DEVELOPMENT OF A REGIONALISED APPROACH TO ESTIMATE AREAL REDUCTION FACTORS AND CATCHMENT RESPONSE TIME PARAMETERS FOR IMPROVED DESIGN FLOOD ESTIMATION IN SOUTH AFRICA

Report to the
Water Research Commission

by

OJ Gericke[1], JPJ Pietersen[1], JC Smithers[2], JA du Plessis[3], CE Allnutt[1] and VH Williams[1]

[1] Central University of Technology, Free State
[2] University of KwaZulu-Natal
[3] Stellenbosch University

May 2023

WATER
RESEARCH
COMMISSION

EXECUTIVE SUMMARY

Event-based deterministic design flood estimation methods are the most commonly used by practitioners in ungauged catchments. In the application of these event-based deterministic methods, it is acknowledged that both the spatial and temporal distribution of runoff, as well as the critical duration of rainfall, are influenced by the catchment response time. Typically, all complex, heterogeneous catchment processes are lumped into a single process to enable the estimation of the expected output (design flood) from causative input (average areal design rainfall and catchment response time).

Design point rainfall estimates are only applicable to a limited area and for larger areas, the average areal design rainfall depth is likely to be less than the maximum design point rainfall depths. 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 (response time) and catchment area. The latter hydrological response of a catchment, i.e. the catchment response time, is normally expressed as a single time parameter, e.g. time of concentration (T_C), lag time (T_L) and/or time to peak (T_P). In other words, when event-based deterministic design flood estimation methods are applied in ungauged catchments, estimates of the peak discharge are based on a single, representative catchment response time parameter. Therefore, average areal design rainfall and catchment response time parameters are regarded as fundamental input to all event-based design flood estimation methods in ungauged catchments; while errors in estimated average areal design rainfall and catchment response time, will directly impact on estimated peak discharges.

In South Africa (SA), 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 latter 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 wrongfully 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.

In terms of catchment response time parameter estimation in South Africa, unfortunately, none of the empirical T_C estimation methods recommended for general use were developed and verified using local data. In small, flat catchments with overland flow being dominant, the use of the Kerby equation is recommended, while the empirical United States Bureau of Reclamation (USBR) equation is used to estimate T_C as channel flow in a defined watercourse. Both the Kerby and USBR equations were developed and calibrated in the United States of America (USA) for catchment areas less than 4 ha and 45 ha, respectively. Consequently, practitioners in South Africa commonly apply these 'recommended methods' outside their bounds, both in terms of areal extent and their original developmental regions, without using any local correction factors. The empirical estimates of T_L used in South Africa are limited to the family of equations developed by the Hydrological Research Unit (HRU), the United States Department of Agriculture Natural Resource Conservation Service (USDA NRCS), formerly known as the USDA Soil Conservation Service (SCS), and SCS-SA equations.

In considering the above status quo in South African flood hydrology, the overall objective of this project is to develop a regionalised approach to estimate ARFs and at-site catchment response time parameters for improved design flood estimation in South Africa. The specific objectives of the project are to: (i) contribute to the establishment of a national catchment variable database, (ii) extract and analyse rainfall and runoff data to provide geographically-centred and probabilistically correct ARFs and observed catchment response time parameters, (iii) refine and update the regionalisation scheme for ARFs in South Africa, (iv) derive regional empirical ARF equations for application throughout South Africa in the different homogeneous rainfall regions, (v) derive a regional time parameter equation for application in Primary Drainage Region X, and (v) assess and verify the derived ARF and time parameter equations.

In terms of ARFs, the primary research aim of this study is to estimate geographically-centred 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).

A total of 2 550 artificial circular catchments with an associated 1 779 rainfall stations with at least 30 years combined areal record lengths, were strategically positioned in 46 homogeneous rainfall regions throughout South Africa. Due to the large number of circular catchments placed in each of the 46 ARF regions, 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 ARF region. In principal this was not regarded as problematic, while it also contributed to the 'smooth' transition between the different regions.

In applying various 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 using Linear Moments (LM) resulted in areal and 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). The estimation of sample ARFs was expressed as the ratio between the areal catchment design rainfall and the design point rainfall estimates for corresponding return periods.

Five (5) ARF regions were deduced from the 46 ARF regions and a single regional empirical ARF equation [Eq. (6.2)], with unique regional calibration coefficients, was assigned to each region. Initially, linear backward stepwise multiple regression analyses with deletion were performed at a 95% confidence level in order to estimate the relationship between the dependent criterion variable (ARF) and the independent predictor variables (catchment area, storm duration and return period) within each of region. Ultimately, the linear regression analyses were outperformed by a second order polynomial non-linear log-transformed empirical ARF equation. The derived regional empirical ARF equation [Eq. (6.2)] performed similarly, and as expected, when compared to a selection of empirical geographically-centred ARF estimation methods currently used in local and/or international practice. The ARFs estimated with Eq. (6.2) decreased within an increase in area, and increased with an increase in both storm duration and return period.

All the above results also confirmed the study assumptions applicable to ARFs, *viz.*: (i) design point rainfall estimates are only representative for a limited area – demonstrated by the differences between areal design rainfall and design point rainfall estimates, (ii) ARFs vary with predominant weather types, storm durations, seasonal factors and return periods – evident in the different ARF regions and hence the reason for having the five (5) ARF regions, and (iii) the current South African ARF estimation methods are only applicable to specific temporal and spatial scales – demonstrated by the absence of any regionalisation, the ARF values exceeding 100% in 'smaller' catchments, the constant ARF values associated with all return periods, and the limited data used.

The ARF methodology used in this study and the subsequent findings are new to the South African flood hydrology research community and practice, *viz.*: (i) ARFs were derived and based on a regionalisation scheme utilising local and up-to-date daily rainfall data, (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 5 ARF regions of South Africa.

In terms of catchment response time, the primary aim is to expand and verify the approach developed by Gericke and Smithers (2017) by estimating observed catchment response time parameters from 51 gauged catchments located in Primary Drainage Region X and to derive a regional empirical time parameter equation.

The conceptual approach developed and refined to derive the time to peak (T_{Px}) using only observed streamflow data at a catchment level, proved to be both practical and objective with consistent results. The combined use of Eqs. (7.2), (7.3) and (7.4) not only ensured/will ensure that the high variability of event-based catchment responses is taken into account, but the estimated catchment T_{Px} values are also well within the range of other uncertainties inherent to all design flood estimation procedures. The high degree of association ($r^2 = 0.84$) between Eqs. (7.2) and (7.3) also confirmed that T_P equals the total net rise (duration) of a multiple-peaked hydrograph in medium to large catchments. The average error bounds between the three different approaches, e.g. net rise duration [Eq. (7.2)], triangular-shaped hydrograph approximation [Eq. (7.3)] and linear response function [Eq. (7.4)] were also limited to ≤ 15%.

It is recommended that for design hydrology and for the calibration of empirical time parameter equations, that the catchment T_{Px} should be estimated using Eq. (7.4). In addition, the conceptual approach used to derive the empirical time parameter equation [Eq. (7.10)], should be adopted when regional time parameter equations are derived at a national-scale in South Africa. It is suggested that the methodology developed (and refined) in this study, should be gradually expanded to Primary Drainage Regions A and B, before deploying it at a national-scale. Approximately 110 gauged catchments covering the whole of the Gauteng, Mpumalanga, Limpopo and the Northern Provinces are situated in regions A, B and X. Typically, these three regions do not only form a continuous geographical region, but the largest percentage of South Africa's population also resides here and are frequently subjected to extreme flooding.

In terms of time parameter proportionality ratios, a pilot case study was conducted in the C5 secondary drainage region in South Africa to investigate and establish the suitability of the currently recommended time parameter definitions and proportionality ratios for small catchments in larger catchment areas exceeding 50 km². The focus was on the development of an automated hyetograph-hydrograph analysis tool to estimate time parameters and average time parameter proportionality ratios at a catchment level.

The Automated Toolkit for hyetograph-hydrograph analyses proved to be very useful in mimicking the typical convolution procedure practitioners would follow to visually inspect and interpret hyetograph-hydrograph data sets. An enhanced methodology was developed, which considered both the impact of the spatial distribution of rainfall events and the distance thereof from the catchment outlet on the resulting runoff and consequently, the derivation of time parameters and proportionality ratios.

The time parameter estimates based on the seven different theoretical time parameter definitions proved to be highly variable due to the spatial and temporal distribution of rainfall events, variation in peak discharges and the distance of the rainfall events from the catchment outlet. In contrast, the time parameter proportionality ratios were characterised by a relatively low variability. In all the sub-catchments under consideration, it was confirmed that $T_C \approx T_L \approx T_P$. In other words, it was demonstrated that the time parameter proportionality ratios currently proposed for small catchments, i.e. $T_C = 1.417T_L$ and $T_C = 1.667T_L$, are not applicable at larger catchment levels.

In terms of stage-discharge relationships above the structural limit of a flow-gauging weir, a pilot scale study was conducted in 10 gauged catchments to evaluate and compare a selection of indirect extension methods (e.g. hydraulic and one-dimensional modelling methods) to direct extension (benchmark) methods (e.g. at-site conventional current gaugings, hydrograph analyses and level pool routing techniques), to establish the best-fit and most appropriate stage-discharge extension method to be used in South Africa.

Overall, the results highlighted that the Stepped Backwater Analysis (SBA) and Slope Area Method (SAM) are the most appropriate indirect estimation methods to reflect the hydraulic conditions during high discharges at a flow-gauging site. It was emphasised that any extension method must be hydraulically correct if it is to be used as a robust approach to extend stage-discharge rating curves beyond the structural limit. The extension of any rating curve is significantly more affected by the site (and river reach) geometry, initial hydraulic conditions, flow regimes and level of submergence at high discharges, than the actual extension method used. Hence, there is no 'one size fits all' approach/method available for the extension of stage-discharge rating curves at a flow-gauging site.

Given the sensitivity of design peak discharges to estimated ARFs and catchment response time parameter values, it is envisaged that the implementation of the specific objectives of this study will contribute fundamentally to the improved estimation of both ARFs and time parameters to ultimately result in improved design flood estimations in South Africa. The methodologies developed could also be adopted internationally to improve the estimation of ARFs and catchment response time parameters to provide more reliable peak discharge and volume estimates as, to date, this remains a constant challenge in flood hydrology.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AEP	Annual Exceedance Probability
AMS	Annual Maximum Series
ANN	Artificial Neural Network
ARF	Areal Reduction Factor
CCP	California Culvert Practice
CN	Curve Number
CPD	Continuous Professional Development
CSS	Cascading Style Sheets
DAWS	Department of Agriculture and Water Supply
DCR	Daily Catchment Rainfall
DDF	Depth-duration-frequency
DEM	Digital Elevation Model
DREU	Daily Rainfall Extraction Utility
DT	Discharge Rating Table
DWS	Department of Water and Sanitation
EA UK	Environmental Agency, United Kingdom
EMA	Expectation Maximisation Algorithm
ESRI	Environmental Systems Research Institute
EV1	Extreme Value Type I
EX-HYD	Flood Hydrograph Extraction Software
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
HAT	Hydrograph Analysis Tool
HRU	Hydrological Research Unit
HTML	Hypertext Mark-up Language
IDE	Integrated Development Environment
IDW	Inverse Distance Weighted
LAT	Latitude
LM	Linear Moments
LN	Log-Normal

LONG	Longitude
LP3$_{MM}$	Log-Pearson Type III using Method of Moments
MacOS	Macintosh Operating System
MAP	Mean Annual Precipitation
MCR	Maximum Catchment Rainfall
MI	Monthly Infilling
MM	Method of Moments
MR	Median Ratio
MSR	Maximum Station Rainfall
NERC	Natural Environment Research Council
NFSP	National Flood Studies Programme
NRCS	Natural Resources Conservation Service
NSCM	National Soil Conservation Manual
PDS	Partial Duration Series
PWM	Probability Weighted Moments
QGIS	Quantum Geographical Information System
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
SAWB	South African Weather Bureau
SAWS	South African Weather Services
SCS	Soil Conservation Service
SRTM	Shuttle Radar Topography Mission
SUH	Synthetic Unit Hydrograph
SVM	Support Vector Machine
T_C	Time of concentration
T_L	Lag time
T_P	Time to peak
TR	Technical Report
UH	Unit Hydrograph
UK	United Kingdom
UK FSR	United Kingdom Flood Studies Report

USA	United States of America
USACE	United States Army Corps of Engineers
USBR	United States Bureau of Reclamation
USDA	United States Department of Agriculture
USGS	United States Geological Survey
USWB	United States Weather Bureau
VBA	Visual Basic Application
VSC	Visual Studio Code
WinXSPRO	Windows Cross-Section Professional
WGS84	World Geodetic System 1984
WRC	Water Research Commission

CHAPTER 1: INTRODUCTION

This chapter provides some background on the estimation of areal design rainfall and catchment response time parameters and the influence these input parameters have on the estimation of design floods in ungauged catchments. This chapter includes the rationale, problem statement, research objectives and the outline of the report 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 South Africa, three basic approaches to design flood estimation are available, e.g. probabilistic, deterministic and empirical methods (Smithers, 2012; Van der Spuy and Rademeyer, 2018). In gauged catchments, despite uncertainties and errors in measurement, observed peak discharges are regarded as the best estimate of the true peak discharge (Gericke and Smithers, 2016b). In terms of design flood estimation in gauged catchments, probabilistic methods are normally used to conduct a frequency analysis of observed flood peak data from a flow-gauging site that are adequate in both length and quality of data (Smithers, 2012). In ungauged catchments, practitioners are required to estimate design floods using either event-based deterministic and/or empirical methods, although, regional probabilistic methods and/or continuous simulation models could also be used to transfer design values from gauged to ungauged sites.

Event-based deterministic design flood estimation methods are the most commonly used by practitioners in ungauged catchments (Van Vuuren et al., 2012). In the application of these event-based deterministic methods, it is acknowledged that both the spatial and temporal distribution of runoff, as well as the critical duration of rainfall, are influenced by the catchment response time. Typically, all complex, heterogeneous catchment processes are lumped into a single process to enable the estimation of the expected output (design flood) from causative input (average areal design rainfall and catchment response time) (Gericke and Du Plessis, 2013; Gericke, 2018).

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 digitised 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 often used to estimate design rainfall. In essence, design rainfall is derived from observed rainfall data and comprises of a depth and duration (which is directly proportional to the catchment response time) 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 ranges between one and seven days and can be estimated from both continuously recorded and daily rainfall data (Smithers and Schulze, 2004). However, design point rainfall estimates are only applicable to a limited area and for larger areas, the average areal design rainfall depth is likely to be less than the maximum design point rainfall depths (Siriwardena and Weinmann, 1996). 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 (response time) and catchment area (Alexander, 2001).

ARFs could be estimated using either analytical or empirical methods (Pietersen et al., 2015). 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 could be classified as either geographically-centred or storm-centred. The geographically-centred approach describes the relationship between average areal design rainfall over a geographically fixed area and a corresponding 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).

The hydrological response of a catchment, i.e. the catchment response time, is normally expressed as a single time parameter, e.g. time of concentration (T_C), lag time (T_L) and/or time to peak (T_P). In other words, when event-based deterministic design flood estimation methods are applied in ungauged catchments, estimates of the peak discharge are based on a single, representative catchment response time parameter (e.g. T_P, T_C and/or T_L), while the catchment is at an 'average condition' and the hazard or risk associated with a specific event is reflected by the joint-probability of the 1: T-year average areal design rainfall and 1: T-year design flood (Rahman et al., 2002; SANRAL, 2013).

The time of concentration (T_C) is not only the most frequently used and required time parameter in event-based methods (SANRAL, 2013; Gericke and Smithers, 2014), but also continues to find application in continuous simulation models (USACE, 2001; Neitsch et al., 2005; Smithers et al., 2013). More specifically, T_C is primarily used to estimate the critical storm duration of a specific design rainfall event used as input to deterministic methods, i.e. the Rational and Standard Design Flood (SDF) methods, while T_L is used as input to the deterministic Soil Conservation Services (SCS) and Synthetic Unit Hydrograph (SUH) methods. The T_P is normally expressed as a function of the critical storm duration and T_L (Mockus, 1957).

Time parameters such as T_C, T_L and T_P serve as indicators of both the catchment storage and the effect thereof on the temporal distribution of runoff. The catchment response time is also directly related to, and influenced by, climatological variables (e.g. meteorology and hydrology), catchment geomorphology, catchment variables (e.g. land cover, soils and storage), and channel geomorphology (Schmidt and Schulze, 1984; Royappen et al., 2002; McCuen, 2005). Amiri et al. (2019a; 2019b) demonstrated that catchment response times, especially T_C and T_L, could be explicitly estimated at a catchment level by applying the average values of typical catchment characteristics, e.g. circumscribing circle, fractal dimension, perimeter-area ratio, and shape indices for the landscape categories and hydrological soil groups.

In medium to large catchments where channel flow in main watercourses generally dominates catchment response time, the estimation of T_C in South Africa is currently based on the length of the longest main watercourse (L_{CH}) and the average main watercourse slope (S_{CH}) as primary catchment descriptors. Typically, catchment descriptors such as the hydraulic length (L_H), centroid distance (L_C), average catchment slope (S), runoff curve numbers (CN) and S_{CH} are used as input to estimate T_L. McCuen (2009) highlighted that, due to differences in the roughness and slope of catchments (overland flow) and main watercourses (channel flow), T_C estimates, based on only the main watercourse characteristics (L_{CH} and S_{CH}), could be underestimated on average by 50%. Consequently, the resulting peak discharges will be overestimated by between 30% and 50% (McCuen, 2009).

Hence, average areal design rainfall and catchment response time are therefore regarded as fundamental input to all event-based design flood estimation methods in ungauged catchments; while errors in estimated average areal design rainfall and catchment response time, will directly impact on 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) 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). 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 period, but the rate of decrease for convective rainfall is noticeably larger than that for 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 type. 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. In catchment areas up to 30 000 km², ARFs are mainly a function of the catchment area and storm duration (Alexander, 2001; SANRAL, 2013).

Internationally, extensive national-scale ARF studies are limited to the United Kingdom (UK; NERC, 1975), USA (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, often fail to successfully incorporate the variation in predominant weather 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 resolutions 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 latter 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 wrongfully 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's assumed that the AEP of both the point and areal rainfall is similar. 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) have conclusively shown that ARFs are dependent on the average AEP of rainfall. According to Pavlovic et al. (2016), the differences between analytical and empirical ARF estimation methods currently in use are also more pronounced for shorter storm durations and larger catchment areas, while being partially dependent on the average return period. Thus, most of these methods are inappropriate to use at a comprehensive set of temporal and spatial scales in larger catchments. Two recognised design point rainfall databases, 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. In order to overcome the limitations of design point rainfall, the Department of Water and Sanitation (DWS) developed an approach 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), this approach eliminates the required use of ARFs; however, there are some limitations associated with this approach: (i) catchment specific, and (ii) no established areal design rainfall database is available. Hence, ARFs still need to be applied to the design point rainfall values obtained from the TR102 and/or RLMA&SI databases.

Based on the shortcomings highlighted above, it is clearly evident that the estimation of ARFs is internationally an on-going research question, particularly in South Africa. Hence, the ARFs in South Africa need to be re-investigated in the light of recent extreme flood events utilising the longer periods of record (± 50 years of additional data since the 1970s) which are now available for analysis. The variation of ARFs with catchment area, return period, duration and rainfall producing mechanisms also needs to be investigated by adopting a regional approach.

In considering observed rainfall and runoff data in gauged catchments, time parameters are normally defined by the difference between two interrelated observed time variables (McCuen, 2009), which represent individual events on either a hyetograph or hydrograph as illustrated in Figure 1.1.
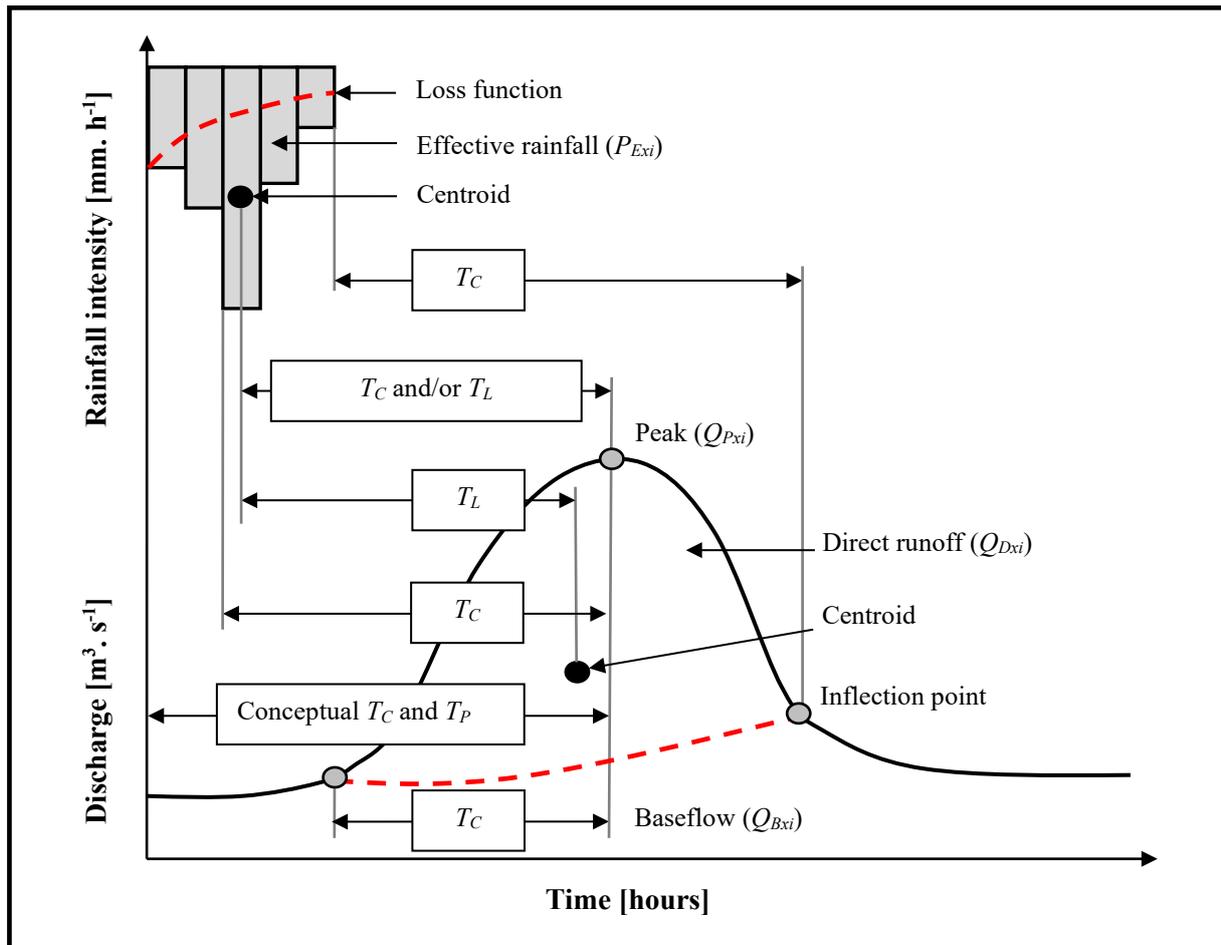
Figure 1.1: Schematic illustrative of the different time parameter relationships (after Gericke and Smithers, 2014)

In small catchment areas (A) up to 20 km², the difference between two interrelated observed time variables is estimated using a simplified convolution process between a single rainfall hyetograph and resulting single-peaked hydrograph as shown in Figure 1.1. In medium to large heterogeneous catchment areas, typically ranging from 20 km² to 30 000 km², a similar convolution process is required where the temporal relationship between a catchment hyetograph, which may be derived from numerous rainfall stations, and the resulting outflow hydrograph, is established (Gericke and Smithers, 2014).

However, several problems are associated with such a simplified convolution procedure at medium to large catchment scales. Conceptually, such a procedure normally assumes that the volume of direct runoff is equal to the volume of effective rainfall, and that all rainfall prior to the start of direct runoff is regarded as initial abstraction, after which the loss rate is assumed to be constant (McCuen, 2005). Therefore, a uniform response to rainfall within a catchment

is assumed, while the spatially non-uniform antecedent soil moisture conditions within the catchment, which are a consequence of both the spatially non-uniform rainfall and the heterogeneous nature of soils and land cover in the catchment, are ignored. Consequently, in contrast to small catchments with single-peaked hydrographs, the variability evident in medium to large catchments typically results in multi-peaked hydrographs.

Furthermore, the use of rainfall data to estimate catchment hyetographs at a medium to large catchment scale, also poses several additional problems as a consequence of the following (Schmidt and Schulze, 1984; Gericke and Smithers, 2014): (i) paucity of rainfall data at sub-daily timescales, both in the number of rainfall gauges and length of the recorded series, (ii) poor time synchronisation between point rainfall data sets from different gauges, (iii) difficulties in measuring time parameters for individual events directly from digitised autographic records owing to difficulties in determining the start time, end time and temporal and spatial distribution of effective rainfall over the catchment, and (iv) poor time synchronised rainfall and streamflow recorders.

In addition to the above-mentioned problems related to a simplified convolution process at medium to large catchment scales, the number of hydrometeorological monitoring stations, especially rainfall stations in South Africa and around the world, has declined steadily over the last few decades. According to Lorenz and Kunstmann (2012), the number of rainfall stations across Europe, declined by nearly 50% between 1989 and 2006, i.e. from 10 000 to less than 6 000 stations, whilst a far more rapid decline occurred in South America, i.e. the nearly 4 300 rainfall stations has reduced to 400. Internationally, the USA has witnessed one of the slowest declines, while large parts of Africa and Asia remain without a single rainfall station (Lorenz and Kunstmann, 2012). South Africa is no exception and the rainfall monitoring network has declined over recent years with the number of stations reducing from more than 2 000 in the 1970s to the current situation where the network is no better than it was as far back as 1920 with less than a 1 000 useful stations open in a specific year (Pitman, 2011). Balme et al. (2006) also showed that a decline in the density of a rainfall monitoring network produces a significant increase in the errors of spatial estimation of rainfall at annual scales and even larger errors at event scales for large catchments. In contrast to rainfall data, streamflow data are generally less readily available internationally, but the data quantity and quality enable it to be used directly to estimate catchment response times at medium to large catchment scales. In South Africa for

example, there are 708 flow-gauging station sites with more than 20 years of record available (Smithers et al., 2014).

In ungauged catchments, catchment response time parameters are estimated using either empirically or hydraulically-based methods, although analytical or semi-analytical methods are also sometimes used (McCuen et al., 1984; McCuen, 2009). Empirical methods are the most frequently used by practitioners to estimate the catchment response time and almost 95% of all the methods developed internationally are empirically-based (Gericke and Smithers, 2014). However, the majority of these methods are applicable to and calibrated for small catchments, with only the research of Thomas et al. (2000) applicable to medium catchment areas of up to 1 280 km² and the research of Johnstone and Cross (1949), Pullen (1969), Mimikou (1984), Watt and Chow (1985), and Sabol (2008) focusing on larger catchments of up to 5 000 km².

In South Africa, unfortunately, none of the empirical T_C estimation methods recommended for general use were developed and verified using local data. In small, flat catchments with overland flow being dominant, the use of the Kerby equation (Kerby, 1959) is recommended, while the empirical United States Bureau of Reclamation (USBR) equation (USBR, 1973) is used to estimate T_C as channel flow in a defined watercourse (SANRAL, 2013). Both the Kerby and USBR equations were developed and calibrated in the USA for catchment areas less than 4 ha and 45 ha, respectively (McCuen et al., 1984). Consequently, practitioners in South Africa commonly apply these 'recommended methods' outside their bounds, both in terms of areal extent and their original developmental regions, without using any local correction factors.

The empirical estimates of T_L used in South Africa are limited to the family of equations developed by the Hydrological Research Unit (HRU; Pullen, 1969); the United States Department of Agriculture Natural Resource Conservation Service (USDA NRCS), formerly known as the USDA Soil Conservation Service (USDA SCS, 1985) and SCS-SA (Schmidt and Schulze, 1984) equations. Both the HRU and Schmidt-Schulze T_L equations were locally developed and verified. However, the use of the HRU methodology is recommended for catchment areas up to 5 000 km², while the Schmidt-Schulze (SCS-SA) methodology is limited to small catchments (up to 30 km²).

The simultaneous use of different time parameter definitions as proposed in literature and the inherent procedural limitations of the traditional simplified convolution process when applied in medium to large catchments, combined with the lack of both continuously recorded rainfall

data and available direct measurements of rainfall and runoff relationships at these catchment scales, has not only curtailed the establishment of unbiased time parameter estimation procedures in South Africa, but also has had a direct impact on design flood estimation. Despite the widespread use of all these time parameters, unique working definitions for each of the parameters are not currently available. Frequently, there is no distinction between these time parameters in the hydrological literature; hence, the question whether they are true hydraulic or hydrograph time parameters, remains unrequited, while some methods as a consequence, are presented in a disparate form. However, the use of several conceptual and computational time parameter definitions are proposed in the literature, as summarised by McCuen (2009) and Gericke and Smithers (2014), some of which are adopted in practice.

Bondelid et al. (1982) indicated that as much as 75% of the total error in peak discharge estimates could be ascribed to errors in the estimation of time parameters. Gericke and Smithers (2014; 2016a) not only demonstrated the inconsistency amongst various time parameter equations applied at a medium to large catchment scale, but also showed that the underestimation of time parameters by 80% or more could result in the overestimation of peak discharges of up to 200%, while the overestimation of time parameters beyond 700% could result in maximum peak discharge underestimations of up to 100%. Consequently, Gericke and Smithers (2016b; 2017) developed a new approach to estimate observed catchment response times using only observed streamflow data to ultimately calibrate and verify empirical time parameter equations in a pilot scale study in four climatologically different regions of South Africa.

However, the biggest limitation of these empirical time parameter equations, in general, is their tendency to provide inconsistent results when applied outside the bounds of their original developmental regions without the use of local correction factors. Therefore, in order to overcome these constraints, the widely used approach of regionalisation in flood hydrology should be adopted. It is therefore necessary to adopt and/or develop a regionalisation scheme for catchment response time estimation in South Africa.

Given the sensitivity of design peak discharges to estimated ARFs and time parameter values as highlighted above, ARFs and catchment response time at a medium to large catchment scale were also identified as potential research projects 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 is supported in principal by the WRC (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 and catchment response time were regarded as high priority research needs.

Consequently, this not only served as a motivation for this project, but also emphasised that the continued use of such inappropriate ARF and time parameter estimation methods in South Africa, not only limits the use of both event-based design flood estimation methods and advanced modelling systems (despite the use of other technologically advanced input parameters in these methods/models), but it also has an indirect impact on hydraulic designs, i.e. underestimated time parameters and associated lower design rainfall depths, although of much higher intensities, would result in over-designed hydraulic structures, while overestimated time parameters would result in under-designs. Not only will hydraulic structures be over- or under-designed, but associated socio-economic implications might render some projects as not being feasible, while any loss of life due excessive flood damages and insufficient infrastructure, is not excluded.

1.3 Research Objectives

The overall objective of this project is to develop a regionalised approach to estimate ARFs and at-site catchment response time parameters for improved design flood estimation in South Africa. The pilot scale studies of Pietersen (2016) and Gericke and Smithers (2016b; 2017; 2018) were used as guiding references, respectively. In the ARF study (Pietersen, 2016), it was recommended that the current methodology 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 areal design rainfall at a national-scale, but it would also result in improved design flood estimation. The same benefits of regionalisation apply to catchment response time parameters; however, given the complexities involved, comprehensive catchment response time analyses beyond the boundaries of the pilot study area, Primary Drainage Region X, were not possible within the project timeframe and budget.

1.3.1 Research aim: Areal Reduction Factors

The primary research aim is to estimate geographically-centred 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 were used, since Pietersen (2016) demonstrated that ARF estimates are influenced by the differences in the catchment shape, orientation and size. Furthermore, the analysis at a catchment level also limits the potential extrapolation of ARFs beyond the catchment boundaries. Therefore, the use of multiple circular catchments with random sizes, determined by the location of rainfall stations, covering a specific homogeneous region (according to the RLMA&SI regionalisation scheme) will enable the estimation of sample ARFs.

The focus is on the development of probabilistically correct sample ARFs, *viz.* the relationships between *T*-year areal rainfall estimates and weighted average *T*-year point rainfall estimates influenced by: (i) catchment area (*A*), (ii) duration (*D*), and (iii) return period (*T*) values. Consequently, this will elucidate how ARF values vary with catchment area, storm duration (rainfall type), and return period throughout South Africa.

In terms of ARFs, the primary research aim is based on the following assumptions:

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

(b) **Assumption 2:** ARFs vary with predominant weather types, storm durations, seasonal factors and return periods.

(c) **Assumption 3:** The current South African ARF estimation methods are only applicable to specific temporal and spatial scales.

1.3.2 Research aim: Catchment response time

The primary aim is to expand and verify the approach developed by Gericke and Smithers (2017) by estimating observed catchment response time parameters from the 51 gauged catchments located in Primary Drainage Region X and to derive a regional empirical time parameter equation. The scope of the study is limited to the latter region, since the flow-gauging stations located in Primary Drainage Region X generally had better and more complete data sets for which the DWS has done some stage-discharge extrapolations.

The approximation of $T_C \approx T_P$ as proposed by Gericke and Smithers (2014; 2017) forms the basis of the approach adopted and is based on the definition that the volume of effective rainfall equals the volume of direct runoff when a hydrograph is separated into direct runoff and baseflow. The separation point on the hydrograph is regarded as the start of direct runoff which coincides with the onset of effective rainfall. In other words, the required extensive convolution process normally required to estimate time parameters (e.g. T_C, T_L and/or T_P) is eliminated, since the time parameters are estimated directly from the observed streamflow data without the need for rainfall data.

The study focuses on the development of semi-automated routines using Microsoft Excel or Visual Basic for Applications (VBA) scripting to extract complete hydrographs to support the future development, testing and verification of a hydrograph extraction utility at catchment level. In other words, the comprehensive results obtained from this study will be used as benchmark to inform the envisaged development, testing and verification of such a software utility.

In terms of catchment response time, the primary research aim is based on the following assumptions:

(a) **Assumption 1:** There is no 'one size fits all' approach/method available for the extension of stage-discharge rating curves at a flow-gauging site, since the extension of a rating curve is significantly more affected by the site characteristics and initial hydraulic conditions, than the actual indirect extension method used.

(b) **Assumption 2:** The time to peak (T_P) equals the total net rise (duration) of a multiple-peaked hydrograph in medium to large catchments.

13

(c) **Assumption 3:** The error bounds between the three different approaches (e.g. net rise duration, triangular-shaped hydrograph approximation and linear response function) to estimate catchment response time parameters from observed streamflow data are within acceptable limits ($\leq 15\%$).

(d) **Assumption 4:** Time parameters proportionality ratios equal unity in medium to large catchments, i.e. $T_C \approx T_L \approx T_P$.

1.3.3 Specific objectives

The specific objectives of the project are to:

(a) Contribute to the establishment of a national catchment variable database;

(b) Extract and analyse rainfall and runoff data to provide geographically-centred and probabilistically correct ARFs and observed catchment response time parameters, respectively;

(c) Refine and update the regionalisation scheme for ARFs in South Africa;

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

(e) Derive a regional empirical time parameter equation for application in Primary Drainage Region X; and

(f) Assess and verify the derived ARF and time parameter equations.

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

1.4 Outline of Report Structure

The estimation of ARFs and catchment response time parameters and the influence thereof on estimates of peak discharge are central to all chapters.

Chapter 2 presents a comprehensive literature review of the different ARF estimation methods used to describe the relationship between point and areal rainfall, while a comprehensive literature review of catchment response time estimation methods is included in Chapter 3.

Chapter 4 presents an overview of the concerns and/or problems encountered, and possible solutions related to the derivation of ARFs. Similarly, Chapter 5 presents an overview of the concerns related to the estimation of catchment response time parameters. In addition, possible solutions are also discussed in this chapter.

Chapter 6 presents the final ARF methodology and results, while the final catchment response time methodology and results are presented in Chapter 7. Chapter 8 presents a synthesis of all the information as discussed in Chapters 6 to 7, as well as some final conclusions.

Chapter 9 contains a summary of: (i) the equipment and resources used, (ii) the project deliverables, (iii) knowledge dissemination, and (iv) the project work plan and achieved milestones.

Chapter 10 contains the list of all the references used.

CHAPTER 2: LITERATURE REVIEW – AREAL REDUCTION FACTORS

The literature review contained in this chapter mainly 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 current status of observed rainfall measurement in South Africa and the associated infilling and averaging techniques commonly applied. Thereafter, design rainfall estimation is detailed. The remaining part of the chapter focusses on ARFs.

2.1   Climate and Rainfall Types

The 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 2.1. 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 2.1.

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, while mesic-subtropical conditions dominate on the KwaZulu-Natal coast of the Escarpment region (Davies and Day, 1998; Alexander, 2010). The mean annual precipitation (MAP) decreases, while potential evaporation increases westwards and northwards across South Africa. The overall MAP is 452 mm, but in many parts of the country, the MAP is much less. 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).
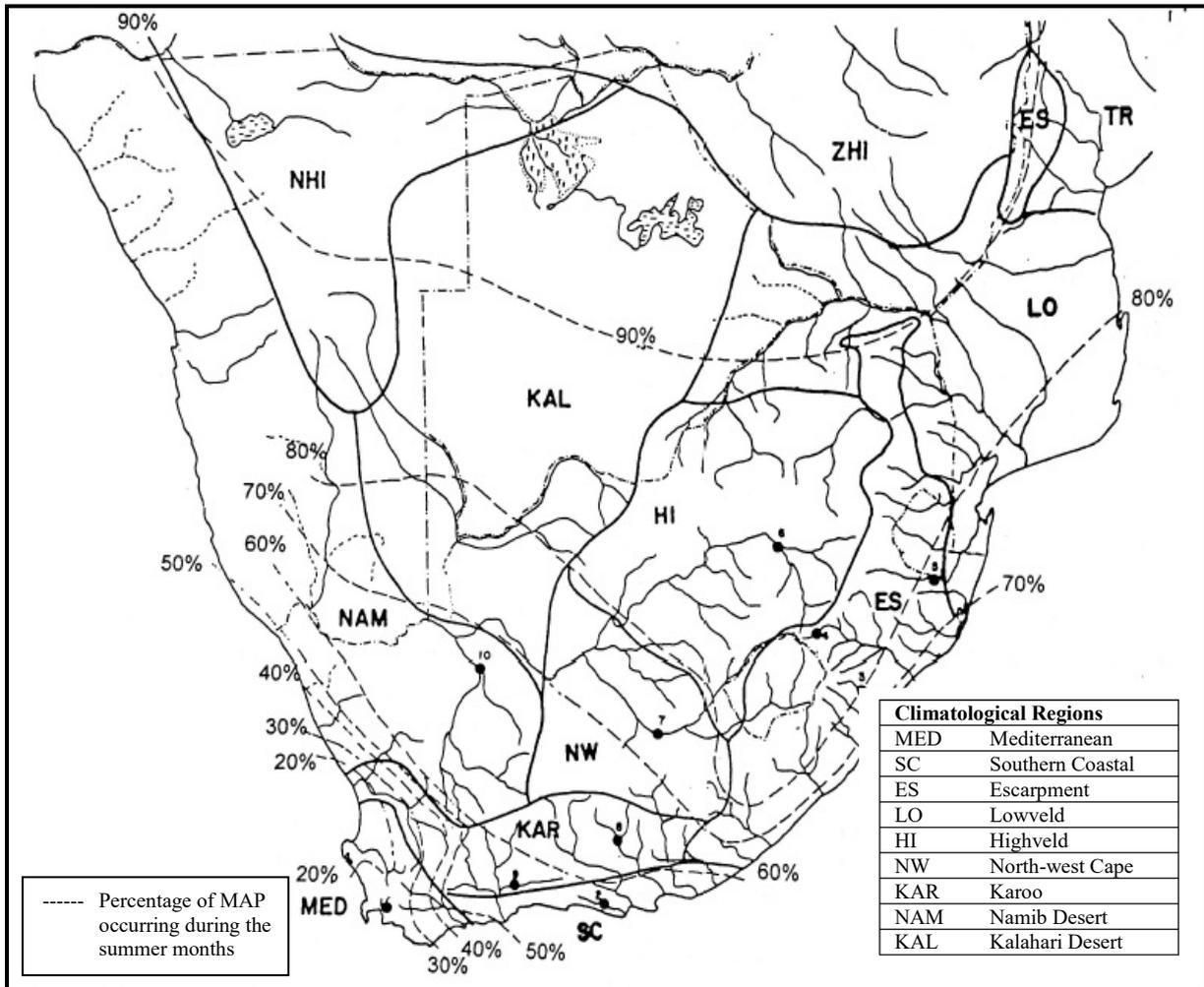
Figure 2.1: Climatological regions for South Africa (Alexander, 2010)

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 (Davies and Day, 1998). In winter, the prevailing north-westerly winds result in high rainfall in the western part 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 (Davies and Day, 1998).

Bárdossy and Pegram (2018) highlighted that a 10-year return period rainfall event across the city of London (UK), will be greater than the same return period event occurring somewhere in Munich, due to their differences in area, i.e. the greater the area, the higher the probability of occurrence of a storm of a given magnitude. Climate does not only affect rainfall distribution, but also rainfall intensity, duration and variability, which are all interdependent.

17

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 in size, allowing moist air to be drawn into the cyclone vortex and allowing mist to be lifted up 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 up 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 to result in condensation and rainfall. The rainfall intensity is normally regarded as moderate and dependent on wind blowing towards the inland areas. Orographic rainfall is characteristic of the coast lines of KwaZulu-Natal and the Western Cape provinces.

The rainfall types listed in (a) to (d) were carefully considered to highlight and describe the direct influence thereof on the estimation of ARFs. The magnitude of ARFs is highly dependent on the different storm mechanisms associated with different rainfall types. In a specific region

with more frequent thunderstorms (convective rainfall) occurring than frontal storms (wide spread 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 the actual rainfall process in that particular 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 rainfall intensity distribution regions. 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. The spatial distribution of these regions can be summarised as follows: (i) Region 1: Eastern Cape, e.g. East London and Port Elizabeth, (ii) Region 2: Western Cape (Karoo) and Free State, (iii) Region 3: Northern Cape, e.g. Upington and Kimberley, as well the Highveld, including Gauteng and Mpumalanga, and (iv) Region 4: The remainder of the country.

2.2   Observed Rainfall Data in South Africa

Observed rainfall data in South Africa can be obtained from daily rainfall stations, which are widespread in space, and are measured as a depth (mm) at a specific time interval on each day. The poor maintenance of rainfall stations in South Africa, under the supervision 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 in the near future (Smithers and Schulze, 2000b; Van Vuuren et al., 2012).

A total of 11 171 daily rainfall record lengths (long duration) are available in the South African database (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).
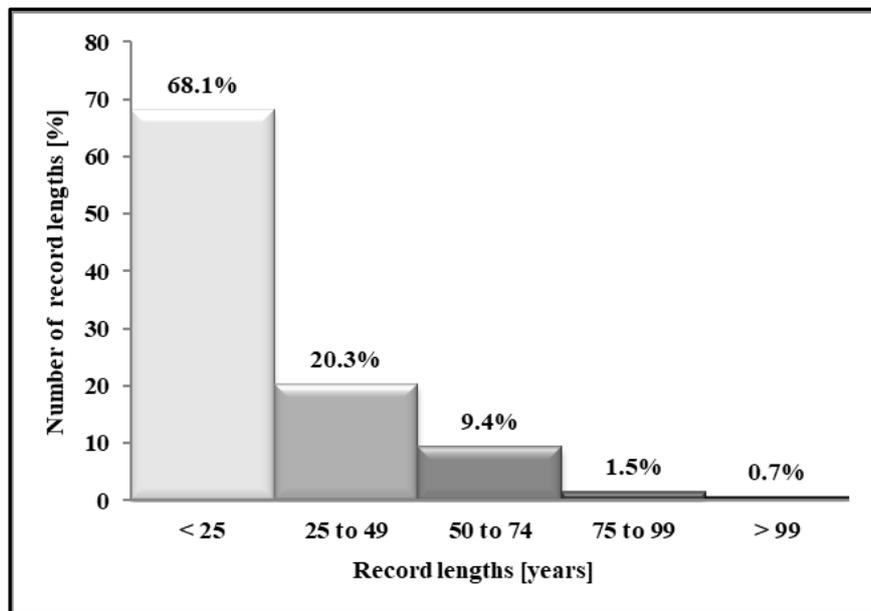
Figure 2.2: Available record lengths for daily rainfall stations in South Africa (after Smithers and Schulze, 2000b)

Short duration rainfall data (less than 24 hours) in South Africa are currently available from 412 stations as shown in Figure 2.3. However, only 49 of these 412 rainfall stations have 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 stations (Smithers and Schulze, 2000a).

Furthermore, Smithers and Schulze (2000a) also highlighted that short duration rainfall data have a low reliability due to several possible errors including missing data and differences (more than 20 mm) between the digitised and standard rain gauge daily totals. It was also noted that the digitised SAWS data are inadequate for the estimation of design storm durations of less than 24 hours. Smithers and Schulze (2000a) developed three approaches based on regional similarities, scaling properties and stochastic simulation of extreme rainfall events to estimate short duration design rainfall values. This is discussed in the next section.
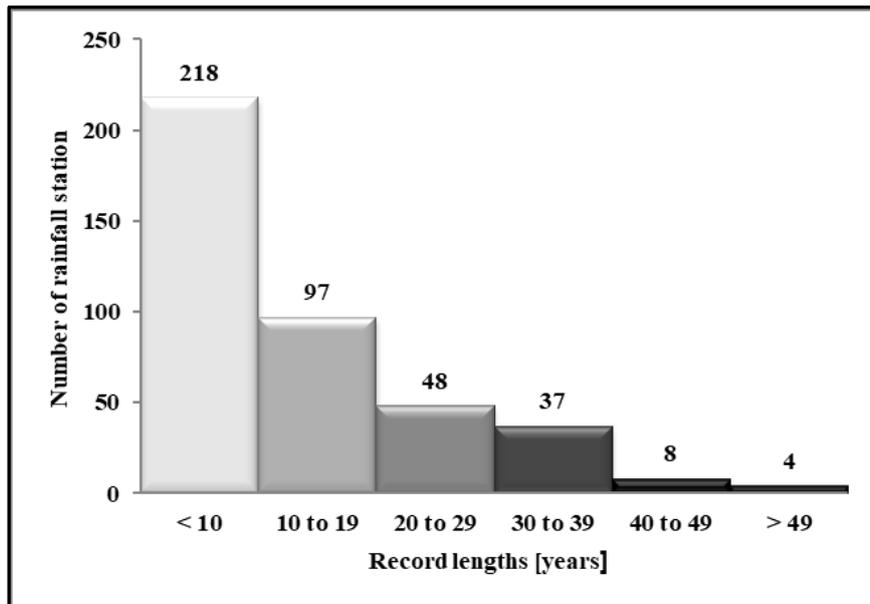
Figure 2.3: Short duration rainfall stations in South Africa (after Smithers and Schulze, 2000a)

2.3 Infilling of Missing Observed Rainfall Data

Rainfall records characterised by missing data are a serious concern when daily hydrometeorological simulation models 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) proposed a number of different infilling techniques based on a categorisation process and developed the Daily Rainfall Extraction Utility (DREU) to determine the best approach to infill any missing data at rainfall station(s). The DREU infilling procedure algorithms are based on one or a combination of the following techniques:

(a) **Inverse Distance Weighting (IDW):** The IDW technique 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 in

order 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 (1997a, 1997b; 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)   **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).

(d)   **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) by Dent (1989; cited by Lynch, 2004) was interrogated and the monthly infilled values of zero and/or ≤ 2 mm were extracted.

The EMA and MR techniques are considered to be the most effective infilling techniques in the DREU (Lynch, 2004). Any missing observed rainfall values not infilled by using the EMA and MR techniques are infilled using the IDW technique. Consequently, 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 has 341 908 152 values (Lynch, 2004).

2.4   Averaging of Observed Rainfall

In the assessment of total quantities of rainfall over large areas, the occurrence of storms and their contribution to single rainfall stations is unknown. Therefore, it is necessary to convert numerous observed point rainfall depths to provide an average rainfall depth over a certain area. The following methods may be used for averaging the 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 sufficient when rainfall stations are uniformly distributed, the topography is relatively flat and spatial variations in rainfall are insignificant.

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

(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.

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

(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 take into account local effects of climate and uneven topography.

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

(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).

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

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

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 areal 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:** The Spline 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 auto-correlation, such as the statistical relationships amongst point rainfall values. Kriging does not only have the capability of producing a rainfall prediction surface, but it also provides some measure of the certainty of the predictions. The variation in the rainfall surface can be explained by the distance or direction between the rainfall stations that present the correlation. 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 Rainfall Estimation

Design rainfall comprises of a depth and duration associated with a given return period 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 and based on daily point rainfall data include 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 in an attempt 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).

In order to develop the depth-duration-frequency (DDF) relationship at every single site, the following steps are of importance (Smithers and Schulze, 2000a):

Selection of the most appropriate data set. This may either be the AMS or partial duration series (PDS) with a sufficient record length;

   (a) Selection of the most appropriate probability distribution; and

   (b) 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 stations 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; cited by Smithers and Schulze, 2000a). This selection is particularly important when estimating extreme events with return periods greater than the record length. Equally important are that, factors such as the type of data in use, data stationarity and the method of fitting the distribution, should also be considered (Cunnane, 1989; cited by Smithers and Schulze, 2000a).

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) propose the use of the Log-Normal (LN), Log-Pearson Type III (LP3), as well as 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.

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 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 (Alexander, 2001; Cunnane, 1989; cited by Smithers and Schulze, 2003). An advantage of this approach is that it can be used to estimate events at ungauged sites where no rainfall data exists. 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 mean 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) established 78 homogeneous long duration rainfall clusters, 15 short duration rainfall clusters, and estimated index values (mean n-hour AMS values) derived from at-site data. Cluster analysis of site characteristics was used to group the 78 long duration rainfall clusters into seven (7) regions with six (6) associated region-specific regression parameters. Firstly, the mean of the 1-day fixed time interval point rainfall AMS was estimated using regional regression relationships. Thereafter, the mean 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 mean of the point rainfall AMS values for durations shorter and longer than 1 day were scaled directly from the mean 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 mean 1-day and continuous 24-hour values, respectively. In the application of the regression relationships to estimate the mean 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.

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. In order to convert the daily design rainfall depth values to independent durations of the same length, conversion and/or scaling factors have to 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 (08h00 to 08h00) 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). In order 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 considered to be 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).

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

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

2.5.3   Other approaches

Apart from the approaches discussed in Sections 2.5.1 and 2.5.2, 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 is 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.

(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.

(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 proposed 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.

2.6 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.

2.6.1 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 air mass storms.

32

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 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) recommends 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 emphasises 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 station network in a given catchment, is that it will not necessarily reflect the actual spatial characteristics of a particular rainfall event.

2.6.2   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 km², 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 catchments with areas of up to 30 000 km², 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 km² from the urban storm centre compared to 67 rural storms. Veneziano and Langousis (2005) concluded that the catchment shape normally has an insignificant effect on ARFs. However, 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.

Bárdossy and Pegram (2018) recently 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. 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; consequently, resulting in poorer areal rainfall estimates. Nevertheless, Allen and DeGaetano (2005) found that topographical rainfall biases appear to be insignificant for the estimation of ARFs.

2.6.3   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.

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 on the catchment's surface 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.

2.7 ARF Estimation Methods

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

2.7.1 Analytical methods

In using analytical methods, derived mathematical relationships 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 a cause for concern. 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).

2.7.2 Empirical methods

Empirical methods can either be based on a geographically-centred or storm-centred approach. Both these approaches are briefly explained below; however, this project (WRC Report K5-2924) focussed only on geographically-centred ARF methods, whereas, WRC Report K5-2923 (Du Plessis et al., 2020) focussed on storm-centred ARF methods.

The **geographically-centred approach** describes the relationship between average areal design rainfall over a geographically fixed area, i.e. catchment, and a corresponding 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 point and areal design rainfall and considers the uniform temporal and spatial distribution of rainfall over the catchment area (Pietersen et al., 2015).

In the **storm-centred 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).

Methods such as the United States Weather Bureau (USWB) method (USWB, 1957; 1958), UK FSR method (NERC, 1975) and Bell's (1976) method are typical examples of empirical methods. The latter method has proved to offer more probabilistically correct ARFs compared to the other methods, since the AEP is incorporated. Each empirical method has different data requirements and would subsequently result in different ARF estimates.

The details of the geographically-centred ARF methods currently used in South Africa and relevant to this project, are summarised below:

(a) **Alexander (1980; 2001):** This is the geographically-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 (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. Estimates of shorter duration rainfall based on extrapolation from longer durations are unreliable when viewed in the light of the storm mechanisms which produce high-intensity rainfall for durations less than 10 minutes (Alexander, 1980). 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.

36
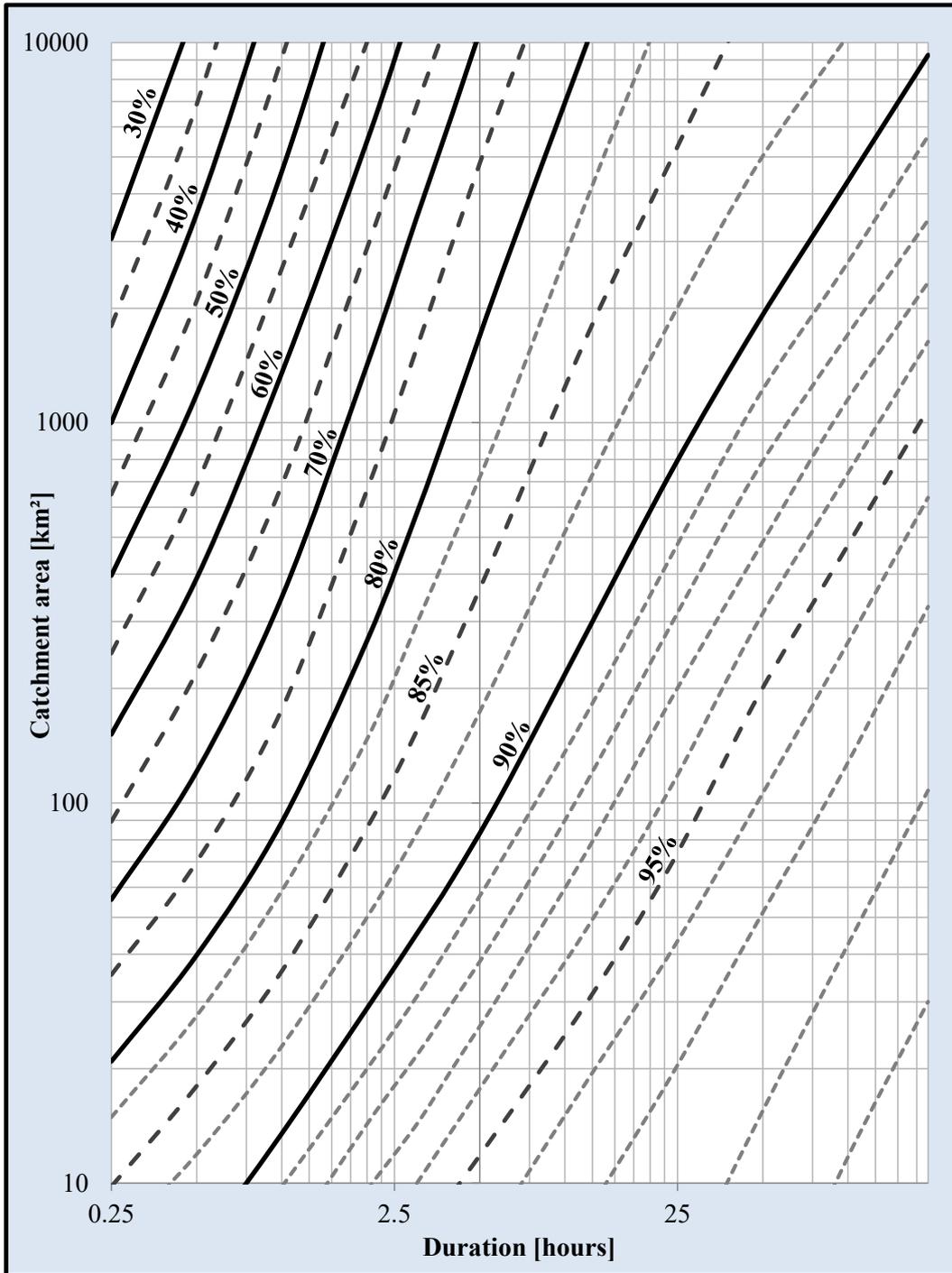
Figure 2.4: UK FSR ARF diagram (NERC, 1975)
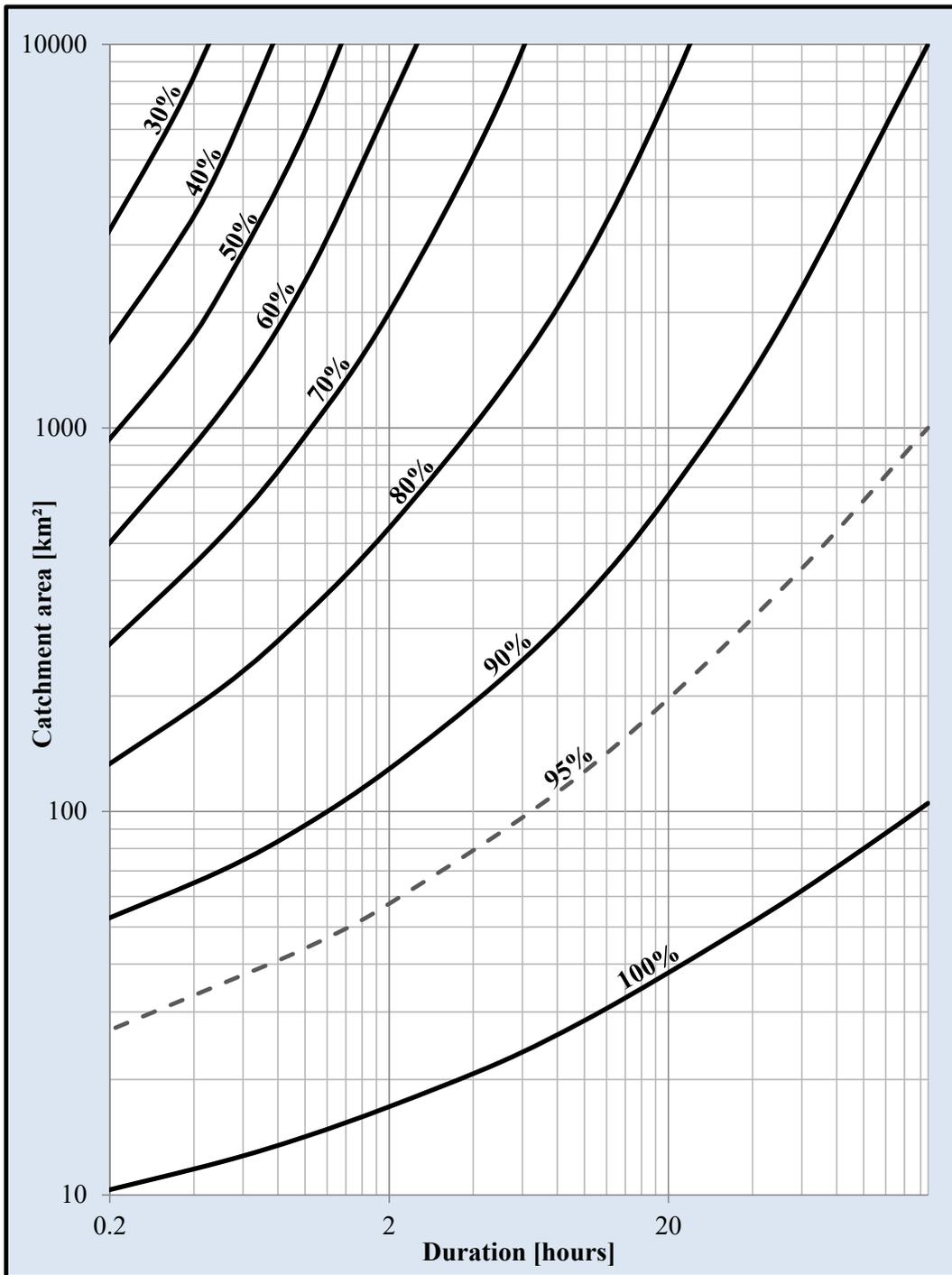
Figure 2.5: Adopted UK FSR ARFs for South Africa (after Alexander, 1980)

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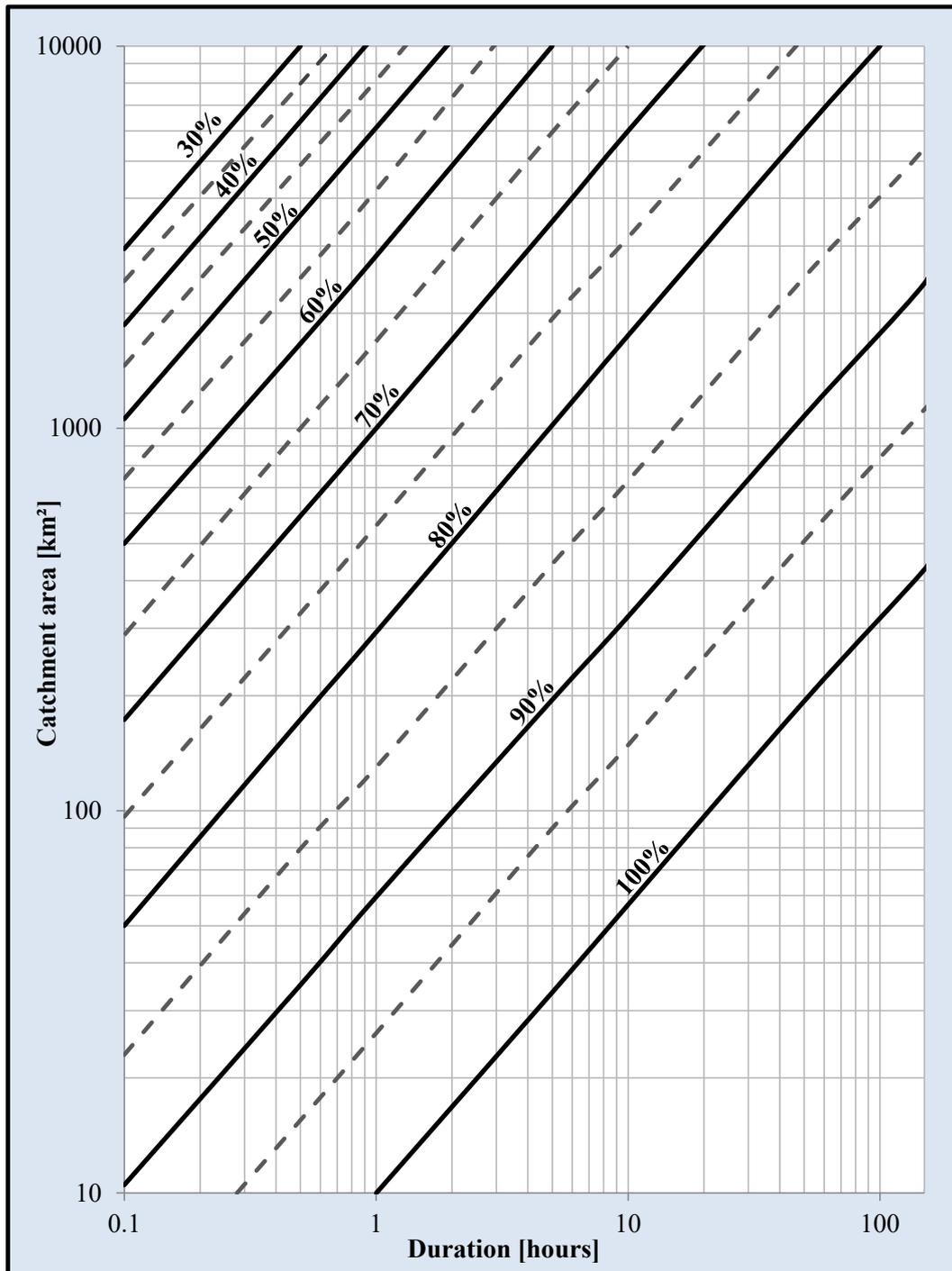
Figure 2.6: Revised ARF diagram for South Africa (Alexander, 2001)

Op ten Noort and Stephenson (1982) converted Figure 2.5 to the mathematical relationship [Eq. (2.5a)], using regression analysis. Alexander (2001) expressed Figure 2.6 in the form of a mathematical relationship as shown in Eq. (2.5b). The use of both Eq. (2.5a) and (2.5b) 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. (2.5b) should be used for Southern African conditions where the average rainfall depth over a catchment has to be established from point rainfall statistics.

$$ARF \quad = [1.306 - 0.0902Ln(A)] + Ln(T_d)[0.0161Ln(A) - 0.0498] \qquad [2.5a]$$

$$ARF \quad = [90000 - 12800Ln(A) + 9830Ln(60T_C)]^{0.4} \qquad [2.5b]$$

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 exists between A, T_C and ARFs. The validity of Eq. (2.5b) 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. (2.6a), which provides a good indication of T_C associated with any catchment area under consideration. Equation (2.6b) resulted from the substitution and simplification of Eq. (2.6a) into Eq. (2.5b).

$$T_C \quad = 0.2284A^{0.596} \qquad [2.6a]$$

$$ARF \quad = [-6944.3Ln(A) + 115731.9]^{0.4} \qquad [2.6b]$$

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. (2.6b) can be used instead of Eq. (2.5b) to convert design point rainfall depths or intensities to an average areal design rainfall depth or intensity.

(b) **Bell (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 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. (2.7).

$$ARF_m = \frac{\sum_{i=1}^{N}\left(w_i\overline{\overline{P_{ij}}}\right)_m}{\sum_{i=1}^{N}\left(w_i P_{ij}\right)_m} \qquad [2.7]$$

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 and UK FSR ARF estimation methods, but the variation of ARFs with return period is also clearly evident when ARFs are directly obtained from the areal and design point rainfall frequency curves. In most of the other ARF estimation methods, e.g. USWB and UK FSR, the variation of ARFs are largely obscured by the regionalisation of data.

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 nowadays available.

(c) **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 the catchment area, storm duration and AEP. Artificial circular catchments in areas with sufficient data were used to generate the areal rainfall time series. This was done by weighting point rainfall values by means of Thiessen polygon areas. Frequency quantiles were

estimated from the areal rainfall time series and were divided by the weighted point frequency quantiles. This process was repeated and subsequently resulted in a number of 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., 2016). 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 in order for the user to select an equation to estimate the corresponding ARFs. The storm duration ranges with corresponding ARF equations are: (i) duration ≥ 12-hour [Eq. (2.8a)], (ii) duration > 24-hour [Eqs. (2.8b) and (2.8d)], and (iii) 24-hour ≤ duration ≤ 168-hour [Eqs. (2.8c) and (2.8d)]. 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., 2016).

Short duration (i), A (1-1 000 km^2):

$$ARF_S = \text{Min}\left\{1, \begin{bmatrix} 1\text{-}0.287\left(A^{0.267} - 0.439\log_{10} D\right)D^{-0.36} + \\ 0.0023A^{0.226}D^{0.125}\left(0.3 + \log_{10} P\right)+ \\ 0.014A^{0.213} \times 10^{-0.021\frac{\left(D-180^2\right)}{1440}}\left(0.3 + \log_{10} P\right) \end{bmatrix}\right\} \qquad [2.8a]$$

Interpolation equation for duration (ii), A (1-30 000 km^2):

$$ARF = ARF_{12\text{hour}} + \left(ARF_{24hour} - ARF_{12hour}\right)\frac{D - 720}{720} \qquad [2.8b]$$

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

$$ARF_L = \text{Min}\left\{1, \begin{bmatrix} 1\text{-}\left(A^{b} - C\log_{10} D\right)D^{-d} + \\ eA^{f}D^{g}\left(0.3 + \log_{10} P\right)+ \\ h10^{iA\frac{D}{1440}}\left(0.3 + \log_{10} P\right) \end{bmatrix}\right\} \qquad [2.8c]$$

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

$$ARF = 1 - 0.6614\left(1 - ARF_{10\text{km}^2}\right)\left(A^{0.4} - 1\right) \qquad [2.8d]$$

where

ARF	= areal reduction factor [fraction],
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 2.8a to 2.8d are respectively the recommended methods for estimating ARFs in Australia and are documented in the Australian Rainfall and Runoff Manual (Ball et al., 2016).

Apart from the methods discussed Sections 2.7.1 and 2.7.2, Pietersen et al. (2015) and Pietersen (2016) contain summaries of all the other additional geographically-centred and storm-centred ARF estimation methods used internationally.

2.8 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 and 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) 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 long duration 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 regions.

The literature review of catchment response time parameters are presented in the next chapter.

CHAPTER 3:   LITERATURE REVIEW – CATCHMENT RESPONSE TIME

This chapter presents a literature review of the time parameters as introduced in Chapter 1, with the emphasis on the catchment response time estimation methods currently used nationally and internationally, and the inconsistencies introduced when using these different time parameter definitions and/or methods. In general, it is assumed that the equations currently used to estimate catchment response time in South Africa have a significant influence on the resulting hydrograph shape and peak discharge as estimated with different design flood estimation methods. Secondly, it is assumed that the most appropriate and best performing time variables and catchment storage effects are not currently incorporated into the methods used in South Africa.

3.1   Review of Catchment Response Time Estimation Methods

It is necessary to distinguish between the various definitions for time variables and time parameters (T_C, T_L and T_P) before attempting to review the various time parameter estimation methods available.

3.1.1   Time variables

Time variables can be estimated from the spatial and temporal distributions of rainfall hyetographs and total runoff hydrographs. In order to estimate these time variables, hydrograph analyses based on the separation of: (i) total runoff hydrographs into direct runoff and baseflow, (ii) rainfall hyetographs into initial abstraction, losses and effective rainfall, and (iii) the identification of the rainfall-runoff transfer function, are required. A convolution process is used to transform the effective rainfall into direct runoff through a synthetic transfer function based on the principle of linear super-positioning, i.e. multiplication, translation and addition (Chow et al., 1988; McCuen, 2005). In this report, 'convolution' refers to the process used to obtain observed time variables from hyetographs and hydrographs respectively, i.e. the transformation of effective rainfall into direct runoff through multiplication, translation and addition, where the volume of effective rainfall equals the volume of direct runoff. Consequently, time parameters are then based on the difference between two related time variables.

45

Effective rainfall hyetographs can be estimated from rainfall hyetographs in one of two different ways, depending on whether observed data are available or not. In cases where both observed rainfall and streamflow data are available, index methods such as the: (i) Phi-index method, where the index equals the average rainfall intensity above which the effective rainfall volume equals the direct runoff volume, and (ii) constant-percentage method, where losses are proportional to the rainfall intensity and the effective rainfall volume equals the direct runoff volume, can be used (McCuen, 2005). However, in ungauged catchments, the separation of rainfall losses must be based on empirical infiltration methods, which account for infiltration and other losses separately. The percentage of direct runoff is normally fixed and based on factors such as soils and land-use, with some possible adjustments based on the antecedent soil moisture conditions and rainfall depth (IH, 1999; Kjeldsen, 2007). The SCS runoff curve number method (CN values associated with specific soils and land-use categories) is internationally the most widely used (Chow et al., 1988).

In general, time variables obtained from hyetographs include the peak rainfall intensity, the centroid of effective rainfall and the end time of the rainfall event. Hydrograph-based time variables generally include peak discharges of observed surface runoff, the centroid of direct runoff and the inflection point on the recession limb of a hydrograph (McCuen, 2009).

3.1.2   Time parameters

Most design flood estimation methods require at least one (1) time parameter (e.g. T_C, T_L or T_P) as input. In the previous sub-section, it was highlighted that time parameters are based on the difference between two time variables, each obtained from a hyetograph and/or hydrograph. In practice, time parameters have multiple conceptual and/or computational definitions, and T_L is sometimes expressed in terms of T_C. Various researchers (e.g. McCuen et al., 1984; Schmidt and Schulze, 1984; Simas, 1996; McCuen, 2005; Jena and Tiwari, 2006; Hood et al., 2007; Fang et al., 2008; McCuen, 2009; Allnutt, 2019) have used the differences between the corresponding values of time variables to define two distinctive time parameters: T_C and T_L. Apart from these two time parameters, other time parameters such as T_P and the hydrograph time base (T_B) are also frequently used.

In the following sub-sections, the conceptual and computational definitions of T_C, T_L and T_P are detailed, and the various hydraulic and empirical estimation methods currently in use in South Africa and their interdependency are reviewed.

3.1.3   Time of concentration

Multiple definitions are used in the literature to define T_C. The most commonly used conceptual, physically-based definition of T_C is defined as the time required for runoff, as a result of effective rainfall with a uniform spatial and temporal distribution over a catchment, to contribute to the peak discharge at the catchment outlet or, in other words, the time required for a 'water particle' to travel from the catchment boundary along the longest watercourse to the catchment outlet (Kirpich, 1940; McCuen et al., 1984; McCuen, 2005; USDA NRCS, 2010; SANRAL, 2013).

Larson (1965) adopted the concept of time to virtual equilibrium (T_{VE}), i.e. the time when response equals 97% of the runoff supply, which is also regarded as a practical measure of the actual equilibrium time. The actual equilibrium time is difficult to determine due to the gradual response rate to the input rate. Consequently, T_C defined according to the 'water particle' concept would be equivalent to T_{VE}. However, runoff supply is normally of finite duration, while stream response usually peaks before equilibrium is reached and at a rate lower than the runoff supply rate. Pullen (1969) argued that this 'water particle' concept, which underlies the conceptual definition of T_C is unrealistic, since streamflow responds as an amorphous mass rather than as a collection of drops. In using such conceptual definition, the computational definition of T_C is thus the distance travelled along the principal flow path, which is divided into segments of reasonably uniform hydraulic characteristics, divided by the mean flow velocity in each of the segments (McCuen, 2009). The current common practice is to divide the principal flow path into segments of overland flow (sheet and/or shallow concentrated flow) and main watercourse or channel flow, after which, the travel times in the various segments are computed separately and totalled. Flow length criteria, i.e. overland flow distances (L_O) associated with specific overland slopes (S_O), are normally used as a limiting variable to quantify overland flow conditions, but flow retardance factors (i_p), Manning's overland roughness parameters (n) and overland conveyance factors (ϕ) are also used (Viessman and Lewis, 1996; Seybert, 2006; USDA NRCS, 2010). Seven typical overland slope-distance classes (based on above-mentioned flow length criteria) and as contained in the National Soil Conservation Manual (NSCM; DAWS, 1986) are listed in Table 3.1.

Table 3.1:   Overland flow distances associated with different slope classes (DAWS, 1986)

Slope class [S_o, %]	Distance [L_o, m]
0-3	110
3.1-5	95
5.1-10	80
10.1-15	65
15.1-20	50
20.1-25	35
25.1-30	20

The NSCM criteria are based on the assumption that the steeper the overland slope, the shorter the length of actual overland flow before it transitions into shallow concentrated flow followed by channel flow. McCuen and Spiess (1995) highlighted that the use of such criteria could lead to less accurate designs, and proposed that the maximum allowable overland flow path length criteria must rather be estimated as $30.48 S_O^{0.5} n^{-1}$. This criterion is based on the assumption that overland flow dominates where the flow depths are of the same order of magnitude as the surface resistance, i.e. roughness parameters in different slope classes.

The commencement of channel flow is typically defined at a point where a regular, well-defined channel exists with either perennial or intermittent flow, while conveyance factors (default value of 1.3 for natural channels) are also used to provide subjective measures of the hydraulic efficiency, taking both the channel vegetation and degree of channel improvement into consideration (Heggen, 2003; Seybert, 2006).

The second conceptual definition of T_C relates to the temporal distribution of rainfall and runoff, where T_C is defined as the time between the start of effective rainfall and the resulting peak discharge. The specific computations used to represent T_C based on time variables from hyetographs and hydrographs are discussed in the next paragraph to establish how the different interpretations of observed rainfall: runoff distribution definitions agree with the conceptual T_C definitions in this section (*cf.* Section 3.1.3). Numerous computational definitions have been proposed for estimating T_C from observed rainfall and runoff data. The following definitions as illustrated in Figure 3.1 are occasionally used to estimate T_C from observed hyetographs and hydrographs (McCuen, 2009):

(a)   The time from the end of effective rainfall to the inflection point on the recession limb of the total runoff hydrograph, i.e. the end of direct runoff. However, this is also the definition used by Clark (1945) to define T_L;

(b)   The time from the centroid of effective rainfall to the peak discharge of total runoff. However, this is also the definition used by Snyder (1938) to define T_L;

(c)   The time from the maximum rainfall intensity to the peak discharge; or

(d)   The time from the start of total runoff (rising limb of hydrograph) to the peak discharge of total runoff.
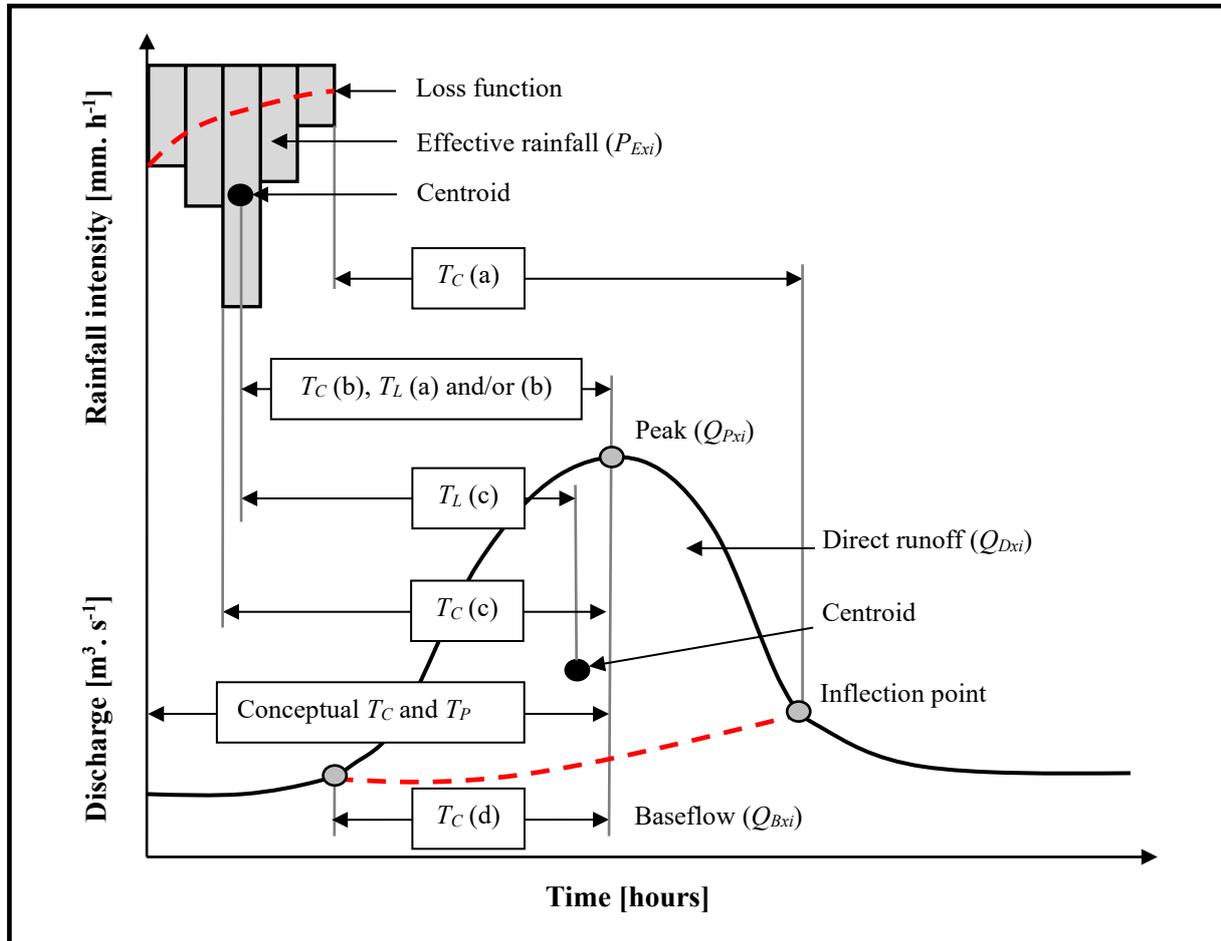
Figure 3.1:   Schematic illustrative of the different time parameter definitions and relationships (after Gericke and Smithers, 2014)

In South Africa, the South African National Roads Agency Limited (SANRAL) recommends the use of T_C definition (d) (SANRAL, 2013), but in essence all these definitions are dependent on the conceptual definition of T_C. It is also important to note that all these definitions listed in (a) to (d) are based on time variables with an associated degree of uncertainty. The 'centroid values' denote 'average values' and are therefore considered to be more stable time variables representative of the catchment response, especially in larger catchments or where flood volumes are central to the design (McCuen, 2009). In contrast to large catchments, the time variables related to peak rainfall intensities and peak discharges are considered to provide

the best estimate of the catchment response in smaller catchments where the exact occurrence of the maximum peak discharge is of more importance.

McCuen (2009) analysed 41 hyetograph-hydrograph storm event data sets from 20 catchment areas ranging from 1 to 60 ha in the USA. The results from floods estimated using the Rational and/or NRCS TR-55 methods indicated that the T_C based on the conceptual definition and principal flow path characteristics significantly underestimated the temporal distribution of runoff and T_C needed to be increased by 56% in order to correctly reflect the timing of runoff from the entire catchment, while the T_C based on T_C definition (b) proved to be the most accurate and was therefore recommended.

The hydraulically-based T_C estimation methods are limited to overland flow, which is derived from uniform flow theory and basic wave mechanics, e.g. the kinematic wave (Henderson and Wooding, 1964; Morgali and Linsley, 1965; Woolhiser and Liggett, 1967), dynamic wave (Su and Fang, 2004) and kinematic Darcy-Weisbach (Wong and Chen, 1997) approximations. The empirically-based T_C estimation methods are derived from observed meteorological and hydrological data and usually consider the whole catchment, not the sum of sequentially computed reach/segment behaviours. Stepwise multiple regression analyses are generally used to analyse the relationship between the response time and geomorphological, hydrological and meteorological parameters of a catchment. The hydraulic and/or empirical methods commonly used in South Africa to estimate the T_C are discussed in the following paragraphs:

(a) **Kerby's method:** This empirical method [Eq. (3.1)] is commonly used to estimate the T_C both as mixed sheet and/or shallow concentrated overland flow in the upper reaches of small, flat catchments. It was developed by Kerby (1959; cited by Seybert, 2006) and is based on the drainage design charts developed by Hathaway (1945; cited by Seybert, 2006). Therefore, it is sometimes referred to as the Kerby-Hathaway method. The South African Drainage Manual (SANRAL, 2013) also recommends the use of Eq. (3.1) for overland flow in South Africa. McCuen et al. (1984) highlighted that this method was developed and calibrated for catchments in the USA with areas less than 4 ha, average slopes of less than 1% and Manning's roughness parameters (n) varying between 0.02 and 0.8. In addition, the length of the flow path is a straight-line distance from the most distant point on the catchment boundary to the start of a fingertip

tributary (well-defined watercourse) and is measured parallel to the slope. The flow path length must also be limited to ± 100 m.

$$T_{C1} = 1.4394 \left(\frac{nL_O}{\sqrt{S_O}} \right)^{0.467} \qquad [3.1]$$

where

T_{C1} = overland time of concentration [minutes],

L_O = length of overland flow path [m], limited to 100 m,

n = Manning's roughness parameter for overland flow, and

S_O = average overland slope [m.m⁻¹].

(b) **SCS method:** This empirical method [Eq. (3.2)] is commonly used to estimate the T_C as mixed sheet and/or concentrated overland flow in the upper reaches of a catchment. The USDA SCS (later NRCS) developed this method in 1962 for homogeneous, agricultural catchment areas up to 8 km² with mixed overland flow conditions dominating (Reich, 1962). The calibration of Eq. (3.2) was based on T_C definition (c) (*cf.* Section 3.1.3) and a T_C: T_L proportionality ratio of 1.417 (McCuen, 2009). However, McCuen et al. (1984) showed that Eq. (3.2) provides accurate T_C estimates for catchment areas up to 16 km².

$$T_{C2} = \frac{L_O^{0.8} \left[\frac{25400}{CN} - 228.6 \right]^{0.7}}{706.9 S^{0.5}} \qquad [3.2]$$

where

T_{C2} = overland time of concentration [minutes],

CN = runoff curve number,

L_O = length of overland flow path [m], and

S = average catchment slope [m.m⁻¹].

(c) **NRCS velocity method:** This hydraulic method is commonly used to estimate T_C both as shallow concentrated overland and/or channel flow (Seybert, 2006). Either Eqs. (3.3a) or (3.3b) can be used to express the T_C for concentrated overland or channel flow. In the case of main watercourse/channel flow, this method is referred to as the NRCS segmental method, which divides the flow path into segments of reasonably uniform hydraulic characteristics. Separate travel time calculations are performed for

each segment based on either Eqs. (3.3a) or (3.3b), while the total T_C is computed using Eq. (3.3c) (USDA NRCS, 2010):

$$T_{C3i} = 0.0167\left(\frac{nL_{O,CH}}{R^{0.667}\sqrt{S_{O,CH}}}\right) \qquad [3.3a]$$

$$T_{C3i} = 0.0167\left(\frac{L_{O,CH}}{18\log\left(\frac{12R}{k_s}\right)\sqrt{RS_{O,CH}}}\right) \qquad [3.3b]$$

$$T_{C3} = \sum_{i=1}^{N} T_{Ci} \qquad [3.3c]$$

where

T_{C3}	= overland/channel flow time of concentration computed using the NRCS method [minutes],
T_{C3i}	= overland/channel flow time of concentration of segment i [minutes],
k_s	= Chézy's roughness parameter [m],
$L_{O,CH}$	= length of flow path, either overland or channel flow [m],
n	= Manning's roughness parameter,
R	= hydraulic radius which equals the flow depth [m], and
$S_{O,CH}$	= average overland or channel slope [m.m^{-1}].

(d) **USBR method:** Equation (3.4) was proposed by the USBR (1973) to be used as a standard empirical method to estimate the T_C in hydrological designs, especially culvert designs based on the California Culvert Practice (CCP, 1955; cited by Li and Chibber, 2008). However, Eq. (3.4) is essentially a modified version of the Kirpich method (Kirpich, 1940) and is recommended by SANRAL (2013) for use in South Africa for defined, natural watercourses/channels. It is also used in conjunction with Eq. (3.1) which estimates overland flow time, to estimate the total travel time (overland plus channel flow) for deterministic design flood estimation methods in South Africa. Van der Spuy and Rademeyer (2018) highlighted that Eq. (3.4) tends to result in estimates that are either too high or too low and recommend the use of a correction factor (τ) as shown in Eq. (3.4a) and listed in Table 3.2.

$$T_{C4} = \left(\frac{0.87L_{CH}^{2}}{1000S_{CH}}\right)^{0.385} \qquad [3.4]$$

$$T_{C4a} = \tau\left(\frac{0.87L_{CH}^{2}}{1000S_{CH}}\right)^{0.385} \qquad [3.4a]$$

where

$T_{C4,4a}$ = channel flow time of concentration [hours],

L_{CH} = length of longest watercourse [km],

S_{CH} = average main watercourse slope [m.m^{-1}], and

τ = correction factor.

Table 3.2:   Correction factors (τ) for T_C (Van der Spuy and Rademeyer, 2018)

Area [A, km²]	Correction factor [τ]
< 1	2
1-00	$2-0.5\log A$
100-000	1
5 000-00 000	$2.42-0.385\log A$
> 100 000	0.5

3.1.4 Lag time

Conceptually, T_L is generally defined as the time between the centroid of effective rainfall and the peak discharge of the resultant direct runoff hydrograph, which is the same as T_C definition (b) as shown in Figure 3.1. Computationally, T_L can be estimated as a weighted T_C value when, for a given storm, the catchment is divided into sub-areas and the travel times from the centroid of each sub-area to the catchment outlet are established by the relationship expressed in Eq. (3.5). This relationship is also illustrated in Figure 3.2 (USDA NRCS, 2010).

$$T_L = \frac{\sum(A_i Q_i T_{Ti})}{\sum(A_i Q_i)} \tag{3.5}$$

where

T_L = lag time [hours],

A_i = incremental catchment area/sub-area [km²],

Q_i = incremental runoff from A_i [mm], and

T_{Ti} = travel time from the centroid of A_i to catchment outlet [hours].

In flood hydrology, T_L is normally not estimated using Eq. (3.5). Instead, either empirical or analytical methods are normally used to analyse the relationship between the response time and meteorological and geomorphological parameters of a catchment. In the following paragraph, the hydrological parameters, as defined by different interpretations of observed rainfall: runoff distribution definitions are explored.
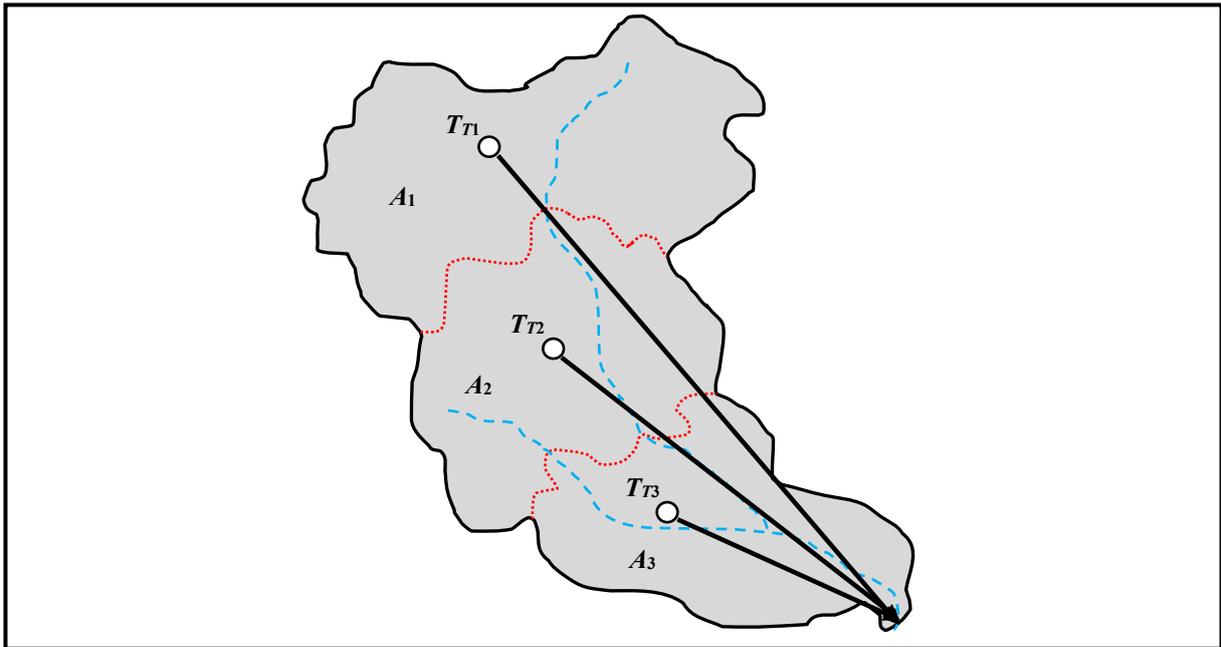
Figure 3.2:   Conceptual travel time from the centroid of each sub-area to the catchment outlet (USDA NRCS, 2010)

Scientific literature often fails to clearly define and distinguish between the T_C and T_L, especially when observed data (hyetographs and hydrographs) are used to estimate these time parameters. The differences between time variables from various points on hyetographs to various points on the resultant hydrographs are sometimes misinterpreted as T_C. The following definitions as illustrated in Figure 3.1 are occasionally used to estimate T_L as a time parameter from observed hyetographs and hydrographs (Heggen, 2003):

(a)   The time from the centroid of effective rainfall to the time of the peak discharge of direct runoff;

(b)   The time from the centroid of effective rainfall to the time of the peak discharge of total runoff; or

(c)   The time from the centroid of effective rainfall to the centroid of direct runoff.

As in the case of the T_C, T_L is also based on uncertain, inconsistently defined time variables. However, T_L definitions (a) to (c) use 'centroid values' and are therefore considered likely to be more stable time variables which are representative of the catchment response in medium to large catchments. Pullen (1969) also highlighted that T_L is preferred as a measure of catchment response time, especially due to the incorporation of storm duration in these definitions. Definitions (a) to (c) are generally used or defined as T_L (Simas, 1996; Hood et al., 2007; Folmar and Miller, 2008; Pavlovic and Moglen, 2008), although T_L definition (b) is also

sometimes used to define T_C. Dingman (2002; cited by Hood et al., 2007) recommended the use of Eq. (3.6) to estimate the centroid values of hyetographs or hydrographs, respectively.

$$C_{P,Q} = \frac{\sum\limits_{i=1}^{N} X_i t_i}{\sum\limits_{i=1}^{N} X_i} \qquad [3.6]$$

where

$C_{P,Q}$ = centroid value of rainfall or runoff [mm or $m^3.s^{-1}$],

t_i = time for period i [hour],

N = sample size, and

X_i = rainfall or runoff for period i [mm or $m^3.s^{-1}$].

The empirical methods commonly used in South Africa to estimate T_L are discussed in the following paragraphs:

(a) **HRU method:** This method was developed by the HRU (Pullen, 1969) in conjunction with the development of Synthetic Unit Hydrographs (SUHs) for South Africa (HRU, 1972). The lack of continuously recorded rainfall data for medium to large catchments in South Africa, forced Pullen (1969) to develop an indirect method to estimate T_L using only observed streamflow data from 96 catchment areas ranging from 21 km² to 22 163 km². Pullen (1969) assumed that the onset of effective rainfall coincides with start of direct runoff, and, that the T_P could be used to describe the time lapse between this mutual starting point and the resulting peak discharge. In essence, it was acknowledged that direct runoff is unable to recede before the end of effective rainfall; therefore, the T_P was regarded as the upper limit storm duration during the implementation of the unit hydrograph theory using the *S*-curve technique. In other words, a hydrograph of 25 mm of direct runoff was initially assumed to be a T_P-hour unit hydrograph. However, due to non-uniform temporal and spatial runoff distributions, possible inaccuracies in streamflow measurements and non-linearity's in catchment response characteristics, the *S*-curves fluctuated about the equilibrium discharge of amplitude. Therefore, the analysis was repeated using descending time intervals of 1-hour until the fluctuations of the *S*-curve ceiling value diminished to within a prescribed 5% range. After the verification of the effective rainfall durations, all the hydrographs of 25 mm of direct runoff were converted to unit hydrographs of relevant duration. In order to facilitate the comparison of these unit hydrographs derived

from different events in a given catchment, all the unit hydrographs for a given record were then converted by the *S*-curve technique to unit hydrographs of standard duration (Pullen, 1969).

Thereafter, the centroid of each unit hydrograph was determined by simple numerical integration of the unit hydrograph from time zero. The T_L values were then simply estimated as the time lapse between the centroid of effective rainfall and the centroid of a unit hydrograph (Pullen, 1969). The catchment-index ($L_C L_H S_{CH}^{-0.5}$), as proposed by the United States Army Corps of Engineers, USACE (Linsley et al., 1988) was used to estimate the delay of runoff from the catchments. The T_L values (dependent variables) were plotted against the catchment indices (independent variables) on logarithmic scales. Least-square regression analyses were then used to derive a family of T_L equations applicable to each of the nine homogeneous veld-type regions with representative SUHs in South Africa, as expressed by Eq. (3.7). The regionalisation scheme of the veld-type regions took into consideration catchment characteristics, e.g. topography, soil types, vegetation and rainfall, which are most likely to influence catchment storage and therefore T_L.

$$T_{L1} = C_{T1}\left(\frac{L_H L_C}{\sqrt{S_{CH}}}\right)^{0.36} \qquad [3.7]$$

where

T_{L1} = lag time [hours],

C_{T1} = regional storage coefficient as listed in Table 3.3,

L_C = centroid distance [km],

L_H = hydraulic length of catchment [km], and

S_{CH} = average main watercourse slope [m.m^{-1}].

Table 3.3: Generalised regional storage coefficients (HRU, 1972)

Veld region	Veld-type description	C_{T1}
1	Coastal tropical forest	0.99
2	Schlerophyllous bush	0.62
3	Mountain sourveld	0.35
4	Grassland of interior plateau	0.32
5	Highland sourveld and Dohne sourveld	0.21
5A	Zone 5, soils weakly developed	0.53
6	Karoo	0.19
7	False Karoo	0.19
8	Bushveld	0.19
9	Tall sourveld	0.13

(b)   **SCS lag method:** In sub-section 3.1.3 it was highlighted that this method was developed by the USDA SCS in 1962 (Reich, 1962) to estimate T_C where mixed overland flow conditions in catchment areas up to 8 km² exists. However, using the relationship of $T_L = 0.6T_C$, Eq. (3.8) can also be used to estimate T_L in catchment areas up to 16 km² (McCuen, 2005).

$$T_{L2} = \frac{L_H^{0.8}\left[\dfrac{25\,400}{CN} - 228.6\right]^{0.7}}{168.862\,S^{0.5}} \qquad [3.8]$$

where

T_{L2}	= lag time [hours],
CN	= runoff curve number,
L_H	= hydraulic length of catchment [km], and
S	= average catchment slope [m.m^{-1}].

(c)   **Schmidt-Schulze (SCS-SA) method:** Schmidt and Schulze (1984) estimated T_L from observed rainfall and streamflow data in 12 agricultural catchments in South Africa and the USA with catchment areas smaller than 3.5 km² by using three different methods to develop Eq. (3.9). This equation is used in preference to the original SCS lag method [Eq. (3.8)] in South Africa, especially when stormflow response includes both surface and subsurface runoff as frequently encountered in areas of high MAP or on natural catchments with good land cover (Schulze et al., 1992).

$$T_{L3} = \frac{A^{0.35}MAP^{1.10}}{41.67\,S^{0.3}i_{30}^{0.87}} \qquad [3.9]$$

where

T_{L3}	= lag time [hours],
A	= catchment area [km²],
i_{30}	= 2-year return period 30-minute rainfall intensity [mm.h^{-1}],
MAP	= mean annual precipitation [mm], and
S	= average catchment slope [%].

The three different methods used to develop Eq. (3.9) are based on the following approaches (Schmidt and Schulze, 1984):

Initially, the relationship between peak discharge and volume was investigated by regressing linear peak discharge distributions (single triangular hydrographs) against

the corresponding runoff volume obtained from observed runoff events to determine the magnitude and intra-catchment variability of T_L. Thereafter, the incremental triangular hydrographs were convoluted with observed effective rainfall to form compound hydrographs representative of the peak discharge and temporal runoff distribution of observed hydrographs. Lastly, the average time response between effective rainfall and direct runoff was measured in each catchment to determine an index of catchment lag time. It was concluded that intra-catchment T_L estimates in ungauged catchments can be improved by incorporating indices of climate and regional rainfall characteristics into an empirical lag equation. The 2-year return period 30-minute rainfall intensity proved to be the dominant rainfall parameter that influences intra-catchment variations in T_L estimates (Schmidt and Schulze, 1984).

3.1.5   Time to peak

T_P, which is used in many hydrological applications, can be defined as the time from the start of effective rainfall to the peak discharge in a single-peaked hydrograph (McCuen et al., 1984; USDA SCS, 1985; Linsley et al., 1988; Seybert, 2006). However, this is also the conceptual definition used for T_C (*cf.* Figure 3.1). T_P is also sometimes defined as the time interval between the centroid of effective rainfall and the peak discharge of direct runoff (Heggen, 2003); however, this is also one of the definitions used to quantify T_C and T_L using T_C definition (b) and T_L definition (c), respectively. According to Ramser (1927), T_P is regarded to be synonymous with the T_C and that both these time parameters, are reasonably constant for a specific catchment. In contrast, Bell and Kar (1969) concluded that these time parameters are far from being constant; in fact, they may deviate between 40% and 200% from the median value.

The empirical methods sometimes used in South Africa to estimate T_P are discussed in the following paragraphs:

(a) **SCS-Mockus method:** This empirical method [Eq. (3.10)] is occasionally used in South Africa to estimate T_P based on the SUH research conducted by Snyder (1938), while Mockus (1957; cited by Viessman et al., 1989) developed the SCS SUHs from dimensionless unit hydrographs as obtained from a large number of natural hydrographs in various catchments with variable sizes and geographical locations. Only the T_P and Q_P values are required to approximate the associated SUHs, while the T_P is

expressed as a function of the storm duration and T_L. Equation (3.10) is based on T_L definition (b), while it also assumes that the effective rainfall is constant with the centroid at $\frac{P_D}{2}$.

$$T_{P1} \qquad = \frac{P_D}{2} + T_L \qquad\qquad\qquad [3.10]$$

where

    T_{P1} = time to peak [hours],

    P_D = storm duration [hours], and

    T_L = lag time based on Eq. (3.7) [hours].

(b) **Gericke-Smithers (G&S) method:** This empirical method [Eq. (3.13)] was developed by Gericke and Smithers (2016b; 2017) in four climatologically different regions in South Africa. In essence, Gericke and Smithers (2017) estimated observed T_{Px} values directly from observed streamflow data using three different methods: (i) duration of total net rise of a multi-peaked hydrograph, (ii) triangular-shaped direct runoff hydrograph approximation, and (iii) a linear catchment response function, Eq. (3.11). The use of the three different methods in combination to estimate individual event (T_{Pxi}) and catchment (T_{Px}) values proved to be both practical and objective with consistent results. Based on the specific results obtained, it was recommended that for design hydrology and for the calibration of empirical equations to estimate catchment response time, the estimation of the observed catchment T_{Px} should be based on the linear catchment response function [Eq. (3.11)]. It is important to note that Eq. (3.11) is only reliant on observed streamflow variables and is therefore not influenced by the limitations and availability of rainfall data in medium to large catchments. De Almeida et al. (2016) also acknowledged the appropriateness of using only streamflow data to estimate the T_C. Equation (3.11) is also regarded as an appropriate 'representative value' which ensures that the averages of individual event-based catchment responses are a good reflection of the catchment conditions and sample-mean.

Consequently, the T_{Px} values estimated with Eq. (3.11) were used as criterion variables in conjunction with a selection of climatological and geomorphological predictor variables to calibrate and derive Eq. (3.13). The following independent predictor variables were retained and included in Eq. (3.13): (i) MAP, (ii) A, (iii) L_C, (iv) L_H, and (v) S. Equation (3.13) not only met the requirement of statistical significance,

consistency and ease of application by practitioners in ungauged catchments, but the interaction between the five retained independent predictor variables, improved the estimation of catchment response times and the resulting peak discharge (Gericke, 2018). However, the inclusion of these variables also proved to be the best combination of 'catchment transfer functions' to estimate the T_{Px} values at a catchment level. Hence, the same equation format, with different regional calibration coefficients was used in each of the four regions. The derived T_{Py} regression is shown in Eq. (3.12) and Eq. (3.13), respectively.

$$T_{Px} = \frac{1}{3600x}\left[\frac{\sum\limits_{i=1}^{N}\left(Q_{Pxi}-\overline{Q_{Px}}\right)\left(Q_{Dxi}-\overline{Q_{Dx}}\right)}{\sum\limits_{i=1}^{N}\left(Q_{Pxi}-\overline{Q_{Px}}\right)^2}\right] \qquad [3.11]$$

$$\ln(T_{Py}) = \ln(x_1)MAP + \ln(x_2)A + \ln(x_3)L_C + \ln(x_4)L_H + \ln(x_5)S \qquad [3.12]$$

In applying some simplification, the final T_{Py} regression is shown in Eq. (3.13):

$$T_{Py} = x_1^{MAP} x_2^{A} x_3^{L_C} x_4^{L_H} x_5^{S} \qquad [3.13]$$

where

T_{Px}	= 'average' catchment time to peak based on a linear catchment response function [hours],
T_{Py}	= estimated time to peak [hours],
A	= catchment area [km²],
L_C	= centroid distance [km],
L_H	= hydraulic length [km],
Q_{Dxi}	= volume of direct runoff for individual flood events [m³],
$\overline{Q_{Dx}}$	= mean of Q_{Dxi} [m³],
Q_{Pxi}	= observed peak discharge for individual flood events [m³.s⁻¹],
$\overline{Q_{Px}}$	= mean of Q_{Pxi} [m³.s⁻¹],
MAP	= mean annual precipitation [mm],
N	= sample size,
S	= average catchment slope [%],
x	= variable proportionality ratio (default $x = 1$), which depends on the catchment response time parameter under consideration, and
x_1 to x_5	= regional calibration coefficients as listed in Table 3.4.

Table 3.4: Regional calibration coefficients applicable to Equation (3.13)

Region	Regional calibration coefficients [* 10^-2]				
	x_1	x_2	x_3	x_4	x_5
Northern Interior	100.280	99.993	99.865	101.612	91.344
Central Interior	100.313	99.984	106.106	98.608	98.081
Southern Winter Coastal region	100.174	99.931	101.805	104.310	99.648
Eastern Summer Coastal region	100.297	99.991	99.594	101.177	97.529

The variable proportionality ratio (x) is included in Eq. (3.11) to increase the flexibility and use of this equation, i.e. with $x = 1$, either T_{Px} or T_{Cx} could be estimated by acknowledging the approximation of $T_C \approx T_P$ (Gericke and Smithers, 2014) and with $x = 1.667$, T_L could be estimated by assuming that $T_L = 0.6T_C$, which is the time from the centroid of effective rainfall to the time of peak discharge (*cf.* Section 3.1.4, definitions (a) and/or (b)).

Apart from the methods discussed Sections 3.1.3 to 3.1.5, an extensive summary of all the other T_C, T_L and T_P estimation methods used internationally can be found in Gericke and Smithers (2014), which reviewed a total of three hydraulic and 44 empirical time parameter (T_C, T_L and T_P) estimation methods.

3.2   Time Parameter Proportionality Ratios

As highlighted in Section 3.1.4, due to the difficulty in estimating the centroid of hyetographs and hydrographs, other T_L estimation techniques have been proposed by introducing the concept of time parameter proportionality ratios. Instead of using T_L as an input for design flood estimation methods, it is rather used as input to the computation of T_C. In using T_L definition (c) (*cf.* Figure 3.1 and Section 3.1.3), T_C and T_L are normally related by $T_C = 1.417T_L$ or $T_L = 0.705T_C$ (McCuen, 2009). In T_L definitions (a) and (b), the proportionality ratio increases to 1.667, i.e. $T_C = 1.667T_L$ or $T_L = 0.6T_C$ (McCuen, 2009).

However, Schultz (1964) established that for small catchments in Lesotho and South Africa, $T_C \approx T_L$, which conflicts with the above proportionality ratios. In addition, Gericke and Smithers (2017) also showed that $T_C \approx T_P$ at medium to large catchment scales in South Africa, but the relevance of the T_L proportionality ratio ($x = 1.667$), i.e. $T_L = 0.6T_C$, was not established.

In considering the inconsistent use of time parameter definitions and the inherent procedural limitations associated with the rainfall-runoff convolution process as introduced in Sections 3.1.3 to 3.1.5 (with the aid of Figure 3.1), Allnutt (2019) and Allnutt et al. (2020), as part of this project (*cf.* Chapter 7, Section 7.4), with specific reference to Study Assumption 4 (*cf.* Chapter 1, Section 1.3.2) investigated and established the suitability of the currently recommended time parameter definitions and proportionality ratios for small catchments in larger catchment areas exceeding 50 km².

Given the fact that the analysis of the hyetograph-hydrograph relationships to obtain time variables and time parameters is often done manually, relying on the visual examination and interpretation, the focus of this part of the project was on the development of an automated hyetograph-hydrograph analysis tool to estimate time parameters and average time parameter proportionality ratios at a catchment level. Apart from saving considerable time, such an automated tool also addressed issues pertained to inconsistency and subjectivity. The latter statement is supported by White and Sloto (1990), which highlighted that automated hydrograph analyses provide objective and consistent results.

It must be noted that most of the former (current) automated tools for hydrograph analyses focus on the selection of hydrograph characteristics and the incorporation of baseflow separation, recession analyses and direct runoff estimation (Arnold et al., 1995; Chapman, 1999; Lim et al., 2005; Piggott et al., 2005; Rutledge, 1998; Sloto and Crouse, 1996). However, the use of automated tools to extract and analyse rainfall hyetographs, is not common practice and most of the rainfall-based time variables are extracted manually. In essence, none of the automated tools developed, include both rainfall hyetograph and streamflow hydrograph characteristics, while the relationship between rainfall-based and runoff-based time variables is not defined (Allnutt et al., 2020).

Given the importance of reliable estimates of both direct runoff and baseflow when time parameters are in question, the separation thereof is discussed in the next section. In addition, direct runoff, both in terms of the peak discharge and runoff volume, could be influenced in cases where flood events exceed the structural limit at a particular gauging weir. Consequently, the extension of stage-discharge relationships is discussed in Section 3.4.

3.3 Baseflow Separation Techniques

A number of methods (e.g. graphical, recursive digital filters, frequency-duration and recession analysis) have been proposed in the literature to separate direct runoff and baseflow (Lyne and Hollick, 1979; Nathan and McMahon, 1990; Arnold et al., 1995; Chapman, 1999; Smakhtin, 2001). The selection or preference for any method will depend on the type and volume of observed data available versus the accuracy required and time constraints. Recursive digital filtering methods are the most frequently used approaches to separate direct runoff and baseflow, despite having no true physical or hydrological basis, but it is objective and reproducible for continuous baseflow separation. Routine tools for signal analysis and processing are used to remove high-frequency signals (direct runoff) to derive the low-frequency (baseflow) signals (Arnold et al., 1995).

According to Smakhtin (2001), the most well-known and widely used recursive digital filtering methods were developed by Lyne and Hollick [1979; Eq. (3.14)], and Nathan and McMahon [1990; Eq. (3.15)], while the filtering method proposed by Chapman [1999; Eq. (3.16)] is also often used. Nathan and McMahon (1990) compared the Lyne and Hollick (1979) methodology of baseflow separation with several other more rigorous algorithms, and concluded that Eq. (3.14) is simpler, more user-friendly and produced as good a result as the alternatives. Smakhtin and Watkins (1997) adopted the methodology as proposed by Nathan and McMahon (1990) with some modifications in a national-scale study in South Africa, while Hughes et al. (2003) and Gericke and Smithers (2017) also adopted the latter methodology on pilot-scale studies in South Africa.

$$Q_{Dxi} = \alpha Q_{Dx(i-1)} + \frac{(1+\alpha)}{2}\left(Q_{Txi} - Q_{Tx(i-1)}\right) \qquad [3.14]$$

$$Q_{Dxi} = \alpha Q_{Dx(i-1)} + \beta(1+\alpha)\left(Q_{Txi} - Q_{Tx(i-1)}\right) \qquad [3.15]$$

$$Q_{Dxi} = \left(\frac{(3\alpha-1)}{(3-\alpha)}\right)Q_{Dx(i-1)} + \frac{2}{(3-\alpha)}\left(Q_{Txi} - Q_{Tx(i-1)}\right) \qquad [3.16]$$

where Q_{Dxi} = filtered direct runoff at time step *i*, which is subject to $Q_{Dx} \geq 0$ for time *i* [m^3.s^{-1}],

            α,β = filter parameters [0 < α < 1; 0 < β < 0.5], and

            Q_{Txi} = total streamflow (i.e. direct runoff + baseflow) at time *i* [m^3.s^{-1}].

Equation (3.14) is the default digital filter algorithm available in the HYBASE extension included in the DWS Hydstra software and three approaches are available for the baseflow separation:

(a) The first approach separates the baseflow by applying the digital filtering algorithm and the output file is a transformed copy of the input file with all peaks smoothed out. The higher the α-parameter value, and the more filter passes applied, the more smoothed is the output.

(b) The second approach is exactly the same as the first, but allows for finer control over the number of passes made on the data. In the first approach, a pass consists of two phases, a forward phase and a reverse phase. In the second approach, a pass consists of just one phase, so three passes would be a forward phase, a reverse phase and a second forward phase.

(c) The third approach is based on a separation value, where a rising slope limit is applied to the hydrograph, and the flows above the slope limit are considered quick flow, those below, are regarded as baseflow.

In applying Eq. (3.15), Smakhtin and Watkins (1997) established that a fixed α-parameter value of 0.995 is suitable for most catchments in South Africa, although in some catchments, α-parameter values of 0.997 proved to be more appropriate. Hughes et al. (2003) also highlighted that a fixed β-parameter value of 0.5 could be used with daily time-step data, since there is more than enough flexibility in the setting of the α-parameter value to achieve an acceptable result.

3.4 Extension of Stage-Discharge Relationships

The first flow-gauging weirs were constructed in 1904 in the Transvaal Province to measure streamflow continuously in South Africa (Menné, 1960). By the year 2007, streamflow data were already recorded continuously at 782 different flow-gauging sites, comprising of Sharp-crested weirs (55%), Crump weirs (35%) and the remaining 10% consisting of Broad-crested weirs, dam spillways and velocity-area gauging sites (Wessels and Rooseboom, 2009). However, streamflow is seldom directly measured; instead, the stage (flow depth) is continuously measured and converted into a discharge utilising a stage-discharge rating curve at a flow-gauging weir. During flood events, flow-gauging weirs might be flooded with the water level beyond the structural limit. Subsequently, the standard calibration of the flow-gauging weir will no longer be a true reflection of the actual discharges that occurred during

the flood events, and the standard stage-discharge rating curve must then be extended beyond the highest stage reading to reflect these high discharges (Petersen-Øverleir and Reitan, 2009).

Direct measurements, e.g. conventional current gaugings, are not always possible due to various practical constraints associated with high discharge events, e.g. high velocities and water depths, danger to staff and equipment to enter a required river reach, and operational difficulties. As a result, different indirect methods for extending stage-discharge rating curves are available, but the impact of using these different methods vary significantly and highlight the need for a robust and reliable extension method, since significant errors could be introduced (Lang et al., 2010).

Various studies were undertaken in attempt to develop robust and reliable extension methods. For example, the Environmental Agency in the United Kingdom (EA UK), undertook a study entitled: *'Extension of rating curves at flow-gauging stations'* to compile a best practices manual using hydraulic and computational modelling techniques (Ramsbottom and Whitlow, 2003). Dymond and Ross (1982) evaluated the accuracy of stage-discharge rating curves by considering both the individual and average discharge measurement errors during a specific period. Petersen-Øverleir and Reitan (2009) examined the joint impact of sample variability and rating curve inaccuracy on at-site flood frequency analysis. Lang et al. (2010) extended stage-discharge rating curves using hydraulic modelling to ultimately improve flood frequency analyses. Shao et al. (2018) extended stage-discharge rating curves using hydrodynamic models and, as a result, quantified the uncertainty associated with the overall process.

However, despite of all the above studies, there is currently no user manual or standard practices available to extend stage-discharge relationships in South Africa (Van der Spuy, 2022). The latter absence of standard practices to extend stage-discharge rating curves is ascribed to the different and unique hydraulic characteristics, topography and discharge conditions present at South African flow-gauging sites; hence, site visits and a fundamental understanding of the at-site hydraulics are crucial to ensure the accurate extension of stage-discharge rating curves. Therefore, it is warranted that the use of appropriate stage-discharge extension methods would not only enhance the estimation of design flood events, i.e. peak discharges, but it would also impact on water resources management. Even if high flood events above the structural limit of flow-gauging weirs might occur only 5% of the time, it could still

constitute 80% or more of the total runoff volume, which is essential in effective water resources management.

Hence, as part of this project, with specific reference to Study Assumption 1 (*cf.* Chapter 1, Section 1.3.2), a selection of indirect extension methods (e.g. hydraulic and one-dimensional modelling methods) was evaluated and compared to direct extension (benchmark) methods (e.g. at-site conventional current gaugings, hydrograph analyses and level pool routing techniques), to establish the best-fit and most appropriate stage-discharge extension method to be used in South Africa.

Please refer to Chapter 7, Section 7.5 for the methodological approach adopted and the results.

3.5   Summary

The use of time parameters based on either hydraulic or empirical estimation methods was evident from the literature review conducted. It was evident that none of these hydraulic and empirical methods are highly accurate or consistent to provide the true value of these time parameters, especially when applied outside their original developmental regions. In addition, many of these methods/equations proved to be in a disparate form and are presented without explicit unit specifications and suggested values for constants.

Heggen (2003), who summarised more than 80 T_C, T_L and T_P estimation methods from the literature, also confirmed the above findings. In addition, Michailidi et al. (2018), highlighted that time parameters, especially T_C due to its general and wide-spread use, has been subjected to severe misuse as a result of the existence of multiple, ambiguous and even illogical definitions, as well as numerous equations providing significantly different estimations. According to Michailidi et al. (2018), assuming that T_C is constant in a particular catchment, is regarded as the biggest shortcoming in all these approaches. Hence, the necessity of implementing a variable T_C concept within hydrological modelling and design flood estimation. Typically, the dynamic unit hydrograph approach, the shape of which follows the variability of the excess rainfall intensity, is an evident consequence of the rainfall-dependence of T_C, and an essential component of this new working paradigm.

CHAPTER 4:   AREAL REDUCTION FACTORS – CONCERNS AND SOLUTIONS

It is essential that the reader is familiar with the ARF literature review as included in Chapter 2. This chapter mainly focusses on the concerns and/or problems encountered during the estimation of sample ARFs, with possible solutions where applicable.

4.1   Homogeneous Rainfall Regions

The original methodology envisaged, with specific reference to the regionalisation of rainfall regions, entailed the implementation of seven (7) long duration rainfall clusters as identified and used by Smithers and Schulze (2000b). However, based on the recommendations made by the Reference Group to adopt the 78 homogeneous rainfall regions, as originally derived by Smithers and Schulze (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 regions. Thereafter, grid points falling within each of the 78 homogeneous groups were plotted and subsequently outlined in ArcGIS^{TM}. 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 region/group. In other words, this step was necessary to ensure that the RLMA&SI grid points indeed fall within the same GIS-based region, seeing that a minor percentage of the plotted grid points are slightly scattered around potential regional borders (*cf.* Figure 4.1).

(b) The regional boundaries were therefore adjusted until the contributing grid points had a similarity ratio ≥ 90%. The latter similarity ratios will not directly affect the estimation of ARFs; however, it would ensure that each region's boundary delineated in GIS mimics the original 78 RLMA&SI homogeneous rainfall regions as close as possible.
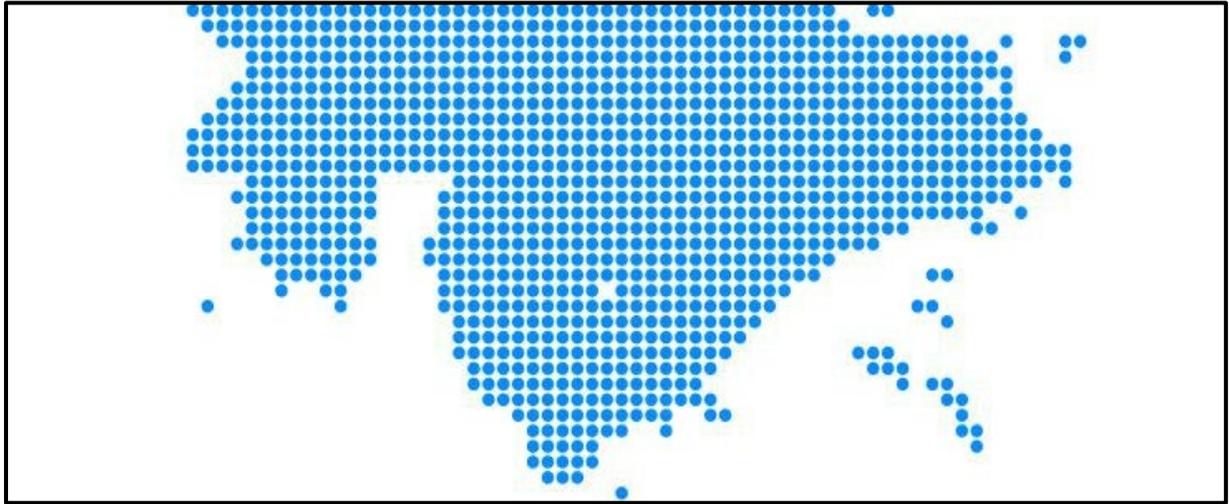
Figure 4.1:   Example of grid points scattered around potential regional borders

(c) The 1 779 selected rainfall stations were plotted in ArcGIS™ to visualise the location of each rainfall station within the 78 homogeneous rainfall regions. Unfortunately, some of the 78 regions are very small and, as a result, included a low number of rainfall stations. Hence, some of the GIS-based homogeneous rainfall regions which are within the same seven (7) long duration rainfall clusters were combined to increase the: (i) size of the region, and (ii) number of rainfall stations per region. Similarly, this was necessary to ensure that the larger circular catchments ($> 8\,000$ km^2) could be positioned within each region. This procedure resulted in the 46 homogeneous rainfall regions as used in this study, with the rainfall stations (*cf.* Table 4.1) being uniformly distributed across all regions (*cf.* Figure 4.2).

Table 4.1:   The 1 779 rainfall stations within each of the 46 homogeneous rainfall regions

Region	Number of stations	Area [km^2]	Region	Number of stations	Area [km^2]	Region	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			
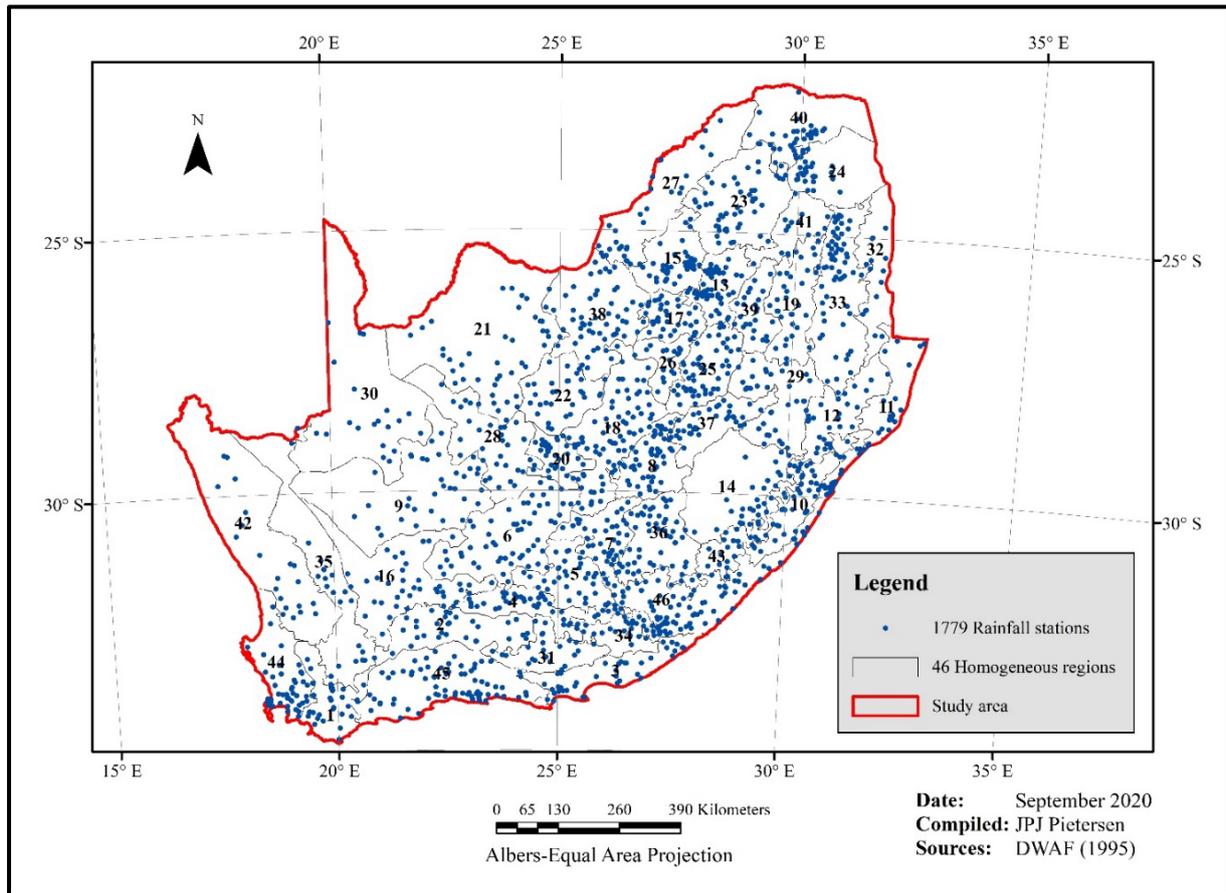
Figure 4.2: Forty-six homogeneous rainfall regions (after Smithers and Schulze, 2004)

4.2 Theoretical Probability Distributions and Method of Moments

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 AMS values were estimated using Linear Moments (LM), Mean Moments (MM) and Probable Weighted Moments (PWM) in 10 circular test catchments (125 km² each). All three methods of moments were fitted to the GEV probability distribution and resulted in areal design rainfall values in each of the 10 circular test catchments. The sample ARFs were estimated in each catchment using the ratio between

69

the GEV-based areal design rainfall and the RLMA&SI design point rainfall. The results are listed in Table 4.2.

Table 4.2: Sample ARFs using different methods of moments

GEV$_{PWM}$				GEV$_{LM}$				GEV$_{MM}$			
Sample ARFs [%]				Sample ARFs [%]				Sample ARFs [%]			
1-d	3-d	5-d	7-d	1-d	3-d	5-d	7-d	1-d	3-d	5-d	7-d
83.8	88.5	89.8	91.3	80.0	85.4	86.7	88.4	81.6	86.6	87.8	89.2
83.1	84.6	85.5	86.0	79.7	84.6	85.3	85.9	81.4	86.1	86.5	86.5
82.8	81.8	82.0	81.9	80.1	84.4	84.5	84.0	80.7	85.2	84.9	84.1
82.5	79.1	78.5	77.8	80.8	84.4	83.8	82.3	79.9	84.1	83.3	81.7
82.3	75.6	73.8	72.4	82.1	84.8	83.1	80.1	78.5	82.6	81.0	78.4
82.3	73.1	70.5	68.5	83.3	85.5	82.7	78.7	77.4	81.5	79.3	76.0
82.3	70.7	67.2	64.8	84.8	86.4	82.6	77.5	76.3	80.5	77.6	73.7

In Table 4.2, the similarities between the different ARF values are evident and highlight that the different methods of estimating moments do not have a significant impact on ARF estimation. Subsequently, based on the above recommendation and results, the estimation of sample ARFs in the subsequent sections is based on the GEV$_{LM}$ probability distribution of areal and design point rainfall values. However, using the RLMA&SI design rainfall database as a possible source of design point rainfall values, is investigated in the next section.

4.3   RLMA&SI Design Point Rainfall

The original methodology proposed entailed the use of the RLMA&SI design rainfall database as the design point rainfall data source for estimating ARFs. As an example, the steps in Section 4.3.1 were primarily followed in a GIS environment to estimate the average and/or median RLMA&SI design point rainfall for the same 10 circular test catchments considered in Section 4.2. In order to simplify the interpretation of the steps followed, all GIS functions are highlighted using *Italic* font.

4.3.1   Estimation of average RLMA&SI design point rainfall

The following steps were followed:
   (a) **Step 1:** Decide on 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 4.3.
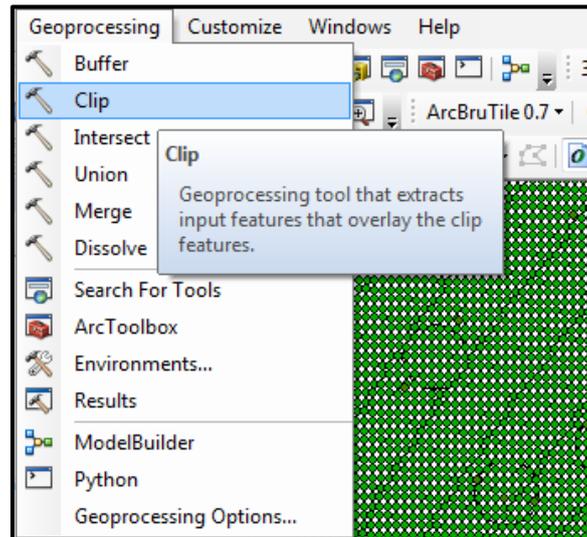
Figure 4.3: Snapshot of the *Clip* function in GIS

(c) **Step 3:** All clipped grid points need to be associated with their respective circular catchments using the *Join* function (*cf.* Figure 4.4), i.e. grid points are joined with a specific circular catchment to result in a summative attribute table for export purposes to MS Excel.
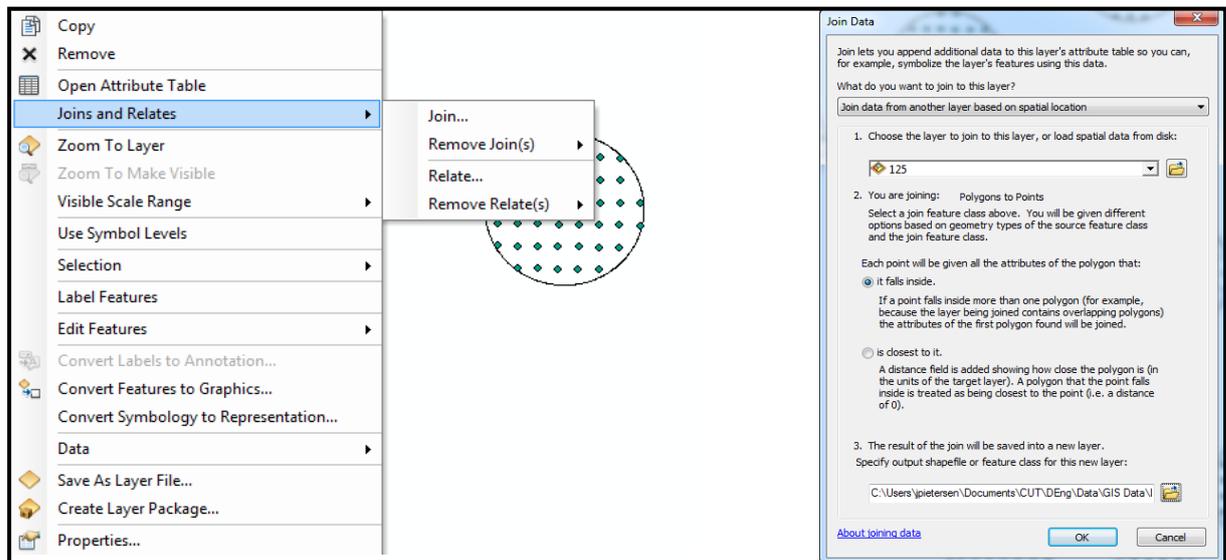
Figure 4.4: Snapshot of the *Join* function in GIS

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

(e) **Step 5:** Insert extra columns (*cf.* Figure 4.5) 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

Figure 4.5: Conversion of decimal degrees to degrees and minutes

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

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

4.3.2 Concerns when deriving ARFs using the RLMA&SI approach

Areal design rainfall values for the same 10 circular test catchments as considered in Sections 4.2 and 4.3, 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 and RLMA&SI design point rainfall values. The results are listed in Table 4.3.

Table 4.3: Sample ARFs (RLMA&SI-based) compared to the sample ARFs estimated using similar international methodologies for an area of 125 km^2

Method	RLMA&SI sample ARFs [%]				Siriwardena & Weinmann (1996) sample ARFs [%]			Podger et al. (2015a; 2015b) sample ARFs [%]			
T [years]	Storm duration [days (d) or hours (h)]										
	1-d	3-d	5-d	7-d	24-h	72-h	120-h	1-d	3-d	5-d	7-d
2	84	89	90	91	91	97	98	96	98	98	98
5	83	85	86	86	91	96	98	96	97	98	98
10	83	82	82	82	91	96	97	95	97	97	98
20	83	79	79	78	91	96	97	95	97	97	97
50	82	76	74	72	91	95	97	95	97	97	97
100	82	73	71	69	90	95	96	95	96	96	97
200	82	71	67	65	90	95	96	94	96	96	96

As shown in Table 4.3, it was established that the latter estimated sample ARFs are in contradiction to the ARFs estimated using various international methodologies, e.g. Siriwardena and Weinmann (1996) and Podger et al. (2015a; 2015b), which highlighted that ARFs typically increase with an increase in both the return period (T) and storm duration. In the case of the 10 circular catchments (125 km² 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 station under consideration.

This contradictive 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, and (ii) GEV$_{LM}$ areal and design point rainfall values originating from identical record lengths for both the observed areal and point rainfall data sets. The results are listed in Table 4.4 and clearly highlight the substantial differences between the two approaches, especially at higher return periods and storm durations.

Table 4.4: ARFs (RLMA&SI-based) compared to ARFs (GEV$_{LM}$-based) using identical record lengths and areas

T [years]	ARFs (identical record lengths) [%]				ARFs (RLMA&SI design point rainfall) [%]			
	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	85.5	89.1	89.6	82.1	83.4	84.6	86.2
10	85.9	87.8		89.1	85.8	82.7		82.5
20	90.4	90.6		88.4	90.9	83.0		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

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 annual maximum series (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 0590028W, having at least 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, strictly speaking, the ratios between the RLMA&SI and GEV_{LM} design point rainfall values (*cf.* Table 4.5) should equate one, i.e. ARF = 100%. However, this was not the case given that the RLMA&SI approach is based on a regionalisation scheme and associated multiple rainfall stations.

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.

Table 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

T [years]	ARFs [%] (GEV_{LM} areal AMS vs. RLMA&SI design point rainfall)				ARFs [%] (GEV_{LM} areal AMS vs. point AMS values)			
	Storm duration [days (d)]							
	1-d	3-d	5-d	7-d	1-d	3-d	5-d	7-d
2	102.1	103.0	102.2	102.4	98.6	100.0	100.0	100.0
5	101.5	101.8	101.2	102.0	98.7	100.0	100.0	100.0
10	100.2	99.8	99.6	101.2	99.1	100.0	100.0	100.0
20	98.4	97.5	97.8	100.3	99.5	100.0	100.0	100.0
50	95.7	94.1	95.3	98.9	100.0	100.0	100.0	100.0
100	93.5	91.5	93.2	97.7	100.0	100.0	100.0	100.0
200	91.2	88.7	91.2	96.6	100.0	100.0	100.0	100.0

As shown in Table 4.5, the latter comparison 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 areal and design point rainfall values using the same rainfall stations and record lengths with mutual starting and ending dates. This will ensure that extreme values, whether floods or droughts, are incorporated in the probabilistic analyses of areal and point rainfall.

Based on the above findings, it was decided (as previously done in Table 4.3), to compare the latter and confirmed ARF methodology with the international ARF methodologies of Siriwardena and Weinmann (1996) and Podger et al. (2015a; 2015b). As a result, sample ARFs were estimated in circular catchments of 60 km^2 and 180 km^2 each and with identical record lengths. The results are listed in Table 4.6.

Table 4.6: Comparison of sample ARFs estimated using identical record lengths and different methodologies

Method	RLMA&SI sample ARFs [%]				Siriwardena & Weinmann (1996) sample ARFs [%]			Podger et al. (2015a; 2015b) sample ARFs [%]			
T	Storm duration [hours (h) or days (d)] @ 60 km^2										
[years]	1-d	3-d	5-d	7-day	24-h	72-h	120-h	1-d	3-d	5-d	7-d
2	88	92	95	97	90	96	98	94	96	97	97
5	90	93	96	98	90	96	97	94	96	97	97
10	91	94	98	98	90	95	97	94	96	96	97
20	91	96	100	100	90	95	96	94	96	96	97
50	92	100	100	100	90	95	96	94	96	96	97
100	92	100	100	100	89	94	96	94	96	96	97
T	Storm duration [hours (h) or days (d)] @ 180 km^2										
[years]	1-d	3-d	5-d	7-day	24-h	72-h	120-h	1-d	3-d	5-d	7-d
2	82	94	95	96	91	95	96	96	98	98	98
5	86	96	97	98	90	94	96	96	97	98	98
10	87	97	98	98	90	90	95	95	97	97	98
20	87	97	98	99	90	94	95	95	97	97	97
50	88	98	99	100	90	94	95	95	97	97	97
100	89	99	99	100	89	93	95	95	96	96	97

It is clearly evident from Table 4.6 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 areal and design point rainfall, associated with specific return periods and storm durations, are within the same range. However, the ARFs in this study tend to increase with increasing return periods, as opposed to the Australian ARFs, which decrease with increasing return periods. The increase of ARFs with increasing return periods are to be expected, since higher magnitudes of rainfall (high *T*-values) will most probably cover a larger portion of the catchment. Furthermore, it is also assumed that a larger portion of a catchment will be covered when subjected to longer storm durations. The latter assumptions are clearly visible from the sample ARFs estimated in this study. 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   Thiessen Polygon Limitations

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. The average point and/or areal rainfall estimated using the Thiessen polygon method becomes questionable when applied on a low-density rainfall-monitoring network. 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 study, 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 used in several international ARF studies, e.g. Bell (1976), Stewart (1989), Siriwardena and Weinmann (1996), and Podger et al. (2015).

4.5   Pre-defined Fixed versus User-defined Circular Areas

The size and shape of a catchment are important variables in the estimation of ARFs and it is therefore important to use representative catchment sizes during the rainfall data analyses. The original methodology proposed 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 sizes is regarded as problematic in areas with a low-density rainfall-monitoring network. As mentioned in Section 4.4, the Thiessen weighted spatial distribution of a rainfall event is not always entirely a true reflection of the actual spatial distribution, especially when applied in areas with less rainfall stations. Therefore, to visualise the true spatial distribution of any rainfall event over a specific area becomes problematic.

Moreover, Figure 4.6 was created to highlight the concerns associated with the use of 'pre-defined fixed areas' as opposed to 'user-defined circular areas'.
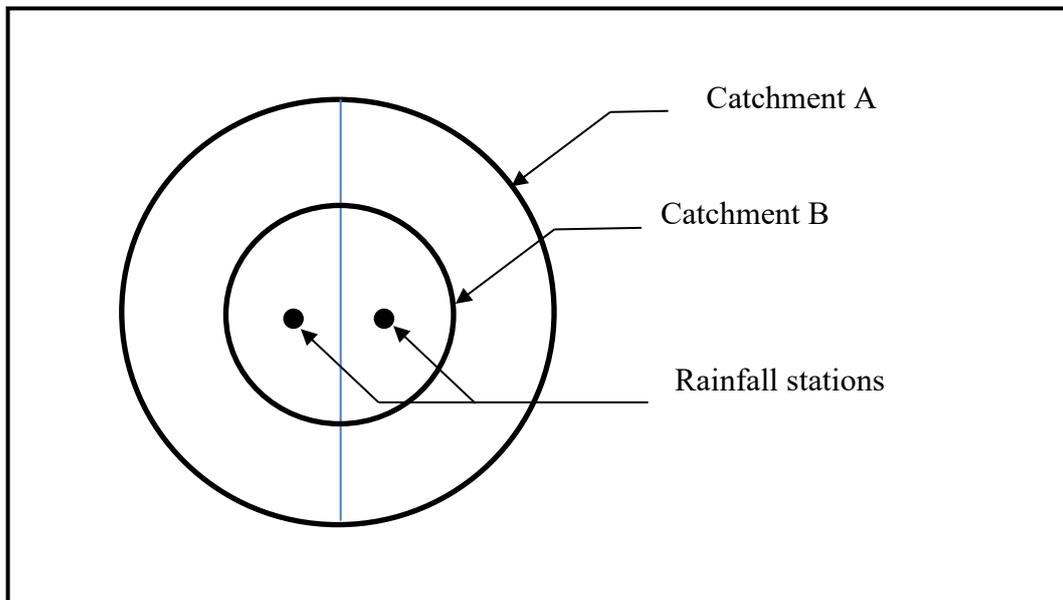
Figure 4.6: Illustration of identical Thiessen weights applicable to different catchment sizes

It is clearly evident from Figure 4.6 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.

Therefore, due to all the concerns associated with the fixed-area circular catchment approach, a 'user-defined' approach, which enables the adjustment of circular catchment areas in such a way that the circumference of a circle will simply overlay the most distant contributing rainfall stations, was adopted.

As shown in Figure 4.7, this ensured that the circular catchment sizes are determined by the spatial location of the contributing rainfall stations, thereby improving the representative spatial distribution of a rainfall event.
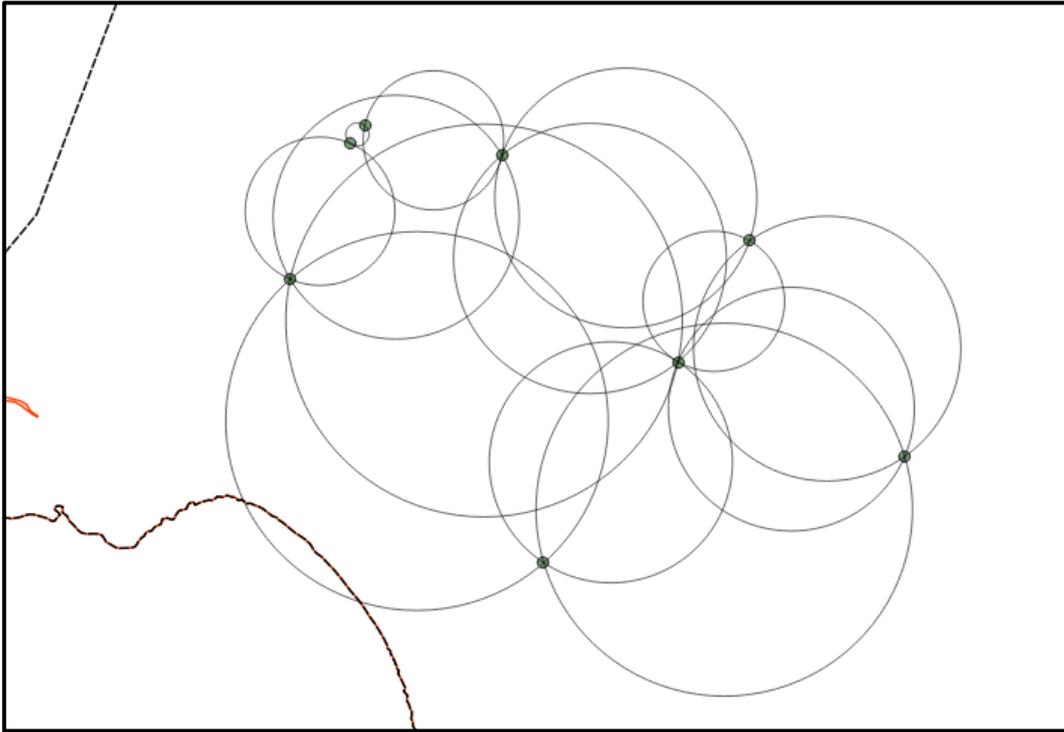
Figure 4.7: Illustration of user-defined circular catchments

4.6 Probabilistic Analyses using Short Record Lengths

The GEV$_{LM}$-based probabilistic analyses of areal and point rainfall values seem acceptable when considered independently; however, when the ratios of higher frequencies (>100 years) between areal and design point rainfall values are considered, then inconsistencies are evident. An example of sample ARFs estimated in a circular catchment of 2 475 km^2 and using a record length of 41 years, is listed in Table 4.7.

Table 4.7: 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
200	104.9	97.9	100.3	98.4

In Table 4.7, it is clearly evident that the sample ARFs increase with both an increase in return period (*T*) and storm duration, while values exceeding 100% are typical at higher frequencies

(T >100 years). Hence, record length does have an impact on the probabilistic analyses and therefore the general rule-of-thumb, i.e. limiting the frequency magnitude to $2N$, where N equals the AMS record length, is recommended. In other words, if the AMS equals 50 years, then T up to 1:100-year could be reasonably estimated.

4.7   Outliers

Excluding data points from a data set should not be taken lightly; however, it might sometimes be necessary to exclude outliers. In order to distinguish outliers from normal data points, requires the estimation of ordinary and subsequent standardised residuals. The ordinary residuals obtained in a regression typically reveals the distance between the observed data points and the predicted data points, whereas standardised residuals, are defined for each observation as an ordinary residual divided by an estimate of its standard deviation. For example, the original linear regression analysis conducted in Region 1 of the study area included 1 960 data points, with the multiple coefficient of determination (R_i^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 $R_i^2 = 0.61$. Thus, the removal of outliers is recommended. By removing outliers, the robustness of a regression model would not necessarily be improved, but it would result in a better fit between the observed and estimated values.

An overview of the concerns related to the estimation of catchment response time parameters are discussed in the next chapter.

CHAPTER 5: CATCHMENT RESPONSE TIME – CONCERNS AND SOLUTIONS

It is essential that the reader is familiar with the comprehensive literature review applicable to catchment response time parameters as included in Chapter 3. This chapter mainly focusses on the concerns and/or problems encountered during the estimation of catchment response time parameters, with possible solutions where applicable.

5.1   Establishment of Catchment Variable Database

As highlighted in the Inception Report (August 2019) and recommended by the Reference Group, information and data sharing took place between this project and WRC Projects K5-2748 (Calitz and Smithers, 2020) and K5-2923 (Du Plessis et al., 2020), respectively.

Dr JP Calitz (Principal Researcher, WRC Project K5-2748) was very helpful in this regard and provided a comprehensive catchment variable/parameter database based on the Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM) data for Southern Africa at 30 metre resolution and applicable to 322 gauged catchments scattered throughout South Africa.

5.2   Extraction and Analysis of Flood Hydrograph Data

In considering the original methodology, it was initially envisaged that a flood database need to be established for the 322 gauged catchments scattered throughout South Africa. However, given the project timeframe and budget, in conjunction with the various problems experienced with the streamflow data acquisition, management and analysis, the Reference Group subsequently recommended that the study scope be reduced. Subsequently, the catchment response time study was reduced to Primary Drainage Region X, given that:

(a) The flow-gauging stations in Region X generally have better and complete data sets for which the DWS have done some stage-discharge extensions;

(b) Complete hydrographs need to extracted and analysed in the 51 gauged catchments to support the development of a Hydstra hydrograph extraction process that can be tested and applied in other catchments; and

(c) DWS should be approached to consider developing the software utility for hydrograph extraction as part of the DWS-Hydstra Maintenance Agreement.

Background information, to clarify the above recommendations, and to elaborate on Points (a) to (c), is provided in the subsequent paragraphs.

At the project proposal stage, the local Hydstra Support Team (hosted by DWS) confirmed their assistance with data analyses and hydrograph extraction at no additional costs. However, as the project evolved, it became evident that DWS does not have the in-house expertise to assist the Project Team. As a result, a consultation process started in May 2020 with the Hydstra Support Team in Australia (Kisters, 2019) to solve most of the issues pertained to automation using a combination of different software scripts.

In essence, Kisters (2019) indicated that they will be able to automate the whole process of extracting complete flood hydrographs as detailed in Chapter 7. However, despite from having insufficient funds available, Kisters (2019) highlighted that the envisaged software tool will not be runnable from the web, at least not until DWS run a Hydstra-based website, while it will never be a non-Hydstra program, as it uses Hydstra tools heavily. Therefore, DWS was requested to take ownership/responsibility and co-fund/fund the envisaged software development in the future. Otherwise, even having the software tool, researchers will always have to rely on DWS (Hydstra) to run the analysis and provide the results. Luckily, DWS agreed in-principle to support the future development of a software utility for hydrograph extraction as part of the DWS-Hydstra Maintenance Agreement with Kisters, Australia.

Consequently, the study focussed on the development of semi-automated routines using Excel/VBA scripting to extract complete hydrographs in Primary Drainage Region X to support the future development, testing and verification of the above hydrograph extraction utility at catchment level. In other words, the comprehensive results obtained from this study will be used as benchmark to inform the envisaged development, testing and verification of the software utility.

Apart from the automation and funding problems highlighted above, several technical problems were encountered as detailed in the subsequent sections. Catchment A2H006 in Region A (prior to the decision to limit the study to Region X), was used to highlight these technical problems typically associated with DWS streamflow data.

5.2.1 Complex and multi-peaked hydrographs

In considering the primary flow hydrograph plot of A2H006 in Figure 5.1, identifying the rising limb and isolating flood events proved to be quite problematic and such multiple-peaked hydrographs were also encountered in Primary Drainage Region X.
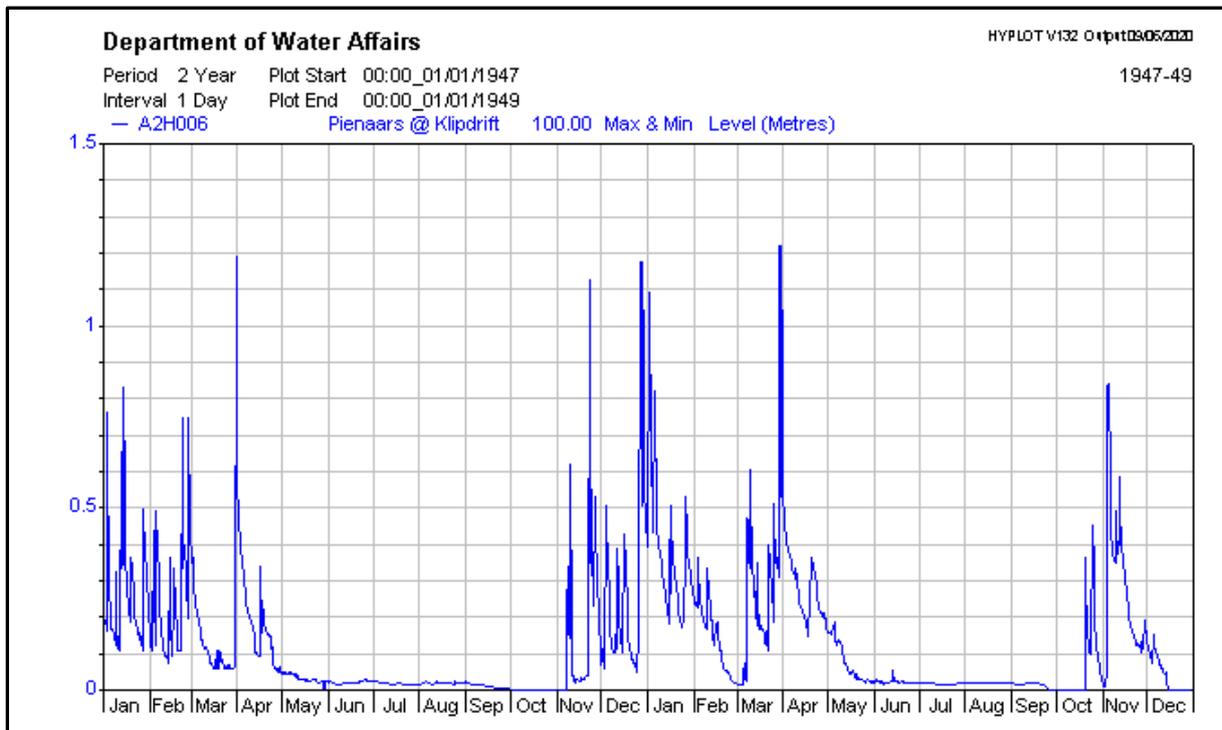
Figure 5.1: Primary flow hydrograph at A2H006

It was proposed that all similar multi-peaked hydrograph cases, with specific reference to Region X, be dealt with as follows:

In considering the flood events from November to end of February and March to the end of May in Figure 5.1, it is evident that all the events are in the same hydrological year and considering their similar peak flows, they would have been most likely also be included in the AMS set for $T > 2$-year. But these events are not necessarily independent, with similar peak flows, but different volumes due to different time to peak (T_P) values as a result of different soil moisture conditions. Hence, the ratio of peak flow: direct runoff volume (Q_{Pxi}/Q_{Dxi}) will be different in each case and also associated with different T_P values. Therefore, if the results using the HYBASE function in Hydstra or Eq. (7.1) in Chapter 7 are plotted on the same graph, then the 'start' and 'end' of each potential flood hydrograph could be established where the total discharge = baseflow discharge. In other words, where the difference between these two

variables = 0, or where these differences are a minimum, not necessarily zero. Hence, then three (3) possible flood hydrographs could be extracted: (i) November-end December, (ii) January-end February, and (iii) March-end May.

In considering the three distinctive flood hydrograph periods above, the cumulative volume of direct runoff under the hydrograph rising limb (Q_{DRi}; refer to Chapter 7, Section 7.2) would be the volume between the 'start' and the peak flow. This latter volume will then be used to determine the 'shape parameter' [Eq. (7.3a), Chapter 7] associated with the actual hydrograph, but converted and simplified into a 'triangular approximated hydrograph' (*cf.* Figure 7.3, Chapter 7). The time to peak (the main output required) is then estimated by using Eq. (7.3).

5.2.2   Stage-discharge rating tables

As highlighted in Chapter 3, s streamflow is seldom directly measured; instead, the stage (flow depth) is continuously measured and converted into a discharge utilising a stage-discharge rating curve at a flow-gauging weir. During flood events, flow-gauging weirs might be flooded with the water level beyond the structural limit. Subsequently, the standard calibration of the flow-gauging weir will no longer be a true reflection of the actual discharges that occurred during the flood events, and the standard stage-discharge rating curve must then be extended beyond the highest stage reading to reflect these high discharges

Based on the above background, the accuracy and relevance of the stage-discharge rating tables (DTs) on the DWS website is/was of a great concern, with the structural limit being exceeded in many cases in Region A. For example, in considering the AMS of catchment A2H006, there are three DTs available: 1905 to 1938, 1939 to 1990 and after 1990. The AMS data are quality coded in some cases as 'exceeding >', but according to the DTs, the limits vary between 5.44 m and 6.4 m. Hence, no exceeding actually took place, but still the '>' quality code.

Various approaches to address the above problems, can be considered, as outlined below:

If the 'Hydstra approach' is to be implemented to correct/extend DTs, the use of HYAUDIT for '*Data within Rating*' must be used to address any problems permanently. As an alternative, DTs can be extended beyond the structural limit using the *'temporarily extend the rating by projecting it up in log-log space as part of the analysis'*. However, such an extension should be carefully considered, because the flow could go beyond 'bankfull conditions' and transitions

into the 'flood plain'. Since this cannot be done in detail for each site, additional verification measures, whether direct or indirect methods, should be used.

If the 'alternative manual approach' is to be implemented to correct/extend DTs, it is suggested that the streamflow data to be extended using a 3rd order polynomial relationship up to 20%. Verification of the extension to +20% (or more) can then be done by considering the hydrograph shape, especially the peakedness as a result of a steep rising limb in relation to the hydrograph base length, and the relationship between individual peak discharge (Q_{Pxi}) and direct runoff volume (Q_{Dxi}) pair values. Typically, in such an event, the additional volume of direct runoff (Q_{DE}) due to the extrapolation should then be limited to 5%, i.e. $Q_{DE} \leq 0.05\ Q_{Dxi}$. Hence, the error made by using larger direct runoff volumes will have little impact on the sample statistics of the total flood volume.

Given that there is no 'one size fits all' approach/method when it comes to the extension of DTs, the 'alternative manual approach' was adopted in a very few cases, given that DWS had already verified and extended the DTs in in Primary Drainage Region X. In addition, a selection of indirect extension methods (e.g. hydraulic and one-dimensional modelling methods) was evaluated and compared to direct extension (benchmark) methods (e.g. at-site conventional current gaugings, hydrograph analyses and level pool routing techniques), to establish the best-fit and most appropriate stage-discharge extension method to be used in South Africa. Please refer to Chapter 7, Section 7.5 for the methodological approach adopted and the results.

5.2.3 Quality code 91 flow data

Quality code 91 (Q91) flow data differ somehow from the DT values being exceeded as discussed in Section 5.2.2. Typically, Q91 is associated with a period of missing data regraded as having a peak flow value which is higher than the last recorded point (minimum recorded height of a potential peak), but nobody knows by how much the peak was exceeded.

Q91 is characteristic of flow-gauging stations where mechanical recorders (charts) are still being used, but could also occur in electronic data sets.

Typically, Q91 occurs when:

(a) The pen of the recorder/chart got stuck; and/or

(b) The gauge plate readings were taken at a flow-gauging weir up to the structure limit. Anything above it, was read at an upstream gauge plate at a bridge or similar site. In these cases, there are two data sets that should be merged. In the past when mechanical recorders were commonly used, DWS considered to further investigate this due to the additional problems being introduced when different data sets are merged, but this never happened.

Based on the above, and for the purpose of this project, all Q91 data periods were excluded from the analyses.

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

CHAPTER 6:   AREAL REDUCTION FACTORS – 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, as well as the recommendations made by the Reference Group members, were considered and implemented while progressing with and adopting the final methodology.

6.1   Homogeneous Rainfall Regions

The revised and recommended methodology involved the 46 homogeneous rainfall regions (*cf.* Figure 6.1) as derived from the current 78 homogeneous rainfall regions as established by Smithers and Schulze (2004). The rationale for utilising the 46 homogeneous rainfall regions was explained and justified in Chapter 4, Section 4.1.
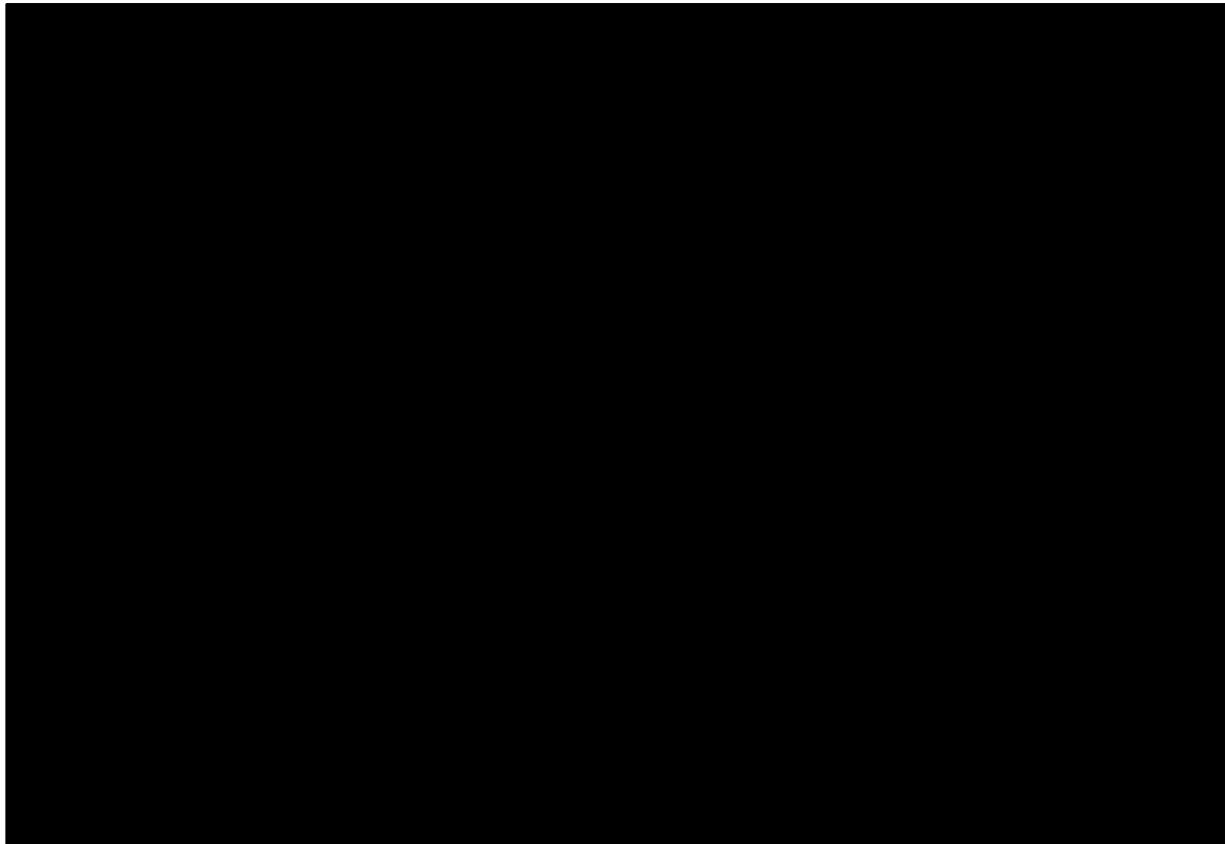
Figure 6.1:   46 Homogeneous regions delineated from 78 homogeneous rainfall regions

6.2   User-defined Circular Catchments

An alternating-area approach was used to generate the artificial circular catchments within each of the 46 homogeneous rainfall regions, which are referred to as the '46 ARF regions' in the remainder of the document. Given the more suitable functionalities available in the Quantum Geographical Information System software (QGIS; 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 catchment depended on the locality of the two most distant rainfall stations. This process required the optimum, manual positioning of circular catchments between the identified rainfall stations, which subsequently resulted in artifilial circles consisting of various sizes.

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 ARF regions, 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 region was limited by and based on the maximum number of rainfall stations available in each region.

(e) The Thiessen polygon grid (Step (b)) and all the regional circular catchments 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.

(g) All the above steps resulted in: (i) circular catchment sizes, (ii) contributing rainfall stations, and (iii) associated Thiessen weights, for the purpose of extracting and infilling daily rainfall data (*cf.* Chapter 4, Section 4.3).

A total of 2 550 artificial circular catchments (*cf.* Figure 6.2), ranging from 0.07 km^2 to 18 985 km^2, were manually placed across South Africa.
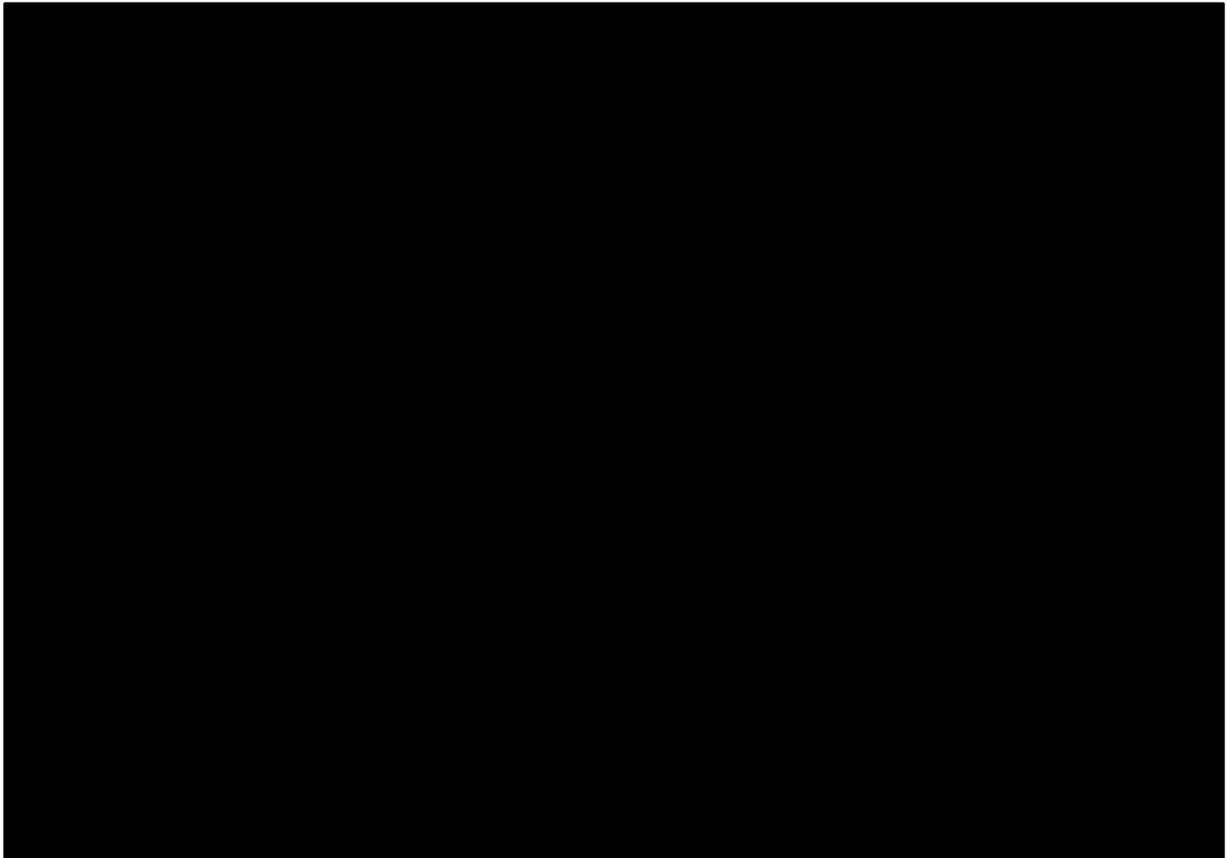
Figure 6.2: Placement of 2 550 artificial circular catchments in South Africa

The 'open spaces' evident between the circular catchment areas in Figure 6.2 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 (*c.f.* Chapter 4, Section 4.3.1) need to be maintained. The total number of rainfall stations present in each of the 46 ARF regions varied from 14 to 102 stations, while the total number of circular catchments per region varied from 23 to 100 (*cf.* Table 6.1).

Table 6.1: Regional circular catchment information

ARF region	Size [km²]	Number of rainfall stations	Number of circles	Circle sizes [km²]	
				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

ARF region	Size [km²]	Number of rainfall stations	Number of circles	Circle sizes [km²]	
				Minimum	Maximum
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		

6.3   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 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 Chapter 4 (*cf.* Section 4.2) resulted in multiple regional 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 ARF regions, the overlapping of circular

catchments was evident. Consequently, this resulted in daily rainfall data from similar rainfall stations being used multiple times within a particular ARF region. In principal, this was not regarded as problematic, while it also contributed to the 'smooth' transition between the different regions.

6.3.1 Criteria for infilling of missing 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 infilling techniques (e.g. inverse distance weighting, expectation maximisation, median ratio and/or monthly infilling) as proposed by Lynch (2004), were used for the infilling of missing daily rainfall data. The rainfall data infilling process was carefully interrogated by considering the following criteria as **Criterion 1**:

(a) Infilling 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 ARF region were removed where 25% or more of the stations required infilling to a minimum of 30 years combined. Figure 6.3 is illustrative of a case where infilling was not considered; two rainfall stations with a combined record length of 11 years where 50% of the stations required infilling.

Nr	Station Name	Station Num	First Year Record	Last Year Record	Record length Low	High	Duration used
3	NELSPOORT	0093005_W	1950	1997	1950	1961	11
4	KROMRIVIER	0116029_W	1903	1961			

Figure 6.3: Example of circular catchment where infilling was rejected

Figure 6.4 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.

Nr	Station Name	Station Num	First Year Record	Last Year Record	Record length		Duration used
					Low	High	
215	HOUD CONSTANT	0095395_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	0119209_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			

Figure 6.4:   Example of circular catchment where infilling was applied

Hence, infilling was not only used to extend the rainfall data series at particular rainfall station(s), but it ensured that there are 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 catchment areal and point rainfall AMS. This approach was followed in all the 46 ARF regions.

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.

Table 6.2 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.

Table 6.2: Number of circular catchments removed for the purpose of probabilistic analyses

ARF region	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**

6.3.2   Averaging of daily rainfall data

The use of the Thiessen polygon method (Wilson, 1990) is justified due to the preferred use thereof in various international ARF studies, e.g. Bell (1976), Stewart (1989), Siriwardena and Weinmann (1996), and Podger et al. (2015a; 2015b).

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 (*cf.* Chapter 4, Section 4.2) 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 ARF regions, associated with different combinations of rainfall stations and Thiessen weights, were generated for South Africa.

6.3.3   Extraction of areal and point AMS

The areal and point AMS for durations of 1-day, 3-day, 5-day and 7-day were extracted for the purpose of the probabilistic analyses. The 1-day fixed time interval point and areal rainfall AMS were firstly obtained from the observed rainfall data. The point AMS for each rainfall station was firstly extracted and then each AMS was individually multiplied with a corresponding Thiessen weight, resulting in one weighted point AMS for each circular catchment. In terms of the areal AMS, the Thiessen weights were applied on a daily basis, i.e. daily point rainfall values from rainfall stations within a particular circular catchment were multiplied with a corresponding Thiessen weight to result in weighted areal daily AMS values.

In order 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. In terms of the areal AMS, the latter *n*-day totals within each hydrological year for a particular circular catchment, were multiplied with an appropriate Thiessen weight, which resulted in the required *n*-day areal AMS values. This procedure resulted in one weighted areal AMS for durations of 1, 3, 5 and 7 days.

6.4   Probabilistic Analyses of Weighted Areal and Point AMS

The probabilistic analyses were conducted on 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 4.5. 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).

6.5   Estimation of Sample ARFs

The estimation of ARFs was based on a 'modified version' of Bell's method (1976), 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. (6.1) and expressed as the ratio between the areal catchment design rainfall and average design point rainfall estimates for corresponding return periods.

$$ARF_{Sample} \;=\; \frac{A_{DR}}{P_{DR}} \times 100 \qquad\qquad [6.1]$$

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 the 46 ARF regions.

6.6 Derivation of Regional Empirical ARF Equations

Initially, linear backward stepwise multiple regression analyses with deletion were performed at a 95% confidence level in order to estimate the relationship between the dependent criterion variable (ARF) and the independent predictor variables within each region. Ultimately, as detailed below, the linear regression analyses were outperformed by a second-order polynomial non-linear log-transformed empirical ARF equation. 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]. Hypothesis testing was performed at each step to ensure that only statistically significant independent variables are retained in the model, while insignificant variables were considered for removal.

Furthermore, the Goodness-of-Fit (GOF) statistics of normal and log-transformed data were tested on each independent and dependant variable using the coefficient of correlation (r^2) and the standard error of the estimate (SE). In each case, the GOF statistics for log-transformed data outperformed the normal data. Therefore, log-transformed independent variables were used in the final regression analyses.

In order to identify and exclude outliers from the sample ARF estimations, standardised residuals were estimated for each of the 46 ARF regions. Estimated standardised residuals, which exceeded a minimum and maximum value of -2.5 and 2.5, respectively, were removed from the list of potential sample ARFs. The number of standardised residuals removed ranged from 6 to 50 (*cf.* Table 6.3) with a total of 1 263 (2.6%) ARF residuals $\geq \pm 2.5$ being removed.

Table 6.3: Standardised residuals removed from each ARF region

RLMA&SI long duration cluster	ARF region	Number of sample ARF estimates	Number of standardised residuals removed	Standardised 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
1	31	624	18	2.9
1	38	1 296	33	2.5
2	3	1 104	21	1.9

RLMA&SI long duration cluster	ARF region	Number of sample ARF estimates	Number of standardised residuals removed	Standardised residuals excluded [%]
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**

Thereafter, partial t-tests were used to test the significance of individual independent variables, while total F-tests were used to determine whether an ARF as a 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.

The log-transformed GOF statistics for the 46 ARF regions were also evaluated by using the r^2 and SE values. As shown in Table 6.4 and applicable to the 46 ARF regions, the r^2 and SE values ranged from 0.59 to 0.84 and 2.1 to 4.9%, respectively. In addition, it was evident that above GOF statistics remain practically the same when the 46 ARF regions within each of the seven RLMA&SI long duration clusters, are selectively merged. Subsequently, the individual ARF regions were merged, providing that the: (i) ARF regions are located within the same

RLMA&SI long duration cluster (*cf.* Figure 6.5), and (ii) GOF statistics of each merged region do not reduce significantly in comparison to the original GOF statistics obtained in each region.

Table 6.4:   GOF statistics before and after merging of ARF regions

RLMA&SI long duration cluster	ARF region	r^2	SE [%]	After merging of ARF regions			
				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

The merging of the regions as listed in Table 6.4 and shown in Figure 6.5, resulted in slightly lower and/or similar r^2 values (0.62 to 0.79) in comparison to those values obtained in the original (individual) ARF regions, i.e. $0.59 \leq r^2 \leq 0.84$. However, after merging, the *SE* values (2.3 to 3.8%) demonstrated even some improvement when compared to the original (individual) *SE* values in each region, i.e. $2.1\% \leq SE \leq 4.9\%$.
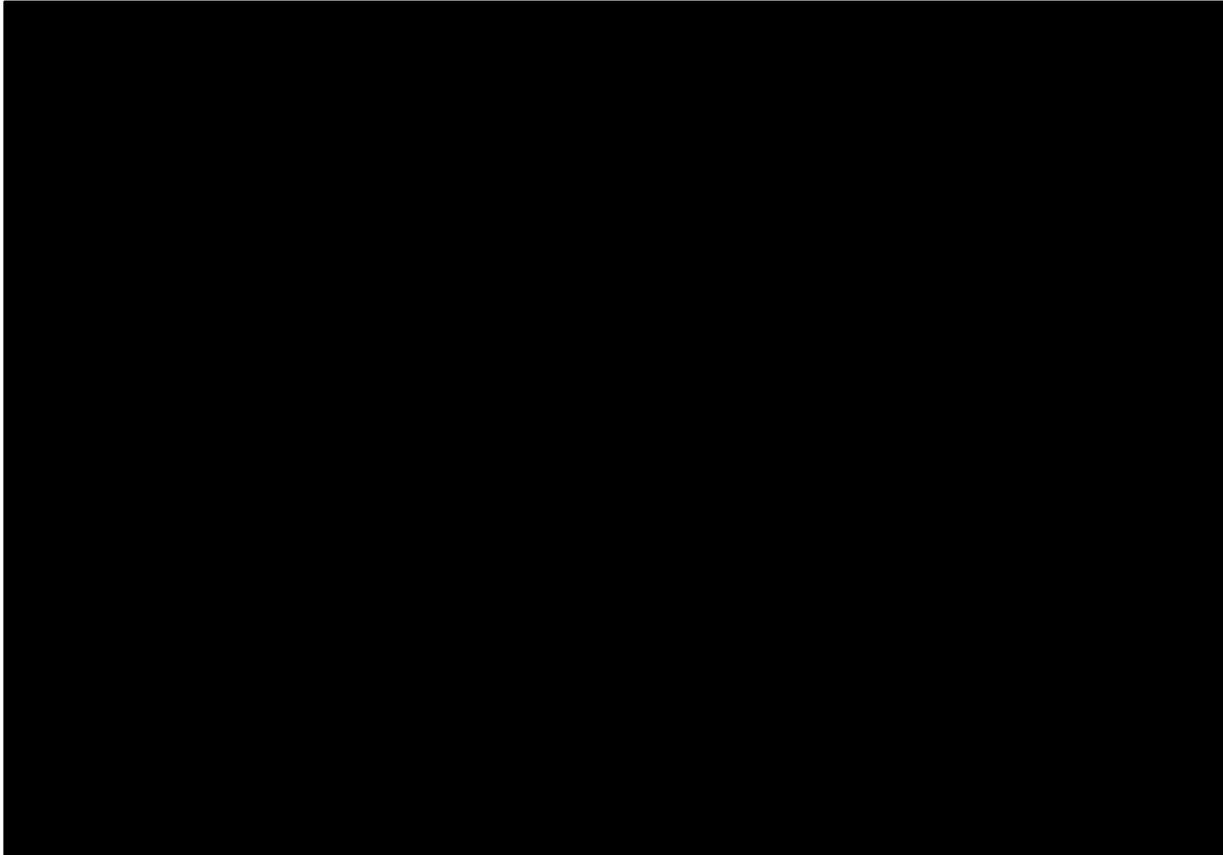
Figure 6.5:   Seven ARF regions associated with the RLMA&SI long duration clusters (after Smithers and Schulze, 2000b)

Given the limited number of original ARF regions (3 each) located within the RLMA&SI long duration Clusters 2 and 3, as well as their close geographical proximity to Clusters 5 and 7, the ARF regions in these clusters were subjected to a second round of merging. Hence, this resulted in the merging of Clusters 2 & 5 and 3 & 7 to result in ARF Regions 2 and 3, respectively. The second round of merging of the latter clusters was justified by the need for 'user-friendliness' and 'consistency'. Typically, Cluster 2 (*cf.* Figure 6.5) covers a very narrow strip along the east and south coast; hence, making it virtually impossible for the practitioner to accurately select an ARF region to estimate ARFs. In the case of Cluster 7 (*cf.* Figure 6.5), the limited number of rainfall stations (ratio of 1 078 km²/station) necessitates the required merging with Cluster

3 to improve the rainfall station density, i.e. 989 km²/station. Overall, the GOF statistics before and after merging proved to be very similar, e.g. r^2 differences of < 0.05 and *SE* differences < 1%.

The final five (5) ARF regions are shown in Figure 6.6.
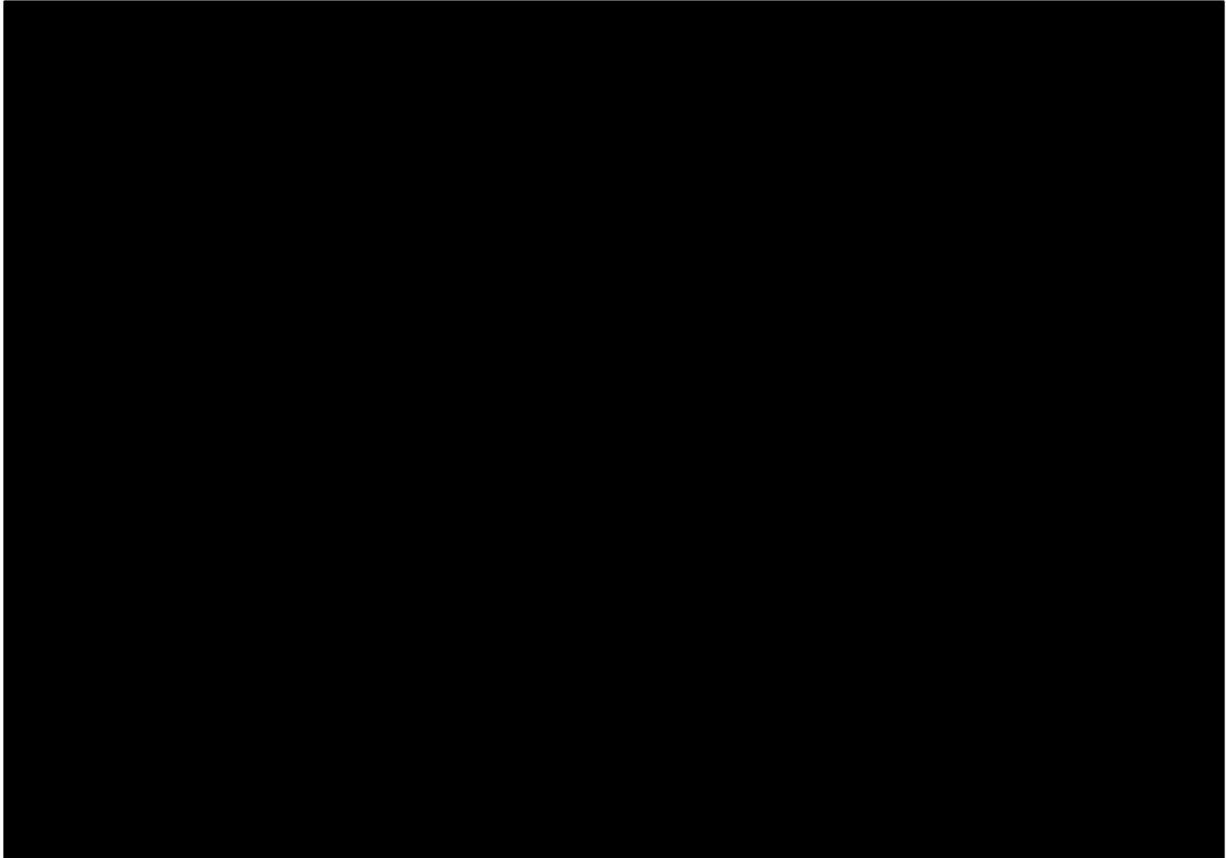
Figure 6.6:   Five (5) ARF regions applicable to South African rainfall

By considering the 'user-friendliness' of any derived empirical equation as very important, the empirical equation initially derived was in a simplistic, linear log-transformed format. However, scatter plots of the ARF_{Sample} [Eq. (6.1)] versus the estimated ARF_y values revealed several 'outliers' deviating from the 1:1 line, while the ARF_y estimates also had a tendency to overestimate the ARF_{Sample} values for ARF values < 80%. The latter overestimation would typically be encountered in larger catchment areas where lower ARF values are to be expected.

Subsequently, a second-order polynomial non-linear log-transformed empirical ARF equation [Eq. (6.2)] with unique regional calibration coefficients, was derived.

$$ARF_y \quad = aX^2 + bX - c \qquad [6.2]$$

$$X = x_1 (\log D)^2 + x_2 (\log D) - x_3 (\log T)^2 + x_4 (\log T) - x_5 (\log A)^2 - x_6 (\log A) + x_7 \quad [6.3]$$

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 6.5].

Table 6.5: 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. (6.2) resulted in improved GOF statistics (*cf.* Table 6.6) when compared to the original GOF statistics based on the linear regressions (*cf.* Table 6.4).

Table 6.6: 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

Scatter plots of the ARF$_{Sample}$ [Eq. (6.1)] and ARF_y [Eq. (6.2)] values associated with all the circular catchments located in ARF Regions 1 to 5 are shown in Figures 6.7 to 6.11 to highlight any differences.
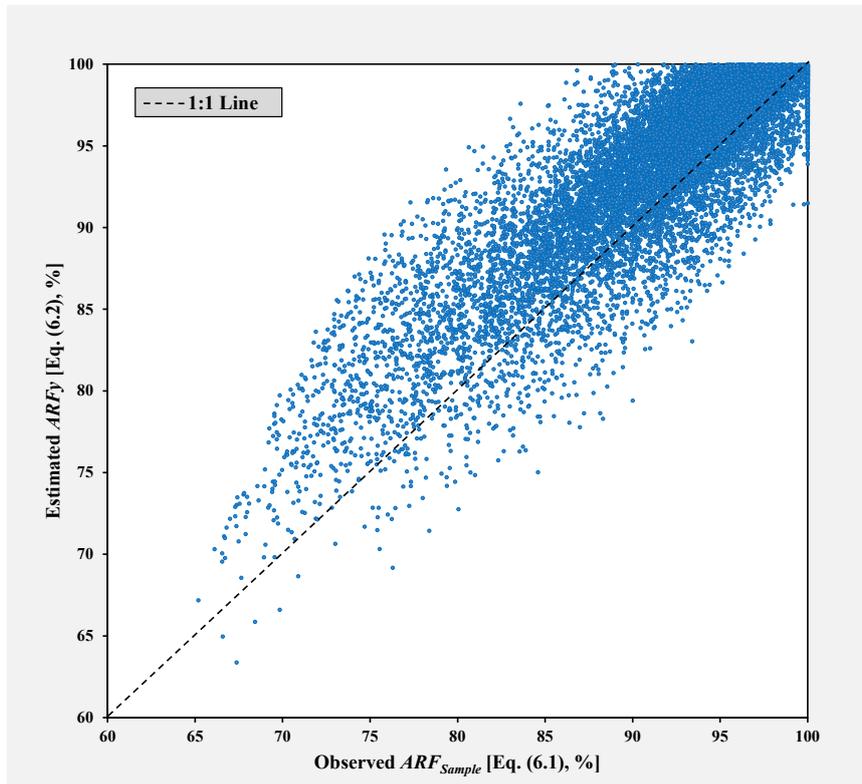
Figure 6.7: Scatter plot of the observed ARF_{Sample} [Eq. (6.1)] and estimated ARF_y [Eq. (6.2)] values in ARF Region 1
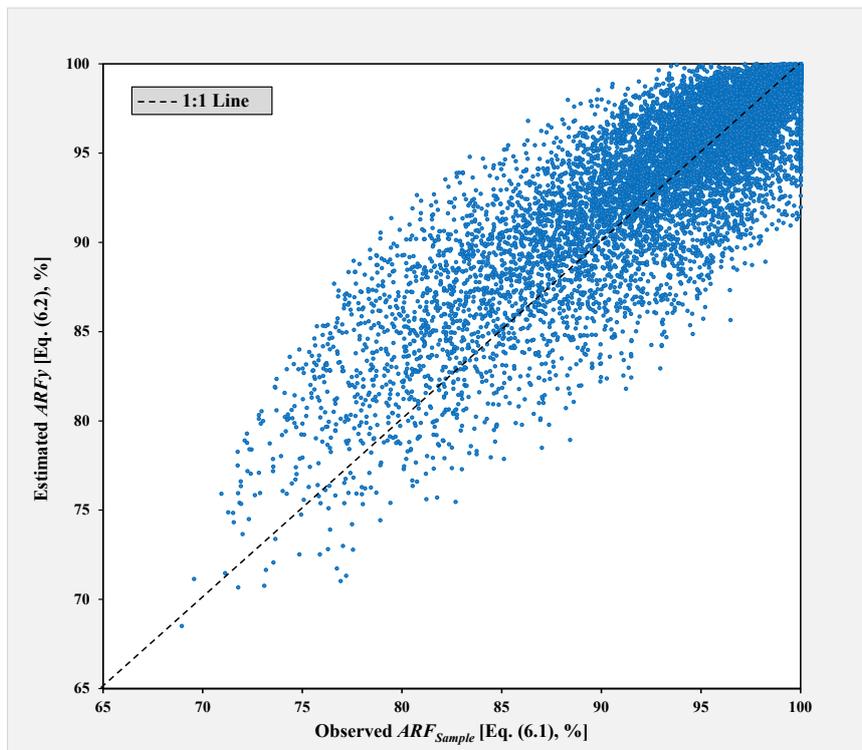
Figure 6.8: Scatter plot of the observed ARF_{Sample} [Eq. (6.1)] and estimated ARF_y [Eq. (6.2)] values in ARF Region 2
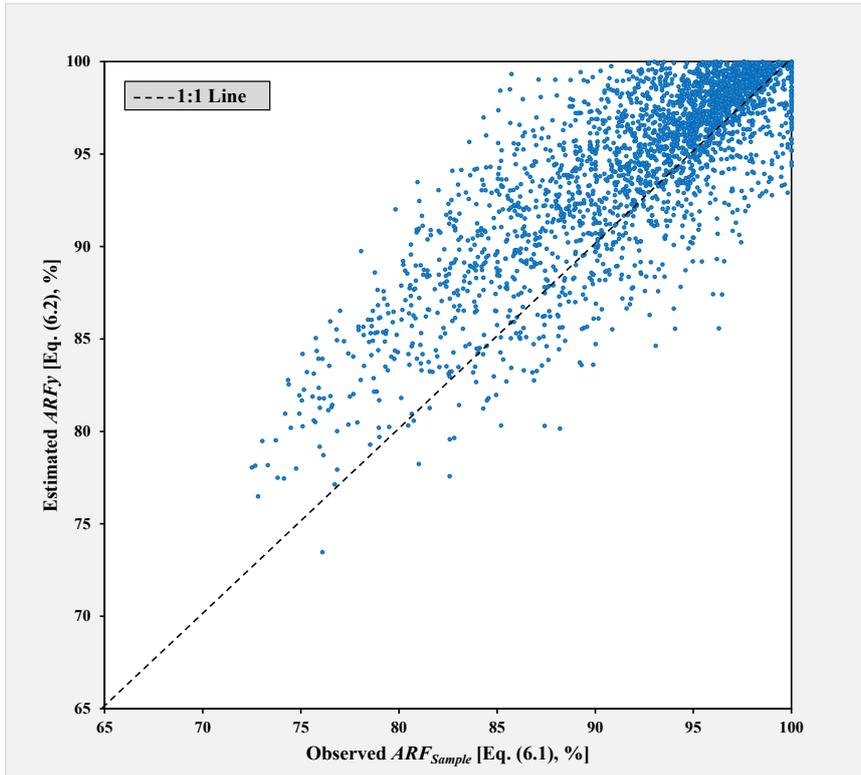
Figure 6.9: Scatter plot of the observed ARF_{Sample} [Eq. (6.1)] and estimated ARF_y [Eq. (6.2)] values in ARF Region 3
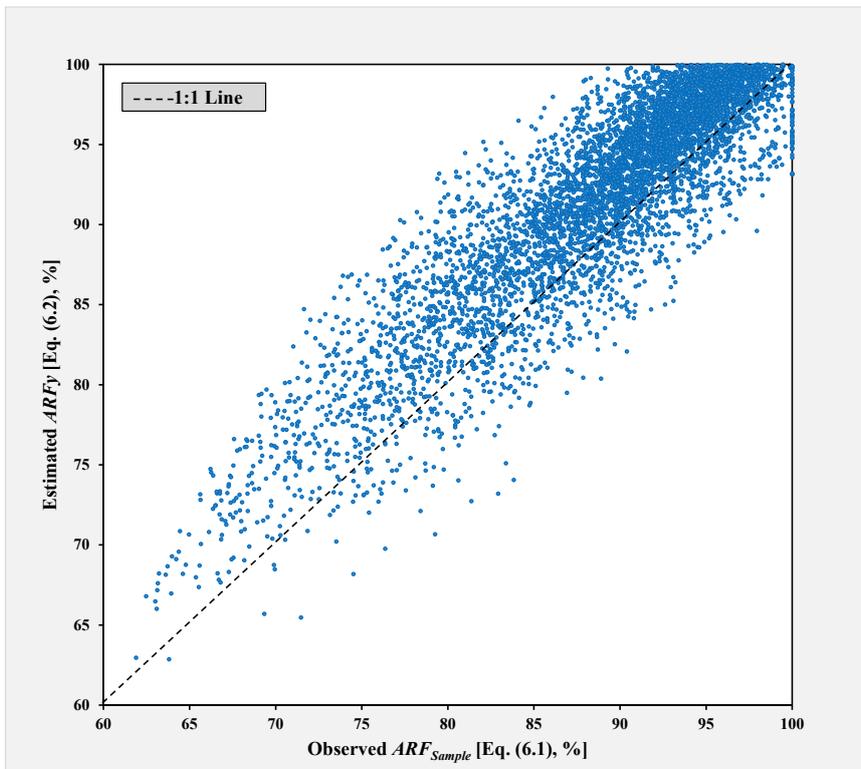
Figure 6.10: Scatter plot of the observed ARF_{Sample} [Eq. (6.1)] and estimated ARF_y [Eq. (6.2)] values in ARF Region 4
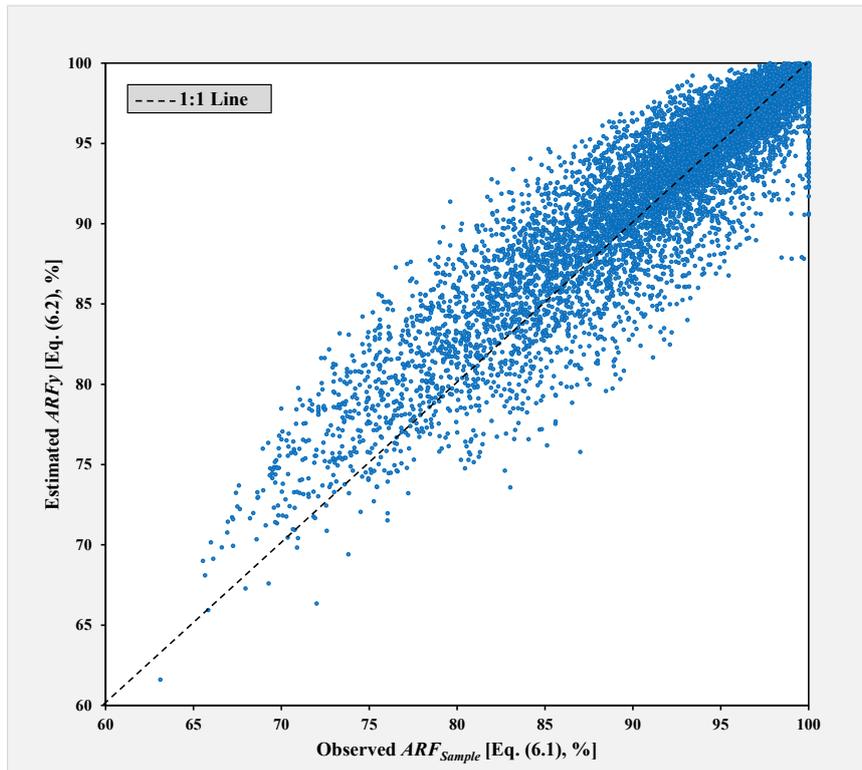
Figure 6.11: Scatter plot of the observed ARF_{Sample} [Eq. (6.1)] and estimated ARF_y [Eq. (6.2)] values in ARF Region 5

Overall, the plotting position and clustering of points in Figures 6.7 to 6.11, are regarded as a significant improvement from the original linear log-transformed equation considered. Furthermore, the distribution of the points in Regions 1 to 5 are also regarded as acceptable, with $0.74 \leq r^2 \leq 0.85$, and $2.9 \leq SE \leq 4.3\%$.

6.7   Comparison of ARF Equations

This section focuses on the comparison of Eq. (6.2) against the selection of geographically-centered ARF estimation methods as included in Chapter 2 and generally used in South Africa. 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 standard input variables and their associated ranges, e.g. catchment area (10 to 30 000 km²), 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. (2.5a)], and Alexander III [2001; Eq. (2.5b)]. All the results are listed in Table 6.7.

Table 6.7: Comparison between geographically-centred ARF estimation methods

Catchment area [km²]	Hours	Days	T [years]	NERC (Fig. 2.4)	Alexander I (Fig. 2.5)	OT&S [Eq. 2.5a)]	Alexander II (Fig. 2.6)	Alexander III [Eq. 2.5b)]	Region 1	Region 2	Region 3	Region 4	Region 5
				Current SA ARF methods [%]					**WRC K5-2924 — ARF Equation (6.2) [%]**				
10	24	1	2						88.1	88.9	90.7	87.2	86.7
			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
	48	2	2						94.3	94.5	96.3	94.6	93.2
			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
	72	3	2						96.4	96.3	98.1	97.2	95.4
			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
50	24	1	2						84.9	86.2	88.2	83.4	84.2
			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
	48	2	2						91.8	92.4	94.5	91.6	91.3
			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
	72	3	2						94.3	94.4	96.6	94.5	93.7
			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
100	24	1	2						83.0	84.6	86.9	81.4	82.7
			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
	48	2	2						90.4	91.1	93.5	90.0	90.1
			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
	72	3	2						93.0	93.4	95.7	93.1	92.7
			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
500	24	1	2						77.4	79.9	83.1	75.7	78.0
			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
	48	2	2						85.8	87.4	90.6	85.4	86.4
			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
	72	3	2						88.9	90.0	93.2	88.9	89.4
			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
1 000	24	1	2						74.3	77.3	81.1	72.8	75.4
			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
	48	2	2						83.3	85.3	89.0	83.0	84.3
			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
	72	3	2						86.6	88.1	91.8	86.7	87.5
			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
5 000	24	1	2						65.2	69.8	75.6	64.8	67.5
			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
	48	2	2						75.7	79.1	84.6	76.3	77.7
			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
	72	3	2						79.6	82.4	87.8	80.5	81.5
			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 [%]					WRC K5-2924 ARF Equation (6.2) [%]				
	Hours	Days		NERC (Fig. 2.4)	Alexander I (Fig. 2.5)	OT&S [Eq. 2.5a]	Alexander II (Fig. 2.6)	Alexander III [Eq. 2.5b]	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 6.7 decreased with an increase in catchment area. Similarly, an increase in ARF values with an increase in storm duration is also evident from the results. The latter is 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 estimates using Eq. (6.2) were similar to the other methods under consideration in the different catchment area and storm duration ranges. However, only Eq. (6.2) considers the variation of ARFs with return periods. ARFs increasing with an increasing return period is to be expected, since higher return period rainfall events as opposed to lower return period events, often tend to cover a larger surface area.

For example, in the 10 to 100 km² range, 2-year return period, and 1-day storm duration, the estimated ARFs [Eq. (6.2)] vary between 80% and 90%, while the ARFs estimated using the current South African methods, varied between 94% to 100%. Given that Eq. (6.2) 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 ($T = 50$-year) and storm durations ($D = 2$-day) are considered.

In Figures 6.12 to 6.16, the percentage differences between the current South African ARF estimation methods and Eq. (6.2) are shown. Similar trends as witnessed in Table 6.7 are evident, i.e. ARFs decrease with catchment area, and increase with an associated increase in both storm duration and return period. However, only Eq. (6.2) 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. Such higher ARF values require a higher degree of similarity between the areal design rainfall and average design point rainfall values. Given that these western parts of South Africa are semi-arid with highly variable rainfall, assuming such a more uniform temporal and spatial rainfall distribution, would be incorrect. Hence, the latter higher ARF values could only be ascribed to the low density of the rainfall-monitoring network in this particular region.
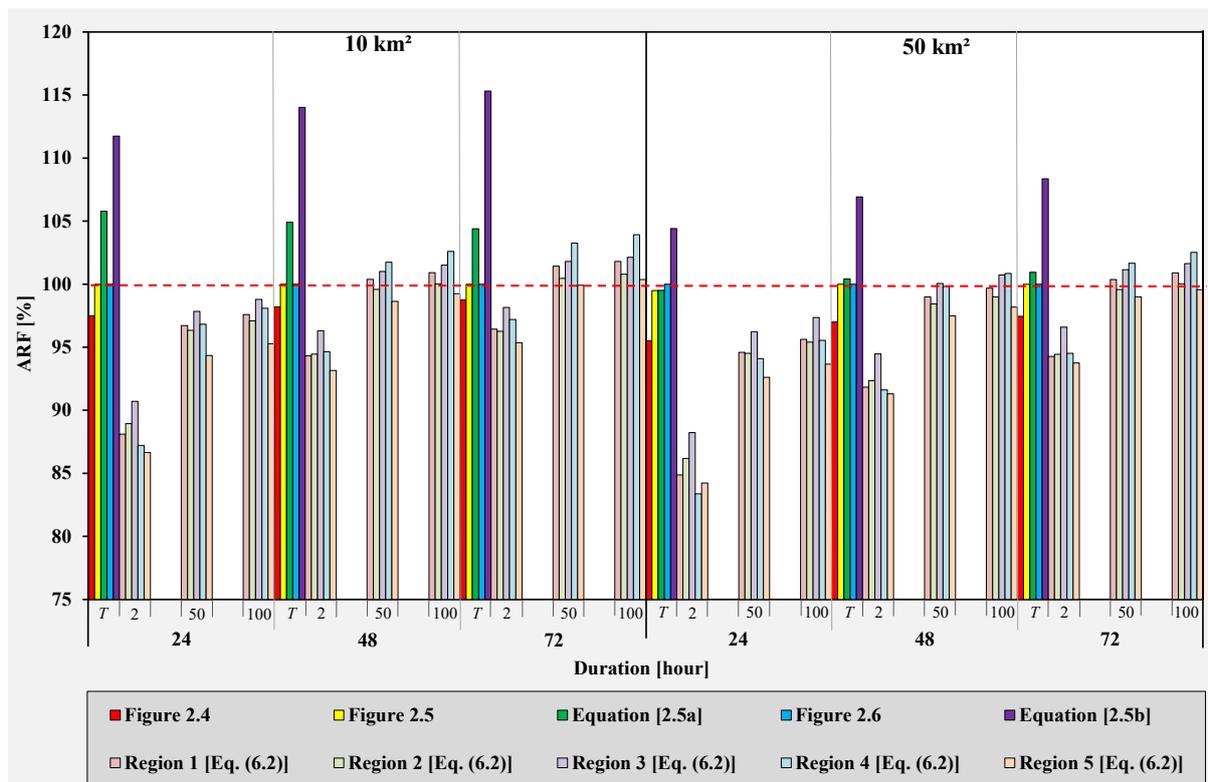
Figure 6.12: Comparison of ARF estimation methods (10 km² to 50 km²)
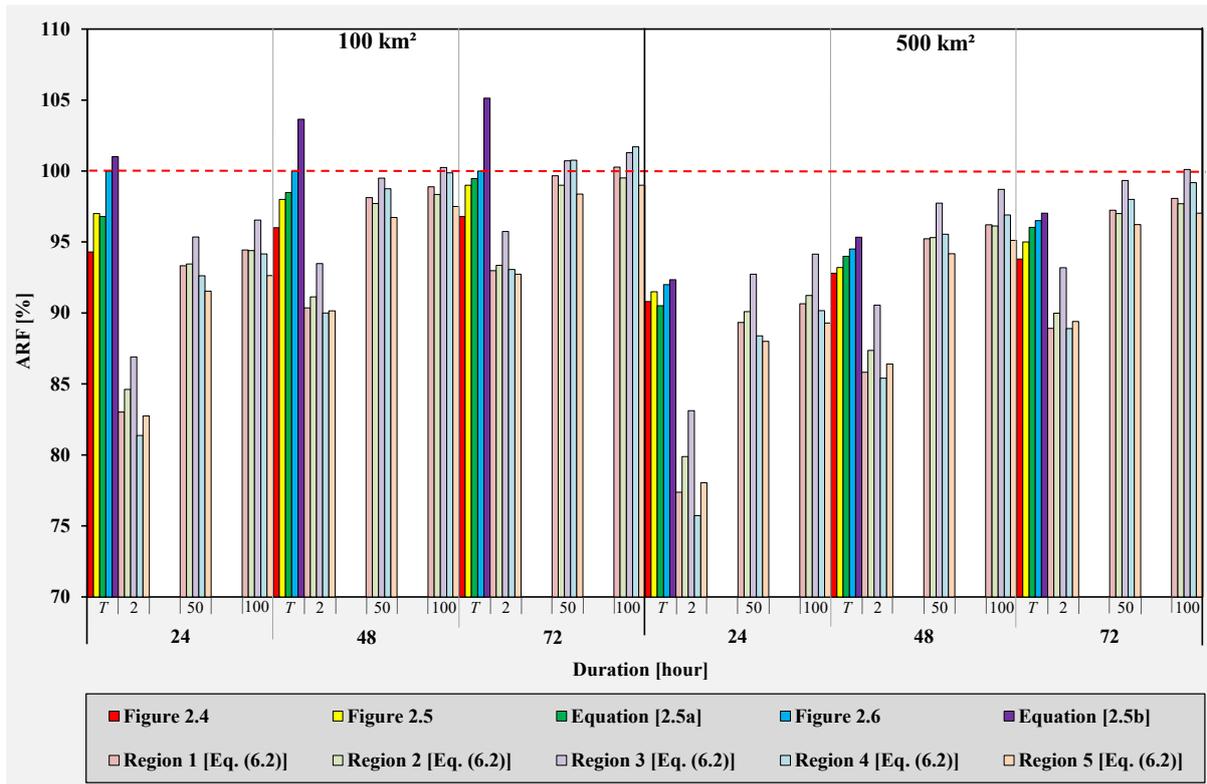
Figure 6.13: Comparison of ARF estimation methods (100 km² to 500 km²)
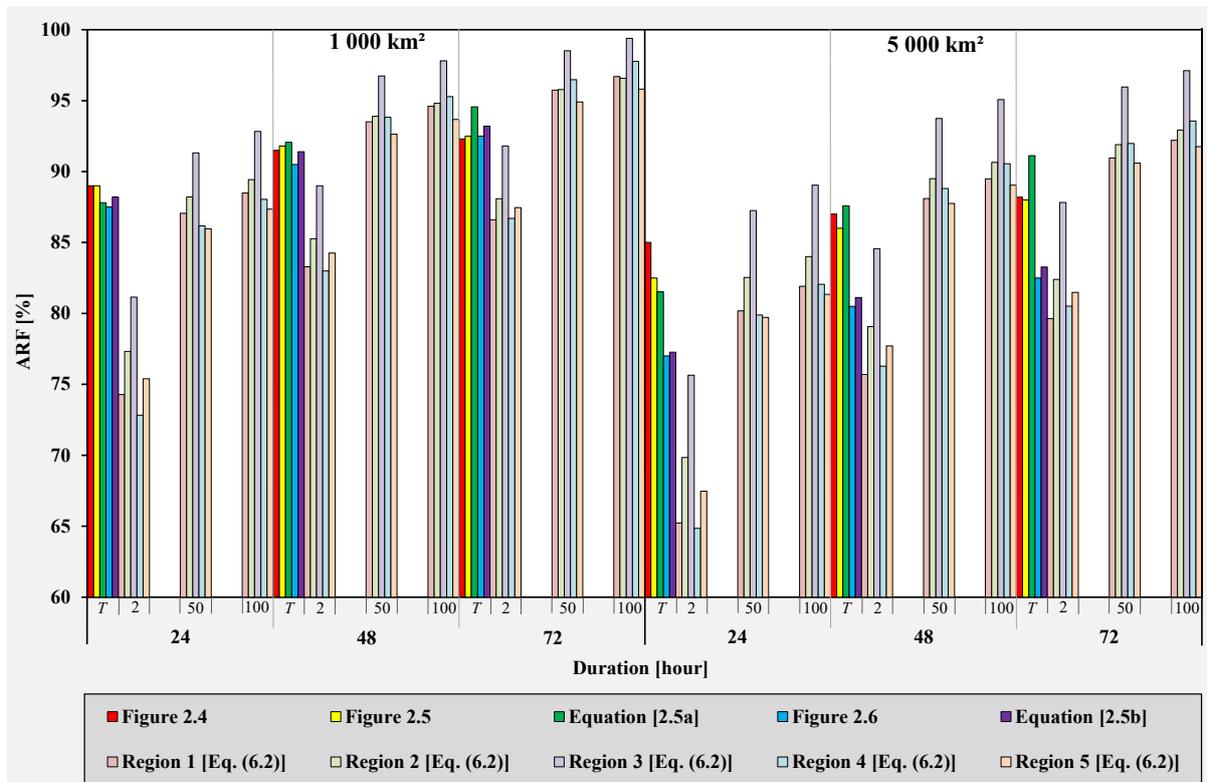
Figure 6.14: Comparison of ARF estimation methods (1 000 km² to 5 000 km²)
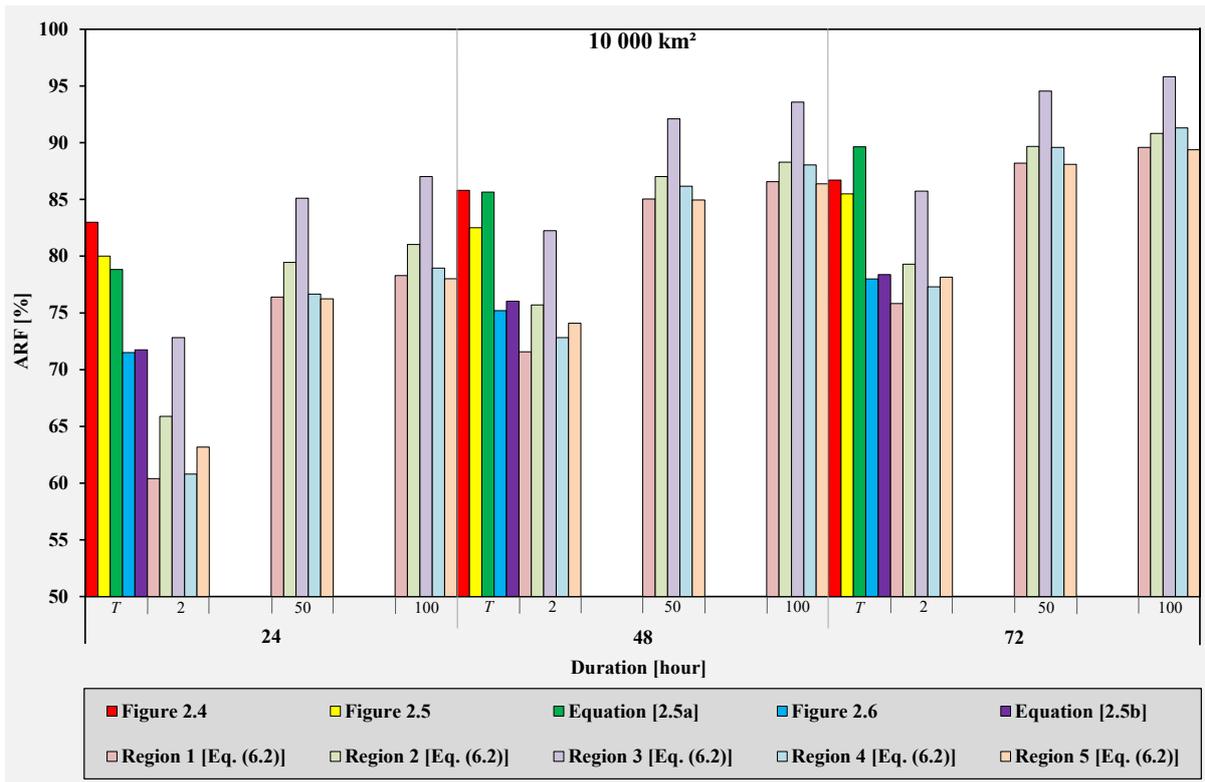
Figure 6.15: Comparison of ARF estimation methods (10 000 km²)
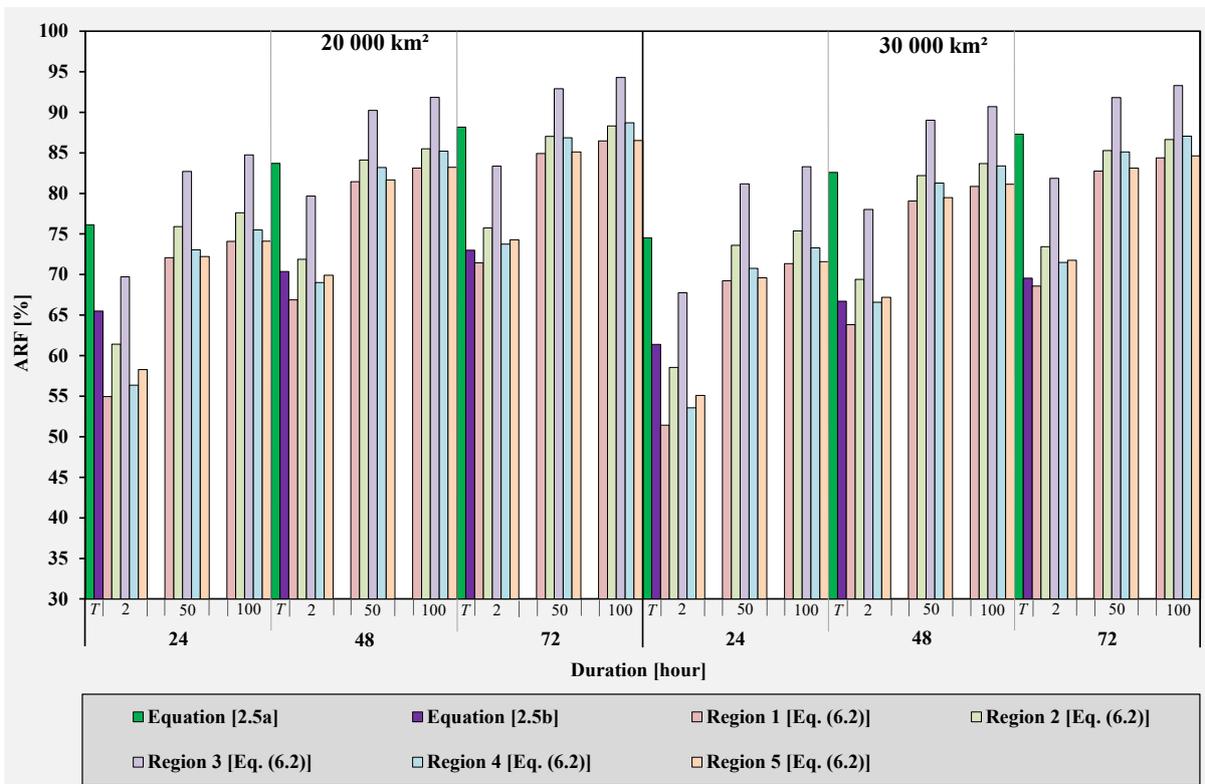
Figure 6.16: Comparison of ARF estimation methods (20 000 km² to 30 000 km²)

6.8   ARF Software Interface

The ARF software interface, as shown in Figure 6.17, 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 basically 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. The software interface neither requires any external database nor contain any cookies.
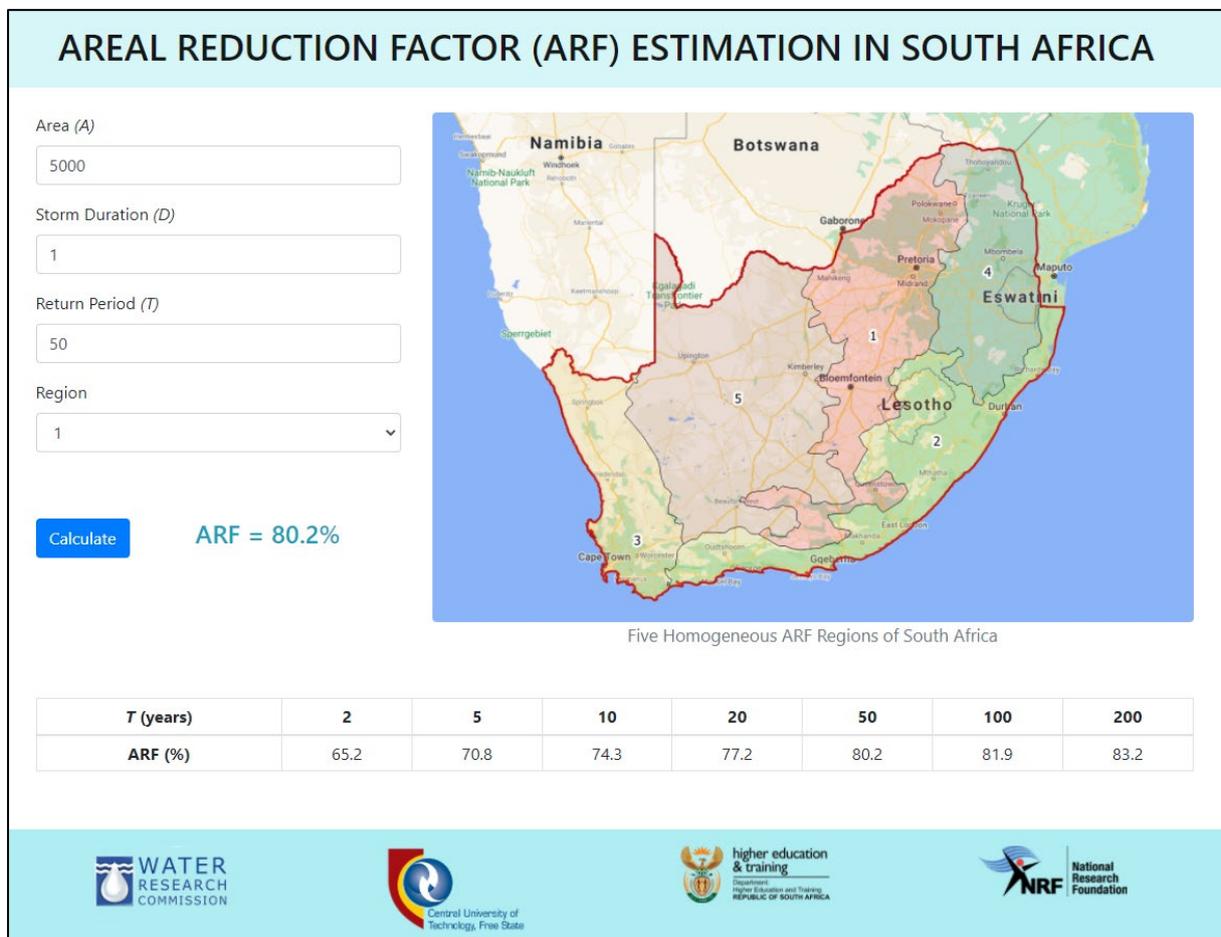
Figure 6.17:   ARF software interface

The results from this chapter also confirmed the study assumptions applicable to ARFs (*cf.* Chapter 1, Section 1.3.1), *viz.*: (i) design point rainfall estimates are only representative for a limited area, which were demonstrated by the differences between areal design rainfall and design point rainfall estimates, (ii) ARFs vary with predominant weather types, storm durations, seasonal factors and return periods, which were evident in the different ARF regions and hence the reason for having the five (5) ARF regions, and (iii) the current South African ARF estimation methods are only applicable to specific temporal and spatial scales, which were demonstrated by the absence of any regionalisation, the ARF values exceeding 100% in 'smaller' catchments, the constant ARF values associated with all return periods, and the limited data used.

The final methodology for estimating catchment response time parameters and the associated results are discussed in the next chapter.

CHAPTER 7:   CATCHMENT RESPONSE TIME – FINAL METHODOLOGY AND RESULTS

This chapter focuses mainly on the final methodological approach followed towards the estimation of catchment response time parameters within Primary Drainage Region X as pilot case study area in South Africa. In addition, all the concerns and possible solutions documented in Chapter 5, as well as the recommendations made by the Reference Group members, were considered and addressed while progressing with and adopting the final methodology.

7.1   Establishment of Catchment Variable Database

A comprehensive catchment variable/parameter database was established for Primary Drainage Region X and typically include the following GIS-based catchment variables/parameters:

(a) **DWS data sources:** Flow-gauging and rainfall station catalogues, drainage regions, discharge rating tables, and annual maximum series (AMS) data sets.

(b) **Catchment geomorphology:** Latitude, longitude, catchment area, catchment perimeter, average catchment slope and centroid distance based on the 30 m corrected SRTM DEM using TauDEM.

(c) **Channel geomorphology:** Length of main watercourse/river, average slope of main watercourse/river based on the 10-85 method, flow accumulations, flow grids, and flow paths based on the 30 m corrected SRTM DEM.

(d) **Climatological variables:** Raster data sets of mean annual precipitation (MAP), and the RLMA&SI-based design rainfall for all durations (5 minutes to 7 days) and return periods (2- to 200-year).

(e) **Catchment variables:** SCS hydrological soil groups, land-use/cover, and Kovács and HRU regionalisation schemes.

The specific catchment variables/parameters required and used as input (independent predictor variables) for the derivation of the regional time parameter equation are discussed in detail in Section 7.3.

7.1.1 Pilot study area characteristics

South Africa, which is located on the most southern tip of Africa, is demarcated into 22 primary drainage regions, i.e. A to X (Midgley et al., 1994), which are further delineated into 148 secondary drainage regions, i.e. A1, A2, to X4.

As shown in Figure 7.1, Primary Drainage Region X covers 31 193 km²; 70% extends across the Mpumalanga Province of South Africa, while the remainder extends into Eswatini (former Swaziland). Primary Drainage Region X is further delineated into four secondary drainage regions, i.e. X1 (11 227 km²), X2 (10 447 km²), X3 (6 322 km²), and X4 (3 197 km²). The 51 gauged catchments to be considered in this study have catchment areas ranging from 6 km² to 21 583 km² (*cf.* Table 7.1). The catchment topography is moderately steep with elevations varying from 112 m to 2 255 m above mean sea level and with average catchment slopes between 3.5% and 36.1% (USGS, 2002). The MAP ranges from 521 mm to 1 325 mm (Lynch, 2004) and the summer rainfall is regarded as highly variable.
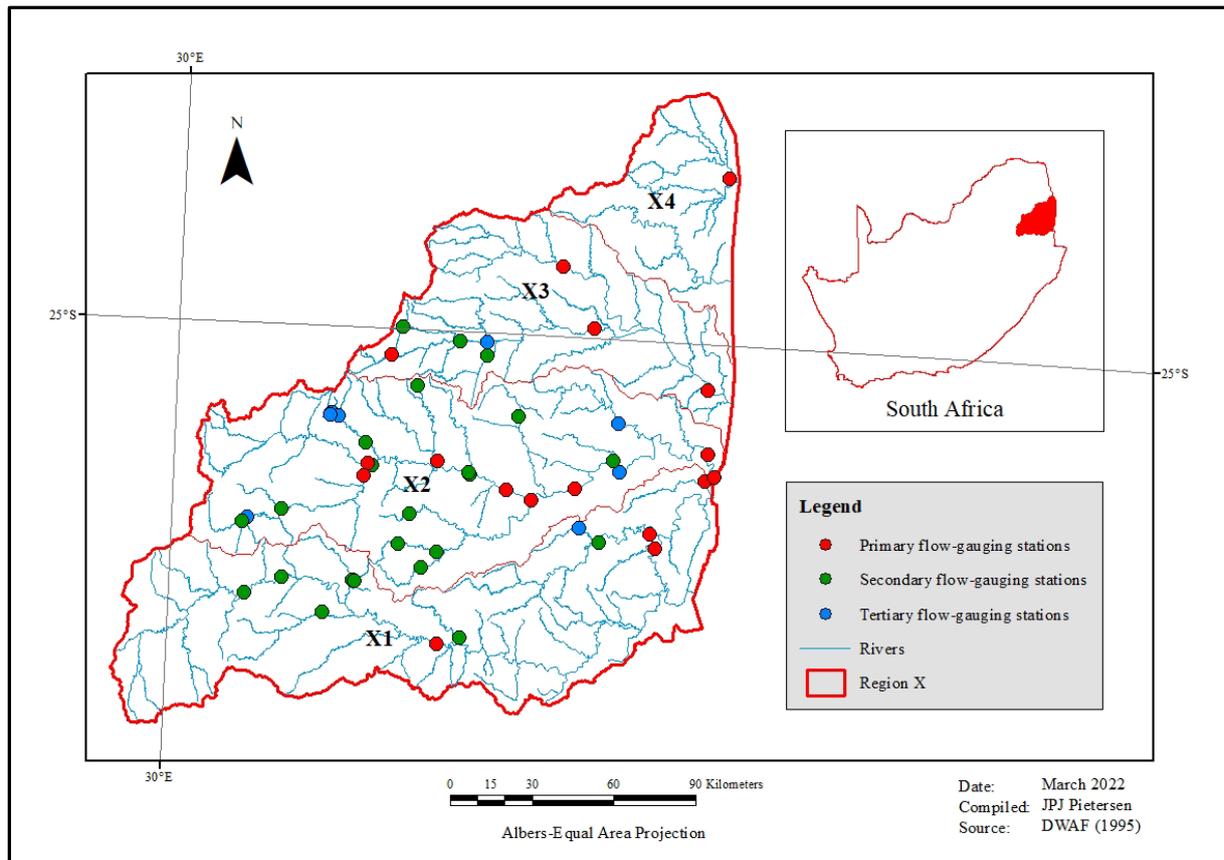
Figure 7.1: Location of the 51 gauged catchments in Primary Drainage Region X

The flow-gauging stations in each catchment were classified by DWS as either primary, secondary or tertiary gauging sites based on the: (i) status (open/closed), (ii) location and importance in the overall monitoring network, (iii) data availability, quality and record length, (iv) type of calibration (standard/extended for above-structural limit conditions), (v) site survey information available (yes/no), and (vi) flood frequency analyses conducted (yes/no).

Table 7.1: The 51 gauged catchments in Primary Drainage Region X

Station/Catchment	A [km²]	Station/Catchment	A [km²]
X1H001	5 504	X2H025	25
X1H003	8 776	X2H026	14
X1H012	118	X2H027	77
X1H014	1 122	X2H028	6
X1H016	585	X2H031	264
X1H017	2 416	X2H032	5 382
X1H018	2 628	X2H035	16
X1H019	186	X2H036	21 583
X1H020	48	X2H046	8 458
X1H021	292	X2H047	111
X1H052	1 457	X2H059	308
X1H053	11 121	X2H072	247
X2H005	640	X2H096	3 089
X2H006	5 090	X2H097	8 164
X2H008	180	X3H001	174
X2H010	127	X3H002	55
X2H011	400	X3H003	48
X2H012	93	X3H004	215
X2H013	1 513	X3H006	771
X2H014	255	X3H008	1 071
X2H015	1 545	X3H011	214
X2H016	10 354	X3H015	5 788
X2H017	8794	X3H021	2 420
X2H018	620	X3H023	679
X2H022	1 642	X4H004	992
X2H024	82		

Primary flow-gauging stations	*Secondary flow-gauging stations*

Tertiary flow-gauging stations

7.2   Extraction and Analysis of Flood Hydrographs to Estimate Time Parameters

The manual/semi-automated procedures to estimate catchment response time parameters within Primary Drainage Region X are detailed in this section. The focus is on the proposed procedure to estimate catchment response time parameters from the observed streamflow data which are continuously monitored at the 51 flow-gauging weirs operated and managed by the DWS. The final methodology, with the aid of functionalities available in the Hydrograph Analysis Tool (HAT), could be summarised as follows:

(a) Evaluation, preparation and extraction of primary streamflow data for the period up to 2020/21 from the DWS streamflow database.

(b) Identification and extraction of the AMS events, i.e. the annual flood peaks at each flow-gauging station. For example, a record length of 50 years would typically contain 50 AMS events. The AMS or yearly peak flow files were also quality controlled and converted into the required *.CSV format for further processing in the HAT.

(c) Verification of the accuracy and relevance of the discharge rating tables (DTs) on the DWS website. In general, all the DTs in Region X were already quality controlled and extended (as required) by the DWS Flood Studies Division. However, in the absence of an extended DT (if required), the AMS data set was extended using a 3rd order polynomial relationship up to 20%. Verification of the extension to +20% was done by considering the hydrograph shape, especially the peakedness as a result of a steep rising limb in relation to the hydrograph base length, and the relationship between individual peak discharge (Q_{Pxi}) and direct runoff volume (Q_{Dxi}) pair values. Typically, in such an event, the additional volume of direct runoff (Q_{DE}) due to the extrapolation were limited to 5%, i.e. $Q_{DE} \leq 0.05\,Q_{Dxi}$. Hence, the error made by using larger direct runoff volumes had little impact on the sample statistics of the total flood volume. In order to address Study Assumption 1 related to the extension of discharge rating curves at a flow-gauging site (*cf.* Section 1.3.2, Chapter 1), Section 7.5 includes detailed information and results pertained to the extension of DTs using indirect estimation methods.

(d) Define and implement user-defined truncation level criteria (Q_{TR}) associated with the record length (N) to extract complete hydrographs. Typically, the following truncation level criteria were implemented: (i) $N \leq 20$ years, use lowest/minimum AMS value, (ii) $20 < N \leq 60$ years, use 25-percentile AMS value, and (iii) $N > 60$ years, use median AMS value. For example, the median AMS value typically has a return period (T) = 2-year or an AEP = 50%. Hence, all complete hydrographs with a peak flow > selected AMS value, i.e. PDS values above a certain discharge threshold, were extracted.

(e) Identification and extraction of complete hydrographs (as shown in Figure 7.2) associated with each AMS event and appropriate truncation level criteria. The extracted hydrographs typically included the following: (i) start/end date/time of flow event, (ii) observed water level [m], (iii) observed discharge [$m^3.s^{-1}$] and total volume of runoff [Q_{Txi}, m^3], (iv) direct runoff discharge [$m^3.s^{-1}$] and total volume of direct runoff [Q_{Dxi}, m^3], (v) baseflow discharge [$m^3.s^{-1}$] and total volume of baseflow [Q_{Bxi}, m^3], and (vi) the cumulative volume of direct runoff under the hydrograph rising limb [Q_{DRi}, m^3].

A summary of the results based on Steps (d) and (e), respectively, is included in Table 7.2.

Table 7.2: Number of extracted and analysed hydrographs based on the truncation level criteria

Catchment	N (years)	Q_{TR} criteria	Implemented Q_{TR} [m³.s⁻¹]	# Total events (extracted/analysed)	# Total events (used)
X1H001	112	Median	183	118	43
X1H003	82	Median	161.1	55	37
X1H012	24	25%	2.5	50	44
X1H014	53	25%	30.6	127	70
X1H016	51	25%	26	93	60
X1H017	50	25%	35.2	124	56
X1H018	48	25%	43.1	110	58
X1H019	47	25%	43.5	80	34
X1H020	47	25%	1	104	93
X1H021	46	25%	10.7	63	20
X1H052	17	Minimum	6.2	93	63
X1H053	14	Minimum	10.1	111	60
X2H005	91	Median	10.2	119	61
X2H006	91	Median	94.4	96	53
X2H008	72	Median	31.4	61	22
X2H010	72	Median	21.4	63	45
X2H011	44	25%	30.2	122	41
X2H012	64	Median	30.3	74	13
X2H013	61	Median	45.3	54	32
X2H014	62	Median	11	32	10
X2H015	61	Median	102.4	67	36
X2H016	60	Median	100	97	69
X2H017	39	25%	100.7	71	43
X2H018	37	25%	11.4	41	27
X2H022	60	Median	25.6	140	91
X2H024	56	25%	5.1	122	50
X2H025	26	25%	2	47	15
X2H026	26	25%	1.1	60	32
X2H027	26	25%	6.1	33	18
X2H028	26	25%	0.2	37	17
X2H031	53	25%	8.1	160	117
X2H032	52	25%	70.8	95	65
X2H035	38	25%	0.5	67	20
X2H036	37	25%	142.7	91	30
X2H046	35	25%	70.9	48	30
X2H047	35	25%	3.2	95	45
X2H059	35	25%	4.1	143	100
X2H072	31	25%	5.1	88	25
X2H096	16	Minimum	18.6	95	37
X2H097	13	Minimum	12	77	55
X3H001	72	Median	10.1	83	43
X3H002	72	Median	3.1	74	17
X3H003	72	Median	5.2	142	27
X3H004	72	Median	20.2	86	15
X3H006	42	25%	25.3	103	50
X3H008	52	25%	25.6	181	148

Catchment	N (years)	Q_{TR} criteria	Implemented Q_{TR} [m^3.s^{-1}]	# Total events (extracted/analysed)	# Total events (used)
X3H011	41	25%	8.1	76	41
X3H015	34	25%	77	66	25
X3H021	30	25%	62.8	78	28
X3H023	18	Minimum	10.1	97	38
X4H004	60	Median	63.3	45	15
Average	49		Total	4 454	2 284

It is evident from Table 7.2 that a total of 4 454 complete hydrographs were extracted and analysed. The record lengths under consideration varied between 13 and 112 years with an overall average record length of 49 years. The Q_{TR} criteria were dominated by the minimum AMS (5 catchments) and 25-percentile AMS (29 catchments) values in 67% of all the catchments under consideration. Therefore, in the latter catchments, at least 75% of all the AMS events were included in the analyses, while it could be argued that 50% or more of the AMS events were discarded in the 17 catchments (33%) remaining where the median AMS criteria were applied. Given that record length is used as the guiding mechanism for the Q_{TR} criteria, the process followed is regarded as consistent, both in terms of the process itself and the results obtained. Subsequently, it is evident that not all the AMS values need to be included in time parameter analyses. In addition, Gericke and Smithers (2017), demonstrated that the smaller T_{Pxi} values [Eqs. (7.2) and/or (7.3)], which occurred more frequently, have a large influence on the average value and consequently result in an underestimated catchment T_{Px} value [Eq. (7.4)]. On the other hand, the longer T_{Px} values have a lower frequency of occurrence, and are reasonable in medium to large catchment scales as the contribution of the whole catchment to peak discharge seldom occurs as a result of the non-uniform spatial and temporal distribution of rainfall in the catchment. In principle, these events should conform to the conceptual definition of T_C ($\approx T_P$), which assumes that T_C is the time required for runoff generated from effective rainfall, with a uniform spatial and temporal distribution over the whole catchment, to contribute to the peak discharge at the catchment outlet.

(f) Separation of complete hydrographs (as shown in Figure 7.2) into direct runoff and baseflow. The recursive digital filtering method [Eq. (7.1) or Eq. (3.15) in Chapter 3] as initially proposed by Lyne and Hollick (1979) and further developed by Nathan and McMahon (1990), was used to separate the direct runoff and baseflow. Since daily/sub-daily time-step data are more appropriate to time parameter estimation and the need for consistency and reproducibility, Eq. (7.1) with a default α-parameter value = 0.995 and a fixed β-parameter value = 0.5 was used in all the catchments under consideration.

$$Q_{Dxi} = \alpha Q_{Dx(i-1)} + \beta(1+\alpha)\left(Q_{Txi} - Q_{Tx(i-1)}\right) \qquad [7.1]$$

where

Q_{Dxi} = filtered direct runoff at time step i, which is subject to $Q_{Dx} \geq 0$ for time i [m^3.s^{-1}],

α, β = filter parameters, and

Q_{Txi} = total streamflow (direct runoff + baseflow) at time i [m^3.s^{-1}].
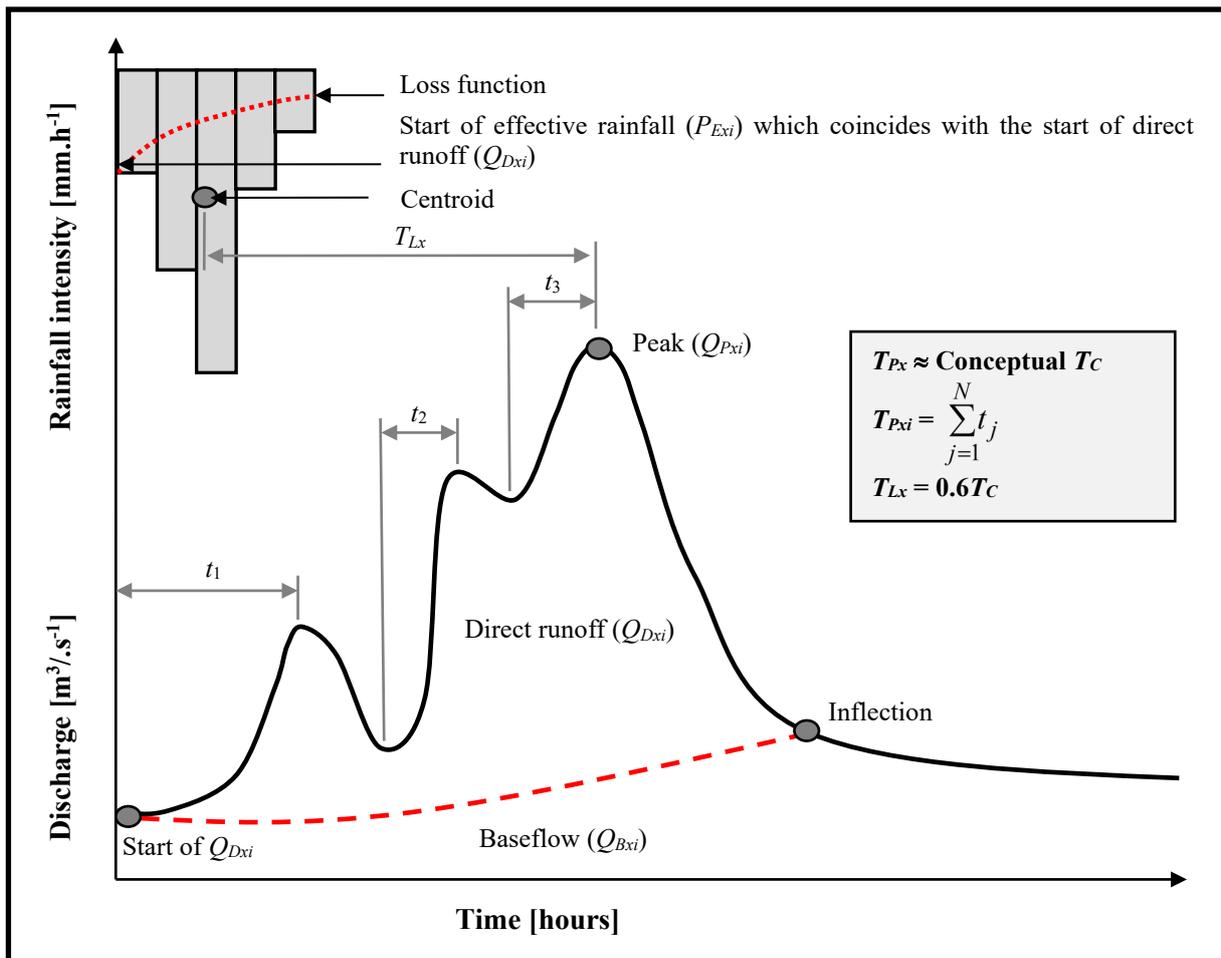
Figure 7.2: Schematic illustrative of the conceptual T_C and T_{Px} relationship for multi-peaked hydrographs (after Gericke and Smithers, 2017; 2018)

(g) Estimation of the time parameter values associated with individual hydrograph/flood events using two different approaches: (i) net rise (duration) of a multi-peaked hydrograph [Eq. (7.2)], and (ii) triangular-shaped direct runoff hydrograph approximation [Eq. (7.3)] and associated variable hydrograph shape parameters [Eqs. (7.3a-c)] as shown in Figure 7.3. This step addresses Study Assumption 2, i.e. T_P equals the total net rise (duration) of a multiple-peaked hydrograph.

$$T_{Pxi} = \sum_{j=1}^{N} t_j \qquad [7.2]$$

$$T_{Pxi} = K \left[\frac{Q_{Dxi}}{3600 x Q_{Pxi}} \right] \qquad [7.3]$$

$$K = 2 \left[\frac{Q_{DRi}}{Q_{Dxi}} \right] \qquad [7.3a]$$

$$T_{Rcxi} = T_{Pxi} \left[\left(\frac{Q_{Dxi}}{Q_{DRi}} \right) - 1 \right] \qquad [7.3b]$$

$$T_{Bxi} = T_{Pxi} + T_{Rcxi} \qquad [7.3c]$$
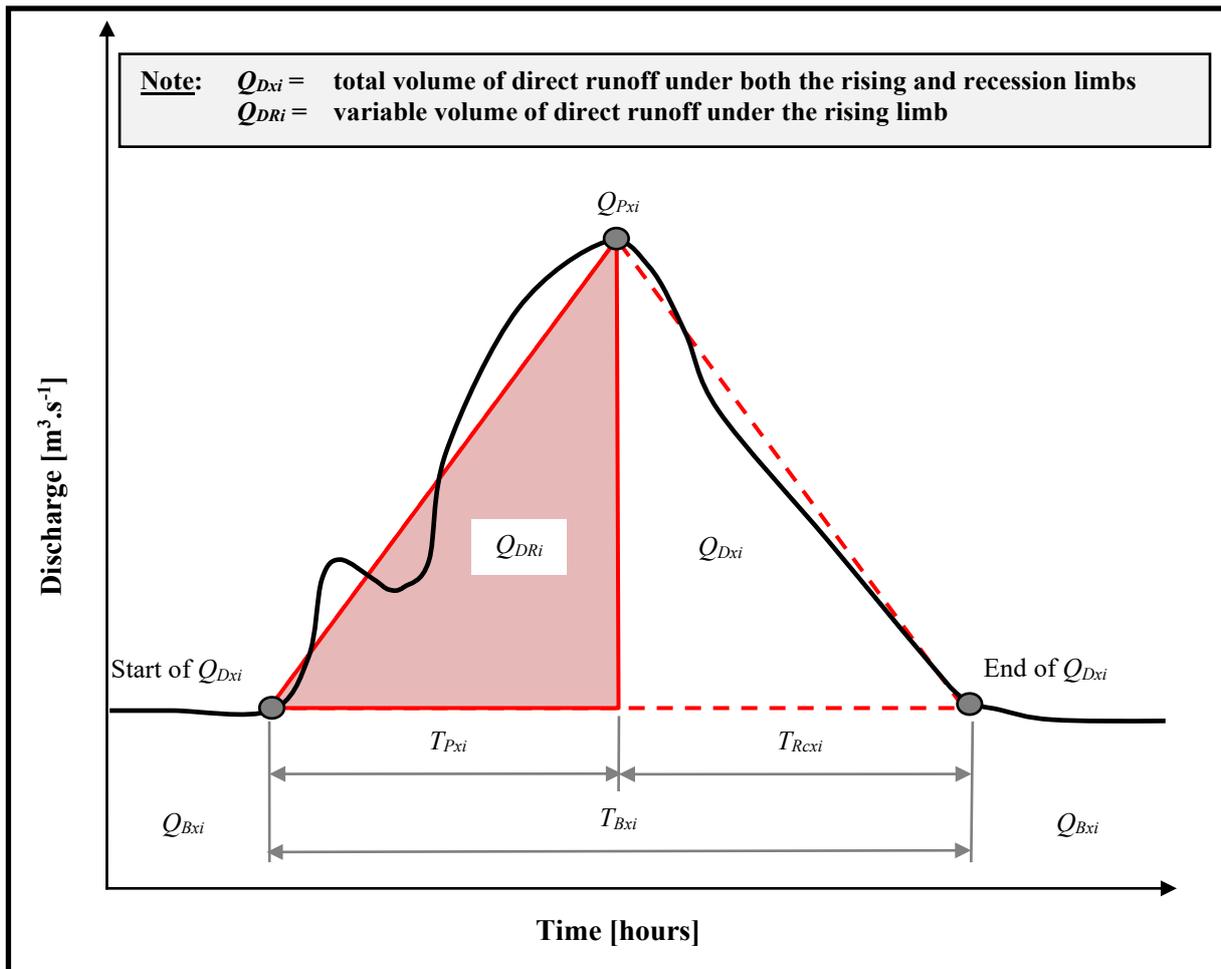
Figure 7.3: Schematic illustrative of the triangular-shaped direct runoff hydrograph approximation [Eq. (7.3)]

A scatter plot of the T_{Pxi} values computed using Eqs. (7.2) and (7.3), respectively for all the catchments under consideration, is shown in Figure 7.4. In comparing Eqs. (7.2) and (7.3) at a catchment level in Region X, the r^2 value of 0.84 (based on the 2 284 flood hydrographs) not only confirmed the relatively high degree of association, but also the usefulness of Eq. (7.3). Taking into consideration the influence catchment area has on response times, the degree of association between these individual T_{Pxi} values could decrease with an increase in catchment area; however, the ultimate goal is to estimate the average catchment T_{Px} by considering the sample-mean of the individual responses based on Eqs. (7.2) and (7.3), respectively.
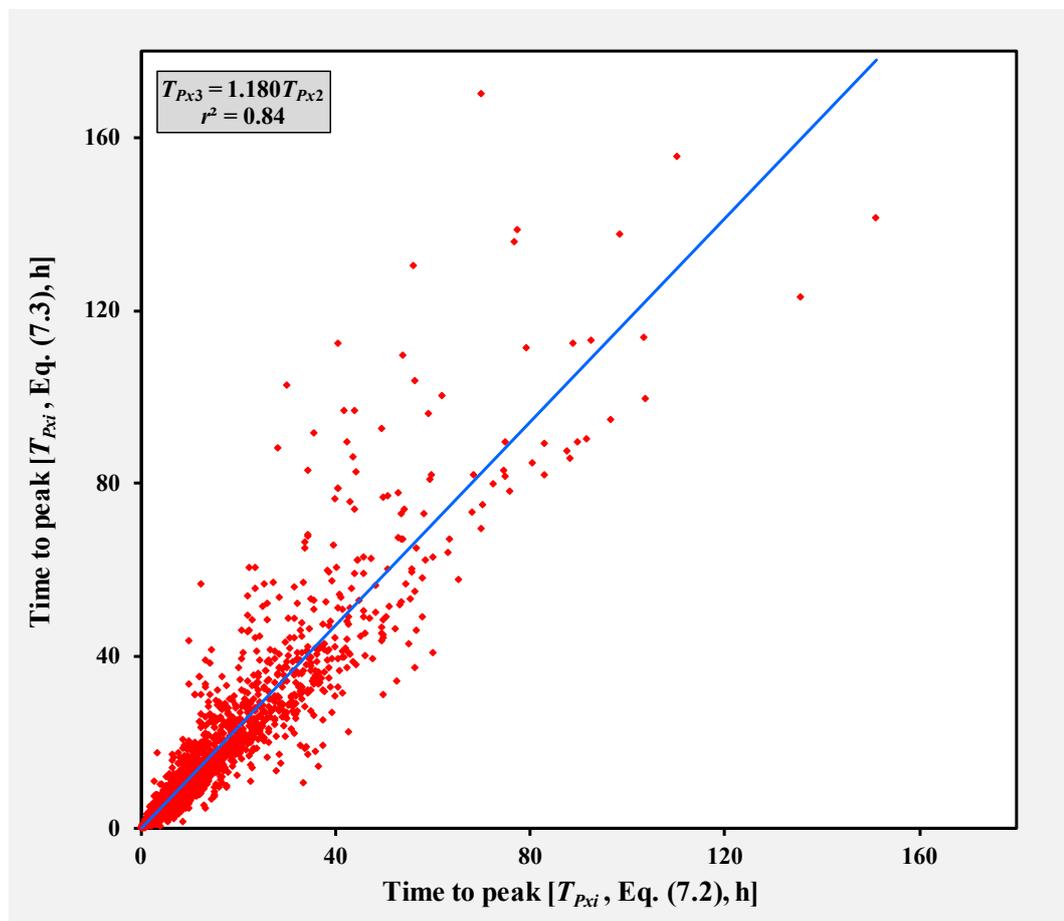
Figure 7.4: Scatter plot of the T_{Pxi} pair values computed using Eqs. (7.2) and (7.3)

In using Eq. (7.3) to estimate the individual T_{Pxi} values by incorporating a triangular approximated hydrograph shape parameter, the variability of Q_{DRi} under the rising limb of individual hydrographs is evident. In Figure 7.5, a frequency distribution histogram of the Q_{DRi} values expressed as a percentage of the total direct runoff volume (Q_{Dxi}) is shown. Taking into consideration that 2 284 (51.3%) of the individual flood

hydrographs extracted were included in the analyses, a few flood events could be characterised by either low (0.4%) or high (92.8%) Q_{DRi} values. However, approximately 35% of the Q_{DRi} values are within the 20 ~ 40 % range. Only 15% of the Q_{DRi} values are within the 30 ~ 40 % range; highlighting some relevance of the conceptual curvilinear unit hydrograph theory (USDA NRCS, 2010) which assigns 37.5% of the direct runoff volume to the hydrograph rising limb.
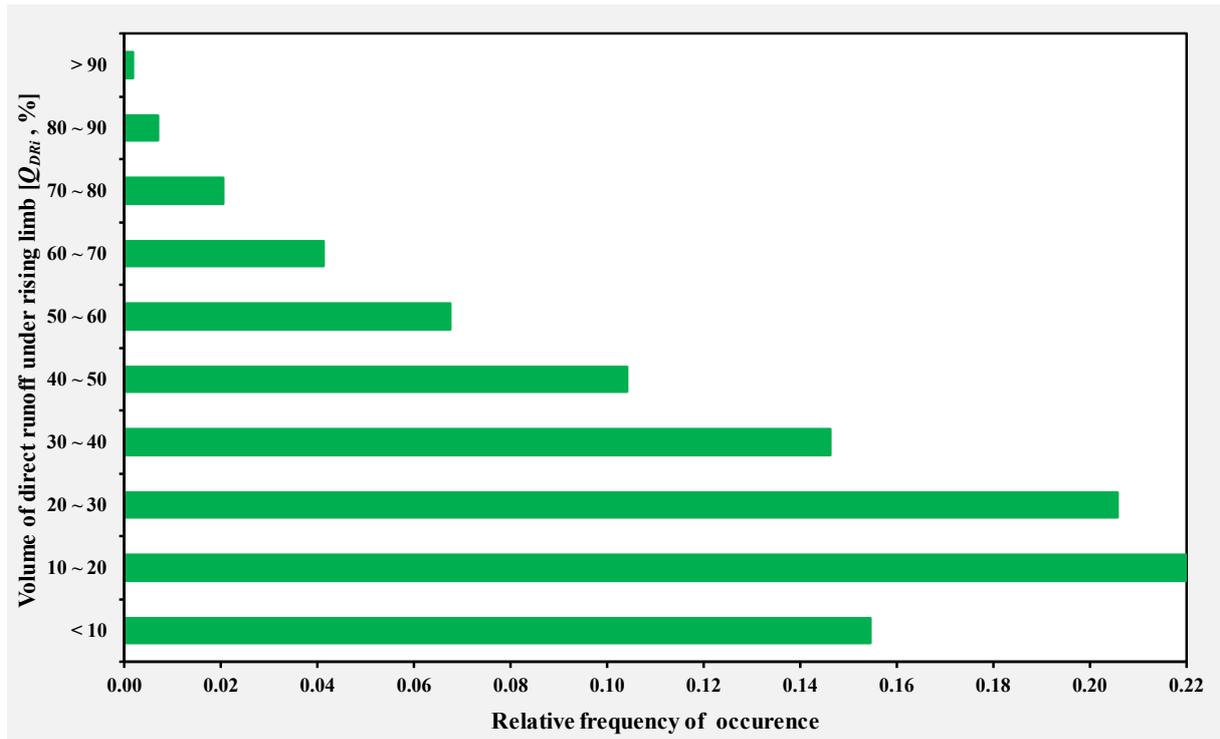
Figure 7.5:   Frequency distribution histogram of the Q_{DRi} values [%] based on the 2 284 analysed flood hydrographs

Thus, by using the above approach as detailed in Step (g), both multi-peaked hydrographs [Eq. (7.2)] and triangular-shaped direct runoff hydrograph approximations [Eq. (7.3)] are included. Ultimately, Eq. (7.3), which reflects the actual percentage of direct runoff under the rising limb of each individual hydrograph, can also be used in future to expand the unit hydrograph theory to larger catchments. In other words, the variable hydrograph shape parameter [Eq. (7.3a)] which reflects the actual percentage of direct runoff under the rising limb of each individual hydrograph can be used instead of the fixed volume of 37.5% normally associated with the conceptual curvilinear unit hydrograph theory.

(h) Estimation of the 'average' catchment response time (T_{Px}) of all the flood events considered in each catchment by using a linear catchment response function [Eq. (7.4)],

i.e. the relationship between individual paired observed peak discharge (Q_{Pxi}) and direct runoff volume (Q_{Dxi}) values.

$$T_{Px} = \frac{1}{3600x}\left[\frac{\sum\limits_{i=1}^{N}\left(Q_{Pxi}-\overline{Q_{Px}}\right)\left(Q_{Dxi}-\overline{Q_{Dx}}\right)}{\sum\limits_{i=1}^{N}\left(Q_{Pxi}-\overline{Q_{Px}}\right)^{2}}\right] \qquad [7.4]$$

where

T_{Bxi} = triangular hydrograph base length for individual hydrograph/flood events [h],

t_j = duration of the total net rise (excluding the in-between recession limbs) of a multiple-peaked hydrograph [h],

T_{Px} = 'average' catchment time to peak based on a linear catchment response function [h],

T_{Pxi} = net rise (duration) or triangular approximated time to peak for individual hydrograph/flood events [h],

T_{Rcxi} = recession time for individual flood events [h],

Q_{Dxi} = volume of direct runoff for individual hydrographs [m^3],

Q_{DRi} = volume of direct runoff under the rising limb for individual hydrographs [m^3],

$\overline{Q_{Dx}}$ = mean of Q_{Dxi} [m^3],

Q_{Pxi} = observed peak discharge for individual hydrographs [m^3/s],

$\overline{Q_{Px}}$ = mean of Q_{Pxi} [m^3.s^{-1}],

K = hydrograph shape parameter,

N = sample size, and

x = a variable proportionality ratio (default $x = 1$), which depends on the catchment response time parameter under consideration.

On average at a catchment level, the averages of Eqs. (7.2) and (7.3) were comparable to those estimates based on Eq. (7.4), with average differences limited to 13.6% and r^2 values ranging from 0.97 to 0.99. Hence, the catchment response times based on an assumed linear catchment response function [Eq. (7.4)] could provide results comparable to the sample-mean of all the individual response times as estimated using Eqs. (7.2) and (7.3), which confirmed Study Assumption 3, i.e. the error bounds between the three different approaches to estimate catchment response time parameters are within acceptable limits ($\leq$ 20%). Therefore, the application of Eq. (7.4) is regarded

to result in a useful 'representative value' to ensure that the average of individual T_{Pxi} values is a good reflection of the catchment conditions and sample-mean. A scatter plot of the average T_{Pxi} values computed using both Eqs. (7.2) and (7.3) in comparison to the catchment T_{Px} values [Eq. (7.4)] for all the catchments under consideration, is shown in Figure 7.6.
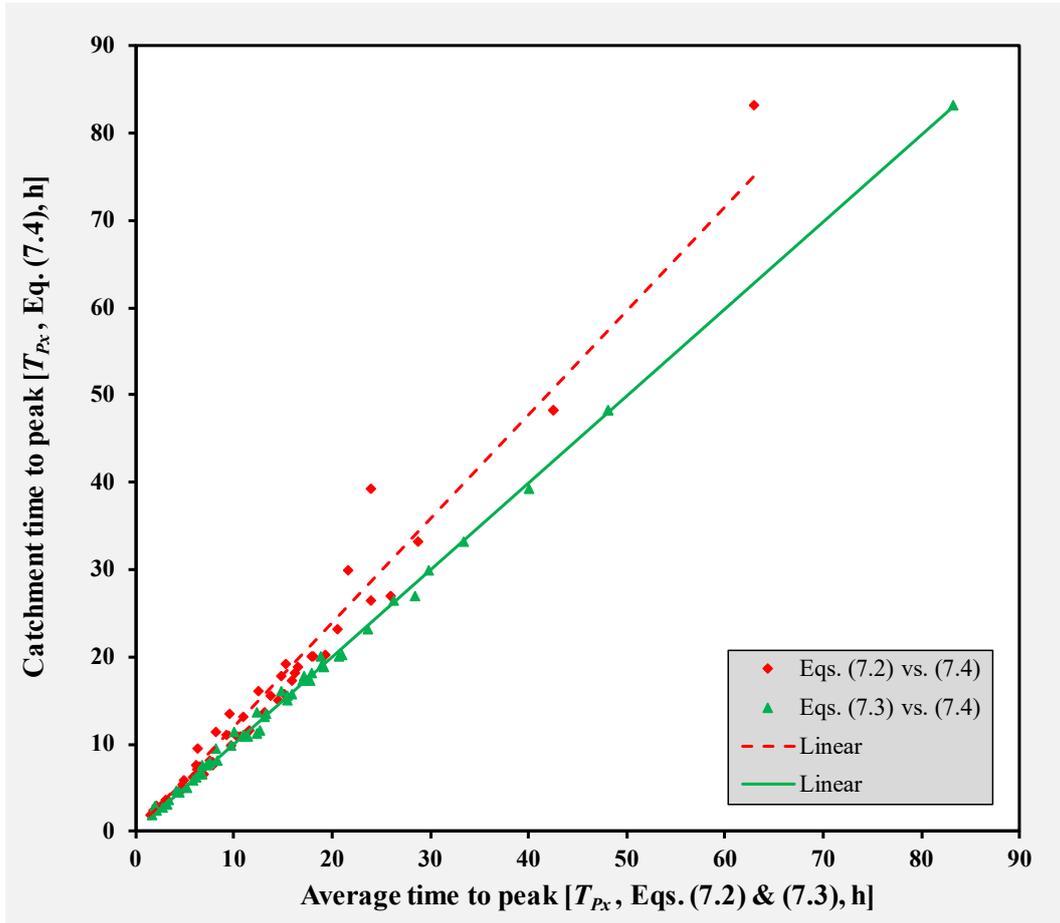
Figure 7.6: Scatter plot of the average T_{Px} values computed using Eqs. (7.2) and (7.3) and the catchment T_{Px} [Eq. (7.4)] values

It is important to note that the variable proportionality ratio (x) is included in Eqs. (7.3) and (7.4) to increase the flexibility and use thereof, i.e. with $x = 1$, either T_{Pxi} or T_{Px} and/or T_{Cxi} or T_{Cx} could be estimated by acknowledging the approximation of $T_C \approx T_P$ (Gericke and Smithers, 2014) and with $x = 1.667$, T_L could be estimated by assuming that $T_L = 0.6T_C$, which is the time from the centroid of effective rainfall to the time of peak discharge (McCuen, 2009). However, to address Study Assumption 4 ($T_C \approx T_P \approx T_L$), Section 7.4 includes detailed information pertained to time parameter proportionality ratios.

Typical output as generated by the HAT and illustrative of the procedures as followed in Steps (a) to (h), is shown in Figures 7.7 and 7.8, respectively.

EXTRACTION OF FLOOD HYDROGRAPHS AND BASEFLOW SEPARATION

GENERAL CATCHMENT INFORMATION	
Catchment	X2H097
Area (km²)	8164
Max structural limit (H, m)	15.14
Max structural limit (Q, m³/s)	6015.5
BF Method: 1- Lyne & Hollick (1978) / 2- Chapman (1999)	1
Alpha-parameter	0.995
Beta-parameter	0.5

Chart legend: Total discharge; Baseflow L&H; Baseflow Chapman. Axes: Discharge (m³/s) vs Number of events.

Flood event number	Start date	Start time	End date	End time	Q_{Pxi} (m³/s)	Q_{Txi} (m³)	Q_{Dxi} (m³)	Q_{DRi} (m³)	Q_{DRi} (%)	Shape factor (K)	1 - T_{Pxi} (hr)	2 - T_{Pxi} (hr)	T_{Rcxi} (hr)	T_{Bxi} (hr)
58	2017/11/15	2:12:00 AM	2017/11/18	12:24:00 AM	42.1	4678091	2385255	1154657	48.4	0.97	15.0	15.2	16.2	31.5

Date	Time	Q_{Txi} (m³/s)	H (m)	$Q_{Dxi\text{-}L\&H}$ (m³/s)	$Q_{Dxi\text{-}Chapman}$ (m³/s)	$Q_{Bxi\text{-}L\&H}$ (m³/s)	$Q_{Bxi\text{-}Chapman}$ (m³/s)	ΔT (sec)	Cum. ΔT (sec)	Q_{Txi} (m³)	$Q_{Dxi\text{-}L\&H}$ (m³)	$Q_{Dxi\text{-}Chapman}$ (m³)	Cum. $Q_{Dx\text{-}L\&H}$ (m³)	Cum. $Q_{Dx\text{-}Chapman}$ (m³)
2017/11/15	02:12:00 AM	1.63	0.693	1.626	1.626	0.004	0.004	0	0	0	0	0	0	0
2017/11/15	02:24:00 AM	1.858	0.697	1.845	1.837	0.013	0.021	720	720	1256	1250	1247	1250	1247
2017/11/15	02:36:00 AM	2.088	0.701	2.065	2.048	0.023	0.040	720	1440	1421	1408	1399	2657	2645
2017/11/15	02:48:00 AM	2.374	0.706	2.340	2.313	0.034	0.061	720	2160	1606	1586	1570	4244	4216
2017/11/15	03:00:00 AM	2.548	0.709	2.502	2.464	0.046	0.084	720	2880	1772	1743	1720	5987	5935
2017/11/15	03:12:00 AM	2.836	0.714	2.777	2.726	0.059	0.110	720	3600	1938	1901	1868	7887	7804

Figure 7.7: Example of an extracted and analysed hydrograph (event #58) in catchment X2H097. Results inclusive of the procedures as followed in Steps (a) to (g).

SUMMARY OF INDIVIDUAL FLOOD HYDROGRAPHS AT FLOW-GAUGING STATION

Catchment	Area (km²)	Max H limit (m)	Max Q limit (m³/s)	BF method	Alpha	Beta	Avg. D_{Ri} (%)	Avg. K-value	Avg. T_{Pxi} (1)	Avg. T_{Pxi} (2)	Catchment T_{Px}	r^2 (Q_{Pxi}/Q_{Dxi})	Comments
X2H097	8164	15.14	6015.5	Lyne & Hollick	0.995	0.5	48.1	0.96	20.6	23.6	23.1	0.94	Completed

Left chart: Peak discharge (Q_{Pxi}, m³/s) vs Volume of direct runoff (Q_{Dxi}, m³); $y = 0.0000x + 4.3143$, $R^2 = 0.9369$.

Right chart: Peak discharge (Q_{Pxi}, m³/s) vs Time to peak (T_{Pxi}, hr); $y = 26.6980x^{0.1902}$, $R^2 = 0.0266$.

Flood event #	Start date	Start time	End date	End time	Q_{Pxi} (m³/s)	Q_{Txi} (m³)	Q_{Dxi} (m³)	Q_{DRi} (m³)	Q_{DRi} (%)	Shape factor (K)	1 - T_{Pxi} (hr)	2 - T_{Pxi} (hr)	T_{Rcxi} (hr)	T_{Bxi} (hr)
4	2006/11/18	01:59:19 AM	2006/11/20	10:47:19 PM	29.6	3608251	1768478	867549	49.1	0.98	16.2	16.3	16.9	33.2
5	2006/11/21	03:35:19 PM	2006/11/26	07:23:19 PM	17.8	5784784	2024630	888915	43.9	0.88	16.8	27.7	35.4	63.2
9	2006/12/28	06:25:45 AM	2007/01/02	08:01:45 AM	122.5	24003958	7166370	2236634	31.2	0.62	11.8	10.1	22.4	32.5
10	2007/01/12	03:25:45 PM	2007/01/17	12:25:45 PM	15.2	3446450	1124948	278943	24.8	0.50	6.4	10.2	30.9	41.1
11	2007/01/19	10:13:45 AM	2007/01/22	10:49:45 PM	12.0	2448132	1076802	338493	31.4	0.63	14.0	15.7	34.2	49.9
12	2007/02/04	11:01:45 PM	2007/02/10	09:49:45 AM	26.9	7253862	2565507	1918887	74.8	1.50	39.2	39.6	13.3	52.9
13	2007/04/13	01:49:45 AM	2007/04/16	05:49:45 AM	21.5	4079575	1924489	1095692	56.9	1.14	17.6	28.3	21.4	49.7
14	2007/04/20	09:50:30 AM	2007/04/23	06:50:30 AM	15.4	1768072	844220	183949	21.8	0.44	4.6	6.7	23.9	30.6
15	2007/11/29	09:04:19 PM	2007/11/29	05:16:19 PM	80.0	14778068	5317109	844236	15.9	0.32	7.0	5.9	31.1	36.9
16	2007/12/02	12:16:19 PM	2007/12/06	11:53:52 AM	12.6	2674523	1256488	639493	50.9	1.02	16.6	28.3	27.3	55.5
18	2007/12/16	01:29:52 PM	2007/12/23	11:17:52 PM	38.5	14409700	4081145	3252733	79.7	1.59	38.6	47.0	12.0	58.9
19	2007/12/26	04:41:52 PM	2008/01/01	08:05:52 AM	29.1	4848565	1538019	456704	29.7	0.59	9.0	8.7	20.7	29.4
20	2008/01/03	10:29:52 PM	2008/01/04	06:05:52 PM	12.0	543831	425650	191872	45.1	0.90	6.0	8.9	10.8	19.7
21	2008/01/10	01:53:52 AM	2008/01/14	04:17:52 AM	79.9	20574472	8798314	3001974	34.1	0.68	25.4	20.9	40.3	61.2
22	2008/02/14	12:27:38 AM	2008/02/19	04:27:38 PM	39.8	8252640	3046992	2224774	73.0	1.46	32.8	31.1	11.5	42.5
23	2008/02/26	12:39:38 AM	2008/03/02	05:27:38 AM	29.2	6750219	2946138	1895848	64.4	1.29	36.4	36.0	20.0	56.0
24	2008/03/16	04:51:38 PM	2008/03/22	07:39:38 AM	41.3	14122080	4556611	2700054	59.3	1.19	37.6	36.3	25.0	61.3
26	2008/11/14	05:39:00 PM	2008/11/17	02:15:00 AM	17.3	2070495	1120834	546202	48.7	0.97	12.0	17.6	18.5	36.1
27	2008/12/05	01:44:00 AM	2008/12/11	05:32:00 PM	38.8	11036910	3270141	2356396	72.1	1.44	24.2	33.7	13.1	46.8
28	2009/01/13	05:20:00 AM	2009/01/19	01:08:00 AM	49.2	12348735	3671059	2460626	67.0	1.34	19.4	27.8	13.7	41.5
29	2009/01/29	11:08:00 PM	2009/02/03	03:20:00 AM	54.8	13482564	5362892	3344344	62.4	1.25	27.4	33.9	20.5	54.4
30	2009/02/27	02:08:00 PM	2009/03/06	05:56:00 AM	83.4	32208091	8772225	5904770	67.3	1.35	34.6	39.3	19.1	58.4
33	2009/11/20	09:14:00 AM	2009/11/26	05:02:00 AM	85.6	23941539	7478128	4467637	59.7	1.19	31.8	29.0	19.5	48.5
35	2009/12/01	02:02:00 AM	2009/12/06	01:50:00 AM	78.3	15644186	4719360	1329083	28.2	0.56	9.2	9.4	24.1	33.5
36	2009/12/17	07:38:00 PM	2009/12/17	11:26:00 PM	69.7	18010554	6397519	4698076	73.4	1.47	20.0	37.5	13.6	51.0
37	2009/12/20	10:50:00 PM	2009/12/25	02:38:00 AM	82.4	13797152	4590573	807230	17.6	0.35	5.6	5.4	25.5	30.9
38	2010/01/03	05:38:00 AM	2010/01/10	05:26:00 PM	37.3	17302637	4239630	3655006	86.2	1.72	40.8	54.4	8.7	63.1
40	2010/03/20	08:00:00 PM	2010/03/23	06:12:00 PM	30.1	4145404	1915887	933134	48.7	0.97	15.6	17.2	18.1	35.3
41	2010/04/02	09:12:00 PM	2010/04/11	12:36:00 AM	60.3	32220522	7294330	6147788	84.3	1.69	54.4	56.7	10.6	67.2
42	2010/04/18	06:48:00 PM	2010/04/25	12:24:00 PM	54.7	20604956	5568375	4042996	72.6	1.45	34.0	41.0	15.5	56.5
43	2012/10/17	12:00:00 AM	2012/10/24	11:36:00 AM	70.9	23026027	7390729	5272746	71.3	1.43	36.6	41.3	16.6	57.9
44	2012/12/21	10:00:00 PM	2012/12/31	04:48:00 PM	121.2	33845265	10474839	6737599	64.3	1.29	27.4	30.9	17.1	48.0
45	2013/01/21	07:36:00 PM	2013/01/24	01:36:00 AM	516.2	149586347	53166820	30490627	57.3	1.15	33.2	32.8	24.4	57.2
46	2013/01/31	01:24:00 AM	2013/02/03	11:00:00 PM	184.7	34294313	13101241	4846035	37.0	0.74	13.2	14.6	24.8	39.4
47	2013/02/08	12:00:00 AM	2013/02/12	05:36:00 AM	128.7	32630873	11614053	3457139	29.8	0.60	9.2	14.9	35.2	50.1
49	2014/03/09	01:48:00 PM	2014/03/16	08:36:00 AM	566.3	151978241	38860379	11269240	29.0	0.58	10.8	11.1	27.1	38.1
50	2014/12/28	07:24:00 PM	2015/01/02	10:48:00 AM	120.1	22332443	8728842	5128007	58.7	1.17	24.4	23.7	16.7	40.4
52	2015/12/15	07:24:00 AM	2015/12/17	07:36:00 AM	16.1	1633488	930870	193595	20.8	0.42	7.0	6.7	25.4	32.1
53	2016/12/15	06:36:00 PM	2016/12/21	06:36:00 PM	34.2	9925555	3341256	2488507	74.5	1.49	28.4	40.4	13.9	54.3
54	2016/12/29	10:36:00 AM	2017/01/04	10:12:00 AM	61.0	15660659	5120869	1429440	27.9	0.56	10.2	13.0	33.6	46.6
55	2017/01/07	06:48:00 PM	2017/01/13	04:48:00 AM	60.9	20309542	6662993	4215984	63.3	1.27	35.0	38.4	22.3	60.8
58	2017/11/15	02:12:00 AM	2017/11/18	12:24:00 AM	42.1	4678091	2385255	1154657	48.4	0.97	15.0	15.2	16.2	31.5
59	2017/12/07	02:00:00 AM	2017/12/15	10:48:00 AM	78.8	28432113	7358840	5870947	79.8	1.60	37.6	41.4	10.5	51.9
61	2018/02/20	11:36:00 AM	2018/02/25	01:12:00 AM	46.7	11600615	4754998	2619035	55.1	1.10	31.6	31.1	25.4	56.5
62	2018/03/21	10:12:00 PM	2018/03/30	02:12:00 PM	29.5	13001980	3770410	3003949	79.7	1.59	48.2	56.5	14.4	70.9
63	2018/04/10	11:24:00 AM	2018/04/17	07:00:00 AM	34.4	12656586	3844056	1738029	45.2	0.90	23.6	28.1	34.0	62.1
64	2018/05/04	12:12:00 AM	2018/05/07	06:24:00 PM	13.3	2073400	858199	266932	31.1	0.62	9.8	11.2	24.7	35.9
65	2018/05/26	05:00:00 PM	2018/05/30	03:12:00 PM	14.9	1919839	972234	413297	42.5	0.85	14.6	15.4	20.9	36.3
68	2018/12/20	10:36:00 PM	2018/12/22	12:24:00 PM	13.9	780702	497009	206246	41.5	0.83	7.6	8.3	11.6	19.9
70	2019/01/03	07:36:00 PM	2019/01/06	10:12:00 PM	51.4	9074763	4095370	878444	21.4	0.43	10.4	9.5	34.8	44.3
71	2019/01/28	05:00:00 AM	2019/02/01	10:36:00 AM	64.6	9863186	3577673	431222	12.1	0.24	3.0	3.7	27.1	30.8
73	2019/02/14	04:24:00 AM	2019/02/17	08:12:00 PM	66.5	12939413	5609219	1261751	22.5	0.45	11.0	10.5	36.3	46.9
74	2019/02/22	11:24:00 PM	2019/02/22	04:48:00 AM	96.2	13427888	4763837	825896	17.3	0.35	3.6	4.8	22.7	27.5
75	2019/03/02	07:12:00 PM	2019/03/05	01:48:00 PM	24.6	3081994	1206090	224281	18.6	0.37	4.2	5.1	22.2	27.3
76	2019/04/05	04:24:00 PM	2019/04/10	08:36:00 AM	36.8	8925311	3285872	706264	21.5	0.43	12.6	10.7	39.0	49.6

r^2: T_{P1} & T_{P2} = 0.91

Figure 7.8: Example of a summary of the 55 flood events used to estimate the catchment T_{Px} in catchment X2H097. The averages of Eqs. (7.2) and (7.3) equal 20.6 h and 23.6 h, respectively. The catchment T_{Px} [Eq. (7.4)] = 23.1 h.

7.3   Derivation and Verification of Regional Empirical Time Parameter Equations

Stepwise multiple regression analyses at a 95% confidence level were performed on the catchment response time parameters and geomorphological catchment characteristics to establish calibrated relationships in the 41 calibration catchments to estimate the catchment T_{Px}. The final calibrated empirical equation was also independently assessed or verified in the 10 verification catchments as listed in Table 7.3. In other words, the T_{Px} values based on Eq. (7.4) were used as dependent criterion variables, while the following independent predictor variables, as listed in Table 7.3, were considered for inclusion: (i) area [A, km²], (ii) perimeter [P, km], (iii) centroid distance [L_C, km], (iv) hydraulic length [L_H, km], (v) average catchment slope [S, %], (vii) average main watercourse slope [S_{CH}, %], (viii) drainage density [D_D, km.km^{-2}], and (ix) MAP [mm].

Table 7.3:   Catchment characteristics considered as potential predictor variables to estimate the catchment T_{Px}

Catchment	Criterion	Standard predictor variables							
Station	T_{Px} [h]	A [km²]	P [km]	L_C [km]	L_H [km]	S [%]	S_{CH} [%]	D_D [km.km^{-2}]	MAP [mm]
X1H001	**20.0**	5 504	608.2	137.7	251.7	12.730	0.416	1.5	790
X1H003	**33.2**	8 776	1040.3	229.0	435.0	14.420	0.391	0.8	809
X1H014	**13.6**	1 122	281.3	49.9	104.7	24.310	0.960	1.0	969
X1H016	**7.5**	585	193.4	30.4	59.1	15.300	1.191	1.9	813
X1H017	**11.1**	2 416	355.3	48.3	112.1	6.040	0.224	2.2	703
X1H018	**16.0**	2 628	400.9	67.6	135.8	7.070	0.306	2.1	709
X1H019	**3.6**	186	112.1	15.9	37.3	22.780	1.905	1.8	895
X1H020	**3.1**	48	46.0	7.0	12.9	19.280	3.461	1.8	923
X1H021	**11.4**	292	115.7	22.0	51.9	19.450	1.419	1.4	892
X1H052	**15.0**	1 457	356.2	81.0	154.8	11.030	0.575	0.8	929
X1H053	**39.2**	11 121	1026.7	253.4	495.3	7.950	0.361	0.7	809
X2H005	**17.8**	640	178.4	31.9	64.0	4.340	1.107	1.6	998
X2H006	**26.4**	5 090	585.2	57.5	122.8	10.290	0.391	1.5	834
X2H008	**4.7**	180	93.7	19.2	33.4	22.680	2.290	1.6	947
X2H010	**9.5**	127	70.4	11.2	24.0	18.630	2.661	1.7	1 017
X2H012	**2.7**	93	68.0	7.4	17.2	4.390	1.132	2.3	782
X2H013	**11.5**	1 513	329.9	61.9	122.4	18.680	1.132	1.9	747
X2H014	**20.2**	255	116.7	20.6	40.0	27.870	1.707	1.9	944
X2H015	**13.1**	1 545	335.6	46.6	110.4	10.990	0.941	1.9	827
X2H016	**48.2**	10 354	927.6	164.7	320.3	16.490	0.495	0.8	762
X2H022	**10.9**	1 642	319.6	61.4	112.3	22.120	0.996	1.3	816
X2H024	**6.2**	82	64.1	10.1	22.7	21.650	3.543	1.7	1 021
X2H031	**8.1**	264	121.4	17.8	42.8	14.600	1.979	1.5	922
X2H032	**29.9**	5 382	635.7	84.4	195.2	19.000	0.698	1.4	828
X2H035	**9.9**	16	23.1	1.9	5.7	23.170	4.442	1.8	1 225
X2H036	**83.3**	21 583	1124.5	265.7	567.8	8.800	0.352	0.7	789
X2H046	**27.0**	8 458	797.1	34.3	73.3	10.550	0.395	1.1	788
X2H047	**7.7**	111	82.7	10.8	28.3	14.110	2.508	2.1	747
X2H059	**11.3**	308	166.7	31.3	54.9	7.200	0.815	1.5	876
X2H072	**6.6**	247	122.8	12.8	35.6	8.630	0.679	1.4	723
X2H096	**15.8**	3 089	444.9	52.1	118.5	10.660	0.907	1.8	788

Catchment	Criterion	Standard predictor variables							
Station	T_{Px} [h]	A [km²]	P [km]	L_C [km]	L_H [km]	S [%]	S_{CH} [%]	D_D [km.km^{-2}]	MAP [mm]
X2H097	23.1	8 164	750.4	51.0	109.0	10.170	0.504	1.1	802
X3H001	20.0	174	82.8	10.3	22.2	30.650	2.944	1.9	1 232
X3H002	6.5	55	43.9	5.0	14.1	20.180	2.870	1.9	1 221
X3H003	13.4	48	49.8	6.9	17.2	8.640	1.587	1.9	1 325
X3H004	7.6	215	135.2	15.6	41.2	9.490	2.164	1.6	963
X3H008	7.9	1 071	225.1	33.1	77.4	5.560	0.550	1.6	887
X3H015	17.2	5 788	639.8	56.5	120.7	5.590	0.203	1.0	800
X3H021	18.1	2 420	396.8	62.4	125.6	8.470	0.753	1.4	980
X3H023	18.8	679	199.1	26.7	58.3	13.410	1.296	1.7	1 245
X4H004	5.0	992	222.5	31.0	52.8	3.480	0.298	1.4	521
X1H012	10.9	118	87.4	14.5	31.1	30.820	1.929	1.1	1 028
X2H011	2.9	400	141.3	13.4	37.7	8.530	1.044	2.2	778
X2H017	48.2	8 794	794.0	118.1	257.0	18.470	0.623	1.0	794
X2H018	5.9	620	174.3	34.6	64.2	5.500	0.388	1.1	597
X2H025	4.4	25	35.4	4.0	11.0	35.640	10.423	2.0	927
X2H026	1.8	14	22.2	3.6	7.9	32.400	10.332	2.0	987
X2H027	19.1	77	58.6	10.3	20.5	30.910	3.472	2.0	1 025
X2H028	2.3	6	14.3	2.1	4.6	36.130	12.984	2.0	868
X3H006	15.6	771	225.1	35.4	70.1	22.810	1.153	1.7	1 218
X3H011	17.2	214	109.7	14.7	33.5	20.340	1.221	1.8	1 239

Calibration catchments   *Verification catchments*

Both normal and log-transformed data sets applicable to the above predictor variables were considered, while a combination of these variables to represent recognised catchment parameters, e.g. shape, circularity, and elongation, was also considered. In addition, power-law transformations ($y = ax^b$), e.g. $a(L_H^2/S_{CH})$ ^{b} and $a(L_C L_H/S_{CH}^{0.5})$ ^{b}, were also considered in cases where these combinations resulted in the highest degree of association when individually plotted against the T_{Px} values. In many catchments, the transformed independent variables performed less satisfactorily when included as part of the multiple regression analyses, while the log-transformations resulted in negative response time values in some of the catchments.

Overall, backward stepwise multiple linear regression analyses using normal data performed the best. Hypothesis testing was performed at each step to ensure that only statistically significant independent predictor variables are retained in the model, while insignificant variables are removed. Partial t-tests were used to test the significance of individual independent predictor variables, while total F-tests were used to determine whether the dependent criterion variables are significantly correlated to the independent predictor variables included in each model. The Goodness-of-Fit (GOF) statistics were assessed by using the coefficient of multiple-correlation [Eq. (7.5)] and the standard error of estimate [Eq. (7.6)]. Equations (7.7) and (7.8) were used as regression diagnostics to identify possible outliers and to estimate standardised residuals.

$$R_i{}^2 \quad = \quad \frac{\sum\limits_{i=1}^{N}\left(y_i - \bar{x}\right)^2}{\sum\limits_{i=1}^{N}\left(x_i - \bar{x}\right)^2} \qquad\qquad\qquad [7.5]$$

$$S_{Ey} \quad = \quad \left[\frac{1}{v}\sum\limits_{i=1}^{N}(y_i - x_i)^2\right]^{0.5} \qquad\qquad [7.6]$$

$$h_{ii} \quad = \quad \frac{1}{N} + \frac{\left(x_i - \bar{x}\right)^2}{\sum\limits_{i=1}^{N}\left(x_i - \bar{x}\right)^2} \qquad\qquad\qquad [7.7]$$

$$e_i \quad = \quad \frac{(y_i - x_i)}{S_{Ey}\sqrt{1 - h_{ii}}} \qquad\qquad\qquad [7.8]$$

where

R_i	= multiple-correlation coefficient for i independent variables,
S_{Ey}	= standard error of estimate [h],
h_{ii}	= i^{th} leverage value,
e_i	= standardised residual,
x_i	= observed value (dependent variable),
$\bar{x}$	= mean of observed values (dependent variables),
y_i	= estimated value of dependent variable (x_i),
N	= number of observations (sample size), and
v	= degrees of freedom (N- i; with y-intercept = 0).

In addition to the steps above, the empirical T_{Py} equation [Eq. (7.9)] as originally developed by Gericke and Smithers (2016b) was also tested in the 51 catchments by incorporating newly derived calibration coefficients. This was done to highlight the possible relevance thereof and/or to highlight the need for an alternative empirical time parameter equation.

$$T_{Py} \quad = \quad x_1{}^{MAP}x_2{}^{A}x_3{}^{L_C}x_4{}^{L_H}x_5{}^{S} \qquad\qquad [7.9]$$

where

T_{Py}	= estimated time to peak [h],
A	= catchment area [km²],
L_C	= centroid distance [km],
L_H	= hydraulic length [km],
MAP	= mean annual precipitation [mm],
S	= average catchment slope (%), and
x_1 to x_5	= calibration coefficients (*cf.* Chapter 3, Table 3.4).

A scatter plot of the T_{Py} [Eq. (7.9)] and catchment T_{Px} [Eq. (7.4)] values for both the calibration and verification catchments are shown in Figure 7.9.

$$T_{Py} = 1.399\,T_{Px}$$
$$r^2 = 0.68$$
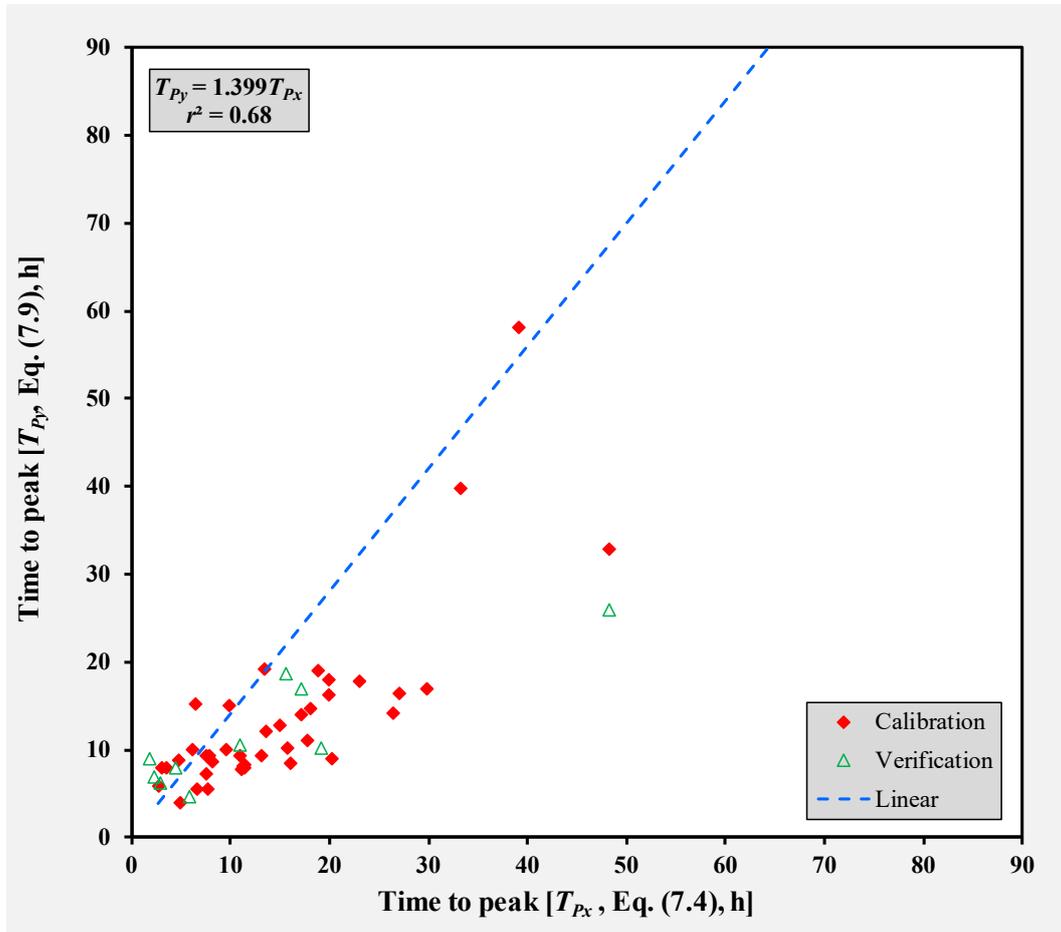
Figure 7.9: Scatter plot of the estimated T_{Py} [Eq. (7.9)] and the catchment T_{Px} [Eq. (7.4)] values

The low to moderate degree of association ($r^2 \leq 0.68$) as depicted in Figure 7.9, highlighted that Eq. (7.9) in its current format would not be useful to estimate the catchment response time in most of the catchments under consideration in Primary Drainage Region X. In addition, many of the standardised residuals computed using Eq. (7.8), exceeded the benchmark standardised residual value of $\pm\,2$.

Subsequently, the backward stepwise multiple linear regression analyses using normal data followed to result in the derivation of the T_{Py} regression in Eq. (7.10). The following statistically significant independent predictor variables were retained and included in the calibrated equation: (i) A, (ii) L_C, (iii) L_H, and (iv) S. At a confidence level of 95%, the above independent variables contributed significantly towards the prediction accuracy. Equation (7.10) was also independently assessed at catchments not used during the calibration

process, i.e. the observed T_{Px} values were compared to the T_{Py} values estimated using the calibrated empirical equation.

$$T_{Py} \qquad = 0.002397A - 0.3585L_C + 0.2122L_H + 0.3882S \qquad\qquad [7.10]$$

where

$\qquad$ T_{Py} $\qquad$ = estimated time to peak [h],

$\qquad$ A $\qquad$ = catchment area [km²],

$\qquad$ L_C $\qquad$ = centroid distance [km],

$\qquad$ L_H $\qquad$ = hydraulic length [km], and

$\qquad$ S $\qquad$ = average catchment slope (%).

A scatter plot of the T_{Py} [Eq. (7.10)] and catchment T_{Px} [Eq. (7.4)] values for both the calibration and verification catchments are shown in Figure 7.10 to highlight any differences.
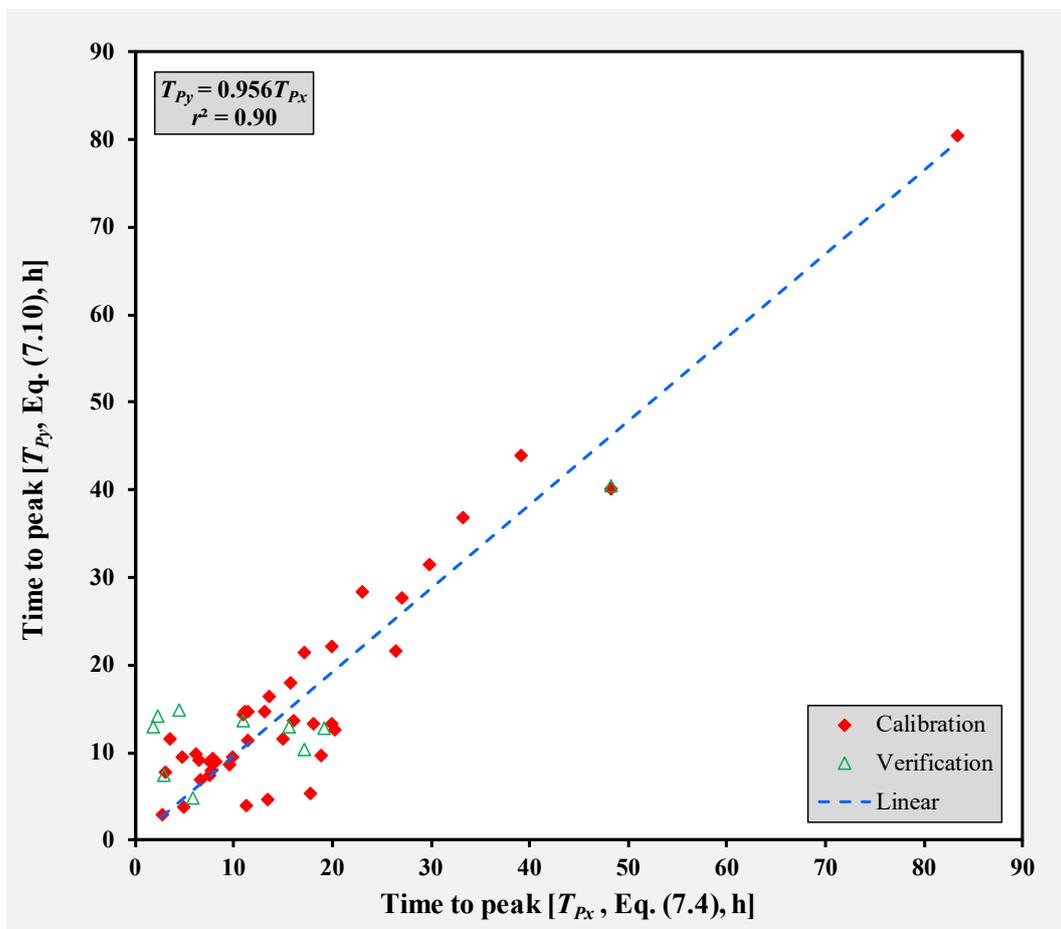
Figure 7.10: Scatter plot of the estimated T_{Py} [Eq. (7.10)] and the catchment T_{Px} [Eq. (7.4)] values

The high degree of association as depicted in Figure 7.10, not only confirmed the good correlation between T_{Px} and T_{Py}, but also the usefulness of Eq. (7.10) to estimate the catchment response time in both the calibration and verification catchments. The overall r^2 value equals 0.90.

In considering the standardised residuals computed using Eq. (7.8) in both the calibration and verification catchments, it was evident that ±94% of the total sample have standardised residuals less than ± 2 (ranging between -1.55 and 1.90), except in the case of calibration catchment X2H005 (-2.27) and the verification catchments X2H026 (2.06) and X2H028 (2.20). However, the latter two catchments have areas ranging between 6 and 14 km²; hence, these catchments are regarded as 'small catchments' and not necessarily 'medium to large catchments', which this study focusses on. According to Chatterjee and Simonoff (2013), it is expected of a reliable regression model to have approximately 95% of the standardised residuals between -2 and +2, while standardised residuals ≥ ± 2 should be investigated as potential outliers. The latter standardised residuals ≥ ± 2 in the four identified catchments are regarded as 'acceptable', given that T_{Py} is consistent with the regression relationship implied by the other T_{Px} values.

A summary of the GOF statistics and hypothesis testing results of the 41 calibration catchments are listed in Table 7.4.

Table 7.4: Summary of GOF statistics and hypothesis testing results applicable to the 41 calibration catchments

Criterion	Value	Criterion/Variable	A	L_C	L_H	S
Multiple-correlation R_I^2 [Eq. (7.5)]	0.95	Coefficients	0.002397	-0.3585	0.2122	0.3882
Coefficient of determination (r^2)	0.96	Coefficient S_E	0.000437	0.1548	0.08397	0.06129
S_{Ey} [Eq. (7.6), h]	4.88	T-statistic	5.49	-2.32	2.53	6.33
Probability of F-statistic	1.80E-24	P-value	3.1E-06	0.026	0.016	2.23E-07
F-Observed value (F-statistic)	198.87	Lower 95%	0.001512	-0.6721	0.04209	0.2640
Critical F-statistic (F_α)	2.63	Upper 95%	0.003281	-0.04497	0.3824	0.5125

In Table 7.4, the standard error results (≈ 4.9 hours) must be clearly understood in the context of the actual travel time associated with the catchment sizes in Primary Drainage Region X, as the impact of such error in the T_{Py} estimates might be critical in a small catchment, while being less significant in a larger catchment. The average regional T_{Px} value ($\overline{T_{Px}}$) equals 16.6 hours and it could be used to justify the latter standard error. In other words, by considering the ratio of S_{Ey}: $\overline{T_{Px}}$ which equals 0.29, the standard error is understood in the correct context.

Furthermore, the rejection of the null hypothesis ($F > F_\alpha$) also confirmed the significant relationship between T_{Px} and the independent predictor variables as included in Eq. (7.10). Given the slightly poorer performance of Eq. (7.10) in the verification catchments, the overall SE_y value increased from 4.9 h to 5.5 h, while the overall r^2 value reduced from 0.90 to 0.86.

The high variability of individual-event observed T_{Pxi} [Eqs. (7.2) & (7.3)] and estimated T_{Py} [Eq. (7.10)] values relative to the catchment T_{Px} values [Eq. (7.4)] in each catchment was estimated using Equation (7.11).

$$\varDelta T_P \quad = \left(\frac{T_{Pxi}, T_{Py}}{T_{Px}} \right) - 1 \qquad\qquad [7.11]$$

where

        $\varDelta T_P$ = catchment response time variability [over/underestimation ($\pm$)],

        T_{Px} = observed catchment response time [Eq. (7.4), h],

        T_{Pxi} = maximum individual-event catchment response time [Eq. (7.2)

                       and/or Eq. (7.3), h], and

        T_{Py} = estimated catchment response time [Eq. (7.10), h].

The latter catchment response time variability at a catchment level in Primary Drainage Region X is shown in Figure 7.11.

The high T_{Pxi} variability as depicted in Figure 7.11 and expressed using Eq. (7.11), highlight that the variability in observed catchment response times is not solely related to catchment area, but the increase in variability is most likely associated with an increase of the spatial and temporal distribution and heterogeneity of other geomorphological catchment characteristics and rainfall as the catchment scale increases. Typically, at these catchment scales, the largest Q_{Pxi} and T_{Pxi} values would be associated with the likelihood of the entire catchment receiving rainfall for the critical storm duration. Smaller T_{Pxi} values could be expected when effective rainfall of high average intensity does not cover the entire catchment, especially when a rainfall event is centered near the catchment outlet. However, these lower T_{Pxi} values are likely to occur more frequently; hence, having a larger influence on the average value and consequently might result in an underestimated representative catchment T_{Px} value. On the other hand, the longer T_{Pxi} values have a lower frequency of occurrence, and are reasonable at medium to large catchment scales, as the contribution of the whole catchment to peak discharge, seldom occurs as a result of the non-uniform spatial and temporal distribution of rainfall. Ultimately, it can be

concluded that catchment response time variability increases as the magnitude (e.g. return period) and spatial distribution of rainfall events decrease.
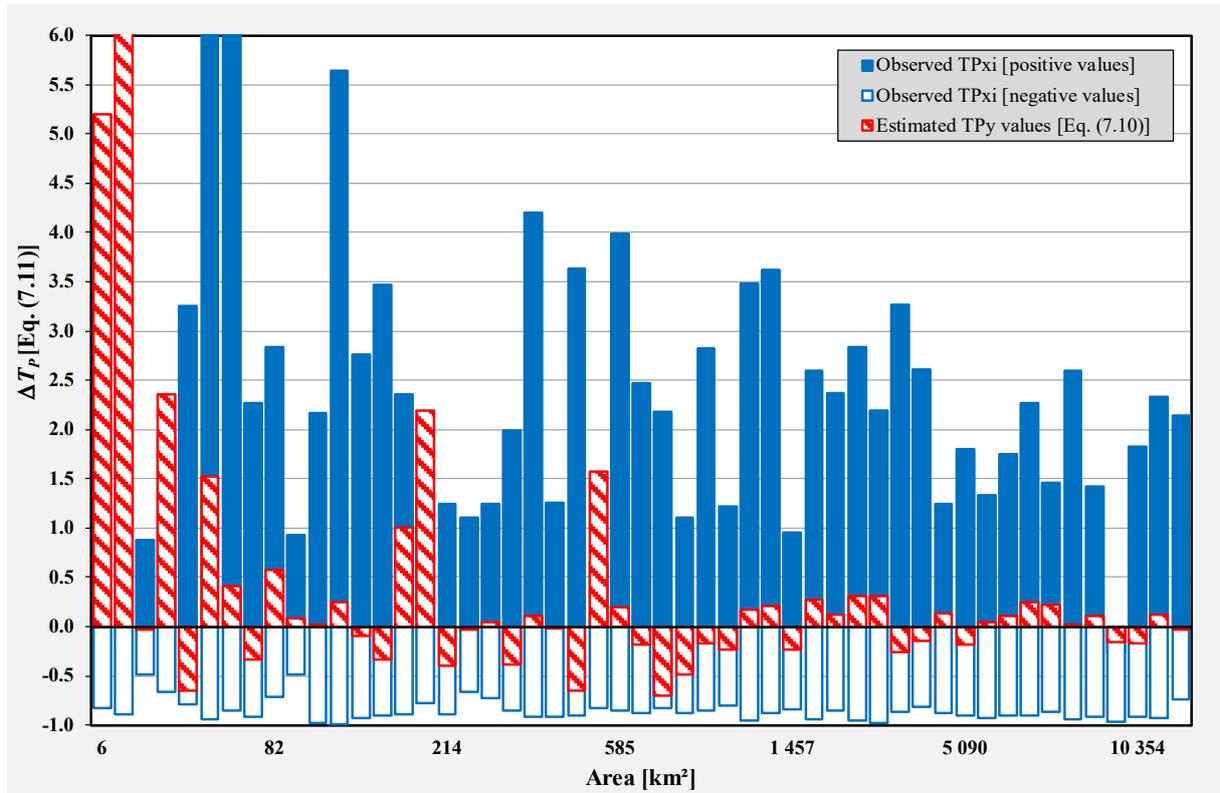
Figure 7.11:   Catchment response time variability [Eq. (7.11)] at a catchment level in Primary Drainage Region X

Furthermore, the validity of the GOF results listed in Table 7.4, is also confirmed by and evident from Figure 7.11, since the T_{Py} estimates are well within the bounds of the maximum individual-event observed T_{Pxi} variability in each catchment. This also confirms that the response of a catchment is most likely to be influenced by a combination of geomorphological catchment characteristics and not by a single catchment characteristic. Hence, the inclusion of a slope predictor (S) in Eq. (7.10) is regarded as essential to ensure that both the size (A) and distance (L_C and L_H) predictors provide a good indication of catchment response times. The distance predictors, in conjunction with the catchment area (A), also proved to be useful in describing the different catchment shapes present in Primary Drainage Region X.

7.4   Estimation of Time Parameter Proportionality Ratios

In considering the inconsistent use of time parameter definitions and the inherent procedural limitations associated with the rainfall-runoff convolution process as discussed in Chapter 3, the overall purpose of this section was to investigate and establish the suitability of the currently recommended time parameter definitions and proportionality ratios for small catchments in larger catchment areas exceeding 50 km². As pilot case study, the 16 gauged sub-catchment areas in the C5 secondary drainage region in South Africa, as shown in Figure 7.12, were considered. Typically, these catchment areas range between 39 km^2 and 33 278 km^2 and there are 185 SAWS daily rainfall stations located within the pilot study area. However, currently, there are only 40 active SAWS rainfall stations available in the C5 region, while only 169 SAWS rainfall stations proved to have adequate historical data both in terms of record length and data quality.
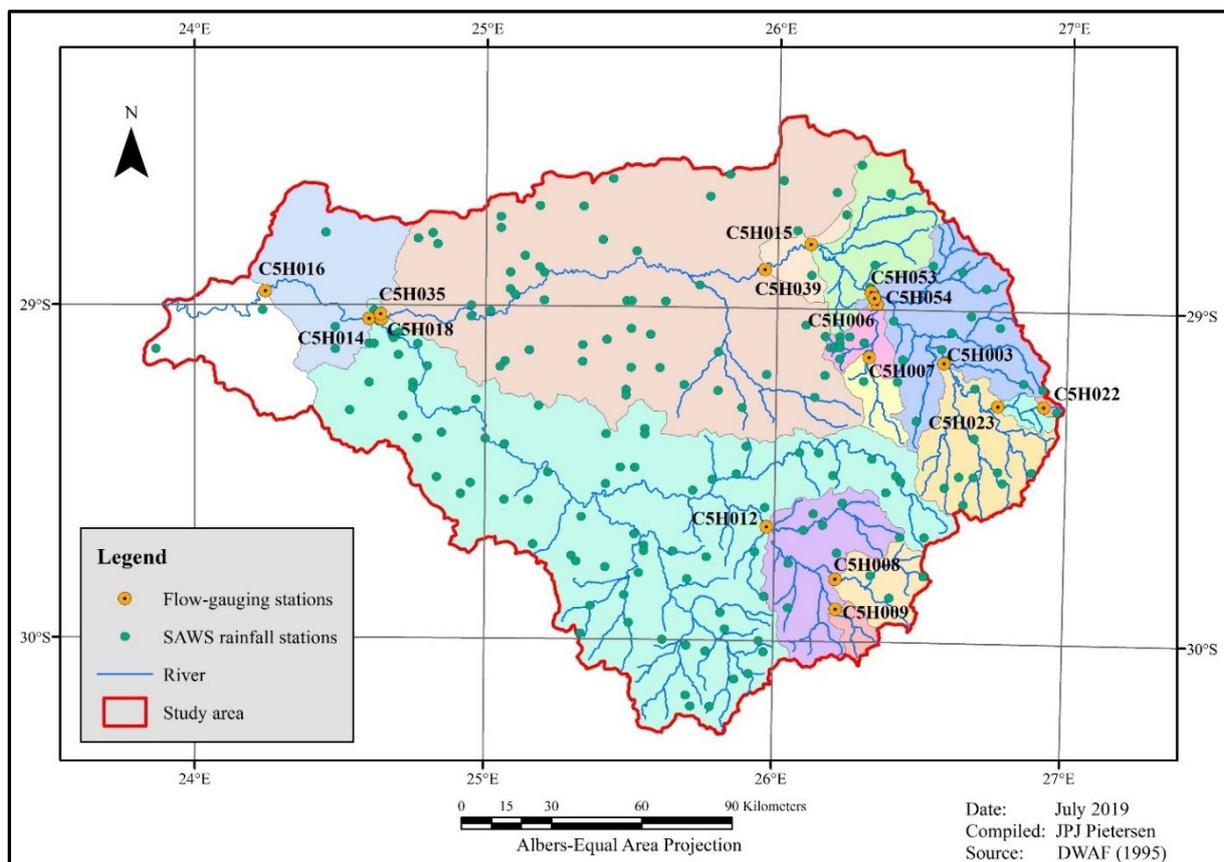
Figure 7.12:   Location of the flow-gauging and daily rainfall stations within the C5 region

The focus was on the development of an automated hyetograph-hydrograph analysis tool to estimate time parameters and average time parameter proportionality ratios at a catchment level. The methodological approach and results are discussed in the subsequent sections.

7.4.1 Synchronisation of rainfall data

The degree of synchronisation between the point rainfall data sets at each rainfall station was established by considering recorded rainfall with mutual time intervals. The rainfall data series at each rainfall station was firstly exported and converted to a Microsoft Excel file (*.xlsx). Thereafter, the rainfall data files were imported to the Automated Toolkit (*cf.* Section 7.4.4). In essence, a number of logic and synchronisation functions are available in the Visual Basic for Applications (imbedded in Microsoft Excel) environment to enable the automatic synchronisation of daily rainfall data, e.g. 'INDEX' and 'MATCH'. The use of the Automated Toolkit ensured that large data sets from numerous rainfall stations within a particular sub-catchment could be synchronised within minutes.

7.4.2 Averaging of observed rainfall data

Given the even spatial distribution of the rainfall stations and the relatively flat topography of the C5 region, in conjunction with the large amount of data and computations required, the Thiessen polygon method [Eq. (2.2), Chapter 2] was used in each sub-catchment to convert the individual point rainfall hyetographs into an average catchment rainfall hyetograph.

7.4.3 Establishment of streamflow database

A total of 1 134 complete hydrographs or runoff events were extracted from the primary flow data sets by using the selection criteria as proposed by Gericke and Smithers (2017; 2018).

7.4.4 Development of automated toolkit

The Automated Toolkit consists of a collection of functions required to estimate the temporal characteristics from rainfall and streamflow records, including: (i) baseflow separation, (ii) time variable identification and estimation, (iii) time parameter estimation, and (iv) the estimation of time parameter proportionality ratios. Typically, the following modules are available in the Automated Toolkit: (i) general catchment information, (ii) processing of observed daily rainfall data, (iii) extracted streamflow data, (iv) analysis and plotting of hyetograph-hydrograph relationships, and (v) exporting of individual hyetograph-hydrograph pairs and summary of results.

The function for baseflow separation is based on Eq. (7.1), while the remaining functions are proposed as a mechanism to extract compounded catchment hyetographs from multiple rainfall

stations with mutual or synchronised events of recorded rainfall. The Automated Toolkit attempts to mimic the typical convolution procedure practitioners would follow to visually inspect and interpret hyetograph-hydrograph data sets. Rainfall and streamflow data are exported to the corresponding modules in the toolkit, followed by the working processes and analyses as summarised in Figure 7.13.
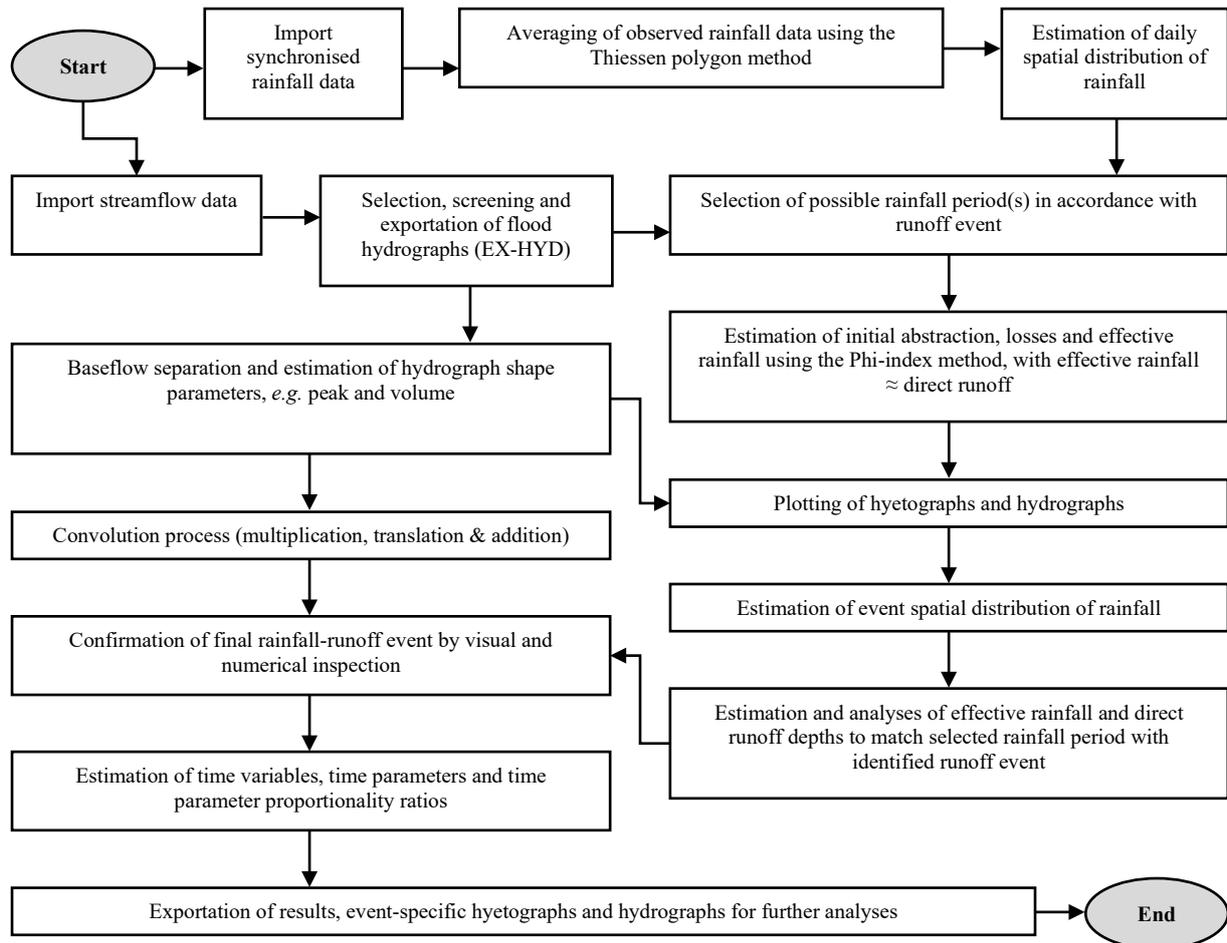
Figure 7.13: Schematic flow diagram of Automated Toolkit (after Allnutt et al., 2020)

7.4.5 Hyetograph analyses

In order to analyse rainfall hyetographs, the associated runoff events need to be identified first. Consequently, a Visual Basic search algorithm was employed to identify the causal rainfall event in a window spanning n days before the start of the identified runoff event to the time of the last streamflow recording, where n is a user-defined time interval. For example, if $n = 12$ days, all rainfall records located in the window 12 days before the start of the runoff event to the last streamflow recording will be identified. The rainfall event starts at the first zero rainfall record in the search window and ends at the last zero recording. Subsequently, after the

averaging of observed rainfall data per rainfall station and the synchronisation of mutual time interval rainfall-runoff events, the daily spatial distribution of any rainfall event was estimated using Eq. (7.12).

$$S_d = \left(\frac{\sum A_T T_{Wi}}{A_T} \right) 100 \qquad [7.12]$$

where

S_d = daily spatial distribution [%],

A_T = total catchment area [km^2], and

T_{Wi} = Thiessen weight of each rainfall station that contributed to the daily rainfall.

During a rainfall event, not all the rainfall contributes to direct runoff. Initial abstractions, e.g. evaporation, transpiration, depression, detention, infiltration and interception by vegetation, reduce the effective runoff producing rainfall that a catchment receives. The Phi-index method [Eq. (7.13)] was used to yield an effective rainfall hyetograph.

$$I = \frac{P_T - Q_D}{t} \qquad [7.13]$$

where

I = Phi-index [mm.h^{-1}],

P_T = total rainfall [mm],

Q_D = direct runoff, which equals the effective rainfall [mm], and

t = time period during which effective rainfall occurred [h].

Hence, Eq. (7.13) enabled the plotting of possible hyetograph-hydrograph combinations to ultimately translate the effective runoff producing rainfall into direct runoff using a simplified convolution process as shown in Figure 7.14.

The selection of an appropriate hyetograph-hydrograph event is characterised by the effective rainfall being equal to the direct runoff (as obtained from the baseflow separation applied to the hydrographs in Section 7.4.6). In cases where the effective rainfall and direct runoff volumes are not in equilibrium, an alternative rainfall period was selected and the process was repeated until equilibrium is reached.
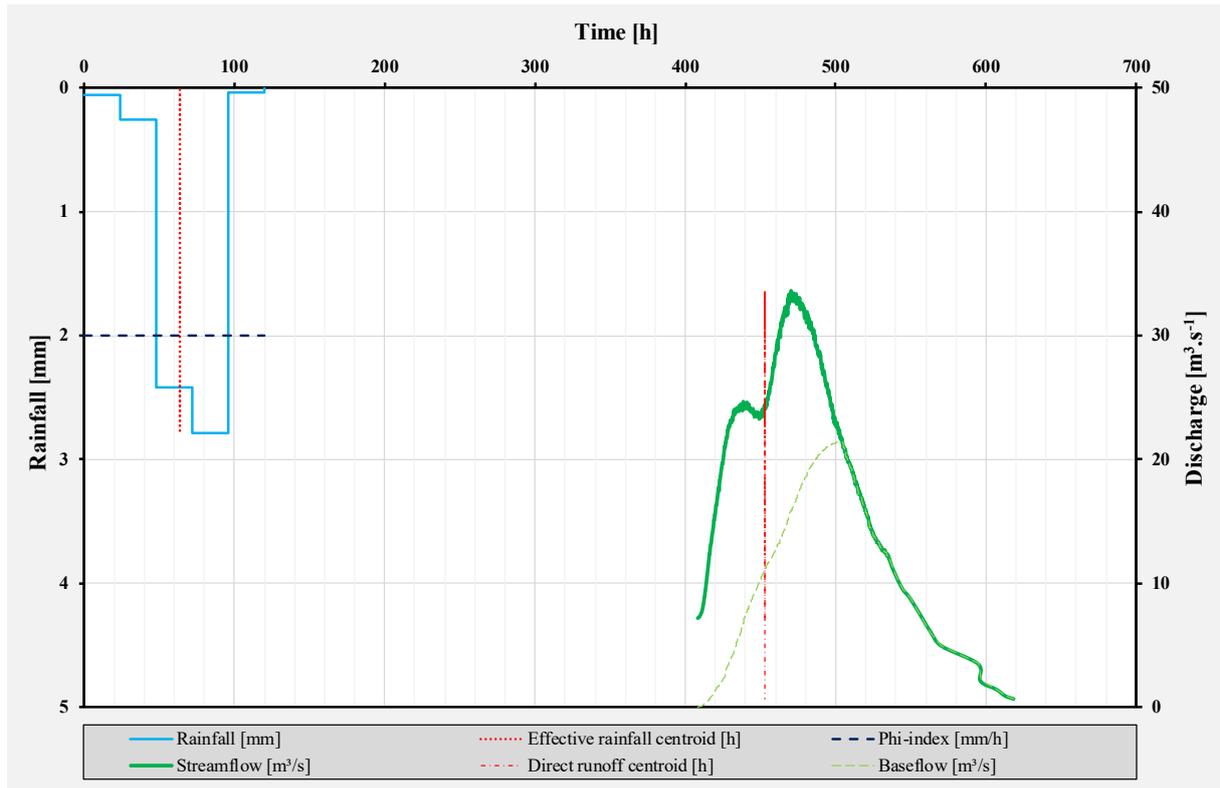
Figure 7.14: Example of a simplified convolution process with a compounded catchment rainfall hyetograph and resulting streamflow hydrograph (after Allnutt et al., 2020)

In each case, the event spatial distribution [Eq. (7.14)] is also automatically estimated for each rainfall period.

$$ S_e = \left[\Sigma \left(\frac{P_i}{\sum\limits_{i=0}^{r-1} P_i} S_{di} \right) \right] 100 \qquad [7.14] $$

where

S_e	= event spatial distribution [%],
i	= number of frequency,
P_i	= weighted daily rainfall [mm],
$\sum\limits_{i=0}^{r-1} P_i$	= cumulative frequency of weighted daily rainfall [mm],
r	= range of frequency, and
S_{di}	= daily spatial distribution [%].

The application of Eq. (7.14) and matching of rainfall-runoff events with corresponding effective rainfall and direct runoff volumes are discussed in the next section. However, it is

important to note that the identification and estimation of time variables e.g. start of effective rainfall (t_{er0}), centroid of effective rainfall (t_{erc}), end of effective rainfall (t_{ere}), and time of maximum rainfall (t_{rmax}) for each rainfall-runoff event, were already possible at this stage.

7.4.6 Hydrograph analyses

The convolution process required to assess the time parameters, e.g. T_C, T_L and T_P, was based on the temporal relationship between an average compounded catchment rainfall hyetograph and a corresponding hydrograph in each sub-catchment. Conceptually, the proposed procedure is based on the definition that the volume of effective rainfall equals the volume of direct runoff when a hydrograph is separated into direct runoff and baseflow. The separation point on the hydrograph is also regarded as the start of direct runoff which coincides with the start of effective rainfall.

As highlighted above, Eq. (7.1) was used for the baseflow separation. As noted in Section 7.4.5, the volumes of effective rainfall and direct runoff need to be in equilibrium when a causal rainfall event of appropriate duration prior to the resulting runoff event is selected. This was done by matching the direct runoff depth (Q_D) with the effective rainfall depth (P_E) in Eq. (7.15).

$$P_E \quad = \frac{\Sigma\left(\dfrac{Q_{Dxi} + Q_{Dx(i-1)}}{2}\Delta T_{xi}\right)}{1000 A_T S_e} \qquad\qquad [7.15]$$

where

P_E = effective rainfall [mm],

A_T = total catchment area [km^2],

Q_{Dxi} = filtered direct runoff as obtained from using Eq. (7.1) [m^3.s^{-1}],

S_e = event spatial distribution [%], and

ΔT_{xi} = absolute change in time at time step i [sec].

As a result, time variables, e.g. start of total runoff (t_{q0}), time of peak discharge (t_{qpk}), centroid of direct runoff (t_{qc}), and time of the inflection point on the recession limb (t_{ip}) were identified and estimated for each rainfall-runoff event at a sub-catchment level.

A total of 394 hyetograph-hydrograph data sets representative of specific rainfall-runoff events were extracted and analysed using the Automated Toolkit. A number of the initially extracted

runoff events (1 134 hydrographs) could not be analysed due to a lack of rainfall data after the year 2001. Consequently, this resulted in a shortfall; however, a number of runoff events could also not be analysed due to the difficulty experienced to identify the inflection point on the recession limb of hydrographs or due to multi-peaked hydrographs. In essence, only 35% of the extracted runoff events could be analysed, i.e. the 394 rainfall-runoff events.

7.4.7 Estimation of time parameters and proportionality ratios

Table 7.5 provides a summary of the different Time Parameter (TP) equations and Time Parameter Proportionality Ratio (TPPR) estimation procedures included in the Automated Toolkit. All these time parameter definitions were introduced in Chapter 3 (*cf.* Figure 3.1), but are included again for the ease of reference. Hence, the letter in brackets () is used to cross-reference to the T_C (a) to (d) and T_L (a) to (c) definitions as shown in Figure 3.1, and defined in Sections 3.1.3 and 3.1.4, respectively.

Table 7.5: Summative description of TP equations and TPPR estimation procedures included in the Automated Toolkit (after Allnutt et al., 2020)

Symbol	Equation	Definition
T_C (a)	t_{ip} - t_{ere}	The time from the end of effective rainfall to the inflection point on the hydrograph recession limb, i.e. the end of direct runoff.
T_C (b) & T_L (a/b)	t_{qpk} - t_{erc}	The time from the centroid of effective rainfall to the peak discharge of the total or direct runoff.
T_C (c)	t_{qpk} - t_{rmax}	The time from the maximum rainfall intensity to the peak discharge.
T_L (c)	t_{qc} - t_{erc}	The time from the centroid of effective rainfall to the centroid of direct runoff.
T_C (d)	t_{qpk} - t_{q0}	The time from the start of direct runoff (rising limb of hydrograph) to the peak discharge.
TPPR 1	$\dfrac{T_c(a)}{T_L(a\,or\,b)}$	Time Parameter Proportionality Ratio (1)
TPPR 2	$\dfrac{T_c(b)}{T_L(a\,or\,b)}$	Time Parameter Proportionality Ratio (2)
TPPR 3	$\dfrac{T_c(c)}{T_L(a\,or\,b)}$	Time Parameter Proportionality Ratio (3)
TPPR 4	$\dfrac{T_c(d)}{T_L(a\,or\,b)}$	Time Parameter Proportionality Ratio (4)
TPPR 5	$\dfrac{T_c(a)}{T_L(c)}$	Time Parameter Proportionality Ratio (5)
TPPR 6	$\dfrac{T_c(b)}{T_L(c)}$	Time Parameter Proportionality Ratio (6)
TPPR 7	$\dfrac{T_c(c)}{T_L(c)}$	Time Parameter Proportionality Ratio (7)
TPPR 8	$\dfrac{T_c(d)}{T_L(c)}$	Time Parameter Proportionality Ratio (8)

In considering the analyses of the 394 hyetograph-hydrograph events, it was quite evident that the seven different time parameter definitions in Table 7.5 contributed to the time parameter variability, which is also influenced by the event spatial distribution (S_e), the variation in peak discharge (Q_P) and the distance (L) between the rainfall station (where the maximum rainfall depth was recorded) and the sub-catchment outlet. Typical results associated with the impact of S_e, Q_P and L values on the estimation of time parameters are listed in Tables 7.6 and 7.7, respectively.

In general, the largest Q_P and direct runoff (Q_D) values were associated with the likelihood of the entire sub-catchment receiving rainfall of a high intensity for the critical storm duration, which in principal, represents the conceptual T_C. Shorter response times, i.e. lower T_C, T_L and T_P values could be expected to occur when the effective rainfall does not cover the entire catchment, especially when a rainfall event is centred near the outlet of a sub-catchment.

Table 7.6: Example of the association between time parameters and the distance (L) of a rainfall event from the catchment outlet in sub-catchment C5H035 ($A = 17\ 359$ km²) (after Allnutt et al., 2020)

Event #	L [km]	Q_P [m³.s⁻¹]	T_C (a) [h]	T_C (b) & T_L (a/b) [h]	T_C (c) [h]	T_L (c) [h]	T_C (d) [h]
17	8.5	11.9	477.9	484.7	484.7	486.6	4.7
20	8.5	197.5	356.8	485.7	498.8	485.8	18.8
21	15.9	155.3	287.2	254.4	238.8	218.1	166.8
9	30.3	178.4	174.0	136.4	97.7	177.1	25.7
2	76.1	12.1	187.8	128.0	128.0	142.3	8.0
18	76.1	17.8	1 115.8	1 109.8	1 109.8	1 112.0	5.8
22	76.1	33.7	384.2	406.5	398.2	389.2	62.2
24	82.3	28.2	388.6	376.2	376.2	385.6	40.2
1	84.0	110.6	688.7	766.2	770.3	750.6	98.3
5	95.0	10.6	463.9	490.8	510.4	493.3	6.4
4	95.0	18.1	933.7	779.4	779.4	784.4	11.4
25	95.0	77.1	845.6	942.5	972.6	942.7	12.6
12	117.0	12.6	587.1	555.9	555.9	560.6	3.9
15	117.0	157.5	526.6	520.8	520.8	522.2	16.8
8	142.1	12.1	198.1	148.3	148.3	155.6	4.3
14	171.3	10.9	457.9	443.9	436.5	447.0	4.5
13	194.8	18.7	232.4	201.1	201.1	206.3	9.1
3	196.8	12.8	397.7	339.6	339.6	344.3	3.6
19	204.6	30.1	950.0	947.6	946.0	948.2	10.0
23	212.6	12.1	347.8	364.4	364.4	366.3	4.4
11	212.6	15.7	616.3	604.9	604.9	610.0	4.9

It is evident from Table 7.6, that an increase in the distance (L) of a rainfall event from the sub-catchment outlet was generally associated with an increase in the time parameter values, while rainfall events which occurred close to the sub-catchment outlets, are more susceptible to

shorter response times. However, in some cases, due to low rainfall intensities resulting in lower Q_P values, the time parameter values are higher and the distance from the sub-catchment outlet has no apparent effect on the response time.

It is evident from Table 7.7 that the largest Q_P and time parameter values are associated with the likelihood of the entire catchment receiving rainfall for the critical storm duration. Lower time parameter values could be expected when effective rainfall of high intensity does not cover the entire catchment, i.e. low S_e values, especially when a storm is centred near the outlet of a sub-catchment. However, in some instances, low rainfall intensities and associated lower peak discharges are ascribed to larger time parameters values, i.e. longer response times due to rainfall events having a low spatial distribution more distant from the sub-catchment outlet.

Table 7.7: Example of the association between time parameters and the spatial distribution of a rainfall event (S_e) from the catchment outlet in sub-catchment C5H035 (A = 17 359 km²) (after Allnutt et al., 2020)

Event #	S_e [%]	Q_P [m³.s⁻¹]	T_C(a) [h]	T_C(b) & T_L(a/b) [h]	T_C(c) [h]	T_L(c) [h]	T_C(d) [h]
18	7.9	17.8	1 115.8	1 109.8	1 109.8	1 112.0	5.8
14	9.3	11.0	457.9	444.0	436.5	447.0	4.5
23	11.8	12.1	347.8	364.4	364.4	366.3	4.4
12	13.2	12.6	587.1	555.9	555.9	560.	3.9
17	15.1	11.8	477.9	484.7	484.7	486.6	4.7
13	16.2	18.7	232.4	201.1	201.1	206.3	9.1
24	17.4	28.1	388.6	376.2	376.2	385.6	40.2
19	19.0	30.1	950.0	947.6	946.0	948.2	10.0
5	20.8	10.6	463.9	490.8	510.4	493.4	6.4
25	21.3	77.1	845.6	942.5	972.6	942.7	12.6
8	21.3	12.1	198.1	148.3	148.3	155.6	4.3
22	21.3	33.7	384.2	406.5	398.2	389.2	62.2
3	24.2	12.8	397.7	339.6	339.6	344.3	3.6
15	24.9	157.5	526.6	520.8	520.8	522.2	16.8
20	27.9	197.5	356.8	485.7	498.8	485.8	18.8
2	28.7	12.1	187.8	128.0	128.0	142.3	8.0
1	32.2	110.6	688.7	766.2	770.3	750.6	98.3
21	33.7	155.3	287.2	254.4	238.8	218.1	166.8
4	35.9	18.1	933.7	779.4	779.4	784.4	11.4
11	37.5	15.7	616.3	604.9	604.9	610.0	4.9
9	46.2	178.4	174.0	136.4	97.7	177.1	25.7

In considering the T_C, T_L and T_P pair values obtained from the 394 hyetograph-hydrograph events, a relatively low variability was evident between the different time parameter proportionality ratios (TPPR 1 to TPPR 8; *cf.* Table 7.5) at a sub-catchment level. In general, where T_L is defined as the time from the centroid of effective rainfall to the peak discharge [T_L definitions (a/b)], T_C and T_L are related by $T_C = 1.003T_L$ (TPPR 1 to TPPR 3). In using T_L

defined as the time from the centroid of effective rainfall to the centroid of direct runoff [T_L definition (c)], the proportionality ratio reduced to 0.992 (TPPR 5 to TPPR 7).

7.4.8   Conclusions

An enhanced methodology was developed which considers both the impact of the spatial distribution of rainfall events and the distance thereof from the catchment outlet on the resulting runoff. The major findings are as follows:

(a) Time parameter estimates based on the seven different theoretical time parameter definitions proved to be highly variable due to the spatial and temporal distribution of rainfall events, variation in peak discharges and the distance of the rainfall events from the catchment outlet.

(b) Time parameter proportionality ratios are characterised by a relatively low variability at a larger catchment level in the C5 secondary drainage region.

(c) In all the sub-catchments under consideration, Study Assumption 4, i.e. $T_C \approx T_L \approx T_P$, was confirmed. In other words, it highlighted that the proportionality ratios currently proposed for small catchments, i.e. $T_C = 1.417T_L$ and $T_C = 1.667T_L$, are not applicable at larger catchment levels.

7.5   Extension of Stage-Discharge Relationships using Indirect Estimation Methods

The purpose of this section is to address Study Assumption 1, i.e. that there is no 'one size fits all' approach/method available for the extension of stage-discharge rating curves. Consequently, a selection of indirect extension methods (e.g. hydraulic and one-dimensional modelling methods) was evaluated and compared to direct extension (benchmark) methods (e.g. at-site conventional current gaugings, hydrograph analyses and level pool routing techniques), to establish the best-fit and most appropriate stage-discharge extension method to be used in South Africa.

As pilot case study, 10 flow-gauging sites in Limpopo, Gauteng, Free State, Mpumalanga, KwaZulu-Natal and the Western Cape provinces were selected based on the range of possible site conditions present, e.g. type of flow-gauging weir, at-site and river geometry, flow conditions, type of hydraulics controls, and data availability.

As shown in Figure 7.15, these 10 flow-gauging sites are located in secondary drainage regions A4, A6, C2, C5, H5, J1, K2, V1 and X3.
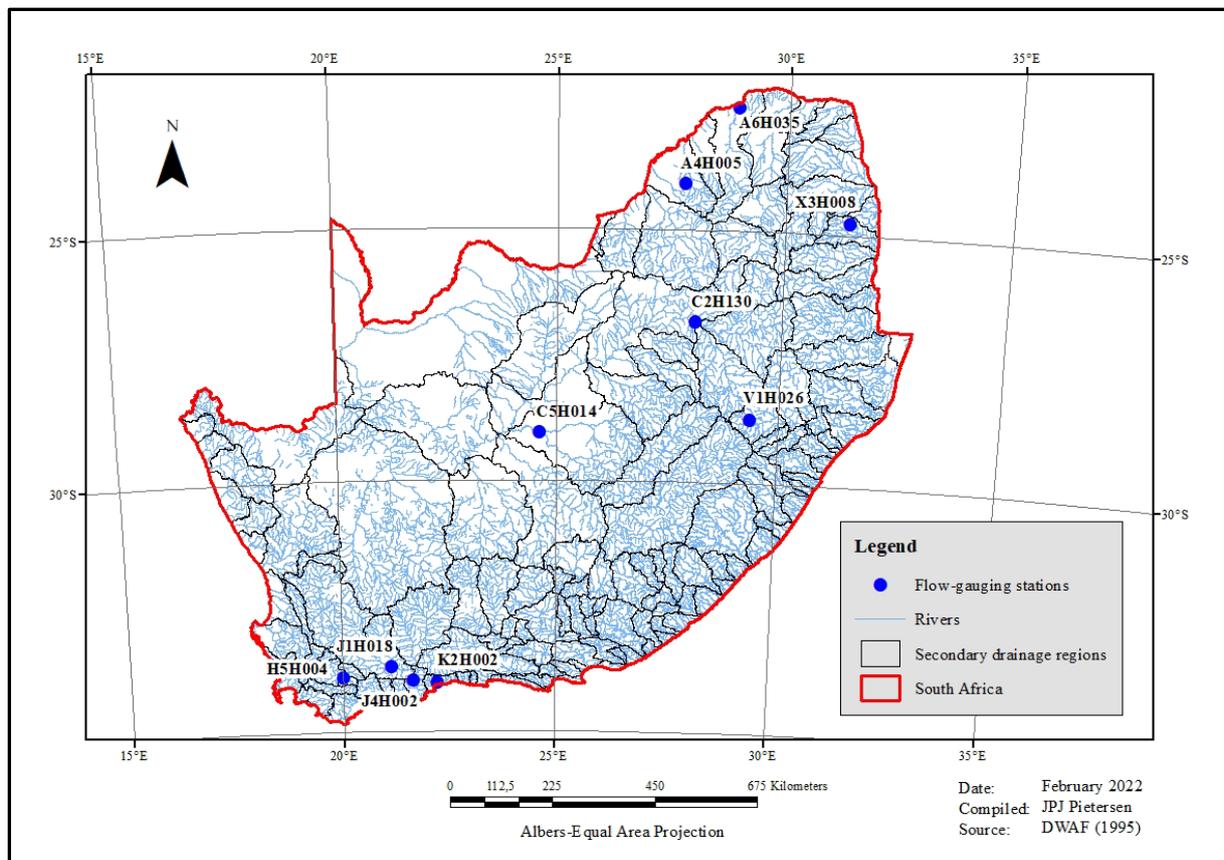
Figure 7.15: Location of the 10 flow-gauging sites in South Africa

7.5.1 Data collection and processing

Data were collected based on the hydrometrical and geometrical requirements for the extension of stage-discharge relationships. Initially, all the stage-discharge extension reports available from DWS, i.e. the *Discharge Table Improvement Reports*, were studied given that a standard stage-discharge rating table is regarded as the first step towards the extension of any stage-discharge relationship. In terms of the hydrometrical and geometrical requirements, the following aspects were considered:

(a) **Streamflow data:** Record length, data quality, flow duration curves to highlight the occurrence and frequency of minimum and maximum flow ranges, and the number of standard and extended stage-discharge relationship tables/curves available.

(b) **Hydraulic conditions:** Modular and/or submerged flow conditions, variable submergence due to backwater effects and vegetative growth, identification of possible

hydraulic controls, sediment transport, unsteady flow conditions, and the influence of in-bank and out-of-bank flow paths.

(c) **Geometrical properties:** Type of flow-gauging weir, e.g. sharp/broad-crested, Crump, flumes, broad-crested flank walls, and flood/natural sections, overall river topography and layout, river channel and flood plain geometry, position of control points within the river system (especially at high flows), availability of survey data, e.g. cross-sections, longitudinal sections, previous flood surveys, and the previous/current allocation of site-specific roughness parameters (typical range possible to account for seasonal variability).

The processing of the geometric data, e.g. wetted perimeter, wetted area and hydraulic radius, was done by using the Windows Cross-Section Professional (WinXSPRO; Hardy et al., 2005), which is essentially a channel cross-section analyser as shown in Figure 7.16. All the stage-discharge extensions were executed in the Microsoft Excel environment using semi-automated tools.
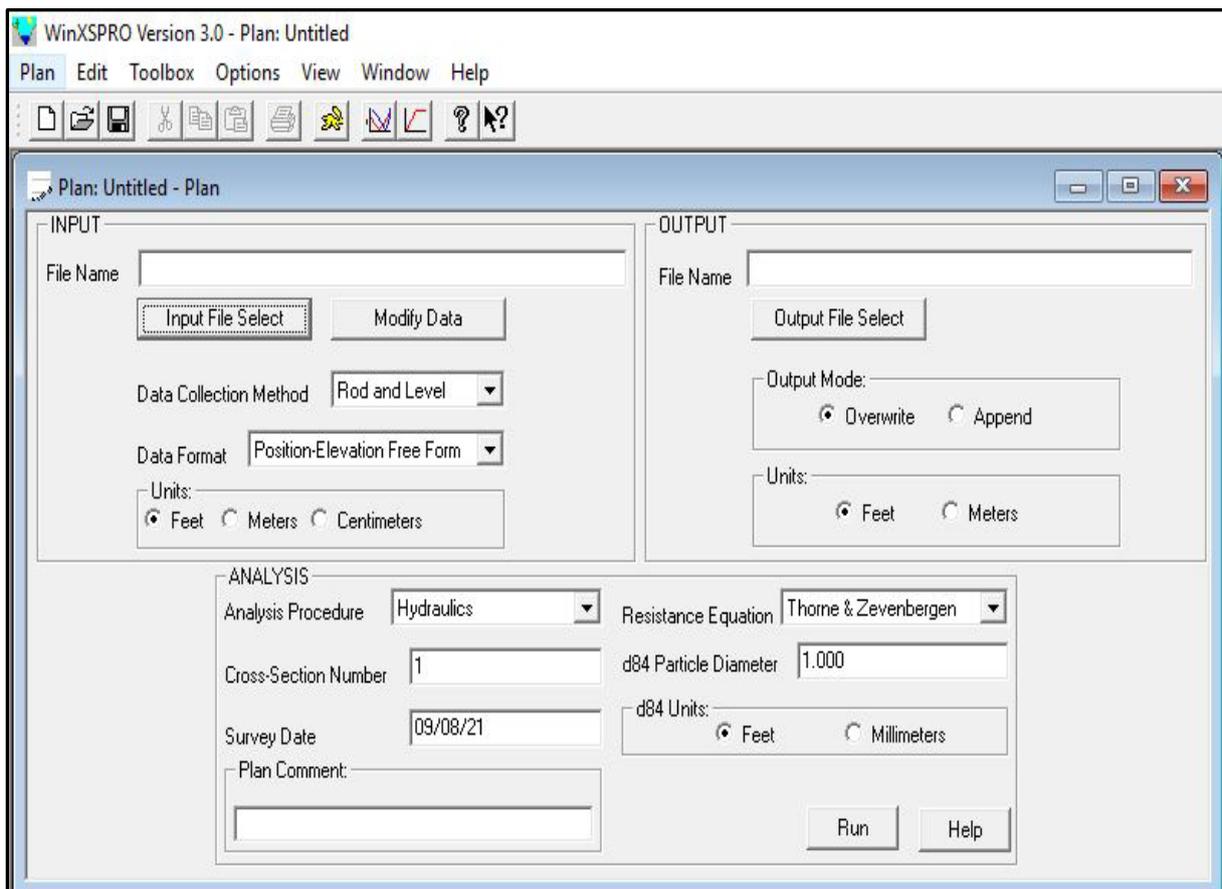
Figure 7.16: WinXSPRO User-interface

7.5.2 Extension of stage-discharge relationships

The following hydraulic methods were considered and applied at each site: (i) Simple Extension (SE), (ii) Log Extension (LE), (iii) Velocity Extension Simple Approach (VE-SA), (iv) Velocity Extension Hydraulic Radius Approach (VE-HRA), (v) Velocity Extension Manning's Approach (VE-MA), (vi) Slope Area Method (SAM), and (vii) Stepped Backwater Analysis (SBA). In addition, one-dimensional modelling was conducted using the Hydrologic Engineering Centre River Analysis System (HEC-RAS).

A ranking-based selection procedure using a selection of GOF criteria, e.g. standard error of estimate (SE_E), Eq. (7.16), mean absolute relative error ($MARE$), Eq. (7.17), root mean square error ($RMSE$), Eq. (7.18), coefficient of determination (r^2), Eq. (7.19), Nash-Sutcliffe coefficient (NSE), Eq. (7.20), and the Z-test [Eq. (7.21)], were used to assess the indirect estimation methods' results to those of the at-site direct extension (benchmark) methods.

$$SE_E = \left[\frac{1}{(N-2)} \left(\sum_{i=1}^{N}\left(Q_{Ei}-\overline{Q_E}\right)^2 - \frac{\left(\sum_{i=1}^{N}\left(Q_{Bi}-\overline{Q_B}\right)\left(Q_{Ei}-\overline{Q_E}\right)\right)^2}{\sum_{i=1}^{N}\left(Q_{Ei}-\overline{Q_B}\right)^2} \right) \right]^{0.5} \quad [7.16]$$

$$MARE = 100\left[\frac{1}{N}\sum_{i=1}^{N}\frac{|Q_{Ei}-Q_{Bi}|}{Q_{Bi}} \right] \quad [7.17]$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(Q_{Bi}-Q_{Ei}\right)^2}{N}} \quad [7.18]$$

$$r^2 = \left[\frac{\sum_{i=1}^{N}\left(Q_{Bi}-\overline{Q_B}\right)\left(Q_{Ei}-\overline{Q_E}\right)}{\sqrt{\sum_{i=1}^{N}\left(Q_{Bi}-\overline{Q_B}\right)^2 \sum_{i=1}^{N}\left(Q_{Ei}-\overline{Q_E}\right)^2}} \right]^2 \quad [7.19]$$

$$NSE = 1 - \left[\frac{\sum_{i=1}^{N}\left(Q_{Bi}-Q_{Ei}\right)^2}{\sum_{i=1}^{N}\left(Q_{Bi}-\overline{Q_B}\right)^2} \right] \quad [7.20]$$

$$Z = \frac{\overline{Q_{Bi}}-\overline{Q_{Ei}}}{\sqrt{\sigma_{Bi}+\sigma_{Ei}}} \quad [7.21]$$

where

SE_E = standard error of estimate [$m^3.s^{-1}$],

$MARE$ = mean absolute relative error [%],

$RMSE$ = root mean square error,

r^2 = coefficient of determination,

NSE = Nash-Sutcliffe coefficient,

Z = Z-test/score,

N = sample size,

Q_{Bi} = at-site benchmark discharge [$m^3.s^{-1}$],

$\overline{Q_B}$ = mean of the at-site benchmark discharge [$m^3.s^{-1}$],

Q_{Ei} = estimated (indirect) discharge using extension methods [$m^3.s^{-1}$],

$\overline{Q_E}$ = mean of the estimated (indirect) discharge [$m^3.s^{-1}$],

σ_{Bi} = standard deviation of the at-site benchmark discharge [$m^3.s^{-1}$], and

σ_{Ei} = standard deviation of the estimated (indirect) discharge [$m^3.s^{-1}$].

Given that all of the indirect extension methods listed above, were deployed at flow-gauging site X3H008, the subsequent section focuses on the latter gauging site, followed by the overall results and performance achieved at the 10 gauging sites under consideration.

7.5.3   Gauging weir X3H008

Gauging weir X3H008 is located in the Sand River, Mpumalanga Province. The weir was constructed in 1967 and consists of a sharp crest, hydro flume and broad-crested flank walls as shown in Figure 7.17.

The structural limit is 1.95 m (gauge plate reading), with an associated discharge of 76 $m^3.s^{-1}$. In the latest DWS Calibration Report, the SBA was used as the preferred stage-discharge extension method. Hence, the SBA was also used as benchmark against which the other indirect extension methods were compared up to a stage of 5.4 m and corresponding discharge of 2 972 $m^3.s^{-1}$. The most recent cross-sectional survey of the gauging weir was conducted on 07 August, but the year is unclear [n.d.]. Five flood sections (surveyed cross-sections up to high watermarks) were used in the SBA. The current (existing) stage-discharge relationship table (DT10) applicable to X3H008 is listed in Table 7.8.

Figure 7.17: Compound gauging-weir X3H008

Table 7.8: DT 10 at X3H008 (Nathanael et al., 2018)

Stage [m]	Stage increments [m]/ Discharge [m³.s⁻¹]									
	0	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
0	0.0	0.2	0.6	1.1	1.8	2.8	4.7	7.1	9.9	13.0
1	16.5	20.3	24.4	28.7	34.1	39.1	45.4	52.7	61.0	70.2
2	84.7	107.6	137.0	172.4	213.8	261.2	315.2	372.1	429.1	488.9
3	552.8	618.7	688.0	760.3	836.0	914.0	995.0	1 080.0	1 167.0	1 258.0
4	1 351.0	1 448.0	1 548.0	1 650.0	1 756.0	1 865.0	1 976.0	2 091.0	2 209.0	2 330.0
5	2 453.0	2 580.0	2 710.0	2 843.0	2 972.0	-	-	-	-	-

Based on the current information available at X3H008, the evaluation and comparison of the indirect stage-discharge extension methods included the SE, LE, VE, SAM, SBA, and HEC-RAS modelling (1-D, and steady and unsteady flow conditions).

The results pertained to the stage-discharge rating curve extension using all the above methods, are shown in Figure 7.18. The GOF statistics applicable to each method are listed in Table 7.9.
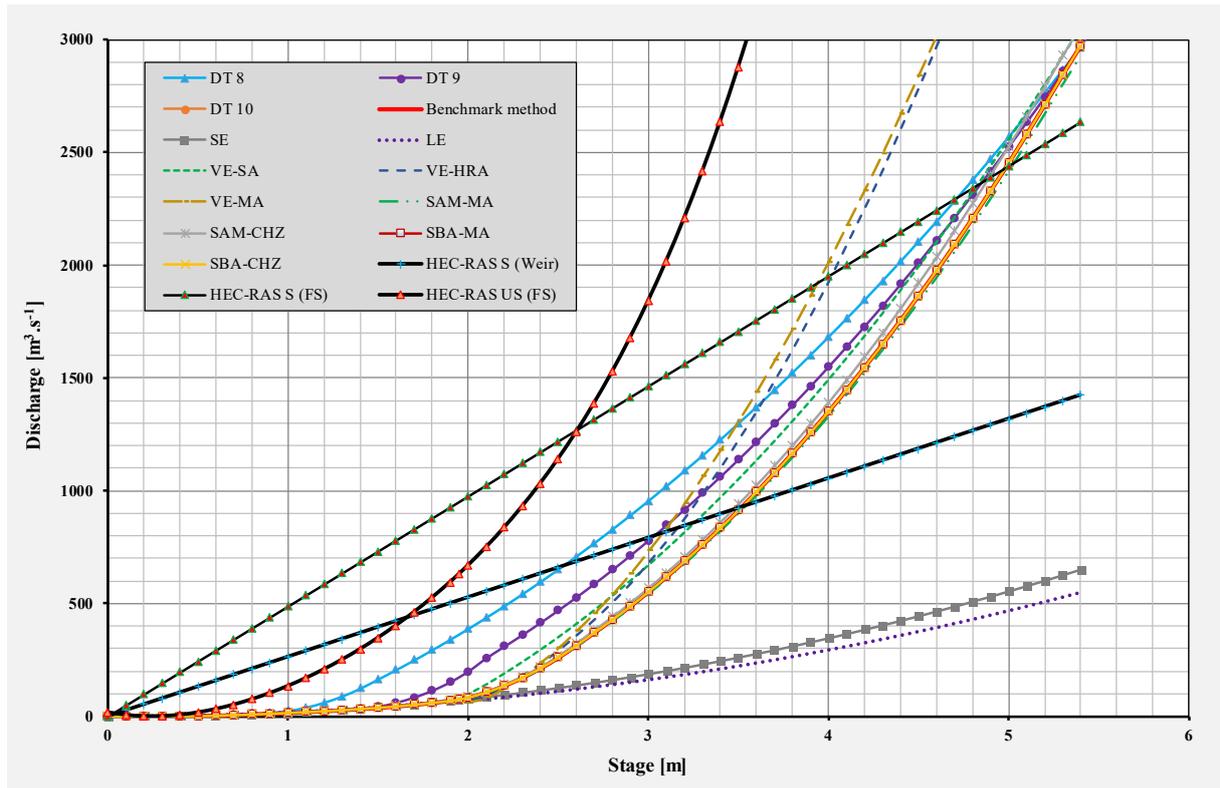
Figure 7.18: Indirect stage-discharge extensions in comparison to the benchmark rating at gauging weir X3H008

Table 7.9: GOF results at X3H008

Criteria	SE	LE	VE			SAM		*SBA		HEC-RAS		
			SA	HRA	MA	MA	CHZ	MA	CHZ	S (Weir)	S (FS)	US (FS)
Sample size (N)	35	35	35	35	35	35	35	35	35	35	35	36
Average	319	271	1 348	1 809	1 839	1 217	1 273	1 235	1 235	975	1 804	4 454
Standard deviation	174	146	899	1 436	1 414	877	917	890	890	270	500	3 580
SE_E [Eq. (7.16); m^3.s^{-1}]	5.7	4.4	31.2	66.3	27.0	0.0	0.2	0.0	0.0	43.8	81.0	519.1
SE_E ranking	**6**	**5**	**8**	**10**	**7**	**3**	**4**	**2**	**1**	**9**	**11**	**12**
$MARE$ [Eq. (7.17); %]	65.4	70.2	16.3	35.0	39.7	1.5	3.0	0.0	0.0	84.4	175.6	254.0
$MARE$ ranking	**8**	**9**	**5**	**6**	**7**	**3**	**4**	**2**	**1**	**10**	**11**	**12**
$RMSE$ [Eq. (7.18)]	1 156.5	1 211.5	116.8	788.4	794.7	22.8	0.0	0.0	0.0	668.9	695.0	3 784.8
$RMSE$ ranking	**10**	**11**	**5**	**8**	**9**	**4**	**3**	**2**	**1**	**6**	**7**	**12**
r^2 [Eq. (7.19)]	0.999	0.999	0.999	0.998	1.000	1.000	1.000	1.000	1.000	0.974	0.974	0.976
r^2 ranking	**7**	**6**	**8**	**9**	**5**	**3**	**4**	**2**	**1**	**11**	**11**	**10**
NSE [Eq. (7.20)]	-0.738	-0.907	0.982	0.192	0.179	0.999	1.000	1.000	1.000	0.419	0.372	-18.146
NSE ranking	**10**	**11**	**5**	**8**	**9**	**4**	**3**	**2**	**1**	**6**	**7**	**12**
Z-test [Eq. (7.21)]	28.1	30.0	2.7	11.9	12.6	0.4	0.9	0.0	0.0	7.7	15.3	48.1
Z-test ranking	**10**	**11**	**5**	**7**	**8**	**3**	**4**	**2**	**1**	**6**	**9**	**12**
Sum of rankings	**51**	**53**	**36**	**48**	**45**	**20**	**22**	**12**	**6**	**48**	**56**	**70**
Overall ranking	**9**	**10**	**5**	**7**	**6**	**3**	**4**	**2**	**1**	**8**	**11**	**12**

*Benchmark method; MA (Manning's approach); CHZ (Chézy's approach); S (Steady flow); US (Unsteady flow) and FS (Flood sections)

In Table 7.9, the various indirect stage-discharge extension methods, except for the SE, LE and HEC-RAS S (Weir) methods, provided estimated discharge (Q_{Ei}) values that exceed the at-site benchmark discharges (Q_{Bi}). The latter underestimations using the SE and LE methods were evident for stages (H) > 2 m (65% ≤ $MARE$ ≤ 70%), while the HEC-RAS S (Weir) underestimations ($MARE$ = 84.4%) are evident at H > 3.6 m, which is almost double the current

structural limit of 1.95 m. Given that the SBA-CHZ method is used as the benchmark method (default ranking = 1) at this site, the SBA-MA demonstrated the second best overall ranking, followed by the SAM-CHZ and SAM-MA methods, respectively. By considering the individual GOF statistics, the latter SAM-CHZ and SAM-MA methods also proved to be equally accurate by being interchangeably ranked at either the 3rd or 4th position. Overall, the HEC-RAS US (FS) method is regarded as the most inappropriate method with the poorest ranking for all the GOF statistics under consideration.

Overall, the methods regarded as being 'hydraulically correct', resulted in the lowest SE_E, *MARE*, *RMSE*, and *z*-test values. All the methods demonstrated a high degree of association, with r^2 values > 0.97; however, this just highlights that there is a high correlative trend between the estimated Q_{Ei} values as suggested by the various extension methods and the benchmark Q_{Bi} values. Hence, the *NSE* results should also be considered, whereas the VE-HRA and VE-MA methods demonstrated the lowest *NSE* values of ±0.19 despite of their r^2 values ≈ 0.99.

7.5.4   Assessment of overall performance

The overall ranking results listed in Table 7.10 highlight that the SBA and SAM performed the best. The other indirect extension methods were characterised by larger statistical differences between the at-site benchmark values (Q_{Bi}) and the modelled values (Q_{Ei}). The HEC-RAS S & US (FS) methods are regarded as the least appropriate methods, respectively ranked in the 11th and 12th position.

Table 7.10:   Summary of the GOF-based ranking of the indirect stage-discharge extension methods applied at the 10 gauging sites

Gauging site	SE	LE	VE			SAM		SBA		HEC-RAS		
			SA	HRA	MA	MA	CHZ	MA	CHZ	S (Weir)	S (FS)	US (FS)
A4H005	4	5	3	1	2	-	-	-	-	6	-	-
A6H035	5	1	1	4	6	-	-	-	-	3	-	-
C3H130	4	1	6	5	3	-	-	-	-	2	-	-
C5H014	12	11	6	9	5	4	3	2	1	8	7	10
H5H004	12	11	8	9	5	3	4	2	1	6	10	7
J1H018	11	7	5	6	12	4	3	2	1	8	9	9
J4H002	6	10	4	7	8	5	1	3	2	-	9	-
K2H002	1	2	3	5	6	-	-	-	-	4	-	-
V1H026	1	2	4	6	5	-	-	-	-	3	-	-
X3H008	9	10	5	7	6	3	4	2	1	8	11	12
Avg. rankings	6.5	6.0	4.5	5.9	5.8	3.8	3.0	2.2	1.2	5.3	9.2	9.5
Overall ranking	10	9	5	8	7	4	3	2	1	6	11	12

The SBA and SAM are therefore regarded as the most appropriate indirect estimation methods to reflect the hydraulic conditions during high discharges at a flow-gauging site. In addition, Chézy's absolute roughness (k_s) version of the latter methods is preferred to the Manning's n-value version and ranked accordingly. This is ascribed to the fact that Manning's n-value is constant and expressed in no-measurable units, i.e. $s.m^{-1/3}$, while k_s is expressed in measurable units (mm or m), which change as changes in the flow depth and subsequently the hydraulic radius (R) occur. As a result, channel profile convergence is normally achieved sooner when Chézy's C-value (function of k_s and R) is used in the SBA.

It is important to note that the starting water surface level in the SBA is regarded as unknown or indefinite; hence, several backwater profiles based on different arbitrary water surface levels should be computed until convergence occurs, i.e. the water levels correspond with the uniform flow profile and energy levels are in equilibrium. Therefore, if convergence is not reached (probably due to insufficient cross-sectional survey data), additional and more distant (up- and/or downstream) cross-sections would be required. In this study at some of the flow-gauging sites under consideration, DWS had already established an arbitrary datum (reference level), and known water surface levels (e.g. flood marks). Subsequently, defining an arbitrary point to reach profile convergence was not required in these cases.

Any extension method must be hydraulically correct if it is to be used as a robust approach to extend stage-discharge rating curves beyond the structural limit. The extension of a rating curve is significantly more affected by the site (and river reach) geometry, initial hydraulic conditions, flow regimes and level of submergence at high discharges, than the actual indirect extension method used.

The next chapter presents a synthesis of all the information as discussed in Chapters 6 to 7, as well as some final conclusions.

CHAPTER 8:   DISCUSSION AND CONCLUSIONS

This chapter contains a discussion and conclusions based on the final methodology and results associated with the derivation of regional ARFs and at-site catchment response time parameters as included in Chapters 6 and 7, respectively.

8.1   ARF Methodology and Results

The main study objective was to estimate geographically-centred 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 6.

A total of 2 550 artificial circular catchments with an associated 1 779 rainfall stations with at least 30 years combined areal record lengths, were strategically positioned in 46 homogeneous rainfall regions throughout South Africa. Due to the large number of circular catchments placed in each of the 46 ARF regions, 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 ARF region. In principal this was not regarded as problematic, while it also contributed to the 'smooth' transition between the different regions.

In applying various screening criteria, only 2 053 circular catchments were used in the probabilistic and regression analyses. The probabilistic analyses based on the GEV_{LM} probability distribution resulted in areal and 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). The estimation of sample ARFs was expressed as the ratio between the areal catchment design rainfall and the design point rainfall estimates for corresponding return periods.

Five (5) ARF regions were deduced from the 46 ARF regions and a single regional empirical ARF equation [Eq. (6.2)], with unique regional calibration coefficients, was assigned to each region. Initially, linear backward stepwise multiple regression analyses with deletion were performed at a 95% confidence level in order to estimate the relationship between the dependent criterion variable (ARF) and the independent predictor variables (catchment area, storm duration and return period) within each of region. Ultimately, the linear regression analyses were outperformed by a second order polynomial non-linear log-transformed empirical ARF equation. The derived regional empirical ARF equation [Eq. (6.2)] performed similarly, and as expected, when compared to a selection of empirical geographically-centred ARF estimation methods currently used in local and/or international practice. The ARFs estimated with Eq. (6.2) decreased within an increase in area, and increased with an increase in both storm duration and return period.

All the above results also confirmed the study assumptions applicable to ARFs, *viz.*: (i) design point rainfall estimates are only representative for a limited area – demonstrated by the differences between areal design rainfall and design point rainfall estimates, (ii) ARFs vary with predominant weather types, storm durations, seasonal factors and return periods – evident in the different ARF regions and hence the reason for having the five (5) ARF regions, and (iii) the current South African ARF estimation methods are only applicable to specific temporal and spatial scales – demonstrated by the absence of any regionalisation, the ARF values exceeding 100% in 'smaller' catchments, the constant ARF values associated with all return periods, and the limited data used.

The ARF methodology used in this study and the subsequent findings are new to the South African flood hydrology research community and practice, *viz.*: (i) ARFs were derived and based on a regionalisation scheme utilising local and up-to-date daily rainfall data, (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 5 ARF regions of South Africa.

8.2   Catchment Response Time Methodology and Results

The primary aim was to expand and verify the approach developed by Gericke and Smithers (2017) by estimating observed catchment response time parameters from 51 gauged catchments

located in Primary Drainage Region X and to derive a regional empirical time parameter equation.

The conceptual approach developed and refined to derive the time to peak (T_{Px}) using only observed streamflow data at a catchment level, proved to be both practical and objective with consistent results. The combined use of Eqs. (7.2), (7.3) and (7.4) not only ensured/will ensure that the high variability of event-based catchment responses is taken into account, but the estimated catchment T_{Px} values are also well within the range of other uncertainties inherent to all design flood estimation procedures. The high degree of association ($r^2 = 0.84$) between Eqs. (7.2) and (7.3) also confirmed Study Assumption 2, i.e. T_P equals the total net rise (duration) of a multiple-peaked hydrograph in medium to large catchments. The average error bounds between the three different approaches, e.g. net rise duration [Eq. (7.2)], triangular-shaped hydrograph approximation [Eq. (7.3)] and linear response function [Eq. (7.4)] were also limited to ≤ 15%; hence, this also served as confirmation of Study Assumption 3.

It is recommended that for design hydrology and for the calibration of empirical time parameter equations, that the catchment T_{Px} should be estimated using Eq. (7.4). In addition, the conceptual approach used to derive the empirical time parameter equation [Eq. (7.10)], should be adopted when regional time parameter equations are derived at a national-scale in South Africa. It is suggested that the methodology developed (and refined) in this study, should be gradually expanded to Primary Drainage Regions A and B, before deploying it at a national-scale. Approximately 110 gauged catchments covering the whole of the Gauteng, Mpumalanga, Limpopo and the Northern Provinces are situated in regions A, B and X. Typically, these three regions do not only form a continuous geographical region, but the largest percentage of South Africa's population also resides here and are frequently subjected to extreme flooding.

8.3   Time Parameter Proportionality Ratios

The overall purpose was to investigate and establish the suitability of the currently recommended time parameter definitions and proportionality ratios for small catchments in larger catchment areas exceeding 50 km². The focus was on the development of an automated hyetograph-hydrograph analysis tool to estimate time parameters and average time parameter proportionality ratios at a catchment level.

The Automated Toolkit for hyetograph-hydrograph analyses proved to be very useful in mimicking the typical convolution procedure practitioners would follow to visually inspect and interpret hyetograph-hydrograph data sets. An enhanced methodology was developed, which considered both the impact of the spatial distribution of rainfall events and the distance thereof from the catchment outlet on the resulting runoff and consequently, the derivation of time parameters and proportionality ratios.

The time parameter estimates based on the seven different theoretical time parameter definitions proved to be highly variable due to the spatial and temporal distribution of rainfall events, variation in peak discharges and the distance of the rainfall events from the catchment outlet. In contrast, the time parameter proportionality ratios were characterised by a relatively low variability. In all the sub-catchments under consideration, Study Assumption 4 was confirmed, i.e. $T_C \approx T_L \approx T_P$. In other words, it was demonstrated that the time parameter proportionality ratios currently proposed for small catchments, i.e. $T_C = 1.417T_L$ and $T_C = 1.667T_L$, are not applicable at larger catchment levels.

8.4 Extension of Stage-Discharge Relationships

In terms of stage-discharge relationships above the structural limit of a flow-gauging weir, a pilot scale study was conducted in 10 gauged catchments to evaluate and compare a selection of indirect extension methods (e.g. hydraulic and one-dimensional modelling methods) to direct extension (benchmark) methods (e.g. at-site conventional current gaugings, hydrograph analyses and level pool routing techniques), to establish the best-fit and most appropriate stage-discharge extension method to be used in South Africa.

Overall, the results highlighted that the Stepped Backwater Analysis (SBA) and Slope Area Method (SAM) are the most appropriate indirect estimation methods to reflect the hydraulic conditions during high discharges at a flow-gauging site. It was emphasised that any extension method must be hydraulically correct if it is to be used as a robust approach to extend stage-discharge rating curves beyond the structural limit, while the extension of any rating curve is significantly more affected by the site (and river reach) geometry, initial hydraulic conditions, flow regimes and level of submergence at high discharges, than the actual extension method used. Hence, the latter also served as confirmation of Study Assumption 1, i.e. there is no 'one

size fits all' approach/method available for the extension of stage-discharge rating curves at a flow-gauging site.

By enhancing the input data available and building on the various methods of extending stage-discharge relationships, the ultimate goal would be to improve the extension of high discharge events to result in more reliable and statistically acceptable estimates. Consequently, the improved rating curves (after extension) will result in improved hydrological data sets, all of which, will contribute towards enhanced operational and water resources planning, management and allocation in South Africa. In terms of future research, the following aspects should be considered: (i) review of the current procedures to determine Manning's *n*-values for flash floods, (ii) development of improved procedures to select Chézy's *C* values, and (iii) alternative methods to extend stage-discharge relationships, e.g. hydrodynamic models, Support Vector Machine (SVM) and Artificial Neural Network (ANN) methods.

8.5   Conclusions

Average areal design rainfall and catchment response time are regarded as fundamental input to all event-based design flood estimation methods in ungauged catchments. Given the sensitivity of design peak discharges to estimated ARFs and catchment response time parameter values, it is envisaged that the implementation of the specific objectives of this study will contribute fundamentally to the improved estimation of both ARFs and time parameters to ultimately result in improved design flood estimations in South Africa. The methodologies developed could also be adopted internationally to improve the estimation of ARFs and catchment response time parameters to provide more reliable peak discharge and volume estimates as, to date, this remains a constant challenge in flood hydrology.

The next chapter contains a summary of: (i) the equipment and resources used, (ii) the project deliverables, (iii) knowledge dissemination, and (iv) the project work plan and achieved milestones.

CHAPTER 9: RESOURCES, DELIVERABLES AND WORK PLAN

This chapter contains a summary of: (i) the equipment and resources used, (ii) the project deliverables, (iii) knowledge dissemination, and (iv) the work plan and milestones achieved, to ultimately enable the successful completion of the project in March 2022.

9.1   Equipment and Resources

Contributions to this research project by the Lead Organisation, Central University of Technology, Free State (CUT) included expertise in flood hydrology and provision of human resources, overall project leadership, project management and administration, office accommodation, office equipment, security, computing facilities, software and library services.

However, this study was a desktop study and no additional equipment has been included in the original budget. The researchers that worked on the project developed capacity in flood hydrology. The human resources that worked on the project are listed in Table 9.1.

Table 9.1:   Project human resources

Name (Institution)	Role	Qualification [studying towards]
Prof OJ Gericke (CUT)	<u>Project Leader</u>: Project coordination, data analysis, catchment response time estimation, report writing, and student supervision.	PhD (Eng.) Pr Eng. and IntPE (SA)
Mr JPJ Pietersen (CUT)	<u>Principal Researcher</u>: Data analysis, ARF estimation, regionalisation, report writing, and post-graduate studies.	M Tech Eng. (Civil) [D Eng. (Civil)] Pr Tech Eng. Envisaged graduation date (Sept. 2022)
Prof JC Smithers (UKZN)	<u>Advisor</u>: Expertise in flood hydrology, quality control, editorial input and student co-supervision.	PhD (Eng.) Pr Eng.
Prof JA du Plessis (SU)	<u>Advisor</u>: Expertise in flood hydrology, quality control, and editorial input.	PhD Pr Eng.
Mr CE Allnutt (CUT)	<u>Post-graduate M Eng. student</u>: Time parameter proportionality ratios.	M Eng. (Civil) Graduated in 2019
Mr VH Williams (CUT)	<u>Post-graduate M Eng. student</u>: Extension of stage-discharge relationships.	B Tech Eng. (Civil) [M Eng.] Envisaged graduation date (Sept. 2022)

9.2 Deliverables

Five deliverables have been identified for the project and are summarised in Table 9.2.

Table 9.2: Project deliverables

Number	Title	Description
1	Advance Payment	20% Advance payment of Year 1.
2	Inception Report	Inception report including the literature review and proposed methodology.
3	Interim Report	Report on progress made and inclusive of the: (i) establishment of national catchment variable database, and (ii) extraction and analysis of rainfall and runoff data.
4	Annual Report	Report on progress made and inclusive of: (i) extraction and analysis of complete flood hydrographs in 51 catchments in Region X to extract time parameters, (ii) regionalisation of homogeneous rainfall regions for ARFs, (iii) the derivation of regional empirical ARF equations, and (iv) comparison and verification of derived empirical ARF equations against ARF methods currently used in South Africa.
5	Final Report	Final report (20% retention) detailing the study findings.

It is envisaged that the deliverables listed in Table 9.2, would ultimately contribute towards:

(a) **WRC Knowledge Tree:** The South African Committee on Large Dams (SANCOLD) has initiated the National Flood Studies Programme (NFSP) to update and modernise methods used for design flood estimation in South Africa.

(b) **Transformation and redress:** One coloured male (Mr Williams; M Eng.), and two white males (Mr Pietersen, D Eng. and Mr Allnutt, M Eng.) were recruited as post-graduate students to work on the project. Staff capacity was also built by working on the project and by supervising students.

(c) **Sustainable development solutions:** This project will contribute to the updating and modernising of methods for ARF and catchment response time estimation in South Africa and will have a direct influence on estimated peak discharges; hence, it will result in the improved design of hydraulic structures in applying the regionalised empirical ARF and time parameter equations. In other words, by improving the current design flood estimation status quo, the risks involved with the design of hydraulic structures will be reduced; reducing not only the financial and sustainability risks, but also the risk associated with the loss of life.

(d) **Informing policy and decision-making:** The most prominent current design flood estimation guideline in South Africa is presented in the Drainage Manual (SANRAL, 2013). By providing improved ARF and catchment response time estimation methodologies, the design flood estimation methods in the Manual can be updated to include the methodologies.

(e) **Human capital development in the Water and Science sectors:** All members of the project team learned new skills and gained knowledge by working on the project. In addition, the project enabled/will enable the graduation of three post-graduate students.

(f) **New products and services for economic development:** Improved regionalisation of ARFs and time parameters and the development of a refined approach to estimate ARFs and catchment response time for South Africa will contribute to the updating of design flood estimation methods in South Africa.

9.3 Knowledge Dissemination

The methodology and results were/will be presented at relevant national and/or international conferences. In addition, journal papers were/will be published in peer-reviewed, accredited journals (e.g. Water SA, SAICE, Hydrological Sciences, Flood Risk Management, Journal of Hydrology, etc.). To date, the following conference proceeding(s) and journal paper(s), either directly or indirectly linked to this project, were published:

Conference Proceeding:

- Pietersen, JPJ and Gericke, OJ. 2019. Areal reduction factors for design rainfall estimation in the Modder-Riet River Basin, South Africa. Proceedings, *10th River Basin Management Conference, WIT Transactions on Ecology and the Environment*, WIT Press (234): 31-40. Alicante, Spain. DOI: 10.2495/RBM190041.

Journal Papers:

- Allnutt, CE, Gericke, OJ and Pietersen, JPJ. 2020. Estimation of time parameter proportionality ratios in large catchments: case study of the Modder-Riet River Catchment, South Africa. *Journal of Flood Risk Management* 13: e12628. DOI: 10.1111/jfr3.12628.
- Gericke OJ and Pietersen JPJ. 2020. Estimation of areal reduction factors using daily rainfall data and a geographically-centred approach. *Journal of the South African*

Institution of Civil Engineering 62 (4): 20-31. DOI: [10.17159/2309-8775/2020/v62n4a3](10.17159/2309-8775/2020/v62n4a3).

The methodology and results were/will also be presented at short courses (e.g. Flood Hydrology courses offered by the Universities of Stellenbosch and Pretoria, respectively) and will be included in future updates of the SANRAL Drainage Manual. Dissemination will also take place in conjunction with activities arranged by the SANCOLD NFSP.

To date, the following papers were presented at Continuous Professional Development (CPD)-accredited short courses:

- Gericke, OJ. 2020. Estimation of catchment response time in medium to large catchments in South Africa. *Proceedings, Flood Hydrology and Urban Runoff Modelling Course*, 2-5 March 2020, University of Pretoria, RSA.

- Gericke, OJ and Pietersen, JPJ. 2020. Evaluation of the Areal Reduction Factor. *Proceedings, Flood Hydrology and Urban Runoff Modelling Course*, 2-5 March 2020, University of Pretoria, RSA.

- Pietersen, JPJ. 2021. Areal Reduction Factors. *Proceedings, 9th Flood Hydrology Course*. Stellenbosch University, Stellenbosch, RSA.

The project's dissemination and uptake objectives are for professional practitioners involved in design flood estimation to utilise the methodologies developed as a standard and preferred design practice. This will be achieved once the methodologies are published in accredited, peer-reviewed journals and endorsed by DWS, SANCOLD and SANRAL.

The target audience/stakeholders that would benefit from the research is professional practitioners involved in design flood estimation. In addition, engagement with other partners, e.g. conference organisers, publishers, and organisers of flood related courses, will ensure that the project's dissemination and uptake objectives are met.

In terms of capacity development and post-graduate supervision, the following students graduated/will be graduating from this project:

Completed research:

- Allnutt, CE. 2019. *Time Parameter Proportionality Ratios in Large Catchments: Case study in the Modder-Riet River catchments, South Africa*. Unpublished M Eng. dissertation, Department of Civil Engineering, Central University of Technology, Free State, Bloemfontein, RSA.

Current research (envisaged to graduate in September 2022):

- **D Eng. (Civil), CUT:** Pietersen, JPJ. *Development and Assessment of Regionalised Areal Reduction Factors for Catchment Design Rainfall Estimation in South Africa.*
- **M Eng. (Civil), CUT:** Williams, VH. *Assessment of Indirect Estimation Methods to extend Observed Stage-discharge Relationships for Above-structure-limit Conditions at Flow-gauging Weirs.*

9.4 Project Concerns

Several project concerns were identified and are either directly or indirectly related to the worldwide COVID-19 pandemic. The concerns identified were included in Chapters 4 and 5, respectively.

9.5 Work Plan and Milestones

A summary of the project milestones and work plan (target dates) is listed in Table 9.3.

Table 9.3: Work plan and milestones

Title	Milestones	Target date
Advance Payment	• Twenty percent (20%) of the annual payment has been claimed from the WRC to initiate the project.	2019/04/01
Inception Report	• Literature review. • Proposed methodologies.	2019/07/01
Interim Report	• Establishment of national catchment variable database. • Extraction and analysis of rainfall and runoff data.	2020/10/01
Annual Report	• Regionalisation of homogeneous rainfall regions. • Derivation of regional empirical ARF and time parameter equations. • Assessment and verification of the regional empirical equations.	2021/12/01
Final Report	• Final report (20% retention) detailing the study findings.	2022/02/28

CHAPTER 10: 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 Managment: Handbook for Prationers 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.

Allnutt, CE. 2019. *Estimation of Time Parameter Proportionality Ratios in Large Catchments: Case Study of the Modder-Riet River Catchment, South Africa*. Unpublished M Eng. dissertation, Department of Civil Engineering, Central University of Technology, Free State, Bloemfontein, RSA.

Allnutt, CE, Gericke, OJ and Pietersen, JPJ. 2020. Estimation of time parameter proportionality ratios in large catchments: case study of the Modder-Riet River Catchment, South Africa. *Journal of Flood Risk Management* 13: e12628. DOI: 10.1111/jfr3.12628.

Amiri, BJ, Gao, J, Fohrer, N, Adamowski, J and Huang, J. 2019a. Examining lag time using the landscape, pedoscape and lithoscape metrics of catchments. *Ecological Indicators* 105 (2019): 36-46. DOI: 10.1016/ecolind.2019.03.050.

Amiri, BJ, Gao, J, Fohrer, N, Adamowski, J and Huang, J. 2019a. Regionalizing time of concentration using landscape structural patterns of catchments. *Journal of Hydrology and Hydromechanics* 67 (2019): 135-142. DOI: 10.2478/johh-2018-0041.

Arnold, JG, Allen, PM, Muttiah, R and Bernhardt, G. 1995. Automated base flow separation and recession analysis techniques. *Ground Water* 33(6): 1010-1018.

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. 2016. *Australian Rainfall and Runoff: A Guide to Flood Estimation*. Institution of Engineers, Australia.

Balme, M, Vischel, T, Level, T, Peugeot, C, and Galle, S. 2006. Assessing the water balance in the Sahel: impact of small scale rainfall variability on runoff. Part 1: Rainfall variability analysis. *Journal of Hydrology* 331 (1-2): 336-348. DOI: 10.1016/j.jhydrol.2006.05.020.

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.

Bell, FC and Kar, SO. 1969. Characteristic response times in design flood estimation. *Journal of Hydrology* 8: 173-196. DOI: 10.1016/0022-1694(69)90120-6.

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

Bondelid, TR, McCuen, RH, and Jackson, TJ. 1982. Sensitivity of SCS models to curve number variation. *Water Resources Bulletin* 20 (2): 337-349.

Chapman, T. 1999. A comparison of algorithms for streamflow recession and baseflow separation. *Hydrological Processes* 13: 701-714.

CCP. 1955. *Culvert Design*. 2nd ed. California Culvert Practice, Department of Public Works, Division of Highways, Sacramento, USA.

Chow, VT, Maidment, DR, and Mays, LW. 1988. *Applied Hydrology*. McGraw-Hill, New York, USA.

Clark, CO. 1945. Storage and the unit hydrograph. *Transactions, American Society of Civil Engineers* 110: 1419-1446.

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.

DAWS. 1986. Estimation of runoff. In: eds. La G Matthee, JF et al., *National Soil Conservation Manual,* Ch. 5, 5.1-5.5.4. Department of Agriculture and Water Supply, Pretoria, RSA.

De Almeida, IK, Almeida, AK and Steffen, JL. 2016. Model for estimating the time of concentration in watersheds. *Water Resources Management* 30: 4083. DOI: 10.1007/s11269-016-1383-x.

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.

Dingman, SL. 2002. *Physical Hydrology*. 2nd ed. Macmillan Press Limited, New York, USA.

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.

Dymond, JR and Christian, R. 1982. Accuracy of discharge determined from a rating curve. *Hydrological Sciences Journal* 27(4): 493-504. DOI: 10.1080/02626668209491128.

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.

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.

Fang, X, Thompson, DB, Cleveland, TG, Pradhan, P, and Malla, R. 2008. Time of concentration estimated using watershed parameters by automated and manual methods. *Journal of Irrigation and Drainage Engineering* 134 (2): 202-211. DOI: 10.1061/(ASCE)0733-9437(2008)134:2(202).

Folmar, ND and Miller, AC. 2008. Development of an empirical lag time equation. *Journal of Irrigation and Drainage Engineering* 134 (4): 501-506. DOI: 10.1061/(ASCE)0733-9437(2008)134:4(501).

Gericke, OJ. 2018. Catchment response time and design rainfall: The key input parameters for design flood estimation in ungauged catchments. *Journal of the South African Institution of Civil Engineering* 60 (4): 51-67. DOI: 10.1759/2309-8775/2018/v60n4a6.

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.

Gericke, OJ and Du Plessis, JA. 2013. Development of a customised design flood estimation tool to estimate floods in gauged and ungauged catchments. *Water SA* 39 (1): 67-94. DOI: 10.4314/wsa.v39i1.9.

Gericke, OJ and Smithers, JC. 2014. Review of methods used to estimate catchment response time for the purpose of peak discharge estimation. *Hydrological Sciences Journal* 59 (11): 1935-1971. DOI: 10.1080/02626667.2013.866712.

Gericke, OJ and Smithers, JC. 2016a. Are estimates of catchment response time inconsistent as used in current flood hydrology practice in South Africa? *Journal of the South African Institution of Civil Engineering* 58 (1): 2-15. DOI: 10.17159/2309-8775/2016/v58n1a1.

Gericke, OJ and Smithers, JC. 2016b. Derivation and verification of empirical catchment response time equations for medium to large catchments in South Africa. *Hydrological Processes* 30 (23): 4384-4404. DOI: 10.1002/hyp.10922.

Gericke, OJ and Smithers, JC. 2017. Direct estimation of catchment response time parameters in medium to large catchments using observed streamflow data. *Hydrological Processes* 3 (5): 1125-1143. DOI: 10.1002/hyp.11102.

Gericke, OJ and Smithers, JC. 2018. An improved and consistent approach to estimate catchment response time: Case study in the C5 drainage region, South Africa. *Journal of Flood Risk Management* 11 (2018): S284-S301. DOI: 10.1111/jfr3.12206.

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.

Hardy, T, Panja, P and Mathias, P. 2005. *WinXSPRO: A Channel Cross-Section Analyzer*. United States Department of Agriculture, Forest Service, Rocky Mountains, USA.

Hathaway, GA. 1945. Design of drainage facilities. *Transactions of American Society of Civil Engineers* 30: 697-730.

Heggen, R. 2003. Time of concentration, lag time and time to peak. [Internet]. In: eds. Shrestha, B and Rajbhandari, R. In: *Proceedings, Regional Training Course: Application of Geo-informatics for Water Resources Management*, 3.1-3.23, [online]. International Centre for Integrated Mountain Development, Kathmandu, Nepal. Available: http://www.hkh-friend.net.np/rhdc/training/lectures/HEGGEN/Tc3.pdf. [Accessed: 30 September 2010].

Henderson, FM and Wooding, RA. 1964. Overland flow and groundwater flow from a steady rainfall of finite duration. *Journal of Geophysical Research* 69 (8): 129-146. DOI: 10.1029/JZ069i008p01531.

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.

Hood, MJ, Clausen, JC, and Warner, GS. 2007. Comparison of stormwater lag times for low impact and traditional residential development. *Journal of the American Water Resources Association* 43 (4): 1036-1046. DOI: 10.1111/j.1752-1688.2007.00085.x.

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

HRU. 1972. *Design Flood Determination in South Africa*. HRU Report No. 1/72. Hydrological Research Unit, University of the Witwatersrand, Johannesburg, RSA.

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.

Hughes, DA, Hannart, P, and Watkins, D. 2003. Continuous baseflow from time series of daily and monthly streamflow data. *Water SA* 29 (1): 43-48.

IEA. 1977. *Australian Rainfall and Runoff: Flood Analysis and Design*. 2nd ed. Institution of Engineers, Canberra, Australia.

IH. 1999. *Flood Estimation Handbook.* Vol. 4, Institute of Hydrology, Wallingford, Oxfordshire, UK.

Jena, SK and Tiwari, KN. 2006. Modelling synthetic unit hydrograph parameters with geomorphologic parameters of watersheds. *Journal of Hydrology* 319: 1-14. DOI: 10.1016/j.AQ10 jhydrol.2005.03.025.

Johnstone, D and Cross, WP. 1949. *Elements of Applied Hydrology*. Ronald Press, New York, USA.

Kerby, WS. 1959. Time of concentration for overland flow. *Civil Engineering* 29 (3): 174.

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.

Kirpich, ZP. 1940. Time of concentration of small agricultural watersheds. *Civil Engineering* 10 (6): 362.

Kisters. 2019. *Hydstra Software, Version* 11, [online]. Kisters Pty Ltd., Weston, ACT, Australia. Available from: http://www.kisters.com.au [15 July 2019].

Kjeldsen, TR. 2007. *The Revitalised FSR/FEH Rainfall-Runoff Method.* FEH Supplementary Report No. l. Centre for Ecology and Hydrology, Wallingford, Oxfordshire, UK.

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.

Lang, M, Pobanz, K, Renard, B, Renouf, E and Sauquet, E. 2010. Extension of rating curves by hydraulic modelling, with application to flood frequency analysis. *Hydrological Sciences Journal* 55(6): 883-898. DOI: 10.1080/02626667.2010.504186.

Larson, CL. 1965. *A Two Phase Approach to the Prediction of Peak Rates and Frequencies of Runoff for Small Ungauged Watersheds.* Technical Report No. 53. Department of Civil Engineering, Stanford University, Stanford, USA.

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.

Li, M and Chibber, P. 2008. Overland flow time on very flat terrains. *Journal of the Transportation Research Board* 2060: 133-140. DOI: 10.3141/2060-15.

Lim, KJ, Engel, BA, Tang, Z, Choi, J, Kim, K, Muthukrishnan, S and Tripathy, D. 2005. Automated web GIS-based hydrograph analysis tool (WHAT). *Journal of the American Water Resource Association* 41(6): 1407-1416.

Linsley, RK, Kohler, MA, and Paulhus, JLH. 1988. *Hydrology for Engineers*. SI Metric ed. McGraw-Hill, Singapore.

Lorenz, C and Kunstmann, H. 2012. The hydrological cycle in three state-of-the-art reanalyses: intercomparison and performance analysis. *Journal of Hydrometeorology* 13: 1397-1420. DOI: 10.1175/JHM-D-11-088.1.

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.

Lyne, V and Hollick, M. 1979. Stochastic time-variable rainfall-runoff modelling.In: *Proceedings, Hydrology and Water Resources Symposium*, Institution of Engineers, Perth, Australia. pp. 89-92.

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.

McCuen, RH. 2005. *Hydrologic Analysis and Design*. 3rd ed. Prentice-Hall, Upper Saddle River, New York, USA.

McCuen, RH. 2009. Uncertainty analyses of watershed time parameters. *Journal of Hydrologic Engineering* 14 (5): 490-498. DOI: 10.1061/(ASCE)HE.1943-5584.0000011.

McCuen, RH and Spiess, JM. 1995. Assessment of kinematic wave time of concentration. *Journal of Hydraulic Engineering* 121 (3): 256-266.

McCuen, RH, Wong, SL, and Rawls, WJ. 1984. Estimating urban time of concentration. *Journal of Hydraulic Engineering* 110 (7): 887-904.

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, South Africa.

Menné, TC .1960. *The Role of Hydrology in Planning the Development of a Country.* Hydrological Research, Technical Report No. 9, Department of Water Affairs, Pretoria, RSA.

Michailidi, EM, Antoniadi, S, Koukouvinos, A, Bacchi, B and Efstratiadis, A. 2018. Timing the time of concentration: shedding light on a paradox. *Hydrological Sciences Journal* 63 (5): 721-740. DOI: 10.1080/02626667.2018.1450985.

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.

Mimikou, M. 1984. Regional relationships between basin size and runoff characteristics. *Hydrological Sciences Journal* 29 (1, 3): 63-73. DOI: 10.1080/02626668409490922.

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.

Mockus, V. 1957. *Use of Storm and Watershed Characteristics in Synthetic Hydrograph Analysis and Application.* United States Department of Agriculture, Soil Conservation Service, Washington, DC, USA.

Morgali, JR and Linsley, RK. 1965. Computer simulation of overland flow. *Journal of Hydraulics Division, ASCE* 91 (HY3): 81-100.

Nathan, RJ and McMahon, TA. 1990. Evaluation of automated techniques for baseflow and recession analyses. *Water Resources Research* 26 (7): 1465-1473.

Nathanael, JJ, Rademeyer, PF and Van der Spuy, D. 2018. *Discharge Rating Improvement of Rating Table and High Flow Data: Sand River at Exeter*. Report No. X3H008DT-2018.12. Department of Water and Sanitation, Pretoria, RSA.

Neitsch, SL, Arnold, JG, Kiniry, JR, and Williams, JR. 2005. *Soil and Water Assessment Tool: Theoretical Documentation*. Agricultural Research Service and Blackland Research Center, Temple, Texas, USA.

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.

Pavlovic, SB and Moglen, GE. 2008. Discretization issues in travel time calculations. *Journal of Hydrologic Engineering* 13 (2): 71-79. DOI: 10.1061/(ASCE)1084-0699(2008)13:2(71).

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 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.

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

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.

Petersen-Øverleir, A and Reitan, T. 2009. Accounting for rating curve imprecision in flood frequency analysis using likelihood-based methods. *Journal of Hydrology* 366 (2009):89-100.

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.

Piggott, AR, Moin, S and Southam, C. 2005. A revised approach to the UKIH method for the calculation of baseflow. *Hydrological Sciences Journal* 50: 911-920.

Pitman, WV. 2011. Overview of water resource assessment in South Africa: Current state and future challenges. *WaterSA* 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.

Pullen, RA. 1969. *Synthetic Unitgraphs for South Africa*. HRU Report No. 3/69. Hydrological Research Unit, University of the Witwatersrand, Johannesburg, RSA.

Rahman, A, Weinmann, PE, Hoang, TMT and Laurenson, EM. 2002. Monte Carlo simulation of flood frequency curves from rainfall. *Journal of Hydrology* 256: 196-210.

Ramsbottom, DM and Whitlow, CD. 2003. *Extension of Rating Curves at Gauging Stations Best Practice Guidance Manual.* Environmental Agency, Bristol, UK.

Ramser, CE. 1927. Runoff from small agricultural areas. *Journal of Agricultural Engineering* 34 (9): 797-823.

Reich, BM. 1962. Soil Conservation Service design hydrographs. *Transactions of South African Institute of Civil Engineers* 4.

Royappen, M, Dye, PJ, Schulze, RE, and Gush, MB. 2002. *An Analysis of Catchment Attributes and Hydrological Response Characteristics in a Range of Small Catchments.* WRC Report No. 1193/01/02. Water Research Commission, Pretoria, RSA.

Rutledge, AT. 1998. *Computer programs for describing the recession of ground-water discharge and for estimating mean ground-water recharge and discharge from streamflow data*. Report No. 984148: 43, USA: United States Geological Survey Water Resources Investigations.

Sabol, GV. 2008. *Hydrologic Basin Response Parameter Estimation Guidelines*. Dam Safety Report, State of Colorado. Tierra Grande International Incorporated, Scottsdale, AZ, USA.

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

Schmidt, EJ and Schulze, RE. 1984. *Improved Estimation of Peak Flow Rates using Modified SCS Lag Equations*. ACRU Report No. 17. University of Natal, Department of Agricultural Engineering, Pietermaritzburg, RSA.

Schultz, GA. 1964. *Studies in Flood Hydrograph Synthesisation*. Unpublished MSc dissertation, University of the Witwatersrand, Johannesburg, RSA.

Schulz, EF and Lopez, OG. 1974. *Determination of Urban Watershed Response Time*. Colorado State University, Fort Collins, Colorado, USA.

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, 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.

Seybert, TA. 2006. *Stormwater Management for Land Development: Methods and Calculations for Quantity Control.* John Wiley and Sons Incorporated, Hoboken, New Jersey, USA.

Shao, Q, Dutta, D, Karim, F and Petheram, C. 2018. A method for extending stage-discharge relationships using a hydrodynamic model and quantifying the associated uncertainty. *Journal of Hydrology* 556(2018):154-172.

Sheridan, JM. 1994. Hydrograph time parameters for flatland watersheds. *Transactions of American Society of Agricultural Engineers* 37 (1): 103-113. DOI: 10.13031/2013.28059.

Simas, MJC. 1996. *Lag Time Characteristics in Small Watersheds in the United States.* Unpublished PhD thesis, School of Renewable Resources, University of Arizona, Tucson, USA.

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.

Sloto, RA and Crouse, MY. 1996. *HYSEP: A computer program for streamflow hydrograph analysis*. Report No. 90-4040: 46, USA: United States Geological Survey Water Resources Investigations.

Smakhtin, VU. 2001. Low flow hydrology: A review. *Journal of Hydrology* 240 (2001): 147-186.

Smakhtin, VU and Watkins, DA. 1997. *Low Flow Estimation in South Africa*. WRC Report No. 494/1/97. Water Research Commission, Pretoria, RSA.

Smithers, JC. 2012. Review of methods for design flood estimation in South Africa. *Water SA* 38 (4): 633-646.

Smithers, JC, Chetty, KT, Frezghi, MS, Knoessen, DM, and Tewolde, MH. 2013. Development and assessment of a daily time step continuous simulation modelling approach for design flood estimation at ungauged locations: ACRU model and Thukela Catchment case study. *Water SA* 39 (4): 449-458. DOI: 10.4314/wsa.v39i4.4.

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.

Snyder, FF. 1938. Synthetic unit hydrographs. *Transactions of American Geophysical Union* 19: 447.

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.

Su, DH and Fang, X. 2004. Estimating travelling time of flat terrain by 2-dimensional overland flow model. In: eds. Jirka, GH and Uijttewaal, WSJ, *Shallow Flows*, 629-635. Balkema, Rotterdam, Netherland.

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

Thomas, WO, Monde, MC, and Davis, SR. 2000. Estimation of time of concentration for Maryland streams. *Journal of the Transportation Research Board* 1720: 95-99.

USACE. 2001. *HEC-HMS Hydrologic Modelling System: User's Manual, Version* 2.2.1. United States Army Corps of Engineers, Vicksburg, Mississippi, USA.

USBR. 1973. *Design of Small Dams*. 2nd ed. United States Bureau of Reclamation, Water Resources Technical Publication, Washington, DC, USA.

USDA NRCS. 2010. Time of concentration. In: eds. Woodward, DE *et al*., *National Engineering Handbook*, Ch. 15 (Section 4, Part 630), 118. United States Department of Agriculture Natural Resources Conservation Service, Washington, DC, USA.

USDA SCS. 1985. Hydrology. In: eds. Kent, KM *et al.*, *National Engineering Handbook*, Ch. 16 (Section 4, Part 630), 1-23. United States Department of Agriculture Soil Conservation Service, Washington, DC, USA.

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.

Viessman, W and Lewis, GL. 1996. *Introduction to Hydrology*. 4th ed. Harper Collins College Publishers Incorporated, New York, USA.

Watt, WE and Chow, KCA. 1985. A general expression for basin lag time. *Canadian Journal of Civil Engineering* 12: 294-300.

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

Wessels, P and Rooseboom, A .2009. Flow-gauging structures in South African rivers. Part 1: An Overview. *Water SA* 35 (1) 1-10.

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

White, K and Sloto, PA. 1990. *Base flow frequency characteristics of selected Pennsylvanian streams*. Report No. 90-4160: 66, USA: United States Geological Survey Water Resources Investigations.

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

Wong, TSW and Chen, CN. 1997. Time of concentration formula for sheet flow of varying flow regime. *Journal of Hydrologic Engineering* 2 (3): 136-139. DOI: 10.1061/(ASCE)1084-0699(1997)2:3(136).

Woolhiser, DA and Liggett, JA. 1967. Unsteady one-dimensional flow over a plane: The rising hydrograph. *Water Resources Research* 3 (3): 753-771. DOI: 10.1029/WR003i003p00753.

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