DEVELOPMENT AND ASSESSMENT OF REGIONALISED APPROACHES TO DESIGN FLOOD ESTIMATION IN SOUTH AFRICA

JP CALITZ

Submitted in fulfilment of the requirements
for the degree of PhD

Bioresources Engineering
School of Engineering
College of Agriculture, Engineering and Science
University of KwaZulu-Natal
Pietermaritzburg
December 2020

PREFACE

The work described in this dissertation was carried out in the Centre for Water Resources Research, School of Agriculture, Earth and Environmental Sciences, University of KwaZulu-Natal, Pietermaritzburg, under the supervision of Professor JC Smithers, Professor TR Kjeldsen and Professor OJ Gericke.

The research represents original work by the author and has not otherwise been submitted in any form for any degree or diploma to any tertiary institution. Where, use has been made of the work of others, it is duly acknowledged in the text.

Signed:
Professor JC Smithers
Date:

Signed:
Professor TR Kjeldsen
Date:

Signed:
Professor OJ Gericke
Date:

DECLARATION 1: PLAGIARISM

I Johannes Pieter Calitz declare that:

(i) The research reported in this thesis, except where otherwise indicated, is my original work.

(ii) This thesis has not been submitted for any degree or examination at any other university.

(iii) This thesis does not contain other persons' data, pictures, graphs, or other information, unless specifically acknowledged as being sourced from other persons.

(iv) This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

- (a) their words have been re-written but the general information attributed to them has been referenced; and
- (b) where their exact words have been used, their writing has been placed inside quotation marks, and referenced.

(v) Where I have reproduced a publication of which I am an author, co-author, or editor, I have indicated in detail which part of the publication was actually written by myself alone and have fully referenced such publications.

(vi) This thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

Signed: [redacted]

Mr JP Calitz

Date: 2021/06/24

DECLARATION 2: PUBLICATIONS

Some aspects of this PhD study are included in a final report submitted to the Water Research Commission (WRC) for WRC Project K5-2748 and constitutes my own work in its entirety. My co-author, Prof JC Smithers was the project leader and provided guidance and reviewed the report prior to submission. Considerable additions and changes have been made subsequent to the final report that was submitted to the WRC and are included in this thesis. The reference for the report submitted to the WRC is:

Calitz, JP and Smithers, JC. 2020. *Development and Assessment of Regionalised Approaches to Design Flood Estimation in South Africa*. Report No. K5-2748. WRC, Pretoria, RSA.

This PhD study also extends on my MSc (Calitz, 2016) study which focused on the development of a probabilistic rational method within three of the primary drainage regions as defined by the South African Department of Water and Sanitation (DWS). The MSc only focussed on the development of a new Probabilistic Rational Method (PRM) for South Africa, as opposed to comparing various Regional Flood Frequency Analysis (RFFA) models. University approval was obtained prior to the award of the MSc degree to convert and extend the study MSc study into a PhD study. In this process, the scope of the PhD study expanded considerably from focussing only on the PRM to other regional approaches.

In addition to the above the following chapters have been prepared as draft papers to be submitted for publication:

Chapter 3

Calitz, JP, Smithers, JC, Kjeldsen, TR and Gericke, OJ. 2021. *Identification of a distribution suitable for at-site flood frequency analysis in South Africa*. Draft paper

In this paper I undertook, what is to my knowledge, the most detailed investigation into the selection of a suitable distribution for use in design flood estimation in South Africa. I prepared the data, designed the selection process and wrote the paper. My co-authors, Prof JC Smithers, Dr TR Kjeldsen and Prof OJ Gericke reviewed the paper and provided guidance.

Chapter 4

Calitz, JP, Smithers, JC, Kjeldsen, TR and Gericke, OJ. 2021. *Formation of hydrologically similar pooling groups for use in design flood estimation in South Africa*. Draft paper

In this paper I undertook regionalisation of the South African streamflow data for design flood purposes. I applied multi-variate regionalisation techniques to assess which method identifies the highest percentage of statistically homogeneous flood producing regions. I collated the required data, undertook the regionalisation studies and wrote the paper. My co-authors, Prof JC Smithers, Dr TR Kjeldsen and Prof OJ Gericke reviewed the paper and provided guidance.

Chapter 5

Calitz, JP, Smithers, JC, Kjeldsen, TR and Gericke, OJ. 2021. *Development and assessment of regional models for design flood estimation in South Africa*. Draft paper

Utilising the results from the previous papers I developed four regional flood models for South Africa. I also compared the results achieved by the flood models to identify the best performing flood model for use in South Africa. I also wrote the paper. My co-authors, Prof JC Smithers, Dr TR Kjeldsen and Prof OJ Gericke reviewed the paper and provided guidance.

ABSTRACT

Engineers rely on design hydrological information for the design of hydraulic structures, such as dams, bridges, and drainage culverts. No single Design Flood Estimation (DFE) method has been identified internationally as the most appropriate method to use and, in many texts and manuals, the use of a combination of these are recommended. In South Africa, some of the currently recommended and widely used methods were developed outside of South Africa with little or no local adaptation or assessment, and most of the recommended methods were developed prior to 1990. The development of new and updated methods can therefore benefit from the use of much longer observed data sets and new and innovative approaches applied internationally. Four Regional Flood Frequency Analysis (RFFA) approaches widely adopted internationally are direct quantile estimation methods, Probabilistic Rational Method (PRM), Index Flood (IF), and Regional Growth Curve (RGC) methods. The Standard Design Flood (SDF) method is a locally developed PRM. However, the method has been recommended for review in a number of studies, and the IF has been shown to have potential for implementation at a national scale in South Africa.

The aim of this study was to develop and assess RFFA approaches for the estimation of design flood quantiles within South Africa utilising the currently available data. This process required the compilation of a hydrological descriptors database, including quality controlled gauged flow data. This data was then utilised to identify a suitable probability distribution for FFA in South Africa, which can be applied at a regional scale through the identification of homogeneous flood producing regions and regional flood models.

DFE methods require a range of catchment descriptors to be determined for use in models. Considering the literature reviewed and the available datasets, 17 catchment descriptors were selected for inclusion in the study. The descriptors range from geographic and catchment descriptors to design rainfall quantiles. After data screening, a total of 383 stations were utilised, in the study. The available record lengths and number of gauges were compared to prominent studies undertaken previously and was found to be comparative to the data availability in Australia and the United Kingdom.

Linear moments (LM) were adopted for the estimation of the distribution parameters. Five distributions were selected for evaluation based on local recommendations as well as recent international developments: (i) General Extreme Value (GEV), (ii) Generalised Pareto (GPA),

(iii) 3-parameter Kappa (KAP3), (iv) Log Pearson Type III (LP3) and (v) Pearson Type III (PE3). The evaluation process relied on an iterative elimination approach, reviewing graphical fits to theoretical distributions, Goodness-of-fit (GoF) criteria, model fit criteria and model uncertainty to identify the most suitable distribution. The graphical fit favoured the GPA, KAP3 and LP3 distributions equally, with the GoF methods ranking LP3 as the most suitable method. Conversely, the GPA was ranked highest for the model fit criterion and displayed the least model uncertainty and is thus recommended as the most suitable distribution for general FFA in South Africa.

Two regionalisation approaches were considered to undertake the formation of the pooling groups, i.e. Clustering, and Region of Influence (RoI). For each regionalisation approach the hydrological descriptors were grouped into parameter sets, that constituted all potential descriptor combinations, which were tested for homogeneity as a selection criterion. Using the RoI approach, a maximum of 51% of the regions identified were relatively homogeneous. The super region approach was also applied to identify five dominant regions within which the RoI was applied in an attempt to refine the RoI approach. Using the combination of super regions and RoI provided little additional benefit, increasing the percentage of relatively homogeneous regions identified to only 52.6%. Conversely, the Clustering approach was able to identify 42 relatively homogeneous clusters in South Africa.

To assess the suitability of Quantile Regression Technique (QRT) and Parameter Regression Technique (PRT) models in South Africa, four models were developed: (i) a QRT model, (ii) IF with equal station weighting (IF1), (iii) IF with station weighting applied (IF2) and (iv) PRM. Regression models were developed at two scales to estimate the required Scaling Factors, i.e. national and regional, with regional models performing best based on the Nash-Sutcliffe model Efficiency (*NSE*) coefficient.

Six key performance indicators were utilised to assess the quantile estimation of the developed models: (i) *NSE*, (ii) Relative Error (*RE*), (iii) Root Mean Square Error (*RMSE*), (iv) Relative *RMSE* (*RMSEr*), (v) *BIAS*, and (iii) *BIASr*. The models that performed best in the *RE* assessment were the IF1 for both regionalisation schemes and the IF2 and PRM models using the RoI. When comparing the *BIAS* and *RMSE* of the four best performing clustering and RoI based models, the IF1 and QRT using Clustering models are the dominant models when considering both the *RMSEr* and the *BIASr*, the models improved on the results of the remaining models by up to a factor of two.

The IF1 and QRT using Clustering models are therefore the best performing models on a national scale. The IF1 however has the added advantage of being able to estimate the entire growth curve as to the predefined QRT models. The IF1 is therefore the recommended model at a national scale, however cognisance needs to be taken when applying the model on the eastern coast due to poor *BIASr* performance.

The new knowledge generated by the study can be divided into data, in the form of potentially the largest database of design flood specific descriptors concentrating on South Africa, and theoretical applications thereof. The theoretical knowledge generated ranges from the investigation into the most suitable frequency distribution to use for FFA in South Africa, to the application of multi-variate regionalisation approaches, which have not been applied in South Africa before. However, one of the key contributions was the development and performance assessment of four DFE models at multiple scales for South Africa for the estimation of peak design flood values.

EXTENDED ABSTRACT

Engineers rely on design hydrological information for the design of hydraulic structures, such as dams, bridges, and drainage culverts (Smithers and Schulze, 2003). This information is often estimated at ungauged sites using models to estimate flood frequencies (Schulze *et al.*, 2004, Smithers *et al.*, 2015). No single Design Flood Estimation (DFE) method has been identified internationally as the most appropriate method to use and, in many texts and manuals, the use of a combination of these are recommended (e.g. Pilgrim and Cordery, 1993, Alexander, 2002b, Chadwick *et al.*, 2004, SANRAL, 2013). In South Africa, some of the currently recommended and widely used methods were developed outside of South Africa with little or no local assessment, and most of the recommended methods were developed prior to 1990. The development of new and updated methods can therefore benefit from the use of much longer observed data sets and new and innovative approaches used internationally.

Internationally there has been a general shift from conventional at-site Flood Frequency Analysis (FFA) to Regional FFA (RFFA), which utilises regional knowledge to supplement temporal knowledge. The RFFA models generally fall into one of two categories, Quantile Regression Technique (QRT), or Parameter Regression Technique (PRT). QRT directly estimates the peak flows of the required Annual Exceedance Probability (AEP), whereas PRT methods estimate descriptive statistics of the regional growth curves to estimate the AEP events.

Four RFFA approaches widely adopted internationally are direct quantile estimation methods, such as the Regional Maximum Flood (Kovács, 1988), Probabilistic Rational Method (PRM), Index Flood (IF), and Regional Growth Curve (RGC) methods. The Standard Design Flood (SDF) method developed by Alexander (2002a) is a locally developed PRM. However, the method has been recommended for review in a number of studies (Görgens, 2002, Smithers and Schulze, 2003, Van Bladeren, 2005, Gericke, 2010, Van Vuuren *et al.*, 2013). Kjeldsen *et al.* (2002) applied the IF in the KwaZulu-Natal (KZN) province in South Africa and showed the potential for implementation at a national scale. Haile (2011) applied the IF approach on a national scale in South Africa but utilised a limited dataset. Nathanael *et al.* (2018) reviewed the national scale RFFA models developed for South Africa and found that none of the current models were satisfactory. The IF approach is also favoured over the PRM approach

internationally as is evident from the recent exclusion of the PRM in the revised Australian Rainfall and Runoff (ARR) guidelines (Rahman *et al.*, 2015b).

Aims and Objectives of Research

The aim of this study was to develop and assess RFFA approaches for the estimation of design flood quantiles within South Africa utilising the currently available data, which required the fulfilment of the following objectives:

(a) Compilation of a hydrological descriptors database.

(b) Collation and quality control of selected gauged flow data in South Africa.

(c) Identification of a suitable probability distribution for FFA in South Africa.

(d) Identify and verify homogeneous flood producing regions.

(e) Regional flood model development and performance assessment.

Hydrological Descriptor Database

DFE methods require a range of catchment descriptors to be determined for use in models. Considering the literature reviewed and the available datasets 17 catchment descriptors were selected for inclusion in the study. The descriptors range from geographic and morphological descriptors to design rainfall quantiles. A hydrologically corrected Digital Elevation Model (DEM) was developed using the Shuttle Radar Topography Mission (SRTM) (NASA-JPL, 2013) data, producing a 30 x 30m DEM for the extraction of various topographic catchment descriptors.

Streamflow Data

The Department of Water and Sanitation (DWS) is the custodians of the flow monitoring network in South Africa and currently has 1 458 streamflow gauging stations within South Africa. The data were screened by considering a minimum record length of 20 years, after which, a total of 383 stations remained and were utilised in the study, totalling 18 349 Annual Maximum Series (AMS) records. The stations are divided into 296 river gauges and 87 synthetic dam inflow records generated by the DWS flood studies group (Naidoo, 2019). The available record lengths and number of gauges were compared to prominent studies undertaken previously and was found to be comparative to the data availability in Australia and the United Kingdom.

Identification of a Distribution Suitable for use in South African FFA

SANRAL (2013) and Van der Spuy and Rademeyer (2018) list methods to undertake FFA and provide the most commonly used probability distributions in South Africa, which are generally applied using the traditional Method of Moments (MM) approach. The distribution selected can, however, provide large differences in flood estimates and hence the selection of the most suitable distribution for use in South Africa is required for FFA. Given the prevalence of Linear moments (LM) (Hosking, 1990) and it's reduced sensitivity to outliers, LMs were adopted for the estimation of the distribution parameters. Five distributions were selected for evaluation based on local recommendations as well as recent international developments: (i) General Extreme Value (GEV), (ii) Generalised Pareto (GPA), (iii) 3-parameter Kappa (KAP3), (iv) Log Pearson Type III (LP3) and (v) Pearson Type III (PE3). The evaluation process relied on an iterative elimination approach, reviewing graphical fits to theoretical distributions, Goodness-of-fit (GoF) criteria, model fit criteria and model uncertainty to identify the most suitable distribution. The graphical fit favoured the GPA, KAP3 and LP3 distributions equally, with the GoF methods ranking LP3 as the most suitable method. Conversely, the GPA was ranked highest for the model fit criterion and displayed the least model uncertainty. Given the overall ranking of the distributions, the GPA is thus recommended as the most suitable distribution for FFA in South Africa.

Regionalisation

Two approaches widely used internationally were considered to undertake the formation of the pooling groups, i.e. Clustering, and Region of Influence (RoI). The size of the pooling groups was motivated by the 5T rule (Robson and Reed, 1999), which requires that pooling groups have a combined record length of five times the desired storm recurrence interval. Where data is limited however, the lower recommended value of 2T was enforced in this study, with emphasis placed on the 1% AEP. This ensures that each group has a combined minimum record length of 200 years, fulfilling the 2T requirement for the 1% AEP. For each regionalisation approach the hydrological descriptors were grouped into parameter sets, that constituted all potential descriptor combinations, which were tested for homogeneity as a selection criterion. The homogeneity test applied was the H-test developed by Hosking and Wallis (1993) and utilised the AMS data for the sites considered.

Using the RoI approach, a maximum of 51% of the regions identified using the Latitude, Longitude, Distance from the Coastline (D_c), and mean runoff percentage were relatively homogeneous. The super region approach (Mostofi Zadeh and Burn, 2019) was also applied to identify five dominant regions within which the RoI was applied in an attempt to refine the RoI approach. The number of super regions was identified using elbow plots and refining the regions using T-distributed Stochastic Neighbour Embedding (TSNE) (van der Maaten and Hinton, 2008), Uniform Manifold Approximation and Projection (UMAP) (McInnes *et al.*, 2018) and geographic distribution. Using the combination of super regions and RoI provided little additional benefit, increasing the percentage of relatively homogeneous regions identified to 52.6%. The RoI approach performed particularly poor in super Region 5, in the Western Cape where a maximum of 28.8% of regions were deemed to be relatively homogeneous.

Conversely, the Clustering approach was able to identify 42 relatively homogeneous clusters in South Africa. Initial clustering was performed using the outlet location (Latitude and Longitude) and the distance from the coast to define 36 clusters, 17 of which were deemed relatively homogeneous. Further refinement was required to identify the final 42 relatively homogeneous clusters and consisted of further clustering within large clusters, merging of clusters and/or exclusion of discordant sites. The final number of sites used to define the 42 relatively homogeneous clusters was 331, as the exclusion of 52 discordant sites was required.

Model Development and Assessment

Regional flood model development generally falls within one of two categories, QRT or PRT. QRT models directly estimate the quantile flows in question, e.g. the 1% AEP flood event, whereas PRT relies on regional growth curves or and estimation of model parameters. Numerous model formulations exist for the development of regional flood models, but the formulation most widely used in the literature are the Index Flood and regional growth curve approach. To assess the suitability of QRT and PRT models in South Africa four models were developed: (i) a QRT model, (ii) Index Flood with equal station weighting (IF1), (iii) Index Flood with station weighting applied (IF2) and (iv) Probabilistic Rational Method (PRM) were developed.

The adopted methods required the estimation of Scaling Factors (SF) and the adopted SFs for the IF1 and IF2 approaches was the Mean Annual Flood (*MAF*) and the 10% AEP C-value

(C_{10}) was used for the PRM. Regression models were developed to estimate the required SFs and Q_T and the development was undertaken at two scales, national and regional, based on the clustering and RoI regionalisation schemes. Hence a total of four combinations of development scale and regionalisation scheme was used: (i) Clustering with cluster-based models, (ii) Clustering with a national scale model, (iii) RoI with super region based models, and (iv) RoI with a national scale model.

Eight catchment descriptors were included as potential predictor variables for regression development and these ranged from outlet elevation to design rainfall values. The p-value, which provides an indication of the correlation between the predictor and the response variables, was used for the selection of significant catchment descriptors and the developed models were limited to the use of three predictor variables. Catchment Area and *MAP* were identified as significant catchment descriptors in the SF regressions. The models developed for *MAF* at a regional scale performed well with an Nash-Sutcliffe model Efficiency (*NSE*) coefficient (Nash and Sutcliffe, 1970) value of up to 0.78 achieved. The national scale models, however, performed poorly only achieving maximum *NSE* values of 0.28. The regional models were therefore used to undertake the quantile estimation assessment.

The *SF* and QRT regression models were used to estimate design peak flows at the sites considered and assessed using six key performance indicators: (i) *NSE*, (ii) Relative Error (*RE*), (iii) Root Mean Square Error (*RMSE*), (iv) Relative *RMSE* (*RMSEr*), (v) *BIAS*, and (iii) *BIASr*. The assessment approach consisted of two steps, initially the four best performing regression models were identified using the model accuracy which used the *NSE* and *RE,* followed by the *RMSE* and *BIAS* assessment, which was used to identify the best performing model. The regional models achieved *NSE* values up to 0.77, but tended to underestimation, which needs to be taken into consideration if the models are applied.

The *RE* assessment relied on the ratio bounds developed by Rahman *et al.* (2012) and Naidoo (2020). The models that performed best in the *RE* assessment were the IF1 for both regionalisation schemes and the IF2 and PRM models using the RoI. The four models were able to estimate the peak flows within the Rahman *et al.* (2012) *RE* bounds at between 64.8 and 75.2% of the sites considered, improving on the models assessed by Nathanael *et al.* (2018). The percentages drop to between 53.1 and 65.3% when applying the more stringent *RE* bounds defined by Naidoo (2020). When considering an overall rank for the model accuracy assessment the four top performing models were the IF1 and QRT using clustering and the IF1 and IF2 using RoI.

The final assessment compared the *BIAS* and *RMSE* of the four best performing clustering and RoI based models. The IF1 and QRT using Clustering models are the dominant models when considering both the *RMSEr* and the *BIASr*, the model achieved median *RMSEr* values ranging between 0.52 and 0.61, improving on the results of the remaining models by up to a factor of two. The *BIASr* values for the IF1 and QRT also improve on the results of the remaining models, in particular for AEPs in excess of 10%. The *BIASr* values for the IF1 using clustering, however, performs poorly in clusters 34 and 37 on the east coast of South Africa.

The IF1 and QRT using Clustering models are therefore the best performing models on a national scale. The IF1 however has the added advantage of being able to estimate the entire growth curve as to the predefined QRT models. The IF1 is therefore the recommended model at a national scale, however cognisance needs to be taken when applying the model on the eastern coast due to the poor *BIASr* performance.

Research Contributions

The following items are considered to be the most prominent unique contributions that have been developed as part of this study:

(i) The database of design flood specific descriptors is potentially the largest database concentrating on South Africa and this study thus provides a basis for further development and refinement of models for DFE in South Africa.

(ii) To the knowledge of the author, it is the first study to perform a detailed investigation into the most suitable probability distribution to use for FFA in South Africa and to recommend the use of the GPA in South Africa.

(iii) The first application of model uncertainty used for the selection of a suitable design flood distribution.

(iv) The first application of the KAP3 methodology and determination of a national Kappa *h* value for South Africa.

(v) The RoI, super-region and multi-variate clustering approaches have also not been applied in South Africa before and previously geographic and morphological maps were used, or reliance was often placed on historically defined homogeneous regions.

(vi) A new and unique set of relatively homogenous clusters for use with DFE have been developed.

(vii) Development and performance assessment of four models using QRT and PRT modelling approaches at multiple scales for South Africa. This included the comparison of the models' predictive ability and identified that the equally weighted IF approach outperformed the record length and Euclidian distance weighted record length weighting, which is in contrast with international findings.

In addition to the above, the study also identified that in the South African context, which was shown to be relatively data rich, the Clustering regionalisation scheme provided the best overall quantile flow estimates, which is in contrast to international findings where in data rich regions, RoI is generally favoured. The result in this study does not exclude the RoI, but only in the current form applied, the inclusion of additional descriptors, weighting schemes and model formulations may improve the RoI performance. Similarly, the equally weighted IF was found to perform best, in contrast to international findings where generally record length weighting has been shown to improve results.

In addition this research has contributed to the following key projects identified by the National Flood Studies Programme (NFSP) as outline by Smithers *et al.* (2014):

(a) A.1.2.2 Guide for AEP distribution for floods
(b) A.1.2.3 Spatial and Temporal distribution of available streamflow data
(c) A.1.2.6 Refined regionalised / pooled Index flood methods
(d) A.1.2.7.1/3 Update and refine the RMF method and its regionalisation
(e) A.1.2.8.2/3 Modernise the Standard Design Flood Method
(f) A.1.2.8.5 and A.1.2.4 Modernise existing synthetic unit hydrographs and related homogeneous flood regions
(g) A.1.2.8.6 Modernise existing empirical methods for small catchments
(h) C.2 Web-based framework of methods on SANCOLD website
(i) C.6 Web-based GIS database/geodatabase

ACKNOWLEDGEMENTS

To the staff at SRK Consulting Durban, Murray Sim, Nic Brien and James Morris, without your knowledge and training large portions of this research would not have been possible.

Thanks to the Department of Water and Sanitation for access to their data in particular Theresa Taylor who gave up a weekend extracting the data.

Special thanks are extended to Prof. OJ Gericke, for your insightful comments and practical knowledge, Dr TR Kjeldsen, the foreign influence and never-ending well of inspiration, and last but not least, Prof. JC Smithers, my endlessly patient supervisor, I could and have sat for hours in your office discussing the ins and outs of the work here presented. Thank you all.

To my parents' in-law, Susan thank you for taking the time to crit my English, which was often illegible. Kevin, thanks for the countless beers and encouragement.

A thank you to my family, my parents who have put countless hours into assuring that I am privileged with a great education. Ek is lief vir julle en waardeur al julle harde werk.

To my dear wife, Kerry, your constant support and motivation is what made this possible and you deserve this more than I. Mason and Ava, my kids, let this be an example of never giving up. Vasbyt!

TABLE OF CONTENTS

Page

LIST OF ABBREVIATIONS

ACRU	Agricultural Catchments Research Unit
AD	Anderson-Darling
ADC	Modified Anderson-Darling Criterion
AEP	Annual Exceedance Probability
AIC	Akaike Information Criterion
AICc	Second Order AIC
AMS	Annual Maximum Series
ARF	Aerial Reduction Factor
ARR	Australian Rainfall and Runoff
AU	Modified Anderson-Darling
BAD	Bootstrap Anderson Darling
BIC	Bayesian Information Criterion
BML	Bayesian Maximum Likelihood
CS	Chi-Squared
CvM	Cramer von Mises
DEM	Digital Elevation Model
DFE	Design Flood Estimation
DWS	Department of Water and Sanitation
EMA	Expected Moment Algorithm
FCC	Filliben Correlation Coefficient
FEH	UK Flood Estimation Handbook
FFA	Flood Frequency Analysis
GAM	Gamma distribution
GEV	Generalised Extreme Value distribution
GLO	Generalised Logistic distribution
GNO	Generalised Normal distribution
GoF	Goodness-of-Fit
GPA	Generalised Pareto distribution
GUM	Gumbel distribution
HRU	Hydrological Research Unit
IAHS	International Association of Hydrological Sciences
IF	Index Flood
JPV	Joint Peak-Volume
KAP	Kappa distribution
KAP3	Three parameter Kappa distribution
KS	Kolmogorov-Smirnov
KZN	KwaZulu-Natal
L-CV	Linear Moment Coefficient of Variation
LHM	LH-moments
LM	Linear moment

LMRD	Linear Moment Ratio Diagram
LNO	Log-Normal
LOO	Leave-one-out
LP3	Log-Pearson Type 3 distribution
LPR	LP3 with a regional skew
MAF	Mean Annual Flood
MAP	Mean Annual Precipitation
MARE	Mean Absolute Relative Error
MC	Monte Carlo
MCMC	Markov Chain Monte Carlo
MEF	Median Annual Flood
ML	Maximum Likelihood
MM	Method of moments
NFSP	National Flood Studies Programme
NOR	Normal distribution
NSE	Nash Sutcliffe Model Efficiency
PCA	Principal Component Analysis
PDS	Partial duration series
PE3	Pearson Type 3
PMD	Product Moment Diagrams
PRM	Probabilistic Rational Method
PRT	Parameter Regression Technique
PWM	Probability weighted moments
QRT	Quantile Regression Technique
RE	Relative Error
RFFA	Regional Flood Frequency Analysis
RFFE	Regional Flood Frequency Estimation
RGC	Regional Growth Curve
RMSE	Root Mean Square Error
RoI	Region of Influence
RR	Rating Ratio
SANCOLD	South African National Commission on Large Dams
SANRAL	South African National Roads Agency Limited
SCS	Soil Conservation Services
SDF	Standard Design Flood
SF	Scaling Factor
SRTM	Shuttle Radar Topography Mission
TauDEM	Terrain analysis using digital elevation models
TCEV	Two Component Extreme Value
THL	Threshold Distance
TSNE	T-distributed Stochastic Neighbour Embedding
UK	United Kingdom
UMAP	Uniform Manifold Approximation and Projection

USA	United Stated of America
USBR	US Bureau of Reclamation
USGS	US Geological Survey
WEI	Weibull
WR2012	Water Resources 2012 study
WRC	Water Research Commission

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

Engineers rely on design hydrological information for the design of hydraulic structures, such as dams, bridges, and drainage culverts (Smithers and Schulze, 2003). This information is often estimated at ungauged sites using models to estimate flood frequencies (Schulze *et al.*, 2004, Blöschl *et al.*, 2013, Smithers *et al.*, 2015). The over or underestimation of design floods could lead to significant economic losses, loss of lives or under or over design of a structure, which results in loss of critical resources if under designed or a waste of capital resources if over designed. The financial impacts of flooding in South Africa have been reported to be up to R1 billion per storm event in regions such as the Western Cape in 2008 by Holloway *et al.* (2010). Table 1.1 contains selected statistics of damage caused by recent flood events in South Africa.

Table 1.1 Social and monetary flood damages of recent flooding events in South Africa

Year	**Region**	**Estimated Damage**	**Reference**
2019	KwaZulu-Natal (KZN)	60 Deaths – R650 Million	Singh (2019)
2017	KZN	3112 households – R576 Million	eNCA (2017)
2016	KZN	7 Deaths	Davies (2017)
	Western Cape	10 000 – 15 000 people displaced	
2011	Northern Cape	R50 Million	Shiceka (2011)
	North West	R6 Million	
	KZN	R300 Million	
2008	Western Cape	R1 Billion	Holloway *et al.* (2010)

Accordingly to Rahman *et al.* (2009), the estimated cost of projects involving the determination of design floods for small to medium-sized rural catchments was approximately AU$ 250 million in 1985 in Australia. This was estimated to increase to the equivalent of AU$ 600 million (approximately R4 billion) in 2009 (Rahman *et al.*, 2009). Stedinger and Griffis (2008) noted that the death toll caused by floods in the United States of America (USA) is approximately 140 per annum, with a financial cost of US$ 6 billion annually, excluding recent events such as Hurricane Katrina.

Design Flood Estimation (DFE) techniques can be broadly categorised as analysis of streamflow data or rainfall-based methods (Smithers and Schulze, 2003), and Figure 1.1 shows the categories of DFE methods available in South Africa. Analysis of streamflow data uses statistics of observed floods to derive estimation techniques such as flood envelopes or empirical formulae. Alternatively, Flood Frequency Analysis (FFA) can be performed to fit a probability distribution to the observed data. Rainfall-based methods use either design or observed rainfall and rainfall-runoff models to estimate design floods, which range from event-based models, which utilise design rainfall as input, to the use of continuous simulation modelling using observed rainfall.

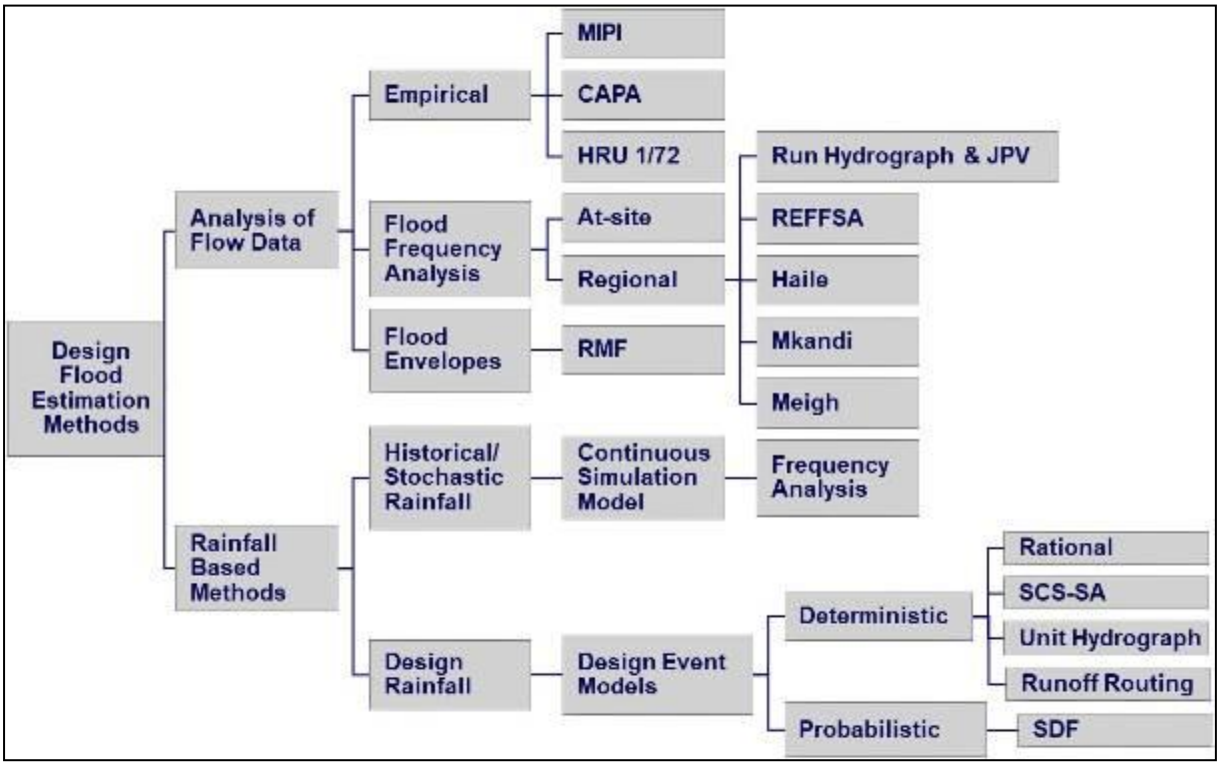


Figure 1.1 Design flood estimation methods available for use in South Africa (after Smithers, 2012)

No single DFE method has been identified as the most appropriate method and, in many texts and manuals, the use of a combination of these are recommended (eg. Pilgrim and Cordery, 1993, Alexander, 2002b, Chadwick *et al.*, 2004, SANRAL, 2013). When estimating design floods for a site, although several methods might be applicable, they may produce vastly different results, which poses the practitioners with the dilemma of which results to use. Pilgrim (1989) identified the following four requirements of a DFE method to ensure the selection of the best possible approach to DFE:

(a) needs to be based on observed flood data,

(b) needs to be simple, lack ambiguity and have familiarity in its application,

(c) should be probabilistic rather than deterministic, and

(d) should incorporate regional differences in hydrological responses.

In the South African context the majority of methods shown in Figure 1.1 were developed in the 1970's and 1980's (Smithers *et al.*, 2014). Additionally, Alexander (2002b), Smithers and Schulze (2003), Görgens (2007a), Smithers (2012) and Van Vuuren *et al.* (2013) highlight the need to revise existing methods and develop new DFE methods for South Africa.

Nathanael *et al.* (2018) assessed the performance of four regional DFE methods applicable to South Africa: (i) Meigh *et al.* (1997), (ii) Mkhandi *et al.* (2000), (iii) Joint Peak-Volume (JPV) (Görgens, 2007a), and (iv) Haile (2011). Nathanael *et al.* (2018) utilised the desirable range of model estimated versus statistical flood quantiles of 0.5 – 2, as defined by Rahman *et al.* (2012), and identified that the percentage of estimated values within the desirable range varied between 48 and 60% for the various Annual Exceedance Probabilities (AEPs). This is further exacerbated when considering catchment areas less than 100 km^2 where the percentage of sites within the desirable range varied between 28 and 57 % for the methods considered. A survey undertaken by Van Vuuren *et al.* (2013) identified that majority of catchments for which DFE are undertaken are relatively small (< 15 km^2), which raises concern regarding the performance of the regional models.

Based on the assessments and recommendations identified in the literature, a National Flood Studies Programme for South Africa (NFSP) was initiated by the South African Committee on Large Dams (SANCOLD) and the Water Research Commission (WRC) (Smithers *et al.*, 2014). The NFSP established working groups which identified 36 research topics grouped into four overarching categories: (i) Rainfall, (ii) Data, (iii) Analysis/Methods, and (iv) Products. 18 of the proposed research topics focus on the analysis of flood data and the refinement of existing or development of new DFE methods. The delineation of new/refined homogeneous flood producing regions (Projects A.1.2.7.1/3. A.1.2.8.5, A.1.2.4, and A.1.3.1), refinement of existing regional DFE methods (Project A.1.2.2) , and a review and guidance on identification of the most suitable frequency distribution/s (Project A.1.2.6) for use in South Africa are identified as key research projects by the NFSP (Smithers *et al.*, 2014) and require the development of a regional flood model. A refinement and extension of the regionalised Index flood (Dalrymple, 1960) based studies undertaken (Mkhandi *et al.*, 2000, Kjeldsen *et al.*, 2001, 2002, Haile, 2011) is also proposed. The availability of extended datasets and improvements or development of new DFE methods applied internationally further substantiates the needs to develop new methods applicable to South Africa.

Where adequate observed flood data are available, FFA is the recommended approach and can be implemented locally or regionally. Even if flow data are available at the site of interest, the augmentation of at-site information can be achieved through applying Regional Flood Frequency Analysis (RFFA) approaches, which can substantially improve the accuracy of the frequency analysis (Kjeldsen *et al.*, 2014, Rahman *et al.*, 2019). In the United Kingdom (UK) (Kjeldsen *et al.*, 2008b), Europe (Castellarin *et al.*, 2012), USA (England *et al.*, 2018) and Australia (Rahman *et al.*, 2019) the use of regionalised approaches to FFA is widely adopted. A general framework for the development of a regional flood model can be divided into three steps: (i) Collation of catchment descriptors and quality-controlled streamflow data, (ii) Selection and application of a regionalisation scheme for the formation of pooling groups, and (iii) regional knowledge transfer model development.

The Department of Water and Sanitation (DWS) are the custodians of hydrological streamflow monitoring in South Africa and the catalogue of stations with data available consists of a network of 1 458 flow gauging stations across the country with a combined record length in excess of 40 000 years. Pitman (2011), however, highlights that the dataset contains numerous inconsistencies and requires critical review prior to use. In addition the fact that the number of useful DWS flow gauges open in each year has been declining since the late 1980's, as shown in Figure 1.2. Kjeldsen *et al.* (2002) also noted the difficulties of the flow monitoring dataset, more specifically referring to the exceedance of gauging station rating curves and the unknown associated uncertainty in the calculation of extreme flood flows exceeding these. Nathanael (2015) assessed 1097 stations to identify the extent of rating table exceedance and found that only 39% of the stations did not contain rating table exceedances.

There exists an abundance of catchment descriptor datasets for use in South Africa, from studies such as the Water Resources 2012 (de Groen *et al.*, 2015) and the South African Agrohydrology and Climatological Atlas (Schulze *et al.*, 1997) having been developed historically. Extracting relevant descriptors from these datasets for model application are, however, continually repeated by practitioners, on an as and when needed basis, at a catchment scale due to a lack of a centralised database. This leads to a duplication of effort and the improvement of data quality and availability forms one of the principal requirements set out by the NFSP (Smithers *et al.*, 2014). The catchment descriptors are imperative for the development of regional flood models and provide both an indication of catchment similarity, or dissimilarity, and are utilised as predictor variables in flood models.

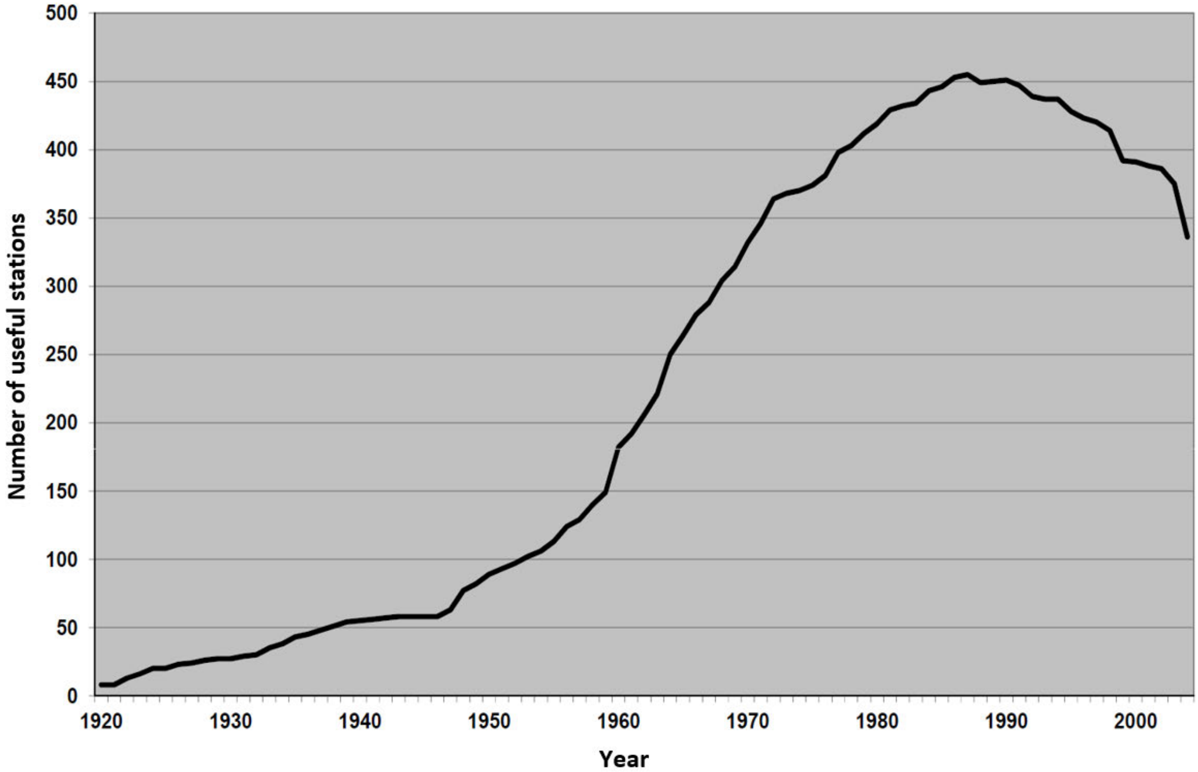


Figure 1.2 Number of useful flow gauges open in each year (after Pitman, 2011)

The use of catchment descriptors to characterise catchment similarity is used for the identification of pooling groups or regions, which allows for the transfer of knowledge from gauged to ungauged sites, which is a key benefit of RFFA. There is no clear consensus on the best method of identifying the pooling groups in hydrology (Oudin *et al.*, 2008, He *et al.*, 2011, Blöschl *et al.*, 2013, Mostofi Zadeh and Burn, 2019), but it has been shown to be dependent on region and climate (Razavi and Coulibaly, 2013), with spatial proximity identified as offering the best solution (Merz and Blöschl, 2005, Oudin *et al.*, 2008, Mostofi Zadeh and Burn, 2019). When considering the forming of pooling groups, Smithers and Schulze (2003) and Van Bladeren (2005) both recommend that a rigorous statistical based approach to regionalisation be adopted. Numerous regionalisation schemes exist for the formation of pooling groups, with the most eminent of the schemes being the Region of Influence (RoI) proposed by Burn (1990) and clustering approaches as adopted by Rahman *et al.* (2019). Regionalisation for DFE purposes in South Africa have been performed by HRU (1972), Kovács (1988), Meigh *et al.* (1997), Mkhandi *et al.* (2000), Kjeldsen *et al.* (2001) for the KZN province, Alexander (2002a), Görgens (2007a) and Haile (2011) but, with the exception of HRU (1972), have not been widely adopted for local use. Mkhandi *et al.* (2000), Kjeldsen *et al.* (2001) and Haile (2011) can be considered the only studies that have performed statistically based flood regionalisation within South Africa. Mkhandi *et al.* (2000) and Haile (2011) grouped catchments through manual inspection of geographic data and verified homogeneity through statistical testing. Kjeldsen *et al.* (2001), however, adopted the clustering procedure outlined by Hosking and Wallis (1993, 1997), but only within the KZN province, and compared the homogeneity of the regions formed to those delineated by Mkhandi *et al.* (2000) and Kovács (1988), and improved the homogeneity in relation to these delineations, thereby warranting further investigation at a national scale.

After suitable pooling groups have been formed, the knowledge of the stations needs to be shared to improve the overall estimates within the group. This is achieved through a number of methods that generally rely on the use of a regional model. As with all models the response variable is estimated through some combination of predictor variables. The RFFA response variables are associated with the estimation of flood frequency curves or flood quantiles at a pooling group scale, which requires the identification of a suitable distribution for the estimation of the flood quantiles. For DFE purposes at ungauged sites, modelling approaches include regional methods (e.g. index flood method), direct regional regression of quantiles, and regional regression of distribution parameters (Aronica and Candela, 2007). Suitable predictor variables are then utilised for the estimation of the required response variables at ungauged sites, e.g. Mean Annual Flood (*MAF*), 1% AEP flood or distribution parameters such as skew and standard deviation.

FFA requires the identification of an appropriate frequency distribution for fitting the historical data and is generally performed at an at-site basis, conversely, the selection of an appropriate distribution for RFFA is undertaken considering the identified pooling groups and is required for the development of the regional knowledge transfer models. The choice of distribution can have a considerable impact on the estimated peaks. For example, Alexander (2002a) demonstrated that the design flood estimate of a 0.5% AEP flood obtained using different distributions fitted to the same data set could result in variations of up to 38%. For conducting

FFA in South African catchments the distributions recommended for application are the Log-Pearson Type 3 (LP3), Log-Normal (LNO), and Generalised Extreme Value (GEV) (Alexander, 1990, 2000, Görgens, 2007b, Gericke, 2010, SANRAL, 2013, Van der Spuy and Rademeyer, 2018), but these texts do not provide guidance or indication for the selection of the most suitable distribution. Kjeldsen *et al.* (2002), Mkhandi *et al.* (2000) and Haile (2011) performed statistical verification of suitable distributions and found that the Generalised Normal (GNO), LNO, Pearson Type 3 (PE3) and Generalised Pareto (GPA) distributions were all suitable candidates. The studies were, however, performed in specific geographic regions or utilised limited datasets. The literature therefore supports the need identified by the NFSP for additional verification of a suitable distributions for use in South Africa.

1.1 Problem Definition

Practitioners are often required to select the most appropriate design flood method to apply, a choice which incorporates the financial and societal risks that affect the design being undertaken. The methods currently available in South Africa were either developed outside of the country or decades before and based on limited datasets. Currently more extensive data sets and modelling approaches are available to provide a better understanding of the South African flood risk regime. Therefore, it is hypothesised that, by utilising the extended data sets, and refined modelling approaches, a regional model which provides results with a higher level of confidence can be developed.

1.2 Research Question

Based on the above overview, the research question this study is addressing is: Can a regional DFE model be developed, utilising the most appropriate statistical distribution, coupled with a statistically developed regionalisation scheme utilising the most recent instantaneous AMS data, that provides improved estimated design flood estimates in relation to the currently available regional flood estimation methods?

The research question can be divided into the following sub-questions:

(a) What is the most suitable distribution for FFA in South Africa on an at-site scale for use on a national scale?

(b) Can South Africa's catchments be divided into statistically homogeneous flood producing regions?
(c) Given the data sparsity in South Africa, which regional DFE model is most suitable?

1.3 Aims and Objectives

The aims of this study are to develop an improved and refined regionalised DFE model for South Africa, through the development and assessment of regional model development approaches. Specific objectives that are required to answer the research questions posed include the following:

(a) Compilation of a hydrological descriptors database.
(b) Collation and quality control of selected gauged flow data in South Africa.
(c) Identification of a suitable probability distribution for use in South Africa.
(d) Identify and verify homogeneous flood producing regions.
(e) Regional flood model development and performance assessment.

1.4 General Methodology and Thesis Structure

RFFA studies can be loosely divided into four distinct steps, as shown in Figure 1.3, within which the basis of FFA also exists. This thesis has been written based on the four steps: i) Data collection, ii) Identification of a suitable frequency distribution, iii) Formation of homogeneous flood producing regions, and iv) Model development, each of which has been dedicated an independent chapter.

Chapter 2, refers to step i) of the RFFA process, and details the development of a national descriptor and streamflow database. The various data sources are outlined for both streamflow data and catchment descriptors. A review of relevant literature will be presented describing the selection of the adopted catchment descriptors. Additionally, the streamflow data screening process is outlined, and an analysis of the available data is presented.

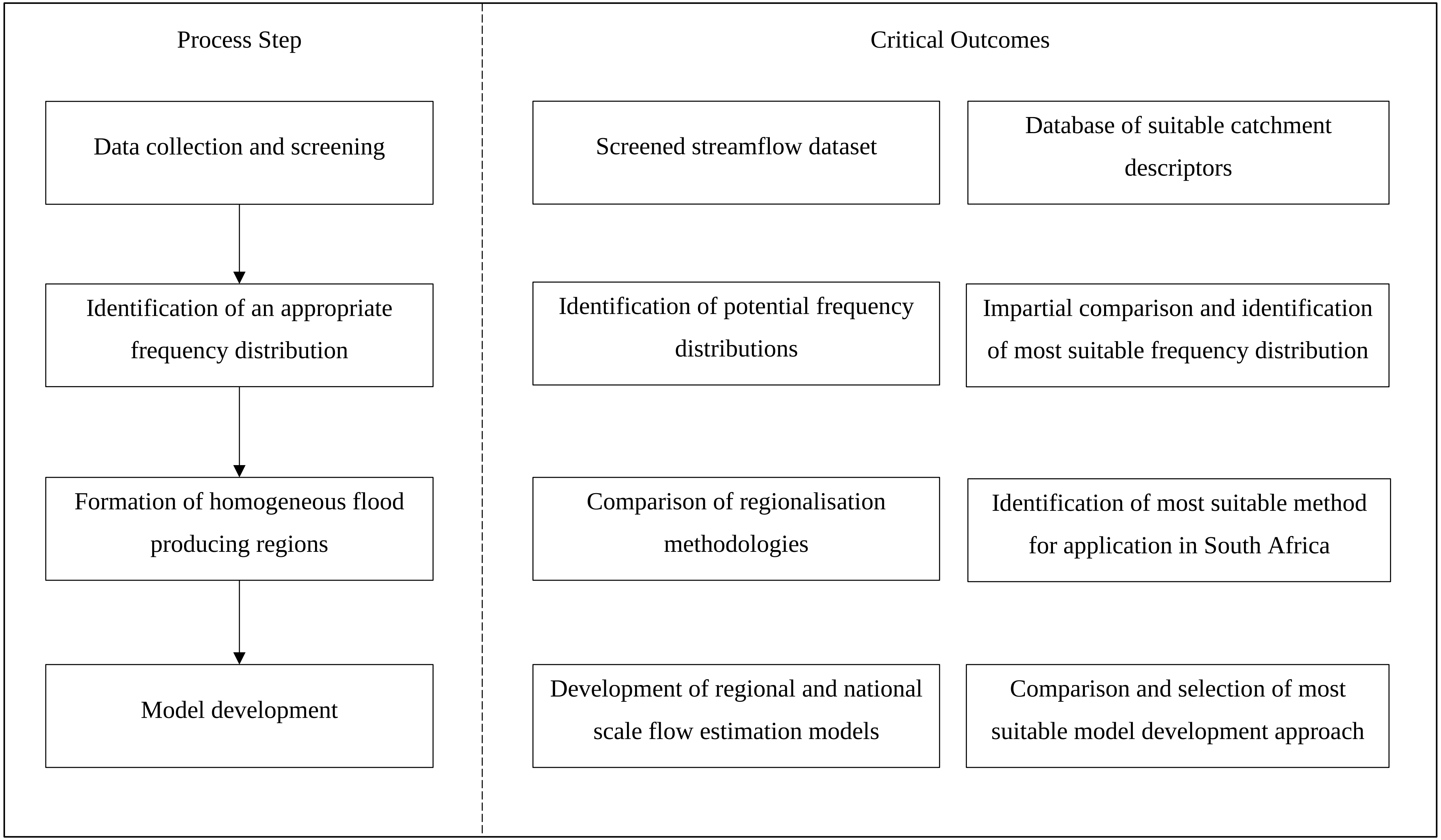


Figure 1.3 RFFA process flow diagram

The identification of the most suitable probability distribution for FFA in South Africa is detailed in Chapter 3. The Chapter reviews existing practise and recommendations in South Africa and internationally for the selection of a suitable distribution. A detailed investigation identified five potentially applicable distributions for South Africa. The selected distributions consisted of distributions commonly applied in South Africa and additional more flexible distributions applied elsewhere. An assessment was then developed considering four selection criteria, i) Graphical methods, ii) Goodness-of-fit, iii) Model selection criterion, and iv) predictive ability. Through ranking the performance of the frequency distributions for the different selection criteria, a recommendation on the most suitable frequency distribution is presented.

Chapter 4 describes the process undertaken for forming the homogeneous pooling groups. A review of existing flood regions developed for South Africa was undertaken, highlighting the adopted methods and potential shortcomings. The two most commonly applied and potentially applicable multi-variate pooling group techniques were applied to the catchment descriptors and streamflow data collated in Chapter 2, for the forming of homogeneous flood producing regions. Additionally, an approach combining the methods was undertaken in an attempt to improve the homogeneity of the developed regions.

Following the formation of the homogeneous flood producing regions, the development and performance assessment of four regional flood models was undertaken and is presented in Chapter 5. The flood models were developed at two scales, national and regional, based on the two regionalisation techniques employed in Chapter 4. Six performance metrics were adopted to assess the predictive ability of the developed models and undertake a comparative assessment. Using the adopted metrics, it was possible to identify the best performing model, as well as identify potential regions where the model performs less adequate than the remaining models.

The overall study results are discussed in Chapter 6 and the conclusions and recommendations for future research are presented in Chapter 7.

Chapters 2 to 5 have been written with the intent of being published as individual papers, and as such may replicate previously stated information. This has, however, been limited to only a few instances and in the opinion of the author improves the readability of the thesis.

2 DEVELOPMENT OF THE HYDROLOGICAL DESCRIPTOR DATABASE

The development of a unified hydrological catchment descriptor database for use across multiple studies is a critical requirement for the development and application of methods for DFE in South Africa. A unified database would move much of the focus of flood studies from the extraction of catchment descriptors to the application of the methods and hydrological judgements, thus reducing duplication of efforts. The NFSP has identified the collation of a hydrological database as a key project, which further justifies the need for its development. The descriptors and flow data are also critical for the development of regional flood models, simultaneously being used for the forming of pooling groups and as predictor variables in the final models.

2.1 Base Data Collation

Estimation of the hydrological descriptors required the collation of the following base data sets:

(a) topographic data,

(b) rainfall data, and

(c) DWS catchment, river network and streamflow data.

2.1.1 Topographic data

The Shuttle Radar Topography Mission (SRTM) (NASA-JPL, 2013) data were utilised for the development of the Digital Elevation Model (DEM). A hydrological conditioning process adjusts the DEM to ensure that flow directions derived from the surface defines the expected flow directions. A common methodology followed for hydrological conditioning is filling (Fernandez *et al.*, 2016), whereby the DEM is assessed for any potential voids or impressions that could prevent the derivation of natural flow lines. After the voids or impressions have been identified, the elevations are increased until the water would flow along a natural pathway. Fernandez *et al.* (2016) identified that, although alternative methods are available for hydrological conditioning, the filling procedure maintained the slope descriptors of the catchment.

The 30 x 30 m SRTM grid was used and, where necessary, infilling was undertaken using the 90 x 90m grid. Infilling was, however, limited to a small region in the Eastern region of the Western Cape province. The infilled DEM was used for the determination of topographic descriptors such as catchment area (A) and elevation at the outlet (E_O).

2.1.2 Rainfall

Three sets of rainfall data were utilised for the study, the Mean Annual Precipitation (*MAP*), the daily rainfall data set, collated and infilled by Lynch (2004) and design rainfall depths (Smithers and Schulze, 2003).

The *MAP* data sets utilised was extracted from the Water Resources 2012 study (de Groen *et al.*, 2015) and the data set developed by Lynch (2004). The data sets consist of national *MAP* depth grids with a minute by minute grid spacing. The minute by minute design rainfall grid developed by Smithers and Schulze (2003) was utilised to derive the design rainfall depths.

2.1.3 DWS data

The DWS data utilised in the study included river networks, primary to quaternary catchment boundaries, gauging station locations, and flow monitoring data. These data were utilised for verification, location purposes and statistical analysis of the instantaneous Annual Maximum Series (AMS). The quality and limitations of the flow data received are further discussed in Section 2.3.

2.2 Descriptor Extraction

As identified in previous flood studies (McDermott and Pilgrim, 1982, Robson and Reed, 1999, Mkhandi *et al.*, 2000, Alexander, 2002a, Van Bladeren, 2005, Görgens, 2007a, Kjeldsen *et al.*, 2008a, Gericke, 2010, Haile, 2011, Rahman *et al.*, 2015b), the geographic location, rainfall intensity, *MAP* and catchment area are potential descriptors used for the regionalisation of the peak flow estimation. Taking both the literature and the requirement of ease of application by practitioners into consideration, the descriptors summarised in Table 2.1 which are readily available, or simple to estimate, were selected for inclusion in the study:

Table 2.1 Catchment descriptors selected for use in the study, including sources where relevant

Descriptor	Unit	Range		Source*
		Min	Max	
Outlet latitude	Decimal degrees	-34.36	-22.63	(DWS, 2011)
Outlet longitude	Decimal degrees	18.69	32.18	
Outlet elevation	masl	11.00	1969.00	(NASA-JPL, 2013)
Catchment area	km^2	0.26	361994.80	
Catchment perimeter	km	2.94	6075.93	
Rainfall region/cluster	Unitless	1	76	(Smithers and Schulze, 2003)
Rainfall seasonality	radians	-3.11	3.08	
Catchment runoff (C_{ro})	Percent	4.00	97.00	(Schulze, 2011)
SCS soil classifications	Unitless	0	7	(Schulze and Schütte, 2020)
Distance from the coastline (D_c)	Decimal Degrees	0.03	6.84	
Hydraulic length	km	0.81	1896.13	
Length to centroid	km	0.07	978.10	
Slope	m/m	0.0004	0.26	
Time of concentration	Hours	0.17	435.72	
Areal reduction factor (*ARF*)	Percent	57.91	100.00	
Mean Annual Precipitation (*MAP*)	mm	60.00	3312.00	(Lynch, 2004, de Groen *et al.*, 2015)
Design rainfall	mm	9.30	416.40	(Smithers and Schulze, 2003)

* Where no source is specified the parameters were estimated as part of the study and are detailed in the following sections

2.2.1 Catchment area (km²)

The catchment areas for each of the DWS gauging stations was programmatically delineated using the Terrain Analysis Using Digital Elevation Models (TauDEM) suite of programs (Tarboton, 2016). A problem was encountered with the delineation of catchment areas, where the gauging stations were not located on the drainage paths defined by the hydrological conditioning, as undertaken in Section 2.1.1. Two approaches were adopted to correct the delineation: (i) allowing for a 200 m clipping radius, which moves the gauging station location to the point of highest flow accumulation within a 200m radius, or (ii) by manual manipulation of the gauging station locations to coincide with the defined drainage paths. On a national scale this will, however, not be significant as the descriptors are derived on a 30 x 30 m grid, and the drainage paths will be available for practitioners to assess the location of the ungauged site relative to the drainage path to ensure the correct catchment descriptors can be extracted. A comparison of the catchment areas calculated automatically from the corrected DEM and the catchment areas from DWS are shown in Figure 2.1, and indicates that the method used to estimate catchment area is similar to the areas from DWS. Due to the values ranging from 1 to 160 000 km^2 the values are presented on a log scale.

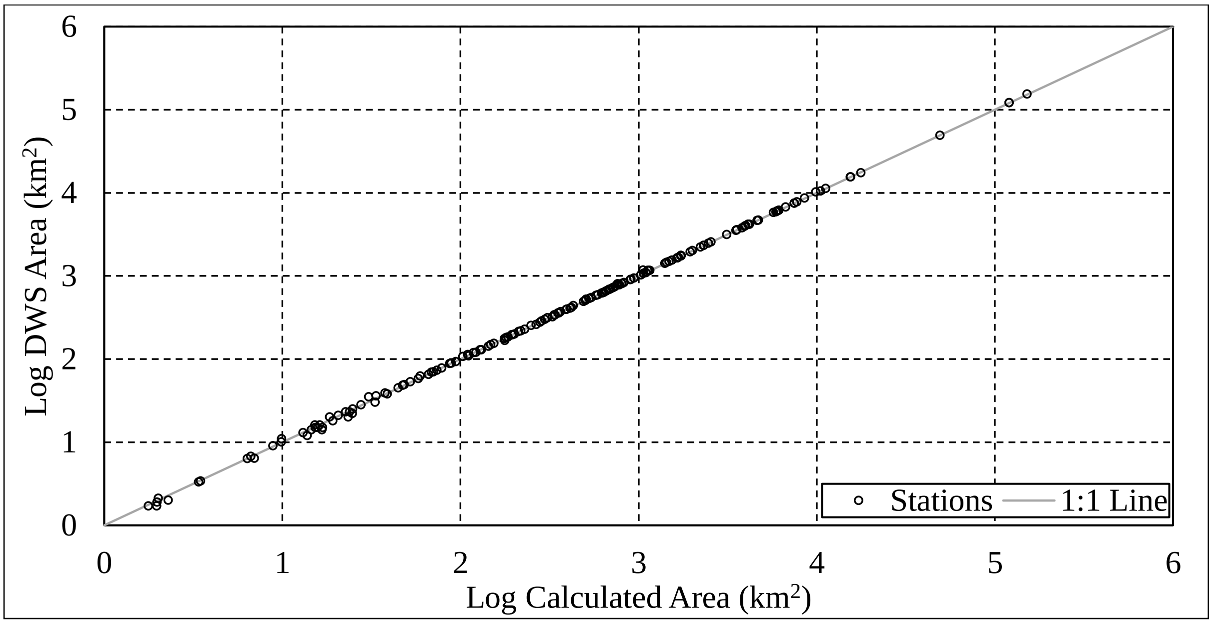


Figure 2.1 Catchment area comparison

2.2.2 Time of concentration (h)

Given the lack of locally developed methods historically SANRAL (2013) and Van der Spuy and Rademeyer (2018) recommend the use of international methods. The method recommended for defined watercourses was developed by the US Bureau of Reclamation (USBR, 1973), which utilises the Hydraulic Length (L), and is shown in Equation 2.1.

$$T_c = \left(\frac{0.87\,L^2}{1000\,S_{10-85}}\right)^{0.385} \tag{2.1}$$

Locally Gericke (2015) developed a regional time to peak (T_p) equation, but the development has been limited to date for selected DWS drainage regions, as shown in Figure 2.2. The T_p is estimated using Equation 2.2 and relates the MAP, A, length to centroid (L_c), L and S_{DEM} to the T_p through the use of calibration coefficients $x_{1\text{-}5}$. provided in Table 2.2.

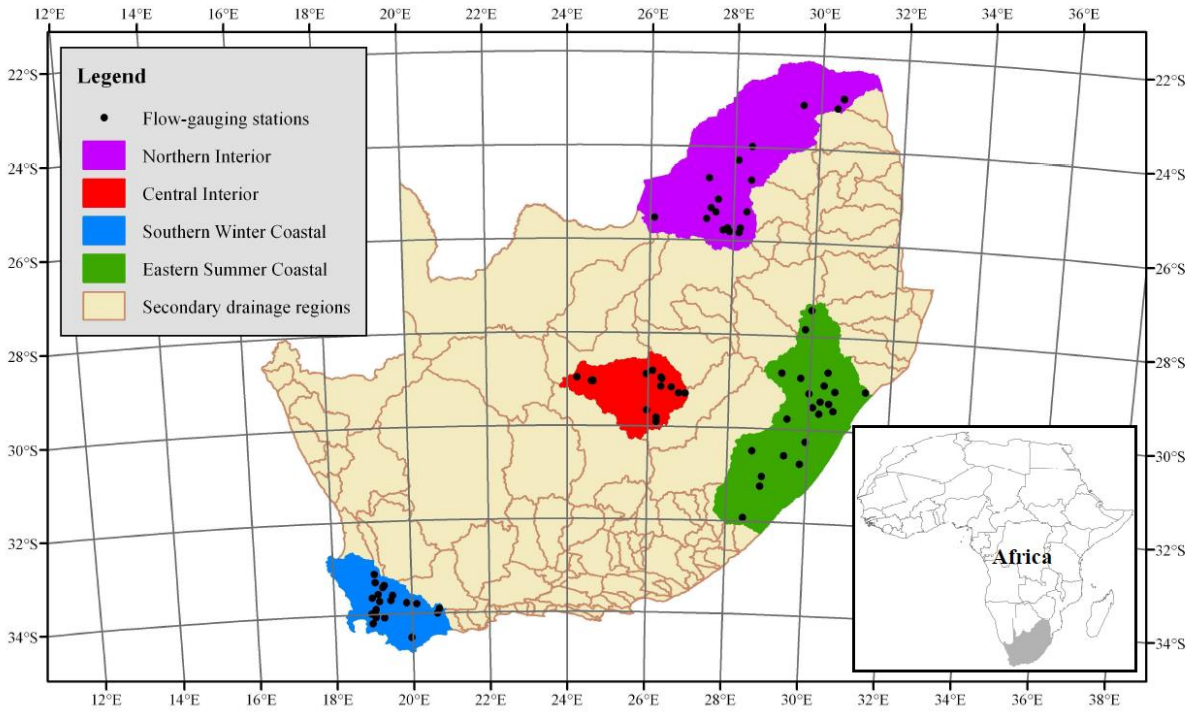


Figure 2.2 South African time to peak (T_p) development catchments (from Gericke, 2015)

$$T_p = {x_1}^{MAP}{x_2}^{A}{x_3}^{L_c}{x_4}^{L}{x_5}^{S_{DEM}} \tag{2.2}$$

Table 2.2 T_p regional calibration coefficients (after Gericke, 2015)

Region	**Regional calibration coefficients**				
	x_1	x_2	x_3	x_4	x_5
Northern Interior	1.00280	0.99993	0.99865	1.01612	0.91344
Central Interior	1.00313	0.99984	1.06106	0.98608	0.98081
Southern Winter Coastal	1.00174	0.99931	1.01805	1.04310	0.99648
Eastern Summer Coastal	1.00297	0.99991	0.99594	1.01177	0.97529

Given the absence of development within the regions surrounding the study areas used by Gericke (2015), the four regions were expanded based on the seasonality and topographic variability to develop expanded regions as shown in Figure 2.3. The expansion of the regions allowed for the estimation of the T_p outside of the original development regions.

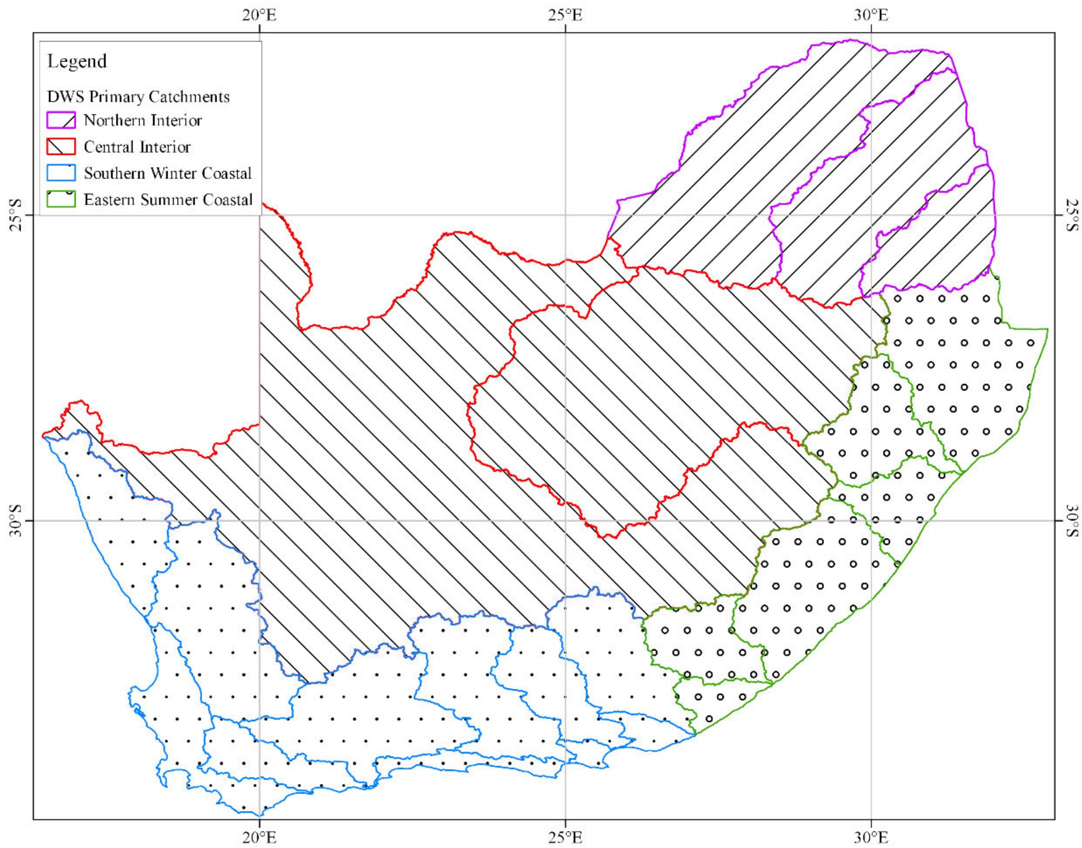


Figure 2.3 Expanded T_p regions

As an additional validity check of the regional expansion, a comparison between the T_p and T_c values was undertaken. When considering the ratio of T_c:T_p a median value of 0.788 and an interquartile range of 0.357 and 1.425 is achieved indicating that the T_c tends to underestimation of the catchment response times in relation to T_p. The validity of the T_p models outside of the catchment size ranges used for development, however, comes into question as, for large catchments, there is a gross under-estimation present regardless of the allocated region. This is particularly evident at stations D7H002, D7H005 and D7H008, where due to the large size of the catchments, the estimated T_p values range between 6.4 x 10^{-67} and 23 hours depending on the region allocation. In contrast the T_c values range between 373 and 435 hours for the same catchments. Given the variability of the T_p estimates outside of the developed catchment sizes and bounds, and the current level of adoption of T_c in practise it has been utilised for this study.

2.2.3 Slope (m/m)

SANRAL (2013) details three methods for the estimation of slope, all of which were calculated for each of the sites investigated:

(a) 10-85 (S_{10-85});

(b) equal area (S_{ea}); and

(c) overland catchment slope (S_{oc}).

Van der Spuy and Rademeyer (2018) also recommend the use of the Taylor-Schwarz method for the estimation of mean river slopes. However, given the adoption of the USBR method for estimation of T_c, the use of S_{10-85} and S_{ea} is required. The above are considered estimations of the average slope calculated along the longest flow paths, utilising variations of the height difference along the flow paths to estimate the average catchment slope. In addition to the above methods, a DEM based catchment slope estimator, S_{DEM}, was calculated and estimates the average catchment slope through determining the average of the maximum slope between neighbouring cells of the DEM. A comparison of the slopes estimated using S_{10-85} and S_{ea}, as shown in Figure 2.4, was undertaken to assess the variability of the estimates.

It is evident from Figure 2.4 that there is a large variation of up to 86% between the estimation methods, with S_{10-85} consistently estimating steeper slopes than S_{ea} by an average of 15%. This will lead to estimates of shorter time of concentration (T_c) and increased peak flows, which could be considered a more conservative approach. Therefore, S_{10-85} was utilised in this study.

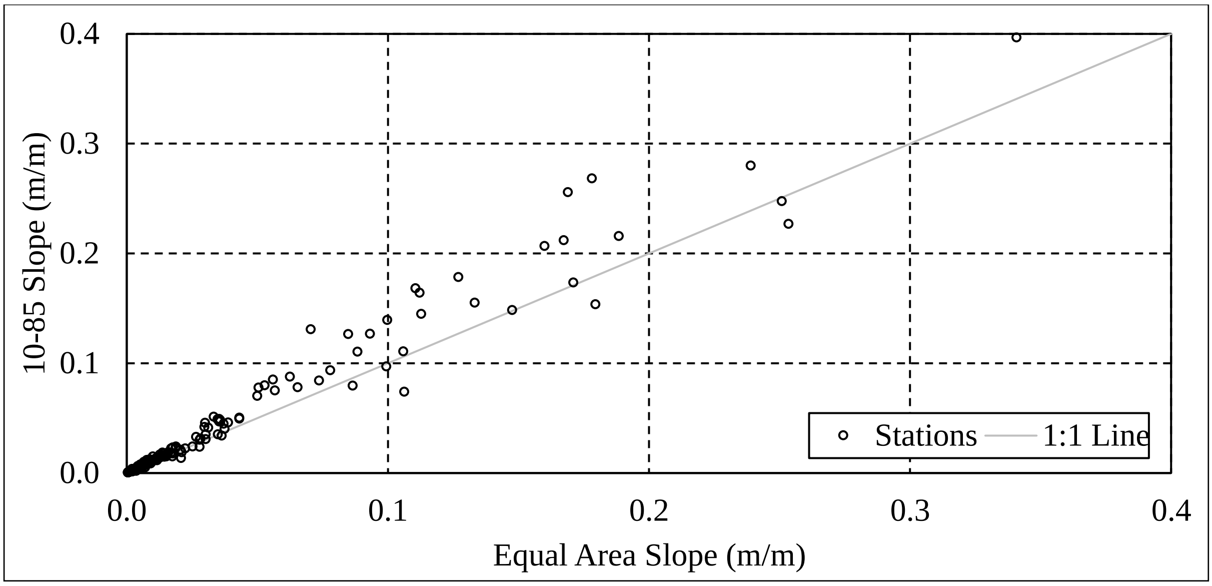


Figure 2.4 Average slope comparison

2.2.4 Areal reduction factor (%)

The method currently recommended by SANRAL (2013) and proposed by Alexander (2001), is shown in Eq. 2.3, and relates the catchment area (*A*) and T_c to the *ARF*.

$$ARF = (90\,000 - 12\,800\ lnA + 9\,830\ ln(60T_c))^{0.4} \tag{2.3}$$

In addition to the above SANRAL (2013) and Van der Spuy and Rademeyer (2018) also provide a number of adjustment curves, but no clear guidance is provided on which method is preferable. As such Eq. 2.3 was adopted for the estimation of ARF.

2.2.5 Rainfall based descriptors

The *MAP* and design rainfall values were calculated at a catchment scale by averaging the gridded values over the catchment. For smaller catchments which contained no grid points within the catchment, the grid point closest to the catchment centroid was utilised to estimate the rainfall descriptors.

Similarly, MAP_{max}, MAP_{min} and MAP_{mean}, which represent the maximum, minimum and mean *MAP* values within the catchment, were derived. The use of the 30 x 30 m grid allows for the identification of the variation of the *MAP* within a catchment.

Rainfall seasonality was estimated using circular statistics as described by Burn (1997), by developing a monthly rain rose, as shown in Figure 2.5, per rainfall station being considered. A national plot of the mean seasonality direction is shown Figure 2.6. For this study the data set developed by Lynch (2004) was adopted due to missing rainfall data being infilled, removing any additional pre-processing.

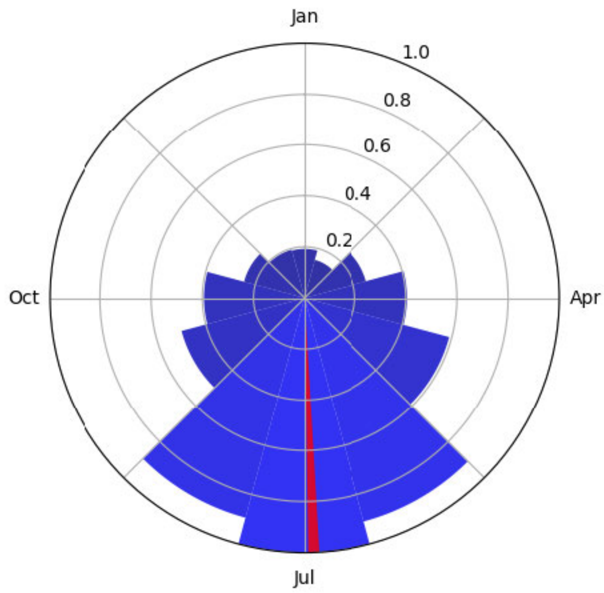


Figure 2.5 Normalised rainfall seasonality for Station 0004816AW indicating the monthly (blue) and average (red) rainfall seasonality

The rainfall regions utilised in the study consisted of the relatively homogeneous daily extreme rainfall (Smithers and Schulze, 2000b) and short duration rainfall clusters (Smithers and Schulze, 2000a). The outlet position dictated the cluster for each gauging station considered.

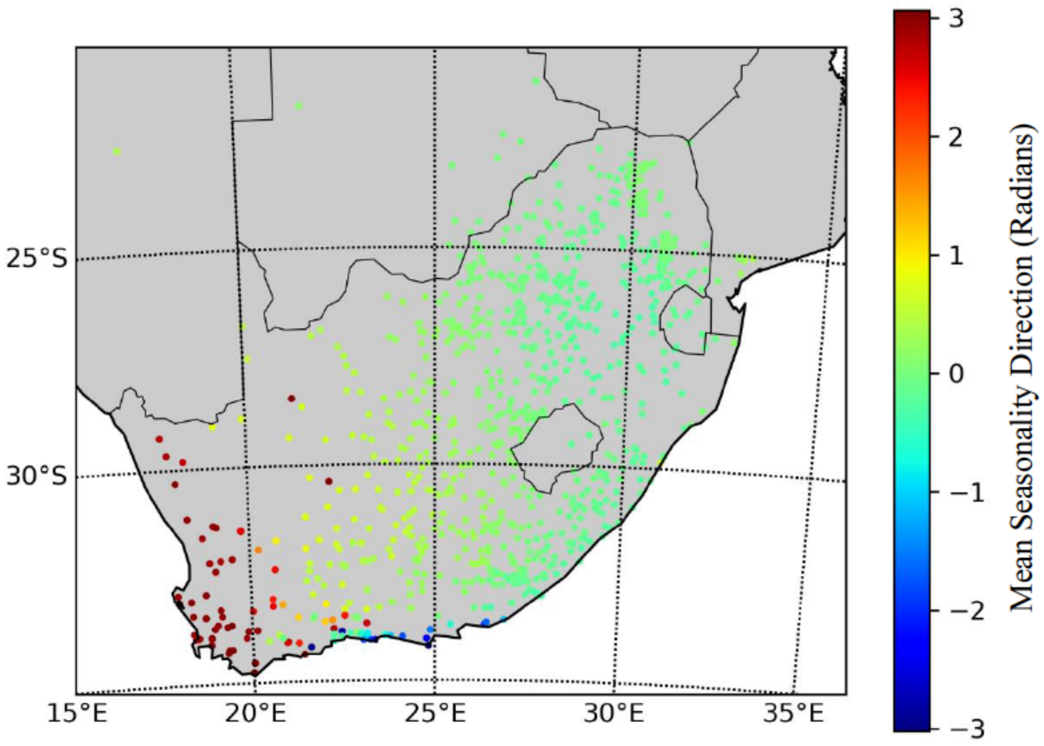


Figure 2.6 National rainfall seasonality indicating the mean direction (radians) of each site investigated

2.2.6 Catchment runoff (%) and soil characteristics

Schulze (2011) developed C_{ro} percentages for naturalised land cover conditions using the Agricultural Catchments Research Unit (ACRU) agrohydrological model and simulating 50 years of runoff from daily rainfall data. The C_{ro} was developed at a quinary level and incorporated into this study as a potential regionalisation parameter. Similarly, Schulze and Schütte (2020) developed SCS soil characteristics map for South Africa at a Terrain unit level, these data were used to estimate the maximum, minimum, average and mode of the SCS characteristics of the catchments.

2.3 Streamflow Data Assessment and Screening

The study utilised the flow gauging stations identified by Nathanael (2015) and extended the data from December 2013 to September 2017, where possible. Primary flow data, derived from flow levels/stage using either breakpoint digitised from autographically recorded levels, or from data logged levels at fixed time intervals, up to 2017 were obtained from the DWS for

411 streamflow gauges. The data consists of instantaneous peak flows and were assessed both in terms of length of record and data quality using a number of criteria for extraction of the AMS. In order to provide reliable design values, long records of data are required. Hence, selecting a minimum record length of 20 years for inclusion in the analysis reduced the number of stations that could potentially be utilised. The second screening process required the identification of human impacts on flow, such as dams, abstractions, and urban development. The registers of dams and abstractions from the WR2012 (de Groen *et al.*, 2015) study, was used for the identification of potentially impacted stations, and secondary manual checks were also performed using aerial imagery to verify and supplement the WR2012 data.

The last criterion considered was the quality of the data. The DWS flow data contain many quality flags ranging from user errors to technical errors. Examples of this include the incorrect manual capturing of data and hardware malfunction. Of the 411 streamflow gauges considered, historical AMS data were available for 160 of the sites as part of the data set prepared by Van Bladeren (1993). The historical peak flow data ware included in the study and was extended to the 2017 hydrological year for the available sites and a combined data set was used in the study. Table 2.3 provides a summary of the screening criteria, similar to the methodology described by Nathanael (2015), and data errors which were used to exclude stations from this study.

As the national available data sets contain many thousands of years of data, it was deemed to be impractical to assess the screening criteria and data errors manually. Therefore, as part of the study the above selection and quality criteria were automated, and recommendations generated based on the primary data received from DWS.

Table 2.3 Data errors and recommendations

Screening Criteria and Data Error	Recommendation
Records shorter than 20 years	Exclusion of site for at-site FFA
Negative/null records	Exclusion of erroneous data
Recorded depth of flow exceeded discharge rating table at flow-gauging station (i.e. “Over-topping”)	Possible extension of the rating tables, otherwise exclusion of erroneous data
Missing periods	Annual records were assessed based on the number of records/days of missing data with possible exclusion of the year. AMS events were immediately excluded if less than 91 days of data were present in a year, otherwise further investigation was undertaken.

In some instances, the recorded river stage exceeded the available discharge rating curves for the flow-gauging stations. Where the rating curve of a station was exceeded, the viability of extending the existing rating curve was assessed. For example, as shown in Figure 2.7, the maximum rated level is 0.96 m, however, the maximum recorded water level for the station is approximately 3.20 m. In such cases simple extension of the rating curve could potentially produce major under or overestimation of peak flow events. Due to the nature of flow gauging weirs, as shown in Figure 2.8, an accurately extended rating curve would require an extensive survey and calibration beyond the structural limit. A general rule was therefore adopted that a rating curve may only be extended up to a maximum of 20% of the original maximum stage, as shown in Figure 2.7. In addition, a limitation of 20% increase in flow discharge exceedance was adopted, similar to Gericke and Smithers (2018), which led to the exclusion of 215 (1%) records. Alternative approaches adopted in other studies include Haddad *et al.* (2010) who adopted a Rating Ratio (RR) which is the ratio of estimated flow to the maximum observed flow. Where a rating curve is extended, a small grouping of 5% of the total number of points located at the upper end of the rating table was considered and a best-fit linear extension was applied. Ninety-five stations had records excluded due to exceedances in excess of the 20% criteria. Although the number of stations that required extensions was not excessive, these values still need to be used with caution due to the uncertainty in the estimation of flow from the recorded stage.

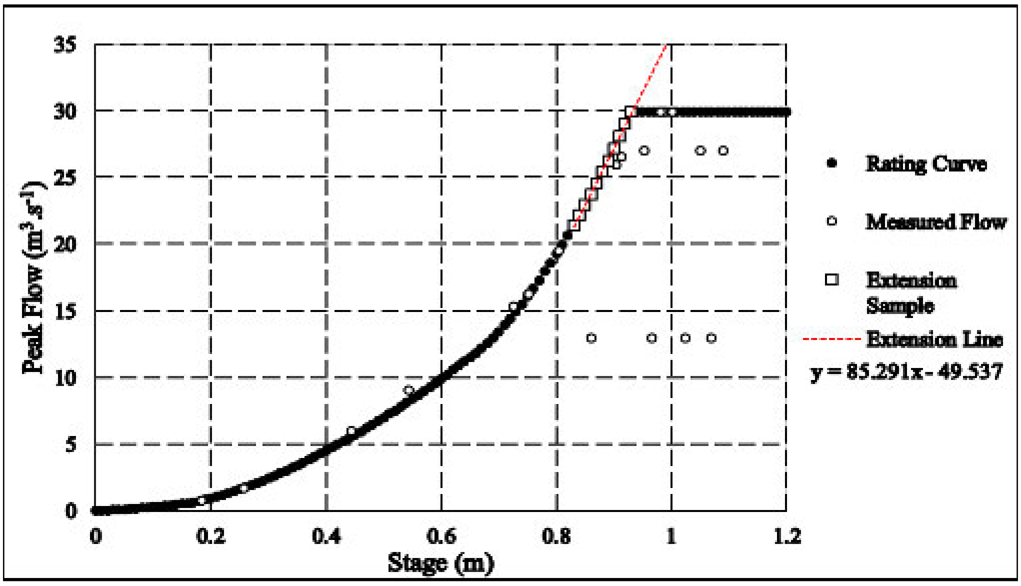


Figure 2.7 Example of a rating curve exceedance and extension the adopted extension methodology

Figure 2.8 Example of a flow gauging weir on the Tongati River at Riet Kuil (DWS, 2015)

The DWS is currently operating 1 458 streamflow gauging stations through-out South Africa. A total of 383 gauging stations, as shown in Figure 2.9, remained after careful data screening, assessment and cleaning and were utilised in the study. The gauging stations are divided into 296 river gauges and 87 synthetic dam inflow records. The synthetic dam inflow records were generated by the DWS Flood Studies section and consists of a combination of at site dam inflow measurements, regional regressions of upstream flow gauges and reservoir routing back-routing from monitored dam outflows and is similar to the flow extension methodology adopted in Bulletin 17C (England *et al.*, 2018).

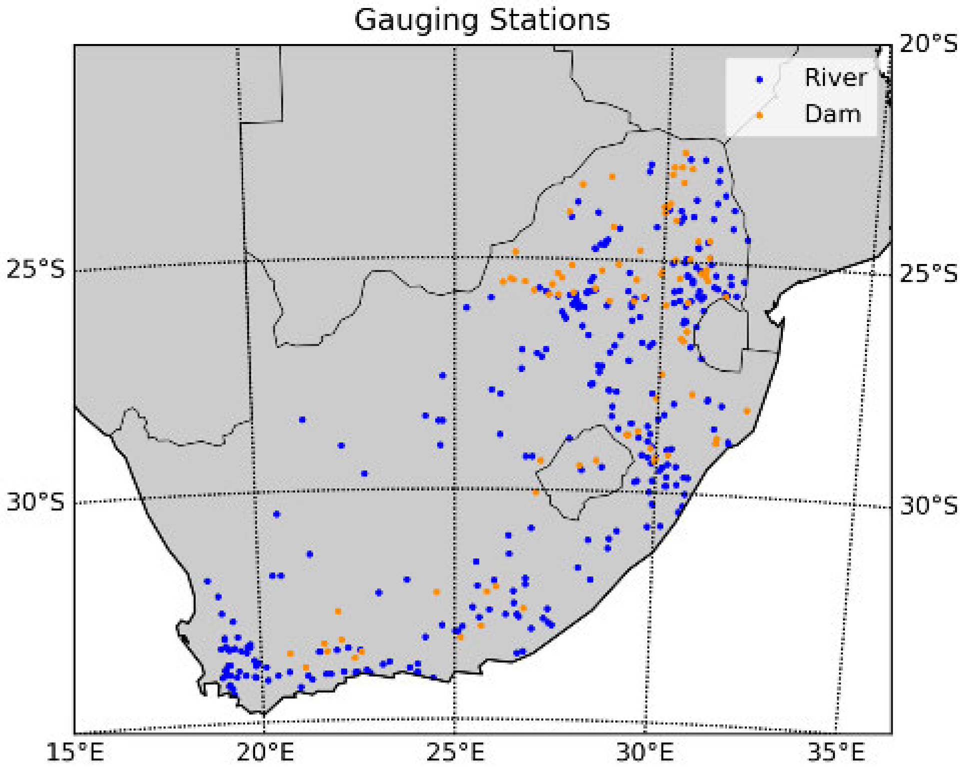


Figure 2.9 Map indicating the DWS gauging stations (blue) and the synthetic dam stations (orange) selected for use in study

2.4 Record lengths of quality controlled annual maximum series

Table 2.4 contains the breakdown of records lengths per DWS drainage region before and after data quality control was undertaken. Figure 2.10 provides a spatial indication of the number of gauging stations after quality control was undertaken. Figure 2.11 shows a histogram of the distribution of the record lengths for the gauging stations considered. The quality-controlled dataset contains a combined total of 18 349 AMS events, with a mean record length of 48 years. These data compares well to the data utilised by Kjeldsen *et al.* (2008a) which contained 602 gauging stations and a total number of 19 679 AMS event, albeit with a lower mean record length of 33 years. Although the data set presents a similar number of AMS events, the data are not evenly distributed across the country, with the centre of the country having a low density of gauging stations. The density of the gauging network relative to the surface area in South

Africa equates to one station per 3 185 km^2, whereas in the UK this drops to one station per 402 km^2. Although this may seem like a poor comparison, Australia holds an even lower density of one station per 9 017km^2 based on the dataset described by Rahman *et al.* (2015c), but includes approximately 31 200 AMS events at the 853 stations. The data available for this study can this be considered a good basis for the development of regional models, given the above comparison.

Table 2.4 Number of DWS flow-gauging stations and record lengths

DWS Drainage Region	**No. of Gauging Stations**	**Cumulative Record Length (years)**		**Mean Record Length (years)**	
		Raw	**Quality Controlled**	**Raw**	**Quality Controlled**
A	62	3347	3314	54	53
B	47	2265	2253	48	48
C	34	1636	1432	48	42
D	20	1147	981	57	49
E	5	290	229	58	46
G	21	868	800	41	38
H	19	791	760	42	40
J	19	1091	1043	57	55
K	10	518	475	52	48
L	5	263	252	53	50
N	4	294	263	74	66
P	2	99	92	50	46
Q	14	635	628	45	45
R	4	176	146	44	37
S	3	179	171	60	57
T	13	782	625	60	48
U	12	619	570	52	48
V	28	1297	1511	46	54
W	19	916	890	48	47
X	42	2039	1914	49	46
TOTAL	383	19252	18349	50	48

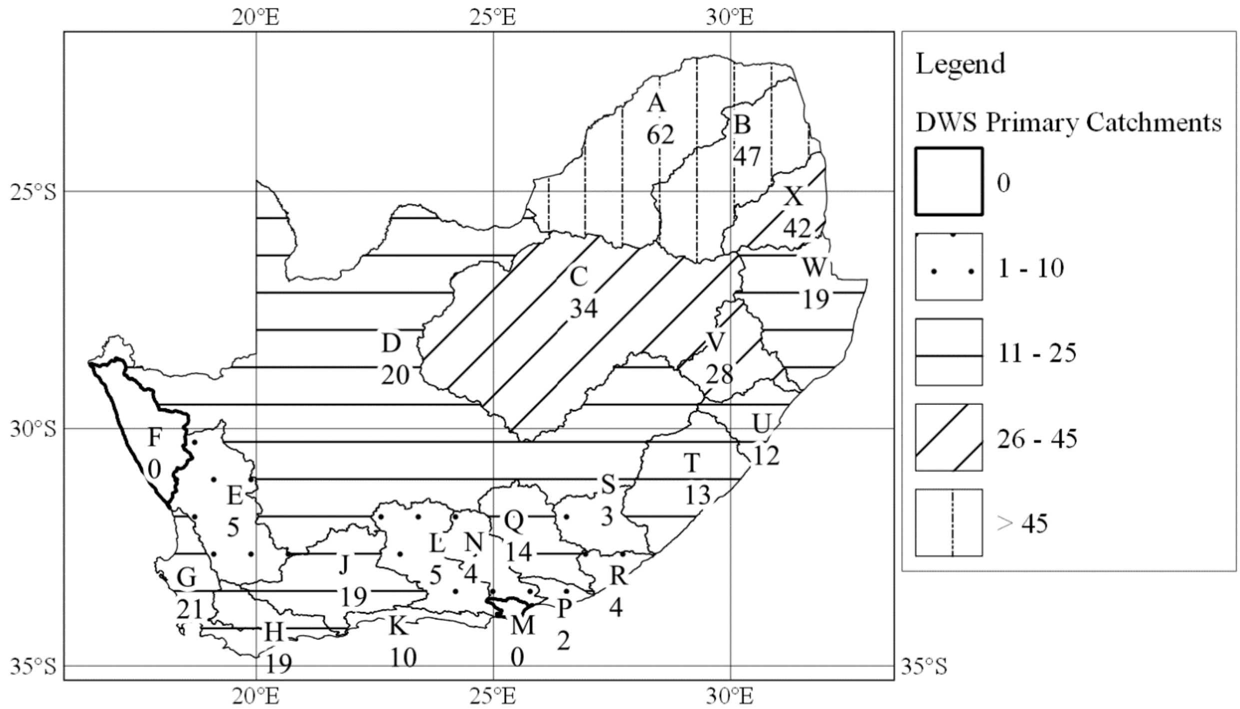


Figure 2.10 Number of gauging stations per DWS primary drainage region across South Africa

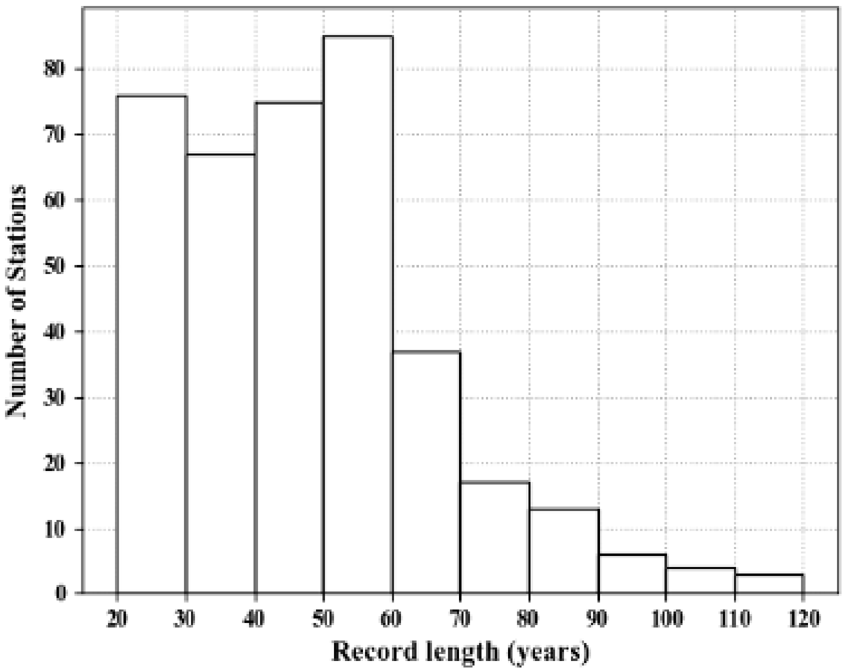


Figure 2.11 Histogram depicting the distribution of the station record lengths for the selected 383 gauging and synthetic dam stations

3 IDENTIFICATION OF A DISTRIBUTION SUITABLE FOR AT-SITE FLOOD FREQUENCY ANALYSIS IN SOUTH AFRICA

3.1 Abstract

Selection of a probability distribution is a critical part of FFA, and potentially affects the estimated magnitudes of the estimated design floods. In South Africa, the LP3, three parameter LNO, GEV, PE3 and GPA distributions are advocated in literature. However, only a few of these recommendations are based on scientific investigation using limited datasets, the remainder rely on subjective experience by practitioners and international studies. In this study, the three parameter Kappa (KAP3) distribution has been included to identify whether the inherent flexibility can describe the South African hydrological conditions. Statistical approaches are utilised to identify the suitability of the distributions for use in South Africa. Goodness-of-Fit (GoF) measures favoured the LP3 method; however, model selection criterion and graphical methods favoured the GPA. The final recommendation was based on the predictive ability of the models, which takes into account the uncertainty associated with estimates derived from the LP3 and GPA distributions. Utilising bootstrapping it was identified that the GPA distribution provided narrower uncertainty bands and is therefore recommended. Future work will focus on verifying the suitability of the GPA on a homogeneous region level.

3.2 Introduction

South Africa experience considerable negative economic impact of floods. For example, Holloway *et al.* (2010) reported flood losses of R1 billion in the Western Cape in 2008, and Davies (2016, 2017) reports that between 10 000 and 15 000 people were displaced in 2016 in the Western Cape alone. These impacts highlight the ongoing challenges for improving flood management in South Africa. Engineers rely on hydrological information, e.g., rainfall and streamflow data, for the design of hydraulic structures, such as dams, bridges and drainage culverts (Maidment, 1993), and to guide spatial planning more generally through risk maps showing areas at risk of inundation during extreme events. The design is guided by the anticipated frequency (e.g. return period or AEP) and magnitude of future floods in the form of design flood estimates.

When adequate streamflow data are available at a site, the relationship between frequency and magnitude of future floods can be established using FFA involving the fitting of a probability distribution to an AMS of peak flow events. The probability distribution selected form the sample of available observed events is then assumed to be the best distribution to fit the entire population of events can be used for estimating the magnitude of a flood that will be exceeded with a specified probability each year, e.g. an AEP of 1% (Pilgrim and Cordery (1993). In South Africa, the Road Drainage Manual (SANRAL, 2013) is considered as one of the authoritative guidance documents on FFA, while procedures applied by the DWS are summarised in a manual by Van der Spuy and Rademeyer (2018). A key aspect of FFA is the identification of a suitable distribution that can describe the probabilistic behaviour of the available flood data.

The choice of distribution can have a considerable impact on the estimated peaks. For example, Alexander (2002a) demonstrated that the design flood estimate of a 0.5% AEP flood obtained using different distributions fitted to the same data set could result in variations of up to 38%. For conducting FFA in South African, Alexander (1990, 2000) recommended using the LP3 distribution. Gericke (2010) proposed that the best distributions for use in South Africa are the three parameter LNO, LP3 and GEV. Van der Spuy and Rademeyer (2018) describe the LNO, LP3 and GEV distributions as the most suitable distributions for FFA but provide no evidence to support these. Görgens (2007b) used both the LP3 and GEV distribution in South Africa, simply stating that the methods are commonly used in practice, and no further motivation for their use is provided. Görgens (2007a) found that the LP3 distribution showed significant variation in its estimation, whereas the GEV provided improved results. In addition to these distributions, Haile (2011) found that the GPA, LNO and PE3 distributions were the best suited distributions in South Africa. However, Haile (2011) only utilised 73 flow-gauging stations within South Africa, where the DWS currently has 1458 registered river gauges. Kjeldsen *et al.* (2002) found that the infrequent occurrence of very extreme events resulting from cyclone activity in the coastal region of KZN resulted in poor performance of standard distributions. However, for the inland region of KZN, the GNO, PE3 and GPA distributions were all suitable candidates. Mkhandi *et al.* (2000) reviewed seven distributions and two parameter estimation methods in southern Africa for their descriptive ability, through the use of LM Ratio Diagrams, and their predictive ability, based on Monte Carlo simulated bias. Mkhandi *et al.* (2000) found that the predictive ability test identified the PE3 as most suitable in 12 of the 13 regions considered, with LP3 being most suitable for the last remaining region, whereas the descriptive ability was divided between the PE3, LNO and GPA distributions and were favoured in six, six and two regions respectively. Mkhandi *et al.* (2000) based the final recommendations on the predictive ability results, however the approach adopted used the average percent bias of the 1%, 0.5% and 0.2% AEP for simulated records with lengths ranging from 15 to 50 years. The majority of the tests therefore exceeded the general "rule-of-thumb" by extrapolating more than twice the available record length. Conversely, Zhang *et al.* (2019) adopted the recommendation by Robson and Reed (1999) when assessing distributions for use in Canada, that a minimum of two times the record length is required for estimates, e.g. the 1% AEP requires 200 years of data, to gauge the predictive ability.

Internationally, numerous scientific studies have been undertaken to validate and substantiate the selection of suitable flood distributions, primarily in Europe, USA, and Australia. Although the hydrological climates and responses vary significantly from prevailing conditions in South Africa, experience can still be drawn from the studies.

Castellarin *et al.* (2012) compiled an inventory of streamflow data and statistical methods used for FFA across Europe. The study compiled data received from 17 countries, which includes distribution selection, FFA and regional FFA procedures. Across Europe, a number of different distributions are recommended, including: GEV, GPA, LP3, LNO, PE3, GUM, Weibull (WEI) and Two Component Extreme Value (TCEV). Salinas *et al.* (2014) investigated the applicability of the GEV distribution as a pan-European distribution and found that the GEV cannot fully describe the differences in flood series characteristics between catchments. However, not enough statistical evidence was found to reject the hypothesis for general applicability of the GEV. Kjeldsen *et al.* (2017) tested the application of the four parameter Kappa (KAP) distribution at a regional scale in the UK, motivated by the fact that several of the commonly used three parameter distributions are special cases of the KAP distribution (Hosking, 1994). Kjeldsen *et al.* (2017) proposed the application of a national KAP distribution, by reducing the KAP distribution to a 3-parameter distribution through estimation of a national shape parameter. The KAP3 improved the description of the regional distribution in the UK compared to both the Generalised Logistic (GLO) and GEV distributions.

In the USA, two predominant studies focussed on the identification of a suitable distribution for DFE. Benson (1968) details testing performed on six different distributions (LP3, GUM, Gamma (GAM), log-GUM, LNO, Hazen), these were tested at 10 stations with record lengths ranging from 40 to 97 years. Recommendations of the methods were based on deviations between design flood estimates, as opposed to statistical methods. From the distributions reviewed, the LNO, LP3 and Hazen methods resulted in the smallest deviations and bias. The LP3 was, however, recommended based on popularity of use, the use of a skew parameter thus increasing its flexibility, and its rigorous mathematical backing. Six years later Beard (1974) tested eight distributions at 300 sites and the two distributions deemed to preform best were the LNO and LP3 with a regional skew (LPR). Apart from these studies there has been little further investigation into the selection of an appropriate distribution in USA. Emphasis has rather been placed on improving the moment estimations for use with the LP3 through the use of moment adjustments as presented by Cohn *et al.* (2013).

South-eastern Australia presents the most climatologically similar region to South Africa. Haddad and Rahman (2008) investigated the performance of 12 distributions and fitting combinations at 18 sites in South-East Australia and concluded that GPA with the use of Linear moment (LM) fitting (Hosking, 1990) (GPA-L) and the GEV with the use of LH-moments (LHM), a generalisation of LM (Wang, 1997), fitting provided the best fits to the data, which was not consistent with the recommendations of the 1987 ARR manual (ARR, 1987). Haddad and Rahman (2011) revisited the assessment of distribution selection in Tasmania to possibly modify the selection criteria. They considered seven distributions and identified that the most suitable model for use in Tasmania was the LNO distribution combined with Bayesian Markov Chain Monte Carlo (MCMC) fitting. The climate and hydrological responses in Tasmania are, however, different to conditions in most parts of South Africa. In the latest revision of the ARR guidelines, it is noted that the GEV and LP3 are reasonable initial choices for FFA, and it is recommended that a single distribution is not prescribed due to the potential sampling variability of the relatively short record lengths (Rahman *et al.*, 2019). In addition, it is recommended that a review of the data at a regional scale can be used to identify the best fit distribution through the use of an LM diagram (Rahman *et al.*, 2019).

From the above summary of the practical and scientific literature it is evident that South African hydrological practise is not well aligned with recent findings of scientific studies regarding choice of flood distributions, and recommendations for particular distributions have generally not been based on any substantial evidence.

This paper aims to identify the distribution most suitable for FFA at an at-site scale taking into consideration the existing recommendations for South Africa. A review of the FFA procedures as applied in South Africa is provided, and through the application of statistical analyses, the most suitable flood distribution applicable to South African streamflow records is identified. This will potentially enhance the credibility of design flood estimates and also provide a more reliable benchmark against which the performance of other methods of estimating design floods at the site could be evaluated.

3.3 Flood Frequency Analysis

FFA requires the selection and fitting of a distribution to a series of data either graphically or analytically (Stedinger *et al.*, 1993, Basson and Pegram, 1994, Smithers and Schulze, 2000a, Alexander, 2002a, Smithers and Schulze, 2003, Gericke, 2010, SANRAL, 2013, Van der Spuy and Rademeyer, 2018). The series of data can consist of AMS or Partial Duration Series (PDS). Karim *et al.* (2017) highlights the main differences between AMS and PDS and notes that PDS requires the investigation of additional complexities, such as the selection of an appropriate threshold. Given the additional complexities inherent with PDS, and in the interest of computational convenience and the assumption of independence between flood events AMS have been utilised in this study. The design value computed from the observed data is then assumed to be the best estimate of the design flood at the site and the performance of the other methods of estimating design floods at the site can be assessed using the design flood estimated from the observed data. In order to perform FFA the following aspects need to be considered:

(a) Selection of a parameter estimation method.

(b) Identification of applicable distributions.

(c) Selection of most suitable distribution.

In this study an investigation was undertaken to identify the most suitable distribution for FFA in South Africa, applying four categories of model selection methods to judge the quality of distribution fits: (i) graphical methods, (ii) GoF tests, (iii) model selection criterion, and (iv) predictive ability. Selection of a suitable distribution is a task that is often subjective, even though, there are numerous methods available to assess the quality of fit for different distributions. Similarly, validation of the appropriateness of the selected distribution is often difficult as multiple distributions may statistically fit the data but may appear less well suited when interpreted graphically and result in vastly different estimates of high return period floods.

3.3.1 Parameter estimation methods

Some of the methods available for parameter estimation include: Method of Moments (MM), LM (Hosking, 1990), LHM (Wang, 1997), Probability Weighted Moments (PWM) (Greenwood *et al.*, 1979) and Maximum Likelihood procedure (ML) (R.A. Fisher, 1912 as referenced in Aldrich, 1997). In South Africa, Görgens (2007b) used both the MM and PWM methods, whereas Alexander (2002b), SANRAL (2013) and Van der Spuy and Rademeyer (2018) recommend the use of MM, which is sensitive to the presence of outliers in the data (Bastianin, 2020).. England *et al.* (2018) prescribes the use of MM in simple cases, where data are not censored, and where data censoring is present the Expected Moment Algorithm (EMA) is recommended in the USA. In Australia, a number of studies were undertaken to identify both the best fit distribution and best fitting procedure. The most notable study was undertaken by Haddad and Rahman (2008), who reviewed twelve distribution/fitting combinations and the LP3, Normal (NOR), LNO, GUM, GEV and GPA were fitted using the MM, LM, LHM and Bayesian Maximum Likelihood (BML) fitting procedures and identified that the three top performing combinations are GPA-LM, GEV-LHM and LP3-BML.

The method of LM (Hosking and Wallis, 1993) is a parameter estimation technique which has gained in popularity and proven successful both locally (e.g. Mkhandi and Kachroo, 1997, Kjeldsen *et al.*, 2002, Smithers and Schulze, 2003, Haile, 2011) and internationally (e.g. Pearson, 1991, Vogel *et al.*, 1993, Zrinji and Burn, 1996, Chen *et al.*, 2007, Borujeni and Sulaiman, 2009, Castellarin *et al.*, 2011, Hassan and Ping, 2012, Rutkowska *et al.*, 2016, Cassalho *et al.*, 2018, Mostofi Zadeh and Burn, 2019). The LM (λ_r) and LM ratios (τ_r), of order *r* and the estimation procedures are detailed in Eqs. 3.1 to 3.3, using an observation dataset *X* of length *n*, with an expected value *E(X)*.

$$\lambda_r = r^{-1} \sum_{j=0}^{r-1} (-1)^j \binom{r-1}{j} E\left(X_{r-j:r}\right) \tag{3.1}$$

$$\tau_r = \frac{\lambda_r}{\lambda_2}, r = 3, 4, \dots \tag{3.2}$$

$$\tau_1 = \frac{\lambda_2}{\lambda_1} \tag{3.3}$$

In addition, the LM technique is theoretically superior to the MM due to lower weighting being applied to the larger values within the dataset and LM are therefore more robust for use in the presence of high outliers (Hosking, 1990). Given that South African flood hydrology is highly variable, which often results in the existence of potential outliers in the datasets LM was therefore adopted for use.

3.3.2 Probability distributions

As is evident from a review of the literature of studies undertaken both in South Africa and internationally, there are numerous differences between the recommendations and applications of distributions and the parameters estimation methods for FFA. The distributions most commonly recommended in literature for use in South Africa are the GEV and LP3. However, the wide-spread use of the GPA and PE3 internationally and the findings by Kjeldsen *et al.* (2002) and Haile (2011), substantiate further investigation into the adoption of these distributions within South Africa. In addition, Kjeldsen *et al.* (2017) provided a methodology to utilise the KAP distribution on a regional scale by determining a regional h shape parameter value, reducing the four-parameter distribution to the three-parameter KAP3.. Eight other distributions are special cases of the KAP distribution for fixed values of the h and k parameters, including the GPA (h = 1), GEV (h = 0) and GLO (h = -1) distributions (Hosking, 1994). The additional flexibility may be better suited to describe the hydrological variability of flood series across the contrasting geographic and climatological regions of South Africa. Thus, the KAP3 distribution utilising a national record length weighted mean h value was included in the assessment.

Hence a total of five distributions (GEV, GPA, KAP3, LP3 and PE3) distributions, shown in Table 3.1, with parameters fitted using LM were assessed to identify the most suitable general distribution for DFE in South African.

It is important to note that this study adopted the use of a T% AEP to represent the probability of flood event occurrence and is related to the Average Recurrence Interval (ARI), where AEP is equivalent to the inverse of ARI.

Table 3.1 Cumulative distribution functions of the statistical distributions selected for assessment

Distribution	Cumulative Distribution Function	Parameters
KAP	$F(x) = \left\{1 - h[1 - k(x - \xi)/\alpha]^{\,1/k}\right\}^{1/h}$	α = scale parameter, ξ th location parameter, and k and h are shape parameters
GPA	$F(x) = [1 - k(x - \xi)/\alpha]^{\,1/k}$	
GEV	$F(x) = exp\{-[1 - k(x - \xi)/\alpha]^{1/k}\}$	
PE3 and LP3	$F(x) = \begin{cases} \frac{G\left(\alpha, \frac{x-\xi}{\beta}\right)}{\Gamma(\alpha)}, \ \gamma < 0 \\ 1 - \frac{G\left(\alpha, \frac{\xi-x}{\beta}\right)}{\Gamma(\alpha)}, \ \gamma \geq 0 \end{cases}$ When $\gamma \neq 0$ let: $\alpha = \frac{4}{\gamma^2}, \beta = 0.5\sigma\lvert\gamma\rvert, \xi = \mu - \frac{2\sigma}{\gamma}$	μ = mean, σ = standard deviation, γ = skewness, G = incomplete gamma function, and Γ = gamma function

3.3.3 Selection of the most suitable distribution

The four most common methods for the selection of a distribution found in the literature was: (i) graphical methods (e.g. Mkhandi *et al.* (2000), Haddad and Rahman (2008)), (ii) GoF tests (e.g. Laio (2004)), (iii) model selection criterion (e.g. Laio *et al.* (2009), Haddad and Rahman (2011)) and, (iv) predictive performance (e.g. Mkhandi *et al.* (2000), Zhang *et al.* (2019)). The most widely used approach for selection of distribution types in South Africa are graphical methods, whereas the literature indicates that internationally graphical, GoF and model selection criterion are commonly applied in combinations. In this study, the model uncertainty, defined by the 95% confidence bands, was used as an additional selection criterion to complement the assessment of the predictive performance. To identify the 95% confidence bands for each of the distributions the bootstrapping methodology was selected as it is not reliant on the assumption that the sample parameters represent the population, and that no distribution needs to be assumed *a priori*. As a final consideration, the data utilised was divided into four distinct sets based on record lengths: (i) less than 40 years of data, (ii) greater than or equal to 40 years, but less than 60 years, (iii) greater than or equal to 60 years, but less than 80 years, and (iv) greater than or equal to 80 years. Given the use of AMS' as short as 20 years in the analysis the record length categories were created to identify whether the length of record available affects the distribution selections. In theory the longer record lengths have a higher probability of representing the true underlying distribution.

3.3.3.1 Graphical methods

Graphical methods are often employed to identify the flood distributions that are least suitable reducing the number of distributions to consider in further detail. The simplest graphical test is the use of plotting positions for observed data on an at-site basis. The observed data are plotted against the calculated distributions to provide a graphical comparison of the distribution to the observed data. The plotting positions identified by DWS (Van der Spuy and Rademeyer, 2018) and SANRAL (2013) are the Weibull, Blom, Gringörten, Cunane, Beard and Greenwood methods. SANRAL (2013) describe these in further detail. Bulletin 17C (England *et al.*, 2018) proposes the use of the plotting positions described by Stedinger *et al.* (1993).

Product Moment Diagrams (PMD) are an additional graphical measure that can be used, however, Vogel and Fennessey (1993) recommend the use of LM Ratio Diagrams (LMRDs) as developed by Hosking (1990) in favour of PMDs due to the LM being nearly unbiased. LMRDs have become a common method for the identification of best fit regional flood distribution and have been used by numerous authors for this purpose (e.g. Vogel *et al.*, 1993, Zafirakou-Koulouris *et al.*, 1998, Peel *et al.*, 2001, Castellarin *et al.*, 2012, Salinas *et al.*, 2014, Kjeldsen *et al.*, 2017). LMRDs are constructed by plotting the L-kurtosis (τ_4) versus the L-skew (τ_3).

Generally the assessment of the most suitable distribution is undertaken using two methods: (i) plotting the mean of the LMs of the region, and (ii) plotting a best fit line and comparing the result to the theoretical distributions for a set of standard three-parameter distributions (GLO, GEV, GNO, LP3). Kjeldsen and Prosdocimi (2015) proposed a modification to the use of LMRD, referred to as KP test hereafter, by applying the assumption that τ_3 and τ_4 share a bivariate normal relationship which allows for the derivation of a 90% confidence ellipse. The confidence ellipse identifies the suitable distributions for the estimated LMs, and a selection is then undertaken through the estimation of the Mahalanobis distance for each distribution.

3.3.3.2 Goodness-of-fit tests

The purpose of a GoF test is to identify, in a statistical manner, the most suitable distribution for the data being fitted. Zeng *et al.* (2015) reviewed the Chi-squared (CS) (Pearson, 1900), Kolmogorov-Smirnov (KS) (Massey, 1951) and Anderson-Darling (AD) (Anderson and Darling, 1952) GoF tests for use in FFA considering the PE3, Uniform, GNO and Weibull distributions. Zeng *et al.* (2015) concluded that most powerful GoF tests for the PE3, GNO and Weibull are the AD, KS and AD. Haddad and Rahman (2008) applied two additional GoF tests, Cramer von-Mises (CvM) (Cramér, 1928, von Mises, 1928) and the Filliben Correlation Coefficient (FCC) test (Filliben, 1975). Laio (2004) tested the power of the AD, CS, CvM, KS, FCC and LM based GoF tests for the GUM, WEI, GNO, GEV, GAM, LNO, and LP3 distributions. For the GEV, GAM and GUM distributions the power of the GoF tests were consistently below 50%, whereas the AD and CvM had power exceeding 80% for the LP3 and LNO distributions. The variation in the power of the GoF tests can be attributed to the fact that the tests apply larger weighting to different components (tail, head or entire curve) of the distribution functions (Kottegoda and Rosso, 2008) and it is therefore recommended that multiple GoF tests be considered simultaneously. AD applies additional weighting to the tails of distributions, favouring the higher or lower observations, whereas CvM weights the centre of the distribution more heavily. Lastly KS can be considered an intermediate test between AD and CvM and weights the entire distribution more evenly. Sinclair *et al.* (1990) noted that when applying FFA, the upper tail of distributions are of more importance, highlighting the use of the 0.99 quantile, or the 1% AEP, for the design for hydraulic structures. Sinclair *et al.* (1990) therefore proposed two modifications to the AD test that place emphasis on the upper or lower tails of the distribution. The upper tails are of particular importance for flood estimation, whereas the lower tails may conversely be beneficial for the analysis of droughts. The GoF tests listed above are generally applied on an at-site scale, but when reviewing regional data, can be used to identify the distribution through identifying the percentage of sites that are accepted for each test (Haddad and Rahman, 2008, Ul Hassan *et al.*, 2019).

The Chi-Squared test, Eq. 3.4, is a measure of the difference between the observed (O) and the expected (E) frequencies of ordered observations (x_i, …, x_n) in a sample of size n. The KS test shown in Eq. 3.5 measures the GoF, in relation to a distribution with a parameter vector θ, through the maximum variance between the hypothetical ($F(x_i, \theta)$) and Empirical Distribution Functions ($F_n(x)$).

$$X^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \tag{3.4}$$

$$KS = max_x |F_n(x) - F(x_i, \theta)| \tag{3.5}$$

Alternatively, quadratic statistics, Eq. 3.6, can be utilised, from which the AD, CvM and upper tail modified AD test (AU), shown in Eqs. 3.7, 3.8 and 3.9 respectively, are derived (Cramér, 1928, von Mises, 1928, Anderson and Darling, 1952).

$$Q^2 = n \int_{all\ x} [F_n(x) - F(x_i, \theta)]^2 \Psi(x) dF(x) \tag{3.6}$$

$$AD = -n - \frac{1}{n} \sum_{i=1}^{n} \left[\left[F(x_i, \theta) - \frac{2i-1}{2n} \right] + (2n + 1 - 2i) ln[1 - F(x_i, \theta)] \right] \tag{3.7}$$

$$AU = \frac{n}{2} - 2 \sum_{i=1}^{n} F(x_i, \theta) - \sum_{i=1}^{n} \left[2 - \frac{2i-1}{2n} \right] \ln [1 - F(x_i, \theta)] \tag{3.8}$$

$$CvM = \sum_{i=1}^{n} \left[F(x_i, \theta) - \frac{2i-1}{2n} \right]^2 + \frac{1}{12n} \tag{3.9}$$

where $\Psi(x)$ is a weighting function, which is 1 for CvM and $[F(x_i, \theta)(1 - F(x_i, \theta))]^{-1}$ for AD.

Hosking and Wallis (1993) also provide a regional GoF measure, *Z*, shown in Eq. 3.10. The test statistics *Z* is a measure of the difference between regional sample ($\bar{t}_4$) and theoretical L-kurtosis (τ_4^D), in relation to the standard deviation of theoretical L-kurtosis (σ_4) estimated using Monte-Carlo simulations. An absolute value of less than 1.64 signifies a suitable distribution, and the distribution (Dist) with the lowest *Z* is often accepted.

$$Z^{Dist} = \left(\bar{t}_4 - \tau_4^{Dist} \right) / \sigma_4 \tag{3.10}$$

The test relies on the assumption that the regional values used are from a homogeneous region. Kjeldsen and Prosdocimi (2015) leverage the assumption that the joint distribution of the L-skew and L-kurtosis is normally distributed to develop confidence ellipses for the selection of candidate distributions. In addition, the distribution that minimises the Mahalanobis distance (D_m), shown in Eqs. 3.11 to 3.13, is deemed to be the best fit. Kjeldsen and Prosdocimi (2015) proved, through undertaking a predictive power test using Monte Carlo simulations, that the new method improved on the results achieved by Hosking and Wallis (1993) even where sites are moderately correlated.

$$U_m = \frac{1}{n}\sum_{i=1}^{n} U_i \tag{3.11}$$

$$S = \frac{1}{n-1}\sum_{i=1}^{n}(U_i - U_m)\,(U_i - U_m)^T \tag{3.12}$$

$$D_m = \sqrt{(U_i - U_m)^T S^{-1}(U_i - U_m)} \tag{3.13}$$

The measure uses the relative proximity of an individual site, *i*, relative to the remaining sites by comparing the site specific LM vectors (U_i) with the regional mean matrix (U_m) and the covariance matrix *S*.

3.3.3.3 Model selection criteria

Laio *et al.* (2009) investigated the use of model selection criterion for use with FFA, which was also adopted by Haddad and Rahman (2011). The criterion chosen were the Akaike Information Criterion (*AIC*, Eq 3.14) (Akaike, 1974), second order *AIC* (*AICc*, Eq 3.15) (Sugiura, 1978), Bayesian Information Criterion (*BIC*, Eq 3.16) (Schwarz, 1978), and a modified Anderson-Darling Criterion (ADC) (Laio *et al.*, 2009). The ADC requires distribution dependent coefficients to be applied, however, these parameters have only been derived for seven of the more commonly used hydrological distributions and has therefore not been included in this study. Model criterion consider relative fit of models to data, of sample size *n*, by measuring the information lost in the process of fitting through the likelihood function ($L(\vartheta)$). Models (*j*) are penalised for the number of parameters (*p*) utilised, and as such a lower value indicates a better model fit.

$$AIC_j = -2ln\left(L_j(\hat{\vartheta})\right) + 2p_j \tag{3.14}$$

$$AICc_j = AIC_j + \frac{2{p_j}^2+2p_j}{n-p_j-1} \tag{3.15}$$

$$BIC_j = -2ln\left(L_j(\hat{\vartheta})\right) + \ln(n)p_j \tag{3.16}$$

The criterion is applied at an at-site level to identify the distribution that provides the best model fit per site. The at-site results are summarised at a regional scale by calculating the percentage of sites where each distribution provides the best fit, thus providing an indication of the regional distribution.

3.3.3.4 Predictive ability

The predictive ability of the distributions was based on the "true" fit as described by Zhang *et al.* (2019) and the uncertainty associated with the distributions. Zhang *et al.* (2019) define the "true" fit of data as the Gringorton plotted AMS and used the 4% AEP for sites with record lengths exceeding 50 years as a test of the predictive ability. The 5% AEP was chosen for this study for sites having record lengths of 50 years or greater, exceeding the minimum requirement of two times the AEP as suggested by Robson and Reed (1999).

In the application of FFA, it is generally assumed that the data being used, after pre-processing of the data, are free of errors. In contrast the model errors can be quantified and are represented by the error introduced by misrepresentation of the actual events by the fitted distribution. Hence it is assumed that the sample data are accurate and that the selected distributions introduce uncertainty into the estimates, which can be determined. Typical approaches used to determine the uncertainty associated with distributions include the following: (i) Analytical methods (e.g. Kjeldsen and Jones, 2006); (ii) Monte Carlo simulations (e.g. Silva *et al.*, 2012); and (iii) Bootstrapping (e.g. Burn, 2003).

Taylor approximation is an analytical method that attempts to approximate non-linear functions with a linear function within a set of known parameters. The performance is linked with the degree of non-linearity of the function in question and a critical assumption for its use it that the known parameters are true reflections of the population parameters.

Similarly, Monte Carlo (MC) simulations assume that the parameters estimated from the sample are a true reflection of the population parameters. However, instead of undertaking an analytical approach, a resampling approach is used. MC simulations resample from a known distribution and generate a number of iterations of random samples with each sample containing the same record length as the original sample. The T% AEP runoff event is then generated for each of the samples generated and the variance in relation to the original dataset calculated and used to estimate the confidence bands.

Bootstrapping refers to a resampling method where *N* number of iterations are considered to identify the variation in estimates, whereby the confidence intervals (uncertainty) can be determined. To determine the variation of a *T*% AEP event, a synthetic record is created from the observed record using resampling with replacement. This process involves the random selection of flood events from the observed records until the synthetic record length matches the observed record length, and FFA of the synthetic record is performed, which is considered a single iteration. Bootstrapping can be undertaken in a balanced or an unbalanced approach. The balanced approach ensures that the mean of the overall sample set is maintained as each sample can only be reproduced *N* number of times, whereas for unbalanced bootstrapping no limitation is applied to the number of occurrences of any sample. When considering 5% and 1% confidence intervals, it is recommended to use *N*=1 000 and *N*=10 000, respectively.

3.4 Results

Given the large number of methods available to identify the suitability of a distribution and the recommendations that all methods be applied, the distributions are to be tested using an iterative approach as indicated in Figure 3.1. The GoF tests being applied are provided in Table 3.2. Given the fact that the graphical methods rely on the theoretical LMRD's to determine the selection of the distributions, it will be used as the initial test to identify the three most suitable methods, after which the GoF, model fit criterion and uncertainty analyses are to be undertaken. The graphical approach will assess the selected distributions, based on the assumption of there is an overarching parent distribution that is most suitable for general use in South Africa, thereafter the at-site analysis will refine the selection through identifying the percentage of sites at which these distributions are deemed suitable. The at-site analysis will rank each of the three most suitable distributions based on the performance in each test. The ranks are combined to identify the most suitable distribution for use, with the distribution scoring the lowest total rankings being most preferred distribution. Prior to undertaking the testing, the National Kappa *h* value was determined for use with the KAP3 distribution.

Table 3.2 List of test categories and the selected methods for selection of an appropriate Distribution

Test Category	Methods
Graphical	KP method
Goodness-of-fit	Modified Anderson-Darling (AU), Chi Squared (CS), Cramer von Mises (CvM), Kolmogorov Smirnov (KS)
Model Selection Criterion	Akaike (AIC), second order Akaike (AICc), and Bayesian Information Criterion (BIC)
Model Uncertainty	Balanced bootstrapping

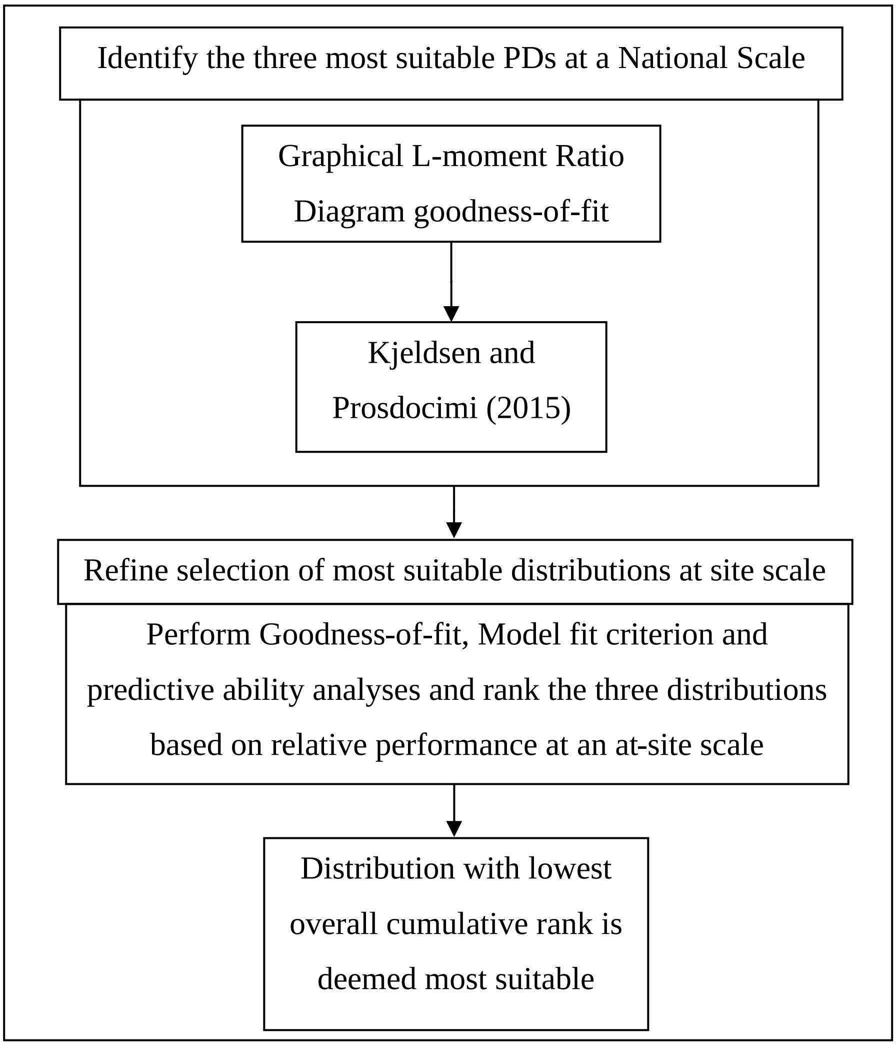


Figure 3.1 Methodology flow diagram for the selection of a suitable distribution

3.4.1 National Kappa *h* Value Estimation

Prior to undertaking the selection process, the estimation of a national KAP3 distribution *h* value was required. Figure 3.2 shows the LM diagram with the KAP *h* value contours, ranging from -1 to 1 at 0.25 intervals, and the record length weighted mean L-skew and L-kurtosis values. From the mean coordinates the *h* value was estimated to be 0.77. Having established the *h* value at a national scale, the number of distributions being assessed increased to five: (i) GEV, (ii) GPA, (iii) KAP3 (*h* = 0.77), (iv) LP3, and (v) PE3.

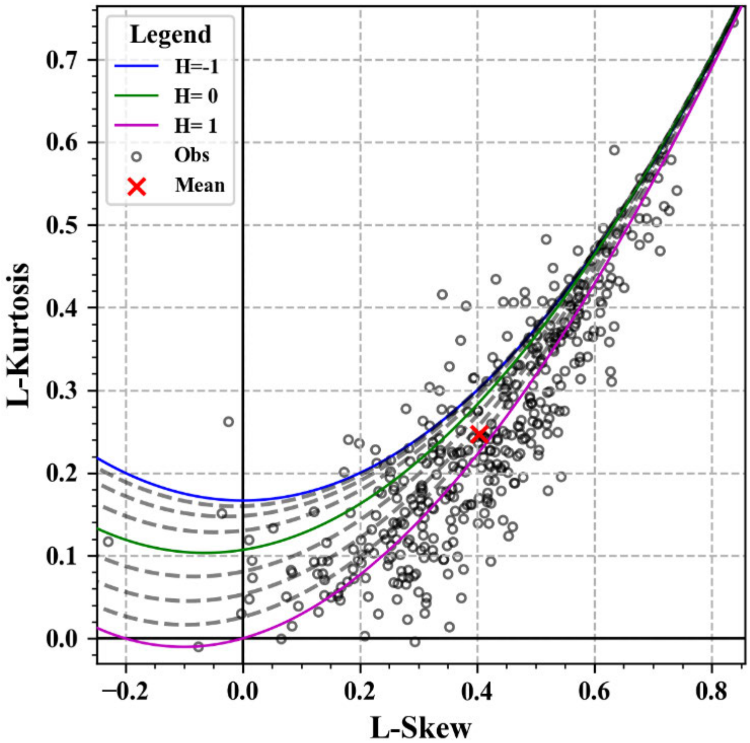


Figure 3.2 LM ratio diagram for all 383 selected sites using untransformed data indicating the KAPPA distribution *h* value contours at 0.25 intervals and the record length weighted mean L-skew and L-kurtosis.

3.4.2 Graphical methods

When considering the graphical approach, all distributions considered were ranked for the performance at a national scale. When considering the graphical data representation using LM diagram, as shown in Figure 3.3, the most suitable regional distributions appear to be the LP3, GPA or KAP3 as the data points tend to cluster around the theoretical lines representing these distributions. Table 3.3 provides the geometric rank of the GEV, GPA, KAP3, LP3 and PE3 distributions based on the KP test (Kjeldsen and Prosdocimi, 2015). The ranked order of selection is the KAP3 (h = 0.77), GPA, LP3, GEV and PE3 distributions.

The average LM within each DWS drainage region is shown in Figure 3.4, and it is evident that the regional averages are largely clustered around the GPA and KAP3 theoretical distributions for untransformed data, whereas the variation of the log transformed regional averages are grouped around an L-skew of -0.6 around the theoretical LP3. Figure 2.10 provides the geographic locations of the DWS drainage regions for reference.

Given that the KAP3, GPA and LP3 distributions are shown to be the graphical best fit distributions, they were further assessed at an at-site level to identify the most suitable distribution.

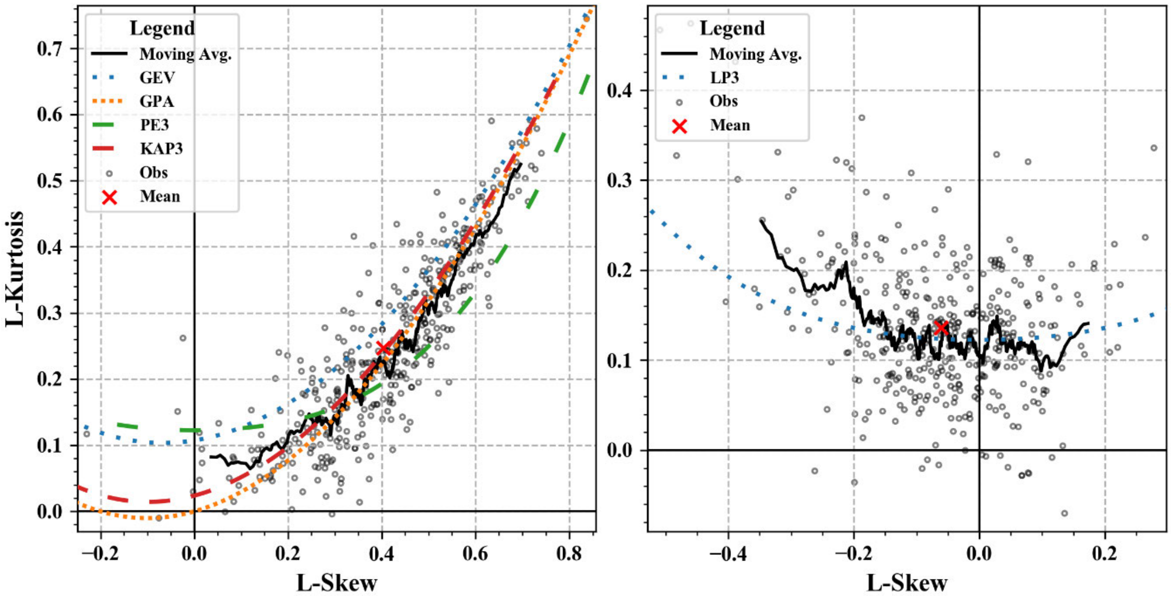


Figure 3.3 LM ratio diagram for all 383 selected sites using untransformed (left) and log transformed (right) data. The record length weighted mean (red cross) and moving average line are indicated (solid) in relation to the GEV, GPA, GNO, PE3, LNO and LP3 distributions.

Table 3.3 Rank of distributions based on the KP test in relation to the theoretical GEV, GPA, KAP3 (h = 0.77), LP3 and PE3 distributions

Distribution	Rank
KAP3	**1**
GPA	2
LP3	3
GEV	4
PE3	5

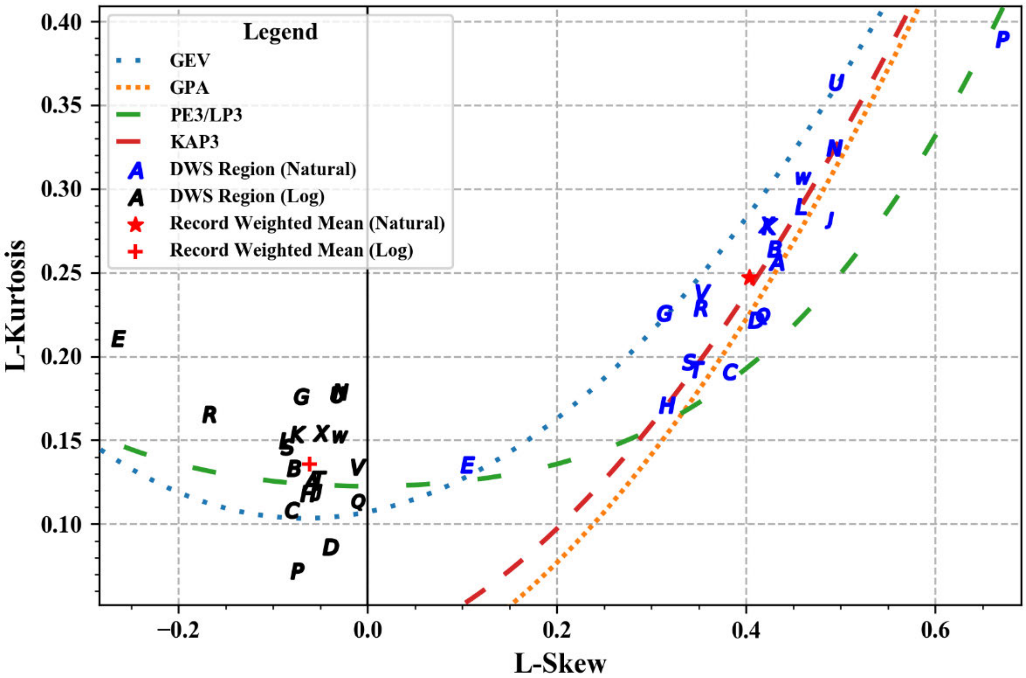


Figure 3.4 LM ratio diagram indicating the position of the record length weighted mean L-skew and L-kurtosis per DWS drainage region, represented by the relevant alphabetic numeral, for natural (blue) and log transformed (black) data as well as the record length weighted means of the entire dataset

3.4.3 Goodness-of-fit

The GoF tests applied include the AU, CS, CvM and KS tests. Table 3.4 contains a summary of the number of sites (%) that were accepted for each GoF test and distribution considered. It is evident from the GoF test acceptance that the distribution that is accepted most frequently is the LP3. However, all distributions under consideration can generally be considered suitable options for the use in South Africa.

The distinction between selection of the different the distributions varies between 1.5 and 6.1% when considering all sites and does not indicate the dominance of a single distribution for use. Even though the gap in the level of acceptance increases to 13.9% for the LP3 for sites with record lengths exceeding 80 years, this dataset is limited to 27 sites and indicates equates to 4 sites. Taking into consideration the rankings for each distribution and record length category the LP3 is deemed to have performed best, followed by the GPA and KAP3 respectively.

Table 3.4 Summary of the modified Anderson-Darling, Chi-squared, Cramer-von-Mises and Kolmogorof-Smirnov GoF test results for 383 sites in South Africa*

Record Length	Distribution	GoF Test Acceptance (% of sites)					Rank
		AU	CS	CvM	KS	Average	
All sites 20 <= x (383 sites)	GPA	83.2	78.1	92.9	97.5	87.9	2
	KAP3	**84.5**	77.1	91.6	97.5	87.7	3
	LP3	80.9	**83.2**	**96.7**	**99**	**90.0**	**1**
20 <= x < 40 (136 sites)	GPA	85.6	82.9	96.6	**100.0**	91.3	2
	KAP3	**88.4**	82.9	**97.3**	**100.0**	**92.2**	**1**
	LP3	81.5	**87.7**	**97.3**	97.9	91.1	3
40 <= x < 60 (162 sites)	GPA	80.2	78.4	93.2	96.3	87.0	2
	KAP3	**80.9**	78.4	90.7	96.9	86.7	3
	LP3	79.0	**82.7**	**96.3**	**99.4**	**89.4**	**1**
60 <= x < 80 (58 sites)	GPA	82.8	72.4	93.1	98.3	86.7	2
	KAP3	**86.2**	65.5	91.4	96.6	84.9	3
	LP3	84.5	**75.9**	**96.6**	**100.0**	**89.3**	**1**
80 <= x (27 sites)	GPA	**88.9**	63.0	70.4	88.9	77.8	2
	KAP3	81.5	63.0	66.7	88.9	75.0	3
	LP3	81.5	**77.8**	**96.3**	**100.0**	**88.9**	**1**

*Distribution with the highest acceptance rate is highlighted in bold

3.4.4 Model selection criteria

The next assessment utilised the model fit criteria: AIC, AICc and BIC, which provide selections based on the relative best fits by comparing the information lost in the model fitting procedure for each distribution. The summary in Table 3.5 show the results of the assessment divided into the four distinct record length categories. Throughout the analysis the KAP3 distribution was the least favoured distribution with a selection rate for being the best PD of 0% for all record length categories and criterion. When considering the entire dataset, the GPA is the most highly ranked distribution, however, only by 12%, which does not suggest that either of the GPA or LP3 distributions outperforms the other at a national scale. Reviewing the analysis based on the record length categories, the GPA distribution is identified as the most selected model for all sites with record lengths less than 80 years and is marginally second (3.8% or one site) for sites with record lengths of 80 years or greater. The GPA and LP3 distributions perform equally well for record lengths less than 40 years and 80 years or greater. The GPA is, however, more dominant in the medium record length ranges being favoured at up to 27.6% more sites.

Table 3.5 Summary of model criterion test selections for South Africa, the distribution with the highest selection rate for each record length category is indicated in bold

Record Length	**Distribution**	**Model Fit Criteria Selection (% of sites)**				**Rank**
		AIC	**AICc**	**BIC**	**Average**	
All sites 20 <= x (383 sites)	GPA	**56.1**	**56.1**	**56.1**	**56.1**	**1**
	KAP3	0	0	0	0	3
	LP3	43.9	43.9	43.9	43.9	2
20 <= x < 40 (136 sites)	GPA	**52.1**	**52.1**	**52.1**	**52.1**	**1**
	KAP3	0	0	0	0	3
	LP3	47.9	47.9	47.9	47.9	2
40 <= x < 60 (162 sites)	GPA	**58**	**58**	**58**	**58**	**1**
	KAP3	0	0	0	0	3
	LP3	42	42	42	42	2
60 <= x < 80 (58 sites)	GPA	**63.8**	**63.8**	**63.8**	**63.8**	**1**
	KAP3	0	0	0	0	3
	LP3	36.2	36.2	36.2	36.2	2
80 <= x (27 sites)	GPA	48.1	48.1	48.1	48.1	2
	KAP3	0	0	0	0	3
	LP3	**51.9**	**51.9**	**51.9**	**51.9**	**1**

An additional review ranking of the distributions for each criterion and record length category is shown in Table 3.6 and Figure 3.5. The ranking of distributions provides additional insight into the performance of the distributions, most prominently that the GPA was in all instances chosen as the first or second choice, whereas the LP3 and KAP3 were divided between second and third choice when the GPA was the first choice. Given these insights, the GPA can be considered the most suitable distribution followed by the LP3 and KAP3 respectively with regards to model selection criterion. An investigation into the relationship between catchment descriptors and selected distribution provided no suitable descriptor to describe the variance in selection.

Table 3.6 Model criterion test selections for South Africa indicating the percentage of sites selected per rank of the distribution and record length category.

Record Length	Rank	GPA			LP3			KAP3		
		AIC	AICc	BIC	AIC	AICc	BIC	AIC	AICc	BIC
All Sites 20 <= x (383 sites)	1	56.1	56.1	56.1	43.9	43.9	43.9	0.0	0.0	0.0
	2	43.7	43.7	43.7	23.9	26.9	34.1	32.4	29.4	22.2
	3	0.2	0.2	0.2	32.2	29.2	22	67.6	70.6	77.8
20 <= x < 40 (136 Sites)	1	52.1	52.1	52.1	47.9	47.9	47.9	0.0	0.0	0.0
	2	47.3	47.3	47.3	24.7	30.8	34.2	28.1	21.9	18.5
	3	0.7	0.7	0.7	27.4	21.2	17.8	71.9	78.1	81.5
40 <= x < 60 (162 Sites)	1	58.0	58.0	58.0	42.0	42.0	42.0	0.0	0.0	0.0
	2	42.0	42.0	42.0	21.0	22.2	31.5	37.0	35.8	26.5
	3	0.0	0.0	0.0	37.0	35.8	26.5	63.0	64.2	73.5
60 <= x < 80 (58 Sites)	1	63.8	63.8	63.8	36.2	36.2	36.2	0.0	0.0	0.0
	2	36.2	36.2	36.2	31.0	32.8	43.1	32.8	31.0	20.7
	3	0.0	0.0	0.0	32.8	31.0	20.7	67.2	69.0	79.3
80 <= x (27 Sites)	1	48.1	48.1	48.1	51.9	51.9	51.9	0.0	0.0	0.0
	2	51.9	51.9	51.9	22.2	22.2	29.6	25.9	25.9	18.5
	3	0.0	0.0	0.0	25.9	25.9	18.5	74.1	74.1	81.5

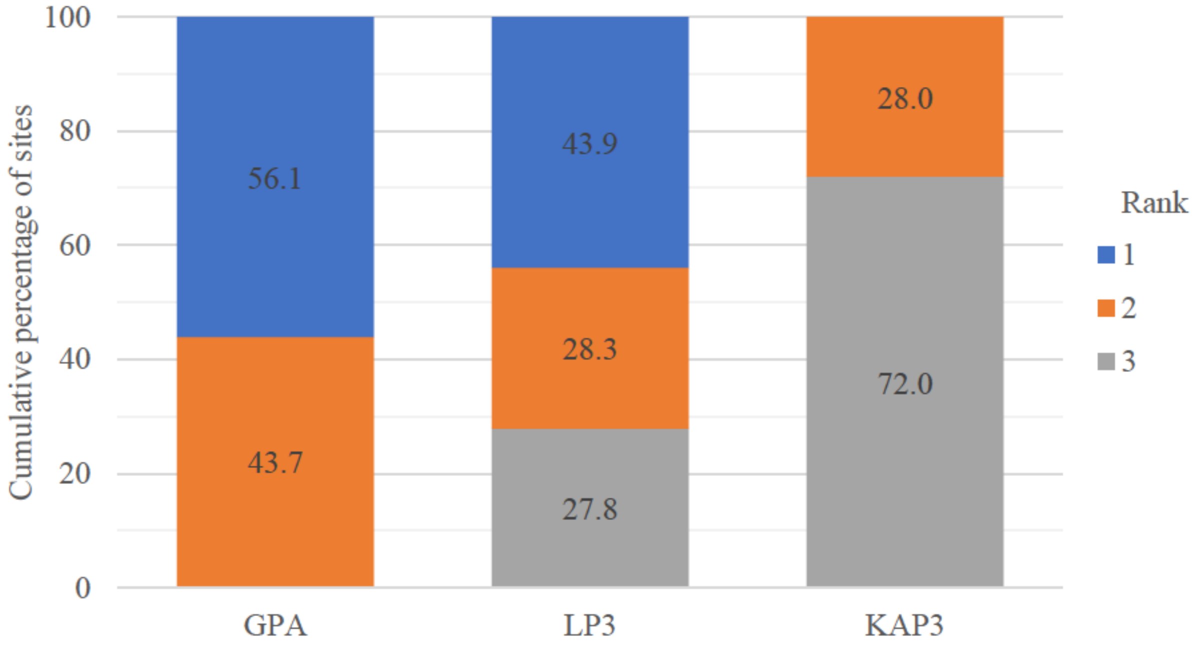


Figure 3.5 Model criterion test selections for South Africa indicating the percentage of sites selected per rank for the GPA, KAP3 and LP3 distributions when considering all sites

3.4.5 Predictive ability

The final consideration adopted for the selection of a suitable distribution was the predictive ability of the distributions and consisted of a comparison of the estimated and AMS plotted using Gringorton plotting position, dubbed the “true” fit, and the estimation of the uncertainty associated with the distributions being considered. The Gringorton plotting position was adopted due to the unbiased nature thereof (Cunnane, 1978). The comparison of the estimates to the “true” fit, shown in Figure 3.6, adopted the 5% AEP as the indicative AEP as this allowed for the use of 148 sites (39%) when only considering sites with 50 years or more of records, which exceeds the requirement prescribed by Robson and Reed (1999). From Figure 3.6 it is evident that the GPA and KAP3 distributions tend to underestimate the 5% AEP event with an interquartile range between 0.87 and 1.04 and median values of 0.97. The LP3 distribution tends to overestimate the 5% AEP with an interquartile range between 0.92 and 1.11 and a median of 1.01. These results however do not indicate a clear favourite for most suitable distribution.

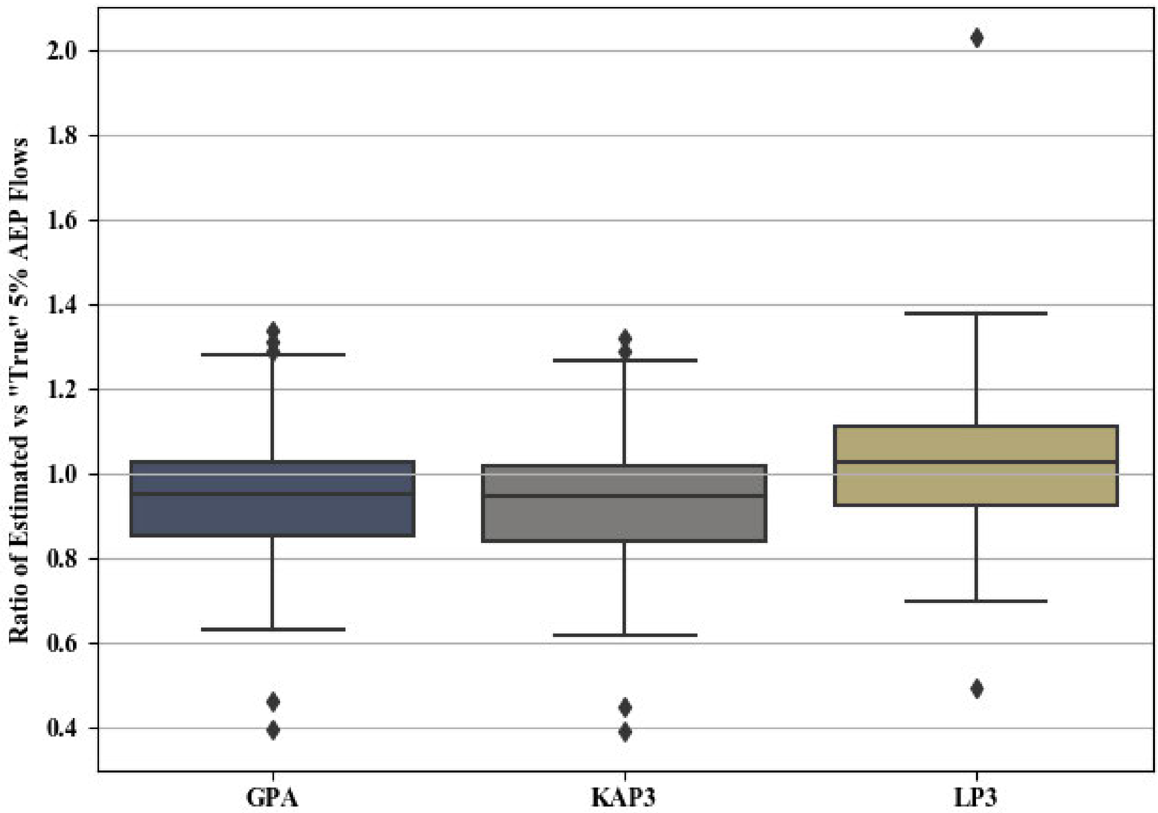


Figure 3.6 Boxplot of the estimated vs the Gringorton plotted AMS data, for the 5% AEP design flows, and the GPA, KAP3 (h=0.77) and the LP3 distributions for 148 sites with record lengths of 50 years or greater

The uncertainty associated with the distributions being considered, shown in Figure 3.7, is based on the 90% confidence limits derived using balanced bootstrapping resampling, which generated 1000 replicates for each sites considered. The variance of the confidence bands is calculated as a percent variance of the FFA of the bootstrap replicates in relation to the FFA results using the original data set. The assessment of the uncertainty associated with the distributions was based on the 1% and 5%s AEPs. Limiting the sites for the assessment to sites with record lengths of 50 years or greater allows for the use of the 5% AEP as discussed above. In addition, when applying the “rule-of-thumb” of not extrapolating beyond two times the record length, an indication of the associated uncertainty can be assessed for the 1% AEP. The 1% AEP is also of importance due to it being commonly applied in practice and being a requirement in regulatory documents in South Africa.

Figure 3.7 shown the variation of the 90% confidence bands for the GPA, KAP3 and LP3 distributions for various AEPs. Across the analysis it is evident that the GPA and KAP3 distributions perform equally well, with the KAP3 marginally outerforming the GPA for the 1 and 0.5% AEPs. Conversely, the LP3 distribution has a much higher associated level of uncertainty, particularlry for AEPs less than 10%. Considering both the predictive ability and uncertainty analyses, although the LP3 may perform better than the GPA and KAP3 when estimating the “true” fit, the associated uncertainty of the distribution brings into question whether this perfomance will be consistent for extended data sets. When considering the predictive ability and uncertainty criterion, the KAP3 has the lowest level of uncertainty, and coupling this with the estimation of the “true” fit, is the most suitable distribution, followed by the GPA and the LP3 respectively.

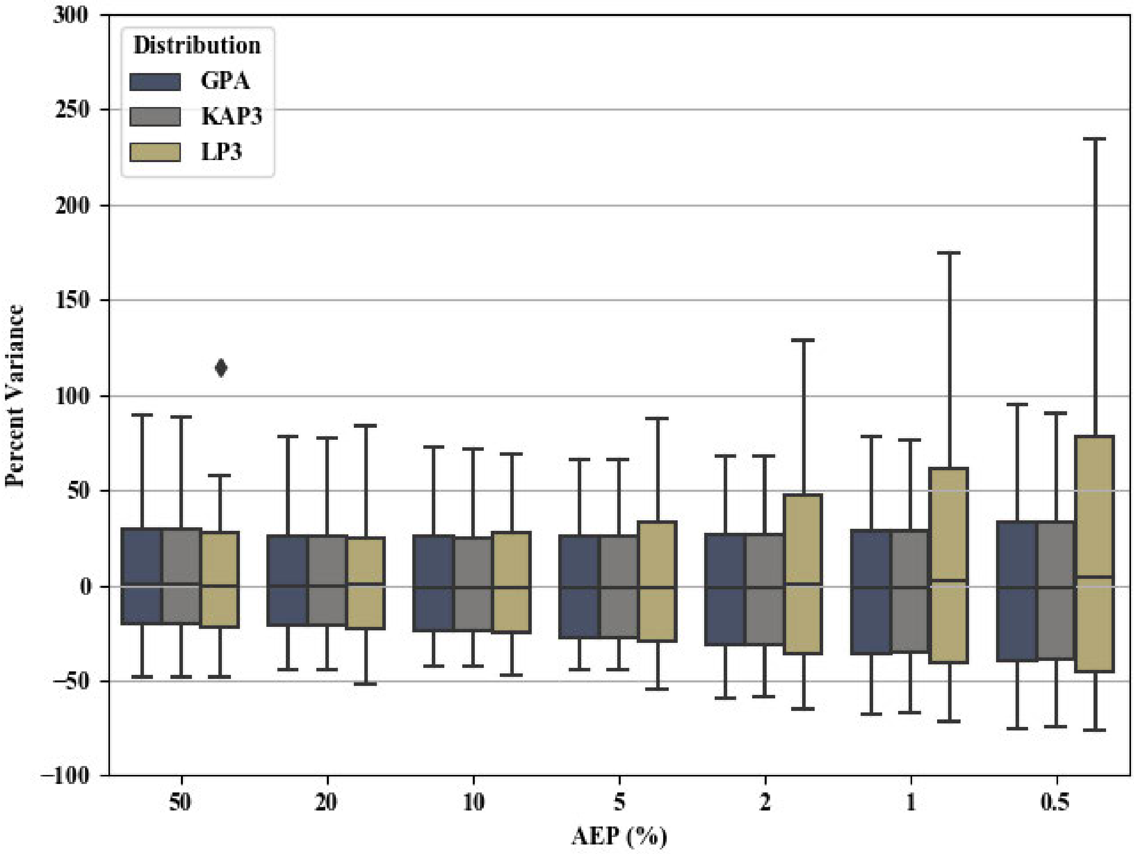


Figure 3.7 Variation of the 90% confidence bands presented as percentage variance of the balanced bootstrap confidence bands for the GPA, KAP3 (h = 0.77), and LP3 distributions for 148 sites with record lengths of 50 years or greater.

3.4.6 Distribution ranking

Each distribution was ranked based on the performance for each test undertaken and is shown in Table 3.7. The KAP3 ranks highest in two of the approaches undertaken, followed by the GPA and LP3 both performing best in one of the two remaining approaches. The combined rank however indicates that the GPA distribution is most suitable with a combined rank of 7, followed by the KAP3 and LP3 with combined ranks of 8 and 9 respectively.

Table 3.7 Rank of distributions for the Goodness-of-fit, model fit criterion, graphical and uncertainty tests. The best performing distribution is highlighted in bold

Distribution	Graphical	GoF	Model Fit Criterion	Predictive Ability	Total
GPA	2	2	**1**	2	**7**
KAP3	**1**	3	3	**1**	8
LP3	3	**1**	2	3	9

3.5 Discussion and Conclusions

The literature indicates that many distributions for FFA are available and are prescribed internationally. However, in South Africa, little scientifically justifiable investigation has been undertaken into the most suitable distributions to use for DFE. There is therefore a need for a detailed scientific investigation to identify the most suitable distribution for use based on South African peak flow data.

Four separate approaches were applied at 383 locations, consisting of 296 river gauges and 87 synthetic dam inflows, for the selection of the most suitable distributions using a combination of graphical methods, GoF tests, model fit criterion and model predictive ability.

The graphical method was utilised at a national scale as the initial screening process to identify the three most suitable distributions. This approach was adopted to ensure that the distributions assessed stemmed from a set of potentially overarching parent distributions as opposed to assessing all potential distributions. LM diagrams and the KP test (Kjeldsen and Prosdocimi, 2015), identified the KAP3 as the most suitable, followed by the GPA and LP3 distributions.

The GoF tests indicated that the distribution with the highest acceptance rate is the LP3. However, given the high acceptance rates of all of the candidate distributions, little clarity was provided as nearly all three candidate distributions considered were accepted at the majority of the sites for all tests applied (AU, CS, CvM, and KS) and for all of the record length categories considered.

The model fit criteria (AIC, AICc, and BIC) provided further guidance on the suitability of the distributions. The results were assessed based on the four record length criteria and the KAP3 distribution was not selected as the most suitable model for any of the criteria or record length categories. The LP3 was chosen as the most suitable distribution for sites with 80 years of records or more only, and the GPA was the chosen model for the remaining categories. The GPA performed particularly well for sites with records between 40 and 80 years, being the highest ranked distribution at up to 27.6% more sites than the LP3. Each distribution was also ranked for each criteria and record length category which gives an indication of the overall performance of the model. The GPA was identified as always being either first or second most suitable, whereas the GPA and LP3 were divided for second most suitable distribution. This led to the GPA being deemed as the most suitable distribution for the model selection criterion, followed by the LP3 and KAP3 distributions.

The final consideration for distribution selection was the predictive ability, which was established based on two indicators: (i) the ratio of the estimated vs the Gringorton plotted 5% AEP, and (ii) the uncertainty of estimated values associated with the selected distributions. The 5% AEP was utilised to adhere to the recommendations of Robson and Reed (1999). To assess the uncertainty of estimates the 1% AEP adopted due to the general "rule-of-thumb" applied in practise the quantiles should not be extrapolated beyond two times the available record length. Although the uncertainty is not traditionally used for the selection of a distribution, it has become an important consideration in hydrological modelling (Montanari *et al.*, 2013), and provides an indication of the reliability of estimates. Balanced bootstrapping was used to determine the 90% uncertainty bands associated with each distribution for each site considered by generating 1000 bootstrap replicates for analysis. No clear most suitable distribution was identified through the "true" fit analysis, although the LP3 marginally outperformed the GPA and KAP3. Similarly the uncertainty analysis was not able to identify a clear favourite between the GPA and KAP3 distributions, however, it did indicate that the use of the LP3 results in high uncertainty and was therefore ranked as the least suitable distribution based on the predictive ability.

Based on the assessment undertaken it is thus recommended that the GPA, which is a special case of the KAP, is the most suitable distribution to use when applying FFA on a national scale in South Africa.

This paper forms part of a larger project identified by the NFSP for the improvement of the status of design flood-based research. The project aims to develop new regional flood frequency models to improve on estimates made by existing models. Future work will focus on the identification of hydrologically similar flood producing regions and develop regional flood models utilising the most suitable distributions for South Africa.

4 FORMATION OF HYDROLOGICALLY SIMILAR POOLING GROUPS FOR USE IN DESIGN FLOOD ESTIMATION IN SOUTH AFRICA

4.1 Abstract

The pooling of hydrological data, through the formation of pooling groups or regions, has been shown to improve the reliability of design flood estimates by supplementing at-site data with spatial knowledge. In South Africa a number of geographic region classifications exist but have been developed either subjectively or through the use of limited datasets. Only three studies have been identified that utilised statistical verification of homogeneity and only a single study performed multivariate classification analysis for the formation of the pooling groups, albeit in a limited geographical region. Since the development of these regions, decades of additional data and new approaches to regionalisation have become available to validate and refine the existing regions. In this study both clustering and the RoI approaches are applied to data from 383 flow gauging stations and were able to identify 42 statistically relatively homogeneous pooling groups within South Africa for use in DFE.

4.2 Introduction

The design of hydraulic structures and water resources management require the determination of the anticipated flow rates for a predetermined AEP. To this end practitioners often rely on the use of statistical analysis of at-site data, however, short record lengths increase the uncertainty of estimates. At-site statistical analysis is also limited to the site in question and transferring of the analysis to ungauged sites is not simple to do.

The use of RFFA addresses the transfer of knowledge from gauged sites to ungauged sites, thus using spatial knowledge to supplement temporal knowledge. The pooling of knowledge between sites has been shown to improve the confidence of estimates, through the formation of adequately similar pooling groups (Burn, 1988, Blöschl *et al.*, 2013). It is, therefore, imperative that the method selected for the identification of suitable donor sites is robust. When performing regionalisation, it is necessary to determine what information is best transferred, how to transfer the information and what catchments are used to derive the information. The selection of catchments to use is generally based on spatial proximity or hydrological similarity, which are often based on catchment descriptors (e.g. catchment size, land use, geology, elevation, soil characteristics as well as climate variables such as *MAP* as surrogates for hydrological response (Merz and Blöschl, 2005).

Contiguous fixed region, non-contiguous fixed region, and hydrological neighbourhood type are approaches used for regionalisation (Gado and Nguyen, 2016). Geographic locations and/or administrative and political boundaries have traditionally been used for regionalisation and more recent techniques include cluster analysis (e.g. Tasker, 1982), discriminant analysis (e.g. Wiltshire, 1986b) and discordancy measures (e.g. Hosking and Wallis, 1993), all of which require some subjectivity in region formation and are dependent on the similarity measures and classification techniques employed (Ilorme and Griffis, 2013, Gado and Nguyen, 2016). Hydrological homogeneity is generally determined by statistical homogeneity (Ilorme and Griffis, 2013). In order to overcome the subjectivity involved, Ilorme and Griffis (2013) introduced a new statistical metric to identify physically discordant sites and a new methodology to identify the physical attributes that are the most indicative of extreme hydrologic responses. Sites which were both hydrologically discordant, as determined by the Hosking and Wallis (1993) *H*-test, and physically discordant as determined using principal component analysis performed on all available physical variables, were excluded from a region. A combination of cluster analyses, principal component analyses, canonical correlation analyses and multiple discriminant analyses applied to flood statistics and physical variables were used as an intermediary step to identify the most relevant physical variables to use in a cluster analysis for the regionalisation process (Ilorme and Griffis, 2013). When this approach was compared to physically-based regionalisation procedures typically employed in practice, it resulted in more homogeneous regions and more efficient quantile estimation at ungauged sites and also enabled the flood regime and estimated quantiles to be inferred at sites outside the extent of the area used for model development (Ilorme and Griffis, 2013).

Merz and Blöschl (2005) evaluated the predictive performance of various flood regionalisation methods in 575 ungauged catchments in Austria and found that spatial proximity is a significantly better predictor of regional flood frequencies than catchment attributes and a combination of spatial proximity and catchment attributes yielded the best predictive performance. When comparing a regression-based approach, an approach based on physical similarity and a spatial proximity approach to regionalisation, it was found that the spatial proximity offers the best solution for regionalisation (Oudin *et al.*, 2008), confirming the findings by Merz and Blöschl (2005). Additionally Gado and Nguyen (2016) identified that generally non-contiguous fixed region, and hydrologic neighbourhood type regionalisation approaches provide more accurate flood estimation than contiguous fixed region approaches.

From a review of the literature in regionalisation in modelling, Razavi and Coulibaly (2013) conclude that variability in catchment physical attributes and climatic variability result in different performances for different regionalisation methods. Razavi and Coulibaly (2013) also confirmed that generally spatial proximity and physical similarity have shown satisfactory performance in arid to warm temperate climate (e.g. Australia) and in cold and snowy regions (e.g. Canada), while spatial and regression-based methods have performed better in warm temperate regions (e.g. most European countries). The performance of regionalisation using three regionalisation approaches was assessed at 57 catchments in Québec, Canada and the results indicate that flood quantiles estimated using a scaling approach were the most accurate and robust (Gado and Nguyen, 2016).

In South Africa there have been eight studies undertaken, which focus on DFE and which developed pooling groups. Five of the studies published spatially defined regions shown in Figure 4.1. The most prominent of the studies, and the only widely adopted region definitions, are the veld type zones (HRU, 1972), K-regions (Kovács, 1988) and the SDF regions (Alexander, 2002a). HRU (1972) and Kovács (1988) performed hydrological analysis on available data sets and derived dimensionless 1-h unit hydrographs and Franco-Rodier *K* values, respectively. These parameters were then utilised in conjunction with physiographical maps to manually delineate homogeneous flood regions.

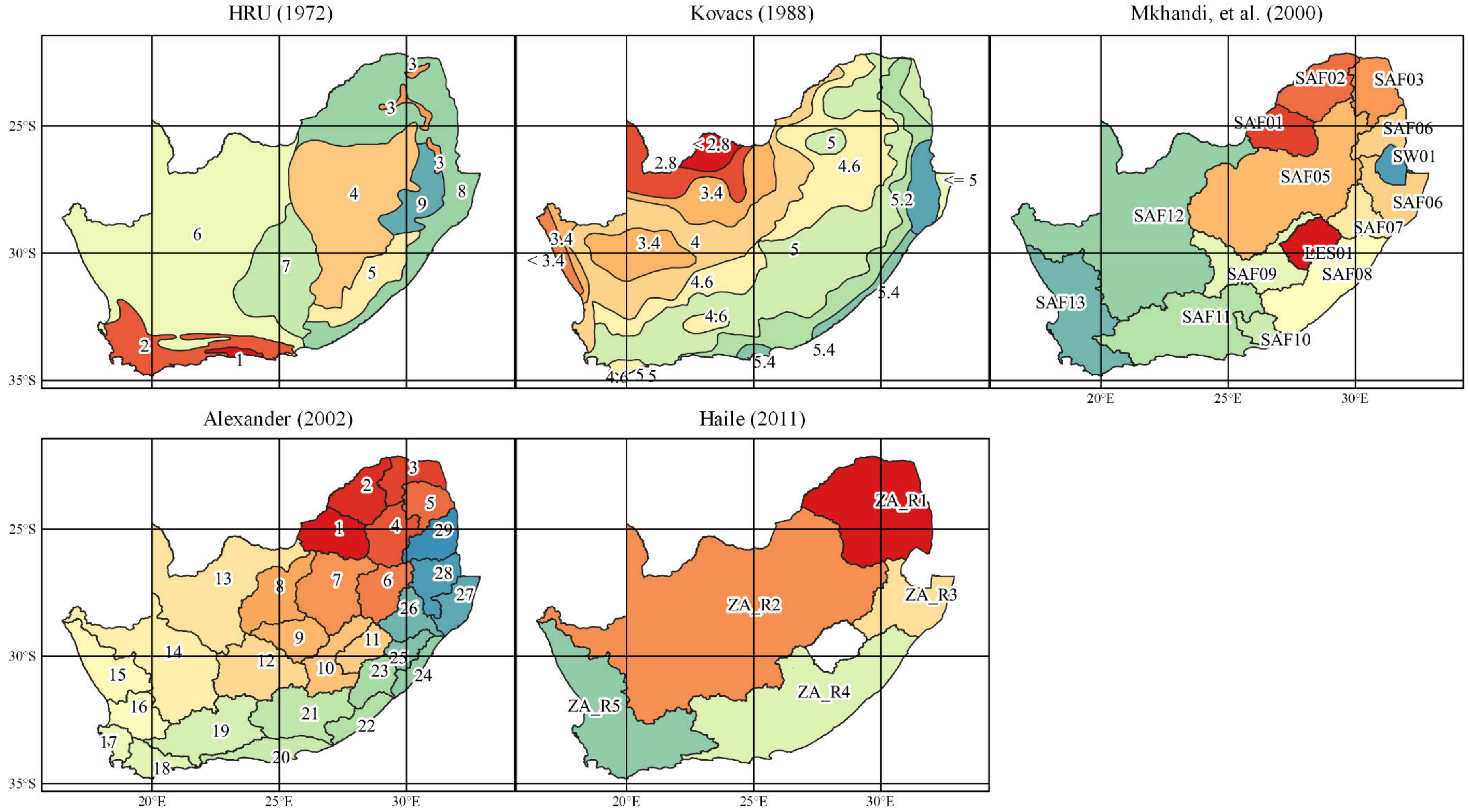


Figure 4.1 National homogeneous region definitions for South Africa as developed by HRU (1972), Kovács (1988), Mkhandi *et al.* (2000), Alexander (2002a), and Haile (2011)

The regions defined as part of the JPV method (Görgens, 2007a) were formed through grouping of the veld type zones into three groups, similarly grouping was undertaken for the K-regions into high, middle and low K-regions as a separate regional definition. Meigh *et al.* (1997), Mkhandi *et al.* (2000), and Haile (2011) relied on geographic and climatological parameters for the delineation of their regions, with Haile (2011) utilising the regions defined by Mkhandi *et al.* (2000) as a basis. Meigh *et al.* (1997) manually assessed the frequency curves of the sites considered and identified two pooling groups that are divided by a *MAP* of 1250 mm, and notes that additional work is required to improve the *MAF* flood model developed. Kjeldsen *et al.* (2001) performed clustering in the KZN province and successfully identified two homogeneous clusters using catchment descriptors and recommended further investigation into suitable regionalisation schemes and descriptors.

From the above, no consistency is evident between studies with regards to the number of regions that South Africa is divided into with the number of regions varying between two (Meigh *et al.*, 1997) and 29 (Alexander, 2002a), with a number of studies recommending that the number of regions be increased to accommodate for the hydrological diversity of South Africa (Van Bladeren, 2005, Gericke, 2010, Smithers, 2012, Van Dijk *et al.*, 2013). In addition, Nathanael *et al.* (2018) assessed the performance of the methods developed by Meigh *et al.* (1997), Mkhandi *et al.* (2000), Görgens (2007a), and Haile (2011) and found that the methods only performed adequately in approximately 57 % of catchments assessed. The methods used for formation of the pooling groups are also largely subjective, which has led to recommendations of refinement of the regions by the NFSP (Smithers *et al.*, 2014).

Conventional regionalisation techniques form groups of fixed pooling groups or regions that only have an interdependence with the catchments within their respective regions (Burn, 1990). However, Burn and Goel (2000) investigated the use of overlapping fixed regions in areas with limited hydrological data availability with promising results. From a literature review on regionalisation, Ridolfi *et al.* (2016) identified the fixed region and RoI as the most widespread approaches to regionalisation. However, there is no clear consensus on the best method of regionalisation in hydrology (Blöschl *et al.*, 2013). Similarly, both Oudin *et al.* (2008) and He *et al.* (2011) concluded that no single method was the best solution to regionalisation, but studies have shown the need to improve both the understanding and quantification of catchment hydrological responses (He *et al.*, 2011).

Rahman *et al.* (2012) noted that the formation of regions can be performed on both geographic or attribute space proximity, indicating that geographic proximity may not equate to hydrological similarity. When considering regionalisation using the attribute space approach, the regional placement of an ungauged catchment may be difficult; however, the catchment can still be placed in a regional grouping based on the attribute space locality (Rahman *et al.*, 2012). This was also identified by Dalrymple (1960) who highlights that within a single state in the USA, there may be a number of homogeneous flood producing regions. However, these regions may be grouped across states and could potentially lead to pockets of homogeneous regions spread across a large area. Non-contiguous regionalisation has also been adopted in the UK (Robson and Reed, 1999, Kjeldsen *et al.*, 2008b) and data rich regions of Australia (Rahman *et al.*, 2019). In data sparse regions of Australia contiguous fixed region approaches have been shown to be the preferred method (Rahman *et al.*, 2019).

This paper describes the application of both the RoI and Clustering multi-variate regionalisation approached in South Africa, with the aim to develop statistically homogeneous flood producing regions. The applicability of the methods across the South Africa is discussed and recommendations are made based on a new regionalisation for DFE within South Africa.

4.3 Methodology

The formation of hydrologically similar pooling groups requires a variety of preliminary assumptions and decisions to be made to take into consideration the measure of similarity, what similarity will be based on, anticipated grouping size or requirements, as well as what metric will be used to verify the validity of the pooling groups. The final consideration is the grouping scheme or multivariate classification technique to be utilised.

4.3.1 Similarity measure

To form pooling groups an indication of the similarity between sites is required. To measure similarity between sites the Euclidian distance (D_{jk}) was used, with the aim being to minimise the combined distance between p number of catchment descriptors (C^i) at different sites (j and k), be it geographic or descriptor related. The Euclidian distance is calculated using Eq. 4.1.

$$D_{jk} = \sqrt{\sum_{i=1}^{p} \left(C_j^i - C_k^i\right)^2} \tag{4.1}$$

The Euclidian distance is used to provide an indication of the catchment similarity or dissimilarity on a multi-dimensional scale and relies on the catchment descriptors to define similarity. The catchment descriptors used are detailed in the following section.

4.3.2 Catchments and descriptors

The study utilised 383 sites, which are divided into 296 river gauges and 87 synthetic dam inflow records as shown in Figure 4.2. Each site has an associated set of catchment descriptors for use in the regionalisation process.

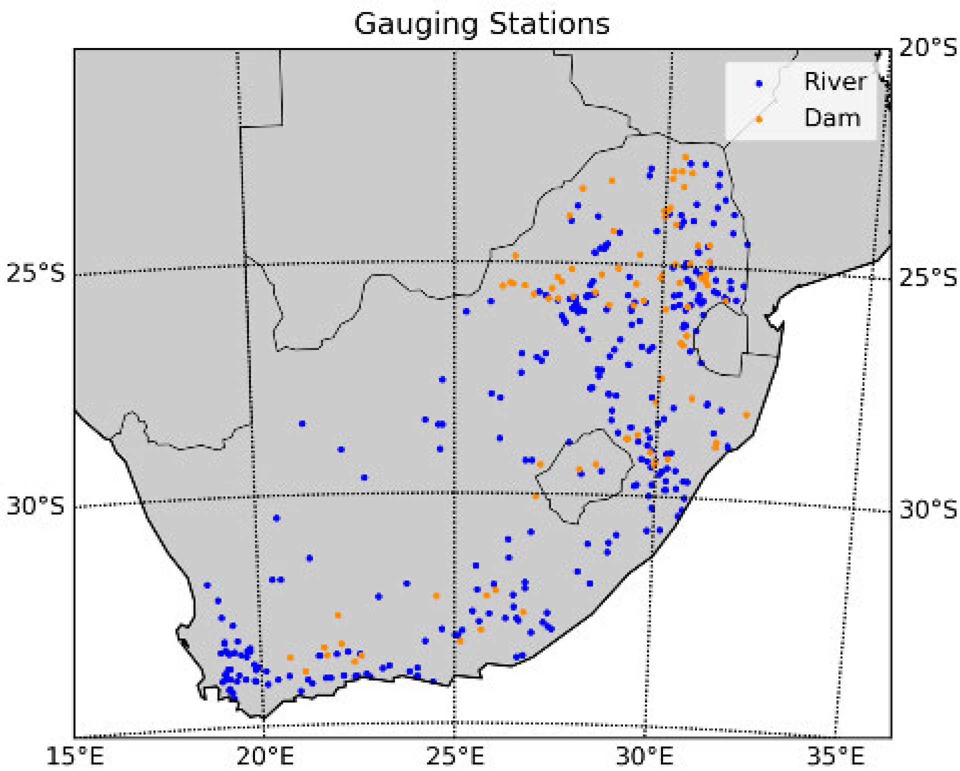


Figure 4.2 Map indicating the DWS gauging stations (blue) and the synthetic dam stations (orange) selected for use in study

Various catchment descriptors have been used for the formation of pooling groups in the literature, but it is important to note that the descriptive statistics of the AMS were excluded from the regionalisation to ensure that the homogeneity testing and pooling group formation remain independent. The catchment descriptors utilised in the formation of the pooling groups is shown in Table 4.1. All possible combinations of catchment descriptors were investigated and the combination that generated the largest number of relatively homogeneous regions without manual intervention was adopted for further investigation.

Table 4.1 Catchment descriptors selected for the formation of pooling groups

Descriptor	Unit	Range		Source
		Min	Max	
Outlet latitude	Decimal degrees	-34.36	-22.63	(DWS, 2011)
Outlet longitude	Decimal degrees	18.69	32.18	(DWS, 2011)
Outlet elevation	masl	11.00	1969.00	(NASA-JPL, 2013)
A	km^2	0.26	361994.80	DEM derived
Mean C_{ro}	Percent	4.00	97.00	(Schulze, 2011)
Areal mean SCS soil classification	Unitless	0	7	(Schulze and Schütte, 2020)
MAP	mm	60.00	3312.00	(Lynch, 2004, de Groen *et al.*, 2015)
10-*year* design rainfall intensity	mm/hr	0.55	148.00	(Smithers and Schulze, 2003)
100 vs 2-year design rainfall depth ratio	Unitless	2.06	4.50	
D_c	Decimal degrees	0.03	6.84	DEM Derived
Catchment Slope	m/m	0.0004	0.26	DEM Derived

Given that the descriptors utilised for the study vary in order of magnitude, normalisation (x_n) was undertaken to reduce the bias towards a single descriptor. The normalisation adopted and shown in Eq. 4.2 ensured that all parameters at site *i* (x_i) were within the range of 0 – 1 by scaling them within the maximum and minimum range of each parameter (x).

$$x_n = \frac{x_i - \min(x)}{\max(x) - \min(x)} \tag{4.2}$$

Merz and Blöschl (2005) found that spatial proximity is the most significant predictor of regional flood frequencies. Hence, the location parameters (latitude and longitude) were double weighted in comparison with the remaining catchment descriptors.

4.3.3 Selection of the number of pooling groups

Selection of the number of pooling groups is often an issue that researchers are faced with when performing RFFA, in particular when using more than three catchment descriptors for regionalisation. This is attributed to the fact that it is not possible to visualise data beyond three dimensions. One approach for the selection of a suitable number of regions is the use of inertia/elbow plots, attributed to Thorndike (1953). Inertia plots compare the within-cluster sum-of-squares with an increasing number of regions. The plots tend to show a significant decrease in inertia with an initial increase in groups, but the trend decreases as the number of groups increases, generating plots that resemble elbow joints. The elbow point in the plots generally indicate an optimal number of pooling groups as increasing the pooling groups beyond this point provides little reduction in the inertia.

Further to the use of inertia plots, dimensionality reduction techniques are often utilised to reduce the multi-dimensional nature of the descriptor space into a two- or three-dimensional space which allows for the visual inspection of the data to define groupings. Additionally, the reduced components can be utilised for the identification of regions. Principal Component Analysis (PCA) (Pearson, 1901) is considered as a traditional dimensionality reduction technique. PCA is a linear reduction technique that aims to visually separate dissimilar data and maintain the global structure of the data but is highly affected by outliers in the data.

More recent advancements in dimensionality reduction techniques include T-distributed Stochastic Neighbour Embedding (TSNE) (van der Maaten and Hinton, 2008) and Uniform Manifold Approximation and Projection (UMAP) (McInnes *et al.*, 2018). These methods are non-linear techniques that preserve the local structures within the data through considering the neighbourhood around the points. McInnes *et al.* (2018) compared the TSNE, UMAP and PCA techniques for the identification of clusters in four multi-dimensional datasets and found that UMAP and TSNE outperformed PCA for the identification of individual data clustering, with UMAP outperforming TSNE for preserving the global structure of the data. The importance of these findings depends on the intended use of the reduced dimension components. The loss of the global structure by TSNE lends itself more to a pure visualisation technique, as opposed to when the global structure is maintained, as is the case with UMAP, allowing the components to be more effectively used for further analysis.

When selecting the number of pooling groups a balance between the number of stations in a pooling group must be found, as identified by Hosking and Wallis (1997), as groups that are too large may bias the data set, whereas too small a group, may add little benefit over at-site FFA. Robson and Reed (1999) recommend as a rule of thumb that the record length of the donor sites be five times the required AEP being estimated, i.e. a 1:20 year (5% AEP) flood estimate requires a donor set that has a combined record length of at least 100 years. This is referred to as the 5*T* rule and, if it is not possible to achieve a data set of five times the return period, then a minimum of two times (2*T*) is recommended.

For this study the elbow plot and dimensionality reduction techniques were utilised for the definition of the Super Regions. The size of the smaller pooling groups was motivated by the 5*T* rule (Robson and Reed, 1999), however, the lower recommended value of 2*T* was enforced when refinement was required, with emphasis placed on the 1% AEP. This ensures that each group has a combined minimum record length of 200 years, fulfilling the 2*T* requirement for the 1% AEP.

4.3.4 Homogeneity testing

Homogeneity testing refers to the calculation of test statistics to validate the assumption of homogeneity for a grouping of donor catchments in a pooling group. Hosking and Wallis (1993) provide test statistics that may be used during homogeneity testing, namely *H* and LM Discordance (D_i). The *H* statistics are derived using LMs (λ_r) and LM ratios (τ_r), of order *r* and the estimation procedures are detailed in Eqs 4.3 – 4.5 using an observation set *X* of length *n*, with an expected value *E(X)*.

$$\lambda_r = r^{-1} \sum_{j=0}^{r-1} (-1)^j \binom{r-1}{j} E\left(X_{r-j:r}\right) \tag{4.3}$$

$$\tau_r = \frac{\lambda_r}{\lambda_2}, r = 3, 4, \ldots \tag{4.4}$$

$$\tau_1 = \frac{\lambda_2}{\lambda_1} \tag{4.5}$$

H calculated using Eqs. 4.6 and 4.7, uses Monte Carlo simulations to create simulated homogeneous regions based on the Kappa distribution with regional record length weighted averaged LMs ($\bar{\tau}_r$), which Hosking and Wallis (1997) use to emulate all distributions. *H* compares the observed weighted standard deviation (*V*) of the at-site (*i*) LM coefficient of variations (L-CV) (τ_1^i) with the mean (μ_V) and standard deviation (σ_V) of the L-CV of the simulated homogeneous regions.

$$V = \frac{\sum_{i=1}^{k} n_i(\tau_1^i - \bar{\tau}_1)^2}{\sum_{i=1}^{k} n_i} \tag{4.6}$$

$$H = \frac{V - \mu_V}{\sigma_V} \tag{4.7}$$

If the value of *H* is less than one, a region is considered to be homogeneous (Hosking and Wallis, 1993). Hosking and Wallis (1993) also classified a value between one and two as possibly homogeneous and a value in excess of two, is considered possibly heterogeneous, however, Guse *et al.* (2010) applied less stringent homogeneity definitions. Guse *et al.* (2010) defined four categories: (i) strong homogeneity ($H < 1$), (ii) possibly homogeneous ($1 < H < 2$), (iii) slightly heterogeneous ($2 < H < 4$), and (iv) strong heterogeneity ($H > 4$). Guse *et al.* (2010) assessed the deterioration of the performance of probabilistic regional envelope curves when varying the homogeneity level of acceptance. The *H* values of 1, 2, and 4 were chosen as the acceptable upper bounds and, based on comparing the mean values of the Mean Absolute Relative Error (MARE) and the standard deviation of absolute relative error, an increase in the *H* value led to an increase of the MARE from 0.54 to 1.12. Based on the results from Guse *et al.* (2010), Ilorme and Griffis (2013) adopted an *H* value of 4 as an indication of a pooling group that, although having heterogeneity present, still improves on at-site estimates. Rahman *et al.* (2019) notes that in several attempts have been made at defining potentially homogeneous regions in Australia with little success and have therefore not enforced the requirement in the model development. Similarly, in the UK homogeneity is not enforced as a requirement in the formation of pooling groups (Robson and Reed, 1999).

The *H* statistic, however, does not provide insight into the homogeneity of individual sites within the proposed region. Therefore, Hosking and Wallis (1997) developed a discordancy measure, shown in Eqs. 4.8 - 4.10, as a means to screen the selected sites. Considering a group of *n* sites, the discordancy measure (D_i) provides a parametric measure of relative proximity of an individual site, *i*, relative to the remaining sites by comparing the site specific LM vectors

(U_i) with the regional mean matrix (U_m) and the covariance matrix S. A site with a D_i in excess of three is considered to be discordant.

$$D_i = \frac{1}{3}(U_i - U_m)^T S^{-1}(U_i - U_m) \tag{4.8}$$

$$U_m = \frac{1}{n}\sum_{i=1}^{n} U_i \tag{4.9}$$

$$S = \frac{1}{n-1}\sum_{i=1}^{n}(U_i - U_m)(U_i - U_m)^T \tag{4.10}$$

Kachroo *et al.* (2000) described a regional "graphical" homogeneity testing methodology used by Mkhandi *et al.* (2000) which, similar to the H statistics, relies on synthetically generated regions to test against. The simulated regions used by Kachroo *et al.* (2000), however, utilised the selected distributions rather than only the Kappa distribution. The "graphical" method identifies whether the regional t_3 falls within the simulated maximum and minimum simulated values, and an additional more stringent check is to identify whether the historical data lie within the approximate 95% bounds, which are estimated using the standard deviation. Kachroo *et al.* (2000) compared the proposed approach to that of Hosking and Wallis (1997) for twelve regions identified in Tanzania and found that the stringent "graphical" approach provided similar results.

Viglione *et al.* (2007) compared some of the common homogeneity tests for RFFA, including H, the Bootstrap Anderson Darling (BAD) test (Scholz and Stephens, 1987) and the Durbin and Knott test (Durbin and Knott, 1972). It was suggested that the homogeneity testing be performed based on the location of a site on the LM t_3 versus t_4 plot. Where the t_3^R value is less than 0.23, the H measure is to be used for homogeneity testing, however, if t_3^R is larger than 0.23, the BAD test is to be used. The H measure was also noted for its performance and its extensive use in hydrology.

In this study, the Hosking and Wallis (1993) H test statistic (Eq. 4.7) was used in conjunction with the discordancy measure D for homogeneity testing. Although the H measure defined by Hosking and Wallis (1993) considers a region to be homogeneous if the value is less than one, it is considered relatively homogeneous with an H value of between one and two, and relatively homogeneous regions are anticipated to provide more accurate DFE than single site FFA. Hence, H values less than 2 were deemed suitable for application in this study.

4.3.5 Classification of pooling groups

The most eminent regionalisation schemes applied in the literature are the RoI and clustering-based methods, or a combination of these, for the identification of pooling groups. Each method has a unique set of benefits, e.g. the RoI method is often used in data rich regions, whereas clustering, is often utilised in data poor regions. It is therefore anticipated that a combination of these methods could be a best suited solution for South Africa.

4.3.5.1 Cluster analysis

Cluster analysis does not require the restriction of contiguous regions, and both Hosking and Wallis (1997) and Blöschl *et al.* (2013) regard it as the most practical method of forming pooling groups. Cluster analysis is used to group catchments that have similar characteristics and hence is performed in the attribute space. Each cluster will therefore contain catchments with similar characteristics, which emphasises the importance of the selection of catchment descriptors and their respective weighting. Cluster analysis aims to minimise the total Euclidian distance for the entire study by ensuring that the Euclidian distance within each defined cluster is minimised. The method can thus be seen as a fixed region approach, which may be non-contiguous.

Clustering aims to group catchments with similar hydrological responses into clusters. Two commonly used approaches are hierarchical (Agglomerative or divisive) and k-means clustering, both of which require the selection of an appropriate number of clusters (k). Agglomerative hierarchical clustering initially assumes that each station being considered is its own cluster, after which clusters are grouped by Euclidian distance until only a single overall cluster remains, while divisive clustering performs this process in reverse order. Due to the Euclidian distances remaining constant between sites, the process is easily reproducible.

Alternatively, k-means clustering requires the definition of the number of clusters prior to undertaking the division into a set of k disjoint clusters (Cl), the initial cluster centroids are randomly generated and iteratively refined until no further reduction in the overall within-cluster sum-of-squares, also known as the inertia, is achieved. Each cluster can be described with a cluster centroid, which is equivalent of the mean location (μ) of the stations in Euclidian space. The inertia minimisation function for a group of n stations (s) is shown in Eq 4.11.

$$\sum_{i=0}^{k} \min_{\mu \in Cl}(\|s_i - \mu\|^2) \tag{4.11}$$

The random nature of the generation of the initial centroids when applying k-means can cause varied results between multiple analyses of the same data set. It is therefore widespread practice to use clusters defined through hierarchical clustering as an initial estimate of the centroids, followed by refining the clusters through k-means clustering, making the process reproducible.

Hosking and Wallis (1997) noted that the results from the clustering analysis should not be considered final and that subjective adjustments may improve the homogeneity of the identified clusters, and listed potential subjective adjustments that can be made. Wiltshire (1986a) used an iterative relocation algorithm to adjust the clusters, which iteratively increases or reduces the number of clusters and adjusts the included stations to achieve the lowest total Euclidian distance. Alternatively, Smithers (1998) and Smithers and Schulze (2003) used a more subjective approach to refine extreme rainfall clusters and Kjeldsen *et al.* (2002) recommends further investigation into the use of clustering for the formation of homogeneous regions in South Africa.

The clustering efficiency was tested by performing the clustering multiple times to assess the impact of the selected catchment descriptors and the methodology is outlined in Figure 4.3.

4.3.5.2 Region of influence and Super Regions

Burn (1990) deviated from conventional fixed regions methods and details the RoI approach which produces a unique region for each catchment or station that is being assessed. The development of the RoI approach has also been attributed to Acreman (1987) and Acreman and Wiltshire (1987). Similar to clustering, the RoI approach uses a similarity metric, for the definition of pooling groups or regions. The RoI, however, forms a unique region for each site investigated based on the similarity between the site in question and each of the donor sites available.

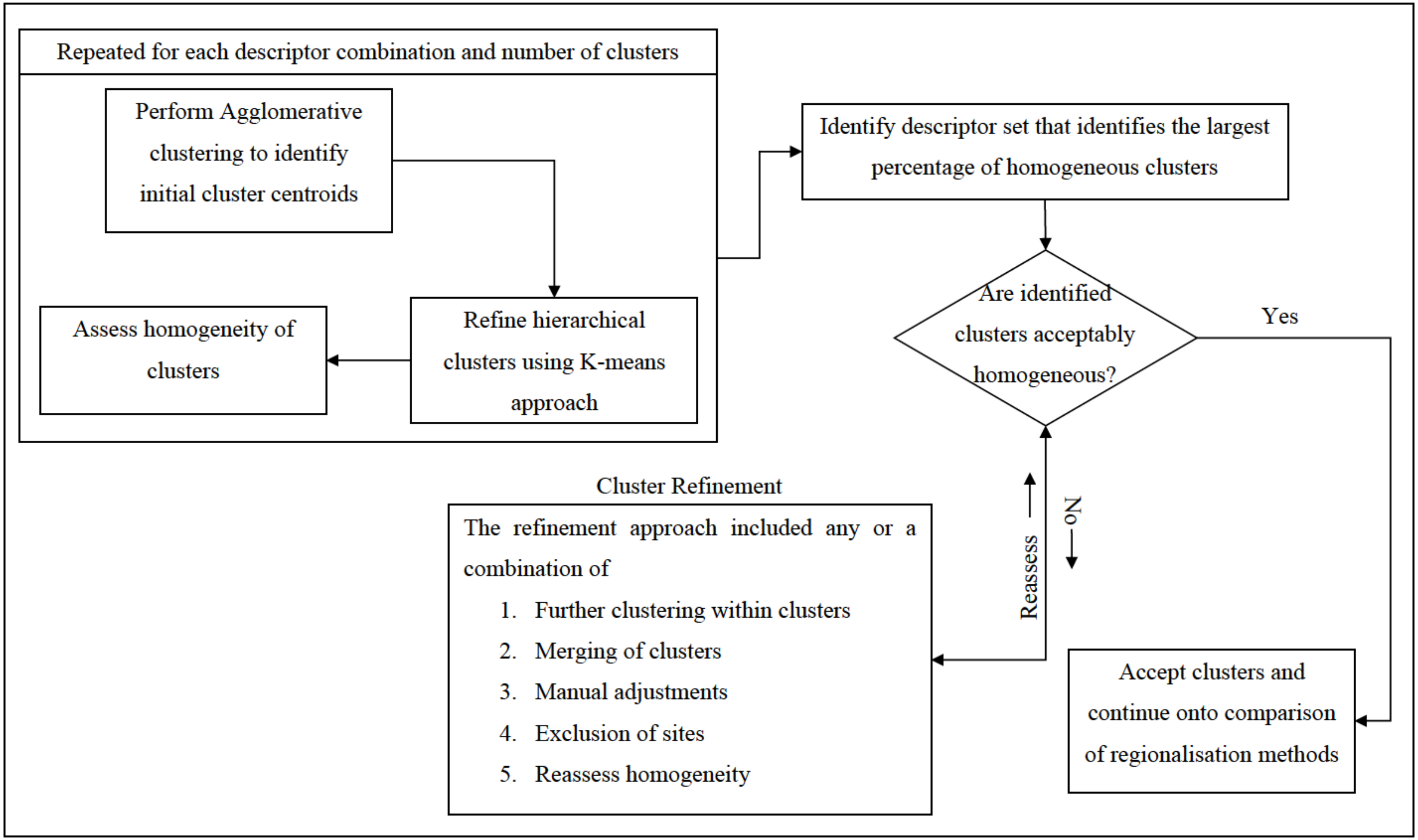


Figure 4.3 Clustering methodology flow diagram

Burn (1990) identifies the need that each identified region requires a threshold distance (*THL*), which provides an upper bound allowable Euclidian distance to be accepted. Burn (1990) highlighted the importance of selecting an appropriate *THL* as it affects the number, size, and homogeneity of the proposed regions. Considering that the donor sites will not all be equally close in Euclidian measure to the site being considered, a weighting factor between sites j and k (WF_{jk}) is proposed by Burn (1990) which considers a D_{jk} based weighting, as shown in Equation 4.12.

$$WF_{jk} = 1 - \left(\frac{D_{jk}}{THL}\right)^n \tag{4.12}$$

The value of WF_{jk} can thus vary between zero and one. The n value can be used to control the rate of decreased influence based on the distance measure. Eng *et al.* (2005) investigated the use of an alternative approach whereby the number of closest stations was predetermined which, in some instances, allows for stations of a distance in excess of the initial *THL* to form part of the region and found that the simple count based cut off provided optimal results. Zrinji and Burn (1996) provide a revised RoI approach, combining it with the Hierarchical approach which uses a number of RoIs per catchment being investigated, depending on the variable being estimated. Haddad *et al.* (2015) compared the RoI approach to fixed region approaches in Tasmania and identified that the RoI methods generally presented improved results over the fixed region approaches. The methodology adopted by Haddad *et al.* (2015) is detailed by Reis *et al.* (2003) and Haddad *et al.* (2012).

Noteworthy studies that use the RoI approach for the formation of homogeneous regions include the development of the UK Flood Estimation Handbook (FEH) (Robson and Reed, 1999, Kjeldsen *et al.*, 2008b) and the ARR Regional Flood Frequency Estimation (RFFE) in Australia (Rahman *et al.*, 2015a), both of which have been adopted in national DFE guidelines.

A more recent development introduces the concept of super regions (Mostofi Zadeh and Burn, 2019), which applies a hierarchical regionalisation methodology. The approach is a more formalised approach to the method used by Eng *et al.* (2005), who applied the RoI to a geographic subregion, rather than a classified grouping. As an initial step clustering is performed, to minimise the hydrological variability, within which secondary regionalisation is undertaken using RoI. Mostofi Zadeh and Burn (2019) describe the super region approach as a hierarchical regionalisation aimed at improving the homogeneity of the secondary regionalisation that is undertaken. The super region approach proved successful in Canada, where the percentage of heterogeneous pooling groups per super region was less than 7% for the worst performing super region.

An iterative process of ensuring all parameter combinations were assessed using the *H* measure, was also applied. The RoI approach proposed by Burn (1990) and used by the UK FEH (1999), was applied, and allows for the determination of regions based on the required record length, using the 5T rule. The RoI process is outlined in a flow diagram shown in Figure 4.4.

Given the success of the use of super region in Canada (Mostofi Zadeh and Burn, 2019), a similar approach was investigated in this study using a combination of clustering and RoI. The applied methodology is outlined in Figure 4.5. Initially the available streamflow gauges were divided into larger clusters, based on the catchment descriptors, to form the initial super regions. The number of super regions was selected based on the use of inertia/elbow plots, attributed to Thorndike (1953), in conjunction with T-distributed Stochastic Neighbour Embedding (TSNE) (van der Maaten and Hinton, 2008) and Uniform manifold approximation and projection (UMAP) (McInnes *et al.*, 2018) dimensionality reduction and geographic assessment. After suitable super regions were identified, the RoI approach in Figure 4.4 was applied to each super region to identify the descriptor set which identifies the largest percentage of homogeneous regions within the super region.

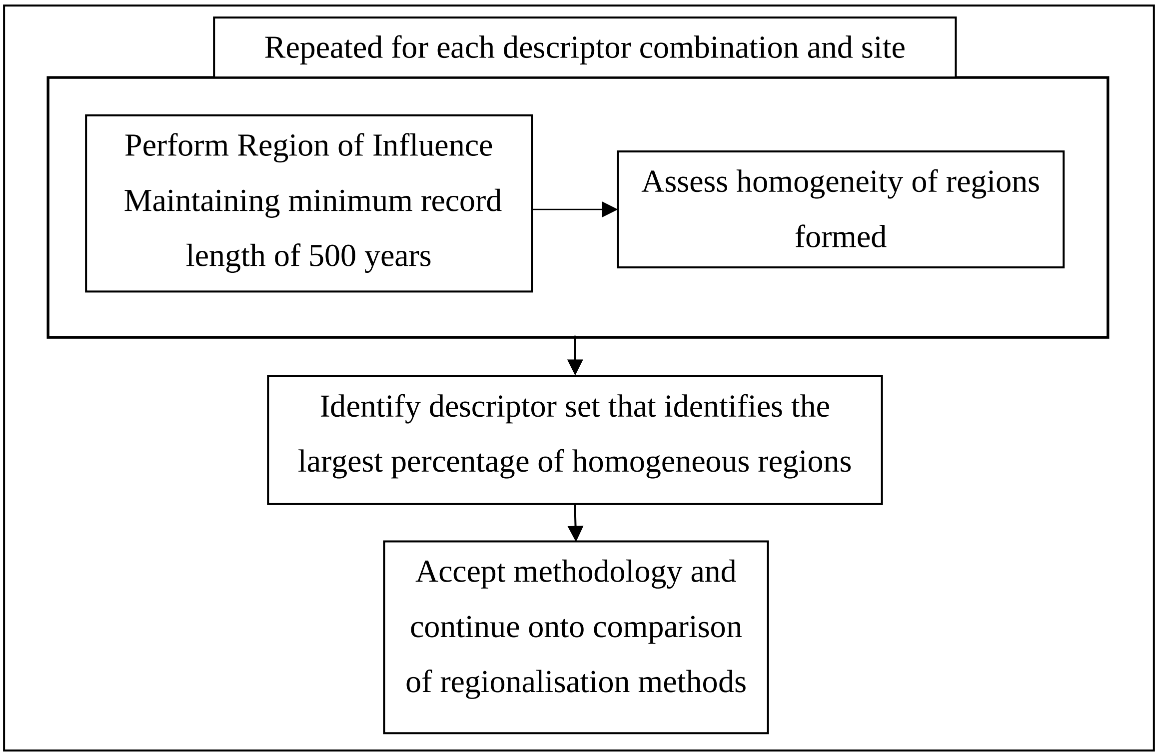


Figure 4.4 Region of Influence methodology flow diagram

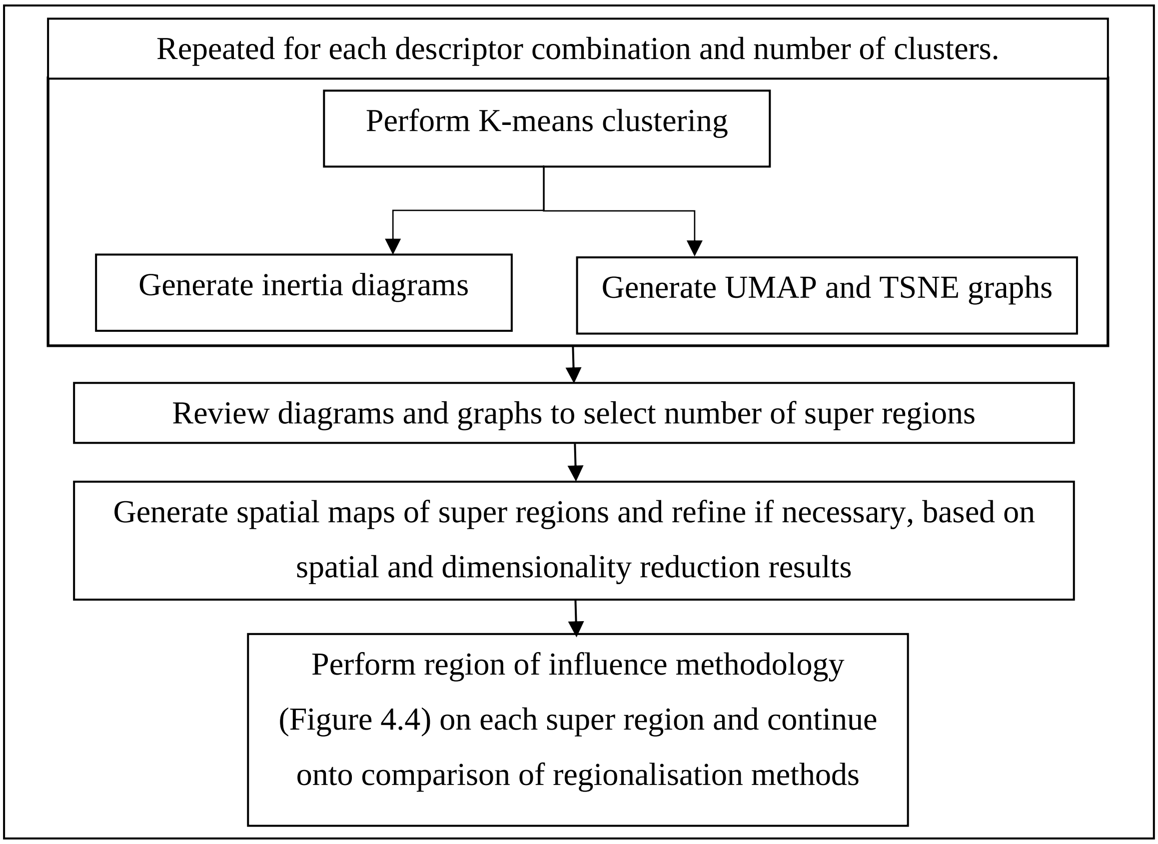


Figure 4.5 Super region methodology flow diagram

4.4 Results

In this section the results of the homogeneity testing of the DWS primary drainage regions is presented and Clustering, RoI and super region approaches are applied to assess the ability of the methods to define homogeneous pooling groups in South Africa.

4.4.1 Homogeneity of the DWS primary drainage regions

As an initial preliminary assessment and staying true to historical approaches of utilising physiographic and/or administrative boundary definitions as homogeneous regions, the DWS Primary Drainage Regions were assessed for homogeneity. The regions divide South Africa into 20 major catchments, as shown in Figure 4.6. Table 4.2 indicates the homogeneity of all stations and stations per regions prior to and after removal of discordant sites. However, it is still evident that even with the removal of the discordant stations the potentially homogeneous ($H < 2$) requirement is generally not met, both national and for 15 of the 20 regions. The number of discordant sites removed was determined by iteration after each exclusion, which required the further exclusion of additional sites.

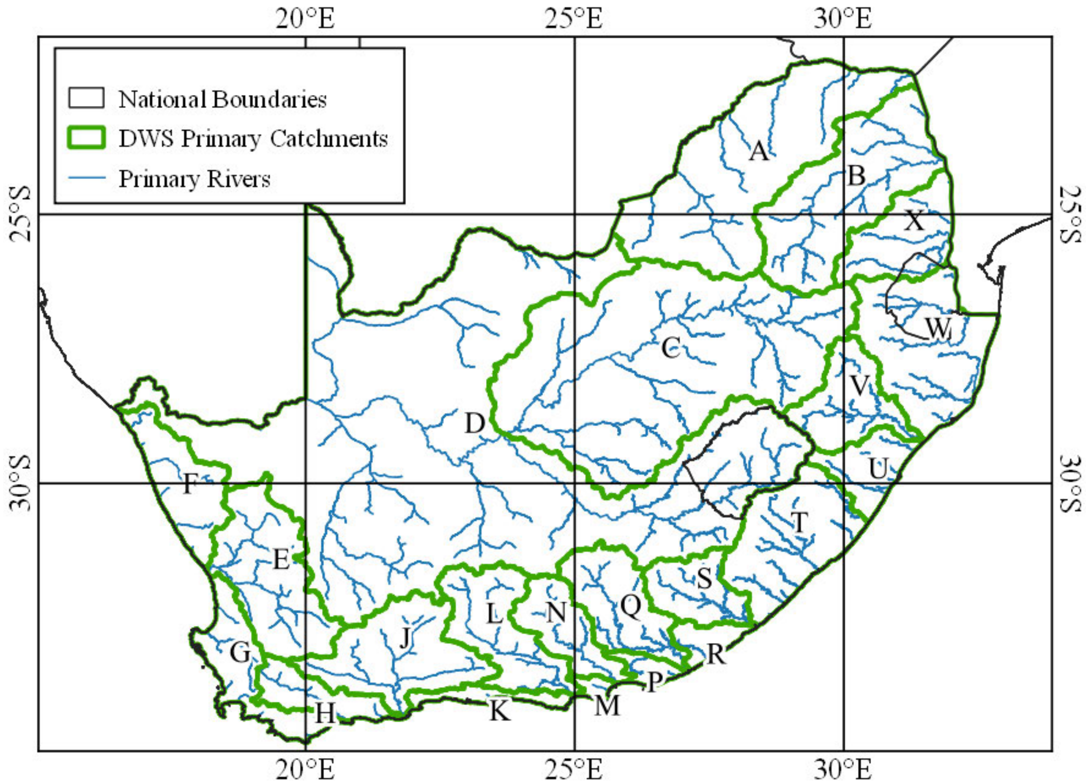


Figure 4.6 DWS primary drainage regions in relation to the primary rivers of South Africa

Table 4.2 Homogeneity testing for national and drainage regions

Region	No. of Stations	Cumulative Record Length (years)	*H* (Including Discordant Sites)	*H* (Excluding Discordant Sites)	Total No. of Discordant Sites Removed
National	411	18965	36.33	28.01	42
A	66	3626	10.87	8.64	4
B	52	2548	9.38	6.89	7
C	37	1679	5.62	5.62	0
D	20	1148	4.66	4.66	0
E	5	247	9.16	9.16	0
G	21	850	10.99	9.13	1
H	23	899	15.44	15.44	0
J	21	1083	6.99	4.95	3
K	10	504	10.18	10.18	0
L	6	295	0.25	0.25	0
N	6	326	0.00	0.00	0
P	2	**100**	0.29	0.29	0
Q	16	722	5.80	5.80	0
R	4	**166**	1.81	1.81	0
S	4	214	0.23	0.23	0
T	13	676	5.37	5.37	0
U	14	636	4.72	4.72	0
V	28	1554	3.36	2.60	2
W	20	957	4.91	4.91	0
X	43	2045	6.38	6.38	0

Removal of the discordant sites from the entire data set and from Regions A and B caused additional sites to become discordant, doubling the required number of sites to be excluded. Apart from Regions L, N, P, R and S, the regions were heterogeneous. Due to the lack of homogeneity in the remaining regions, and the need to exclude sites to achieve homogeneity, further division of the groupings is required.

4.4.2 Formation of Super Regions

As an initial step the formation of Super Regions in South Africa was undertaken, as per Figure 4.5. The descriptor combination which displayed the best performance in terms of TSNE, UMAP and spatial variability included seven descriptors: (i) Latitude, (ii) Longitude, (iii) *A*, (iv) *MAP*, (v) D_c, (vi) Catchment Slope (10-85) and (vii) 24-hour 10% AEP design rainfall. The Inertia/elbow plot shown in Figure 4.7 shows a comparison of the overall within cluster

Euclidian distance, or inertia, with an increasing number of clusters. It is evident that the elbow point in the graph is at five clusters given the sharp decline in the slope up to this point and was used as an initial selection for the number of super regions.

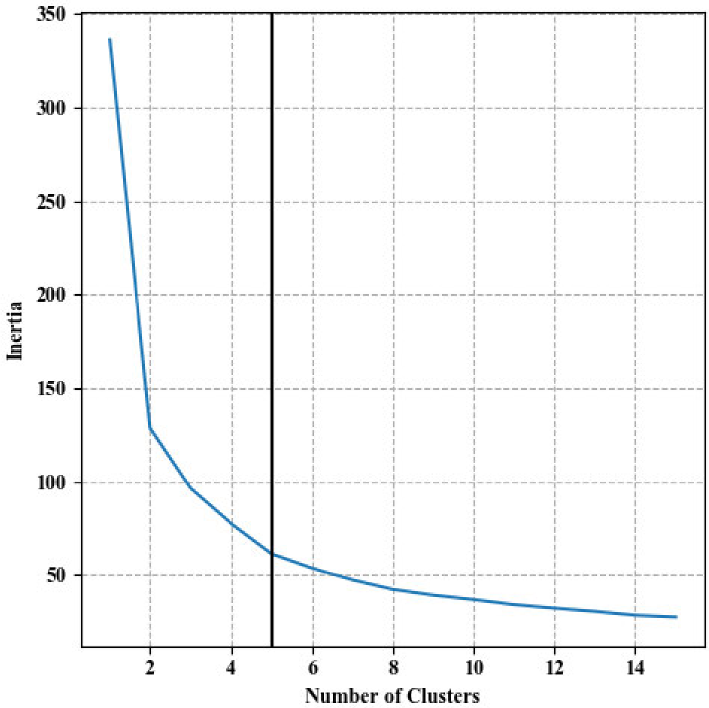


Figure 4.7 Inertia plot of the chosen super region descriptor set indicating the chosen number of clusters (five) with the black line using Latitude, Longitude, Catchment Area, *MAP*, D_c, Catchment Slope (10-85) and 24-hour 10% AEP design rainfall

Following the identification of the number of clusters, further refinement was undertaken through the use of the TSNE, UMAP unitless two dimensional components and spatial variation, shown in Figure 4.8 and Figure 4.9. From the dimensionality reduction and spatial plots, adjustments were made to the regions based on the TSNE results. Refinement of the TSNE clusters improved the spatial variation, however, the UMAP results still indicated potential improvements through the addition of an additional region. The inclusion of an additional region and improvement of the UMAP results also defined acceptable TSNE results, however, the spatial variation in DWS drainage regions V, W, and X deteriorated with overlapping super regions. Given the TSNE, elbow plot and spatial variation, the five super regions were adopted for use as opposed to the six regions defined by UMAP as the spatial overlap may lead to spatial inconsistencies of estimates.

Table 4.3 contains a summary of the characteristics of the five identified Super Regions. Super Region 1, located on the eastern seaboard of South Africa consists of 74 sites and is characterised by catchment altitudes ranging from nearly sea level to the highest catchment altitudes in the study. The catchments in the region are limited to a maximum A of 6 905 km^2.

Table 4.3 Catchment characteristics of the five identified super regions

Parameter	Super Region	Minimum	Median	Maximum
Outlet Elevation (masl)	1	35	1074	1969
	2	44	671	1584
	3	732	1247	1625
	4	142	846	1626
	5	11	349	1180
A (km^2)	1	0.68	640.15	6905.58
	2	0.61	1292.11	16812.53
	3	4.63	1126.58	361994.77
	4	5.59	263.44	13373.77
	5	0.56	52.28	43413.86
MAP (mm)	1	565	899	1107
	2	232	468	1206
	3	423	637	737
	4	477	842	1457
	5	173	573	1793
C_{ro} (%)	1	0.19	0.50	0.72
	2	0.44	0.56	0.72
	3	0.27	0.43	0.64
	4	0.22	0.45	0.65
	5	0.37	0.74	0.94
Catchment Slope (m/m)	1	0.0019	0.0085	0.0702
	2	0.0022	0.0055	0.1292
	3	0.0004	0.0036	0.0460
	4	0.0015	0.0122	0.1298
	5	0.0012	0.0372	0.2642

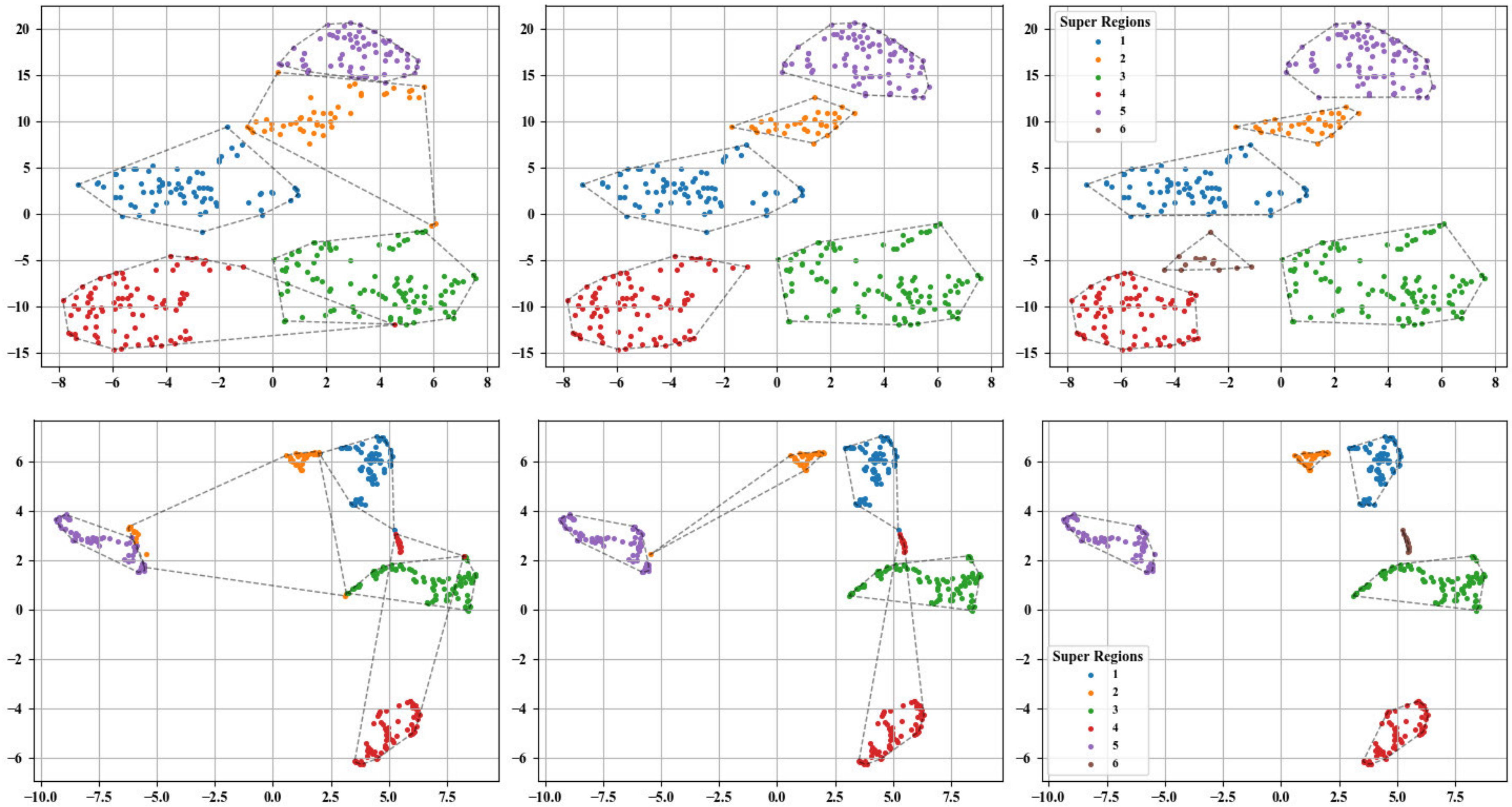


Figure 4.8 Verification of super region selection indicating the original (left), refined 5 regions (middle) and refined 6 regions (right) using the unitless two dimensional components of the TSNE (top) and UMAP (bottom) dimensionality reduction techniques

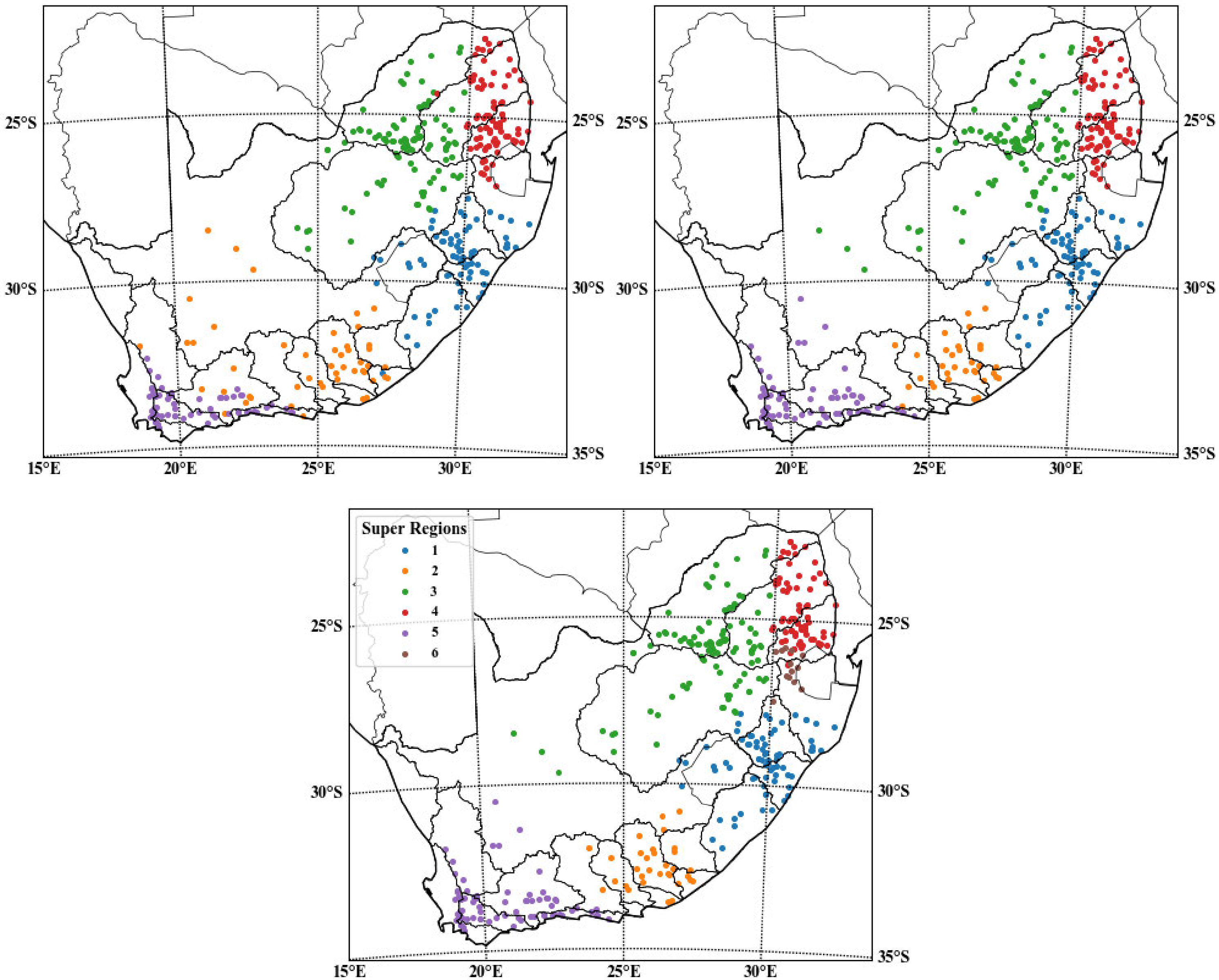


Figure 4.9 Geographic validation of super region selection indicating the original (top-left), refined 5 regions (top-right) and refined 6 regions (bottom)

As shown in Table 4.4 none of the super regions identified are considered relatively homogeneous, however the regions will be subjected to further regionalisation through the use of RoI. Super Region 2 represents the portion of the country where the catchment runoff variation is the least throughout the country. The 107 sites in the interior section of South Africa encompass Region 3 and have the mildest slopes, with a maximum slope of 0.046 m/m. Region 3 is also the driest of the super regions with a maximum MAP of 737 mm. Region 5 is predominantly located in the Western Cape province, where the mountainous catchments lead to the steepest catchments located within this region. Similarly, the catchments also present the highest runoff percentages coupled with the highest MAP.

Table 4.4 Homogeneity of defined super regions

Super Region	1	2	3	4	5
H	27.0	11.5	11.5	7.9	15.2

4.4.3 Classification of pooling groups

Classification of the pooling groups was undertaken using both RoI and Clustering approaches and both schemes were investigated due to the variation in the flow measuring network densities and data availability. As identified in the literature, clustering is often the preferred approach in data sparse regions as per the inland region of South Africa, whereas the RoI is often preferred in data rich regions.

4.4.3.1 Region of Influence

At each site investigated the full set of potential descriptor combinations were tested in order to identify the best descriptors for use in the selection of donor catchments in the RoI approach. Enforcing the 5T rule reduced the homogeneity of the regions in many instances and resulted in an increase of H beyond the adopted maximum value of 2. The best performing set of descriptors included Latitude, Longitude, D_c, and mean runoff percentage. Using these four descriptors resulted in 16% and 51% of the regions being relatively homogeneous for 500- and 200-year minimum record lengths approaches, respectively.

Having identified the five super regions, the RoI approach was also applied within each of the super regions to identify the descriptor combinations that identify the largest percentage of homogeneous regions, with each site forming a unique region, and the results are summarised in Table 4.5. The use of super regions has not significantly increased the percentage of relatively homogeneous regions on a national scale, yielding, from a homogeneity perspective, little additional benefit for the additional complexity.

Table 4.5 Descriptors and percentage of homogeneous regions identified per super region

Super Region	**No. of sites**	**Descriptors**	**Number of relatively homogeneous regions (%)**
1	80	Lat, Lon, A, MAP, D_c, $H24_{10\%}$	82.4
2	89	Lat, Lon	54.5
3	74	Lat, Lon, A, D_c	47.7
4	33	Lat, Lon, A, MAP, $H24_{10\%}$	62.9
5	107	Lat, Lon, MAP	28.8
All	**383**		**52.6**

Super Regions 1 and 4, located in the Eastern part of the country, contain the highest number of relatively homogeneous regions, 82.4 and 62.9 % respectively. In contrast, Super Region 5 performed poorly with only 28.8 % of relatively homogeneous regions identified. Super regions one and four are climatologically similar and affected by similar rainfall generating conditions, as opposed to super region five which is climatologically disparate, which may indicate that the descriptors investigated are better suited for the identification of hydrologically similar catchments in the eastern regions of South Africa. The use of the super regions is, however, anticipated to provide improvements for model development as the refined regions may better capture hydrological variations across South Africa.

4.4.3.2 Clustering

The clustering efficiency was tested by performing the clustering multiple times to assess the impact of the catchment descriptors, as per Figure 4.3. The descriptors were tested in an iterative fashion ensuring that every possible unique combination of descriptors was assessed. The entire station catalogue was used for the clustering for each iteration by initially dividing the data set into a maximum of 36 clusters, which allows for an average cluster size exceeding 500 years. The catchment descriptor combination that provided the highest percentage of relatively homogeneous clusters was: (i) Latitude, (ii) Longitude and (iii) D_c, which generated relatively homogeneous clusters in 44.4% of clusters that contain 35% of sites used as indicated in Table 4.6.

Table 4.6 Preliminary 36 clusters formed using location and distance from sea as the best clustering parameters

Cluster No.	No. of Stations*	H**	Record Length (years)
1	29 (1)	9.1	1445
2	4	4.6	172
3	8	1.1	340
4	8	2.9	457
5	12	1.5	540
6	4	2.4	230
7	9	1.4	437
8	14	6.3	685
9	17 (1)	10.6	775
10	5	7.3	217
11	7	1.6	393
12	7	2.9	264
13	13	1.8	660
14	8	1.8	405
15	15 (1)	1.3	777
16	4	0.3	199
17	10	0.8	424
18	7	1.2	296

Cluster No.	No. of Stations*	H**	Record Length (years)
19	6	0.4	254
20	18	4.2	959
21	14 (1)	7.8	634
22	11	4	482
23	19	10.7	648
24	2	0.9	124
25	3	1.5	219
26	12	3	718
27	9	1.1	407
28	15	9	841
29	19	4.5	1043
30	28	6.7	1279
31	10	1.5	371
32	9	5.2	456
33	8	8.2	297
34	5	0.9	136
35	6	1.9	368
36	8	2.3	397

* Station numbers in brackets indicate the number of discordant sites

** Shaded cells indicate relatively homogeneous clusters

Clusters were then adjusted using a combination of manual adjustments such as: (i) further clustering within clusters, (ii) merging of clusters, (iii) manual adjustment of clusters to improve spatial variations, and (iv) exclusion of discordant sites. The final number of sites utilised in the relatively homogeneous clusters is 332. The formation of the relatively homogeneous clusters therefore required the exclusion of 51 sites (13% of sites), and further investigation into the excluded sites did not present any clear spatial distribution or trends. Table 4.7 lists the final accepted clusters.

A total of 42 relatively homogeneous clusters were created that all satisfy the $H < 2$ requirement, however, in some instances the minimum record length of 200 years was relaxed with a minimum accepted record length of 129 in Cluster 18. The spatial distribution of the accepted clusters is shown in Figure 4.10. In addition, it must be noted that three of the relatively homogeneous clusters contain fewer than 5 sites, which may limit the predictive ability in these clusters.

Table 4.7 Accepted 42 relatively homogeneous clusters

Cluster No.	**No. of Stations**	H_1	**Record Length (years)**		**Cluster No.**	**No. of Stations**	H_1	**Record Length (years)**
1	10	1.5	357		22	8	1.9	263
2	15	1.6	860		23	5	1.3	182
3	5	1.0	325		24	5	1.0	192
4	11	1.9	530		25	6	1.9	368
5	12	1.5	540		26	5	1.5	299
6	6	1.2	329		27	6	1.9	280
7	5	1.8	184		28	5	0.2	252
8	7	1.5	431		29	8	2.0	400
9	10	0.8	424		30	5	1.9	241
10	8	1.9	424		31	8	0.1	374
11	9	1.8	389		32	7	1.5	317
12	6	1.3	198		33	11	1.4	507
13	8	1.8	405		34	9	2.0	387
14	12	1.3	517		35	18	0.7	1015
15	8	1.1	340		36	6	1.6	301
16	7	0.6	276		37	9	1.1	407
17	3	0.6	191		38	3	0.2	192
18	4	1.0	129		39	8	2.0	440
19	7	1.2	296		40	9	1.4	437
20	6	0.4	240		41	13	1.8	660
21	8	1.8	230		42	10	1.5	371

The geographic distribution of the clusters presents an even spread of clusters across the country, with data sparse regions in the inland DWS primary drainage Regions D, E and F. These regions have a small number of stations that met the selection criteria of the study and future studies may need to consider a relaxation of the criteria to improve the spatial coverage in these regions. Conversely, the data rich DWS drainage Regions B and H have clusters that overlap where Region B has a single cluster which overlaps spatially with three other clusters. Figure 4.11 provides the spatial bounds of each cluster at a scale of 0.1 degrees across the country, based on the Euclidian distance using the location and distance from sea of each of the points.

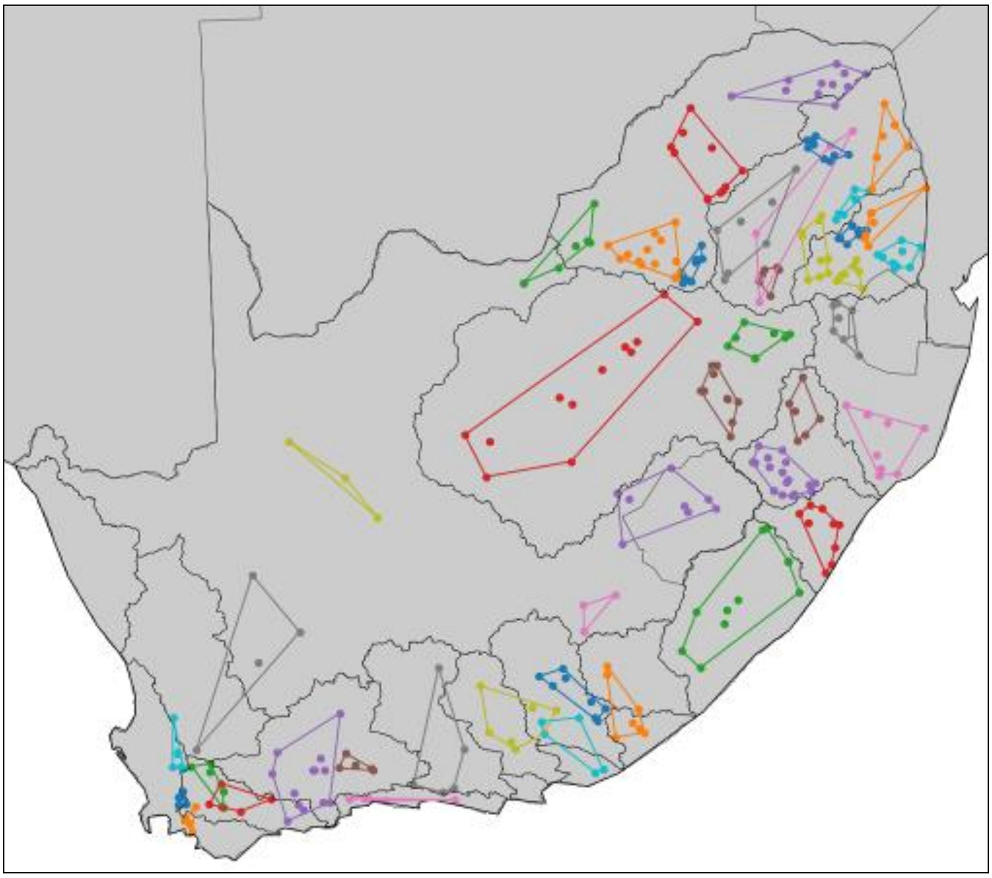


Figure 4.10 Distribution of the 42 relatively homogeneous clusters within South Africa identified enclosed by the convex hulls

When taking into consideration the physiographical catchment boundaries further refinement was undertaken as shown in Figure 4.12. The application of the physiographical boundaries is only used to determine the cluster membership of the sites and did not restrict the clusters from containing sites from neighbouring catchments. When developing the physiographical cluster membership map Cluster 7 was removed, as the cluster overlapped with three other clusters.

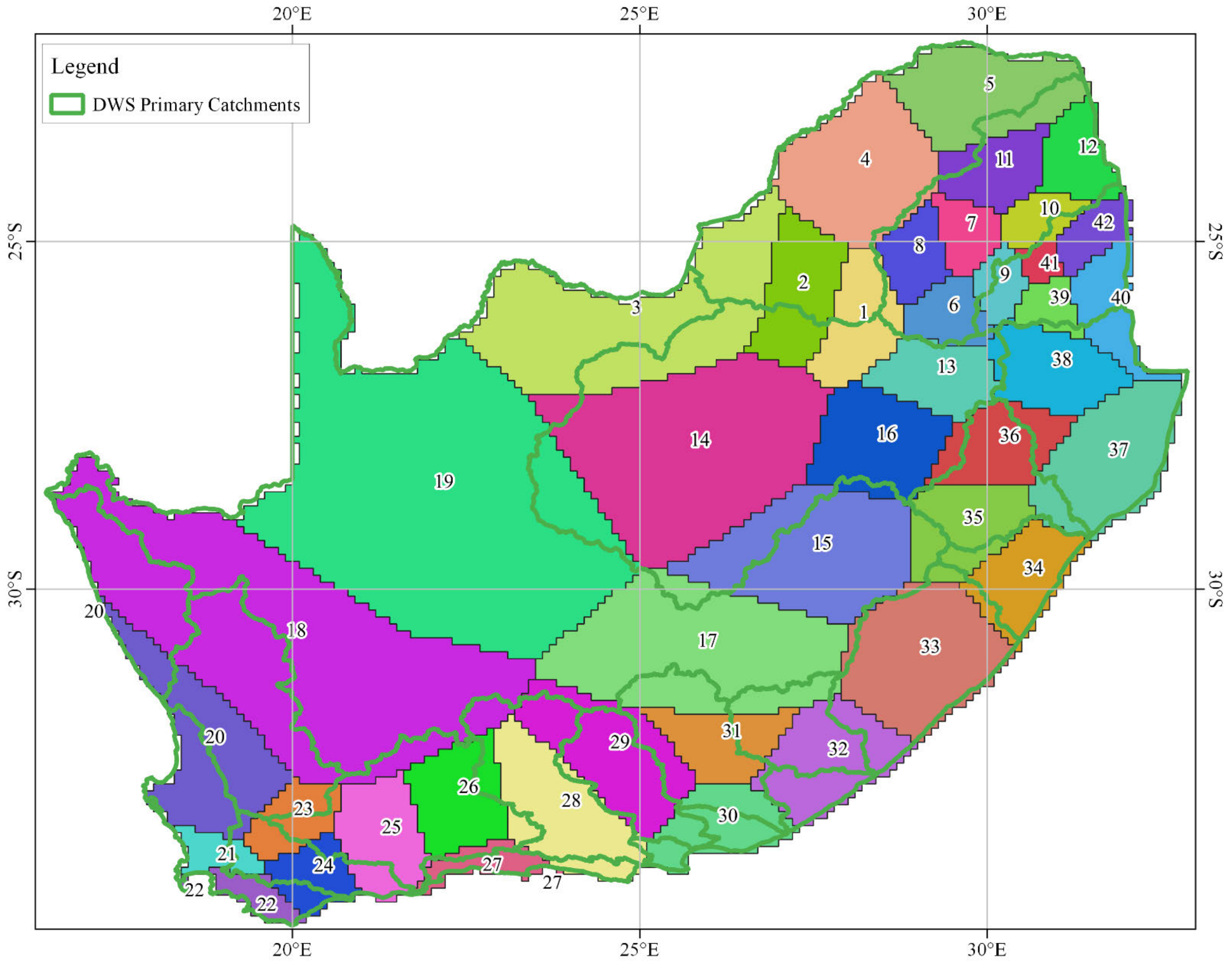


Figure 4.11 Delineation of national cluster association based on the Euclidian distance using the location and distance to coastline in relation to the DWS primary drainage regions (green) at a scale of 0.1 degrees

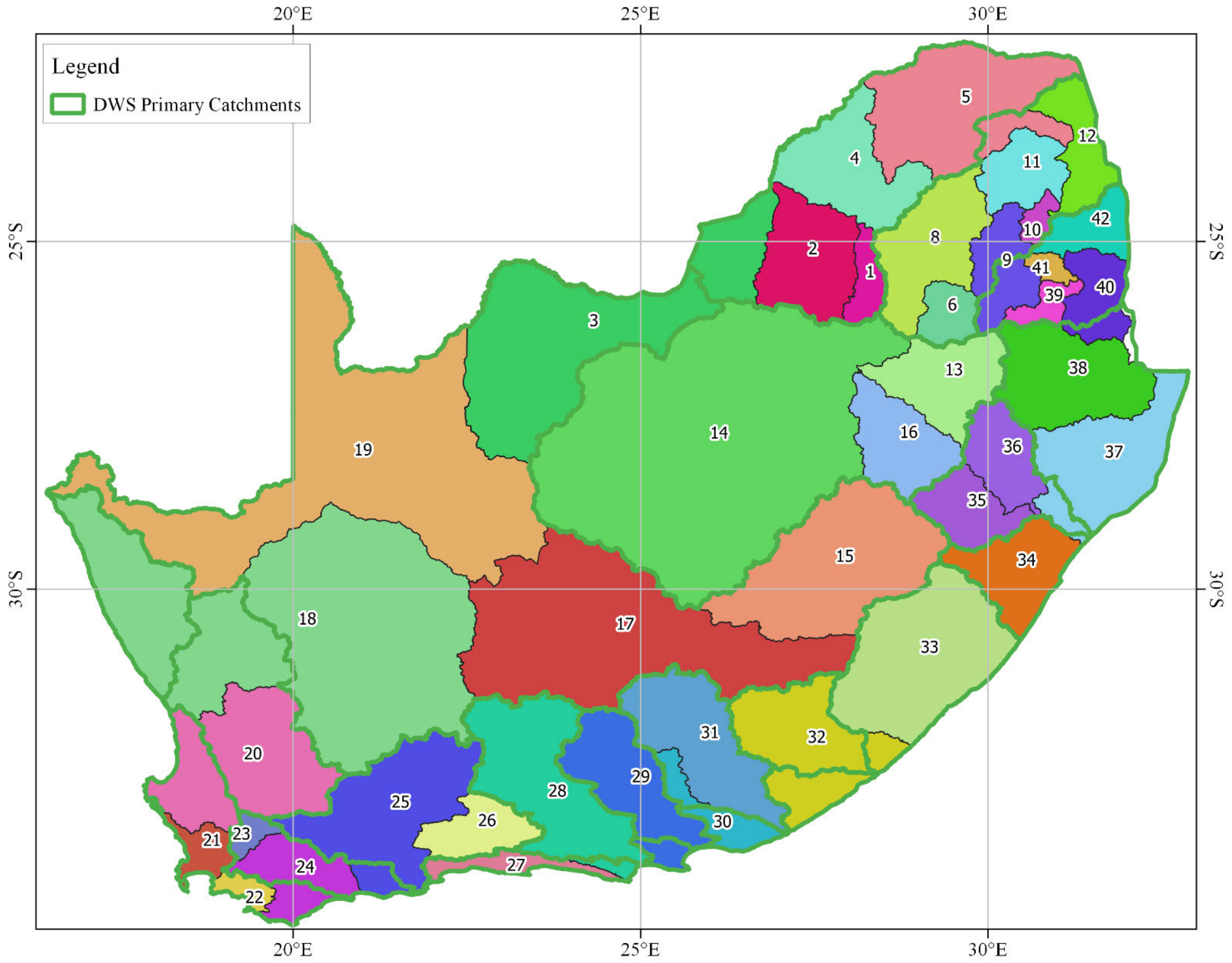


Figure 4.12 Delineation of national cluster association based on the Euclidian distance using the location and distance to coastline taking into consideration the physiographical catchments, shown in relation to the DWS primary drainage regions (green)

4.5 Discussion and Conclusions

An assessment of the national and DWS drainage regions as potential homogeneous flood producing regions was undertaken, utilising the 383 sites, to verify whether further division of the country was required. It was identified that majority of the drainage regions, except for Regions L, N, P, R and S were deemed heterogeneous. Of the five homogeneous regions, two had cumulative record lengths less than 200 years, therefore only three of the twenty regions were acceptable and substantiated the need for additional regionalisation to be undertaken. Eleven catchment descriptors were investigated for use and were applied in an iterative fashion with all of the regionalisation schemes investigated. This ensured that all potential combinations of catchment descriptors were assessed for its ability to identify homogeneous pooling groups.

The RoI approach is flexible to the needs of the user; however, once the rules of application are defined, adjusting regions becomes difficult due to each site defining a unique region. Latitude, longitude, D_c, and mean C_{ro} were the catchment descriptors that identified the largest percentage of homogeneous pooling groups, albeit only for 51% of sites. The rigid nature, coupled with the low level of identification of homogeneous regions could limit the uptake of the developed models, although statistical homogeneity is not a prerequisite of the RoI approach.

The super region approach, when applied in South Africa, adds an additional layer of complexity to the estimation for little additional benefit with regards to the identification of additional homogeneous regions (52.6%). The super regions, however, provide an indication that the RoI approach performs relatively well along the Eastern coast of South Africa, identifying 82.4% and 62.9% homogeneous regions in Super Regions 1 and 4, respectively. The performance is, however, particularly poor in Region 5, located in the south western region of South Africa. The inability of the RoI to form relatively homogeneous regions does, however, not prevent the development of flood estimation models. Both the ARR (Rahman *et al.*, 2019) and the FEH (Kjeldsen *et al.*, 2008a) do not rely on the formation of statistically homogeneous regions for model development and application.

In contrast to RoI, clustering forms fixed pooling groups and has been noted as the most practical approach to regionalisation. The catchment descriptor combination that identified the highest number of homogeneous clusters was: (i) Latitude, (ii) Longitude, and (iii) D_c. Initial clustering formed 36 clusters of which 44% were considered relatively homogeneous. Although the initial performance of RoI outperforms Clustering for 36 clusters, clustering allows the flexibility of further refinement. Manual adjustments and modifications led to the identification of 42 relatively homogeneous regions that are distributed geographically resulting in fixed regions. The use of physiographical catchment boundaries in the definition of the regions has led to a quasi-fuzzy clustering scheme as some regions incorporate sites located outside of the defined regions. The simpler definition and geographic distribution of the regions provide an approach that has a higher probability of acceptance with practitioners. The 42 relatively homogeneous clusters therefore, from a regionalisation perspective, is the recommended approach.

It is also important to note that the ability of the methods to form homogeneous regions is directly dependent on whether the catchment descriptors selected are well suited for this purpose. The inclusion of additional parameters may therefore lead to an improved ability to define homogeneous regions where the methods performed poorly. International guidelines also do not consistently require statistical homogeneity to be satisfied, in particular when applying the RoI, but emphasis is rather placed on the accuracy of the models developed based on the regionalisation approach adopted.

Future work will focus on the development of regional flood models for the regionalisation schemes investigated to further assess the viability of these in South Africa.

5 DEVELOPMENT AND ASSESSMENT OF REGIONAL MODELS FOR DESIGN FLOOD ESTIMATION IN SOUTH AFRICA

5.1 Abstract

The ideal situation for DFE modelling has been considered to be when flow data are available at the site in question. This is, however, not possible in many instances and the development of regional flood models has become prominent and allows for the estimation of design floods at ungauged sites. Regional flood models pool temporal and spatial data, thus supplementing data sparse regions with information from neighbouring regions. Through the use of dimensionless growth curves and regional regressions regional flood responses can be applied to an ungauged site analysis. In the South African context, a number of regional flood models have been developed, but have been shown to perform inadequately. Utilising the latest available streamflow data and homogeneous flood regions, four DFE models have been developed that improve on previously observed modelling results. Adopting an ordinary least square modelling approach in conjunction with both Quantile Regression Technique and Parameter Regression Technique model development, four model formulations were developed. Applying the models at both a regional scale and a national scale revealed that the regional models provide improved results, with the equally weighted Index Flood approach applied within a clustering regionalisation scheme performing best. Using a leave-one-out performance assessment the models are shown to estimate design floods within a desirable ratio at 76% of sites considered, improving on previously developed models by up to 28%.

5.2 Introduction

Design flood practitioners are often required to estimate design floods at sites where no hydrological flow monitoring has been undertaken. In situations such as these, practitioners are faced with a choice of which method to implement from a wide variety of DFE methods available. In South Africa, it is generally recommended that all appropriate techniques are applied and to use the results from the most appropriate technique based on professional judgement (Van der Spuy and Rademeyer, 2018). The estimation of floods in ungauged catchments has been a challenge faced by hydrologists for many years, so much so that the

International Association of Hydrological Sciences (IAHS) dedicated a decade of research from 2003 – 2012 to the reduction of uncertainty in hydrological predictions (Sivapalan, 2003).

DFE techniques can be broadly categorised as based on an analysis of streamflow data or rainfall-based methods (Smithers and Schulze, 2003), and Figure 5.1 shows the relevant categories of the DFE methods available in South Africa. Rainfall-based methods rely on rainfall data to estimate design floods, which range from event-based models, which utilise design rainfall as input, to the use of continuous simulation modelling, which requires long rainfall records. Streamflow analysis uses the statistics of observed floods to derive estimation techniques such as flood envelopes or empirical formulae. Alternatively, FFA can be performed to fit a distribution to the observed data and, when adequate periods of good quality data are available, is the recommended option for DFE. FFA is performed at an at-site scale, where the data from a single site is analysed to determine the design floods and is only applicable to the site being considered.

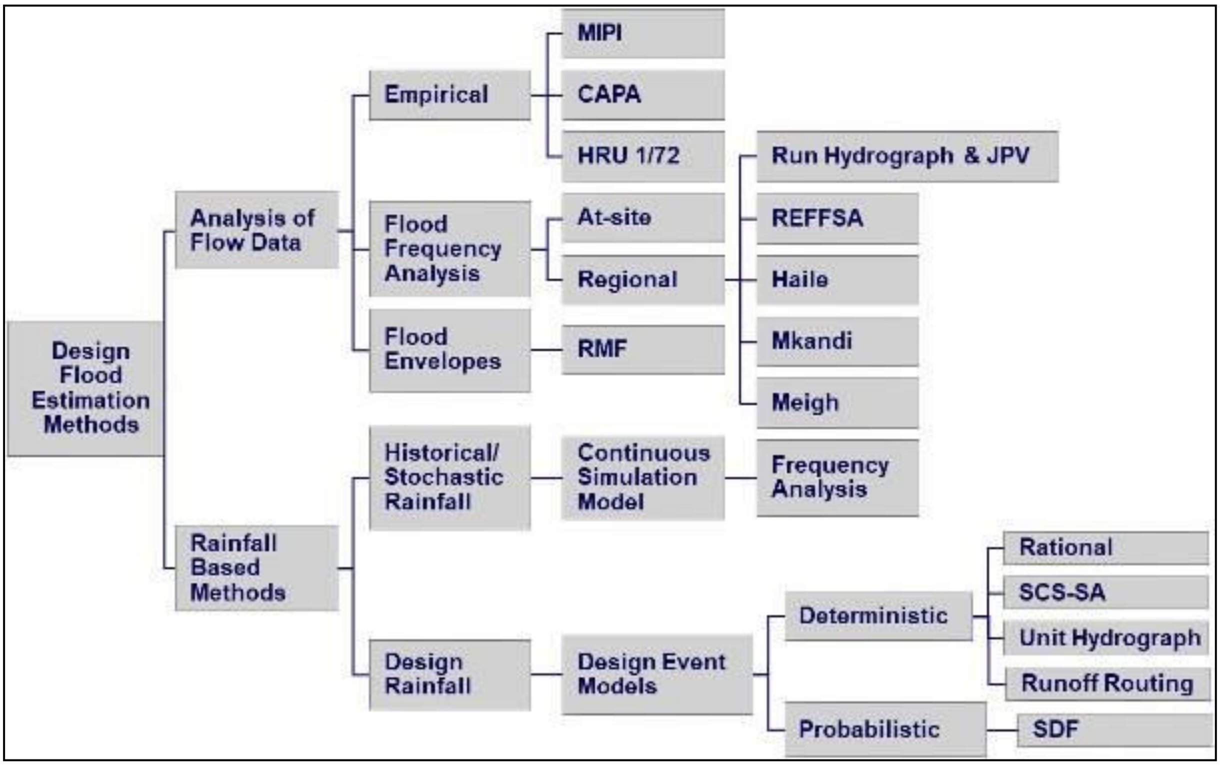


Figure 5.1 Design flood estimation methods available for use in South Africa (after Smithers, 2012)

Regional FFA (RFFA) has become more prominent in recent years as it has been shown to reduce the uncertainty associated with design flood estimates through the pooling of data

(Burn, 1990). This is generally applied in hydrological pooling groups, identified through regionalisation techniques such as RoI or Clustering. Once a set of pooling groups have been identified, regional models need to be developed to estimate the design floods at ungauged sites.

From a literature review the most common method of regional flood information transfer is the use of regression analysis. Weisberg (2005) describes regressions as the “study of dependence”, i.e. the dependence of the response variable on predictor variables. Regional flood model development generally falls within one of two categories, Quantile Regression Techniques (QRT) or Parameter Regression Techniques (PRT). QRT models directly estimate the quantile flows in question, e.g. the 1% AEP flood event, whereas PRT relies on regional growth curves and estimation of model parameters. Numerous model formulations exist for the development of regional flood models, but the formulation most widely reported in the literature is the Index Flood approach.

Rahman *et al.* (2012) notes that QRT may lead to an inconsistent growth curve, which can be avoided through the use of PRT which estimate parameters used to generate a growth curve at an ungauged site, thus providing a smooth increase with increased AEP. In addition, PRT can estimate floods for any AEP and is not limited to the derived QRT relationships.

The foremost PRT modelling methods currently implemented are the Index Flood (IF) (Dalrymple, 1960, Robson and Reed, 1999, Rahman *et al.*, 2019), and Probabilistic Rational Method (PRM) (McDermott and Pilgrim, 1982) methods. The IF relies on the estimation of Scaling Factors (SFs) to scale dimensionless growth curves to estimate the at-site quantile growth curves, whereas the PRM approach scales the dimensionless runoff coefficient curves for application with the RM. Each of these methods have been implemented internationally in well-established design guidelines (e.g. Robson and Reed, 1999, Rahman *et al.*, 2019).

Comparisons between QRT and PRT was undertaken on 53 catchments in Tasmania (Haddad *et al.*, 2012). *A* and design rainfall intensity were found to be the most important predictor variables in the QRT and four predictor variables were used in the PRT (Haddad *et al.*, 2012). The QRT was found to provide more accurate flood quantile estimates for the higher return periods while the PRT resulted in relatively better flood estimates for smaller return periods (Haddad *et al.*, 2012).

A similar comparison between QRT and PRT was undertaken on 237 catchments in North-Eastern USA and the PRT is recommended due to its accuracy, computational simplicity, and ability to estimate design floods for any return period, even though the QRT gave a slightly better performance for all return periods (Ahn and Palmer, 2016). From a study in 1 535 catchments in France, Odry and Arnaud (2017) found that inconsistencies between floods estimated for different return periods were possible when the QRT approach was used and therefore recommend the use of the PRT.

In South Africa studies have been undertaken to develop PRT flood models, however, it has been identified that the availability of catchment descriptor data was a limitation in these studies, with the studies mostly relying on only the *A* and MAP for the development of the regression models (Kjeldsen *et al.*, 2001). More recently, Görgens (2007a) developed the JPV approach which expanded the number of descriptors utilised. In addition, Nathanael *et al.* (2018) assessed the performance of the methods developed by Meigh *et al.* (1997), Mkhandi *et al.* (2000), Görgens (2007a), and Haile (2011) and found that the methods only performed adequately in approximately 57 % of catchments assessed.

The aim of the study is to develop and assess the performance of regional flood models utilising the currently available data. Specific objectives include the following:

(a) The development of PRT and QRT models, developed using Clustering and RoI regionalisation schemes, and considering national and regional scale model development for South Africa.

(b) Assessment and comparison of the developed models to identify the best performing model formulation in combination with regionalisation scheme and model development scale.

5.3 Model Formulations

The development of a regional flood model can be divided into three processes: (i) development of regional growth curves for the hydrological pooling groups, (ii) selection of a suitable response variable/s used to describe and/or scale the regional growth curve, and (iii) developing relationships between catchment descriptors and the response variable/s selected. As per Section 3.3.1, LM (Hosking, 1990) have been adopted as the parameter estimation method to fit the data to the most suitable distribution. Considering the results presented in Chapter 3, the GPA distribution was utilised for the estimation of the at-site quantiles using the observed AMS data. The at-site quantiles were then used to derive at-site and regional growth curves using the clustering and RoI regionalisation schemes presented in Chapter 4.

Four distinct modelling approaches were investigated for the development of models for the estimation of design floods at ungauged sites: (i) QRT, (ii) IF with equal weighting (IF1), (iii) IF with varied weighting (IF2), and (iv) PRM.

5.3.1 Quantile Regression Technique

The US Geological Survey (USGS) adopted a QRT approach that uses the catchment descriptors as predictor variables to estimate the T% AEP peak flow (Q_T) event using predictor variables (*B, C, D, …*) and regression parameters (*a, b, c, d, …*) (Benson, 1962, 1964, Cruff and Rantz, 1965, Riggs, 1973). The regression equations generally take the form of Eqs. 5.1 or 5.2.

$$Q_T = a\,B^b\,C^c\,D^d \ldots \tag{5.1}$$

$$Q_T = a + b*B + c*C + d*D \ldots \tag{5.2}$$

Benson (1962) noted that when developing flood models, using statistical methods such as IF and distribution fitting assumes a relationship between all AEP flood events. This assumption, however, does not hold for the QRT approach as each AEP is considered separately allowing for the identification of significant factors for each AEP. Cruff and Rantz (1965) compared the QRT approach for the 50% and 1% AEP against five other statistical DFE methods, including the IF, in the coastal region of California. The IF and QRT were found to perform best.

More recently investigations have been undertaken to develop regional QRT flood models in Australia (Rahman, 2005, Rahman *et al.*, 2011, Haddad and Rahman, 2012, Haddad *et al.*, 2012). Rahman (2005) developed a set of QRT models for South-East Australia using 88 catchments and a combination of hydrological and climatological descriptors. The models achieved median relative errors ranging between 15 and 39%, proving the viability of the approach. Rahman *et al.* (2011) compared the performance of the PRM and a newly developed QRT. The QRT developed used the same formulation as the PRM and parameter to provide a sound basis for comparison and the QRT models outperformed the PRM approach at the 107 catchments considered. Haddad *et al.* (2012) compared the performance of PRT and QRT approaches at 53 sites in Tasmania and the models provided mixed results, with the PRT performing best for higher AEPs and the QRT for lower AEPs.

5.3.2 Index Flood Method

Dalrymple (1960) describes the methodology for the IF method and divides the approach into two distinct parts:

(a) the development of a dimensionless scaled growth curve for a hydrologically homogeneous region, which relates scaled at-site flood peaks to exceedance probability or return period, and

(b) determining relationships between catchment descriptors and the scaling variable used, e.g. the *MAF*.

The IF method assumes a constant coefficient of variation within pooling groups and the dimensionless growth curve is derived by scaling the at-site AMS values by a SF, referred to as the index flood (Q_{IND}). The scaling in the original IF approach used the *MAF* and another commonly used index flood is the Median Annual Flood (*MEF*) (Robson and Reed, 1999, Kjeldsen *et al.*, 2001, 2002, Nobert *et al.*, 2014). Traditionally the growth curve is applied as shown in Eq. 5.3, whereby the desired T% AEP flood event (Q_T) is related to the index flood by means of a *T*% AEP linked growth factor (GF_T), derived from the dimensionless growth curve.

$$Q_T \quad = Q_{IND} \text{ x } GF_T \tag{5.3}$$

The IF approach has, since the original development, been implemented in international guidelines by Robson and Reed (1999) and Rahman *et al.* (2019) . The implementation has also been modified to develop a site-specific growth curve, through the development of models to estimate the descriptive statistics of the adopted distribution. For example, when considering the LP3 distribution, the standard deviation, mean and coefficient of skewness are required and each of these parameters are individually estimated for a selected site using separate regressions. In theory, estimating each of these parameters allows the models the flexibility to more closely mimic the observed growth curves as opposed to relying on a regionally averaged growth curve.

Relating Q_{IND} to catchment descriptors enables the user to estimate Q_{IND} at an ungauged site. The IF method has been successfully applied in a number of studies including the UK Flood Engineering Handbook (Kjeldsen *et al.*, 2008b), South Africa (HRU, 1972, Kovács, 1988, Van Bladeren, 1993, Mkhandi and Kachroo, 1997, Kachroo *et al.*, 2000, Mkhandi *et al.*, 2000, Kjeldsen *et al.*, 2001, 2002, Görgens, 2007a, Haile, 2011) and Australia for data poor regions (Rahman *et al.*, 2015a).

5.3.2.1 Regional weighting scheme

The regional growth curves are derived by pooling all of the data within the identified pooling groups, but given the variation in record length and data quality within pooling groups, a weighting scheme is generally applied. Three weighting schemes were applied as part of this study to assess the impacts thereof. For both regionalisation schemes an equally weighted IF approach was applied, where each station in the pooling group was applied equal weighting and is referred to as IF1. In the second approach which was applied for both sets of regionalisation, a station weighting scheme was applied (IF2). When using the 42 homogeneous clusters a record length weighting within the clusters was used, thereby applying a higher weighting to stations with a longer available dataset. For the RoI and super region approaches the regional LM were estimated by applying a Euclidian distance (*D*) scaled record length weighting, calculated using Eq. 5.4, for the station weighting scheme.

$$RW_i = \left(\frac{N_i}{\sum N} * \left(1 - \frac{D_i}{\sum D}\right)\right) \Big/ \sum\left(\frac{N_i}{\sum N} * \left(1 - \frac{D_i}{\sum D}\right)\right) \tag{5.4}$$

where

RW_i = regional weighting for site *i*,

N_i = record length for site *i*, and

D_i = Euclidian distance between site *i* and the ungauged estimation location.

This allowed for the integration of catchment similarity, through D_i, into the weighting scheme. The D_i of each site is weighted based on the total D of the pooling group and inverted by subtracting the weighting from one. The D based weighting is then further utilised to scale the traditional record length weighting, which is then finally normalised to provide the final weighting of the site within the pooling group considered.

5.3.2.2 Scaling factor selection

The SFs refers to the variable(s) required to scale the regional flood model and varies based on the model formulation being applied. Some of the SFs used in flood studies include, but are not limited to:

(a) Distribution descriptive statistics (Rahman *et al.*, 2015b),
(b) *MEF* (Robson and Reed, 1999),
(c) *MAF* (Dalrymple, 1960, Kjeldsen *et al.*, 2001), and
(d) average rainfall intensity (McDermott and Pilgrim, 1982).

A critical assumption in the development of regional flood models is that at-site flood responses are assumed to be similar within a pooling group after scaling the at-site growth curve. Hence the choice of a suitable SF is important.

For the development of the IF, the use of the *MAF* or *MEF* are widely reported in the literature. The Institute of Hydrology (IH, 1999) utilised the *MEF* to minimise the impact of outliers, but models developed for South Africa have largely been based on *MAF*. Haile (2011) highlights the existence of outliers in the South African flow datasets when reviewing the data at an at-site basis, but also notes that when considering the data at a regional scale only two observations were considered unacceptably high. Kjeldsen *et al.* (2002) also highlights the AMS records in the South African dataset that generated uncharacteristically high flows, given the short record length, and affecting nearly all South African catchments. Hence, both the *MAF* and *MEF* values were assessed for use in the study, however, the *MAF* value proved to be most representative and reduced the spread of the dimensionless growth curves within the clusters, as shown in Figure 5.2. The *MAF* was therefore adopted for use.

Similarly, the at-site growth curves, LMs and the site *MAF* were used to derive at-site dimensionless growth curves and LMs for the development of the IF1 and IF2 approaches. The growth curves developed for all clusters are included in Appendix B. Reference can be made to Figure 4.12 to identify the location of the clusters. Appendix C and Appendix D contain the IF1 dimensionless growth factors and IF2 LMs, respectively.

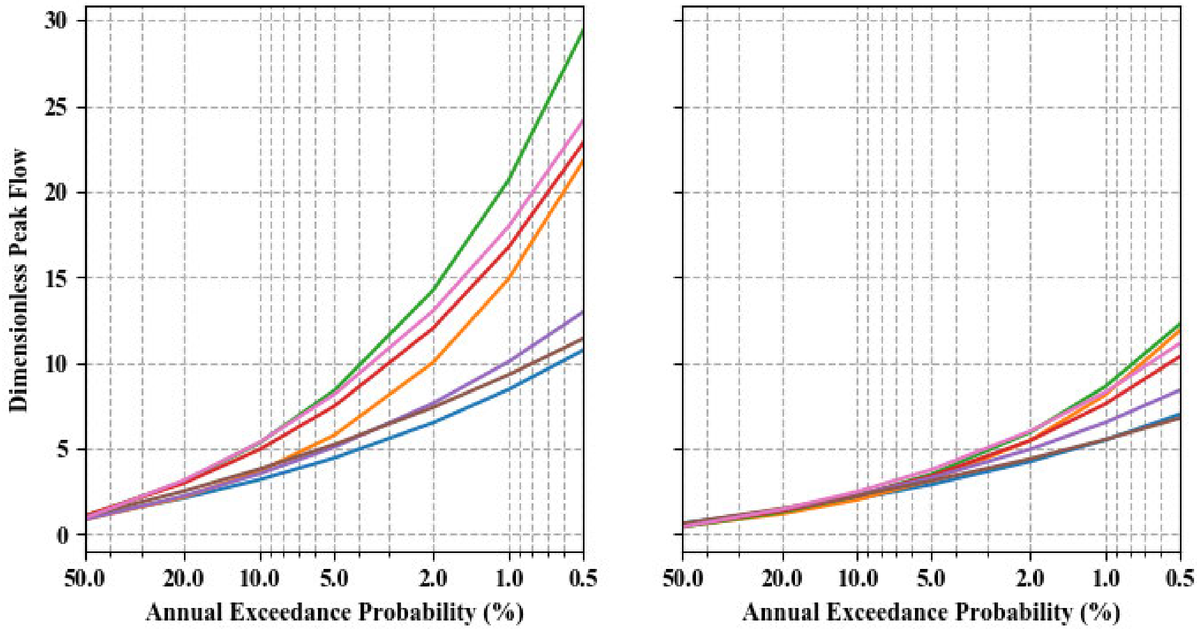


Figure 5.2 Comparison of the dimensionless growth curves for sites within Cluster 42 derived using the *MEF* (left) vs the *MAF* (right) annual floods for selection of a suitable scaling factor

5.3.3 Probabilistic Rational Method

The PRM relies on the traditional Rational Method (Mulvaney, 1850, cited by Stephenson, 1981, Shaw, 1994, Thompson, 2007), which takes the form shown in Eq 5.5 and relates the T% AEP peak flow (Q_T) to the A (km^2), design rainfall intensity (mm/h) for a known critical duration T_c ($I_{(Tc,\ T)}$), and the dimensionless runoff coefficient C_T. A unit factor (U_F) is also incorporated to convert peak flows to the desired units (m^3/s). To calibrate the RM, the relationship between the runoff coefficient C_T and the remaining parameters needs to be defined for each AEP and is provided in Eq 5.6.

$$Q_T = U_F\, C_T\, I_{(Tc,T)}\, A \qquad (5.5)$$

$$C_T = \frac{Q_T}{U_F\, A\, I_{(Tc,T)}} \tag{5.6}$$

5.3.3.1 PRM applications

McDermott and Pilgrim (1982) developed a PRM for Australia, which utilises regional runoff coefficient (*C* value) curves scaled by the 10% AEP *C* value (C_{10}). Similar to the original IF approach the *T*% AEP *C* value is estimated through the use of GF_T derived from the regional *C* value curves. McDermott and Pilgrim (1982) developed regional C_{10} maps by mapping the at-site C_{10} values and manually drawing isopleths. The approach developed by McDermott and Pilgrim (1982) estimated C_{10} values with an accuracy of 30% at 42% of sites and was adopted in the Australian Rainfall and Runoff manual as one of the recommended methods. Alsuwaidi *et al.* (2015) developed a new PRM method, through assessing twelve different forms of PRM, in New South Wales, Australia, and achieved good results in small to medium sized catchments. The best performing PRM form was linking the inverse distance weighted C_{10} value of the three nearest sites coupled with the median GF_T for all 106 sites considered. The latest revision of the ARR (Rahman *et al.*, 2019) has, however, recommended an alternative approach to the PRM and which is based on the PRT approach which yields improved results.

In the South African context, the SDF method developed by Alexander (2002a) is a locally developed PRM. However, the method has been recommended for revision in a number of studies (Görgens, 2002, Smithers and Schulze, 2003, Van Bladeren, 2005, Gericke, 2010, Van Vuuren *et al.*, 2013).

Görgens (2002) found that when estimating the 2% AEP floods the SDF estimates could be up to 210% in excess of the observed estimates. In the development of the SDF, Alexander (2002a) does state that conservative “upper envelope” coefficients were derived, which could cause the overestimation, but are within the uncertainty levels related to hydrological estimation. Smithers and Schulze (2003) expressed the need to assess the SDF method and provide further refinement. Van Bladeren (2005) proposed modifications to the SDF method, but in the DWS C5 secondary drainage region these only resulted in improved estimates in 26% of the catchments assessed. Gericke (2010) reviewed the SDF method and found that the SDF overestimated design floods by up to 230% in the C5 secondary drainage region. Gericke (2010) also determined correction factors for the SDF method. The corrected SDF method provided the most accurate results in the majority of the study area (Gericke, 2010). The ratios of calibrated SDF: FFA ranged between 0.85 and 1.15, and resulted in a major improvement on the standard SDF results. Van Vuuren *et al.* (2013) identified inconsistencies in the estimation of catchment parameters during the development of the SDF as one of the potential problems that needs further research and refinement.

Calitz (2016) developed a PRM for selected regions in South Africa and successfully regionalised the flood distributions in ten homogeneous regions, which were used to calibrate PRM C value relationships for the estimation of the design floods. Calitz (2016) investigated the use of mapping and regression development for estimation of C_{10}, coupled with GF_T estimated as the regional median or through regressions. The best performing combination was the use of regression estimated C_{10} values coupled with the median GF_T, and as such has been adopted in this study for application of the PRM.

5.3.3.2 C_T value growth curves

By applying Equation 5.6 at-site C_T value growth curves were developed for each site being considered. The calibrated C_T values were found to have some inconsistencies, as summarised in Appendix A for each pooling group. Similar to the findings of Parak and Pegram (2006), the calibrated C_T values were found at some sites to not be consistent with the assumption that the C_T values should increase with a decrease in AEP. This occurred at 79 of the 383 sites investigated. Parak and Pegram (2006) identified that the calibrated C_T values used in their study were within reasonable bounds when compared to Chow *et al.* (1988) and hence tentatively included the inconsistent results for the remainder of their study. Similarly, the sites with a decrease of C_T with AEP have been tentatively included in this study. It should also be noted that AEP's lower than 5% contain C_T values in excess of 1 at six sites for the 0.5% AEP and at a single site for 5% AEP. These values could have been the result of design rainfall estimates being restricted to use of the median values and the use of a catchment based average design rainfall. Smithers and Schulze (2003) do, however, provide upper and lower 90% confidence bounds for estimates and these bounds could be investigated to restrict the C_T values to not exceed 1. Similarly, the use of the median, or an alternative percentile of point rainfall within the catchment could be used to reduce the estimated C_T values. Attaining C_T values below 1 are, however, not critical to the results of the study, and, is only anticipated to affect the uptake of the model, should it be identified as the most suitable approach in terms of quantile estimation.

Having calibrated the RM at each site, regional dimensionless growth curves were developed using the C_{10} value to scale the individual at-site curves. An example of the dimensionless *C* value curves for Cluster 42 is shown in Figure 5.3and the *C* value curves for all clusters are included in Appendix B. Reference can be made to Figure 4.12 to identify the location of the clusters. Appendix E contains the detailed cluster-based C_T.

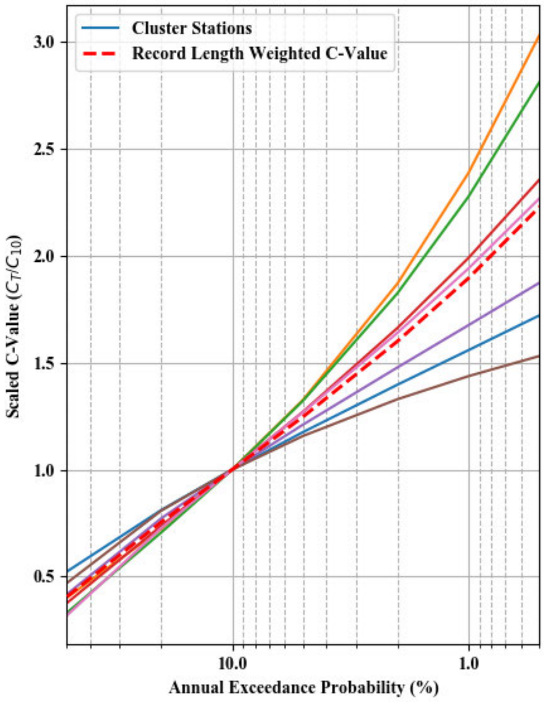


Figure 5.3 Scaled PRM growth curves for Cluster 42 indicating the record length weighted average curve (red dash) in relation to the sites within the cluster (coloured solid lines)

5.4 Development of regression models to estimate scaling factors

Having identified the C_{10} and *MAF* as suitable SFs for use, the required set of response variables was complete and are shown in Table 5.1. These variables are now required to be estimated at ungauged sites, which was achieved through the use of regression models developed using the ordinary least squares framework. When considering the development of the QRT models emphasis was placed in the 1% AEP value for the selection of important catchment descriptors as predictor variables.

Table 5.1 Scaling Factors for each of the selected model formulations

Model Formulation	Scaling Factor / Response Variable
QRT	Q_T
PRM	C_{10}
IF1 (equal weighting)	*MAF*
IF2 (varied weighting)	*MAF*

Kjeldsen *et al.* (2008b) reported on the development of a national scale regression model in the UK linking the IF to four different catchment descriptors, thereby providing a model for estimating the IF in ungauged catchments. In contrast, Rahman *et al.* (2019), developed a set of regional regression models for Australia as a result of hydrological variability and variation in station density across the country. The advantage of using a national scale SF regression is that more data are utilised for the development, reducing the effect of potential outliers, however, any regional trends may be lost in the overall analysis. Given the similarity of climate variability between South Africa and Australia, both national and regional scale models were developed and assessed for application in South Africa.

The clustering and RoI approaches described in Chapter 4 were adopted for the identification of pooling groups, within which the SF regression models were developed at two scales, national and regional. This resulted in four different SF model development scenarios being adopted: (i) Clustering with cluster-based models, (ii) Clustering with a national scale model, (iii) RoI with super region based models, and (iv) RoI with a national scale model.

The pooling groups vary in size and contain as few as three sites. Developing regressions fitted to a small number of sites could severely impact the ability of the models to be used. In an attempt to improve the robustness of the models a minimum number of required sites was imposed for regression development. Where pooling groups contain less than 30 stations (approximately 10 % of the available data), the closest geographic pooling groups were included until a minimum of 30 sites was reached. Geographic proximity was defined by the distance between pooling group centroids. The minimum number of stations imposed is a consideration that can be refined in future research.

5.4.1 Catchment descriptors considered for regression development

McDermott and Pilgrim (1982) divided catchment descriptors into two groupings, natural and introduced. Natural variables refer to descriptors such as area, soil, rainfall, and topography, whereas introduced variables, consider man-made effects such as land use and urbanisation. Since introduced variables change more rapidly than natural variables, they are often difficult to quantify. It is evident from the literature that the use of natural variables is widely adopted (Dalrymple, 1960, Riggs, 1973, McDermott and Pilgrim, 1982, Pilgrim, 1989, Mkhandi and Kachroo, 1997, Mkhandi *et al.*, 2000, Kjeldsen *et al.*, 2001, 2002, Smithers and Schulze, 2003, Merz and Blöschl, 2005, Görgens, 2007a, Rao and Srinivas, 2008, Rahman *et al.*, 2015b). Rahman *et al.* (2009, 2012) identified that increasing the number of predictor variables does not necessarily increase the accuracy of the flood model and has a diminishing returns effect. Out of a pool of ten potential predictor variables, five were used for the final Australian Regional Flood Frequency Estimation model and the selected variables consisted solely of natural variables, such as A, shape and rainfall intensity (Rahman *et al.*, 2015b).

Historically the availability of catchment descriptors have been identified as a limitation on the development of regional models in South Africa as only limited data was available (Kjeldsen *et al.*, 2001). Based on a review of the literature and the descriptors extracted as described in Chapter 2, eight descriptors were considered as potential regression predictor variables, shown in Table 5.2.

Table 5.2 Catchment descriptors identified as potential predictor variables for the development of the SF regression models

Descriptor	Unit	Range		Source
		Min	Max	
Outlet elevation	masl	11.00	1969.00	(NASA-JPL, 2013)
A	km^2	0.26	361994.80	DEM Derived
C_{ro}	Percent	4.00	97.00	(Schulze, 2011)
D_c	Decimal Degrees	0.03	6.84	(Smithers and Schulze, 2003)?
Slope	m/m	0.0004	0.26	DEM Derived
MAP	mm	60.00	3312.00	(Lynch, 2004, de Groen *et al.*, 2015)
T_c linked 10% AEP design rainfall depth	mm	15.60	241.50	(Smithers and Schulze, 2003)
24 hour 1% AEP deign rainfall depth	mm	72.50	524.30	(Smithers and Schulze, 2003)

The selection of the catchment descriptors for use as predictor variables in the regressions models was based on the p-value of the model parameters. The p-value estimates the probability of the assumption that the inclusion of the descriptor parameter has no effect on the model. This is termed the null hypothesis. A level of significance of 0.05 has been selected for the study, hence any parameter with a p-value in excess of 0.05 is rejected. A p-value of 0.05 signifies a 5% probability of the null hypothesis being correct. Rahman *et al.* (2011) noted that models containing a large number of predictor variables become difficult to apply and highlight that acceptably accurate results can be achieved with as few predictor variables as two or three. Given the number of models to be developed in the study and the findings of Rahman *et al.* (2011) a maximum number of three parameters has been adopted. The selection process of the most significant catchment descriptors was undertaken by including all the chosen descriptors as an initial model, thereafter descriptors with the highest p-value or with a p-value in excess of 0.05 were eliminated until a maximum of three descriptors remained. Further refinement of the best performing model is recommended as further research.

5.4.2 Model development within the clustering scheme

For both the C_{10} and *MAF* the most significant descriptors was identified on a national scale, the adopted descriptors were the *A* (km^2), *MAP* (mm) and the D_c (decimal degree), whereas the most significant descriptors for $Q_{1\%}$ were *A*, 10% AEP design rainfall ($DR_{10\%}$) and the D_c. Equation 5.7 provides the equation for the estimation of the SFs (C_{10} or *MAF*) and Equation 5.8 provides the adopted model for the QRT model based on the estimation of the 1% AEP flood quantile.

$$Ln(SF) = a * Ln(A) + b * Ln(MAP) + c * Ln(D_c) + Const \tag{5.7}$$

$$Ln(Q_T) = a * Ln(A) + b * Ln(DR_{10\%}) + c * Ln(D_c) + Const \tag{5.8}$$

where

SF = Scaling Factor (*MAF* or C_{10}),

a, b, c = model coefficients, and

Const = intercept (constant).

The adopted SF model coefficients, using the minimum station requirement of 30 sites, are contained in Table 5.3 and a comparison between the observed and estimated values are shown in Figure 5.4 for the C_{10} and *MAF* models at both National and regional scales. The regression statistics for the *MAF* and C_{10} models are listed in Appendix F. In addition, the same approach was adopted for the estimation of the QRT approach and the model coefficients for each cluster and AEP are provided in Appendix G.

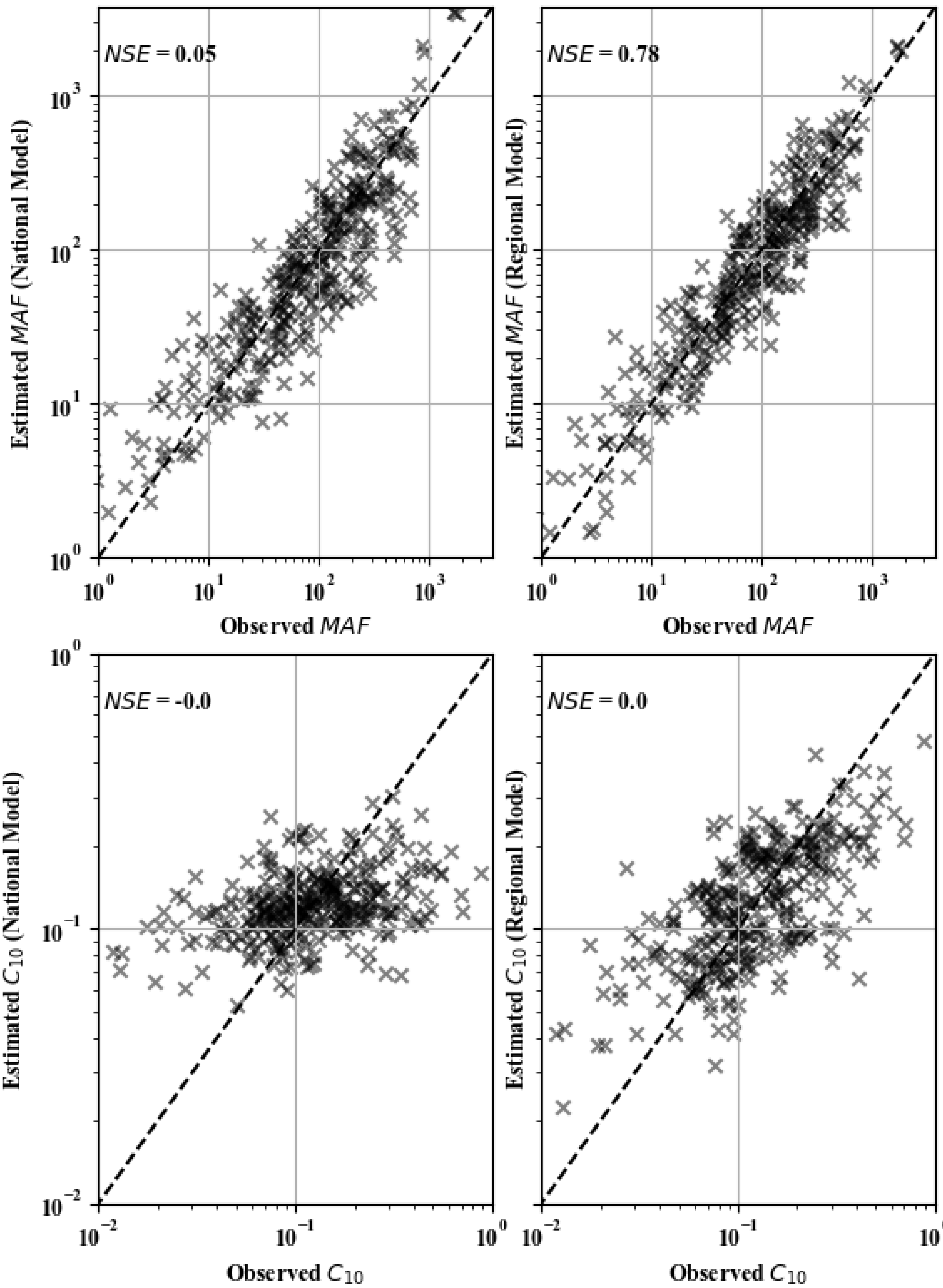


Figure 5.4 Observed versus estimated *MAF* (top) and C_{10} (bottom) for the national- (left) and cluster based (right) models for the clustering regionalisation

Table 5.3 Scaling Factor model parameter coefficients for estimation of the *MAF* and C_{10} for application with the IF1/IF2 and PRM models using the 42 relatively homogeneous clusters

Scale	No. of Sites	*MAF*					C_{10}				
		Constant	Area	*MAP*	Distance from Coast	R^2	Constant	Area	*MAP*	Distance from Coast	R^2
National	332	-5.88	0.69	0.90	-0.24	0.79	-4.26	0.11	0.24	-0.26	0.13
Cluster 1	32	-18.32	0.49	3.19	-0.82	0.76	-12.99	-0.11	2.09	-1.39	0.33
Cluster 2-3	30	2.34	0.50	0.43	-2.52	0.68	12.02	-0.14	-1.37	-2.82	0.45
Cluster 4	43	-39.34	0.63	6.03	0.23	0.83	-39.39	0.06	5.67	-0.10	0.34
Cluster 5	35	-6.01	0.64	1.00	-0.30	0.64	-5.78	0.13	0.51	-0.71	0.13
Cluster 6	35	-21.52	0.66	3.09	1.13	0.85	-8.92	0.08	0.85	0.65	0.11
Cluster 7	35	-18.34	0.70	2.60	0.79	0.86	-11.35	0.10	1.21	0.46	0.15
Cluster 8	37	-35.46	0.64	5.10	1.96	0.85	-18.84	0.02	2.32	1.24	0.29
Cluster 9	31	-6.40	0.80	0.76	0.52	0.83	3.67	0.19	-1.08	0.02	0.40
Cluster 10	33	-8.67	0.85	1.08	0.28	0.79	-6.82	0.29	0.29	1.02	0.30
Cluster 11	35	-6.01	0.64	1.00	-0.30	0.64	-5.78	0.13	0.51	-0.71	0.13
Cluster 12	30	-5.23	0.74	0.73	0.16	0.73	-3.47	0.16	0.05	-0.12	0.16
Cluster 13	30	-10.59	0.51	1.77	0.20	0.59	-5.55	-0.03	0.51	0.65	0.05
Cluster 14	32	-8.95	0.62	1.41	-0.17	0.90	4.82	0.04	-0.79	-1.52	0.21
Cluster 15	47	-8.45	0.63	1.37	-0.15	0.75	-0.17	0.09	-0.29	-0.17	0.09
Cluster 16	42	-8.38	0.57	1.40	0.02	0.69	1.06	0.04	-0.44	-0.06	0.05
Cluster 17	31	-11.82	0.70	1.83	-0.03	0.74	-7.39	0.07	0.82	-0.01	0.03
Cluster 18	33	-9.27	0.82	1.37	-0.32	0.80	-9.08	0.24	0.90	-0.40	0.28
Cluster 19	42	-3.36	0.66	0.54	-0.52	0.89	-4.76	0.06	0.38	-0.35	0.22
Cluster 20 - 24	32	-13.04	1.02	1.87	0.39	0.88	-11.03	0.36	1.16	0.08	0.45
Cluster 25	34	-5.17	0.74	0.77	-0.32	0.85	-4.81	0.14	0.27	-0.50	0.37
Cluster 26 - 27	34	-2.16	0.69	0.30	-0.50	0.82	-2.56	0.08	-0.03	-0.55	0.29
Cluster 28	34	-3.21	0.67	0.50	-0.38	0.84	-3.97	0.08	0.22	-0.38	0.30
Cluster 29	42	0.15	0.47	0.17	-0.41	0.71	3.53	-0.21	-0.65	-0.36	0.15

Scale	No. of Sites	*MAF*					C_{10}				
		Constant	**Area**	***MAP***	**Distance from Coast**	R^2	**Constant**	**Area**	***MAP***	**Distance from Coast**	R^2
Cluster 30 - 32	33	-7.48	0.68	1.16	-0.16	0.84	-3.44	-0.01	0.26	-0.13	0.02
Cluster 33	31	-2.45	0.67	0.45	-0.05	0.74	4.72	0.10	-1.03	-0.05	0.13
Cluster 34	46	4.42	0.68	-0.55	-0.05	0.80	8.71	0.09	-1.60	-0.08	0.21
Cluster 35	37	2.34	0.61	-0.20	-0.02	0.70	10.03	0.03	-1.75	-0.02	0.13
Cluster 36	33	-6.21	0.60	1.06	0.10	0.70	2.85	0.04	-0.71	-0.01	0.10
Cluster 37	32	4.15	0.66	-0.50	-0.06	0.79	8.23	0.07	-1.52	-0.08	0.16
Cluster 38 - 39	43	-4.41	0.77	0.53	0.05	0.82	5.48	0.17	-1.30	-0.34	0.35
Cluster 40	34	-7.92	0.78	1.05	-0.09	0.85	-6.23	0.20	0.36	-0.26	0.32
Cluster 41	33	-10.56	0.90	1.38	-0.22	0.79	-9.58	0.34	0.75	0.06	0.32
Cluster 42	32	-11.46	0.85	1.48	0.32	0.83	-10.36	0.27	0.88	0.20	0.28

5.4.3 Model development within the Region of Influence scheme

When considering the RoI approach on a national scale, the *MAF* and QRT regressions developed in Section 5.4.2 were unchanged. The regressions for the remaining modelling approaches varied and the significant descriptors were also estimated at a Super Region scale and are provided in Table 5.4. The regression parameters for the estimation of the *MAF* and C_{10} are provided in Table 5.5, and the QRT regression parameters are provided in Appendix H. Figure 5.5 shows the estimated versus observed plots for the estimation of the *MAF* and C_{10}, from the plots it can be seen that the models perform poorly for the estimation of C_{10}, achieving a maximum *NSE* of 0.28, and similarly the estimation of the *MAF* using the national model also performs poorly. The regional *MAF* model however performs well, achieving an *NSE* of 0.69. This is further substantiated by the R^2 values achieved by the models, shown in Table 5.5, where the *MAF* regressions achieve values between 0.78 and 0.83, and conversely the C_{10} regressions achieve values between 0.18 and 0.45.

Table 5.4 Super region *SF* predictor variables per modelling approach

Model	Dependent Variable	Super Region	Predictor Variables
IF1 and IF2	*MAF*	National	A, MAP, D_c
		1	A, MAP
		2	A, $S_{10\text{-}85}$, $DR_{10\%}$
		3	A, E_O, D_c
		4	A, 24-hour 1% AEP design rainfall depth ($H24_{1\%}$)
		5	A, MAP, $S_{10\text{-}85}$
PRM	C_{10}	National	D_c, MAP, $S_{10\text{-}85}$
		1	A, MAP, $H24_{1\%}$
		2	$DR_{10\%}$, E_O, $H24_{1\%}$
		3	E_O, C_{ro}, $H24_{1\%}$
		4	$DR_{10\%}$, C_{ro}, $H24_{1\%}$
		5	A, MAP, E_O
QRT	$Q_{1\%}$	National	A, $DR_{10\%}$, D_c
		1	A, $DR_{10\%}$, $H24_{1\%}$
		2	A, $S_{10\text{-}85}$, $DR_{10\%}$
		3	A, E_O, $H24_{1\%}$
		4	A, E_O, MAP
		5	A, MAP, $S_{10\text{-}85}$

Table 5.5 Scaling Factor model parameter coefficients and regression statistics for estimation of the *MAF* and C_{10} for application with the IF1, IF2 and PRM models using the 5 super regions

Scaling Factor	**Super Region**	**Predictor Variables**	**Predictor Variable Coefficients**				**Performance Metrics**				**Descriptor Standard Error**			
			Const	**1**	**2**	**3**	**AIC**	**BIC**	**R^2**	***RMSE***	**Const**	**1**	**2**	**3**
MAF	National	A, *MAP*, D_c	-5.88	0.69	0.90	-0.24	848.99	864.78	0.79	214.86	0.78	0.02	0.11	0.04
	1	A, *MAP*	-21.02	0.66	3.25	-0.18	227.22	237.91	0.82	122.49	5.22	0.03	0.74	0.37
	2	A, $S_{10\text{-}85}$, $DR_{10\%}$	-3.99	0.70	0.65	-0.06	151.07	160.28	0.76	136.91	3.96	0.05	0.56	0.09
	3	A, E_O, D_c	-11.02	0.90	1.59	-0.05	183.78	193.31	0.83	165.83	1.60	0.05	0.22	0.10
	4	A, $H24_{1\%}$	-8.45	0.80	1.07	0.71	178.62	188.58	0.80	67.62	2.63	0.05	0.35	0.22
	5	A, *MAP*, $S_{10\text{-}85}$	-4.02	0.61	0.69	-0.19	35.71	41.70	0.79	64.21	1.90	0.06	0.26	0.10
C10	National	D_c, *MAP*, $S_{10\text{-}85}$	-5.17	0.12	0.36	-0.26	824.86	840.66	0.16	0.12	0.76	0.02	0.11	0.04
	1	A, *MAP*, $H24_{1\%}$	-20.01	0.11	2.74	-0.52	215.15	225.84	0.26	0.08	4.93	0.03	0.70	0.35
	2	$DR_{10\%}$, E_O, $H24_{1\%}$	2.43	0.12	-0.72	-0.08	147.46	156.67	0.14	0.14	3.87	0.05	0.55	0.09
	3	E_O, C_{ro}, $H24_{1\%}$	-8.64	0.24	0.84	-0.19	179.00	188.53	0.29	0.14	1.56	0.05	0.22	0.10
	4	$DR_{10\%}$, C_{ro}, $H24_{1\%}$	-4.64	0.20	0.11	0.43	169.89	179.85	0.23	0.08	2.50	0.05	0.34	0.21
	5	A, *MAP*, E_O	-5.02	0.05	0.40	-0.15	60.37	66.35	0.10	0.07	2.75	0.09	0.38	0.15

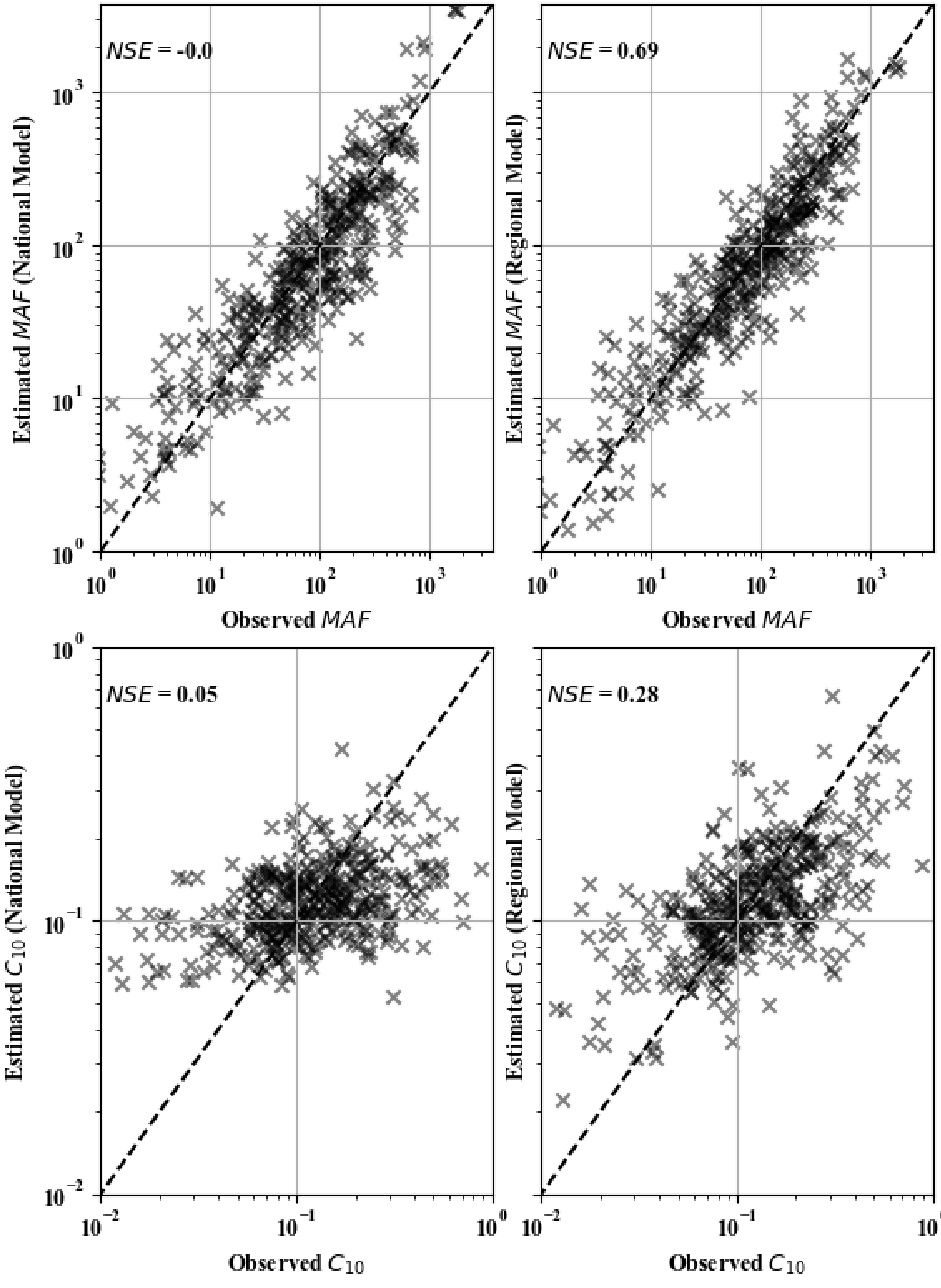


Figure 5.5 Observed versus estimated *MAF* (top) and C_{10} (bottom) for the national- (left) and super region based (right) models for the RoI and super region regionalisation

5.5 Q_T Estimation Performance Assessment

Prior to application of the methods on a national scale, the method that yields the best performance statistics needs to be identified. The performance statistics utilised to assess the performance of the regionalisation methods are the *BIAS*, relative *BIAS* ($BIAS_r$), Root Mean Square Error (*RMSE*) relative *RMSE* ($RMSE_r$) as adopted by Gado and Nguyen (2016). Additional measures used for the estimation of model accuracy were the Nash-Sutcliffe model efficiency coefficient (*NSE*) (Nash and Sutcliffe, 1970) and Relative Error (*RE*). The *NSE* is generally used for the estimation of the efficiency of continuous models and when considering regression analysis is equivalent to the coefficient of determination (R^2). The relative errors of the models give an indication of the performance relative to the growth curve derived peak flows. Rahman *et al.* (2012) and Naidoo (2020) utilised the ratio of modelled vs estimated values as an indication of model performance. Rahman *et al.* (2012) categorised ratios of 0.5 – 2 as acceptable, whereas Naidoo (2020) provided additional categories of potentially acceptable for over and under-estimation, the entire range of potentially acceptable ratios ranges between 0.5 and 1.5. Rahman *et al.* (2012) notes that the limits provided are arbitrary limits but provide a guide with regards to relative accuracy between models. The *RE* of -50% - + 100% and -50% - +50% are equivalent to the Rahman *et al.* (2012) and Naidoo (2020) ratios respectively. Eqs. 5.9 - 5.14 provide the six performance statistics adopted, which compare the modelled (Q_m) with the observed (Q_o) design flows for a set of *n* sites.

$$RMSE^{PG,T} = \sqrt{\frac{1}{n}\sum(Q_m - Q_o)^2} \tag{5.9}$$

$$RMSE_r^{PG,T} = \sqrt{\frac{1}{n}\sum\frac{(Q_m - Q_o)^2}{(Q_o)}} \tag{5.10}$$

$$BIAS^{PG,T} = \frac{1}{n}\sum|Q_m - Q_o| \tag{5.11}$$

$$BIASr^{PG,T} = \frac{1}{n}\sum\left(\frac{|Q_m - Q_o|}{Q_o}\right) \tag{5.12}$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}\left(Q_m^i - Q_o^i\right)^2}{\sum_{i=1}^{n}\left(Q_o^i - \bar{Q}_o\right)^2} \tag{5.13}$$

$RE^{i,T} = 100 \times \frac{Q_{m-}Q_o}{Q_o}$ (5.14)

where,

$RMSE^{PG,T}$	= root mean squared error ($m^3.s^{-1}$) for each pooling group (*PG*) and AEP% (*T*),
$BIAS^{PG,T}$	= bias for each PG and AEP% (*T*),
$RE^{i,T}$	= relative error for each site (*i*) and AEP% (*T*),
$\bar{Q}_o$	= mean of the observed design flows, and
r	= indicates relative values (*BIASr* and $RMSE_r$)

Rahman *et al.* (2012) utilised a "Leave-one-out" (LOO) assessment approach to assess the performance of models. This approach "hides" each gauging station from the model development for a single iteration, hence creating a number of models equal to the number of stations being considered, plus an iteration including all gauging stations. This facilitates a statistical test of proof of concept, which thereafter allows for the use of all sites in the final model development. After the LOO resampling, the evaluation statistics can be computed for the final Q_T values estimated using the developed models.

The chosen performance metrics can be categorised into two groups: individual performance and relative performance, respectively. The *NSE* and *RE* are classified as individual performance metrics that indicate a model's ability to estimate the observed design peak discharges, with an ideal result for the metrics being 1 and 0, respectively. *BIAS* and *RMSE*, however, are generally used as relative performance metrics, providing an indication of which model, from a number of models, performs best. As such, the *NSE* and RE will first be employed to refine the number of models to the four best performing models. The top four models will then be compared through the use of *RMSE* and *BIAS* to identify the best performing single model.

5.5.1 Model accuracy and relative error

Figure 5.6 shows the modelled versus the observed Q_T for the IF1, IF2, PRM, and QRT approaches when using regional scale regressions, respectively, in the 42 relatively homogeneous clusters. It is evident that the PRM model does not perform adequately, and IF2 performs adequately for AEPs greater than 2%. The best performing cluster based modes are

IF1 and QRT and nearly equally well. The IF1 and QRT models achieve *NSE* values between 0.54 and 0.77 and slopes between 0.87 and 1.07.

Figure 5.7 provides the estimated versus observed plots using regionally based regressions for each modelling approach and AEP considered using RoI regionalisation. In comparison to the cluster based models, the RoI models do not perform as well, with *NSE* values ranging between -0.34 and 0.69, and the slopes ranging between 0.76 and 1.13. The best performing RoI models are the IF1 and the IF2 and the models perform nearly identically well. The model performance, however, deteriorates for AEPs below 5% with *NSE* values dropping below 0.55 and as low as 0.35 in these ranges.

Table 5.6 provides the ranking of the modelling approaches based on the *NSE* achieved for all AEPs and regionalisation type. The top ranked modelling approaches are the QRT, IF1 and IF2, respectively, within the clustering framework and regional scale model development. These models performed well when considering the slope between estimated and observed values, with majority of the slopes ranging between 0.9 and 1.1.

A breakdown of the percentage of estimates that are within, under or in excess of the desirable *RE*s prescribed by Rahman *et al.* (2012) and Naidoo (2020) are provided in Figure 5.8 and Figure 5.9. When comparing the percentages of sites within the desirable *RE*s between modelling frameworks it is evident that, even though some of the models perform better, the improvement is limited to a maximum of 12.2%, which when considering the uncertainty inherent in hydrological estimations is not a significant improvement.

Reviewing the results for the models developed in the 42 homogeneous clusters the IF1 and IF2 models perform best. The performance of the IF2, however reduces for AEPs lower than 2%, the QRT model also indicates a similar trend of reduced accuracy for AEPs less than 5%. The IF1 and PRM models show a similar level of consistency of estimates across all AEPs, however, the PRM model tends more to over estimation than the IF1. Similar trends in the models are observed in the models developed using the RoI approach. The best performing models are the IF1 and PRM, however the PRM accuracy reduces and tends to overestimate for AEPs less than 2%.

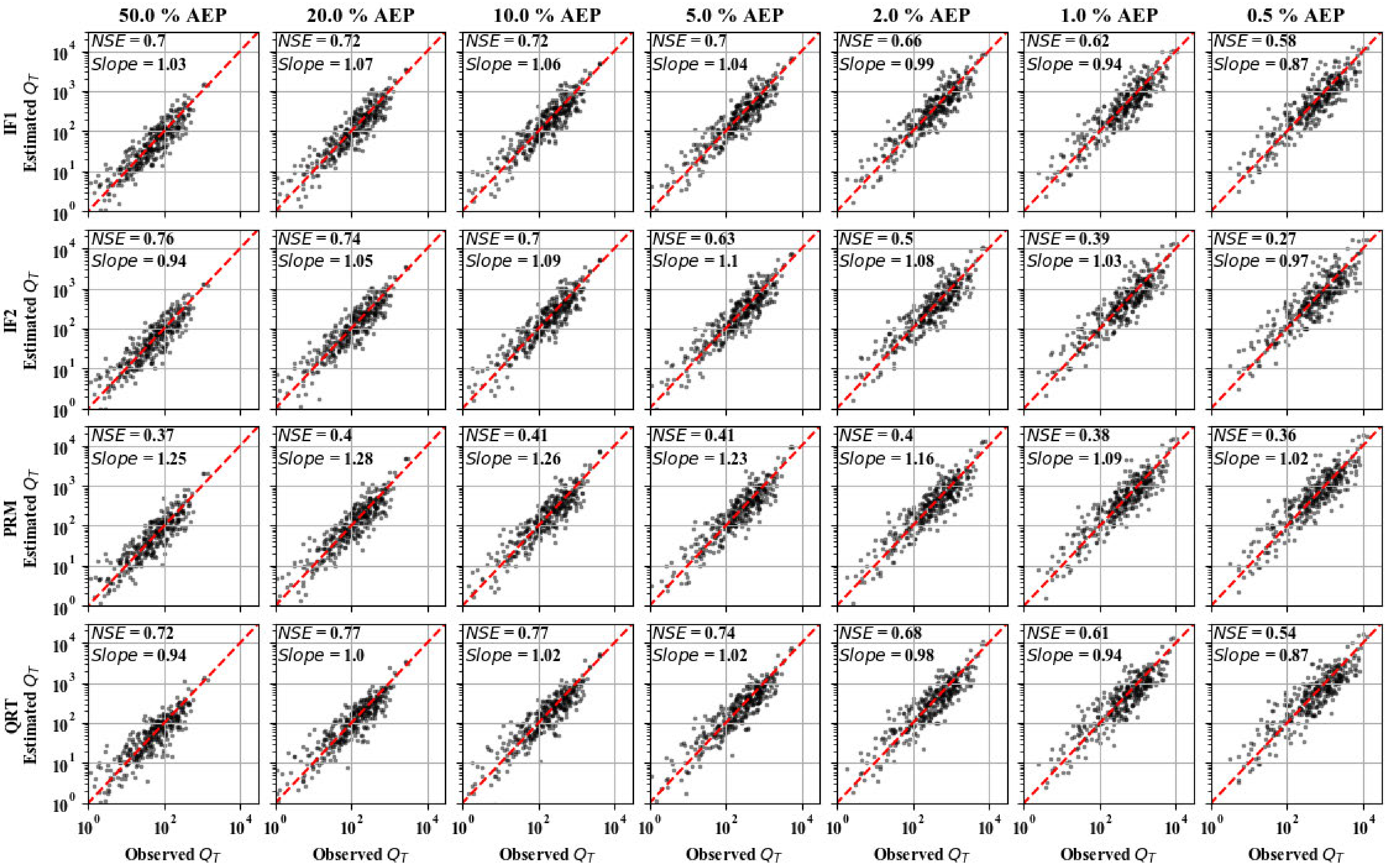


Figure 5.6 Estimated vs observed Q_T for the IF1, IF2, PRM and QRT approaches for the 50, 20, 10, 5, 2, 1 and 0.5% annual exceedance probabilities utilising the 42 homogeneous clusters

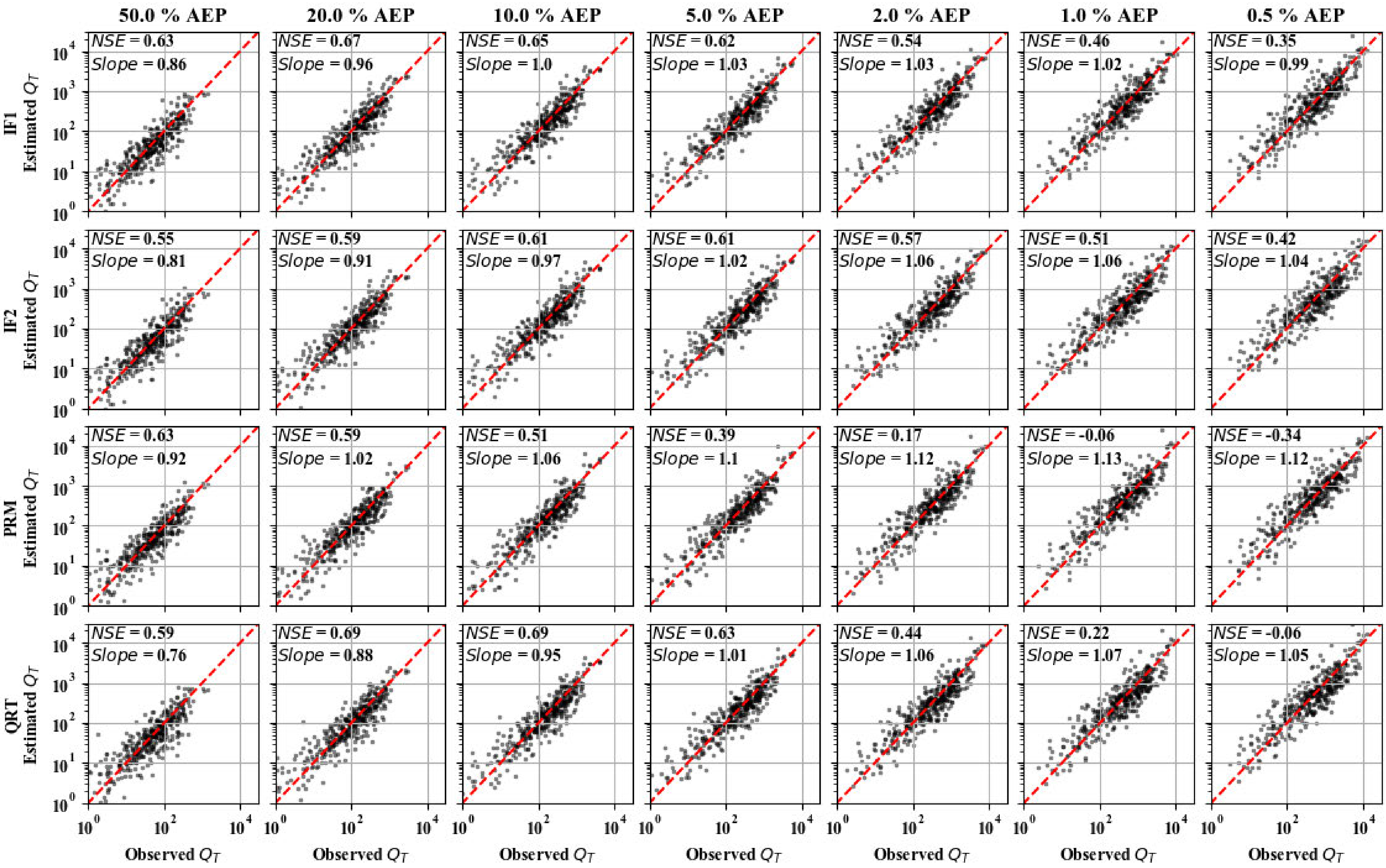


Figure 5.7 Estimated vs observed Q_T for the IF1, IF2, PRM and QRT approaches for the 50, 20, 10, 5, 2, 1 and 0.5% annual exceedance probabilities utilising the region of influence regionalisation methodology

Table 5.6 Ranking of the IF1, IF2, PRM and QRT models based on the *NSE* scores achieved through the application of national and regional scale regressions

Scale	Region Type	Model	Rank per AEP (%)*							Overall Rank
			50	20	10	5	2	1	0.5	
Regional	Cluster	IF1	**3**	**3**	**2**	**2**	**2**	**1**	**1**	**2**
		IF2	**1**	**2**	**3**	**3**	5	5	6	**3**
		PRM	11	8	8	7	7	6	4	7
		QRT	**2**	**1**	**1**	**1**	**1**	**2**	**2**	**1**
	RoI	IF1	4	5	5	5	4	4	5	4
		IF2	7	6	6	6	**3**	**3**	**3**	5
		PRM	4	6	7	8	8	9	11	8
		QRT	6	4	4	3	6	7	8	6

* Numbers in bold indicate the top three performing models

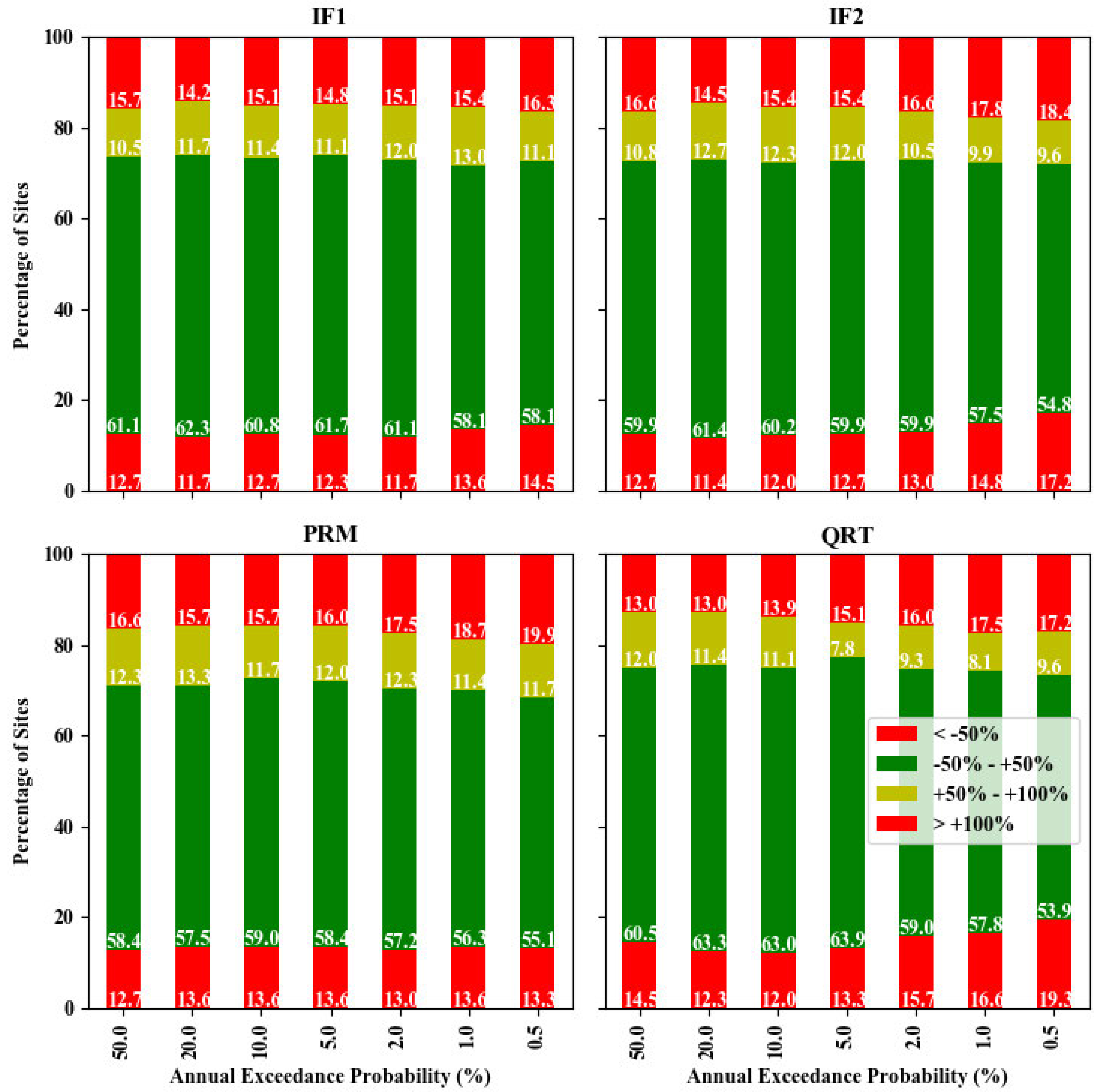


Figure 5.8 Percentage of Q_T estimates within the desirable *RE* ranges defined by Rahman *et al.* (2012) and Naidoo (2020) for the IF1, IF2, PRM and QRT approaches for both regional scale regressions and the 42 homogeneous clusters

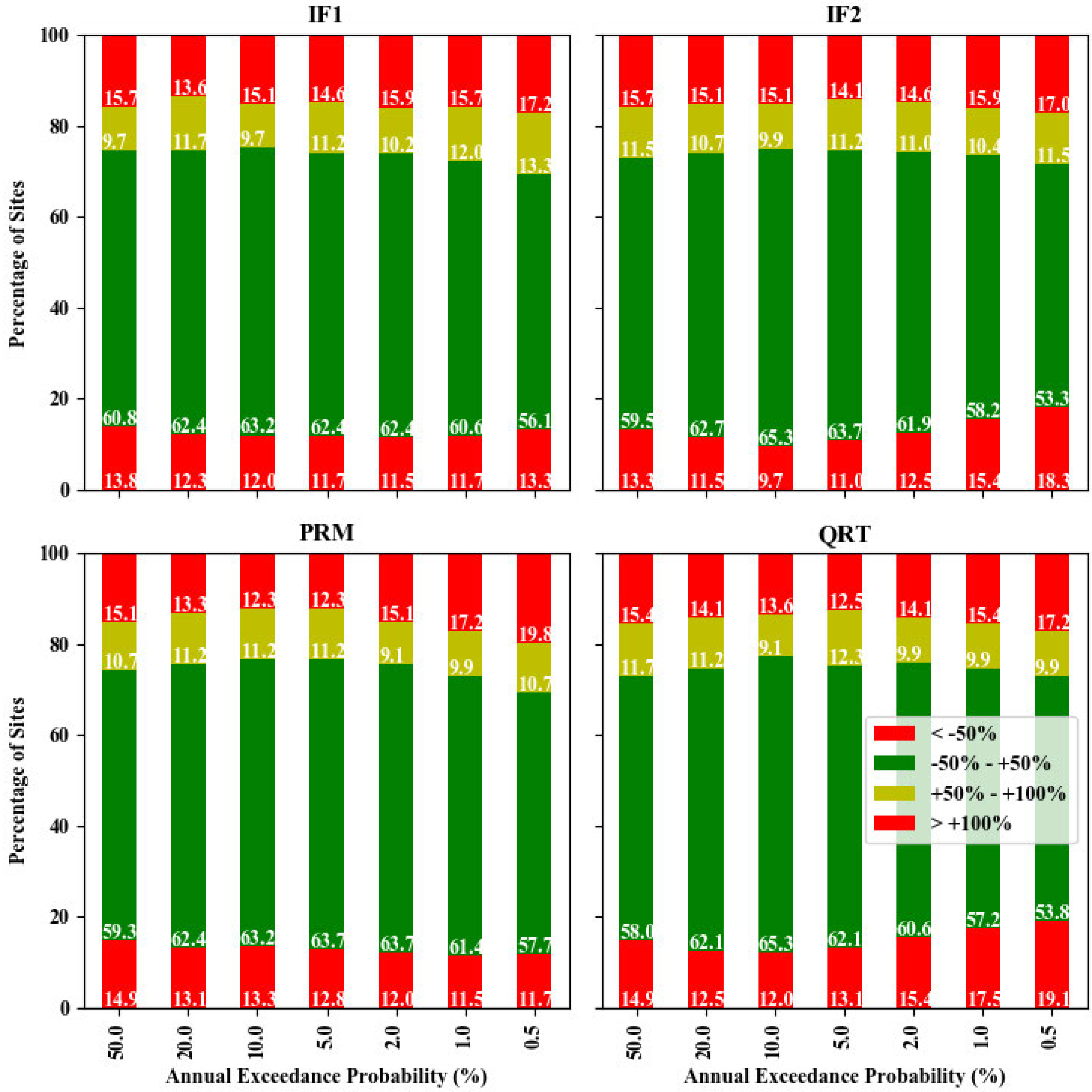


Figure 5.9 Percentage of Q_T estimates within the desirable *RE* defined by Rahman *et al.* (2012) and Naidoo (2020) using the IF1, IF2, PRM and QRT approaches for regional scale regressions and the region of influence regionalisation approach

As an additional an analysis of the performance of the *RE* was compared relative to the catchment areas as shown in Figure 5.10 and Figure 5.11. The results from the 42 homogeneous cluster models indicate a higher level of accuracy when estimating peak flows for small (<100 km^2) and large (>10 000 km^2) catchments relative to the RoI models. In particular, the RIF model developed in the 42 homogeneous clusters indicates a more balanced level of accuracy across catchment areas in relation to the other models developed.

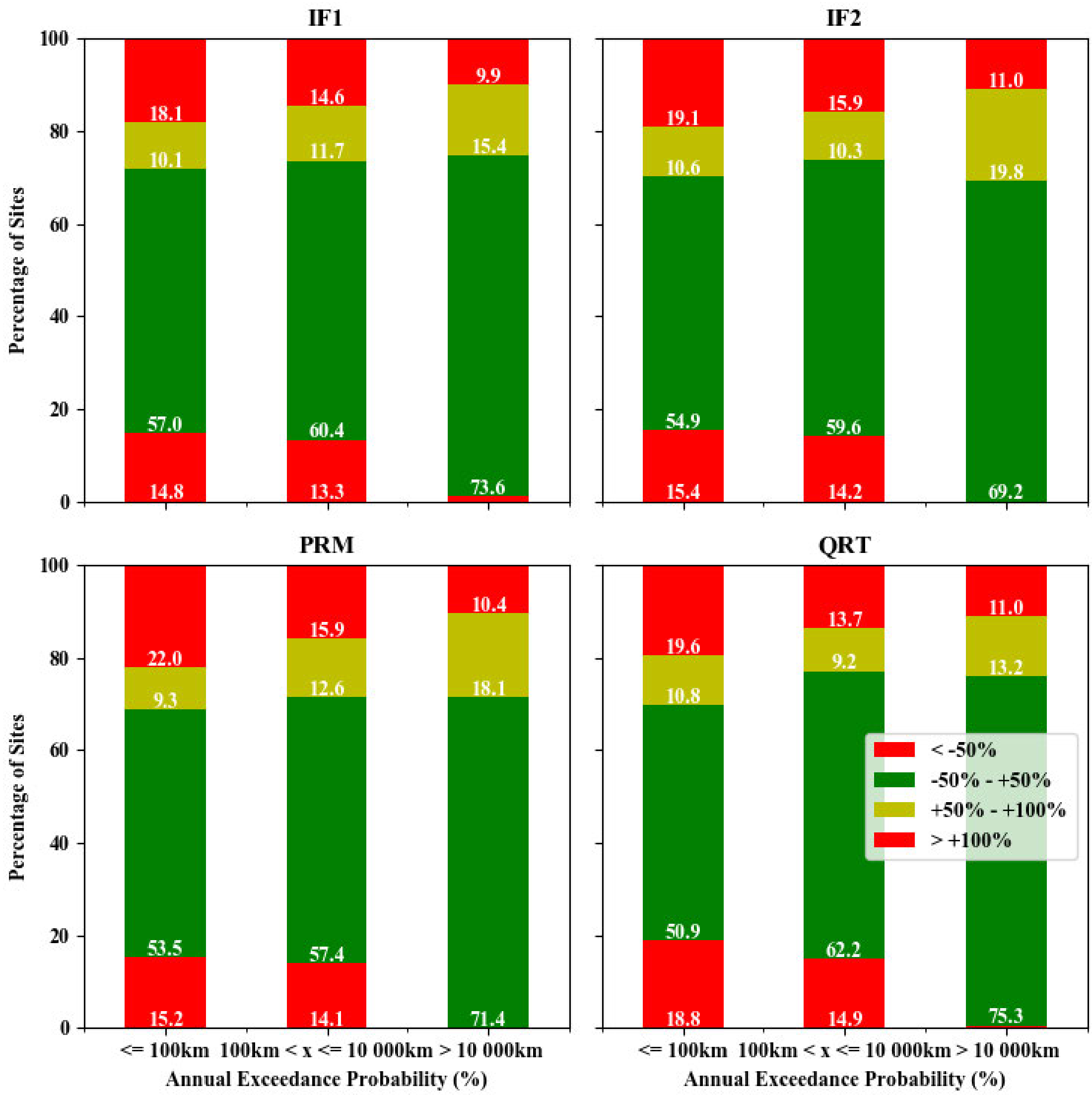


Figure 5.10 Percentage of Q_T estimates within the desirable *RE* defined by Rahman *et al.* (2012) and Naidoo (2020) using the IF1, IF2, PRM and QRT approaches for regional scale regressions and the clustering regionalisation approach at different catchment area ranges

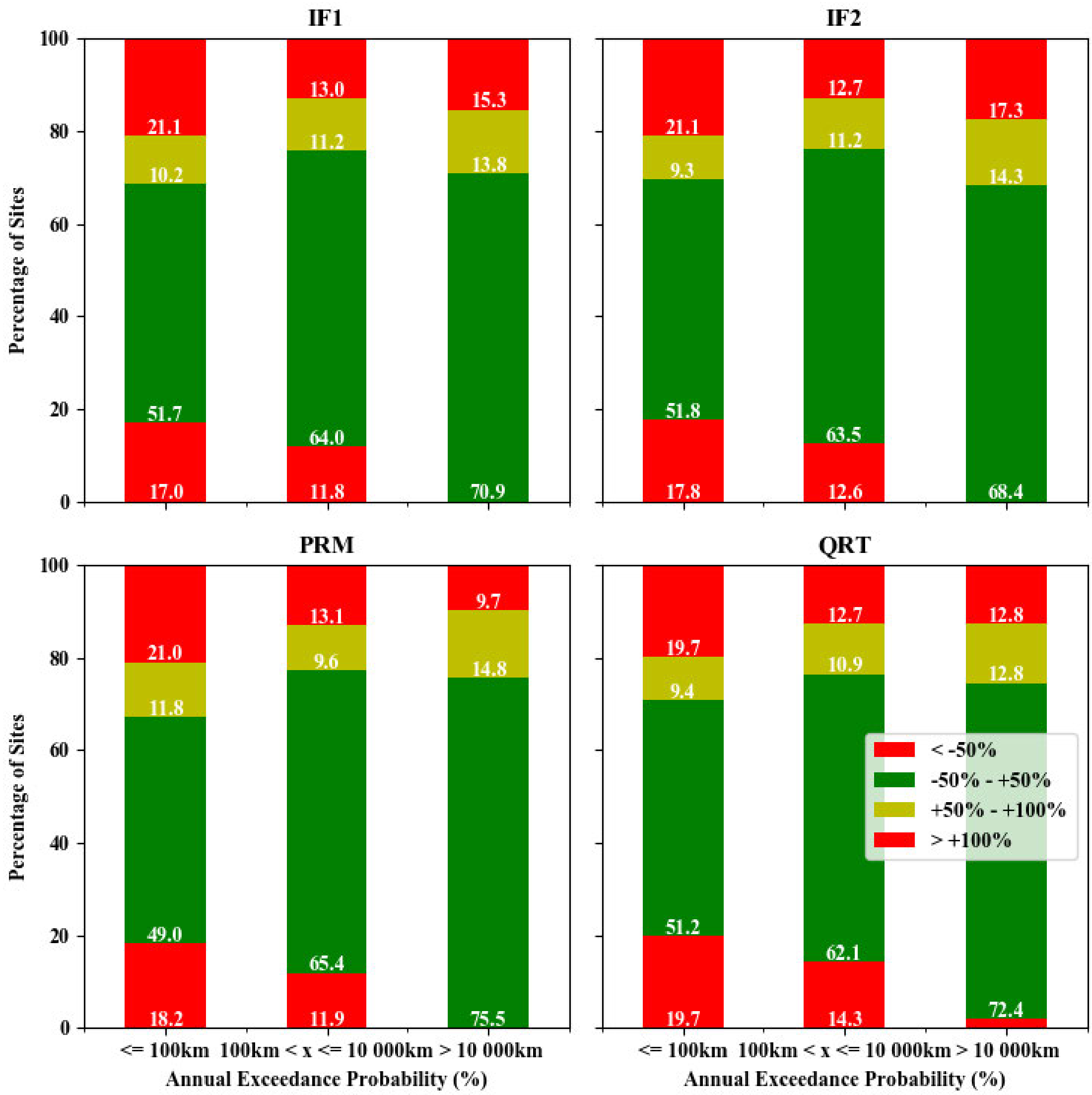


Figure 5.11 Percentage of Q_T estimates within the desirable *RE* defined by Rahman *et al.* (2012) and Naidoo (2020) using the IF1, IF2, PRM and QRT approaches for regional scale regressions and the region of influence regionalisation approach at different catchment area ranges

To undertake a holistic comparison Table 5.7 provides a ranking of the percentage of sites that are within acceptable or potentially acceptable *RE* ranges, for each modelling approach and AEP using the regional regression development.

Table 5.7 Ranking of models based on the percentage of sites within acceptable or potentially acceptable *RE* ranges for each AEP and modelling approach at a regional scale.

Ratio Category	**Region Type**	**Model**	**Model Rank per AEP (%)**							**Overall Rank**
			50	**20**	**10**	**5**	**2**	**1**	**0.5**	
Rahman *et al.* (2012)	Cluster	IF1	**2**	**3**	7	5	**1**	**3**	**2**	**3**
		IF2	4	**3**	6	6	5	6	6	5
		PRM	4	8	8	8	7	5	4	8
		QRT	**1**	**1**	4	7	8	8	8	6
	RoI	IF1	6	**2**	5	4	4	**1**	**1**	**3**
		IF2	**3**	6	**1**	**1**	**2**	4	5	**1**
		PRM	7	5	**2**	**1**	**2**	**2**	**3**	**1**
		QRT	8	6	**2**	**3**	5	7	7	7
Naidoo (2020)	Cluster	IF1	**1**	**3**	6	6	4	4	**1**	**3**
		IF2	4	7	7	7	6	6	5	7
		PRM	7	8	8	8	8	8	4	8
		QRT	**3**	**1**	5	**1**	7	5	6	5
	RoI	IF1	**2**	**3**	**3**	4	**2**	**2**	**3**	**2**
		IF2	5	**2**	**2**	**2**	**3**	**3**	8	**3**
		PRM	6	**3**	**3**	**2**	**1**	**1**	**2**	**1**
		QRT	8	6	**1**	5	5	7	7	6

When combining the ranks achieved for the NSE and the average RE rank as shown in Table 5.8, the IF1 and QRT approaches perform best for the clustering regionalisation, however, when considering the RoI approach, the IF1 and IF2 approaches perform the best. The IF1 and QRT developed in the homogeneous clusters and the IF1 and IF2 models developed using RoI were therefore assessed further to identify the best performing model.

Table 5.8 Combined model accuracy ranking

Region Type	**Model**	***NSE* Rank**	***RE* Average Rank**	**Combined Rank**
Cluster	IF1	**2**	**3**	**5**
	IF2	**3**	6	9
	PRM	7	8	15
	QRT	**1**	5.5	**6.5**
RoI	IF1	4	2.5	**6.5**
	IF2	5	**2**	**7**
	PRM	8	**1**	9
	QRT	6	6.5	12.5

5.5.2 *RMSE* and *BIAS*

The full set of *RMSE, RMSEr, BIAS* and *BIASr* results are provided in Appendices I to K. When considering the summarised *RMSEr* and *BIASr* values presented in Table 5.9 and Table 5.10, respectively, it is evident that the IF1 model developed at a clustering scale is the dominant model. The IF1 model developed using RoI produces the largest maximum *RMSEr* values (3.56). The IF2 and IF1 using RoI scale presents similar estimation performance with the *RMSEr*, with the IF1 only performing inconsistently for AEPs less than 5%. The IF1 *RMSEr* values reach maximums of 3.18 and 3.56 for the 1 and 0.5% AEPs, but the median values do not increase significantly, caused by the poor performance of Super Region 5, on the South Western coast of South Africa, for all the RoI models developed. The IF1 using clustering *RMSEr* values range between 0.19 and 2.17 and the median values range between 0.54 and 0.59, improving on the results of the IF1 and IF2 using RoI models by up to a factor of two. The QRT model developed at a cluster scale performs similarly well achieving median RMSEr values ranging between 0.52 and 0.61 and performing best for AEPs ranging between 1 and 20%.

Table 5.9 *RMSE*r statistics for the best performing model formulations

Model Form.	**Statistic**	**AEP (%)**							**Avg. Rank**
		50	**20**	**10**	**5**	**2**	**1**	**0.5**	
IF1 (RoI)	Minimum	0.61	0.54	0.52	0.53	0.56	0.60	0.60	4
	Median	0.74	0.69	0.70	0.74	0.81	0.87	0.94	
	Mean	1.01	1.03	1.07	1.12	1.21	1.29	1.39	
	Maximum	1.92	2.03	2.19	2.43	2.82	3.18	3.56	
IF2 (RoI)	Minimum	0.62	0.52	0.51	0.52	0.54	0.57	0.60	3
	Median	0.83	0.74	0.74	0.78	0.86	0.95	1.06	
	Mean	1.07	1.10	1.11	1.13	1.18	1.22	1.28	
	Maximum	2.06	2.37	2.48	2.53	2.51	2.44	2.33	
QRT (Clustering)	Minimum	0.19	0.20	0.19	0.19	0.20	0.23	0.26	1
	Median	0.58	0.56	0.56	0.52	0.53	0.57	0.61	
	Mean	0.72	0.63	0.62	0.63	0.66	0.71	0.78	
	Maximum	2.69	1.80	1.56	1.43	1.60	1.81	2.04	
IF1c (Clustering)	Minimum	0.25	0.23	0.19	0.20	0.20	0.20	0.19	1
	Median	0.55	0.56	0.54	0.54	0.54	0.56	0.59	
	Mean	0.70	0.69	0.69	0.69	0.70	0.72	0.74	
	Maximum	1.76	2.08	2.13	2.11	2.03	2.09	2.17	

The *BIASr* values for the RoI models are consistent across all of the models with values ranging between 0.37 and 1.06. The median values range between 0.64 and 0.84 and do not significantly deviate from the mean values. The IF1 using clustering models shows similar consistencies, but lower median *BIASr* values are observed, the maximum *BIASr* values are however significantly higher with values between 1.07 and 1.76. The QRT using clustering model again performs similarly to the IF1 model, except for the 2% AEP, where a maximum BIASr value of 3.17 is observed. The difference in the model performance per regionalisation scheme observed is believed to be caused by the regionalisation scheme differences, i.e. fixed region versus hydrological neighbourhood. The IF1 using RoI *BIASr* performs poorly in Clusters 34 and 37 on the east coast of South Africa.

Table 5.10 *BIASr* statistics for the best performing model formulations

Model Form.	Statistic	AEP (%) 50	20	10	5	2	1	0.5	Avg. Rank
IF1 (RoI)	Minimum	0.44	0.38	0.38	0.41	0.48	0.57	0.68	3
	Median	0.66	0.64	0.64	0.66	0.71	0.74	0.80	
	Mean	0.70	0.64	0.63	0.64	0.67	0.73	0.80	
	Maximum	1.03	0.87	0.80	0.75	0.78	0.83	0.90	
IF2 (RoI)	Minimum	0.41	0.37	0.38	0.42	0.48	0.55	0.63	4
	Median	0.73	0.68	0.66	0.66	0.69	0.75	0.84	
	Mean	0.73	0.66	0.64	0.65	0.68	0.74	0.81	
	Maximum	1.06	0.89	0.82	0.78	0.79	0.83	0.90	
QRT (Clustering)	Minimum	0.19	0.19	0.22	0.25	0.24	0.22	0.23	2
	Median	0.49	0.50	0.48	0.53	0.51	0.54	0.59	
	Mean	0.68	0.59	0.58	0.59	0.64	0.69	0.78	
	Maximum	3.17	1.91	1.48	1.32	1.37	1.41	1.54	
IF1 (Clustering)	Minimum	0.23	0.20	0.20	0.21	0.19	0.19	0.23	1
	Median	0.50	0.51	0.50	0.57	0.65	0.69	0.68	
	Mean	0.59	0.57	0.58	0.59	0.62	0.66	0.71	
	Maximum	1.46	1.23	1.11	1.07	1.29	1.50	1.76	

When combining the rank of the models, shown in Table 5.11, it is evident that the clustering regionalisation scheme produced the best performing models. This may be attributed to the size of the pooling groups. In contrast to findings internationally the IF1, which utilises equal weighting of the pooling group as opposed to applying a record length or combined record length and Euclidian distance weighting.

Table 5.11 Combined model *RMSEr* and *BIASr* ranking

Region Type	Model	*RMSEr* Rank	*BIASr* Rank	Combined Rank
Cluster	IF1	**1**	**1**	**2**
	QRT	**1**	2	3
RoI	IF1	4	3	7
	IF2	3	4	7

5.6 Discussion and Conclusions

Regional flood model development tends to fall within one of two categories, QRT or PRT. QRT models directly estimate the quantile flows in question, e.g. the 1% AEP flood event, whereas PRT relies on regional growth curves. PRT has the advantage of allowing the estimation of the entire growth curve as opposed to QRT, which develops individual models for each AEP being considered. PRT growth curves are generally scaled using a SF, most commonly the *MAF* or *MEF*, to ensure that the regional values are scale independent allowing for estimation at ungauged sites within a homogenous region.

Numerous model formulations exist for the development of regional flood models, but the formulation most prominent in the literature is the IF. To assess the suitability of QRT and PRT models in South Africa, a QRT model as well as three PRT model formulations: (i) IF with equal weighting (IF1), (ii) IF with varied weighting (IF2) and (iii) PRM were developed. The SF adopted for the IF approach was the MAF and the 10% AEP C-value (C_{10}) was used for the PRM.

Regression models were developed to estimate the required SFs and Q_T and the development was undertaken at two scales, national and regional, based on two regionalisation schemes. The regionalisation schemes adopted were Clustering, which consisted of 42 relatively homogeneous clusters, and RoI, which included the use of five Super Regions to refine the regression model development. As such a total of four combinations of development scale and regionalisation scheme was used: (i) Clustering with cluster-based models, (ii) Clustering with a national scale model, (iii) RoI with super region based models, and (iv) RoI with a national scale model.

A key assumption for the development of the SF and Q_T regression models is the use of a minimum number of 30 sites. This restriction was imposed due to some of the 42 clusters containing as few as three sites and to improve the robustness of the models developed. Although the use of a minimum number of 30 sites has proven to develop models that perform adequately it is considered an aspect of the study where further research can be undertaken to improve estimates.

Eight catchment descriptors were included as potential predictor variables for regression development that ranged from outlet elevation through to design rainfall values. Due to the number of models developed, the p-value was used for the selection of significant catchment descriptors and the developed models were limited to the use of three predictor variables and is an aspect of the study where further refinement could be considered. *A*, *MAP* and D_c were identified as significant catchment descriptors in the SF regressions, whereas the Q_T regression favoured the use of the 10% AEP design rainfall above the *MAP*. The models developed for *MAF* at a regional scale performed well with an *NSE* value of 0.78 achieved. The *NSE* values for the national *MAF* both C_{10} models performed poorly achieving 0.05 and 0.00 respectively.

The regression models developed for the RoI approach varied significantly between model formulation and Super Region, with *A* and *MAP* being dominant descriptors. Similar to the cluster based models, the RoI *MAF* regional models performed relatively well achieving an *NSE* of 0.69. The regional C_{10} value performs better than the clustering, albeit still poorly, with an *NSE* of only 0.28. The national models again only achieve values of up to 0.05. Due to the poor performance of the national scale models, only the regional models were assessed further.

The *SF* and Q_T regression models were used to estimate the anticipated design peak flows at the sites considered and assessed using six key performance indicators: (i) *NSE*, (ii) *RE*, (iii) *RMSE*, (iv) *RMSEr*, (v) *BIAS*, and (iii) *BIASr*. The assessment approach consisted of three steps, initially the best performing regression development scale was identified using the *NSE*, followed by the identification of the four best performing models with regards to *RE* and, finally the *RMSE* and *BIAS* was used to identify the best performing model.

The accuracy of the regional models was measured through the use of *NSE* and *RE*. The cluster based IF1 and QRT models and the RoI based IF1 and IF2 models were the best performing models for the regional scale models when considering the NSE, significantly improving on the results from the national models. The regional models achieved *NSE* values up to 0.77, but tended to underestimate, which needs to be taken into consideration if the models are applied.

The *RE* assessment relied on the ratio bounds developed by Rahman *et al.* (2012) and Naidoo (2020). Rahman *et al.* (2012) notes that the acceptable ratio bounds, termed "desirable", are arbitrary, but provide a guide for model assessment and adopted a ratio bound of 0.5 – 2 as acceptable. Naidoo (2020) further refined the bounds of acceptable ratios through the inclusion of potentially acceptable bounds and reducing the strictly acceptable bounds. For comparative purposes the potentially acceptable and acceptable bounds defined by Naidoo (2020) were used. The two ratio bounds considered acceptable were therefore 0.5 – 2 and 0.5 – 1.5. The models that performed best in the *RE* assessment were the IF1 for both regionalisation schemes and the IF2 and PRM models using the RoI. The four models were able to estimate the peak flows within the Rahman *et al.* (2012) *RE* bounds at between 64.8 and 75.2% of the sites considered. The percentages drop to between 53.1 and 65.3% when applying the more stringent *RE* bounds defined by Naidoo (2020).

The four best performing models from the model accuracy assessment were the IF1 and QRT using clustering and the IF1 and IF2 using RoI. The final assessment compared the *BIAS* and *RMSE* of the above models. The IF1 and QRT using clustering are the best performing models when considering both the *RMSEr* and the *BIASr*, the models achieved median *RMSEr* values ranging between 0.52 and 0.61, improving on the results of the remaining models by up to a factor of two. The RIF using RoI *BIASr* values also improve on the results of the remaining models, in particular for AEPs in excess of 10%. The IF1using Clustering *BIASr*, however, performs poorly in clusters 34 and 37 on the east coast of South Africa. When ranking the models based on the *RMSEr* and *BIASr* the IF1 and QRT using clustering were the top two performing models. This result is in contrast to international recommendations, where it has generally been found that using a record length weighting improves the results of developed models. In addition alternative weighting schemes have also been shown to benefit the quantile estimates, as such additional investigation into an appropriate weighting scheme may also prove beneficial.

The IF1 and QRT using Clustering models are therefore the best performing models on a national scale. The IF1 however has the added advantage of being able to estimate the entire growth curve whereas QRT models only estimated for pre-defied AEPs. The IF1 using

Clustering is therefore the recommended model at a national scale, however cognisance needs to be taken when applying the model on the eastern coast due to the poor *BIASr* performance. Practitioners can also use the provided performance metrics to compare the models at a regional scale and select the best performing model.

6 DISCUSSION

The main purpose of this study was to develop and assess the performance of regional DFE models for application in South Africa, utilising currently available streamflow data. Three main research questions were identified that needed to be answered:

(a) What is the most suitable distribution for FFA in South Africa on an at-site scale for use on a national scale?
(b) Can South Africa's catchments be regionalised into statistically homogeneous flood producing regions?
(c) Given the data sparsity in South Africa, which regional DFE model is most suitable?

Specific objectives that were achieved through the study were:

(a) Compilation of a hydrological descriptors database.
(b) Collation and quality control of selected gauged flow data in South Africa.
(c) Identification of a suitable distribution for use in South Africa.
(d) Identify and verify homogeneous flood producing regions.
(e) Regional flood model development and assessment of the performance.

6.1 Data Collation and Screening

The collection of the required data for the study was considered to be a two-phased approach, whereby both the catchment specific descriptors, and the hydrological streamflow data were collated.

6.1.1 Catchment specific descriptors

The methods used in this study are based on internationally accepted DFE procedures. This was used to ensure familiarity in the approach and ease of application. These methods require nearly identical catchment parameters to be estimated for use in the four DFE methods developed and range from meteorological parameters to topographic and land use parameters.

DFE methods require a range of catchment descriptors to be determined for use in models. Considering the requirement of ease of application by practitioners, 17 catchment descriptors that are readily available, or easy to estimate, were selected for inclusion in the study. The descriptors ranged from geographic and morphological to design rainfall descriptors. A hydrologically corrected DEM was developed using the SRTM (NASA-JPL, 2013) data, producing a 30 x 30m DEM for the extraction of various catchment descriptors. The derived catchment parameters were verified against DWS data sets.

6.1.2 Development of a quality-controlled streamflow data set

The DWS is the custodian for streamflow data for all the sites across South Africa. In addition, the data set compiled by Van Bladeren *et al.* (2007) was used to supplement the available data from DWS. The entire data set received consisted of 474 flow-gauging stations with a total record length exceeding 15 000 years. The data were screened to identify potential errors and to summarise the primary data into annual, monthly, weekly, and daily peak values. This screening process allowed for the inclusion of additional years that would conventionally have been excluded due to the extent of missing information and the time at which these occurred.

The screening process considered multiple steps. Firstly, the selection of a minimum record length of 20 years, secondly, identifying stations impacted by upstream developments and, lastly, data quality assessment. The quality assessment included the identification of missing data, verification of the regional occurrence of floods, error identification and quality control. Some of the stations considered also required rating curve extensions to be performed, and this was limited to a maximum of 20% increase in the currently maximum rated stage and flow. After the screening, a total number of 383 sites remained for further processing and development of the models.

6.2 Selection of a suitable Distribution for Flood Frequency Analysis

At-site design peak discharges were required to form the basis of the model development and performance assessment processes. It was identified that several parameter estimation methods have been suggested to fit probability distributions to the data in South Africa, ranging from standard MM estimation techniques to LM. The wide use of LM both locally and internationally resulted in the adoption of LM to fit probability distribution to the data used in this study.

An integral part of the FFA is the selection of an appropriate distribution. The distributions that have been recommended in the literature for South Africa are the GEV, LP3, GPA and PE3 distributions. More recently Kjeldsen *et al.* (2017) reviewed the KAP3 distribution for application in RFFA, which provides improved fits to regional data. Given the recommendations listed above, and the higher level of flexibility of the KAP3 five distributions were selected for evaluation: (i) GEV, (ii) GPA, (iii) KAP3, (iv) LP3 and (v) PE3. The national KAP3 *h*-value for South Africa was estimated to be 0.77. The evaluation process relied on an iterative elimination approach, reviewing graphical fit to theoretical distributions, GoF, model fit criterion and model predictive ability for identification of the most suitable distribution.

The graphical fit test, which utilises LM diagrams and the Mahalanobis distance (Kjeldsen and Prosdocimi, 2015) to identify the most suitable distributions, identified the KAP3 as the most suitable, followed by the GPA and LP3 distributions. The GoF tests included the modified Anderson Darling, Chi-Squared, Cramer von Mises and Kolmogorov Smirnov tests and were applied in four record length categories. For the record lengths less than 80 years no dominant distribution was identified as the maximum selection rate difference between distributions was only 4.4%. For sites with record lengths exceeding 80 years, the LP3 distribution on average accepted at 13.9% more sites, however the number of stations within this category is only 27, therefore the higher level of acceptance only equates to 4 sites. The LP3 was deemed most favourable, followed by the GPA and KAP3 when strictly applying the GoF test acceptance percentages to determine rankings

When applying the model fit criterion, AIC, AICc and BIC, record length categories were also applied. The model fit criterion identified the KAP3 as the worst performing distribution, never being selected as the most suitable distribution. However, the GPA was selected as the most suitable distribution for record lengths less than 80 years, with the most dominant category being for sites having between 60 and 80 years of records, where the GPA was selected at 27.6% mor sites. When considering sites with record lengths in excess of 80 years, the LP3 was selected at only one site more than the GPA. The GPA was therefore deemed as the best performing distribution followed by the LP3 and KAP3 respectively.

The final consideration adopted for the selection of a suitable distribution was the predictive ability of the distributions and consisted of a comparison of the estimated and Gringorton plotted AMS, dubbed the “true” fit, and the estimation of the uncertainty associated with the distributions being considered. The 5% AEP was adopted to test the "true" fit as the indicative AEP and this allowed for the use of 148 sites (39%) when only considering sites with 50 years or more of records, which exceeds the 2T requirement prescribed by Robson and Reed (1999). It was identified that the GPA and KAP3 distributions tend to underestimate the 5% AEP event with an interquartile range between 0.87 and 1.04 and median values of 0.97, whereas LP3 distribution tends to overestimate with an interquartile range between 0.92 and 1.11 and a median of 1.01. These results however did not clearly identify any of the distributions as the most suitable distribution.

The uncertainty associated with the distributions considered, was based on the 90% confidence limits derived using balanced bootstrapping resampling. The variance of the confidence bands was calculated as a percent variance of the FFA of the bootstrap replicates in relation to the FFA results using the original AMS data set. The assessment of the uncertainty associated with the distributions was based on the 1 and 5%s AEPs, when restricting the sites for the assessment to sites with record lengths of 50 years or greater. This restriction adhered to the 2T rule and when applying the "rule-of-thumb" of extrapolating to two times the record length an indication of the associated uncertainty was assessed for the 1% AEP. It was identified that the uncertainty associated with the GPA and KAP3 was very similar, and was anticipated due to the GPA being a special case of the Kappa distribution. The KAP3 marginally outperformed the GPA for the 1 and 0.5% AEPs. Conversely, the LP3 distribution has a much higher associated level of uncertainty for AEPs less than 10%, with uncertainty bands reaching up to 240% as opposed to 95% for the 0.5% AEP. Considering both the predictive ability analyses, although the LP3 may perform marginally better than the GPA and KAP3 when estimating the "true" fit, the associated uncertainty of the distribution brings into question whether this perfomance will be consistent for extended data sets. The KAP3 has the lowest level of uncertainty, and coupling this with the estimation of the "true" fit has been deemed the most suitable distribution when considering the predictive ability, followed by the GPA and the LP3 respectively.

Taking into consideration the tests undertaken, the LP3, KAP3 and GPA distributions were ranked to identify the most suitable distribution. The final ranking indicated that the GPA was most suitable followed by the KAP3 and the LP3 respectively.

6.3 Regionalisation

Clustering and RoI regionalisation approaches were applied, with both methods requiring multiple adjustments and further refinement of the regional delineation. The regionalisation, modification and recommendations are discussed below.

6.3.1 Clustering

K-means clustering aims to estimate the minimum overall Euclidian distance for all clusters being considered. The identified clusters were also required to adhere to homogeneity requirements as stipulated by Hosking and Wallis (1997) as well as the 5T rule. The homogeneity measures adopted in the study were the *H* statistic and the discordancy measure (*D*).

Using the national streamflow dataset, the homogeneity testing identified that the dataset contained discordant sites, which would need to be moved, replaced, or excluded to improve the homogeneity. After removing the discordant sites for the entire data set, homogeneity was still not achieved. Following the same approach further discordant sites were excluded when the primary drainage regions were considered independently of each other. Removal of the discordant sites did not improve the homogeneity of the primary drainage regions and only 5 of the regions were deemed relatively homogeneous, two of which did not adhere to the 2T requirement. Therefore, a re-clustering approach was adopted to identify relatively homogeneous clusters within the entire data set.

The clustering was performed in the attribute space and the catchment descriptors were normalised to a range of 0 to 1. This reduced the bias of large values such as the *MAP* that may unduly influence the clustering. The longitude and latitude were double weighted in relation to other parameters due to the literature identifying that spatial proximity is a key indicator of homogeneity. An iterative cluster analysis process was followed, whereby each potential descriptor combination was included for a cycle of cluster analyses. The homogeneity measures were also calculated for each iteration and each iteration ranked based on the level of homogeneity achieved. Clustering was performed using site descriptors, whereas the homogeneity of the clusters, was assessed using the site specific quantitative FFA parameters.

The descriptor sets that were able to identify the highest percentage of homogeneous clusters were largely meteorological and geographic parameters, with the combination of latitude longitude and D_c being deemed as the most suitable parameters for clustering. It should be noted that the additional rainfall descriptive statistics can be used, such as the growth curve slope, rainfall seasonality, and rainfall clusters, and further investigation is recommended into the validity of their use.

The preferred number of clusters was determined using the 2/5*T* approach adopted by Robson and Reed (1999) as a minimum criterion for the sizing of clusters. This specifies the absolute minimum required record length for RFFA as two times the design event being estimated, with five times being preferable. As an initial estimate a maximum of 36 clusters was adopted, which provides an average of approximately 500 years per cluster. The initial clustering identified 17 homogeneous clusters. The remaining heterogeneous clusters were further analysed using the same clustering approach to ensure continuity and prevent any potential subjectivity. The initial selection of the number of clusters was such that the clusters varied in size from two to 29 sites. Hosking and Wallis (1997) noted that there is no set standard for the selection of the cluster sizes and that the size will affect the model's capabilities to identify regional differences or bias.

A total of 42 relatively homogeneous clusters were identified through adjusting the initial 36 clusters. The process, however, required the exclusion of 51 sites due to discordancy and inconsistencies with geographic variance. To the knowledge of the author this is the first complete multi-variate clustering analysis that has been undertaken on a national scale for DFE in South Africa. Previous studies relied on historically defined homogeneous regions and amalgamated morphological and historical homogeneous region definitions for the definition of the homogeneous regions.

6.3.2 Region of Influence

A RoI approach which enforces a minimum required record length was investigated, which allowed for an assessment of whether the enforcement of the 2/5T rule generates homogeneous flood regions. The descriptor set that identified the highest percentage of relatively homogeneous regions consisted of Latitude, Longitude, D_c, and mean runoff percentage, which identified 16% and 51% homogeneous regions for 500- and 200-year minimum record lengths approaches, respectively.

To improve the homogeneity achieved through RoI, the use of Super Regions was investigated as developed by Mostofi Zadeh and Burn (2019). Through the use of elbow plots, TSNE, UMAP and geographic plots, five super regions were identified based on the Latitude, Longitude, *A*, *MAP*, D_c, Catchment Slope (10-85) and 24-hour 10% AEP design rainfall. Although the UMAP plots indicated the potential for an additional super region, the additional super region affected the geographic plots negatively. Within each of the super regions, the parameter set identifying the highest percentage of relatively homogeneous clusters was identified and utilised for the model development steps. The parameter sets, however, only identified 52.6% relatively homogeneous sites on a national scale, yielding slight improvement over the conventional RoI approach. One key finding was that the RoI performed particularly poor in super region 5, located on the South Western coastline of South Africa, only identifying homogeneous regions for 28% of sites.

6.4 Model Development

Four distinct modelling approaches were adopted for the estimation of Q_T, i.e. RIF, PGC, CRM and the DQE approaches, and each approach was developed on a national and a cluster-based scale. The national and regional scale models refer to the sites used for the formation of the regressions to estimate the *SF*, i.e. for the national models all sites were utilised, whereas regional models relied only on the sites within the defined clusters or super region. As such a total of four combinations of development scale and regionalisation scheme was used: (i) Clustering with cluster-based models, (ii) Clustering with a national scale model, (iii) RoI with super region based models, and (iv) RoI with a national scale model. Prior to application of the modelling approaches the at-site RM calibration and selection of suitable SFs was undertaken.

6.4.1 RM calibration

The calibrated C_T-values were derived from the design rainfall determined by Smithers and Schulze (2003) and design peak discharges determined from the observed flow data using the GPA distribution. Some inconsistencies in the C_T values were identified resulted as some *C* values exceeded one, with a maximum and minimum of 1.569 and 0.002, respectively. In theory the C_T value will increase as the AEP decreases, however, 20 sites, had negative C_T value growth curves after the regional calibration process was undertaken. These inconsistencies have been identified in previous studies both locally and internationally (McDermott and Pilgrim, 1982, Parak, 2007) Although, the stations potentially introduce errors into the derived models, it is anticipated that due to the use of regional analysis that the effect will be limited. However, further investigation is necessary to resolve these inconsistencies.

6.4.2 Scaling factor and dimensionless growth curves

Both the *MAF* and *MEF* were assessed for viability for use as a SF for the developed growth curves. The *MAF* reduced the spread of the dimensionless growth curves in relation to the *MEF*, particularly for AEPs less than 5% and was therefore adopted. The C_{10} value, as adopted by Pilgrim (2001) and Calitz and Smithers (2020), was used as the SF for the at-site *C* value growth curves.

Two sets of regional growth curves were developed, the *C* value and Quantile, additionally regional LMs were derived, the latter allows for the derivation of growth curves from additional distributions for comparative purposes.

6.4.3 Regression development

For the development of the regressions to estimate the required SFs the predictor variable p-values were used to identify the three most suitable descriptor variables. Three variables were chosen as a means to reduce model complexity and to aid in the adoption of the model by practitioners.

The regressions developed for the clustering regionalisation scheme and estimation of the SFs utilised a combination of *A*, *MAP*, D_c and the 10% AEP design rainfall depth. Given the small size of some of the defined clusters, the regressions were found to be extremely sensitive when applying LOO resampling performance assessment and a minimum requirement of 30 sites was adopted for the development of the cluster scale regressions. The developed models to estimate the *MAF* performed well with R^2 values ranging between 0.59 and 0.9, however the C_{10} models only achieved R^2values between 0.05 and 0.45. Additionally, when comparing the estimated vs. the observed values, the regional *MAF* model achieved an *NSE* of 0.78, as opposed to values between 0 and 0.05 for the remaining models. The QRT modelling approach regressions also achieved well when considering the R^2 values achieving values ranging between 0.49 and 0.91 with a median value of 0.79.

The RoI and Super Region regressions were unable to improve on those developed for the clustering scheme. Regressions were developed, including the estimation of the most suitable parameters, at a Super Region scale in an attempt to improve estimates. The developed regressions achieved R^2 values ranging between 0.78 and 0.83 for estimation of the *MAF*, but the *NSE* values indicated that the national based models were again a poor fit, however the regional *MAF* performed adequately achieving an *NSE* of 0.69. Additionally, the C_{10} regressions were improved, however still performed poorly, only achieving maximum *NSE* and R^2 values of 0.28 and 0.45, respectively. The developed QRT models performed well with R^2 values ranging between 0.66 and 0.86 and a median value of 0.78.

6.4.4 Model performance assessment and application recommendations

The regional models were assessed using the *BIAS*, *RMSE*, *BIASr*, *RMSEr*, *RE* and the *NSE* values. The *NSE*s were used to identify the best performing modelling scale, after which the four best performing models were selected based in the *RE*, the final assessment was undertaken using the *BIAS*, *RMSE*, *BIASr*, and *RMSEr*.

The model accuracy was assessed using the *NSE* and *RE*. The regional models achieved *NSE* values up to 0.77 but tended to underestimation. The *RE* assessment relied on the ratio bounds developed by Rahman *et al.* (2012) and Naidoo (2020). Rahman *et al.* (2012) notes that the acceptable ratio bounds, termed "desirable", are arbitrary, but provide a guide for model assessment and adopted a ratio bound of 0.5 – 2 as acceptable. Naidoo (2020) further refined the bounds of acceptable ratios, taking into consideration the type of error, e.g. over or under estimation, and introduced potentially acceptable bounds and firm acceptance bounds. The refined bounds reduce the potentially acceptable upper and lower bounds to range between 0.5 and 1.5. For comparative purposes the potentially acceptable and acceptable bounds defined by Naidoo (2020) were considered as acceptable as no guidance is provided on how to identify whether an estimate within the potentially acceptable bounds is approved or rejected. The two ratio bounds considered were therefore 0.5 – 2 and 0.5 – 1.5. The models that performed best in the *RE* assessment were the IF1 for both regionalisation schemes and the IF2 and PRM models using the RoI. The four models were able to estimate the peak flows within the Rahman *et al.* (2012) *RE* bounds at between 64.8 and 75.2% of the sites considered. The percentages drop to between 53.1 and 65.3% when applying the more stringent *RE* bounds defined by Naidoo (2020). When considering an overall rank for the model accuracy assessment the four top performing models were the IF1 and QRT using clustering and the IF1 and IF2 using RoI.

The final assessment compared the *BIAS* and *RMSE* of the Clustering based (IF1, QRT) and the RoI based (IF1 and IF2) models. The IF1 and QRT using Clustering models are the best performing models when considering both the *RMSEr* and the *BIASr*, the models achieved median *RMSEr* values ranging between 0.52 and 0.61, improving on the results of the remaining models by up to a factor of two. The *BIASr* values for the IF1 and QRT also improve on the results of the remaining models, in particular for AEPs in excess of 10%. The *BIASr* values for the IF1, however, performs poorly in clusters 34 and 37 on the east coast of South Africa.

The IF1 and QRT using Clustering models are therefore the best performing models on a national scale. However, the IF1 has the added advantage of being able to estimate the entire growth curve and is not limited to predefined AEP quantiles with the QRT models. The IF1 is therefore the recommended model at a national scale, however cognisance needs to be taken when applying the model on the eastern coast due to the poor *BIASr* performance. Practitioners can also use the provided performance metrics to compare the models at a regional scale and select the best performing model.

7 CONCLUSIONS AND RECOMMENDATIONS

The main findings of the study can be summarised into the following:

(i) The South African streamflow database can be considered equally rich in data to other countries such as Australia and the United Kingdom.

(ii) Out of five potential frequency distributions utilised both locally and internationally, the most appropriate distribution for application in South Africa, based on four selection criteria, was the GPA distribution.

(iii) Two methods for the formation of pooling groups were assessed for application in South Africa, and through applying strict homogeneity requirements 42 homogeneous clusters were identified. The homogeneous clusters were also mapped out geographically to ease the application process.

(iv) The super region approach, which applied dimensionality reduction methods, for the identification of five super regions, combines RoI and Clustering showing potential for application in South Africa.

(v) When developing models for the estimation of the SFs at an at-site basis regional scale models outperformed the use of national scale models.

(vi) Of the modelling approaches adopted, the Index flood approach, applying equal station weighting, and the Quantile Regression Techniques were deemed best performing models, however, due to the IF1 estimating the entire growth curve it is the recommended approach.

7.1 Unique Contributions from This Study

The following items are considered to be the most prominent unique contributions that have been developed as part of this study:

(i) The database of design flood specific descriptors is potentially the largest database concentrating on South Africa and this study thus provides a basis for further development and refinement of models for DFE in South Africa.

(ii) To the knowledge of the author, it is the first study to perform a detailed investigation into the most suitable probability distribution to use for FFA in South Africa and to recommend the use of the GPA in South Africa.

(iii) The first application of model uncertainty used for the selection of a suitable design flood distribution.

(iv) The first application of the KAP3 methodology and determination of a national Kappa *h* value for South Africa.
(v) The RoI, super-region and multi-variate clustering approaches have also not been applied in South Africa before and previously geographic and morphological maps were used, or reliance was often placed on historically defined homogeneous regions.
(vi) A new and unique set of relatively homogenous clusters for use with DFE have been developed.
(vii) Development and performance assessment of four models using QRT and PRT modelling approaches at multiple scales for South Africa. This included the comparison of the models' predictive ability and identified that the equally weighted IF approach outperformed the record length and Euclidian distance weighted record length weighting, which is in contrast with international findings.

In addition to the above, the study also identified that in the South African context, which was shown to be relatively data rich, the Clustering regionalisation scheme provided the best overall quantile flow estimates, which is in contrast to international findings where in data rich regions, RoI is generally favoured. The result in this study does not exclude the RoI, but only in the current form applied, the inclusion of additional descriptors, weighting schemes and model formulations may improve the RoI performance. Similarly, the equally weighted IF was found to perform best, in contrast to international findings where generally record length weighting has been shown to improve results.

This research has contributed to the following key projects identified by the NFSP as outlined by Smithers *et al.* (2014):
(i) A.1.2.2 Guide for AEP distribution for floods
(ii) A.1.2.3 Spatial and Temporal distribution of available streamflow data
(iii) A.1.2.6 Refined regionalised / polled Index flood methods
(iv) A.1.2.7.1/3 Update and refine the RMF method and its regionalisation
(v) A.1.2.8.2/3 Modernise the Standard Design Flood Method
(vi) A.1.2.8.5 and A.1.2.4 Modernise existing synthetic unit hydrographs and related homogeneous flood regions
(vii) A.1.2.8.6 Modernise existing empirical methods for small catchments
(viii) C.2 Web-based framework of methods on SANCOLD website
(ix) C.6 Web-based GIS database/geodatabase

7.2 Recommendation for Future Research

The following recommendations for future research and development are made:

(a) Investigation into the impacts of non-stationarity on regionalisation,

(b) Potential relaxation of data requirements in data sparse regions to improve regionalisation coverage in these regions,

(c) Further improvement of the RoI approach within South Africa, such as an improved weighting scheme or additional parameters for region formation,

(d) Investigation into the applicability of Bayesian statistics in South African hydrology,

(e) Compilation of a national hydrological descriptor database, beyond flood estimation, similar to the databases implemented internationally,

(f) Refinement of the estimation of *ARF*s,

(g) Further refinement of the developed models to improve estimates, e.g. through application of Bayesian models, inclusion of additional catchment descriptors, weighting scheme modifications, etc.,

(h) Update of the National Design Rainfall Database, and

(i) Improvement of time of concentration estimation at a national scale.

8 REFERENCES

Acreman, MC. 1987. *Regional Flood Frequency Analysis in the Uk: Recent Research - New Ideas*. Report No. None. Institute of Hydrology, Wallingford, UK.

Acreman, MC and Wiltshire, SE. 1987. Identification of Regions for Regional Flood Frequency Analysis. *EOS* 68(44):1262

Ahn, K-H and Palmer, R. 2016. Regional Flood Frequency Analysis Using Spatial Proximity and Basin Characteristics: Quantile Regression Vs. Parameter Regression Technique. *Journal of Hydrology* 540(Supplement C):515-526

Akaike, H. 1974. A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control* 19(6):716-723

Aldrich, J. 1997. R.A. Fisher and the Making of Maximum Likelihood 1912-1922. *Statist. Sci.* 12(3):162-176

Alexander, WJR. 1990. *Flood Hydrology for Southern Africa*. SANCOLD, Pretoria, RSA.

Alexander, WJR. 2000. *Flood Risk Reduction Measures*. University of Pretoria, Pretoria, South Africa.

Alexander, WJR. 2002a. The Standard Design Flood. *Journal of the South African Institution of Civil Engneering* 44(1):26-30

Alexander, WJR. 2002b. Statistical Analysis of Extreme Floods. University of Pretoria, Pretoria, South Africa.

Alsuwaidi, O, Haque, MM, Rahman, A and Haddad, K. 2015. A New Probabilistic Rational Method for Design Flood Estimation in Ungauged Catchments for the State of New South Wales in Australia. In: eds. *21st International Congress on Modelling and Simulation*, 2158-2164. Modelling and Simulation Society of Australia and New Zealand Inc., Gold Coast, Australia.

Anderson, TW and Darling, DA. 1952. Asymptotic Theory of Certain "Goodness of Fit" Criteria Based on Stochastic Processes. *Ann. Math. Statist.* 23(2):193-212

Aronica, GT and Candela, A. 2007. Derivation of Flood Frequency Curves in Poorly Gauged Mediterranean Catchments Using a Simple Stochastic Hydrological Rainfall-Runoff Model. *Journal of Hydrology* 347(1):132-142

Basson, MS and Pegram, GGS. 1994. *Probabalistic Management of Water Resource and Hydropower Systems*. Water Resources Publications, Highlands Ranch, Colo., USA.

Bastianin, A. 2020. Robust Measures of Skewness and Kurtosis for Macroeconomic and Financial Time Series. *Applied Economics* 52(7):637-670

Beard, LR. 1974. *Flood Flow Frequency Techniques*. Report No. CRWR-119. University of Texas, Texas, USA.

Benson, MA. 1962. *Evolution of Methods for Evaluating the Occurrence of Floods*. Report No. Water Supply Paper 1580-A. US Geological Survey, Washington DC, USA.

Benson, MA. 1964. *Factors Affecting the Occurence of Floods in the Southwest*. Report No. Water Supply Paper 1580-A. US Geological Survey, Washington DC, USA.

Benson, MA. 1968. Uniform Flood-Frequency Estimating Methods for Federal Agencies. *Water Resources Research* 4(5):891-908

Blöschl, G, Sivapalan, M, Wagener, T, Viglione, A and Savenije, H. 2013. *Runoff Prediction in Ungauged Basins: Synthesis across Processes, Places and Scales*. Cambridge University Press, Cambridge, UK.

Borujeni, SC and Sulaiman, WNA. 2009. Development of L-Moment Based Models for Extreme Flood Events. *Malaysian Journal of Mathematical Sciences* 3(2):281-296

Burn, D. 1997. Catchments Similarity for Regional Flood Frequency Analysis Using Seasonality Measures. *Journal of Hydrology* 202(1):212-230

Burn, DH. 1988. Delineation of Groups for Regional Flood Frequency Analysis. *Journal of Hydrology* 104(1):345-361

Burn, DH. 1990. An Appraisal of the "Region of Influence" Approach to Flood Frequency Analysis. *Hydrological Sciences Journal* 25(2):149-165

Burn, DH. 2003. The Use of Resampling for Estimating Confidence Intervals for Single Site and Pooled Frequency Analysis / Utilisation D'un Rééchantillonnage Pour L'estimation Des Intervalles De Confiance Lors D'analyses Fréquentielles Mono Et Multi-Site. *Hydrological Sciences Journal* 48(1):25-38

Burn, DH and Goel, NK. 2000. The Formation of Groups for Regional Flood Frequency Analysis. *Hydrological Sciences Journal* 45(1):97-112

Calitz, JP. 2016. Development and Assessment of a Probabilistic Rational Method for Design Flood Estimation in Selected Primary Drainage Regions in South Africa. Unpublished Bioresources Engineering, University of KwaZulu Natal, Pietermaritzburg, RSA.

Calitz, JP and Smithers, JC. 2020. *Development and Assessment of Regionalised Approaches to Design Flood Estimation in South Africa*. Report No. K5-2748. WRC, Pretoria, RSA.

Cassalho, F, Beskow, S, de Mello, CR and de Moura, MM. 2018. Regional Flood Frequency Analysis Using L-Moments for Geographically Defined Regions: An Assessment in Brazil. *Journal of Flood Risk Management* 2018(e12453):1-13

Castellarin, A, Domeneghetti, A and Brath, A. 2011. Identifying Robust Large-Scale Flood Risk Mitigation Strategies: A Quasi-2d Hydraulic Model as a Tool for the Po River. *Physics and Chemistry of the Earth, Parts A/B/C* 36(7–8):299-308

Castellarin, A, Kohnova, S, Gaal, L, Fleig, A, Salinas, JL, Toumazis, A, Kjeldsen, TR and Macdonald, N. 2012. *Review of Applied-Statistical Methods for Flood-Frequency Analysis in Europe*. Report No. ES0901. Centre for Ecology & Hydrology, Wallingford, UK.

Chadwick, A, Morfett, J and Borthwick, M. 2004. *Hydraulics in Civil and Environmental Engineering*. Spon, London, England.

Chen, Y, Wang, Q, Wang, L, Pieter, VG and Chen, M. 2007. Regional Flood Frequency Analysis by L-Moments for the Middle and Lower Reaches Region of the Yangtze River. *Water Resources and Environment* 1(2007):204-209

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

Cohn, TA, England, JF, Berenbrock, CE, Mason, RR, Stedinger, JR and Lamontagne, JR. 2013. A Generalized Grubbs-Beck Test Statistic for Detecting Multiple Potentially Influential Low Outliers in Flood Series. *Water Resources Research* 49(8):5047-5058

Cramér, H. 1928. On the Composition of Elementary Errors. *Scandinavian Actuarial Journal* 1928(1):13-74

Cruff, RW and Rantz, SE. 1965. *A Comparison of Methods Used in Flood-Frequency Studies for Coastal Basins in California*. United States Geological Survey, Washington, USA.

Cunnane, C. 1978. Unbiased Plotting Positions — a Review. *Journal of Hydrology* 37(3):205-222

Dalrymple, T. 1960. *Manual of Hydrology: Part 3 Flood Flow Techniques. Flood Frequency Analyses.* Report No. 1543-A. US Geol. Survey Water Supply, USA.

Davies, R. 2016. *Updated: South Africa – 7 Dead, Thousands Affected by Floods in Cape Town and Durban*. Date Accessed: 08/10/2016, Floodlist, http://floodlist.com/africa/south-africa-floods-cape-town-durban-july-2016.

Davies, R. 2017. *South Africa – Storm Leaves 8 Dead, Record Rain in Durban*. Date Accessed: 27/08/2018, Floodlist, http://floodlist.com/africa/kwazulu-natal-durban-flood-october-2017.

de Groen, M, Bailey, A and Commission, WR. 2015. *Water Resources of South Africa 2012*. Date Accessed: 2 February 2016, Water Research Commission, http://waterresourceswr2012.co.za.

Durbin, J and Knott, M. 1972. Components of Cramer-Von Mises Statistics. I. *Journal of the Royal Statistical Society. Series B (Methodological)* 34(2):290-307

DWS. 2011. *Hydrology - Department of Water Affairs*. Date Accessed: 2017-03-24, http://www.dwa.gov.za/Hydrology/Verified/hymain.aspx.

DWS. 2015. *U3h001*. Date Accessed: 01/05/2015, Department of Water and Sanitation, https://www.dwa.gov.za/Hydrology/HyImage.aspx?Station=U3H001.

eNCA. 2017. Durban Storm Damage Costs near R600-Million. eNCA, https://www.enca.com/south-africa/durban-storm-damage-costs-near-r600-million.

Eng, K, Tasker, GD and Milly, PCD. 2005. An Analysis of Region-of-Influence Methods for Flood Regionalisation in the Gulf-Atlantic Rolling Plains. *Journal of the American Water Resources Association* 41(1):135-143

England, JF, Cohn, TA, Faber, BA, Stedinger, JR, Thomas Jr, WO, Veilleux, AG, Kiang, JE and Mason Jr, RR. 2018. *Guidelines for Determining Flood Flow Frequency—Bulletin 17c*. Report No. 4-B5. Reston, VA, USA.

Filliben, JJ. 1975. The Probability Plot Correlation Coefficient Test for Normality. *Technometrics* 17(1):111-117

Gado, T and Nguyen, V. 2016. Comparison of Homogenous Region Delineation Approaches for Regional Flood Frequency Analysis at Ungauged Sites. *Journal of Hydrologic Engineering* 21(3):441-495

Gericke, OJ. 2010. Evaluation of the Sdf Method Using a Customised Design Flood Estimation Tool. Unpublished MScEng Thesis, Department of Civil Engineering, Stellenbosch University, Stellenbosch, RSA.

Gericke, OJ. 2015. Estimation of Catchment Response Time in Medium to Large Catchments in South Africa. Unpublished PhD Thesis, Bioresources Engineering, University of KwaZulu-Natal, Pietermaritzburg.

Gericke, OJ and Smithers, JC. 2018. An Improved and Consistent Approach to Estimate Catchment Response Time Parameters: Case Study in the C5 Drainage Region, South Africa. *Journal of Flood Risk Management* 11(S1):S284-S301

Görgens, AHM. 2002. Design Flood Hydrology. In: ed. Basson, G, *Design and Rehabilitation of Dams*, 460-524. Institute for Water and Environmental Engineering, Department of Civil Engineering, University of Stellenbosch, South Africa.

Görgens, AHM. 2007a. *Joint Peak-Volume (Jpv) Design Flood Hydrographs for South Africa*. Report No. 1420/3/07. Water Research Commission, Gezina, South Africa.

Görgens, AHM. 2007b. *Modernised South African Design Flood Practice in the Context of Dam Safety*. Report No. 1420/2/07. Water Research Commission, Gezina, South Africa.

Greenwood, JA, Landwehr, JM, Matalas, NC and Wallis, JR. 1979. Probability Weighted Moments: Definition and Relation to Parameters of Several Distributions Expressable in Inverse Form. *Water Resources Research* 15(5):1049-1054

Guse, B, Thieken, AH, Castellarin, A and Merz, B. 2010. Deriving Probabilistic Regional Envelope Curves with Two Pooling Methods. *Journal of Hydrology* 380(1):14-26

Haddad, K and Rahman, A. 2008. Investigation on at-Site Flood Frequency Analysis in South-East Australia. *The institution of engineers, Malaysia* 69(3):59-64

Haddad, K and Rahman, A. 2011. Selection of the Best Fit Flood Frequency Distribution and Parameter Estimation Procedure: A Case Study for Tasmania in Australia. *Stochastic Environmental Research and Risk Assessment* 25(3):415-428

Haddad, K and Rahman, A. 2012. Regional Flood Frequency Analysis in Eastern Australia: Bayesian Gls Regression-Based Methods within Fixed Region and Roi Framework – Quantile Regression Vs. Parameter Regression Technique. *Journal of hydrology* 430-431(2012):142-162

Haddad, K, Rahman, A and Ling, F. 2015. Regional Flood Frequency Analysis Method for Tasmania, Australia: A Case Study on the Comparison of Fixed Region and Region-of-Influence Approaches. *Hydrological Sciences Journal* 60(12):2086-2101

Haddad, K, Rahman, A and Stedinger, JR. 2012. Regional Flood Frequency Analysis Using Bayesian Generalized Least Squares: A Comparison between Quantile and Parameter Regression Techniques. *Hydrological Processes* 26(7):1008-1021

Haddad, K, Rahman, A, Weinmann, PE, Kuczera, G and Ball, J. 2010. Streamflow Data Preparation for Regional Flood Frequency Analysis: Lessons from Southeast Australia. *Australasian Journal of Water Resources* 14(1):17-32

Haile, AT. 2011. Regional Flood Frequency Analysis in Southern Africa. Unpublished MSc Thesis, Department of Geosciences, University of Oslo, Oslo, Norway.

Hassan, BGH and Ping, F. 2012. Formation of Homogenous Regions for Luanhe Basin – by Using L-Moments and Cluster Techniques. *International Journal of Environmental Science and Development* 3(2):205-210

He, Y, Bárdossy, A and Zehe, E. 2011. A Review of Regionalisation for Continuous Streamflow Simulation. *Hydrol. Earth Syst. Sci.* 15(11):3539-3553

Holloway, A, Fortune, G and Chasi, V. 2010. *Risk and Development Annual Review: Western Cape*. Report No. 2010. University of Cape Town, Cape Town, South Africa.

Hosking, JRM. 1990. L-Moment – Analysis and Estimation of Distributions Using Linear-Combinations of Order-Statistics. *Journal of the Royal Statistical Society* 52(1):105-124

Hosking, JRM. 1994. The Four-Parameter Kappa Distribution. *IBM Journal of Research and Development* 38(3):251-258

Hosking, JRM and Wallis, JR. 1993. Some Statistics Useful in Regional Frequency Analysis. *Water Resources Research* 29(2):271-281

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

HRU. 1972. *Design Flood Determination in South Africa*. Report No. 1/72. University of the Witwatersrand, Department of Civil Engineering, Johannesburg, South Africa.

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

Ilorme, F and Griffis, VW. 2013. A Novel Procedure for Delineation of Hydrologically Homogeneous Regions and the Classification of Ungauged Sites for Design Flood Estimation. *Journal of Hydrology* 492(2013):151-162

Kachroo, RK, Mkhandi, SH and Parida, BP. 2000. Flood Frequency Analysis of Southern Africa: I. Delineation of Homogeneous Regions. *Hydrological Sciences Journal* 45(3):437-447

Karim, F, Hasan, M and Marvanek, S. 2017. Evaluating Annual Maximum and Partial Duration Series for Estimating Frequency of Small Magnitude Floods. *Water* 9(481

Kjeldsen, T, Jones, D, Bayliss, A, Spencer, P, Surendran, S, Stefan, L, Webster, P and McDonald, D. 2008a. Improving the Feh Statistical Method. Environment Agency/Defra, University of Manchester, UK.

Kjeldsen, TR, Ahn, H and Prosdocimi, I. 2017. On the Use of a Four-Parameter Kappa Distribution in Regional Frequency Analysis. *Hydrological Sciences Journal* 62(9):1354-1363

Kjeldsen, TR and Jones, DA. 2006. Prediction Uncertainty in a Median-Based Index Flood Method Using L Moments. *Water Resources Research,* 42. doi: 10.1029/2005WR004069.

Kjeldsen, TR, Jones, DA and Bayliss, AC. 2008b. *Improving the Feh Statistical Procedures for Flood Frequency Estimation*. Report No. SC050050. Environment Agency, Bristol, UK.

Kjeldsen, TR, Lamb, R and Blazkova, SD. 2014. Uncertainty in Flood Frequency Analysis. In: ed. Beven, K and Hall, J, *Applied Uncertainty Analysis for Flood Risk Management*, 684. IMPERIAL COLLEGE PRESS, London, UK.

Kjeldsen, TR and Prosdocimi, I. 2015. A Bivariate Extension of the Hosking and Wallis Goodness-of-Fit Measure for Regional Distributions. *Water Resources Research* 51(2):896-907

Kjeldsen, TR, Smithers, JC and Schulze, RE. 2001. Flood Frequency Analysis at Ungauged Sites in the Kwazulu-Natal Province, South Africa. *Water SA* 27(3):315-324

Kjeldsen, TR, Smithers, JC and Schulze, RE. 2002. Regional Flood Frequency Analysis in the Kwazulu-Natal Province, South Africa, Using the Index-Flood Method. *Journal of hydrology* 2002(255):194-211

Kottegoda, NT and Rosso, R. 2008. *Applied Statistics for Civil and Environmental Engineers*. McGraw-Hill, West Sussex, UK.

Kovács, ZP. 1988. *Regional Maximum Flood Peaks in Southern Africa*. Department of Water Affairs, Pretoria, South Africa.

Laio, F. 2004. Cramer–Von Mises and Anderson-Darling Goodness of Fit Tests for Extreme Value Distributions with Unknown Parameters. *Water Resources Research*, 40. doi: 10.1029/2004WR003204.

Laio, F, Di Baldassarre, G and Montanari, A. 2009. Model Selection Techniques for the Frequency Analysis of Hydrological Extremes. *Water Resources Research*, 45. doi: 10.1029/2007WR006666.

Lynch, SD. 2004. *Development of a Raster Database of Annual, Monthly and Daily Rainfall for Southern Africa*. Report No. 1156-1-04. Water Research Commission, Gezina, South Africa.

Maidment, DR. 1993. *Handbook of Hydrology*. McGraw-Hill, New York, N.Y., USA.

Massey, FJ. 1951. The Kolmogorov-Smirnov Test for Goodness of Fit. *Journal of the American Statistical Association* 46(253):68-78

McDermott, GE and Pilgrim, DH. 1982. *Design Flood Estimation for Small Catchments in New South Wales*. Department of National Development and Energy, Canberra, Australia.

McInnes, L, Healy, J and Melville, J. 2018. Umap: Uniform Manifold Approximation and Projection for Dimension Reduction. Tutte Institute for Mathematics and Computing, arXiv Pre-print.

Meigh, JR, Farquharson, FAK and Sutcliffe, JV. 1997. A Worldwide Comparison of Regional Flood Estimation Methods and Climate. *Hydrological Sciences Journal* 42(2):225-244

Merz, R and Blöschl, G. 2005. Flood Frequency Regionalisation - Spatial Proximity Vs. Catchment Attributes. *Journal of hydrology* 302(1):283-306

Mkhandi, SH and Kachroo, RK. 1997. *Regional Flood Frequency Analysis for Southern Africa. Southern African Friend, Technical Documents in Hydrology 15*. UNESCO, Paris, France.

Mkhandi, SH, Kachroo, RK and Gunsdeksrs, TAG. 2000. Flood Frequency Analysis for Southern Africa: Ii. Identification of Regional Distributions. *Hydrological Sciences Journal* 45(3):449-464

Montanari, A, Young, G, Savenije, HHG, Hughes, D, Wagener, T, Ren, LL, Koutsoyiannis, D, Cudennec, C, Toth, E, Grimaldi, S, Blöschl, G, Sivapalan, M, Beven, K, Gupta, H, Hipsey, M, Schaefli, B, Arheimer, B, Boegh, E, Schymanski, SJ, Di Baldassarre, G, Yu, B, Hubert, P, Huang, Y, Schumann, A, Post, DA, Srinivasan, V, Harman, C, Thompson, S, Rogger, M, Viglione, A, McMillan, H, Characklis, G, Pang, Z and Belyaev, V. 2013. "Panta Rhei—Everything Flows": Change in Hydrology and Society—the Iahs Scientific Decade 2013–2022. *Hydrological Sciences Journal* 58(6):1256-1275

Mostofi Zadeh, S and Burn, DH. 2019. A Super Region Approach to Improve Pooled Flood Frequency Analysis. *Canadian Water Resources Journal / Revue canadienne des ressources hydriques*, doi: 10.1080/07011784.2018.1548946.

Mulvaney, TJ. 1850. On the Use of Self-Registering Rain and Flood Gauges in Making Observations on the Relationship of Rainfall and Runoff and of Flood Discharges in a Given Catchment. In: eds. *Transactions, Institute of Civil Engineers (Ireland)*, 10.

Naidoo, J. 2019. Personal communication, Department of water and sanitation, Pretoria, South Africa,

Naidoo, J. 2020. An Assessment of the Performance of Deterministic and Empirical Design Flood Estimation Methods Used in South Africa. Draft MSc. UKZN, Pietermaritzburg, RSA.

NASA-JPL. 2013. Nasa Shuttle Radar Topography Mission Global 1 Arc Second. Distributed by NASA EOSDIS Land Processes DAAC, https://doi.org/10.5067/MEaSUREs/SRTM/SRTMGL1.003. Date Accessed: 21/07/2017.

Nash, JE and Sutcliffe, JV. 1970. River Flow Forecasting through Conceptual Models Part I — a Discussion of Principles. *Journal of Hydrology* 10(3):282-290

Nathanael, JJ. 2015. Assessing the Performance of Regional Flood Frequency Analysis Methods in South Africa. Unpublished MSc Thesis, Centre for Water Resources Research, University of KwaZulu-Natal, Durban, South Africa.

Nathanael, JJ, Smithers, JC and Horan, MJC. 2018. Assessing the Performance of Regional Flood Frequency Analysis Methods in South Africa. *Water SA* 44(3 July):

Nobert, J, Mugo, M and Gadain, H. 2014. Estimation of Design Floods in Ungauged Catchments Using a Regional Index Flood Method. A Case Study of Lake Victoria Basin in Kenya. *Physics and Chemistry of the Earth, Parts A/B/C* 67–69(2014):4-11

Odry, J and Arnaud, P. 2017. Comparison of Flood Frequency Analysis Methods for Ungauged Catchments in France. *Geosciences* 7(3):88

Oudin, L, Andréassian, V, Perrin, C, Michel, C and Le Moine, N. 2008. Spatial Proximity, Physical Similarity, Regression and Ungaged Catchments: A Comparison of Regionalization Approaches Based on 913 French Catchments. *Water Resources Research*, 44. doi: 10.1029/2007WR006240.

Parak, M. 2007. Flood Estimation for Roads, Bridges and Dams. Unpublished Thesis MScEng, Department of Civil Engineering, University of KwaZulu-Natal, Durban, RSA.

Parak, M and Pegram, GGS. 2006. Rational Formula from Runhydrograph. *Water SA* 32(2):163-180

Pearson, CP. 1991. New Zealand Regional Flood Frequency Analysis Using L-Moments. *Journal of Hydrology (New Zealand)* 30(2):53-64

Pearson, K. 1900. On the Criterion That a Given System of Deviations from the Probable in the Case of a Correlated System of Variables Is Such That It Can Be Reasonably Supposed to Have Arisen from Random Sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 50(302):157-175

Pearson, K. 1901. Liii. On Lines and Planes of Closest Fit to Systems of Points in Space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 2(11):559-572

Peel, MC, Wang, QJ, Vogel, RM and McMahon, TA. 2001. The Utility of L-Moment Ratio Diagrams for Selecting a Regional Probability Distribution. *Hydrological Sciences Journal* 46(1):147-155

Pilgrim, D and Cordery, I. 1993. Flood Runoff. In: ed. Maidment, D, *Handbook of Hydrology*, 9, McGraw-Hill, New York, USA.

Pilgrim, DH. 1989. Regional Methods for Estimation of Design Floods for Small to Medium Sized Drainage Basins in Australia. In: eds. *Baltimore Symposium*, 14. Baltimore, USA.

Pilgrim, DH. 2001. *Australian Rainfall and Runoff : A Guide to Flood Estimation*. Institution of Engineers, Barton, A.C.T., Australia.

Pitman, VW. 2011. Overview of Water Resource Assessment in South Africa: Current State and Future Challenges. *Water SA* 37(5):659-664

Pratchett, T. 1988. *Sourcery*. Victor Gollancz, UK.

Rahman, A. 2005. A Quantile Regression Technique to Estimate Design Floods for Ungauged Catchments in South-East Australia. *Australasian Journal of Water Resources* 9(1):81-89

Rahman, A, Haddad, K, Haque, MM, Kuczera, G and Weinmann, E. 2015a. *Australian Rainfall and Runoff Revision Project 5: Regional Flood Methods (Stage 3)*. Report No. P5/S3/025. Engineers Australia, Australia.

Rahman, A, Haddad, K, Kuczera, G and Weinmann, E. 2009. *Australian Rainfall and Runoff Revision Project 5: Regional Flood Methods (Stage 1)*. Report No. P5/SI/003. University of Western Sydney, Sydney, Australia.

Rahman, A, Haddad, K, Kuczera, G and Weinmann, E. 2015b. Regional Flood Methods. In: ed. Ball, J, *Australian Rainfall and Runoff: Book 3 Draft for Comment*, Engineers Australia, Australia.

Rahman, A, Haddad, K, Kuczera, G and Weinmann, E. 2019. Book 3: Peak Flow Estimation. In: ed. Ball, J, *Australian Rainfall and Runoff: A Guide to Flood Estimaiton*, © Commonwealth of Australia (Geoscience Australia), Australia.

Rahman, A, Haddad, K, Rahman, AS and Haque, MM. 2015c. *Australian Rainfall and Runoff Revision Project 5: Regional Flood Methods: Database Used to Develop Arr Rffe Technique 2015 (Stage 3)*. Report No. P5/S3/026. University of Western Sydney, Sydney, Australia.

Rahman, A, Haddad, K, Zaman, M, Ishak, E, Kuczera, G and Weinmann, E. 2012. *Australian Rainfall and Runoff Revision Project 5: Regional Flood Methods (Stage 2)*. Report No. P5/S2/015. University of Western Sydney, Sydney, Australia.

Rahman, A, Haddad, K, Zaman, M, Kuczera, G and Weinmann, PE. 2011. Design Flood Estimation in Ungauged Catchments: A Comparison between the Probabilistic Rational Method and Quantile Regression Technique for Nsw. *Australasian Journal of Water Resources* 14(2):127-139

Rao, AR and Srinivas, VV. 2008. *Regionalization of Watersheds : An Approach Based on Cluster Analysis*. Springer Science+Business Media B.V., Dordrecht, Netherlands.

Razavi, T and Coulibaly, P. 2013. Streamflow Prediction in Ungauged Basins: Review of Regionalization Methods. *Journal of Hydrologic Engineering* 18(8):958-975

Reis, DS, Stedinger, JR and Martins, ES. 2003. Bayesian Gls Regression with Application to Lp3 Regional Skew Estimation. *Water Resources Research* 41(10):1049-1058

Ridolfi, E, Rianna, M, Trani, G, Alfonso, L, Di Baldassarre, G, Napolitano, F and Russo, F. 2016. A New Methodology to Define Homogeneous Regions through an Entropy Based Clustering Method. *Advances in Water Resources* 96(Supplement C):237-250

Riggs, HC. 1973. *Techniques of Water-Resources Investigations of the United States Geological Survey: Regional Analyses of Streamflow Characteristics*. United States Geological Survey, Washington, USA.

Robson, A and Reed, D. 1999. *Flood Estimation Handbook Volume 3: Statistical Procedures for Flood Frequency Estimation*. Institute of Hydrology, Oxfordshire, UK.

Rutkowska, A, Żelazny, M, Kohnová, S, Łyp, M and Banasik, K. 2016. Regional L-Moment-Based Flood Frequency Analysis in the Upper Vistula River Basin, Poland. *Pure and Applied Geophysics,* 174. doi: 10.1007/s00024-016-1298-8.

Salinas, JL, Castellarin, A, Viglione, A, Kohnová, S and Kjeldsen, TR. 2014. Regional Parent Flood Frequency Distributions in Europe – Part 1: Is the Gev Model Suitable as a Pan-European Parent? *Hydrol. Earth Syst. Sci.* 18(11):4381-4389

SANRAL. 2013. *Drainage Manual (6th Edition)*. South African National Roads Agency Limited, Pretoria, South Africa.

Scholz, FW and Stephens, MA. 1987. K-Sample Anderson-Darling Tests. *Journal of the American Statistical Association* 82(399):918-924

Schulze, RE. 2011. *A 2011 Perspective on Climate Change and the South African Water Sector*. Report No. WRC Report TT 518/12. Water Research Commission, Pretoria, RSA.

Schulze, RE, Schmidt, EJ and Smithers, JC. 2004. *Visual Scs-Sa User Manual : Pc-Based Scs Design Flood Estimates for Small Catchments in Southern Africa*. University of KwaZulu-Natal, Pietermaritzburg, South Africa.

Schulze, RE and Schütte, S. 2020. Mapping Scs Hydrological Soil Groups over South Africa at Terrain Unit Resolution. Draft Paper. UKZN, Pietermaritzburg, RSA.

Schulze, RE, South, A and Water Research, C. 1997. South African Atlas of Agrohydrology and -Climatology. Water Research Commission, Pretoria.

Schwarz, G. 1978. Estimating the Dimension of a Model. *The Annals of Statistics* 6(2):461-464

Shaw, EM. 1994. *Hydrology in Practice*. Chapman & Hall, London, England.

Shiceka, S. 2011. *Media Statement on the Declaration of a National Disaster as a Result of Flooding and Other Natural Disasters in the South Africa.* Ministry for cooperative governance and traditional affairs, Pretoria, South Africa.

Silva, A, Portela, M, Baez, J and Naghettini, M. 2012. Construction of Confidence Intervals for Extreme Rainfall Quantiles. *WIT Transactions on information and Communication technologies,* 44. doi: 10.2495/RISK120251.

Sinclair, CD, Spurr, BD and Ahmad, MI. 1990. Modified Anderson Darling Test. *Communications in Statistics - Theory and Methods* 19(10):3677-3686

Singh, K. 2019. Durban Floods Damage Estimated at over R650m. News24, https://www.news24.com/SouthAfrica/News/durban-floods-damage-estimated-at-over-r650m-20190426.

Sivapalan, M. 2003. Prediction in Ungauged Basins: A Grand Challenge for Theoretical Hydrology. *Hydrological Processes* 17(15):3163-3170

Smithers, JC. 1998. Development and Evaluation of Techniques for Estimating Short Duration Design Rainfall in South Africa. Unpublished PhD, Agricultural Engineering, Univeristy of Natal, Pietermaritzburg, South Africa.

Smithers, JC. 2012. Review of Methods for Design Flood Estimation in South Africa. *Water SA* 38(4):633-646

Smithers, JC, Görgens, A, Gericke, OJ, Jonker, V and Roberts, CPR. 2014. The Initiation of a National Flood Studies Programme for South Africa. SANCOLD, Birchwood, Gauteng.

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

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

Smithers, JC and Schulze, RE. 2003. *Design Rainfall and Flood Estimation in South Africa*. Report No. 1060/1/03. Water Research Commission, Gezina, South Africa.

Smithers, JC, Streatfield, J, Gray, RP and Oakes, EGM. 2015. Performance of Regional Flood Frequency Analysis Methods in Kwazulu-Natal, South Africa. *Water SA* 41(3):390-397

Stedinger, JR and Griffis, VW. 2008. Flood Frequency Analysis in the United States: Time to Update. *Journal of Hydrologic Engineering* 13(4):199-204

Stedinger, JR, Vogel, RM and Foufoula-Georgiou, E. 1993. Frequency Analysis of Extreme Events. In: ed. Maidment, DR, *Handbook of Hydrology*, 18, McGraw-Hill, Inc., New York.

Stephenson, D. 1981. *Stormwater Hydrology and Drainage*. Elsevier Scientific, Amsterdam, Netherlands.

Sugiura, N. 1978. Further Analysts of the Data by Akaike' S Information Criterion and the Finite Corrections. *Communications in Statistics - Theory and Methods* 7(1):13-26

Terrain Analysis Using Digital Elevation Models (Taudem): 5.3.7, 2016. USA, http://hydrology.usu.edu/taudem/.

Tasker, GD. 1982. Comparing Methods of Hydrologic Regionalization. *Water Resources Bulletin* 18(6):965-970

Thompson, DB. 2007. *The Rational Method*. R. O. Anderson Engineering, Minden, Nevada, USA.

Thorndike, RL. 1953. Who Belongs in the Family? *Psychometrika* 18(4):267-276

Ul Hassan, M, Hayat, O and Noreen, Z. 2019. Selecting the Best Probability Distribution for at-Site Flood Frequency Analysis; a Study of Torne River. *SN Applied Sciences* 1(12):1629

Van Bladeren, D. 1993. Application of Historical Flood Data in Flood Frequency Analysis, for the Natal and Transkei Region. In: eds. *Proceedings of Sixth South African National Hydrological Symposium*, 359-366. Department of Agricultural Engineering, University of Natal.

Van Bladeren, D. 2005. *Verification of the Proposed Standard Design Flood (Sdf)*. Report No. 344512/1. SRK Consulting (Pty) Ltd, Pretoria, South Africa.

Van Bladeren, D, Zawada, PKZ and Mahlangu, D. 2007. *Statistical Based Regional Flood Frequency Estimation Study for South Africa Using Systematic, Historical and Paleoflood Data - Pilot Study - Catchment Management Area 15*. Report No. 1260/1/07. WRC, Pretoria, RSA.

van der Maaten, LJP and Hinton, GE. 2008. Visualizing High-Dimensional Data Using T-Sne. *Journal of Machine Learning Research* 9(nov):2579-2605

Van der Spuy, D and Rademeyer, PF. 2018. Flood Frequency Estimation Methods as Applied in the Department of Water Affairs. Department of Water Affairs, Pretoria, South Africa.

Van Dijk, M, van Vuuren, SJ and Smithers, JC. 2013. Chapter 3: Flood Calculations. In: ed. Kruger, E, *Drainage Manual (Version 6)*, 3, 3.1-3.69. The South African Roads Agency SOC Limited, Pretoria, South Africa.

Van Vuuren, SJ, Van Dijk, M and Coetzee, GL. 2013. *Status Review and Requirements of Overhauling Flood Determination Methods in South Africa*. Report No. TT563/13. WRC, Pretoria, RSA.

Viglione, A, Laio, F and Claps, P. 2007. A Comparison of Homogeneity Tests for Regional Frequency Analysis. *Water Resources Research,* 43. doi: 10.1029/2006WR005095.

Vogel, RM and Fennessey, NM. 1993. L Moment Diagrams Should Replace Product Moment Diagrams. *Water Resources Research* 29(6):1745-1752

Vogel, RM, McMahon, TA and Chiew, FHS. 1993. Floodflow Frequency Model Selection in Australia. *Journal of Hydrology* 146(1993):421-449

von Mises, RE. 1928. *Wahrscheinlichkeit, Statistik Und Wahrheit*. Julius Springer, Vienna, Austria.

Wang, QJ. 1997. Lh Moments for Statistical Analysis of Extreme Events. *Water Resources Research* 33(12):2841-2848

Weisberg, S. 2005. *Applied Linear Regression*. John Wiley and Sons, Inc., New Jersey, USA.

Wiltshire, SE. 1986a. Regional Flood Frequency Analysis Ii: Multivariate Classification of Drainage Basins in Britain. *Hydrological Sciences Journal* 31(3):335-346

Wiltshire, SE. 1986b. Regional Flood Frequency Analysis, I, Homegeneity Statistics. *Hydrological Sciences Journal* 31(3):321-333

Zafirakou-Koulouris, A, Vogel, RM, Craig, SM and Habermeier, J. 1998. L Moment Diagrams for Censored Observations. *Water Resources Research* 34(5):1241-1249

Zeng, X, Wang, D and Wu, J. 2015. Evaluating the Three Methods of Goodness of Fit Test for Frequency Analysis. *Journal of Risk Analysis and Crisis Response* 5(3):178-187

Zhang, Z, Stadnyk, TA and Burn, DH. 2019. Identification of a Preferred Statistical Distribution for at-Site Flood Frequency Analysis in Canada. *Canadian Water Resources Journal / Revue canadienne des ressources hydriques* 1-16

Zrinji, Z and Burn, H. 1996. Regional Flood Frequency with Hierarchical Region of Influence. *Journal of Water Resources Planning and Management* 122(4):245-252

APPENDIX A: CALIBRATED *C* VALUES SUMMARY PER AEP FOR THE 42 HOMOGENEOUS CLUSTERS

Cluster	Statistic	AEP (%)						
		50	20	10	5	2	1	0.5
1	Max	0.326	0.327	0.345	0.377	0.402	0.412	0.415
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.167	0.217	0.234	0.237	0.205	0.184	0.178
	Average	0.166	0.211	0.219	0.218	0.210	0.202	0.192
2	Max	0.052	0.094	0.124	0.152	0.212	0.273	0.350
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.030	0.060	0.082	0.109	0.137	0.166	0.201
	Average	0.031	0.060	0.081	0.104	0.138	0.168	0.202
3	Max	0.030	0.063	0.097	0.143	0.230	0.326	0.462
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.028	0.048	0.072	0.092	0.120	0.144	0.171
	Average	0.025	0.049	0.070	0.095	0.140	0.187	0.251
4	Max	0.036	0.080	0.118	0.161	0.230	0.294	0.372
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.025	0.063	0.078	0.096	0.109	0.117	0.123
	Average	0.023	0.050	0.069	0.091	0.123	0.152	0.188
5	Max	0.135	0.247	0.299	0.391	0.533	0.663	0.818
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.030	0.062	0.087	0.114	0.165	0.215	0.266
	Average	0.047	0.090	0.120	0.152	0.201	0.248	0.307
6	Max	0.084	0.174	0.237	0.300	0.385	0.454	0.528
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.054	0.099	0.129	0.157	0.196	0.220	0.251
	Average	0.055	0.107	0.141	0.173	0.215	0.248	0.283
7	Max	0.146	0.182	0.176	0.162	0.157	0.158	0.156
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.082	0.112	0.115	0.111	0.102	0.094	0.088
	Average	0.089	0.120	0.123	0.121	0.114	0.108	0.101
8	Max	0.062	0.119	0.159	0.199	0.254	0.298	0.345
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.040	0.072	0.100	0.130	0.172	0.207	0.247
	Average	0.044	0.080	0.104	0.127	0.158	0.183	0.211
9	Max	0.123	0.181	0.194	0.195	0.186	0.178	0.213
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.050	0.078	0.090	0.111	0.142	0.162	0.164
	Average	0.055	0.088	0.104	0.117	0.130	0.138	0.146
10	Max	0.226	0.361	0.413	0.440	0.451	0.446	0.434
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.060	0.106	0.134	0.152	0.173	0.186	0.198
	Average	0.070	0.118	0.145	0.165	0.188	0.202	0.216
11	Max	0.071	0.150	0.207	0.268	0.358	0.437	0.655
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018

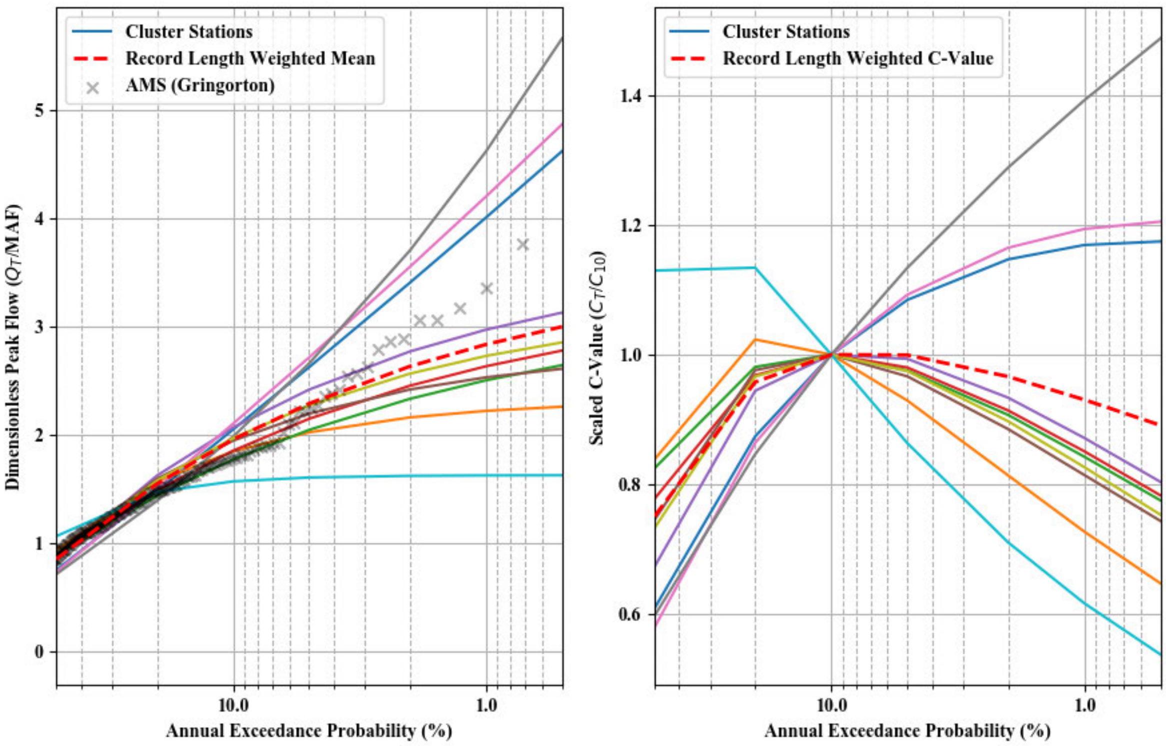
Cluster	Statistic	AEP (%)						
		50	20	10	5	2	1	0.5
	Median	0.024	0.053	0.074	0.104	0.140	0.171	0.208
	Average	0.030	0.060	0.084	0.112	0.158	0.206	0.270
12	Max	0.160	0.314	0.434	0.567	0.772	0.960	1.185
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.065	0.123	0.161	0.198	0.251	0.292	0.322
	Average	0.071	0.138	0.184	0.231	0.296	0.353	0.417
13	Max	0.238	0.397	0.462	0.499	0.519	0.521	0.515
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.080	0.151	0.198	0.239	0.266	0.307	0.377
	Average	0.116	0.196	0.238	0.271	0.308	0.334	0.361
14	Max	0.123	0.252	0.343	0.434	0.561	0.663	0.776
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.045	0.083	0.101	0.118	0.143	0.170	0.186
	Average	0.054	0.106	0.138	0.168	0.205	0.233	0.263
15	Max	0.435	0.632	0.712	0.786	0.885	0.945	0.996
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.115	0.180	0.211	0.233	0.258	0.274	0.299
	Average	0.172	0.276	0.329	0.368	0.408	0.432	0.454
16	Max	0.114	0.239	0.343	0.459	0.638	0.801	0.995
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.081	0.139	0.177	0.213	0.263	0.309	0.360
	Average	0.082	0.146	0.192	0.238	0.304	0.361	0.426
17	Max	0.150	0.266	0.365	0.478	0.660	0.830	1.036
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.084	0.145	0.170	0.208	0.290	0.366	0.458
	Average	0.095	0.173	0.230	0.290	0.381	0.463	0.562
18	Max	0.068	0.135	0.184	0.236	0.316	0.387	0.469
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.021	0.049	0.074	0.105	0.159	0.216	0.293
	Average	0.033	0.070	0.098	0.129	0.178	0.226	0.286
19	Max	0.041	0.065	0.075	0.089	0.107	0.120	0.132
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.032	0.058	0.074	0.080	0.091	0.098	0.103
	Average	0.035	0.060	0.073	0.083	0.094	0.100	0.106
20	Max	0.425	0.534	0.552	0.546	0.519	0.494	0.467
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.166	0.204	0.205	0.198	0.183	0.172	0.161
	Average	0.208	0.265	0.276	0.274	0.263	0.251	0.239
21	Max	0.388	0.501	0.544	0.585	0.673	0.733	0.787
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.098	0.124	0.141	0.153	0.170	0.184	0.197
	Average	0.162	0.226	0.258	0.283	0.306	0.321	0.332

Cluster	Statistic	AEP (%)						
		50	20	10	5	2	1	0.5
22	Max	0.059	0.134	0.187	0.242	0.321	0.387	0.460
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.025	0.049	0.073	0.098	0.122	0.160	0.209
	Average	0.030	0.057	0.078	0.101	0.139	0.174	0.219
23	Max	0.107	0.161	0.174	0.177	0.172	0.165	0.157
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.047	0.063	0.069	0.071	0.072	0.071	0.069
	Average	0.062	0.085	0.089	0.089	0.086	0.082	0.077
24	Max	0.159	0.251	0.312	0.391	0.534	0.665	0.819
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.117	0.218	0.284	0.303	0.312	0.310	0.304
	Average	0.089	0.153	0.192	0.229	0.277	0.316	0.359
25	Max	0.225	0.401	0.479	0.528	0.565	0.578	0.682
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.050	0.102	0.142	0.185	0.228	0.301	0.379
	Average	0.072	0.130	0.166	0.199	0.245	0.282	0.325
26	Max	0.112	0.203	0.266	0.328	0.412	0.481	0.602
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.045	0.118	0.177	0.221	0.238	0.245	0.248
	Average	0.059	0.118	0.156	0.195	0.249	0.297	0.352
27	Max	0.264	0.388	0.453	0.500	0.546	0.571	0.588
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.140	0.212	0.240	0.253	0.259	0.256	0.249
	Average	0.167	0.236	0.262	0.275	0.280	0.279	0.274
28	Max	0.392	0.707	0.879	1.021	1.176	1.276	1.365
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.123	0.216	0.267	0.310	0.358	0.390	0.420
	Average	0.148	0.272	0.345	0.411	0.491	0.551	0.611
29	Max	0.074	0.114	0.155	0.213	0.312	0.414	0.547
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.046	0.085	0.115	0.136	0.157	0.173	0.194
	Average	0.050	0.087	0.111	0.136	0.171	0.203	0.241
30	Max	0.072	0.164	0.248	0.345	0.506	0.662	0.859
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.045	0.106	0.153	0.219	0.325	0.390	0.518
	Average	0.048	0.115	0.170	0.235	0.343	0.451	0.590
31	Max	0.060	0.110	0.143	0.173	0.210	0.238	0.267
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.039	0.064	0.083	0.098	0.126	0.146	0.163
	Average	0.039	0.067	0.085	0.103	0.127	0.145	0.165
32	Max	0.128	0.180	0.235	0.289	0.362	0.419	0.481
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018

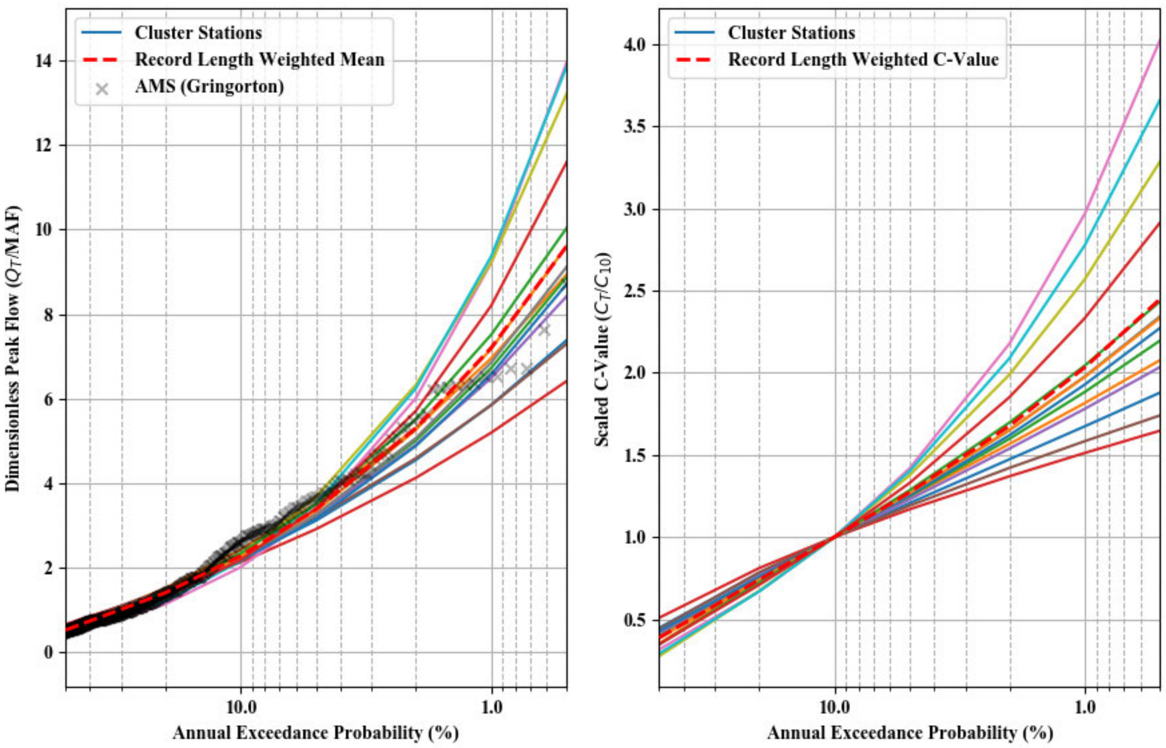
Cluster	Statistic	AEP (%)						
		50	20	10	5	2	1	0.5
	Median	0.081	0.149	0.142	0.164	0.199	0.227	0.255
	Average	0.081	0.130	0.154	0.174	0.197	0.213	0.230
33	Max	0.293	0.413	0.461	0.487	0.498	0.502	0.583
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.148	0.200	0.211	0.209	0.213	0.216	0.235
	Average	0.151	0.219	0.250	0.270	0.287	0.295	0.301
34	Max	0.137	0.294	0.429	0.589	0.853	1.109	1.430
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.053	0.101	0.108	0.124	0.154	0.214	0.286
	Average	0.075	0.135	0.176	0.217	0.278	0.333	0.400
35	Max	0.256	0.361	0.397	0.411	0.410	0.471	0.574
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.108	0.177	0.220	0.245	0.268	0.307	0.339
	Average	0.116	0.173	0.206	0.233	0.266	0.291	0.318
36	Max	0.179	0.266	0.300	0.318	0.325	0.323	0.317
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.104	0.152	0.176	0.194	0.211	0.213	0.211
	Average	0.109	0.158	0.178	0.190	0.198	0.201	0.201
37	Max	0.217	0.445	0.614	0.793	1.060	1.296	1.569
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.086	0.139	0.174	0.215	0.280	0.360	0.461
	Average	0.093	0.173	0.234	0.300	0.401	0.494	0.606
38	Max	0.103	0.157	0.177	0.187	0.190	0.187	0.189
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.050	0.078	0.094	0.106	0.121	0.129	0.138
	Average	0.053	0.081	0.095	0.106	0.117	0.125	0.132
39	Max	0.124	0.216	0.297	0.394	0.558	0.715	0.914
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.049	0.082	0.110	0.129	0.165	0.194	0.226
	Average	0.057	0.096	0.125	0.155	0.202	0.246	0.302
40	Max	0.099	0.159	0.207	0.256	0.320	0.368	0.419
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.036	0.065	0.093	0.124	0.149	0.167	0.186
	Average	0.045	0.081	0.105	0.128	0.159	0.186	0.216
41	Max	0.048	0.063	0.071	0.078	0.085	0.090	0.093
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.015	0.023	0.025	0.025	0.030	0.034	0.039
	Average	0.020	0.028	0.032	0.035	0.038	0.040	0.042
42	Max	0.049	0.096	0.131	0.166	0.218	0.260	0.308
	Min	0.003	0.008	0.012	0.015	0.019	0.019	0.018
	Median	0.033	0.070	0.091	0.118	0.164	0.194	0.227
	Average	0.035	0.067	0.089	0.112	0.145	0.173	0.204

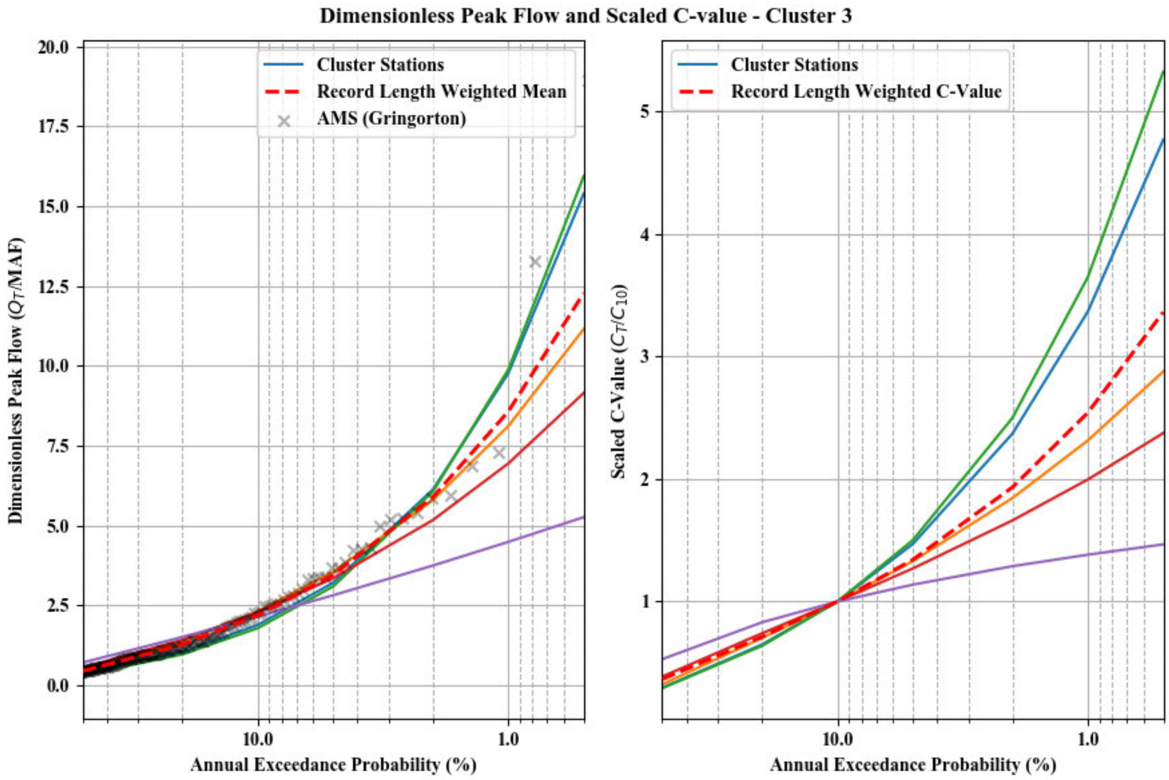
APPENDIX B: DIMENSIONLESS GROWTH CURVES FOR EACH OF THE 42 RELATIVELY HOMOGENEOUS CLUSTERS

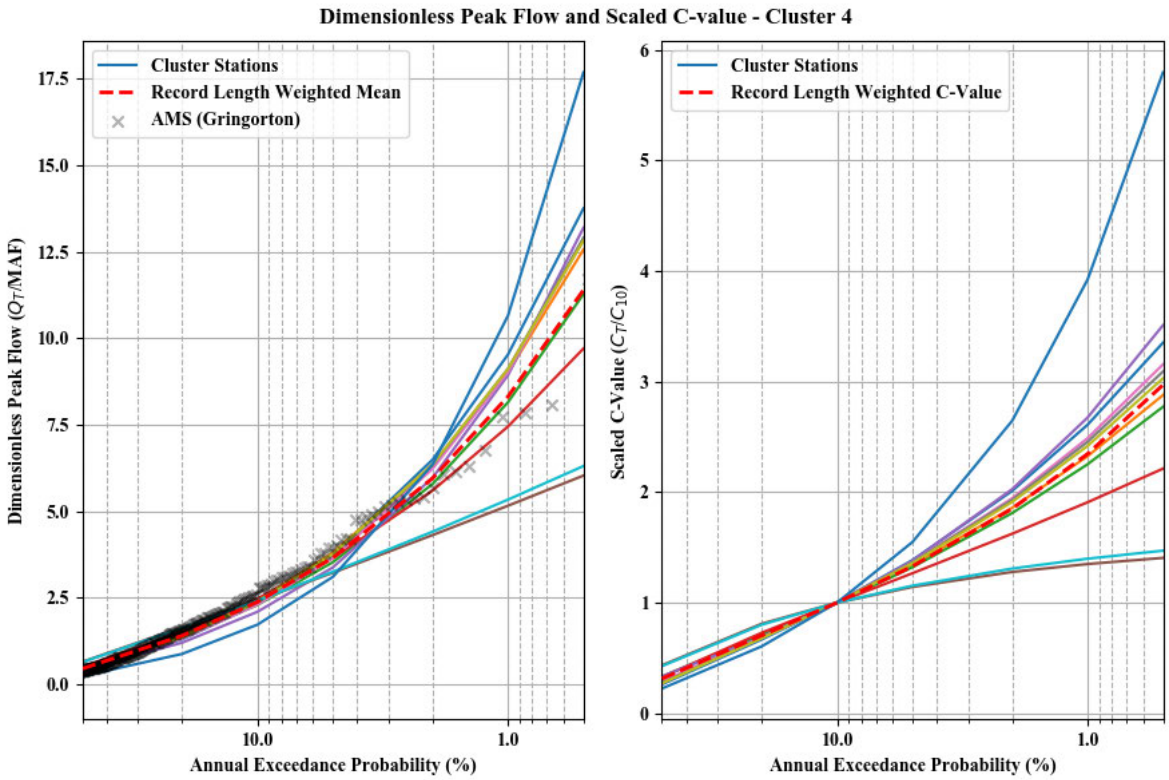
Dimensionless Peak Flow and Scaled C-value - Cluster 1



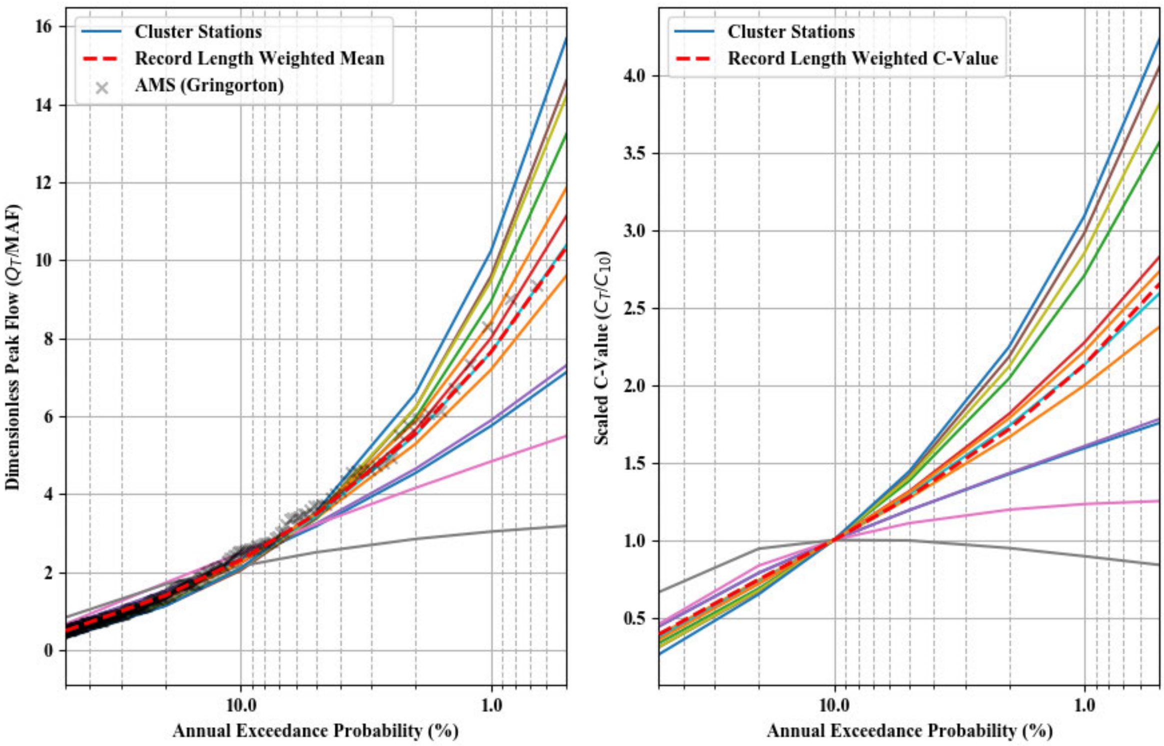
Dimensionless Peak Flow and Scaled C-value - Cluster 2



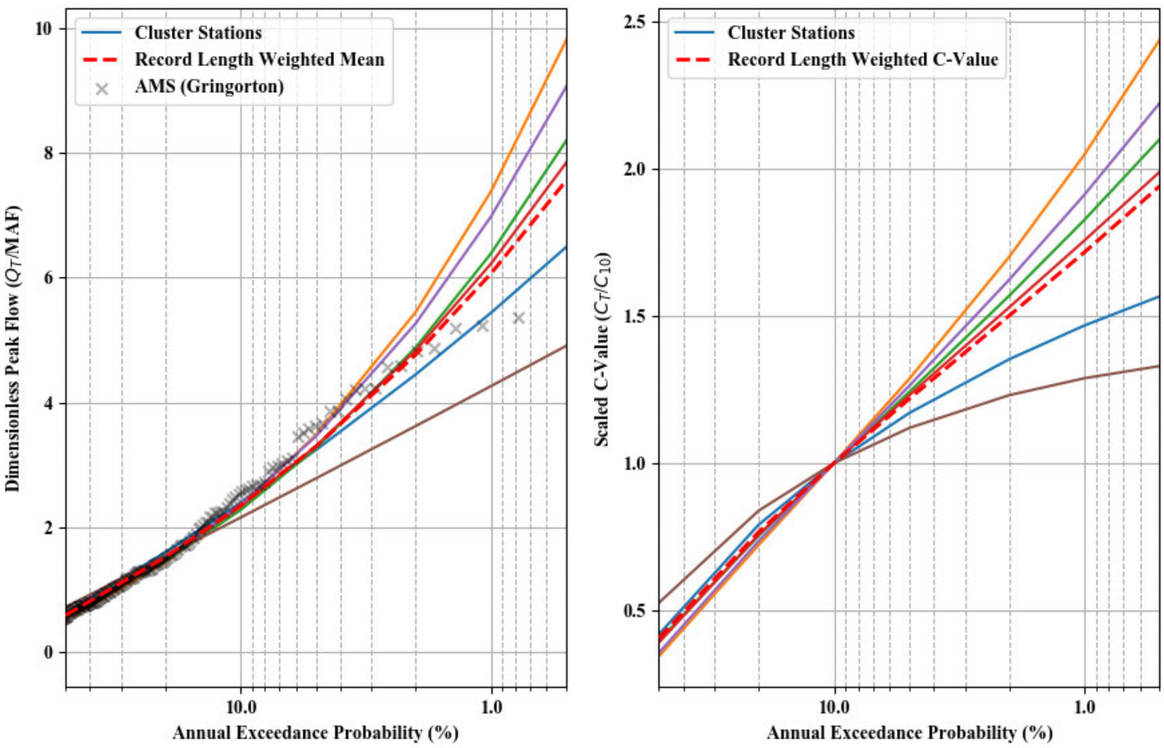

Dimensionless Peak Flow and Scaled C-value - Cluster 3
Cluster Stations
Record Length Weighted Mean
AMS (Gringorton)
Dimensionless Peak Flow (Q_T/MAF)
Annual Exceedance Probability (%)
Cluster Stations
Record Length Weighted C-Value
Scaled C-Value (C_T/C_{10})
Annual Exceedance Probability (%)


Dimensionless Peak Flow and Scaled C-value - Cluster 4
Cluster Stations
Record Length Weighted Mean
AMS (Gringorton)
Dimensionless Peak Flow (Q_T/MAF)
Annual Exceedance Probability (%)
Cluster Stations
Record Length Weighted C-Value
Scaled C-Value (C_T/C_{10})
Annual Exceedance Probability (%)

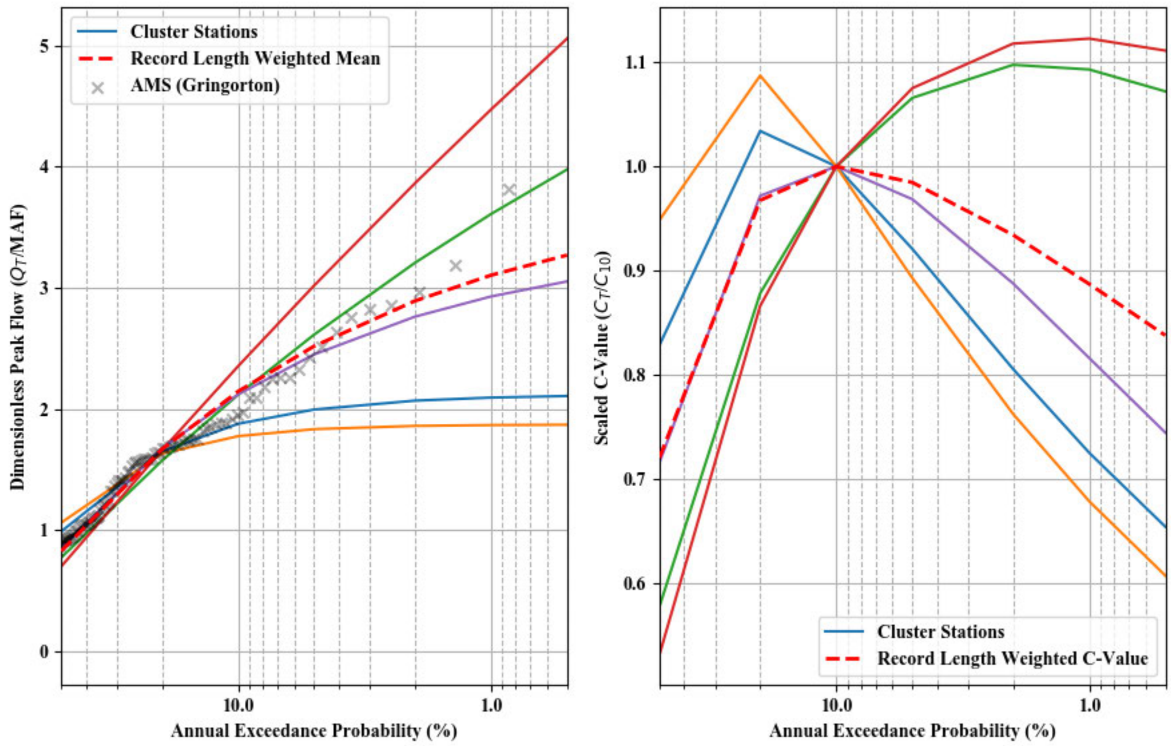
Dimensionless Peak Flow and Scaled C-value - Cluster 5



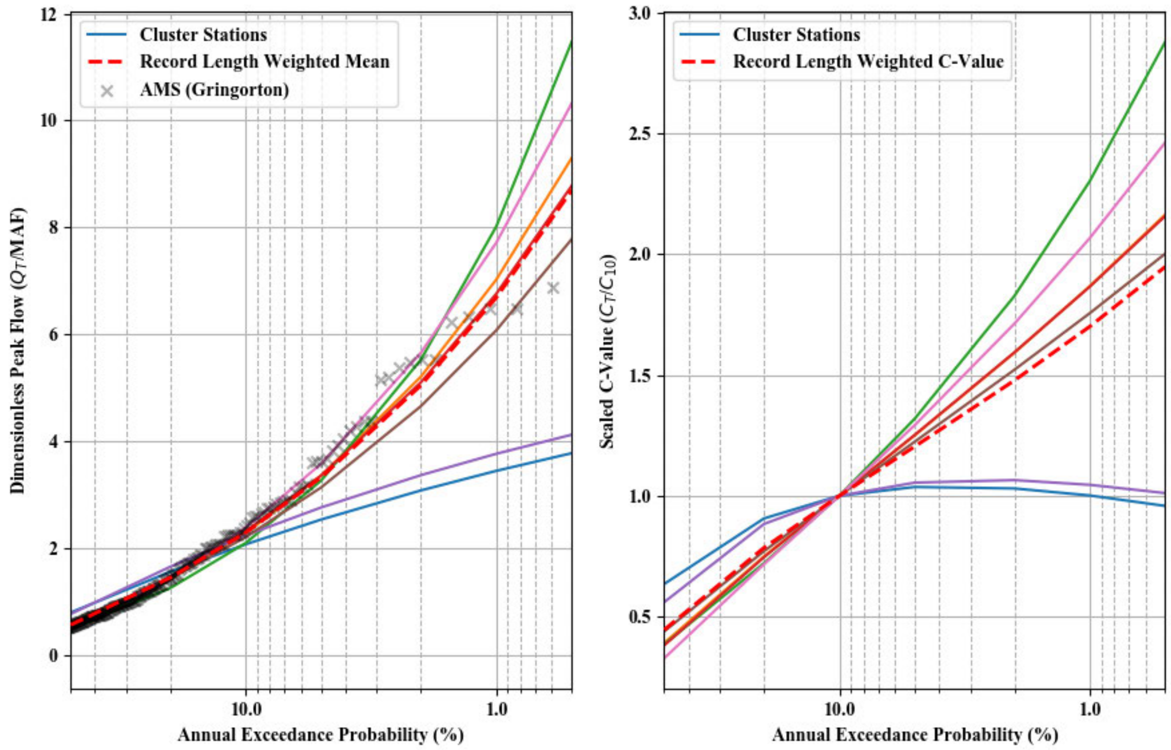
Dimensionless Peak Flow and Scaled C-value - Cluster 6

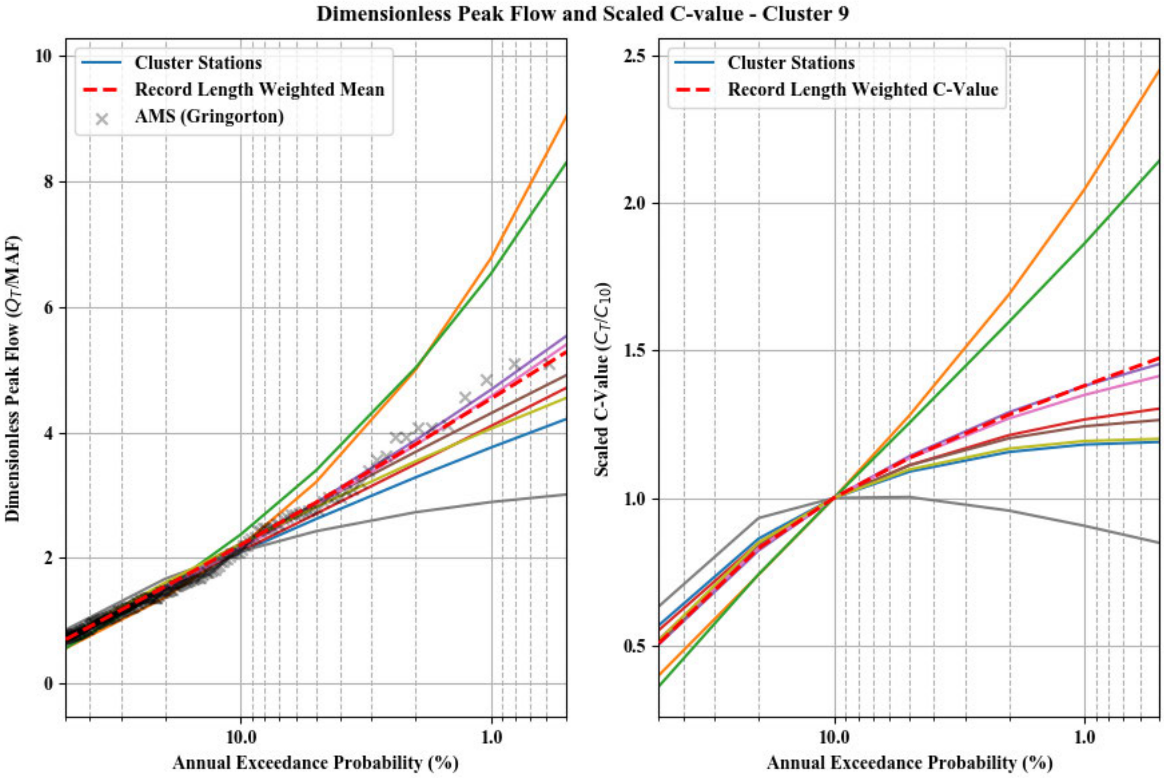


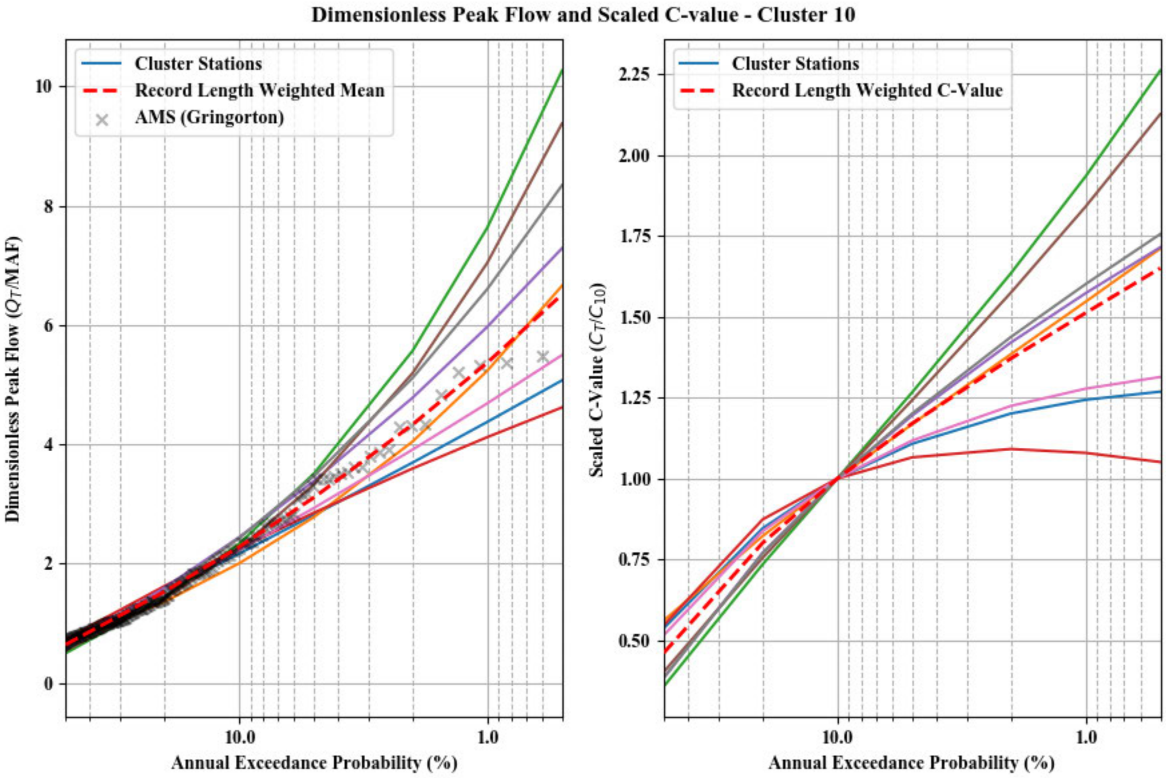
Dimensionless Peak Flow and Scaled C-value - Cluster 7



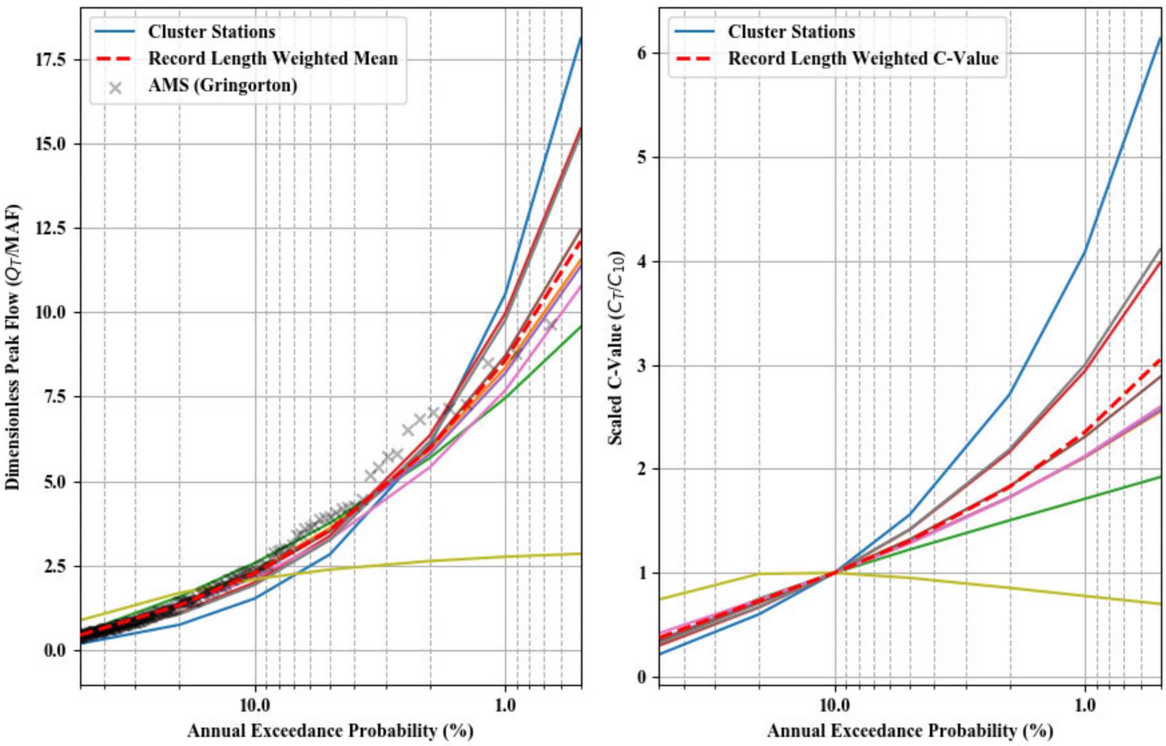
Dimensionless Peak Flow and Scaled C-value - Cluster 8



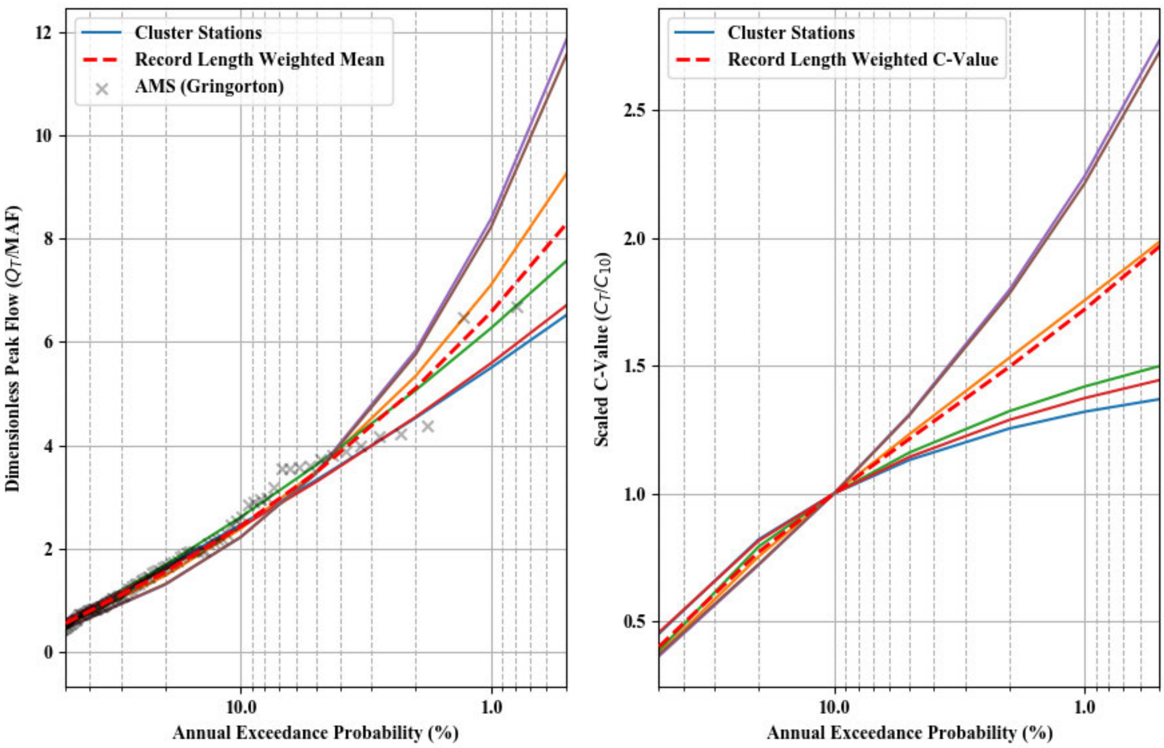

Dimensionless Peak Flow and Scaled C-value - Cluster 9
Cluster Stations
Record Length Weighted Mean
AMS (Gringorton)
Dimensionless Peak Flow (Q_T/MAF)
Annual Exceedance Probability (%)
Cluster Stations
Record Length Weighted C-Value
Scaled C-Value (C_T/C_{10})
Annual Exceedance Probability (%)


Dimensionless Peak Flow and Scaled C-value - Cluster 10
Cluster Stations
Record Length Weighted Mean
AMS (Gringorton)
Dimensionless Peak Flow (Q_T/MAF)
Annual Exceedance Probability (%)
Cluster Stations
Record Length Weighted C-Value
Scaled C-Value (C_T/C_{10})
Annual Exceedance Probability (%)

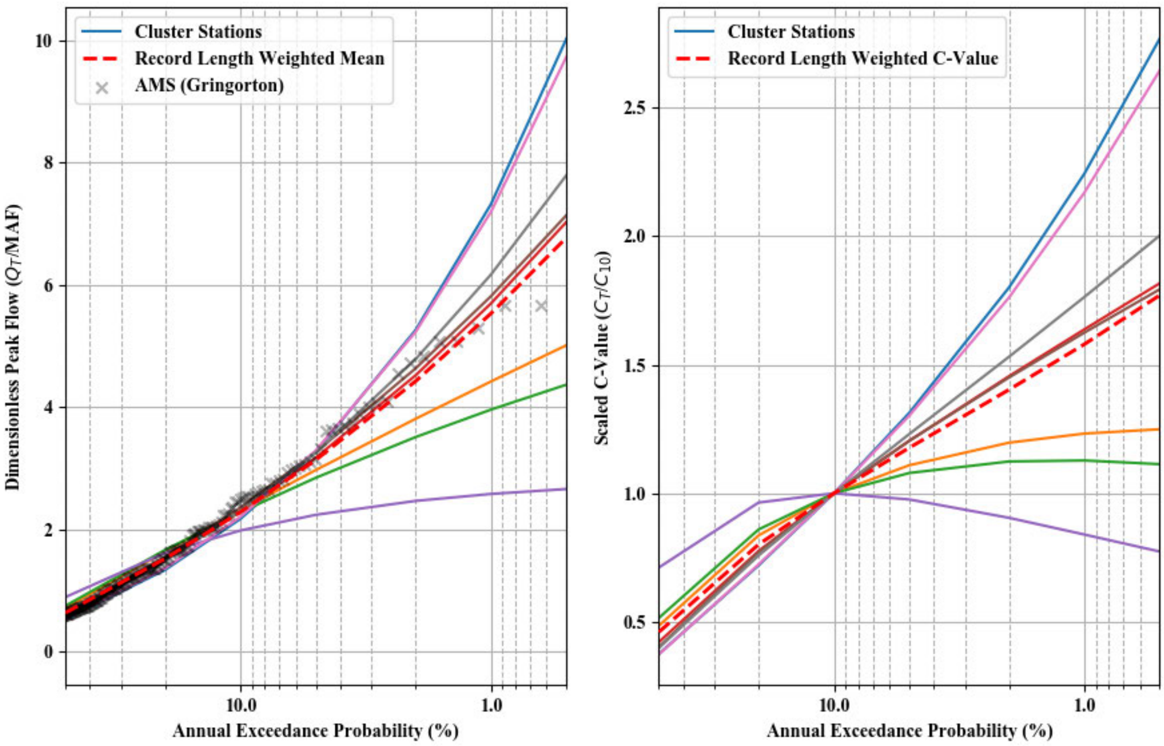
Dimensionless Peak Flow and Scaled C-value - Cluster 11



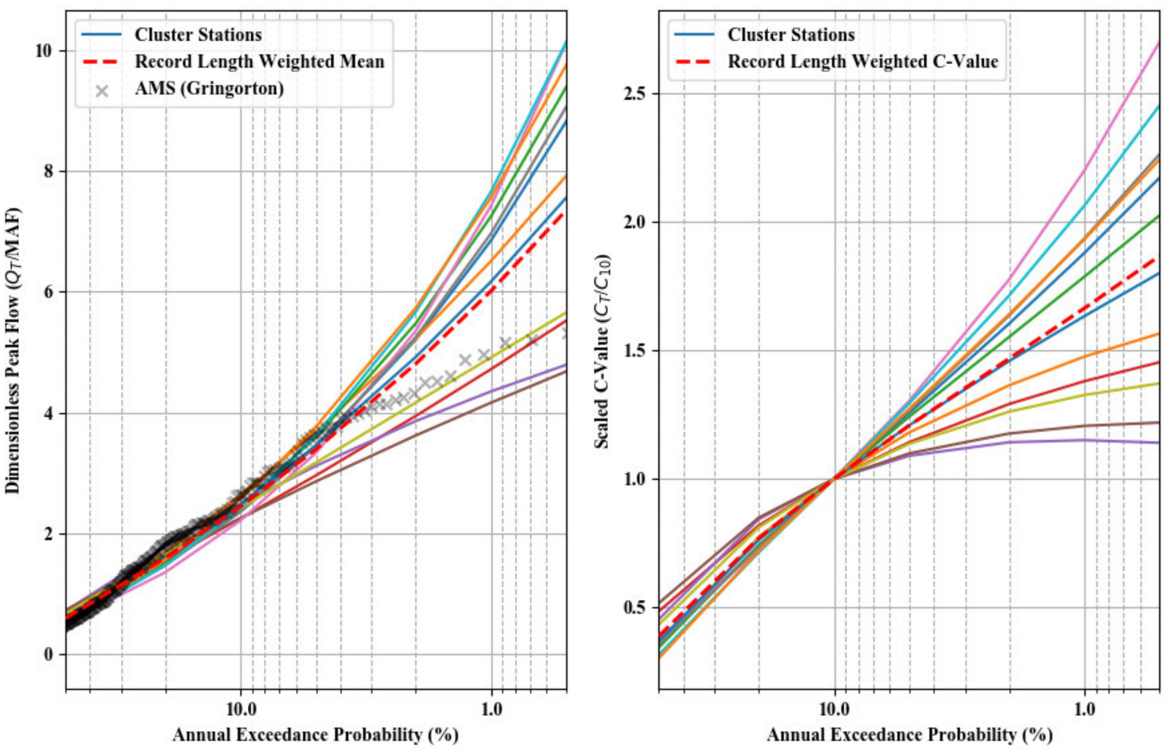
Dimensionless Peak Flow and Scaled C-value - Cluster 12



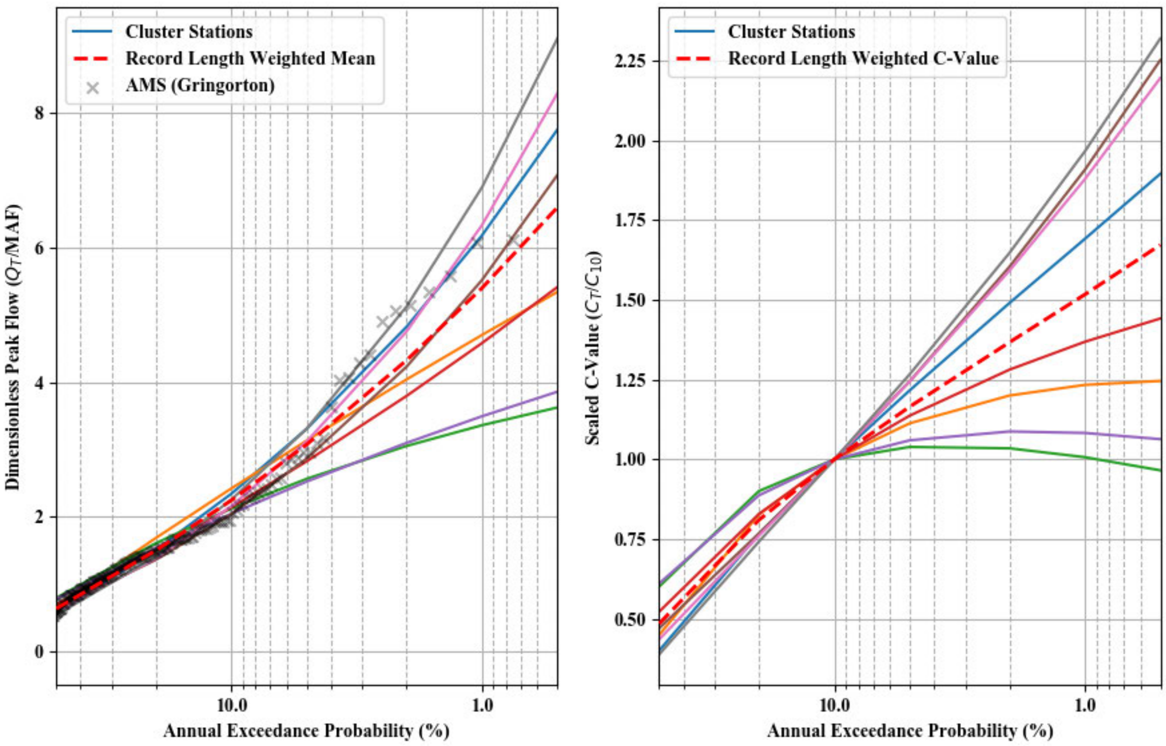
Dimensionless Peak Flow and Scaled C-value - Cluster 13



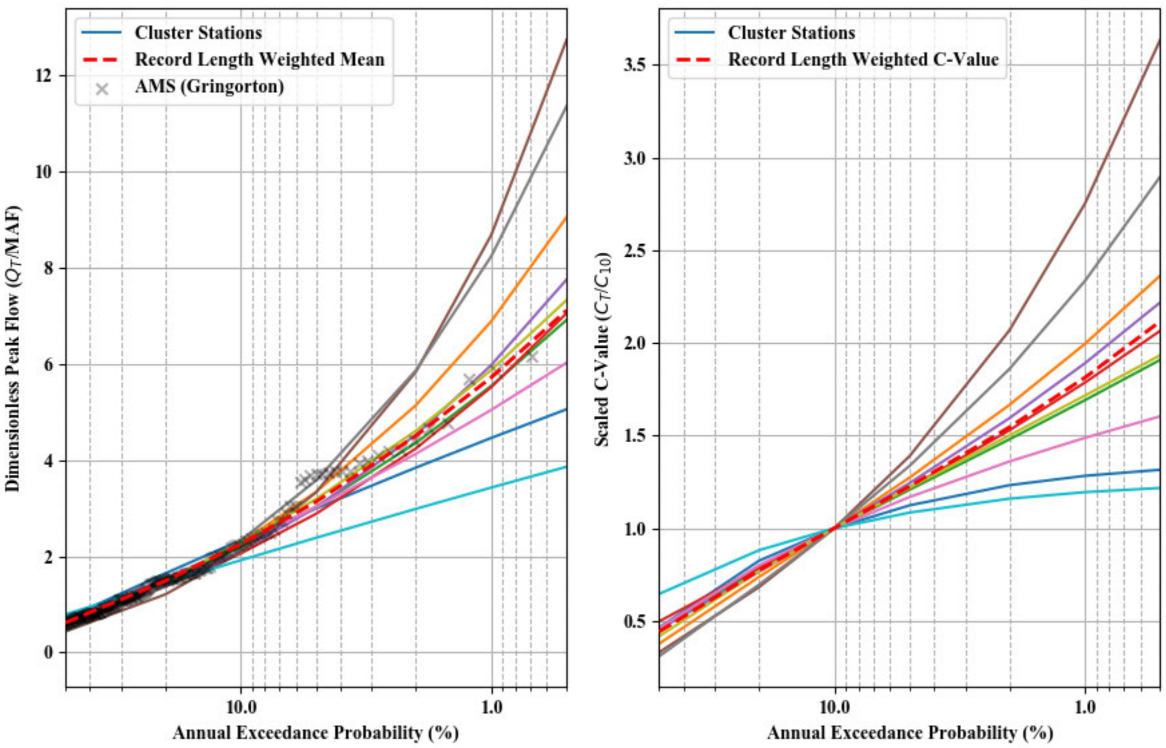
Dimensionless Peak Flow and Scaled C-value - Cluster 14

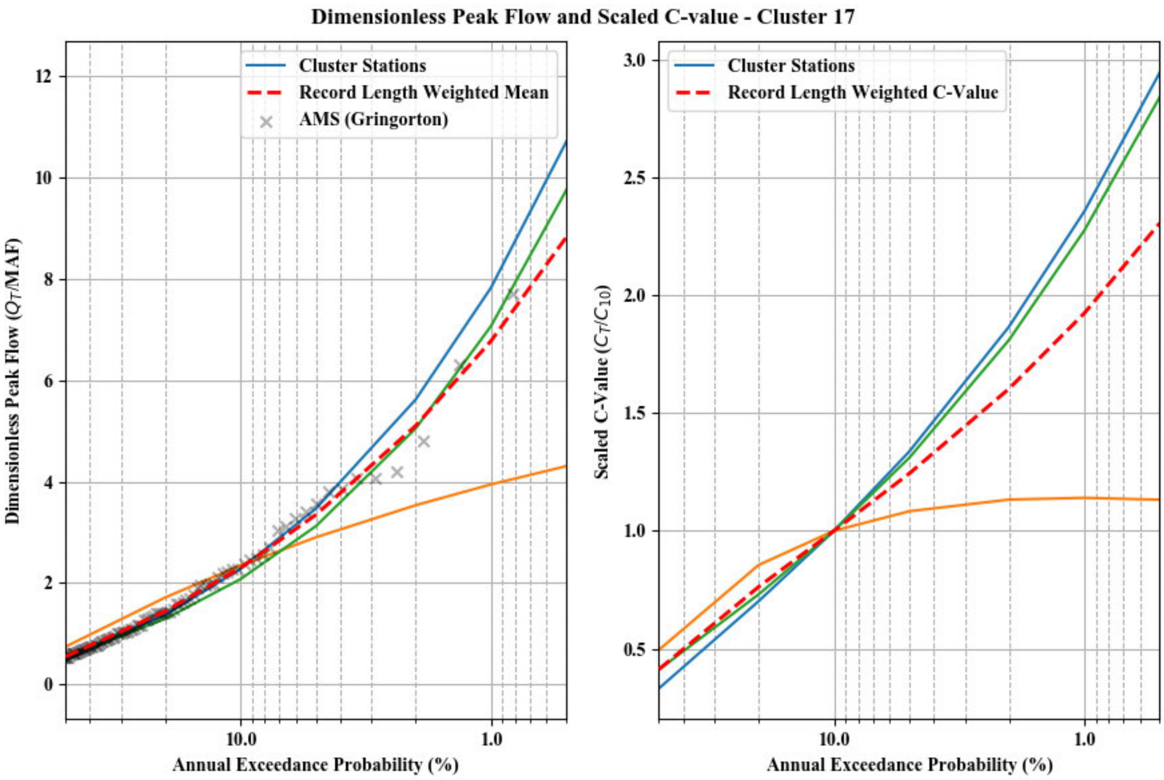


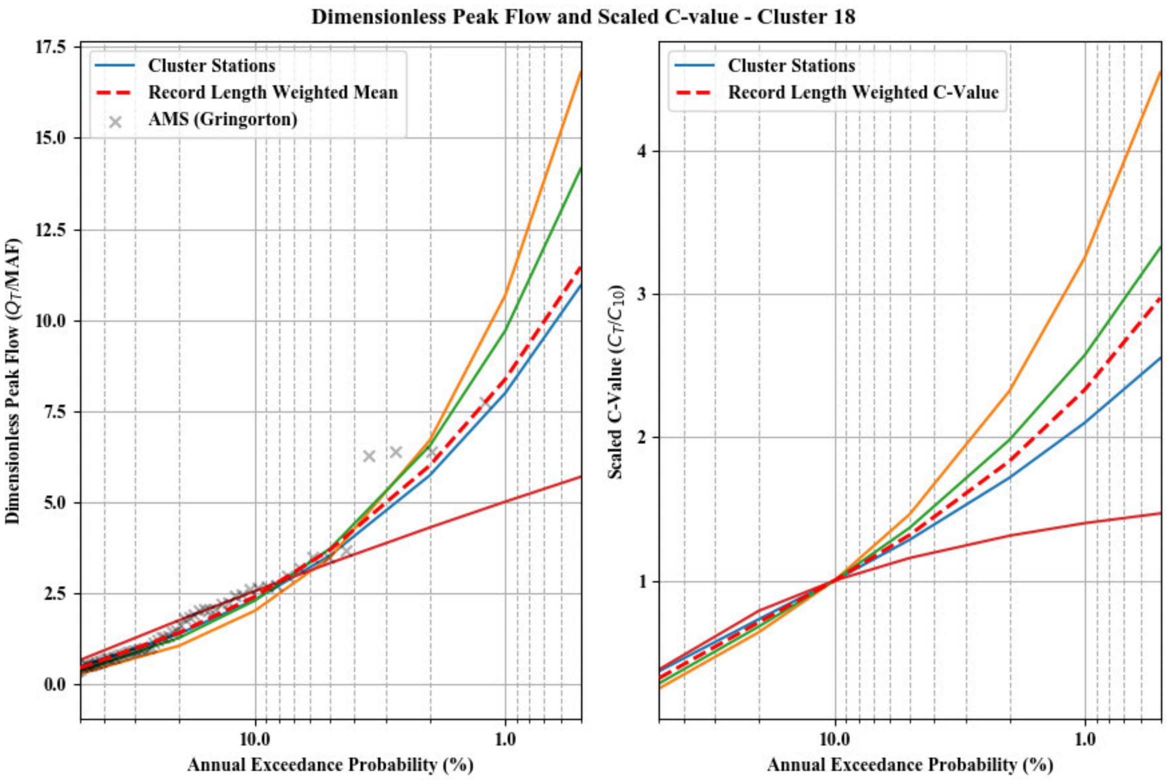
Dimensionless Peak Flow and Scaled C-value - Cluster 15



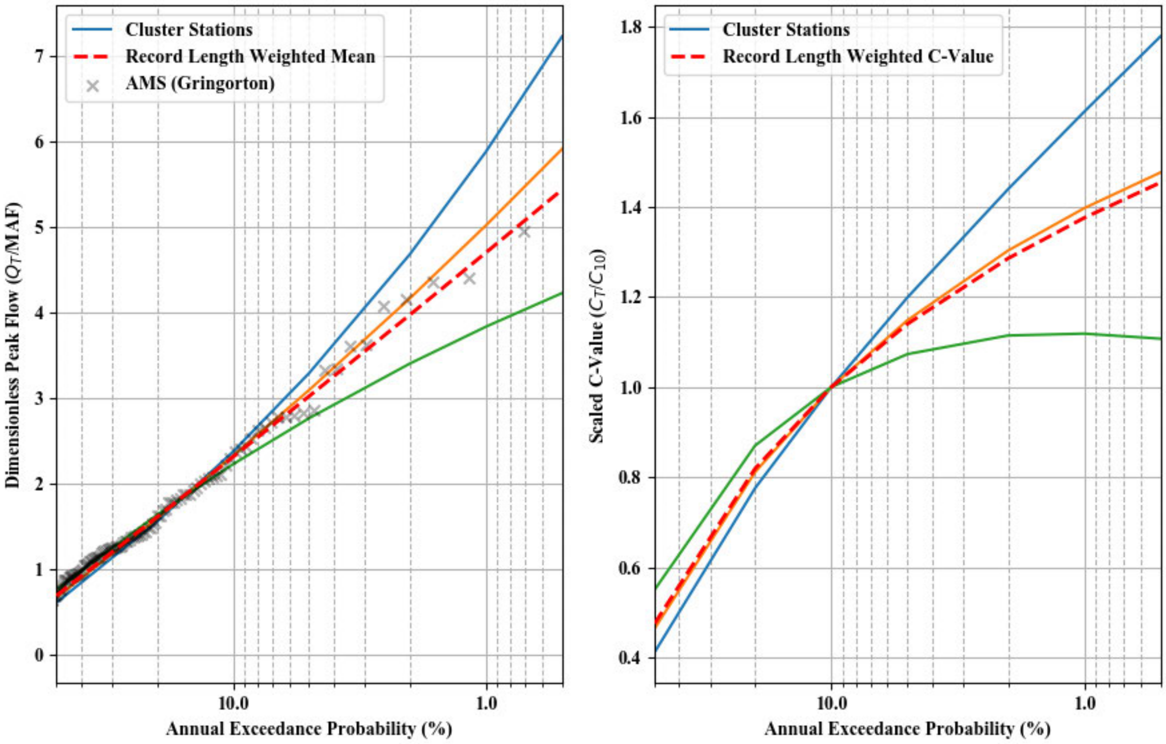
Dimensionless Peak Flow and Scaled C-value - Cluster 16



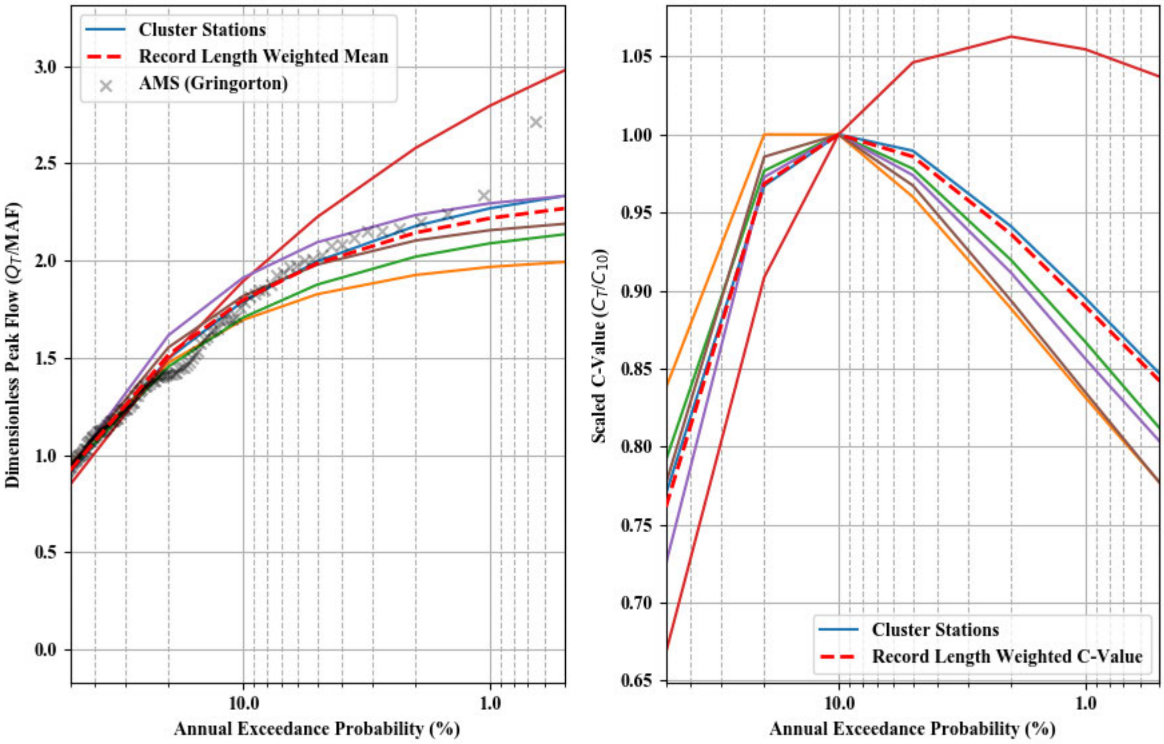

Dimensionless Peak Flow and Scaled C-value - Cluster 17
Cluster Stations
Record Length Weighted Mean
AMS (Gringorton)
Dimensionless Peak Flow (Q_T/MAF)
Annual Exceedance Probability (%)
Cluster Stations
Record Length Weighted C-Value
Scaled C-Value (C_T/C_{10})
Annual Exceedance Probability (%)

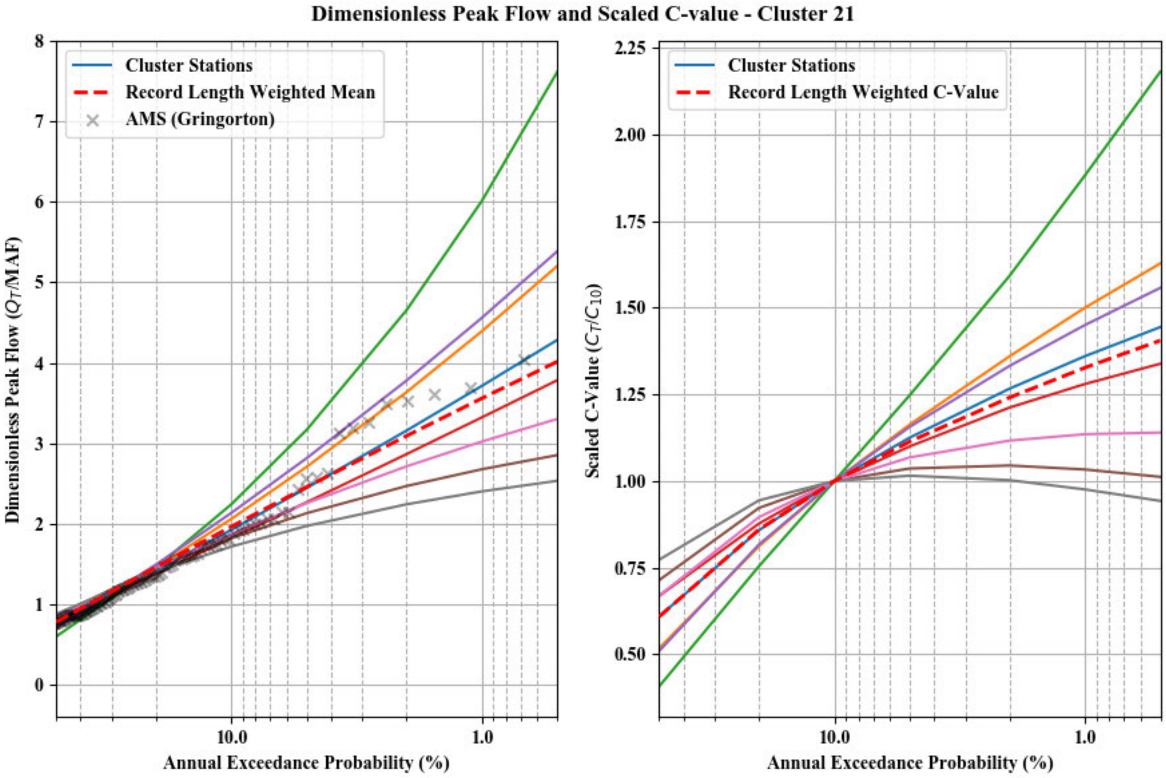

Dimensionless Peak Flow and Scaled C-value - Cluster 18
Cluster Stations
Record Length Weighted Mean
AMS (Gringorton)
Dimensionless Peak Flow (Q_T/MAF)
Annual Exceedance Probability (%)
Cluster Stations
Record Length Weighted C-Value
Scaled C-Value (C_T/C_{10})
Annual Exceedance Probability (%)

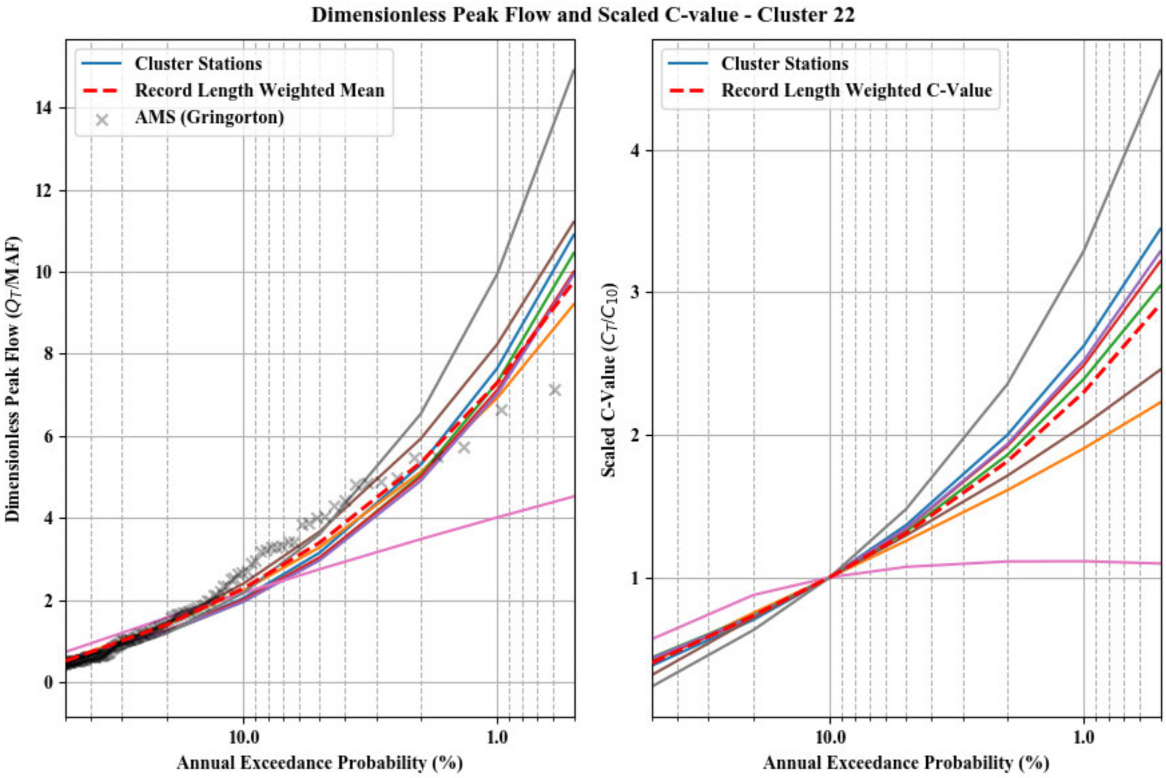
Dimensionless Peak Flow and Scaled C-value - Cluster 19



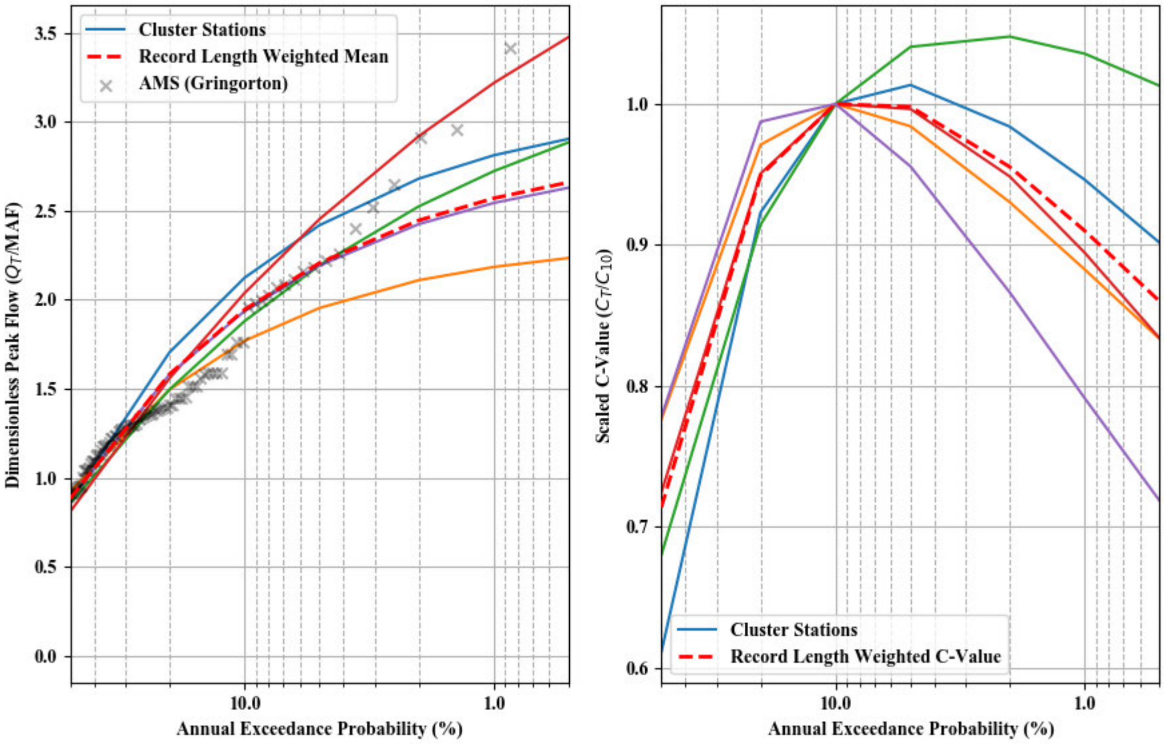
Dimensionless Peak Flow and Scaled C-value - Cluster 20



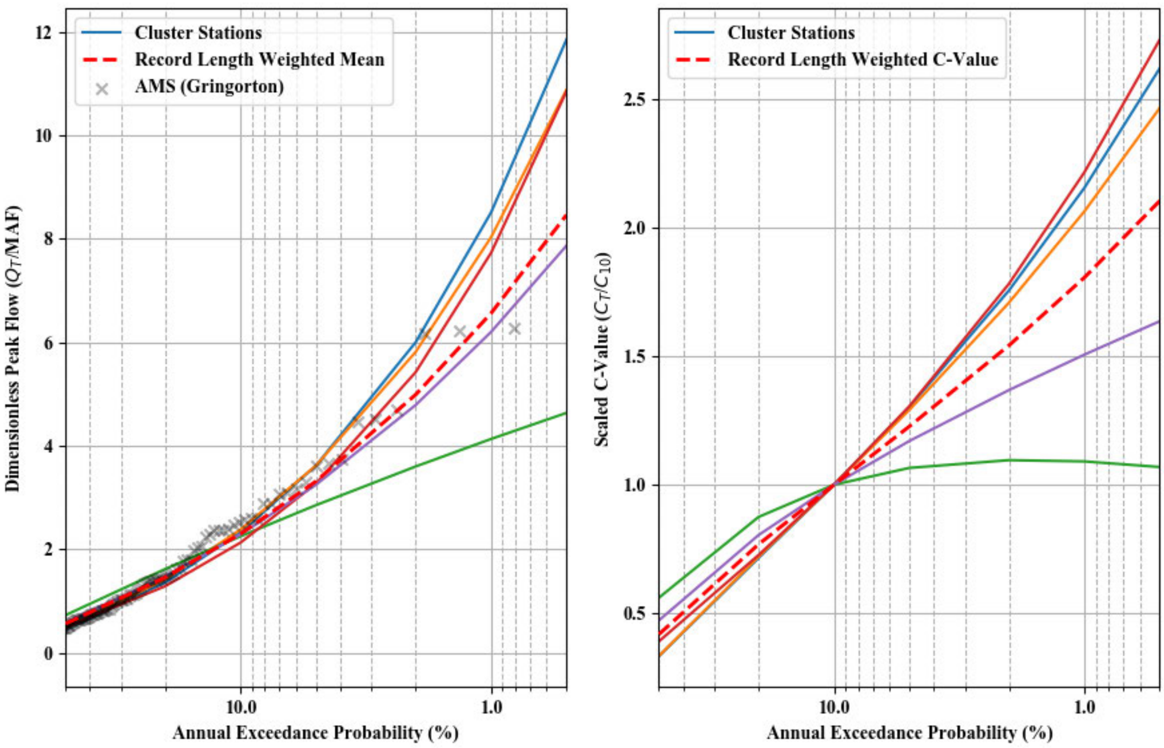

Dimensionless Peak Flow and Scaled C-value - Cluster 21
Cluster Stations
Record Length Weighted Mean
AMS (Gringorton)
Dimensionless Peak Flow (Q_T/MAF)
Annual Exceedance Probability (%)
Cluster Stations
Record Length Weighted C-Value
Scaled C-Value (C_T/C_{10})
Annual Exceedance Probability (%)


Dimensionless Peak Flow and Scaled C-value - Cluster 22
Cluster Stations
Record Length Weighted Mean
AMS (Gringorton)
Dimensionless Peak Flow (Q_T/MAF)
Annual Exceedance Probability (%)
Cluster Stations
Record Length Weighted C-Value
Scaled C-Value (C_T/C_{10})
Annual Exceedance Probability (%)

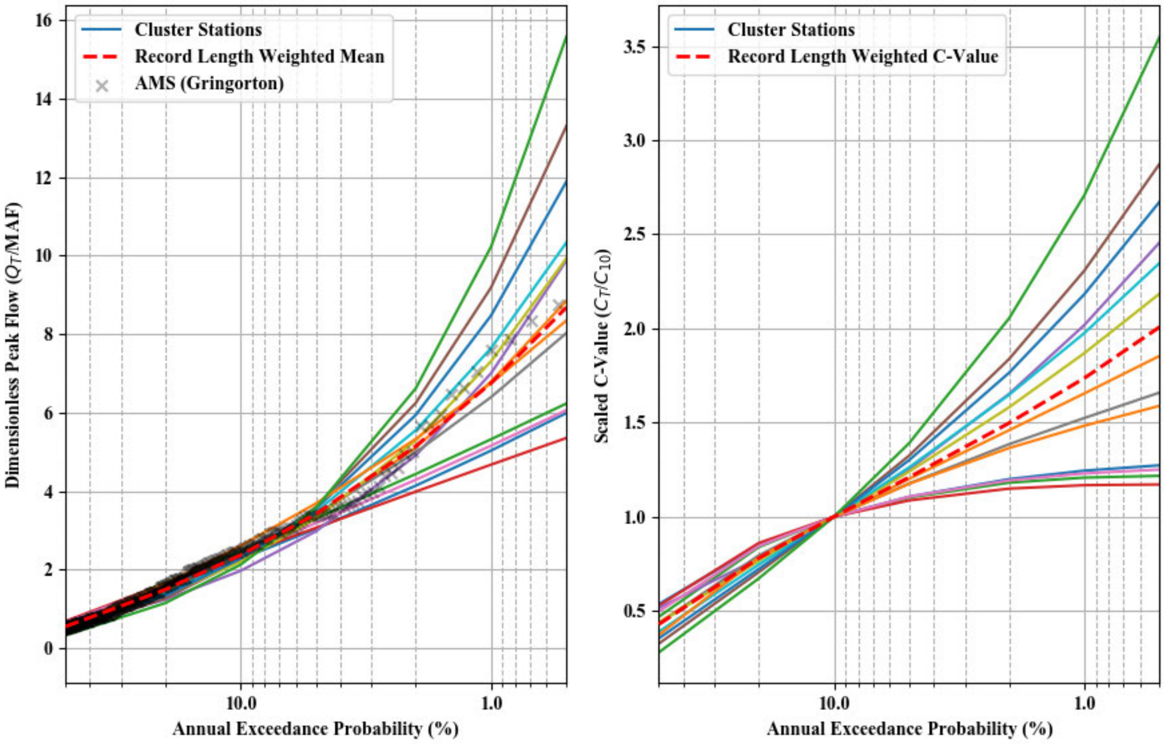
Dimensionless Peak Flow and Scaled C-value - Cluster 23



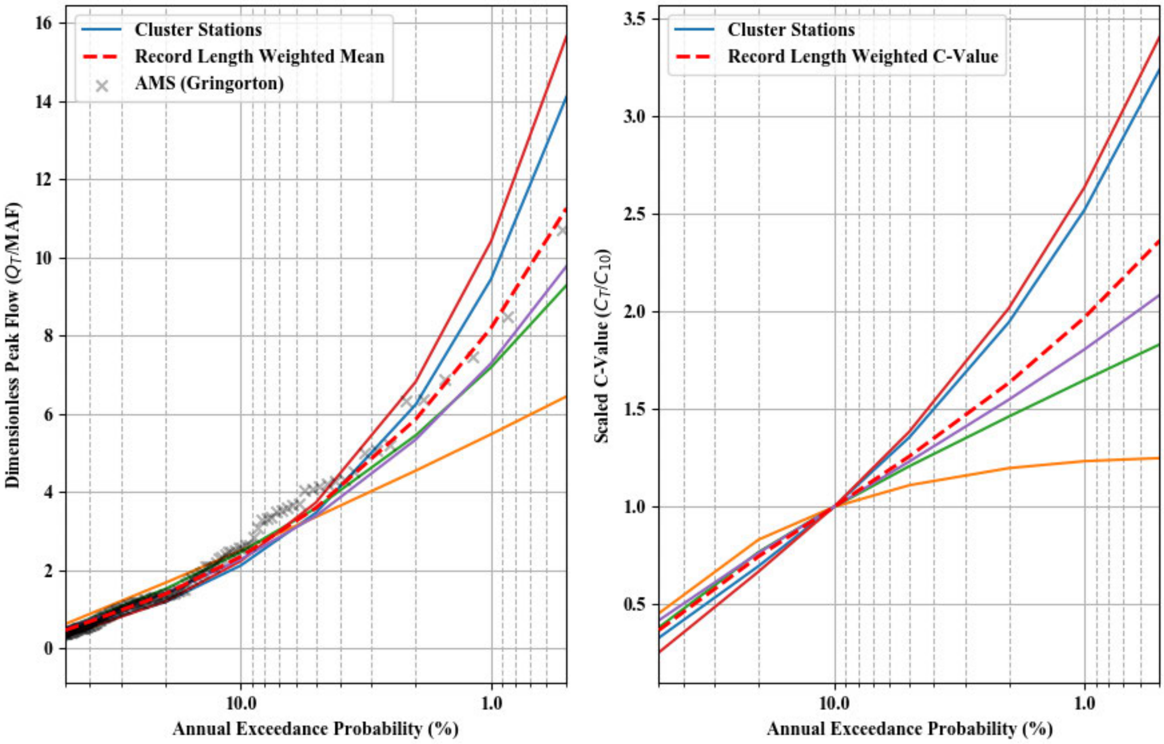
Dimensionless Peak Flow and Scaled C-value - Cluster 24



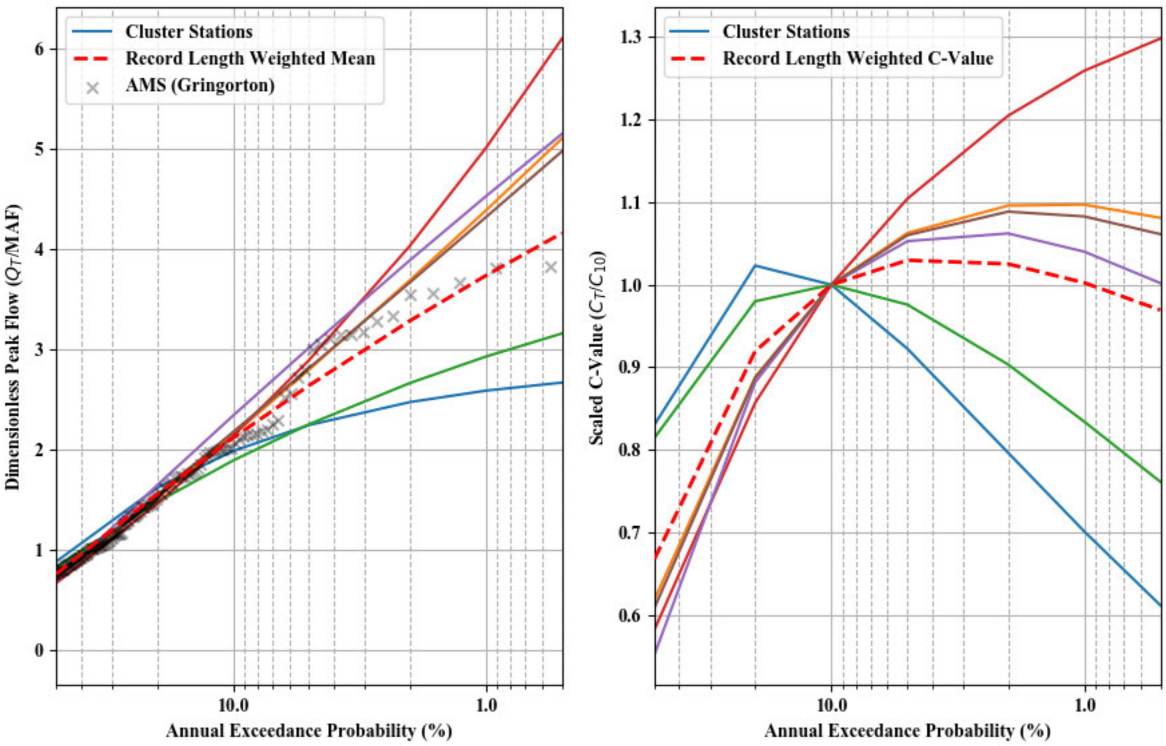
Dimensionless Peak Flow and Scaled C-value - Cluster 25



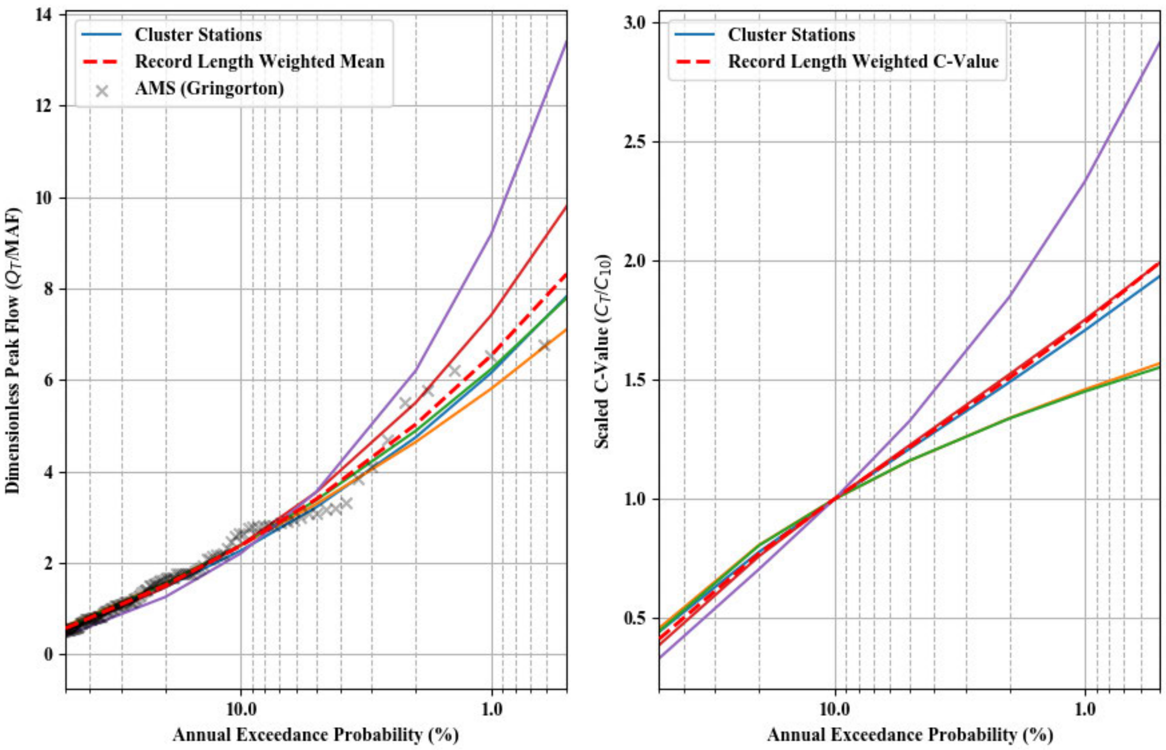
Dimensionless Peak Flow and Scaled C-value - Cluster 26



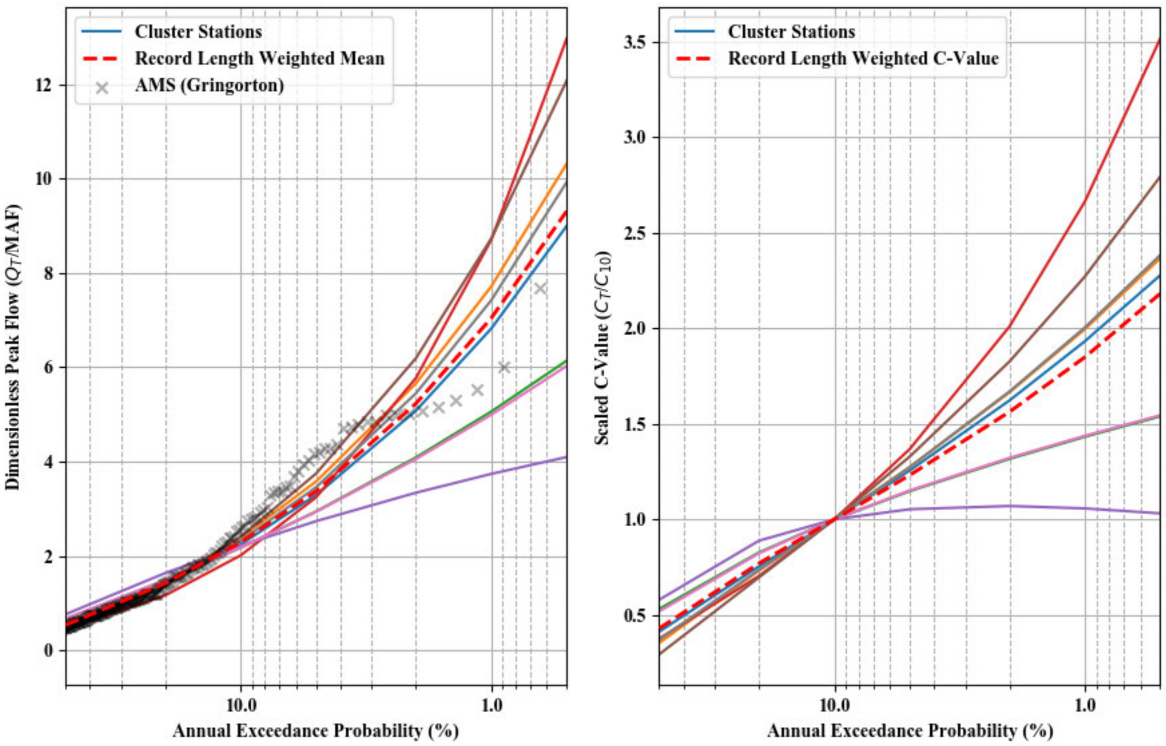
Dimensionless Peak Flow and Scaled C-value - Cluster 27



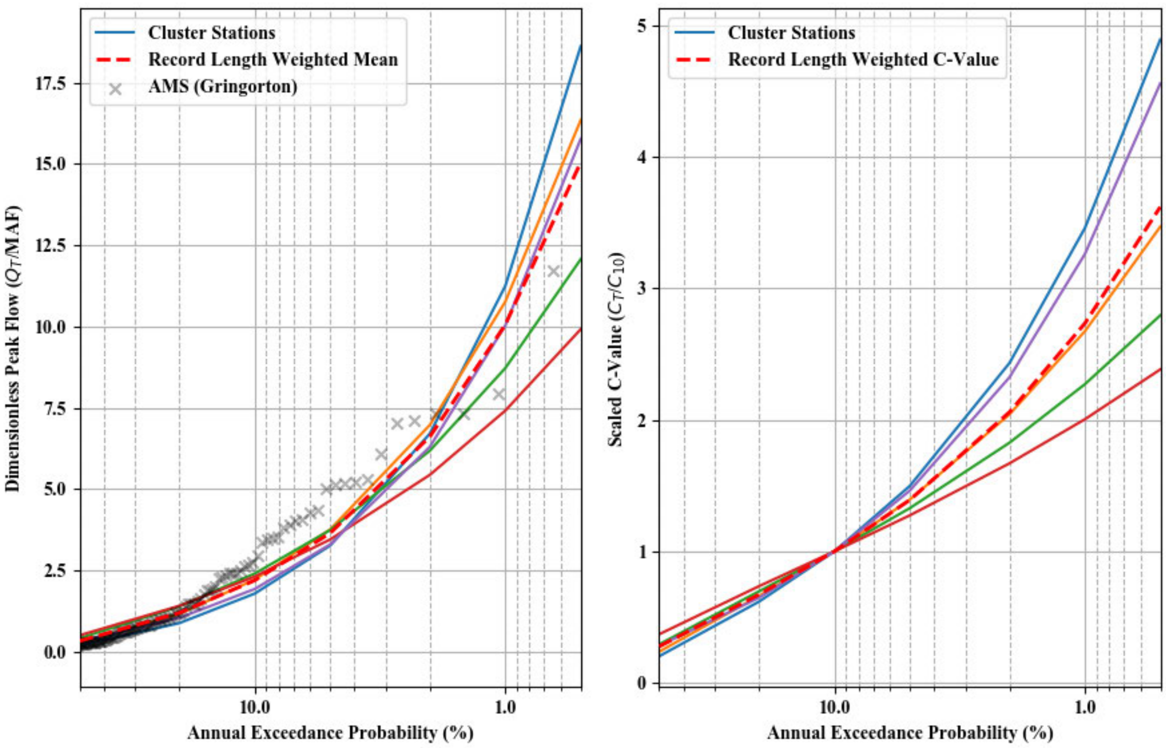
Dimensionless Peak Flow and Scaled C-value - Cluster 28

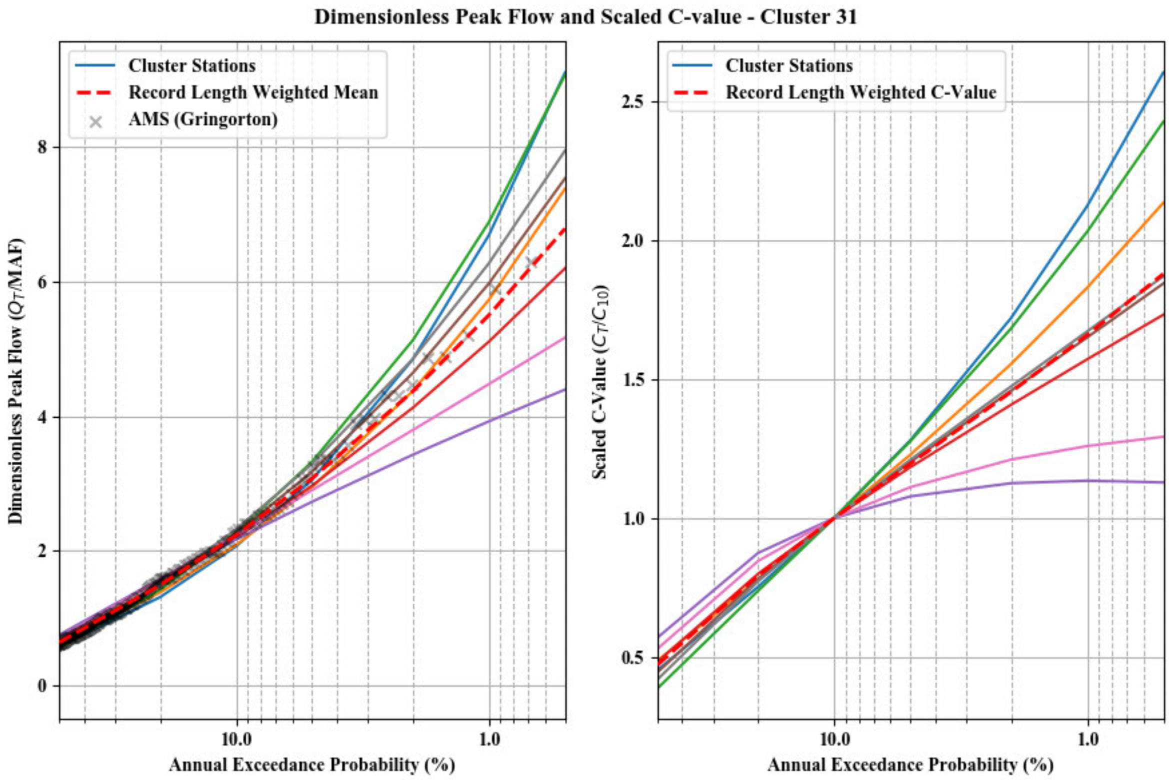


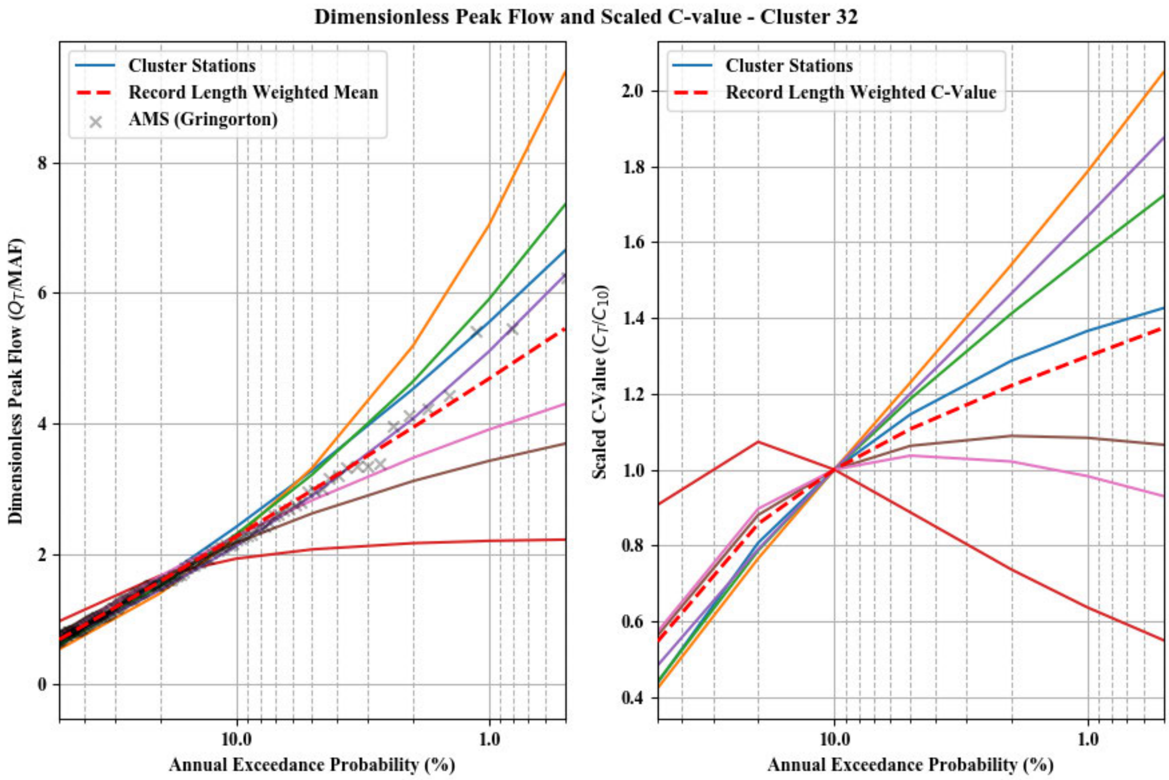
Dimensionless Peak Flow and Scaled C-value - Cluster 29



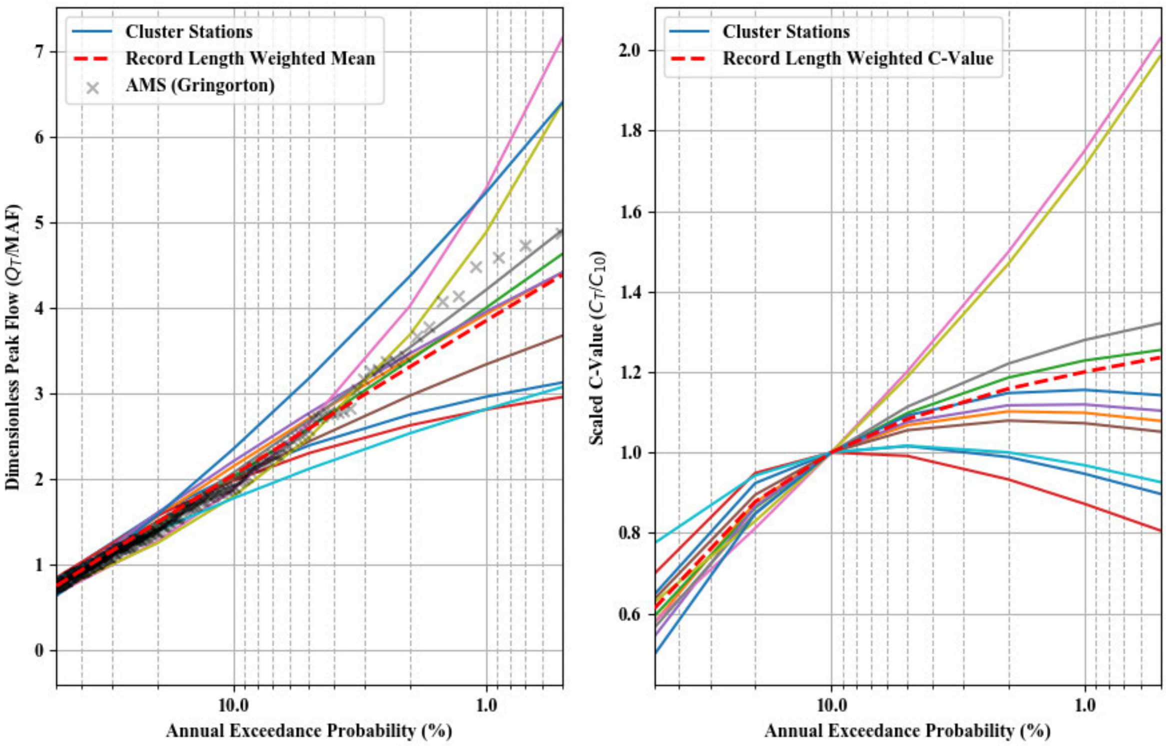
Dimensionless Peak Flow and Scaled C-value - Cluster 30



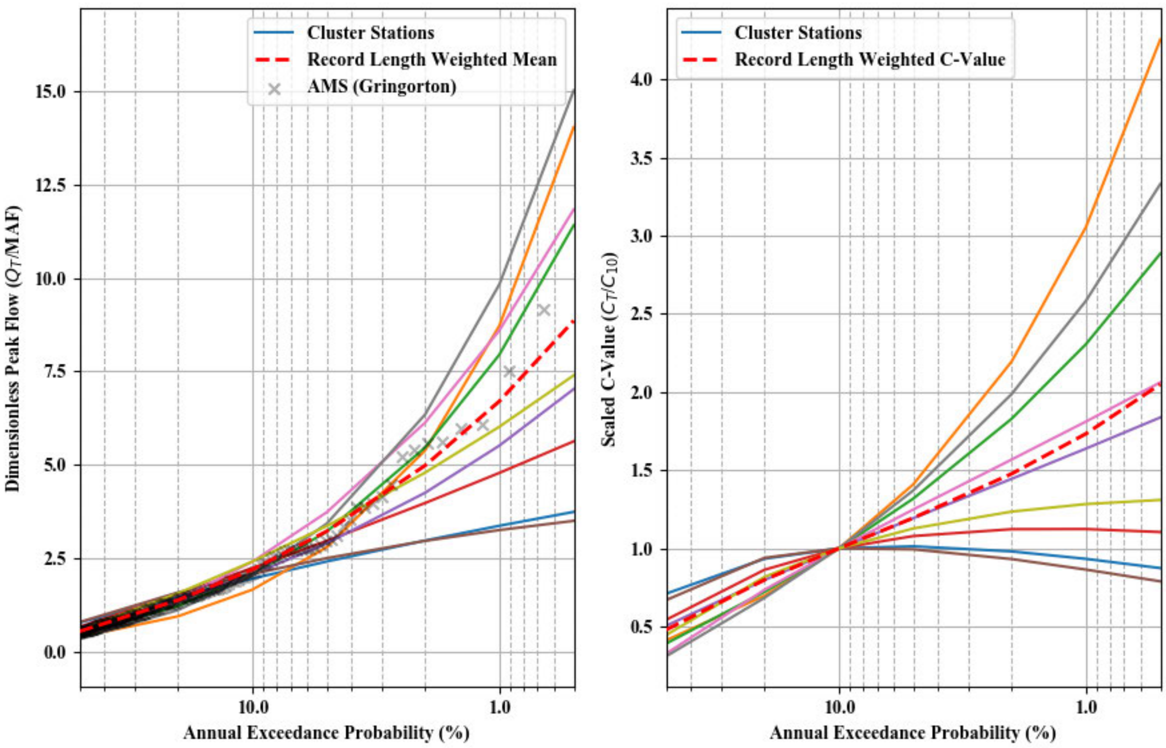

Dimensionless Peak Flow and Scaled C-value - Cluster 31
Cluster Stations
Record Length Weighted Mean
AMS (Gringorton)
Dimensionless Peak Flow (Q_T/MAF)
Annual Exceedance Probability (%)
Cluster Stations
Record Length Weighted C-Value
Scaled C-Value (C_T/C_{10})
Annual Exceedance Probability (%)

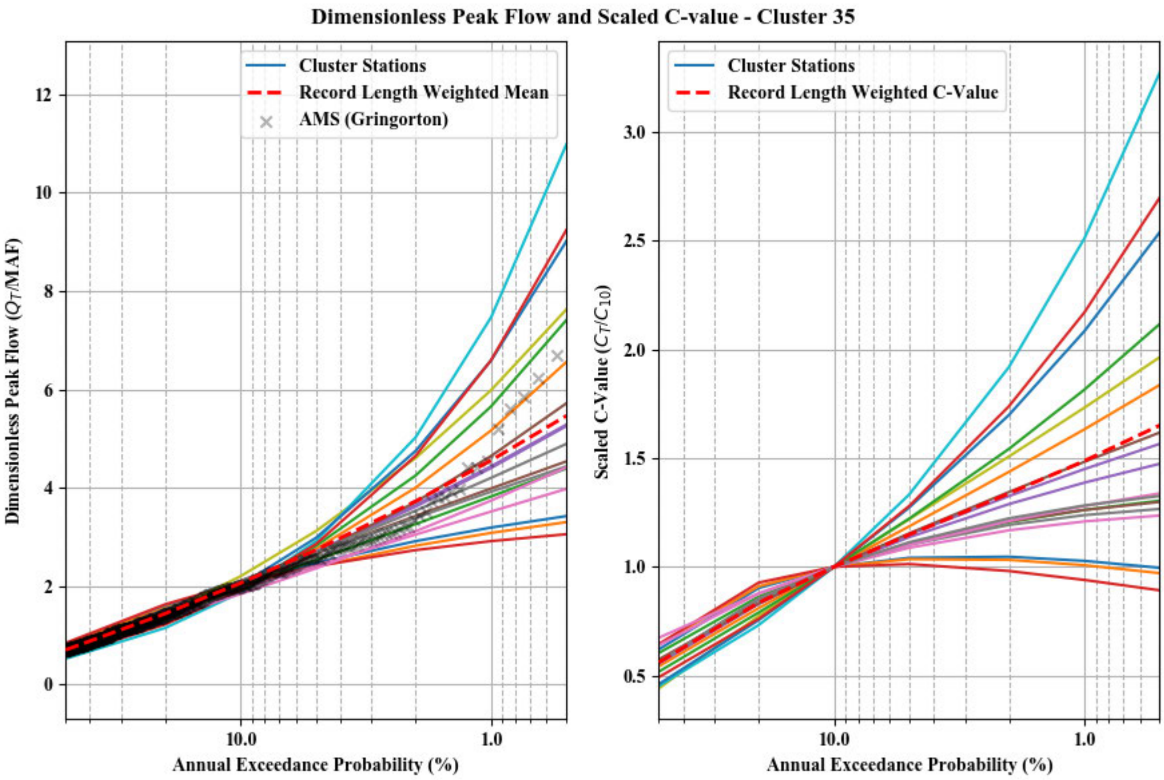

Dimensionless Peak Flow and Scaled C-value - Cluster 32
Cluster Stations
Record Length Weighted Mean
AMS (Gringorton)
Dimensionless Peak Flow (Q_T/MAF)
Annual Exceedance Probability (%)
Cluster Stations
Record Length Weighted C-Value
Scaled C-Value (C_T/C_{10})
Annual Exceedance Probability (%)

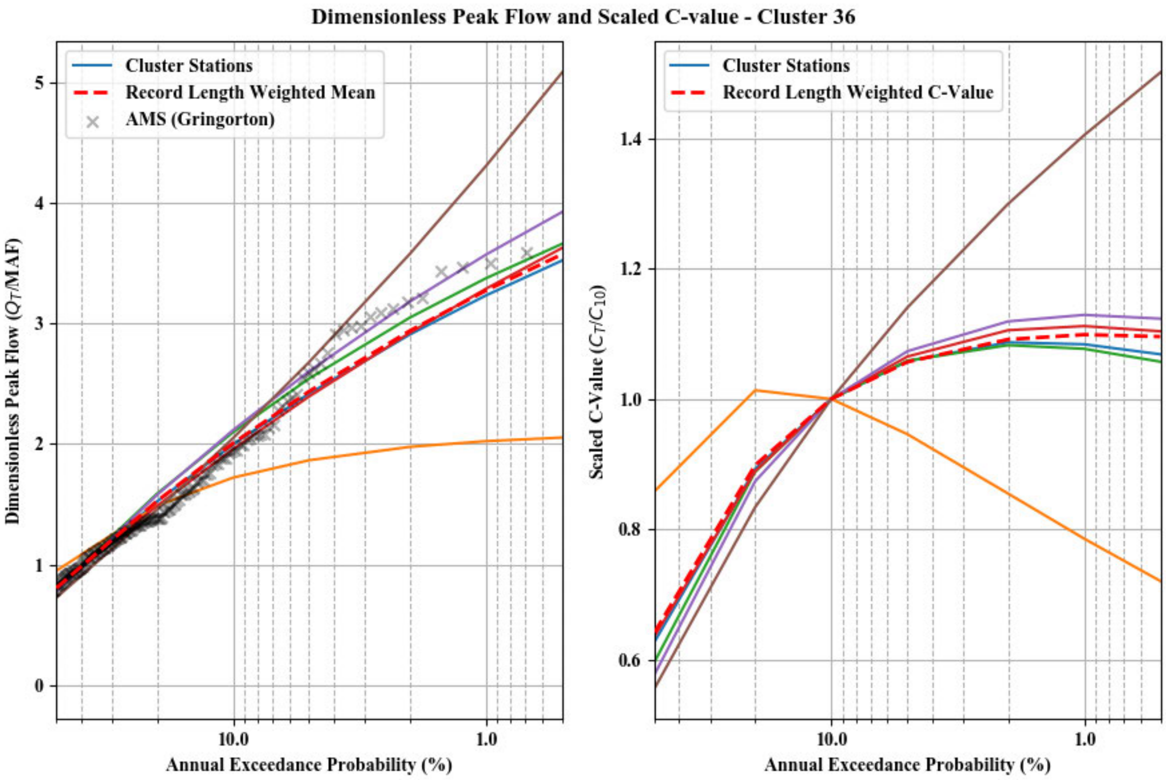
Dimensionless Peak Flow and Scaled C-value - Cluster 33



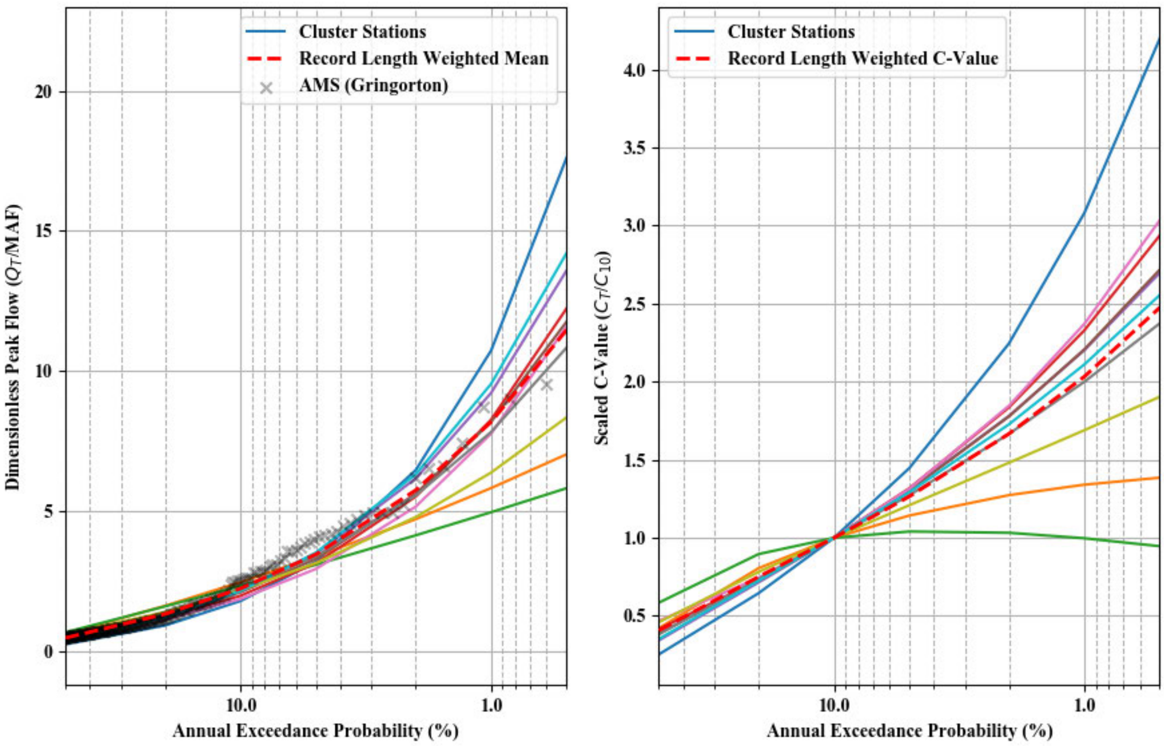
Dimensionless Peak Flow and Scaled C-value - Cluster 34



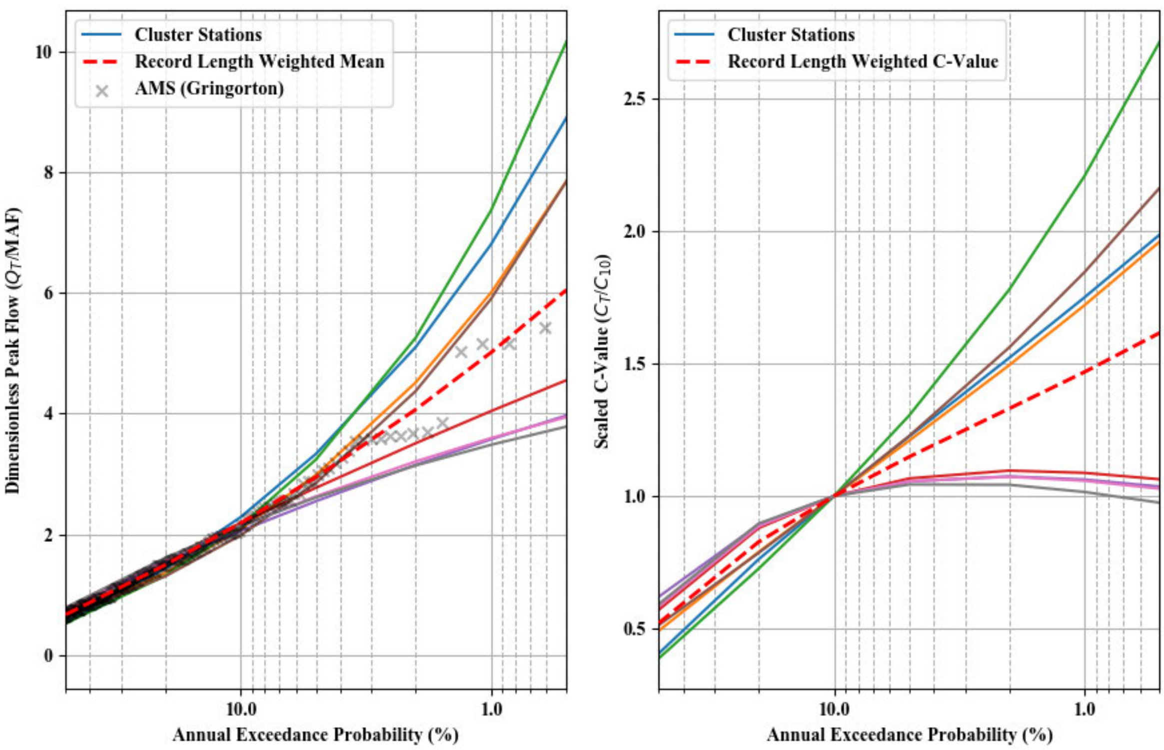

Dimensionless Peak Flow and Scaled C-value - Cluster 35
Cluster Stations
Record Length Weighted Mean
AMS (Gringorton)
Dimensionless Peak Flow (Q_T/MAF)
Annual Exceedance Probability (%)
Cluster Stations
Record Length Weighted C-Value
Scaled C-Value (C_T/C_{10})
Annual Exceedance Probability (%)


Dimensionless Peak Flow and Scaled C-value - Cluster 36
Cluster Stations
Record Length Weighted Mean
AMS (Gringorton)
Dimensionless Peak Flow (Q_T/MAF)
Annual Exceedance Probability (%)
Cluster Stations
Record Length Weighted C-Value
Scaled C-Value (C_T/C_{10})
Annual Exceedance Probability (%)

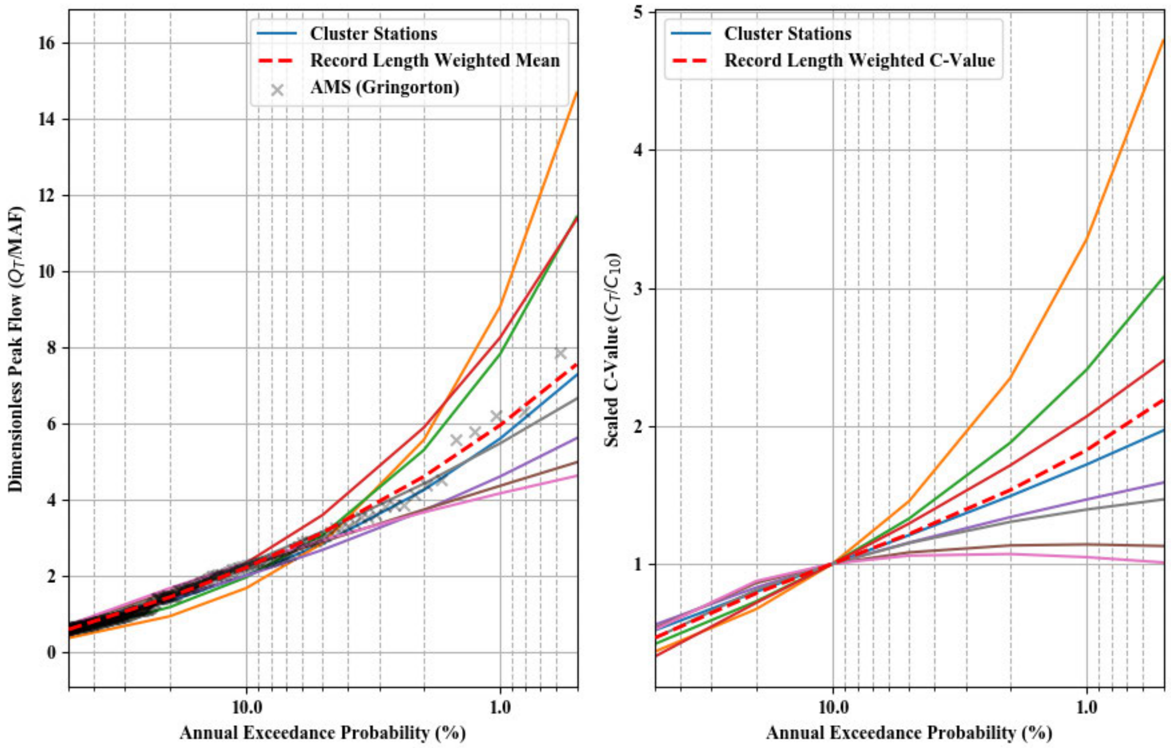
Dimensionless Peak Flow and Scaled C-value - Cluster 37



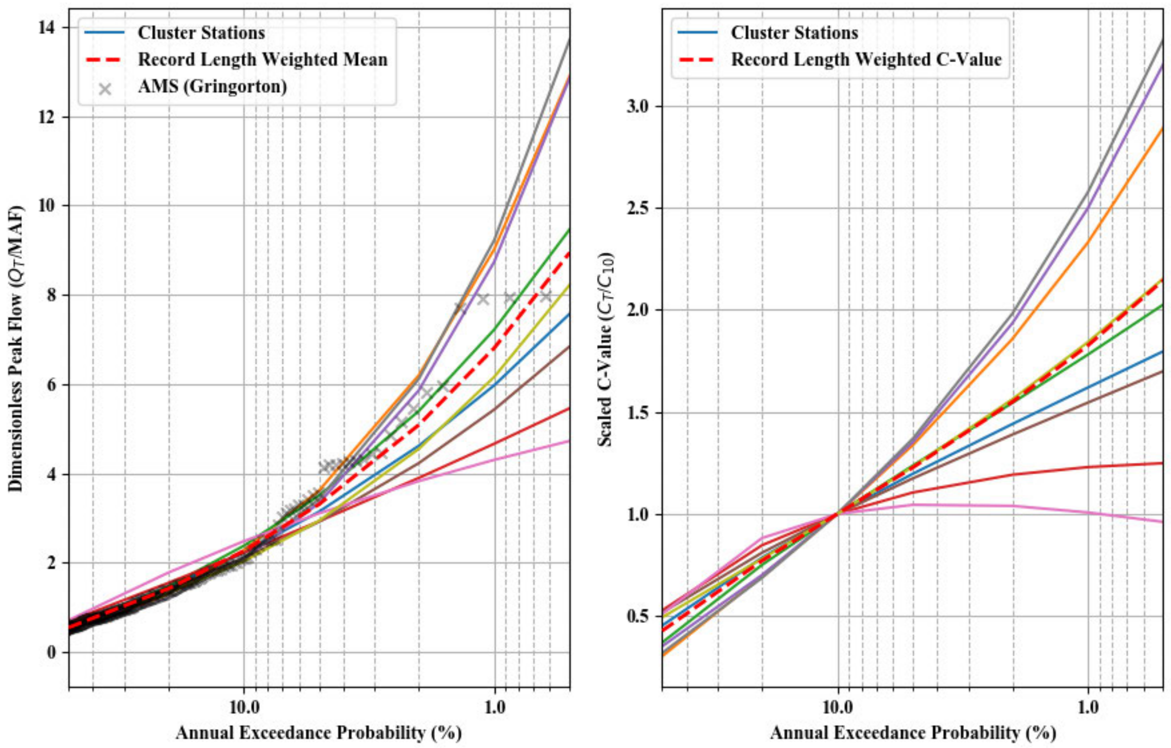
Dimensionless Peak Flow and Scaled C-value - Cluster 38



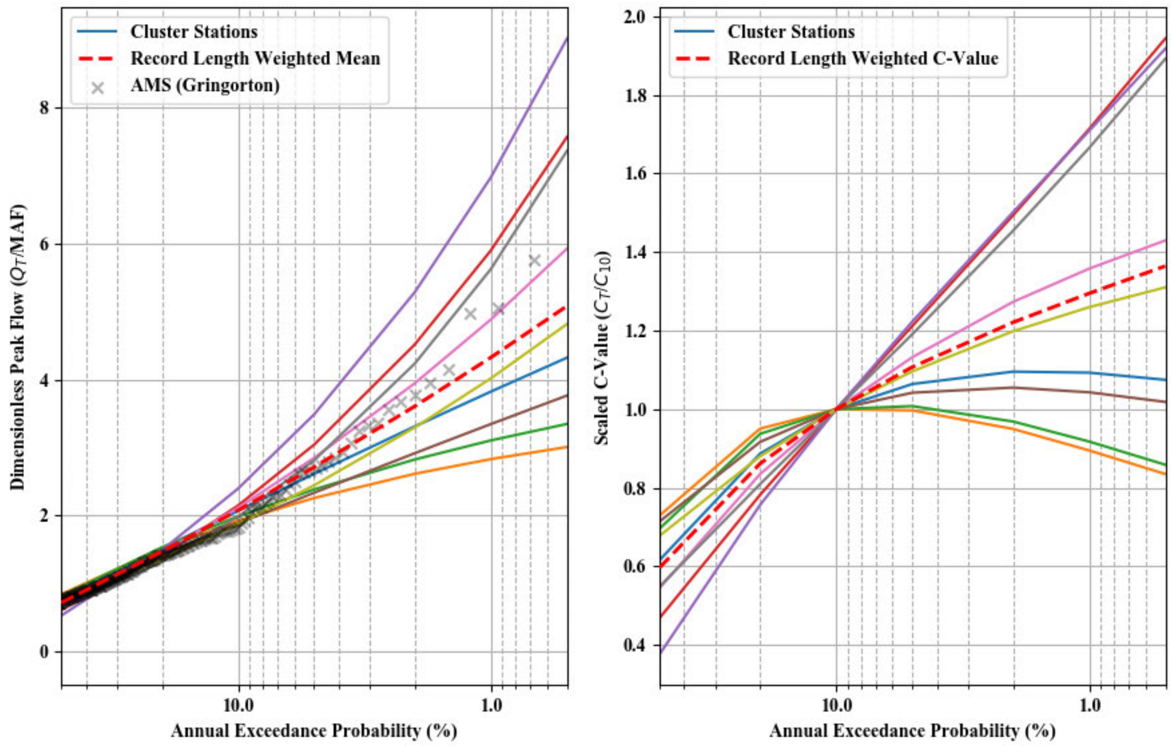
Dimensionless Peak Flow and Scaled C-value - Cluster 39



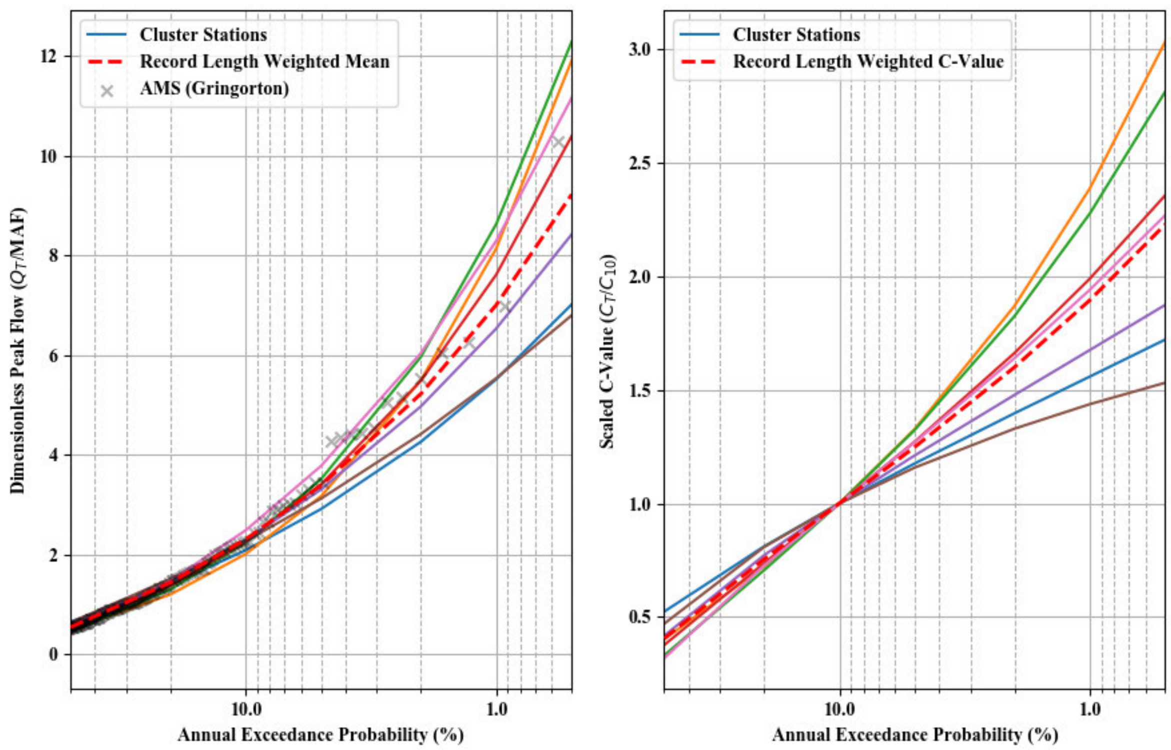
Dimensionless Peak Flow and Scaled C-value - Cluster 40



Dimensionless Peak Flow and Scaled C-value - Cluster 41



Dimensionless Peak Flow and Scaled C-value - Cluster 42



APPENDIX C: DIMENSIONLESS IF1 GROWTH CURVES FOR EACH OF THE 42 RELATIVELY HOMOGENEOUS CLUSTERS

Cluster	RIF Growth Factor (GF_T) per AEP (%)						
	50	**20**	**10**	**5**	**2**	**1**	**0.5**
1	0.85	1.54	1.95	2.28	2.63	2.83	2.99
2	0.53	1.41	2.28	3.38	5.30	7.20	9.61
3	0.44	1.27	2.19	3.45	5.89	8.57	12.28
4	0.45	1.40	2.38	3.66	5.96	8.32	11.40
5	0.50	1.40	2.31	3.48	5.56	7.65	10.33
6	0.59	1.52	2.35	3.30	4.76	6.07	7.56
7	0.83	1.67	2.15	2.52	2.89	3.10	3.27
8	0.56	1.44	2.29	3.32	5.04	6.69	8.69
9	0.70	1.55	2.21	2.89	3.81	4.54	5.29
10	0.64	1.53	2.28	3.11	4.33	5.37	6.51
11	0.44	1.32	2.27	3.56	5.98	8.57	12.09
12	0.56	1.55	2.44	3.47	5.11	6.58	8.29
13	0.63	1.52	2.28	3.14	4.42	5.53	6.77
14	0.59	1.59	2.44	3.39	4.81	6.01	7.35
15	0.64	1.51	2.25	3.08	4.33	5.40	6.59
16	0.62	1.49	2.25	3.13	4.50	5.72	7.11
17	0.55	1.45	2.30	3.35	5.10	6.77	8.82
18	0.44	1.41	2.40	3.69	6.01	8.38	11.46
19	0.69	1.61	2.31	3.02	3.97	4.70	5.44
20	0.92	1.52	1.80	1.99	2.14	2.22	2.27
21	0.78	1.46	1.97	2.46	3.09	3.56	4.02
22	0.52	1.40	2.29	3.40	5.35	7.30	9.76
23	0.88	1.58	1.94	2.21	2.45	2.57	2.66
24	0.57	1.46	2.30	3.32	4.99	6.56	8.46
25	0.55	1.49	2.37	3.42	5.14	6.75	8.68
26	0.46	1.39	2.34	3.59	5.86	8.20	11.26
27	0.76	1.56	2.12	2.64	3.29	3.74	4.16
28	0.56	1.50	2.37	3.39	5.03	6.54	8.32
29	0.54	1.42	2.29	3.37	5.23	7.04	9.31
30	0.34	1.20	2.20	3.65	6.62	10.06	15.05
31	0.64	1.49	2.23	3.07	4.37	5.50	6.79
32	0.69	1.58	2.27	2.97	3.94	4.68	5.45
33	0.76	1.50	2.05	2.60	3.32	3.86	4.39
34	0.56	1.40	2.21	3.23	4.99	6.71	8.86
35	0.71	1.45	2.06	2.74	3.73	4.56	5.47
36	0.80	1.52	2.01	2.44	2.94	3.27	3.57
37	0.47	1.32	2.23	3.45	5.74	8.17	11.45
38	0.67	1.49	2.19	2.95	4.07	5.01	6.05
39	0.61	1.45	2.23	3.14	4.61	5.96	7.56
40	0.55	1.42	2.27	3.31	5.09	6.81	8.94
41	0.72	1.47	2.08	2.71	3.61	4.32	5.08
42	0.54	1.43	2.30	3.38	5.22	7.00	9.22

APPENDIX D: REGIONAL L-MOMENTS FOR THE 42 RELATIVELY HOMOGENEOUS CLUSTERS FOR USE IN THE IF2 APPROACH

Regional LMs per relatively homogeneous cluster						
Cluster	**λ_2**	**τ_3**		**Cluster**	**λ_2**	**τ_3**
1	0.359	0.200		22	0.588	0.532
2	0.590	0.513		23	0.360	0.167
3	0.642	0.580		24	0.571	0.473
4	0.670	0.534		25	0.593	0.471
5	0.619	0.513		26	0.651	0.534
6	0.562	0.436		27	0.423	0.296
7	0.430	0.172		28	0.586	0.455
8	0.568	0.475		29	0.583	0.495
9	0.471	0.353		30	0.738	0.635
10	0.514	0.410		31	0.510	0.430
11	0.657	0.558		32	0.491	0.341
12	0.610	0.452		33	0.402	0.345
13	0.520	0.395		34	0.545	0.493
14	0.587	0.413		35	0.426	0.392
15	0.509	0.414		36	0.378	0.261
16	0.522	0.438		37	0.628	0.570
17	0.573	0.473		38	0.476	0.394
18	0.678	0.523		39	0.518	0.449
19	0.506	0.342		40	0.567	0.494
20	0.310	0.120		41	0.424	0.380
21	0.372	0.329		42	0.589	0.511

APPENDIX E: DIMENSIONLESS *C* VALUE GROWTH CURVES FOR EACH OF THE 42 RELATIVELY HOMOGENEOUS CLUSTERS

Cluster	C_T Growth Factor (GF_T) per AEP (%)						
	50	**20**	**10**	**5**	**2**	**1**	**0.5**
1	0.75	0.96	1.00	1.00	0.97	0.93	0.89
2	0.39	0.74	1.00	1.27	1.68	2.03	2.45
3	0.37	0.71	1.00	1.34	1.93	2.54	3.36
4	0.31	0.70	1.00	1.33	1.85	2.34	2.97
5	0.39	0.74	1.00	1.28	1.72	2.13	2.65
6	0.40	0.77	1.00	1.22	1.50	1.72	1.94
7	0.72	0.97	1.00	0.98	0.93	0.89	0.84
8	0.44	0.78	1.00	1.20	1.48	1.70	1.95
9	0.51	0.83	1.00	1.14	1.28	1.38	1.47
10	0.46	0.80	1.00	1.17	1.37	1.51	1.65
11	0.38	0.73	1.00	1.31	1.83	2.35	3.05
12	0.40	0.77	1.00	1.21	1.50	1.72	1.97
13	0.46	0.80	1.00	1.18	1.40	1.58	1.77
14	0.39	0.77	1.00	1.21	1.47	1.66	1.87
15	0.49	0.81	1.00	1.17	1.37	1.52	1.67
16	0.44	0.77	1.00	1.23	1.54	1.81	2.12
17	0.41	0.76	1.00	1.24	1.60	1.92	2.30
18	0.32	0.71	1.00	1.32	1.83	2.33	2.97
19	0.47	0.82	1.00	1.14	1.29	1.38	1.45
20	0.76	0.97	1.00	0.99	0.94	0.89	0.84
21	0.61	0.86	1.00	1.11	1.24	1.33	1.41
22	0.40	0.73	1.00	1.31	1.81	2.30	2.92
23	0.71	0.95	1.00	1.00	0.96	0.91	0.86
24	0.42	0.77	1.00	1.23	1.54	1.81	2.11
25	0.43	0.78	1.00	1.21	1.50	1.74	2.01
26	0.37	0.75	1.00	1.26	1.64	1.97	2.36
27	0.67	0.92	1.00	1.03	1.03	1.00	0.97
28	0.41	0.77	1.00	1.22	1.51	1.74	1.99
29	0.43	0.77	1.00	1.23	1.56	1.85	2.18
30	0.28	0.67	1.00	1.39	2.06	2.73	3.62
31	0.47	0.79	1.00	1.20	1.45	1.66	1.88
32	0.55	0.86	1.00	1.11	1.22	1.30	1.37
33	0.61	0.88	1.00	1.08	1.16	1.20	1.24
34	0.48	0.80	1.00	1.19	1.48	1.73	2.05
35	0.56	0.83	1.00	1.15	1.34	1.49	1.65
36	0.64	0.90	1.00	1.06	1.09	1.10	1.10
37	0.41	0.75	1.00	1.27	1.67	2.03	2.47
38	0.52	0.83	1.00	1.15	1.33	1.47	1.62
39	0.46	0.79	1.00	1.22	1.53	1.82	2.19
40	0.43	0.77	1.00	1.23	1.55	1.82	2.14
41	0.60	0.86	1.00	1.11	1.22	1.29	1.36
42	0.41	0.75	1.00	1.25	1.60	1.89	2.23

APPENDIX F: *MAF* AND C_{10} REGRESSION STATISTICS FOR THE 42 RELATIVELY HOMOGENEOUS CLUSTERS

Scale	No. Sites	*MAF*								C_{10}							
		AIC	BIC	R^2	RMSE	Descriptor Standard Error				AIC	BIC	R^2	RMSE	Descriptor Standard Error			
						Const	*A*	*MAP*	D_c					Const	*A*	*MAP*	D_c
National	332	848.99	864.78	0.79	214.86	0.78	0.02	0.11	0.04	745.82	761.05	0.13	1.18	0.87	0.02	0.12	0.04
Cluster 1	32	55.01	60.87	0.76	58.36	13.59	0.07	1.96	0.75	49.54	55.41	0.33	0.07	12.48	0.07	1.80	0.69
Cluster 2 - 3	30	51.37	56.98	0.68	53.94	19.26	0.07	2.58	1.62	43.84	49.45	0.45	0.06	16.98	0.06	2.28	1.43
Cluster 4	43	79.85	86.90	0.83	56.43	8.83	0.05	1.31	0.65	80.64	87.68	0.34	0.07	8.91	0.05	1.32	0.65
Cluster 5	35	81.60	87.82	0.64	98.18	4.88	0.11	0.62	0.54	74.45	80.67	0.13	0.10	4.40	0.10	0.56	0.49
Cluster 6	35	51.11	57.33	0.85	66.67	11.16	0.05	1.62	0.65	53.77	59.99	0.11	0.09	11.59	0.06	1.68	0.67
Cluster 7	35	52.45	58.68	0.86	52.02	5.17	0.05	0.69	0.61	49.71	55.93	0.15	0.06	4.97	0.05	0.67	0.59
Cluster 8	37	47.86	54.31	0.85	51.68	9.45	0.05	1.38	0.36	45.87	52.32	0.29	0.06	9.20	0.05	1.35	0.35
Cluster 9	31	62.27	68.00	0.83	38.82	6.76	0.08	0.90	0.76	61.83	67.56	0.40	0.06	6.71	0.08	0.89	0.75
Cluster 10	33	69.22	75.21	0.79	38.32	4.44	0.08	0.58	0.83	71.06	77.05	0.30	0.07	4.57	0.09	0.60	0.86
Cluster 11	35	81.60	87.82	0.64	98.18	4.88	0.11	0.62	0.54	74.45	80.67	0.13	0.10	4.40	0.10	0.56	0.49
Cluster 12	30	62.38	67.98	0.73	54.71	4.31	0.11	0.56	0.50	58.67	64.28	0.16	0.09	4.05	0.10	0.52	0.47
Cluster 13	30	44.70	50.31	0.59	71.64	9.69	0.09	1.36	0.64	41.94	47.54	0.05	0.09	9.25	0.08	1.30	0.62
Cluster 14	32	45.70	51.56	0.90	107.76	7.31	0.04	0.99	0.84	42.51	48.38	0.21	0.05	6.96	0.04	0.94	0.80
Cluster 15	47	76.79	84.19	0.75	129.51	3.64	0.06	0.52	0.15	77.63	85.03	0.09	0.14	3.68	0.06	0.52	0.15
Cluster 16	42	65.63	72.58	0.69	101.45	4.57	0.07	0.64	0.28	59.05	66.00	0.05	0.09	4.22	0.06	0.59	0.26
Cluster 17	31	49.09	54.82	0.74	98.34	2.90	0.08	0.40	0.11	99.96	105.70	0.03	3.85	6.59	0.18	0.91	0.25
Cluster 18	33	77.31	83.30	0.80	190.12	3.20	0.10	0.45	0.23	79.59	85.57	0.28	0.12	3.31	0.10	0.46	0.23
Cluster 19	42	78.71	85.66	0.89	171.99	1.51	0.04	0.23	0.10	87.98	94.93	0.22	0.13	1.69	0.05	0.26	0.11
Cluster 20 - 24	32	69.96	75.82	0.88	92.84	2.74	0.08	0.39	0.18	72.46	78.33	0.45	0.10	2.85	0.08	0.40	0.18
Cluster 25	34	69.81	75.91	0.85	150.94	3.01	0.08	0.43	0.14	75.86	81.97	0.37	0.11	3.30	0.09	0.47	0.15
Cluster 26 - 27	34	76.23	82.34	0.82	167.73	3.54	0.09	0.53	0.19	82.17	88.27	0.29	0.16	3.86	0.10	0.58	0.21
Cluster 28	34	77.83	84.78	0.84	104.65	2.57	0.06	0.39	0.14	79.48	86.43	0.30	0.14	2.62	0.06	0.40	0.14
Cluster 29	42	39.72	45.71	0.71	43.14	2.18	0.07	0.30	0.12	98.57	104.56	0.15	3.73	5.31	0.17	0.73	0.29
Cluster 30 - 32	33	26.06	31.80	0.84	53.94	2.04	0.06	0.28	0.10	95.90	101.64	0.02	3.85	6.29	0.19	0.86	0.31
Cluster 33	31	81.97	89.29	0.74	133.99	3.89	0.06	0.56	0.09	85.32	92.64	0.13	0.15	4.03	0.06	0.58	0.10

Scale	No. Sites	*MAF*								C_{10}							
		AIC	BIC	R^2	RMSE	Descriptor Standard Error				AIC	BIC	R^2	RMSE	Descriptor Standard Error			
						Const	*A*	*MAP*	D_c					Const	*A*	*MAP*	D_c
Cluster 34	46	76.72	83.17	0.80	132.96	7.05	0.07	1.00	0.12	70.64	77.08	0.21	0.11	6.49	0.06	0.92	0.11
Cluster 35	37	59.16	65.15	0.70	120.60	6.89	0.08	0.99	0.11	55.27	61.26	0.13	0.09	6.49	0.07	0.94	0.10
Cluster 36	33	55.07	60.93	0.70	115.65	5.13	0.08	0.72	0.32	49.63	55.49	0.10	0.10	4.71	0.07	0.66	0.29
Cluster 37	32	86.09	93.14	0.79	141.62	6.49	0.07	0.92	0.11	80.36	87.40	0.16	0.11	6.07	0.06	0.86	0.10
Cluster 38 - 39	43	73.38	79.48	0.82	31.38	8.15	0.08	1.06	1.27	71.90	78.01	0.35	0.06	7.97	0.08	1.03	1.25
Cluster 40	34	65.00	70.98	0.85	89.49	4.32	0.09	0.59	0.54	65.93	71.91	0.32	0.06	4.38	0.09	0.60	0.55
Cluster 41	33	68.25	74.11	0.79	64.71	5.14	0.09	0.71	1.02	71.08	76.94	0.32	0.08	5.37	0.10	0.75	1.07
Cluster 42	32	70.10	76.08	0.83	87.15	4.50	0.09	0.62	0.52	72.40	78.38	0.28	0.08	4.66	0.09	0.65	0.54

APPENDIX G: QRT MODEL COEFFICIENTS AND REGRESSION STATISTICS FOR THE 42 RELATIVELY HOMOGENEOUS CLUSTERS

Cluster	AEP (%)	Predictor Variables				Performance Metrics				Descriptor Standard Error			
		Constant	**Area**	$DR_{10\%}$	D_c	**AIC**	**BIC**	**R2**	***RMSE***	**Const**	**Area**	$DR_{10\%}$	D_c
National	50	-2.13	0.52	0.62	-0.25	818.47	833.69	0.73	112.44	0.51	0.03	0.14	0.05
	20	-1.60	0.52	0.70	-0.26	736.98	752.20	0.79	272.97	0.45	0.03	0.12	0.04
	10	-1.37	0.53	0.74	-0.26	703.23	718.45	0.81	442.56	0.43	0.02	0.12	0.04
	5	-1.20	0.53	0.77	-0.27	687.22	702.44	0.82	670.32	0.42	0.02	0.12	0.04
	2	-1.05	0.53	0.82	-0.28	695.31	710.53	0.82	1090.68	0.43	0.02	0.12	0.04
	1	-0.97	0.53	0.86	-0.28	722.50	737.72	0.81	1528.21	0.44	0.02	0.12	0.04
	0.5	-0.90	0.53	0.89	-0.29	764.14	779.36	0.79	2103.49	0.47	0.03	0.13	0.05
1	50	10.14	0.46	-1.36	-2.13	69.51	75.37	0.63	54.47	6.55	0.18	1.71	0.86
	20	10.00	0.54	-1.47	-1.50	58.91	64.78	0.73	97.92	5.55	0.15	1.45	0.73
	10	9.73	0.59	-1.53	-1.10	52.59	58.45	0.78	127.40	5.03	0.14	1.31	0.66
	5	9.38	0.63	-1.57	-0.72	48.42	54.28	0.82	162.20	4.71	0.13	1.23	0.62
	2	8.85	0.68	-1.62	-0.22	47.44	53.31	0.84	232.91	4.64	0.13	1.21	0.61
	1	8.40	0.72	-1.65	0.15	50.21	56.07	0.84	319.84	4.84	0.13	1.26	0.63
	0.5	7.93	0.76	-1.67	0.52	55.22	61.08	0.82	449.85	5.24	0.14	1.37	0.69
2 - 3	50	12.43	0.60	-1.37	-3.97	59.49	65.10	0.61	44.20	7.22	0.22	1.93	0.90
	20	11.95	0.64	-1.46	-3.09	50.48	56.09	0.69	81.28	6.21	0.19	1.66	0.78
	10	11.41	0.67	-1.49	-2.51	45.96	51.56	0.73	107.00	5.76	0.18	1.54	0.72
	5	10.76	0.69	-1.50	-1.95	44.21	49.81	0.76	137.32	5.60	0.17	1.49	0.70
	2	9.77	0.71	-1.49	-1.23	46.61	52.22	0.76	197.58	5.83	0.18	1.56	0.73
	1	8.97	0.73	-1.48	-0.69	51.30	56.90	0.75	274.49	6.30	0.19	1.68	0.79
	0.5	8.13	0.74	-1.46	-0.14	57.39	62.99	0.74	401.22	6.97	0.22	1.86	0.87

Cluster	AEP (%)	Predictor Variables				Performance Metrics				Descriptor Standard Error			
		Constant	**Area**	***$DR_{10\%}$***	***D_c***	**AIC**	**BIC**	**R2**	***RMSE***	**Const**	**Area**	***$DR_{10\%}$***	***D_c***
4	50	-0.57	0.49	0.55	-0.96	122.32	129.36	0.61	46.48	6.21	0.21	1.76	1.10
	20	1.10	0.55	0.16	-0.58	104.30	111.34	0.71	83.65	5.04	0.17	1.43	0.89
	10	1.91	0.59	-0.07	-0.32	93.79	100.84	0.76	111.14	4.46	0.15	1.26	0.79
	5	2.57	0.63	-0.28	-0.06	84.82	91.86	0.80	145.89	4.02	0.14	1.14	0.71
	2	3.31	0.68	-0.55	0.29	76.61	83.66	0.84	216.98	3.65	0.13	1.03	0.65
	1	3.82	0.71	-0.74	0.55	74.54	81.58	0.85	302.38	3.56	0.12	1.01	0.63
	0.5	4.29	0.74	-0.93	0.81	76.59	83.63	0.85	428.99	3.65	0.13	1.03	0.65
5	50	-5.04	0.38	1.53	-0.86	76.29	82.51	0.69	45.46	2.38	0.08	0.55	0.48
	20	-4.76	0.37	1.68	-0.77	72.34	78.56	0.72	124.18	2.25	0.08	0.52	0.45
	10	-4.58	0.36	1.73	-0.68	72.73	78.95	0.72	209.21	2.26	0.08	0.52	0.46
	5	-4.45	0.36	1.77	-0.58	75.13	81.35	0.71	330.63	2.34	0.08	0.54	0.47
	2	-4.35	0.36	1.81	-0.44	80.81	87.03	0.67	579.06	2.54	0.09	0.59	0.51
	1	-4.32	0.35	1.85	-0.33	86.43	92.65	0.64	868.95	2.75	0.09	0.64	0.55
	0.5	-4.33	0.35	1.88	-0.21	92.69	98.91	0.60	1293.24	3.01	0.10	0.70	0.61
6	50	-4.60	0.43	1.25	0.07	61.02	67.24	0.79	48.48	2.12	0.09	0.58	0.59
	20	-3.31	0.49	1.05	0.14	55.57	61.80	0.83	100.62	1.97	0.09	0.54	0.55
	10	-2.50	0.53	0.88	0.18	52.64	58.86	0.85	140.54	1.89	0.08	0.52	0.52
	5	-1.77	0.57	0.70	0.23	52.03	58.25	0.86	187.55	1.87	0.08	0.51	0.52
	2	-0.87	0.63	0.46	0.30	54.86	61.09	0.86	272.69	1.95	0.08	0.54	0.54
	1	-0.21	0.67	0.28	0.37	59.18	65.40	0.85	365.16	2.07	0.09	0.57	0.57
	0.5	0.42	0.71	0.10	0.45	64.67	70.89	0.83	492.76	2.24	0.10	0.62	0.62

Cluster	AEP (%)	Predictor Variables				Performance Metrics				Descriptor Standard Error			
		Constant	Area	$DR_{10\%}$	D_c	AIC	BIC	R2	*RMSE*	Const	Area	$DR_{10\%}$	D_c
7	50	-5.80	0.42	1.64	-0.41	56.53	62.76	0.84	33.68	1.69	0.08	0.42	0.48
	20	-4.79	0.44	1.59	-0.46	52.54	58.77	0.86	85.11	1.59	0.07	0.39	0.45
	10	-4.28	0.45	1.55	-0.51	51.91	58.13	0.86	142.05	1.58	0.07	0.39	0.45
	5	-3.85	0.47	1.50	-0.55	54.03	60.25	0.86	220.35	1.63	0.07	0.40	0.46
	2	-3.37	0.49	1.45	-0.59	60.16	66.38	0.84	366.26	1.78	0.08	0.44	0.50
	1	-3.06	0.50	1.40	-0.61	66.28	72.50	0.82	518.10	1.94	0.09	0.48	0.55
	0.5	-2.76	0.51	1.36	-0.62	72.94	79.17	0.79	716.38	2.13	0.10	0.53	0.60
8	50	-4.45	0.37	1.03	1.26	69.69	76.13	0.73	45.88	2.23	0.09	0.62	0.43
	20	-3.20	0.43	0.89	1.06	60.62	67.06	0.80	86.43	1.98	0.08	0.54	0.38
	10	-2.46	0.47	0.76	0.96	56.31	62.75	0.82	121.38	1.86	0.08	0.51	0.36
	5	-1.80	0.51	0.64	0.89	55.08	61.52	0.84	169.29	1.83	0.08	0.51	0.35
	2	-1.01	0.56	0.47	0.81	58.04	64.49	0.84	264.39	1.91	0.08	0.53	0.36
	1	-0.44	0.59	0.34	0.76	62.88	69.32	0.83	368.92	2.04	0.09	0.56	0.39
	0.5	0.11	0.63	0.21	0.72	69.00	75.45	0.81	510.12	2.21	0.09	0.61	0.42
9	50	-5.72	0.59	1.17	0.67	56.81	62.54	0.87	25.55	1.94	0.10	0.51	0.48
	20	-4.31	0.60	1.05	0.47	58.15	63.89	0.85	49.38	1.98	0.10	0.52	0.49
	10	-3.35	0.61	0.94	0.23	60.01	65.74	0.84	70.47	2.04	0.11	0.54	0.51
	5	-2.46	0.62	0.82	-0.02	63.09	68.82	0.83	98.94	2.14	0.11	0.56	0.54
	2	-1.33	0.64	0.67	-0.34	68.68	74.42	0.79	158.11	2.35	0.12	0.62	0.59
	1	-0.51	0.66	0.54	-0.57	73.64	79.38	0.76	231.56	2.54	0.13	0.67	0.63
	0.5	0.29	0.67	0.42	-0.80	78.90	84.64	0.73	347.19	2.77	0.14	0.73	0.69

Cluster	AEP (%)	Predictor Variables				Performance Metrics				Descriptor Standard Error			
		Constant	Area	$DR_{10\%}$	D_c	AIC	BIC	R2	*RMSE*	Const	Area	$DR_{10\%}$	D_c
10	50	-10.61	0.45	2.09	2.74	56.96	62.95	0.85	22.11	2.36	0.11	0.51	0.90
	20	-9.71	0.45	2.12	2.51	58.73	64.71	0.85	51.00	2.43	0.11	0.53	0.93
	10	-9.47	0.43	2.19	2.40	59.69	65.68	0.84	74.43	2.46	0.11	0.54	0.94
	5	-9.35	0.41	2.27	2.30	61.25	67.24	0.84	102.09	2.52	0.11	0.55	0.96
	2	-9.27	0.39	2.38	2.16	64.54	70.53	0.82	150.93	2.65	0.12	0.58	1.01
	1	-9.25	0.36	2.47	2.04	67.91	73.90	0.81	202.71	2.79	0.13	0.61	1.07
	0.5	-9.24	0.34	2.56	1.92	71.88	77.86	0.79	273.32	2.96	0.13	0.64	1.13
11	50	-5.04	0.38	1.53	-0.86	76.29	82.51	0.69	45.46	2.38	0.08	0.55	0.48
	20	-4.76	0.37	1.68	-0.77	72.34	78.56	0.72	124.18	2.25	0.08	0.52	0.45
	10	-4.58	0.36	1.73	-0.68	72.73	78.95	0.72	209.21	2.26	0.08	0.52	0.46
	5	-4.45	0.36	1.77	-0.58	75.13	81.35	0.71	330.63	2.34	0.08	0.54	0.47
	2	-4.35	0.36	1.81	-0.44	80.81	87.03	0.67	579.06	2.54	0.09	0.59	0.51
	1	-4.32	0.35	1.85	-0.33	86.43	92.65	0.64	868.95	2.75	0.09	0.64	0.55
	0.5	-4.33	0.35	1.88	-0.21	92.69	98.91	0.60	1293.24	3.01	0.10	0.70	0.61
12	50	-1.54	0.57	0.41	-0.20	65.09	70.69	0.70	31.12	2.77	0.11	0.67	0.52
	20	-1.23	0.58	0.52	-0.07	61.40	67.01	0.74	70.62	2.60	0.11	0.63	0.49
	10	-1.16	0.58	0.59	-0.01	61.49	67.10	0.74	114.18	2.61	0.11	0.63	0.49
	5	-1.18	0.58	0.67	0.05	63.13	68.73	0.73	184.73	2.68	0.11	0.65	0.51
	2	-1.30	0.57	0.77	0.13	67.32	72.93	0.71	348.76	2.87	0.12	0.69	0.54
	1	-1.44	0.57	0.86	0.20	71.66	77.26	0.69	555.34	3.09	0.13	0.75	0.58
	0.5	-1.62	0.56	0.95	0.26	76.62	82.23	0.65	870.73	3.36	0.14	0.81	0.63

Cluster	AEP (%)	Predictor Variables				Performance Metrics				Descriptor Standard Error			
		Constant	Area	$DR_{10\%}$	D_c	AIC	BIC	R2	*RMSE*	Const	Area	$DR_{10\%}$	D_c
13	50	-8.74	0.24	2.61	-0.26	49.63	55.23	0.52	55.49	7.10	0.18	1.71	0.61
	20	-6.76	0.26	2.24	0.10	46.21	51.81	0.56	116.49	6.71	0.17	1.62	0.57
	10	-5.35	0.29	1.93	0.26	45.07	50.68	0.59	166.75	6.58	0.16	1.59	0.56
	5	-3.97	0.32	1.60	0.38	45.39	50.99	0.61	228.90	6.62	0.16	1.60	0.56
	2	-2.14	0.38	1.15	0.52	48.01	53.61	0.62	346.98	6.91	0.17	1.67	0.59
	1	-0.76	0.42	0.81	0.62	51.37	56.97	0.62	481.10	7.31	0.18	1.76	0.62
	0.5	0.62	0.46	0.46	0.70	55.54	61.14	0.61	673.89	7.84	0.19	1.89	0.67
14	50	2.06	0.61	-0.15	-1.28	46.33	52.19	0.90	56.89	2.64	0.11	0.79	0.80
	20	2.25	0.59	0.05	-1.23	43.72	49.58	0.91	128.53	2.54	0.10	0.75	0.77
	10	1.97	0.57	0.20	-1.11	45.17	51.04	0.91	203.03	2.59	0.10	0.77	0.79
	5	1.55	0.54	0.36	-0.96	48.27	54.13	0.90	311.68	2.72	0.11	0.81	0.83
	2	0.85	0.51	0.58	-0.74	54.34	60.20	0.88	543.71	2.99	0.12	0.89	0.91
	1	0.25	0.49	0.75	-0.55	59.94	65.80	0.85	819.72	3.27	0.13	0.97	0.99
	0.5	-0.39	0.46	0.92	-0.36	66.01	71.87	0.83	1219.50	3.59	0.14	1.07	1.09
15	50	-2.66	0.48	0.94	-0.29	94.33	101.73	0.67	109.57	3.25	0.11	0.84	0.19
	20	-1.68	0.50	0.85	-0.21	87.60	95.00	0.70	215.81	3.02	0.10	0.78	0.18
	10	-1.14	0.51	0.79	-0.18	82.37	89.77	0.72	295.18	2.86	0.10	0.74	0.17
	5	-0.64	0.52	0.73	-0.15	77.94	85.34	0.74	381.31	2.73	0.09	0.71	0.16
	2	0.00	0.53	0.64	-0.12	74.77	82.17	0.76	523.04	2.64	0.09	0.68	0.16
	1	0.47	0.54	0.56	-0.10	75.24	82.64	0.75	669.31	2.65	0.09	0.69	0.16
	0.5	0.95	0.55	0.48	-0.08	78.39	85.79	0.74	870.04	2.74	0.09	0.71	0.16

Cluster	AEP (%)	Predictor Variables				Performance Metrics				Descriptor Standard Error			
		Constant	**Area**	$DR_{10\%}$	D_c	**AIC**	**BIC**	**R2**	*RMSE*	**Const**	**Area**	$DR_{10\%}$	D_c
16	50	-12.95	0.09	3.75	-0.03	64.30	71.25	0.70	65.74	3.23	0.11	0.83	0.24
	20	-10.38	0.17	3.22	0.05	61.27	68.22	0.73	137.92	3.12	0.10	0.80	0.23
	10	-8.32	0.24	2.75	0.01	59.86	66.81	0.74	197.84	3.07	0.10	0.79	0.23
	5	-6.21	0.31	2.24	-0.06	60.02	66.98	0.75	269.08	3.07	0.10	0.79	0.23
	2	-3.35	0.41	1.54	-0.17	63.01	69.96	0.74	393.81	3.18	0.10	0.82	0.24
	1	-1.15	0.49	0.99	-0.26	67.16	74.11	0.73	524.73	3.35	0.11	0.86	0.25
	0.5	1.09	0.57	0.43	-0.35	72.54	79.49	0.72	703.77	3.57	0.12	0.92	0.26
17	50	1.42	0.45	-0.10	0.17	73.64	79.38	0.49	106.58	1.73	0.11	0.44	0.18
	20	1.88	0.46	-0.01	0.05	69.28	75.01	0.51	222.02	1.62	0.10	0.41	0.17
	10	2.02	0.47	0.04	0.01	67.54	73.27	0.53	314.02	1.57	0.10	0.40	0.16
	5	2.06	0.49	0.08	-0.03	67.10	72.84	0.55	417.38	1.56	0.10	0.39	0.16
	2	2.02	0.51	0.15	-0.08	68.71	74.44	0.56	590.48	1.60	0.10	0.40	0.16
	1	1.93	0.53	0.20	-0.12	71.43	77.17	0.55	771.84	1.67	0.10	0.42	0.17
	0.5	1.82	0.54	0.25	-0.16	75.18	80.92	0.54	1026.93	1.78	0.11	0.45	0.18
18	50	-0.64	0.47	0.30	-0.55	98.35	104.34	0.60	83.23	2.38	0.10	0.68	0.30
	20	-0.42	0.52	0.38	-0.56	86.79	92.77	0.73	163.95	2.00	0.08	0.57	0.25
	10	-0.48	0.55	0.46	-0.56	81.03	87.02	0.78	259.29	1.83	0.08	0.52	0.23
	5	-0.62	0.58	0.53	-0.56	76.94	82.92	0.82	413.97	1.72	0.07	0.49	0.21
	2	-0.85	0.61	0.63	-0.56	74.48	80.47	0.85	755.26	1.66	0.07	0.47	0.21
	1	-1.05	0.64	0.70	-0.56	75.07	81.06	0.86	1158.70	1.67	0.07	0.47	0.21
	0.5	-1.26	0.66	0.76	-0.55	77.62	83.60	0.86	1736.17	1.74	0.07	0.49	0.22

Cluster	AEP (%)	Predictor Variables				Performance Metrics				Descriptor Standard Error			
		Constant	**Area**	***$DR_{10\%}$***	***D_c***	**AIC**	**BIC**	**R2**	***RMSE***	**Const**	**Area**	***$DR_{10\%}$***	***D_c***
19	50	-2.97	0.53	0.76	-0.47	88.46	95.41	0.87	93.88	1.25	0.07	0.38	0.11
	20	-1.60	0.53	0.66	-0.47	82.47	89.42	0.88	220.79	1.17	0.07	0.36	0.10
	10	-0.92	0.54	0.60	-0.47	81.10	88.05	0.88	358.49	1.15	0.07	0.35	0.10
	5	-0.32	0.54	0.54	-0.48	81.38	88.33	0.88	550.21	1.15	0.07	0.35	0.10
	2	0.39	0.55	0.45	-0.50	84.28	91.23	0.87	913.58	1.19	0.07	0.36	0.11
	1	0.90	0.56	0.38	-0.51	88.23	95.18	0.85	1296.15	1.25	0.07	0.38	0.11
	0.5	1.40	0.57	0.31	-0.52	93.41	100.37	0.83	1800.72	1.33	0.08	0.41	0.12
20 - 24	50	1.54	1.04	-0.59	0.48	99.29	105.15	0.74	90.16	2.56	0.25	0.86	0.30
	20	1.23	0.96	-0.27	0.32	91.09	96.95	0.78	163.69	2.25	0.22	0.76	0.26
	10	0.96	0.91	-0.09	0.21	86.87	92.73	0.80	197.70	2.11	0.21	0.71	0.24
	5	0.68	0.86	0.07	0.10	83.47	89.33	0.81	213.18	2.00	0.20	0.67	0.23
	2	0.28	0.79	0.28	-0.05	80.28	86.15	0.81	203.67	1.90	0.19	0.64	0.22
	1	-0.02	0.75	0.42	-0.16	79.04	84.90	0.81	173.92	1.87	0.18	0.63	0.22
	0.5	-0.32	0.70	0.57	-0.26	78.92	84.78	0.80	138.64	1.86	0.18	0.62	0.22
25	50	-1.42	0.53	0.36	-0.48	81.81	87.91	0.77	37.41	1.61	0.08	0.48	0.15
	20	-0.74	0.57	0.38	-0.46	72.89	79.00	0.84	112.59	1.41	0.07	0.42	0.13
	10	-0.53	0.59	0.41	-0.45	69.85	75.96	0.86	208.97	1.35	0.07	0.40	0.12
	5	-0.41	0.61	0.44	-0.44	69.05	75.16	0.87	353.16	1.33	0.07	0.40	0.12
	2	-0.34	0.63	0.49	-0.42	71.31	77.41	0.87	644.49	1.38	0.07	0.41	0.13
	1	-0.32	0.64	0.53	-0.41	75.11	81.22	0.86	967.86	1.45	0.07	0.43	0.13
	0.5	-0.33	0.66	0.57	-0.39	80.18	86.28	0.85	1410.15	1.57	0.08	0.47	0.14

Cluster	AEP (%)	Predictor Variables				Performance Metrics				Descriptor Standard Error			
		Constant	Area	$DR_{10\%}$	D_c	AIC	BIC	R2	*RMSE*	Const	Area	$DR_{10\%}$	D_c
26 - 27	50	-0.31	0.66	-0.14	-0.67	79.20	85.31	0.80	93.69	1.45	0.09	0.45	0.15
	20	0.15	0.65	0.01	-0.60	76.39	82.50	0.82	220.61	1.39	0.08	0.43	0.14
	10	0.20	0.64	0.14	-0.56	76.50	82.61	0.82	330.08	1.40	0.08	0.43	0.14
	5	0.16	0.63	0.26	-0.51	77.52	83.63	0.82	455.21	1.42	0.09	0.44	0.15
	2	0.02	0.61	0.43	-0.45	80.07	86.17	0.82	653.45	1.47	0.09	0.45	0.15
	1	-0.13	0.60	0.57	-0.40	82.78	88.88	0.81	838.89	1.53	0.09	0.47	0.16
	0.5	-0.30	0.59	0.70	-0.36	86.04	92.15	0.80	1069.61	1.61	0.10	0.50	0.17
28	50	0.55	0.64	-0.32	-0.53	88.12	95.07	0.77	63.80	1.28	0.09	0.42	0.14
	20	0.38	0.62	0.01	-0.49	81.69	88.64	0.82	134.36	1.18	0.09	0.39	0.13
	10	0.19	0.60	0.20	-0.46	81.17	88.12	0.83	193.99	1.18	0.09	0.39	0.13
	5	-0.06	0.58	0.39	-0.43	82.43	89.38	0.83	264.66	1.19	0.09	0.40	0.13
	2	-0.48	0.56	0.64	-0.38	86.48	93.43	0.83	395.58	1.25	0.09	0.42	0.14
	1	-0.84	0.54	0.84	-0.35	91.07	98.02	0.82	552.31	1.32	0.10	0.44	0.14
	0.5	-1.23	0.52	1.03	-0.31	96.66	103.61	0.81	799.62	1.41	0.10	0.47	0.15
29	50	3.28	0.43	-0.55	-0.28	45.89	51.88	0.60	21.12	0.90	0.07	0.26	0.15
	20	3.64	0.48	-0.50	-0.46	40.41	46.39	0.67	54.23	0.83	0.07	0.24	0.13
	10	3.77	0.51	-0.49	-0.55	41.38	47.37	0.70	87.93	0.84	0.07	0.24	0.14
	5	3.80	0.55	-0.47	-0.62	44.97	50.96	0.70	133.01	0.89	0.07	0.25	0.14
	2	3.76	0.59	-0.45	-0.73	52.40	58.39	0.70	222.02	0.99	0.08	0.28	0.16
	1	3.68	0.63	-0.44	-0.81	59.08	65.07	0.68	327.31	1.10	0.09	0.31	0.18
	0.5	3.57	0.67	-0.42	-0.89	66.07	72.05	0.67	488.15	1.22	0.10	0.35	0.20

Cluster	AEP (%)	Predictor Variables				Performance Metrics				Descriptor Standard Error			
		Constant	Area	$DR_{10\%}$	D_c	AIC	BIC	R2	*RMSE*	Const	Area	$DR_{10\%}$	D_c
30 - 32	50	2.03	0.43	-0.26	0.00	46.10	51.83	0.65	44.56	1.14	0.07	0.30	0.16
	20	2.28	0.48	-0.18	-0.16	42.52	48.25	0.70	106.12	1.08	0.07	0.28	0.15
	10	2.33	0.52	-0.15	-0.23	44.06	49.80	0.71	172.74	1.10	0.07	0.29	0.15
	5	2.29	0.55	-0.12	-0.30	47.87	53.61	0.72	264.57	1.17	0.08	0.31	0.16
	2	2.15	0.61	-0.09	-0.39	55.16	60.89	0.71	443.55	1.32	0.08	0.35	0.18
	1	1.99	0.65	-0.06	-0.46	61.51	67.24	0.69	643.88	1.46	0.09	0.38	0.20
	0.5	1.80	0.69	-0.03	-0.54	68.06	73.79	0.67	928.10	1.62	0.10	0.43	0.23
33	50	-5.21	0.58	1.33	0.22	87.00	94.31	0.75	87.46	3.11	0.10	0.80	0.16
	20	-4.20	0.55	1.32	0.15	84.00	91.32	0.74	181.47	3.01	0.10	0.77	0.15
	10	-3.55	0.53	1.28	0.12	81.00	88.31	0.74	257.19	2.92	0.10	0.75	0.15
	5	-2.91	0.51	1.23	0.08	79.14	86.45	0.74	343.91	2.86	0.09	0.73	0.14
	2	-2.09	0.49	1.15	0.04	80.01	87.33	0.72	491.86	2.89	0.10	0.74	0.14
	1	-1.46	0.47	1.08	0.00	83.60	90.92	0.69	646.68	3.00	0.10	0.77	0.15
	0.5	-0.82	0.46	1.01	-0.03	89.43	96.74	0.64	860.90	3.20	0.11	0.82	0.16
34	50	-3.74	0.58	1.00	0.20	76.55	82.99	0.82	62.54	3.32	0.14	0.90	0.16
	20	-5.12	0.47	1.62	0.16	73.16	79.61	0.82	134.83	3.17	0.14	0.86	0.16
	10	-5.57	0.43	1.86	0.14	72.16	78.61	0.82	208.98	3.13	0.13	0.85	0.15
	5	-5.92	0.39	2.06	0.12	73.09	79.53	0.82	323.37	3.17	0.14	0.86	0.16
	2	-6.36	0.35	2.29	0.09	77.64	84.08	0.79	580.22	3.37	0.15	0.91	0.17
	1	-6.69	0.32	2.46	0.08	83.22	89.66	0.76	896.93	3.63	0.16	0.98	0.18
	0.5	-7.03	0.29	2.63	0.06	90.02	96.46	0.72	1371.64	3.98	0.17	1.08	0.20

Cluster	AEP (%)	Predictor Variables				Performance Metrics				Descriptor Standard Error			
		Constant	**Area**	***$DR_{10\%}$***	***D_c***	**AIC**	**BIC**	**R2**	***RMSE***	**Const**	**Area**	***$DR_{10\%}$***	***D_c***
35	50	-10.25	0.32	2.79	0.43	50.35	56.34	0.80	61.37	3.22	0.12	0.85	0.13
	20	-10.01	0.26	2.97	0.33	50.97	56.96	0.78	130.39	3.25	0.12	0.85	0.13
	10	-8.90	0.27	2.81	0.26	51.94	57.93	0.76	189.16	3.30	0.13	0.87	0.14
	5	-7.56	0.28	2.56	0.18	54.45	60.44	0.73	259.30	3.43	0.13	0.90	0.14
	2	-5.60	0.30	2.16	0.07	60.19	66.17	0.67	384.30	3.74	0.14	0.98	0.15
	1	-4.04	0.32	1.83	-0.02	65.88	71.87	0.62	521.30	4.08	0.16	1.07	0.17
	0.5	-2.42	0.35	1.49	-0.11	72.22	78.21	0.56	719.58	4.49	0.17	1.18	0.18
36	50	-13.94	0.07	4.00	0.05	53.54	59.40	0.73	68.85	4.07	0.14	1.07	0.27
	20	-11.26	0.16	3.43	0.10	51.38	57.25	0.75	142.82	3.94	0.14	1.04	0.26
	10	-9.09	0.23	2.92	0.04	49.89	55.75	0.76	201.34	3.85	0.13	1.01	0.26
	5	-6.85	0.31	2.38	-0.04	49.26	55.12	0.76	266.78	3.81	0.13	1.00	0.25
	2	-3.81	0.42	1.63	-0.18	50.10	55.96	0.76	370.55	3.86	0.13	1.02	0.26
	1	-1.45	0.51	1.03	-0.29	51.98	57.84	0.75	467.84	3.98	0.14	1.05	0.27
	0.5	0.96	0.60	0.42	-0.40	54.76	60.63	0.74	588.23	4.15	0.14	1.09	0.28
37	50	-4.31	0.53	1.18	0.23	84.20	91.25	0.82	67.81	3.04	0.13	0.82	0.14
	20	-5.48	0.44	1.74	0.17	81.49	88.53	0.81	148.38	2.95	0.12	0.79	0.14
	10	-5.76	0.40	1.93	0.13	81.57	88.62	0.80	229.78	2.95	0.12	0.79	0.14
	5	-5.94	0.37	2.08	0.09	84.00	91.05	0.79	347.41	3.03	0.13	0.81	0.14
	2	-6.11	0.34	2.24	0.04	90.79	97.84	0.75	599.18	3.28	0.14	0.88	0.15
	1	-6.23	0.31	2.35	0.00	97.99	105.04	0.71	905.00	3.57	0.15	0.96	0.17
	0.5	-6.36	0.29	2.45	-0.05	106.23	113.27	0.67	1362.87	3.93	0.16	1.05	0.18

Cluster	AEP (%)	Predictor Variables				Performance Metrics				Descriptor Standard Error			
		Constant	**Area**	$DR_{10\%}$	D_c	**AIC**	**BIC**	**R2**	*RMSE*	**Const**	**Area**	$DR_{10\%}$	D_c
38 - 39	50	-4.20	0.64	0.80	0.31	74.59	80.69	0.82	20.38	3.63	0.14	0.86	1.10
	20	-2.34	0.65	0.56	0.14	73.35	79.46	0.82	41.78	3.57	0.14	0.85	1.08
	10	-1.30	0.67	0.41	-0.03	72.88	78.99	0.82	60.38	3.54	0.14	0.84	1.07
	5	-0.34	0.69	0.27	-0.21	73.48	79.58	0.81	86.00	3.57	0.14	0.85	1.08
	2	0.87	0.71	0.07	-0.47	76.21	82.32	0.80	142.06	3.72	0.14	0.88	1.13
	1	1.77	0.73	-0.08	-0.68	79.64	85.75	0.78	213.80	3.91	0.15	0.93	1.18
	0.5	2.67	0.75	-0.23	-0.89	83.99	90.09	0.76	327.32	4.17	0.16	0.99	1.26
40	50	-5.93	0.58	1.23	0.42	56.73	62.71	0.88	30.28	1.88	0.08	0.44	0.49
	20	-5.32	0.55	1.35	0.14	57.62	63.60	0.88	76.63	1.91	0.08	0.44	0.50
	10	-5.00	0.56	1.38	0.07	60.51	66.50	0.87	122.32	2.00	0.09	0.46	0.52
	5	-4.75	0.56	1.40	0.03	64.73	70.72	0.86	183.46	2.13	0.09	0.49	0.56
	2	-4.46	0.57	1.41	0.00	71.86	77.85	0.84	305.20	2.37	0.10	0.55	0.62
	1	-4.27	0.59	1.41	-0.02	77.96	83.94	0.82	449.36	2.60	0.11	0.60	0.68
	0.5	-4.09	0.60	1.41	-0.03	84.31	90.30	0.80	667.02	2.86	0.12	0.66	0.75
41	50	-7.77	0.59	1.53	1.39	60.44	66.30	0.83	17.67	2.15	0.10	0.49	0.95
	20	-6.86	0.58	1.57	1.12	63.11	68.97	0.82	43.23	2.24	0.11	0.51	0.99
	10	-6.33	0.59	1.55	0.95	65.48	71.34	0.81	68.45	2.32	0.11	0.53	1.03
	5	-5.84	0.59	1.53	0.78	68.50	74.36	0.80	103.49	2.43	0.12	0.55	1.08
	2	-5.23	0.61	1.50	0.55	73.59	79.45	0.78	177.04	2.64	0.13	0.60	1.17
	1	-4.78	0.62	1.46	0.37	78.10	83.97	0.75	266.75	2.83	0.14	0.64	1.25
	0.5	-4.34	0.63	1.43	0.18	83.00	88.86	0.73	403.25	3.05	0.15	0.69	1.35

Cluster	**AEP (%)**	**Predictor Variables**				**Performance Metrics**				**Descriptor Standard Error**			
		Constant	**Area**	***$DR_{10\%}$***	***D_c***	**AIC**	**BIC**	**R2**	***RMSE***	**Const**	**Area**	***$DR_{10\%}$***	***D_c***
42	50	-8.67	0.54	1.84	0.82	58.54	64.52	0.88	29.74	1.76	0.09	0.43	0.44
	20	-8.18	0.52	1.97	0.64	59.06	65.05	0.88	71.07	1.77	0.10	0.43	0.44
	10	-7.95	0.51	2.03	0.55	59.95	65.94	0.88	107.26	1.79	0.10	0.44	0.45
	5	-7.79	0.51	2.08	0.46	61.60	67.59	0.88	153.25	1.84	0.10	0.45	0.46
	2	-7.63	0.50	2.15	0.35	65.25	71.24	0.87	246.97	1.94	0.11	0.48	0.48
	1	-7.54	0.50	2.19	0.27	69.04	75.03	0.86	364.78	2.06	0.11	0.50	0.51
	0.5	-7.46	0.49	2.24	0.18	73.51	79.50	0.85	550.80	2.20	0.12	0.54	0.55

APPENDIX H: QRT MODEL COEFFICIENTS AND REGRESSION STATISTICS FOR THE FIVE SUPER REGIONS

Super Region	Predictor Variables	AEP (%)	Descriptor Number				Performance Metrics				Descriptor Standard Error			
			CONST	1	2	3	AIC	BIC	R2	*RMSE*	Const	1	2	3
National	A, $DR_{10\%}$, D_c	50	-1.90	0.53	0.56	-0.28	989.42	1005.22	0.70	108.79	0.52	0.03	0.14	0.05
		20	-1.49	0.53	0.66	-0.29	896.05	911.85	0.76	257.69	0.46	0.03	0.13	0.04
		10	-1.33	0.53	0.72	-0.29	858.49	874.29	0.78	412.86	0.44	0.02	0.12	0.04
		5	-1.25	0.53	0.77	-0.30	841.66	857.45	0.79	620.66	0.43	0.02	0.12	0.04
		2	-1.19	0.53	0.84	-0.31	852.36	868.15	0.79	1004.62	0.44	0.02	0.12	0.04
		1	-1.18	0.53	0.89	-0.32	883.62	899.42	0.79	1405.64	0.45	0.02	0.13	0.04
		0.5	-1.19	0.53	0.95	-0.33	930.59	946.38	0.77	1935.36	0.48	0.03	0.13	0.04
1	A, $DR_{10\%}$, $H24_{1\%}$	50	2.83	0.61	-0.32	-1.26	271.88	282.57	0.74	71.44	2.62	0.10	0.75	0.41
		20	3.55	0.63	-0.39	-1.03	247.61	258.30	0.79	130.79	2.34	0.09	0.67	0.37
		10	3.91	0.65	-0.45	-0.90	237.92	248.61	0.81	191.71	2.24	0.08	0.64	0.35
		5	4.18	0.67	-0.50	-0.78	233.96	244.66	0.81	306.24	2.20	0.08	0.63	0.34
		2	4.47	0.68	-0.57	-0.62	237.46	248.16	0.81	581.01	2.23	0.08	0.64	0.35
		1	4.65	0.70	-0.61	-0.51	246.02	256.72	0.80	905.31	2.32	0.09	0.67	0.36
		0.5	4.81	0.71	-0.66	-0.39	258.52	269.21	0.79	1354.95	2.46	0.09	0.71	0.39
2	A, $S_{10\text{-}85}$, $DR_{10\%}$	50	-4.10	0.61	1.03	0.16	162.41	171.62	0.77	85.80	2.48	0.09	0.65	0.13
		20	-3.79	0.55	1.22	0.10	153.56	162.77	0.77	179.31	2.34	0.09	0.62	0.12
		10	-3.61	0.52	1.31	0.07	149.12	158.34	0.76	263.76	2.27	0.08	0.60	0.12
		5	-3.47	0.49	1.38	0.04	147.92	157.14	0.76	374.52	2.25	0.08	0.59	0.12
		2	-3.33	0.45	1.48	0.01	152.47	161.69	0.73	595.44	2.32	0.09	0.61	0.12
		1	-3.26	0.42	1.56	-0.02	160.35	169.56	0.69	855.57	2.45	0.09	0.64	0.13
		0.5	-3.20	0.39	1.63	-0.05	171.13	180.35	0.64	1240.72	2.63	0.10	0.69	0.14

Super Region	Predictor Variables	AEP (%)	Descriptor Number				Performance Metrics				Descriptor Standard Error			
			CONST	1	2	3	AIC	BIC	R2	*RMSE*	Const	1	2	3
3	A, E_O, $H24_{1\%}$	50	-4.08	0.38	1.34	-0.19	240.26	249.79	0.66	66.04	1.22	0.08	0.38	0.13
		20	-3.17	0.44	1.24	-0.24	214.99	224.52	0.75	114.66	1.04	0.07	0.33	0.11
		10	-2.68	0.47	1.16	-0.27	202.96	212.49	0.79	155.31	0.96	0.06	0.30	0.10
		5	-2.28	0.50	1.09	-0.30	194.85	204.38	0.81	217.19	0.92	0.06	0.29	0.10
		2	-1.82	0.54	1.00	-0.34	190.84	200.37	0.83	367.24	0.89	0.06	0.28	0.10
		1	-1.51	0.56	0.94	-0.37	193.25	202.78	0.83	562.78	0.91	0.06	0.29	0.10
		0.5	-1.22	0.59	0.88	-0.40	199.95	209.48	0.82	863.48	0.95	0.06	0.30	0.10
4	A, E_O, *MAP*	50	-4.01	0.58	0.84	0.25	175.69	185.64	0.80	35.08	1.05	0.06	0.28	0.21
		20	-4.30	0.54	1.14	0.29	167.74	177.70	0.82	83.87	1.01	0.06	0.27	0.20
		10	-4.45	0.52	1.29	0.29	168.14	178.10	0.82	137.79	1.01	0.06	0.27	0.20
		5	-4.63	0.50	1.43	0.29	173.57	183.52	0.81	219.24	1.04	0.06	0.28	0.21
		2	-4.91	0.48	1.60	0.29	187.64	197.59	0.79	395.54	1.13	0.06	0.30	0.22
		1	-5.16	0.46	1.73	0.29	202.04	212.00	0.77	609.08	1.22	0.07	0.32	0.24
		0.5	-5.42	0.45	1.85	0.29	218.29	228.25	0.74	928.37	1.34	0.07	0.35	0.26
5	A, *MAP*, $S_{10\text{-}85}$	50	1.86	0.44	-0.23	-0.04	48.25	54.23	0.65	43.17	1.09	0.07	0.29	0.15
		20	2.22	0.49	-0.18	-0.21	44.54	50.53	0.69	104.19	1.03	0.07	0.27	0.14
		10	2.29	0.53	-0.15	-0.28	46.17	52.15	0.71	170.15	1.06	0.07	0.28	0.14
		5	2.26	0.57	-0.13	-0.36	50.11	56.09	0.71	260.85	1.12	0.07	0.30	0.15
		2	2.12	0.62	-0.10	-0.45	57.66	63.65	0.71	436.98	1.26	0.08	0.33	0.17
		1	1.97	0.66	-0.07	-0.52	64.28	70.26	0.69	633.46	1.39	0.09	0.37	0.19
		0.5	1.79	0.70	-0.04	-0.60	71.13	77.11	0.68	911.48	1.54	0.10	0.41	0.21

APPENDIX I: *RMSE* AND *RMSE*$_R$ PERFORMANCE METRICS FOR MODELS DEVELOPED USING CLUSTERING

Model Scale	RMSE							RMSEr						
	50	20	10	5	2	1	0.5	50	20	10	5	2	1	0.5
QRT														
National	115.17	280.87	456.70	693.38	1131.02	1587.00	2186.60	1.17	1.21	1.31	1.43	1.59	1.71	1.81
Cluster 1	58.63	92.64	103.95	109.39	115.54	124.46	140.05	0.74	0.66	0.59	0.54	0.50	0.49	0.52
Cluster 2	29.35	75.44	119.51	175.22	276.74	386.64	540.85	0.64	0.62	0.61	0.60	0.60	0.61	0.63
Cluster 3	19.36	48.55	80.11	128.53	246.14	411.66	693.53	0.42	0.38	0.38	0.39	0.44	0.50	0.58
Cluster 4	17.00	31.42	46.13	71.86	141.78	239.78	397.75	0.68	0.41	0.35	0.35	0.41	0.48	0.57
Cluster 5	58.17	167.84	287.20	456.96	799.59	1192.25	1755.27	0.81	0.86	0.91	0.97	1.06	1.12	1.19
Cluster 6	40.01	97.52	153.47	225.91	358.57	497.60	681.38	0.55	0.51	0.51	0.52	0.55	0.58	0.61
Cluster 7	54.74	88.24	113.48	141.78	182.88	215.39	248.20	0.56	0.45	0.45	0.48	0.53	0.57	0.61
Cluster 8	77.81	156.43	233.22	335.31	529.34	739.40	1023.88	0.48	0.38	0.36	0.36	0.38	0.40	0.42
Cluster 9	30.31	56.80	73.64	91.05	120.15	150.13	189.20	0.68	0.60	0.56	0.54	0.55	0.59	0.64
Cluster 10	37.64	87.98	123.05	155.53	195.49	225.35	258.10	0.68	0.70	0.66	0.62	0.57	0.53	0.50
Cluster 11	21.63	60.82	100.55	155.65	277.54	447.69	751.41	1.00	0.90	0.85	0.81	0.80	0.85	0.93
Cluster 12	41.95	107.16	212.09	408.16	873.61	1447.09	2289.16	0.29	0.27	0.33	0.44	0.60	0.74	0.88
Cluster 13	81.18	163.21	214.85	261.41	322.73	380.03	464.59	0.50	0.43	0.39	0.35	0.31	0.29	0.29
Cluster 14	126.15	254.13	362.58	495.22	737.43	997.16	1347.47	0.58	0.43	0.39	0.38	0.39	0.41	0.44
Cluster 15	158.86	336.91	469.06	600.39	775.99	912.83	1055.38	0.88	0.88	0.87	0.87	0.86	0.86	0.86
Cluster 16	54.39	126.28	203.71	312.69	531.38	780.02	1132.24	0.36	0.36	0.38	0.41	0.47	0.52	0.59
Cluster 17	133.52	313.97	513.37	789.15	1318.93	1894.50	2679.49	0.89	0.85	0.87	0.90	0.95	0.98	1.01
Cluster 18	87.85	198.48	276.75	345.55	429.15	540.67	821.78	1.82	1.21	0.94	0.71	0.50	0.42	0.43
Cluster 19	226.03	610.93	1079.00	1746.28	2981.26	4216.41	5743.52	0.19	0.22	0.27	0.33	0.42	0.50	0.57
Cluster 20	85.80	126.70	151.28	171.64	193.43	206.78	217.91	0.69	0.61	0.60	0.60	0.60	0.61	0.62
Cluster 21	63.53	105.15	130.57	151.43	172.91	185.10	194.22	1.42	1.34	1.29	1.24	1.18	1.13	1.08
Cluster 22	9.95	13.16	13.38	11.66	15.17	33.62	68.28	2.06	0.93	0.57	0.33	0.27	0.43	0.64
Cluster 23	8.38	13.38	13.51	11.30	7.90	11.91	22.31	0.71	0.61	0.49	0.36	0.22	0.32	0.57

Model Scale	RMSE							RMSEr						
	50	20	10	5	2	1	0.5	50	20	10	5	2	1	0.5
Cluster 24	16.04	25.48	35.18	48.31	74.66	105.78	152.17	0.55	0.34	0.29	0.27	0.26	0.26	0.27
Cluster 25	226.49	424.63	591.26	777.02	1071.58	1357.83	1735.38	2.69	1.80	1.56	1.41	1.27	1.21	1.17
Cluster 26	15.71	37.29	57.06	83.98	142.32	214.84	323.52	0.29	0.26	0.24	0.24	0.27	0.30	0.34
Cluster 27	14.77	36.59	59.29	88.64	140.18	191.07	254.55	0.71	0.79	0.90	1.03	1.23	1.40	1.59
Cluster 28	22.56	57.87	91.72	134.36	210.87	291.23	399.63	0.39	0.36	0.35	0.35	0.35	0.35	0.36
Cluster 29	21.81	45.80	72.16	108.03	185.93	290.48	465.85	0.26	0.20	0.19	0.19	0.20	0.23	0.26
Cluster 30	20.00	54.80	94.15	157.59	312.80	525.34	876.32	0.48	0.40	0.38	0.39	0.44	0.48	0.54
Cluster 31	39.58	99.23	152.86	218.19	331.84	448.04	601.51	0.67	0.73	0.74	0.76	0.80	0.84	0.88
Cluster 32	30.05	69.26	106.46	154.12	241.46	334.75	462.09	0.57	0.56	0.59	0.63	0.71	0.79	0.88
Cluster 33	80.02	140.63	205.84	305.39	512.98	743.17	1052.69	0.33	0.30	0.32	0.38	0.49	0.59	0.71
Cluster 34	51.73	138.03	235.98	375.57	654.57	972.96	1432.78	0.72	0.83	0.94	1.06	1.22	1.34	1.46
Cluster 35	81.61	155.89	200.04	240.86	308.73	390.83	519.74	0.50	0.48	0.44	0.41	0.39	0.41	0.45
Cluster 36	107.01	228.56	344.65	479.52	686.81	866.90	1068.98	0.86	0.99	1.16	1.34	1.60	1.81	2.04
Cluster 37	47.51	153.08	300.00	534.03	1048.47	1673.73	2606.40	0.54	0.65	0.75	0.85	0.96	1.04	1.11
Cluster 38	22.13	45.15	61.38	78.83	108.38	139.54	181.89	0.53	0.52	0.50	0.49	0.50	0.52	0.56
Cluster 39	19.04	60.62	115.95	201.04	383.73	603.17	929.89	0.31	0.45	0.57	0.69	0.85	0.98	1.11
Cluster 40	49.87	123.15	192.03	285.84	488.47	748.11	1157.11	0.42	0.42	0.42	0.42	0.46	0.52	0.59
Cluster 41	6.72	15.40	23.91	34.34	51.66	67.98	87.64	0.82	0.97	1.08	1.20	1.37	1.49	1.63
Cluster 42	16.43	35.66	58.74	93.13	164.69	247.60	365.91	0.49	0.39	0.39	0.41	0.46	0.50	0.55
IF1														
National	152.40	355.17	516.07	688.07	945.62	1181.06	1478.38	1.55	1.53	1.48	1.42	1.33	1.27	1.22
Cluster 1	42.39	69.25	80.57	89.05	99.91	109.38	120.47	0.53	0.49	0.46	0.44	0.43	0.43	0.45
Cluster 2	28.03	71.09	111.03	159.94	245.60	335.33	459.03	0.62	0.59	0.57	0.55	0.53	0.53	0.53
Cluster 3	19.79	43.85	63.71	92.12	174.04	307.72	549.93	0.43	0.35	0.30	0.28	0.31	0.38	0.46
Cluster 4	13.69	35.05	60.50	98.90	181.61	282.26	433.71	0.55	0.46	0.46	0.48	0.52	0.57	0.62

Model Scale	RMSE							RMSEr						
	50	**20**	**10**	**5**	**2**	**1**	**0.5**	**50**	**20**	**10**	**5**	**2**	**1**	**0.5**
Cluster 5	79.08	229.95	387.30	599.87	1004.16	1444.60	2052.67	1.11	1.18	1.23	1.28	1.33	1.36	1.39
Cluster 6	28.42	84.61	141.41	214.85	346.47	481.08	655.87	0.39	0.44	0.47	0.49	0.53	0.56	0.59
Cluster 7	39.91	82.49	110.85	138.85	176.47	205.42	234.59	0.41	0.43	0.44	0.47	0.51	0.55	0.58
Cluster 8	76.74	157.74	230.50	315.76	452.39	578.20	728.23	0.47	0.38	0.36	0.34	0.32	0.31	0.30
Cluster 9	30.24	61.37	81.68	100.98	126.86	147.66	170.08	0.68	0.65	0.62	0.60	0.59	0.58	0.58
Cluster 10	38.35	84.04	117.65	151.08	197.04	235.59	280.53	0.70	0.66	0.63	0.60	0.57	0.55	0.54
Cluster 11	20.74	61.58	103.90	161.73	282.39	440.94	716.28	0.96	0.92	0.87	0.84	0.82	0.83	0.89
Cluster 12	64.72	194.91	317.37	476.09	785.53	1146.97	1682.96	0.44	0.49	0.50	0.51	0.54	0.59	0.64
Cluster 13	119.88	257.58	358.98	461.01	599.67	709.98	829.55	0.73	0.68	0.65	0.62	0.58	0.55	0.52
Cluster 14	117.69	250.68	357.12	478.18	684.58	900.40	1193.75	0.54	0.42	0.39	0.37	0.36	0.37	0.39
Cluster 15	160.38	316.31	428.86	540.56	693.12	817.02	952.35	0.89	0.82	0.80	0.78	0.77	0.77	0.77
Cluster 16	51.71	137.15	224.20	337.37	546.64	773.32	1089.33	0.34	0.39	0.42	0.44	0.48	0.52	0.57
Cluster 17	81.56	164.28	268.79	427.19	758.22	1141.70	1689.19	0.54	0.45	0.46	0.49	0.54	0.59	0.64
Cluster 18	52.41	148.26	234.97	336.15	503.40	680.69	956.53	1.09	0.90	0.79	0.70	0.58	0.52	0.50
Cluster 19	333.70	674.75	896.07	1111.78	1443.97	1780.00	2230.50	0.28	0.24	0.22	0.21	0.20	0.21	0.22
Cluster 20	81.04	139.75	176.56	207.79	241.77	262.82	280.50	0.65	0.67	0.70	0.72	0.75	0.78	0.80
Cluster 21	30.09	46.42	54.64	60.56	66.82	71.82	78.13	0.67	0.59	0.54	0.50	0.46	0.44	0.43
Cluster 22	2.07	5.56	9.75	15.50	26.58	38.72	55.46	0.43	0.39	0.41	0.44	0.47	0.49	0.52
Cluster 23	4.32	9.60	12.41	14.55	16.69	17.96	19.00	0.37	0.44	0.45	0.46	0.47	0.48	0.48
Cluster 24	7.48	17.36	22.73	35.48	88.85	171.41	306.59	0.25	0.23	0.19	0.20	0.30	0.42	0.54
Cluster 25	132.08	327.32	525.36	780.57	1234.76	1696.30	2294.05	1.57	1.39	1.39	1.42	1.47	1.51	1.55
Cluster 26	28.82	59.34	80.29	98.72	120.75	141.30	177.29	0.54	0.41	0.34	0.29	0.23	0.20	0.19
Cluster 27	21.62	40.51	55.14	70.70	93.71	113.71	136.55	1.04	0.88	0.84	0.82	0.82	0.83	0.85
Cluster 28	30.74	72.12	104.78	144.57	233.51	360.97	577.57	0.53	0.45	0.40	0.38	0.39	0.44	0.52
Cluster 29	22.98	60.10	110.42	198.04	415.37	706.77	1173.51	0.27	0.26	0.30	0.35	0.45	0.55	0.66

Model Scale	RMSE							RMSEr						
	50	**20**	**10**	**5**	**2**	**1**	**0.5**	**50**	**20**	**10**	**5**	**2**	**1**	**0.5**
Cluster 30	18.54	50.44	83.61	128.98	222.81	340.84	531.59	0.45	0.37	0.34	0.32	0.31	0.31	0.33
Cluster 31	44.51	100.36	141.88	181.52	228.40	257.26	278.56	0.75	0.73	0.69	0.63	0.55	0.48	0.41
Cluster 32	20.08	51.16	85.08	130.38	213.09	298.52	410.06	0.38	0.42	0.47	0.54	0.63	0.70	0.78
Cluster 33	88.52	166.86	242.49	344.62	543.04	760.20	1054.96	0.37	0.36	0.38	0.43	0.52	0.61	0.71
Cluster 34	75.51	229.56	400.78	633.22	1065.32	1520.18	2127.92	1.05	1.38	1.59	1.78	1.98	2.09	2.17
Cluster 35	102.91	194.13	245.12	283.20	321.27	352.95	408.73	0.64	0.60	0.54	0.48	0.41	0.37	0.36
Cluster 36	156.64	326.78	451.08	568.14	709.77	806.75	895.40	1.26	1.42	1.51	1.59	1.65	1.69	1.71
Cluster 37	75.76	197.27	342.62	557.34	1004.54	1531.57	2305.61	0.86	0.84	0.86	0.88	0.92	0.95	0.98
Cluster 38	21.22	47.28	67.19	88.73	122.21	153.59	192.83	0.50	0.54	0.55	0.55	0.56	0.57	0.59
Cluster 39	108.51	279.42	435.22	618.99	916.77	1197.92	1548.72	1.76	2.08	2.13	2.11	2.03	1.94	1.84
Cluster 40	111.15	302.58	483.54	704.68	1079.85	1451.23	1932.81	0.94	1.03	1.05	1.04	1.02	1.00	0.99
Cluster 41	9.71	19.26	27.00	35.20	46.76	56.06	65.88	1.18	1.21	1.23	1.23	1.24	1.23	1.22
Cluster 42	15.79	51.48	88.51	137.34	227.15	321.75	448.31	0.47	0.56	0.58	0.61	0.63	0.65	0.67

APPENDIX J: *BIAS* AND *BIASR* PERFORMANCE METRICS FOR MODELS DEVELOPED USING CLUSTERING

Model Scale	BIAS							BIASr						
	50	**20**	**10**	**5**	**2**	**1**	**0.5**	**50**	**20**	**10**	**5**	**2**	**1**	**0.5**
QRT														
National	53.25	116.85	175.30	247.97	386.80	540.20	749.47	0.83	0.70	0.66	0.64	0.66	0.71	0.77
Cluster 1	38.95	63.02	74.85	82.37	87.73	90.44	96.67	0.46	0.44	0.42	0.40	0.38	0.38	0.41
Cluster 2	16.90	43.78	72.32	112.77	191.05	276.37	393.80	0.38	0.40	0.41	0.43	0.47	0.50	0.53
Cluster 3	13.62	35.90	65.94	116.19	232.07	381.98	620.55	0.33	0.39	0.51	0.66	0.92	1.18	1.50
Cluster 4	11.89	22.01	34.70	55.21	95.03	160.86	260.26	3.17	1.91	1.48	1.18	0.90	0.76	0.66
Cluster 5	31.67	94.39	169.22	276.36	488.88	728.87	1096.03	0.34	0.35	0.41	0.49	0.62	0.75	0.95
Cluster 6	32.14	80.70	133.40	199.49	316.88	439.73	595.63	0.88	0.85	0.87	0.89	0.93	0.98	1.03
Cluster 7	43.08	75.26	96.01	114.75	137.70	153.98	169.36	0.46	0.39	0.41	0.44	0.47	0.50	0.52
Cluster 8	70.69	133.49	190.72	274.50	438.06	604.95	827.13	0.97	0.79	0.72	0.69	0.69	0.70	0.74
Cluster 9	19.08	36.76	49.92	63.74	84.58	104.12	127.03	0.37	0.37	0.39	0.41	0.45	0.49	0.53
Cluster 10	23.70	60.43	88.02	115.47	152.44	181.55	211.70	0.30	0.39	0.42	0.42	0.43	0.44	0.45
Cluster 11	14.70	39.31	63.25	92.73	190.27	317.71	522.96	0.72	0.58	0.60	0.65	0.83	1.01	1.24
Cluster 12	40.77	93.88	202.48	381.64	773.56	1236.60	1901.55	0.74	0.71	0.80	0.93	1.13	1.32	1.54
Cluster 13	75.83	144.98	179.56	207.43	277.55	339.08	414.45	0.63	0.52	0.43	0.37	0.35	0.35	0.35
Cluster 14	90.54	189.10	263.77	351.25	485.83	645.20	891.02	0.39	0.33	0.32	0.32	0.34	0.36	0.41
Cluster 15	118.86	253.04	344.30	426.80	521.58	582.23	658.65	1.17	0.97	0.85	0.75	0.63	0.54	0.47
Cluster 16	49.32	101.43	158.55	239.60	409.68	597.27	858.40	0.41	0.35	0.34	0.36	0.42	0.48	0.55
Cluster 17	79.19	203.98	341.51	532.32	900.46	1301.90	1852.04	0.27	0.37	0.40	0.42	0.44	0.44	0.46
Cluster 18	60.83	136.94	188.33	224.88	327.88	472.42	680.08	1.23	0.95	0.88	0.86	0.96	1.08	1.22
Cluster 19	222.98	517.13	1007.32	1659.96	2805.63	3916.25	5267.24	0.19	0.19	0.25	0.32	0.41	0.49	0.59
Cluster 20	54.16	76.76	87.06	93.19	97.25	101.99	115.76	0.49	0.52	0.54	0.57	0.62	0.68	0.77
Cluster 21	38.03	63.41	79.24	92.45	106.21	113.97	119.54	0.69	0.61	0.58	0.55	0.51	0.47	0.44
Cluster 22	4.94	6.89	6.89	7.29	8.71	20.26	41.04	1.47	1.10	0.93	0.87	0.90	1.03	1.23
Cluster 23	6.91	10.00	10.00	9.13	6.24	8.64	17.47	0.56	0.40	0.32	0.29	0.27	0.36	0.59

Model Scale	BIAS							BIASr						
	50	20	10	5	2	1	0.5	50	20	10	5	2	1	0.5
Cluster 24	12.40	22.82	32.46	44.53	66.43	90.06	123.18	1.45	1.31	1.31	1.32	1.37	1.41	1.45
Cluster 25	84.87	171.92	255.67	366.85	563.18	764.40	1031.49	1.01	0.86	0.84	0.86	0.91	0.97	1.05
Cluster 26	12.33	29.16	45.01	64.32	97.65	134.58	195.43	0.51	0.39	0.34	0.30	0.24	0.22	0.23
Cluster 27	10.00	23.89	37.13	53.20	79.62	107.30	141.89	0.42	0.53	0.60	0.68	0.80	0.91	1.03
Cluster 28	18.16	50.33	82.53	123.12	193.13	262.10	349.08	0.28	0.31	0.33	0.34	0.36	0.37	0.38
Cluster 29	18.29	39.09	61.13	96.66	164.26	237.67	364.99	0.38	0.37	0.37	0.40	0.44	0.47	0.51
Cluster 30	14.36	36.67	75.52	138.09	272.96	434.52	672.94	0.81	0.50	0.45	0.42	0.38	0.37	0.35
Cluster 31	28.87	73.45	113.43	159.62	232.07	296.87	410.46	0.47	0.53	0.53	0.53	0.50	0.48	0.51
Cluster 32	21.00	47.03	77.94	117.90	182.38	240.41	345.02	0.31	0.30	0.40	0.55	0.79	1.03	1.36
Cluster 33	60.80	103.06	139.81	205.24	333.05	466.69	656.65	0.23	0.22	0.22	0.25	0.30	0.35	0.43
Cluster 34	40.95	105.63	173.79	262.38	454.75	701.41	1077.20	0.65	0.66	0.68	0.72	0.84	1.03	1.29
Cluster 35	54.07	99.31	127.65	156.49	208.96	285.97	399.43	0.40	0.37	0.36	0.36	0.39	0.43	0.48
Cluster 36	67.30	144.47	210.58	281.88	392.59	494.91	609.87	0.39	0.50	0.62	0.75	0.98	1.20	1.45
Cluster 37	29.09	90.89	176.76	316.94	625.98	1005.62	1574.43	1.15	1.08	1.11	1.16	1.25	1.37	1.49
Cluster 38	17.67	37.02	50.07	61.45	72.99	86.03	107.10	1.41	1.20	1.06	0.93	0.77	0.68	0.63
Cluster 39	13.10	43.52	83.46	143.39	271.37	428.67	661.87	0.30	0.36	0.40	0.42	0.46	0.51	0.58
Cluster 40	34.08	84.40	151.01	244.54	429.79	641.52	946.58	0.36	0.33	0.37	0.41	0.48	0.55	0.65
Cluster 41	4.35	9.66	15.01	21.52	32.00	42.05	55.71	0.52	0.62	0.67	0.75	0.88	1.02	1.18
Cluster 42	13.40	29.59	45.55	71.07	122.98	180.76	273.41	0.55	0.49	0.48	0.49	0.51	0.52	0.56
IF1														
National	55.54	129.18	191.56	263.46	385.69	509.55	671.82	0.71	0.69	0.69	0.70	0.73	0.77	0.82
Cluster 1	29.95	51.53	62.67	71.38	80.71	86.78	93.53	0.44	0.43	0.41	0.39	0.37	0.36	0.36
Cluster 2	17.06	43.70	67.69	98.62	161.06	230.78	326.67	0.38	0.39	0.39	0.42	0.46	0.50	0.53
Cluster 3	12.81	33.10	55.79	88.91	159.98	270.53	499.52	0.28	0.36	0.46	0.57	0.75	0.92	1.16
Cluster 4	8.94	21.84	35.31	56.74	103.84	160.42	243.43	1.00	0.83	0.78	0.73	0.69	0.70	0.71

Model Scale	BIAS							BIASr						
	50	**20**	**10**	**5**	**2**	**1**	**0.5**	**50**	**20**	**10**	**5**	**2**	**1**	**0.5**
Cluster 5	52.68	152.83	262.59	411.71	692.11	993.64	1404.94	0.66	0.70	0.75	0.81	0.91	1.00	1.12
Cluster 6	22.74	71.87	122.75	188.47	304.43	422.07	572.25	0.66	0.74	0.79	0.84	0.91	0.96	1.02
Cluster 7	35.98	69.66	89.26	105.31	123.32	135.49	146.89	0.49	0.43	0.44	0.45	0.45	0.45	0.44
Cluster 8	52.93	112.93	174.68	249.19	368.98	476.83	600.08	0.44	0.46	0.52	0.59	0.70	0.82	0.96
Cluster 9	19.65	39.72	53.59	66.95	84.91	99.63	115.37	0.39	0.40	0.41	0.42	0.43	0.45	0.47
Cluster 10	24.44	58.86	85.27	112.21	149.08	178.34	211.78	0.42	0.43	0.43	0.42	0.41	0.40	0.41
Cluster 11	13.13	35.98	58.95	99.17	193.38	314.67	508.33	0.59	0.52	0.57	0.67	0.85	1.04	1.27
Cluster 12	59.14	174.40	277.76	394.69	633.93	974.08	1451.93	0.78	0.77	0.80	0.83	0.90	0.98	1.07
Cluster 13	115.49	243.57	328.29	415.78	544.41	633.09	708.34	0.85	0.76	0.68	0.62	0.58	0.55	0.53
Cluster 14	80.46	183.54	272.45	374.09	535.94	685.08	857.09	0.36	0.32	0.32	0.32	0.34	0.36	0.38
Cluster 15	113.63	225.19	300.70	392.69	516.36	610.29	703.27	0.91	0.90	0.85	0.81	0.76	0.72	0.68
Cluster 16	44.14	118.06	189.26	282.92	463.06	647.79	889.01	0.36	0.39	0.41	0.44	0.50	0.55	0.60
Cluster 17	55.84	118.39	191.53	298.57	521.48	798.53	1193.10	0.34	0.32	0.30	0.27	0.25	0.34	0.44
Cluster 18	37.59	101.84	155.32	214.82	371.80	550.31	808.30	0.72	0.68	0.71	0.77	0.96	1.14	1.37
Cluster 19	273.07	619.57	863.48	1079.37	1303.58	1469.06	2009.42	0.25	0.23	0.22	0.21	0.19	0.19	0.23
Cluster 20	49.78	85.57	107.93	126.69	146.77	158.96	169.03	0.34	0.32	0.32	0.33	0.33	0.33	0.34
Cluster 21	16.36	26.21	32.68	37.99	43.62	47.06	49.99	0.49	0.41	0.38	0.37	0.35	0.35	0.34
Cluster 22	1.30	3.95	6.56	9.69	14.91	19.87	25.92	0.91	0.95	0.94	0.91	0.85	0.81	0.76
Cluster 23	3.40	7.33	9.64	11.52	13.49	14.68	15.68	0.27	0.31	0.33	0.35	0.37	0.38	0.39
Cluster 24	6.98	16.63	19.88	34.48	70.10	115.07	183.39	1.46	1.20	1.08	1.00	0.91	0.84	0.77
Cluster 25	58.89	151.00	245.35	374.46	612.70	864.75	1200.43	0.84	0.81	0.83	0.86	0.93	1.00	1.09
Cluster 26	19.33	42.82	59.85	75.38	91.63	103.15	151.20	0.41	0.41	0.40	0.38	0.36	0.34	0.39
Cluster 27	12.12	24.24	34.38	45.40	61.46	74.89	89.56	0.50	0.54	0.59	0.65	0.72	0.77	0.82
Cluster 28	25.91	58.71	79.97	116.63	205.59	307.83	455.05	0.41	0.37	0.34	0.35	0.38	0.39	0.41
Cluster 29	18.95	48.91	82.40	137.48	264.33	444.05	727.02	0.23	0.20	0.20	0.24	0.31	0.41	0.53

Model Scale	BIAS							BIASr						
	50	**20**	**10**	**5**	**2**	**1**	**0.5**	**50**	**20**	**10**	**5**	**2**	**1**	**0.5**
Cluster 30	15.94	42.90	68.58	99.65	165.54	283.35	480.56	0.64	0.46	0.40	0.33	0.27	0.28	0.30
Cluster 31	32.40	74.22	106.05	137.28	175.75	200.63	219.29	0.52	0.52	0.50	0.47	0.42	0.39	0.36
Cluster 32	15.44	36.15	59.46	90.98	147.52	201.48	274.30	0.31	0.34	0.42	0.51	0.65	0.75	0.87
Cluster 33	67.08	125.34	183.08	257.98	384.70	514.72	675.97	0.27	0.29	0.30	0.32	0.34	0.38	0.41
Cluster 34	48.88	150.41	262.53	414.00	694.00	986.35	1372.04	0.69	0.84	0.95	1.07	1.29	1.50	1.76
Cluster 35	75.16	138.98	173.92	195.24	198.03	193.90	260.89	0.55	0.51	0.47	0.41	0.35	0.31	0.33
Cluster 36	89.20	185.05	250.56	308.94	381.34	437.63	490.84	0.51	0.61	0.67	0.73	0.82	0.90	0.97
Cluster 37	46.93	119.09	203.69	327.47	584.52	895.58	1357.97	0.98	0.91	0.93	0.98	1.09	1.22	1.39
Cluster 38	16.81	39.04	54.42	67.93	81.51	92.22	108.83	1.40	1.23	1.11	0.98	0.83	0.73	0.68
Cluster 39	50.97	134.70	214.04	314.89	495.74	683.92	939.20	0.48	0.57	0.59	0.62	0.66	0.69	0.75
Cluster 40	91.12	242.35	381.58	550.50	825.03	1083.67	1396.13	0.67	0.71	0.72	0.73	0.73	0.75	0.78
Cluster 41	7.40	14.12	19.50	25.08	32.69	38.63	45.21	1.16	1.08	1.05	1.03	1.02	1.02	1.05
Cluster 42	11.01	33.85	58.33	91.34	153.46	220.11	312.76	0.28	0.32	0.35	0.37	0.40	0.42	0.45

APPENDIX K: *RMSE* AND *RMSE*$_R$ PERFORMANCE METRICS FOR MODELS DEVELOPED USING REGION OF INFLUENCE

Model Scale	*RMSE*							*RMSEr*						
	50	**20**	**10**	**5**	**2**	**1**	**0.5**	**50**	**20**	**10**	**5**	**2**	**1**	**0.5**
IF2														
National	133.26	329.02	512.95	736.71	1112.36	1477.66	1940.74	1.43	1.51	1.58	1.63	1.68	1.71	1.73
Super 1	97.47	207.26	303.47	423.58	650.23	909.4	1290.89	0.7	0.72	0.74	0.78	0.85	0.94	1.06
Super 2	44.21	94.95	148.05	220.39	362.42	524.87	761.31	0.62	0.52	0.51	0.52	0.54	0.57	0.6
Super 3	118.82	232.86	297.85	369.41	567.61	882.08	1402.49	0.95	0.75	0.64	0.57	0.6	0.72	0.9
Super 4	39.46	102.4	165.71	252.89	429.57	638.43	950.24	0.65	0.7	0.73	0.78	0.87	0.96	1.07
Super 5	112.92	291.66	451.09	634.71	921.46	1179.74	1486.98	2.06	2.37	2.48	2.53	2.51	2.44	2.33
IF1														
National	130.54	348.2	539.51	758.71	1103.82	1422.55	1814.55	1.4	1.6	1.66	1.68	1.67	1.65	1.62
Super 1	97.75	199.78	291.23	404.77	610.45	832.4	1143.35	0.7	0.69	0.71	0.74	0.8	0.86	0.94
Super 2	43.98	97.07	150.48	223.51	372.92	553.59	829.69	0.61	0.54	0.52	0.53	0.56	0.6	0.66
Super 3	97.08	209.75	300.58	403.13	568.93	729.14	934.4	0.78	0.68	0.64	0.62	0.6	0.6	0.6
Super 4	39.66	97.73	157.64	239.36	398.17	576.87	834.25	0.66	0.67	0.7	0.74	0.81	0.87	0.94
Super 5	105.51	250.15	398.72	607.91	1036.72	1539.72	2274.79	1.92	2.03	2.19	2.43	2.82	3.18	3.56

APPENDIX K: *BIAS* AND *BIASR* PERFORMANCE METRICS FOR MODELS DEVELOPED USING CLUSTERING

Model Scale	BIAS							BIASr						
	50	**20**	**10**	**5**	**2**	**1**	**0.5**	**50**	**20**	**10**	**5**	**2**	**1**	**0.5**
IF2														
National	54.52	123.12	183.33	257.42	390.26	528.46	713.79	0.88	0.79	0.76	0.75	0.79	0.83	0.9
Super 1	59.78	124.6	181.57	252.48	384.97	533.09	738.13	0.68	0.66	0.65	0.66	0.7	0.75	0.83
Super 2	25.04	55.87	87.37	130.27	217.53	319.94	463.75	0.41	0.37	0.38	0.42	0.48	0.55	0.63
Super 3	54.29	115.42	158.67	202.33	315.63	480.13	720.69	0.79	0.7	0.67	0.66	0.68	0.71	0.77
Super 4	24.72	61.67	97.47	143.73	236.36	342.06	506.22	0.58	0.56	0.58	0.61	0.68	0.76	0.85
Super 5	44.6	103.38	151.05	202.98	282.13	362.64	475.03	1.06	0.89	0.82	0.78	0.78	0.81	0.88
IF1														
National	51.42	122.53	185.39	260.78	391.09	524.4	703.75	0.81	0.75	0.74	0.75	0.78	0.83	0.9
Super 1	58.76	121.03	175.31	239.85	356.73	484.55	662.19	0.61	0.61	0.63	0.65	0.71	0.78	0.88
Super 2	25.99	57.59	86.15	128.79	213.19	312.64	462.22	0.44	0.38	0.38	0.41	0.48	0.57	0.68
Super 3	47.15	106.48	156.32	212.99	313.29	416.52	552.77	0.71	0.66	0.66	0.67	0.71	0.76	0.82
Super 4	24.78	60.35	94.61	138.05	221.64	316.65	454.28	0.58	0.56	0.57	0.58	0.63	0.7	0.78
Super 5	43.27	95.77	140.68	193.9	283.96	374.83	504.76	1.03	0.87	0.8	0.75	0.72	0.73	0.76

I couldn't quite find my own words to describe how it feels completing this work, so I am relying on somebody else's:

"It was quite impossible to describe.

Here is what it looked (read: feels) like.

It looked like a piano sounds shortly after being dropped down a well. It tasted yellow, and it felt Paisley. It smelled like the total eclipse of the moon." - Pratchett (1988)

Thank you for taking the time to read this work.