DEVELOPMENT AND ASSESSMENT OF REGIONALISED AREAL REDUCTION FACTORS FOR CATCHMENT DESIGN RAINFALL ESTIMATION IN SOUTH AFRICA

by

JACOBUS PETRUS JOHANNES PIETERSEN

A thesis submitted in fulfilment of the requirements for the degree

Doctor of Engineering in Civil Engineering

in the

Department of Civil Engineering

of the

Faculty of Engineering, Built Environment, and Information Technology

of the

Central University of Technology, Free State, South Africa

Promoter: Prof. OJ Gericke[1] IntPE (SA) Pr Eng.

Co-promoter: Prof. JC Smithers[2] Pr Eng.

[1] Central University of Technology, Free State

[2] University of KwaZulu-Natal

January 2023

DECLARATION

I, the undersigned, declare that the thesis hereby submitted by me for the degree *Doctor of Engineering in Civil Engineering* at the Central University of Technology, Free State, is my own independent work and has not been submitted by me to another university and/or faculty to obtain a degree. I further cede copyright of this thesis in favour of the Central University of Technology, Free State.

Signature: ………………

Date: 02 January 2023

Bloemfontein, South Africa

ABSTRACT

Design point rainfall depths converted to an average areal design rainfall depth using Areal Reduction Factors (ARFs) are regarded as fundamental input to various design flood estimation methods. The ARF estimation methods currently used in South Africa are regarded as being outdated and not being developed and/or verified using local data. The primary research objective is to estimate geographically-centred and probabilistically correct ARFs representative of the different rainfall regions associated with the Regional Linear Moment Algorithm and Scale Invariance (RLMA&SI) regionalisation scheme in South Africa. Merging of the 78 homogeneous RLMA&SI rainfall clusters was necessary to increase the size of the clusters and the number of rainfall stations within a particular cluster to meet the minimum required number of rainfall stations/km² criteria. The latter merging resulted in 46 delineated clusters. Long duration geographically-centred and probabilistically correct ARFs were estimated using a total of 2 053 artificial circular catchments and 1 779 daily rainfall stations located within the 46 clusters. Random combinations of the 46 clusters were used in an alternating fashion for calibration and/or verification purposes until all possible combinations were considered. Ultimately, it was noted that whether a dedicated set of clusters or all clusters are assigned to calibration, differences are regarded as insignificant, given that all ARF data sets, whether used for calibration or verification, remain only estimated sample values. Subsequently, five (5) ARF regions were deduced from the 46 clusters and all clusters in a particular ARF region were used for the final derivation of a non-linear (second-order polynomial) log-transformed empirical ARF equation. The new regional ARF equation performed similarly, and as expected, when compared to a selection of geographically-centred ARF estimation methods currently used in local and/or international practice in a range of catchment sizes. The estimated ARFs decreased with an increase in area and increased with an increase in both storm duration and return period. The ARF methodology developed in this research and the subsequent findings are new to the South African flood hydrology research community and practice: (i) ARFs were derived and are based on a regionalisation scheme utilising the daily rainfall data in the Daily Rainfall Extraction Utility (DREU) database, (ii) ARFs are probabilistically correct, *i.e.,* vary with return period, and (iii) a web-based software application was developed to enable the consistent estimation of ARFs within the five (5) ARF regions of South Africa.

EXTENDED ABSTRACT

Design point rainfall estimates are only applicable to a limited spatial area, and for larger areas, the average areal design rainfall depth is likely to be less than the average design point rainfall depth. Areal Reduction Factors (ARFs) are used to describe this relationship between point and areal rainfall, *i.e.,* design point rainfall depths are converted to an average areal design rainfall depth for a catchment-specific critical storm duration and catchment area. Average areal design rainfall is a fundamental input parameter to several design flood estimation methods and errors in these estimated values will thus directly impact on peak discharge estimates which are used in the design of hydraulic structures.

In South Africa, the estimation of ARFs is limited to the storm-centred approaches of Van Wyk (1965) and Wiederhold (1969), and the geographically-centred approach of Alexander (2001). These methods are only applicable to specific temporal and spatial scales and do not account for any regional differences. Only the method proposed by Van Wyk (1965) is regarded as being probabilistically correct, *i.e.*, ARFs vary with return period. However, both the methods of Van Wyk (1965) and Wiederhold (1969) are storm-centred approaches, which are currently incorrectly applied by practitioners in a geographically-centred manner. Alexander's geographically-centred method (2001) was transposed from methods developed in the United Kingdom (UK) with little local verification and it is also regarded as being probabilistically incorrect, *i.e.*, ARFs remain constant irrespective of the return period under consideration. The variation of ARFs with return period has previously been confirmed by several international studies, *e.g.*, Bell (1976), Omolayo (1993), Stewart (1989), Siriwardena and Weinmann (1996), and Podger *et al*. (2015a; 2015b).

The research aim is to develop a regionalised approach to estimate ARFs for improved design flood estimation in South Africa. The primary research objective is to develop regionalised, geographically-centred and probabilistically correct ARFs for South Africa by considering the relationship between average areal design rainfall and average design point rainfall estimates. Hence, the research focus is to develop regionalised and probabilistically correct ARFs representative of the different rainfall producing mechanisms in South Africa at a 'circular catchment level' using: (i) daily rainfall data to estimate areal design and design point rainfall, (ii) a modified version of Bell's method (1976), and (iii) the current regionalisation scheme

associated with the Regional Linear Moment Algorithm and Scale Invariance (RLMA&SI) approach (Smithers and Schulze, 2004).

Merging of the 78 homogeneous rainfall clusters within the seven (7) long duration rainfall regions associated with the RLMA&SI regionalisation scheme was necessary to increase the size of the clusters and the number of rainfall stations within a particular cluster to meet the minimum required number of rainfall stations/km² criteria. Merging only took place between clusters within the same long duration rainfall region, which resulted in 46 delineated clusters. A total of 2 550 artificial circular catchments associated with 1 779 daily rainfall stations with at least 30 years combined areal record lengths were strategically positioned in the 46 clusters throughout South Africa. Due to the large number of circular catchments placed in each of the 46 clusters, an overlapping of circular catchments was evident. Consequently, this resulted in daily rainfall data from similar rainfall stations being used multiple times within a particular cluster. In principle, this was not regarded as problematic, while it also contributed to the 'smooth' transition between the different clusters. In applying the selected screening criteria, only 2 053 circular catchments were used in the probabilistic and regression analyses. The probabilistic analyses based on the General Extreme Value (GEV) distribution, fitted to the Annual Maximum Series (AMS), using Linear Moments (LM), resulted in areal and design point rainfall values for a range of storm durations (*e.g.*, 1, 3, 5 and 7-day), and return periods (*e.g.*, 2, 5, 10, 20, 50 and 100-year). ARFs based on the long duration rainfall were subsequently estimated and expressed as the ratio between the average areal catchment design rainfall and the average design point rainfall estimates for corresponding return periods.

Initially, linear regression analyses were considered in each of the 46 clusters. In each case, the Goodness-of-Fit (GOF) statistics of normal and log-transformed data were evaluated for the independent criterion and dependant predictor variables. The GOF statistics of the log-transformed data outperformed the normal data and resulted in only log-transformed independent predictor variables being used. The 46 clusters were subjected to further merging providing that the clusters are still located within the same RLMA&SI long duration rainfall region and that the GOF statistics of each merged region do not decrease with more than 15% when compared to the original average GOF statistics in the 46 clusters. Subsequently, five (5) ARF regions were deduced from the 46 clusters and a non-linear (second-order polynomial) log-transformed empirical ARF equation, with unique regional calibration coefficients, was derived for each region using backward stepwise multiple regression analyses with deletion at a 95% confidence level.

Verification requires the use of data sets not being used during the calibration process. During the developmental phases of evaluating the suitability of linear versus non-linear log-transformed regression analyses, the concept of calibration and verification was evaluated in each of the 46 clusters. In other words, random combinations of the 46 clusters (with their own unique set of circular catchments and rainfall stations) were used in an alternating fashion for calibration and/or verification purposes until all possible combinations were considered. The process was applied in all clusters. Ultimately, it was noted that whether a dedicated set of clusters or all clusters are assigned to calibration, differences are regarded as insignificant. Hence, all clusters in a particular ARF region were used for the final derivation of the non-linear (second-order polynomial) log-transformed empirical ARF equation. The latter approach was justified by the limited number of rainfall stations available in many of the clusters, which is also affected by the number of rainfall stations/km² criteria to be met in each cluster. Furthermore, it must be noted that 'assessment' of verification results can only apply if the benchmark data set is an observed data set. In the case of ARFs, all data sets, whether used for calibration or verification, remain only estimated sample values.

Subsequently, the derived ARF equation was compared to a selection of geographically-centred ARF estimation methods currently used in local and/or international practice in a range of catchment sizes to establish the consistency and/or possible biases of the newly derived ARF equation. The new regional ARF equation performed similarly, and as expected, when compared to the other methods. The estimated ARFs decreased with an increase in area and increased with an increase in both storm duration and return period.

Based on the research findings in the different ARF regions, it was evident that ARFs, derived from local South African data, are proportional to or influenced by the RLMA&SI regionalisation scheme, rainfall types, storm durations, and return periods. All the research assumptions were also confirmed: (i) design point rainfall estimates are only representative for a limited area, which was demonstrated by the differences between the average areal design rainfall and average design point rainfall estimates, and (ii) the current geographically-centred South African ARF estimation methods are only applicable to specific temporal and spatial scales, do not account for any regionalisation, provide ARF estimates exceeding 100% in ‘smaller’ catchments, provide constant ARF values for all return periods, and are based on limited/no local data.

The ARF methodology developed in this research and the subsequent findings are new to the South African flood hydrology research community and practice: (i) ARFs were derived and are based on a regionalisation scheme utilising the daily rainfall data in the Daily Rainfall Extraction Utility (DREU) database (Lynch, 2004), (ii) ARFs are probabilistically correct, *i.e.*, vary with return period, and (iii) a web-based software application was developed to enable the consistent estimation of ARFs within the five (5) ARF regions of South Africa.

ACKNOWLEDGEMENTS

The research was funded by the: (i) Water Research Commission (WRC) through a project entitled: *Development of a Regionalised Approach to Estimate Areal Reduction Factors and Catchment Response Time Parameters for Improved Design Flood Estimation in South Africa (WRC Project K*5-2924), (ii) Central University of Technology, Free State (CUT) in collaboration with the Department of Higher Education and Training (DHET), and (iii) National Research Foundation (NRF) of South Africa.

Without the above funding, none of the research would have been possible; hence, I am extremely grateful!

Several individuals also deserve specific mention:

(a) The Rectorate and relevant functionaries from the Central University of Technology, Free State, for the opportunity to complete this research.
(b) Prof. OJ Gericke, my promoter, for his positive attitude, expertise, direction, and supervision throughout my research.
(c) Prof. JC Smithers, my co-promoter, for his expertise, direction, and guidance.
(d) Mr F Sokolic (GIS Solutions) for technical assistance in compiling all the relevant Geographical Information System (GIS) data sets.
(e) My wife and our children for their love and support.
(f) My family, friends, and colleagues, for their patience and understanding throughout this research.

Above all, acknowledgement to my Heavenly Father for setting my feet on a rock and making my steps secure (Psalm. 40).

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

AEP	Annual Exceedance Probability
AMS	Annual Maximum Series
ARC	Agricultural Research Council
ARF	Areal Reduction Factor
CAPPI	Constant Altitude Plan Position Indicator
CEH	Centre for Ecology and Hydrology
CPD	Continuous Professional Development
CSS	Cascading Style Sheets
CUT	Central University of Technology, Free State
DCR	Daily Catchment Rainfall
DDF	Depth-duration-frequency
DHET	Department of Higher Education and Training
DREU	Daily Rainfall Extraction Utility
DWS	Department of Water and Sanitation
EMA	Expectation Maximisation Algorithm
ESRI	Environmental Systems Research Institute
EV1	Extreme Value Type I
GEV_{LM}	General Extreme Value using Linear Moments
GEV_{PWM}	General Extreme Value using Probable Weighted Moments
GIS	Geographical Information System
GOF	Goodness-of-Fit
HTML	Hypertext Mark-up Language
IDW	Inverse Distance Weighted
IDE	Integrated Development Environment
ISCW	Institute for Soil, Climate and Water
LAT	Latitude
LM	Linear Moments
LN3	3 Parameter Log Normal
LN	Log-Normal
LONG	Longitude
$LP3_{MM}$	Log-Pearson Type III using Method of Moments
MacOS	Macintosh Operating System
MAE	Mean Annual Evaporation
MAP	Mean Annual Precipitation

MATLAB	Matrix Laboratory
MCR	Maximum Catchment Rainfall
MI	Monthly Infilling
MM	Method of Moments
MML	Method of Maximum Likelihood
MR	Median Ratio
MSR	Maximum Station Rainfall
NERC	Natural Environment Research Council
NFSP	National Flood Studies Programme
NRF	National Research Foundation
PDS	Partial Duration Series
PWM	Probability Weighted Moments
QGIS	Quantum Geographical Information System
RL	Record Length
RLMA&SI	Regional Linear Moment Algorithm and Scale Invariance
RSA	Republic of South Africa
SA	South Africa
SANCOLD	South African National Committee on Large Dams
SANRAL	South African National Roads Agency Limited
SASA	South African Sugar Association
SASEX	South African Sugar Association Experiment Station
SAWB	South African Weather Bureau
SAWS	South African Weather Services
SE	Standard Error
SRTM	Shuttle Radar Topography Mission
SRR	Smithers Regional Rainfall
T_C	Time of Concentration
TR	Technical Report
UK	United Kingdom
UK FSR	United Kingdom Flood Studies Report
USA	United States of America
USGS	United States Geological Survey
USWB	United States Weather Bureau
VBA	Visual Basic for Applications
VSC	Visual Studio Code
WGS84	World Geodetic System 1984
WRC	Water Research Commission

CHAPTER 1: INTRODUCTION

This chapter provides some background on the estimation of areal design rainfall, through the application of an Areal Reduction Factor (ARF), and the influence this input parameter has on the estimation of design floods in South Africa. This chapter includes the rationale, problem statement, research objectives, and an outline of the thesis structure.

1.1 Rationale

The estimation of design flood events, *i.e.*, floods characterised by a specific magnitude-frequency relationship, at a particular site in a specific region is necessary for the planning, design, and operation of hydraulic structures, *e.g.*, culverts, bridges, spillways, *etc.* (Pegram and Parak, 2004). In general, observed rainfall data can be obtained from continuously recording rainfall stations or from daily rainfall stations. In South Africa, daily rainfall data are recorded at a fixed daily interval and are more abundant, reliable, and generally have longer record lengths than the continuously recorded sub-daily rainfall data (Smithers and Schulze, 2000b; 2004). Hence, due to the availability and quality of daily rainfall data, these data sets are more frequently used to estimate design rainfall. In essence, design rainfall is derived from observed rainfall data and comprises of a depth and duration associated with a given Annual Exceedance Probability (AEP) or return period (*T*) (Gericke and Du Plessis, 2011).

Design rainfall for durations < 24-hour is normally classified as 'short duration' design rainfall and generally estimated directly from continuously recorded rainfall. 'Long duration' design rainfall typically exceeds one day in duration and can be estimated from both continuously recorded and daily rainfall data (Smithers and Schulze, 2004). However, design point rainfall estimates derived from a single rainfall station are only applicable to a limited area and for larger areas, the average areal design rainfall depth is likely to be less than the average design point rainfall depth (Siriwardena and Weinmann, 1996). ARFs are used to describe the relationship between point and areal rainfall, *i.e.*, design point rainfall depths are converted to an average areal design rainfall depth for a catchment-specific critical storm duration and catchment area (Alexander, 2001).

ARFs are estimated using either analytical or empirical methods as shown in Figure 1.1 (Pietersen *et al.*, 2015).

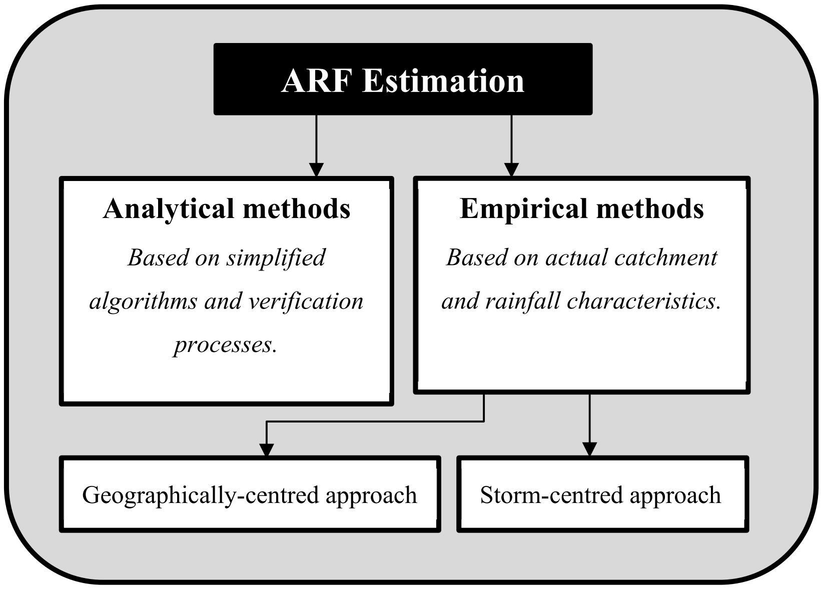

Figure 1.1: Typical ARF estimation approaches

The first analytical methods were based on simplified algorithms and limited verification processes (Siriwardena and Weinmann, 1996; Svensson and Jones, 2010); hence, several new analytical methods have been proposed during the last four decades, *e.g.*, storm movement (Bengtsson and Niemczynowicz, 1986), crossing properties (Bacchi and Ranzi, 1996), spatial correlation structure (Sivapalan and Blöschl, 1998), scaling relationships (De Michéle *et al.*, 2001), and temporal-spatial rainfall dependence (Mineo *et al.*, 2018). Empirical methods are classified as either geographically-centred or storm-centred approaches. The geographically-centred approach describes the relationship between average areal design rainfall over a fixed geographical area and a corresponding average design point rainfall value representative of the area under consideration. In the storm-centred approach, the estimation of average areal design rainfall is not limited to a fixed geographical area, but rather associated with the extent of individual storm rainfall events and the way in which the rainfall intensity decreases with distance from the central maximum rainfall core (Alexander, 2001; Svensson and Jones, 2010).

Hence, average areal design rainfall, which is derived using an ARF, is regarded as fundamental input to all rainfall-runoff event-based design flood estimation methods, while errors in estimated average areal design rainfall will directly impact on the estimated peak discharges.

1.2 Problem Statement

Numerous factors could have a significant impact on the estimation of ARFs, *e.g.*, geographical location, rainfall types, catchment geomorphology, methodological approaches, climatological regions, storm duration and AEP (Asquith and Famiglietti, 2000; Svensson and Jones, 2010; Li *et al.*, 2015; Kim *et al*., 2019). In terms of geographical location, it was established that the 1-day ARFs in the United States of America (USA) exceed the equivalent ARF estimates used in Australia, while the ARFs in the semi-arid south-western parts of the USA, decline more rapidly than in the rest of the USA (Svensson and Jones, 2010). Different rainfall-producing mechanisms, *e.g.*, convective versus frontal rainfall, will produce different spatial rainfall patterns and consequently result in different ARF values (Eggert *et al.*, 2015). For example, Skaugen (1997) established that ARFs for both convective and frontal rainfall decrease with increasing return periods, but the rate of decrease for convective rainfall is noticeably larger than that of frontal rainfall.

In the USA, areal rainfall was found to decrease in comparison with the corresponding point rainfall with increasing return periods (Asquith and Famiglietti, 2000; Allen and DeGaetano, 2005). In contrast, Grebner and Roesch (1997) demonstrated that ARFs in Switzerland (catchment areas > 4 500 km²) are independent of the return period. Most research conducted on the estimation of ARFs concluded that catchment geomorphology (*e.g.*, area, shape, and topography) and topographical rainfall biases (*e.g.*, leeward, and windward effects) have an insignificant influence on ARFs (Allen and DeGaetano, 2005; Svensson and Jones, 2010). However, Singh *et al.* (2018) highlighted that ARF differences in New Zealand are ascribed to differences in topography and rainfall types. Kim *et al.* (2019) also showed that storm-centred ARF values obtained from storms of a different shape, *i.e.*, elliptical versus circular, could be different by up to 20%. In catchment areas less than 800 km², ARFs are mainly a function of the point and areal rainfall intensity since the relationship between rainfall intensity and the infiltration rate of the soil is dominant. The South African National Roads Agency Limited (SANRAL, 2013) indicated that for catchment areas up to 30 000 km², ARFs are mainly a function of the area and storm duration.

Internationally, the extensive national-scale ARF studies reported in the literature, are limited to the United Kingdom Flood Studies Report (UK; FSR, 1975), United States Weather Bureau (USWB, 1957; 1958) and Australia (Siriwardena and Weinmann, 1996; Podger *et al.*, 2015a; 2015b). Due to insufficient rainfall-monitoring networks and a lack of short duration rainfall data, most of the data-intensive analytical and empirical methods developed internationally, often fail to successfully incorporate the variation in predominant weather/rainfall types, storm durations, seasonal factors and return periods (Skaugen, 1997; Asquith and Famiglietti, 2000; Allen and DeGaetano, 2005, Pavlovic *et al.*, 2016). In recent years, radar information has also become more readily available in many parts of the world and assists in improving the spatial and temporal resolution to estimate ARFs, *e.g.*, Peleg *et al.* (2018), Kim *et al.* (2019), and Du Plessis *et al.* (2020).

In South Africa, the estimation of ARFs is limited to the storm-centred approaches of Van Wyk (1965) and Wiederhold (1969), and the geographically-centred approach of Alexander (2001). The methods of Van Wyk (1965) and Wiederhold (1969) are based on small-scale, pilot case studies (*cf.* Chapter 2, Section 2.9.2), which are therefore only applicable to specific temporal and spatial scales, and do not account for any regional differences. This is a concern since the temporal and spatial distributions of rainfall in South Africa are highly variable on a seasonal and annual basis given that the rainfall is produced by different weather systems in different regions and at contrasting times of the year (Davies and Day, 1998; Mambo and Faccer, 2017). Only the method proposed by Van Wyk (1965) is regarded as being probabilistically correct, *i.e.*, ARFs vary with return period. However, both the methods of Van Wyk (1965) and Wiederhold (1969) are storm-centred approaches, which are currently incorrectly applied by practitioners in a geographically-centred manner. Alexander's geographically-centred method (2001) was transposed from methods developed in the UK with little local verification and it is also regarded as being probabilistically incorrect, *i.e.*, ARFs remain constant irrespective of the return period under consideration.

The empirical (storm-centred) and analytical (correlation-based and annual maxima-centred) methods (*cf.* Section 1.1) do not provide probabilistically correct areal design rainfall estimates since it is assumed that the AEP of both the point and areal rainfall is the same. Most of these methods are also based on a limited amount of observed rainfall data and use assumptions that are not entirely true descriptions of the actual rainfall process (Svensson and Jones, 2010). Moreover, some studies (*e.g.*, Omolayo, 1993; Siriwardena and Weinmann, 1996; Podger *et al*., 2015a; 2015b; Du Plessis and Loots, 2019) have conclusively shown that ARFs are dependent on the average AEP of rainfall. According to Pavlovic *et al.* (2016), the differences between the analytical and empirical ARF estimation methods currently in use are also more pronounced for shorter storm durations (< 24-hour) and larger catchment areas

(> 750 km^2), while some ARFs depend on the average return period. Thus, with the focus on larger catchments, most of the methods listed above are inappropriate to use at a comprehensive set of temporal and spatial scales.

Two recognised methods to estimate design point rainfall, *i.e.*, Technical Report (TR) 102 (Adamson, 1981) and the Regional Linear Moment Algorithm and Scale Invariance (RLMA&SI) (Smithers and Schulze, 2004), are commonly used by practitioners in South Africa. To estimate catchment design rainfall and overcome the limitations of design point rainfall, the Department of Water and Sanitation (DWS) developed an approach (*cf*. Chapter 2, Section 2.6) to estimate areal design rainfall known as the Daily Catchment Rainfall (DCR) approach (Van der Spuy and Rademeyer, 2018). According to Van der Spuy and Rademeyer (2018), the latter approach eliminates the required use of ARFs; however, there are some limitations associated with the latter approach: (i) it is catchment specific and needs to be derived for each catchment under consideration, and (ii) no established areal design rainfall database is available. Hence, ARFs still need to be applied to the design point rainfall values, which are the primary source of available design rainfall values in South Africa.

Based on the shortcomings highlighted above, it is evident that the estimation of ARFs is internationally an on-going focus of research. South African researchers only recently studied ARFs (*e.g.*, Pietersen *et al*., 2015; Pietersen, 2016; Du Plessis and Loots, 2019; Du Plessis *et al.,* 2020), despite the on-going need for locally developed ARF estimation methods. Hence, the ARFs in South Africa need to be re-investigated in the light of the recent rainfall events utilising the longer periods of record (± 50 years of additional data since the last South African derived ARF method in the 1970s), which are now available for analysis. The variation of ARFs with area, return period, duration and rainfall producing mechanisms also needs to be investigated by adopting a regional approach.

Given the impact of estimated ARF values on design flood estimation, as highlighted above, the updating of ARFs was also identified as a potential research project to be included in the National Flood Studies Programme (NFSP) (Smithers *et al.*, 2014). The NFSP was initiated by the South African Committee on Large Dams (SANCOLD) and the Water Research Commission (WRC) and is supported in principle by the DWS and SANRAL (Smithers *et al.*, 2014). A wide range of issues have been highlighted for research by the four Working Groups of the NFSP, of which, the estimation of ARFs was regarded as one of the high priority research areas.

Consequently, the above not only served as a motivation for this research, but also emphasised that the continued use of inappropriate ARF estimation methods in South Africa may have a negative direct impact when design flood estimation methods are used in the design of hydraulic designs. Potentially, the hydraulic structures might not only be over- or under-designed, but any loss of life, due to excessive flood damages and under-designed infrastructure, is also not excluded.

1.3 Research Aim

The research aim is to develop a regionalised approach to estimate ARFs for improved design flood estimation in South Africa. The pilot scale study of Pietersen (2016) was used as guiding reference, which recommended that the current ARF methodologies should be improved and expanded throughout South Africa to enable the development of a regional approach. A regional approach will not only improve the robustness and accuracy of the ARFs at a national scale, but it will also result in improved design flood estimation.

1.3.1 Research objective

The primary research objective is to develop regionalised, geographically-centred and probabilistically correct ARFs for South Africa by considering the relationship between average areal design rainfall and average design point rainfall estimates. Consequently, this will elucidate how ARF values vary with catchment area, storm duration, rainfall types, and return period within the different rainfall regions of South Africa. Hence, the research focus is to develop regionalised and probabilistically correct ARFs representative of the different rainfall producing mechanisms in South Africa at a 'circular catchment level' using: (i) daily rainfall data to estimate areal design and design point rainfall, (ii) a modified version of Bell's method (1976), and (iii) the current regionalisation scheme associated with the Regional Linear Moment Algorithm and Scale Invariance (RLMA&SI) approach (Smithers and Schulze, 2004).

Artificial 'circular catchments' as opposed to actual catchments are used, since as demonstrated by Pietersen (2016), catchment shape, orientation, and size have a negative impact on ARF estimates, which could typically lead to incorrect and/or inconsistent ARF sample values being derived. Various Australian ARF studies (Siriwardena and Weinmann, 1996; Podger *et al.,* 2015a; 2015b) also demonstrated and confirmed that circular catchments are preferred to actual catchments, given that catchment shape and size have a negative impact on rainfall distribution and uniformity. Hence, the rainfall distribution within a circular catchment of any given size, will not only be more uniform, but circular shapes also simplify the processing and extraction of huge data sets in a Geographical Information System (GIS) environment, all factors which contribute towards consistency. Hence, the use of multiple circular catchments with random sizes, determined by the location of rainfall stations, covering a specific homogeneous rainfall region (according to the RLMA&SI regionalisation scheme) will enable the estimation of sample ARFs in a consistent fashion.

Based on the research conducted in various international studies, and as outlined in Chapter 2 of this thesis, a research question, applicable to South African derived ARFs, applies: Does: (i) rainfall types, (ii) storm durations, (iii) climatological/regional factors, and (iv) return periods affect the estimation of ARFs? The following two hypotheses apply:

(a) **Null Hypothesis (H_0):** ARFs, derived from local South African data, are not proportional to ($\not\propto$) nor influenced by the RLMA&SI regionalisation scheme, rainfall types, storm durations and return periods.

(b) **Alternative Hypothesis (H_1):** ARFs, derived from local South African data, are proportional to ($\propto$) or influenced by the RLMA&SI regionalisation scheme, rainfall types, storm durations and return periods.

The primary research objective is furthermore based on the following two assumptions:

(a) **Assumption 1:** Design point rainfall estimates are only representative for a limited catchment area and for larger catchment areas, the average areal design rainfall depth or intensity is likely to be less than the average design point rainfall depth or intensity.

(b) **Assumption 2:** The current geographically-centred South African ARF estimation methods are only applicable to specific temporal and spatial scales and do not account for regional variations or return periods.

1.3.2 Specific objectives

The specific objectives of the research are to:

(a) Conduct a comprehensive literature review on the ARF estimation methods currently used internationally.

(b) Contribute to the establishment of an improved and updated rainfall database for the purpose of design rainfall estimation.

(c) Identify homogeneous rainfall regions in South Africa by adopting a regional approach.

(d) Extract and analyse rainfall data to provide geographically-centred and probabilistic-correct ARFs.

(e) Derive regional empirical ARF equations for application throughout South Africa in different homogeneous rainfall regions.

(f) Evaluate and verify the derived empirical ARF equation(s).

(g) Develop a web-based software interface to enable users to utilise the empirical ARF equations at a national-scale in South Africa.

It is envisaged that the implementation of the specific objectives will contribute fundamentally to the improved estimation of ARFs to ultimately result in improved design flood estimations in South Africa.

1.4 Outline of Thesis Structure

Chapter 2 presents a comprehensive literature review of the various rainfall driving mechanisms, typical factors which influence the estimation of ARFs, and the different ARF estimation methods used to describe the relationship between average design point and average areal design rainfall. Chapter 3 provides an overview and a general description of the study area. Chapter 4 presents an overview of the concerns encountered while progressing with and developing the methodological approach, as well as the implementation of solutions to formalise the final methodology. Chapter 5 comprises of the final methodology and subsequent results. Chapter 6 presents a synthesis of all the information as discussed in Chapters 1 to 5, as well as some final conclusions and recommendations. Chapter 7 contains the list of the references used.

All the supplementary information is included in Appendices A to E.

CHAPTER 2: LITERATURE REVIEW

The literature review contained in this chapter focuses on the methods developed nationally and internationally to estimate ARFs. Chapter 1 provided a general overview and insight on the current circumstances related to South African ARFs. Therefore, the climate of South Africa and associated rainfall types are discussed first, followed by an overview of the status of observed rainfall measurement in South Africa and the associated infilling and averaging techniques which are commonly used. Thereafter, design rainfall estimation is detailed, while the remaining part of the chapter focusses on the current national and international ARF estimation methods.

2.1 Climate and Rainfall Types

Climate does not only affect rainfall distribution, but also rainfall intensity, duration, and variability, which are all interdependent. However, the four major rainfall processes occurring in South Africa will also affect this interdependency and are most likely to have different influences on the estimation of ARFs. The four major rainfall processes occurring in South Africa can be summarised as follows (Haarhoff and Cassa, 2009; Van der Spuy and Rademeyer, 2018):

(a) **Convective rainfall:** This process typically occurs during the summer season when air layers (closest to the earth's surface) saturated with water vapour are heated and subsequently tend to rise and cool down, resulting in cloud formation and rainfall. The rainfall intensity is normally high to very high with associated thunder activity. Convective rainfall is characteristic of the Highveld region which covers the Free State, Gauteng, and Mpumalanga provinces.

(b) **Cyclonic rainfall:** This rare process typically occurs over the open sea and is formed when cyclones (large circular patterns) are growing, allowing moist air to be drawn into the cyclone vortex and allowing mist to be lifted into the centre, resulting in very strong winds and extremely high rainfall intensities.

(c) **Frontal rainfall:** This inland process typically occurs when cold or warm fronts are moving across the country and interact with one another. The cold air has the tendency to move underneath the warm air, and the warm air is deflected upwards by the trailing edge of the cold air. In both cases, the warm air is lifted into the colder region, resulting in rainfall.

(d) **Orographic rainfall:** This process usually occurs near coast lines and typically develops when wind blows over the open sea towards land carrying air saturated with water vapour until it reaches a mountain range. At these geographical barriers, the saturated air is forced upwards resulting in condensation and rainfall. The rainfall intensity is normally regarded as moderate and depends on wind blowing towards the inland areas. Orographic rainfall is characteristic of the coast lines of the KwaZulu-Natal and the Western Cape provinces.

The rainfall types listed in (a) to (d) is expected to have an influence on the estimation of ARFs, since the magnitude of ARFs is expected to be highly dependent on the different storm mechanisms associated with different rainfall types (Skaugen, 1997; Van der Spuy and Rademeyer, 2018). In a specific region with more frequent thunderstorms (convective rainfall) occurring than frontal storms (widespread rainfall), the typical observed point rainfall Annual Maximum Series (AMS) for that specific region would likely consist of rainfall values associated with convective activity (rainfall with rapidly changing intensity); whereas the frontal rainfall values could have been more representative of widespread rainfall within a catchment or region. This may result in much lower probabilistically correct ARFs (thunderstorms with high intensities), as opposed to the probabilistically higher ARFs represented by the frontal activity (Siriwardena and Weinmann, 1996).

In recognition of the above-mentioned interdependencies, Weddepohl (1988; cited by Schulze *et al.*, 1992) demarcated South Africa into four distinctive daily design rainfall intensity distribution regions (*cf.* Figure 2.1). Typically, Region 1 is associated with a Type 1 design rainfall intensity distribution which is regarded as the lowest, while Type 4 is associated with the highest rainfall intensity.

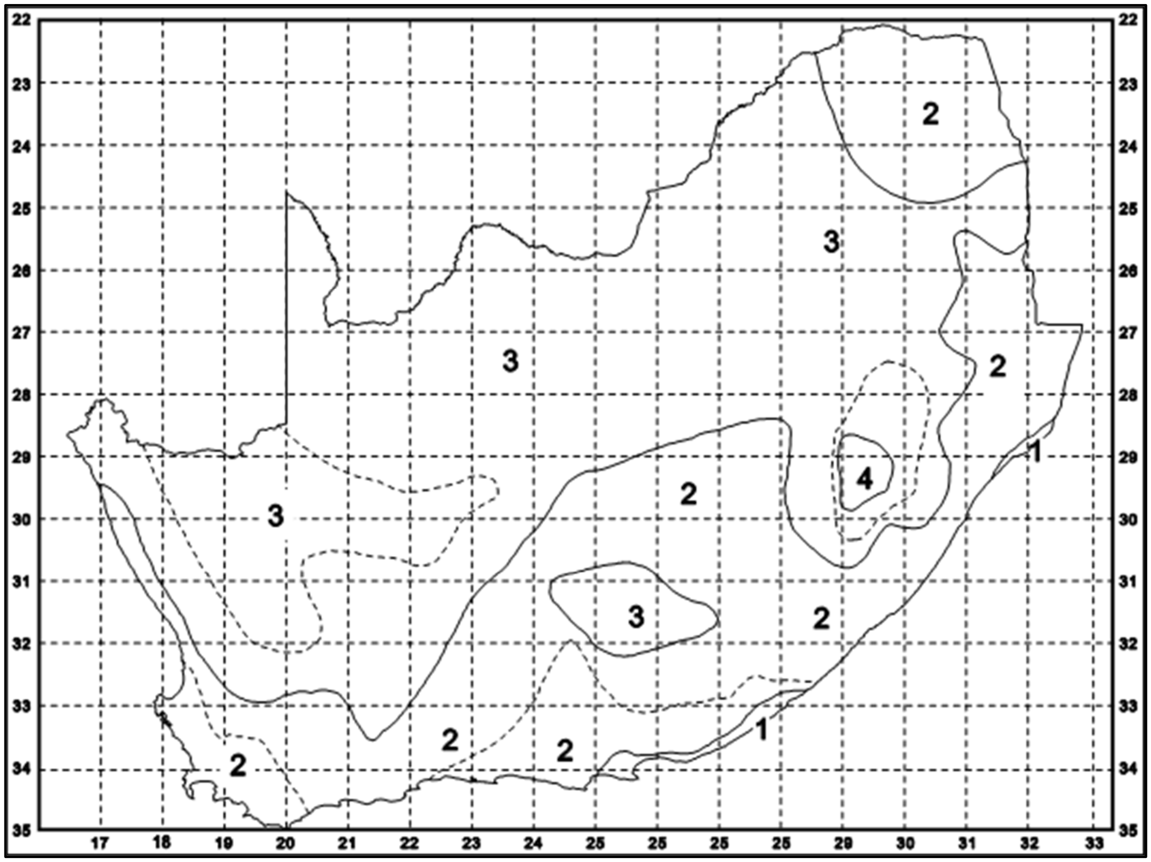

Figure 2.1: Design rainfall intensity distributions over South Africa (Weddepohl, 1988)

2.2 Observed Rainfall Data in South Africa

Observed daily rainfall data in South Africa are obtained from daily rainfall stations, which are widespread in space, and are measured as a depth (mm) at a specific time interval (8:00) on each day. The poor maintenance of rainfall stations in South Africa, under the responsibilities of the South African Weather Services (SAWS), was highlighted by Smithers and Schulze (2000b) and confirmed in a more recent study by Van Vuuren *et al.* (2012). The recent survey highlighted that approximately 1 200 rainfall stations are currently out of service, whereas most of these stations were operational in the late 1960s. Unfortunately, the current number of operational rainfall stations is less than in the 1920s and, considering this trend, South Africa might have even fewer operational rainfall stations soon (Smithers and Schulze, 2000b; Van Vuuren *et al*., 2012).

A total of 11 171 daily rainfall record lengths (long duration), obtained from available daily rainfall stations, were available in the South African database in the late 1990s (Smithers and Schulze, 2000b). This is illustrated in Figure 2.2. The SAWS contributed 78.9%

of the data, followed by the Institute for Soil, Climate and Water (ISCW) (7.7%), joint SAWS-ISCW (3.3%), South African Sugar Association Experiment Station (SASEX) (1.4%) and the remaining 8.8% by private entities (Smithers and Schulze, 2000b). However, more than 20% of all daily rainfall stations with record lengths exceeding 20 years have more than 10% of their data missing (Smithers and Schulze, 2000b).

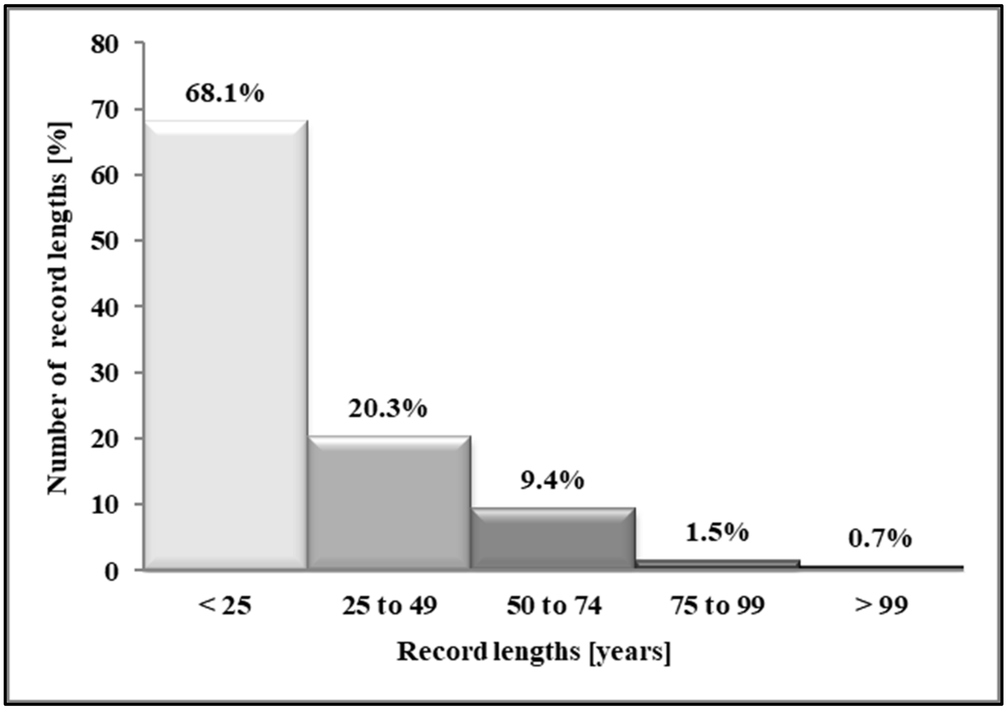

Figure 2.2: Available record lengths for daily rainfall stations in South Africa (after Smithers and Schulze, 2000b)

As reported by Smithers and Schulze (2000a) and shown in Figure 2.3, digitised short duration (less than 24 hours) rainfall data were only available from 412 stations in South Africa in the late 1990s. However, only 49 of these 412 rainfall stations had record lengths exceeding 30 years or longer (Smithers and Schulze, 2000a). The SAWS was the largest contributor to this sub-daily rainfall database, *i.e.*, 81% of all the rainfall stations (Smithers and Schulze, 2000a).

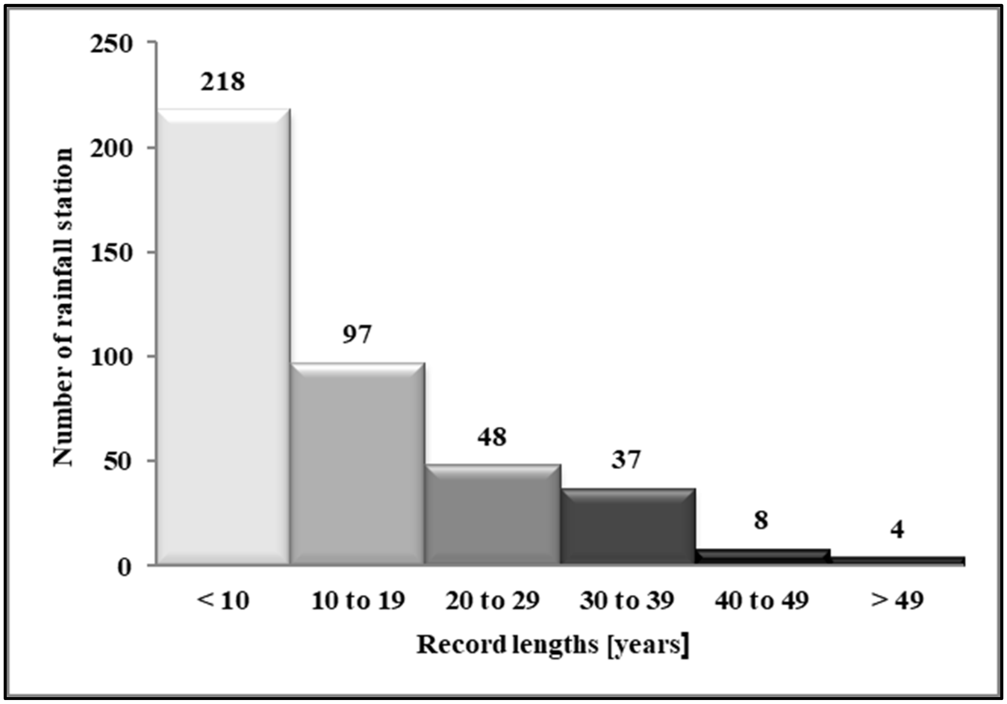

Figure 2.3: Short duration rainfall stations in South Africa (after Smithers and Schulze, 2000a)

Furthermore, Smithers and Schulze (2000a) also highlighted that the digitised short duration rainfall data have a low reliability due to several possible frequent errors including missing data and differences (*e.g.,* more than 20 mm) between the digitised and standard rain gauge daily totals located at similar sites. It was concluded that the digitised SAWS rainfall data are inadequate for the estimation of design storm durations of less than 24 hours. Furthermore, Smithers and Schulze (2000a) developed specific approaches which are based on regional similarities, scaling properties and stochastic simulation of extreme rainfall events to estimate short duration design rainfall values. These approaches are discussed in the subsequent sections.

2.3 Infilling of Missing Observed Rainfall Data

Rainfall records characterised by missing data are a serious concern when daily/sub-daily rainfall-runoff modelling systems are used, since all these models are reliant on a continuous rainfall data series input (Pitman, 2011).

Lynch (2004) highlighted the importance of rainfall data infilling and emphasised that a missing day implies an incomplete month and consequently an incomplete year. Lynch (2004) developed the Daily Rainfall Extraction Utility (DREU) to extract infilled/patched quality-controlled rainfall data, which are based on a hierarchical process, by utilising the infilling suite consisting of different infilling techniques as explained in (a) to (d) below. The infilling procedure algorithms are based on one or a combination of the following techniques:

(a) **Inverse Distance Weighting (IDW):** The IDW technique (Meier, 1997; cited by Lynch, 2004) inversely weights the rainfall records from rainfall stations surrounding the rainfall station under consideration, depending on the distance of those rainfall stations from the rainfall station under consideration. Meier (1997; cited by Lynch, 2004) established a procedure for selecting neighbouring rainfall stations from each quadrant around the rainfall station under consideration. This approach ensured that a certain number of rainfall stations are selected from each of the four quadrants surrounding the station to minimise the uncertainty introduced when the closest few rainfall stations are all in the same direction from the rainfall station under consideration (Meier, 1997; cited by Lynch, 2004).

(b) **Expectation Maximisation Algorithm (EMA):** The EMA technique was adopted and refined by Makhuva *et al.* (1997; cited by Lynch, 2004) and Pegram (1997; cited by Lynch, 2004) to infill missing rainfall data on a monthly basis. The EMA technique revolves around a recursive action of substituting missing data in a multiple linear regression relationship to re-estimate the values between the data at the rainfall station under consideration and the data from the nearby control rainfall stations. Smithers and Schulze (2000b) highlighted that the EMA technique requires the selection of suitable control rainfall stations to be valuable in determining the suitability of using the selected target and control rainfall stations for the simultaneous infilling of missing data.

(c) **Monthly Infilling (MI) technique:** A regression approach was used to infill the non-existing missing monthly rainfall data by using the surrounding control rainfall stations as described by Zucchini (1984; cited by Lynch, 2004). The monthly database (observed and infilled) identified by Dent (1989; cited by Lynch, 2004) was interrogated whereby the monthly infilled values of zero, as well as monthly total values less than 2mm, were extracted. Infilled monthly totals of zero implied that infilled monthly totals of the month under consideration could be assumed to be equal to zero. Furthermore, the first day of each month was assigned the monthly total, in cases where the infilled monthly totals were less than 2 mm, and the rainfall amounts for the rest of the month were then set to a value of zero.

(d) **Median Ratio (MR) technique:** The MR technique depends on the median values between the rainfall station under consideration and the nearest control rainfall station to estimate a proportionality ratio. The latter proportionality ratio is used to correct the data from the rainfall station under consideration and to infill the missing daily data series. The advantage of the MR technique is that the closest control rainfall station with non-existing data will be replaced by the second closest control rainfall station (Lynch, 2004).

The EMA and MR techniques are the most effective infilling techniques in the DREU (Lynch, 2004). Any missing observed rainfall values not infilled by using the EMA and MR techniques, were infilled using the IDW technique. Subsequently, zero and less than 2 mm rainfall values, as derived by Dent (1989; cited by Lynch, 2004), are then used to infill any remaining missing values that have not been infilled. The South African daily rainfall database has more than doubled in size with the infilling techniques described above. The rainfall database consists of 105 753 218 daily observed values with 236 154 934 infilled values (Lynch, 2004). The observed and infilled rainfall database therefore contains 341 908 152 values (Lynch, 2004).

2.4 Estimation of Average Point Rainfall

In the assessment of total quantities of rainfall over large catchment areas, the occurrence of storms and their contribution to single rainfall stations are unknown. Therefore, it is necessary to convert observed point rainfall depths to provide the best estimate of the average point rainfall depth over a certain catchment area. The following methods may be used for averaging the point rainfall depth over an area (Wilson, 1990):

(a) **Arithmetic mean method:** This method [Eq. (2.1)] is defined as the sum of all the point rainfall information divided by the number of rainfall stations within the catchment area. This method is only adequate when rainfall stations are uniformly distributed in the catchment of interest, the topography is relatively flat and spatial variations in rainfall are insignificant.

$$\overline{P} = \sum \frac{P_i}{N_i} \qquad [2.1]$$

(b) **Thiessen polygon method:** This method [Eq. (2.2)] defines the zone of influence of each rainfall station by drawing lines between pairs of stations, bisecting the lines with perpendiculars. The total area enclosed within the boundary formed by these intersecting perpendiculars has rainfall of the same amount as the enclosed rainfall station. This method is not suitable for mountainous areas due to the orographic influences.

$$\overline{P} = \sum \frac{A_s P_i}{A_T} \qquad [2.2]$$

(c) **Isohyetal method:** This method [Eq. (2.3)] is based on the interpolation between rainfall stations to produce isohyets or contours of equal rainfall depth. The areal average of the weighted rainfall depths between the isohyets is then used to determine the average rainfall. This method is possibly the most accurate with an added advantage that the isohyets may be drawn to consider local effects of climate and uneven topography.

$$\overline{P} = \frac{\sum P_i A_i}{\sum N_i} \qquad [2.3]$$

(d) **IDW method:** This method [Eq. (2.4a)] is based on deterministic interpolation and takes the geographical position of each rainfall station relative to the other rainfall stations into consideration. A rainfall station which is geographically distant/close to other stations will have a larger/smaller weighting factor [Eq. (2.4b)] and will therefore contribute more/less to the estimation of the average areal rainfall. In essence, the sum of all point rainfall information is multiplied with individual weighting factors and divided by the total number of rainfall stations within the catchment under consideration (ESRI, 2006; Dyson, 2009).

$$\overline{P} = \frac{\sum P_i W_i}{N} \qquad [2.4a]$$

$$\overline{P} = \frac{\sum_{m=1}^{N-1} r_{\max}}{(N-1) r_{\max}} \qquad [2.4b]$$

where

$\overline{P}$ = spatial average rainfall depth [mm],

A_i = area [km²],

A_s = area of the sub-catchment contributing to the rainfall station [km²],

A_T = total catchment area [km²],

m = rank value of individual weighting factors,

N = total number of rainfall stations,

N_i = number of rainfall stations within area,

W_i = individual weighting factor,

P_i = point rainfall depth [mm], and

r_{max} = maximum distance between the specific rainfall station and any other rainfall station [m or km].

(e) **Grid point method:** For the grid point method, a uniform grid is superimposed over a catchment area containing the spatial location of each rainfall station (and associated rainfall depths). Rainfall is estimated at each corner of the grid and then multiplied with the representative grid-area to obtain the average rainfall volume. The sum of all the estimated volumes divided by the total catchment area equals the average areal rainfall depth (Patra, 2008).

(f) **Isopercental method:** This method is very similar to the Isohyetal method but is preferred when dealing with orographic and other topographical differences in mountainous areas. For this method, a catchment map containing the spatial location of each rainfall station and associated rainfall depths (daily, monthly, and annually) must be available. The rainfall data (daily or monthly) are expressed as a percentage of the annual rainfall values to produce isopercental lines. The isopercental lines with the same percentage value and intervals should join each other over the catchment. Isohyetal lines, representative of the annual rainfall values in the region, must be drawn to overlay the isopercental lines. The intersecting points between the isopercental and isohyetal lines are then used to estimate the rainfall. Consequently, the average rainfall over the catchment could be estimated in a similar fashion to the Isohyetal method. However, it should be noted that this method is difficult to implement and is regarded as data intensive (Patra, 2008).

(g) **Spline method:** This method is also based on deterministic interpolation, which provides a smooth rainfall surface based on the point rainfall values as primary input. In other words, it fits a mathematical function to a specified number of nearest input points, while passing through the sample points. This method is recommended for generating gently varying rainfall surfaces, such as frontal rainfall distributed over larger areas as opposed to highly variable, localised convective rainfall (ESRI, 2006).

(h) **Kriging method:** Kriging is based on a geostatistical interpolation process utilising autocorrelation, such as the statistical relationships between point rainfall values. Kriging does not only have the capability of producing a rainfall prediction surface, but also provides some measure of the probability of the predictions. The variation in the surface rainfall can be explained, not only by the distance or direction between the rainfall stations, but also by considering the overall spatial arrangement of the rainfall stations under consideration. The average rainfall for each location is determined by a mathematical function applied to the number of rainfall stations within a catchment or specified radius. The use of Kriging is recommended when the rainfall information is characterised by a spatially correlated distance or directional bias (ESRI, 2006).

2.5 Design Point Rainfall Estimation

Design rainfall comprises of a depth and duration associated with a given return period (T) or AEP (Smithers and Schulze, 2004). Short and long duration design rainfall estimations can either be based on point or regionalised data. Rainfall durations less than 24 hours are generally classified as short, while long durations typically range from 1 to 7 days (Smithers and Schulze, 2004). Several regional and national scale studies in South Africa based on short durations and point data were conducted between 1945 and 2001. Studies focusing on long durations based on daily point rainfall data included studies done by the SAWB (South African Weather Bureau), Schulze (1980), Adamson (1981), Pegram and Adamson (1988), and Smithers and Schulze (2000b). Smithers and Schulze (2000a; 2000b) also used a regionalised approach to increase the reliability of the design values at gauged sites, as well as for the estimation of design values at ungauged sites (Smithers and Schulze, 2003).

2.5.1 Single site approach

A single site approach requires that each rainfall station within the relevant catchment be investigated to determine the record length, data quality (errors, missing data, and outliers) and topographical position (Smithers and Schulze, 2000a). To develop the depth-duration-frequency (DDF) relationship at every single site, the following steps are of importance (Smithers and Schulze, 2000a):

(a) Selection of the most appropriate data set. This may either be the AMS or Partial Duration Series (PDS) with a sufficient record length.
(b) Selection of the most appropriate probability distribution.
(c) Selection of a suitable parameter and quantile method.

These steps are discussed in more detail in the following paragraphs:

A probabilistic analysis needs to be conducted at each rainfall station and it is thus advisable not to use rainfall stations with short record lengths. Furthermore, it is impossible to conclusively select a distribution that could consistently provide adequate rainfall frequency estimates for return periods greater than the period of record. On the other hand, small samples may define a distribution which is markedly different from the parent population (Smithers and Schulze, 2000a).

According to Viessman *et al*. (1989), a minimum record length of 10 years is required, while Schulze (1984) questioned the significance of the record length for extreme events recorded and hence the design values. Hogg (1992) demonstrated that even 20 years of data are not stable enough to estimate the 10-year return period event. Hogg (1992) indicated that the assumptions of stationarity and homogeneity of the AMS of rainfall are seldom valid. It is suggested that a regional approach be used to improve the frequency analysis of extreme rainfall events.

According to Weddepohl (1988), the malfunctioning of rainfall gauges and processing errors are inherent in rainfall data. The spatial density and distribution of rainfall stations, sporadic rainfall events as opposed to the continuous digitised data in use, the length of available records, and the presence of outliers are all problems associated with these errors (Weddepohl, 1988).

The selection of the most suitable probability distribution resembling the probability distribution of the population must be made according to the theoretical basis, consistency, acceptance, user-friendliness, and applicability thereof (Cunnane, 1989). This selection is particularly important when estimating extreme events with return periods greater than the length of record. Equally important factors, such as the type of data in use, data stationarity and the method of fitting the distribution, should also be considered (Cunnane, 1989). The Extreme Value Type I (EV1) distribution has been extensively used in rainfall DDF studies in South Africa since 1963, while the use of the integrated General Extreme Value (GEV) distribution is growing in the application of frequency analysis. Van der Spuy and Rademeyer (2018) proposed the use of the Log-Normal (LN), Log-Pearson Type III (LP3), as well as the GEV using the Method of Moments (MM), Probable Weighted Moments (PWM) or Linear-Moments (LM) to estimate the required design rainfall depths in South Africa. Similarly, Smithers (1996) established that the 3-parameter Log-Normal (LN3) and GEV probability distributions provide consistent short duration design rainfall estimates and therefore recommended the future use thereof in South Africa.

The Technical Report 102 (TR102; Adamson, 1981) is an example of a design point rainfall database based on a single site approach and is commonly used in South Africa. Adamson (1981) estimated the 1, 2, 3 and 7-day extreme design point rainfall depths for return periods of 2, 5, 10, 20, 50, 100 and 200 years using approximately 1 946 rainfall stations. A censored LN distribution based on the PDS was used to estimate the design point rainfall depths at a single site.

2.5.2 Regional approach

Regional frequency analysis is generally based on the assumption that the standardised variate distributions of rainfall data are similar at every single site in a region and that the data from various single sites in a region can thus be combined to generate a single regional rainfall frequency curve representative of any site in the specific region with appropriate site-specific scaling. An advantage of this approach is that it can be used to estimate events at ungauged sites where no rainfall data exists (Alexander, 2001; Cunnane, 1989). In nearly all practical situations, a regional approach is preferred to a single site approach primarily based on the efficiency and accuracy of the rainfall quantile estimation and where statistical homogeneity or heterogeneity might exist (Hosking and Wallis, 1997; cited by Smithers and Schulze, 2003). The large degree of uncertainty introduced in the extrapolation of AEPs beyond the record length of data can also be reduced by regionalisation, since the observed rainfall at a single site is then related to the hydrological response at a regional scale by making use of an extended or combined record length of data (Smithers and Schulze, 2003).

In considering the limitations of a single site approach and the paucity of sub-daily rainfall data in South Africa, *i.e.*, 412 sub-daily rainfall stations and only 49 of these rainfall stations having record lengths exceeding 30 years, Smithers and Schulze (2000a; 2000b; 2003; 2004) developed a regional scale invariance approach to estimate the average point rainfall AMS for any duration and associated 'scaling factors' as an alternative for the 'conversion factors' proposed by Adamson (1981). These 24-hour to 1-day continuous rainfall measurement 'scaling factors' range between 1.14 and 1.30 in South Africa (Smithers and Schulze, 2003).

Smithers and Schulze (2003; 2004) conducted a cluster analysis, based of seven (7) site characteristics, to identify homogeneous regions for extreme daily rainfall in South Africa. The regions identified in the cluster analysis were subsequently tested for homogeneity using a heterogeneity approach developed by Hosking and Wallis (1993; cited by Smithers and Schulze, 2003). This procedure resulted in 15 short and 78 long duration homogenous rainfall clusters, and subsequent regressions to estimate index values (mean *n*-hour AMS values) derived from at-site data. Furthermore, cluster analysis of site characteristics was used to group the 78 long duration homogeneous rainfall clusters into seven (7) regions with six (6) associated region-specific regression parameters applicable to each of the seven (7) regions. Firstly, the average of the 1-day fixed-time interval point rainfall AMS was estimated using regional regression relationships. Thereafter, the average of the 24-hour continuously recorded point rainfall AMS was estimated directly from the 1-day value for the specific site under consideration. Lastly, the average of the point rainfall AMS values for durations shorter and longer than 1-day were scaled directly from the average of the continuous 24-hour and 1-day values, respectively, using the established regression parameters.

The up- and downscaling were found to scale linearly as a function of the average 1-day and continuous 24-hour values, respectively. In the application of the regression relationships to estimate the average of the AMS for durations shorter and longer than 1 day, inconsistencies in the growth curves derived from the 24-hour continuously recorded and daily rainfall data were evident due to the quality and non-concurrent periods of the digitised rainfall data, as well the differences in the AMS extracted from: (i) continuously recorded data using a sliding window, and (ii) daily rainfall data using a fixed period window. However, Smithers and Schulze (2003) found that the best estimate of the growth curve, for all durations, are the values derived from the daily rainfall database for the 1-day duration.

As a result, a scale invariance approach was introduced to the Regional Linear Moment Algorithm and termed the RLMA&SI approach to address the inconsistencies evident in the above-mentioned growth curves (Smithers and Schulze, 2003). In South Africa, the RLMA&SI approach is the preferred method for design rainfall estimation and is automated and included in the software program, *Design Rainfall Estimation in South Africa* (Smithers and Schulze, 2003; 2004). The latter software facilitates the estimation of design rainfall depths at a spatial resolution of 1-arc minute, for any location in South Africa, for durations ranging from 5 minutes to 7 days and for return periods of 2 to 200 years.

Irrespective of whether a single site or regional approach is adopted, the design rainfall depth to be used in design flood estimation, especially in the deterministic methods, must be based on the critical storm duration or time of concentration (T_C) of a catchment. Thus, depending on the T_C, the daily design rainfall depth used in flood estimations must either be increased or decreased to adapt to the T_C requirement. To convert the daily design rainfall depth values to independent durations of the same length, conversion and/or scaling factors must be used. The conversion factors are dependent on the duration in question and various values have been proposed.

The use of conversion factors (Adamson, 1981) is generally accepted in South Africa to convert 1-day fixed time interval rainfall (08:00 to 08:00) to continuous measures of *n*-hour rainfall associated with T_C. Adamson (1981) proposed the use of a conversion factor of 1.11 to convert daily rainfall depths recorded at fixed 1-day intervals to continuous 24-hour rainfall depths. At an international level, similar conversion factors have been proposed to convert daily fixed time interval rainfall depths to continuous 24-hour maxima, *e.g.*, 1.13 in the USA (Hershfield, 1962), 1.06 in the UK (NERC, 1975), and 1.13 in South Africa (Alexander, 1978). To convert continuous 24-hour rainfall series to critical storm or T_C durations ranging between 0.10 hour and 24 hours, Adamson (1981) proposed the use of the conversion factors as listed in Table 2.1.

Table 2.1: Conversion of continuous 24-hour rainfall depths to T_C-hour rainfall depths (Adamson, 1981)

T_C [hours]	Conversion factor Summer/inland region	Conversion factor Winter/coastal region
0.10	0.17	0.14
0.25	0.32	0.23
0.50	0.46	0.32
1	0.60	0.41
2	0.72	0.53
3	0.78	0.60
4	0.82	0.67
5	0.84	0.71
6	0.87	0.75
8	0.90	0.81
10	0.92	0.85
12	0.94	0.89
18	0.98	0.96
24	1.00	1.00

The conversion factors listed in Table 2.1 are independent of return period but are influenced by regional climatological differences as evident in the summer rainfall/inland and winter rainfall/coastal regions of South Africa (Midgley and Pitman, 1978). Converting daily rainfall depths to durations longer than 1-day simply entails the conversion of fixed time interval rainfall to continuous measures of rainfall (*e.g.*, 2-days to 48-hour, 3 days to 72-hour, *etc.*), and interpolating between the different T_C durations as listed in Table 2.2. The conversion factors listed in Table 2.2 are normally used in practice (Van der Spuy and Rademeyer, 2018); however, no literature is available as to how these conversion factors were derived.

Table 2.2: Conversion of fixed time interval rainfall to continuous estimates of *n*-hour rainfall (Van der Spuy and Rademeyer, 2018)

Duration		Conversion factor
From [days]	To [hours]	
1	24	1.11
2	48	1.07
3	72	1.05
4	96	1.04
5	120	1.03
7	168	1.02
> 7	> 168	1

However, the latter South African approaches as listed in Tables 2.1 and 2.2 are regarded as outdated. Smithers and Schulze (2000a) developed 15 individual regionalised factors, varying from 1.15 to 1.28, for the 15 relatively homogeneous short duration rainfall clusters in South Africa, with a national average of 1.21.

2.6 DWS Catchment Design Rainfall Estimation

Catchment design rainfall, also known as areal design rainfall, is usually obtained when an ARF is applied to a design point rainfall value associated with a specific storm duration and return period. In addition to the approaches discussed in Sections 2.5.1 and 2.5.2, to estimate design rainfall at a point, four different approaches are also used by the Department of Water and Sanitation (DWS) to estimate catchment design rainfall (Van der Spuy and Rademeyer, 2018). The first approach, referred to as the 'Smithers Regional Rainfall (SRR)' approach, is in essence the RLMA&SI approach as discussed in Section 2.5.2.

The remaining three approaches are summarised as follows (Van der Spuy and Rademeyer, 2018):

(a) **Maximum Station Rainfall (MSR) approach:** The rainfall data at a single rainfall station are probabilistically analysed by using either the observed or infilled rainfall data series. This approach is similar to a conventional single site approach as discussed in Section 2.5.1 and would therefore still require an ARF to convert the estimated MSR design rainfall values to a catchment design rainfall value.

(b) **Maximum Catchment Rainfall (MCR) approach:** The weighted AMS catchment rainfall data based on either the full observed or infilled record length (if applicable) are probabilistically analysed. The use of an infilled record length ensures that the longest possible record length is utilised in the analysis. The probabilistically analysed weighted AMS catchment rainfall will thus result in catchment design rainfall values.

(c) **Daily Catchment Rainfall (DCR) approach:** This approach requires the weighted point rainfall at a daily time interval within a specific catchment (Van der Spuy and Rademeyer, 2018). The weighted daily catchment rainfall is then probabilistically analysed to obtain areal design rainfall which incorporates the temporal and spatial variation of predominant weather types in a catchment. Van der Spuy and Rademeyer (2018) also highlighted that ARFs are not applicable to this approach since the areal design rainfall is already representative of a geographically-fixed area. A similar approach was followed by Dyrrdal *et al.* (2016) in Norway where a direct probabilistic analysis was performed on the average areal 24-hour rainfall from a gridded data set for the period from 1957 to 2016.

2.7 Factors Influencing ARFs

Numerous factors can have a significant influence on the estimation of ARFs, *e.g.*, climatological variables, catchment geomorphology, methodological approaches, climatological regions, return periods, storm durations and/or a combination of these (Asquith and Famiglietti, 2000; Svensson and Jones, 2010; Kim *et al*., 2019). All these factors are discussed in this section to highlight their individual influences.

2.7.1 Climatological variables

Geographical location within different climatological regions has a direct influence on ARFs. It was established that the 1-day ARFs in the USA exceeded the equivalent ARF estimates in Australia, while the ARFs decline more rapidly in the semi-arid south-western USA than in the rest of the USA (Svensson and Jones, 2010). Similar trends were also confirmed by Asquith and Famiglietti (2000), who established that the ARFs are higher in the eastern USA than in Texas. ARFs are also influenced by seasonal variability, *e.g.*, higher values are obtained in winter than in summer. This could be ascribed to the response to higher convective activity in summer (Allen and DeGaetano, 2005).

Different rainfall-producing mechanisms, *e.g.*, convective versus frontal rainfall, will produce different spatial rainfall patterns. Typically, the spatial averages for large-scale frontal rainfall do not reduce much in magnitude with increasing area, whereas this is the case for small-scale convective rainfall events (Skaugen, 1997). Skaugen (1997) also established that ARFs for both convective and frontal rainfall decrease with an increasing return period, but the rate of decrease for convective rainfall is noticeably larger than that for frontal rainfall. The decrease in ARFs with increasing return periods may also reflect the importance of convection in producing very high point rainfalls. Huff and Shipp (1969) highlighted that the spatial correlation decay pattern of low-pressure centred storms is smaller compared to fronts associated with mid-latitude cyclones, while it is the greatest in weaker storms with a shorter duration.

In the USA, areal rainfall was found to decrease with the corresponding point rainfall and with increasing return periods (Asquith and Famiglietti, 2000; Allen and DeGaetano, 2005). In contrast, Grebner and Roesch (1997) demonstrated that ARFs in Switzerland (A > 4 500 km²) are independent of the return period. The ARFs contained in the Centre for Ecology and Hydrology (CEH; Faulkner, 1999), as originally documented in the UK FSR (NERC, 1975), decrease more rapidly (with increasing catchment areas) for shorter storm durations than for longer storm durations. It was also confirmed that the ARFs derived using a storm-centred approach are independent of the return period and geographical location (Svensson and Jones, 2010). Alexander (2001) recommended a geographically-centred approach when assuming a uniform spatial and temporal rainfall distribution for the total storm duration over the whole catchment area.

Alexander (2001) also emphasised that practitioners using storm-centred data to derive ARFs, should not assume uniform rainfall intensity distribution over the catchment. Kim *et al.* (2019) highlighted that one of the disadvantages of a geographically-centred approach based on a rainfall-monitoring network in a given catchment, is that it will not necessarily reflect the actual spatial characteristics of a particular rainfall event.

2.7.2 Catchment geomorphology

Most research conducted on the estimation of ARFs concluded that catchment geomorphology (*e.g.*, area, shape, and topography) has an insignificant influence on ARFs (Svensson and Jones, 2010). In catchments with areas less than 800 km^2, ARFs are mainly a function of the area and point intensity since the relationship between rainfall intensity and the infiltration rate of the soil is predominant. In catchment areas up to 30 000 km^2, ARFs are mainly a function of the area and storm duration (Alexander, 2001; SANRAL, 2013). Lambourne and Stephenson (1986) demonstrated that the ARF will decrease from unity with an increasing catchment area.

ARFs could also vary between urban areas and the surrounding rural areas. Huff (1995) showed that eight storms in Chicago, USA, had a slower ARF decreasing rate within 500 km^2 from the urban storm centre compared to the 67 rural storms. Bárdossy and Pegram (2018) questioned the application of ARFs in urban areas, since towns and cities are not fixed catchments with a single outlet. Furthermore, they argued that urban areas have many sub-catchments where the drainage system will determine the stormwater drainage.

Veneziano and Langousis (2005) highlighted that empirical ARFs are, to a degree, influenced by the catchment shape. Different ARF estimates could be expected in catchments with an elongated shape where the rainfall distribution patterns, and direction of movement, are aligned along the catchment or perpendicular to it. Topography (*e.g.*, hills and mountains) has leeward and windward effects on rainfall and may affect ARFs. Rainfall-monitoring networks also tend to be sparser at higher altitudes; hence, resulting in poorer areal rainfall estimates. Singh *et al.* (2018) highlighted that ARF differences in New Zealand are ascribed to differences in topography and rainfall types. However, Allen and DeGaetano (2005) found that topographical rainfall biases are insignificant for the estimation of ARFs.

2.7.3 Methodological approaches

The record length of rainfall data and frequency of data collection may influence ARFs due to temporal rainfall variability. Asquith and Famiglietti (2000) showed that three overlapping rainfall-monitoring networks around Houston in Texas, USA did not yield the same ARFs due to different rainfall-monitoring networks that cannot be indiscriminately combined. However, Allen and DeGaetano (2005) showed that the density of rainfall-monitoring networks and the use of different interpolation methods have an insignificant influence on the estimation of ARFs in North Carolina and New Jersey, USA. Asquith and Famiglietti (2000) estimated probabilistically correct ARFs and proved that the return period has a significant influence on the relationship between the ratio of the annual maxima to concurrent rainfall depth and on the separation distance from the annual maxima point rainfall. Sugawara (1992) highlighted that Thiessen polygons, when applied to a low-density rainfall-monitoring network, may not represent the actual rainfall processes in a specific polygon as they interrupt the spatial transitions of the natural occurring process.

Unfortunately, the two recognised approaches, namely, the storm-centred and geographically-centred approaches, used to estimate ARFs, generally provide inconsistent results. In using a storm-centred approach, the isohyets of a complete storm are analysed without considering the geographical location thereof (Alexander, 2001). In the case of a geographically-centred approach, storms occurring over a fixed area or collection of rainfall stations within a catchment are considered (Alexander, 2001). Bell (1976) highlighted that the theoretical significance of the geographically-centred approach is more statistical than physical and is therefore best interpreted in terms of average areal point rainfall frequency curves, which simply provides the ratios of areal to point rainfall with the same AEP. It is thus quite evident that the use of different methodologies to estimate ARFs is likely to result in different ARF estimates. Previous studies, *e.g.*, Bell (1976), Stewart (1989), and Kim *et al.* (2019) have shown that ARFs will reduce with an increasing return period.

ARF estimation methods can be grouped into two broad categories, *i.e.*, analytical, and empirical methods, as discussed in the subsequent sections.

2.8 Analytical ARF Estimation Methods

In using analytical methods, derived mathematical algorithms are used to characterise the spatial and temporal rainfall variability by incorporating simplified assumptions that are not entirely true descriptions of the actual rainfall process (Siriwardena and Weinmann, 1996; Svensson and Jones, 2010). The fact that the actual rainfall processes are partially ignored is regarded as problematic. In response to these inherent shortcomings, several new analytical methods to estimate ARFs have been proposed during the last three decades, *e.g.*, storm movement (Bengtsson and Niemczynowicz, 1986), crossing properties (Bacchi and Ranzi, 1996), spatial correlation structure (Sivapalan and Blöschl, 1998), and scaling relationships (De Michéle *et al*., 2001).

2.9 Empirical ARF Estimation Methods

The empirical methods can either be based on a geographically-centred or storm-centred approach. Both these approaches are explained below, while the methods currently used in national and/or international practice are also listed.

2.9.1 Geographically-centred methods

This approach describes the relationship between average areal design rainfall over a geographically fixed area, *i.e.*, catchment, and a corresponding average design point rainfall value representative of the area under consideration. In other words, the ARF is used for percentage reduction, which relates to the statistics of design point and areal design rainfall and considers the uniform temporal and spatial distribution of rainfall over the catchment area (Pietersen *et al*., 2015).

(a) **United States Weather Bureau (USWB, 1957; 1958):** Observed rainfall records (10 to 15 years) were obtained from dense rainfall-monitoring networks in catchment areas ranging from 250 km² to 1 000 km². The rainfall record lengths were regarded as insufficient to establish the effect of return period on the point-area rainfall relationships. The areal rainfall of each event and associated duration was estimated using Thiessen weights. The average of the AMS was estimated, while the highest point rainfall measurement at each station in a particular year was selected for the derivation of Eq. (2.5).

$$ARF_T = \frac{N \sum_{i=1}^{N} \sum_{j=1}^{n} w_i \overline{\overline{P_{ij}}}}{\sum_{i=1}^{N} \sum_{j=1}^{n} P_{ij}} \qquad [2.5]$$

where

ARF = areal reduction factor,

N = number of stations within the catchment area,

n = record length [years],

$\overline{\overline{P_{ij}}}$ = point rainfall for station i on the day of the annual maximum areal rainfall in year j [mm],

P_{ij} = annual maximum point rainfall of station i in year j [mm], and

w_i = Thiessen weighted factor for station i.

(b) **Bell's method (1976):** Probabilistic rainfall analyses at single rainfall stations (14-year record length) situated in circular catchment areas of 1 000 km² each were conducted in the UK to estimate areal and average design point rainfall frequency curves and to estimate ARFs. The ARFs were expressed as the ratio of areal to average design point rainfall with associated AEPs.

A modified Thiessen weighting procedure was used to estimate the daily areal rainfall values, after which, these values were ranked to obtain the 20 highest independent values for each sample area (Bell, 1976). In other words, a PDS using equally ranked observations curtailed to a common base period, were used, and fitted to an exponential distribution with parameters estimated by the Method of Maximum Likelihood (MML). The average design point rainfall frequency curves were estimated using the 20 highest daily rainfall values at each rainfall station (Bell, 1976). Instead of deriving separate frequency curves for each rainfall station to estimate weighted averages, a simpler equivalent procedure was adopted. Each ranked weighted average point rainfall value was determined using the same modified Thiessen weighting procedure, followed by the exponential distribution curve fitting to provide estimates of the average design point rainfalls for return periods from 2 to 20 years. The ARFs were then estimated directly using the corresponding areal and average design point rainfall values associated with each return period or AEP (Siriwardena and Weinmann, 1996).

Bell (1976) concluded that this procedure is probabilistically more correct due to the inclusion of AEPs, while the derived ARFs based on daily (24-hour) and sub-daily rainfall (1-hour to 2-hours), proved to vary between 5% and 10%, respectively. The ARF estimates also compared reasonably well with the 2 to 20-year return period ARF estimates contained in the UK FSR (NERC, 1975), while, for the higher return periods (50- to 100-year), the ARF estimates were significantly lower (Siriwardena and Weinmann, 1996). The mathematical relationship representative of Bell's (1976) method is shown in Eq. (2.6).

$$ARF_m = \frac{\sum_{i=1}^{N}\left(w_i \overline{\overline{P_{ij}}}\right)_m}{\sum_{i=1}^{N}\left(w_i P_{ij}\right)_m} \qquad [2.6]$$

where

ARF_m = areal reduction factor [ratio of the areal rainfall of rank m to the Thiessen weighted average point rainfall of the same rank [%],

m = rank value,

N = number of rainfall stations within the catchment area,

$\overline{P}$ = point rainfall for station i on the day of the annual maximum areal rainfall in year j [mm],

P_{ij} = annual maximum point rainfall of station i in year j [mm], and

w_i = ratio of the areal rainfall of rank m to the Thiessen weighted average point rainfall of the same rank.

The ARF estimation methodology proposed by Bell (1976), is not only probabilistically more correct than the USWB (1957; 1958) and United Kingdom Flood Studies Report (UK FSR; NERC, 1975) ARF estimation methods, but the variation of ARFs with return period is also evident when ARFs are directly obtained from the areal and design point rainfall frequency curves.

In South Africa, Du Plessis and Loots (2019) evaluated and compared Bell's method (1976) against the geographically and storm-centred approaches currently recommended for general use. Nineteen fixed catchment areas ($1\,010\ \text{km}^2 \leq A \leq 9\,270\ \text{km}^2$), three to 10 rainfall stations per catchment ($35 \leq N \leq 83$ years), and at-site probabilistic analyses (GEV_{MM} probability distribution and Gringörten plotting position), were used. It was established that ARFs do vary with return period, while Bell's method is likely to result in higher ARFs and subsequently more conservative areal design rainfall estimates. It was recommended that Bell's approach should be further developed by using the longer record lengths and more advanced computing power currently available.

(c) **Alexander's method (1980; 2001):** This is the geographical-centred method currently used and recommended for general use in South Africa (SANRAL, 2013). Alexander (1980) developed a geographically-centred ARF relationship based on the ARF diagrams contained in the UK FSR (*cf.* Figure 2.4; NERC, 1975). This ARF diagram (Figure 2.4) had an adjustment made to account for short duration rainfall over small catchment areas, which are mostly characterised by severe storm mechanisms producing very high intensity rainfall with cell core areas exceeding 10 km^2 and durations exceeding 10 minutes.

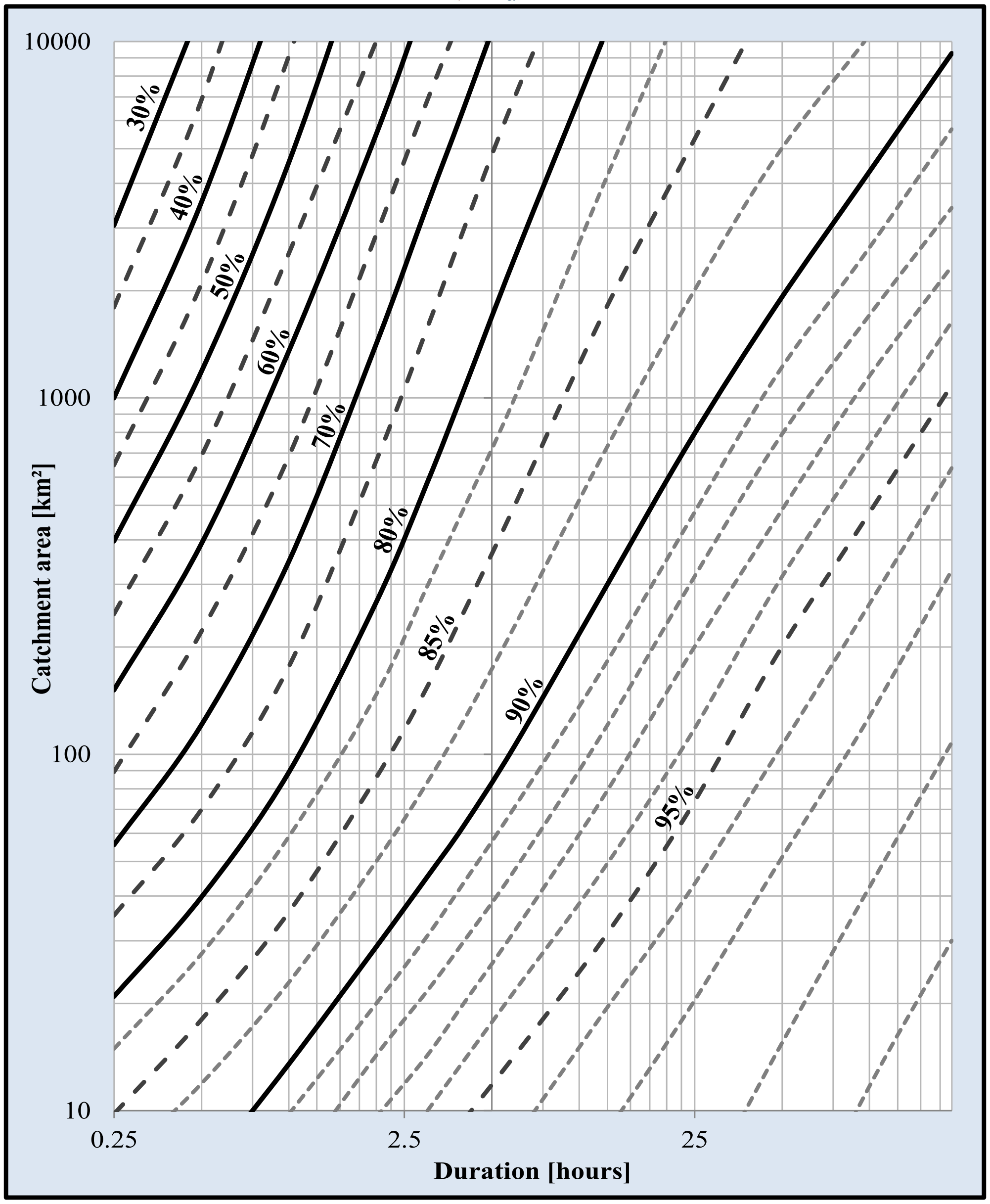

Figure 2.4: UK FSR ARF diagram (NERC, 1975)

According to Alexander (1980), estimates of shorter duration rainfall based on the extrapolation from longer durations, are regarded as unreliable when viewed in the light of the storm mechanisms which produce high-intensity rainfall for durations less than 10 minutes. Thus, there is little justification in assuming ARFs less than 100% in these area and duration regions; consequently, the UK FSR values were adjusted accordingly.

The UK FSR ARFs adopted for South African conditions are shown in Figure 2.5.

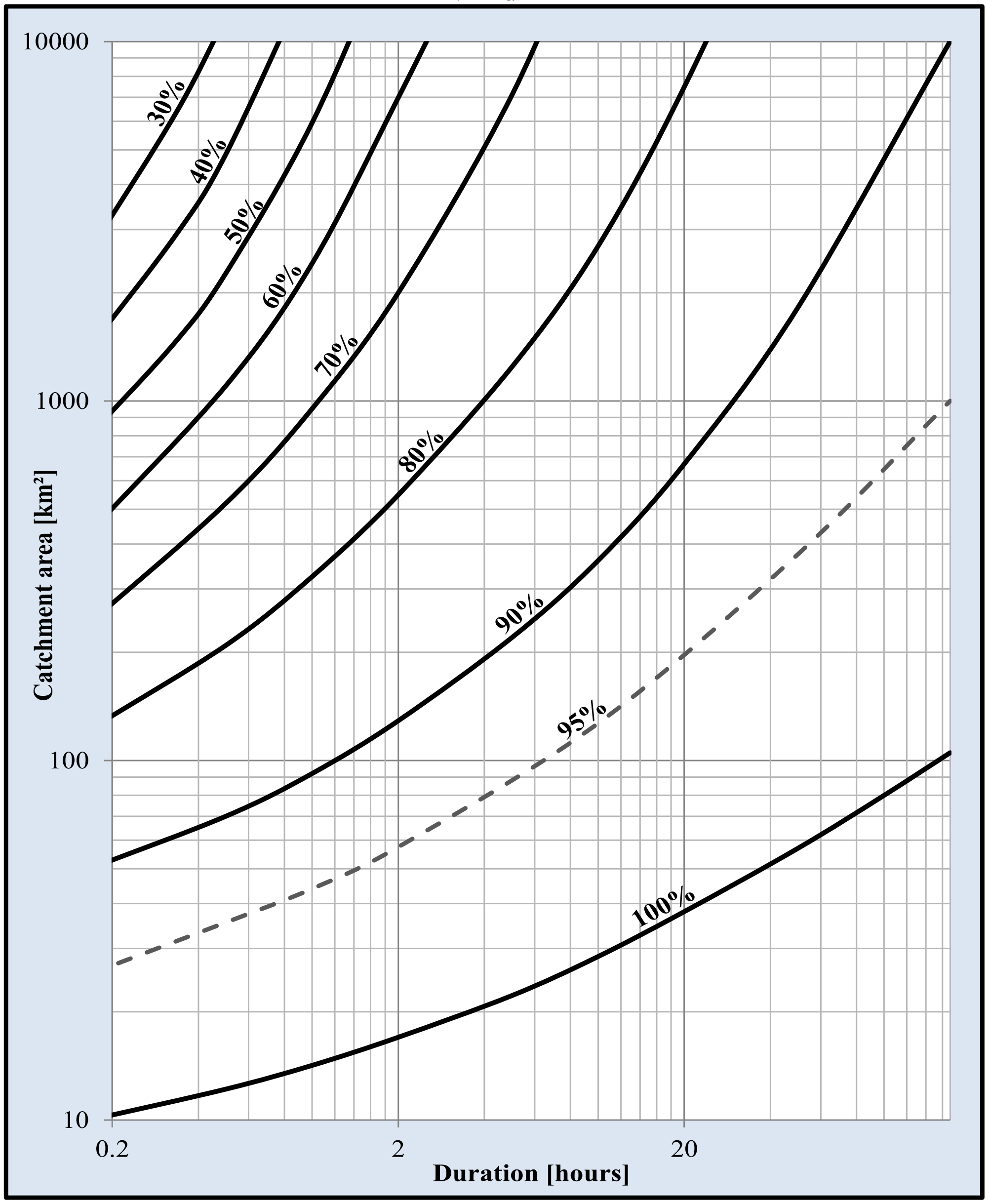

Figure 2.5: Adopted UK FSR ARFs for South Africa (after Alexander, 1980)

Alexander (2001) revised the ARF diagram in Figure 2.5 to a more reliable and user-friendly diagram that is currently used by practitioners (SANRAL, 2013).

The revised version of Figure 2.5 is shown in Figure 2.6.

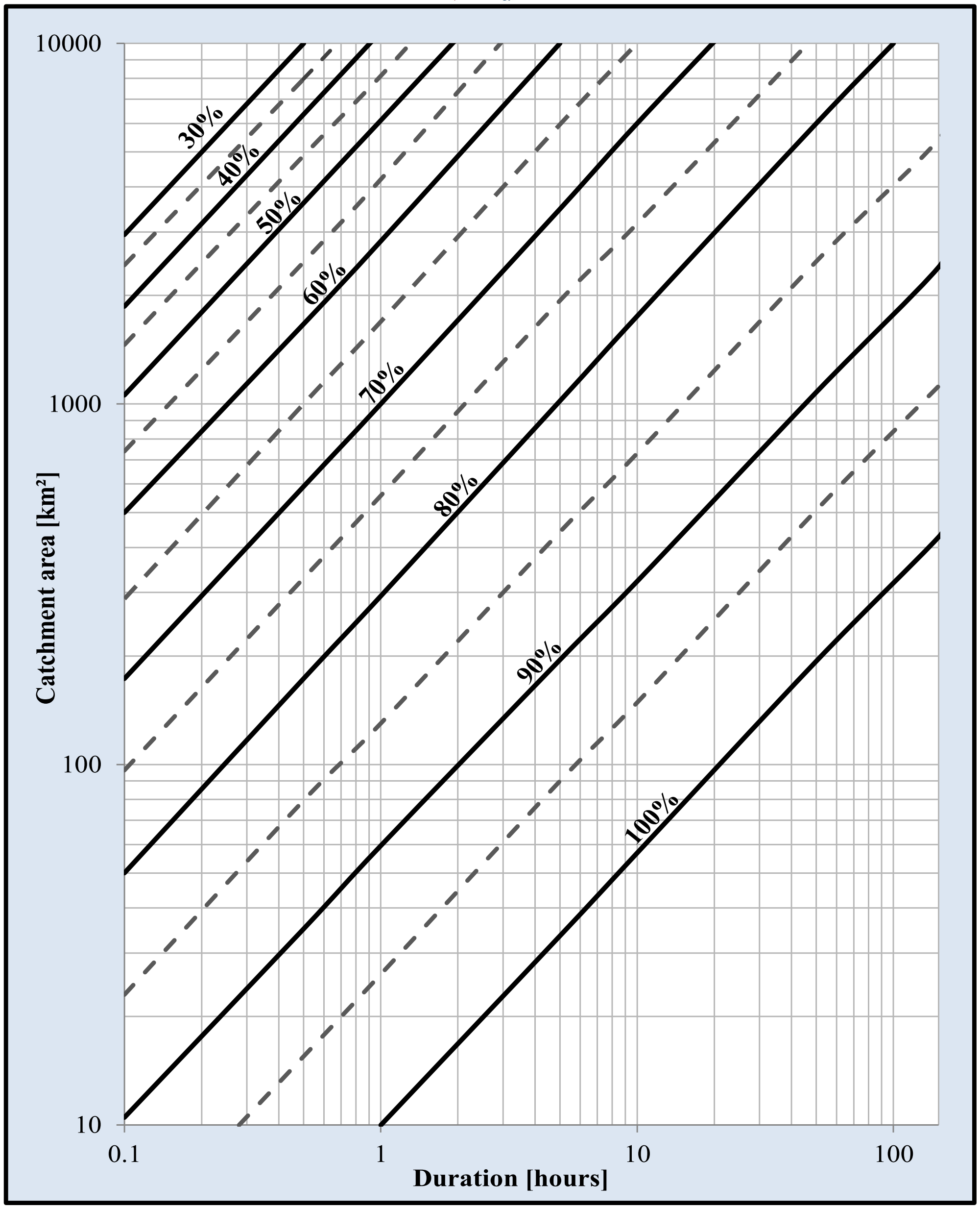

Figure 2.6: Revised ARF diagram for South Africa (Alexander, 2001)

Op ten Noort and Stephenson (1982) converted Figure 2.5 to a mathematical relationship [Eq. (2.7a)], using regression analysis. Alexander (2001) expressed Figure 2.6 in the form of a mathematical relationship as shown in Eq. (2.7b). The use of both Eqs. (2.7a) and (2.7b) resulted in slightly more conservative results when compared to the UK FSR and USWB values, respectively. Alexander (2001) recommended that the ARF relationship shown in Eq. (2.7b) should be used in South Africa where the average rainfall depth over a catchment must be established from point rainfall statistics.

$$ARF = [1.306 - 0.0902Ln(A)] + Ln(T_d)[0.0161Ln(A) - 0.0498] \quad [2.7a]$$

$$ARF = [90000 - 12800Ln(A) + 9830Ln(60T_C)]^{0.4} \quad [2.7b]$$

where

ARF = areal reduction factor [%],

A = catchment area [km²],

T_C = time of concentration/critical storm duration [hours], and

T_d = storm duration [hours].

Gericke and Du Plessis (2011) established that a relationship between the A, T_C and ARFs exists. The validity of Eq. (2.7b) was assessed by plotting T_C within each catchment under consideration against the catchment area, after which a power-law curve fitted through the data points was superimposed on Figure 2.6 and the original ARF diagram as published in the UK FSR (NERC, 1975). The fitted power-law relationship was expressed as Eq. (2.8a), which provides a good indication of T_C associated with any catchment area under consideration. Equation (2.8b) resulted from the substitution and simplification of Eq. (2.8a) into Eq. (2.7b).

$$T_C = 0.2284A^{0.596} \quad [2.8a]$$

$$ARF = [-6944.3Ln(A) + 115731.9]^{0.4} \quad [2.8b]$$

where

ARF = areal reduction factor [%],

A = catchment area [km²], and

T_C = time of concentration [hours].

The results obtained from this study clearly indicated that the power-law curve yielded a constant ARF range of between 87% and 88% across the original UK FSR ARF diagram for durations exceeding three hours. Gericke and Du Plessis (2011) concluded that the ARF relationship expressed by Eq. (2.8b) can be used instead of Eq. (2.7b) to convert design point rainfall depths or intensities to an average areal design rainfall depth or intensity.

(d) **Stewart (1989):** This method [Eq. (2.9)] is based on Bell's method (1976) using daily rainfall data from north-west England. A total of 834 rainfall stations with at least

25 years of data were used. This method was amongst the first methods to combine gauged and radar-based rainfall data. A total of 544 sample catchments, ranging from 25 km² to 10 000 km², and storm durations ranging from 1 to 8 days, were analysed. The ARFs were expressed as a function of the geographical location and AEP and decreased with an increasing catchment area and AEP. The ARF estimates proved to be significantly lower than those based on the UK FSR method (NERC, 1975).

$$ARF_T = \frac{P_{AS(T)}\overline{P_A}}{P_{PS(T)}\overline{P_P}} \quad [2.9]$$

where

ARF_T = areal reduction factor for specific AEP [%],

$\overline{P_A}$ = mean of annual maximum areal rainfall [mm],

$\overline{P_P}$ = mean of annual maximum point rainfall [mm],

$P_{AS(T)}$ = standardised T-year areal rainfall [mm], and

$P_{PS(T)}$ = standardised T-year point rainfall [mm].

(e) **Omolayo (1993):** This data intensive method [Eq. (2.10)] is based on daily rainfall data exceeding 30 years of record, while the point rainfall values being used were assumed to be representative for the entire catchment areas under consideration. The ARF was defined as the ratio between the average areal and point rainfall of the same return period/AEP; hence, the ARF estimates are probabilistically correct, given the variation with return period. For comparison purposes, the 1-day ARFs applicable to the USA were transposed to Australia, assuming that the climatological variables are similar.

$$ARF = \frac{P_A(A, T_d, T)}{\frac{1}{\sum_i w_i}\sum_i\left[w_i\overline{P_P}(T_d, T)\right]} \quad [2.10]$$

where

ARF = areal reduction factor,

A = catchment area under consideration [km²],

P_A = T-year areal rainfall [mm],

$\overline{P_P}$ = average T-year point rainfall [mm],

T = return period [years],

T_d = storm duration [hours], and

w_i = weighted average P_P of the gauges i in the same region.

(f) **Siriwardena & Weinmann (1996):** This method [Eq. (2.11)] is also regarded as a modified version of Bell's (1976) method, since the AMS of areal and point rainfall were used instead of the PDS curtailed to a common base period. Over 2 000 daily rainfall stations in Victoria, Australia were used and the ARF values were estimated for 'circular sample catchment areas' distributed through areas characterised by a high-density rainfall-monitoring network. ARF values were estimated for storm durations (1 to 3 days), catchment areas (125, 250, 500, 1 000, 4 000 and 8 000 km²) and return periods (2 to 200 years).

$$ARF_m = 1 - 0.4\left(A^{0.14} - 0.7\log T_d\right)T_d^{\;0.48} + 0.002\,A^{0.4}T_d^{\;0.41}\left[0.3 + \log\left(\frac{1}{T}\right)\right] \quad [2.11]$$

where

ARF_m = areal reduction factor,

A = catchment area [1–10 000 km²],

T = return period [20–2 000 years], and

T_d = storm duration [18–120 hours].

(g) **Yoo *et al.* (2007):** This method [Eq. (2.12)], a variant of the geographically-centred approach, is based on a mixed gamma distribution and the concept of rainfall intermittency, *i.e.*, wet and dry periods, with a continuous gamma distribution fitted to the wet periods. A catchment area of 9 843 km² and 25 rainfall stations, with a minimum record length of 30 years each, were used. Return periods ranging from 2 to 1 000 years were considered, while the ARF estimates were based on the 1-day storm durations.

$$ARF_{(A,T)} = 1 - Me^{-\left(aA^{b}\right)^{-1}} \quad [2.12]$$

where

ARF = areal reduction factor,

A = rainfall storm areas [km²],

e = mathematical constant,

M, a, b = parameters associated with each return period, and

T = return period [years].

(h) **Podger *et al*. (2015a; 2015b):** This method is based on a modified version of Bell's method (Bell, 1976; Siriwardena and Weinmann, 1996). It considers catchment area, storm duration and AEP. Artificial circular catchments in areas with sufficient data were used to generate areal rainfall time series. This was done by weighting point rainfall values by means of Thiessen polygon areas. Frequency quantiles were calculated from areal rainfall time series and were divided by the weighted point frequency quantiles. This process was repeated and subsequently resulted in several ARF estimates. All the estimated ARF values were then averaged across the regions and an equation was fitted to result in a prediction model for the selected regions (Ball *et al*., 2019).

The research conducted by Podger *et al*. (2015a) focused on deriving ARFs associated with long duration rainfall (> 24-hour), while Podger *et al*. (2015b) focused on deriving ARFs associated with short duration rainfall (≤ 24-hour). Podger *et al*. (2015a; 2015b) pre-defined different ranges of storm durations for the user to select an equation to estimate the corresponding ARFs. The storm duration ranges with corresponding ARF equations are: (i) duration ≥ 12-hours [Eq. (2.13a)], (ii) duration > 24-hours [Eqs. (2.13b) and (2.13d)], and (iii) 24-hours ≤ duration ≤ 168-hours [Eqs. (2.13c) and (2.13d)]. These equations can be used to estimate ARFs for catchments areas up to 30 000 km^2, storm durations ≤ 168 hours and AEP > 0.05% (Ball *et al*., 2019).

<u>Short duration (i), *A* (1 – 1 000 km^2):</u>

$$ARF_S = Min\left\{1, \left[\begin{array}{l} 1\text{-}0.287\left(A^{0.267} - 0.439\log_{10} D\right)D^{-0.36} + \\ 0.0023A^{0.226}D^{0.125}\left(0.3 + \log_{10} P\right) + \\ 0.014A^{0.213} \times 10^{-0.021\frac{\left(D-180^2\right)}{1440}}\left(0.3 + \log_{10} P\right) \end{array}\right]\right\} \quad [2.13a]$$

<u>Interpolation equation for duration (ii), *A* (1 – 30 000 km^2):</u>

$$ARF = ARF_{12hour} + \left(ARF_{24hour} - ARF_{12hour}\right)\frac{D - 720}{720} \quad [2.13b]$$

Long duration (iii), A (1 – 30 000 km^2):

$$ARF_L = Min\left\{1, \left[\begin{array}{l} 1-\left(A^{b} - C\log_{10} D\right)D^{-d} + \\ eA^{f} D^{g}\left(0.3+\log_{10} P\right)+ \\ h10^{iA\frac{D}{1440}}\left(0.3+\log_{10} P\right) \end{array}\right]\right\} \quad [2.13c]$$

Interpolation equation for durations (i), (ii) and (iii), A (1 – 10 km^2):

$$ARF = 1 - 0.6614\left(1 - ARF_{10km^2}\right)\left(A^{0.4} - 1\right) \quad [2.13d]$$

where

ARF = areal reduction factor [factor],

A = catchment area [km^2],

D = duration [minutes],

P = AEP, fraction between [0.5 and 0.0005], and

a–i = regional coefficients applicable to Australia.

Equations (2.13a) to (2.13d) are respectively the recommended methods for estimating ARFs in Australia and are documented in the Australian Rainfall and Runoff Manual (Ball *et al.,* 2019).

2.9.2 Storm-centred methods

In this approach, the region over which the areal design rainfall is estimated, is not fixed, but changes for each storm (Alexander, 2001; Svensson and Jones, 2010). The centre point for the approach is characterised by the maximum rainfall, which also changes for each storm. In other words, the ARF relates to the way in which rainfall intensity decreases with distance from the central core of individual storm events, with the average areal design rainfall intensity being estimated (Alexander, 2001; Svensson and Jones, 2010). Sivapalan and Blöschl (1998) noted that storm-centred ARFs are usually somewhat lower than geographically-centred ARFs as the ARFs are either derived from convective storms with heavy design point rainfall and a limited areal extent, or the highest design point rainfall might be located outside the catchment boundary. The methods currently used in national and/or international practice, are listed below:

(a) **Van Wyk (1965):** The first South African attempt to analyse ARFs based on a storm-centred approach was conducted by Van Wyk (1965; cited by Lambourne and

Stephenson, 1986) on a small-scale (catchment areas ≤ 800 km²) in Pretoria, Gauteng Province. In addition, a few rainfall storm areas from the USA and Canada were also analysed for comparison purposes. Isohyetal maps of several storms were plotted based on the average areal rainfall depths in catchments ranging from 10 km² to 800 km² centred on the maximum point rainfall and expressed as a percentage of point rainfall at the storm centre (Van Wyk, 1965; cited by Lambourne and Stephenson, 1986). As a result, depth-intensity-area envelope diagrams were developed, as shown in Figure 2.7.

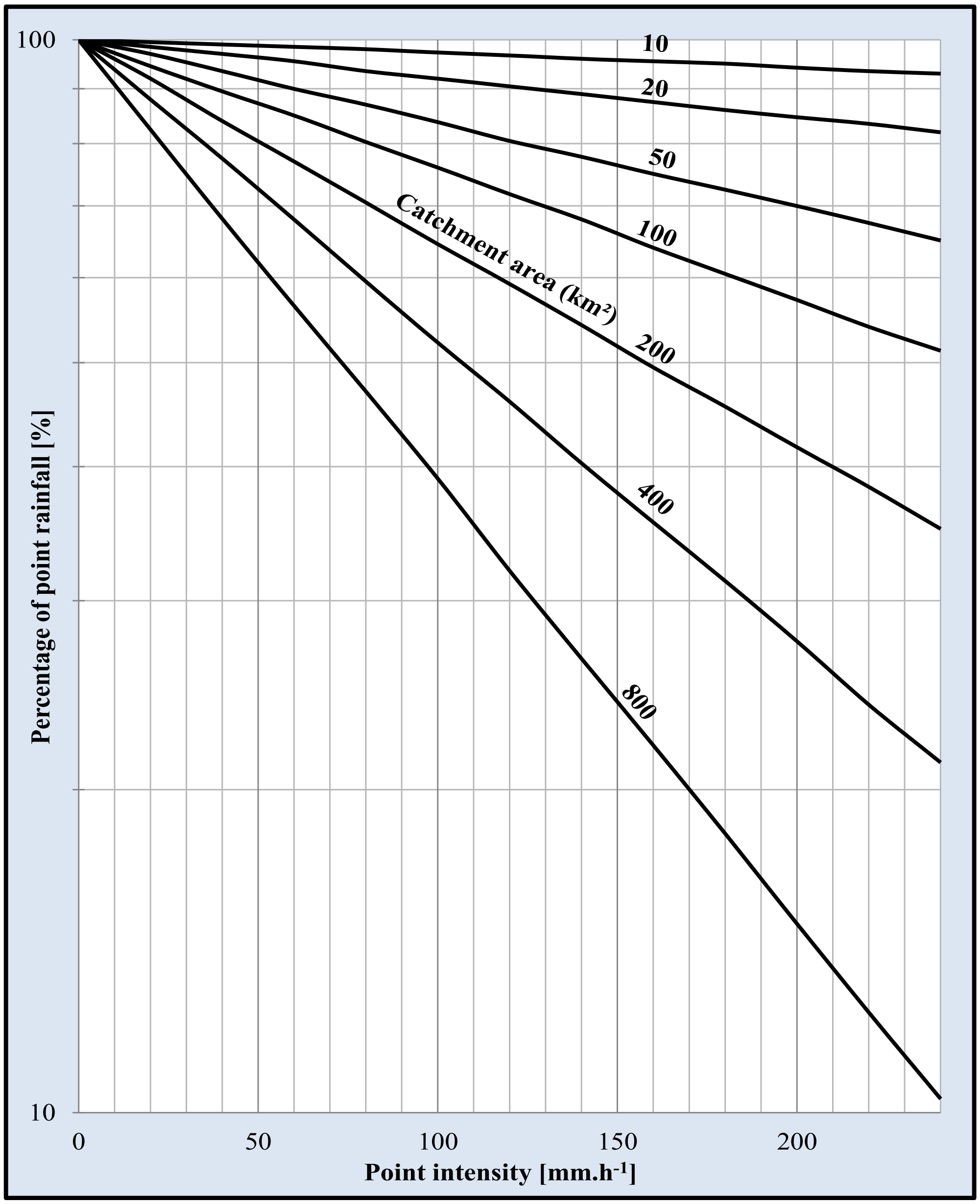

Figure 2.7: Expected percentage of runoff as a function of point rainfall intensity (SANRAL, 2013)

The ARFs were also expressed as a function of the point rainfall intensity, particularly an average intensity over the storm duration at the storm centre (Van Wyk, 1965; cited by Lambourne and Stephenson, 1986). The impact of return periods on ARFs is clearly evident when considering the same catchment area, *i.e.*, ARFs decrease with an increase in rainfall intensity (which is associated with a decrease in storm duration) for the same catchment area.

Op ten Noort and Stephenson (1982) used regression analyses to convert Figure 2.7 into a mathematical relationship [Eq. (2.14)]. In small to medium catchment areas ($\leq$ 800 km²), the ARF is mainly a function of the area and design point rainfall intensity, since the relationship between rainfall intensity and the infiltration rate of the soil is predominant (Alexander, 2001; SANRAL, 2013).

$$ARF = Exp(-0.000068 i A^{0.77}) \quad [2.14]$$

where

ARF = areal reduction factor for point rainfall [%],

A = catchment area [km²], and

i = point rainfall intensity at the storm centre [mm.h^{-1}].

(b) **Wiederhold (1969):** In the late 1960s, Wiederhold (1969, cited by Lambourne and Stephenson, 1986) used a variable location, storm-centred approach which is a modified version of Van Wyk's (1965) method to establish ARFs for 170 storms over large catchment areas between 500 km² and 30 000 km² within 18 regions delineated for South Africa. In these medium to large catchment areas ($A \leq$ 30 000 km²), the ARF is mainly a function of the area and storm duration, since the quantity of rainfall relative to the number of storage areas, is of great importance (Alexander, 2001; SANRAL, 2013).

The large area storms were delineated, while the point rainfall depths at each rainfall station were used to fit a sixth-degree polynomial surface to enable the plotting of isohyets. Regionalised depth-area curves were produced for each storm at a daily interval resulting in co-axial diagrams to estimate the rainfall equalled or exceeded for storm durations of 1-day or longer. The developed depth-duration-area envelope diagram is shown in Figure 2.8.

In the case of large area storms associated with storm durations less than 24 hours, the average areal rainfall over increasing areas (durations of 1 to 6 days) within each of the 18 regions were expressed as percentages of the maximum observed point rainfall. Depth-area diagrams were produced for durations of 1 to 6 days. The upper envelope diagrams (of individual durations) were then re-plotted to produce depth-duration-area diagrams. Thereafter, the 24-hour to 1-hour durations, were linearly extrapolated to express the rainfall associated with a given area as a proportion of the point rainfall between one and 72 hours (Lambourne and Stephenson, 1986).

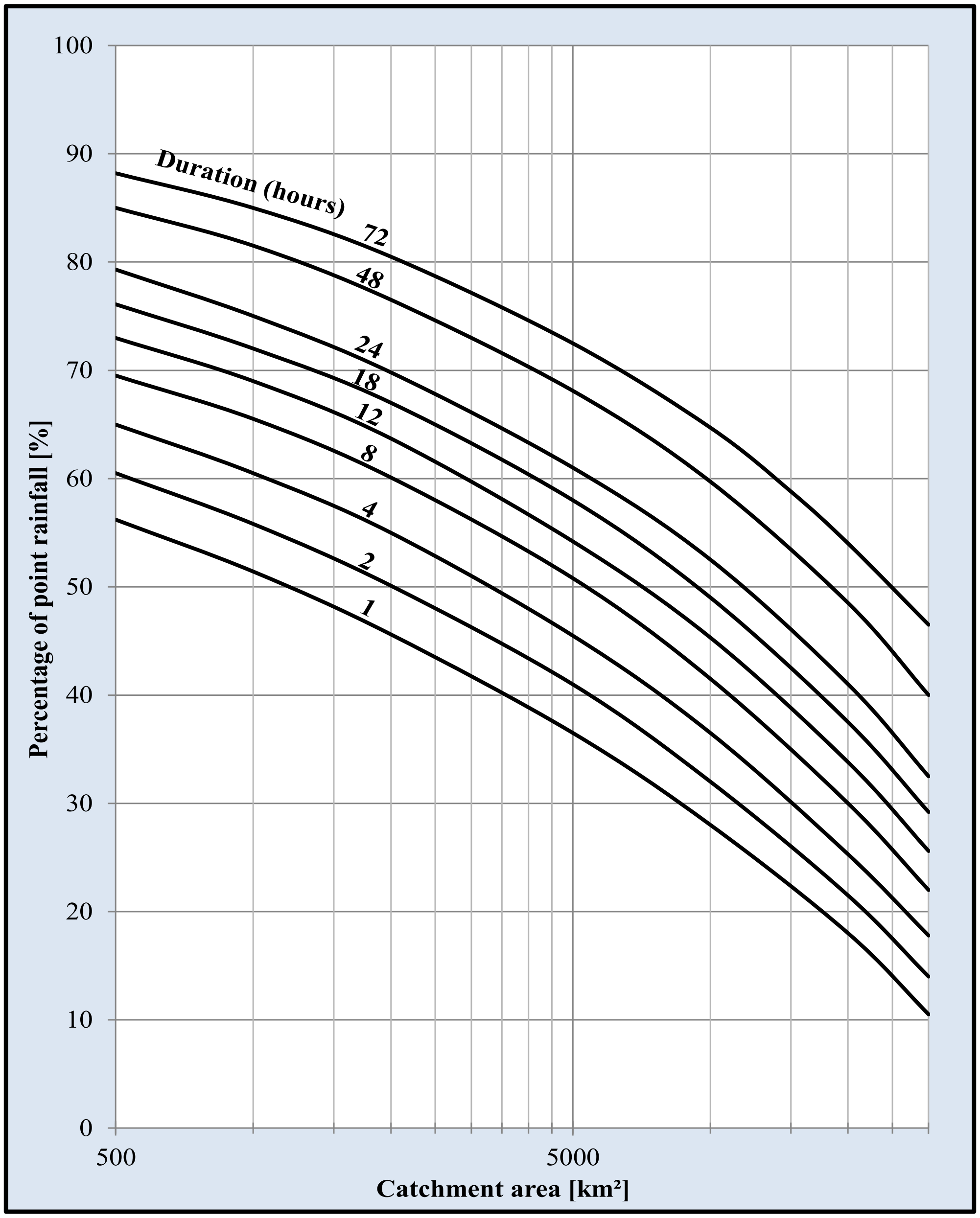

Figure 2.8: Expected percentage of runoff as a function of storm duration (SANRAL, 2013)

Op ten Noort and Stephenson (1982) used regression analyses to convert Figure 2.8 to a mathematical relationship [Eq. (2.15)].

$$ARF = [1.343 - 0.09\ln(A)]T_d^{0.03A^{0.19}} \qquad [2.15]$$

where

ARF = areal reduction factor for point intensity [%],

A = catchment area [km²], and

T_d = storm duration [hours].

Op ten Noort and Stephenson (1982) compared Eqs. (2.14) and (2.15) and established that the use thereof could cause a discontinuity in design flood estimations. Consequently, Figure 2.8 was extrapolated as such that the ARFs approach unity at short durations. This relationship is expressed by Eq. (2.16).

$$ARF = [1.04 - 0.08\ln(A)]T_d^{0.02A^{0.28}} \qquad [2.16]$$

where

ARF = areal reduction factor for point intensity [%],

A = catchment area [km²], and

T_d = storm duration [hours].

(c) **Rodriquez-Iturbe & Mejia (1974):** This method [Eq. (2.17)] is based on a spatial correlation structure, using either an exponentially decaying function or a Bessel-type correlation structure. All observed rainfall data, *i.e.*, the primary data, and not only the AMS, are required. The ‘design storm’ areal rainfall distributions were not included in the analyses.

$$ARF = \sqrt{E(\rho(d))} \qquad [2.17]$$

where

ARF = areal reduction factor, and

$E(\rho(d))$ = expected correlation coefficient for the characteristic correlation distance.

(d) **Bengtsson & Niemczynowicz (1986):** This method [Eqs. (2.18) & (2.19)] represents the relationship between rainfall movement and associated ARFs. The ARFs are based on the limited extension of rain cells, movement and spacing between rain cells, and the effect of rain cells on each other. Typically, the ARFs were estimated from point rainfall hyetographs and storm moving rates. Relationships were established between moving, storm-derived ARFs and ARFs estimated by a dense rainfall-monitoring network. The ARFs proved to be constant in Norway.

$$ARF = \frac{L_P}{L} = \frac{vT_d}{L} \quad \text{if } L_P < 0.5 \qquad [2.18]$$

$$ARF = 1 - \frac{0.25L}{L_P} = 1 - \left(\frac{0.25L}{vT_d}\right) \quad \text{if } L_P \geq 0.5 \qquad [2.19]$$

where

ARF = areal reduction factor,

L = catchment length [km],

L_P = extension of block rain cell [km],

T_d = storm duration [hours], and

v = storm speed [$m.s^{-1}$].

(e) **Bacchi & Ranzi (1996):** This method [Eq. (2.20)] is based on the analyses of the crossing properties of spatial and temporal rainfall processes. Sixteen Constant Altitude Plan Position Indicator (CAPPI) maps were recorded and analysed from the C-band weather radar and subsequently compared with a corresponding rainfall data set consisting of 17 rainfall stations. The estimated ARFs were expressed as the ratio of areal and point rainfall intensity values associated with the same duration and return period. It was found that the estimated ARFs are dependent on the return period and catchment area.

$$AR_{(A,\, Td,\, F)} = \frac{T_{A,Td}(F')}{T_A(F')} \qquad [2.20]$$

where

ARF = areal reduction factor,

A = area under consideration [km²],

F' = *F*-quantile of the corresponding probability distribution,

T_d	= duration within the space-time domain where the rainfall process can be assumed uniform [hours], and
T	= return period [years].

(f) **Sivapalan & Blöschl (1998):** This method [Eq. (2.21)] is based on a complex spatial correlation structure using both extreme value and/or parent distributions. It was found that the ARF values are dependent on the catchment area, storm duration (spatial correlation structure) and return period, while the ARF values are independent of the rainfall regime. The ARF values decreased with an increasing catchment area and return period.

$$ARF\left[k^2\left(\frac{A}{\lambda^2}\right),T_d,T\right] = \frac{b(T_d)c(T_d)k^2F_2\left(k^{-2}\right)-\frac{k^2}{F_1\left(k^{-2}\right)}\ln\left[\ln\left(\frac{T}{T-1}\right)\right]}{b(T_d)c(T_d)-\ln\left[\ln\left(\frac{T}{T-1}\right)\right]} \qquad [2.21]$$

where

ARF	= areal reduction factor,
A	= catchment area [km²],
b	= function of duration, where $b(T_d) = -0.05 + 0.25{T_d}^{0.49}$
c	= function of duration, where $c(T_d) = 0.2 + 20{T_d}^{-0.7}$
$F_1(k^{-2})$	= generic properties of the gamma distribution,
$F_2(k^{-2})$	= generic properties of the gamma distribution,
k^2	= rainfall correlation structure,
T	= return period [years],
T_d	= storm duration [hours], and
λ	= spatial correlation length [km].

(g) **Asquith & Famiglietti (2000):** This method [Eq. (2.22)] was developed for the Austin (15 600 km²), Dallas (21 000 km²), and Houston (35 800 km²) regions in the USA with a dense rainfall-monitoring network (103 to 193 rainfall stations per region). In general, the rainfall data record lengths used, exceeded 80 years. This method focused on the analysis of the areal rainfall distribution and resulted in ARFs which decreased with an increase in AEPs.

$$ARF = \frac{\int_0^R 2rS_{T(r)}\Delta r}{R^2} \qquad [2.22]$$

where

ARF	= areal reduction factor,
A	= rainfall storm areas [km²],
R	= maximum radius of circular catchment or integration limit [km],
r	= radius of concentric circle within the catchment [km), and
$S_{T(r)}$	= ratio between rainfall depth at a specific location, distance *r* from the point of the design storm and the annual maxima rainfall.

(h) **Lombardo *et al.* (2006):** The ARF values were estimated by using radar reflectivity maps collected with Polar 55C, 15 km south-east of Rome, Italy. Rainfall intensities over the radar scanning region (allowing a single radar image to last for one minute), were estimated for durations of 1, 5, 10, 60 and 120 minutes and return periods of 2, 10, 25 and 50 years by using the arithmetic mean and Thiessen polygon methods. The ARFs exceeded unity in small areas associated with relative longer storm durations. The ARF relationship is expressed by Eq. (2.23).

$$ARF_{(Td,\,T)} = \frac{i_A(T_d, T)}{i_{A=1}(T_d, T)} \qquad [2.23]$$

where

ARF	= areal reduction factor,
A	= area under consideration [km²],
i	= rainfall intensity [mm.h^{-1}],
T	= return period [years], and
T_d	= storm duration [hours].

(i) **Kim *et al.* (2019):** This method [Eq. (2.24)] considered the building of a time series at every radar pixel for six (6) durations. Storm centres that contained specific time and location information were used. Specific elliptical shapes (aspect ratio, orientation, and sizes) with a maximising rainfall volume of 13 areas, were identified, while additional filtering was performed to verify the spatial-temporal proximity of the storm events. After the filtering process, the total number of storms analysed, reduced from 2 218 905 to 54 758.

$$ARF = \frac{\frac{\Delta A}{A_0}\sum R_i}{R_0} \qquad [2.24]$$

where

ARF	= areal reduction factor,
A_0	= area of an elliptical storm [m²],
ΔA	= area of radar cells [*i.e.*, 1 km * 1 km],
R_0	= maximum rainfall depth of an ellipse [mm], and
R_i	= rainfall volume of the i^{th} cell within the storm [mm].

2.10 Regionalisation Methods

The RLMA&SI regionalisation scheme developed by Smithers and Schulze (2000a; 2000b; 2003; 2004) and as discussed in Section 2.5.2, are based on a cluster analysis of site characteristics. In essence, a vector of site characteristic was associated with each site and standard multivariate probabilistic analyses were performed to group the sites according to the similarity that exists between the vectors (Smithers and Schulze, 2004). Site characteristics that were used during the regionalisation are: (i) latitude, longitude, and altitude, (ii) index of rainfall concentration, (iii) mean annual precipitation (MAP), (iv) seasonal rainfall index, and (v) the distance from the ocean. As highlighted in Section 2.5.2, above procedures resulted in 78 homogeneous rainfall clusters (Smithers and Schulze, 2000b). The site characteristics for each cluster were normalised using cluster analysis and subsequently grouped into seven (7) long duration rainfall regions.

2.11 Goodness-of-Fit Statistics

In some cases, the visual comparison of simulated results against observed data can be subjective. Statistical model performance measures are therefore used to evaluate a model's performance against observed data using pre-determined statistical criteria expressed as Goodness-of-Fit (GOF) statistics. In short, GOF statistics are mathematical functions which express the most desirable characteristics between a model's simulated output and the actual observations (Schulze, 1995). The following GOF statistics are considered in this research:

(a) **Coefficient of correlation (r^2):** This statistic measures the degree of association between the simulated values and the simulated values as predicted by a regression model. The objective function is to maximise the correlation coefficient to unity (1.0), since higher r^2 values are indicative of a higher degree of association (Schulze, 1995).

(b) **Standardised residuals:** According to Field (2000), the following general rules apply to the estimation of standardised residuals: (i) standardised residuals lower or greater than -3 or 3 respectively, are regarded as problematic given that their presence in a data set is not by chance, (ii) when more than 1% of the sample size have standardised residuals with values lower or greater than -2.5 or 2.5 respectively, sufficient evidence is present to confirm that the level of error within the model is unacceptable, and (iii) when more than 5% of the sample size have standardised residuals with a value lower or greater than -2 or 2 respectively, we can reason, with evidence, that the model is a poor representation of the observed data.

(c) **The *t*-test:** This statistic is used to determine whether there is a statistical significance between the averages of the observed and regression model values associated with a range of predictor variables (Schulze, 1995).

(d) **Standard Error of the Mean (*SE*):** This statistic represents the percentage difference between standard errors of the average observed and simulated values. The *SE* value should ideally be minimised to represent a good simulation (Schulze, 1995).

(e) ***F*-test:** The total *F*-test is used to determine whether the dependent criterion variable, such as an ARF, is significantly correlated or different to the independent predictor variables under consideration (Schulze, 1995).

2.12 Conclusions

Currently, the acquisition of reliable observed rainfall data (*e.g.*, daily and/or sub-daily) in South Africa (and internationally) remains problematic, despite the requirement thereof in the development and/or updating of rainfall-runoff modelling systems and design flood estimation. Internationally, researchers from various countries and climatological regions, tend to differ when factors having an influence on ARFs, are considered, For example, as included in this chapter, many researchers concluded that catchment geomorphology has an insignificant influence on the derivation of ARFs (Svensson and Jones, 2010), while other researchers have shown that catchment geomorphology does play a role in the derivation of ARFs (Veneziano and Langousis, 2005; Pietersen, 2016; Singh *et al.*, 2018). Hence, any findings from the literature would be subjected to and are dependent on the location of the study area, data availability, and/or method of analyses (*e.g.*, geographical or storm-centred analysis).

Typical ARFs, based on the literature review conducted, are regarded as dynamic factors influenced by numerous other input parameters, *e.g.*, climatological variables, catchment geomorphology, methodological approaches, return periods, storm durations, and available record lengths and/or a combination of these. Hence, during the estimation of ARFs, whether an analytical or empirical approach is followed, the latter factors/input parameters should be incorporated to result in representative sample ARFs applicable to a specific region/country. However, given the complexities involved, the estimation of ARFs remains an ongoing worldwide research question. This is especially valid in South Africa where South African researchers (*e.g.*, Pietersen *et al*., 2015; Pietersen, 2016; Du Plessis and Loots, 2019; Du Plessis *et al.,* 2020) only recently studied ARFs again by considering the longer periods of record which are now available for analysis (± 50 years of additional data since the last South African derived ARF method in the 1970s).

The various concerns and shortcomings associated with the current national and international ARF estimation methods as identified and discussed in this chapter, will be further interrogated in Chapters 4 and 5 to provide a viable ARF solution for South Africa.

The location and characteristics of the study area are presented in the next chapter.

CHAPTER 3: STUDY AREA

This chapter provides an overview of the location and characteristics of the study area.

3.1 Location and General Characteristics

South Africa is located at the most southern tip of Africa with a coastline stretching over more than 3 000 km. Botswana, Eswatini (former Swaziland), Lesotho, Mozambique, Namibia, and Zimbabwe are neighbouring countries of South Africa. As shown in Figure 3.1, South Africa is demarcated into 22 primary drainage regions (A to X), with catchment areas ranging from 2 592 km^2 to 409 721 km^2. The study area, inclusive of South Africa, Lesotho and Eswatini covers a surface area of approximately 1 267 480 km^2 and is located between 22° and 35° S and 17° and 33° E (DWAF, 1995).

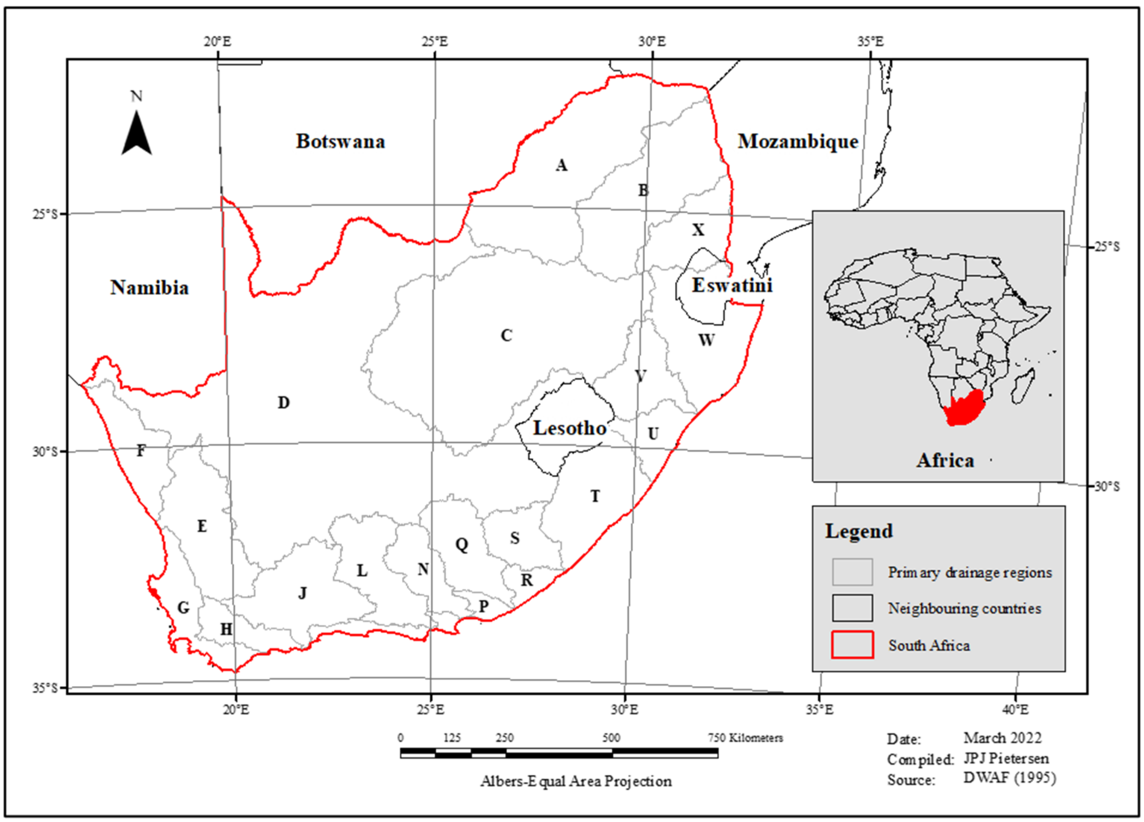

Figure 3.1: Location of the study area

The study area consists primarily of deserts, bushveld, grasslands, forests, mountains, beaches, and coastal wetlands. As highlighted in Chapter 2 (*cf.* Chapter 2, Section 2.7), land cover/use proved to have an influence on the derivation of ARFs. Hence, Table 3.1 contains information regarding the spatial distribution of the different land cover categories in South Africa (SSA, 2004).

Table 3.1: Spatial distribution of land cover categories in South Africa

Sector	Land cover [km^2]	Land cover [%]
Agricultural land (cultivation)	147 532	12.1
Forest land	17 617	1.5
Major waterbodies	3 392	0.4
Natural vegetation	949 993	77.9
Nature reserve and conservation	55 393	4.5
Urban build-up and structures	13 790	1.1
Wetlands	5 409	0.4
Other	25 952	2.1
Total	**1 219 078**	**100**

The Shuttle Radar Topography Mission (SRTM) elevation data for Southern Africa at 90-metre resolution (USGS, 2002) are used to illustrate the variation in topography in Figure 3.2, which varies from gentle to steep, with elevations ranging from 0 m to 3 482 m above mean sea level (USGS, 2002).

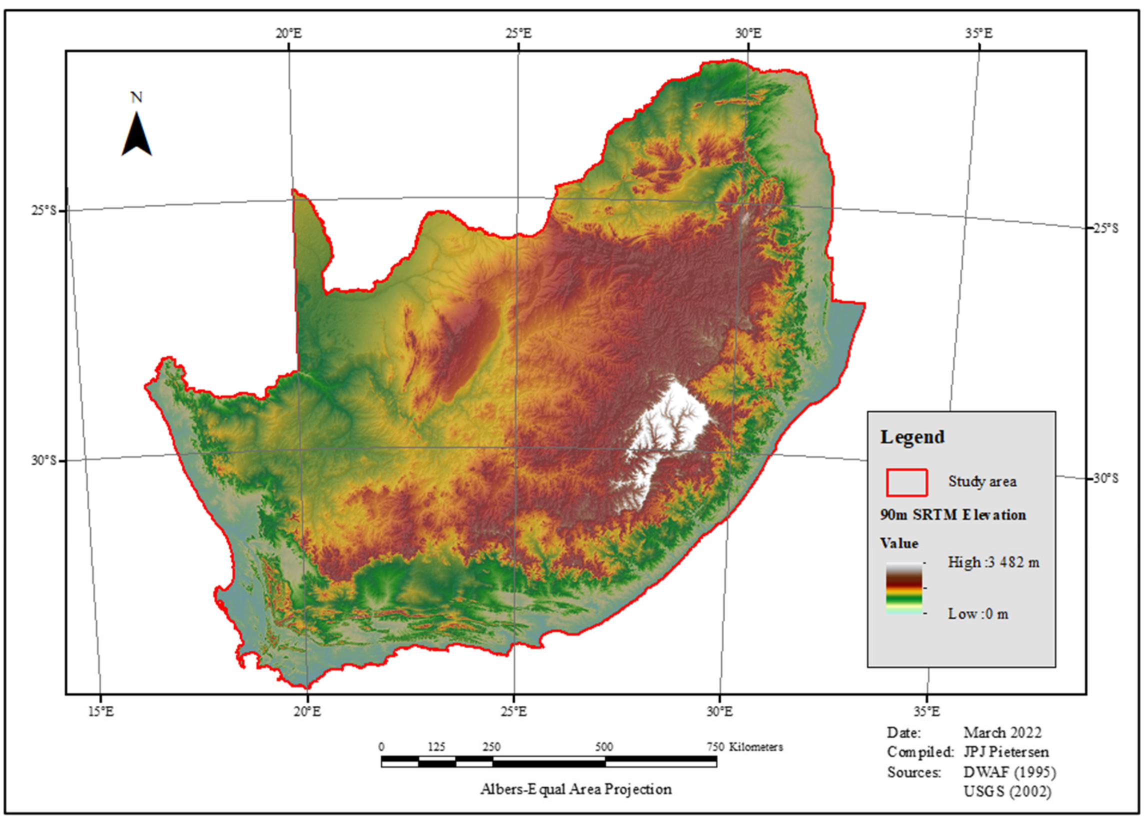

Figure 3.2: Elevation map of the study area based on a 90-metre resolution SRTM data set

It was further decided to establish the elevation distribution, since Singh *et al.* (2018), as highlighted in Section 2.7.2, indicated that ARFs in New Zealand are influenced by the elevation profile and rainfall types. Therefore, the elevation distribution in the study area, as derived from the SRTM 90 data set, is summarised in Table 3.2.

Table 3.2: Elevation distribution in study area (USGS, 2002)

Elevation ranges [m]	Distribution [%]
< 310	8.7
310 – 860	20.9
860 – 1 411	50.3
1 411 – 1 985	18.3
> 1 985	1.8
Total	**100**

3.2 Climate

Climate is highly variable in South Africa. Hence, hydrological and climatological information were used by Alexander (2010) to define nine distinctive climatological regions in South Africa, as illustrated in Figure 3.3.

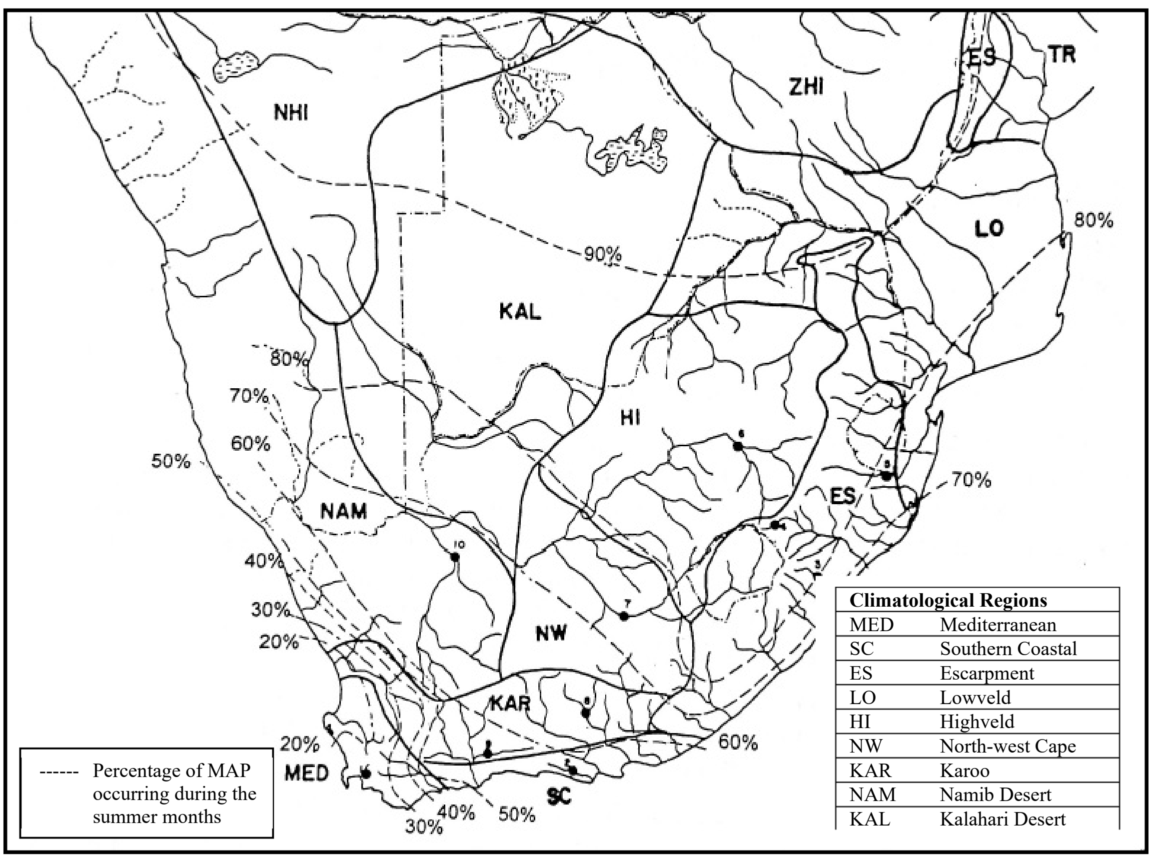

Figure 3.3: Climatological regions for South Africa (Alexander, 2010)

Typically, apart from climate, other factors such as geographical location, altitude above mean sea level, rainfall type (convective, frontal and/or orographic), rainfall seasonality (summer, winter and/or all year) and average catchment slope classes (flat, moderate, or steep) were also considered to define the various regions as shown in Figure 3.3.

Typically, in the south-western Cape (Mediterranean, and Southern Coastal regions), the climate is characterised by winter rainfall and warm windy summers, while highly variable, non-seasonal rainfall and extreme temperatures occur in the Karoo (KAR) region. Hot summers with convective thunderstorms and cold winters are typical on the Highveld, with the average long-term minimum and maximum summer temperatures varying between 12°C and 30°C, while the winter months are characterised by average long-term minimum and maximum temperatures ranging between 3°C and 18°C. Mesic-subtropical conditions dominate on the KwaZulu-Natal coast of the Escarpment region (Midgley *et al.*, 1994; Davies and Day, 1998; Alexander, 2010).

The temporal and spatial distributions of rainfall are highly variable on a seasonal and annual basis, since the rainfall is produced by different weather systems in different regions and at different times of the year. In winter, the prevailing north-westerly winds result in high rainfall in the western parts of the country, while the southern interior and Karoo remain dry. Summer rainfall is normally higher in the north and east, but due to dry high-pressure air masses that persist for long periods, the rainfall is low in the western parts of the country. Furthermore, wide-spread flooding in South Africa is known to be caused by intense cut-off low pressure systems and/or by occasional tropical cyclones (Davies and Day, 1998; Mambo and Faccer, 2017).

In South Africa, the MAP decreases, while the Mean Annual Evaporation (MAE) increase westwards and northwards. The overall MAP is 452 mm, but in many parts of the country, as shown in Figure 3.4, the MAP is much less; ranging from as low as 50 mm in the western parts, to MAP values exceeding 1 500 mm in the eastern parts (Mambo and Faccer, 2017). Evaporation exceeds rainfall throughout the country, except in the mountainous Escarpment and Mediterranean regions. In the central parts of South Africa, evaporation is approximately twice the rainfall, while in the western parts of the country, evaporation exceeds the rainfall by a factor of ten (Davies and Day, 1998; Mambo and Faccer, 2017).

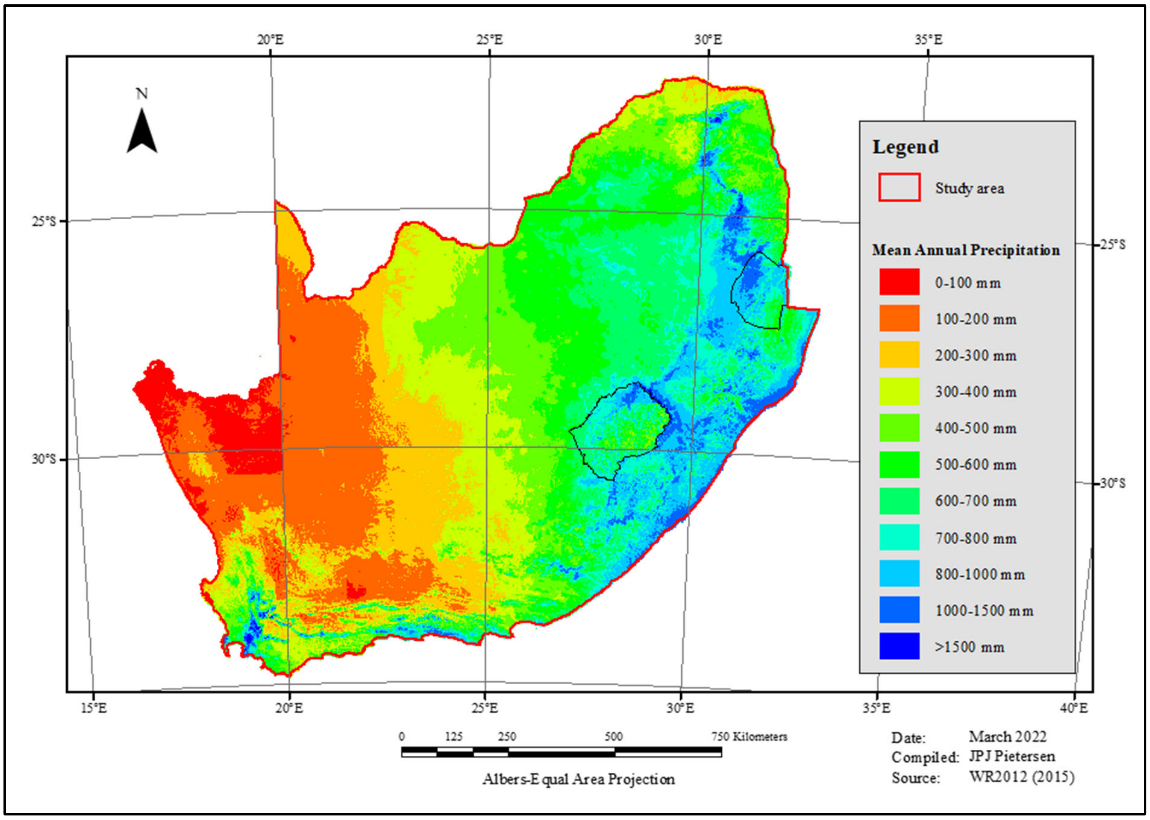

Figure 3.4: MAP distribution over study area

3.3 Rainfall Monitoring Network

A total of 1 779 daily rainfall stations, as obtained from the: (i) South African Weather Services (SAWS), (ii) Agricultural Research Council (ARC), and (iii) South African Sugar Association (SASA), with individual record lengths exceeding 40 years, are located within the boundaries of the study area (*cf.* Figure 3.5). It is evident from Figure 3.5 that the rainfall-monitoring network is generally denser along the southern and central parts of the study area as opposed to the north-western and north-eastern parts.

The overall distribution and location of the available individual rainfall stations are regarded as sufficient for the purpose of this study. However, when point rainfall depths are converted to rainfall depths over an area using averaging techniques, such as the Thiessen polygon method (*cf.* Chapter 2, Section 2.4), denser rainfall-monitoring networks are preferred.

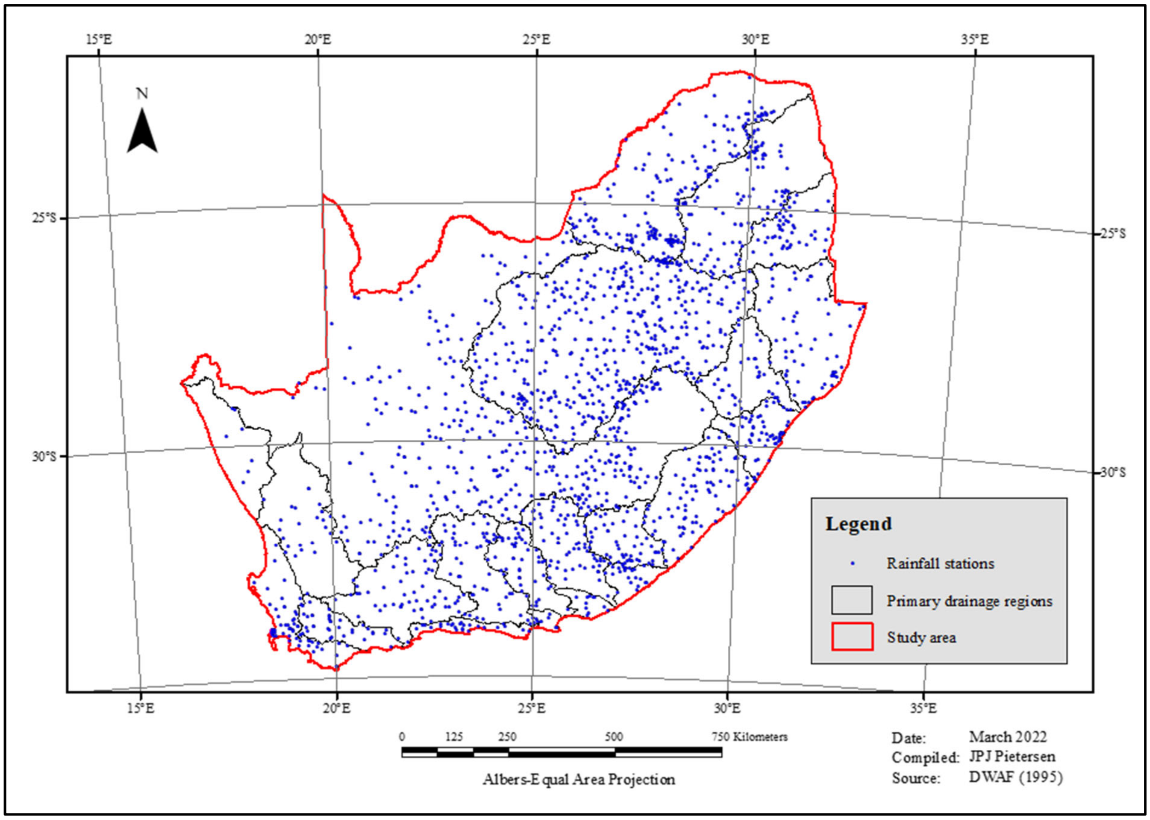

Figure 3.5: Location of the 1 779 daily rainfall stations within the study area

The summary of the number of daily rainfall stations located within each of the 22 primary drainage regions, with their associated minimum and maximum record lengths, are listed in Table 3.3. A comprehensive list, inclusive of the 1 779 daily rainfall stations and their associated record lengths, is also included in Appendix A.

Table 3.3: Minimum and maximum record lengths for daily rainfall stations

Primary drainage region	Number of daily rainfall stations	Record length [years]	
		Minimum	Maximum
A	174	40	100
B	91	44	95
C	410	41	119
D	340	43	121
E	26	46	129
F	7	48	119
G	61	43	148
H	37	46	120
J	62	50	121
K	24	49	120
L	57	46	121
M	5	48	117
N	45	44	115
P	11	58	121
Q	63	46	121

Primary drainage region	Number of daily rainfall stations	Record length [years]	
		Minimum	Maximum
R	22	49	121
S	46	52	113
T	87	47	109
U	57	41	127
V	40	44	86
W	63	41	94
X	51	53	95
Study area	**1 779**	**40**	**148**

It is evident from Table 3.3 that all the individual rainfall stations have a record length of at least 40 years. However, when point rainfall values need to be averaged to representative catchment values, the record lengths of all the rainfall stations under consideration should have the same starting and ending dates. In addition, the minimum combined record length, consisting of two or more rainfall stations, should also not be less than 30 years. Subsequently, the infilling of missing rainfall data would be required in some cases. The latter infilling of daily rainfall data and/or screening procedures are detailed in Chapters 4 and 5, respectively.

The evaluation of different methodological approaches to be adopted when probabilistically correct ARFs need to be estimated, is included in the next chapter.

CHAPTER 4: INITIAL EVALUATION OF DIFFERENT METHODOLOGICAL APPROACHES

This chapter focuses on the identification and evaluation of various methodological approaches and/or data sources to establish the suitability thereof for the estimation of local ARFs. These include the: (i) influence of artificial catchment size and placement on the estimation of ARFs, (ii) identification/selection of the most suitable theoretical probability distribution(s) to estimate areal and design point rainfall, (iii) most suitable design point rainfall data source, and (iv) minimum allowable record lengths required to estimate design rainfall values.

Each of the above factors are detailed below, while the final methodology adopted and the associated results are presented in Chapter 5.

4.1 Pre-defined Fixed versus User-defined Circular Areas

The size of a catchment is an important variable in the estimation of ARFs, and it is therefore important to use representative catchment areas during the rainfall data analyses. The original envisaged methodological procedure entailed the use of fixed, artificial circular catchment areas in pre-defined area ranges of 125, 250, 500, 1 000, 2 000, 4 000, 8 000 and 16 000 km^2.

However, the use of pre-defined catchment areas is regarded as problematic in catchments with a low-density rainfall-monitoring network. The average point and/or areal rainfall estimated using the Thiessen polygon method becomes questionable, as highlighted by Sugawara (1992), when applied to a low-density rainfall-monitoring network, since the Thiessen polygons may not represent the actual rainfall processes in a specific polygon, as they interrupt the spatial transitions of the natural occurring process. Therefore, to visualise the true spatial distribution of any rainfall event over a specific area becomes problematic. Figure 4.1 highlights the concerns associated with the use of 'pre-defined fixed areas' as opposed to 'user-defined circular areas'.

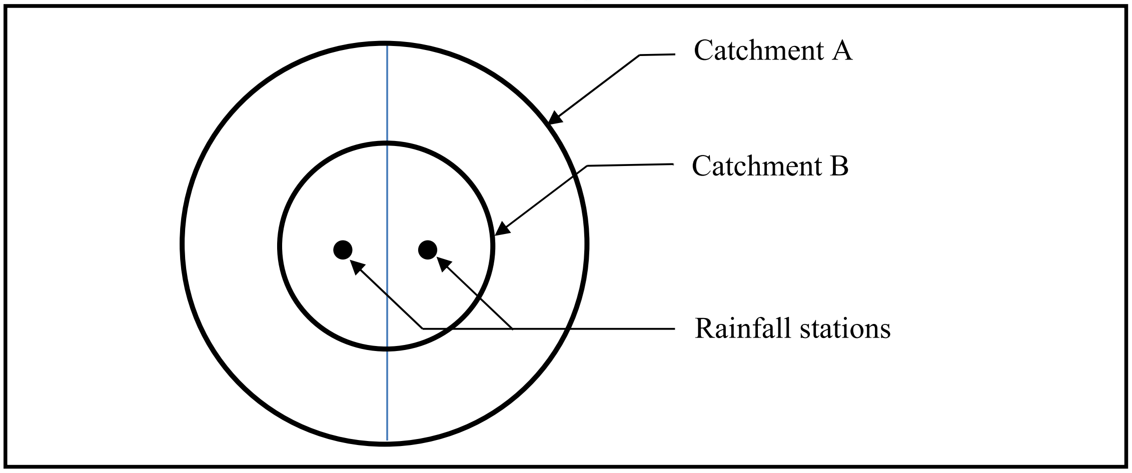

Figure 4.1: Illustration of identical Thiessen weights applicable to different catchment sizes

It is evident from Figure 4.1 that the delineated Thiessen weights would be identical for both Catchments A and B, irrespective of the catchment sizes under consideration. Hence, this scenario could have a significant impact when ARFs are estimated, *i.e.*, identical sample ARF values would be estimated for both Catchments A and B, although they are completely different in size. Therefore, due to the concerns associated with the fixed-area circular catchment approach, a 'user-defined' approach, which enables the adjustment of circular catchment areas, as shown in Figure 4.2, was adopted in this study.

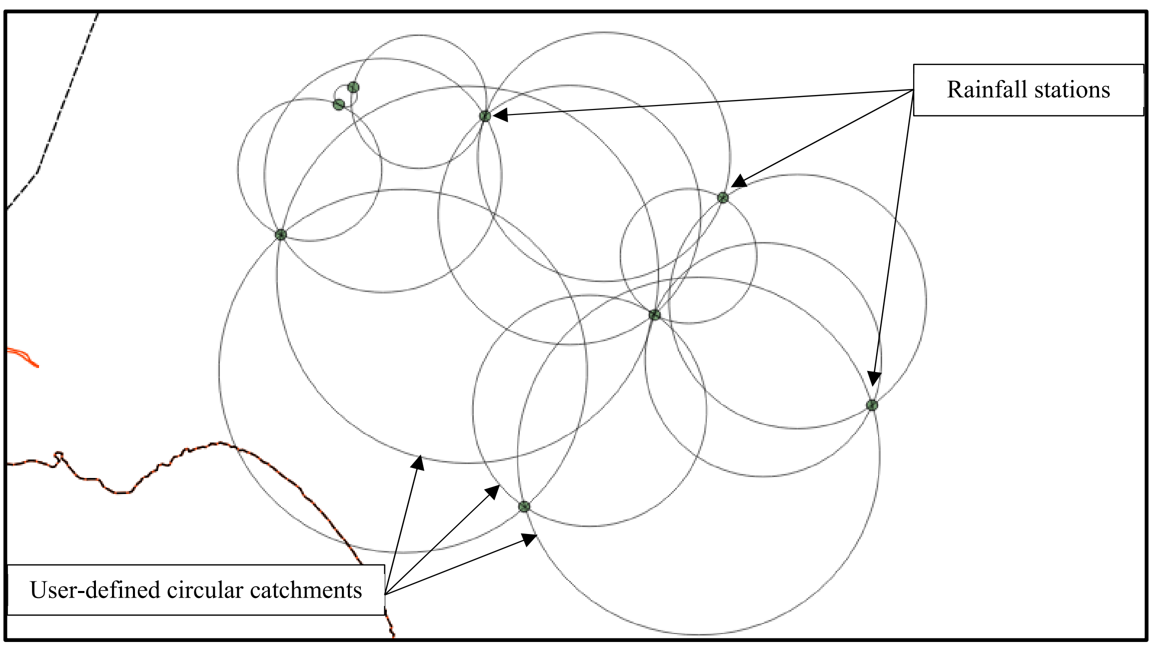

Figure 4.2: Illustration of user-defined circular catchments

The 'user-defined' approach enables the adjustment of circular catchment areas in such a way that the circumference of any particular circle (representative of a circular catchment; black circles in Figure 4.2) will intersect the most distant contributing rainfall stations (green dots in Figure 4.2), and subsequently, include all the rainfall stations located within this circumference for further analyses. This ensured that the circular catchment areas are determined by the spatial location of the contributing rainfall stations (green dots), thereby improving the representative spatial distribution of a rainfall event.

4.2 Probability Distribution and Method of Moments for Areal Design Rainfall

Smithers and Schulze (2000a) highlighted that the General Extreme Value (GEV) probability distribution using Linear Moments (LM) is regarded as the most suitable distribution to estimate 1-day design rainfall values in South Africa. In addition, the GEV probability distribution was also used to estimate ARFs in various other international studies, *e.g.*, Dyrrdal *et al.* (2016) and Peleg *et al.* (2018).

Hence, apart from using the most suitable theoretical probability distribution, the potential influence which the different methods of estimating moments could have on the estimation of sample ARFs, was investigated. Subsequently, the areal design rainfall values were estimated in 10 circular test catchments (125 km² each) located within Region 1 of the study area using the Thiessen weighted areal AMS and Linear Moments (LM), Method of Moments (MM), and Probable Weighted Moments (PWM). The GEV probability distribution was fitted using all three methods of moments and subsequently resulted in areal design rainfall values in each of the 10 circular test catchments. Thereafter, the sample ARFs were estimated in each circular catchment using the ratio between the GEV-based average areal design rainfall and the average RLMA&SI design point rainfall.

The average sample ARFs, associated with the different methods of moments, are presented in Figure 4.3.

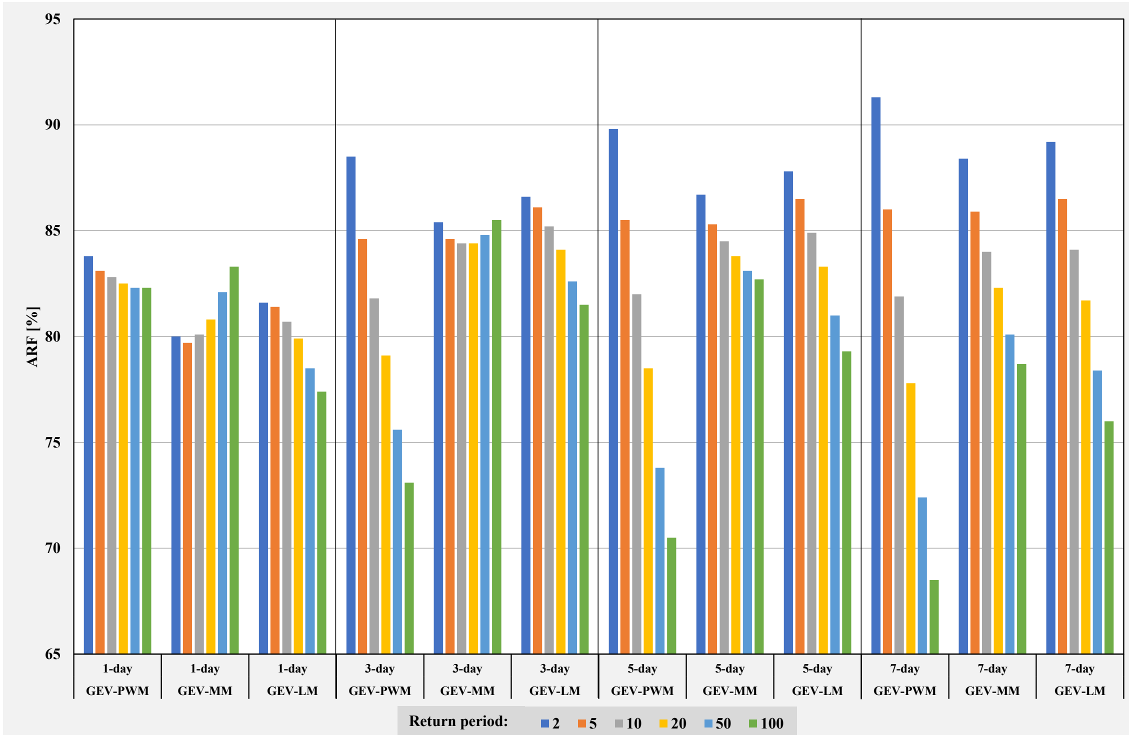

Figure 4.3: Average sample ARFs using different methods of moments

In Figure 4.3 the differences, although similar trends between the different ARF values, are evident and highlight that the different methods of estimating moments might have a significant impact on ARF estimation, especially when shorter durations are considered. The 1-day GEV_{LM} ARFs increase with return period, while the GEV_{PWM} and GEV_{MM} ARFs decrease with return period. In addition, the 3-day GEV_{LM} ARFs resulted in equivalent ARF values for the 2- and 100-year, 5- and 50-year, and 10 to 20-year return periods. The ARFs derived using the three selected method of moments also tend to decrease with return periods for durations exceeding 5-days. Subsequently, based on the above results, the estimation of sample ARFs in the subsequent sections will be based on the GEV_{LM} probability distribution of areal and design point rainfall values.

The RLMA&SI design rainfall database as a possible (alternative) source of design point rainfall values, is discussed in the next section.

4.3 Data Source of Design Point Rainfall to Estimate Sample ARFs

The purpose of this section is to establish the advantages and disadvantages associated with different design point rainfall data sources (*e.g.*, RLMA&SI or GEV_{LM} derived) to estimate sample ARFs in the same 10 circular test catchments considered in Section 4.4. The original methodological procedure proposed the use of the RLMA&SI design rainfall database as the design point rainfall data source for estimating ARFs. The first section (Section 4.5.1) focusses on the concerns associated with the RLMA&SI design rainfall, when considered as the design point rainfall data source to estimate sample ARFs, while Section 4.5.2 highlights the use of the GEV_{LM} design point rainfall when GEV_{LM} areal and design point rainfall values are derived from identical record lengths, *i.e.*, derived from observed areal and point rainfall data.

4.3.1 Concerns when deriving ARFs using the RLMA&SI approach

ArcGIS was used to extract the design point rainfall values, inclusive of all the points within a particular circular catchment, for durations of 1, 3, 5 and 7 days associated with return periods (*T*) of 2, 5, 10, 20, 50, 100 and 200 years. A step-by-step procedure to estimate the average RLMA&SI design point rainfall values, based on the 10 circular test catchments, is included in Appendix B. Areal design rainfall values for the same 10 circular test catchments as considered in Section 4.2 and Appendix B, were estimated by means of the GEV_{LM} probability distribution. The sample ARFs were estimated as the ratios between the latter GEV_{LM}-based areal (*cf.* Section 4.2) and RLMA&SI design point rainfall values (*cf.* Appendix B).

These estimated sample ARFs were compared, as shown in Figure 4.4, with the Australian ARFs (*cf.* Chapter 2, Section 2.9.1), *e.g.*, Siriwardena and Weinmann (1996) and Podger *et al.* (2015a; 2015b), which demonstrated that ARFs vary with return period (*T*) and storm duration. The Australian ARF methodologies were selected for this comparison, since both countries, Australia and South Africa, are predominantly located between 20°S and 40°S within the Southern Hemisphere, sharing areas with comparable climatological conditions (Milewski, 1981).

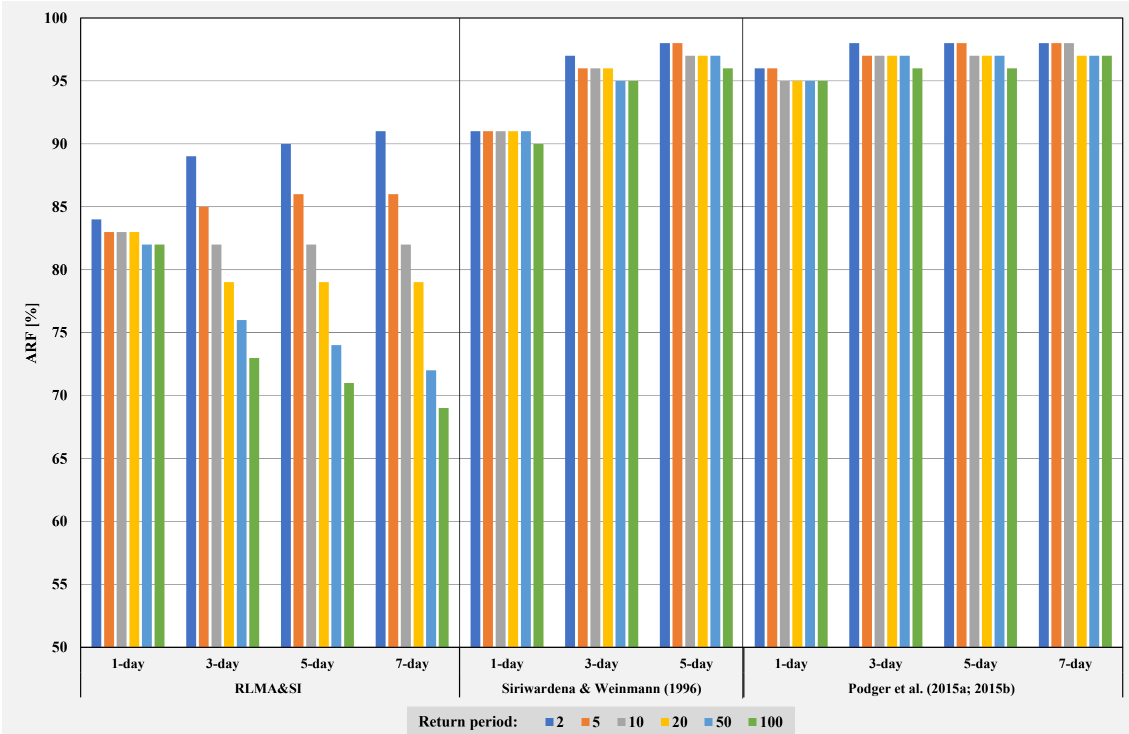

Figure 4.4: Sample ARFs (RLMA&SI-based) compared to the sample ARFs estimated using similar international methodologies for an area of 125 km^2

Based on the results shown in Figure 4.4, it is evident that the RLMA&SI-based ARF estimates increase with an increase in storm durations (2 and 5-year return periods only) and are in contradiction to those ARFs estimated using the respective Australian methodologies, which decrease with an increase in return period and increase with an increase in storm duration. In the case of the 10 circular catchments (125 km^2 each), the estimated sample ARFs decreased substantially with an associated increase in both return period and storm duration. The latter trend could possibly be ascribed to the fact that the RLMA&SI design rainfall values are based on a regional approach, *i.e.*, design rainfall values estimated using multiple rainfall stations in proximity of the rainfall station under consideration.

This unexpected trend was further investigated and confirmed by comparing ARFs estimated from the ratios of: (i) GEV_{LM} areal design rainfall and RLMA&SI design point rainfall (*cf.* Sections 4.2 and Appendix B), and (ii) GEV_{LM} areal and design point rainfall values derived from identical record lengths for both the observed areal and point rainfall data sets (*cf.* Section 4.3.2).

4.3.2 Areal and Point Design Rainfall Derived from Observed Data

The results, by comparing ARFs estimated from the ratios of averaged: (i) GEV_{LM} areal design rainfall and RLMA&SI design point rainfall (*cf.* Sections 4.2 and Appendix B), and (ii) GEV_{LM} areal and design point rainfall values derived from identical record lengths for both the observed areal and point rainfall data sets, are listed in Table 4.1. The substantial differences between the two approaches, especially at higher return periods and storm durations, are evident from Table 4.1.

Table 4.1: ARFs (RLMA&SI-based) compared to ARFs (GEV_{LM}-based) using identical record lengths and areas

	ARFs (identical record lengths) [%]				ARFs (RLMA&SI design point rainfall) [%]			
T **[years]**	**Storm duration [days (d)]**							
	1-d	**3-d**	**5-d**	**7-d**	**1-d**	**3-d**	**5-d**	**7-d**
2	80.7	86.2	89.5	90.2	80.2	85.6	88.6	90.9
5	82.6	[illegible]	89.1	89.6	82.1	[illegible]	84.6	86.2
10	85.9	87.8	[illegible]	89.1	85.8	82.7	[illegible]	82.5
20	90.4	90.6	[illegible]	88.4	90.9	83.0	[illegible]	78.8
50	97.8	95.5	90.6	87.5	99.6	84.1	80.2	74.0
100	100.0	100.0	91.3	86.6	100.0	85.3	74.3	70.4
200	100.0	100.0	92.1	85.7	100.0	86.9	72.1	67.0

Increase

Decrease

Therefore, ARF estimates based on the RLMA&SI approach might be challenging, since the design rainfall values are based on different record lengths with different start and end dates. In other words, to estimate representative ARFs, both the areal and point rainfall AMS values should originate from the same rainfall record length. Furthermore, when estimating areal design and design point rainfall from one station's AMS, it should result in the same design rainfall values, *i.e.*, ARF = 100%. This is to be expected when estimating ARFs for smaller catchments, *i.e.*, the point rainfall should be 100% representative for the catchment. As an example, a probabilistic analysis was conducted using the extracted AMS values from Rainfall Station 0590028 W, having 91 years of observed data. Thereafter, the RLMA&SI design point rainfall values were obtained for the same rainfall station and compared to the GEV_{LM}-based design rainfall. Hence, the ratios between the RLMA&SI and GEV_{LM} design point rainfall values (*cf.* Figure 4.5) should equate one, *i.e.*, ARF = 100%. However, this is not the case given that the RLMA&SI approach uses a constant growth curve for all sites in a specific region and an estimated at-site index value. Similarly, the same comparison was conducted using the GEV_{LM} areal and point rainfall values obtained from two rainfall stations within a circular catchment of 0.243 km^2, *i.e.*, rainfall stations in very close proximity. The two rainfall stations had the same record lengths, *i.e.*, AMS values based on the same start and end dates.

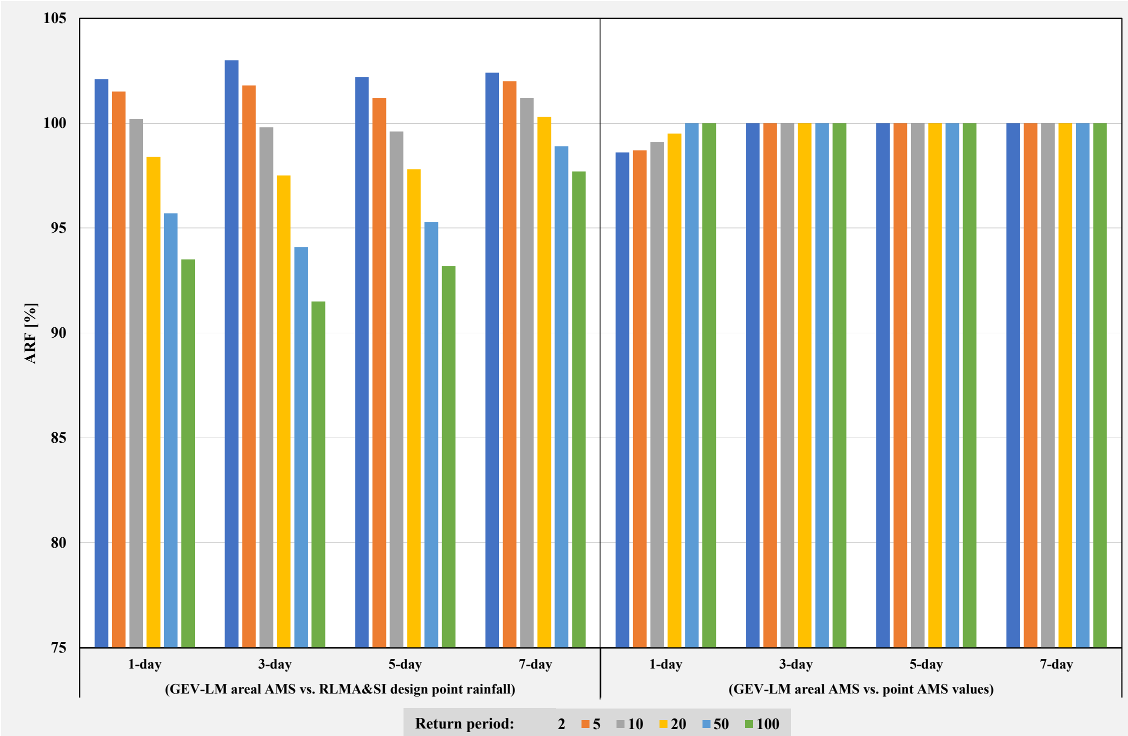

Figure 4.5: Comparison of ARFs estimated using GEV_{LM} areal AMS values and GEV_{LM} point AMS values and/or RLMA&SI design point rainfall values

As shown in Figure 4.5, the latter comparison of the GEV_{LM} areal and design point rainfall values derived from identical record lengths for both the observed areal and point rainfall data sets, resulted in improved ARF values, which confirm that design point rainfall and areal design rainfall are comparable (very similar) in smaller catchments. Hence, this also confirmed that ARFs should be estimated and expressed as the ratios between average areal and average design point rainfall values using the same rainfall stations and record lengths with mutual starting and ending dates.

Based on the above findings, it was decided (as previously done in Figure 4.4), to compare the latter GEV_{LM} areal and design point rainfall values derived from identical record lengths for both the observed areal and point rainfall data sets, with the international ARF methodologies of Siriwardena and Weinmann (1996) and Podger *et al.* (2015a; 2015b). As a result, sample ARFs using identical areal and point rainfall record lengths, were estimated in circular test catchments of 60 km^2 and 180 km^2, respectively. The results are shown in Figures 4.6 and 4.7.

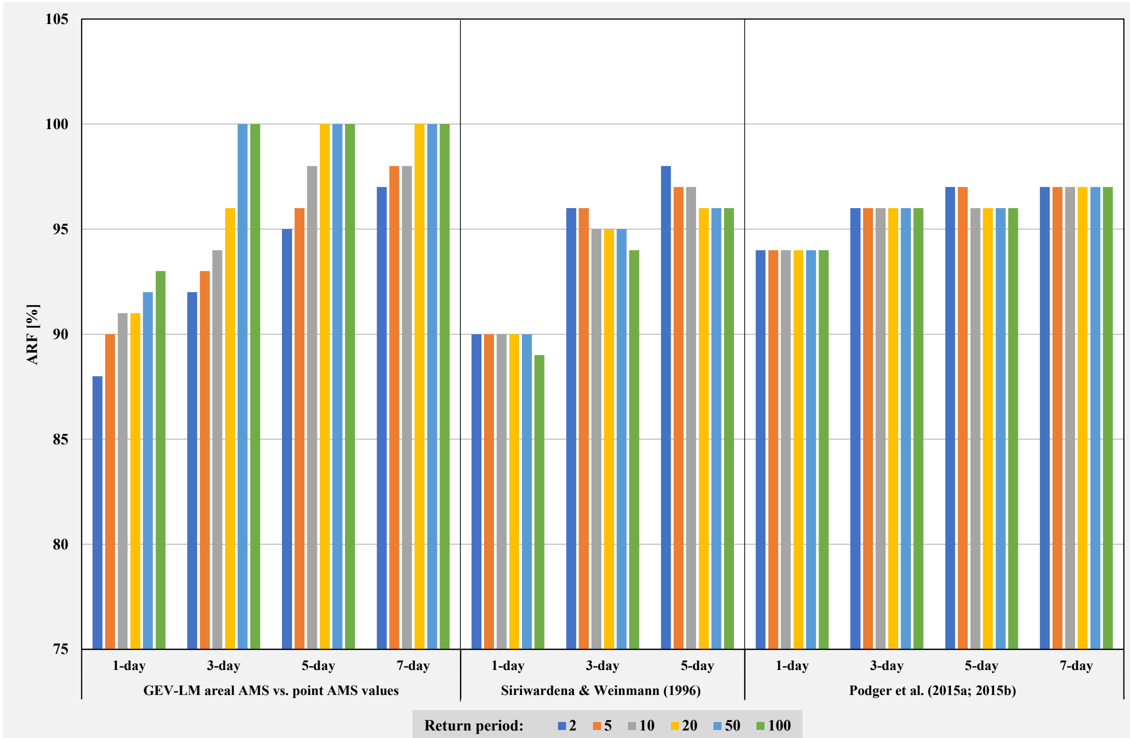

Figure 4.6: Comparison of sample ARFs estimated using identical record lengths and different methodologies for an area of 60 km^2

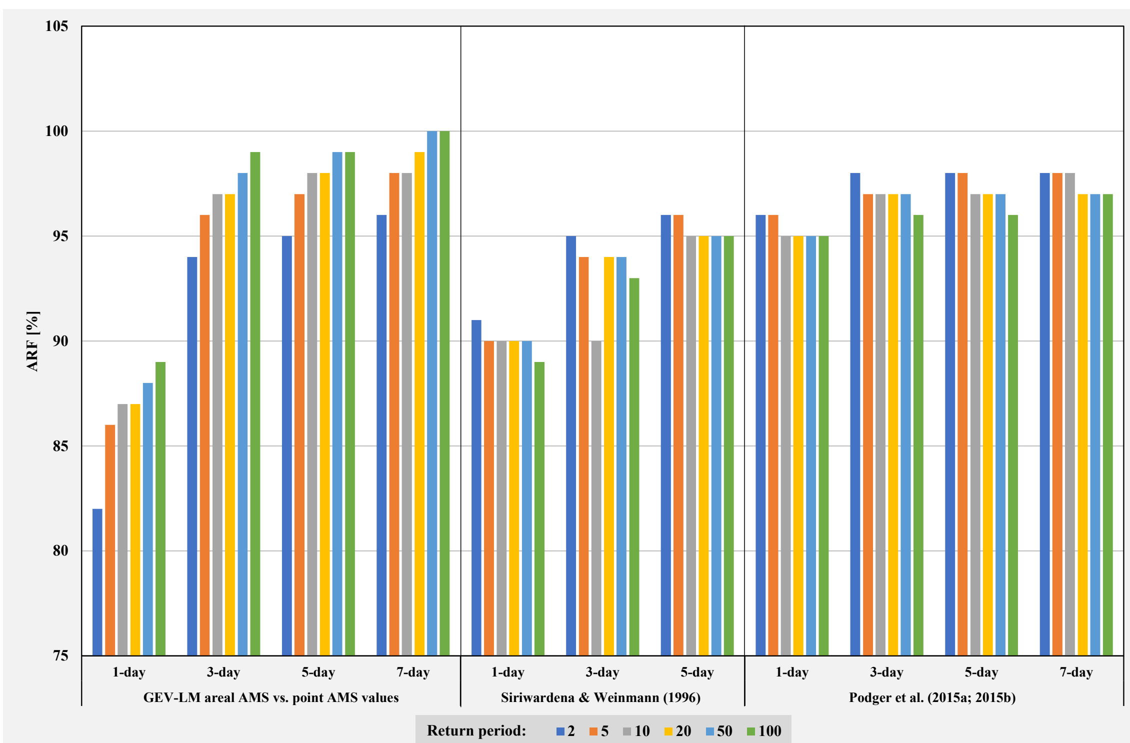

Figure 4.7: Comparison of sample ARFs estimated using identical record lengths and different methodologies for an area of 180 km^2

It is evident from Figures 4.6 and 4.7 that a similar trend between the ARFs estimated in this study, and those estimated for Australian climatological conditions, exists. Furthermore, the ratios (ARFs) established between the average areal and average design point rainfall, associated with specific return periods and storm durations, are within the same range. However, the ARFs in this research tend to increase with increasing return periods, as opposed to the Australian ARFs, which decrease with increasing return periods and may be attributed to the: (i) variation in methodological approaches used, and/or (ii) availability, robustness and sources of data used. Furthermore, the increase of ARFs with increasing return periods is expected, since it is assumed that higher magnitudes of rainfall (higher T-values) will most probably cover a larger portion of a catchment. The variations between the Australian return periods and corresponding storm durations, are insignificant when compared to the ARFs derived in this research. Furthermore, it is also assumed that a larger portion of a catchment will be covered when subjected to longer storm durations.

The latter assumptions, *i.e.*, (i) higher magnitudes of rainfall will probably cover a larger portion of a catchment, and (ii) that a larger portion of a catchment will be covered when subjected to longer storm durations, are clearly visible from the sample ARFs estimated in this research. It is therefore recommended that identical record lengths should be used in estimating sample ARFs to ensure that all potential extreme rainfall events are incorporated when areal and design point rainfall values are estimated.

4.4 Probabilistic Analyses using Short Record Lengths

The GEV_{LM}-based probabilistic analyses of areal and point rainfall values are acceptable when considered independently; however, when the ratios of higher return periods (T >100 years) between areal and design point rainfall values are considered, then inconsistencies are evident.

In Table 4.2, an example of estimated sample ARFs is listed for a circular test catchment (A = 2 475 km^2), located in the study area, and using a combined record length consisting of the same period and record length of 41 years.

Table 4.2: Sample ARFs estimated in a 2 475 km^2 circular catchment using a record length of 41 years

T [years]	Storm duration [days] and ARFs [%]			
	1-d	3-d	5-d	7-d
2	81.0	90.0	91.2	92.2
5	84.9	93.5	93.9	95.2
10	88.1	95.0	95.4	96.4
20	91.6	96.1	96.7	97.3
50	96.6	97.0	98.1	98.3
100	100.6	97.5	99.3	98.3

In Table 4.2, it is evident that the sample ARFs increase with both an increase in return period (T) and storm duration, while values exceeding 100% are typically associated with higher return periods (T >100 years). Hence, record length does have an impact on the probabilistic analyses, which agrees with the findings of Meier *et al.* (2015), and therefore the general rule of thumb, *i.e.*, limiting the frequency magnitude to 2N, where N equals the AMS record length, is recommended.

The latter general rule of thumb was adopted in the rest of this research, where the frequency magnitude was adjusted for each circular catchment, based on the available record lengths.

The final methodology for estimating ARFs and the associated results are discussed in the next chapter.

CHAPTER 5: FINAL METHODOLOGY AND RESULTS

This chapter contains the final methodology and results applicable to the development of a geographically-centred and probabilistically correct approach to estimate ARFs which are representative of the different rainfall producing mechanisms in South Africa. In addition, all the concerns and possible solutions documented in Chapter 4, were considered and implemented, while progressing with and adopting the final methodology.

5.1 Homogeneous Rainfall Regions

The original methodological procedure envisaged, with specific reference to the regionalisation of rainfall regions, entailed the use of the seven (7) RLMA&SI long duration rainfall regions as identified and used by Smithers and Schulze (2000b). In addition, Smithers and Schulze (2003; 2004) evaluated the 78 long duration rainfall clusters for homogeneity by means of a cluster analysis, and subsequently grouped these clusters into seven (7) long duration rainfall regions (*cf.* Chapter 2, Section 2.5.2). It was confirmed that these homogeneous rainfall regions do account for the variation in predominant rainfall types, seasonal, and regional factors. Furthermore, based on the recommendations made by the Reference Group members of WRC Project K5-2924 to adopt the 78 homogeneous rainfall clusters, as originally derived by Smithers and Schulze (2003; 2004) for the RLMA&SI design rainfall database, the following approach was adopted:

(a) Approximately 420 000 1' x 1' latitude and longitude grid points, as used in the RLMA&SI design rainfall database, were extracted, and grouped in the 78 homogeneous rainfall clusters. Thereafter, grid points falling within each of the 78 homogeneous clusters were plotted and subsequently outlined in an ArcGIS environment. The latter step enabled spatial orientation, *i.e.*, the visualisation and inspection of each outlined region's size (km^2) and the position of all grid points within and surrounding a specific cluster. In other words, this step was necessary to ensure that as many RLMA&SI grid points as possible fall within the same GIS-based cluster, seeing that a minor percentage of the plotted grid points fall beyond potential regional borders (*cf.* Figure 5.1). This is expected, according to Smithers and Schulze (2000a), since the RLMA&SI approach is based on the identification of potentially homogeneous regions by a cluster analysis of site characteristics which are then tested using the statistics of the sites in the region.

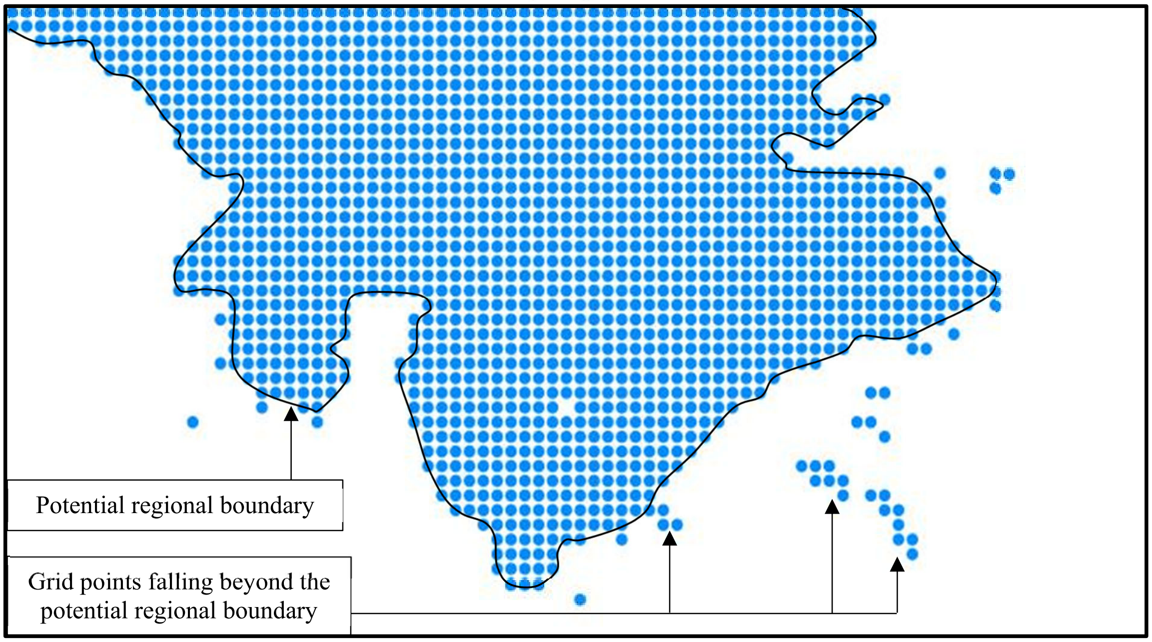

Figure 5.1: Example of grid points scattered around potential regional borders

(b) The regional boundaries, based on the individual 78 homogeneous clusters, were therefore adjusted until the contributing grid points, located within a particular cluster boundary, covers ≥ 90% of that region. The ratios of contributing grid points, located within a particular cluster boundary, will not directly affect the estimation of ARFs; however, it would ensure that each cluster's boundary delineated in ArcGIS mimics the original 78 homogeneous rainfall clusters as closely as possible.

(c) The position of the 1 779 selected rainfall stations (*cf.* Appendix A) were plotted in ArcGIS to visualise the location of each rainfall station within the 78 homogeneous rainfall clusters. Unfortunately, some of the 78 clusters are very small and, as a result, included a low number of rainfall stations. Hence, some of the GIS-based homogeneous rainfall clusters which are within the same seven (7) long duration rainfall regions were combined to increase the: (i) size of the cluster, and (ii) number of representative rainfall stations located within a cluster. Similarly, this was necessary to ensure that larger circular catchments (> 8 000 km^2) could be positioned within each cluster (*cf.* Table 5.1).

The above procedure resulted in the 46 relatively homogeneous rainfall clusters to be used in this research, with the rainfall stations being uniformly distributed across all 46 clusters (*cf.* Figure 5.2).

Table 5.1: The 1 779 rainfall stations within each of the 46 relatively homogeneous rainfall clusters

Cluster	Number of stations	Area [km^2]	Cluster	Number of stations	Area [km^2]	Cluster	Number of stations	Area [km^2]
1	33	16 315	17	36	11 492	33	59	36 560
2	33	24 564	18	71	37 038	34	22	10 562
3	39	19 576	19	30	21 742	35	24	40 098
4	49	16 612	20	33	10 202	36	51	23 707
5	34	18 302	21	42	80 478	37	36	16 370
6	56	40 073	22	30	21 244	38	53	38 065
7	38	15 304	23	55	36 188	39	24	12 802
8	49	17 347	24	30	24 099	40	28	28 671
9	37	62 327	25	50	19 089	41	16	13 265
10	33	15 758	26	18	9 448	42	21	65 540
11	22	16 517	27	30	38 523	43	34	13 378
12	35	20 276	28	32	32 129	44	56	20 577
13	66	14 484	29	35	26 847	45	70	40 315
14	28	37 279	30	20	71 390	46	102	38 371
15	41	18 221	31	26	15 036			
16	38	41 448	32	14	19 807			

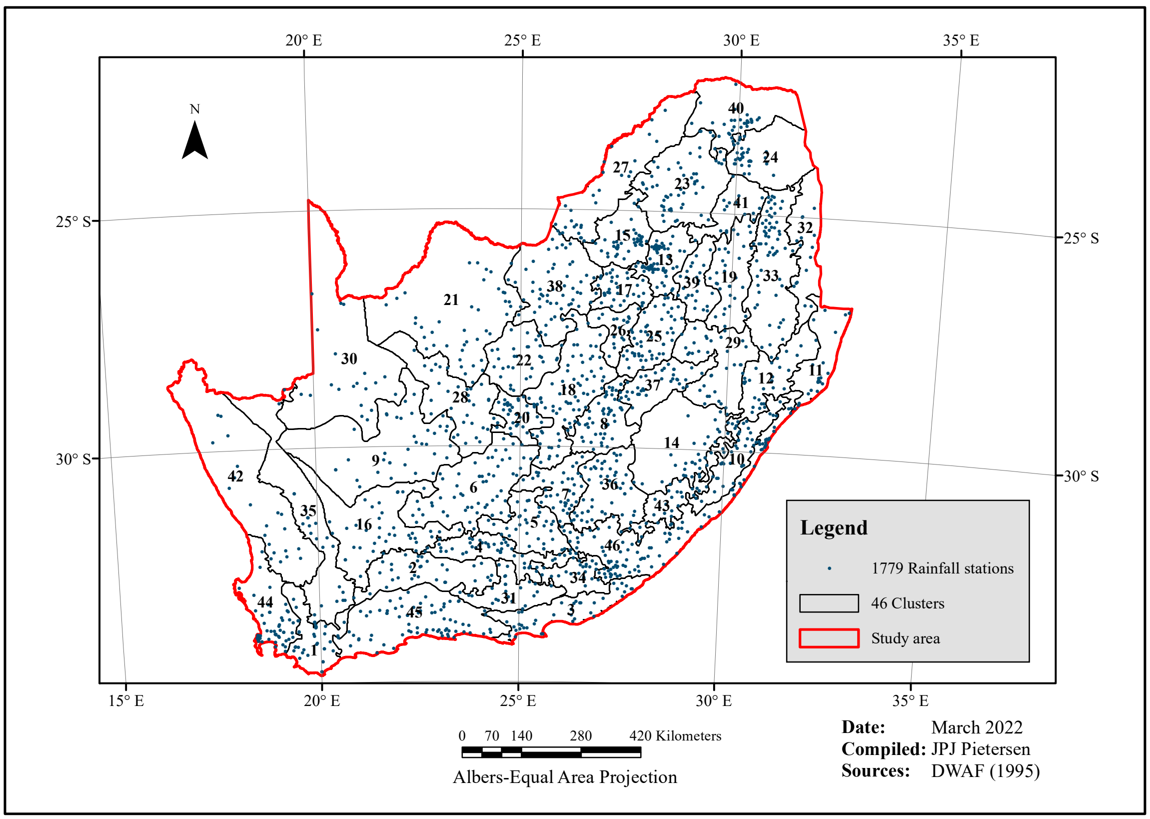

Figure 5.2: Forty-six relatively homogeneous clusters delineated from the 78 homogeneous rainfall clusters

5.2 User-defined Circular Catchments

An alternating-area approach was used to generate the artificial circular catchments within each of the 46 clusters. Given the more suitable functionalities available in the Quantum Geographical Information System (QGIS) software (open-source software available in public domain), it was used instead of ArcGIS to create the various artificial circular catchments. In all cases, the location and size of each circular catchment depended on the locality of any two pre-selected rainfall stations, which overlay the circumference of a circular catchment. This process required the optimum, manual positioning of circular catchments between the identified rainfall stations, which subsequently resulted in artificial circles consisting of various sizes (*cf.* Chapter 4, Section 4.1).

The following steps were necessary to achieve the successful positioning of the artificial circular catchments in a QGIS environment:

(a) In each of the 46 clusters, a temporary polygon scratch layer was created using the World Geodetic System 1984 (WGS 84) coordinate system and Pseudo-Mercator projection, respectively.

(b) One Thiessen polygon grid for South Africa, including all the 1 779 rainfall stations, was generated.

(c) The *Shape Digitising Toolbar* was activated, and thereafter, the tool *Add Circle from 2 Points* were utilised. A manual process, *i.e.*, computer mouse, was used to manually place the artificial circles between the most distant rainfall stations available.

(d) The maximum number of artificial circles per cluster were limited to: (i) one hundred (100) circles, which are based on the maximum number of available inputs possible in the customised Microsoft Excel processing tool developed, and (ii) the number of rainfall stations available in each cluster, which necessarily did not include, depending on criteria (i), all possible combinations of available rainfall stations within each cluster.

(e) The Thiessen polygon grid (Step (b)) and all the circular catchments within a particular cluster were used to *Clip* (extract) the contributing Thiessen polygon areas.

(f) All circular catchments and Thiessen polygon areas clipped in the previous step were intersected (*Join*) to ensure that the overlapping circles consist of a contributing Thiessen weight set.

All the above steps resulted in: (i) circular catchment sizes, (ii) contributing rainfall stations, and (iii) associated Thiessen weights, for the purpose of extracting infilled daily rainfall data.

The maximum circular catchment sizes, associated with each cluster, were overall limited by the size of the clusters initially delineated. In addition, the maximum circular catchment sizes were also limited by the quality criteria as listed in Section 5.3.1. As a result, a total of 2 550 artificial circular catchments (*cf.* Figure 5.3), ranging from 0.07 km^2 to 18 985 km^2, were manually placed across South Africa.

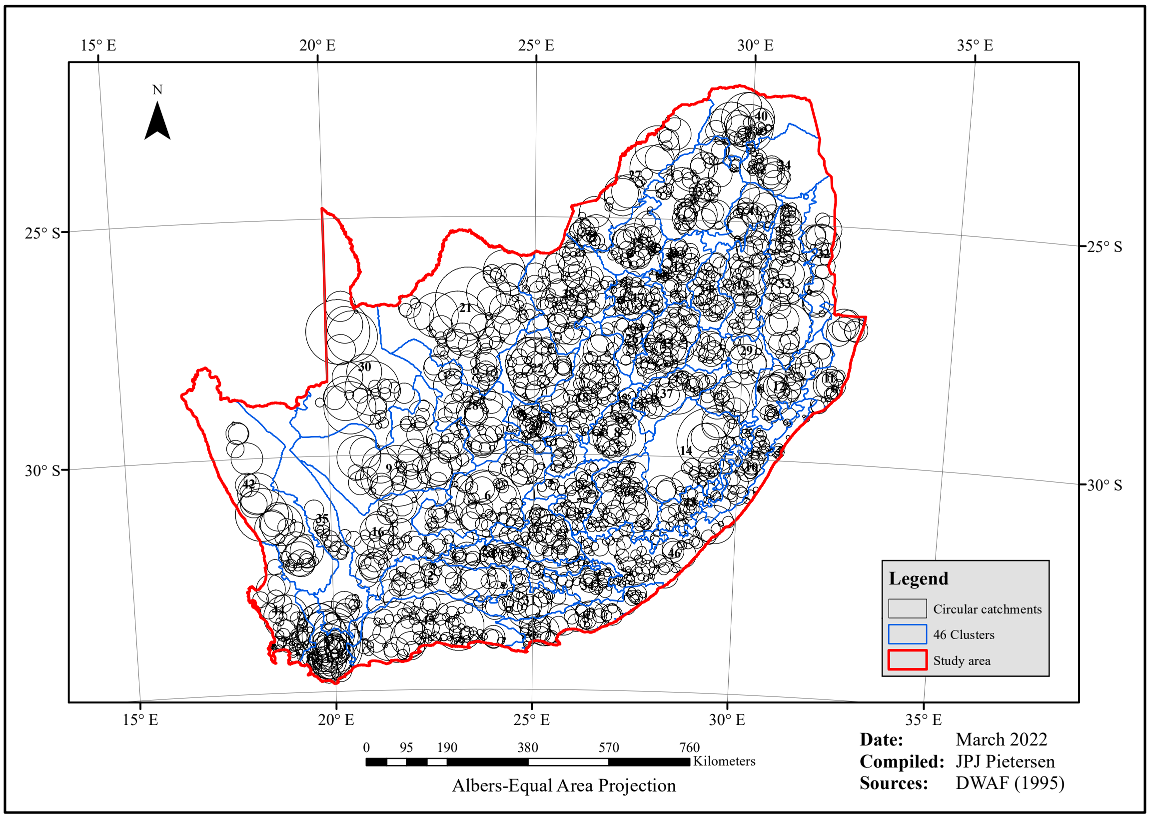

Figure 5.3: Placement of 2 550 artificial circular catchments in South Africa

The 'open spaces' evident between the circular catchment areas in Figure 5.3 are ascribed to the lack of adequate rainfall stations. 'Larger' circular catchments cannot simply be placed to close these gaps, since the minimum number of rainfall stations per circular catchment criterion (*cf.* Section 5.3.1) needs to be maintained. The total number of rainfall stations present in each of the 46 clusters varied from 14 to 102 stations, while the total number of circular catchments per cluster varied from 23 to 100 (*cf.* Table 5.2).

Table 5.2: Circular catchment information in each of the 46 clusters

Cluster	Size [km^2]	Number of rainfall stations	Number of circles	Circle size ranges [km^2]	
				Minimum	Maximum
1	16 315	33	72	0.30	17 513
2	24 565	33	43	1.91	7 030
3	19 577	39	59	5.47	4 273
4	16 620	49	70	2.70	3 631
5	18 302	34	53	4.67	12 600
6	40 073	56	76	0.51	18 654
7	15 304	38	56	0.44	5 501
8	17 347	49	67	1.01	10 207
9	62 327	37	55	5.84	16 143
10	15 757	33	45	1.79	2 255
11	16 517	22	33	0.68	4 473
12	20 276	35	52	1.63	7 109
13	14 484	66	100	0.24	4 332
14	37 279	28	39	0.5	18 711
15	18 221	41	64	1.19	8 462
16	41 448	38	53	0.56	6 457
17	11 492	36	55	0.33	5 118
18	37 038	71	95	2.07	6 137
19	21 742	30	46	2.16	11 025
20	10 202	33	51	1.05	5 402
21	80 478	42	70	10.51	18 985
22	21 244	30	46	7.23	17 821
23	36 188	55	83	2.53	8 140
24	24 099	30	43	0.99	3 479
25	19 089	50	71	0.08	12 632
26	9 448	18	29	8.02	2 023
27	38 523	30	43	2.91	8 709
28	32 129	32	43	1.90	16 113
29	26 848	35	48	2.10	7 174
30	71 390	20	29	47.40	18 035
31	15 036	26	34	4.62	7 628
32	19 807	14	23	2.06	7 065
33	36 560	59	82	19.47	11 923
34	10 553	22	32	2.67	3 478
35	40 098	24	30	0.73	2 938
36	23 707	51	70	0.13	10 381
37	16 369	36	57	0.37	8 712
38	38 065	53	82	17.43	10 919
39	12 802	24	35	41.41	8 877
40	28 671	28	40	10.56	12 216
41	13 265	16	23	8.91	6 521
42	65 548	21	37	44.74	12 138
43	13 378	34	41	10.5	1 407
44	20 577	56	77	0.07	8 645
45	40 316	70	98	0.08	15 674
46	38 371	102	100	0.73	6 384
Total	**1 267 446**	**1 779**	**2 550**		

5.3 Extraction, Infilling and Analyses of Observed Rainfall Data

A daily rainfall database was established by evaluating, preparing, and extracting daily rainfall data from 1 779 daily rainfall stations in South Africa by using the data available in the Daily Rainfall Extraction Utility (DREU; Lynch, 2004). These identified rainfall stations have at least 30 years of individual data and have been previously used by Smithers and Schulze (2000b; 2003; 2004) to estimate design rainfall values for durations of 1 to 7 days.

The steps explained in Section 5.2 resulted in multiple regional/cluster files comprising of various circular catchment sizes, rainfall station numbers and Thiessen polygon areas necessary for the extraction and infilling of daily rainfall data. Due to the large number of circular catchments placed in each of the 46 clusters, the overlapping of circular catchments was evident. Consequently, this resulted in daily rainfall data from the same rainfall stations being used multiple times within a particular cluster. In principle, this was not regarded as problematic, while it also contributed to the 'smooth' transition between the different clusters.

5.3.1 Criteria for selecting infilled rainfall data

In considering the impact that an incomplete month and consequently an incomplete year could have on the record length of a particular rainfall station, the default DREU infilling techniques, *e.g.*, inverse distance weighting, expectation maximisation, median ratio, and/or monthly infilling (*cf.* Chapter 2, Section 2.3) for the extraction of quality controlled infilled daily rainfall data, were used. The rainfall data infilling process was carefully interrogated by considering the following criteria as **Criterion 1**:

(a) Infilling, as applicable to certain rainfall stations within a particular circular catchment, was limited to periods within the observed record (*N*) under consideration, *i.e.*, no backward extrapolation of the observed record in time.

(b) Circular catchments within a particular cluster were removed where 25% or more of the stations required infilling to a minimum of 30 years combined. Table 5.3 is an example of a case where infilling was not considered; two rainfall stations with a combined record length of 11 years where 50% of the stations that required infilling.

Table 5.3: Example of circular catchment where infilling was rejected

No	Station Name	Station number	First year recorded	Last year recorded	Record used		Combined record used [years]
					Start	End	
3	NELSPOORT	0093005 W	1950	1997	**1950**	**1961**	**11**
4	KROMRIVIER	0116023 W	1903	1961			

Table 5.4 is illustrative of a typical circular catchment consisting of 18 rainfall stations. Infilling was required to ensure a minimum combined record length of 30 years (1934 to 1964). Only one rainfall station did not meet the above criteria with its record ending in 1964. Infilling was therefore applied to the next lowest station (1976), which resulted in a record length of 42 years.

Table 5.4: Example of circular catchment where infilling was applied

No	Station Name	Station number	First year recorded	Last year recorded	Record used		Combined record used [years]
					Start	End	
215	HOUD CONSTANT	009539 W	1927	1986	**1934**	**1976**	**42**
216	VAN DER WALTSHOEK	0095428 W	1927	1997			
217	WATERFALL	0095635 W	1913	1976			
218	WINTERHOEK	0095823 W	1890	1997			
219	VAN RHYNESVELD'S DAM	0096044 W	1934	1977			
220	GROOTHOEK	0096094 W	1920	1987			
221	ROODEBLOEM	0096101 W	1888	1997			
222	GRAAF-REINET TNK	0096045 A	1894	1980			
223	ABERDEEN (TNK)	0095119 W	1881	1991			
224	NIEU-BETHESDA (POL)	0119082 W	1885	1997			
225	KENDREW ESTATES	0073871 W	1890	1997			
226	WELLWOOD	0119208 W	1874	1982			
227	BLOEMHOF	0096272 W	1899	1997			
228	GLEN HARRY	0096366 W	1900	1955			
229	EXCELSIOR	0096551 W	1925	1987			
230	GROENKLOOF	0096680 W	1906	1997			
231	KLIPFONTEIN	0074363 W	1892	1997			
232	BETHESDAWEG (SAR)	0119444 W	1900	1988			

Hence, the infilling of rainfall data within a particular circular catchment and applicable to only certain rainfall stations, not only ensured that combined record lengths exceeding 30 years could be achieved, but it also ensured sufficient synchronisation and overlapping between the various rainfall station recordings to extract a complete areal and point rainfall AMS from each circular catchment. This process resulted in the longest possible rainfall record length with mutual starting and ending dates to extract the areal catchment and point rainfall AMS. This approach was followed in all 46 clusters.

Furthermore, an additional filtering criterion, *i.e.*, **Criterion 2,** was applied to the circular catchments. Circular catchments were removed where the minimum number of rainfall stations in each circular catchment was less than the revised criteria, as recommended by Siriwardena and Weinmann (1996), namely, a minimum of two rainfall stations for catchment areas up to 100 km^2, thereafter, a minimum of three stations for catchment areas up to 500 km^2, plus one additional station for every 500 km^2 thereafter.

Figure 5.4 is illustrative of the circular catchment's relative frequency distribution within the 46 clusters, while Table 5.5 contains the total number of circular catchments removed from the 2 550 circular catchments for the purpose of probabilistic analyses. A total of 2 053 (80.5%) circular catchments were used in the final analyses.

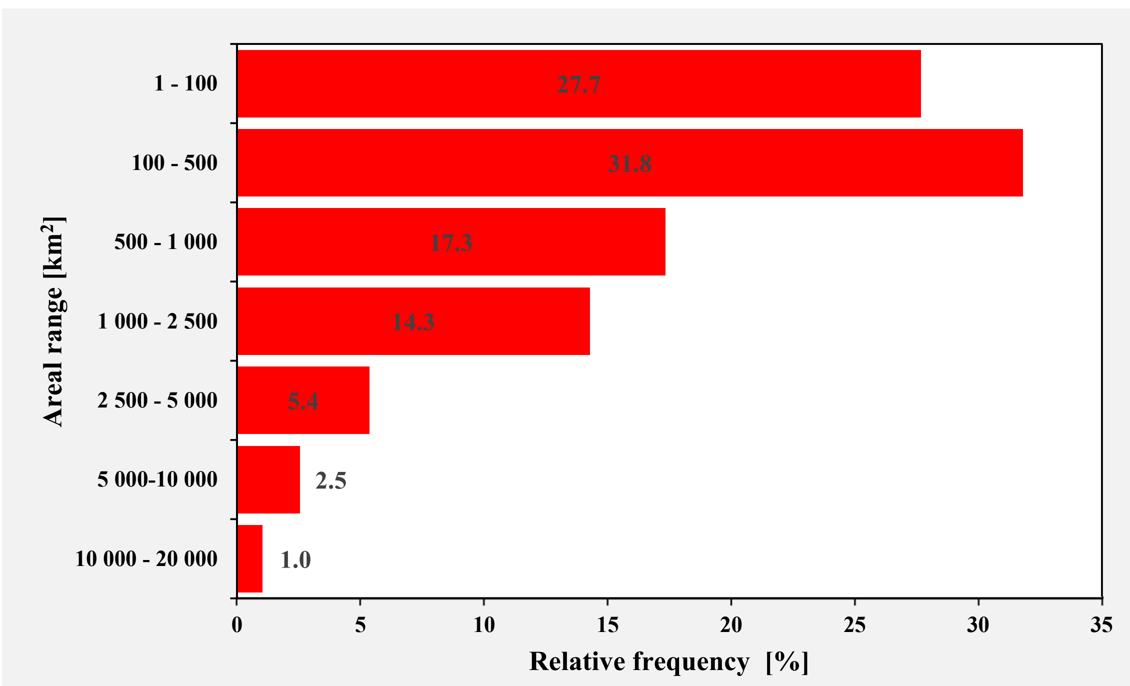

Figure 5.4: Relative frequency distribution of the 2 053 circular catchments within the 46 clusters

Table 5.5: Number of circular catchments removed for the purpose of probabilistic analyses

Cluster	Number of circles	Number of circles removed (Criterion 1)	Number of circles removed (Criterion 2)	Circles excluded [%]
1	72	1	9	13.9
2	43	0	7	16.3
3	59	0	13	22.0
4	70	8	5	18.6
5	53	1	4	9.4
6	76	0	22	28.9
7	56	0	7	12.5
8	67	0	5	7.5
9	55	0	28	50.9
10	45	2	4	13.3
11	33	0	11	33.3
12	52	0	9	17.3
13	100	7	1	8.0
14	39	0	9	23.1
15	64	0	4	6.3
16	53	2	14	30.2
17	55	1	9	18.2
18	95	3	14	17.9
19	46	3	8	23.9
20	51	5	1	11.8
21	70	3	24	38.6
22	46	1	7	17.4
23	83	5	10	18.1
24	43	0	4	9.3
25	71	0	8	11.3
26	29	0	3	10.3
27	43	0	15	34.9
28	43	0	12	27.9
29	48	0	9	18.8
30	29	0	18	62.1
31	34	2	6	23.5
32	23	0	9	39.1
33	82	0	17	20.7
34	32	0	5	15.6
35	30	0	6	20.0
36	70	2	8	14.3
37	57	1	4	8.8
38	82	11	17	34.1
39	35	0	6	17.1
40	40	0	3	7.5
41	23	0	4	17.4
42	37	0	19	51.4
43	41	1	6	17.1
44	77	1	6	9.1
45	98	0	17	17.3
46	100	0	11	11.0
Total	**2 550**	**497**		**Avg.: 20.8**

5.3.2 Averaging of daily rainfall data

This section highlights the limitations associated with the Thiessen polygon method (Wilson, 1990), especially when the actual spatial distribution of a rainfall event is considered. These limitations became evident during the weighting procedure of areal and point rainfall AMS values. An irregular spatial distribution of actual rainfall patterns can originate over larger distances, and under such circumstances, the Thiessen polygon method can yield erroneous results. In other words, the Thiessen polygon method will most probably neglect certain rainfall events, especially when rainfall stations are distant. According to Sugawara (1992), and as discussed in Chapter 2, the average point and/or areal rainfall estimated using the Thiessen polygon method becomes questionable when applied to a low-density rainfall-monitoring network, since the Thiessen polygons may not represent the actual rainfall processes in a specific polygon as they interrupt the spatial transitions of the natural occurring process. Unfortunately, this is the status quo (or dilemma) in South Africa, especially with a declining number of operational SAWS rainfall stations since the 1960s. In terms of this research, the 1 779 rainfall stations selected, cover a total surface area of 1.22 million km^2, which equates approximately 1 rainfall station for every 686 km^2.

Despite these shortcomings, the Thiessen polygon method is still recommended and was used in several international ARF studies, *e.g.*, Bell (1976), Stewart (1989), Siriwardena and Weinmann (1996), and Podger *et al.* (2015). Therefore, the Thiessen polygon method was adopted as the preferred method to estimate weighted point and areal AMS data sets.

Therefore, the GIS feature classes (shape files) containing the spatial features of the 1 779 daily rainfall stations and location of the 2 053 artificial circular catchments were generated in the QGIS environment. The large amount of data and repetitive computations required the use of the *Create Voronoi (Thiessen) Polygons* extension under the *Geoprocessing* tools in QGIS to generate representative Thiessen (Voronoi) polygon weights for each of the 2 053 circular catchments. Therefore, multiple circular catchments, within each of the 46 clusters associated with different combinations of rainfall stations and Thiessen weights, were generated for South Africa.

5.3.3 Extraction of areal and point AMS

The areal and point AMS extracted for durations of 1-day, 3-day, 5-day, and 7-day were used for the probabilistic analyses. The 1-day fixed time interval point and areal rainfall AMS were firstly obtained from the observed rainfall data. To estimate the point AMS for each circular catchment, the point AMS for each rainfall station was extracted first, and then each AMS was individually multiplied with a corresponding Thiessen weight [Eq. [(2.2)]. This resulted in one weighted point AMS for each circular catchment, representative of all rainfall stations in and surrounding a particular circular catchment. To estimate the areal AMS, the Thiessen weights [Eq. [(2.2)] were applied daily at each rainfall station, *i.e.*, daily point rainfall values, from each rainfall station within a particular circular catchment, were multiplied with their corresponding Thiessen weight to estimate the daily catchment rainfall, after which, the areal AMS was extracted from the daily catchment rainfall.

To obtain the 3-day, 5-day, and 7-day fixed time interval areal and point rainfall AMS, a 'moving window' was applied to the 1-day fixed time interval point rainfall to provide the accumulated 3-day, 5-day and 7-day totals, respectively. The point AMS was extracted from the *n*-day highest accumulated values within each hydrological year and subsequently used as the 3-day, 5-day or 7-day fixed time interval point rainfall AMS values. To estimate the areal *n*-day AMS, the Thiessen weights [Eq. [(2.2)] were applied to the *n*-day totals at each rainfall station, *i.e.*, *n*-day totals from each rainfall station within a particular circular catchment were multiplied with their corresponding Thiessen weights to estimate the weighted *n*-day totals for the circular catchment, after which, the *n*-day areal AMS was established. This procedure resulted in one weighted areal AMS for durations of 1, 3, 5, and 7 days.

5.4 Probabilistic Analyses of Weighted Areal and Point AMS

The probabilistic analyses were conducted for each of the 2 053 circular catchments' (*cf*. Chapter 4, Section 4.2) areal and point rainfall AMS for a range of storm durations (*e.g.*, 1, 3, 5 and 7-day), and return periods (*e.g.*, 2, 5, 10, 20, 50 and 100-year) by using the GEV_{LM} probability distribution. The rationale for using the GEV_{LM} probability distribution was explained and justified in Chapter 4 (Section 4.2). This procedure resulted in 98 496 equally distributed areal and design point rainfall values associated with various storm durations (*e.g.*, 1, 3, 5 and 7-day) and return periods (*e.g.*, 2 to 100-year).

5.5 Estimation of Sample ARFs

The estimation of ARFs was based on a 'modified version' of Bell's method (1976; *cf*. Chapter 2, Section 2.9.1), since the AMS of point and areal rainfall were used as opposed to the PDS used by Bell (1976). This modification will reflect the variation of ARFs with return period, instead of using equally ranked observations curtailed to a common base period. Sample ARF values applicable to the 2 053 circular catchments were estimated using Eq. (5.1) and expressed as the ratio between the average areal catchment design rainfall and the average design point rainfall estimates for corresponding return periods.

$$ARF_{Sample} = \frac{A_{DR}}{P_{DR}} \times 100 \qquad [5.1]$$

where

ARF_{Sample} = circular-area sample ARF [%],

A_{DR} = average areal design rainfall [mm], and

P_{DR} = average design point rainfall [mm].

This procedure resulted in a total of 49 248 sample ARF values representative of all 46 clusters.

5.6 Derivation of Regional Empirical ARF Equations

This section primarily focuses on the establishment of a regression model that will allow the prediction of ARFs by considering a pre-defined set of independent predictor variables. Firstly, the impact of potential outliers and normal/log-transformed data on the regression analyses is evaluated. Secondly, the further merging of the 46 clusters to coincide with the original seven (7) RLMA&SI long duration rainfall regions and to meet the minimum World Meteorological Organisation (WMO) rainfall stations/km² criteria, is discussed. Lastly, the derivation of a non-linear (second-order polynomial) log-transformed empirical ARF equation associated with the final five (5) ARF Regions deduced from the clusters, is discussed.

Initially, linear backward stepwise multiple regression analyses with deletion were performed at a 95% confidence level to estimate the relationship between the dependent criterion variable (ARF) and the independent predictor variables within each of the 46 clusters. The following independent predictor variables were considered for inclusion: (i) catchment area [*A*, km²], (ii) storm duration [*D*, days], and (iii) return period [*T*, years].

In each case, the Goodness-of-Fit (GOF) statistics of normal and log-transformed data were evaluated for the independent criterion and dependant predictor variables using the coefficient of correlation (r^2) and the standard error of the estimate (*SE*). Standardised residuals were also estimated for each of the 46 clusters to identify and exclude potential outliers. Excluding data points from a data set is not always recommended; however, it may sometimes be necessary to exclude outliers, should the model integrity allow this (*cf.* Chapter 2, Section 2.11). As recommended by Field (2000), the standardised residuals which exceeded the ± 2.5 margin were removed from the data sets.

As an example of the exclusion of outliers, a linear regression analysis based on 1 960 data points (normal data) in Cluster 1 (*cf.* Figure 5.2) was conducted and resulted in a coefficient of correlation (r^2) equal to 0.45. The standardised residuals were estimated and all values exceeding ±2.5 were identified and removed accordingly. Subsequently, 652 data points having standardised residual values > ±2.5, were removed and the updated linear regression improved with $r^2 = 0.61$. In other words, by using a linear regression based on normal data, approximately 33% of the data points are regarded as outliers, which is a very high percentage. Subsequently, a linear regression analysis based on log-transformed data was conducted on the same data set. The standardised residuals were estimated and all values exceeding ±2.5 were identified and removed. As a result, only 37 data points (2.5%) were removed, while the r^2 value increased to 0.72. Hence, the log-transformed data resulted in a much better fit with less data points being regarded as outliers. Overall, the total number of outliers being removed was insignificant compared to the size of the original data set in the 46 clusters, while using only log-transformed data resulted in improved GOF statistics. As summarised in Table 5.6, the number of outliers removed ranged from 6 to 50 with a total of 1 263 (2.6%) ARF residuals ≥ ±2.5 being removed.

Table 5.6: Outliers removed from each cluster

RLMA&SI long duration region	Cluster	Number of sample ARF estimates	Number of outliers removed	ARF residuals excluded [%]
1	4	1 368	32	2.3
1	7	1 176	34	2.9
1	8	1 488	49	3.3
1	13	2 208	58	2.6
1	15	1 440	40	2.8
1	17	1 080	35	3.2
1	18	1 872	50	2.7
1	23	1 632	41	2.5
1	25	1 512	45	3.0
1	26	624	6	1.0
1	27	672	8	1.2

RLMA&SI long duration region	Cluster	Number of sample ARF estimates	Number of outliers removed	ARF residuals excluded [%]
1	31	624	18	2.9
1	38	1 296	33	2.5
2	3	1 104	21	1.9
2	11	528	18	3.4
2	45	1 944	38	2.0
5	10	936	27	2.9
5	14	720	18	2.5
5	34	648	19	2.9
5	36	1 440	40	2.8
5	37	1 248	32	2.6
5	43	816	20	2.5
5	46	2 136	44	2.1
3	1	1 488	37	2.5
3	35	576	13	2.3
7	42	432	12	2.8
7	44	1680	37	2.2
4	12	1 032	32	3.1
4	19	840	24	2.9
4	24	936	24	2.6
4	29	936	26	2.8
4	32	336	10	3.0
4	33	1 560	43	2.8
4	39	696	8	1.1
4	40	888	23	2.6
4	41	456	11	2.4
6	2	864	22	2.5
6	5	1 152	31	2.7
6	6	1 296	30	2.3
6	9	648	19	2.9
6	16	888	21	2.4
6	20	1 080	32	3.0
6	21	1 032	23	2.2
6	22	912	28	3.1
6	28	744	24	3.2
6	30	264	7	2.7
Total		**49 248**	**1 263**	**Avg.: 2.6**

In general, it is accepted that individual clusters could demonstrate better results in terms of GOF criteria. Although GOF criteria are important, the further merging of the 46 clusters to coincide with the original seven (7) RLMA&SI long duration rainfall regions was necessary to increase the size of the clusters and the number of rainfall stations within a particular cluster to meet the minimum required number of rainfall stations/km² criteria. This step was regarded as crucial, since the derivation of sample ARFs, especially for larger circular catchments, depends on the number of available rainfall stations within a particular homogeneous region*, i.e.,* a minimum of two rainfall stations for catchment areas up to 100 km^2, thereafter, a minimum of three rainfall stations for catchment areas up to 500 km^2, plus one additional rainfall station for every 500 km^2 thereafter. In addition, as highlighted in Chapter 2 (Section 2.5.2), Smithers and Schulze (2003; 2004) evaluated the 78 long duration rainfall clusters for homogeneity, by means of a cluster analysis, and subsequently grouped these clusters into seven (7) long duration rainfall regions and confirmed that these homogeneous rainfall regions do account for the variation in predominant rainfall types, seasonal, and regional factors in South Africa.

Hence, the 46 clusters were merged, providing that the: (i) clusters are located within the same RLMA&SI long duration rainfall regions, and (ii) GOF statistics of each merged region do not decrease with more than 15% when compared to the original average GOF statistics obtained in the individual clusters. It was evident that the GOF statistics, from individual clusters, remain practically the same when the 46 clusters, located within each of the seven (7) RLMA&SI long duration rainfall regions are selectively grouped and merged as shown in Figure 5.5. Typically, percentage differences ranged between $1.99\% \leq r^2 \leq 8.35\%$ and $4.65\ \% \leq SE \leq 13.59\%$. Acceptable GOF criteria results (*cf.* Table 5.7) were also achieved after merging, *e.g.*, individual clusters ($0.59 \leq r^2 \leq 0.84$ and $2.1\% \leq SE \leq 4.9\%$) versus merged regions ($0.62 \leq r^2 \leq 0.79$ and $2.3\% \leq SE \leq 3.8\%$).

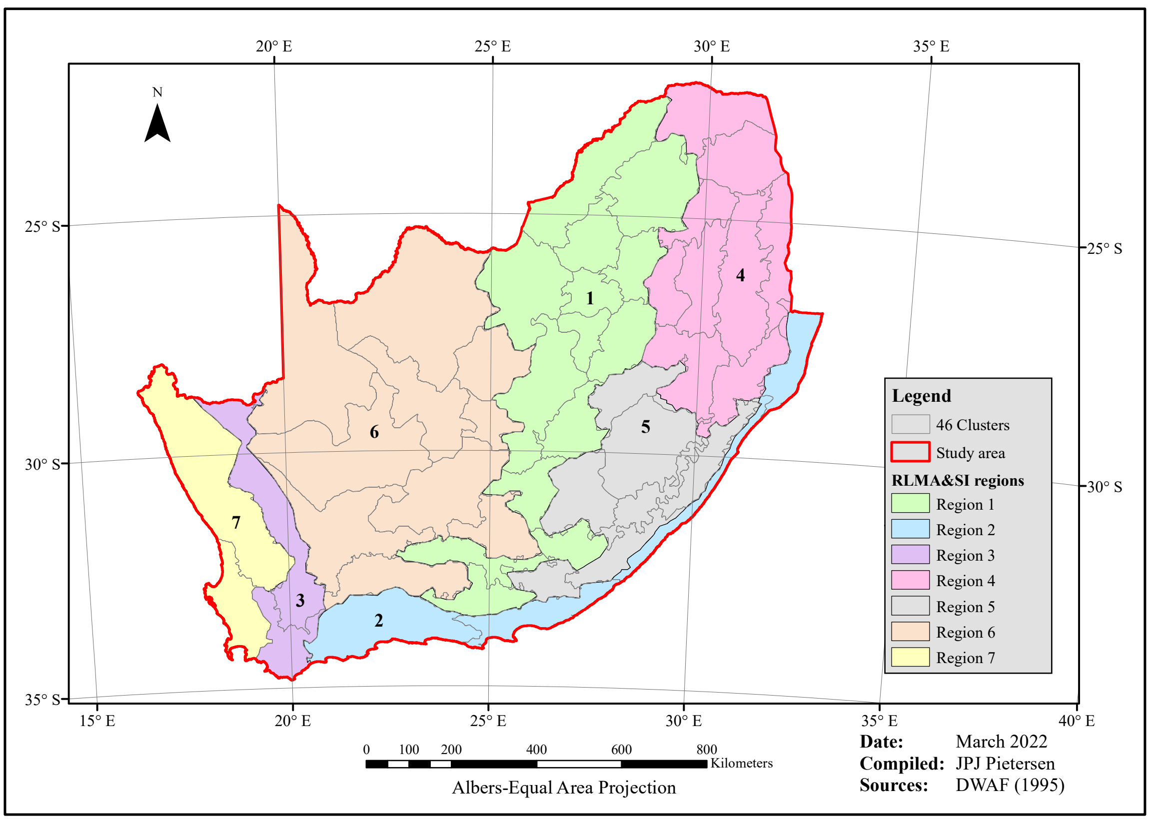

Figure 5.5: Forty-six ARF regions located within the seven RLMA&SI long duration rainfall regions (after Smithers and Schulze, 2000b)

Table 5.7: GOF statistics before and after merging of the 46 clusters

RLMA&SI long duration region	Cluster	r^2	*SE* [%]	After merging of clusters			
				r^2	*SE* [%]	r^2	*SE* [%]
1	4	0.759	2.882	0.722 *0.748	3.445 *3.292		
1	7	0.742	2.530				
1	8	0.749	2.718				
1	13	0.739	2.583				
1	15	0.700	2.737				
1	17	0.702	3.834				
1	18	0.779	3.224				
1	23	0.775	3.845				
1	25	0.761	3.387				
1	26	0.747	3.383				
1	27	0.712	4.927				
1	31	0.746	3.652				
1	38	0.815	3.098				
2	3	0.755	2.830	0.622 *0.646	3.573 *3.143	0.671 *0.720	3.402 *2.995
2	11	0.588	2.727				
2	45	0.607	3.872				
5	10	0.694	3.493	0.701 *0.750	3.250 *2.931		
5	14	0.807	3.436				
5	34	0.800	2.065				
5	36	0.744	3.111				
5	37	0.688	3.321				
5	43	0.800	2.379				
5	46	0.719	2.713				
3	1	0.718	3.463	0.685 *0.727	3.751 *3.423	0.648 *0.707	3.302 *2.926
3	35	0.735	3.382				
7	42	0.732	2.685	0.660 *0.688	2.273 *2.430		
7	44	0.643	2.175				
4	12	0.789	3.357	0.787 *0.803	3.458 *3.274		
4	19	0.824	3.284				
4	24	0.747	3.364				
4	29	0.775	3.779				
4	32	0.802	3.144				
4	33	0.801	2.883				
4	39	0.828	2.944				
4	40	0.835	3.220				
4	41	0.829	3.493				
6	2	0.823	3.040	0.769 *0.790	3.253 *3.042		
6	5	0.796	2.597				
6	6	0.795	2.725				
6	9	0.755	3.532				
6	16	0.779	3.823				
6	20	0.785	2.590				
6	21	0.793	3.134				
6	22	0.791	3.261				
6	28	0.828	3.195				
6	30	0.751	2.524				

* Averages based on individual ARF regions

Given the limited number of original clusters located within the RLMA&SI long duration rainfall Region 2 (3 clusters) and Region 3 (2 clusters), as well as their close geographical proximity to Regions 5 and 7, the clusters in these RLMA&SI regions were subjected to a second round of merging. The second round of merging of the latter clusters was justified by the need for 'practicality for application' and 'uniformity'. Typically, RLMA&SI Region 2 (*cf.* Figure 5.5) covers a very narrow strip along the east and south coast; hence, making it virtually impossible for a practitioner to accurately select an ARF region to estimate ARFs. In the case of RLMA&SI Region 7 (*cf.* Figure 5.5), the limited number of rainfall stations (ratio of 1 078 km^2/station) necessitates the required merging with Region 3 to improve the rainfall station density, *i.e.*, 989 km^2/station, which is closer to the recommendation made by the World Meteorological Organization (1994), *i.e.*, 575 – 900 km^2/station for coastal to interior plain areas. Hence, this resulted in the merging of Regions 2 & 5 and 3 & 7 to result in ARF Regions 2 and 3, respectively. Overall, the GOF statistics before and after merging proved to be very similar, *e.g.*, r^2 differences < 0.05 and *SE* differences < 1%.

The final five (5) ARF regions are shown in Figure 5.6.

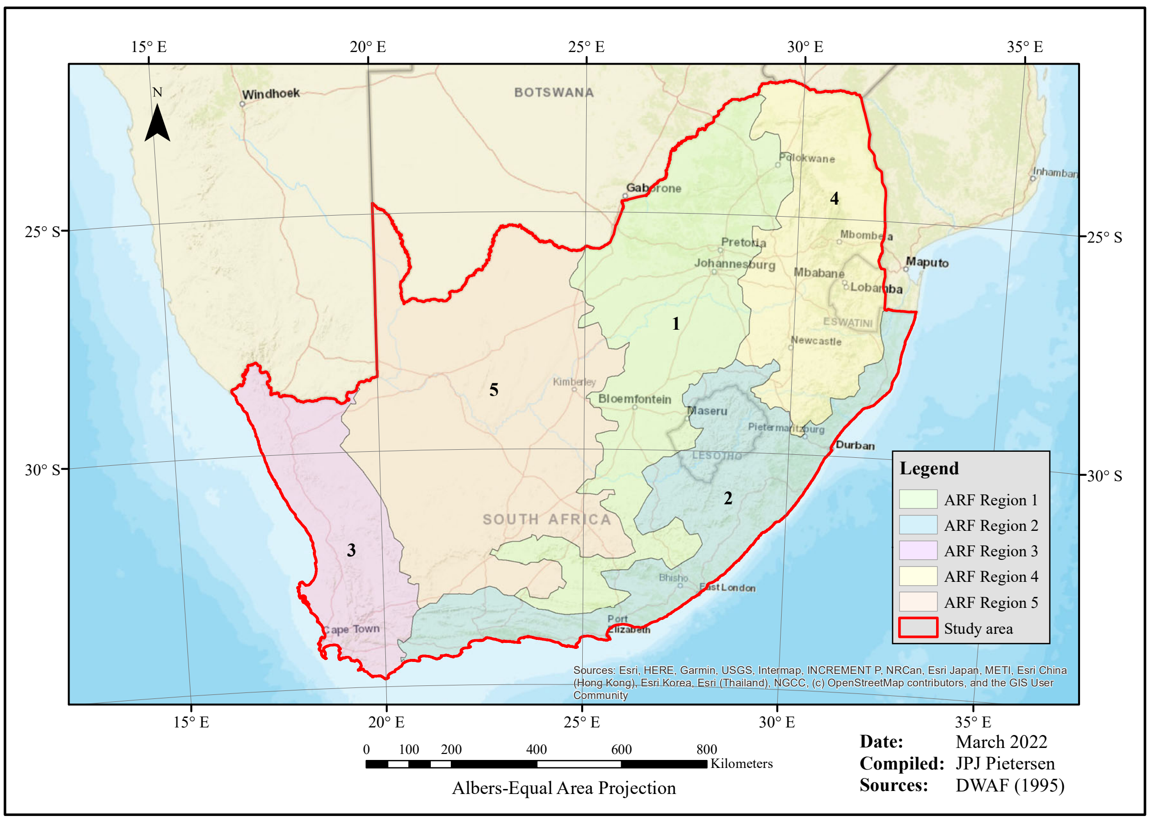

Figure 5.6: Five (5) ARF regions applicable to South African rainfall

By considering the 'user-friendliness' of any derived empirical equation as particularly important, the empirical equation initially derived was in a simplistic, linear log-transformed format with the GOF statistics shown in Table 5.7. However, scatter plots of the ARF_{Sample} [Eq. (5.1)] versus the afore-mentioned linear log-transformed estimates (ARF_y values) revealed several 'outliers' deviating from the 1:1 line, while the ARF_y estimates also tended to overestimate the ARF_{Sample} values for ARF < 80%. The latter overestimation would typically be encountered in larger catchment areas where lower ARF values are to be expected. The latter deficiencies were further investigated to improve the scatter plots between the ARF_{Sample} and ARF_y values.

Subsequently, a non-linear (second-order polynomial) log-transformed empirical ARF equation [Eq. (5.2)], with unique regional calibration coefficients, was derived for each of the five (5) ARF regions using backward stepwise multiple regression analyses with deletion at a 95% confidence level. Therefore, a leave-one-out cross-correlation matrix was implemented to provide a more objective assessment of the predictive ability, while maintaining the 95% (or higher) confidence level. Typically, the MATLAB® software, which is a programming platform used to analyse data and develop algorithms/models, was used to simplify and reduce the equation in its 'raw format' (with 16 constants and/or coefficients) to the current format as expressed in Eq. (5.2). The details pertained to the MATLAB® software running sequence used to derive Eq. (5.2) are included in Appendix C. Partial *t*-tests were used to establish the statistical significance between the regression model and the subset of predictor variables, while total *F*-tests were used to determine whether an ARF (dependent criterion variable) is significantly correlated to the independent predictor variables included in the model. A rejected null hypothesis [*F*-statistic of observed value (F) > critical *F*-statistic ($F\alpha$)] was used to identify the significant contribution of one or more of the independent variables towards the prediction accuracy.

$$ARF_y = aX^2 + bX - c \qquad [5.2]$$

$$X = x_1(\log D)^2 + x_2(\log D) - x_3(\log T)^2 + x_4(\log T) - x_5(\log A)^2 - x_6(\log A) + x_7 \qquad [5.3]$$

where

$ARFy$ = estimated ARF [%],

A = catchment area [km^2],

D = storm duration [days],

T = return period [years],

X = major expression variable,

a to c = major expression constants, and

x_1 to x_7 = regional calibration coefficients [Table 5.8].

Table 5.8: Calibration coefficients associated with the five (5) ARF regions

Region	*a*	*b*	*c*	x_1	x_2	x_3	x_4	x_5	x_6	x_7
1	-0.034	7.286	287.648	-9.415	19.494	1.164	7.666	0.754	1.081	86.067
2	-0.037	7.896	319.770	-9.527	18.229	1.042	6.816	0.629	1.058	88.019
3	-0.055	11.395	487.770	-7.608	15.724	0.330	4.562	0.330	1.216	89.190
4	-0.024	5.391	196.710	-12.363	24.372	0.817	7.660	0.540	2.436	85.056
5	-0.025	5.502	200.890	-11.957	23.453	0.896	7.037	0.953	0.129	84.444

Overall, Eq. (5.2) resulted in improved GOF statistics (*cf.* Table 5.9) when compared to the original GOF statistics based on the linear log-transformed regressions (*cf.* Table 5.7).

Table 5.9: Improved GOF statistics between linear and non-linear equations

GOF		Region				
		1	2	3	4	5
Original	r^2	0.72	0.67	0.65	0.79	0.77
Improved		0.79	0.74	0.75	0.85	0.83
Original	*SE* [%]	3.44	3.40	3.30	3.46	3.25
Improved		3.80	3.11	3.62	4.25	2.90

Note: In applying Eq. (5.2) in Regions 1 to 5, the estimated ARF values could exceed 100%. Hence, all estimated ARF values >100% should be capped at 100%. The latter would typically occur in smaller catchment areas associated with longer storm durations and higher return periods. An ARF value >100% simply implies that the average design point rainfall equals the areal design rainfall for a specific catchment area, storm duration and return period. However, the above capping of ARF values at 100% (if required), is incorporated in the ARF software interface (*cf.* Section 5.8).

Verification requires the use of data sets not being used during the calibration process. During the developmental phases of evaluating the suitability of linear versus non-linear log-transformed regression analyses as discussed above, the concept of calibration and verification was evaluated in each of the 46 clusters. In other words, random combinations of the 46 clusters (with their own unique set of circular catchments and rainfall stations) were used in an alternating fashion for calibration and/or verification purposes until all possible combinations were considered. The process was applied in all clusters. Ultimately, it was noted that whether a dedicated set of clusters or all clusters are assigned to calibration, differences are regarded as insignificant. Hence, all clusters in a particular ARF region were used for the final derivation of Eq. (5.2).

The latter approach was justified by the limited number of rainfall stations available in many of the clusters, which is also affected by the number of rainfall stations/km² criteria to be met in each cluster. Furthermore, it must be noted that 'assessment' of verification results can only apply if the benchmark data set is an observed data set. In the case of ARFs, all data sets, whether used for calibration or verification, remain only estimated sample values.

Scatter plots of the ARF_{Sample} [Eq. (5.1)] and ARF_y [Eq. (5.2)] values associated with all the circular catchments located in ARF Regions 1 to 5 are shown in Figures 5.7 to 5.11 to highlight any differences.

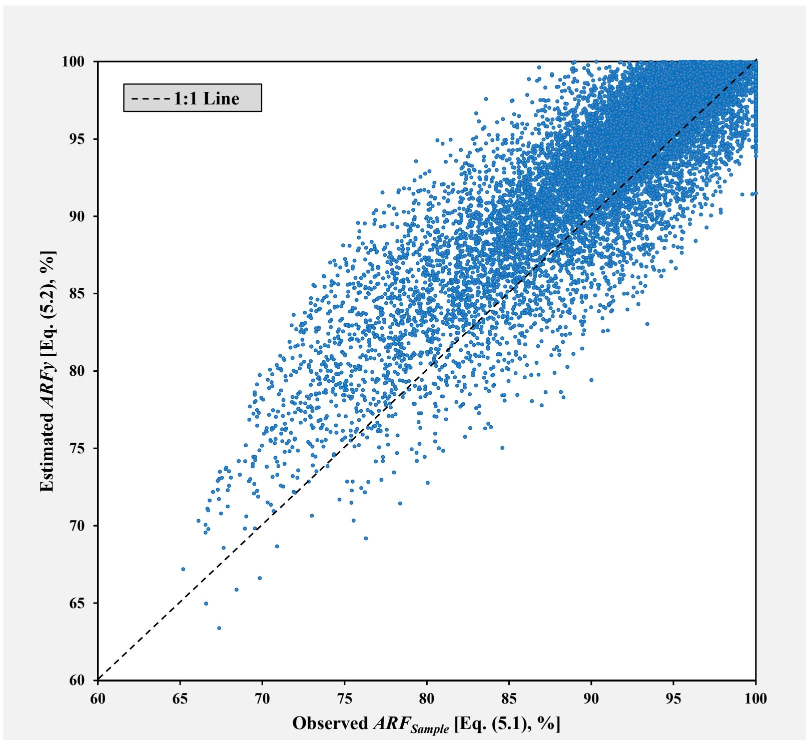

Figure 5.7: Scatter plot of the observed ARF_{Sample} [Eq. (5.1)] and estimated ARF_y [Eq. (5.2)] values in ARF Region 1

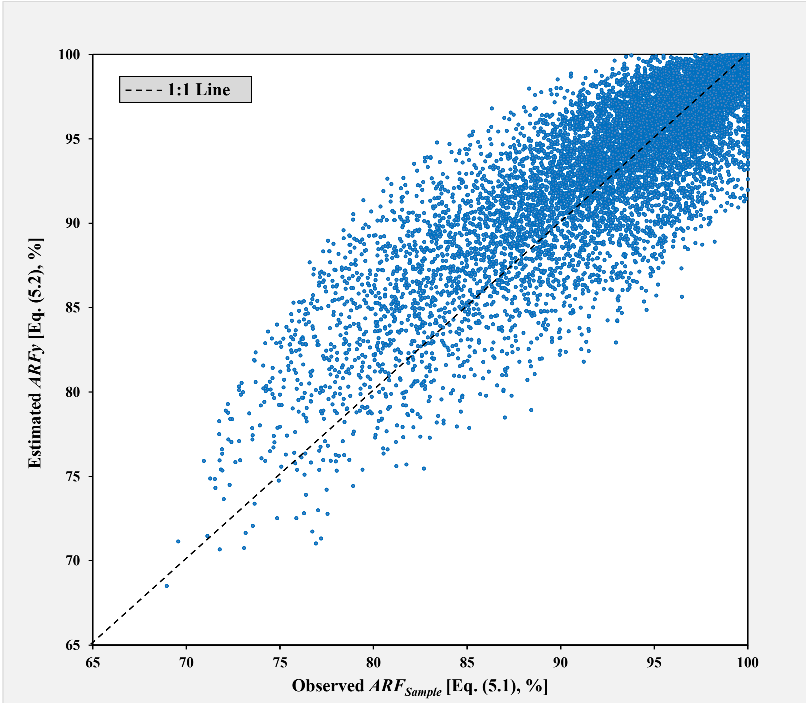

Figure 5.8: Scatter plot of the observed ARF_{Sample} [Eq. (5.1)] and estimated ARF_y [Eq. (5.2)] values in ARF Region 2

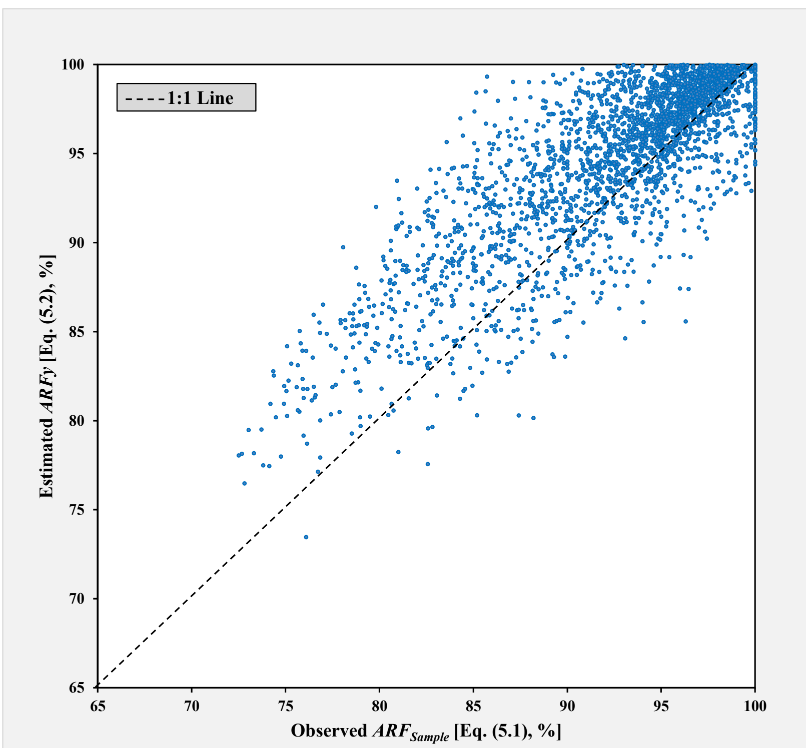

Figure 5.9: Scatter plot of the observed ARF_{Sample} [Eq. (5.1)] and estimated ARF_y [Eq. (5.2)] values in ARF Region 3

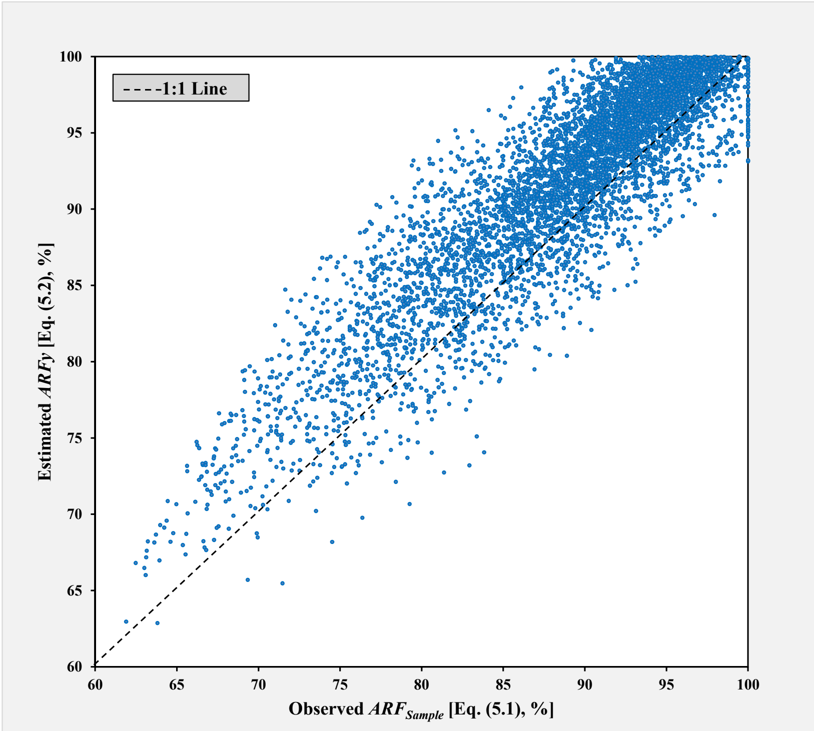

Figure 5.10: Scatter plot of the observed ARF_{Sample} [Eq. (5.1)] and estimated ARF_y [Eq. (5.2)] values in ARF Region 4

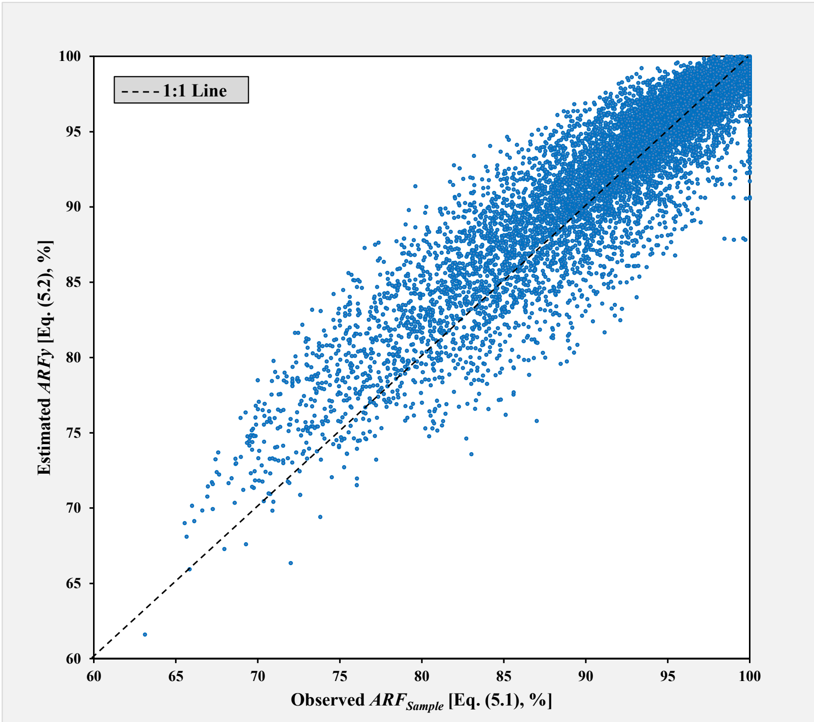

Figure 5.11: Scatter plot of the observed ARF_{Sample} [Eq. (5.1)] and estimated ARF_y [Eq. (5.2)] values in ARF Region 5

In all the ARF regions, Eq. (5.2) tends to overestimate the ARFs associated with larger catchment areas. The latter can be ascribed to: (i) the larger circular catchments automatically having shorter combined rainfall record lengths (≤ 30-years), (ii) the minimum number of rainfall stations per circular catchment, as recommended by Siriwardena and Weinmann (1996), which may not be adequate for South African conditions, and (iii) an overall low rainfall monitoring-network density. Overall, the plotting position and clustering of points in Figures 5.7 to 5.11, are regarded as a significant improvement from the original linear log-transformed equation considered. Furthermore, the distribution of the points in ARF Regions 1 to 5 are also regarded as acceptable, with $0.74 \leq r^2 \leq 0.85$, and $2.9 \leq SE \leq 4.3\%$.

5.7 Comparison of ARF Equations

This section focuses on the comparison of Eq. (5.2) against a selection of geographically-centered ARF estimation methods currently used in local and/or international practice in a range of catchment sizes to establish the consistency and/or possible biases of the newly derived ARF equation. Typically, standard input variables, *e.g.*, catchment area, storm duration, and return period, were used to evaluate the consistency between all the different methods. The storm duration input variable is provided in hours for the selected geographically-centered ARF estimation methods, while Eq. (5.2) requires the storm duration variable in days.

This is the reason for including both durations (days and hours) in Table 5.10 and Figures 5.12 to 5.20. The standard input variables and their associated ranges, *e.g.*, catchment area (10 to 30 000 km^2), storm duration (24, 48 and 72 hours) and return periods (2, 50 and 100 years) were used as input to the various methods, *e.g.*, NERC (1975; Figure 2.4), Alexander I (1980; Figure 2.5), Alexander II (2001; Figure 2.6), Op Ten Noort and Stephenson (OT&S) [1982; Eq. (2.7a)], and Alexander III [2001; Eq. (2.7b)]. All the results are listed in Table 5.10.

Table 5.10: Comparison of geographically-centred ARF estimation methods

Catchment area [km²]	Duration		*T* [years]	Current SA ARF methods [%]					Eq. (5.2) [%]				
	Hours	Days		NERC (Fig. 2.4)	Alexander I (Fig. 2.5)	OT&S [Eq. 2.7a)]	Alexander II (Fig. 2.6)	Alexander III [Eq. 2.7b)]	Region 1	Region 2	Region 3	Region 4	Region 5
			2						88.1	88.9	90.7	87.2	86.7
	24	1	50	97.5	100.0	105.8	100.0	111.8	96.7	96.3	97.8	96.8	94.3
			100						97.6	97.1	98.8	98.1	95.3
			2						94.3	94.5	96.3	94.6	93.2
10	48	2	50	98.2	100.0	104.9	100.0	114.0	100.4	99.6	101.0	101.7	98.6
			100						100.9	100.0	101.5	102.6	99.2
			2						96.4	96.3	98.1	97.2	95.4
	72	3	50	98.8	100.0	104.4	100.0	115.3	101.4	100.5	101.8	103.2	99.9
			100						101.8	100.8	102.1	103.9	100.4
			2						84.9	86.2	88.2	83.4	84.2
	24	1	50	95.5	99.5	99.5	100.0	104.4	94.6	94.5	96.2	94.1	92.6
			100						95.6	95.4	97.4	95.5	93.7
			2						91.8	92.4	94.5	91.6	91.3
50	48	2	50	97.0	100.0	100.4	100.0	106.9	99.0	98.4	100.1	99.8	97.5
			100						99.7	99.0	100.7	100.9	98.2
			2						94.3	94.4	96.6	94.5	93.7
	72	3	50	97.5	100.0	101.0	100.0	108.4	100.4	99.6	101.1	101.7	99.0
			100						100.9	100.0	101.6	102.5	99.6
			2						83.0	84.6	86.9	81.4	82.7
	24	1	50	94.3	97.0	96.8	100.0	101.0	93.3	93.4	95.3	92.6	91.5
			100						94.4	94.4	96.5	94.2	92.6
			2						90.4	91.1	93.5	90.0	90.1
100	48	2	50	96.0	98.0	98.5	100.0	103.6	98.1	97.7	99.5	98.7	96.7
			100						98.9	98.3	100.3	99.9	97.5
			2						93.0	93.4	95.7	93.1	92.7
	72	3	50	96.8	99.0	99.5	100.0	105.1	99.7	99.0	100.7	100.8	98.4
			100						100.3	99.5	101.3	101.7	99.0
			2						77.4	79.9	83.1	75.7	78.0
	24	1	50	90.8	91.5	90.5	92.0	92.3	89.3	90.1	92.7	88.4	88.0
			100						90.7	91.2	94.1	90.2	89.3
			2						85.8	87.4	90.6	85.4	86.4
500	48	2	50	92.8	93.2	94.0	94.5	95.3	95.2	95.3	97.7	95.6	94.2
			100						96.2	96.1	98.7	96.9	95.1
			2						88.9	90.0	93.2	88.9	89.4
	72	3	50	93.8	95.0	96.0	96.5	97.0	97.2	97.0	99.3	98.0	96.2
			100						98.1	97.7	100.1	99.2	97.0
			2						74.3	77.3	81.1	72.8	75.4
	24	1	50	89.0	89.0	87.8	87.5	88.2	87.1	88.2	91.3	86.2	86.0
			100						88.5	89.4	92.8	88.0	87.3
			2						83.3	85.3	89.0	83.0	84.3
1 000	48	2	50	91.5	91.8	92.1	90.5	91.4	93.5	93.9	96.7	93.8	92.6
			100						94.6	94.8	97.8	95.3	93.7
			2						86.6	88.1	91.8	86.7	87.5
	72	3	50	92.3	92.5	94.6	92.5	93.2	95.7	95.8	98.5	96.5	94.9
			100						96.7	96.6	99.4	97.8	95.8
			2						65.2	69.8	75.6	64.8	67.5
	24	1	50	85.0	82.5	81.5	77.0	77.3	80.2	82.5	87.3	79.9	79.7
			100						81.9	84.0	89.0	82.1	81.3
			2						75.7	79.1	84.6	76.3	77.7
5 000	48	2	50	87.0	86.0	87.6	80.5	81.1	88.1	89.5	93.7	88.8	87.7
			100						89.5	90.7	95.1	90.6	89.0
			2						79.6	82.4	87.8	80.5	81.5
	72	3	50	88.2	88.0	91.1	82.5	83.3	91.0	91.9	96.0	92.0	90.6
			100						92.2	92.9	97.1	93.6	91.8

Catchment area [km²]	Duration		*T* [years]	Current SA ARF methods [%]					Eq. (5.2) [%]				
	Hours	Days		NERC (Fig. 2.4)	Alexander I (Fig. 2.5)	OT&S [Eq. 2.7a)]	Alexander II (Fig. 2.6)	Alexander III [Eq. 2.7b)]	Region 1	Region 2	Region 3	Region 4	Region 5
10 000	24	1	2	83.0	80.0	78.8	71.5	71.7	60.4	65.9	72.8	60.8	63.2
			50						76.4	79.5	85.1	76.7	76.2
			100						78.3	81.0	87.0	79.0	78.0
	48	2	2	85.8	82.5	85.6	75.2	76.0	71.6	75.7	82.3	72.8	74.1
			50						85.0	87.0	92.1	86.2	85.0
			100						86.6	88.3	93.6	88.0	86.4
	72	3	2	86.7	85.5	89.6	78.0	78.4	75.8	79.3	85.7	77.3	78.2
			50						88.2	89.7	94.6	89.6	88.1
			100						89.6	90.8	95.8	91.3	89.4
20 000	24	1	2			76.1		65.5	54.9	61.4	69.7	56.4	58.3
			50						72.1	75.9	82.7	73.1	72.2
			100						74.1	77.6	84.8	75.5	74.1
	48	2	2			83.7		70.4	66.9	71.9	79.7	69.0	69.9
			50						81.4	84.1	90.2	83.2	81.7
			100						83.1	85.5	91.8	85.2	83.2
	72	3	2			88.2		73.0	71.4	75.7	83.4	73.7	74.3
			50						84.9	87.0	92.9	86.9	85.1
			100						86.5	88.3	94.3	88.7	86.5
30 000	24	1	2			74.5		61.4	51.4	58.5	67.7	53.6	55.1
			50						69.2	73.6	81.2	70.8	69.6
			100						71.4	75.4	83.3	73.3	71.6
	48	2	2			82.6		66.7	63.8	69.4	78.0	66.6	67.2
			50						79.1	82.2	89.0	81.3	79.5
			100						80.9	83.7	90.7	83.4	81.1
	72	3	2			87.3		69.5	68.6	73.4	81.9	71.5	71.7
			50						82.8	85.3	91.8	85.1	83.1
			100						84.4	86.6	93.3	87.0	84.6

As expected, all the ARF estimates in Table 5.10 decreased with an increase in catchment area. Similarly, the ARF estimates based on Eq. (5.2) increased with an increase in storm duration in all the five (5) ARF Regions. For example, in the areal range (10 to 1 000 km^2), the estimated 50-year return period, 1-day ARFs are between 5.6% (Region 3) and 8.4% (Region 4) smaller than the 3-day ARFs, while in the areal range (5 000 to 30 000 km^2), the estimated 1-day ARFs are between 8.8% (Region 3) and 12.2% (Region 4) smaller than the 3-day ARFs.

The latter differences are ascribed to the fact that rainfall events of a longer duration are also more likely to be evenly distributed over the catchment area under consideration. Overall, the ARF trends estimated using Eq. (5.2), and applicable to certain return periods, were similar to the other methods under consideration in the different catchment area and storm duration ranges. However, only Eq. (5.2) considers the variation of ARFs with return period, and the ARFs increased with an increase in return period.

For example, in the 10 to 100 km^2 areal range, 2-year return period, and 1-day storm duration, the estimated ARFs [Eq. (5.2)] ranged between 80% and 90%, while the ARFs estimated using the current South African methods ranged between 94% to 100%. Given that Eq. (5.2) varies with catchment area, storm duration and return period, it is regarded as a better reflection of the actual rainfall distribution than the other methods. On the other hand, all the ARF estimation methods seem to converge at between 97% and 100% when higher return periods ($T \geq 50$-year) and storm durations ($D \geq 2$-day) are considered.

In Figures 5.12 to 5.20, the percentage differences between the current South African geographically-centred ARF estimation methods and Eq. (5.2), as shown in Table 5.10, are presented. However, only Eq. (5.2) provided results which vary with return period. The ARFs in Region 3 are also slightly higher when compared to the other regions for catchment areas > 100 km^2. Essentially, such higher ARF values require a higher degree of similarity between the average areal design rainfall and average design point rainfall values. However, given that rainfall is highly variable in semi-arid areas such as South Africa (especially in the western parts; Region 3), assuming a more uniform temporal and spatial rainfall distribution (to result in similar areal design rainfall and average design point rainfall values), would be incorrect. Hence, the latter higher ARF values could only be ascribed to the low density of the rainfall-monitoring network.

In general, such higher ARF values (> 100%) are typically associated with higher return periods ($T > 100$ years), and shorter storm durations occurring more frequently in smaller catchment areas.

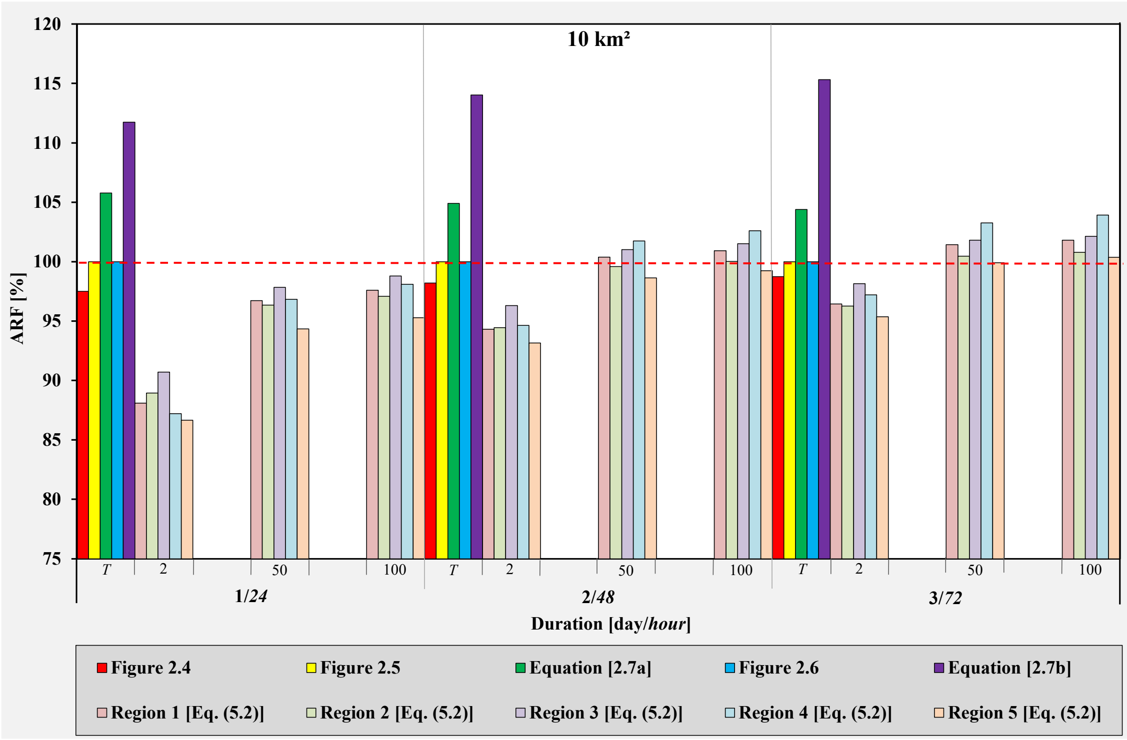

Figure 5.12: Comparison of ARF estimation methods for an area of 10 km², return periods (T) of 2, 50 and 100-year and storm durations of 1/*24*, 2/*48* and 3/*72*-days/*hours*

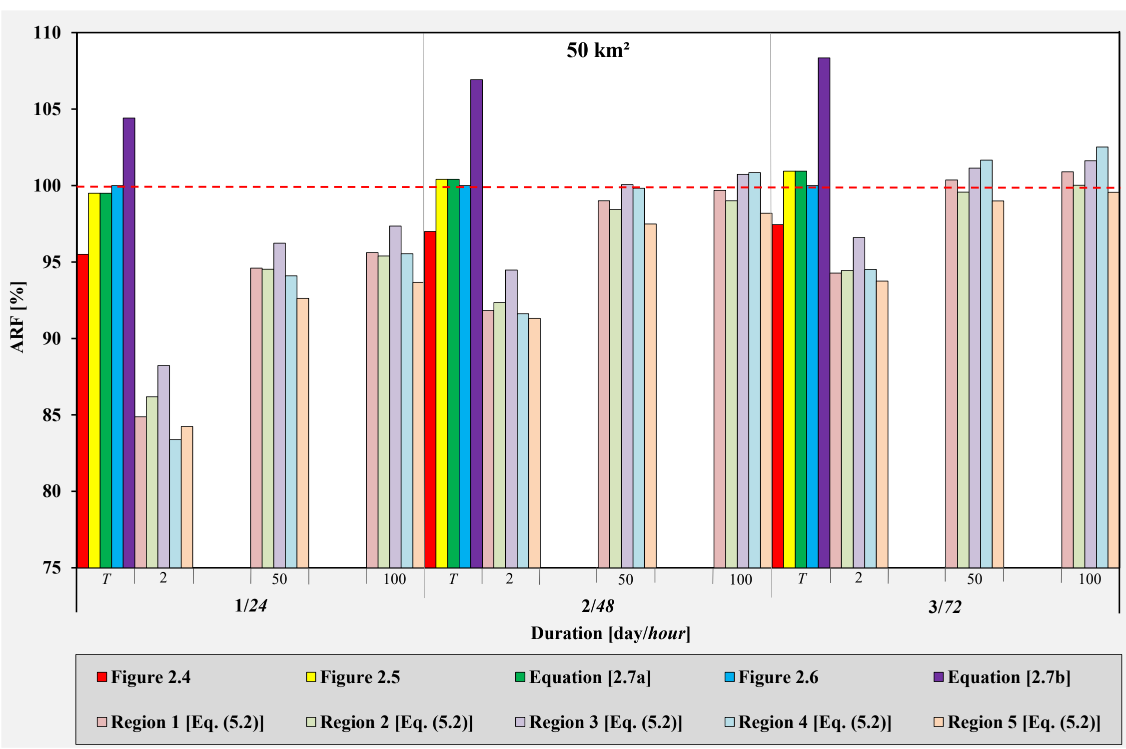

Figure 5.13: Comparison of ARF estimation methods for an area of 50 km², return periods (T) of 2, 50 and 100-year and storm durations of 1/*24*, 2/*48* and 3/*72*-days/*hours*

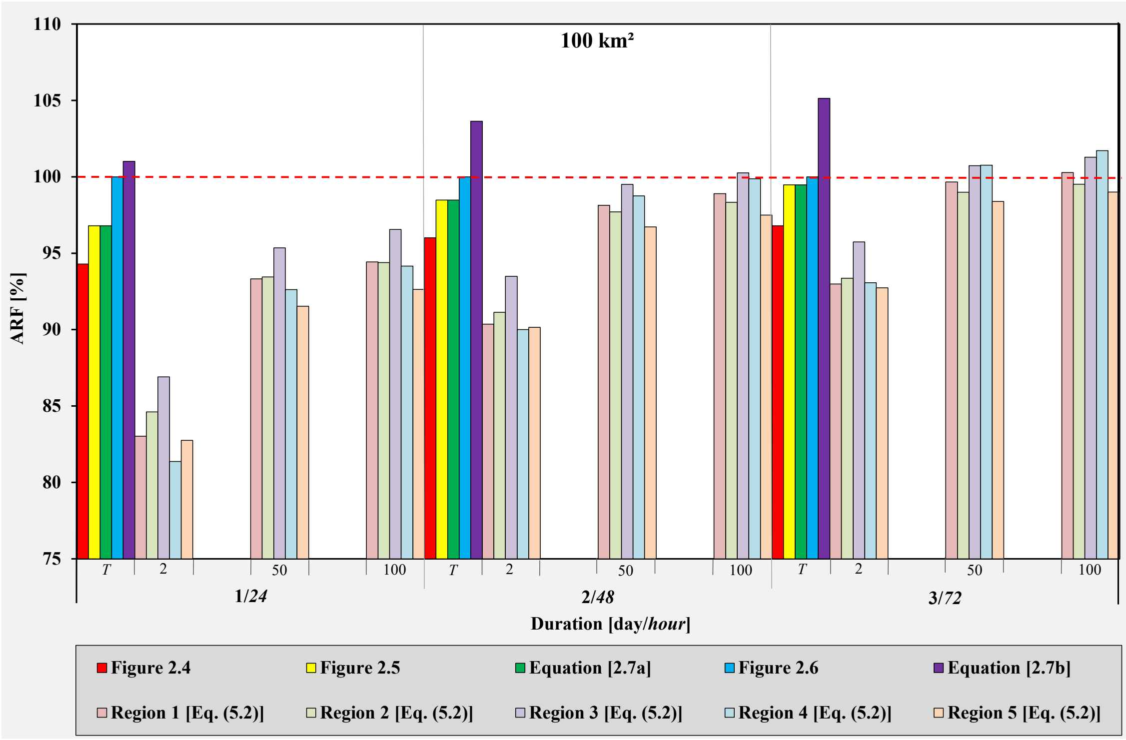

Figure 5.14: Comparison of ARF estimation methods for an area of 100 km², return periods (*T*) of 2, 50 and 100-year and storm durations of 1/*24*, 2/*48* and 3/*72*-days/*hours*

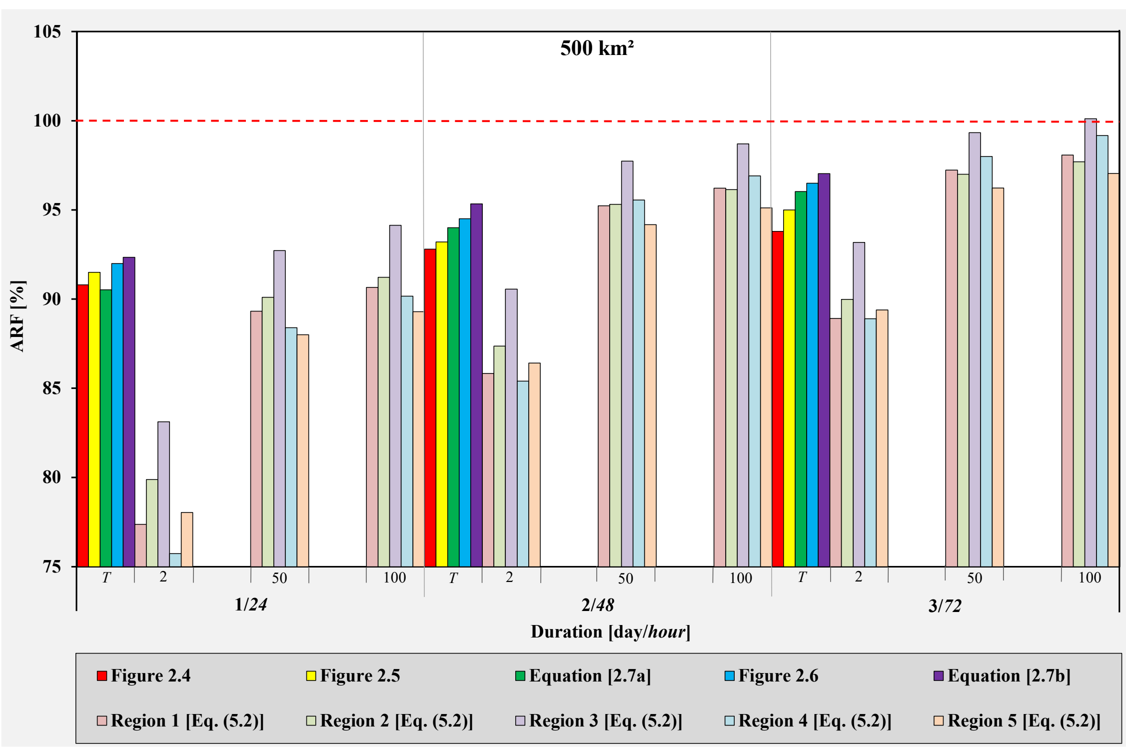

Figure 5.15: Comparison of ARF estimation methods for an area of 500 km², return periods (*T*) of 2, 50 and 100-year and storm durations of 1/*24*, 2/*48* and 3/*72*-days/*hours*

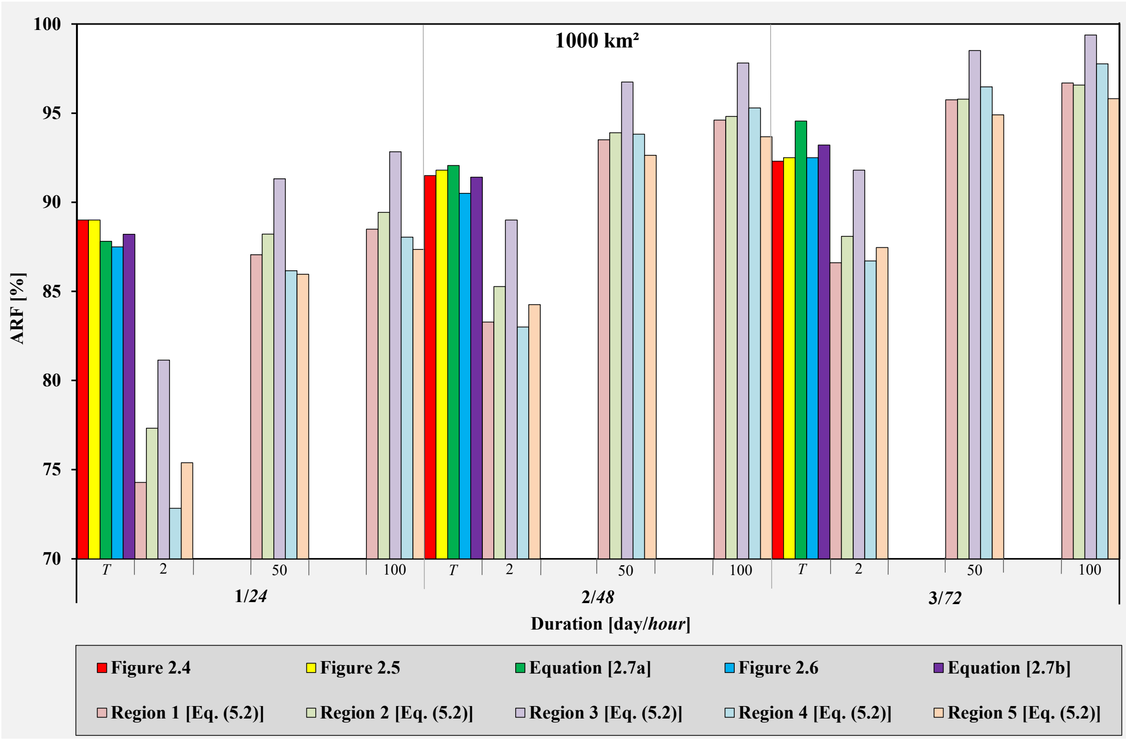

Figure 5.16: Comparison of ARF estimation methods for an area of 1 000 km², return periods (*T*) of 2, 50 and 100-year and storm durations of 1/*24*, 2/*48* and 3/*72*-days/*hours*

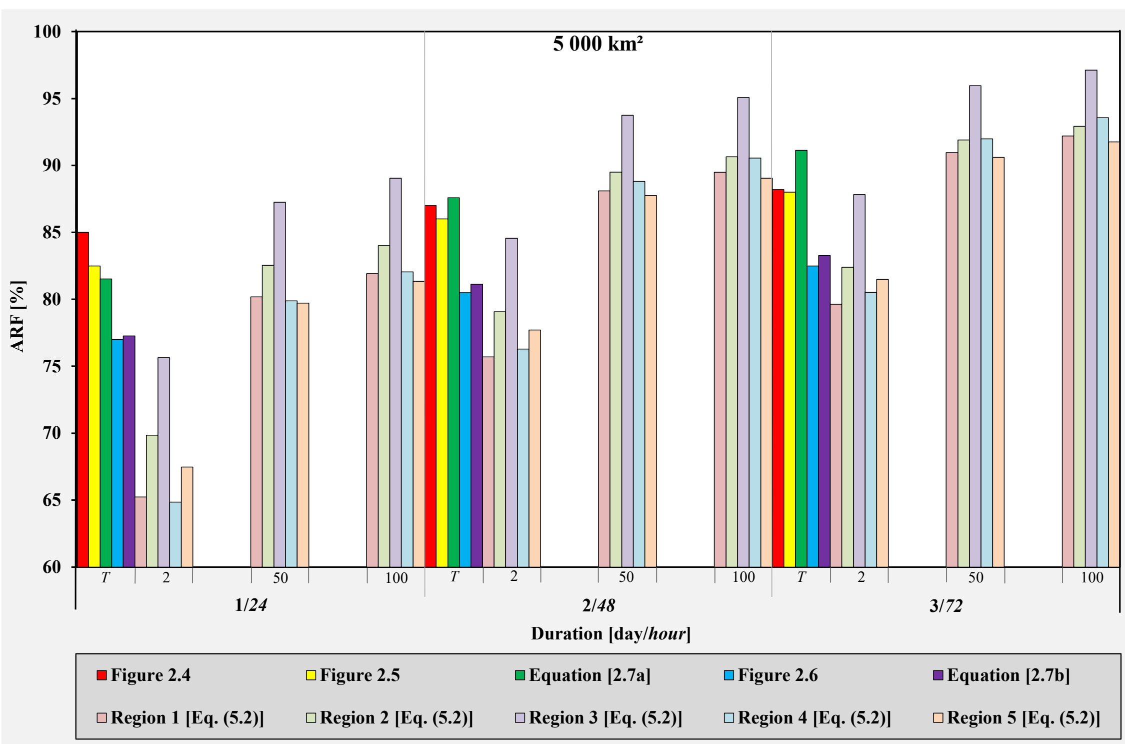

Figure 5.17: Comparison of ARF estimation methods for an area of 5 000 km², return periods (*T*) of 2, 50 and 100-year and storm durations of 1/*24*, 2/*48* and 3/*72*-days/*hours*

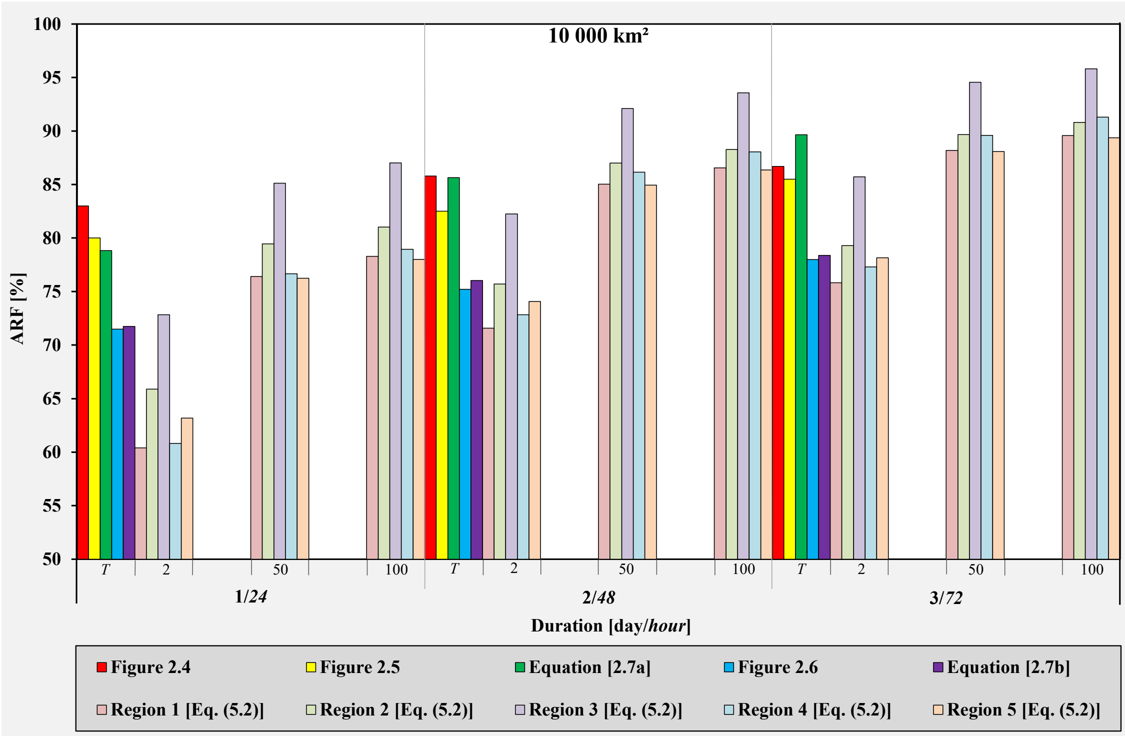

Figure 5.18: Comparison of ARF estimation methods for an area of 10 000 km², return periods (*T*) of 2, 50 and 100-year and storm durations of 1/*24*, 2/*48* and 3/*72*-days/*hours*

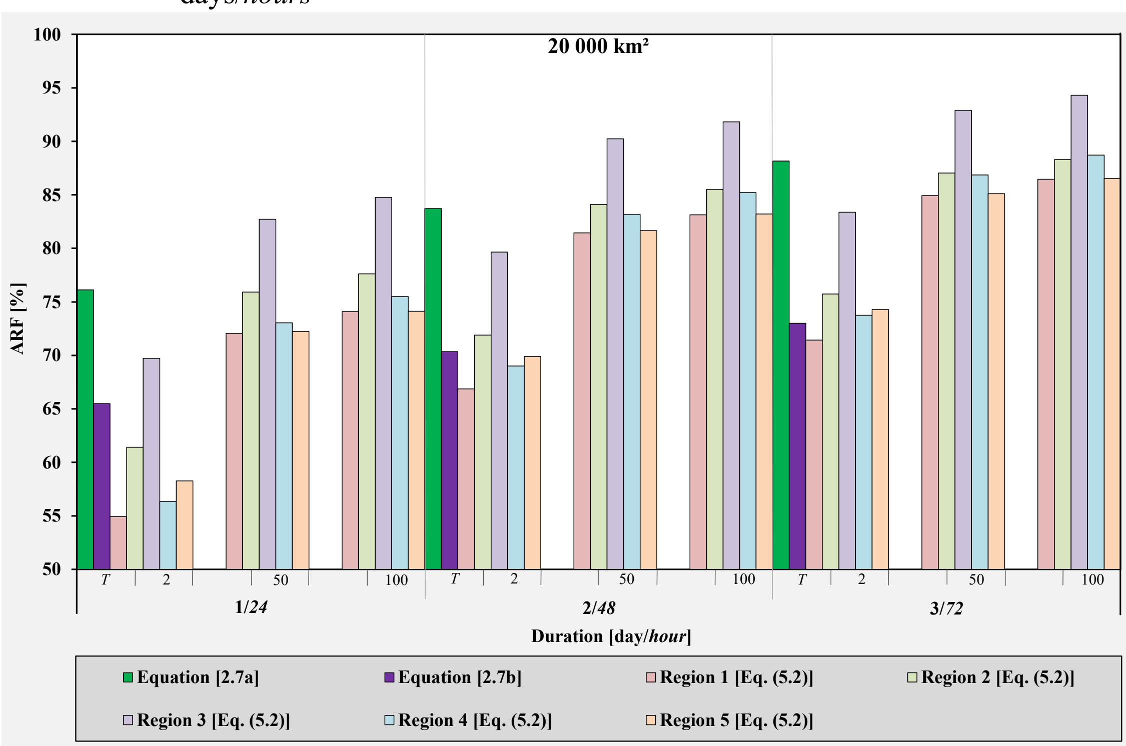

Figure 5.19: Comparison of ARF estimation methods for an area of 20 000 km², return periods (*T*) of 2, 50 and 100-year and storm durations of 1/*24*, 2/*48* and 3/*72*-days/*hours*

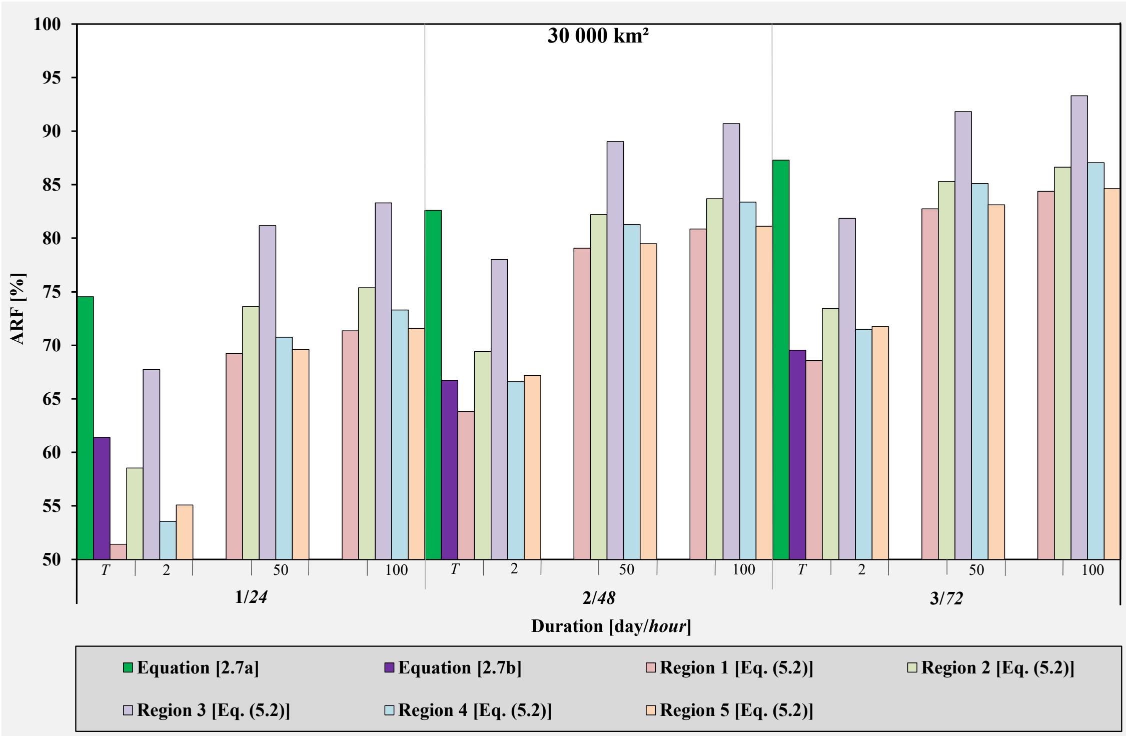

Figure 5.20: Comparison of ARF estimation methods for an area of 30 000 km², return periods (*T*) of 2, 50 and 100-year and storm durations of 1/*24*, 2/*48* and 3/*72*-days/*hours*

The above comparisons shown in Figures 5.12 to 5.20 were further investigated by examining any trends evident in the ARF percentage differences. Given that the current geographically-centred ARF estimation methods used in South Africa provide constant ARF values irrespective of the return period under consideration, the average ARF values estimated using Eq. (5.2) and associated with the 2, 50, and 100-year return periods, were regarded as the 'most representative' ARF values for comparison purposes. As a result, the ARF percentage differences are summarised and discussed below:

(a) **Areal range 10 – 100 km^2 (Figures 5.12 to 5.14 & Appendix D, Table D.1):** Overall, the average ARF percentage differences were generally lower for the 1-day/24-hour storm durations and well spread for the 3-day/72-hour storm durations. Typically, on average, the 1-day/24-hour storm duration ARF percentage differences ranged from 1.6% to 19.4% lower, compared to the 3-day/72-hour storm duration ARF percentage differences ranging from 15.7% lower to 2.7% higher. The NERC method (1975; Figure 2.4) resulted in the lowest average percentage differences in comparison to Eq. (5.2), *i.e.*, 0.1% (Region 4) to 2.5% (Region 5). The Alexander III method [2001;

Eq. (2.7b)] performed the worst with average percentage differences ranging between 6.7% (Region 3) and 17.7% (Region 5).

(b) **Areal range 500 – 1 000 km^2 (Figures 5.15 to 5.16 & Appendix D, Table D.2):** Overall, the average ARF percentage differences were generally well spread for all the storm durations under consideration. Typically, on average, the 1-day/24-hour storm duration ARF percentage differences ranged from 8.8% lower to 0.9% higher, compared to the 3-day/72-hour storm duration ARF percentage differences ranging from 3% lower to 4.5% higher. For catchment areas up to 500 km^2, the NERC method (1975; Figure 2.4) resulted in the lowest average percentage differences in comparison to Eq. (5.2), *i.e.*, 1.1% (Region 2) to 2.4% (Region 5), while the Alexander III method [2001; Eq. (2.7b)] performed the worst, *i.e.*, 0.7% (Region 3) to 5% (Region 5). For catchment areas up to 1 000 km^2, the Alexander II method (2001; Figure 2.6) resulted in the lowest average percentage differences in comparison to Eq. (5.2), *i.e.*, 0.4% (Region 2) to 3.1% (Region 3), while the Op Ten Noort and Stephenson (OT&S) method [1982; Eq. (2.7a)] performed the worst with average percentage differences ranging between 1.7% (Region 3) and 3.4% (Region 5).

(c) **Areal range 5 000 – 10 000 km^2 (Figures 5.17 to 5.18 & Appendix D, Table D.3):** Overall, the average ARF percentage differences were generally well spread for all the storm durations under consideration. Typically, on average, the 1-day/24-hour storm duration ARF percentage differences ranged from 15.2% lower to 12.9% higher, compared to the 3-day/72-hour storm duration ARF percentage differences ranging from 5.2% lower to 16.4% higher. The Alexander I (1980; Figure 2.5) resulted in the lowest average percentage differences in comparison to Eq. (5.2), *i.e.*, 1.2% (Region 2) to 5.6% (Region 3). For catchment areas up to 5 000 km^2, the Alexander III method [2001; Eq. (2.7b)] performed the worst with average percentage differences ranging between 2% (Region 1) and 10.4% (Region 3). For catchment areas up to 10 000 km^2, the Alexander II method (2001; Figure 2.6) performed the worst with average percentage differences ranging between 4.8% (Region 1) and 15.4% (Region 3).

(d) **Areal range 20 000 – 30 000 km^2 (Figures 5.19 to 5.20 & Appendix D, Table D.4):** Overall, the average ARF percentage differences were generally well spread for all the storm durations under consideration. Typically, on average, the 1-day/24-hour storm duration ARF percentage differences ranged from 13.6% lower to 22.6% higher, compared to the 3-day/72-hour storm duration ARF percentage differences ranging from 9.2% lower to 24.4% higher. The Op Ten Noort and Stephenson (OT&S) method [1982; Eq. (2.7a)] resulted in the lowest average percentage differences in comparison to Eq. (5.2), *i.e.*, 2.9% (Region 3) to 10.5% (Region 5), while the Alexander III method [2001; Eq. (2.7b)] performed the worst, *i.e.*, 6.7% (Region 1) to 24% (Region 3).

The implication of the ARF percentage differences highlighted in (a) to (d) above, especially when these methods are applied in practice, needs to be emphasised. According to SANRAL (2013), the Alexander III method [2001; Eq. (2.7b)] is currently the recommended geographically-centred ARF estimation method in South Africa. Hence, practitioners should typically be made aware of any differences between the latter method and Eq. (5.2). In the areal range of 10 to 500 km^2, Eq. (5.2) underestimated the ARF values in comparison to the Alexander III method [Eq. (2.7b)], *i.e.*, underestimations of between 3.3% and 25.3%. In the areal ranges (500 to 10 000 km^2 and 10 000 to 30 000 km^2), both underestimations (17.7% to 19.8%) and overestimations (15.8% to 30.5%) are evident.

In other words, the underestimated ARF values will typically result in lower average areal (catchment) design rainfall values, which ultimately will result in lower peak discharge estimates when event-based design flood estimation methods are used. The opposite is also true, *i.e.*, larger ARF values would result in higher catchment design rainfall and peak discharge estimates. However, given the fact that Alexander's method is not probabilistically correct and not validated using local and up-to-date rainfall data applicable throughout South Africa, it is recommended that Eq. (5.2) should be incorporated as the standard and preferred ARF estimation procedure in South African flood hydrology practice.

5.8 ARF Software Interface

The ARF software interface, as shown in Figure 5.21, was developed to support the estimation of ARFs in South Africa.

It is a web-based application developed through Visual Studio Code (VSC), an Integrated Development Environment (IDE). Microsoft developed VSC IDE for Windows, Linux, and the Macintosh Operating System (MacOS). VSC is a source-code editor compatible with a variety of programming languages, *e.g.*, Java, JavaScript, Go, Nodejs, Python and C++.

The ARF software interface was developed by incorporating a combination of different programme languages, *e.g.*, Hypertext Mark-up Language Version 5 (HTML 5), Cascading Style Sheets Version 3 (CSS 3), Bootstrap 4 and JavaScript.

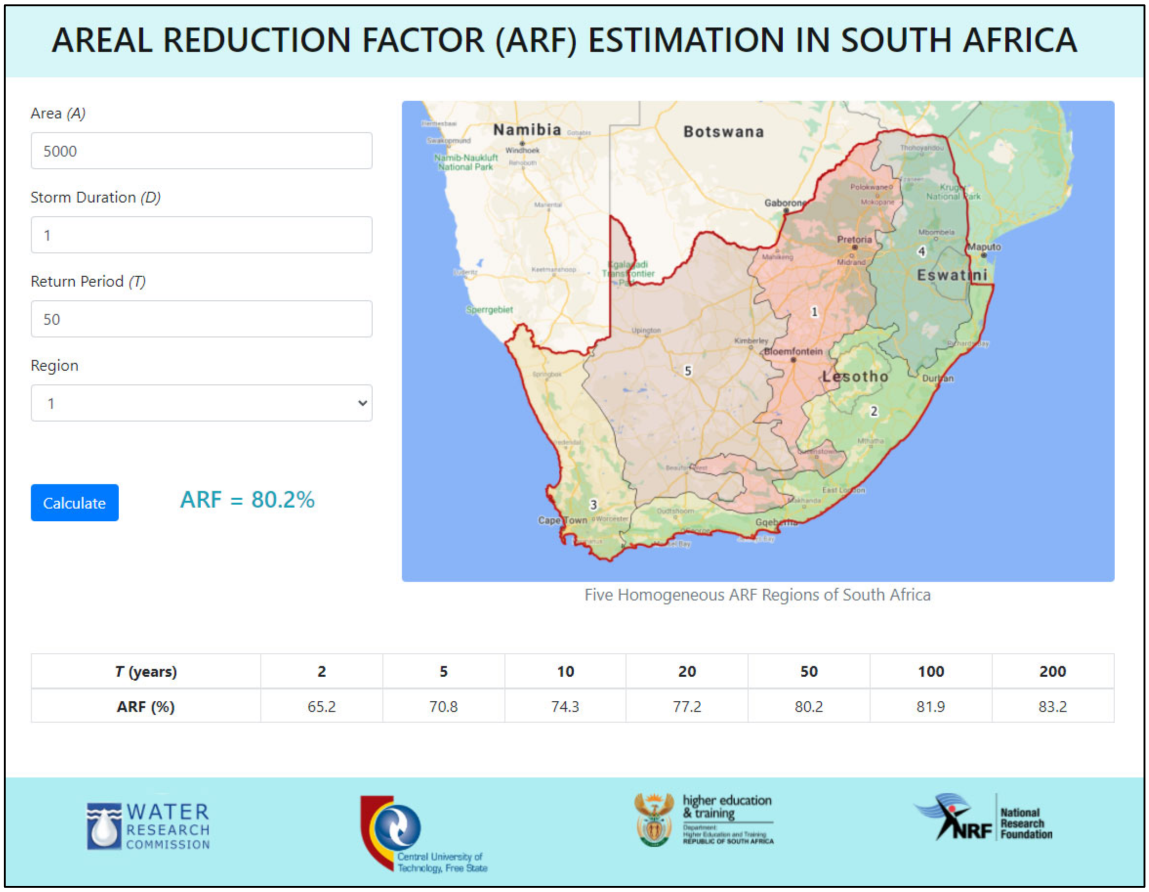

Figure 5.21: ARF software interface

The ARF software script sequences, using HTML 5 and JavaScript, are included in Appendix E. The ARF software (*.zip format), as developed during this research, is also included. The software interface neither requires any external database nor contain any cookies.

Based on the findings obtained from this research, the Alternative Hypothesis (H_1), *i.e.,* ARFs, derived from local South African data, are proportional to or influenced by the RLMA&SI regionalisation scheme, rainfall types, storm durations, and return periods, was accepted due to the sufficient evidence obtained in the different ARF regions, after which, the Null Hypothesis (H_0), was rejected.

Furthermore, the results from this chapter also confirmed the research assumptions applicable to ARFs, (*cf.* Chapter 1, Section 1.3.1): (i) design point rainfall estimates are only representative for a limited area, which was demonstrated by the differences between the average areal design rainfall and average design point rainfall estimates, and (ii) the current geographically-centred South African ARF estimation methods are only applicable to specific temporal and spatial scales, do not account for any regionalisation, provide ARF estimates exceeding 100% in 'smaller' catchments, provide constant ARF values for all return periods, and are based on limited/no local data.

The next chapter presents a synthesis of all the information as discussed in Chapters 4 and 5, as well as some final conclusions and recommendations.

CHAPTER 6: DISCUSSION AND CONCLUSIONS

This chapter contains a discussion and conclusions based on the final methodology and results associated with the derivation of regional ARFs as included in Chapters 4 and 5, respectively.

6.1 ARF Methodology and Results

The research aim was to develop a regionalised approach to estimate ARFs for improved design flood estimation in South Africa. The primary research objective was to develop regionalised, geographically-centred and probabilistically correct ARFs for South Africa by considering the relationship between average areal design rainfall and average design point rainfall estimates. Hence, the research focus was to develop regionalised and probabilistically correct ARFs representative of the different rainfall producing mechanisms in South Africa at a 'circular catchment level' using: (i) daily rainfall data to estimate areal design and design point rainfall, (ii) a modified version of Bell's method (1976), and (iii) the current regionalisation scheme associated with the RLMA&SI approach. Several concerns were documented in Chapter 4; however, solutions were identified and subsequently implemented in Chapter 5.

Merging of the 78 homogeneous rainfall clusters within the seven (7) long duration rainfall regions associated with the RLMA&SI regionalisation scheme was necessary to increase the size of the clusters and the number of rainfall stations within a particular cluster to meet the minimum required number of rainfall stations/km² criteria. Merging only took place between clusters within the same long duration rainfall region, which resulted in 46 delineated clusters. A total of 2 550 artificial circular catchments associated with 1 779 daily rainfall stations with at least 30 years combined areal record lengths, were strategically positioned in the 46 clusters throughout South Africa. Due to the large number of circular catchments placed in each of the 46 clusters, an overlapping of circular catchments was evident. Consequently, this resulted in daily rainfall data from similar rainfall stations being used multiple times within a particular cluster. In principle, this was not regarded as problematic, while it also contributed to the 'smooth' transition between the different clusters.

In applying the selected screening criteria, only 2 053 circular catchments were used in the probabilistic and regression analyses. The probabilistic analyses based on the GEV distribution, fitted to the AMS, using LM, resulted in areal and design point rainfall values for a range of storm durations (*e.g.*, 1, 3, 5 and 7-day), and return periods (*e.g.*, 2, 5, 10, 20, 50 and

100-year). ARFs based on the long duration rainfall were subsequently estimated and expressed as the ratio between the average areal catchment design rainfall and the average design point rainfall estimates for corresponding return periods.

Initially, linear regression analyses were considered in each of the 46 clusters. In each case, the GOF statistics of normal and log-transformed data were evaluated for the independent criterion and dependant predictor variables. The GOF statistics of the log-transformed data outperformed the normal data and resulted in only log-transformed independent predictor variables being used. The 46 clusters were subjected to further merging providing that the clusters are still located within the same RLMA&SI long duration rainfall region and that the GOF statistics of each merged region do not decrease with more than 15% when compared to the original average GOF statistics in the 46 clusters. Subsequently, five (5) ARF regions were deduced from the 46 clusters and Eq. (5.2), a non-linear (second-order polynomial) log-transformed empirical ARF equation with unique regional calibration coefficients, was derived for each region using backward stepwise multiple regression analyses with deletion at a 95% confidence level.

Verification requires the use of data sets not being used during the calibration process. During the developmental phases of evaluating the suitability of linear versus non-linear log-transformed regression analyses, the concept of calibration and verification was evaluated in each of the 46 clusters. In other words, random combinations of the 46 clusters (with their own unique set of circular catchments and rainfall stations) were used in an alternating fashion for calibration and/or verification purposes until all possible combinations were considered. The process was applied in all clusters. Ultimately, it was noted that whether a dedicated set of clusters or all clusters are assigned to calibration, differences are regarded as insignificant. Hence, all clusters in a particular ARF region were used for the final derivation of Eq. (5.2). The latter approach was justified by the limited number of rainfall stations available in many of the clusters, which is also affected by the number of rainfall stations/km^2 criteria to be met in each cluster. Furthermore, it must be noted that 'assessment' of verification results can only apply if the benchmark data set is an observed data set. In the case of ARFs, all data sets, whether used for calibration or verification, remain only estimated sample values.

Subsequently, Eq. (5.2) was compared to a selection of geographically-centred ARF estimation methods currently used in local and/or international practice in a range of catchment sizes to establish the consistency and/or possible biases of the newly derived ARF equation. Equation (5.2) performed similarly, and as expected, when compared to the other methods. The estimated ARFs decreased with an increase in area and increased with an increase in both storm duration and return period.

Based on the research findings in the different ARF regions, it was evident that ARFs, derived from local South African data, are proportional to or influenced by the RLMA&SI regionalisation scheme, rainfall types, storm durations, and return periods. All the research assumptions were also confirmed: (i) design point rainfall estimates are only representative for a limited area, which was demonstrated by the differences between the average areal design rainfall and average design point rainfall estimates, and (ii) the current geographically-centred South African ARF estimation methods are only applicable to specific temporal and spatial scales, do not account for any regionalisation, provide ARF estimates exceeding 100% in 'smaller' catchments, provide constant ARF values for all return periods, and are based on limited/no local data.

Given the rejection of the Null Hypothesis and the subsequent confirmation of the research assumptions, the primary research objective, *i.e.*, the development of a regionalised and probabilistic approach to estimate ARFs using an empirical equation [Eq. (5.2)] for South Africa, was also achieved. This new ARF methodology clearly elucidates how ARFs, within the five (5) ARF regions of South Africa, vary with catchment area, storm duration, and return period.

6.2 Contribution to New Knowledge

The ARF methodology developed in this research and the subsequent findings are new to the South African flood hydrology research community and practice: (i) ARFs were derived and are based on a regionalisation scheme utilising the daily rainfall data available in the DREU database, (ii) ARFs are probabilistically correct, *i.e.*, vary with return period, and (iii) a web-based software application was developed to enable the consistent estimation of ARFs within the five (5) ARF regions of South Africa.

6.3 Conclusions and Recommendations

Several international ARF studies, as included in Chapter 2, demonstrated that empirically derived ARFs vary with, but are not limited to, catchment area, storm duration, and return period. The latter factors are primarily influenced by predominant rainfall types, seasonal/regional factors, the density of rainfall-monitoring networks, and both the quantity and quality of rainfall data. Internationally, ARF research and the refinement of existing methodologies are ongoing to better accommodate all the factors having an influence on the estimation of ARFs. For example, Podger *et al.* (2015a; 2015b) revised the previous version of the Australian ARF methodology as proposed by Siriwardena and Weinmann (1996), to a more comprehensive approach as included in the current Australian Rainfall and Runoff (ARR) Manual (Ball *et al.*, 2019). In contradiction, the UK is still using the ARF diagrams as published in the 1975 UK FSR (Faulkner, 1999), which provide constant ARF values for all return periods; hence, probabilistically incorrect ARFs. To date, the situation in South Africa is very similar to the UK; however, given that none of the current geographically-centred ARF methods used in South Africa are based on local rainfall data, the situation is even worse. Therefore, the regionalised ARF methodology developed and adopted in this research, with the aid of Eq. (5.2), overcame the limitations highlighted above by enabling the estimation of geographically-centred and probabilistically correct ARFs representative of the different rainfall producing mechanisms in South Africa. Thus, the ARFs estimated with Eq. (5.2) do not only vary with return period, but the approach is also compatible with the current and recommended RLMA&SI and/or TR102 design rainfall databases available in South Africa, given that both these databases contain design point rainfall information which need to be converted to catchment design rainfall by applying an ARF.

It is important to note that the ARFs derived in this research are regarded as long duration ARFs and apply to storm durations exceeding 1-day. However, Eq. (5.2) could also be applied to storm durations less than 1-day given that plausible short duration ARF estimates were obtained during the testing of the web-based ARF software. Hence, until the ARF methodology is further refined to accommodate short durations using appropriate sub-daily rainfall data sets, the general application of Eq. (5.2) in South Africa is recommended. Furthermore, for the interim, short duration ARFs should not be estimated using scaling and/or conversion factors to disaggregate daily rainfall values to sub-daily rainfall values (*cf.* Chapter 2, Section 2.5.2). In doing so, the ratio of design point and areal design rainfall values (short duration ARFs) would render values equal to the 1-day ARFs. It is recommended that short duration ARFs for South Africa should be developed by either using an updated sub-daily (continuous) rainfall database, or by developing an improved approach to disaggregate daily rainfall data into sub-daily rainfall values. Given the importance of climate change and the potential effect thereof on rainfall patterns, it would also be worthwhile to consider different rainfall data periods for the same rainfall stations, *e.g.*, 1900 to 1950 vs. 1950 to 2000, to establish whether any significant differences exist in the AMS data sets to ultimately result in different ARF estimates. Any notable differences will not only highlight that sample ARFs are dependent on the period of record, but it would also underpin that the current South African ARF estimation methods (based on using outdated data), should be used with caution or updated regularly. Ultimately, the above would confirm that an ARF is a dynamic factor which needs to be updated regularly as changes in rainfall patterns and processes become more prevalent, while the non-stationary of rainfall data cannot be ignored. However, given that an ARF is based on the ratio between areal and design point rainfall, it is envisaged that the latter influences will only be evident on the long term.

The ARFs derived in this research are based on a specific methodology, as explained in Chapter 5; however, deviating from this methodology may result in completely different ARF estimates. The methodology adopted was carefully formulated to eliminate potential problems, as outlined in Chapter 4, and to incorporate all the data sets and resources available. The methodology, in its current form without using any local, country-specific correction factors, may not necessarily be suitable when applied in other countries. Furthermore, it must be noted that Eq. (5.2) was derived from estimated ARF values, *i.e.*, the ratio between average areal and design point rainfall values. Thus, the fact that 'estimated' as opposed to 'observed' values were used, highlights the inherent uncertainty involved when Eq. (5.2) is used to convert design point rainfall to average areal (catchment) design rainfall representative of the spatial and temporal rainfall distribution in a catchment.

Average areal design rainfall is regarded as fundamental input to all event-based design flood estimation methods used in South Africa. Given the sensitivity of design peak discharges to estimated ARFs, it is envisaged that the implementation of the specific objectives of this research will contribute fundamentally to the improved estimation of ARFs to ultimately result in improved design flood estimations in South Africa. The methodology developed, could also be tested, and subsequently adopted internationally to improve the estimation of ARFs to provide more reliable peak discharge and volume estimates as, to date, this remains a constant challenge in flood hydrology.

CHAPTER 7: REFERENCES

Adamson, PT. 1981. *Southern African Storm Rainfall*. Technical Report TR102. Department of Environmental Affairs and Tourism. Pretoria, RSA.

Alexander, WJR. 1978. *Depth-Area-Duration-Frequency Properties of Storm Rainfall in South Africa*. Technical Report TR83. Department of Water Affairs, Pretoria, RSA.

Alexander, WJR. 1980. *Depth-Area-Duration-Frequency Properties of Storm Rainfall in South Africa*. Technical Report TR103. Department of Water Affairs, Pretoria, RSA.

Alexander, WJR. 2001. *Flood Risk Reduction Measures: Incorporating Flood Hydrology for Southern Africa*. Department of Civil and Biosystems Engineering, University of Pretoria, Pretoria, RSA.

Alexander, WJR. 2010. *Analytical Methods for Water Resource Development and Management: Handbook for Practitioners and Decision Makers*, [online]. Department of Civil and Biosystems Engineering, University of Pretoria, Pretoria, RSA. Available from: http://www.droughtsandfloods.com [5 February 2013].

Allen, RJ and DeGaetano, AT. 2005. Areal reduction factors for two eastern United States regions with high rain-gauge density. *Journal of Hydrologic Engineering* 10 (4): 327–335.

Asquith, WH and Famiglietti, JS. 2000. Rainfall areal-reduction factor estimation using annual-maxima centred approach. *Journal of Hydrology* 230 (1, 2): 55–69.

Bacchi, B and Ranzi, R. 1996. On the derivation of the areal reduction factor of storms. *Atmospheric Research* 42: 123–135.

Ball, JE, Babister, MK, Nathan, R, Weinmann, PE, Weeks, W, Retallick, M and Testoni, I. 2019. *Australian Rainfall and Runoff: A Guide to Flood Estimation*. Institution of Engineers, Australia.

Bárdossy, A and Pegram, G. 2018. Intensity–duration–frequency curves exploiting neighbouring extreme precipitation data. *Hydrological Sciences Journal* 63 (11): 1593–1604.

Bell, FC. 1976. *The Areal Reduction Factor in Rainfall Frequency Estimation*. Institute of Hydrology, Report No. 35. Natural Environment Research Council, UK.

Bengtsson, L and Niemczynowicz, J. 1986. Areal reduction factors from rain movement. *Nordic Hydrology* 17 (2): 65–82.

Cunnane, C. 1989. *Statistical Distributions for Flood Frequency Analysis*. WMO Report 718. World Meteorological Organisation, Geneva, Switzerland.

Davies, B and Day, J. 1998. *Vanishing Waters.* 1st ed. University of Cape Town Press, Cape Town, RSA.

De Michéle, C, Kottegoda, NT and Rosso, R. 2001. The derivation of areal reduction factor of storm rainfall from its scaling properties. *Water Resources Research* 37 (12): 3247–3252.

Dent, MC, Lynch, SD, and Schulze, RE. 1989. *Mapping Mean Annual and other Rainfall Statistics over Southern Africa*. University of Natal, Department of Agricultural Engineering, ACRU Report 27, WRC Report No. 109/89. Water Research Commission, Pretoria, RSA.

Du Plessis, JA, Burger, RP, Gericke, OJ, Todd, TJ and Marebane, TS. 2020. *The Applicability of the use of Radar Data to Develop Areal Reduction Factors in South Africa.* WRC Report No. K5-2923. Water Research Commission, Pretoria, RSA.

Du Plessis, JA, and Loots, W. 2019. Re-evaluation of area reduction factors and their impact on floods in South Africa. *Arabian Journal of Geosciences* 12 (11): 1–16.

Dyrrdal, AV, Skaugen, T, Stordal, F and Førland, EJ. 2016. Estimating extreme areal precipitation in Norway from a gridded dataset. *Hydrological Sciences Journal* 61 (3): 483–494.

Dyson, LL. 2009. Heavy daily-rainfall characteristics over the Gauteng Province. *Water SA* 35 (5): 627–638.

DWAF. 1995. *GIS Data: Drainage Regions of South Africa.* Department of Water Affairs and Forestry, Pretoria, RSA.

Eggert, B, Berg, P, Haerter, JO, Jacob, D and Moseley, C. 2015. Temporal and spatial scaling impacts on extreme precipitation. *Atmospheric Chemistry and Physics* 15 (10): 5957–5971.

ESRI. 2006. *ArcGIS Desktop Help: Geoprocessing.* Environmental Systems Research Institute, Redlands, CA, USA.

Faulkner, D., 1999. *Flood Estimation Handbook, Volume 2: Rainfall Frequency Estimation*. Wallingford, Institute of Hydrology, UK.

Field, AP. 2000. *Discovering Statistics using SPSS for Windows: Advanced Techniques for Beginners.* SAGE Publications, ISBN 0761957553, 9780761957553.

Gericke, OJ and Du Plessis, JA. 2011. Evaluation of critical storm duration rainfall estimates used in flood hydrology in South Africa. *Water SA* 37 (4): 453–470. DOI: 10.4314/wsa.v37i4.4

Grebner, G and Roesch, T. 1997. Regional dependence and application of DAD relationships. *International Association of Hydrological Sciences* 246: 223–230.

Haarhoff, J and Cassa, AM. 2009. *Introduction to Flood Hydrology*. Juta and Company Limited, Cape Town, RSA.

Hershfield, DM. 1962. Extreme rainfall relationships. In: *Proceedings*, *American Society of Civil Engineers* HY6 (11):73–92.

Hogg, WD. 1992. *Inhomogeneities in Time Series of Extreme Rainfall*. 5th International Meeting on Statistical Climatology. The Steering Committee for International Meetings on Statistical Climatology, Toronto, Canada 481–484.

Hosking, JRM and Wallis, JR. 1993. Some statistics useful in a regional frequency analysis. *Water Resources Research* 29 (2): 271–281.

Hosking, JRM and Wallis, JR. 1997. *Regional Frequency Analysis: An Approach Based on L-Moments.* Cambridge University Press, Cambridge, UK. 254 pp.

Huff, FA. 1995. Characteristics and contributing causes of an abnormal frequency of flood-producing rainstorms at Chicago. *Journal of the American Water Resources Association* 31 (4): 703–714.

Huff, FA and Shipp, WL. 1969. Spatial correlations of storm, monthly, and seasonal rainfall. *Journal of Applied Meteorology* 8 (4): 542–550.

Kim, J, Lee, J, Kim, D and Kang, B. 2019. The role of rainfall spatial variability in estimating areal reduction factors. *Journal of Hydrology* 568: 416–426.

Lambourne, JJ and Stephenson, D. 1986. *Research in Urban Hydrology and Drainage: Factors affecting Storm Runoff in South Africa*. Water Systems Research Programme Report No. 2/86. University of Witwatersrand. Johannesburg, RSA.

Li, J, Sharma, A, Johnson, F and Evans, J. 2015. Evaluating the effect of climate change on areal reduction factors using regional climate model projections. *Journal of Hydrology* 528: 419–434.

Lombardo, F, Napolitano, F and Russo, F. 2006. On the use of radar reflectivity for estimation of the areal reduction factor. *Natural Hazards and Earth System Sciences* 6 (3): 377–386.

Lynch, SD. 2004. *Development of a Raster Database of Annual, Monthly and Daily Rainfall for Southern Africa.* WRC Report No. 1156/1/04. Water Research Commission, Pretoria, RSA.

Makhuva, T, Pegram, GGS, Sparks, R and Zucchini, W. 1997a. Patching rainfall data using regression methods. 1. Best subset selection, EM and pseudo-EM methods: Theory. *Journal of Hydrology* 198: 289–307.

Makhuva, T, Pegram, GGS, Sparks, R and Zucchini, W. 1997b. Patching rainfall data using regression methods. 2. Comparisons of accuracy, bias, and efficiency. *Journal of Hydrology* 198: 308–318.

Mambo, J and Faccer, K. 2017. *South African Risk and Vulnerability Atlas: Understanding the Social and Environmental Implications of Global Change*. African SUN Media, Pretoria, RSA.

Meier, CI, Moraga, JS, Pranzini, G and Molnar, P. 2015. Effects of record length and resolution on the derived distribution of annual precipitation. *Hydrology and Earth System Sciences Discussions* 12:12987–13018.

Meier, KB. 1997. Development of a Spatial Database for Agrohydrological Model Applications in Southern Africa. Unpublished MSc Eng. dissertation, Department of Agricultural Engineering, University of Natal, Pietermaritzburg, RSA.

Midgley, DC and Pitman, WV. 1978. *A Depth-Duration-Frequency Diagram for Point Rainfall in Southern Africa*. HRU Report 2/78. University of Witwatersrand, Johannesburg, RSA.

Midgley, DC, Pitman, WV and Middleton, BJ. 1994. *Surface Water Resources of South Africa*. WRC Report No. 298/2/94. Water Research Commission, Pretoria, RSA.

Milewski, AV. 1981. A comparison of vegetation height in relation to the effectiveness of rainfall in the Mediterranean and adjacent arid parts of Australia and South Africa. *Journal of Biogeography* 8 (2): 107–16.

Mineo, C, Ridolfi, E, Napolitano, F and Russo, F. 2018. The areal reduction factor: A new analytical expression for the Lazio Region in central Italy. *Journal of Hydrology* 560: 471–479.

NERC. 1975. *Flood Studies Report*. Natural Environment Research Council, London, UK.

Omolayo, AS. 1993. On the transposition of areal reduction factors for rainfall frequency estimation. *Journal of Hydrology* 145 (1–2): 191–205.

Op ten Noort, TH and Stephenson, D. 1982. *Flood Peak Calculation in South Africa.* Water System Research Programme Report No. 2/1982. University of the Witwatersrand, Johannesburg, RSA.

Patra, KC. 2008. *Hydrology and Water Resources Engineering*. 2nd Ed. Narosa Publishing House Limited. New Delhi, India.

Pavlovic, S, Perica, S, St Laurent, M and Mejia, A. 2016. Intercomparison of selected fixed-area areal reduction factor methods. *Journal of Hydrology* 537: 419–430.

Peleg, N, Marra, F, Fatichi, S, Paschalis, A, Molnar, P and Burlando, P. 2018. Spatial variability of extreme rainfall at radar subpixel scale. *Journal of Hydrology* 556: 922–933.

Pegram, GGS. 1997. Patching rainfall data using regression methods. 3. Grouping, patching, and outlier detection. *Journal of Hydrology* 198: 319–334.

Pegram, GGS. 1997. *Patching Rainfall Data*. Report H 6/6/0194. Department of Water Affairs and Forestry, Directorate of Project Planning, Pretoria, RSA.

Pegram, GGS and Parak, M. 2004. A review of the Regional Maximum Flood and Rational formula using geomorphological information and observed floods. *Water SA* 30 (3): 377–392.

Pegram, GGS and Adamson, PT. 1988. Revised risk analysis for extreme storms and floods in Natal/KwaZulu. *The Civil Engineer in South Africa*: 331–336.

Pietersen, JPJ. 2016. Areal Reduction Factors for Design Rainfall Estimation in the C5 Secondary Drainage Region of South Africa. Unpublished M Tech Eng. dissertation, Department of Civil Engineering, Central University of Technology, Free State. Bloemfontein, RSA.

Pietersen, JPJ, Gericke, OJ, Smithers, JC and Woyessa, YE. 2015. Review of current methods for estimating areal reduction factors applied to South African design point rainfall and preliminary identification of new methods. *Journal of the South African Institution of Civil Engineering* 57 (1): 16–30.

Pitman, WV. 2011. Overview of water resource assessment in South Africa: Current state and future challenges. *Water SA* 37 (5): 659–664. DOI: 10.4314/wsa.v37i5.3

Podger, S, Green, J, Jolly, C and Beesley, C. 2015a. Creating long duration areal reduction factors. In: *Proceedings, 36th Hydrology and Water Resources Symposium: The Art and Science of Water.* Institution of Engineers, Australia. p. 39.

Podger, S, Green, J, Steynsmyr, P and Babister, M. 2015b. Combining long and short duration areal reduction factors. In: *Proceedings, 36th Hydrology and Water Resources Symposium: The Art and Science of Water.* Institution of Engineers, Australia. p. 210.

Rodriguez-Iturbe, I and Mejía, JM. 1974. On the transformation of point rainfall to areal rainfall. *Water Resources Research* 10 (4): 729–735.

SANRAL. 2013. *Drainage Manual*. 6th Ed. South African National Roads Agency Limited, Pretoria, RSA.

Schulze, RE. 1980. *Potential Flood Producing Rainfall for Medium and Long Duration in Southern Africa*. Report to Water Research Commission, Pretoria, RSA.

Schulze, RE. 1984. *Depth-duration-frequency Studies in Natal based on Digitised Data*. South African National Hydrology Symposium, Technical Report TR119. Department of Environment Affairs, Pretoria, RSA.

Schulze, RE. 1995. *Hydrology and Agrohydrology: A Text to Accompany the ACRU 3.00 Agrohydrological Modelling System.* WRC Report No. TT 69/95. Water Research Commission, Pretoria, RSA.

Schulze, RE, Schmidt, EJ, and Smithers, JC. 1992. *SCS-SA User Manual: PC-Based SCS Design Flood Estimates for Small Catchments in Southern Africa*. ACRU Report No. 40. Department of Agricultural Engineering, University of Natal, Pietermaritzburg, RSA.

Singh, SK, Griffiths, GA and McKerchar, AI. 2018. Towards estimating areal reduction factors for design rainfalls in New Zealand. *Journal of Hydrology* (New Zealand) 57 (1): 25–33.

Siriwardena, L and Weinmann, PE. 1996. *Derivation of Areal Reduction Factors for Design Rainfalls in Victoria*. Report No. 96/4. Cooperative Research Centre for Catchment Hydrology, Victoria, Australia.

Sivapalan, M and Blöschl, G. 1998. Transformation of point rainfall to areal rainfall: Intensity-duration-frequency curves. *Journal of Hydrology* 204 (1–4): 150–167.

Skaugen, T. 1997. Classification of rainfall into small- and large-scale events by statistical pattern recognition. *Journal of Hydrology* 200 (1–4): 40–57.

Smithers, JC. 1996. Short-duration rainfall frequency model selection in Southern Africa. *Water SA* 22 (3): 211–217.

Smithers, JC, Görgens, AHM, Gericke, OJ, Jonker, V, and Roberts P. 2014. *The Initiation of a National Flood Studies Programme for South Africa.* SANCOLD, Pretoria, RSA.

Smithers, JC and Schulze, RE. 2000a. *Development and Evaluation of Techniques for Estimating Short Duration Design Rainfall in South Africa*. WRC Report No. 681/1/00. Water Research Commission, Pretoria, RSA.

Smithers, JC and Schulze, RE. 2000b. *Long Duration Design Rainfall Estimates for South Africa.* WRC Report No. 811/1/00. Water Research Commission, Pretoria, RSA.

Smithers, JC and Schulze, RE. 2003. *Design Rainfall and Flood Estimation in South Africa*. WRC Report 1060/01/03. Water Research Commission, Pretoria, RSA.

Smithers, JC and Schulze, RE. 2004. The estimation of design rainfall for South Africa using a regional scale invariant approach. *Water SA* 30 (4): 435–444.

Statistics South Africa (SSA). 2004. *Land accounts: Land-use and Land-cover for South Africa, 1994/1995.* Pretoria, RSA.

Stewart, EJ. 1989. Areal reduction factors for design storm construction: Joint use of rain gauge and radar data. *Association of Hydrological Sciences* 181: 31–40.

Sugawara, M. 1992. On the weights of precipitation stations. *European Geophysical Society Series on Hydrological Sciences, Advances in Theoretical Hydrology*, Elsevier 59–74.

Svensson, C and Jones, DA. 2010. Review of methods for deriving areal reduction factors. *Journal of Flood Risk Management* 3 (2010): 232–245.

USGS. 2002. *SRTM Topography* [online]. United States Geological Survey. Available from: http://dds.cr.usgs.gov/srtm/version2.1/SRTM/Document/Topo.pdf [2 June 2020].

USWB. 1957. *Rainfall Intensity-Frequency Regime, The Ohio Valley*. Technical Paper 29. United States Weather Bureau, Washington, DC, USA.

USWB. 1958. *Rainfall Intensity-Frequency Regime, South-eastern United States*. Technical Paper 29. United States Weather Bureau, Washington, DC, USA.

Van der Spuy, D and Rademeyer, PF. 2018. *Flood Frequency Estimation Methods as Applied in the Department of Water and Sanitation*. DWS, Pretoria, RSA.

Van Vuuren, SJ, Van Dijk, M and Coetzee, GL. 2012. *Status Review and Requirements of Overhauling Flood Determination Methods in South Africa*. Department of Civil Engineering, University of Pretoria, Pretoria, RSA.

Van Wyk, W. 1965. *Aids to the Prediction of Extreme Floods from Small Watersheds*. Department of Civil Engineering, University of the Witwatersrand, Johannesburg, RSA.

Veneziano, D and Langousis, A. 2005. *The Areal Reduction Factor a Multifractal Analysis*. Water Resources Research Report 41, W07008. DOI: 10.1029/2004WR003765

Viessman, W, Lewis, GL, and Knapp, JW. 1989. *Introduction to Hydrology*. 3rd Ed. Harper and Row Publishers Incorporated, New York, USA.

Weddepohl, JP. 1988. Design Rainfall Distributions for Southern Africa. Unpublished MSc dissertation, Department of Agricultural Engineering, University of Natal. Pietermaritzburg,

Wiederhold, JFA. 1969. *Design Storm Determination in South Africa*. HRU Report No. 1/1969. Hydrological Research Unit, University of the Witwatersrand. Johannesburg, RSA.

Wilson, EM. 1990. *Engineering Hydrology.* 4th Ed. Macmillan Press Limited, London, UK.

World Meteorological Organization. 1994. *Guide to Hydrological Practices*. 5th Ed. WMO Report No. 168. Geneva, Switzerland.

WR2012. 2012. *Water Resources of South Africa, 2012 Study* [online]. Available from: http://waterresourceswr2012.co.za/ [4 March 2022].

Yoo, C, Kim, K, Kim, HS and Park, MJ. 2007. Estimation of areal reduction factors using a mixed gamma distribution. *Journal of Hydrology* 335 (3–4): 271–284.

Zucchini, W and Hiemstra, LAV. 1984. *Part 1: Augmenting Hydrological Records*. WRC Report 91/3/84.Water Research Commission, Pretoria, RSA.

APPENDIX A: DAILY RAINFALL STATIONS

Table A.1: 1 779 Daily rainfall stations and their associated record length (RL) within the study area

Station number	RL	Station number	RL	Station number	RL	Station number	RL
Primary Drainage Region A							
0474751W	57	0512545W	79	0550612W	92	0677818W	75
0475456A	70	0512552W	70	0585056W	67	0678023W	80
0475881W	100	0512580W	79	0586341W	46	0678132W	63
0476012W	63	0512602W	63	0586441W	70	0678144W	92
0476031W	47	0512613W	93	0586545W	66	0678297W	59
0476040W	83	0512787W	56	0588230W	65	0678381W	50
0476071W	59	0513135W	91	0588385W	86	0678722W	64
0476100W	62	0513227W	71	0588406W	91	0718772W	47
0476128W	58	0513255W	74	0588721W	81	0718874W	66
0476129W	84	0513285W	82	0589503AW	57	0720727W	64
0476160W	64	0513312W	60	0589586W	49	0721197W	59
0476246W	76	0513337W	72	0589594A	55	0721257W	69
0476396W	92	0513345W	86	0589628W	73	0721618W	68
0508721W	67	0513347W	76	0589670W	82	0721665W	51
0508825W	94	0513350W	93	0589732AW	49	0721772W	55
0509078W	55	0513382W	82	0589877W	76	0722082W	75
0509283W	70	0513404AW	68	0590307W	81	0722277W	79
0509695W	55	0513405A	49	0590361W	94	0722497W	53
0509726W	68	0513437W	81	0590486W	66	0722529W	92
0509759W	88	0513464W	57	0629702W	55	0722571W	82
0510308W	91	0513465A	50	0630556W	65	0722614W	78
0510433W	66	0513524W	70	0630886W	61	0722653W	56
0510712W	87	0513550W	49	0631011W	64	0722700W	71
0510817W	68	0513643W	81	0631564W	67	0722721W	85
0511291W	74	0513677W	66	0631596W	68	0722749W	63
0511310W	62	0513827W	94	0632044W	47	0723070W	95
0511400W	87	0514010W	94	0632089W	50	0723155W	75
0511437W	47	0545499W	55	0632274W	63	0723331W	66
0511467W	74	0545626W	73	0632297W	64	0723334W	71
0511469W	77	0545839W	87	0632726W	51	0723338AW	65
0511477W	66	0546082W	69	0633393W	79	0723363W	94
0511524W	81	0546272W	70	0633463W	59	0723513W	91
0511527W	45	0546314W	79	0633796W	75	0762532W	50
0511672W	70	0546412W	94	0633881AW	94	0764161W	63
0512144W	57	0546525W	60	0634131W	83	0765707W	65
0512246W	70	0546630A	40	0672748W	88	0766030W	50
0512280W	70	0546712W	51	0673128W	57	0766324W	61
0512344W	55	0546740W	51	0674207W	59	0766480W	75
0512395W	64	0548165W	79	0675117W	65	0766509W	66
0512399W	64	0548747W	73	0675182W	90	0766779W	91
0512451W	70	0549130W	73	0676237W	57	0766837W	87
0512481W	75	0549354W	69	0676705W	52		
0512488W	50	0549358W	64	0677188W	49	0809706A	58
0512545A	83	0550545W	61	0677259W	77		
Primary Drainage Region B							
0477191W	61	0551120W	83	0593419W	89	0679019W	74
0477309W	91	0551281W	70	0593778W	62	0679139W	75
0477501W	91	0552363W	44	0594141W	94	0679141W	67
0477762W	78	0552700W	52	0594217W	60	0679156W	56
0478008W	74	0553351W	53	0594444W	86	0679164W	85

Station number	RL	Station number	RL	Station number	RL	Station number	RL
0478093W	91	0553593W	55	0594457W	90	0679197W	75
0478292W	80	0553651W	75	0594493W	65	0679221W	77
0478386W	57	0553859W	77	0594590W	72	0679227W	72
0478406W	88	0554614W	65	0594609W	62	0679268W	60
0478546W	70	0554786W	89	0594635W	70	0679401W	62
0478862W	51	0589867W	95	0594696W	63	0679441W	65
0479225W	70	0589896W	60	0634050W	74	0679456W	69
0479552W	72	0589897W	51	0634084W	70	0679508W	92
0513836W	93	0590028W	92	0634140W	72	0679532W	66
0514408W	91	0590444W	52	0634417W	68	0679592W	57
0514618W	93	0590500W	74	0635763W	49	0680225W	66
0515155W	59	0591125W	81	0636308W	63	0680280W	71
0515234W	54	0592371W	61	0636518W	74	0680354W	47
0515826W	95	0592474W	92	0637609W	69	0722900W	65
0516285A	62	0592560W	63	0678776W	95	0723055W	61
0516480W	64	0592615W	58	0678836W	58	0723080W	71
0516554W	93	0593015W	78	0678858W	83	0723231W	76
0551103W	45	0593126W	70	0678883W	52		
Primary Drainage Region C							
0201361W	85	0295001W	112	0366117W	75	0436575W	60
0201373W	53	0295116W	87	0366303W	69	0436747W	76
0201482W	85	0295139W	51	0366359W	51	0436855A	88
0228571W	68	0295174W	74	0366371W	63	0436855W	94
0228725W	52	0295254W	46	0366529W	59	0436887A	50
0228783W	56	0295405W	57	0366662W	54	0436887W	56
0229124W	59	0295408W	72	0366710W	48	0437100W	80
0229344W	52	0295413W	46	0366735W	78	0437104W	95
0229556A	41	0295539W	74	0366872W	56	0437134A	78
0229737W	108	0295558W	78	0367066W	68	0437194W	73
0230011W	54	0295760A	74	0367177W	65	0437383W	54
0230027W	80	0295760W	86	0367219W	83	0437452W	78
0230073W	75	0295770W	81	0367484W	49	0437517W	74
0230764W	91	0296157W	61	0367600W	68	0437555W	73
0230774W	62	0298301W	73	0367768A	88	0437660W	92
0230810W	92	0298481W	73	0367768W	94	0437834W	94
0231279W	92	0298512W	91	0367780W	49	0438047W	55
0231361W	64	0298638W	56	0368003W	73	0438315W	82
0231395W	71	0298871W	65	0368222W	59	0438514W	49
0231754W	55	0322071W	78	0368243W	77	0438588W	45
0232018W	85	0322172W	72	0368634W	89	0438716W	77
0232123W	87	0322329W	78	0368831W	79	0438734W	93
0232211W	88	0322648W	62	0369030W	46	0438744W	82
0232275W	93	0323102A	51	0369117W	78	0438750W	66
0256453W	115	0323144W	73	0369238W	91	0439389W	76
0256638W	78	0323649W	96	0369284W	67	0439396W	78
0257391A	63	0323820W	59	0369505W	83	0439688W	46
0257391W	66	0324202W	68	0369785W	67	0439755W	62
0257845W	84	0324379W	85	0369819W	77	0439764W	78
0258164W	70	0324561W	55	0369896W	51	0439769W	94
0258182W	84	0324607W	87	0370101W	92	0440018W	93
0258218W	50	0324725W	83	0396284W	80	0440129W	94
0258306W	68	0325471W	78	0396454W	49	0440157W	85
0258335W	49	0325816W	59	0396813W	85	0440435W	93
0258339W	50	0325877W	74	0396853W	43	0440449W	66
0258380W	63	0326073W	48	0397581W	66	0440637W	44
0258399W	50	0326668W	66	0397784W	64	0440767W	84
0258434AW	50	0326670W	55	0398177W	87	0440885W	89
0258458W	113	0327257W	86	0398409W	69	0441104W	61

Station number	RL	Station number	RL	Station number	RL	Station number	RL
0258467W	50	0327264W	79	0398556W	92	0441261W	81
0258581W	56	0327426W	64	0398876W	88	0441270W	71
0258740W	50	0327784W	90	0398889W	89	0441309W	79
0258812W	70	0327883W	87	0399208W	62	0441777W	73
0258894W	119	0327899W	53	0399241W	88	0442068W	72
0259278W	66	0328308A	60	0399404W	66	0442393W	61
0259348W	70	0328384W	87	0399662W	81	0442660W	93
0259578W	57	0328425W	73	0400203W	93	0442781A	50
0259609W	49	0328726W	66	0400647W	68	0442811A	41
0259727W	93	0328800W	61	0400792W	64	0442853W	50
0259743W	64	0329215W	91	0401272W	49	0442867W	81
0260004W	88	0329783W	77	0401407W	68	0443196W	91
0260030W	79	0330098W	79	0401658W	70	0443451W	70
0260163W	73	0330199W	93	0401798W	83	0471259W	69
0260519W	66	0330363W	76	0402081W	85	0471435W	45
0260555W	50	0330421W	85	0402104W	86	0471490W	60
0260678W	88	0330573W	81	0402140W	54	0472175W	71
0260882W	56	0330601W	58	0402166W	79	0472281W	76
0261146W	85	0330699W	67	0402300W	52	0472455W	44
0261183W	94	0330753W	72	0402543W	67	0473083W	55
0261275W	60	0330797W	72	0402788W	78	0473416W	62
0261365W	73	0331068W	85	0402827W	107	0473686W	74
0261523W	93	0331271W	84	0402866W	84	0473713W	55
0261597W	53	0331275W	71	0403054W	73	0474020W	48
0261722W	93	0331455W	64	0403646W	74	0474198W	85
0261733W	69	0331474W	71	0404007W	77	0474255W	95
0261750W	55	0331740W	70	0404132W	69	0474270W	88
0262129W	70	0331828W	58	0404152W	93	0474295W	51
0262314W	54	0332103W	68	0404177W	67	0474502W	93
0262353W	52	0332326W	60	0404316W	87	0475019W	57
0262479W	93	0332349A	85	0404614W	90	0475056W	64
0262694W	72	0332349W	91	0405295AW	70	0475370W	93
0287885W	63	0333226W	114	0405632W	71	0475611W	94
0288378W	95	0333682W	76	0405753W	60	0475669W	94
0288416W	60	0358263W	52	0406221W	85	0475717W	73
0288528W	79	0358268W	64	0406607W	92	0475736W	86
0288610W	58	0358349W	50	0431820W	66	0475761W	78
0289102W	58	0358825W	47	0432136W	81	0475852W	59
0290032W	114	0359304W	66	0432196W	88	0476012AW	63
0290463W	52	0359458W	69	0432237A	72	0476044W	89
0290468A	41	0359569W	58	0432387W	106	0476072W	99
0290560W	81	0359808W	112	0432633W	73	0476163W	93
0290810W	64	0360512W	100	0433115W	73	0476227W	67
0291148W	90	0360597A	54	0433494W	65	0476252W	61
0291245W	50	0361277W	45	0433512W	65	0476283AW	83
0291360W	44	0361295W	94	0433791W	76	0476283W	91
0291392W	119	0361354W	67	0433804W	90	0476403W	94
0291899A	79	0361736W	64	0433858W	74	0476433W	95
0291899W	85	0361760W	65	0434020W	72	0476630W	92
0292461W	89	0361832W	68	0434228W	84	0476641W	59
0292833W	47	0361846W	46	0434236W	68	0476644W	66
0293007W	66	0362159W	67	0434359W	76	0476736W	95
0293106W	71	0362862W	56	0434431W	54	0476766W	79
0293514W	73	0363470W	49	0434599W	73	0476802W	65
0293597A	70	0363571W	66	0434888W	87	0477772W	92
0293700W	66	0363651W	74	0435036W	61	0478360W	64
0293792W	90	0364322W	92	0435359W	59	0479238W	79
0294052W	53	0364562W	50	0435400W	69	0479298W	76

Station number	RL	Station number	RL	Station number	RL	Station number	RL
0294154W	60	0364728W	50	0435442W	45	0480170W	49
0294233W	84	0365400W	66	0435615W	47	0511084W	93
0294481W	50	0365409W	68	0435721W	64	0511120W	87
0294500W	75	0365444W	72	0435735W	94	0511418W	75
0294826W	71	0365731W	84	0436297W	56		
0294847W	69	0365855W	64	0436495W	78		
Primary Drainage Region D							
0087635W	65	0148525W	75	0199435W	68	0263041W	92
0088293W	116	0149082A	109	0200058W	57	0263047W	75
0088293A	113	0149082W	120	0200166W	64	0263228W	53
0088667W	65	0149204W	83	0200466W	119	0263280W	94
0089038W	48	0149490W	54	0200764W	68	0263314W	69
0089385W	64	0163882W	64	0201020W	93	0263316W	68
0089643W	63	0164141W	53	0201376W	75	0263373W	82
0089693W	68	0165003W	66	0201701W	54	0263499W	83
0091763W	78	0165580W	51	0202505W	66	0263567W	81
0091782W	46	0165898W	64	0202680W	55	0263760A	99
0110385W	108	0165898A	61	0203043W	111	0263760W	115
0110512W	64	0166238AW	79	0203175W	60	0263792W	91
0110649W	57	0166755W	86	0203478W	60	0264399P	67
0111373W	74	0167264W	70	0203483W	58	0267137P	59
0111789W	61	0167665W	95	0203595W	92	0279497W	47
0112346W	60	0168066W	60	0203657W	87	0280351W	48
0113025W	115	0168250W	68	0203737W	66	0280772W	65
0113043W	53	0168800W	63	0204050W	66	0281760W	53
0113321W	57	0169005W	111	0204058W	71	0282166W	62
0113330W	54	0169090W	79	0204138W	110	0282823W	99
0113673W	84	0170009AW	86	0204418W	70	0283098W	72
0113711W	110	0170009A	104	0204518W	63	0283466W	46
0114505W	58	0170099W	68	0204616W	61	0284008W	99
0114597W	45	0170137W	65	0204640W	66	0284361W	53
0114649W	75	0170315W	45	0204819P	67	0284832W	60
0114747W	88	0170427W	58	0207337P	76	0285742W	49
0114815W	66	0170639W	80	0220820W	50	0287041W	72
0118395W	83	0171117W	74	0223707W	46	0287441W	115
0118640W	86	0171546W	75	0223834W	55	0295720W	72
0119315W	91	0171756W	83	0224208W	71	0295807W	77
0122514W	86	0172163W	121	0224389W	71	0296115W	92
0122603W	57	0172580W	78	0224430W	115	0296169W	90
0137337W	68	0172724W	78	0224734W	64	0296238W	91
0137614W	49	0172814W	85	0225065W	62	0296379W	87
0138030W	47	0172818W	75	0225311W	52	0296583W	79
0138041W	85	0173134W	75	0225395W	48	0296622A	88
0138204W	66	0173266W	49	0225540W	64	0296682W	93
0139658W	75	0173497W	115	0225679W	98	0296682AW	74
0139850W	63	0174076W	55	0225792W	48	0296767W	57
0140020W	65	0174093W	71	0226327W	92	0296825W	81
0140207W	46	0174312W	94	0226572W	54	0296886W	52
0140582W	60	0174387W	47	0226771W	66	0297083P	74
0140616W	112	0174550W	114	0227054W	73	0297721W	82
0141066W	63	0174590W	51	0227127W	120	0298244W	63
0141204W	75	0174600W	111	0227368W	94	0316061W	55
0141329W	85	0174600A	91	0227756W	67	0317447A	109
0141664W	64	0174753W	71	0227811W	64	0320348W	105
0142153W	82	0174877W	47	0228170W	74	0320654W	68
0142262W	58	0175022W	73	0228458W	75	0320828W	57
0142379W	78	0175371W	91	0228495W	91	0321110W	75
0142497W	72	0176015W	68	0228567W	89	0321441W	58

Station number	RL	Station number	RL	Station number	RL	Station number	RL
0142805W	121	0176038W	47	0229089W	68	0330750W	68
0142853W	60	0176277W	73	0232083W	58	0331058W	62
0143258W	67	0176372W	111	0232383W	59	0355808W	45
0143345W	78	0176575W	90	0232673W	65	0356285W	79
0143579W	87	0176631W	59	0232823W	84	0356417W	80
0143598A	97	0176631AW	75	0232895W	83	0356636W	73
0143598W	93	0176735W	53	0232898W	75	0356712W	86
0143784W	113	0176740W	51	0233049W	81	0357413W	59
0144085W	66	0176891W	53	0233103W	66	0358307W	45
0144250W	49	0177045W	110	0233154W	53	0387240W	76
0144791W	112	0177178A	108	0233211W	67	0390514W	61
0145310W	69	0177178W	114	0233211A	71	0391834W	61
0145399W	95	0177184W	78	0233239W	74	0391857W	67
0146071W	53	0177441W	59	0233439P	69	0392148W	72
0146274W	45	0177552W	72	0247242W	104	0393083W	67
0146453W	73	0177885W	50	0251261W	120	0394574W	68
0146751W	57	0190868A	105	0251430W	50	0394878W	66
0147010W	74	0190868W	98	0251476W	55	0423044W	91
0147170W	69	0191526W	52	0251674W	52	0424357W	51
0147222W	80	0193339W	48	0252470A	59	0424509W	50
0147225W	78	0193347W	74	0252470W	47	0427083W	59
0147253W	53	0193561W	108	0252894W	66	0427469W	43
0147338W	65	0193561A	107	0253174W	82	0467487W	100
0147409W	73	0194323W	104	0253363A	56	0467818W	52
0147416W	67	0194454W	71	0253648W	70	0468210W	64
0147524W	81	0195730W	58	0254572W	55	0468318W	85
0147654W	109	0196375W	84	0254589W	84	0468500W	67
0147700W	53	0197713W	54	0255202W	97	0469359W	50
0147777W	88	0198435W	68	0255552W	61	0469459W	69
0148083W	82	0198524W	78	0255626W	80	0470196W	80
0148127W	55	0198726W	48	0256240W	45	0470516W	69
0148352A	107	0198836W	120	0256381W	63	0508261W	83
0148352W	113	0199107W	75	0257655W	78	0508428W	58
0148517W	119	0199275W	105	0263014W	58	0508649W	83
Primary Drainage Region E							
0042581W	54	0086079W	89	0109550W	46	0134378W	64
0084558W	72	0087186W	54	0109588W	62	0134478A	110
0084701W	129	0107318W	65	0109785W	70	0134478W	116
0085112W	90	0107396W	113	0109835W	65	0160807A	82
0085309W	100	0107510W	85	0131639W	73	0160807W	90
0085376W	89	0108311W	71	0133050W	84		
0086007W	84	0109215W	58	0133202W	90		
Primary Drainage Region F							
0157647W	48	0186139W	100	0214670W	115	0244405W	119
0157874W	107	0213888W	50	0244283W	66		
Primary Drainage Region G							
0002885W	87	0006733W	116	0020866W	148	0022113W	79
0002885A	75	0006836W	71	0021130A	72	0022113A	64
0003032A	111	0007050W	60	0021260W	93	0040604W	91
0003032W	117	0007106W	62	0021330W	89	0040653W	107
0004722A	102	0007263W	64	0021441W	56	0041060W	50
0004723A	100	0008136W	52	0021591A	87	0041347A	59
0004723W	109	0008136A	66	0021625A	48	0041417W	121
0004874W	46	0008470W	65	0021655W	108	0041417A	103
0005605A	101	0020689W	92	0021655A	105	0042227W	121
0005611A	43	0020716W	60	0021778W	74	0060864W	71
0006039A	84	0020719BW	90	0021795A	48	0062444W	119
0006167W	55	0020746W	100	0021809W	62	0084059W	62

Station number	RL	Station number	RL	Station number	RL	Station number	RL
0006192W	68	0020746AW	111	0021823W	109	0084159W	74
0006527W	80	0020747W	106	0021860A	51		
0006612W	78	0020776W	91	0022004W	56		
0006612A	70	0020839W	72	0022038W	94		
Primary Drainage Region H							
0006214W	64	0010456W	105	0023597W	101	0024197W	115
0006332W	66	0010742W	74	0023602W	56	0025414W	73
0007698A	58	0022368W	46	0023611A	62	0025599W	100
0007699W	55	0022504W	64	0023611W	66	0026510W	62
0008367W	67	0022521W	97	0023619W	63	0042325W	94
0008751W	68	0022539W	74	0023629W	69	0042357A	65
0008782W	110	0022759W	118	0023674W	65	0042802W	54
0009430W	53	0022789A	47	0023678W	110		
0009783W	78	0022803A	74	0023678AW	74		
0009815W	120	0022803W	86	0024146W	59		
Primary Drainage Region J							
0011132W	74	0044050W	112	0048624W	88	0069483W	78
0011617W	61	0044765W	77	0049050W	88	0069559W	73
0025162W	57	0045134W	112	0049060W	84	0069674W	53
0026215W	50	0045184W	79	0049372W	86	0069856W	85
0027302A	106	0045224W	56	0049562W	58	0070033W	66
0027302W	121	0045611W	97	0050058W	68	0070093W	73
0027876W	78	0046058W	67	0050205W	55	0090176W	86
0028335W	120	0046457W	91	0050230W	88	0090196W	86
0028335BW	52	0046479W	120	0050527W	62	0090600W	67
0028407W	72	0046898W	67	0066027W	73	0091288W	86
0028771W	74	0047359W	75	0067074W	66	0091835W	71
0028842W	81	0047436W	64	0068010W	88	0091890W	69
0029542W	83	0047716W	70	0068329W	110	0092141W	101
0029692W	85	0048043W	121	0068547W	85	0092283W	58
0030219W	102	0048083W	68	0068589W	67		
0030493W	63	0048275W	56	0068857W	110		
Primary Drainage Region K							
0014393W	110	0028536W	71	0029863W	89	0030446W	63
0014633W	107	0028838W	116	0030088W	70	0030775W	65
0017452W	120	0029294W	74	0030090W	107	0031237W	117
0017723W	65	0029297W	61	0030265W	108	0031507W	102
0028150W	70	0029624W	70	0030323W	54	0032209W	106
0028415W	79	0029805W	110	0030390W	49	0032507W	71
Primary Drainage Region L							
0030764W	46	0051430W	69	0073377W	59	0095006W	66
0031167W	47	0051580W	77	0073501W	48	0095123W	69
0031438W	53	0052270W	68	0092369W	79	0115528W	113
0032275W	63	0052571W	91	0092386W	77	0115595W	72
0032503W	66	0052590W	106	0093005W	48	0116029W	59
0033283W	50	0052765W	54	0093070W	59	0116083W	119
0033680W	106	0052886W	64	0093074W	108	0117047W	81
0033774W	67	0053588W	58	0093314W	108	0117447W	120
0034047W	68	0070770W	54	0093580W	84	0117487W	72
0034052W	68	0071121W	48	0094167W	71	0117675W	84
0034138W	63	0071264W	83	0094316W	57	0117749W	66
0034231W	69	0071264A	74	0094513W	63	0118029W	66
0050887W	121	0071294A	61	0094578W	110		
0050887A	115	0071337W	51	0094730A	58		
0051005W	61	0072712W	48	0094730W	66		
Primary Drainage Region M							
0034381W	117	0034762W	84	0035148W	48	0035209W	109
0034706W	60						

Station number	RL	Station number	RL	Station number	RL	Station number	RL
Primary Drainage Region N							
0033871W	71	0073871AW	90	0095119W	111	0096272W	99
0034121W	72	0074256W	73	0095273W	48	0096366W	56
0035605W	50	0074285W	60	0095395W	60	0096551W	63
0053274W	47	0074296AW	49	0095428W	71	0096680W	92
0053432W	61	0074296A	104	0095635W	64	0097140W	48
0054177W	54	0074296W	102	0095823W	108	0097239W	71
0054293W	98	0074363W	106	0096044W	44	0119082W	113
0054670W	49	0075090W	71	0096045A	87	0119209W	109
0054785W	57	0075215W	115	0096045BW	91	0119444W	89
0072662W	44	0075366W	72	0096045W	115		
0073555W	101	0075473W	57	0096094W	68		
0073871W	108	0075483W	70	0096101W	110		
Primary Drainage Region P							
0036605W	58	0037541A	62	0056597W	66	0057048AW	121
0036642W	79	0037696W	116	0056709W	106	0057108W	90
0036729W	113	0056139W	121	0056737W	67		
Primary Drainage Region Q							
0057737W	60	0078227W	121	0099496W	81	0121574W	60
0058192W	114	0078272W	102	0099622W	81	0121598W	84
0075745W	70	0078279W	63	0099735W	75	0121875W	73
0075759W	62	0078296W	52	0100025W	105	0122071W	75
0075784W	71	0078453W	111	0100060W	80	0122197W	54
0076133W	120	0097427W	58	0100329W	116	0122469W	65
0076215W	69	0098184W	65	0119736W	74	0122480W	121
0076555AW	90	0098190A	101	0120010W	75	0144449W	68
0076567W	94	0098190W	116	0120243W	51	0144900W	97
0076884W	95	0098595W	72	0120338W	113	0145029A	76
0077309W	96	0098713W	68	0120351W	68	0145261W	84
0077522W	107	0099020W	84	0120838W	56	0145810W	91
0077881W	93	0099102W	65	0121191W	46	0146017W	65
0078153W	85	0099221W	82	0121275W	84	0146052W	53
0078226W	69	0099229W	100	0121497W	48	0146588W	121
0078227A	115	0099492W	87	0121518W	100		
Primary Drainage Region R							
0059158W	49	0079215W	99	0079683AW	72	0080355W	115
0059243W	65	0079251W	97	0079730W	89	0080457W	81
0078587W	120	0079316W	84	0079809W	79	0080569W	79
0078755W	113	0079396W	99	0079823W	61	0080629W	115
0078859W	54	0079524W	93	0080072W	121		
0078879W	99	0079551W	65	0080143W	101		
Primary Drainage Region S							
0079485W	89	0101192W	99	0102542W	90	0123512W	70
0079490W	113	0101257W	110	0102762W	109	0123654W	108
0079632W	113	0101444W	82	0102840W	87	0123654A	108
0079782W	98	0101447W	66	0103081W	68	0123793W	63
0079811A	53	0101535W	113	0103139W	74	0124402W	113
0080694W	108	0101719W	98	0103230W	109	0124510W	84
0081007W	82	0101804W	106	0103397W	74	0125047W	75
0099789W	52	0102042W	107	0122806W	101	0125150W	101
0099811W	85	0102150W	72	0123063W	113	0149598W	91
0100594W	68	0102300W	72	0123111W	52	0150085W	108
0100779W	76	0102352W	55	0123304W	96		
0101097W	113	0102369W	78	0123466W	64		
Primary Drainage Region T							
0103377W	75	0128032W	80	0179713W	79	0182535A	83
0103516W	104	0128040W	82	0179790W	72	0182535W	87
0103570W	109	0128314W	78	0179864W	79	0182794W	74

Station number	RL	Station number	RL	Station number	RL	Station number	RL
0103886W	109	0129007W	75	0180030W	74	0207560W	95
0104206W	70	0150595W	50	0180032W	66	0208083W	93
0104762W	97	0150620W	92	0180123W	80	0208107A	55
0125432W	75	0151604W	107	0180394A	66	0208406A	85
0125454W	64	0151623W	57	0180439W	84	0208406W	93
0125868W	75	0152190W	75	0180537W	100	0208528A	57
0125880W	108	0152259W	70	0180648AW	49	0208635W	82
0126082W	78	0152468W	75	0180712W	81	0208733W	71
0126245W	73	0152475W	79	0180721A	60	0208743A	51
0126724W	91	0152482W	79	0180722AW	99	0209039W	97
0127197W	75	0152792W	75	0181073W	84	0209139W	83
0127298W	106	0153631W	82	0181423W	66	0209332W	68
0127406W	78	0153875W	98	0181561W	82	0209766W	96
0127414W	69	0154142W	87	0181604W	75	0209766AW	61
0127426W	54	0154354W	84	0181664W	51	0209825A	68
0127438W	105	0154796W	73	0182013W	62	0209886A	71
0127485A	61	0178881W	107	0182379A	68	0237471W	54
0127789W	78	0179344W	72	0182430W	47	0238022A	50
0127833W	75	0179353W	72	0182439W	65		
Primary Drainage Region U							
0182730A	53	0239421W	65	0240649W	71	0269114A	53
0182730W	61	0239472W	82	0240716W	78	0269295A	54
0183005W	88	0239566A	55	0240738W	57	0269477A	89
0210002W	85	0239577W	49	0240775W	61	0269532A	69
0210285A	50	0239585A	65	0240862W	63	0269774A	71
0210301W	54	0239585W	71	0240883A	61	0270021W	84
0210798A	59	0239605P	83	0240887W	63	0270119W	66
0210826W	75	0240022W	75	0240891W	127	0270219A	55
0211437W	73	0240073W	84	0241019W	65	0270243A	42
0238132W	74	0240185W	57	0241042AW	65	0270544W	70
0238468W	78	0240269W	61	0241042W	69	0271099W	75
0238543A	45	0240284W	70	0241042S	71	0271357S	41
0238636W	70	0240381W	74	0241103W	81		
0238837A	57	0240564W	69	0241131W	83		
0239097A	85	0240586W	66	0241302W	68		
Primary Drainage Region V							
0268441W	61	0299788A	55	0302320W	70	0335550A	61
0268640A	81	0299900W	71	0302503W	58	0335620W	67
0268845W	68	0300051A	58	0302628W	66	0335746W	60
0269043A	64	0300141A	52	0302687W	50	0336059W	50
0270722W	66	0300345W	46	0302699W	59	0370486W	74
0271252A	44	0300567A	86	0334174W	84	0370834W	59
0271402S	64	0300690W	58	0334175A	54	0371579W	82
0298791W	50	0301692W	75	0334825W	79	0371706W	71
0299008W	61	0301692A	74	0335250W	65	0406658W	68
0299614W	68	0301751W	79	0335520A	52	0406682AW	73
Primary Drainage Region W							
0272002S	67	0337148W	56	0373329A	54	0444277W	94
0272121W	86	0337431W	65	0373329W	56	0445100W	80
0272127W	77	0337795W	82	0373485W	79	0446471W	60
0272272W	57	0339065W	67	0373680W	70	0446741S	73
0303127W	79	0339326S	41	0374264W	77	0447446W	79
0303534S	66	0339352W	70	0374402W	67	0448450W	65
0303667P	45	0339354A	78	0375124W	82	0448597W	47
0304446W	82	0339357W	65	0407639W	69	0481167W	88
0304475S	68	0339415W	70	0408798W	74	0481239W	74
0304822W	72	0339441W	72	0409320W	50	0482229W	92
0305037W	79	0339483W	59	0409375W	94	0482357W	70

Station number	RL	Station number	RL	Station number	RL	Station number	RL
0305308W	66	0339538W	69	0409460W	74	0482867W	71
0336283W	62	0339734W	70	0410878W	79	0483053W	61
0337006W	58	0372496W	65	0412052W	64	0483082W	65
0337119A	46	0372852W	83	0444176W	64	0483426W	73
0337143W	66	0373058W	85	0444203W	79		
Primary Drainage Region X							
0479545W	70	0518759W	68	0555441W	83	0556143W	53
0481310W	60	0518859W	88	0555445W	65	0556898W	60
0516708W	78	0518886W	81	0555455W	67	0557115W	66
0517235W	63	0519134W	55	0555473W	55	0557712W	60
0517430W	95	0519310W	71	0555567W	93	0594539W	79
0517762W	79	0519732W	65	0555579W	65	0594764W	58
0517816W	84	0520125W	55	0555588W	67	0594806W	63
0518088W	87	0520589W	70	0555631W	56	0595110W	92
0518186W	76	0520636W	54	0555673W	61	0595161W	74
0518215W	65	0554175W	88	0555794W	54	0596179A	61
0518455W	82	0554682W	75	0555878W	65	0596179W	87
0518589W	64	0555280A	55	0556088W	68	0596647W	63
0518676W	67	0555405W	64	0556110W	58		

APPENDIX B: STEP-BY-STEP PROCEDURE

ArcGIS was used to extract design point rainfall values, inclusive of all the points within a particular circular catchment, for durations of 1, 3, 5 and 7 days associated with return periods of 2, 5, 10, 20, 50, 100 and 200 years. To simplify the interpretation of the steps followed, all GIS functions are highlighted using *Italic* font.

The following steps were followed:

(a) **Step 1:** Select the artificial circular catchment size, *e.g.*, 125 km^2.

(b) **Step 2:** Use the selected circular catchments of 125 km^2 to *Clip* all the RLMA&SI grid points falling within the circumference of the circular catchment as illustrated in Figure B.1.

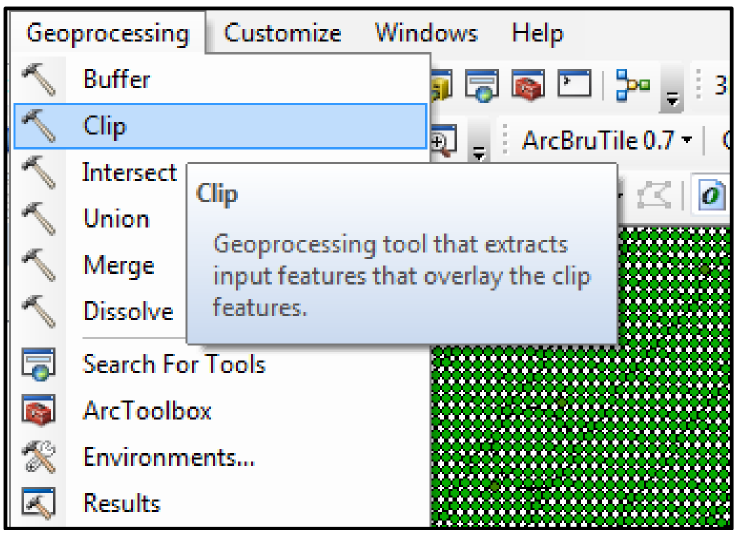

Figure B.1: Snapshot of the *Clip* function in GIS

(c) **Step 3:** All clipped grid points need to be associated with their respective circular catchments using the *Join* function (*cf.* Figure B.2), *i.e.*, grid points were joined with a specific circular catchment to result in a summative attribute table for export purposes to Microsoft Excel.

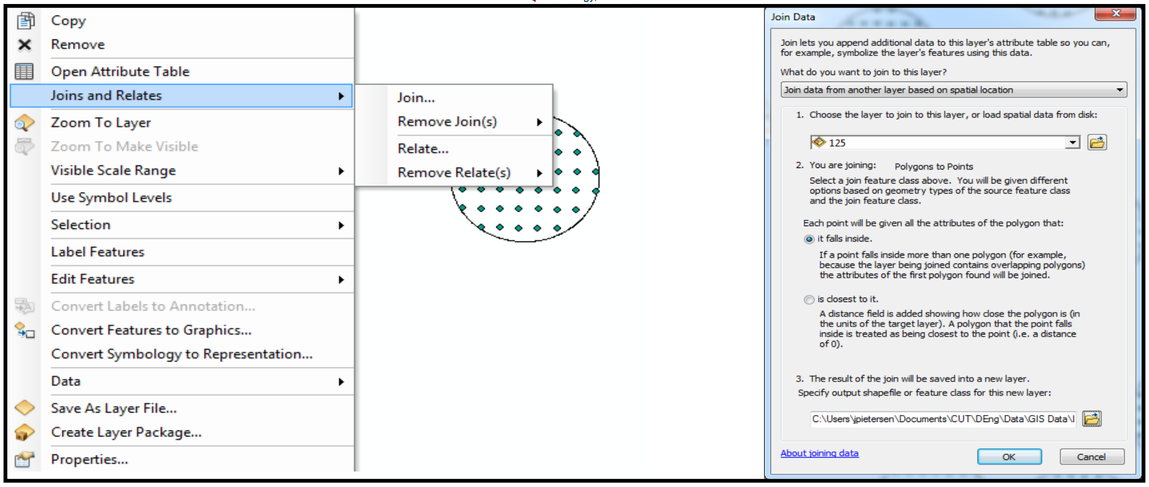

Figure B.2: Snapshot of the *Join* function in GIS

(d) **Step 4:** Copy all the information contained in the GIS attribute table to Microsoft Excel.

(e) **Step 5:** Insert extra columns (*cf.* Figure B.3) in the worksheet to estimate the latitude (LAT) and longitude (LONG) information (degrees and minutes) derived from decimal degrees.

D	E	F	G	H	I	J	K	L	M	N
Lat		Lo		LAT_M_	LONG_M_	LAT		LONG		MAP
-23	0.233333	29	0.116667	1394	1747	23	14	29	7	438
-23	0.233333	29	0.133333	1394	1748					422
-23	0.233333	29	0.15	1394	1749					417
-23	0.25	29	0.1	1395	1746					436
-23	0.25	29	0.116667	1395	1747					436
-23	0.25	29	0.133333	1395	1748					422
-23	0.25	29	0.15	1395	1749					426
-23	0.25	29	0.166667	1395	1750					415
-23	0.266667	29	0.1	1396	1746					442
-23	0.266667	29	0.116667	1396	1747					419
-23	0.266667	29	0.133333	1396	1748					414
-23	0.266667	29	0.15	1396	1749					412

Figure B.3: Conversion of decimal degrees to degrees and minutes

(f) **Step 6:** The calculated LAT and LONG coordinate values, for each independent circular catchment, can now be copied to the DRE_SAG design rainfall extraction utility, which is the executable application used in the RLMA&SI software.

(g) **Step 7:** The extracted design point rainfall results, inclusive of all the points within a particular circular catchment, are now available for durations of 1, 3, 5 and 7 days. The average or median design point rainfall values, for return periods of 2, 5, 10, 20, 50, 100, and 200 years can now be estimated.

APPENDIX C: MATLAB SCRIPTING

The MATLAB ® script running sequence is outlined below:

Simplification of Original Equation

The originally derived non-linear (second-order polynomial) log-transformed equation in its 'raw format':

Equation with three (3) constants:

$$ARF = A\gamma^2 + B\gamma - C$$

Region	A	B	C
1	-0.034	7.286	287.648
2	-0.037	7.896	319.770
3	-0.055	11.395	487.770
4	-0.024	5.391	196.710
5	-0.025	5.502	200.890

Sub-equation with thirteen (13) constants:

$$\gamma = a(b(\log D)^2 + c\log D + d) + e(f(\log T)^2 + g\log T + h) + i(j(\log A)^2 - k\log A + l) - m$$

Region	*a*	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*I*	*j*	*k*	*l*	*m*
1	1.06	- 8.89	18.43	86.36	1.03	-1.13	7.46	85.87	1.07	- 0.71	1.01	99.18	199.69
2	1.05	-9.11	17.43	88.07	1.03	-1.02	6.65	87.70	1.06	-0.60	1.00	99.71	199.18
3	1.05	-7.22	14.93	89.33	1.04	-0.32	4.37	89.88	1.04	-0.32	1.17	99.00	201.49
4	1.05	-11.81	23.28	83.25	1.03	-0.80	7.45	83.61	1.06	-0.51	2.29	100.13	194.50
5	1.04	-11.52	22.59	83.85	1.04	-0.87	6.80	84.89	1.07	-0.89	0.12	98.380	195.32

Define layout for improved equation (Region 1 used as an example):

$$X = \log(D),\ Y = \log(T),\ Z = \log(A)$$

$$= a(bX^2 + cX + d) + e(fY^2 + gY + h) + i(jZ^2 + kZ + l) - m$$

$$= abX^2 + acX + ad + efY^2 + egY + eh + ijZ^2 - ikZ + il - m$$

$$= x_1X^2 + x_2X + x_3Y^2 + x_4Y + x_5Z^2 + x_6Z + (ad + eh + il - m);$$

$where\ (ad + eh + il - m) = x_7$

$Let{:}\ a = -0.034,\ b = 7.286,\ c = -287.648$

Equation reduced to:

$$ARF = aX^2 + bX + c$$

MATLAB ® Running Sequence

Region 1:

syms X Y Z

%Data 1:

b=-8.899;c=18.425;d=86.36;f=-1.132;g=7.457;h=85.869;j=-0.705;k=1.01;l=99.176.

a=1.058;e=1.028;i=1.070;m=199.694.

y1(X)=a*(b*X^2 +c*X+d); %express as x1*X^2 +x2*X +a*d

vpa(y1,5)

ans(X) = $-9.415\, X^2 + 19.494X + 91.369$

y2(Y)=e*(f*Y^2 +g*Y+h); %express as x3*Y^2 +x4*Y +e*h

vpa(y2,5)

ans(Y) = $-1.164\, Y^2 + 7.666Y + 88.273$

y3(Z)=i*(j*Z^2 -k*Z+l); %express as x5*Z^2 +x6*Z +i*l

vpa(y3,5)

ans(Z) = $-0.754\, Z^2 + 1.081Z + 106.12$

x_7=a*d+e*h+i*l-m % add all the constants together

x_7 = 86.0665

Region 1 Test:

Duration = 1 hour

Return period = 50 year

Area = 5 000 km^2

D=1; T=50; A=5 000; X=log10(D); Y=log10(T);Z=log10(A).

a=-0.034; b=7.293; c=287.780.

X=x(1)*X^2 +x(2)*X +x(3)*Y^2 +x(4)*Y +x(5)*Z^2 +x(6)*Z +x(7).

ARF=a*X^2 + b*X - c

ARF = 80.17 %

Region 2:

syms X Y Z

%Data 2

a=1.046;b=-9.108;c=17.427;d=88.068;e=1.025;f=-1.017;g=6.650;h=87.694;i=1.055.

j=-0.596;k=1.003;l=99.711;m=199.182.

y1(X)=a*(b*X^2 +c*X+d); %express as x1*X^2 +x2*X +a*d

vpa(y1,5)

ans(X) = $-9.527\,X^2 + 18.2294X + 92.119$

y2(Y)=e*(f*Y^2 +g*Y+h); %express as x3*Y^2 +x4*Y +e*h

vpa(y2,5)

ans(Y) = $-1.042\,Y^2 + 6.816Y + 89.886$

y3(Z)=i*(j*Z^2 -k*Z+l); %express as x5*Z^2 +x6*Z +i*l

vpa(y3,5)

ans(Z) = $-0.629\,Z^2 + 1.058Z + 105.20$

x7=a*d+e*h+i*l-m % add all the constants together

x7 = 88.0186

Region 2 Test:

Duration = 1 hour

Return period = 50 year

Area = 5 000 km^2

D=1; T=50; A=5 000; X=log10(D); Y=log10(T);Z=log10(A).

a=-0.037; b=7.896; c=319.770.

Q=x(1)*X^2 +x(2)*X +x(3)*Y^2 +x(4)*Y +x(5)*Z^2 +x(6)*Z +x(7).

ARF=a*X^2 + b*X - c

ARF = 82.54 %

Region 3:

syms X Y Z

%Data 3:

a=1.053;b=-7.225;c=14.933;d=89.332;e=1.043;f=-0.316;g=4.374;h=89.881;i=1.039.

j=-0.317;k=1.170;l=99.001;m=201.485.

y1(X)=a*(b*X^2 +c*X+d); %express as x1*X^2 +x2*X +a*d

vpa(y1,5)

ans(X) = $-7.608\,X^2 + 15.7244X + 94.067$

y2(Y)=e*(f*Y^2 +g*Y+h); %express as x3*Y^2 +x4*Y +e*h

vpa(y2,5)

ans(Y) = $-0.330\,Y^2 + 4.562Y + 93.746$

y3(Z)=i*(j*Z^2 -k*Z+l); %express as x5*Z^2 +x6*Z +i*l

vpa(y3,5)

ans(Z) = $-0.330\,Z^2 + 1.216Z + 102.86$

x7=a*d+e*h+i*l-m % add all the constants together

x7 = 89.1895

Region 3 Test:

Duration = 1 hour

Return period = 50 year

Area = 5 000 km^2

D=1; T=50; A=5 000; X=log10(D); Y=log10(T);Z=log10(A).

a=-0.055; b=11.395; c=487.77.

Q=x(1)*X^2 +x(2)*X +x(3)*Y^2 +x(4)*Y +x(5)*Z^2 +x(6)*Z +x(7).

ARF=a*X^2 + b*X - c

ARF = 87.27 %

Region 4:

syms X Y Z

%Data 4:

a=1.047;b=-11.808;c=23.278;d=83.250;e=1.028;f=-0.795;g=7.451;h=83.609;i=1.063.

j=-0.508;k=2.292;l=100.13;m=194.495.

y1(X)=a*(b*X^2 +c*X+d); %express as x1*X^2 +x2*X +a*d

vpa(y1,5)

ans(X) = $-12.363\,X^2 + 24.372X + 87.163$

y2(Y)=e*(f*Y^2 +g*Y+h); %express as x3*Y^2 +x4*Y +e*h

vpa(y2,5)

ans(Y) = $-0.817\,Y^2 + 7.660Y + 85.95$

y3(Z)=i*(j*Z^2 -k*Z+l); %express as x5*Z^2 +x6*Z +i*l

vpa(y3,5)

ans(Z) = $-0.540\,Z^2 - 2.436Z + 106.44$

x7=a*d+e*h+i*l-m % add all the constants together

x7 = 85.056

Region 4 Test:

Duration = 1 hour

Return period = 50 year

Area = 5 000 km^2

D=1; T=50; A=5 000; X=log10(D); Y=log10(T);Z=log10(A).

a=-0.024; b=5.391; c=196.71.

Q=x(1)*X^2 +x(2)*X +x(3)*Y^2 +x(4)*Y +x(5)*Z^2 +x(6)*Z +x(7).

ARF=a*X^2 + b*X - c

ARF = 79.88 %

Region 5:

syms X Y Z

%Data 5:

a=1.038;b=-11.519;c=22.594;d=83.847;e=1.035;f=-0.866; g=6.799;h=84.888;i=1.066.

j=-0.894;k=0.121;l=98.375;m=195.316.

y1(X)=a*(b*X^2 +c*X+d); %express as x1*X^2 +x2*X +a*d

vpa(y1,5)

ans(X) = $-11.957\ X^2 + 23.453X + 87.035$

y2(Y)=e*(f*Y^2 +g*Y+h); %express as x3*Y^2 +x4*Y +e*h

vpa(y2,5)

ans(Y) = $-0.896\ Y^2 + 7.037Y + 87.859$

y3(Z)=i*(j*Z^2 -k*Z+l); %express as x5*Z^2 +x6*Z +i*l

vpa(y3,5)

ans(Z) = $-0.953\ Z^2 - 0.129Z + 104.87$

x7=a*d+e*h+i*l-m % add all the constants together

x7 = 84.4440

Region 5 Test:

Duration = 1 hour

Return period = 50 year

Area = 5 000 km^2

D=1; T=50; A=5 000; X=log10(D); Y=log10(T);Z=log10(A).

a=-0.025; b=5.502; c=200.89.

Q=x(1)*X^2 +x(2)*X +x(3)*Y^2 +x(4)*Y +x(5)*Z^2 +x(6)*Z +x(7).

ARF=a*X^2 + b*X - c

ARF = 79.71 %

APPENDIX D: ARF PERCENTAGE DIFFERENCES

As highlighted in Chapter 5 (Section 5.7), the comparisons shown in Figures 5.12 to 5.20, were further investigated by examining any trends evident in the ARF percentage differences. Given that the current geographically-centred ARF estimation methods used in South Africa provide constant ARF values irrespective of the return period under consideration, the average ARF values estimated using Eq. (5.2) and associated with the 2, 50, and 100-year return periods, were regarded as the 'most representative' ARF values for comparison purposes. The results are listed in Tables D.1 to D.4.

In Tables D.1 to D.4, the positive values (highlighted in green) represent the cases where the ARF estimates based on Eq. (5.2) exceed the ARF estimates based on the current geographically-centred ARF estimation methods used in South Africa. All the underestimations (negative values) are highlighted in red.

Table D.1: Percentage differences between the estimated ARFs using Eq. (5.2) and the current South African ARF estimation methods (Areal range 10 – 100 km^2)

Catchment area [km²]		10									50									100								
Duration	Hours	24			48			72			24			48			72			24			48			72		
	Days	1			2			3			1			2			3			1			2			3		
T [years]		2	50	100	2	50	100	2	50	100	2	50	100	2	50	100	2	50	100	2	50	100	2	50	100	2	50	100
Region 1	Fig 2.4	-10.1	-0.8	0.1	-4.0	2.2	2.7	-2.4	2.7	3.1	-11.8	-1.0	0.1	-5.5	2.0	2.7	-3.3	2.9	3.5	-12.7	-1.0	0.1	-6.1	2.2	3.0	-4.0	2.9	3.5
	Fig 2.5	-12.7	-3.3	-2.4	-5.9	0.4	0.9	-3.6	1.4	1.8	-15.9	-5.1	-4.0	-8.5	-1.0	-0.3	-5.9	0.4	0.9	-15.5	-3.9	-2.7	-8.1	0.1	0.9	-6.3	0.7	1.3
	Eq 2.7a	-18.2	-9.3	-8.6	-10.6	-5.2	-4.8	-7.9	-3.8	-3.5	-15.9	-5.1	-4.2	-8.9	-2.0	-1.4	-6.8	-1.4	-0.9	-15.3	-3.5	-2.5	-8.6	-0.8	-0.1	-6.7	-0.5	0.0
	Fig 2.6	-12.7	-3.3	-2.4	-5.9	0.4	0.9	-3.6	1.4	1.8	-16.4	-5.6	-4.5	-8.5	-1.0	-0.3	-5.9	0.4	0.9	-18.6	-6.9	-5.7	-10.1	-1.9	-1.1	-7.3	-0.3	0.3
	Eq 2.7b	-23.7	-14.4	-13.5	-18.9	-12.7	-12.2	-17.8	-12.8	-12.4	-20.7	-9.9	-8.8	-15.2	-7.7	-7.0	-13.9	-7.7	-7.1	-19.6	-7.9	-6.7	-13.7	-5.5	-4.7	-12.3	-5.3	-4.7
Region 2	Fig 2.4	-9.2	-1.2	-0.4	-3.9	1.4	1.8	-2.5	1.7	2.0	-10.3	-1.0	-0.1	-4.9	1.5	2.0	-3.1	2.2	2.6	-10.8	-0.9	0.1	-5.2	1.8	2.4	-3.6	2.2	2.8
	Fig 2.5	-11.7	-3.7	-3.0	-5.7	-0.4	0.0	-3.8	0.5	0.8	-14.3	-5.1	-4.2	-7.9	-1.6	-1.0	-5.7	-0.4	0.0	-13.6	-3.7	-2.7	-7.3	-0.3	0.3	-5.9	0.0	0.5
	Eq 2.7a	-17.3	-9.3	-8.6	-10.5	-5.2	-4.8	-8.1	-3.8	-3.5	-14.3	-5.1	-4.2	-8.4	-2.0	-1.4	-6.7	-1.4	-0.9	-13.4	-3.5	-2.5	-7.8	-0.8	-0.1	-6.3	-0.5	0.0
	Fig 2.6	-11.7	-3.7	-3.0	-5.7	-0.4	0.0	-3.8	0.5	0.8	-14.8	-5.6	-4.7	-7.9	-1.6	-1.0	-5.7	-0.4	0.0	-16.7	-6.8	-5.8	-9.3	-2.3	-1.7	-6.9	-1.0	-0.5
	Eq 2.7b	-22.7	-14.8	-14.0	-18.8	-13.5	-13.1	-18.0	-13.8	-13.5	-19.1	-9.9	-9.0	-14.6	-8.3	-7.7	-13.7	-8.4	-8.0	-17.7	-7.8	-6.8	-12.8	-5.9	-5.2	-11.9	-6.0	-5.5
Region 3	Fig 2.4	-7.2	0.3	1.3	-1.9	2.8	3.3	-0.6	3.1	3.4	-7.9	0.8	1.9	-2.6	3.1	3.8	-0.9	3.7	4.2	-8.2	1.1	2.4	-2.7	3.6	4.3	-1.1	4.0	4.5
	Fig 2.5	-9.7	-2.2	-1.2	-3.8	1.0	1.5	-1.9	1.8	2.1	-12.0	-3.3	-2.2	-5.7	0.1	0.7	-3.5	1.1	1.6	-11.0	-1.7	-0.5	-4.7	1.5	2.3	-3.4	1.7	2.3
	Eq 2.7a	-15.3	-7.8	-6.8	-8.5	-3.8	-3.3	-6.2	-2.5	-2.2	-12.0	-3.3	-2.2	-6.1	-0.4	0.3	-4.4	0.2	0.7	-10.8	-1.5	-0.3	-5.2	1.0	1.8	-3.8	1.3	1.8
	Fig 2.6	-9.7	-2.2	-1.2	-3.8	1.0	1.5	-1.9	1.8	2.1	-12.5	-3.8	-2.7	-5.7	0.1	0.7	-3.5	1.1	1.6	-14.0	-4.8	-3.5	-6.7	-0.5	0.3	-4.4	0.7	1.3
	Eq 2.7b	-20.8	-13.3	-12.3	-16.8	-12.1	-11.6	-16.1	-12.4	-12.1	-16.8	-8.2	-7.0	-12.4	-6.6	-6.0	-11.5	-6.9	-6.4	-15.0	-5.8	-4.5	-10.3	-4.1	-3.3	-9.4	-4.3	-3.7
Region 4	Fig 2.4	-11.1	-0.7	0.6	-3.7	3.5	4.4	-1.6	4.5	5.1	-13.6	-1.5	0.0	-5.7	2.9	3.9	-3.1	4.2	5.1	-14.7	-1.8	-0.1	-6.4	2.8	4.0	-3.9	4.0	4.9
	Fig 2.5	-13.7	-3.2	-1.9	-5.5	1.7	2.6	-2.8	3.2	3.8	-17.6	-5.6	-4.1	-8.8	-0.2	0.8	-5.6	1.7	2.5	-17.5	-4.6	-3.0	-8.5	0.8	1.9	-6.2	1.8	2.7
	Eq 2.7a	-19.2	-8.8	-7.5	-10.3	-3.1	-2.2	-7.1	-1.1	-0.4	-17.6	-5.6	-4.1	-9.2	-0.6	0.4	-6.6	0.7	1.5	-17.3	-4.4	-2.8	-9.0	0.3	1.4	-6.7	1.3	2.2
	Fig 2.6	-13.7	-3.2	-1.9	-5.5	1.7	2.6	-2.8	3.2	3.8	-18.1	-6.1	-4.6	-8.8	-0.2	0.8	-5.6	1.7	2.5	-20.5	-7.7	-6.0	-10.5	-1.3	-0.1	-7.2	0.8	1.7
	Eq 2.7b	-24.7	-14.3	-13.0	-18.6	-11.4	-10.6	-17.1	-11.0	-10.4	-22.4	-10.4	-8.9	-15.4	-6.9	-5.8	-13.6	-6.4	-5.5	-21.5	-8.7	-7.0	-14.1	-4.8	-3.7	-12.2	-4.3	-3.3
Region 5	Fig 2.4	-11.8	-3.3	-2.3	-5.3	0.4	1.1	-3.5	1.2	1.6	-12.5	-3.1	-1.9	-6.1	0.5	1.2	-3.9	1.6	2.1	-13.1	-3.0	-1.8	-6.3	0.8	1.5	-4.3	1.6	2.3
	Fig 2.5	-14.3	-5.8	-4.8	-7.1	-1.4	-0.8	-4.8	-0.1	0.4	-16.6	-7.2	-6.0	-9.1	-2.5	-1.8	-6.5	-1.0	-0.4	-15.9	-5.8	-4.6	-8.4	-1.3	-0.5	-6.5	-0.6	0.0
	Eq 2.7a	-19.9	-11.4	-10.5	-11.9	-6.2	-5.5	-9.0	-4.4	-3.9	-16.6	-7.2	-6.0	-9.5	-3.0	-2.2	-7.4	-2.0	-1.4	-15.7	-5.6	-4.4	-8.8	-1.8	-1.0	-7.0	-1.1	-0.5
	Fig 2.6	-14.3	-5.8	-4.8	-7.1	-1.4	-0.8	-4.8	-0.1	0.4	-17.1	-7.7	-6.5	-9.1	-2.5	-1.8	-6.5	-1.0	-0.4	-18.9	-8.8	-7.6	-10.4	-3.3	-2.5	-7.5	-1.6	-1.0
	Eq 2.7b	-25.3	-16.9	-15.9	-20.1	-14.5	-13.9	-19.0	-14.3	-13.9	-21.4	-12.0	-10.9	-15.8	-9.2	-8.5	-14.5	-9.0	-8.5	-19.9	-9.8	-8.6	-13.9	-6.9	-6.1	-12.5	-6.6	-6.0

Table D.2: Percentage differences between the estimated ARFs using Eq. (5.2) and the current South African ARF estimation methods (Areal range 500 – 1 000 km^2)

Catchment area [km²]		500									1000								
Duration	Hours	24			48			72			24			48			72		
	Days	1			2			3			1			2			3		
T [years]		2	50	100	2	50	100	2	50	100	2	50	100	2	50	100	2	50	100
Region 1	Fig 2.4	-16.0	-1.6	-0.2	-7.8	2.6	3.6	-5.4	3.6	4.5	-18.0	-2.2	-0.6	-9.4	2.2	3.3	-6.4	3.7	4.7
	Fig 2.5	-16.7	-2.4	-0.9	-8.2	2.2	3.2	-6.6	2.3	3.2	-18.0	-2.2	-0.6	-9.7	1.8	3.0	-6.6	3.5	4.4
	Eq 2.7a	-15.7	-0.5	0.8	-9.1	1.4	2.2	-7.7	1.0	1.7	-16.7	0.5	1.8	-10.0	2.0	2.9	-8.8	1.3	2.1
	Fig 2.6	-17.3	-3.0	-1.5	-9.6	0.8	1.8	-8.2	0.8	1.6	-16.3	-0.5	1.1	-8.3	3.3	4.4	-6.6	3.5	4.4
	Eq 2.7b	-17.6	-3.3	-1.8	-10.5	-0.1	0.9	-8.7	0.2	1.1	-17.1	-1.3	0.3	-9.3	2.3	3.4	-7.3	2.7	3.7
Region 2	Fig 2.4	-12.8	-0.8	0.5	-6.0	2.7	3.5	-4.2	3.4	4.1	-14.0	-0.9	0.5	-7.0	2.6	3.6	-4.7	3.7	4.5
	Fig 2.5	-13.5	-1.5	-0.3	-6.5	2.2	3.1	-5.4	2.1	2.8	-14.0	-0.9	0.5	-7.4	2.3	3.2	-4.9	3.5	4.3
	Eq 2.7a	-12.5	-0.5	0.8	-7.3	1.4	2.2	-6.5	1.0	1.7	-12.7	0.5	1.8	-7.7	2.0	2.9	-7.1	1.3	2.1
	Fig 2.6	-14.1	-2.1	-0.8	-7.8	0.9	1.7	-7.0	0.5	1.2	-12.4	0.8	2.2	-5.9	3.7	4.7	-4.9	3.5	4.3
	Eq 2.7b	-14.5	-2.5	-1.2	-8.7	0.0	0.8	-7.5	0.0	0.7	-13.2	0.0	1.4	-6.9	2.7	3.7	-5.6	2.7	3.6
Region 3	Fig 2.4	-8.8	2.1	3.6	-2.4	5.2	6.2	-0.7	5.7	6.5	-9.2	2.6	4.2	-2.8	5.6	6.7	-0.5	6.5	7.4
	Fig 2.5	-9.6	1.3	2.8	-2.9	4.8	5.7	-1.9	4.5	5.2	-9.2	2.6	4.2	-3.1	5.2	6.3	-0.8	6.3	7.2
	Eq 2.7a	-8.5	2.4	3.9	-3.7	3.9	4.9	-3.0	3.4	4.1	-7.9	3.9	5.6	-3.4	5.0	6.1	-3.0	4.1	5.0
	Fig 2.6	-10.1	0.8	2.3	-4.3	3.4	4.4	-3.5	2.9	3.7	-7.5	4.3	5.9	-1.7	6.7	7.8	-0.8	6.3	7.2
	Eq 2.7b	-10.5	0.4	1.9	-5.1	2.5	3.5	-4.0	2.3	3.1	-8.3	3.5	5.1	-2.7	5.7	6.8	-1.5	5.5	6.4
Region 4	Fig 2.4	-18.1	-2.7	-0.7	-8.3	2.9	4.3	-5.4	4.4	5.6	-20.0	-3.2	-1.1	-9.7	2.5	4.1	-6.3	4.4	5.7
	Fig 2.5	-18.9	-3.5	-1.5	-8.7	2.5	3.9	-6.6	3.1	4.3	-20.0	-3.2	-1.1	-10.1	2.2	3.7	-6.5	4.2	5.5
	Eq 2.7a	-17.8	-2.4	-0.4	-9.6	1.6	3.0	-7.7	2.0	3.2	-18.7	-1.9	0.3	-10.4	1.9	3.4	-8.7	2.0	3.3
	Fig 2.6	-19.4	-4.0	-2.0	-10.1	1.1	2.5	-8.2	1.5	2.7	-18.3	-1.5	0.6	-8.6	3.6	5.2	-6.5	4.2	5.5
	Eq 2.7b	-19.8	-4.4	-2.4	-11.0	0.2	1.6	-8.8	1.0	2.2	-19.1	-2.3	-0.2	-9.6	2.6	4.2	-7.2	3.4	4.8
Region 5	Fig 2.4	-15.1	-3.1	-1.7	-7.1	1.5	2.5	-4.8	2.6	3.4	-16.6	-3.5	-1.9	-8.2	1.2	2.3	-5.4	2.8	3.7
	Fig 2.5	-15.9	-3.9	-2.4	-7.6	1.0	2.0	-6.1	1.3	2.1	-16.6	-3.5	-1.9	-8.6	0.9	2.0	-5.6	2.6	3.5
	Eq 2.7a	-14.8	-2.8	-1.4	-8.4	0.2	1.2	-7.2	0.2	1.0	-15.2	-2.1	-0.5	-8.9	0.6	1.7	-7.8	0.4	1.3
	Fig 2.6	-16.4	-4.4	-3.0	-9.0	-0.4	0.7	-7.6	-0.3	0.6	-14.9	-1.8	-0.2	-7.1	2.3	3.4	-5.6	2.6	3.5
	Eq 2.7b	-16.8	-4.8	-3.4	-9.8	-1.2	-0.2	-8.2	-0.8	0.0	-15.7	-2.6	-1.0	-8.1	1.3	2.5	-6.4	1.8	2.8

Table D.3: Percentage differences between the estimated ARFs using Eq. (5.2) and the current South African ARF estimation methods (Areal range 5 000 – 10 000 km^2)

Catchment area [km²]		5000									10000								
Duration	Hours	24			48			72			24			48			72		
	Days	1			2			3			1			2			3		
T [years]		2	50	100	2	50	100	2	50	100	2	50	100	2	50	100	2	50	100
Region 1	Fig 2.4	-26.3	-5.8	-3.7	-13.9	1.3	2.8	-10.2	3.1	4.4	-31.5	-8.3	-5.8	-18.1	-0.9	0.9	-13.4	1.7	3.3
	Fig 2.5	-23.4	-2.9	-0.7	-12.7	2.4	4.0	-10.0	3.3	4.7	-27.9	-4.6	-2.2	-14.2	3.0	4.8	-12.0	3.1	4.7
	Eq 2.7a	-22.2	1.2	3.0	-14.6	2.2	3.4	-13.4	0.8	2.0	-26.5	0.8	2.8	-17.9	1.6	3.0	-16.7	0.0	1.3
	Fig 2.6	-16.6	4.0	6.2	-6.2	9.0	10.6	-3.5	9.8	11.1	-16.8	6.6	9.1	-4.9	12.3	14.1	-2.8	12.3	13.8
	Eq 2.7b	-16.9	3.7	5.8	-6.9	8.2	9.8	-4.4	8.8	10.2	-17.2	6.3	8.7	-6.0	11.2	13.0	-3.3	11.8	13.3
Region 2	Fig 2.4	-19.6	-2.9	-1.2	-9.6	2.8	4.1	-6.8	4.1	5.2	-23.0	-4.4	-2.4	-12.5	1.4	2.9	-8.9	3.4	4.6
	Fig 2.5	-16.6	0.1	1.8	-8.4	4.0	5.3	-6.6	4.3	5.5	-19.4	-0.7	1.3	-8.6	5.3	6.8	-7.5	4.8	6.0
	Eq 2.7a	-15.4	1.2	3.0	-10.2	2.2	3.4	-10.1	0.8	2.0	-17.9	0.8	2.8	-12.3	1.6	3.0	-12.3	0.0	1.3
	Fig 2.6	-9.7	6.9	8.7	-1.8	10.6	11.9	-0.1	10.8	11.9	-8.2	10.5	12.5	0.7	14.6	16.0	1.6	13.9	15.2
	Eq 2.7b	-10.1	6.6	8.4	-2.6	9.8	11.1	-1.0	9.9	11.0	-8.5	10.2	12.2	-0.4	13.5	14.9	1.2	13.4	14.7
Region 3	Fig 2.4	-11.6	2.6	4.6	-2.8	7.5	8.9	-0.4	8.4	9.6	-13.0	2.5	4.7	-4.2	7.1	8.7	-1.1	8.7	10.0
	Fig 2.5	-8.7	5.6	7.6	-1.7	8.6	10.0	-0.2	8.7	9.8	-9.4	6.2	8.4	-0.3	11.0	12.6	0.3	10.1	11.4
	Eq 2.7a	-7.5	6.8	8.8	-3.5	6.8	8.2	-3.7	5.2	6.4	-7.9	7.7	9.9	-4.0	7.3	8.8	-4.5	5.3	6.7
	Fig 2.6	-1.8	12.5	14.5	4.9	15.2	16.6	6.3	15.1	16.3	1.8	17.4	19.6	9.0	20.2	21.8	9.4	19.2	20.5
	Eq 2.7b	-2.1	12.1	14.2	4.2	14.4	15.8	5.3	14.2	15.4	1.5	17.1	19.3	7.9	19.1	20.7	9.0	18.7	20.0
Region 4	Fig 2.4	-26.9	-6.2	-3.5	-13.1	2.1	4.0	-9.1	4.2	5.9	-30.9	-7.9	-5.0	-16.4	0.4	2.6	-11.5	3.3	5.2
	Fig 2.5	-24.0	-3.2	-0.5	-12.0	3.2	5.2	-8.9	4.4	6.1	-27.3	-4.3	-1.3	-12.5	4.3	6.5	-10.1	4.7	6.6
	Eq 2.7a	-22.8	-2.0	0.6	-13.8	1.4	3.3	-12.4	0.9	2.6	-25.8	-2.8	0.2	-16.2	0.6	2.8	-14.8	-0.1	1.8
	Fig 2.6	-17.1	3.7	6.4	-5.4	9.8	11.8	-2.4	10.9	12.6	-16.2	7.0	9.9	-3.2	13.6	15.7	-0.9	13.8	15.7
	Eq 2.7b	-17.5	3.3	6.0	-6.2	9.0	11.0	-3.4	10.0	11.6	-16.5	6.6	9.6	-4.3	12.5	14.6	-1.4	13.3	15.2
Region 5	Fig 2.4	-23.0	-6.4	-4.4	-11.3	0.9	2.3	-7.9	2.7	4.0	-27.1	-8.5	-6.2	-14.7	-1.0	0.7	-10.4	1.6	3.0
	Fig 2.5	-20.1	-3.4	-1.4	-10.1	2.0	3.5	-7.7	2.9	4.2	-23.5	-4.8	-2.5	-10.7	2.9	4.6	-9.0	3.0	4.4
	Eq 2.7a	-18.9	-2.3	-0.2	-12.0	0.2	1.7	-11.2	-0.6	0.7	-22.0	-3.3	-1.0	-14.5	-0.8	0.8	-13.7	-1.7	-0.3
	Fig 2.6	-13.2	3.5	5.5	-3.5	8.6	10.1	-1.2	9.4	10.6	-12.4	6.4	8.7	-1.5	12.2	13.8	0.2	12.2	13.6
	Eq 2.7b	-13.5	3.1	5.2	-4.3	7.8	9.3	-2.2	8.4	9.7	-12.7	6.1	8.4	-2.6	11.1	12.7	-0.3	11.7	13.1

Table D.4: Percentage differences between the estimated ARFs using Eq. (5.2) and the current South African ARF estimation methods (Areal range 20 000 – 30 000 km^2)

Catchment area [km²]		20000									30000								
Duration	Hours	24			48			72			24			48			72		
	Days	1			2			3			1			2			3		
T [years]		2	50	100	2	50	100	2	50	100	2	50	100	2	50	100	2	50	100
Region 1	Fig 2.4																		
	Fig 2.5																		
	Eq 2.7a	-32.3	-0.3	1.9	-22.4	0.5	2.1	-21.0	-1.3	0.2	-36.7	-1.3	1.1	-25.7	-0.5	1.3	-24.0	-2.3	-0.8
	Fig 2.6																		
	Eq 2.7b	-17.5	9.5	12.3	-5.1	14.6	16.6	-2.2	15.1	16.9	-17.7	12.0	15.0	-4.4	17.0	19.2	-1.4	17.4	19.3
Region 2	Fig 2.4																		
	Fig 2.5																		
	Eq 2.7a	-21.4	-0.3	1.9	-15.2	0.5	2.1	-15.2	-1.3	0.2	-24.0	-1.3	1.1	-17.3	-0.5	1.3	-17.3	-2.3	-0.8
	Fig 2.6																		
	Eq 2.7b	-6.4	14.7	16.9	2.1	17.8	19.4	3.7	17.5	19.0	-4.8	18.1	20.5	4.0	20.8	22.6	5.4	20.3	21.9
Region 3	Fig 2.4																		
	Fig 2.5																		
	Eq 2.7a	-8.8	8.3	10.7	-5.0	7.5	9.2	-5.6	5.2	6.7	-9.5	8.5	11.1	-5.7	7.5	9.4	-6.4	5.1	6.7
	Fig 2.6																		
	Eq 2.7b	6.2	23.2	25.6	12.4	24.7	26.5	13.3	24.0	25.5	9.8	27.8	30.3	15.6	28.7	30.5	16.3	27.6	29.2
Region 4	Fig 2.4																		
	Fig 2.5																		
	Eq 2.7a	-29.8	-4.1	-0.8	-19.3	-0.6	1.8	-17.8	-1.5	0.6	-32.7	-5.2	-1.7	-21.5	-1.6	1.0	-19.9	-2.5	-0.3
	Fig 2.6																		
	Eq 2.7b	-15.0	10.9	14.2	-2.0	16.7	19.1	1.0	17.4	19.4	-13.6	14.2	17.7	-0.2	19.7	22.2	2.7	20.1	22.4
Region 5	Fig 2.4																		
	Fig 2.5																		
	Eq 2.7a	-26.6	-5.2	-2.7	-18.0	-2.5	-0.6	-17.1	-3.5	-1.9	-30.0	-6.8	-4.0	-20.6	-3.8	-1.8	-19.6	-4.9	-3.1
	Fig 2.6																		
	Eq 2.7b	-11.7	9.8	12.4	-0.7	14.9	16.7	1.8	15.3	17.0	-10.8	12.5	15.3	0.7	17.5	19.5	3.1	17.8	19.6

APPENDIX E: HTML 5 AND JAVASCRIPT CODING

The HTML 5 and JavaScript coding scripts used to develop the ARF software, are outlined below:

HTML 5

```
<!DOCTYPE html>
<html lang="en">
<head>
  <meta charset="UTF-8">
  <meta http-equiv="X-UA-Compatible" content="IE=edge">
  <meta name="viewport" content="width=device-width, initial-scale=1.0">
  <link rel="stylesheet" href="includes/css/bootstrap.css">
  <title>ARF Estimation SA</title>
  <style>
    #answer {
      font-size: 1.5rem;
    }
  </style>
</head>
<body>
<!-- As a heading -->
<nav class="navbar navbar-light" style="background-color: #D6F6F8;">
  <div class="container justify-content-center">
  <span class="navbar-brand mb-0" style="font-size: 35px; font-weight:
500;">AREAL REDUCTION FACTOR (ARF) ESTIMATION IN SOUTH AFRICA</span>
  </div>
</nav>

<!-- Body -->
<div class="container mb-5 pb-2">

<div class="row mt-4">
  <div class="col-md-5 col-lg-4" id="parentForAlert">

    <form id="FormDiv">
      <div class="form-group">
        <label for="area">Area <em>(A)</em></label>
        <input type="number" step="any" id="area" class="form-control"
placeholder="Catchment area in km&sup2;">
        </div>

        <div class="form-group">
        <label for="durationDays">Storm Duration <em>(D)</em></label>
        <input type="number" id="durationDays" class="form-control"
placeholder="Storm duration in days">
        </div>

        <div class="form-group">
        <label for="returnPeriodYears">Return Period <em>(T)</em></label>
        <input type="number" id="returnPeriodYears" class="form-control"
placeholder="Return period: 2&le; years &le;200">
        </div>

        <div class="form-group">
```

```
        <label for="regionNr">Region</label>
        <select id="regionNr" class="form-control">
          <option value="1">1</option>
          <option value="2">2</option>
          <option value="3">3</option>
          <option value="4">4</option>
          <option value="5">5</option>
        </select>
      </div>

      <div class="row">
        <div class="col-4">
          <button id="calculate" class="btn btn-primary mt-
5">Calculate</button>
        </div>
        <div class="col-7 text-center">
          <h4 class="text-info mt-5" id="answer"><!-- JS Output --></h4>
        </div>
      </div>
      </form>


  </div>

  <div class="col-md">
    <img src="images/MAP1.png" class="rounded w-100" alt="South Africa
Map">
    <figcaption class="figure-caption">Five Homogeneous ARF Regions of
South Africa.</figcaption>
  </div>
</div>


<table class="table table-sm mt-5 mb-0 table-bordered">
  <thead>
    <tr class="text-center">
      <th scope="col">T (years)</th>
      <th scope="col">2</th>
      <th scope="col">5</th>
      <th scope="col">10</th>
      <th scope="col">20</th>
      <th scope="col">50</th>
      <th scope="col">100</th>
      <th scope="col">200</th>
    </tr>
  </thead>
  <tbody>
    <tr class="text-center">
      <th scope="row">ARF (%)</th>
      <td id="answerLoop_2"><!-- JS OUTPUT --></td>
      <td id="answerLoop_5"><!-- JS OUTPUT --></td>
      <td id="answerLoop_10"><!-- JS OUTPUT --></td>
      <td id="answerLoop_20"><!-- JS OUTPUT --></td>
      <td id="answerLoop_50"><!-- JS OUTPUT --></td>
      <td id="answerLoop_100"><!-- JS OUTPUT --></td>
      <td id="answerLoop_200"><!-- JS OUTPUT --></td>
    </tr>
  </tbody>
```

```
</table>


</div> <!-- Eind of div.container -->

<!-- Footer -->
<footer class="text-dark" style="background-color: #AEEDF2;">
  <div class="container">
    <div class="row justify-content-center">
      <div class="col-md">
        <img src="images/WRC2.png" class="w-75 my-3 pl-5" alt="WRC Logo">
      </div>
      <div class="col-md">
        <img src="images/CUT1.png" class="w-75 my-3 pl-5" alt="CUT Logo">
      </div>
      <div class="col-md">
        <img src="images/DHEAT.png" class="w-75 my-3 pl-5" alt="DHeat
Logo">
      </div>
      <div class="col-md">
        <img src="images/NRF3.png" class="w-75 my-3 pl-5" alt="NRF Logo">
      </div>
    </div>

    <div class="text-center pb-3">
      <small>This program implements the procedures to estimate Areal
Reduction Factors (ARFs) in South Africa, as developed by JPJ Pietersen and
OJ Gericke. Funding for this project was obtained from the: (i) Water
Research Commission through a project entitled "DEVELOPMENT OF A
REGIONALISED APPROACH TO ESTIMATE AREAL REDUCTION FACTORS AND CATCHMENT
RESPONSE TIME PARAMETERS FOR IMPROVED DESIGN FLOOD ESTIMATION IN SOUTH
AFRICA (WRC Project K5-2924)”, (ii) Central University of Technology, Free
State (CUT) in collaboration with the Department of Higher Education and
Training (DHET), and (iii) National Research Foundation (NRF) of South
Africa.</small>

      <br><br>
      <small class="text-muted">&copy; Developed by JPJ Pietersen (2022)
</small>
    </div>
</div>
</footer>

<script src="includes/js/app.js"></script>
</body>
</html>
```

JavaScript

```
document.getElementById('calculate').addEventListener('click', calculate);
  function calculate(e) {
  // Values from form
  var area = document.getElementById('area').value;
  var durationDays = document.getElementById('durationDays').value;
  var returnPeriodYears =
document.getElementById('returnPeriodYears').value;
  var regionNr = document.getElementById('regionNr').value;

  if(area == '' || durationDays == '' || returnPeriodYears == '') {

    // Create Alarm if fields are empty
    createAlarm('Please complete all compulsary fields.');

    // Remove alarm after 5s
    setTimeout(removeAlert, 5000);
  }

  else if (returnPeriodYears < 2 || returnPeriodYears > 200) {
    createAlarm('Please provide a return period between 2 and 200 years');
    setTimeout(removeAlert, 5000);
  }

  else {

  // Standard year periods
  const periods = [2, 5, 10, 20, 50, 100, 200];

  switch (regionNr) {
  case '1':
    var x1 = -9.415;
    var x2 = 19.494;
    var x3 = -1.164;
    var x4 = 7.666;
    var x5 = -0.754;
    var x6 = -1.081;
    var x7 = 86.067;

    var A = -0.034;
    var B = 7.286;
    var C = 287.648;
        var eQLong = longEquation(x1, durationDays, x2, x3, x4,
returnPeriodYears, x5, x6, area, x7);
    // print function
    calculateARF(A, B, C, eQLong);

        // Loop throug years and calculate into table
    for (var i = 0; i < periods.length; i++) {

      var years = periods[i];
      document.getElementById('answerLoop_' + years).innerHTML = '';

      var eQLongL = longEquation(x1, durationDays, x2, x3, x4, years, x5,
x6, area, x7);
      // Print answer
      calculateARFloop(A, B, C, years, eQLongL);
```

```
    }
    break;

  case '2':
    var x1 = -9.527;
    var x2 = 18.229;
    var x3 = -1.042;
    var x4 = 6.816;
    var x5 = -0.629;
    var x6 = -1.058;
    var x7 = 88.019;

    var A = -0.037;
    var B = 7.896;
    var C = 319.770;

    var eQLong = longEquation(x1, durationDays, x2, x3, x4,
returnPeriodYears, x5, x6, area, x7);
    // print function
    calculateARF(A, B, C, eQLong);

    // Loop throug years and calculate into table
    for (var i = 0; i < periods.length; i++) {

      var years = periods[i];
      document.getElementById('answerLoop_' + years).innerHTML = '';

      var eQLongL = longEquation(x1, durationDays, x2, x3, x4, years, x5,
x6, area, x7);
      // Print answer
      calculateARFloop(A, B, C, years, eQLongL);
    }

    break;

  case '3':
    var x1 = -7.608;
    var x2 = 15.724;
    var x3 = -0.330;
    var x4 = 4.562;
    var x5 = -0.330;
    var x6 = -1.216;
    var x7 = 89.190;

    var A = -0.055;
    var B = 11.395;
    var C = 487.770;

    var eQLong = longEquation(x1, durationDays, x2, x3, x4,
returnPeriodYears, x5, x6, area, x7);
    // print function
    calculateARF(A, B, C, eQLong);

    // Loop throug years and calculate into table
    for (var i = 0; i < periods.length; i++) {

      var years = periods[i];
      document.getElementById('answerLoop_' + years).innerHTML = '';
```

```
      var eQLongL = longEquation(x1, durationDays, x2, x3, x4, years, x5,
x6, area, x7);
      // Print answer
      calculateARFloop(A, B, C, years, eQLongL);
    }

    break;

  case '4':
    var x1 = -12.363;
    var x2 = 24.372;
    var x3 = -0.817;
    var x4 = 7.660;
    var x5 = -0.540;
    var x6 = -2.436;
    var x7 = 85.056;

    var A = -0.024;
    var B = 5.391;
    var C = 196.710;

        var eQLong = longEquation(x1, durationDays, x2, x3, x4,
returnPeriodYears, x5, x6, area, x7);
    // print function
    calculateARF(A, B, C, eQLong);

    // Loop throug years and calculate into table
    for (var i = 0; i < periods.length; i++) {

      var years = periods[i];
      document.getElementById('answerLoop_' + years).innerHTML = '';

      var eQLongL = longEquation(x1, durationDays, x2, x3, x4, years, x5,
x6, area, x7);
      // Print answer
      calculateARFloop(A, B, C, years, eQLongL);
    }

    break;

  case '5':
    var x1 = -11.957;
    var x2 = 23.453;
    var x3 = -0.896;
    var x4 = 7.037;
    var x5 = -0.953;
    var x6 = -0.129;
    var x7 = 84.444;

    var A = -0.025;
    var B = 5.502;
    var C = 200.890;

        var eQLong = longEquation(x1, durationDays, x2, x3, x4,
returnPeriodYears, x5, x6, area, x7);
    // print function
    calculateARF(A, B, C, eQLong);
```

```
    // Loop throug years and calculate into table
    for (var i = 0; i < periods.length; i++) {

      var years = periods[i];
      document.getElementById('answerLoop_' + years).innerHTML = '';

      var eQLongL = longEquation(x1, durationDays, x2, x3, x4, years, x5,
x6, area, x7);
      // Print answer
      calculateARFloop(A, B, C, years, eQLongL);
    }

    break;

  default:
    createAlarm('Some error occurred.');
    setTimeout(removeAlert, 5000);
}
  } // End of else

    // Prevent default
    e.preventDefault();
  } // End of calculate function

  // Create Alarm Div
  function createAlarm(message) {
    const div = document.createElement('div');
    div.className = 'alert alert-danger';
    // Add text
    div.appendChild(document.createTextNode(message));
    // Get parent
    const ParentContainer = document.querySelector('#parentForAlert');
    // Kry die form div
    const formDIV = document.querySelector('#FormDiv');
    // Inset alert
    ParentContainer.insertBefore(div, formDIV);
  }

  // Remove alert function
  function removeAlert() {
    const divAlert = document.querySelector('.alert');

    if(divAlert) {
      divAlert.remove();
    }
  }

  // Lang formule
  function longEquation(x1, durationDays, x2, x3, x4, returnPeriodYears,
x5, x6, area, x7) {
    // Part durationDays
    var a1 = x1 * Math.pow((Math.log10(durationDays)), 2);
    var a2 = x2 * Math.log10(durationDays);
    var aTotal = a1 + a2;

    // Part returnPeriodYears
    var b1 = x3 * Math.pow((Math.log10(returnPeriodYears)), 2);
```

```
    var b2 = x4 * Math.log10(returnPeriodYears);
    var bTotal = b1 + b2;

    // Part area
    var c1 = x5 * Math.pow((Math.log10(area)), 2);
    var c2 = x6 * Math.log10(area);
    var cTotal = c1 + c2;

    return aTotal + bTotal + cTotal + x7;
  }

  // Final ARF
  function calculateARF(A, B, C, eQLong) {

    var ARFlong = A * Math.pow(eQLong, 2) + B * eQLong - C;

    var ARF = ARFlong.toFixed(1);

    if (ARF >= 100) {
      ARF = 100;
    }

    return document.getElementById('answer').innerHTML = `ARF = ${ARF}%`;

  }

  // Loop final ARF
  function calculateARFloop(A, B, C, year, eQLong) {

    var ARFlong = A * Math.pow(eQLong, 2) + B * eQLong - C;

    var ARF = ARFlong.toFixed(1);

    if (ARF >= 100) {
      ARF = 100;
    }

    return document.getElementById('answerLoop_' + year).innerHTML +=
`${ARF}`;

  }
```