Hollinger, S. E., 2005: Storm Precipitation in the United States. Part II: soil erosion characteristics. *J. Appl. Meteor.*, **44**: 947 – 959.
- Arkin, P. A., 1979: The relationship between fractional coverage of high cloud and rainfall accumulations during the GAGE over the B-scale array. *Mon. Wea. Rev.*, **107**, 1382-1387.
- Arkin, P. A., Meisner, B. N., 1987: The relationship between large scale convective rainfall and cold cloud over the Western Hemisphere during 1982-84. *Mon. Wea. Rev.*, **115**, 51-74.
- Arnaud P., Bouvier, C., Cisneros, L., Dominguez, R., 2002: Influence of rainfall spatial variability on flood prediction. *J. Hydrol.*, **260**, 216 – 230.

References

- Arnaud, P., Fine, J. A., Lavabre, J., 2007: An hourly rainfall generation model applicable to all types of climate. *Atmos. Res.*, **85**, 230–242.
- Ba, M. B., Nicholson, S. E., 1998: Analysis of convective activity and its relationship to the rainfall over the Rift Valley lakes of East Africa during 1983-90 using Meteosat Infrared channel. *J. Appl. Meteor.*, **37**, 1250–1264.
- Bárdossy, A., Das, T. 2008: Influence of rainfall observation network on model calibration and application. *Hydrol. Earth Syst. Sci.*, **12**, 77 – 89.
- Barrett, E. C. Martin, D. W., 1991: The use of satellite data in rainfall monitoring. Academic press, 340 pp.
- Barrett, E. C., Beaumont, M. J., 1994: Satellite rainfall monitoring: An overview. *Remote Sens. Rev.*, **11**, 23-48.
- Barros, A. P., Joshi, M., Putkonen, J., Burbank, D. W., 2000: A study of the 1999 monsoon rainfall in a mountainous region in central Nepal using TRMM products and rain gauge observations. *Geophys. Res. Lett.*, **27**: 3683 – 3686.
- Barros, A. P., Kim, G., Williams, E., Nesbitt, S. W., 2004: Probing orographic controls in the Himalayas during the monsoon using satellite imagery. *Nat. Hazards Earth Syst. Sci.*, **4**, 29–51.
- Basist, A., Bell, G. D., Meentemeyer, V., 1994: Statistical relationships between topography and precipitation patterns. *J. Climate*, **7**, 1305 – 1315.
- BCEOM and Associates, 1999: Abbay River Integrated Development Master Plan Project.
- Bell V. A., Moore, R., 2000a: The sensitivity of catchments runoff models to rainfall data at different spatial scales, *Hydrol. Earth Syst. Sci.*, **4**, 653-667.
- Bell, V. A., Moore, R. J., 2000b: Short period forecasting of catchment-scale precipitation. Part II: a water-balance storm model for short-term rainfall and flood forecasting. *Hydrol. Earth System Sci.*, **4**, 635–651.
- Beven K. J., Hornberger G. M. 1982: Assessing the effect of spatial pattern of precipitation in modeling stream-flow hydrographs. *Water Resour. Bull.*, **18**, 823–829.

References

- Bewket, W., Conway, D., 2007: A note on the temporal and spatial variability of rainfall in the drought-prone Amhara region of Ethiopia. *Int. J. Climatol.*, **27**, 1467–1477.
- Bidin, K., Chappell, N. A., 2006: Characteristics of rain events at an inland locality in northeastern Borneo, Malaysia. *Hydrol. Process.*, **20**: 3835 – 3850.
- Boi, P., Marrocu, M., Giachetti, A., 2004 : Rainfall estimation from infrared data using an improved version of the Auto-estimator technique. *Int. J. Remote Sens.*, **25**, 4657-4673.
- Bowman, K. P., Collier, J. C., North, G. R., Wu, Q., Ha, E., 2005: Diurnal cycle of tropical precipitation in Tropical Rainfall Measuring Mission (TRMM) satellite and ocean buoy rain gauge data. *J. Geophys. Res.*, **110**, D21104, doi:10.1029/2005JD005763.
- Bras R. L., Rodríguez-Iturbe, I., 1976: Rainfall network design for runoff prediction. *Water Resour. Res.*, **12**, 1197 – 1208.
- Brath, A., Montanari, A., Toth, E., 2004: Analysis of different scenarios of historical data availability on the calibration of a spatially-distributed hydrological model. *J. Hydrol.*, **291**, 232 – 253.
- Brown, B. G., Katz, R. W., Murphy, A. H., 1985: Exploratory analysis of precipitation events with implications for stochastic modelling. *J. Clim. Appl. Meteorol.*, **24**: 57 – 67.
- Buytaert, W., Celleri, R., Willems, P., Bièvre, B. D., Wyseure, G., 2006: Spatial and temporal rainfall variability in mountainous areas: A case study from the south Ecuadorian Andes. *J. Hydrol.*, **329**, 413–421.
- Camberlin, P., 1997: Rainfall anomalies in the Source Region of the Nile and their connection with the Indian Summer Monsoon. *J. Climate*, **10**, 1380–1392.
- Chaubey, I., Haan, C. T., Grunwald, S., Salisbury, J. M. 1999. Uncertainty in the model parameters due to spatial variability of rainfall. *J. Hydrol.*, **220**, 48 – 61.
- Ciach, G. J., Krajewski, W. F., 2006: Analysis and modelling of spatial correlation structure in small-scale rainfall in Central Oklahoma. *Adv. Water Resour.*, **29**, 1450–1463.

References

- Conway, D., 1997: A water balance model of the Blue Nile in Ethiopia. *Hydrolog. Sci. J.*, **42**, 265–286.
- Conway, D., Hulme, M., 1993: Recent fluctuations in precipitation and runoff over the Nile sub-basins and their impact on main Nile discharge. *Climatic change*, **25**, 125–151.
- Curtis, S., Salahuddin, A., Adler, R. F., Huffman, G. J., Gu, G., Hong, Y., 2007: Precipitation extremes estimated by GPCP and TRMM: ENSO relationships. *J. Hydrometeorol.*, **8**, 678 – 689.
- Dai, A. 2001: Global precipitation and thunderstorm frequencies. Part II: Diurnal variations. *J. Climate*, **14**, 1112–1128.
- Dairaku, K., Emori, S., Oki, T., 2004: Rainfall amount, intensity, duration and frequency relationships in the Mae Chaem watershed in Southeast Asia. *J. Hydrometeorol.*, **5**: 458 – 470.
- Dawdy D. R., Bergmann J. M. 1969. Effect of rainfall variability on streamflow simulation. *Water Resour. Res.*, **5**, 958–966.
- De Lima, M. I. P., Grasman, J., 1999: Multifractal analysis of 15-min and daily rainfall from a semi-arid region in Portugal. *J. Hydrol.*, **220**, 1–11.
- de Vos N. J., Rientjes, T. H. M., 2008: Multi - objective training of artificial neural networks for rainfall - runoff modelling. *Water Resour. Res.*, **44**, W08434, doi:10.1029/2007WR006734.
- Demaria, E. M., Nijssen, B., Wagener, T., 2007: Monte Carlo sensitivity analysis of land surfaces using the variable infiltration capacity model. *J. Geophys. Res.*, **112**, D11113, doi: 10.1029/2006JD007534.
- Diodato, N., Bellocchi, G., 2007: Estimating monthly (R)USLE climate input in a Mediterranean region using limited data. *J. Hydrol.*, **345**, 224 – 236.
- Dodov B., Foufoula-Georgiou, E., 2005: Incorporating the spatio-temporal distribution of rainfall and basin geomorphology into nonlinear analyses of streamflow dynamics. *Adv. Water Resour.*, **28**, 711-528.
- Dolciné, L., Andrieu, H., French, M. N., 1998: Evaluation of a conceptual rainfall forecasting model from observed and simulated rain events. *Hydrol. Earth System Sci.*, **2**(2-3), 173–182.

References

- Doswell, C. A., Davies-Jones, R., Keller, D. L., 1990: On summary measures of skill in rare event forecasting based on contingency tables. *Wea. Forecasting*, **5**, 576 – 585.
- Dunkerley, D., 2008a: Rain event properties in nature and in rainfall simulation experiments: a comparative review with recommendations for increasingly systematic study and reporting. *Hydrol. Process.*, **22**: 4415 – 4435.
- Dunkerley, D., 2008b: Identifying individual rain events from pluviograph records: a review with analysis of data from an Australian dryland site. *Hydrol. Process.*, **22**: 5024 – 5036.
- Eklundh, L., Pilesjö, P., 1990: Regionalization and spatial estimation of Ethiopian mean annual rainfall. *Int. J. Climatol.*, **10**, 473–494.
- Feidas, H., 2003: A software tool for monitoring the features of convective cloud systems with the use of Meteosat images. *Environ. Modell. Softw.*, **18**, 1 – 12.
- Freer, J., Beven, K., Ambroise, B., 1996: Bayesian estimation of uncertainty in runoff prediction and the value of data: An application of the GLUE approach. *Water Resour. Res.*, **32**(7), 2161–2173.
- French, M. N., Krajewski, W. F., 1994: A model for real-time quantitative rainfall forecasting using remote sensing 1. Formulation. *Water Resour. Res.*, **30**(4), 1075–1083.
- French, M. N., Andrieu, H., Krajewski, W. F., 1994: A model for real-time quantitative rainfall forecasting using remote sensing 2. Case studies. *Water Resour. Res.*, **30**(4), 1085–1097.
- Fritz, S., Laszlo, I., 1993: Detection of water vapor in the stratosphere over very high clouds in the tropics. *J. Geophys. Res.*, **98** (D12), 22959–22967.
- Gandin L. S., 1970: The planning of meteorological station networks, WMO *Tech. Note* No. 111, 35 p.
- Garcia-Guzman, A., Aranda-Oliver, E., 1993: A Stochastic Model of Dimensionless Hyetograph. *Water Resour. Res.*, **29**: 2363 – 2370.
- Gebremichael, M., Krajewski, W. F., 2004: Characterization of the temporal sampling error in space–time-averaged rainfall estimates from satellites. *J. Geophys. Res.*, **109**, D11110, doi:10.1029/2004JD004509.

References

- Gebremichael, M., Krajewski, W. F., Morrissey, M., Langerud, D., Huffman, G., Adler, R., 2003: Error uncertainty analysis of GPCP monthly rainfall products: A data-based simulation study. *J. Appl. Meteor.*, **42**, 1837–1848.
- Gebremichael, M., Krajewski, W. F., Over, T. M., Takayabu, Y. N., Arkin, P., Katayams, M., 2008: Scaling of tropical rainfall as observed by TRMM precipitation radar. *Atmos. Res.*, **88**, 337 – 354.
- Gebremichael, M., Vivoni, E. R., Watts, C. J., Rodríguez, J. C., 2007: Submesoscale spatiotemporal variability of North American monsoon rainfall over complex terrain. *J. Climate*, **20**, 1751–1773.
- Georgakakos, K. P., Bras, R. L., 1984a: A hydrologically useful station precipitation model 1. Formulation. *Water Resour. Res.*, **20**(11), 1585–1596.
- Georgakakos, K.P., Bras, R.L., 1984b: A hydrologically useful station precipitation model 1. Case studies. *Water Resour. Res.*, **20**(11), 1597–1610.
- Gieske, A. S. M., Hendrikse, J. H. M., Retsios, V., van Leeuwen, B., Maathuis, B. H. P., Romaguera, M., Sobrino, J. A., Timmermans, W. J., Su, Z., 2005: Processing of MSG - 1 SEVIRI data in the thermal infrared - algorithm development with the use of the SPARC2004 data set. In: Proceedings of the ESA WPP-250: SPARC final workshop, 4-5 July, 2005. Enschede : ESA, 2005. 8 p.
- Griffith, C. G., Woodley, W. L., Grube, P. G. , Martin, D. W. , Stout, J., Sikdar, D. N., 1978: Rain estimation from geosynchronous satellite imagery – visible and infrared studies. *Mon. Wea. Rev.*, **106**, 1153 – 1171.
- Haile A. T., Rientjes, T., Gieske, M., Gebremichael, M., 2009a: Rainfall Variability over mountainous and adjacent lake areas: the case of Lake Tana basin at the source of the Blue Nile River. *J. Appl. Meteor. Climatol.*, **48**, 1696 – 1717.
- Haile, A. T., Rientjes, T. H., Gieske, A., Gebremichael, M., 2009b: Rainfall estimation at the source of the Blue Nile: A multispectral remote sensing approach. *Int. J. Appl. Earth Obs. Geoinf.* In Press.
- Haile, A. T., Rientjes, T., Jetten, V., 2009c: Rain event properties and dimensionless rain event hyetographs at the source of the Blue Nile River. *Hydrol. Process. Prepared for submission.*

References

- Haile, A. T., Rientjes, T., Reggiani, P., 2009d: Model sensitivity to rainfall representation: the representative elementary watershed mode. *Water Resour. Res.*, Revised and to be finalized for re-submission.
- Hingray, B., Haha, M. B., 2005: Statistical performance of various deterministic and stochastic models for rainfall series disaggregation. *Atmos. Res.*, **77**: 152 – 175.
- Hong, Y., Hsu, K.L., Sorooshian, S., Gao, X., 2005: Improved representation of diurnal variability of rainfall retrieved from the tropical rainfall measurement mission MW imager adjusted precipitation estimation from remotely sensed information using artificial neural networks PERSIANN system. *J. Geophys. Res.*, **110**, 1 – 13.
- Hossain, F., Huffman, J. G., 2008: Investigating error metrics for satellite rainfall data at hydrologically relevant scales. *J. Hydrometeor.*, **9**, 563 – 575.
- Hsu K.L., Gao, X., Sorooshian S., Gupta H.V., 1997: Precipitation estimation from remotely sensed imagery using an artificial neural network. *J. Appl. Meteorol.*, **36**, 1176-1190.
- Huff, F. A., 1967: Time distribution of rainfall in heavy storms. *Water Resour. Res.*, **3**: 1007 – 1019.
- Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., Hong, Y., Stocker, E. F., Wolff, D. B., 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeor.*, **8**, 38–55.
- Hurst, H. E., Black, R. P., Simaika, Y. M., 1959: The Nile Basin, Vol IX, Government Printing, Cairo Egypt, 206 pp.
- Iguchi, T., Kozu, T., Meneghini, R., Awaka, J., Okamoto, K., 2000: Rain-Profiling Algorithm for the TRMM Precipitation Radar. *J. Appl. Meteorol.*, **39**(12), 2038–2052.
- Imaoka, K., Spencer, R. W., 2000: Diurnal variation of precipitation over the tropical oceans observed by TRMM/TMI combined with SSM/I. *J. Climate*, **13**, 4149 – 4158.
- Inoue, T., 1987: A cloud type classification with NOAA 7 split window measurements. *J. Geophys. Res.*, **92**, 3991 – 4000.

References

- Johnson, G. L., Hanson, C. L., 1995: Topographic and atmospheric influences on precipitation variability over a mountain watershed. *J. Appl. Meteor.*, **34**, 68 – 86.
- Kebede, E.W., 2009: Hydrological responses to land cover changes in Gilgel Abbay catchment, Ethiopia. Enschede, ITC, 2009. 67 p.
- Kebede, S., Travi, Y., Alemayehu, T., Marc, V., 2006: Water balance of Lake Tana and its sensitivity to fluctuations in rainfall, Blue Nile basin, Ethiopia. *J. Hydrol.*, **316**, 233–247.
- Kidder, S. Q., Vonder Haar, T., H., 1995: Satellite Meteorology: An introduction. Academic press, 466 pp.
- Koutsoyiannis, D., Onof, C., 2001: Rainfall disaggregation using adjusting procedures on a poisson cluster model. *J. Hydrol.*, **246**: 109 – 122.
- Kuligowski, R. J., 2002: A self-calibrating real-time GOES rainfall algorithm for short-term rainfall estimates. *J. Hydrometeor.*, **3**, 112-130.
- Kusumastuti, D. I., Stuthers, I., Sivaplan, M., Reynolds, D. A., 2007: Threshold effects in catchment storm response and the occurrence and magnitude of flood events: implications for flood frequency. *Hydrol. Earth Syst. Sci.*, **11**: 1515 – 1528.
- Laing, A. G., Fritsch, J. M., Negri, A. J., 1999: Contribution of mesoscale convective complexes to rainfall in Sahelian Africa: Estimates from Geostationary Infrared and Passive microwave data. *J. Appl. Meteor.*, **38**, 957 – 964.
- Lamb, R., Beven, K., Mayrabø, S., 1997: Discharge and water table predictions using a generalized top model formulation. *Hydrological processes*, **11**, 1145 – 1167.
- Lee, M., Schubert, S. D., Suarez, M. J., Bell, T. L., Kim, K.-M., 2007: Diurnal cycle of precipitation in the NASA Seasonal to International Prediction Project atmospheric general circulation model. *J. Geophys. Res.*, **112**, D16111, doi:10.1029/2006JD008346.
- Levizzani, V., R. Amorati, and F. Meneguzzo, 2002: A review of satellite-based rainfall estimation methods. *European Commission Project MUSIC Report (EVK1-CT-2000-00058*, <http://www.isac.cnr.it/~meteosat/pub.html>), 66 pp.

References

- Loukas, A., Quick, M. C., 1996: Spatial and temporal distribution of storm precipitation in southwestern British Columbia. *J. Hydrol.*, **174**: 37 – 56.
- Maathuis, B. H. P., Gieske, A. S. M., Retsios, V., van Leeuwen, B., Mannaerts, C. M., Hendrikse, J. H. M., 2006: Meteosat - 8 : from temperature to rainfall. In: ISPRS 2006: ISPRS mid-term symposium 2006 remote sensing: from pixels to processes, 8-11 May 2006, Enschede, the Netherlands. Enschede: ITC, 2006. 1 – 5.
- Menabde M, Sivapalan M. 2001: Linking space-time variability of runoff and rainfall fields: a dynamic approach. *Adv. Water Resour.*, **24**, 1001 – 1014.
- Meneghini, R., Iguchi, T., Kozu, T., Liao, L., Okamoto, K., Jones, J.A., Kwiatkowski, J., 2000: Use of the Surface Reference Technique for Path Attenuation Estimates from the TRMM Precipitation Radar. *J. Appl. Meteor.*, **39**(12), 2053–2070.
- Milly P, Eagleson P. 1987a: Effect of Storm Scale on Surface Runoff Volume. *Water Resour. Res.*, **24**: 620–624.
- Milly, P., Eagleson P. 1987b: Effects of Spatial Variability on Annual Average Water Balance. *Water Resour. Res.*, **23**: 2135-2143.
- Morin, E., Goodrich, D. C., Maddox, R. A., Gao, X., Gupta, H. V., Sorooshian, S. 2006: Spatial patterns in thunderstorm rainfall events and their coupling with watershed hydrological response. *Adv. Water Resour.*, **29**, 843 – 860.
- Naden P. S., 1992: Spatial variability in flood estimation for large catchments: the exploitation of channel network structure. *Hydrol. Sci. J.*, **37**, 53-71.
- Negri, A. J., Adler, R. F., Wetzel, P. J., 1984: Rain estimation from satellite: An examination of the Griffith-Woodley technique. *J. Clim. Appl. Meteorol.*, **23**, 102-116.
- Nykanen, D. K., 2008: Linkages between orographic forcing and the scaling properties of convective rainfall in mountainous regions. *J. Hydrometeor.*, **9**, 327 – 347.
- Obled C., Wendling, J., Beven, K., 1994: The sensitivity of hydrological models to spatial rainfall patterns: an observation using observed data. *J. Hydrol.*, **159**, 305 – 333.

References

- Ohsawa, T., Ueda, H., Hayashi, T., Watanabe, A., Matsumoto, J., 2001: Diurnal variations of convective activity and rainfall in Tropical Asia. *J. Meteor. Soc. Japan*, **79**, 333–352.
- Olsson, J., Berndtsson, R., 1998: Temporal rainfall disaggregation based on scaling properties. *Water Sci. Technol.*, **37** (11), 73–79.
- Over, T. M., Gupta, V. K., 1994: Statistical analysis of mesoscale rainfall: Dependence of a random cascade generator on large-scale forcing. *J. Appl. Meteorol.*, **33**, 1526–1542.
- Palecki, M. A., Angel, J. R., Hollinger SE. 2005: Storm Precipitation in the United States. Part I: Meteorological characteristics. *J. Appl. Meteorol.*, **44**: 933 – 946.
- Pedder, M. A., Haile, M., Thorpe, A. J., 2000: Short period forecasting of catchment-scale precipitation. Part I: the role of Numerical Weather Prediction. HYREX Special Issue, *Hydrol. Earth System Sci.*, **4**, 627 – 633.
- Petty, G. W., 1995: The status of satellite-based rainfall estimation over land. *Remote Sens. Environ.*, **51**, 125-137.
- Rappold, G. D., 2005: Precipitation analysis and agricultural water availability in the southern highlands of Yemen. *Hydrol. Process.*, **19**, 2437–2449.
- Reggiani P. R., Rientjes, T. H. M., 2005: Flux parameterization in the Representative Elementary Watershed (REW) approach: application to a natural basin. *Water Resour. Res.*, **41**, W04013, doi:10.1029/2004WR003693.
- Reggiani P., Hassanizadeh, S. M., Sivapalan, M., Gray, W. G., 1999: A unifying framework of watershed thermodynamics: Constitutive relationships. *Adv. Water Resour.*, **23**, 15-39.
- Reggiani P., Sivapalan, M., Hassanizadeh, S. M., 1998: A unifying framework of watershed thermodynamics: 1 Balance equations for mass, momentum, energy and entropy and the second law of thermodynamics. *Adv. Water Resour.*, **22**, 367-398, doi:10.1016/S0309-1708(98)00012-8.
- Reggiani, P. R., Rientjes, T. H. M., 2005: Flux parameterization in the Representative Elementary Watershed (REW) approach: application

References

- to a natural basin, *Water Resour. Res.*, **41**: W04013, doi:10.1029/2004WR003693.
- Renard, K. G., Freimund, J. R., 1994: Using monthly precipitation data to estimate the R factor in the revised USLE. *J. Hydrol.*, **157**: 287 – 306.
- Rientjes T., de Neef, M., Hassanizadeh, S. M., Reggiani, P., 2003: Development of the representative elementary watershed modeling code, In: van Loon, E.E., Troch, P.A. (Eds.), Final Report of the DAUFIN Project. Section Waterhuishouding, Wageningen University, The Netherlands, 7–24.
- Robinson J. S, Sivapalan M. 1997 : An investigation into the physical causes of scaling and heterogeneity of regional flood frequency. *Water Resour. Res.*, **33**: 1045-1049.
- Robinson P. J. 1994: Precipitation regime changes over small watersheds. In Statistics for the environment 2 Water related Issues. Barnett, V. and Turkman, K. F. (Ed.). John Wiley & Sons 43 – 59.
- Rodriguez-Iturbe, I., Bras, R. L., 1976: Rainfall network design for runoff prediction. *Water Resour. Res.*, **12**, 1197–2208.
- Segond M., Wheater, H. S., Onof, C., 2007: The significance of spatial rainfall representation for flood runoff representation: A numerical evaluation based on the Lee catchment, UK, *J. Hydrol.*, **347**, 116 – 131.
- Segond, M-L., Onof, C., Wheater, H. S., 2006: Spatial-temporal disaggregation of daily rainfall from a generalized linear model *J. Hydrol.*, **331**: 674–689.
- Şen, Z., Habib, Z., 2001: Monthly spatial correlation functions and interpretations for Turkey. *Hydrolog. Sci. J.*, **46(4)**, 525–535.
- Sinclair, M. R., 1994: A diagnostic model for estimating orographic precipitation. *J. Appl. Meteor.*, **33**, 1163–1175.
- SMEC, 2007: Hydrological study of the Tana-Beles sub-Basins. *Draft Inception Report*, 85pp.
- Smith M. B., Koren, V. I., Zhang, Z., Reed, S. M., Pan, J., Moreda, F., 2004: Runoff response to spatial variability of in precipitation: an analysis of observed data. *J. Hydrol.*, **298**, 267 – 286.

References

- Sorooshian, S., Hsu, K. -L., Gao, X., Gupta, H. V., Imam B., Braithwaite, D., 2000: Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bull. Am. Meteor. Soc.*, **81**, 2035 - 2046.
- Spear, R. C., Hornberger, G. M., 1980: Eutrophication in Peel Inlet, II. Identification of critical uncertainties via Generalised Sensitivity Analysis. *Water Resour. Res.*, **14**, 43 – 49.
- Stedinger, J. R., Vogel, R. M., Foufoula-Georgiou, E., 1993: Frequency analysis of extreme events. Handbook of Hydrology, Maidment D. R. (ed.). McGraw-Hill New York; 18.1–18.66.
- Stephens, G. L., Kummerow, C. D., 2007: The remote sensing of clouds and precipitation from space: A review. *J. Atmos. Sci.*, **64**, 3742 -3765.
- Stull, R. B., 2000: Meteorology for Scientists and Engineers. 2nd Edition. Brooks/Cole. ISBN 0-534-37214-7.
- Syed, K.H., Goodrich, D.C., Myers, D.E., Sorooshian, S. 2003: Spatial characteristics of thunderstorm rainfall and their relation to runoff. *J. Hydrol.*, **271**, 1 – 21.
- Todd, M. C., Barrett, E. C., Beaumont M. J., Green, J. L., 1995: Satellite Identification of Rain Days over the Upper Nile River Basin Using an Optimum IR Rain/No-Rain Threshold Temperature Model. *J. App. Meteorol.*, **34** (12), 2600–2611.
- Tsubo, M., Walker, S., Hensley, M., 2005: Quantifying risk for water harvesting under semi-arid conditions. Part I. Rainfall intensity generation. *Agricultural water management* **76**: 77 – 93.
- Turk, F. J., Miller, S. D., 2005: Toward improving estimates of remotely-sensed precipitation with MODIS/AMSR-E blended data techniques. *IEEE Trans. Geosci. Rem. Sens.*, **43**, 1059-1069.
- van Dijk, A. I. J. M., Meesters, A. G. C. A., Schellekens, J., Bruijnzeel, L. A., 2005: A two-parameter exponential rainfall depth-intensity distribution applied to runoff and erosion modeling. *J. Hydrol.*, **300**: 155 – 171.
- Varado N., Braud, I., Galle, S., Le Lay, M., Séguis, L., Kamagate, B., Depraetere, C., 2006: Multi-criteria assessment of the Representative Elementary Watershed approach on the Donga catchment (Benin) using a downward approach of model complexity. *Hydrol. Earth Syst. Sci.*, **10**, 427 – 442.

References

- Vicente, G. A., Scofield, R. S., Menzel, W. P., 1998: The operational GOES infrared rainfall estimation technique. *Bull. Am. Meteor. Soc.*, **79**, 1883-1898.
- Vrugt J. A., Gupta, H. V., Dekker, S. C., Sorooshian, S., Wagener, T., Bouten, W., 2006: Application of stochastic parameter optimization to the Sacramento soil moisture accounting model. *J. Hydrol.*, **325**, 288 – 307.
- Wagener T., Wheater, H. S., 2006: Parameter estimation and regionalization for continuous rainfall-runoff models including uncertainty. *J. Hydrol.*, **320**, 132 – 154.
- Wagener, T., Boyle, D. P., Lees, M. J., Wheater, H. S., Gupta, H. V., Sorooshian, S., 2001: A framework for development and application of hydrological models. *Hydrol. Earth System Sci.*, **5**(1), 13–26.
- Wale, A., Rientjes, T. H. M., Gieske, A. S. M., Getachew, H. A., 2009: Ungauged catchment contributions to Lake Tana's water balance. *Hydrol. Process.*, In Press. DOI: 10.1002/hyp.7284.
- Wardah, T., Abu Bakar, S. H., Bardossy, A., Maznorizan, M., 2008: Use of geostationary meteorological satellite images in convective rain estimation for flash-flood forecasting. *J. Hydrol.*, **356**, 283-298.
- Weiss, M., Smith, E. A., 1987: Precipitation discrimination from satellite infrared temperatures over the CCOPE Mesonet region. *J. Clim. Appl. Meteorol.*, **26**, 687 – 697.
- Wilson CB, Valdes JB, Rodriguez-Iturbe I. 1979: Influence of the spatial distribution of rainfall on storm runoff. *Water Resour. Res.*, **15**, 321–328.
- Woods R., Sivapalan, M., 1999: A synthesis of space-time variability in storm response, runoff generation and routing. *Water Resour. Res.*, **35**, 2469 – 2489.
- Woolhiser, D. A., Osborn, H. B., 1985: A stochastic model of dimensionless thunderstorm rainfall. *Water Resour. Res.*, **21**: 511 – 522.
- Young, C. B., Bradley, A. A., Krajewski, W. F., Kruger, A., Morrissey, M. L., 2000: Evaluating NEXRAD multisensor precipitation estimates for operational hydrologic forecasting. *J. Appl. Meteor.*, **1**, 241–254.

References

- Younger, P. M., Freer, J.E., Beven, K. J., 2009: Detecting the effects of spatial variability of rainfall on hydrological modeling within an uncertainty analysis framework. *Hydrol. Process.*, **23**, 1988 – 2003.
- Zar, J. H., 1996. Biostatistical Analysis. (3rd ed.) Upper Saddle River, NJ: Prentice Hall.
- Zeng, N., Shuttleworth, J. W., Gash, J. H. C., 2000: Influence of temporal variability of rainfall on interception loss. Part I. Point analysis. *J. Hydrol.*, **228**: 228 – 241.
- Zhang G. P., Fenicia, F., Rientjes, T. H. M., Reggiani, P., Savenije, H. H. G. , 2005: Modeling runoff generation in the Geer river basin with improved model parameterizations to the REW approach. *Phys. Chem. Earth.*, **30**, 285-296.

Curriculum Vitae

Alemseged Tamiru Haile was born on February 28, 1978 in Addis Ababa, Ethiopia. After completing high school in 1996, he joined Arba Minch Water Technology institute (AWTI) of Arba Minch University to pursue his study for B.Sc. degree. In 2001, Alemseged received his B.Sc. degree in Irrigation Engineering with distinction and was offered a lecturer position in AWTI. He has taught several undergraduate courses including principles of hydraulics, irrigation and hydrology.

Alemseged joined the International Institute for Geo-information Science and Earth Observation (ITC), Enschede, The Netherlands in 2003 where he received his M. Sc. degree in 'Geo-information science and Earth Observation with specialization Watershed management, Conservation and River Basin Planning' in 2005. His M.Sc. thesis was selected as the best thesis by the scientific jury of ITC and he has received the 'van de Klaas Jan Beek Award of 2005'. The title of his M.Sc. thesis is: 'Integrating hydrodynamic models and high resolution DEM (LIDAR) for flood modeling'.

Since March 2006, Alemseged has been a Ph.D. candidate at the ITC faculty of the University of Twente. His research interest includes hydrologic modeling, flood modeling, and hydrological change detection and variability using ground based and remote sensing observations.

Publications

ISI Journal Articles

- **Haile, A.T.**, Rientjes, T.H.M. and Gieske, A. S. M. 2009: Rainfall variability over mountainous and adjacent lake areas: the case of Lake Tana basin at the source of the Blue Nile river. *Journal of applied meteorology and climatology*, 48, 1696–1717.
- **Haile, A. T.**, Rientjes, T., Gieske, A., Gebremichael, M. 2009: Multispectral remote sensing for rainfall detection and estimation at the source of the Blue Nile River, *International Journal of Applied Earth Observation and Geoinformation*, <http://dx.doi.org/10.1016/j.jag.2009.09.001>. (In press).
- Abdo, K.S., Fiseha, B.M., Rientjes, T.H.M., Gieske, A.S.M. and **Haile, A.T.** 2009: Assessment of climate change impacts on the hydrology of

Gilgel Abbay catchment in Lake Tana basin, Ethiopia. *Hydrological processes*, 23, 3661-3669.

- **Haile, A. T.**, Rientjes, T., Reggiani, P., 2009: Model sensitivity to rainfall representation: the representative elementary watershed model. *Water Resources Research*, Revised and re-submitted.
- **Haile, A. T.**, Rientjes, T., Gieske, A., Jetten, V. G., Gebremichael, M., 2009: Satellite remote sensing and conceptual cloud modelling for convective rainfall simulation. *Advances in Water Resources*. Conditionally accepted.
- Tarekegn, T. A., **Haile, A. T.**, Rientjes, T., Alkema, D., 2009: Two-Dimensional hydrodynamic flood modelling using ASTER DEM. *International Journal of Applied Earth Observation and Geoinformation*, Conditionally accepted.
- **Haile, A. T.**, Rientjes, T., Jetten, V., 2009: Rain event properties and dimensionless rain event hyetographs at the source of the Blue Nile River. *Hydrological Processes*. Prepared for submission.

Book Chapters

- Rientjes, T.H.M. and **Haile, A.T.** 2009: The effectiveness of high - resolution LIDAR DSM for two dimensional hydrodynamic flood modeling in an urban area. In: *Quality aspects in spatial data mining.* / ed. by. A. Stein, W. Shi and W. Bijker. Boca Raton : CRC, 2009. ISBN978-1-4200-6926-6.pp.221-238.

Conference papers (with peer review)

- **Haile, A.T.** and Rientjes, T.H.M. 2007: Spatio - temporal rainfall mapping from space: setbacks and strengths. In: Proceedings of the 5th International symposium on Spatial Data Quality SDQ 2007, Modelling qualities in space and time, ITC, Enschede, The Netherlands, 13-15 June, 2007. Enschede : ITC, 2007. 9 p.
- **Haile, A.T.** and Rientjes, T.H.M. 2007: Uncertainty issues in hydrodynamic flood modeling. In: Proceedings of the 5th International symposium on Spatial Data Quality SDQ 2007, Modelling qualities in space and time, ITC, Enschede, The Netherlands, 13-15 June, 2007. Enschede : ITC, 2007. 6 p.
- **Haile, A.T.** and Rientjes, T.H.M. 2005: Effects of LIDAR DEM resolution in flood modelling: a model sensitivity study for the city of Tegucigalpa,

Conference Proceedings (full papers)

- Gieske, A.S.M., Rientjes, T.H.M., **Haile, A.T.**, Abeyou, W.W. and Getachew, H.A. 2008: Determination of Lake Tana evaporation by the combined use of SEVIRI, AVHRR and IASI. In: Proceedings of the 2008 EUMETSAT Meteorological Satellite Conference, Session 10. hydrometeorology, 8-12 September 2008, Darmstad, Germany. 7 p.
- Gieske, A.S.M., Abeyou, W.W., Getachew, H.A., **Haile, A.T.** and Rientjes, T.H.M. 2008: Non - linear parametrization of Lake Tana's flow system. In: Proceedings of the workshop on hydrology and ecology of the Nile river basin under extreme conditions : Addis Ababa, Ethiopia, June 16-19, 2008 / editor W. Abtew, A.M. Melesse. - Miami : Florida International University, 2008. ISBN 978-1-4276-3150-3. pp. 127-144.
- Fiseha, B.M., **Haile, A.T.**, Rientjes, T.H.M. and Gieske, A.S.M. 2008: Rainfall estimation using satellite remote sensing and ground truth for hydrologic modelling over the Upper Blue Nile region. In: Proceedings of the workshop on hydrology and ecology of the Nile river basin under extreme conditions : Addis Ababa, Ethiopia, June 16-19, 2008 / editor W. Abtew, A.M. Melesse. - Miami : Florida International University, 2008. ISBN 978-1-4276-3150-3. pp. 228-248.

MSc thesis

- **Haile, A.T.** 2005: Integrating hydrodynamic models and high resolution LIDAR DEM for flood modelling. Enschede, ITC, 2005. 75 p.