DEVELOPMENT OF INTENSITY – DURATION REQUENCY RELATIONSHIPS FOR BURUNDI COUNTRY

BY

GAHUNGU CHRISTOPHE

A THESIS SUBMITTED

TO

THE UNIVERSITY OF ARBAMINCH IN PARTIAL FULFILLMENT OF THE REQUIREMENTS OF THE DEGREE OF MASTER OF SCIENCE IN HYDROLOGY AND WATER RESOURCES MANAGEMENT

ARBAMINCH UNIVERSITY

 SEPTEMBER, 2008

PDF created with pdfFactory Pro trial version www.pdffactory.com

CERTIFICTION

The undersigned certify that they have read the dissertation entitled: **Development of Intensity –Duration-Frequency Relationships for Burundi country** in partial fulfillment of the requirement of the degree of Masters of Science in Hydrology and water Resources management.

Dr. Semu Ayalew Moges

__

Advisor

Date **Date**

PDF created with pdfFactory Pro trial version www.pdffactory.com

DECLARATION AND COPYRIGHT

I, Christophe Gahungu, declare that this is my own original work and that it has not been presented and will not be presented to any other University for similar or other degree award.

Signature ……………………………………..

Date…………………………………………..

This thesis is copy righted material protected under the Berne Convection, the Copy Right Act 1999 and other international and national enactments, in that behalf, on the intellectual property. It may not be reproduced by any in full or in parts, except for short extracts in fair dealing, for research or private study, critical scholarly review or discourse with an acknowledgement, without written permission of the school of Graduate studies, on behalf of both the author and the Arba Minch University.

DEDICATION

TO the memory of my Family

L

Acknowledgment

I would like to express special thanks to indebted to my God, through him that I had my well being and passed every hurdle in my study period and in my life at all.

A special thanks is also owed to ATP – Nile Basin Program for the Scholarship and supports I obtained

I owe a great deal of gratitude to Dr Semu Ayalew Moges, my Advisor. Without him, I surely would have been a lost ball in tall grass. I appreciate all the did to help me get where I am today. I am gratefully for all his encouragement, guidance and support.

I express my gratitude to all Post Graduate Instructors for the knowledge and experiences shared with me.

Also, I would like to thank my uncle, Mrs. Sinarinzi. E, he have provided me with strength and encouragement in times when I need it most.

Lastly, I would like to thanks my fiancée, N. Justine. Without her understanding and support, I would have been lost. Her constant words of encouragements and her ever- ready shoulder to cry on were a saving grace.

ABSTRACT

The rainfall-Intensity-Duration-frequency (IDF) relationships is one of the most commonly used tools in water resources engineering , either for planning , designing and operating of water resource projects. The IDF allows for the estimation of the return period of an observed rainfall events for different time durations conversely, it may be used to estimate rainfall amount corresponding to a given return period for different aggregation times. The objective of the research is therefore, to develop operational IDF relationships for the Burundi whole country based on nineteen first class stations.

The annual maximum rainfall magnitudes of varying durations were collected from rainfall charts and fitted to the probability distributions after which quantiles estimated for different return periods based on the best fitted distributions. Then, the rainfall intensities are computed and the parameters of the general mathematical form of IDF were generated for each station. Three different methods expressing rainfall intensities were established for the study area: the general mathematical form, curves relating Intensity-Duration-Frequency of rainfall and IDF maps.

Burundi area has been regionalized and four different regions were established based pooled quantiles of the 24-hour durations. The regional IDF parameters, IDF curves and regression equations were developed for each region. This helps to extract the intensity of rainfall of any durations and frequency at areas farthest from the principal stations.

The result of this research can be used to all water professionals and designers on the activities of water resources and other related disciplines with in the region by supplying important information's on rainfall Intensity, Duration and Frequency relationships.

TABLE OF CONTENTS

LIST OF APPENDICES

LIST OF TABLES

LiST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

CHAPTER ONE: INTRODUCTION AND BACKGROUND

1.1.INTRODUCTION

The rainfall – Intensity – Duration – Frequency (IDF) relationship is one of the most commonly used tools in water resources engineering, either for planning, designing, or operating of water resource projects, or for various engineering projects against floods. It gives an idea about the frequency or return period of rainfall intensity or rainfall volume that can be expected with in a certain period, i.e., the storm duration (*Pilgrim, 2001*).

Rainfall Intensity – Duration – Frequency (IDF) curves are graphical representations of the amount of water that falls with in a given period of time. It is usually presented as a graph, with duration (D) plotted on the horizontal axis, intensity (I) on the vertical axis and a series of curves, one for each design return period (T) (*chow, 1988*).The IDF curves represent, for a given non- exceedence probability (or usually) expressed in terms of the return period in years, the variation of the maximum annual rainfall intensity with in a specific time interval.

The development of intensity duration frequency IDF curves for precipitation remains a powerful tool in the risk analysis of natural hazards. Indeed the IDF curves allow for the estimation of the return period of an observed rainfall event or conversely of the rainfall amount corresponding to a given return period for different aggregation times.

The purpose of this study is manly to produce IDF relationships for precipitation for nineteen different first order recording climatologicall stations found in Burundi country because the IDF relationships have not been developed for this area. These stations are Ruvyironza, Muyinga, Gisozi, musasa, Ruyigi, Rwegura, Karuzi, Gitega, Mparambo, Bujumbura, Cankuzo, Nyanza-lac, Makamba, Kirundo, Kinyinya, Tora, Rumonge, Teza and Nyamuswaga,

The basic information for the selected different first order recording climatologically stations of the region and some peripheral stations have been described under section 3.1 in table 3.3.

1.2. Description of study Area.

1.2.1. Location

Burundi is a landlocked country located in east central Africa at 3^0 3^{S} under the equator; bordering Rwanda to its north, Tanzania to the east and south and to the west by the former Democratic Republic of the Congo. Burundi's general locate is defined as between 2⁰s and 4⁰ 30[']s of latitude; between 2.9 ⁰ E and 31⁰ E of longitude (Kabundege,.G., 2007).

Figure 1.1: Location map of the study area (Burundi location).

1.2.2. Topography

Burundi with tropical climate is characterized by mountainous relief which extends over a limited area of 27834 km^2 . The major areas of this relief are: the Imbo plains over looking the Tanganyika Lake, west of Burundi. This plain make up the natural region of Imbo (Ntiburumusi, 2000).They constitute the part of the African rift valley. The width of these plains ranges from 2 km to 25 km. The average annual temperature is 23 $\rm ^{0}$ c and the average annual precipitation is 800mm.

The Congo- Nile crests which the average height is around 2300 m. The highest peak in Burundi ranges from 2650 m to 2670 m of altitude for Heha Mountain. This mountain range separates the waters of the Burundi into 2 basins: The Nile basin and the Congo basin.

1.2.3. Climate

Burundi's general climate is defined as tropical highland, but differences in altitude from region to region cause temperature variations. The equatorial; high plateau with considerable altitude variation (772 m to 2,760m); average annual temperature varies with altitude from 23 to 17 degrees centigrade but is generally moderate as the average altitude is about 1700 m; average annual rainfall is about 150 cm . There are four clear seasons; the short dry season (December-January) and the long dry season (June-August); the long wet season (February- May) and finally the short wet season (September-November) (Ntiburumusi, 2000).

1.3. Problem statement

 Engineers must often consider storm run-off when planning new water projects. One of the first steps in many hydrologic design projects such as in urban drainage design, risk analysis of natural hazards is the determination of the rainfall events or events that involve a relationship between rainfall intensity (and depth), duration and frequency or return period appropriate for the facility and site location.

Risk evaluations and mitigations necessitate statistical information in order to plan appropriate infrastructure related to sewerage, dikes … in order to project effectively the population and goods.

To effectively protect populations and ensure the longevity of infrastructures, it is indispensable to accurately estimate the risks associated with extreme event and mitigation necessitates statistical information. Consequently to supply to engineers, governments, insurance and risk management companies, every key statistical elements necessary to build reliable, safe and adequately positioned infrastructures.

The main problem in Burundi, the hydrological information like IDF, being the principal input of design of water resources and other similar sectors, is not yet well developed and not yet readily available in a systematic relationships to the concerned users.

In this context and taking into account of the variability of rainfall in the country, it is necessary to produce and regionalize Intensity-Duration-Frequency (IDF) curves for each station and for each region.

1.4. Objectives of the study.

1.4.1 Global objective

The global objective of the study is to produce operational Intensity – Duration-Frequency relationships for Burundi country.

1.4.2. Specific objectives

Taking in to account available information on rainfall intensities

- Computing the IDF parameters,
- Constructing IDF curves, and IDF maps covering the country,
- Grouping together the homogeneous regions based on 24 hours durations of annual maximum rainfall depth,

1.5. Scope of the study.

This study is limited to the development of Intensity $-$ Duration $-$ Frequency relationships, construction of IDF maps covering the County, grouping homogenous regions together, developing regional IDF curves relationships

1.6. Thesis organization

This thesis is categorized into six main chapters. Chapter one describes introduction and back ground of the study area, problem description, objectives and significance of the study. Chapter two presents the literature review. Chapter three deals with material and methods used data source and availability for the study. Chapter four presents data analysis, results and discussions for the establishments of at-site IDF relationships. Chapter five presents regionalization of homogeneous regions based on annual maximum rainfall data of 24hr durations and conducting homogeneity tests. Chapter six includes conclusion and recommendations

1.7. Materials and methods

The first step in this study was setting station selection criteria. The stations selected are all with first class self recording gauges. The locations of these stations are in such a way that they can represent the region's different geographical coverage. Annual maximum rainfalls of different durations from 19 (nineteen) stations were considered for the study.

Generally the following procedures were employed in carrying out the thesis work

- 1. Collection of annual maximum rainfall data for 0.5,1, 2, 3, 5, 6, 12, 24 hrs durations using the annual maximum series model.
- 2. Carrying out data quality control
- 3. Selection and evaluation of frequency distributions
- 4. Selection of methods of parameter estimations methods
- 5. Estimation of parameters, Quantiles, and Standard error of estimate
- 6. Estimation of IDF parameters and evaluations
- 7. Construction of IDF curves and IDF maps
- 8. Regionalization and identifying homogeneous regions based on the 24 hr duration annual maximum rainfall data.
- 9. Test of homogeneous regions, delineations of homogeneous regions and employing graphical evaluations of regional stations.

The other methodology employed to meet the reach objectives includes literature review, formulating data collection format, developing data processing and presentation methods.

CHAPTER TWO: LITERATURE REVIEW

2.1. Introduction

The establishment of Intensity-Duration-Frequency (IDF) curves for precipitation remains a powerful tool in the risk analysis of natural hazards. Indeed the IDF-curves allow for the estimation of the return period of an observed rainfall event or conversely of the rainfall amount corresponding to a given return period for different aggregation times.

There is a high need for IDF-curves in Africa, and especially in Burundi. One of the first step in many hydrologic design projects, such as in urban drainage design is the determination of the rainfall event or events to be used. The most common approach is to use a design storm or event that involves a relationship between rainfall intensity (and depth), duration, and the frequency or return period appropriate for the facility and site location. In many cases, the hydrologist has standard intensity duration frequency (IDF) curves available for the site and does not have to perform this analysis. The IDF is usually presented as a graph, with duration plotted on the horizontal axis, intensity on the vertical axis, and a series of curves, one for each design return period (Chow, 1988).

2.2. Equations for IDF Relationships

IDF Curves have also been expressed as equations to avoid having to read the design rainfall intensity from a graph. For example, Wenzel (1982) provided coefficients from a number of cities in the United States for an equation of the form

$$
i = \frac{C}{T^e + f} \quad \dots \tag{2.1}
$$

Where i is the design rainfall intensity, T_d is the duration, and C, e, and f are coefficients varying with location and return period (Chow 1988).

It is also possible to extend the above equation to include the return period T using the equation

$$
i = \frac{CT^m}{T_d^e + f}
$$
 (2.2)

Wenzel, (1982) has also proposed a relationship between intensity–Duration– Frequency which is applicable in most locations by the equation of the form

$$
I = \frac{A}{\left(D+C\right)^{B}} \dots \tag{2.3}
$$

Where: I is intensity, D is duration, A is a constant for a given return period, B and C are constants that do not depend on return period.

The 'A' coefficient

The value of the 'A' coefficient depends on (i) the return interval in years of the storm and (ii) the system of units being used.

$$
A=aT^m
$$

$$
I = \frac{aT^m}{(D+B)^c}
$$
............2.4

The 'B' constant

This constant in minutes is used to make the log-log correlation as linear as possible. Typical values range from 2 to 12 minutes. A value of zero for this parameter represents a special case of the IDF equation where

$$
i = \frac{A}{D^c}
$$

In general, this results in poor agreement between observed values of intensity and duration and those represented by the IDF equation.

The 'C' exponent

This parameter is usually less than 1.0 and is obtained in the process of fitting the data to the power expression. Values are usually in the range of 0.75 to 1.0

These equations have no theoretical basis; they are purely empirical devices that are some times useful for expressing relations such as depth–exceedence probability and return period. The constants in the above equation have a strong geographic variation and must be determined by analysis of data for the location of interest.

After determining the numerical value of the IDF parameters, rainfall intensity for any duration and recurrence interval can be computed. Based on the estimated parameters of the IDF relationships of the general equation of the form

$$
i = \exp[(\ln(A) - C \ln(B + D))]
$$
 (2.5)

is developed to calculate the intensities for all durations at each station.

2.3. IDF Analysis of Point Rainfall

The basic data used for intensity- duration-frequency analysis of point rainfall consists of the largest events of selected durations in each year (e.g., the largest 30 minute of rainfall of each year or the largest 6-hour of rainfall of each year), which is known to be the annual maximum series and is a sample of the population of all annual extremes at the measuring stations.

The analysis procedure consists of the estimation of quantiles for this time series following the method of probability distributions of extreme events. The analysis begins with a review of the history of the weather station to assure that measurement conditions have not changed significantly during the period of record.

Assuming that conditions have been stable, it needs to examine the rainfall records to determine the annual maximum rainfall records for each duration of interest for the period of record. In practice one would use the complete record of each rainfall from stations.

The next step is to compute the estimated quantile for each value. To determine the depths associated with the return periods of interest is usually done on a graph with depth or intensity plotted on a logarithmic or arithmetic scale (which ever gives a smoother and more nearly straight–line pattern) and exceedence probability on a probability scale (Dingman, 2002)

Different studies on IDF analysis have been made at different regions of the world. DuPont, B.S (2000) revised the rainfall intensity duration curves for the Common Wealth of Kentucky based on nine first order weather stations. The purpose of the study was to revise and update the existing rainfall intensity duration frequency (IDF) curves for the Common Wealth of Kentucky. Data was used from first order and cooperative weather stations. Four steps were followed in the process. Determining the area of influence, gather data from those areas, analyze the data: and produce the curves. Thiessen polygon and similar climatological zones are used to determine the areas of influence. The study result came up with extremely steep curves for short durations and as a result, linear regression was applied to the curves to produce usable values.

Grenney (2005) has also developed a comparative analysis of IDF curves at selected sites in Utah. In this study analyzing the magnitude of the Orographic effect of precipitation on IDF curve in four strategic regions and comparison of IDF curves from two different sources were done.

Lam, K.H (2004) has updated the short duration rainfall IDF curves for recent climates in Quebec, Canada. Ninety-five active stations equipped with tipping bucket rain gauges distributed throughout the province of Quebec were used for the study. A classic statistical method takes in to account & the GEV type I was applied to describe the frequency of the extreme rains. Values of maximum fallen rains for different laps of time on a daily base were evaluated and used to calculate firstly, series of annual maximum intensities and secondly, the IDF tables using the GEV/ Gamble method.

2.4. Components and selection of rainfall frequency analysis

The primary objective of frequency analysis is to relate the magnitude of extreme events to their frequency of occurrence through the use of probability distributions (Chow, V.T. 1988). Data observed over an extended period of time are analyzed in frequency analysis and are assumed to be independent and identically distributed.

In practice, the true probability distribution of the data at a site or a region is unknown. The assumption that data in a given system arise from simple parent distribution may be questionable when data from large watersheds are analyzed. In such cases more than one type of rainfall may contribute to extreme events in a region. However, for the analysis to be of practical use, simpler distributions are often used to characterize the relation between magnitudes and their frequencies (Rao and Hamed (2000). In general the chosen distribution should be (Cunnane, 1989) widely accepted, simple and convenient to apply, consistent, flexible, or robust, theoretically well based, or documented in the guide.

There are many distributions that have been suggested for AM series models and recommended by WMO. (Cunnane, 1989)

These are:

- i) Normal distribution (N)
- ii) Two parameter Lognormal distribution (LN2)
- iii) Three parameter Lognormal distribution (LN 3)
- iv) Exponential distribution (EXP)
- v) Two parameter Gamma distribution (G 2)
- vi) Pearson III distribution (P-III)
- vii) Log Pearson III distribution (LP-III)
- viii) Generalized Extreme value distribution (GEV)
- ix) Extreme value Type I distribution (EV1)
- x) Five parameters Wake by distribution (WAK 5)
- xi) Four parameters Wake by distribution (WAK 4)
- xii) Generalized Pareto distribution (GPAR)
- xiii) Log Logistic distribution (LLg)
- xiv) Generalized Logistic distribution (GLg)

The list and mathematical form of this distribution of presented in Appendix B.

2.5. Tests on Hydrologic Data

2.5.1. Test for independence and stationarity.

Given a sample of size N, the Wald- Wolfowitz (1943) (wwtest) test is used to test for the independence of a dataset and to test for the existence of trends in it. For a data set x_1, x_2, \ldots, x_N the statistic R is calculated from Equation 2.5

∑ − = = ⁺ + 1 1 1 1 ..2.5 *N i ⁱ ⁱ ^N R x x x x*

When the elements of the sample are independent, R follows a normal distribution with mean and variance given by equations 2.2 and 2.3,

 () ² ² . ⁶ 2 1 − 1 − = *N s s R*

 ...2.7 (1)(2) (4 4 2) 1 () 4 2 2 1 3 2 2 1 4 4 2 1 2 2 − − − + + − − + − − = − *N N s s s s s s s R N s s Var R*

Where $s_r = Nm_r$ and m_r is the rth moment of the sample about the origin. The statistic $u = (R - \bar{R})/(\text{var}(R))^{1/2}$ is approximately normally distributed with mean zero and variance unity and is used to test the hypothesis of independence at significance level α , by comparing the statistic u with the standard normal variate $u_{a/2}$ corresponding to a probability of exceedence $\alpha/2$. The program wwtest is used to analyze the data and when the value of statistic u is less than the critical level $u_{0.025}$ = 1.96. Thus we can accept the hypothesis of independence and stationarity.

2.5.2. Test for outliers

An outlier is an observation that deviates significantly from the bulk of the data, which may be due to errors in data collection, or recording, or due to natural causes. The presence of outliers in the data causes difficulties when fitting a distribution to the data. Low and high outliers are both possible and have different effects on the analysis.

The Grubbs and Beck (1972) test (G-B) may be used to detect outliers. In this test the quantities X_H and X_L are analyzed using the following equations.

X ^H = exp(*X* + *KNS*)...2.8 *X ^L* = exp(*X* − *KNS*)..2.9

Where : \overline{X} and S are the mean and standard deviations of the logarithm of the annual rainfall peaks, respectively, and K_n , is detected and K_n , is the G-B statistic tabulated for various sample sizes and significant levels by Grubbs and Beck(1972). At 10% significant level, the following approximation proposed by Pylon et al.(1985) is used, where N is the sample size.

$$
K_N = -3.62201 + 6.28446 N^{1/4} - 2.49835 N^{1/2} + 0.49146 N^{3/4} - 0.037911 N \dots (2.10)
$$

Sample values greater than x_H are considered to be high outliers, while those less than x_L are considered to be low outliers.

2.6. Selection and evaluation of parent distributions

2.6.1 Conventional moments.

Moment about the origin or about the mean are used to characterize probability distributions. For a distribution with a probability density function *f(x),* the *r th* moment about the origin is given by

x f () *x dx mean ^r r* ′ ⁼ ′ ⁼ ⁼ ∫ ∞ −∞ *m* , *m*¹ *m* …………………………. (2.11)

The Central moments m_r are computed by

 = (− ¹ ′) () , ¹ ⁼ ⁰ ∫ ∞ −∞ *m x m f x dx m r ^r* …………………………..(2.12)

Sample moments m_r , and m_r , on the other hand, are calculated as

 ⁼ ∑ *^x ^m*′ ⁼ *^X* ⁼ *n m ⁱ r n i r* 1 , 1 ' Sample mean ……………………………(2.13) (*X X*) *m o n m r i n i ^r* ⁼ ∑ [−] ⁼ = 1 1 , 1

These moments are often biased and may be corrected by (Cunnane, 1989)

$$
\hat{m}_2 = \frac{1}{N-1} \sum (X_i - \overline{X})^2
$$
\n
$$
\hat{m}_3 = \frac{N}{(N-1)(N-2)} \sum (X_i - \overline{X})^3 \qquad \qquad (2.14)
$$
\n
$$
\hat{m}_4 = \frac{N^2}{(N-1)(N-2)(N-3)} \sum (X_i - \overline{X})^4
$$

The conventional moment ratios are defined as;

The coefficient of variation,
$$
C_v = \frac{\hat{m}_2^{1/2}}{\hat{m}_1}
$$

The coefficient of skewness, $C_s = \frac{\hat{m}_3}{\hat{m}_2^{3/2}}$ (2.15)
The coefficient of kurtosis, $C_K = \frac{\hat{m}_4}{\hat{m}_2^2}$

2.6.2. Probability weighted moments

Probability weighted moments (PWM) are defined by Green wood et al. (1979) as

$$
M_{p,r,s} = E\Big(x^p F^r (1 - F)^s\Big) = \int\limits_{O}^{1} (x(F))^p F^r (1 - F)^s dF \cdots (2.16)
$$

In particular, the following two moments $M_{1,0,s}$ and $M_{1,r,0}$ are often considered

$$
M_{1,0,s} = a_s = \int_{0}^{1} x(F)(1-F)^s dF
$$

$$
M_{1,r,0} = b_r = \int_{0}^{1} x(F)F^r dF
$$
 (2.17)

Where: p, r, and s are real numbers

The plotting position estimates for sample PWM_S are given by

$$
a_{s} = \hat{a}_{s} = \hat{M}_{1,0,s} = \frac{1}{N} \sum_{i=1}^{N} (1 - F)^{s} x_{i}
$$

\n
$$
b_{r} = \hat{b}_{r} = \hat{M}_{1,r,0} = \frac{1}{N} \sum_{i=1}^{N} F_{i}^{r} x_{i}
$$
\n(2.18)

On the other hand, L – moments are defined by Hosking in terms of the PWM's *a* and *b* as

$$
I_{r+1} = (-1)^r \sum_{k=0}^r p^*_{r,k} a_k = \sum_{k=0}^r P^*_{r,k} b_k \quad \dots \tag{2.19}
$$

L- Moment ratios, which are analogous to conventional moment ratios, are defined by Hosking (1990) as

$$
t = l_2 / l,
$$

\n
$$
t_r = l_r / l_2, \quad r \ge 3
$$

Where: I_1 is a measure of location, t is a measure of scale and dispersion \qquad (L-Cv), t_3 is a measure of skewness (L-Cs) and t_4 is a measure of kurtosis (Lck). Sample L moment rations (t and t_r) are calculated by replacing l_r by their sample estimates L_r.

The first few L-moments are

$$
L_1 = M_{1,0,0}
$$

\n
$$
L_2 = M_{1,0,0} - 2M_{1,0,1}
$$

\n
$$
L_3 = M_{1,0,0} - 6M_{1,0,1} + 6M_{1,0,2}
$$

\n
$$
L_4 = M_{1,0,0} - 12M_{1,0,1} + 30M_{1,0,2} - 20M_{1,0,3}
$$
\n(2.20)

The L-moment ratio diagrams are based on the relations between the L-moment ratios. A diagram based on L-C_s (τ₃) versus L-C_k (τ₄) is used to identify appropriate distributions that best fits the rainfall data. For each station, the sample L-moment ratios t_3 and t_4 are plotted on the L-moment ratio diagrams. A suitable parent distribution is that which the average value of (t_3, t_4) gets close to it (Rao et. al. 2000).

The L-moment ratio diagrams are based on unbiased sample quantities in contrast to C_s and C_k which have to be corrected for bias. It was shown by Hosking (1990) that Cs and C_k values from several samples drawn from three different distributions lay close to a single line on the graph and overlaps each other offering little hope of identifying the population distribution. In contrast, the sample L-moment ratios plot as fairly well separated groups and permit better discrimination between the distributions.

2.6.3. Probability plots and Goodness- of –fit tests.

Probability plots are used to visually evaluate the agreement between distribution and observed data and also extremely useful for visually revealing the character of a data set. If the fitted distribution is the exact parent distribution, this relationship should appear as a straight line through the origin with a 45° slope. Plots are an effective way to see what the data looks like and to determine if fitted distributions appear consistent with the data. Analytical goodness to fit criteria are useful for gaining an appreciation for whether the lack of fit is likely to be due to sample to sample variability, or whether a particular departure of the data from a model is statistically significant. In most cases several distributions will provide statistically acceptable fits to the available data so that goodness of fit tests is unable to identify the "true" or "best" distribution to use. Such tests are valuable when they can demonstrate that some distributions appear inconsistent with the data (*Rao et. al*., *2000*).

The graphical evaluation of the adequacy of the fitted distribution is generally performed by plotting the observations so that they would fall approximately on a straight line if a postulated distribution were the true distribution from which the observations were drawn. This can be done with the use of special commercially available probability papers for some distributions.

- **Extreme Value Type I distribution**

An ordered observations X_i is plotted vs. the reduced variate Yi of the distribution

$$
Yi = -LN\left(-LN\left(\frac{T}{T-1}\right)\right)
$$
 (2.21)

The *Cunnane* Plotting position is applied with the relation.

$$
T = \frac{N + 0.20}{m - 0.40}
$$
 (2.22)

Where; i is the rank in ascending order=N-m+1

m: is the rank in descending order =N-i+1

N: is the number of observations

- **Pearson Type III Distribution**

An ordered observations X_i is plotted versus the Standard Normal Variate, u of the distribution

$$
u = W - \frac{C_o + C_1 W + C_2 W^2}{1 + d_1 W + d_2 W^2 + d_3 W^3} + e(p)
$$
 (2.23)

$$
W = \sqrt{-2\ln(p)} \text{ for } p < 0.5
$$
\nWhere: $p = 1 - F$

\n
$$
p = 1 - p \text{ for } p > 0.5
$$

PDF created with pdfFactory Pro trial version www.pdffactory.com

Co, C1, C2, d1, d2, and d3 are constants and $e(p)$ is the error term.

2.7. Parameter and quantile estimations

2.7.1. Introduction.

After a distribution or a number of distributions are selected to fit the data, their parameters must be estimated. The estimated parameters are used to calculated quantile estimates for different return periods or, conversely, to calculate the return period for a given flood magnitude. This is achieved by using the distribution function, in which the parameters of the distribution are replaced by their estimates and the relationship between return period (T) and probability of non-exceedence (F) in the form $F= 1-1/T$ is used. Different errors are associated with quantile estimates.

2.7.2. Parameter Estimation.

A number of methods that can be used for parameter estimation. These include the method of moments (MOM), the maximum likelihood method (MLM), the probability weighted moment method (PWM), the least squares method, maximum entropy, mixed moments Three of the more commonly used methods are considered here, namely: Method of ordinary moments (MOM), Method of maximum likelihood (ML), Method of probability weighted moments (PWM).

According to Rao, et. al. (2000) the maximum likelihood method is considered the most efficient method since it provides the smallest sampling variance of the estimated parameters, and hence of the estimated quantiles compared to the other methods. However, for some particular cases, such as the Pearson type III distribution, the optimality of the ML method is only asymptotic and small sample estimates may lead to estimates of inferior quality (Bobbie, et. al., 1991).

The method of moments (MOM) is a natural & relatively easy parameter estimation method. However, MOM estimates are usually inferior in quality and generally are not as efficient as the ML estimates, especially for distributions with large number of parameters (three or more), because higher order moments are more likely to be highly biased in relatively small samples.

The PWM method (Green wood et al., 1979; Hosking, 1986 a) gives parameter estimates comparable to the ML estimates, yet in some cases the estimation procedures are much less complicated and the computations are simpler.

2.7.3. Quantile Estimation

After the parameters of distribution are estimated, quantile estimates (X_T) which correspond to the different return periods may be computed. The relation between return period and the probability of non-expedience (F) is given by

$$
F = 1 - \frac{1}{T}
$$
 (2.24)

Where; F= $F(x_T)$ is the probability of having a flood of magnitude x_T or smaller. The problem thus reduces to evaluating X_T for a given value of F. Chow (1964) proposed a general for calculating X_T as follows.

$$
X_T = u_1' + K_T \sqrt{m_2}
$$
 (2.25)

Where; K_T is the frequency factor which is a function of the return period and of the parameters of the distribution, u'_1 *and* m_2 are the moments of the distribution

2.7.4. Standard error of estimate (SEE).

It is clear that a point estimate of a certain quantile corresponding to a return period may be of no real significance unless there is an indication of the accuracy of the estimate. A measure of the variability of the estimated value is the standard error of estimate S_T which is defined as (Cunnane, 1989)

$$
S_T = \sqrt{E\{\hat{X}_T - E(\hat{X}_T)\}^2}
$$
 (2.26)

The standard error of estimate accounts for the error due to small samples, but not the error due to the choice of inappropriate distribution. The standard error of estimate depends in general on the method of parameter estimation method (MOM, LM, WM), is that which gives the smallest standard error of estimate (Rao, 2000).

2.8. Regionalization.

2.8.1. Introduction.

Regional analysis is based on the concept of regional homogeneity which assumes that annual maximum flow population at several sites in a region are similar in statistical characteristics and are not dependent on catchment size (Cunnane, 1989).

Regionalization serves two purposes. For sites where data are not available, the analysis is based on regional data. For sites with available data, the joint use of the data measured at a site, called at site data, and regional data from a number of stations in a region provides sufficient information to enable a probability distribution to by use with greater rehabilitee.

2.8.2. Regional Homogeneity Tests

Hosking and Wallis (1991) give two statistics which are used to test regional homogeneity.

The *first statistic* is a discordancy measure, intended to identify those sites that are grossly discordant whit the group as whole. The discordance measure, D estimates how far a given site is from the center of a group. If $U_{\scriptscriptstyle i} = \left[t^{\scriptscriptstyle (i)}, t^{\scriptscriptstyle (i)}_3, t^{\scriptscriptstyle (i)}_4 \right]^T$ 4 (i) 3 $=$ $|t^{(i)}, t_3^{(i)}, t_4^{(i)}|$ is the vector containing the t, t_3 , t_4 values for site (i), then the group average for NS sites is given by

∑= = *NS i Ui NS U* 1 ¹ ………………………………….……………... …..……(2.26)

The sample covariance matrix is given by

$$
S = (NS - 1)^{-1} \sum_{i=1}^{NS} (U_i - \overline{U}) (U_i - \overline{U})^T \dots (2.27)
$$

The discordance measure is defined by

$$
D_i = \frac{1}{3} (U_i - \overline{U})^T S^{-1} (U_i - \overline{U}) \dots (2.28)
$$

A site (i) is declared to be unusual if D_i is large. A suitable criterion to classify a station as a discordant is that D_i should be greater than or equal to 3.

The *second statistic* is a heterogeneity measure intended to estimate the degree of heterogeneity in a group of sites and to asses whether they might reasonably be treated as homogenous. Specifically, the heterogeneity measure compares the between-site variations in sample L-moments for the group of sites with that expected for a homogenous region (Rao, et. al. 2000). Three measures of variability V_1 , V_2 and V_3 are available.

1. Based on LCv (t), the weighted standard deviation of (t) is given by

() ∑ () ∑ ⁼ ⁼ = − *NS i i NS i ⁱ V Nⁱ tt N* 1 1 2 1 / …………………………………………...………….(2.29)

Where: $\overline{\mathsf{NS}}$ is the number of sites, N_i is the record length at each site and \overline{t} is the average value of $t^{(i)}$

() ∑ ∑ ⁼ ⁼ = *NS i i NS i i t Nⁱ t N* 1 1 / ………………………………………………………….(2.30)

 $t^{(i)}$ =L-moment ratio at site i =L₂/L₁

2. Based on LCv and LCs, the weighted average distance from the site to the group weighted mean on a t vs t_3 graph is computed.

() / ..2.31 1 1/ 2 3 3 2 1 ² ∑ ∑⁼ − − ⁼ − = − *NS i i i i NS i V Nⁱ t t t t N*

3. Based on L-skewness (t₃) and L – kurtosis (t₄), the weighted average distance from the site to site to the group weighted mean on a t_3 vs t_4 graph is computed in eq. 2.32.

() / ..2.32 1 1/ 2 4 4 2 3 3 1 ³ ∑ ∑⁼ − − ⁼ − = − *NS i i i i NS i V Nⁱ t t t t N*
From the simulated data the mean m_{ν} *and* \mathbf{s}_{ν} the standard deviation of the N_{sim} value of *Vi* are determined. The heterogeneity measure is then defined

() *Hⁱ* = *Vⁱ* − *m^V s ^V* ……………………………………………………………….(2.33)

The region is declared to be *heterogeneous* if *Hi* is sufficiently large. Hosking, et. al. (1991b) sited in Rao, (2000) suggested that a region be regarded as acceptably homogenous if H_i is less than one, possibly heterogeneous if it is between 1 and 2 and definitely heterogeneous if *Hⁱ* is greater than 2.

2.8.3. Homogeneity tests

Different tests are available to examine regional homogeneity in terms of the hydrologic response of the stations in a region. Hosking et. al., (1991) gave a statistic which is used to test regional homogeneity. The statistic is a discordance measure, intended to identify those sites that are grossly discordant with the group as a whole. The discordance measure D, estimates how far a given site is from the center of the group.

Wiltshire, (1986a) developed a homogenous test based on the regional variability in the sites coefficient of variations $(C_v's)$ (Rao, et. al., 2000). Hosking, et. al., (1991) are also proposed a homogeneity test based on L- moments which provide to be efficient. For each site in a region calculate mean, standard deviation and coefficient of variation C_v

$$
\overline{R}_{i} = \sum_{j=1}^{n_{i}} R_{ij} / n_{i}
$$
\n(2.15)

$$
S_i = \sqrt{\frac{\sum_{j=1}^{i} (R_{ij} - \overline{R}_i)^2}{n_i - 1}}
$$
 (2.16)

Where: R_{ii} is the rainfall of station j in region i

 \overline{R}_i is the mean annual maximum rainfall for station i

 s_i is the standard deviation of R_{ij} for station i

 C_{vi} is the coefficient of variation of station i

The LCv can be calculated as:

 $LC_{vi} = L_{2i}/L_{1i}$ …………………………………………………………………………………(2.18)

Where L1 and L2 are as described in section 2.7.2

For each region, the CC value is calculated as:

$$
C\overline{V} = \sum_{i=1}^{N} CV_i / N
$$
 (2.19)

$$
S_{CV} = \sqrt{\sum_{i=1}^{N} (CV_i - \overline{C}_V)^2 / N}
$$
 (2.20)

$$
CC = S_{CV}/\overline{C}_V
$$
 (2.21)

The same procedure is followed for the corresponding L-moment values. The criteria for the region to be homogeneous is CC<0.3

2.8.4. Goodness of fit tests

Hosking, et. al. (1991) give a goodness of fit measure based on \bar{t}_r , the regional average of the sample L-kurtosis, mainly for three parameter distributions. Since all three parameter distributions fitted to the data will have the same t_3 on the LCs, vs LCk diagram, the quality of fit can be judged by the difference between regional average \bar{t}_4 and the value of t_4^{Dist} for the fitted distribution. The statistic () ⁴ ⁴ ⁴ *t* /*s Dist Dist Z* = *t*− …………………………………………………………...……….(2.22)

Where: $\boldsymbol{s}_\text{\tiny 4}$ is the standard deviation of $\bar{t}_\text{\tiny 4}$

Dist t is L-kurtosis value of the distribution

 $t₄$ is the average L-kurtosis value computed from the available stations data with in the region

() ∑ ∑ ⁼ ⁼ = *NS i i NS i i t nⁱ t n* 1 1 ⁴ ⁴ …………………………………………..………………..(2.23)

() () 0.5 1 2 4 2 4 4 1 4 1 ⁼ [−] ∑ [−] [−] = − *Nsim m sim ^m s Nsim tt N b* ……………….……....………….(2.24)

() () *sim N m m t t N sim* ∑= = − 1 *b*⁴ ⁴ ⁴ ……………………………………………….………..(2.25)

Nsim is the simulation of large number of regions

 b_4 is the bias in the regional average L-kurtosis for regions for the same number of sites and the same record lengths as the observed data.

A fit is adequate if Z^{Dist} is sufficiently close to zero, a reasonable criterion being $/Z^{Dist}/$ \leq 1.64. For small samples (N \leq 20) or large L-skewness (t_{3} \geq 4) a correction of \bar{t}_4 is required, that is instead of \bar{t}_4 , \bar{t}_4 - b_4 is used (Rao, 2000).

2.8.5. Forecast Accuracy

Forecast accuracy is a measure of the forecast error, that is, the difference between the amount forecasted, and the value that actually occurs. Forecast errors can be either systematic (recurring), or random. Forecast errors are best assed by retrospective comparison of forecasts actually made or that might have been made, and the values observed during the forecast period. Let I_c be the computed intensity and I_0 be the observed intensity during the same period and, \overline{I}_c & \overline{I}_o the mean of the computed and observed intensities for the same period.

∑= = *n j c c j I n I* 1 , 1 And ∑⁼ = *n i o o i I n I* 1 , ¹ ……………………………………………(2.26)

The following are widely used measures of forecast errors.

The square of the correlation coefficient between the observed and computed values,

$$
R^{2} = \left[\frac{\frac{1}{n}\sum_{i=1}^{n}I_{o,i}I_{c,i} - \bar{I}_{c}\bar{I}_{o}}{\left(\frac{1}{n}\sum_{i=1}^{n}I_{o}^{2} - \bar{I}_{o}^{2}\right)\left(\frac{1}{n}\sum_{i=1}^{n}I_{c}^{2} - \bar{I}_{c}^{2}\right)}\right]^{2}
$$
(2.31)

Mean square error, root mean square error, and forecast efficiency are all measures that incorporate both the systematic and random errors. Bias is a measure of systematic error while the variance is a measure of the variability, or scatter, of a number of forecasts about the true value, and is therefore, a measure of the random error (Maidment, 1992).

CHAPTER THREE: DATA SOURCE AND AVAILABILITY

3.1. AVAILABILITY OF DATA

The first step of the IDF relationship development consists in identifying all first class automatic recording stations which has sufficient length of record with in the region to retrieve intensities from the available charts. According to IGEBU (Burundi geographical Institution), the first class stations consists of both manual and automatic recording rain gauges, evaporation pan, screen Thermometer, Wind vane, Sunshine Hours and intensity recording and staffed with well trained personnel. From among the stations available nineteen stations are selected which have relatively better data length and are believed to represent the regions different climate characteristics. By the way, the data length of some stations which was less than ten years has been extended using the regression equation include the station Cankuzo (9 years), Kinyinya (8) and Makamba (7).The regression equations are established between two neighboring station. For example Cankuzo is near Ruyigi, Makamba near Nyanza and Kinyinya with Musasa (Figure 1.2). The study area includes: Ruvyironza, Gitega, Gisozi, Bujumbura, Ruyigi, Cankuzo, Musasa, Muyinga, Rwegura, Mparambo, Karuzi, Makamba, Nyanza-Lac stations.

……… table continued

3.2. Source of Data

The data used to develop the IDF relationships consisted of recorded rainfall charts of IGEBU from which maximum annual rainfall values for 0.5, 1,2,3,5,6,12 and 24 hours from the selected 19 stations with in the region. The data for indicated durations is directly read from daily recorded rainfall charts. The charts are traced by a float type gauge in which the rainfall collected by a funnel shaped collector is led in to a float chamber causing a float to rise. As the float rises, a pen attached to the float through a lever system records the elevation of the float on a rotating drum driven by a clock work mechanism. A siphon arrangement empties the float chamber when the float has reached a preset maximum level which in most cases is 10mm for the entire gauges. A typical weekly chart of date November. 28th, 1987 to november.28th, 1987 from Cankuzo station is shown in figure below.

Figure 3.1. The weekly rainfall chart from cankuzo station

This chart shows a rainfall depth maximum of 37.1 mm in 2hour the vertical lines in the pen trace correspond to the sudden emptying of the float chamber by siphon action which resets the pen to zero level. However, these rainfall charts are fairly available in the area.

3.3. Data collection

Rainfall data was collected from those charts starting from the time instant which provides the greatest reading for one hour duration after which the rest of the duration was read continuously from the chart. Table 3.2 shows the maximum depth of rainfall recorded for each of the durations of 30-minute, 1-hour, 2-hour, 3-hour, 5 hour, 6-hour, 12- hour and 24-hour occurred in the different months of the year 1987 at Cankuzo station.

		Observed	rainfall	(mm)	for	the	indicated	Duration(hr)	
Year	Data of record	0.5	1	$\overline{2}$	3	5	6	12	24
1987	24-Jan	13.9	23	23	23	23	23	23	23
1987	2-Feb	22	34	34	34	36	36	36	36
1987	3-Mar	6	10.5	14	15.6	21	21	21	21
1987	15-Mar	6.7	6.7	6.7	6.7	6.7	17	17	17
1987	6-Apr	5.8	8.1	8.1	8.5	13	14	14	14
1987	23-Apr	6	10.5	14	15.6	21	21	34	34
1987	1-May	4.7	4.7	4.7	4.7	8.7	9.8	8.7	8.7
1987	12-Sep	8.9	8.9	13.4	17	21	21	21	21
1987	15-Oct	24.9	24.9	24.9	24.9	24.9	24.9	36	36
1987	23-Oct	32	32	32	32	32	32	34	34
1987	19-Nov	23.7	23.7	23.7	32.9	45.6	45.6	45.6	45.6
1987	28-Nov	27	28.9	35.9	35.9	45.9	45.9	45.9	45.9
1987	7-Dec	26.3	26.3	26.3	26.3	26.3	26.3	26.3	26.3
1987	23-Dec	22	22.5	34.1	34.1	34.1	34.1	34.1	34.1
	Max	32	34	35.9	35.9	45.9	45.9	45.9	45.9

Table 3.2. Samples of data collected from rainfall charts for 1987 at Cankuzo station.

The annual maximum series model is used to determine the maximum of peak rainfall for each year of data for a specific station.

The derived annual maximum rainfall depths occurring in different durations for Cankuzo station are indicated in table 3.3. While for the rest of the stations is tabulated in appendix A

Table 3.3. Annual Maximum Rainfall depths in different durations for Gitega station

3.4. Testing for outliers

After collecting the data for each station, the outliers test was been checked. According to the equations 3.1; 3.2, and 3.3, all observations data are greater than upper (X_H) and therefore it is considered a high outlier. No low outliers were detected. The illustration is the Ruvyironza station and the results of the outlier test for different duration of rainfall depths are shown in table 3.4.

			Limiting Value		Data range		
Duration							
(h)	Mean	STEDV	Upper	_ower	Max	Min	
0.5	34.3	5.8	43.7	24.9	43.7	27.0	
	40.4	5.6	49.5	31.3	49.5	28.7	
$\overline{2}$	41.1	4.9	49.0	33.2	49.0	31.6	
3	43.6	4.5	50.8	36.3	50.8	36.0	
5	46.9	4.3	53.9	40.0	53.9	36.0	
6	49.0	1.7	51.7	46.2	51.7	46.5	
12	49.3	1.5	51.7	46.9	51.7	46.5	
24	49.8	1.5	52.2	47.4	52.2	47.8	

Table 3.4: Outlier test for Ruvyironza station

3.5. Tests for independence and Stationarity

It is usually assumed that all the peak magnitudes in the annual maximum (AM) series are mutually independent in the statistical sense. This assumption is usually justified.

The statistical analysis for dependence and stationarity is carried out for all the durations of rainfall record with in each station. A FORTRAN program is used for the analysis based on Wald – Wolfowitz (W –W) test and Lag – one serial correlation coefficient test is used for the analysis.

Accordingly to the result indicated in appendix C, illustrated in table 3.5 for Bujumbura station, the statistics value (u) are less than the critical value at 5 % significance level $u_{0.025}$ (=1.96). Thus we can accept the hypothesis of independence and stationarity. All stations are concluded to be independent and stationary at the 5% significance level.

Station					L1 correlation	Upper limit	Lower limit	
			Critical test statistics					
				Remark				Remark
					coefficient			
	Duration(h)	Statistic						
	0.5							
			1.96	Independent				
		0.05			-0.32107	0.44633	-0.62815	Random
	1		1.96	Independent				
		0.054			-0.32522	0.44633	-0.62815	Random
	$\overline{2}$		1.96	Independent				
		0.093			-0.32331	0.44633	-0.62815	Random
	3		1.96	Independent				
		0.023			-0.28713	0.44633	-0.62815	Random
	5		1.96	Independent				
		0.190			-0.44534	0.44633	-0.62815	Random
	6		1.96	Independent				
		0.109			-0.41743	0.44633	-0.62815	Random
	12		1.96	Independent				
		0.539			0.12166	0.44633	$-0.62815-$	Random
	24		1.96	Independent				
Bujumbura		0.071			-0.50748	0.44633	-0.62815	Random

Table 3.5. : Test of independent and stationarity for Bujumbura station

CHAPTER FOUR: ANALYSIS, RESULTS AND DISCUSION

4.1. Analysis procedure

Fitting an appropriate probability distribution involves three steps namely (i) the selection of a distribution, (ii) testing its goodness of fit to the observed data and (iii) the estimation of its parameters, quantiles and Standard error of estimate.

One may be tempted to conclude that a proper procedure for selection of a distribution could be to consider a wide variety of distribution functions that are described in section 2.6, estimate their parameters using the testing procedure in section 2.7.

4.2. Selection and Evaluation of best fitted statistical parent distribution of

Rainfall data.

4.2.1. L-Moment Ratio Diagrams Method

The identification of a parent distribution can be achieved much more easily by using L-moment ratio diagrams described in section 2.6.2. Figure 4.1 and table 4.1 indicates the graph of L-MRD and the candidate distributions for different durations of annual maximum rainfall respectively at Muyinga station. The same analysis could be applied to fit the best distribution to each station data.

In figure 4.1, it is observed from the L-MRD that the muyinga station cluster around GP &WLB/P3 and therefore GP distribution may be expected to give the best regional fit for the data.

Duration				
(h)		t_3	t4	Candidate Distribution
0.5	0.067	0.07	0.006	GP/WLB/GPIII
1	0.057	-0.008	0.069	GEV/L2)(3)/G&PIII
2	0.05	-0.011	0.237	GL/L(2)(3)/GEV
3	0.048	-0.394	0.271	GL/L(2)(3)/GEV
5	0.044	-0.534	0.417	GL/L(2)(3)/GEV
6	0.019	-0.052	0.157	GL/L(2)(3)/GEV
12	0.012	0.103	0.003	GP/WLB/G&PIII
24	0.018	0.263	0.246	GL/L(2)(3)/GEV

Table 4.1.candidate distributions based on the L-MRD for Muyinga station

4.2.2. Probability plots and Goodness- of –fit tests.

Probability plots are used to visually evaluate the agreement between distribution and observed data and also extremely useful for visually revealing the character of a data set. The graphical evaluation of the adequacy of the fitted distribution is generally performed by plotting the observations so that they would fall approximately on a straight line if a postulated distribution were the true distribution from which the observations were drawn. This can be done with the use of special commercially available probability papers for some distributions. Accordingly to the equations in section 2.6 (2.21, 2.22, 2.23), the following two distributions are compared for their fitness for two hour and Twenty-four hour annual maximum rainfall . The results are illustrated as following for Ruvyironza station graphically.

From the relative comparison of t (coefficient of determination) the two figures indicated above, it is observed:

Ø For each distribution, the graphs show the best fit for 24 hours of annual maximum rainfall if we consider the coefficient of determination shown into the graphs.

Ø For each duration, Pearson III gives the best fit for 24 hours durations. Otherwise, it is observed the EVI best fitted the annual maximum rainfall of 2 hours .

4.2.3. Standard error of estimate (SEE)

Parameters of the best fitted distributions are estimated based on the methods described in section 2.7. The standard error of estimate (SEE) is given by Eq. 2.7.4, for annual maximum rainfall of 6-hour durations in the different return periods for Bujumbura station are shown in tables (4.2).

Table 4.2 Standard Error of Estimate of the Candidate distributions for 6-hour rainfall at Bujumbura station

Distributions	$\overline{2}$	5	10	25	50	100	Average	REMARK
EV1/MOM	2.02	2.1	2.55	3.3	3.92	4.57	3.08	
EV1/ML	2.1	2.95	3.65	4.66	5.47	6.29	4.19	
EV1/PWM	2.02	2.17	2.66	3.45	4.1	4.77	3.20	
LN/MOM	2.01	2.32	2.82	3.55	4.14	4.74	3.26	
P3/MOM	2.14	2.07	2.39	2.95	3.43	3.92	2.82	
P3/PWM	3.33	1.77	1.84	2.52	3.26	4.07	2.80	
LP3/MOM	2.27	2.13	2.53	3.29	3.96	4.66	3.14	
G2/MOM	1.99	2.11	2.47	3	3.42	3.84	2.81	
G2/ML	1.99	2.11	2.47	3	3.42	3.84	2.81	
G2/PWM	$\overline{2}$	2.13	2.51	3.06	3.51	3.95	2.86	
GEV/PWM	2.19	2.14	2.51	3.27	4	4.81	3.15	
GEV/MOM	2.17	2.11	2.41	2.91	3.32	3.72	2.77	Min SEE
LLG/PWM	2.16	2.09	2.48	3.4	4.38	5.61	3.35	
EXP/MOM	2.11	2.1	2.68	3.66	4.48	5.33	3.39	

The best candidate distribution for 6 hour rainfall at Bujumbura is GEV for the MOM method because, it have the small value of SEE greater than others distributions and by either the PWM or the ML method.

The best fitted candidate distributions of different rainfall duration's for all stations are shown in table 4.3.

Table 4.3 Best Fitted Distributions for the indicated durations depending on the smallest Standard Error

PDF created with pdfFactory Pro trial version www.pdffactory.com

4.3. Estimate Quantiles

Based on the selected distributions the estimated quantiles for different rainfall durations at Bujumbura station is shown in table 4.4. The estimated quantiles for the rest of stations are tabulated in appendix E.

Return Period		Estimated Quantiles for the indicated durations of rainfall (mm) Bujumbura station						
(years)	0.5 _{hr}	1hr	2hr	3hr	5hr	6hr	12hr	24hr
2	25.67	35.57	40.4	43.68	44.98	45.89	49.8	53.14
5	32.61	42.4	47.6	50.4	51.98	52.12	54.6	57.5
10	34.7	44.14	53.71	57.9	58.92	59.67	61.12	63.45
25	37.1	45.1	55.56	59.96	61.39	63.67	65.56	68.43
50	38.77	45.54	57.8	61	62.88	65.53	69.78	72.12
100	40.31	46.89	59.78	63.4	64.11	66.23	71.12	74.35

Table 4.4. Estimated Quantiles for Bujumbura station

Figure 4. 3: Graph showing estimated quantiles versus durations

The quantile estimated increase when the duration increase for all returns period and increase when the return periods increase also (Figure 4.4 and table 4.4).

4.4. Computation of rainfall Intensity

4.4.1. Classical method.

The intensity of rainfall, i is calculated based on the relation

i T D X Duration of rain fall mi nutes $i = \frac{Rainfall \ depth (mm)}{}}$ (minutes) (mm)

Illustrative Example: computation of rainfall Intensity

Required*:* To find the intensity of rainfall for 1 hour duration and 50 years return period for Bujumbura area.

Solution: From the estimated quantiles for the 50 year frequency and 6-hour duration (Table 4.4) rainfall depth, $X_T = 65.53$ mm and the intensity of rainfall, is: i= 65.53mm/6-hr= 10.9 mm/hr. Table. 4.5. Shows the intensity for different durations and frequencies of rainfall for Bujumbura station.

Table. 4.5. Intensity of rainfall for different durations and frequencies for Bujumbura station

Duration		Intensity of rainfall for the indicated durations, mm/hr-Bujumbura station				
(minutes)		5	10	25	50	100
0.5	51.3	65.2	69.4	74.2	77.5	80.6
	35.6	42.4	44.1	45.1	45.5	46.9
$\mathbf{2}$	20.2	23.8	26.9	27.8	28.9	29.9
3	14.6	16.8	19.3	20.0	20.3	21.1
5	9.0	10.4	11.8	12.3	12.6	12.8
6	7.6	8.7	9.9	10.6	10.9	11.0
$12 \,$	4.2	4.6	5.1	5.5	5.8	5.9
24	2.2	2.4	2.6	2.9	3.0	3.1

For each return period, the intensity decrease when the duration of rain increase and for each duration, the intensity of rainfall increase when the return period increase.

4.4.2. Estimation of the IDF Parameters

The IDF-Curve Fit Software (version 2.07) is employed to solve the parameters of the IDF equation 2.3 discussions in section 2. The IDF Curve Fit tool manipulates data describing an Intensity-Duration-Frequency relates for a particular geographical locality and can be used in two modes:

- 1. To compute the 'A', 'B' and 'C' parameters that most closely approximates a set of observed rainfall data.
- 2. To compute the IDF curve for user-supplied values of the three coefficients

and compare this with observed data.

For any time interval the rainfall can be defined either as a total depth of rainfall or as an average intensity over the time interval. Table 4.6 shows the computed parameters A, B, C of the IDF of various frequencies for some stations.

		$T=2$ Years			T=5 Years			T=10 Years			T=25 Years			T=50 Years		T=100 Years		
Station Name	A	B	C	Α	B	C	A	B	C	A	B	C	A	B	C	Α	B	C
Gisozi	1298.55	20.21	0.8879	2091.2	15.098	0.8682	2299.3	15.325	0.8752	2588	15.92	0.885	2768.8	15.918	0.89'	2982	16.328	0.8975
Tora	1089.06	4.146	0.7954	1656.1	8.597	0.8336	2121.2	7.472	0.8614	4124	7.472	0.951	5309.	19.944	0.982	5576	23.202	0.9825
Ruvyironza	639.09	0.014	0.7565	931.24	2	0.7819	1106.7	4.472	0.7998	1347	7.472	0.821	1615.2	10.91	0.842	1963	15.09	0.8644
Makamba	2763.73	29.04	0.9702	2849.8	17.474	0.9463	3120.6	15.09	0.9556	4464	18.8	1.003	6282.7	23.27	1.047	3862	11.395	0.9636
nyanza lac	2521.03	23.22	0.9636	2491.2	16.328	0.9353	2883.8	17.18	0.9533	3998	23.27	0.99	3632.2	19.944	0.974	3649	20.618	0.9625
Musasa	1363.42	14.25	0.9007	2339.3	16.85°	0.9379	2941.2	19.944	0.9595	3438	20.56	0.96	4135	23.218	0.988	4486	23.27	0.9918
kinyinya	2243.09	9.213	0.9503	3498.4	18.798	0.97	4310	23.18	0.9901	5355	30.01	1.004	6525.1	35.854	1.023	8085	42.416	.0445
Ruyigi	983.06	9.238	0.8574	1564.8	6.703	0.8924	1897.7	7.472	0.9123	2439	10.09	0.936	2709	10.103	0.946	356'	16.328	0.9686
Cankuzo	2514.77	11.79	0.9662	2988.3	10.09	0.9704	2974.3	9.236	0.9609	3065	7.472	0.966	2807.5	4.471	0.951	2794	3.202	0.9493
Muyinga	529.75	0.036	0.7241	946.68	4.472	0.781	1117	5.597	0.7979	1408	7.472	0.82.42	1798.2	7.472	0.856	2372	12.692	0.8904
Kirundo	946.21	0.528	0.8223	1660.2	7.472	0.8759	2000.4	10.09	0.8953	2446	12.76	0.916	2771.5	4.236	0.929	3201	16.326	0.9444
Nyamuswaga	1994.6	4.472	0.9493	2003.1	0.528	0.9083	2080	0.528	0.9018	2244	0.528	0.899	2279.8	0.148	0.892	2338	0.055	0.8863
Karuzi	3497.37	37.51	1.0326	3755.5	28.035	1.0253	3797.8	25.797	1.0217	3867	23.27	1.02	3661	19.944	1.008	3545	17.966	0.9996
Rwegura	1298.89	13.42	0.8217	1781.7	14.252	0.8376	1949.7	14.252	0.8425	2191	15.098	0.847	2356.6	16.326	0.856	2468	16.326	0.8559
Teza	977	7.472	0.7976	1386	10.093	0.8331	1718.3	13.467	0.8608	2171	16.33	0.893	2358.1	16.944	0.901	1654	10.09	0.6114
Bujumbura	2238.68	21.33	0.9516	3007.9	18.798	0.9817	4134.7	27.563	1.0103	3545	23.22	0.978	2929.9	17.952	0.943	2969	16.851	0.9421
Rumonge	2156.01	10.91	0.957	4419.3	23.27	0.9992	3366.5	19.125	0.9307	5532	30.01	0.991	5674.5	29.035	0.987	6431	300.03	1.006
Mparambo	2045.02	20.54	0.9477	3023.1	23.215	0.9752	3955.9	25.888	1.0009	4383	25.79	1.013	3445.5	19.944	0.969	3339	18.798	0.9562
Gitega	1395.39	6	0.8999	2020.3	9.215	0.9163	2256.	10.652	0.9223	2533	12.1	0.928	2760.5	13.418	0.933	3219	16.326	0.9488

Table 4.6. Summary of the Estimated IDF Parameters of some stations for the indicated frequencies

After determining the numerical value of the IDF parameters, rainfall intensity for any duration and recurrence interval can be determined with the equation 2.4.The equation is developed to calculate the intensities for all durations at each station. There is a different equation for different return period at each station. The resulting six equations for each station can be used for intensity calculations in the area represented by that station. Listed below are the six equations for the IDF relationships for Gisozi station.

2 Year return period,
$$
i = \exp[\ln(1298.55) - 0.8879 * \ln(20.21 + D)]
$$

5 Year return period, *i* = exp[ln(2091.21) − 0.8682*ln(15.098+ *D*)]

10 Year return period, *i* = exp[ln(2299.34) − 0.8752*ln(15.325+ *D*)]

25 Year return period, *i* = exp[ln(2588.48) − 0.8854*ln(15.923+ *D*)]

50 year return period, *i* = exp[ln(2768.16) − 0.8906*ln(15.918+ *D*)]

100 Year return period, *i* = exp[ln(2982.39) − 0.8975*ln(16.328+ *D*)]

Illustrative Example: Calculation of rainfall intensity using the IDF relationships

Required: To find the intensity of rainfall for 100 years return period and duration of 1440 minutes for Gisozi station.

Solution: 100 Year return period, $i = \exp[\ln(2982.39) - 0.8975 * \ln(16.328 + D)]$

$$
i = \exp[\ln(2982.39) - 0.8975 * \ln(16.328 + D)] = 4.3
$$
mm/hr

Intensity Values generated from those equations at Gisozi station for different return periods are listed in table 4.7 below. The Intensities for the rest of stations are tabulated in appendix F.

Return	Computed Intensity of rainfall(mm/hr) for the indicated durations -												
Period	Gisozi Station												
(years)	0.5 _{hr}	1hr	2 hrs	3hrs	5hrs	6hrs	12hrs	24hrs					
$\overline{2}$	40.114	26.465	16.118	11.748	7.7424	6.6474	3.6792	2.0127					
5	76.605	49.201	29.551	21.478	14.166	12.177	6.7895	3.753					
10	81.654	52.349	31.349	22.737	14.952	12.838	7.1262	3.9209					
25	87.394	55.997	33.437	24.19	15.846	13.585	7.4947	4.0964					
50	91.648	58.567	34.865	25.175	16.45	14.09	7.7462	4.2189					
100	95.383	60.933	36.205	26.098	17.009	14.554	7.9683	4.3206					

Table 4.7: Estimated Intensity values for Gisozi station

4.4.3 Evaluation of the method of parameter estimation

4.4.3.1 Graphical/Visual verification

The graphical evaluation of the goodness of fit is performed by plotting the observed versus the computed intensities of rainfall. The result of the graph indicated that, the plot fall approximately on a straight line and the efficiency (R^2) is approaching to 100% for all of the frequency of rainfall.

The percentage difference between computed and observed intensities is plotted versus duration of rainfall for different return periods. Figure 4.4 shows the graphical comparisons of the computed and observed intensities with the percentage difference of estimate from the observed value and the comparison of rainfall depths for the same area.

Figure 4.4 Comparison of observed versus computed rainfall depths for 2 and 10 years recurrence interval & different durations of Bujumbura station.

The graph of percentage difference indicates that relatively higher difference between computed and observed intensity showed at lower rainfall durations, especially less than 120 minutes. The general percentage difference increases the

frequency increases and less than 1.5% for 2 years return period and less than 2.5% for 100 years return period of Bujumbura station. In general, from the graphs of observed versus estimated values of intensities and rainfall depths with their percentage difference, it can be concluded that the estimated values using parameters describe the observed values.

4.4.3.2. Forecast Accuracy

Forecast accuracy is a measure of the forecast error, that is, the difference between the amount forecasted, and the value that actually occurs (section 2.8.5).

Table 4.8. Comparison of observed and computed rainfall intensity for 25 years return period at Gisozi station

The Mean Square Error, $MSE = \frac{1}{n} \sum_{i,j} [I_{c,i} - I_{o,i}]^2$

$$
E = \frac{1}{n} \sum_{i=1}^{n} \left[I_{c,i} - I_{o,i} \right]^2 = 1/8^* 6 = 0.75
$$

-Root Mean Square Error, $RMSE = MSE^{0.5} = (0.75)$ $\sqrt{0.5} = 0.86$

-Bias,
$$
B = \overline{I}_c - \overline{I}_o = 28.36-28.37 = -0.01
$$

-Variance,
$$
V = MSE - B^2 = 0.75 - (-0.01)^2 = 0.75
$$

From the values calculated above, all the measure of accuracy is with in the

limit value. The symmetric error (bias) which is the measure of the degree to which the estimation is consistently above or below the actual value, is too small in this case. The variance which is the measure of the random error is also small. Therefore it can be concluded that the estimated intensity described the observed value**.**

4.4.3.3. Sensitivity of the IDF parameters

Sensitivity of the IDF on intensity of rainfall was done by increasing the parameters by 10% and computing the intensity of rainfall with increased parameters. Comparison between the intensity of rainfall obtained from the optimized IDF parameters and the other from increased parameters is made.

Figure 4.5. Results of the sensitivity test on the IDF parameters of Bujumbura station for the indicated frequencies.

From the test result, most sensitive parameter is found to be the "C" exponent. An increase in "C" exponent by 10% resulted in a difference of more than 33% between the two values of rainfall intensity mostly for larger return periods and short durations. Increasing the "A" coefficient by 10% resulted in increased intensity by approximately 8% which indicated that the rate of increase or decrease in "A" coefficient results in

the same rate of increase or decrease of the intensity of rainfall, respectively. An increase in the "B" constant has resulted in insignificant decrease on the intensity of rainfall mainly for larger return periods. The parameter 'c' influences significantly the result of intensity and the 'C' exponent depends on the relative increase or decrease of the 'A' coefficient.

4.5. Construction of the IDF curves

The IDF curves were plotted on a log-log graph the duration D as abscissa and The intensity I as ordinate. Figure 4.6 and 4.7 shows the IDF curves plotted on Double logarithmic papers and normal papers, respectively for Gisozi station. The rest of the IDF curves for the rest of the stations are compiled in appendix C

Figure 4.6. IDF curves on double logarithmic scale for Gisozi station

4.6. Construction of the IDF maps

IDF maps are drawn for each station based on some frequency and duration to show the spatial distribution of the intensity of rainfall with in the region. Arc View GIS is used for this analysis. Figure 4.8 shows the constructed IDF maps for 12-hour 10 years rainfall intensity map covering the study area. The rest of the IDF maps for different durations are compiled in Appendix D.

IDF maps help to interpolate intensities for areas where there has no intensity data. It is also possible to interpolate rainfall intensities for various rainfall durations and frequencies by making use of these maps.

Figure 4.8: IDF maps for 12 hours and 10 years return period rainfall intensities.

CHAPTER FIVE : REGIONALIZATION OF THE ANNUAL MAXIMUM RAINFALL DEPTH OF 24-HOUR DURATION WITH IN BURUNDI

5.1. INTRODUCTION

Regionalization in the case of this study refers to the grouping of homogenous region that contain stations having similar climatic characteristics.

The primary identification of homogenous regions was done by using L-MRD. The sample L-moment ratios L-Cs and L-Ck for each station based on specific duration data as well as their regional averages are plotted on L-moment ratio diagrams. It is assumed that (LCs, LCk) values of one station varies linearly with (LCs, LCk) values of the neighboring station. A suitable parent distribution is that which averages the scattered data and around which the data spread consistently and considered as the same region. The delineation result indicated that *four regions* were delineated. On the digitized map of the region, (on Arc View GIS software) the distance between one station and its neighboring station was determined and (LCs, LCk) values were interpolated to fix the boundary between two stations of different regions. Two boundaries are fixed, one from the LCs and the other from the LCk values. The final boundary between regions is fixed between the mid ways of the two boundaries.

5.2. Results of Regionalization

5.2.1. IDENTIFICATION OF REGIONS

Regionalization was made on the statistical values (LCs, LCk) of maximum rainfall of the selected duration for each station based on the concept that stations from the same region, their data series come from the same parent distribution.

The LCs-LCk and Cs-Ck moment ratio diagrams for durations of twenty four hour data are shown in figure 5.1 with various distributions. The best fitted theoretical probability distributions according to their priority of closeness are shown in table 5.1 and those distributions based on which the primary classification of the regions are made and common for (L-Cs,L-Ck) are shown in table 5.1.

Figure 5.1: L -MRD of 24-hours rainfall depth

Table 5.1 illustrates the prioritization of distribution based on closeness to stations on L-MRD method.

Table 5.1 Prioritized distribution based on closeness to stations on L-MRD

Station	Distribution	Station	Distribution	Station	Distribution	Station	Distribution
Region 1	G&PIII	Region2	GEV	Region3		Region4	GP
Bujumbura		Rwegura		Muyinga	G&PIII	Karuzi	
Mparambo		Tora		Kirundo		Cankuzo	
Rumonge		Gisozi		Nyamuswaga		kinyinya	
Nyanza		Ruvyironza				Ruyigi	
		Teza					
		Gitega					
		Makamba					

Table 5.2. Classified regions based on closeness to distributions

The table 5.2 shows the result of identification of region based on closeness to distributions from the figure 5.1. Four regions are graphically identified. For example region one is visible with the red symbol and is near the G&PIII distribution.

5.2.2. Discordance test

Discordance test is done based on the methods described in section 5.4 and the result of the test is shown in table 5.3.

A suitable criterion to classify a station as a discordant is that D should be greater than or equal to 3 so on, all discordance measure are less than 3.

5.2.3. Cv- based homogeneity test

A FORTRAN program is used for this test. The program is developed based on the method described in sections 2.8.2 and 2.8.3. From the test result, all stations with in a region satisfy homogeneity criteria for L-moment C_v -based homogeneity tests. One station (Musasa) is found to be heterogeneous to the rest of the stations and do not satisfy the criteria for homogeneity test to lie in the regions. Therefore this station is considered as a discordant region. Table 5.4 and table 5.5 show the summarized result of this test.

Table 5.4. The CC values for the delineated regions

The CC values are less than 0.3 (eq. 2.21, section 2.8.3) for that four regions are homogeneous.

Figure 5.2. Established Homogeneous Regions

5.3. Selection of best fitted distributions for the delineated regions

Figure 5.3 L-MRD of the average values of L-Cs and L-Ck for the delineated regions.

Based on the L-MRD the average L-moments candidate distributions for the regions delineated for the three durations of annual maximum rainfall is shown in table 5.6. From the candidate distributions the best fitted distribution is selected using the goodness of fit measure Z ^{Dist}. A fit is adequate if Z^{Dist} is sufficiently close to zero, a reasonable criterion being $/Z^{Dist}/\leq 1.64$ and the best fitted distribution result are shown in table 5.6.

5.4. REGIONAL QUANTILES

The quantiles for the classified regions are estimated based on the selected best fit distribution described in table 2.6.The quantiles for each station grouped in a region are estimated using the regional best fitted distribution by the methods described in section 2.8.4. The estimated quantiles are then pooled together to calculate the mean of those stations with in the region for each return period and duration. Then these mean quantiles are used for the estimation of the regional parameters for the specified region. The pooled regional quantiles for 0.5,1,2,3,5,6,12 and 24hr durations I shown in appendix G and table 5.7 show the regional quantile for the region 1.

Table 5.7. Regional Quantiles for region 1

5.5. Regional IDF Parameters

The IDF parameters for each classified regions with in the study area are estimated based on the following equation $I = \frac{aI}{(D+B)^c}$ *m D B* $I = \frac{aT}{a}$ + $=\frac{u}{\sqrt{2}}$. Estimated parameters of the IDF shown in table 5.8.

Table 5.8. Estimated IDF Parameters of some regions.

5.6. Regional IDF curves

IDF curves are constructed for the classified regions based on the regional intensity on a double logarithmic scale using the IDF curve fit tool as in figure 5.5. These curves can be used for intensity determinations for ungauged areas with in the regions except for the limitations described in the previous sections.

Figure 5.5. **IDF** curves for region two

Figure 5.6. IDF curves for region three

Figure 5. 7. IDF curves for region Four
CHAPTER S IX : CONCLUSION AND RECOMMENDATION

6.1. Summary

The establishment of Intensity-Duration-Frequency (IDF) curves for precipitation remains a powerful tool in the risk analysis of natural hazards. Indeed the IDF-curves allow for the estimation of the return period of an observed rainfall event or conversely of the rainfall amount corresponding to a given return period for different aggregation times. There is a high need for IDF-curves in Burundi. The present paper assesses IDF-curves for precipitation for nineteen stations selected. The IDF-curves for Burundi are an interesting tool to be used in sewer system design to combat the frequently occurring inundations in semi-urbanized and urbanized areas.

The purpose of this study is mainly to produce IDF-curves for precipitation for nineteen different climatological stations. These stations are respectively: Ruvyironza, Muyinga, Gisozi, musasa, Ruyigi, Rwegura, Karuzi, Gitega, Mparambo, Bujumbura, Cankuzo, Nyanza-lac, Makamba, Kirundo, Kinyinya, Tora, Rumonge, Teza and Nyamuswaga.

The methodology used for the analysis of IDF curves involve the following steps : Identification of data series, tests of the data, identification of theoretical parent probability distribution, estimation of distribution, selection of distribution, estimation quantile distribution, analysis of IDF parameters, computation of intensity and construction of IDF curves and maps. Respectively for the selected stations.

From the available charts of the IGEBU, the rainfall depths for the durations of 30 minutes, 1, 2, 3, 5, 6, 12 and 24 hours were collected. The annual maximum series model was employed to select the maximum annual rainfall values from records of each year of the durations. The selected annual maximum rainfall data were checked outliers and independence. The data that showed outliers were discarded from the IDF analysis. After checking the quality of data, parent probability distributions for the annual maximum rainfall depths for all the durations have been selected based on two commonly used methods: L-moment ratio diagrams and minimum standard error of estimate. A FORTRAN program is applied to estimate the parameters and quantiles based on different probability distributions. Estimated Quantiles are selected based on the most robust and best fitted probability distribution to the annual maximum rainfall depths.

The IDF parameters are estimated and the IDF curves are constructed for the nineteen stations using the IDF curve Fit tool (Miduss software). The intensities are computed based on the estimated IDF parameters for any durations and recurrence interval of each station. The adequacy of the computed intensities are evaluated by making use of graphical/visual verification and the goodness of fit tests using statistical method between observed and computed intensities and the result of these tests indicated that the estimated value of IDF parameters and the computed rainfall intensity are adequate for all stations.

The IDF maps are developed using ARC-VIEW GIS soft ware for each station based on some frequencies and durations to show the spatial distribution of the rainfall intensity with in the region and can be used to interpolate intensities for areas where there is no rainfall intensity data.

Moreover, regionalization has been done based on annual maximum rainfall depth of 24-hour durations and the best fitted distributions for each homogenous region which were identified. Regionalization was made based on the statistical values of (LCs and LCk) of annual maximum rainfall depths of the selected durations for all stations. Four different groups of stations satisfying the homogeneity tests were identified; one station (Musasa) became heterogeneous to all delineated regions.

The best fitted probability distributions for the regions are identified after which parameters and quantiles are estimated and pooled together to obtain their mean for the IDF analysis with in the region. The estimated regional IDF parameters are evaluated for the adequacy of representing the specified regions.

6.2. CONCLUSION.

Regional and at-site IDF relationships have been developed for the study area.

It was shown by applying different statistical tests that the annual maximum values of rainfall for nineteen stations, outliers were discarded from the IDF analysis and confirm the independence of the data used. Generally three different methods of computing to obtain the intensity of rainfall have been developed for this work.

The first method is the general IDF mathematical form that relates the intensity (I), durations and frequency of rainfall, developed in the form of $I = (A/(D+B)^C)$ for each station in the study area based on the optimum IDF parameters(A,B,C) estimated. In general, the value of 'A" coefficient increase with an increase in return period for most of the stations considered. The parameter 'B' is constant on the other hand the 'C' exponent depend on the relative increase or decrease of the 'A' coefficient. For most of the return periods these two parameters increases with an increase of 'A' coefficient and vice versa.

, The *second method* is plotting IDF curves. It is plotted on a double logarithmic scale with intensity as an abscissa and duration as ordinate. The best fit IDF curves are developed based on the optimum parameters for each station by making use of the IDF curve fit tool in which all curves show similar shape.

The second method is plotting IDF maps. Generally, the above methods are developed to estimate intensity for a certain return period and duration for areas with in 25 km radius from the principal stations.

To extract the intensity of rainfall of any duration and frequency at areas farthest from the principal stations, Burundi has been regionalized and the regional IDF curves. From the established four regions and one heterogeneous station, the value of regional intensity under region four showed some divergence from the intensity values of some stations with in the region for shorter durations. For such cases, the use of IDF relationships developed for the region jointly with the IDF maps gives better results. Therefore, planners and designers in Burundi can effectively utilize one or all of the procedures to derive the IDF value in any part of the Country.

6.3. Recommendations

The following recommendations are based on the results obtained:

- 1. The IDF relationships developed for all stations can be used for areas close to the principal stations with in 25km radius according to the WMO guide line where as for areas farthest from those stations the regional IDF relationships developed can give better results for intensity determinations with in the Burundi country.
- 2. The developed IDF relationships for whole country can be used to extract the intensity or depth of rainfall of a specific duration for practical applications water resources engineering activities with in the study area.
- 3. The availability of first class recording stations in the region is limited. Some of the existing stations are intermittently recording or stopped recording. To process basic information which is related to water resources developments it highly essential if additional recording stations are established and get improved the functionality of the existing stations.
- 4. The development of the IDF relationships is a powerful tool for practical purposes of water resource developments. Therefore it is advisable to develop the standardized comprehensive IDF relationships for the country in general**.**

BIBLIOGRAPHY

- 1. Chow, V.T. (1988). 'Applied hydrology'. McGraw-Hill. New York, NY
- *2.* Cunnane, C. (1989).'Statistical Distributions for Flood Frequency Analysis', *Word Meteorological Organization Operational Hydrology*, report No.33, WMO-No. 718, Geneva, Switzerland.
- 3. Dingman, L.S (2002) .'Physical Hydrology', 2nd edition. Prentice–Hall. Inc. New **Jersey**
- 4. DuPont, B.S& Allen, D.L. (2000). 'Revision of the rainfall Intensity curves for the Common Wealth of Kentucky'. Research Report. Kentucky.
- 5. Green W.S (1986), 'Time series Analysis. In Advances in hydrosciences vol 5', edited by V.T. Chow, Mc-Graw Hill, New York.
- 6. Grenney, W,J, Nataraj.R, (2005), 'Comparative Analysis of IDF curves at Selected Sites in Utah', Research Report No. UT-05.02, Utah state University.
- 7. Grubbs and Beck (1972). 'Mathematical Models for surface Water Hydrology'. Research report No UK- 09- 05, University of surrey, UK.
- 8. Kabundege. G (2007). 'Assessment of water resources Potential in Ruvubu river Basin'. Arbaminch University, Msc Thesis.
- 9. Lam, K.H (2004). 'Update of the Short Duration Rainfall IDF Curves for Recent Climate in Quebec'. 57th Canadian Water Resources Association Annual Congress.
- 10.Maidment, D.R. (1993). Hand book of hydrology, McGRAW-Hill. Inc.USA.
- 11.NtiburumusiI. F (2000) "Etat Actuel des Connaissances sur l'Etude Pedologique et la Fertilite des marais au Burundi", PNUD/FAO, Rapport BDI/96/001.
- 12.Rao, K.H, & Hamed, A.R. (2000). 'Flood Frequency Analysis', CRC press LLC, Florida.
- 13.Reddy, P.J. (1997). 'Stochastic hydrology', Prentice Hall of India, New Delhi.
- 14.Wenzel, H.G. (1982). 'Rainfall for Urban Storm Design'. American Geological Union. Washington DC.

Appendix

Appendix A: Annual maximum rainfall data for the selected stations Station: Ruvyironza

Station: Muyinga

Station: Gisozi

Station: Musasa

Station: Ruyigi

Station: Rwegura

Station: Karuzi

Station: Gitega

Station: Mparambo

Station: Bujumbura

Station: Cankuzo

Station: Nyanza lac

Station: Makamba

Station: Kirundo

Station: Rumonge

Station: Tora

Station: Nyamuswaga

Station : Kinyinya

Station Teza

Appendix B: Mathematical expression of Probability Distributions for annual maximum series.

Tora.

Ruvyironza

Musasa

Kinyinya

Ruyigi

Cankuzo

Muyinga

Kirundi

Nyamuswaga

Appendix D : IDF maps for some Durations and frequencies

Appendix E: Estimated Quantiles

1. Gisozi

2.Karuzi

3. Kinyinya

4. Musasa

5. Kirundo

6. Cankuzo

7. Bujumbura

8. Makamba

9.Muyinga

10. Nyanza lac

11. Nyamuswaga

12. Rwegura

13. Rumonge

14. Ruvyironza

15. Ruyigi

16. Teza

17. Gitega

18. Mparambo

19. Tora

20. Region 1

21. Region 2

22. Region 3

23. Region 4

Appendix F: Intensity of rainfall for the selected durations and frequencies

Gisozi

Tora

Ruvyironza

Makamba

Nyanza lac

Musasa

Kinyinya

Ruyigi

Cankuzo

Muyinga

Kirundo

Nyamuswaga

Karuzi

Rwegura

Teza

Bujumbura

Rumonge

Mparambo

Gitega

Region 1.

Region 2

Region 3

Region 4

