DETECTING AND ASSESSING THE IMPACTS OF OUTLIER EVENTS AND DATA AVAILABILITY ON DESIGN RAINFALL AND FLOOD ESTIMATION IN SOUTH AFRICA

KR Singh

Submitted in partial fulfilment of the

requirements for the degree of MSc Hydrology

School of Agriculture, Earth and Environmental Sciences

Centre for Water Resources Research

University of KwaZulu-Natal

Pietermaritzburg

January 2021

ABSTRACT

Accurate Design Rainfall Estimation (DRE) and Design Flood Estimation (DFE) require long periods of quality-controlled data for the planning, design, operation, and improved flood risk assessment of hydraulic structures. However, observed hydrological data frequently include outlier events and there is a decline of hydrological monitoring in South Africa which may impact DRE and DFE. It is therefore necessary to assess the impact of outlier events and reduced data availability on DRE and DFE. The aims of this study were to: (a) assess the impact of outlier events on DRE and DFE in South Africa, (b) assess the performance of outlier detection methods under South African conditions, and (c) assess the impact of reduced data availability on DRE and DFE in South Africa. The impact of synthetic Low Outlier (LO) and High Outlier (HO) events on DRE and DFE from observed and synthetically generated data series were assessed. The performance of the BoxPlot, Modified Z-Score (MSZ) and Multiple Grubbs-Beck Test (MGBT) outlier detection methods were assessed. Record length and network density were reduced to assess the impact of reduced data availability on DRE and DFE. Results from the analysis of observed data show that design rainfall is impacted by up to 22% and design floods by up to 45% in the presence of LOs. Design rainfall is impacted by up to 16% and design floods by up to 46% in the presence of HOs. For synthetically generated data series, design rainfall and floods are impacted by up to 2% and 1% respectively in the presence of LOs and by up to 13% in the presence of HOs. At best, LOs in observed rainfall and streamflow data are under-detected by up to 6% and 30% respectively by the MGBT method, whereas HOs are over-detected up to 50% and 150% respectively by the MZS method. Design rainfall and flood events are impacted by up to 4% and 24% respectively by reduced record lengths, and by up to 4.5% and 60% respectively from a reduced gauged network. This study indicates that outlier detection be adopted as regular practice in South Africa and that additional national resources must be directed towards maintaining and improving the hydrological monitoring networks in South Africa.

DECLARATION

I, Keanu Reeve Singh, declare that:

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

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

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

(iv) This dissertation 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;

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

Signed: 

Keanu Reeve Singh

Date: 28/01/2021

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 and Ms KA Johnson.

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: …[redacted]…….
Professor JC Smithers
Date:

Signed: …[redacted]…….
Ms KA Johnson
Date: 29/01/2021

ACKNOWLEDGEMENTS

First, to my Heavenly Father whose promise's hold true for supplying every need of mine according to the riches of His glory in Christ Jesus (Philippians 4:19) and for guiding me throughout this process (Psalms 32:8).

I would also like to acknowledge and express my sincere thanks and appreciation to the following individuals for their help and support throughout this study.

(a) Prof JC Smithers (Supervisor)
(b) Ms KA Johnson (Co-Supervisor)
(c) Centre for Water Resources Research (CWRR) and University of KwaZulu-Natal (UKZN) for administrative support
(d) Udhav Maharaj
(e) My family: Navin, Geetha and Cameron
(f) Mackyla Govender

The financial assistance of the National Research Foundation, National Flood Studies Programme, South African National Council on Large Dams, and Water Research Commission towards this research is hereby acknowledge.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

LIST OF ABBREVIATIONS

Symbol/Abbreviation	Description
AMS	Annual Maximum Series
ARC	Agricultural Research Committee
Avg.	Average
AWBM	Australian Water Balance Model
BP	BoxPlot method
CDF	Cumulative Density Function
CSAG	Climate System Analysis Group
CWRR	Centre for Water Resources Research
DFE	Design Flood Estimation
DRE	Design Rainfall Estimation
DWS	Department of Water and Sanitation
ESC	Eastern Summer Coastal
FFA	Flood Frequency Analysis
GBT	Grubbs-Best Test
GEV	Generalised Extreme Value
HDes	Hydrometric Desert Method
HO	High Outlier
HRand	Random Hydrometrical Reduction
IFM	Index Flood Method
IQR	Inter Quartile Range
IRM	Index Rainfall Method
Kappa	3 parameter Kappa
L-moments	Method of Linear Moments
LN	Log-Normal
LNO	3 Parameter Log Normal
LO	Low Outliers
LP3	Log-Pearson Type III
MAP	Mean Annual Precipitation
MARD	Mean Absolute Relative Differences
MGBT	Multiple Grubbs-Beck Test
MOM	Method of Moments
MRD	Mean Relative Differences

Symbol/Abbreviation	Description
MZS	Modified Z-Score method
NFSP	National Flood Studies Programme
NI	Northern Interior
NSE	Nash Sutcliffe Efficiency
PBIAS	Percent Bias
PD	Probability Distribution
PDF	Probability Density Function
PDSI	Palmer's Drought Severity Index
PILF	Potentially Influential Low Flows
POT	Peaks Over Threshold
QC	Quality Control
RFA	Rainfall Frequency Analysis
SANRAL	The South Africa National Roads Agency Limited
SASA	South African Sugar Association
SAWS	South African Weather Services
SCS-SA	South African Adaptation of the SCS Model
SWC	Southern Winter Coastal
UKZN	The University of KwaZulu-Natal
WRC	Water Research Commission

1. INTRODUCTION

Floods are naturally occurring events which may result in the loss of life, severe economic loss and environmental hazards (Smithers, 2012). The nature of extreme flood events worldwide are changing over time influenced by changes in climate driven by anthropogenic activities such as industrialisation and urbanization (Liu *et al.*, 2017). Climate change has intensified the natural variability, magnitude and frequency of extreme weather events resulting in flood events (Liu *et al.*, 2017). Furthermore, poverty, environmental degradation and an increase in the demand for natural resources has resulted in an increased number of people that are vulnerable to the disastrous impacts of floods (CGaTA, 2009). It is therefore essential to improve the accuracy of infrastructural design which can be achieved through accurate Design Flood Estimation (DFE).

Design floods are flood events with a given probability of exceedance. Design flood events are used in, *inter alia*, estimating environmental or ecological flows, managing water rights and transboundary water issues, the planning, design and operation of hydraulic structures, the development of flood forecasts and early warning systems aimed at protecting lives and property, and for educational and research purposes (USGS, 2006; Van Bladeren *et al.*, 2007).There are numerous DFE methods available for application in practice in South Africa (Smithers, 2012; SANRAL, 2013; Van Vuuren *et al.*, 2013; Van der Spuy and Rademeyer, 2018). These methods were derived from rainfall and streamflow data which are the two primary sources of hydrological data (Dent, 1994). Accurate DFE are required to improve the accuracy of infrastructural design by limiting the risk of failure, to limit the risk to loss of life and to limit over-expenditure on hydraulic structures.

Data screening and quality control are necessary to ensure that reliable input data are available for DFE. Data screening and quality control is regular practice as described in international literature, *e.g.* in United States of America (USA): Bulletin 17B and Bulletin 17C (England Jr *et al.*, 2019), Australia: Australian Rainfall and Runoff (Ball *et al.*, 2016), United Kingdom (UK): Flood Estimation Handbook (Robson and Reed, 1999) and in many European countries: FloodFreq Cost Action ES0901 (Madsen, 2013). Guidelines for data screening and quality control include the detection and treatment of outlier events prior to Flood Frequency Analysis (FFA) (England Jr *et al.*, 2019).

Outlier events are observations which significantly depart from the trend of the remaining dataset (Lamontagne *et al.*, 2013; Lamontagne *et al.*, 2016; England Jr *et al.*, 2019). Outlier events affect sample statistics and are potentially influential on Design Rainfall Estimation (DRE) and DFE (Lamontagne *et al.*, 2013; England Jr *et al.*, 2019). It is therefore necessary to assess the impact of outlier events on DRE and DFE and to assess the performance of outlier detection methods for potential use in South Africa (SA).

Observed data with sufficient quantity increases confidence in DRE and DFE therefore enhancing the value of information as understood by the public and private sectors (Van Bladeren *et al.*, 2007). However, there is a decline of hydrological monitoring both internationally (Lorenz and Kunstmann, 2012; Muller *et al.*, 2015; Stewart, 2015; Sunilkumar *et al.*, 2016) and in South Africa (Pitman, 2011; Pegram *et al.*, 2016) which is a pressing concern for many practitioners and policy developers (Stewart, 2015). The urgent need to assess the impact of outlier events and data availability on DFE in SA is highlighted by the National Flood Study Program (NFSP) (Smithers *et al.*, 2014).

1.1 Research Questions

This study will address the following questions:

(a) What is the impact of Low Outlier (LO) and High Outlier (HO) events on DRE and DFE and should outliers be excluded from DRE and DFE in SA?
(b) What is the performance of outlier detection methods in detecting LOs and HOs under South African conditions and should outlier detection be regular practice in DRE and DFE in SA?
(c) What is the impact of declining data availability, *i.e.*, rainfall and streamflow record lengths and monitoring network density, on DRE and DFE in SA?

1.2 Aims and Objectives

Aims of the study are the following:

(a) Assess the impact of LO and HO events on DRE and DFE in SA.
(b) Assess the performance of outlier detection methods under South African conditions.
(c) Assess the impact of reduced data availability on DRE and DFE in SA.

Specific objectives to meet the aims include the following:

(a) Undertake a comprehensive review of the relevant literature.
(b) Estimate design rainfall and floods using observed data and synthetically generated data series both with and without outliers [Aim (a)].
(c) Determine the impact of the introduction of LO and HO events on DRE and DFE [Aim (a)].
(d) Apply the BoxPlot (BP), Modified Z-Score (MZS) and Multiple Grubbs-Beck Test (MGBT) methods and assess their performance in detecting outlier events [Aim (b)].
(e) Evaluate the impact of reduced rainfall and streamflow record length and for different periods of time on DRE and DFE [Aim (c)].
(f) Evaluate the impact of a reduced rainfall and streamflow gauged density and proximity on DRE and DFE by means of a random and systematic reduction of gauges [Aim (c)].

1.3 Dissertation Structure

Chapter 2 contains a literature review on the assessment of data availability, outlier events and outlier detection methods. The study area, detailed data collation and pre-processing are presented in Chapter 3. Chapter 4 includes the method, results and discussion on assessing the impact of outlier events on DRE and DFE. Chapter 5 includes the method, results and discussion for the performance of outlier detection methods. Chapter 6 provides the method, results and discussion on assessing the impact of reduced record length on DRE and DFE. Chapter 7 provides the method, results and discussion on assessing the impact of reduced network density on DRE and DFE. The discussion, conclusions, and recommendations arising from this study are provided in Chapter 8. A list of references is presented in Chapter 9, and Appendices in Chapter 10 to Chapter 13.

2. LITERATURE REVIEW

This chapter contains a review of relevant literature on assessing hydrological data availability, outlier events and outlier detection methods.

2.1 Assessment of Data Availability

An assessment and evaluation of available water resources is essential for water resource management. The socio-economic and political history of South Africa has presented many challenges in collecting and maintaining rainfall and streamflow data (Hughes, 2008). However, despite these challenges, an adequately functioning hydrological monitoring network is required nationwide to, *inter alia*, provide information to enable well informed investment decisions in water resource management infrastructure (*e.g.* water supply and irrigation schemes) and to provide accurate and timely warning for floods and drought events (Sene and Farquharson, 1998). The use of observed data, with sufficient quantity and quality, from available networks also increases confidence in DRE and DFE, therefore enhancing the value of information as understood by the public and private sectors (Van Bladeren *et al.*, 2007). An evaluation of the currently available hydrological monitoring networks is therefore required prior to its application (Muller *et al.*, 2015). A review of the current rainfall and streamflow monitoring networks in SA and factors, such as network density, record length and gauge proximity, which affect the quantity and quality of observed data, are detailed in Section 2.1.1 and Section 2.1.2, respectively.

2.1.1 Rainfall data

The South African Weather Services (SAWS) holds the primary responsibility for collecting rainfall data in SA. Rainfall data is also collected by organisations such as the Agricultural Research Committee (ARC), the South African Sugar Association (SASA), the South African Environmental Observation Network (SAEON) and many private individuals. Figure 2.1 shows the annual active daily rainfall gauges in each year from a daily rainfall database maintained by Climate System Analysis Group (CSAG) at the University of Cape Town. There is a declining trend of active rainfall gauges from 1980 as shown in Figure 2.1. An average decline of 25% per decade over three decades (*i.e.*, 1979 to 2009) and a total decline of approximately 60% between 1979 and 2009 was calculated from Figure 2.1. The SAWS rainfall gauge network currently has approximately 1 200 rainfall gauges open which is

approximately equivalent to the number of rainfall gauges open in 1930 (Pegram *et al.*, 2016). The decline in the South African rainfall monitoring networks has had severe impacts on water resource management such as the calculation of irrigation demands and losses from reservoirs and wetlands (Pitman, 2011).

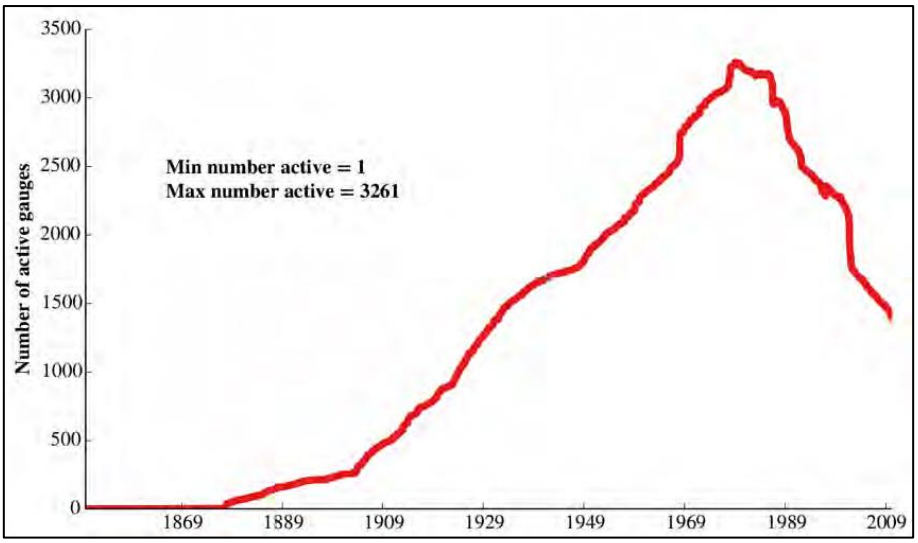

Figure 2.1 Annual active daily rainfall gauges in each year from the daily rainfall database maintained by CSAG (Pegram *et al.*, 2016)

Regarding the importance of record length, Boughton (2007) showed through the use of the Australian Water Balance Model (AWBM) that over-estimation of long-term runoff substantially reduces when 10 or more years of calibrated simulated runoff are available, although the under-estimation of runoff is still high even when 20 years of simulated runoff was used. Further, results have shown that an increase in record length improves the estimated rainfall characteristics and decreases errors in model estimates (Boughton, 2007). The use of a rainfall-runoff model to extend the records of less than 15 year will improve the estimates of long-term runoff (Boughton, 2007). Knisel *et al.* (1979) suggests that using short term records creates biased MAP estimates, especially in semi-arid regions. Dyler and Tyson (1977) highlighted that the period of available records influences its applicability as quasi-periodic fluctuations of rainfall generally occurs throughout recorded observations and thus Dent *et al.* (1987) states that a record of sufficient length is necessary to cover these quasi-periodic fluctuations, irrespective of the period when the observations were recorded. Dent *et al.* (1987) concludes that minimum lengths of record varies between regions with longer periods required in arid regions than in wetter regions. Smithers and Schulze (2000b) developed a short duration (*i.e.* sub-daily) rainfall database consisting of 412 stations throughout SA, and Smithers and

Schulze (2000a) developed a long duration (*i.e.* daily) rainfall database consisting of 11 171 stations throughout Southern Africa. Figure 2.2 shows the distribution of record lengths within the short duration rainfall database and Figure 2.3 shows the distribution of record lengths within the long duration rainfall database in Southern Africa. It is shown in Figure 2.2 that the majority of short duration rainfall is less than 20 years, and in Figure 2.3 that the majority of long duration rainfall is less than 25 years, which indicates a lack of longer-term monitoring.

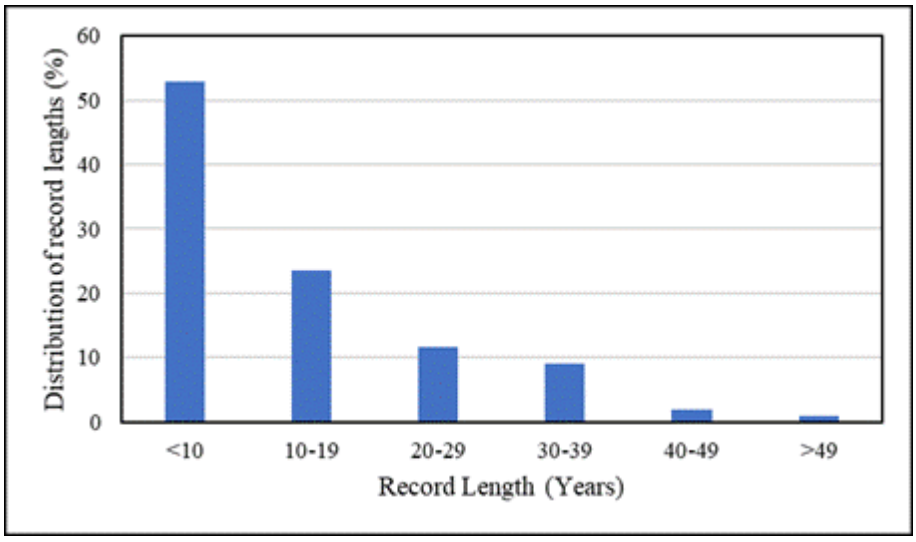

Figure 2.2 Distribution of record lengths in the short duration rainfall database for South Africa (after Smithers and Schulze, 2000a)

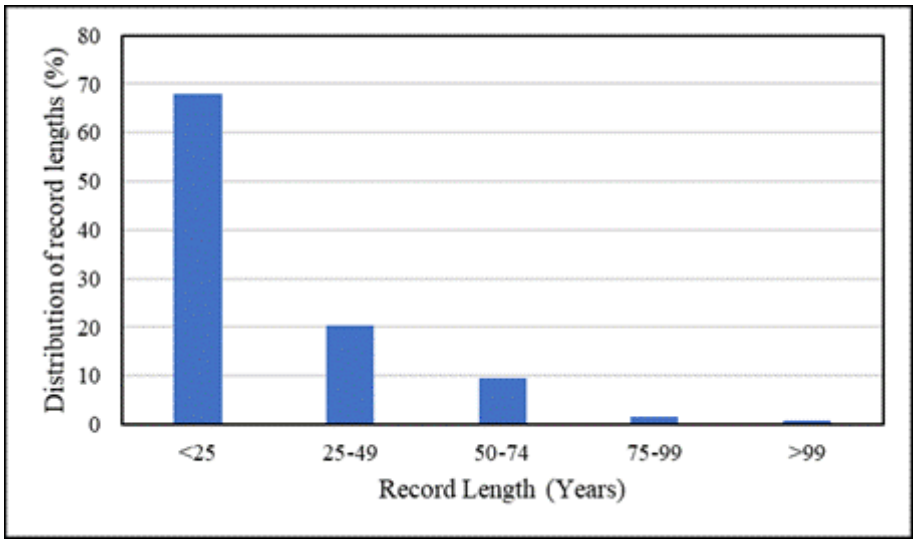

Figure 2.3 Distribution of record lengths in the long duration rainfall database for Southern Africa (after Smithers and Schulze, 2000b)

Regarding rainfall gauge density, rainfall events vary spatially within a catchment and a dense rainfall gauge network is able to better capture rainfall characteristics such as the magnitude and distribution of a rainfall event (Krajewski *et al.*, 2003). A denser rainfall gauge network improved total simulated streamflow (St-Hilaire *et al.*, 2003), reduced errors in simulated peaks (Bárdossy and Das, 2008), improved areal estimates of rainfall and reduced under-estimation of cumulative rainfall (St-Hilaire *et al.*, 2003; Bárdossy and Das, 2008). Xu *et al.* (2013) also showed that a larger sample of rainfall gauges improves the estimation of the Mean Annual Precipitation (MAP). The influence of rainfall gauge density is, however, dependent on factors such as rainfall type, topography, seasonality of precipitation and land use (Stewart, 2015).

Rainfall gauge location is another important consideration in DRE and DFE. Runoff is better estimated with an improved distribution of rainfall gauge locations rather than a more dense, but unevenly distributed rainfall gauge network (Xu *et al.*, 2013). St-Hilaire *et al.* (2003) concluded that the quality of data from a network is further improved with well-located rainfall stations and not only by a denser network. An analysis of the percentile, variances and means can be performed to identify gaps in a rainfall gauge network, and hence, gauges may be strategically placed in order to obtain the maximum usefulness of data (St-Hilaire *et al.*, 2003). Pegram *et al.* (2016) highlighted that gauges sparser than a radius of 35 km^2 per gauge show very small spatial correlation at a daily scale and could be treated as independent gauges. Elevation influences the distribution of meteorological variables such as temperatures, precipitation mechanisms and the rate of evaporation and is therefore another important consideration in rainfall gauge location (Bárdossy and Das, 2008).

Innovative technologies such as satellite-based remotely sensed data have been developed as a supplementary method of collecting rainfall data. However, without the presence of a dense gauge network, there is no meaningful manner of ground truthing remotely sensed data (Pegram *et al.*, 2016). Weather radar has also been used to measure rainfall, however, the use of raw radar data alone has introduced many errors in flood estimates (Sun *et al.*, 2000). The usefulness of radar data is increased when calibrated and combined with *in situ* rainfall gauge data (Sun *et al.*, 2000). A further concern is the lack of overlapping data periods between *in situ* and remotely sensed data (Hughes, 2008).

2.1.2 Streamflow data

The first long-term daily streamflow stage measurement in South Africa started in 1865 on the Van Stadens River and was initiated by the Port Elizabeth Town Council (Wessels and Rooseboom, 2009). The Department of Water and Sanitation (DWS) is currently responsible for monitoring water flow in SA (Pitman, 2011). Figure 2.4 shows the number of useful streamflow gauges, as determined by the project team, open in each year in SA as derived from the flow data set used in the WR2005 water resource assessment study (Middleton and Bailey, 2008; Pitman, 2011; Bailey and Pitman, 2016). There was a relatively slow start in developing the gauging network before 1950, a rapid growth after World War II, a peak in the 1980's, and a steady decline in the number of useful gauges thereafter as shown in Figure 2.4.

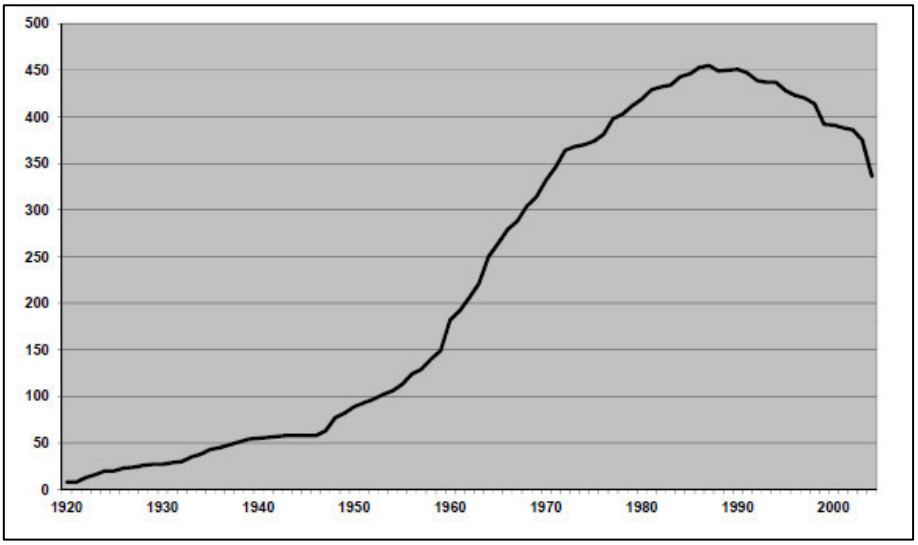

Figure 2.4 Number of useful streamflow gauges open in each year in South Africa (Pitman, 2011)

Data for 1 458 streamflow gauging stations and 89 synthesised dam inflow records located throughout South Africa were obtained from the DWS in a study by Nathanael (2015). A set of criteria were applied to filter these records and a final total of 1 097 stations were selected for further analysis. The criteria used by Nathanael (2015) were as follows:

(a) The available record length must have a minimum of 20 years of data.
(b) The station of interest must not be located at an outlet of a dam or significantly influenced by the presence of an upstream dam.
(c) The gauging station must be a river gauging station and not a natural spring, canal or pipeline.

(d) The percentage of the rating table exceedance must not be greater than or equal to 20% of the record.

The distribution of record lengths of the 1 097 stations, of which the majority of flow gauging weirs have less than 40 years of record, are shown in Figure 2.5.

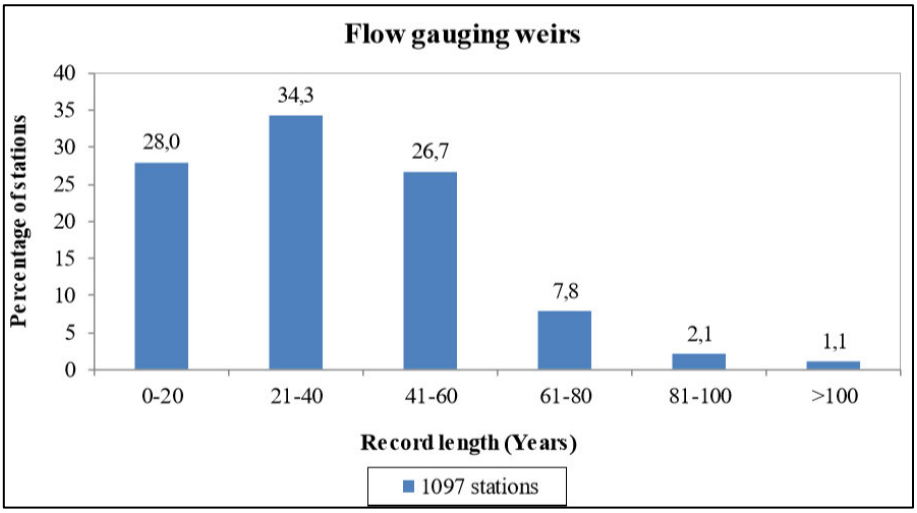

Figure 2.5 Distribution of record lengths for all flow gauging weirs across South Africa (Nathanael, 2015)

Related to streamflow record length, Pool *et al.* (2019) highlighted the importance of short record lengths in parameter regionalisation for design flood prediction in ungauged catchments. Results by Pool *et al.* (2019) showed that datasets with a minimum of three observations from a single hydrological year improved regionalisation for the majority of catchments and further highlighted that datasets with three observations have a similar value in design flood prediction compared to datasets with 24 observations. Pool *et al.* (2019) concluded that datasets with short record lengths are valuable for decision making in ungauged catchments and these findings are consistent with results from Viviroli and Seibert (2015) and Rojas-Serna *et al.* (2016).

In a statistical approach to DFE *i.e.* FFA, the assumption of stationarity of extreme events are often made irrespective of available record length and period of record which may be invalid in watersheds that are sensitive to climate variability and/or anthropogenic factors such as, *inter alia*, urbanisation, deforestation, land degradation (Koutsoyiannis *et al.*, 2009). Studies by Zhang *et al.* (2001) and Montanari and Koutsoyiannis (2014) have questioned the assumption of stationarity. Šraj *et al.* (2016) assessed the impacts of assuming stationarity on FFA by using one stationary and three non-stationary models *i.e.*, models which had varying Probability

Distribution (PD) parameters. Results indicated that all three non-stationary models fitted maximum annual floods better than the stationary model. The results have also highlighted the difference in quantile estimates between stationary and non-stationary models, wherein stationary model generally underestimated flood quantiles in more recent records. Šraj *et al.* (2016) concluded that the unjustified assumption of stationarity in FFA could lead to an under-estimation of extreme floods and therefore suggests applying non-stationary models in FFA. Chen and Ramachandra (2002) examined the stationarity of data series from four hydrological data sources *i.e.* streamflow, temperature, precipitation, and Palmer's Drought Severity Index (PDSI) series between 1900 and 1980 in the mid-western United States including Iowa, Indiana, Michigan and Wisconsin. Chen and Ramachandra (2002) concluded that, *inter alia*, more streamflow and PDSI data series were identified as non-stationary compared to temperature and precipitation data series, and only two periods of time between 1900 and 1980 were commonly identified as non-stationary across all four data sources indicating a difference in the detection of stationarity between hydrological data series. The detection of stationarity of streamflow data has a further complexity as streamflow discharges are dependent on climatic factors and other variables such as antecedent soil moisture and landuse change which is difficult to predict under future conditions (Bezak *et al.*, 2015).

Streamflow gauge network density is another important consideration in DFE as information such as hydrological response patterns can be transferred from gauged to ungauged catchments (Hrachowitz *et al.*, 2013). Despite this importance, Lebecherel *et al.* (2016) states that the impact of hydrometric data density is not a prioritised concern in literature and efforts to assess the impacts of gauge density are necessary. Lebecherel *et al.* (2016) tested the regionalisation efficiency, which is a measure used to evaluate the robustness of the regionalisation method, of the GR4J model on 609 catchments using the Random Hydrometrical Reduction (HRand) method and the Hydrometric Desert Method (HDes) of gauge reduction. Lebecherel *et al.* (2016) defined robustness as the degree at which the performance of the regionalisation method degrades when the hydrometric network becomes sparser from either gauge reduction method. HRand aims to randomly reduce the available gauge network of donor catchments by 10% to 90%. Furthermore, the number of donor catchments remain the same but are located on average further from the receiver catchment (Lebecherel *et al.*, 2016). HDes aims to progressively exclude the closest donor catchments (Lebecherel *et al.*, 2016). Two conclusions were drawn from this study *i.e.* (a) a decrease in regionalisation efficiency occurred when applying both the HRand and HDes methods, and (b) there is a more abrupt decrease in regionalisation

efficiency when applying the HDes method compared to the HRand method. Lebecherel *et al.* (2016) recommends using the HDes method to assess regionalisation efficiency to generate worst-case scenario outcomes.

Regarding spatial proximity of gauges, Tobler's (Tobler, 1970) first law of Geography states that "everything is related to everything else, but near things are more related than distant things" which is particularly important in the context of gauge proximity in a gauged network as factors such as, *inter alia*, climate, geology, soil, topography and land cover vary over space (Lebecherel *et al.*, 2016). Pool *et al.* (2019) highlighted that spatial proximity of gauges within a relatively dense streamflow gauging network is an important consideration accounting for relevant climate and catchment related attributes that influence hydrograph generation.

2.2 Outlier Events and Outlier Detection Methods

Raw at-site flow data are usually associated with several common issues such as the exceedance of discharge tables and presence of outlier events which requires data screening and pre-processing before use (Van Vuuren *et al.*, 2013; England Jr *et al.*, 2019). Data screening and pre-processing requires skill and experience, and is often the responsibility of the analyst to further examine data afterwards (Asikoglu, 2017). Data screening and pre-processing within this study addresses LO and HO events in Section 2.2.1 and Section 2.2.2 respectively, and outlier detection methods are reviewed in Section 2.2.3.

2.2.1 Low outlier events

Outlier events are data points which significantly depart from the trend of the remaining dataset (Lamontagne *et al.*, 2013; Lamontagne *et al.*, 2016; England Jr *et al.*, 2019). LOs are significantly small events which may be the result of catchment characteristics such as evaporation exceeding annual rainfall or channel infiltration (Paretti *et al.*, 2014; Lamontagne *et al.*, 2016; England Jr *et al.*, 2019).

In particular, Potentially Influential Low Flows (PILF) are LOs which affects sample statistics such as the mean, standard deviation and coefficient of skewness, and results in biased parameter estimates (Asikoglu, 2017; England Jr *et al.*, 2019). PILFs may also distort the exceedance probabilities of large events and therefore the detection and treatment of PILFs allows for an improved DFE (Lamontagne *et al.*, 2016). The Log-Pearson Type III (LP3) and Log-Normal (LN) PD are sensitive to PILFs whereas the Generalised Extreme Value (GEV)

PD is not as sensitive (Plavšić *et al.*, 2014). PILFs may be managed in numerous ways which include: (a) selecting a fitting technique which places less weight on small events, such as the Linear Moments (Hosking and Wallis, 1997; Rowinski *et al.*, 2002) and real-space Method Of Moments (MOM) techniques, (b) the use of a mixed distribution, and (c) adoption of a Peaks Over Threshold (POT) method (Lamontagne *et al.*, 2016). However, effort should first be on the detection of PILFs (Lamontagne *et al.*, 2016).

Graphical and statistical methods were developed through numerous studies to detect LOs (Whitacre *et al.*, 2006; Asikoglu, 2017). Grubbs and Beck (1972) developed the Grubbs-Beck Test (GBT) which is prescribed in Bulletin 17B for use in the USA. A regression analysis to detect outliers was also used and highlighted the importance of detecting outliers in a data series (Pegram, 1997). The Spatial Consistency test was developed to detect outliers in rainfall gauge measurements which can further be verified using remotely sensed data (Kondragunta, 2001). The Multiple Grubbs-Beck Test (MGBT) was developed to overcome the shortfall of the GBT, *i.e.* it's poor performance in detecting multiple LOs (Cohn *et al.*, 2013) and is prescribed by Bulletin 17C for use in the USA (England Jr *et al.*, 2019).

Asikoglu (2017) applied the Modified Z-Score method (MZS) (Iglewicz and Hoaglin, 1993), Quality Control (QC) test (Kondragunta, 2001), BoxPlot method (BP), GBT and the Stedinger test (Stedinger *et al.*, 1993) for the detection of outliers. Each of these five methods assign a critical value to each observation. An observation is classified as an outlier if it's assigned critical value is above or below a predetermined threshold value. Results from Asikoglu (2017) show that: (a) each of the above five methods show different precisions in the detection of outliers, (b) the BP method was the most precise method followed by the MZS, and (c) the GBT and Stedinger Tests flagged the least number of outliers compared to the other tests. Rahman *et al.* (2014) assessed the performance of the GBT and MGBT for six gauging stations. For three of these stations the GBT did not detect any PILF's whereas the MGBT detected 46% to 57% of the annual maximum flood peaks as PILF's (Rahman *et al.*, 2014). For the other three stations, the GBT detected one PILF's in each station whereas the MGBT identified 45% to 51% of the events as PILF's (Rahman *et al.*, 2014). For the six stations, there was an estimated 61% difference in flood quantile estimates for these two methods. It was concluded that the MGBT be used rather than the GBT (Rahman *et al.*, 2014). Van der Spuy and Rademeyer (2018) recommend using the standardised Z-Score method for detecting outliers in

South Africa, however, treatment of identified outliers is under the discretion of the analyst. Details of the BP, standardised Z-score, MZS and MGBT are provided in Section 2.2.3.

2.2.2 High outlier events

HO events are large magnitude events which significantly depart from the trend of the remaining dataset (Costa and Jarrett, 2008; England Jr *et al.*, 2019). HO events provide valuable information and are of interest when estimating design event magnitudes and frequency as such events have a low probability of exceedance and have a direct influence on the fitting of frequency distributions (Costa and Jarrett, 2008; England Jr *et al.*, 2019). HOs can be identified using the BP method, MZS, regional flood-peak envelops, time series plots or flood peak ratios (Costa and Jarrett, 2008; Asikoglu, 2017). It is recommended that HOs be retained and used in frequency analysis however judgement by an analyst is ultimately required on whether to include or exclude HOs from further analysis (England Jr *et al.*, 2019)

2.2.3 Outlier detection methods

Numerous outlier detection methods are available as described in Section 2.1.1, however, only the BP, MGBT, MZS and standardised Z-Score are detailed in this section. The BP and MZS were chosen due to their detection performance as indicated by Asikoglu (2017) and the relative ease of use. The MGBT was also chosen due to detection performance indicated by Rahman *et al.* (2014) and is part of the guidelines for FFA (England Jr *et al.*, 2019). The standardised Z-Score was chosen as it is the suggested method by the DWS (Van der Spuy and Rademeyer, 2018).

2.2.3.1 BoxPlots

The length of the box is first calculated, which is the difference between the third (Q_3) and first (Q_1) quartiles of the dataset, *i.e.* the Inter-Quartile Range (IQR). Any observation greater than 1.5 box lengths from the sample's maximum observation or less than 1.5 box lengths from the sample's minimum observation is classified as an outlier (Asikoglu, 2017). Figure 2.6 depicts the classification of an outlier using the box plots method.

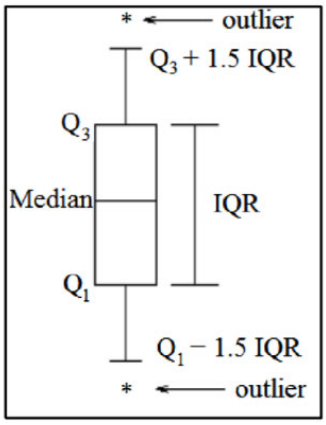

Figure 2.6 Outlier classification using the BoxPlot method (Asikoglu, 2017)

2.2.3.2 Standardised Z-Score

Van der Spuy and Rademeyer (2018) recommends using the standardised Z-Score in SA. A Z-Score for each observation in an Annual Maximum Series (AMS) is calculated using Equation 2.1 and observations with Z-Scores greater than 3 or less than -3 are classified as outliers.

$$Z_i = \frac{x_i - \bar{x}}{s} \quad (2.1)$$

where:

Z_i = Z-Score of the i^{th} observation,

x_i = i^{th} observation,

$\bar{x}$ = mean of sample AMS, and

S = standard deviation of the sample AMS.

2.2.3.3 Modified Z-Score

The standardized process aims to convert each data observation into a unit of standard deviation in order to express the distance of each observation from the mean of the sample in comparable units. This is problematic as outliers affect both the mean and standard deviation, hence calculated standardised Z-Scores are often not larger than the pre-determined threshold, therefore they will not be identified as outliers (Iglewicz and Hoaglin, 1993). The standardised Z-Score was therefore modified by Iglewicz and Hoaglin (1993) using robust statistics, which have a greater resistance to the influence of outliers than standard statistics, in the Z-Score formula. The MZS test is shown in Equation 2.2. An outlier is defined as having an absolute modified Z-Score greater than 3.5 (Iglewicz and Hoaglin, 1993).

$$Z_i = 0.675(x_i - x_{0.5})/MAD \quad (2.2)$$

where:

Z_i = Z-score,

x_i = observation of interest,

$x_{0.5}$ = sample median, and

MAD = calculated using Equation 2.3.

$$MAD = \frac{1}{N} \sum_{i=1}^{N} \left| x_i - x_{0.50} \right| \quad (2.3)$$

where:

MAD = Median Absolute Deviation, and

N = sample size.

2.2.3.4 Multiple Grubbs-Beck test

The Multiple Grubbs-Beck Test (MGBT) is a generalization of the GBT (Lamontagne *et al.*, 2013). The GBT, as shown in Equation 2.4, classifies any event less than X_{crit} as a LO (Grubbs and Beck, 1972).

$$X_{crit} = \hat{\mu} - K_n\hat{\sigma} \quad (2.4)$$

where:

X_{crit} = low outlier threshold,

$\hat{\mu}$ = sample mean,

K_n = 10% significance value for an independent sample of *n* normal variates, and

$\hat{\sigma}$ = sample standard deviation.

The MGBT is preferred to the GBT because it is able to consistently identify multiple LOs and it is shown to improve the extreme quantile estimators of negatively skewed data (Lamontagne *et al.*, 2013). The MGBT involves sequentially evaluating a test statistic, $\hat{\omega}$ shown in Equation 2.5, for each flood peak in a dataset (Lamontagne *et al.*, 2013; England Jr *et al.*, 2019). The MGBT systematically tests the hypothesis that *k* samples in the left-hand tail are from the same sample of the remaining population which is normally distributed (Lamontagne *et al.*, 2013; England Jr *et al.*, 2019).

$$\widehat{\omega}_{[k,\mathrm{N}]} = \frac{X_{[\mathrm{k},N]} - \hat{\mu}_{\mathrm{k}}}{\widehat{\sigma_{\mathrm{k}}}} \quad (2.5)$$

where:

$\widehat{\omega}.$ = test statistic,

k = kth smallest flood peak,

N = sample size,

$X_{[\mathrm{k},N]}$ = the logarithm of the kth smallest flood peak in an ordered sample,

$\hat{\mu}_{\mathrm{k}}$ = partial mean for all flood peaks larger than $X_{[\mathrm{k},N]}$, and

$\hat{\sigma}_{\mathrm{k}}$ = partial standard deviation for all flood peaks larger than $X_{[\mathrm{k},N]}$.

The critical value, η, shown in Equation 2.6, is then compared to the test statistic $\widehat{\omega}$. The critical value is the probability given a null hypothesis (H_0) of obtaining a value of $\widehat{\omega}_{[k,N]}$ that is less than that observed in the sample (Lamontagne *et al.*, 2013; England Jr *et al.*, 2019). The critical value can be calculated by integrating the distributions of $Z_{[\mathrm{k},N]}$, $\hat{\mu}_{Z,k}$ and $\hat{\sigma}_{Z,k}$ (Lamontagne *et al.*, 2013; England Jr *et al.*, 2019).

$$p\left[\widehat{\omega}_{[k,N]} < \eta\right] = p\left[\frac{Z_{[\mathrm{k},N]} - \hat{\mu}_{Z,k}}{\hat{\sigma}_{Z,k}} < \eta\right] \quad (2.6)$$

where:

p = probability,

$Z_{[\mathrm{k},N]}$ = kth-order statistic,

N = sample size,

$\hat{\mu}_{Z,k}$ = partial mean for all flood peaks larger than $Z_{[\mathrm{k},N]}$, and

$\hat{\sigma}_{Z,k}$ = partial standard deviation for all flood peaks larger than $Z_{[\mathrm{k},N]}$.

The MGBT consists of two steps, *i.e.* sweeping outward and sweeping inward. Sweeping outward involves iterating starting at the median and progressing down to the smallest observation (Lamontagne *et al.*, 2013; England Jr *et al.*, 2019). Each observation, $X_{[\mathrm{k},N]}$, is tested by comparing $\widehat{\omega}_{[k,N]}$ to a predetermined significance level ($\alpha_{out} = 0.005$). If $\widehat{\omega}_{[k,N]} < \alpha_{out}$, then the observation is classified as a LO (Lamontagne *et al.*, 2013; England Jr *et al.*, 2019). Breaks in the data are identified using outward sweeping which implies the presence of several PILF's (Lamontagne *et al.*, 2013; England Jr *et al.*, 2019).

Sweeping inward involves iterating inward to the median starting with the smallest observation (Lamontagne *et al.*, 2013; England Jr *et al.*, 2019). Each observation, $X_{[\mathrm{k},N]}$, is tested by

comparing *p(k:n)*, probability (*p*) of k^{th} observation in data sample (N) to a significance level ($\alpha_{in} = 0.10$). If $p(k:n) < \alpha_{in}$, then the observation is classified as a LO (Lamontagne *et al.*, 2013; England Jr *et al.*, 2019). The value of the significance levels α_{out} and α_{in} are prescribed through extensive research, testing and evaluation by Lamontagne *et al.* (2013), Lamontagne *et al.* (2016) and Cohn *et al.* (2013). The MGBT is, however, prone to the mechanisms referred to as masking and swamping. Regarding masking, potential LOs may not be detected in the context of other small values which results in an under-detection of potential LOs (Cohn *et al.*, 2013; Lamontagne *et al.*, 2016). Regarding swamping, the smallest observation may cause the second and subsequent small observations to be identified as outliers which results in an over-detection of events as LOs (Cohn *et al.*, 2013; Lamontagne *et al.*, 2016). The prescribed significance value accounts for potential masking and swamping, however, these mechanisms may still occur. PILF's which have been subjectively identified by hydrologists have also been successfully detected using the MGBT (Lamontagne *et al.*, 2013; England Jr *et al.*, 2019).

2.3 Chapter Discussion and Conclusions

The use of data of sufficient quantity and quality from available networks may also increase the confidence of DRE and DFE and therefore enhance the value of information as understood by the public and private sectors (Van Bladeren *et al.*, 2007). However, there is a decline of hydrological monitoring in SA (Pitman, 2011). The decline in rainfall monitoring networks has been shown to have severe impacts on water resource management such as the estimation of irrigation demands and losses from reservoirs and wetlands (Pitman, 2011) and the lack of overlapping data periods between observed and remotely sensed data is a concern (Hughes, 2008). Gauge density is important to capture rainfall characteristics from spatially and temporally varying events (St-Hilaire *et al.*, 2003). Increased gauge density has improved simulated streamflow estimates, areal estimates of rainfall, MAP estimates and reduced under-estimation of cumulative rainfall (Krajewski *et al.*, 2003; St-Hilaire *et al.*, 2003; Bárdossy and Das, 2008; Xu *et al.*, 2013). Dense rainfall gauge networks also provide meaningful data for groundtruthing and calibrating satellite and lidar based remotely sensed data (Sun *et al.*, 2000; Pegram *et al.*, 2016). Xu *et al.* (2013) highlighted the importance of rainfall gauge location in estimating runoff and St-Hilaire *et al.* (2003) concluded that the quality of data is further improved with well-located rainfall stations and not only by a denser network.

Streamflow gauge network density is an important consideration in DFE due to possible information transfer from gauged to ungauged catchments (Hrachowitz *et al.*, 2013).

Lebecherel *et al.* (2016) confirm the importance of streamflow gauge network density and showed a decrease in regionalization efficiency when using a reduced gauged network density. Pool *et al.* (2019) further highlighted that spatial proximity of gauges within a relatively dense streamflow gauging network is an important consideration as it accounts for relevant attributes influencing hydrograph generation.

Regarding record length, Boughton (2007) concluded that rainfall record length has an influence on estimated rainfall characteristics and on the performance of rainfall-runoff modelling. Dent *et al.* (1987) concludes that the minimum length of record required varies between regions. It is, however, important that an analyst uses a record length sufficiently long enough to cover quasi-periodic fluctuations irrespective of when the observations were recorded (Dyler and Tyson, 1977). Using rainfall datasets with shorter record lengths has resulted in over-estimated simulated streamflow and an increase in the errors of the model estimates (Boughton, 2007), however, these datasets are valuable for decision making in ungauged catchments (Pool *et al.*, 2019). The validity of the assumption of stationarity, irrespective of available record length and period of record in FFA, has been questioned (Zhang *et al.*, 2001; Koutsoyiannis *et al.*, 2009). Šraj *et al.* (2016) concluded that an unjustified assumption of stationarity in FFA could lead to an under-estimation of extreme floods, and furthermore, the detection of stationarity of streamflow data has an additional complexity in that streamflow discharges are dependent on climatic factors and other variables such as antecedent soil moisture which is difficult to predict under future conditions (Bezak *et al.*, 2015).

Data screening and quality control are necessary to ensure reliable DFE and are a common practice as reported in the international literature, e.g. in USA, Australia, UK and European countries. Guidelines for data screening and quality control for FFA includes the detection and treatment of outliers prior to FFA (England Jr *et al.*, 2019). There are no prescribed guidelines for data screening and quality control of rainfall or streamflow data for regular practice in SA, apart from the standardised Z-Scores in outlier detection (Van der Spuy and Rademeyer, 2018) which has been shown to be problematic (Iglewicz and Hoaglin, 1993).

Outliers are events which significantly depart from the trend of the remaining dataset (Lamontagne *et al.*, 2013; Lamontagne *et al.*, 2016; England Jr *et al.*, 2019). LOs affect sample statistics resulting in biased parameter estimates (Asikoglu, 2017; England Jr *et al.*, 2019) and a distortion of exceedance probabilities of large design events(Lamontagne *et al.*, 2013;

England Jr *et al.*, 2019). HOs have a low probability of exceedance therefore providing valuable information in the estimation of design events (Costa and Jarrett, 2008; England Jr *et al.*, 2019). With outlier detection being included in many international guidelines and with numerous available outlier detection methods, the performance of various methods under South African conditions needs to be investigated.

It is thus necessary to assess the impact of outliers on DRE and DFE and to assess the performance of outlier detection methods to determine if outlier detection should be recommended for regular practice in SA. In addition, the impact of a declining hydrological monitoring network, in terms of gauged density and record length on DRE and DFE, needs to be evaluated. This will aid in assessing if additional national resources should be directed towards maintaining and improving the hydrological monitoring network in SA.

3. STUDY AREA AND DATA COLLATION

A total of six catchments in SA spanning three climatologically varying regions and having varying catchment area were selected for use in this study. The climatologically varying regions are the Eastern Summer Coastal (ESC), Southern Winter Coastal (SWC) and Northern Interior (NI) which were used by (Gericke, 2015). ESC is within the eastern escarpment (Alexander, 2002) and is predominantly characterised with all year and/or summer rainfall (Gericke, 2015). NI is within the subtropical lowveld and highveld region (Alexander, 2002) and is characterised as a summer rainfall region (Gericke, 2015). SWC is within the Mediterranean region (Alexander, 2002) and is characterised as a winter rainfall region (Gericke, 2015). In this study, a catchment is defined as the watershed upstream of a streamflow gauge. Catchments were labelled according to the DWS (2011) flow-gauging station number located at the outlet of the catchment. Catchment Areas (A) varied between small ($A \leq 100$ km^2), medium ($100\ km^2 < A \leq 1\ 000\ km^2$) and large catchments ($A > 1000\ km^2$).

Streamflow gauges used by Nathanael (2015) which were completely within the three climatologically varying regions were shortlisted. Thereafter, six streamflow gauges with record lengths of greater than 40 years and with no AMS records having missing or suspect flags were selected. Catchment boundaries for these six gauges were delineated using QGIS 10.3.1. Table 3.1 provides a summary of attributes of the streamflow gauges used in this study.

Table 3.1 Streamflow gauges and summary attributes

Streamflow Gauge ID	**Gericke (2015) Climate Zone**	**Record Length (Years)**	**Area (km^2)**
U2H013	ESC	46	295.70
V2H004	ESC	40	269.13
A2H012	NI	54	2579.65
A6H011	NI	40	73.66
G1H008	SWC	43	396.07
H7H004	SWC	42	25.60

Rainfall stations from the Lynch (2004) database was used in this study. A driver rainfall gauge within each catchment was selected based on record length (minimum of 40 years of unedited observed data, *i.e.* no patching or infilling), Mean Annual Precipitation (MAP), altitude and proximity to the catchment centroid. Table 3.2 provides a summary of attributes of the driver rainfall gauges used in this study. Details of the selected streamflow gauges and corresponding driver rainfall gauges are provided in Table A.1 and Table A.2 in Appendix A. The selected streamflow gauges and driver rainfall gauges within the climatologically varying regions are shown in Figure 3.1.

Table 3.2 Driver rainfall gauges and associated attributes

Rain Gauge Number	**Respective Streamflow Gauge ID**	**Gericke (2015) Climate Zone**	**Record Length (Years)**
0239097	U2H013	ESC	46
0268640	V2H004	ESC	82
0476031	A2H012	NI	49
0589670	A6H011	NI	82
0042227	G1H008	SWC	99
0025414	H7H004	SWC	74

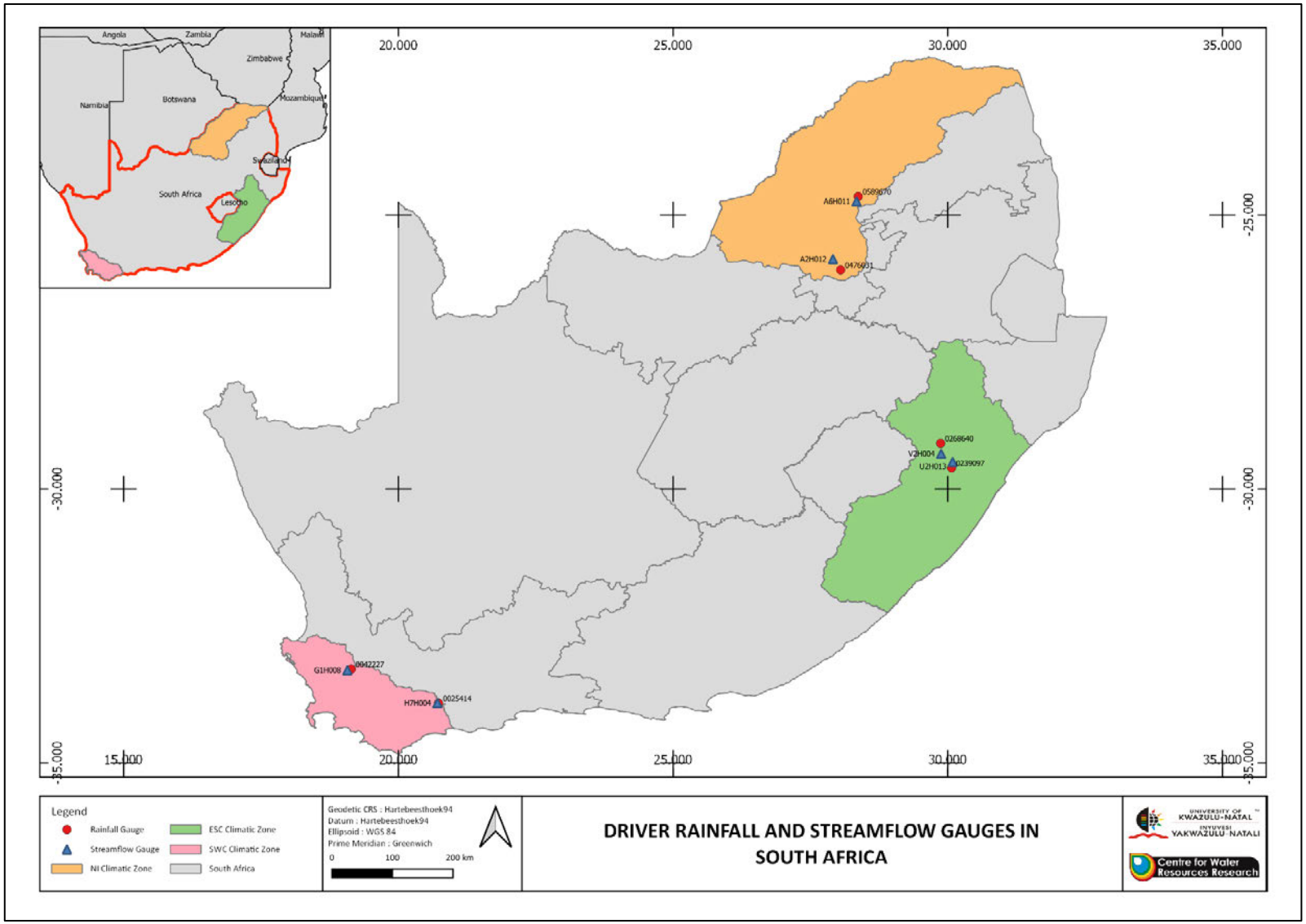

Figure 3.1 Location of selected driver rainfall and streamflow gauges in South Africa used in study

The AMS extracted from rainfall and streamflow observed data were then screened for outlier events prior to further analysis using a simple visual inspection, *i.e.*, a time series plot, and the standardised Z-Score test as detailed in Section 2.2.3. As an example, Figure 3.2 shows the AMS plot for driver Rainfall Gauge 0239097 with outliers identified using the standardised Z-Score test.

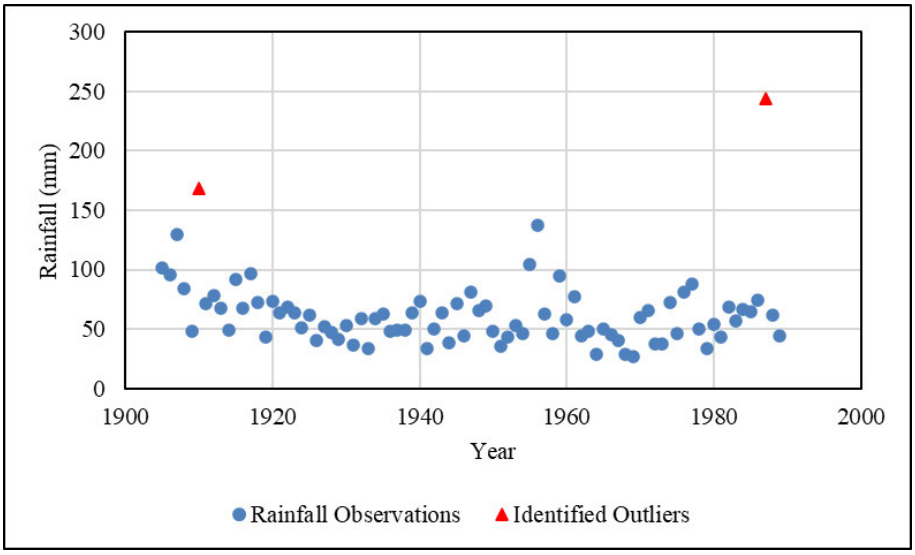

Figure 3.2 AMS plot for driver Rainfall Gauge 0239097

Events classified as outliers from both the visual inspection of time series plot and standardised Z-Score method were then compared to observations from neighbouring gauges within the same year, assuming that these potential outlier events were caused by the same event. Events were then identified as outliers using the combination of time series plot, standardised Z-Scores and the comparison to neighbouring gauges. A maximum of 2% of observations were identified as outliers within rainfall datasets throughout all study catchments and 3% within streamflow datasets. Identified outlier observations were removed from the AMS and it was thereafter assumed that no outlier events were present in the dataset.

.

4. IMPACT OF OUTLIER EVENTS ON DESIGN RAINFALL AND FLOODS ESTIMATED USING DIFFERENT PROBABILITY DISTRIBUTIONS

The impact of outlier events on DRE and DFE in SA are detailed in this chapter. This analysis was performed using at-site observed and synthetically generated AMS data series for selected PDs, as described in Section 4.1 and Section 4.2, respectively. A chapter summary and conclusion is provided in Section 4.3. A schematic of the structure of Chapter 4 is shown in Figure 4.1.

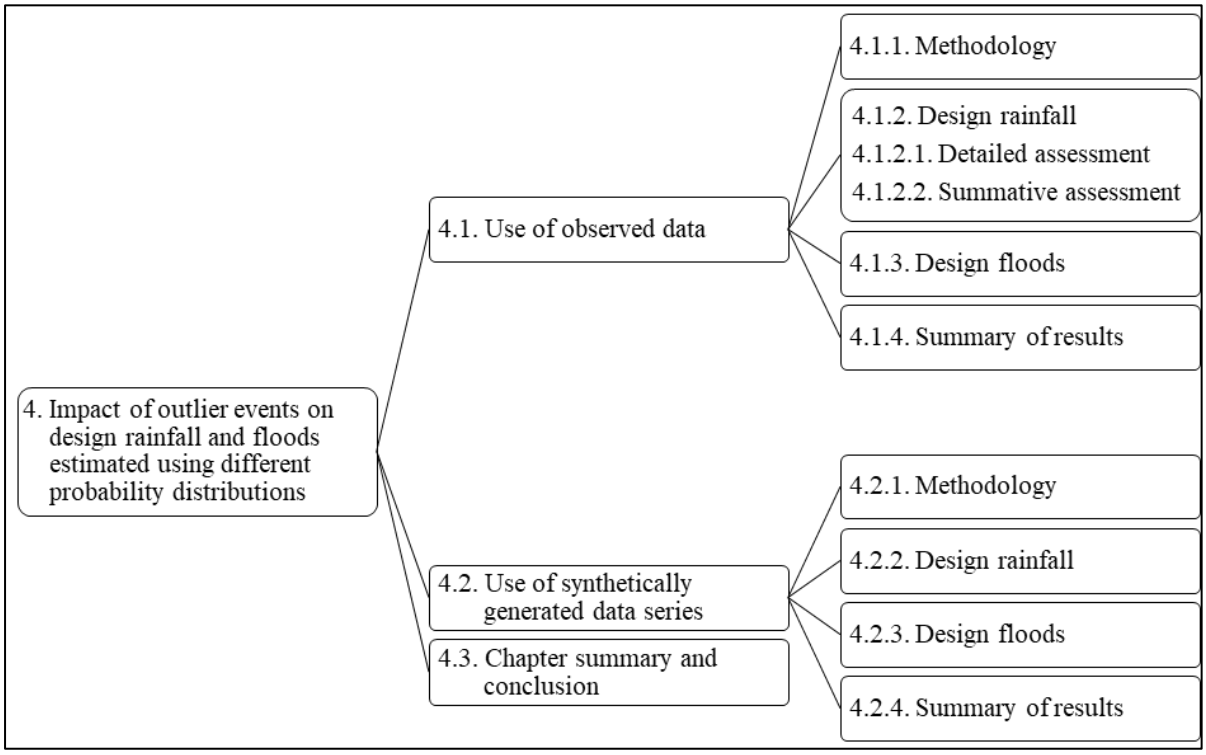

Figure 4.1 Schematic of the structure of Chapter 4

4.1 Use of Observed Data

This section details the methodology, results and discussion to assess the impact of outlier events on estimated DRE and DFE using observed datasets.

4.1.1 Methodology

The 1-, 2-, 5-, 10-, 20-, 50-, 100- and 200-year return period design rainfall and flood events were estimated by the GEV, Generalised Pareto (GPA), 3 parameter Kappa (Kappa), LN, and

LP3 PDs from an initial observed AMS of rainfall and streamflow data. The initial observed AMS refers to observed data after data screening and pre-processing as detailed in Chapter 3. The *evd* (Stephenson, 2018), *extRemes* (Gilleland, 2018) and *extraDistr* (Wolodzko, 2018) packages were used in R-Studio to estimate design rainfall and floods. Linear Moments (L-Moments) were chosen to estimate the required parameters due to its robust nature which provide an unbiased parameter estimation in the presence of outliers (Hosking and Wallis, 1997; Rowinski *et al.*, 2002). Version 2.6 of the *lmom* package was used to estimate L-Moments within R-Studio (Hosking, 2017).

The above PDs were chosen as:

(a) Krasovskaia *et al.* (2001) advocates the use of the GPA and GEV PDs for DFE.
(b) Görgens (2007) advocates the use of the GEV and LP3 PDs for DFE in SA.
(c) Haile (2011) applied the GPA PD for DFE in SA.
(d) The GEV, LP3, GPA and Gumbel Maximum PDs were found to be the best fit in the UK and Australia by Rahman *et al.* (2013).
(e) The GEV PD is used in national guidelines in Malaysia and Indonesia (Zalina *et al.*, 2002; Liu *et al.*, 2015).
(f) Kjeldsen *et al.* (2017) advocates the Kappa PD for regional FFA.
(g) Van der Spuy and Rademeyer (2018) identified the LNO, LP3 and GEV PDs as the suitable PDs for FFA in SA.
(h) Rulfova *et al.* (2016) showed that the GEV PD is applicable for RFA.
(i) The GEV PD was identified as the most appropriate for design rainfall estimation in SA (Smithers, 1996; Smithers and Schulze, 2000b; Smithers and Schulze, 2000a).
(j) Calitz and Smithers (2020) concluded that LP3 PD has the largest uncertainty bands in FFA with the GPA and GEV PD having the smallest uncertainty bands. Calitz and Smithers (2020) further advocated the use of the GPA PD for FFA on a national scale in SA.

Synthetically generated LOs and HOs were thereafter substituted into each of the initial observed rainfall and streamflow AMS dataset. The number of substituted synthetically generated outliers were based on three fractions of the AMS record length *i.e.*, equivalent to 1%, 2% and 3% as shown in Table 4.1. A maximum substitution of 3% was chosen on the basis of the maximum percentage of outliers identified from the initial observed AMS as detailed in Chapter 3. This substitution process was used to provide consistency in the study and to ensure that the initial observed AMS dataset length remained unchanged.

Table 4.1 Number of generated synthetic outliers substituted in observed rainfall and streamflow AMS data

		Record Length (Year)	Number of Substituted Outliers		
			1% of AMS	2% of AMS	3% of AMS
Rainfall Gauge ID	0476031	49	0	1	2
	0589670	82	1	2	3
	0042227	99	1	2	3
	0025414	74	1	1	2
	0239097	85	1	2	3
	0268640	82	1	2	3
Streamflow Gauge ID	A2H012	54	0	1	2
	A6H011	40	0	1	2
	G1H008	43	0	1	2
	H7H004	42	0	1	2
	U2H013	46	0	1	2
	V2H004	40	0	1	2

The magnitude of the synthetically generated LOs were calculated as fractions, *i.e.,* 1%, 2% and 3%, of the estimated 1-year return period event and as multiples, *i.e.,* 150%, 160% and 170%, of the 100-year return period event for synthetically generated HOs. The outlier scenarios are explained in Table 4.2, and an example of the substitution process is provided in Table 4.3. The 2-, 5-, 10-, 20-, 50-, 100- and 200-year return period rainfall and flood events were estimated by the GEV, GPA, Kappa, LN and LP3 PDs with observed data containing LOs and HOs, and thereafter compared to design events estimated from datasets without substituted outliers.

Table 4.2 Outlier substitution process and scenarios when using observed data

Dataset Type	Scenario ID	Explanation
Observed Datasets with LOs	*Obs.L1*	All observations ≤ 1st percentile value in observed AMS are substituted by synthetic LOs equivalent to 1% of the AMS (Table 4.1)
	Obs.L2	All observations ≤ 2nd percentile value in observed AMS are substituted by synthetic LOs equivalent to 2% of the AMS (Table 4.1)
	Obs.L3	All observations ≤ 3rd percentile value in observed AMS are substituted by synthetic LOs equivalent to 3% of the AMS (Table 4.1)
Observed Datasets with HOs	*Obs.H1*	All observations ≥ 99th percentile value in observed AMS are substituted by synthetic HOs equivalent to 1% of the AMS (Table 4.1)
	Obs.H2	All observations ≥ 98th percentile value in observed AMS are substituted by synthetic HOs equivalent to 2% of the AMS (Table 4.1)
	Obs.H3	All observations ≥ 97th percentile value in observed AMS are substituted by synthetic HOs equivalent to 3% of the AMS (Table 4.1)

Table 4.3 Example of the outlier substitution process at Rainfall Gauge 0042227

Rank	**Input (mm)**	***Obs.L1* (mm)**	***Obs.L2* (mm)**	***Obs.L3* (mm)**	***Obs.H1* (mm)**	***Obs.H2* (mm)**	***Obs.H3* (mm)**
1	4.50	0.33*	0.33*	0.33*	4.50	4.50	4.50
2	22.50	22.50	0.65*	0.65*	22.50	22.50	22.50
3	22.90	22.90	22.90	0.98*	22.90	22.90	22.90
4	23.90	23.90	23.90	23.90	23.90	23.90	23.90
…..							
96	83.50	83.50	83.50	83.50	83.50	83.50	83.50
97	83.80	83.80	83.80	83.80	83.80	83.80	119.79*
98	84.00	84.00	84.00	84.00	84.00	119.79*	127.78*
99	88.00	88.00	88.00	88.00	119.79*	127.78*	135.77*

* Substituted synthetic outliers

It was assumed that the initial observed AMS resulted in the most accurate and representative design events compared to events estimated with substituted outliers. In addition, it was

assumed that all the PDs were appropriate to use in all initial observed AMS and AMS with substituted outliers. It was also assumed that substituted outlier events belonged to the respective sample dataset populations. It is acknowledged that any uncertainties in the initial observed data have not been taken into consideration, and a violation of the assumption that all the selected PDs fitted the initial observed data will result in inherent errors in the estimated design events.

Four statistics were computed to assess the impact of outliers *i.e.*, Mean Relative Differences (*MRD*), Mean Absolute Relative Differences (*MARD*), Nash Sutcliffe Efficiency (*NSE*) and Percent Bias (*PBIAS*). Relative Differences (*RD*) were calculated using Equation 4.1. *MRD* were calculated as the arithmetic mean of the *RD* across all return periods per catchment and PDs. A positive *RD* indicates an over-estimation of design events from datasets with substituted outliers and the converse is true for a negative *RD*. *MARD* was calculated as the arithmetic mean of the absolute *RD* across all return periods per catchment and PDs.

$$RD_{outl.} = \frac{[E_{Si,PD,T,D} - E_{0,PD,T,D}]}{E_{0,PD,T,D}} \times 100 \quad (4.1)$$

where:

$RD_{outl.}$ = RD calculated between design rainfall or flood events estimated using a dataset scenario (*Obs.L1* to *Obs.H3*) with outliers and the initial observed dataset (%),

E_{Si} = design rainfall or flood estimate from *Obs.L1* to *Obs.H3* (mm or $m^3.s^{-1}$),

PD = probability distributions (GEV, GPA, Kappa LN and LP3),

T = return period (2-, 5-, 10-, 20-, 50-, 100- and 200-year),

D = data type (observed or synthetic), and

E_0 = rainfall or flood estimated from initial observed dataset (mm or $m^3.s^{-1}$).

The *NSE* (Nash and Sutcliffe, 1970) calculated using Equation 4.2 quantifies the fit of estimated design events from an outlier scenario against the 1:1 line which represents a perfect fit scenario. *NSE* values range from negative infinity to 1, with *NSE* = 1 representing a perfect fit against the 1:1 line. *NSE* values are categorized as follows: $0.75 < NSE \leq 1.0$ indicates a very good fit against the 1:1 line, $0.65 < NSE \leq 0.75$ indicates a good fit, $0.50 < NSE \leq 0.65$ indicates a satisfactory fit, $NSE \leq 0.5$ indicates an unsatisfactory fit (Santhi *et al.*, 2001; Lim *et al.*, 2006; Moriasi *et al.*, 2007; Parajuli *et al.*, 2009).

$$NSE = \left(1 - \frac{\sum_{T=1}^{7}\left(E_{0,PD,T,D} - E_{S_i,PD,T,D}\right)^2}{\sum_{T=1}^{7}(E_{0,PD,T,D} - \bar{E}_D)^2}\right) \quad (4.2)$$

where:

NSE = statistic quantifying the fit of estimated design events against the 1:1 line, and

$\bar{E}$ = average rainfall or flood from initial observed dataset (mm or $m^3.s^{-1}$).

PBIAS calculated using Equation 4.3 indicates if design events from a scenario are constantly over- or under-estimated. A positive PBIAS value indicates an over-estimation whereas a negative PBIAS value indicates an under-estimation. *PBIAS* values are categorized as: PBIAS < ± 10% indicates a very good fit between estimated events, ± 10% < PBIAS ≤ ± 15% indicates a good fit, ± 15% < PBIAS ± ≤ 25% indicates a satisfactory fit and PBIAS > 25% indicates an unsatisfactory (Van Liew *et al.*, 2003; Singh *et al.*, 2004; Moriasi *et al.*, 2007; Archibald *et al.*, 2014).

$$PBIAS = \sum_{T=1}^{7}\left[\frac{\left(E_{S_i,PD,T,D} - E_{0,PD,T,D}\right)\times 100}{E_{0,PD,T,D}}\right] \quad (4.3)$$

The following two assessments were performed to evaluate the impact of outliers:

(a) Detailed assessment: Statistics (*MRD, MARD, NSE* and *PBIAS*) were computed per scenario (*Obs.L1* to *Obs.H3*) and for each PD for each catchment. This was to highlight the impact of outliers on design estimates per scenario within each catchment.

(b) Summative assessment: Computed statistics from the detailed assessment were averaged across scenarios with LOs (*Obs.L1, Obs.L2 and Obs.L3*) and across scenarios with HO (*Obs.H1, Obs.H2 and Obs.H3*) per PD for each catchment and thereafter averaged across all catchments. This provides an indication of the overall impact of LOs and HOs on DRE and DFE.

Design rainfall results are provided in Section 4.1.2 with design floods in Section 4.1.3.

4.1.2 Design rainfall

As an example, results for the detailed assessment for Rainfall Gauge 0042227 are presented in Section 4.1.2.1 with results for all rainfall gauges using observed data provided in Figure B1.1 to Figure B1.5 in Appendix B1. Results from the summative assessment are provided in Section 4.1.2.2. Results from the detailed assessment are used to inform the summative assessment, therefore a discussion of all of the results from the detailed assessment are not provided. For ease of reference, the four computed statistics used in the detailed

assessment are referred to as *MRD*, *MARD*, *NSE* and *PBIAS* whereas the four computed statistics used in the summative assessment are referred to as *Avg.$_s$ MRD*, *Avg.$_s$ MARD*, *Avg.$_s$ NSE* and *Avg.$_s$ PBIAS*.

4.1.2.1 Detailed assessment

As evident in Figure 4.2, design rainfall events are over-estimated in the presence of LOs when estimated by the LN PD as indicated by positive *MRD* and *PBIAS* values and under-estimated by the GEV, GPA and LP3 PDs as indicated by negative *MRD* and *PBIAS* values.

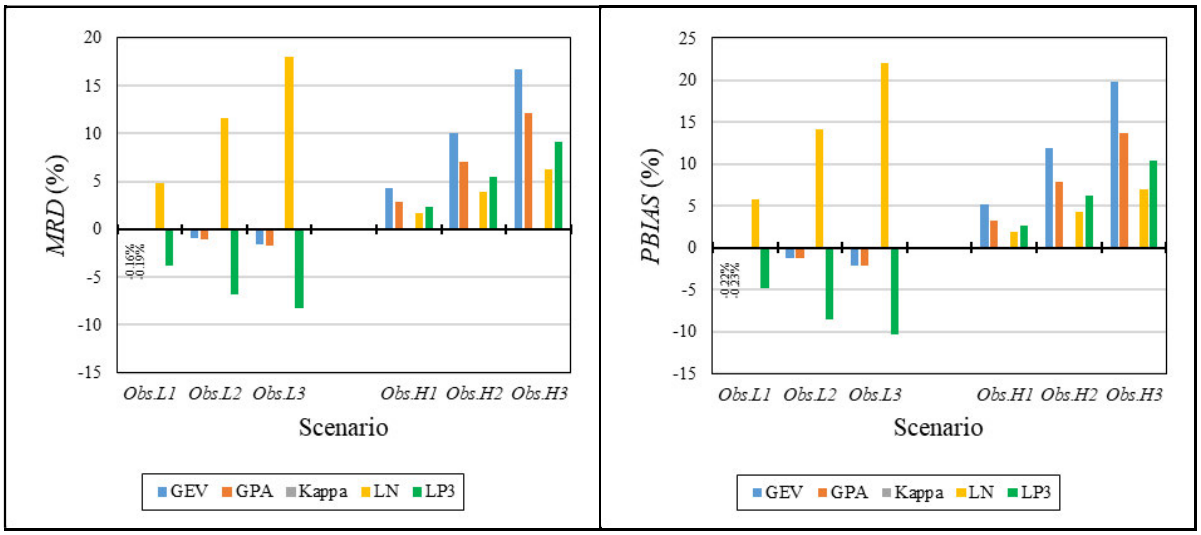

Figure 4.2 *MRD* and *PBIAS* for design rainfall events estimated using observed data for Rain Gauge 0042227 for LO (*Obs.L1*, *Obs.L2* and *Obs.L3*) and HO (*Obs.H1*, *Obs.H2* and *Obs.H3*) scenarios

It is indicated by *MARD* and *NSE* values, as shown in Figure 4.3, and *MRD* and *PBIAS* values, as shown in Figure 4.2, that events estimated by the LN PD are the most impacted in the presence of all LO scenarios. Design rainfall events estimated by the GEV and GPA PDs are impacted the least in the presence of LOs as indicated by the smallest *MARD* and largest *NSE* (Figure 4.3), and smallest *MRD* and *PBIAS* values (Figure 4.2). There is an increased impact on design rainfall estimates with an increased presence of LOs especially events estimated by the LN and LP3 PDs indicated by *MARD* and *NSE* values (Figure 4.3) and *MRD* and *PBIAS* values (Figure 4.2). Observed rainfall for Rainfall Gauges 0268640, 0042227, 0025414 and the streamflow gauge at Catchment V2H004 were unable to be fitted to the Kappa PD as estimated parameters were not consistent with the Kappa PD due to a violation of the defined parameter space (Hosking and Wallis, 1997).

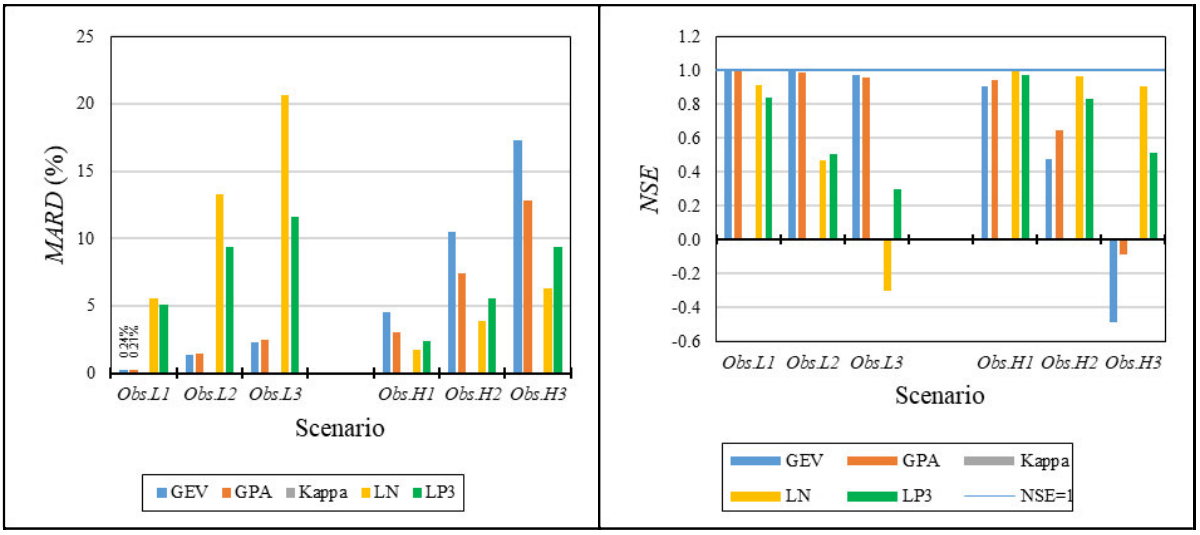

Figure 4.3 *MARD* and *NSE* for design rainfall events estimated using observed data for Rainfall Gauge 0042227 for LO (*Obs.L1*, *Obs.L2* and *Obs.L3*) and HO (*Obs.H1*,*Obs.H2* and *Obs.H3*) scenarios

The presence of HOs results in an over-estimation of design rainfall events across all PDs for Rainfall Gauge 0042227 as indicated by a positive *MRD* and *PBIAS* value as shown in Figure 4.2. Design rainfall events estimated by the GEV PD are the most impacted by the presence of HOs as indicated with the largest *MARD* and smallest *NSE* (Figure 4.3), and by the largest *MRD* and *PBIAS* (Figure 4.2). Design events estimated by the LN PD are least impacted by HOs as indicated by the smallest *MARD* and largest *NSE* (Figure 4.3), and by the smallest *MRD* and *PBIAS* (Figure 4.2).

4.1.2.2 Summative assessment

For the summative assessment, computed statistics were averaged across LO scenarios and across scenarios with HOs per PD for each catchment and thereafter averaged across all catchments to obtain averaged values, i.e. *Avg.$_S$ MRD*, *Avg.$_S$ MARD*, *Avg.$_S$ NSE* and *Avg.$_S$ PBIAS.*

Identical trends observed from the detailed assessment for Rainfall Gauge 0042227 are reflected in the summative assessment for LO and HO scenarios. Design rainfall events are generally over-estimated by the LN PD in the presence of LOs as indicated by a positive *Avg.$_S$ MRD* and *Avg.$_S$. PBIAS* value as shown in Figure 4.4. Design rainfall events are generally under-estimated by the GEV, GPA and LP3 PDs in the presence of LOs as indicated by the negative *Avg.$_S$ MRD* and *Avg.$_S$ PBIAS* values as shown in Figure 4.4. It is further noted that

events estimated by the LP3 PD are most under-estimated in the presence of LOs as indicated by the largest *Avg.s MRD* values as shown in Figure 4.4.

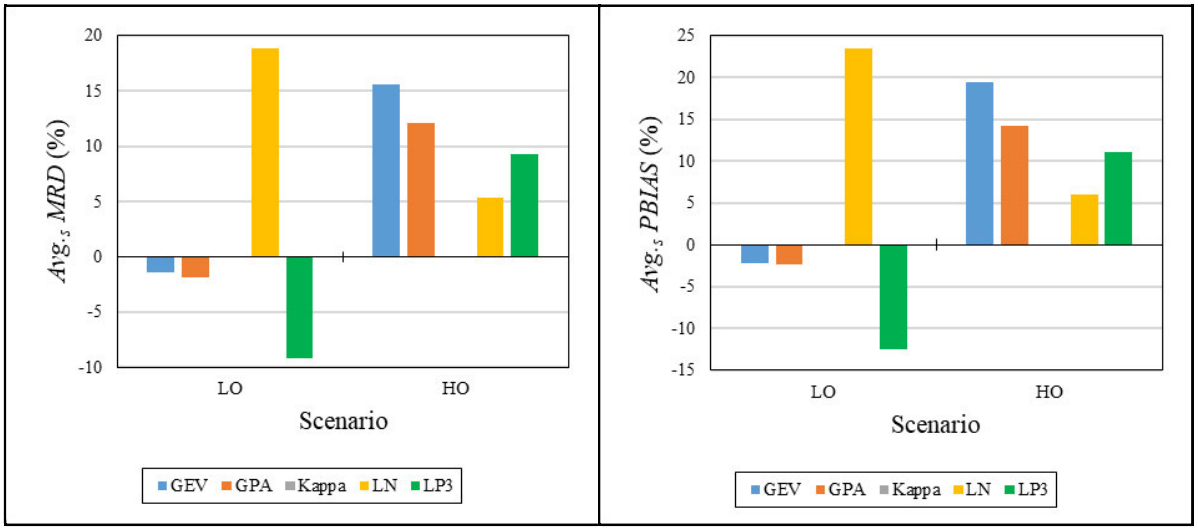

Figure 4.4 *Avg.s MRD* and *Avg.s PBIAS* for design rainfall events estimated using observed data with LOs and HOs

Design rainfall events estimated by the LN PD are the most impacted *i.e.* by up to 22% across catchments in the presence of LOs as indicated by the largest *Avg.s MARD* value and the smallest *Avg.s NSE* value as shown in Figure 4.5, and the largest *Avg.s MRD* and *Avg.s PBIAS* as shown in Figure 4.4. Design estimates by the GEV and GPA PDs were the least impacted to LOs *i.e.* by to 3% across catchments as indicated by the smallest *Avg.s MARD* value and the largest *Avg.s NSE* value (Figure 4.5), and by the smallest *Avg.s MRD* and *Avg.s PBIAS* (Figure 4.4). Based on the computed *Avg.s NSE* value (*Avg.s NSE* < 0.5), design rainfall events with LOs estimated by the LN and LP3 PDs are classified as unsatisfactory.

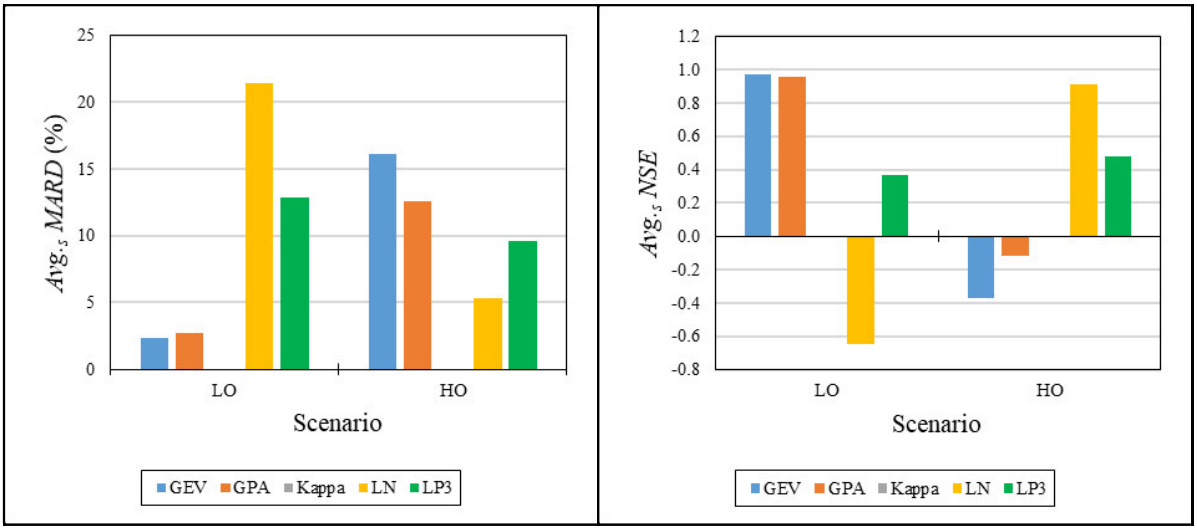

Figure 4.5 *Avg.s MARD* and *Avg.s NSE* for design rainfall events estimated using observed data with LOs and HOs

Design rainfall events are generally over-estimated by all PDs across catchments in the presence of HOs as indicated by the positive *Avg.*$_s$ *MRD* and *Avg.*$_s$ *PBIAS* values as shown in Figure 4.4. Design rainfall events estimated by the GEV PD are the most impacted *i.e.* by up to 16% across catchments in the presence of HOs across catchments as indicated by the *Avg.*$_s$ *MARD* and the smallest *Avg.*$_s$ *NSE* values as shown in Figure 4.5, and by the largest *Avg.*$_s$ *MRD* and *Avg.*$_s$ *PBIAS* as shown in Figure 4.4. Design estimates by the LN PD are the least impacted *i.e.* by up to 6% across catchments in the presence of HOs across catchments as indicated by the *Avg.*$_s$ *MARD*. Estimated design rainfall by the GEV, GPA and LP3 PDs in the presence of HOs are classified by *Avg.*$_s$ *NSE* values (< 0.5), as shown in Figure 4.5, as unsatisfactory.

4.1.3 Design floods

MRD, *MARD*, *PBIAS* and *NSE* results for the detailed assessment (*cf.* Section 4.1.1) for all catchments are provided in Figure B1.6 to Figure B1.11 in Appendix B1. Results for the summative assessment on design floods estimated using observed data are provided in this section.

Design floods are generally over-estimated in the presence of LOs by the LN PD as indicated by a positive *Avg.*$_s$ *MRD* and *Avg.*$_s$ *PBIAS* value, and under-estimated by the GEV, GPA and LP3 PDs as indicated by a negative *Avg.*$_s$ *MRD* and *Avg.*$_s$ *PBIAS* value as shown in Figure 4.6.

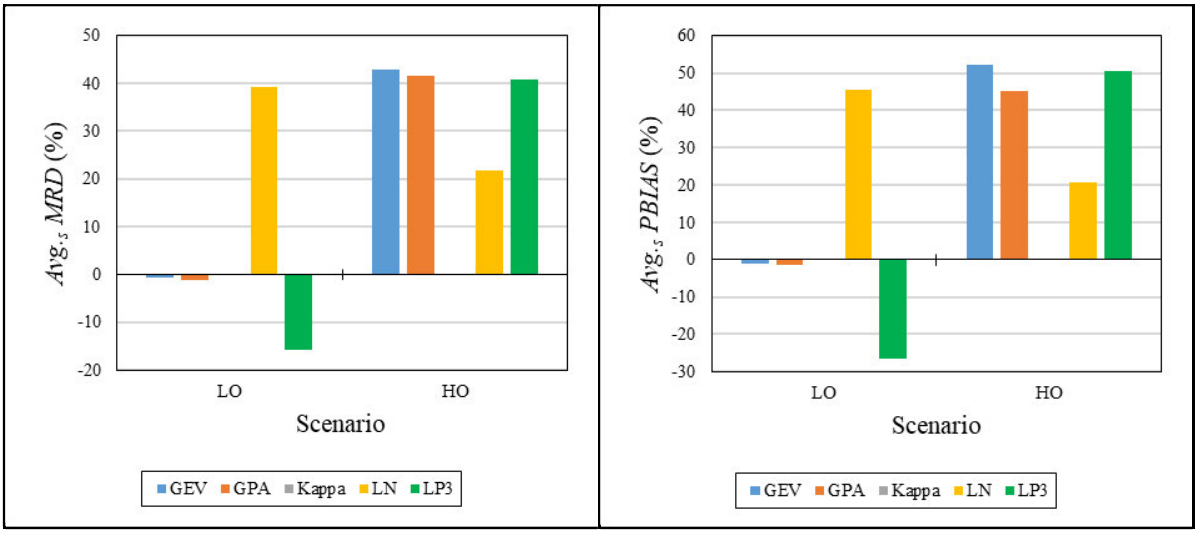

Figure 4.6 *Avg.s MRD* and *Avg.s PBIAS* between design flood events estimated using observed data with LOs and HOs

Design floods estimated by the GEV and GPA PDs are the least impacted to LOs *i.e.* by up to 2% across catchments as indicated by *Avg.s MARD* values and by *Avg.s NSE* value which are close to 1 as shown in Figure 4.7, and by the smallest *Avg.s MRD* and *Avg.s PBIAS* as shown in Figure 4.6. Floods estimated by the LN PD are most impacted to LOs *i.e.* by up to 45% across catchments as indicated by *Avg.s MARD* values and *Avg.s NSE* (< -0.5) as shown in Figure 4.7, and by the largest *Avg.s MRD* and *Avg.s PBIAS*, as shown in Figure 4.6. Furthermore, estimated design floods by the LN and LP3 PDs in the presence of LOs are classified as unsatisfactory by *Avg.s PBIAS* (> ±25%) and *Avg.s NSE* (< 0.5), as shown in Figure 4.6 and Figure 4.7 respectively.

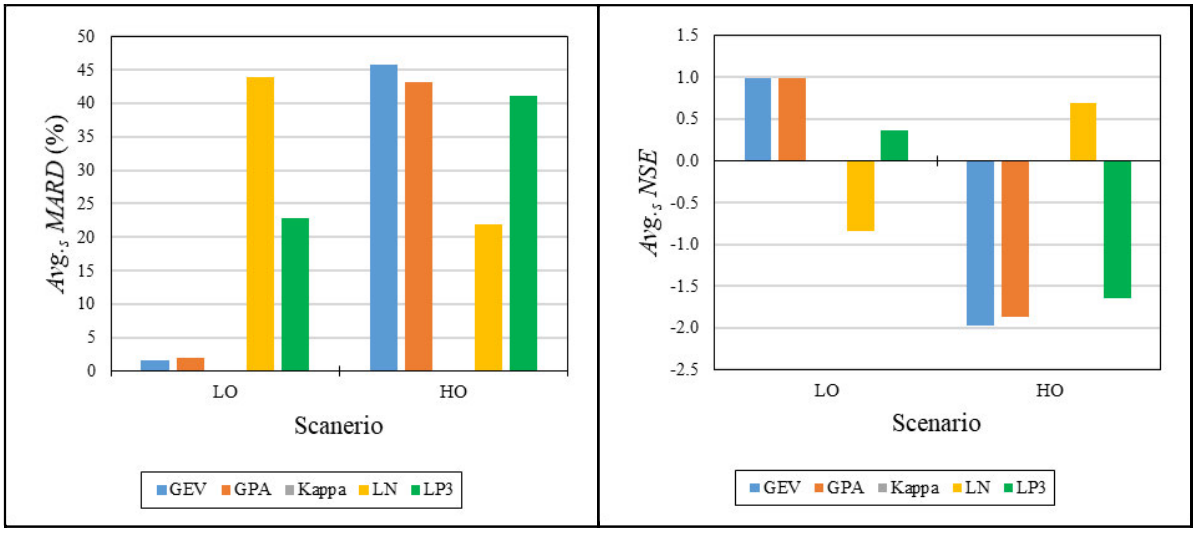

Figure 4.7 *Avg.s MARD* and *Avg.s NSE* for design flood events estimated using observed data with LOs and HOs

Design floods are generally over-estimated by all PDs in the presence of HOs across all catchments as indicated by the positive *Avg.s MRD* and *Avg.s PBIAS* values as shown in Figure 4.6. Design floods estimated by the LN PD are the least impacted by HOs *i.e.* by up to 22% across catchments as indicated with the smallest *Avg.s MARD* values and by the largest *Avg.s NSE* value (0.69) as shown in Figure 4.7, and by the smallest *Avg.s MRD* and *Avg.s PBIAS* values, as shown in Figure 4.6, respectively. Design floods estimated using the GEV PD are the most impacted by HOs *i.e.* by up to 46% across catchments as indicated by the largest *Avg.s MARD* values and smallest *Avg.s NSE*, and the largest *Avg.s MRD* and *Avg.s PBIAS* values, as shown in Figure 4.7 and Figure 4.6, respectively.

4.1.4 Summary of results

Identical trends were obtained between the detailed assessment for design rainfall at Rainfall Gauge 0042227 and the summative assessment when using observed data and are summarised as follows:

(a) Design rainfall and flood events estimated by the GPA and GEV PDs are the least impacted by LOs *i.e.*, by up to 3% for design rainfall and 2% for design floods across catchments.
(b) Design rainfall and flood events estimated by the LN PD are the most impacted by LOs *i.e.*, by up to 22% for design rainfall and 45% for design floods across catchments.
(c) Design events estimated by the GEV PD are the most impacted by HOs *i.e.*, by up to 16% for design rainfall and 46% for design floods across catchments.

(d) Design events estimated by the LN PD are the least impacted by HOs *i.e.,* by up to 6% for design rainfall and 22% for design floods across catchments.

4.2 Use of Synthetically Generated Data Series

The actual PD of each observed dataset is not known, therefore the analysis on observed data may be biased for or against a particular PD. Synthetic dataset were then generated to improve confidence of the analysis by creating AMS datasets from a defined PD. Synthetic datasets were only generated for the GEV and GPA PD as these were advocated as the most appropriate for national scale use in South Africa, as highlighted in Section 4.1.1. A detailed methodology and results on the application of synthetically generated AMS data series to assess the impact of LO and HO events on design rainfall and flood estimates are provided in Section 4.2.1, Section 4.2.2 and Section 4.2.3.

4.2.1 Methodology

A total of 100 AMS each having 100 observations were synthetically generated for both the GEV and GPA PDs using R-Statistical software (Gilleland, 2018; Stephenson, 2018; Wolodzko, 2018), as detailed in Section 4.1.1, and statistics calculated of the observed rainfall and streamflow AMS after being screened for outliers, as detailed in Chapter 3. Generated AMS datasets were thereafter screened for outliers by using the standardised Z-Score method and identified outliers were then removed. DRE and DFE were computed by the GEV and GPA PDs for the 1-, 2-, 5-, 10-, 20-, 50-, 100- and 200-year return period event on each of the 100 AMS which had no outliers. Thereafter, synthetic LOs and HOs were generated and substituted as detailed in Section 4.1.1 to each of the 100 synthetically generated AMS creating scenarios. *Syn.L1, Syn.L2 and Syn.L3* which represent all observations $\leq 1^{st}$, 2^{nd} and 3^{rd} percentile values in synthetic AMS respectively, and *Syn.H1, Syn.H2 and Syn.H3* which represent all observations $\geq 99^{th}$, 98^{th} and 97^{th} percentile values in synthetic AMS respectively.

Design rainfall and floods were then estimated for each scenario dataset from the 100 synthetically generated AMS. A comparison was thereafter made between design events estimated with and without substituted outliers. *MRD, MARD, NSE* and *PBIAS* detailed in Section 4.1.1 were calculated for each scenario dataset from the 100 synthetically generated data series and used to assess the impact of design events to the presence of LOs and HOs.

It was assumed that the substituted outlier events belonged to the respective PDs used to generate the sample datasets. It is acknowledged that any uncertainties in each synthetically generated data series, and variations between observed and synthetically generated data series have not been taken into consideration. As an example, the Probability Density Function (PDF) and Cumulative Distribution Function (CDF) for the observed and 100 synthetically generated rainfall data series using the GEV PD at Rainfall Gauge 0268640 are shown in Figure 4.8 and Figure 4.9 respectively, to highlight the variations between observed and synthetically generated data series.

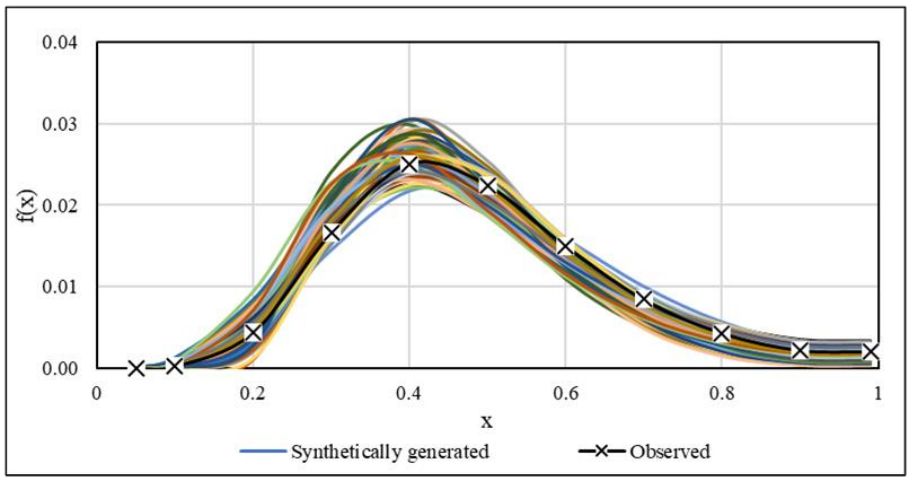

Figure 4.8 PDF for the observed and 100 synthetically generated rainfall data series at Rain Gauge 0268640

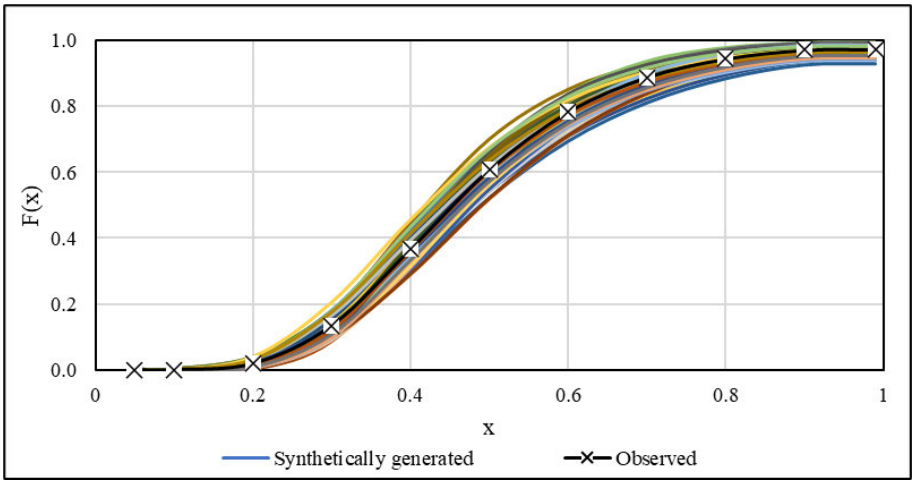

Figure 4.9 CDF for the observed and 100 synthetically generated rainfall data series at gauge 0268640

The detailed and summative assessments as described in Section 4.1.1 were performed on design rainfall and flood events which are in Section 4.2.2 and Section 4.2.3 respectively.

4.2.2 Design rainfall

For the detailed assessment, *MRD*, *MARD*, *NSE* and *PBIAS* were calculated for each scenario dataset from the 100 synthetically generated data series and then an average ($Avg._d$) across all 100 synthetically generated series per scenario were calculated. Results for the detailed assessment for all rainfall gauges when using synthetically generated data series are provided in Figure B2.1 to Figure B2.6 in Appendix B2. Design rainfall results for the summative assessment are provided in this section. Results from the detailed assessment are used to inform the summative assessment, therefore a discussion of all of the results from the detailed assessment are not provided here.

For the summative assessment, results for the computed statistics were averaged across LO scenarios *i.e.*, Syn.*L1, Syn.L2 and Syn.L3* and across HO scenarios *i.e., Syn.H1, Syn.H2 and Syn.H3* per PD for each rainfall gauge and thereafter averaged across all rainfall gauges to obtain $Avg._s$ *MRD*, $Avg._s$ *MARD*, $Avg._s$ *NSE* and $Avg._s$ *PBIAS.* Design rainfall events estimated by the GEV and GPA PDs are generally under-estimated by up to 2% in the presence of LOs as indicated by negative $Avg._s$ *MRD* and $Avg._s$ *PBIAS*, as shown in Figure 4.10.

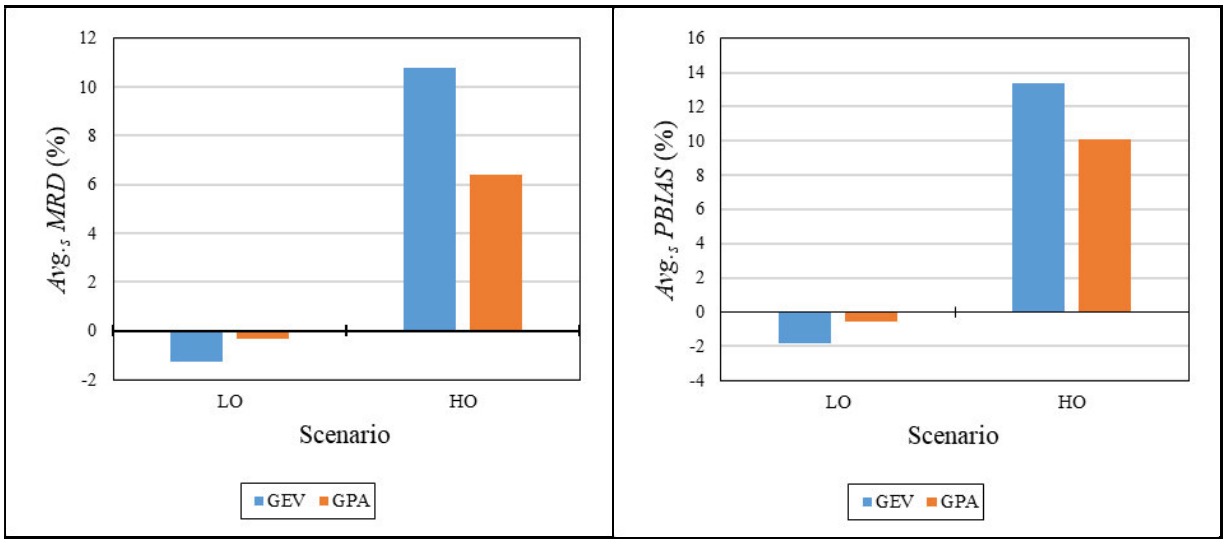

Figure 4.10 *Avg.s MRD* and *Avg.s PBIAS* for design rainfall events estimated using synthetically generated data with LOs and HOs

Events estimated by the GEV and GPA PDs are impacted by up to 2% across catchments in the presence of LOs indicated by *Avg.$_s$ MARD* as shown in Figure 4.11. Furthermore, acceptable design events are estimated in the presence of LOs from both the GEV and GPA PDs as indicated by *Avg.$_s$ PBIAS* values (< 10%) and *Avg.$_s$ NSE* values (> 0.75), as shown in Figure 4.10 and Figure 4.11 respectively.

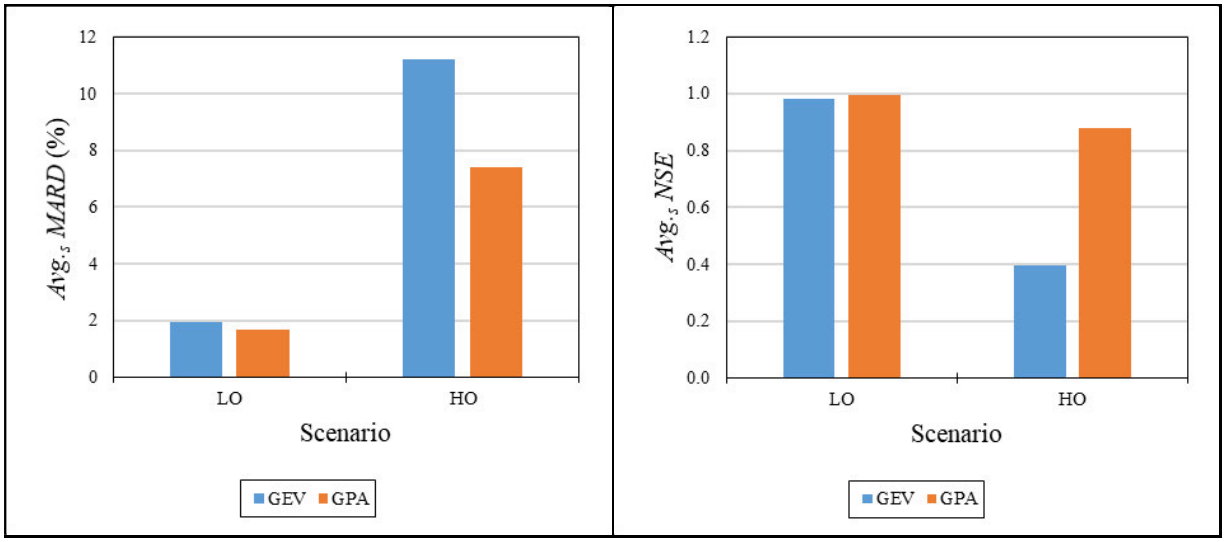

Figure 4.11 *Avg.$_s$ MARD and Avg.$_s$ NSE* for design rainfall events estimated using synthetically generated data with LOs and HOs

Rainfall events estimated by the GEV and GPA PDs are over-estimated in the presence of HOs as indicated by positive *Avg.$_s$ MRD* and *Avg.$_s$ PBIAS*, as shown in Figure 4.10. Rainfall events are impacted by up to 12% in the presence HOs across catchments indicated by *Avg.$_s$ MARD* values shown in Figure 4.11, with events estimated from the GEV PD being more impacted than estimated events using the GPA PD indicated by larger *Avg.$_s$ MARD* and smaller *Avg.$_s$ NSE* values, as shown in Figure 4.11, and by larger positive *Avg.$_s$ MRD* and *Avg.$_s$ PBIAS* as shown in Figure 4.10.

4.2.3 Design floods

Computed *Avg.$_d$ MRD*, *Avg.$_d$ MARD*, *Avg.$_d$ PBIAS* and *Avg.$_d$ NSE* for the detailed assessment in all catchments are provided in Figure B2.7 to Figure B2.12 in Appendix B2. Results for the summative assessment on design floods estimated using synthetically generated data series are provided in this section.

Design floods are under-estimated from both GEV and GPA PDs in the presence of LOs as indicated by negative *Avg.s MRD* and *Avg.s PBIAS* values as shown in Figure 4.12. Flood events estimated by the GEV and GPA PDs are impacted by up to 1% across catchments in the presence of LOs as indicated by *Avg.s MARD* as shown in Figure 4.13. Estimated flood events are classified as acceptable by *Avg.s PBIAS* (< -10%) and *Avg.s NSE* (> 0,75) as shown in Figure 4.12 and Figure 4.13 respectively.

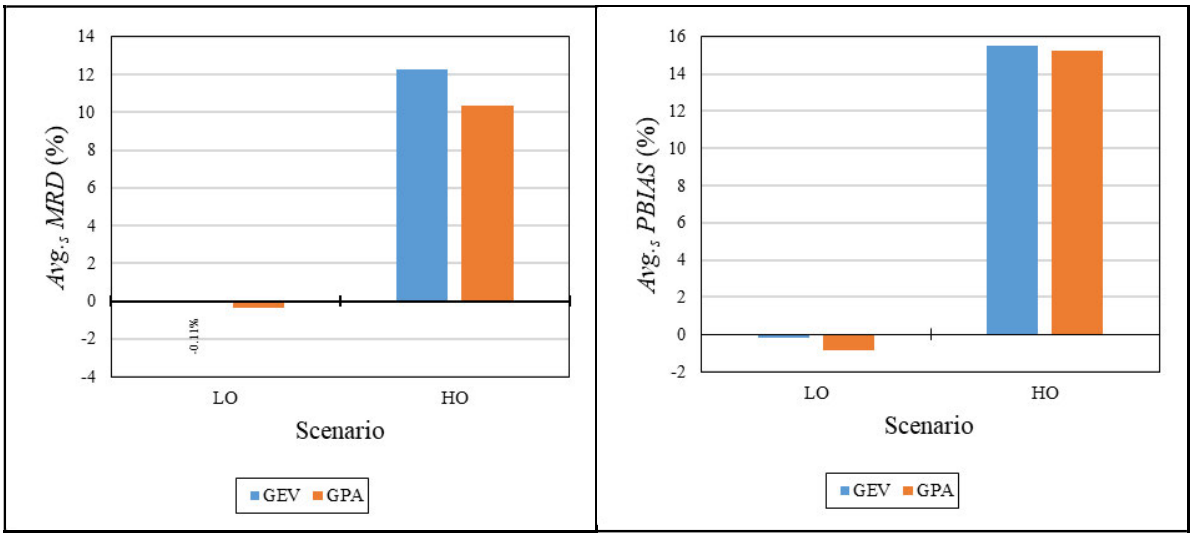

Figure 4.12 *Avg.s MRD* and *Avg.s PBIAS* for design flood events estimated using synthetically generated data with LOs and HOs

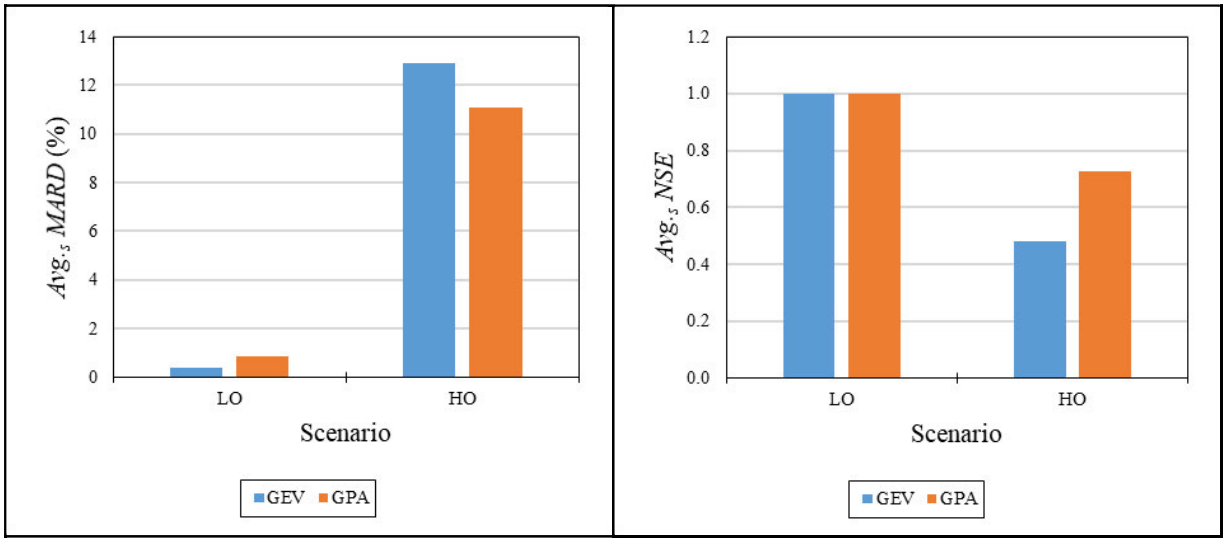

Figure 4.13 *Avg.s MARD* and *Avg.s NSE* for design flood events estimated using synthetically generated data with LOs and HOs

Design flood events are over-estimated in the presence of HOs by the GEV and GPA PDs with estimates from the GEV PD being impacted by up to 13% across catchments whereas estimates from the GPA PD are impacted by up to 12% as indicated by $Avg._s$ *MARD,* The greater impact of floods estimated by the GEV PD in the presence of HOs compared to estimates by the GPA PD are also indicated by smaller $Avg._s$ *NSE* values as indicated in Figure 4.13, and by larger $Avg._s$ *MRD* and $Avg._s$ *PBIAS* as shown in Figure 4.12.

4.2.4 Summary of results

The following trends are summarised for design rainfall and flood events estimated with synthetically generated data series:

(a) Design rainfall events estimated by the GEV and GPA PDs are impacted by up to 2% in the presence of LOs across catchments and design floods are impacted by up to 1%.
(b) Design rainfall and flood, events estimated by the GEV and GPA PDs in the presence of HOs, are impacted by up to 12% and 13% respectively across catchments.

4.3 Chapter Summary and Conclusions

The impact of outlier events on DRE and DFE were assessed using both at-site observed AMS data and synthetically generated AMS data series for selected PDs. Synthetic datasets were generated to improve confidence in the analysis by creating AMS datasets from a defined PD, *i.e.,* to ensure that the fitted PD is the correct PD for the dataset. The GEV, GPA, Kappa, LN and LP3 PDs were used in analysing observed datasets, whereas only the GEV and GPA PD were used in analysing synthetically generated datasets. Synthetic LO and HO events were calculated and substituted into the each observed and each of the 100 synthetically generated AMS per rainfall gauge and streamflow gauge which created six dataset scenarios of outliers. The 2-, 5-, 10-, 20-, 50-, 100- and 200-year return period design rainfall and flood events were estimated from observed data and synthetically generated data series with and without synthetic outliers. Comparisons between estimated design rainfall and floods with and without substituted outliers were undertaken.

From the analysis of observed datasets, design events estimated by the LN PD are the most impacted by LOs *i.e.,* by up to 22% for design rainfall and 45% for design floods across catchments, and the least impacted by HOs *i.e.,* by up to 6% for design rainfall and 22% for design floods across catchments. Design events estimated by the GEV and GPA PDs are the least impacted by LOs *i.e.,* by up to 3% for design rainfall and 2% for design floods across

catchments, and the most impacted by HOs *i.e.*, by up to 16% for design rainfall and 46% for design floods across catchments.

Regarding the analysis of synthetically generated data series, design rainfall events estimated by the GEV and GPA PDs are impacted by up to 2% in the presence of LOs across catchments and up to 1% for design flood events. Design rainfall and flood events estimated by the GEV and GPA PDs in the presence of HOs are impacted by up to 12% and 13% respectively across catchments. The difference in results between using observed and synthetically generated datasets may be a result of the incorrect PD being applied on the observed datasets.

For both practice and research purposes, the impact of outlier events are highlighted in this study, which indicates that outlier events must not be ignored in DRE and DFE, and caution needs to be taken when applying the Kappa, LN and GEV PDs for DRE and DFE in SA due to the impact of outliers when using these PDs. Within practice, the presence of LOs generally results in an under-estimation of design rainfall and floods thereby reducing the accuracy of infrastructural design which increases the risk of failure and increases risk to the safety of lives and severe economic, environmental, and social consequences. The presence of HOs generally results in an over-estimation of design rainfall and floods thereby resulting in an over-design of infrastructure which provides a conservative approach, however, the economic viability of the design may be questioned. The LN and GEV PD are most impacted by LOs and HOs respectively, therefore special care should be taken in their application.

It is recommended from results in this study that LOs be excluded and HOs should not be excluded from DRE and DFE in SA after such events have been verified against events from neighboring stations. LOs should be excluded from DRE and DFE as these events affect the estimation of sample statistics resulting in biased parameter estimates (Asikoglu, 2017; England Jr *et al.*, 2019) and a distortion of exceedance probabilities of large design events (Lamontagne *et al.*, 2013; England Jr *et al.*, 2019). HOs should not be excluded because these events have a low probability of exceedance therefore providing valuable information in the estimation of design events (Costa and Jarrett, 2008; England Jr *et al.*, 2019). Judgement from the analyst is ultimately required on whether to include or exclude HOs from further analysis, as also recommended by (England Jr *et al.*, 2019). It is recommended from results in this study that special caution be taken when applying the Kappa, LN and GEV PDs for DRE and DFE in SA due to the impact of outliers when using these PDs. It is also recommended that this

study be expanded to other regions in SA to have more confidence in the findings and thereafter be used in a South African national guideline.

5. PERFORMANCE OF OUTLIER DETECTION METHODS

The performance of the BP, MZS and MGBT outlier detection methods (*cf.* Section 2.2.3) in detecting substituted outliers on observed data and synthetically generated rainfall and streamflow AMS records are detailed in Section 5.1 and Section 5.2 respectively. A schematic of the structure of Chapter 5 is shown in Figure 5.1.

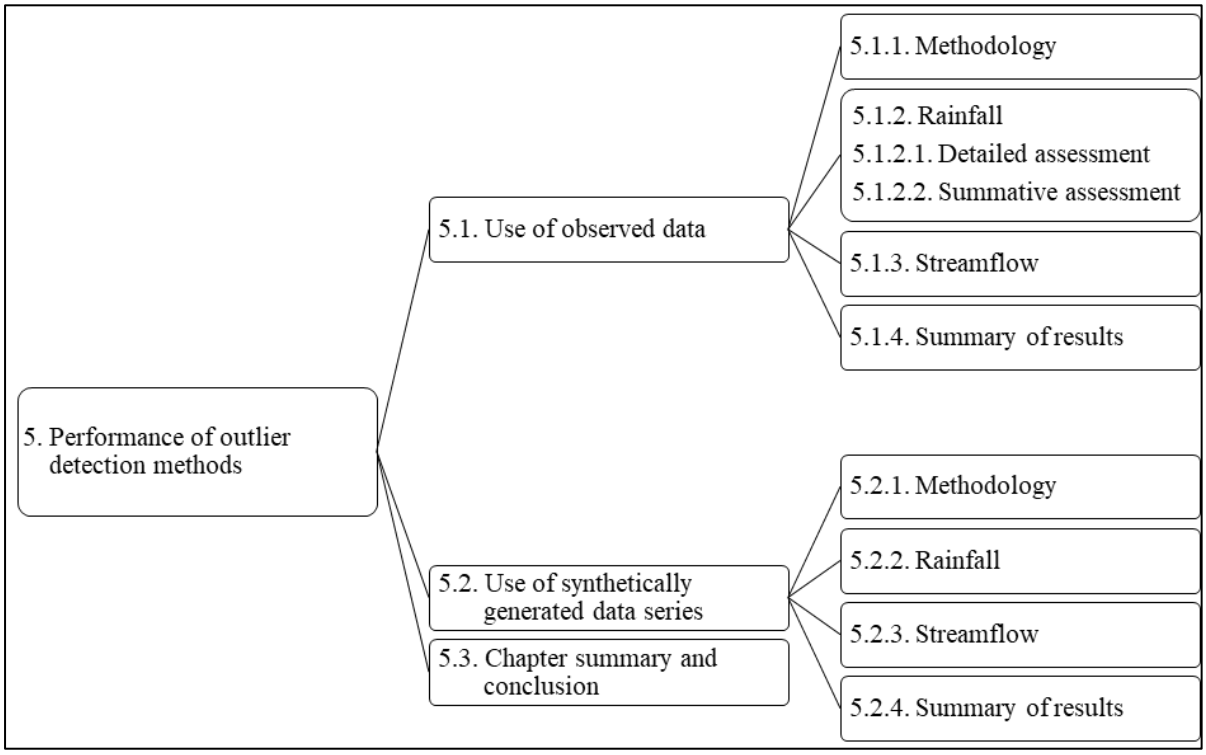

Figure 5.1 Schematic of the structure of Chapter 5

5.1 Use of Observed Data

The applied method, results and summary of results on the performance of outlier detection methods on observed rainfall and flow data are provided in Section 5.1.1, Section 5.1.2 and Section 5.1.3, and Section 5.1.4, respectively.

5.1.1 Methodology

Outlier events were generated and substituted within observed rainfall and streamflow data creating six scenarios *i.e. Obs.L1, Obs.L2, Obs.L3, Obs.H1, Obs.H2* and *Obs.H3* (*cf.* Section 4.1.1). The BP and MZS were applied to detect both LO and HO events in data from each rainfall and streamflow gauge for all scenarios. The MGBT was applied in R-Studio by the MGBT package (Asquith *et al.*, 2018) with default parameters for LO scenarios as it is only

designed to detect LOs. It was assumed that detected outlier events by each outlier detection method are true outlier events within each sample dataset.

The following two assessments were conducted to evaluate the performance of outlier detection methods on observed data:

(a) Detailed assessment: Draws a comparison between the percentage of outliers substituted and detected within each scenario across all rainfall and streamflow gauges.
(b) Summative assessment: Average Substituted (*Avg. Sub*) and Average Detected (*Avg. Det*) outliers were calculated across *Obs.L1, Obs.L2* and *Obs.L3* and across *Obs.H1, Obs.H2* and *Obs.H3* for each rainfall and streamflow gauge, and thereafter averaged across all rainfall and streamflow gauges. *RD* were then calculated between *Avg. Sub* and *Avg. Det* using Equation 5.1 to provide a *RD* of Avg. detection (%). A positive *RD* indicates over-detection whereas a negative *RD* indicates an under-detection.

$$RD_{det.obs} = \frac{[Avg.Det_{O,D} - Avg.Sub_{O,D}]}{Avg.Sub_{O,D}} \times 100 \quad (5.1)$$

where:

$RD_{det.obs}$ = *RD* of Avg. detection on observed data series (%),
Avg.Det = Average detected outliers (%),
O = Outlier type (low or high),
D = Data type (rainfall or streamflow), and
Avg.Sub = Average substituted outliers (%).

The performance of outlier detection methods on observed rainfall data are provided in Section 5.1.2 and for streamflow data in Section 5.1.3.

5.1.2 Rainfall

Results for the detailed assessment is first presented in Section 5.1.2.1 followed by the summative assessment in Section 5.1.2.2.

5.1.2.1 Detailed assessment

As shown in Figure 5.2, the performance of the BP method in detecting outliers varies across all scenarios and rainfall gauges. There are more LOs detected than substituted for scenarios *Obs.L1, Obs.L2* and *Obs.L3* for Rainfall Gauges 0476031, 0589670, 0239097 and 0268640 with an equal number of substituted and detected outliers in Rainfall Gauge 0042227, and no

outliers detected for Rainfall Gauge 0025414. For scenarios *Obs.H1*, *Obs.H2* and *Obs.H3*, there are generally more HOs detected than substituted across all rainfall gauges except for *Obs.H3* at Rainfall Gauge 0239097, as shown in Figure 5.2.

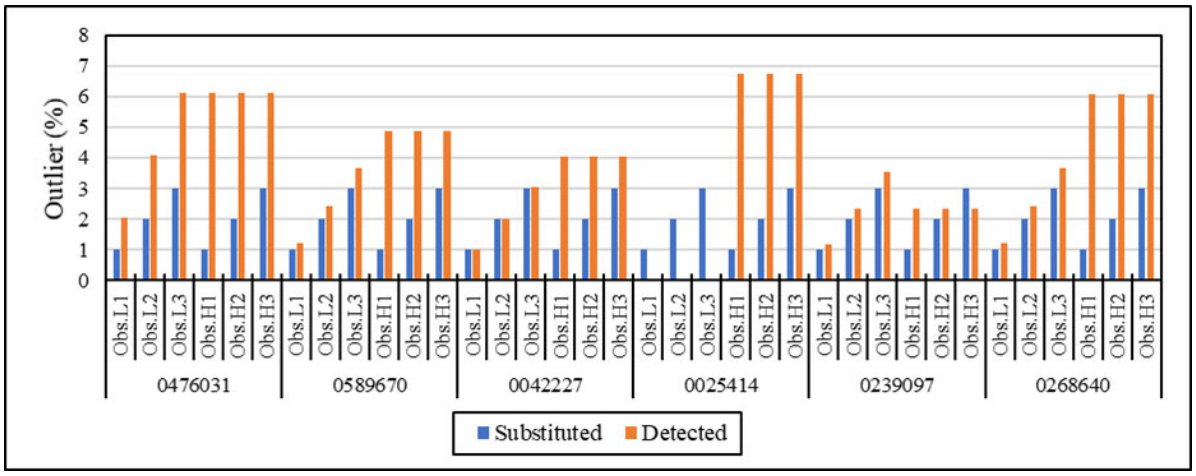

Figure 5.2 Percentage of substituted and detected outliers using the BP method on observed rainfall data with LO (*Obs.L1*, *Obs.L2* and *Obs.L3*) and HO (*Obs.H1*, *Obs.H2* and *Obs.H3*) scenarios

As shown in Figure 5.3, the MZS method was unable to detect any substituted LOs across all LO scenarios and rainfall gauges as shown in Figure 5.3. No HOs were detected for Rainfall Gauge 0042227 and 0268640 with an over-detection at Rainfall Gauge 0589670 and 0025414 and under-detection in Rainfall Gauge 0239097 across all HO scenarios.

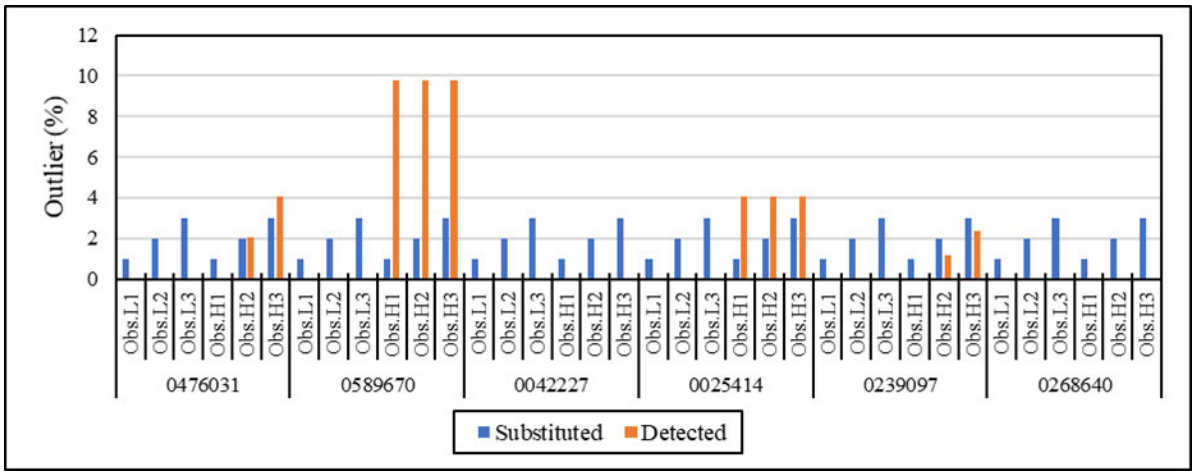

Figure 5.3 Percentage of substituted and detected outliers using the MZS method on observed rainfall data with LO (*Obs.L1*, *Obs.L2* and *Obs.L3*) and HO (*Obs.H1*, *Obs.H2* and *Obs.H3*) scenarios

The MGBT performed well in detecting the correct number of substituted outliers across all rainfall gauges for all LO scenarios except for *Obs.L2* at Rainfall Gauge 0476031 and *Obs.L3* at Rainfall Gauge 0025414 at which outliers were under-detected as shown in Figure 5.4.

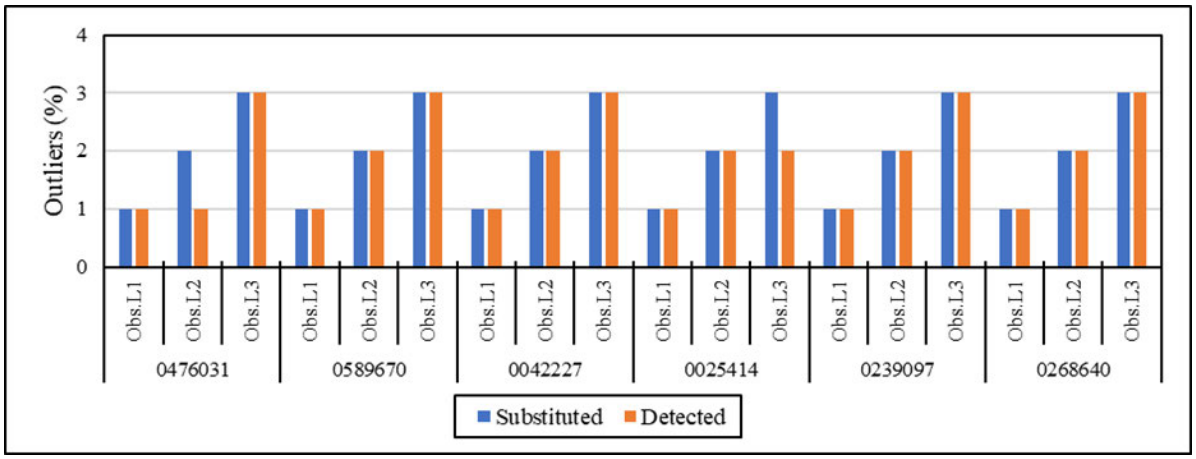

Figure 5.4 Percentage of substituted and detected outliers using the MGBT on observed rainfall data with LO (*Obs.L1*, *Obs.L2* and *Obs.L3*) scenarios

5.1.2.2 Summative assessment

The MZS and MGBT generally under-detects LOs by up to 100% and 6% respectively as indicated by negative *RD* value, as shown in Figure 5.5. The BP generally over-detects LOs up to 15% as shown in Figure 5.5. HOs are generally over-detected by the MZS and BP methods with the BP over-detecting up to 150% compared to an over-detection of up to 50% by the MZS as indicated by positive *RD* values as shown in Figure 5.5.

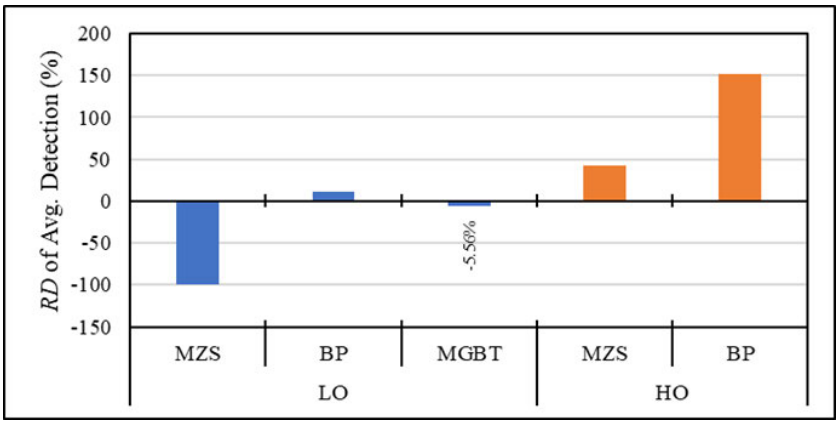

Figure 5.5 *RD* of Avg. Detection of LOs and HOs in observed rainfall data using the BP, MZS and MGBT

5.1.3 Streamflow

Results for the detailed assessment on synthetically generated streamflow data series are provided in Figure C.1 to Figure C.3 in Appendix C. Results for the summative assessment are provide in this section. LOs within observed streamflow data are under-detected by up to 100% by the MZS and BP methods and by up to 30% by the MGBT indicated by negative *RD* values as shown in Figure 5.6. There is a greater over-detection of HOs by the BP method compared to the MZS as HOs are over-detected by up to 300% from the BP method and by up to 100% from the MZS method as indicated by positive *RD* values, as shown in Figure 5.6.

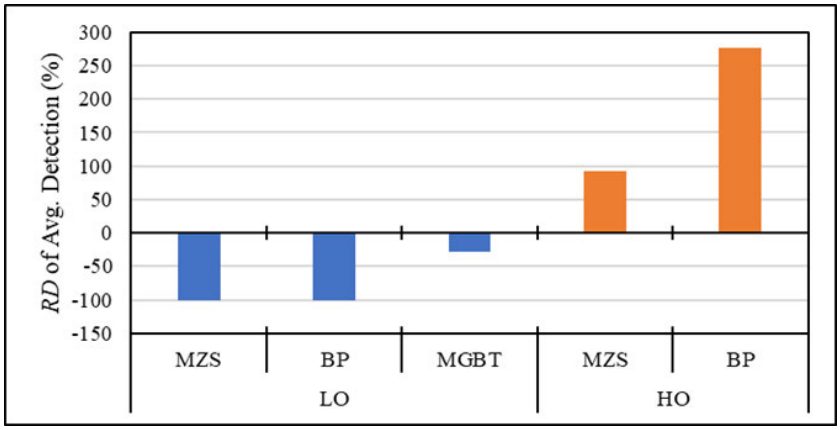

Figure 5.6 *RD* of Avg. Detection of LOs and HOs in observed streamflow data using the BP, MZS and MGBT

5.1.4 Summary of results

The following summarises the performance of outlier detection methods when using observed rainfall and streamflow data:

(a) The MGBT performs the best in detecting LOs in observed rainfall (-6%) and streamflow data (-30%) compared to the BP (15%) and MZS (-100%) methods for observed rainfall data, and compared to the BP (-100%) and MZS (-100%) methods for observed streamflow data.
(b) The MZS method outperforms the BP method in detecting HOs in observed rainfall and streamflow datasets. This is indicated by a *RD* of Avg. detection of up to 50% and 150% by the MZS and BP methods respectively for observed rainfall data, and up to 100% and 300%, respectively, for observed streamflow data.

(c) The BP was better at detecting LOs than detecting HOs for both rainfall and streamflow data. This is indicated by lower *RD* of Avg. LO detection compared to HOs in observed rainfall of up to -15% and 50% respectively.

5.2 Use of Synthetically Generated Data Series

A brief methodology, results and summary of results on the application of BP, MZS and MGBT outlier detection methods on synthetically generated AMS data series are provided in Section 5.2.1, Section 5.2.2 and Section 5.2.3, and Section 5.2.4 respectively.

5.2.1 Methodology

Synthetic outlier scenarios (*i.e. Syn.L1, Syn.L2, Syn.L3, Syn.H1, Syn.H2* and *Syn.H3*) for each of the 100 synthetically generated rainfall and streamflow data series generated from the GEV PD per rainfall and streamflow gauge as detailed in Section 4.2.1 was used in this analysis. Datasets generated using the GEV PD was chosen due its widespread use in South Africa as highlighted in Section 4.1.1. The BP, MZS and MGBT were applied to each of the 100 synthetically generated rainfall and streamflow data series per scenario and gauge. The BP and MZS were used to detect both LOs and HOs whereas the MGBT is designed to detect LOs only. It was assumed that no outliers were present in each of the 100 synthetically generated data series prior to outlier substitution.

The following two assessments were conducted to evaluate the performance of outlier detection methods on synthetically generated data series:

(a) Detailed assessment: Outliers detected in each of the 100 synthetically generated data series were averaged ($Avg._d$) per scenario to provide an average outlier detection per scenario. This process was repeated for each rainfall and streamflow gauge and outlier detection method. A comparison between the $Avg._d$ percentage of outliers substituted and detected per scenario for all rainfall and streamflow gauges was then drawn.
(b) Summative assessment: $Avg._d$ outliers per scenario were further averaged across *Syn.L1*, *Syn.L2* and *Syn.L3* and across *Syn.H1*, *Syn.H2* and *Syn.H3* for each rainfall and flow gauges, and thereafter averaged across all rainfall and streamflow gauges. This resulted in averaged ($Avg._s$) substituted ($Avg._s$ *Sub*) and detected ($Avg._s$ *Det*) outliers per detection method. *RD* were then calculated between $Avg._s$ *Sub* and $Avg._s$ *Det* using Equation 5.2 to provide a *RD* of $Avg._s$ detection (%). A positive *RD* indicates as over-detection whereas a negative *RD* indicates an under-detection.

$$RD_{det\ syn} = \frac{[Avg._s Det_{O,D} - Avg._s Sub_{O,D}]}{Avg._s Sub_{O,D}} \times 100 \quad (5.2)$$

where:

$RD_{det\ syn}$ = *RD* of Avg. detection on synthetically generated data series (%),

$Avg._s\ Det$ = Average detected outliers (%),

O = Outlier type (low or high),

D = Data type (rainfall or streamflow), and

$Avg._s\ Sub$ = Average substituted outliers (%).

The performance of outlier detection methods on synthetically generated rainfall data series are provided in Section 5.2.2 and for streamflow data series in Section 5.2.3.

5.2.2 Rainfall

Results for the summative assessment are provided in this section with results for the detailed assessment provided in Figure C.4 to Figure C.6 in Appendix C. Substituted LOs are under-detected by up to 30% from the MZS as indicated by a negative *RD*, with the BP and MGBT over-detecting LOs by up to 50% and 100% respectively as shown in Figure 5.7. HOs are over-detected from both MZS and BP, with the BP method over-detecting more outliers (by up to 70%) than MZS (by up to 20%), as shown in Figure 5.7.

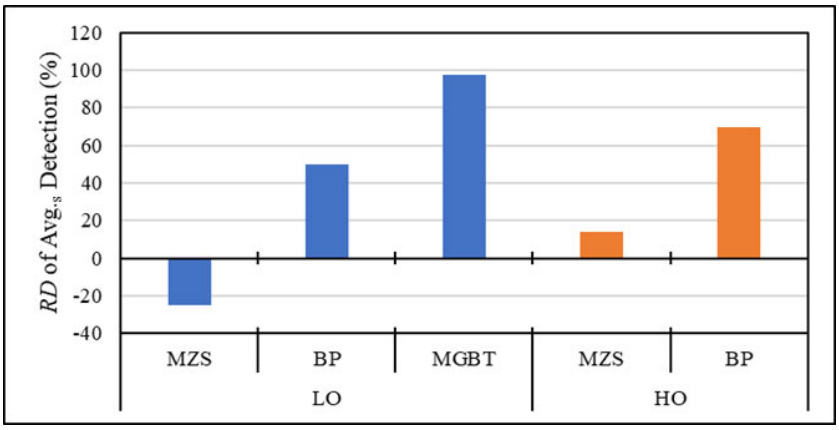

Figure 5.7 *RD* of Avg. Detection of LOs and HOs in synthetically generated rainfall data series using the BP, MZS and MGBT

5.2.3 Streamflow

Results for the detailed assessment on synthetically generated streamflow data series are provided in Figure C.7 to Figure C.9 in Appendix C. The MZS and BP generally under-detects

LOs, with the MZS under-detecting by up to 60% and BP by up to 10%, as shown in Figure 5.8. The MGBT significantly over-detects LOs up to 1550% as shown in Figure 5.8.

This significant over-detection may be attributed to the swamping mechanism described in Section 2.2.3.4. Swamping was also observed in a study by Rahman *et al.* (2014). HOs are generally over-detected by the MZS and BP methods with the BP over-detecting by up to 40% indicated by positive *RD* values as shown in Figure 5.5.

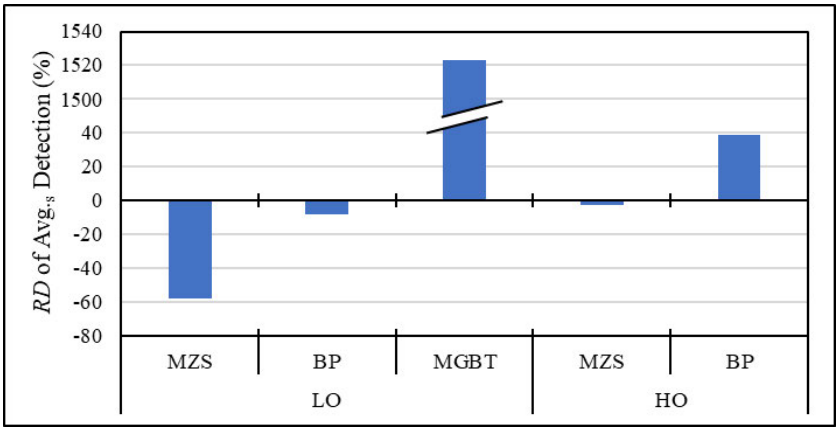

Figure 5.8 *RD* of Avg. Detection of LOs and HOs in synthetically generated streamflow data series using the BP, MZS and MGBT

5.2.4 Summary of results

The following summarises the performance of outlier detection methods when using synthetically generated rainfall and streamflow data:

(a) The MGBT performs the worst in detecting LOs within synthetically generated rainfall and streamflow data series by over-detecting up to 100% and 1550%, respectively, with the BP performing the best in detecting LOs.
(b) MZS outperforms the BP method in detecting HOs within synthetically generated rainfall and streamflow datasets. This is indicated by a *RD* of Avg. detection of up to 20% and 80% respectively for synthetically generated rainfall data series and -5% and 40% respectively for synthetically generated streamflow data series.
(c) There was an improved performance of the BP in detecting LOs compared to HOs for both rainfall and streamflow data. This is indicated by lower *RD* of Avg. LO detection compared to HOs in synthetically generated rainfall of up to 60% and 80% respectively, and streamflow of up to -20% and 40% respectively.

5.3 Chapter Summary and Conclusion

The BP, MZS and MGBT were applied to observed rainfall and streamflow data and to synthetically generated rainfall and streamflow data series containing substituted LOs (*Obs.L1, Obs.L2, Obs.L3, Syn.L1, Syn.L2 and Syn.L3*) and HOs (*Obs.H1, Obs.H2, Obs.H3, Syn.H1, Syn.H2 and Syn.H3*). The BP and MZS were used to detect both LOs and HOs whereas the MGBT is only designed to detect LO.

From the analysis of observed data the MGBT outperforms the BP and MZS in detecting LOs with a *RD* of Avg. detection of by up to -6% and -30% in observed rainfall and streamflow data respectively. The MZS outperforms the BP method in detecting HOs with a *RD* of Avg. detection of by up to 50% and 100% in observed rainfall and streamflow data respectively. From the analysis of synthetically generated data series, the MZS outperforms the BP and MGBT in detecting LOs in rainfall datasets of by up to -30%. The BP outperforms the MGBT and MZS in detecting LOs in streamflow datasets of by up to -15%. The MZS outperforms the BP method in detecting HOs. *RD* of Avg. detection of by up to 20% and -3% in synthetically generated rainfall and streamflow data respectively.

It is recommended from these results that the MGBT be used to detect LOs and the MZS be used to detect HOs in both rainfall and streamflow data. It is acknowledged, in being recommended, that the MGBT method is prone to under- and over-detection of possible LO events due to the mechanisms referred to as masking and swamping, respectively, as described in Section 2.2.3.4. Swamping is noticed on the use of synthetically generated data series. Results from this study will also inform future application in both practice and research and may reduce over- or under-estimation DRE and DFE.

6. IMPACT OF PERIOD OF RECORD AND REDUCED RECORD LENGTH ON DESIGN RAINFALL AND FLOOD ESTIMATION

This chapter contains the detailed methodology, results, and conclusions on the impact of period of record and reduced record length on DRE and DFE in SA. A schematic of the structure of Chapter 6 is shown in Figure 6.1.

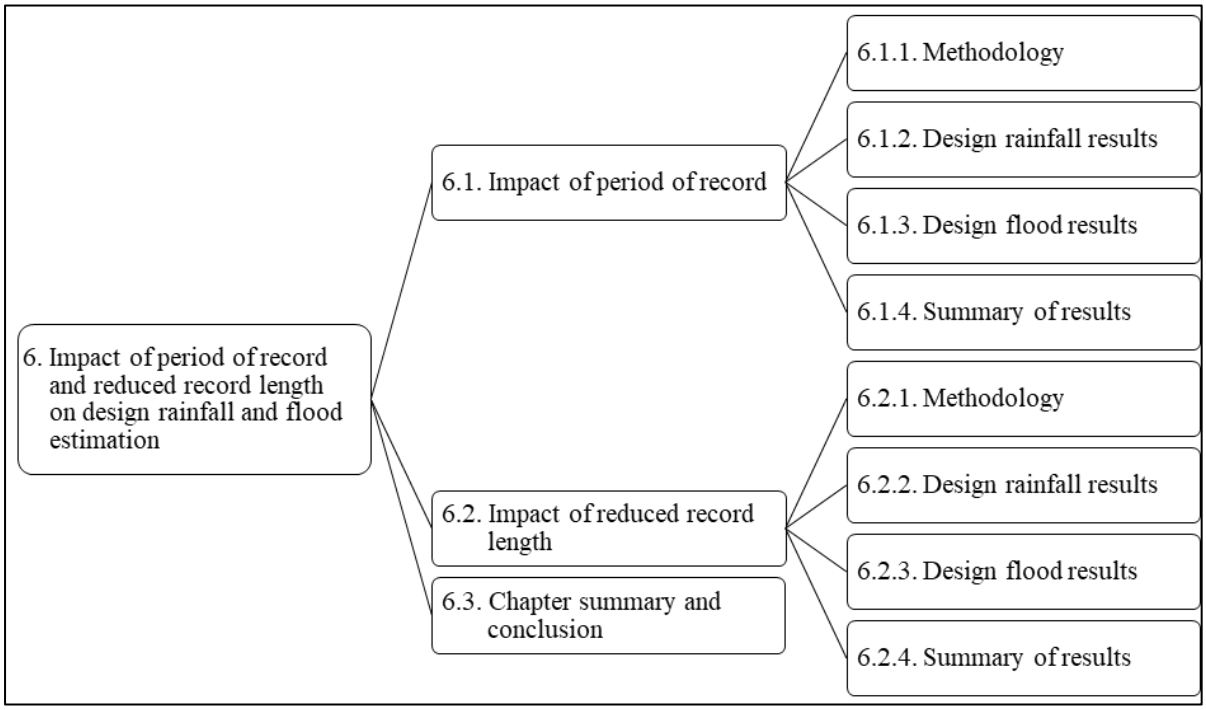

Figure 6.1 Schematic of the structure of Chapter 6

6.1 Impact of Period of Record

A detailed methodology, results and summary of results on the impact of period of record on design rainfall and flood events are provided in Section 6.1.1, Section 6.1.2 and Section 6.1.3, and Section 6.1.4, respectively.

6.1.1 Methodology

The length of the initial observed AMS record of each driver rainfall and streamflow gauge listed in Chapter 3 was reduced to 75% and 50% by using a moving window approach. A total of three windows each representing a chronological period of time, for both 75% and 50% of the AMS, were chosen thus creating six scenarios as explained in Table 6.1. A moving window

approach was adopted to increase confidence in this analysis and to account for the impact of different time periods on DRE and DFE.

Table 6.1 Explanation of reduced period of record scenarios

Dataset Type	Scenario ID	Explanation
Datasets using 75% of record length	*f.75*	First 75% of observations, *i.e.* 1 to 75% of initial AMS
	m.75	Middle 75% of observations, *i.e.* 13 to 88% of initial AMS
	l.75	Last 75% of observations, *i.e.* 25 to 100% of initial AMS
Datasets using 50% of record length	*f.50*	First 50% of observations, *i.e.* 1 to 50% of initial AMS
	m.50	Middle 50% of observations, *i.e.* 25 to 75% of initial AMS
	l.50	Last 50% of observations, *i.e.* 50 to 100% of initial AMS

Design rainfall and flood events were estimated using the GEV PD for each scenario of reduced record length and compared to design estimates computed using the entire period, *i.e.,* the initial AMS. The GEV PD was chosen to provide a conservative approach due to the degree of impact, as detailed in Chapter 4, and due its widespread use as highlighted in Section 4.1.1. The initial observed AMS datasets, *i.e.,* observed data after data screening and pre-processing as detailed in Chapter 3, were used in this analysis and not the synthetically generated data series. It was assumed that the initial observed AMS record resulted in the most accurate and representative design events compared to events estimated from a reduced record length. It was also assumed that all observed rainfall and streamflow data fitted the GEV PD. It is acknowledged that stationarity within then rainfall and streamflow datasets was assumed and that any uncertainties in the observed data had not been taken into consideration.

MRD, *MARD*, *PBIAS* and *NSE* were calculated per scenario *i.e., f.75, m.75, l.75, f.50, m.50, l.50* for each catchment and used to evaluate the impact of period of record on design estimates within each catchment. *RD* as described in Section 4.1.1 was calculated using Equation 6.1. *MRD* were then calculated as the arithmetic mean of the *RD* across all return periods. *MARD* was calculated as the arithmetic mean of the absolute *RD* across all return periods. *NSE* and *PBIAS* as detailed in Section 4.1.1 were calculated using Equation 6.2 and Equation 6.3 respectively.

$$RD_{rec} = \frac{[E_{rec,T} - E_{100,T}]}{E_{100,T}} \times 100 \quad (6.1)$$

where:

RD_{rec} = *RD* calculated between design rainfall or flood events estimated using a scenario of reduced record length and entire record length (%),

E_{rec} = Design rainfall or flood estimate from *f.75*:*l.50* (mm or $m^3.s^{-1}$), and

E_{100} = Rainfall or flood estimated using entire record length (mm or $m^3.s^{-1}$).

$$NSE = \left(1 - \frac{\sum_{T=1}^{7}(E_{100,T} - E_{rec,T})^2}{\sum_{T=1}^{7}(E_{100,T} - \bar{E})^2}\right) \quad (6.2)$$

where:

NSE = Statistic quantifying the fit of estimated design events against the 1:1 line, and

$\bar{E}$ = Avg. rainfall or flood estimated by the entire record length (mm or $m^3.s^{-1}$).

$$PBIAS = \sum_{T=1}^{7}\left[\frac{(E_{100,T} - E_{rec,T}) \times 100}{E_{100,T}}\right] \quad (6.3)$$

Design rainfall results are presented in Section 6.1.2 followed by design floods in Section 6.1.3.

6.1.2 Design rainfall

Positive and negative *MRD* and *PBIAS* values are computed across all scenarios and rainfall gauges indicating an over- and under-estimation respectively of design events as shown in Figure 6.2 with no clear trend that an individual period of time results in the largest over- or under-estimation. For example, the largest three under-estimations were calculated from *f.75*, *m.75* and *l.50* for Rainfall Gauges 0239097 and 0589670, and the largest three over-estimations were calculated from *f.75*, *l.75* and *f.50* within for Rainfall Gauges 0268640 and 0239097 as indicated by having the largest *MRD* and *PBIAS* values as shown in Figure 6.2.

There is no definite trend of consistent over- or under-estimation of rainfall events from an individual period of time across all gauges as assessed from *MRD* and *PBIAS* values as shown in Figure 6.2. For example, *f.75* results in an over-estimation for Gauge 0476031 indicated by positive *MRD* and *PBIAS* values whereas *f.75* results in an under-estimation for Gauge 0589670 indicated by negative *MRD* and *PBIAS* values as shown in Figure 6.2.

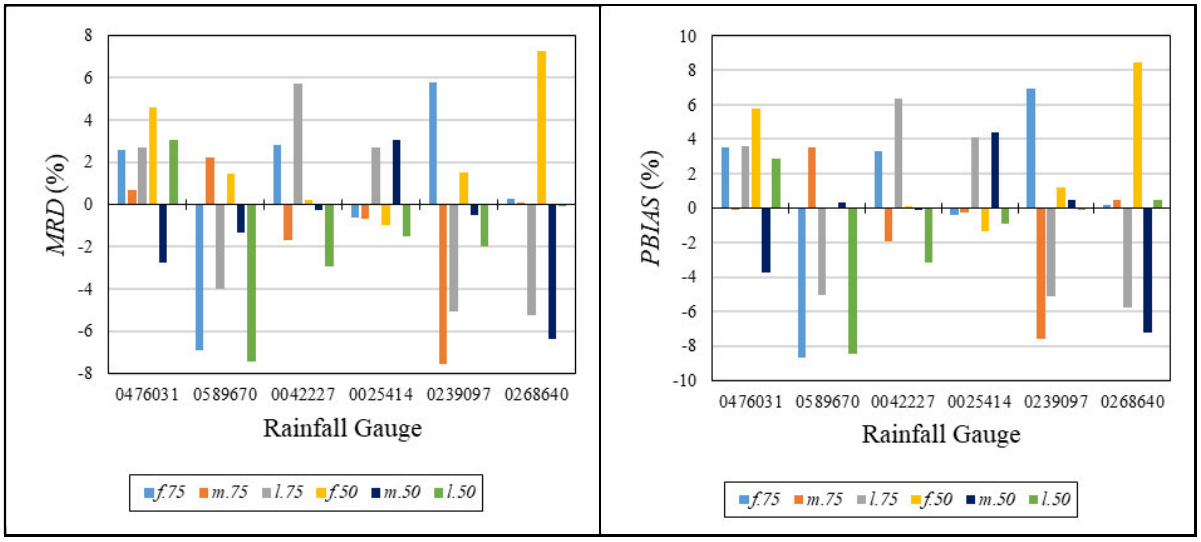

Figure 6.2 *MRD* and *PBIAS* for design rainfall events estimated using scenarios of 75% (*f.75, m.75, l.75)* and 50% (*f.50, m.50, l.50)* of record length

Design rainfall estimates are impacted by up to 8% across all scenarios as indicated by *MARD* values, as shown in Figure 6.3. No individual scenario consistently results in the largest impact on rainfall estimates across all gauges as indicated by varying *MARD* and *NSE* values shown in Figure 6.3 and by varying *MRD* and *PBIAS* values shown in Figure 6.2. For example, *f.50* results in the largest *MARD* for Rainfall Gauge 0476031 whereas *m.75* results in the largest *MARD* for Rainfall Gauge 0239097. This highlights that any scenario can result in the largest impact at a particular gauge. Furthermore, acceptable design rainfall events are estimated when using scenarios *f.75, m.75, l.75, f.50, m.50 or l.50* across all gauges as indicated by *PBIAS* values (< ±10%) and *NSE* values ($0.75 \leq NSE \leq 1.0$) as shown in Figure 6.2 and Figure 6.3 respectively.

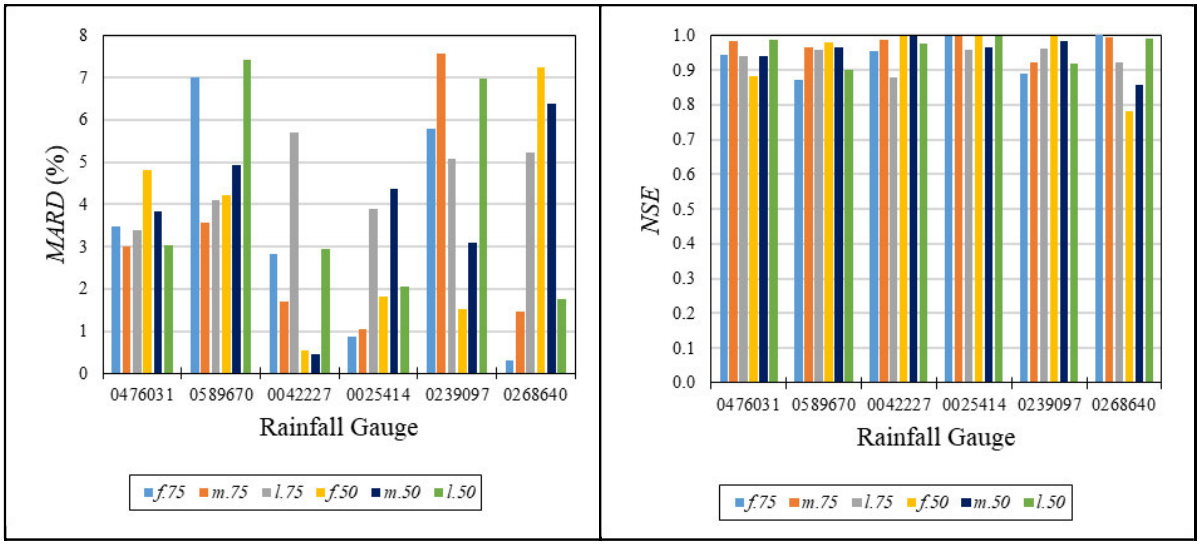

Figure 6.3 *MARD* and *NSE* for design rainfall events estimated using scenarios of 75% (*f.75, m.75, l.75)* and 50% (*f.50, m.50, l.50)* of record length

6.1.3 Design floods

Design floods events are both over- and under-estimated with no definitive trend of consistent over- or under-estimation from an individual scenario across all catchments as indicated by *MRD* and *PBIAS* values shown in Figure 6.4. For example, the largest three over- and under-estimations are not from an individual scenario or catchment indicated by *MRD* and *PBIAS* as evident in in Figure 6.4.

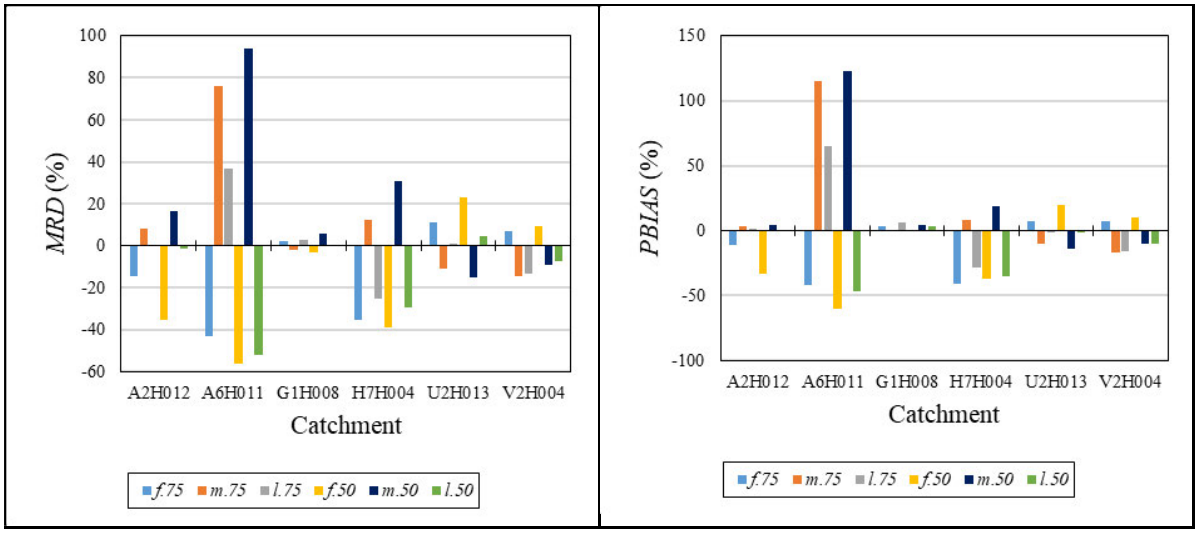

Figure 6.4 *MRD* and *PBIAS* for design flood events estimated using scenarios of 75% (*f.75, m.75, l.75*) and 50% (f.*50, m.50, l.50*) of record length

Design floods are impacted by up to 95% by different periods of reduced record length scenarios as indicated by *MARD* values as shown in Figure 6.5. No scenario consistently results in the largest impact on flood estimates across all catchments as indicated by varying *MARD* and *NSE* values shown in Figure 6.5, and by varying *MRD* and *PBIAS* values shown in Figure 6.4. *NSE* values further indicate that estimated design floods using different periods or record range between a classification of unsatisfactory (NSE < 0.5) to acceptable (0.75 ≤ NSE ≤ 1.0).

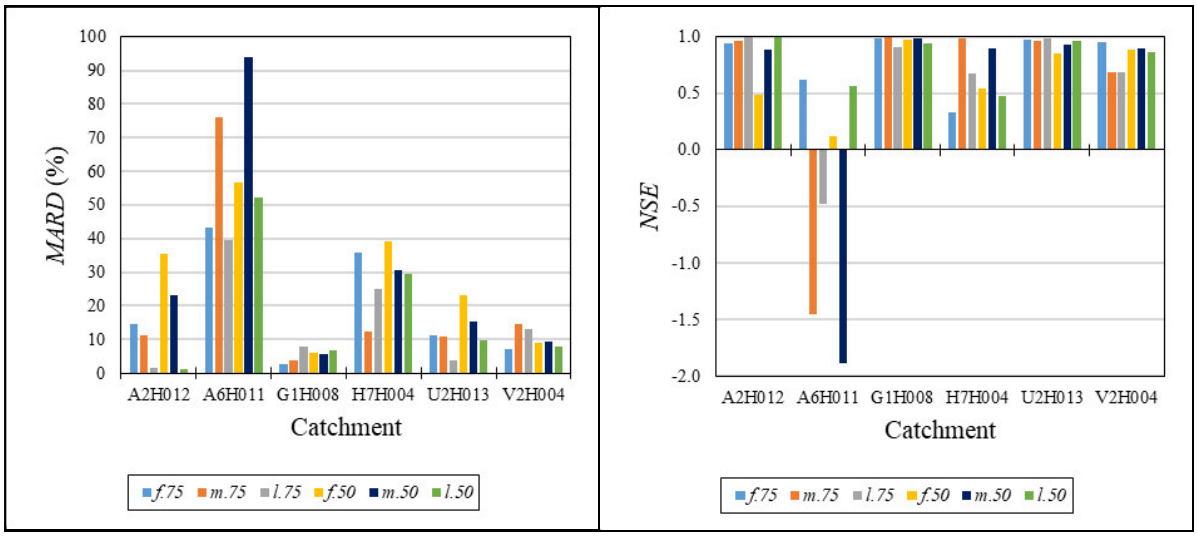

Figure 6.5 *MARD* and *NSE* for design flood events estimated using scenarios of 75% (*f.75, m.75, l.75*) and 50% (*f.50, m.50, l.50*) of record length

6.1.4 Summary of results

The following results are summarised for design rainfall and flood events estimated with scenarios of different periods of record:

(a) Estimated design rainfall are impacted up to 8% by different periods of record and which are classified as acceptable estimates, whereas design floods are impacted by up to 95% when using different periods of record with estimates ranging from unsatisfactory to acceptable.
(b) No specific period of record consistently results in the largest impact on design rainfall and flood events.

This significant difference between the impacts of different periods of record on design rainfall and floods may be attributed to the period of available rainfall (on average between 1880 and 2000) and streamflow (on average between 1954 and 2013) data *i.e.*, rainfall data covers a greater time period and possibly different climate regimes compared to streamflow data. The

identification of climatic regimes is beyond the scope of this study. The difference in impacts may also be attributed to non-stationarity in land use and other aspects associated to the assumption of stationarity as detailed in Section 2.1.2. Assessing stationarity within rainfall and streamflow data is beyond the scope of this study.

6.2 Impact of Reduced Record Length

A brief methodology, detailed results and summary of results on the impact of reduced record length on design rainfall and flood events are provided in Section 6.2.1, Section 6.2.2 and Section 6.2.3, and Section 6.2.4, respectively.

6.2.1 Methodology

Computed statistics (*MRD, MARD, NSE* and *PBIAS*) were averaged across scenarios using 75% of record and across scenarios using 50% of record for each catchment and thereafter averaged across all catchments to provide *Avg. MRD, Avg. MARD, Avg. NSE* and *Avg. PBIAS*. These averaged statistics were used to evaluate and draw comparisons between the overall impact of 75% and 50% of reduced record length on DRE and DFE. Design rainfall results are presented in Section 6.2.2 followed by design floods in Section 6.2.3.

6.2.2 Design rainfall

Design rainfall events are generally under-estimated from the scenarios reduced to 75% and 50% of AMS record length as indicated by negative *Avg. MRD* and *Avg. PBIAS* values as shown in Figure 6.6. Design rainfall events estimated from the 75% and 50% scenarios are impacted by up to 4% indicated by *Avg. MARD* values as shown in Figure 6.7. It is also observed that there is a negligible difference between events estimated by the 75% and 50% scenarios. Estimated rainfall events from the 75% and 50% scenarios are classified as acceptable estimates based on *Avg. PBIAS* ($< \pm 10\%$) and *Avg. NSE* (> 0.75) as shown in Figure 6.6 and Figure 6.7 respectively.

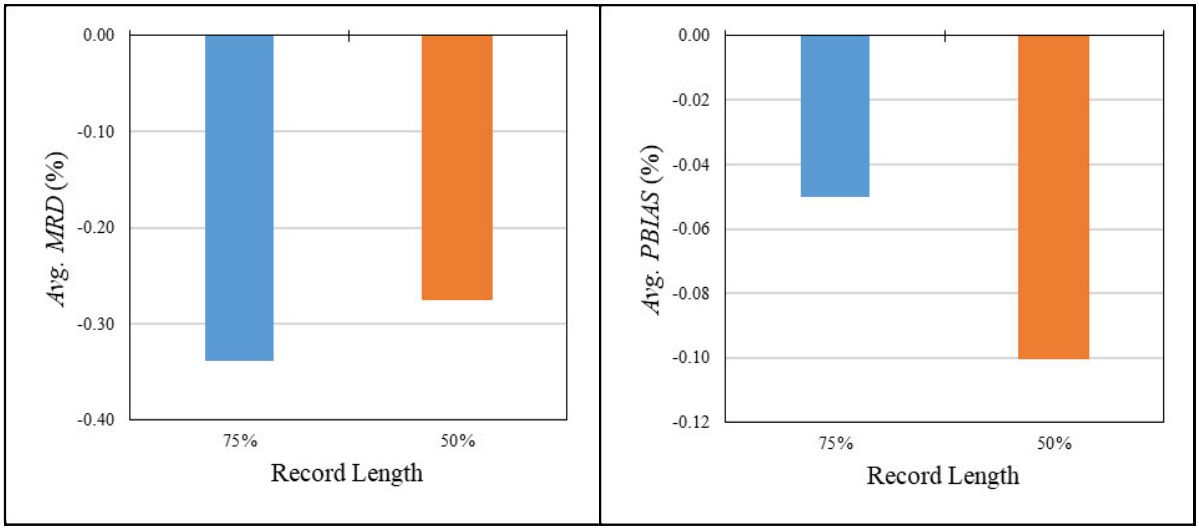

Figure 6.6 *Avg. MRD* and *Avg. PBIAS* for design rainfall events estimated using 75% and 50% of AMS records

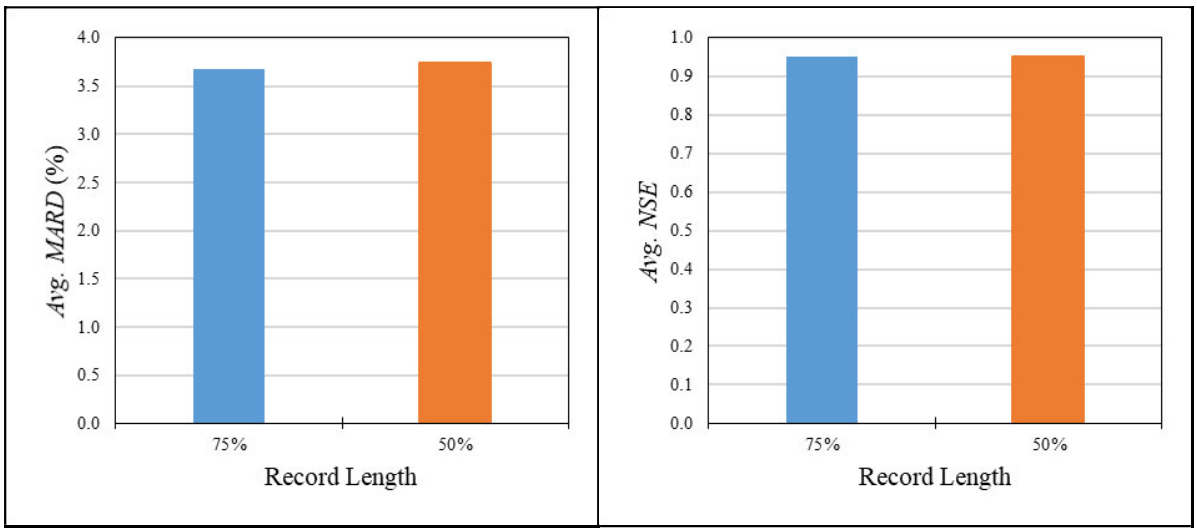

Figure 6.7 *Avg. MARD* and *Avg. NSE* for design rainfall events estimated using 75% and 50% of AMS records

6.2.3 Design floods

Design floods are under-estimated from both the 75% and 50% scenarios, as indicated by negative *Avg, MRD* and *Avg. PBIAS* values as shown in Figure 6.8. Design floods are impacted by up to 24% from reduced AMS record lengths, with the impact being greater from the 50% scenario than the 75% reduction scenario, as indicated by a larger *Avg. MARD* and smaller *Avg. NSE* value as shown in Figure 6.9. Estimates are classified as satisfactory from the 75% and 50% scenarios as indicated by *Avg. NSE* values ($0.50 < \text{NSE} \leq 0.65$) as shown in Figure 6.9.

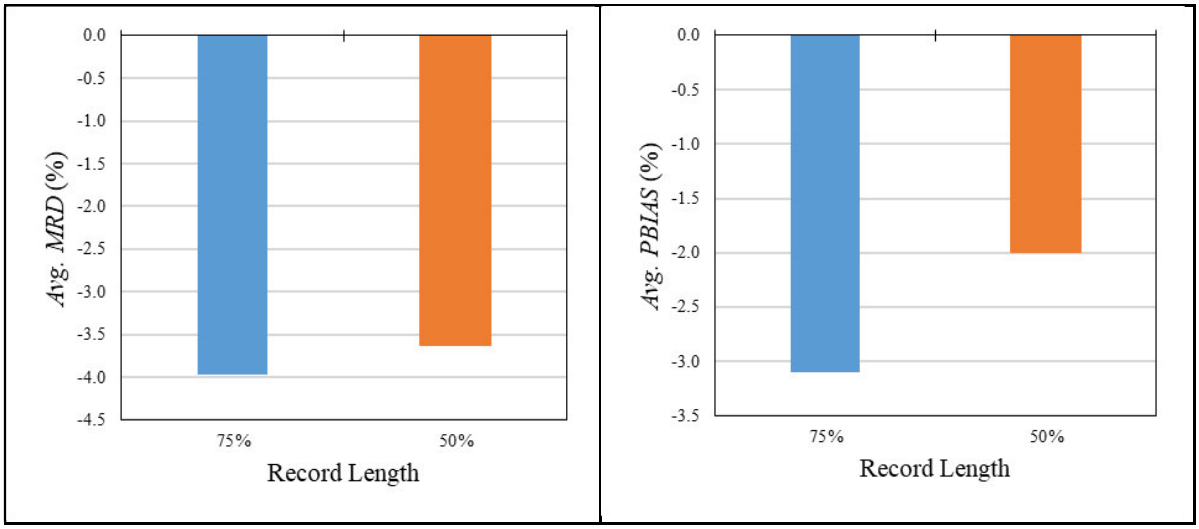

Figure 6.8 *Avg. MRD* and *Avg. PBIAS* for design flood events estimated using 75% and 50% of AMS records

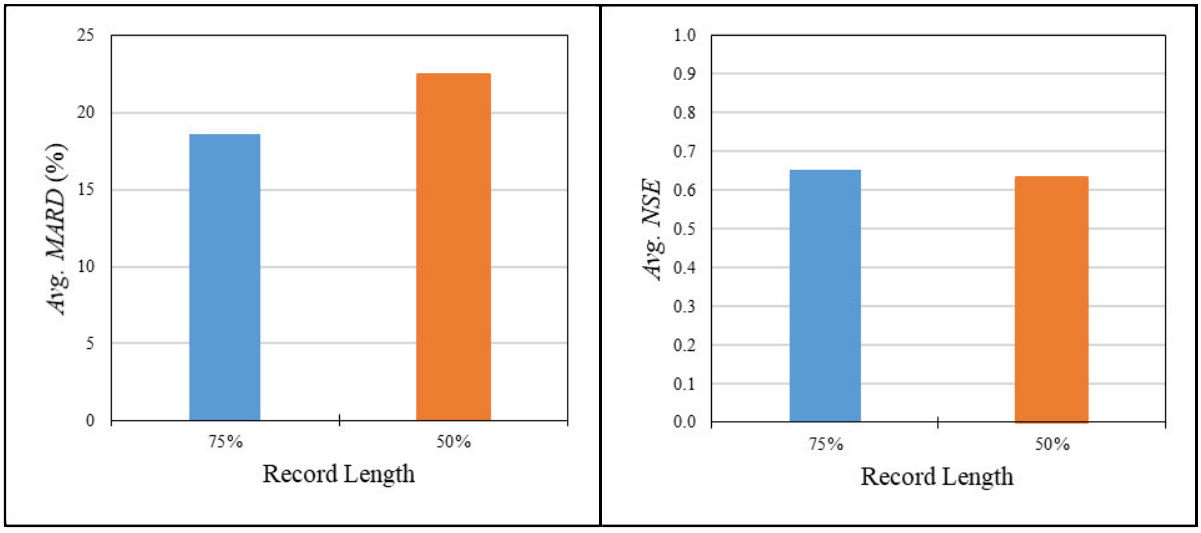

Figure 6.9 *Avg. MARD and Avg. NSE* for design flood events estimated using 75% and 50% of AMS records

6.2.4 Summary of results

The following results are summarised for design rainfall and flood events estimated from 75% and 50% of AMS records:

(a) Design rainfall events are impacted by up to 4% with design floods being much more impacted by up to 24% from reduced AMS record lengths.

(b) There is a negligible difference in impact between design rainfall estimated using the 75% and 50% of AMS records scenario. Design floods are impacted more when using the 50% of AMS record scenario than the 75% of the AMS record scenario.

This significant difference between the impacts of different record lengths on design rainfall and floods may be attributed to period of available rainfall (on average between 1880 and 2000) and streamflow (on average between 1954 and 2013) and the assumption of stationarity as detailed in Section 6.1.4.

6.3 Chapter Summary and Conclusion

Three windows each representing a chronological period of time within 75% and 50% of AMS records were used to evaluate the impact of different periods of available record and the impact of reduced record length on DRE and DFE. Six scenarios (*f.75, m.75, l.75, f.50, m.50, l.50*) were created and comparisons were made between design events estimated from a scenario of reduced record length and the entire record length to evaluate the impact of period of record. Computed statistics were averaged across scenarios using 75% of record (*f.75, m.75, l.75*) and across scenarios using 50% of record (*f.50, m.50 or l.*50) for each catchment and thereafter averaged across all catchments to evaluate the overall impact of 75% and 50% of reduced record length on DRE and DFE.

From the analysis of the impact of period of record, estimated design rainfall events are both over- and under-estimated and are impacted by up to 8% whereas design floods are impacted by up 95% when using various scenarios (*f.75, m.75, l.75, f.50, m.50, l.50*) of periods of record.

From the analysis of the impact of reduced record length, design rainfall events are impacted by up to 4% and design floods impacted by up to 24% from either the 75% and 50% scenario with both design rainfall and floods under-estimated. There is a negligible difference in impact between design rainfall estimates from using the 75% and 50% scenario whereas for design floods, 50% of AMS record length has a greater impact than 75%.

In practice, the use of longer periods of record will improve the applicability of the data for water management purposes as longer periods may cover quasi-periodic fluctuations irrespective of when the observations were recorded. In addition, the use of reduced record length generally results in an under-estimation of design rainfall by up to 4% and floods by 24% thereby reducing the accuracy of infrastructural design which increases the risk of failure

and increases risk to the safety of lives and severe economic, environmental, and social consequences. In research, the use of longer periods of record will improve the accuracy of DRE and DFE thereby increasing the value of research aimed for, *inter alia*, decision and policy making purposes. It is also recommended that this study be expanded to other regions in SA to have more confidence in the findings.

7. IMPACT OF A REDUCED NETWORK DENSITY ON DESIGN RAINFALL AND FLOOD ESTIMATION

This chapter contains a general approach, detailed methodology, discussion of results, and conclusion to assess the impact of a reduced network density on DRE and DFE. The general approach is detailed in Section 7.1 followed by a detailed methodology, discussion of results and summary of results for both systematic and random removal of gauges in Section 7.2 and Section 7.3 respectively. A chapter summary and conclusion is provided in Section 7.4. A schematic of the structure of Chapter 7 is shown in Figure 7.1

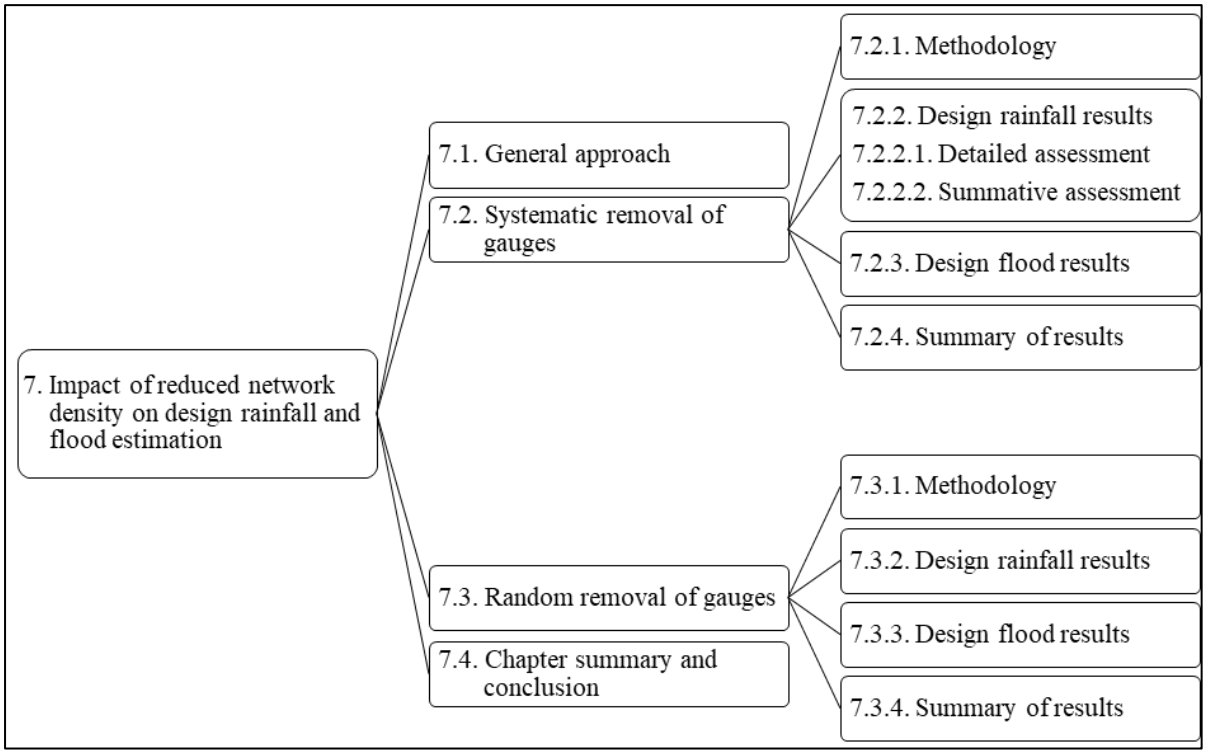

Figure 7.1 Schematic of the structure of Chapter 7

7.1 General Approach

The Index Flood Method (IFM) introduced by Dalrymple (1960) and an index storm-based approach (Hosking and Wallis, 1993; Hosking and Wallis, 1997), referred to as an Index Rainfall Method (IRM) in this study, were used to estimate design rainfall and floods at a predetermined location using different gauge network densities. The IFM estimates design floods at an ungauged site by using hydrological data from gauged sites within a homogenous region and by scaling flood data with a scaling factor referred to as the index flood or index rainfall. The IFM assumes that all scaled flood data within a homogenous region follows the

same PD. The IRM follows the same principles as the IFM. The application of the IFM or IRM involves identifying a homogenous region and then developing regional growth curves within the homogenous region. Regional growth curves describe the relationship between the design estimate for a given return period which is normalised by a scaling factor and the exceedance probability or return period. The scaling factor may be the median or mean of an AMS (Hosking and Wallis, 1997; Chebana and Ouarda, 2009) or the 10-year design estimate (Farquharson *et al.*, 1992; Pilgrim, 2001; Calitz and Smithers, 2020).

In this study, the IFM and IRM were used to estimate design events at a predetermined Point of Interest (PoI) which is a gauged location within the total available gauged network using reduced network densities, based on the assumption that the PoI is ungauged. Design events estimated at the PoI using a reduced network density were then compared to design events estimated at the PoI using the total available network. A gauged network reduction to 75% and 50% of the total available network were chosen to follow the averaged per decade decline of rainfall gauges in South Africa between 1979 and 2009 as calculated from Figure 2.4

The scaling factor required for the IFM and IRM chosen in this study is the 2-year design event for rainfall (R_2) and floods (Q_2). The 2-year design event scaling factor is an adaptation of the 10-year design event scaling factor as used by Farquharson *et al.* (1992) and Pilgrim (2001). The 10-year design event scaling factor has proved to be successful by Farquharson *et al.* (1992), Pilgrim (2001) and Calitz and Smithers (2020), and due to its simplicity of use has been adopted in this study. A study on the possible impacts of the choice of the index value on the performances of the estimated design events was not an objective of this study. Homogenous regions required for the IFM as shown in Figure 7.2 were identified as overlapping regions between the HRU (1972) veld types, RMF K-regions (Kovacs, 1988)and the climatic zones used by Gericke (2015). Homogenous regions for the IRM were defined as overlapping regions between the clusters developed by Smithers (1996) and climate zones used by Gericke (2015). Three homogenous regions were chosen for the IFM and for the IRM. Streamflow gauges and rainfall gauges within these homogenous regions were then selected and used in the index flood and rainfall approaches. The rainfall gauges and associated attributes are provided in Table D1.1 and mapped in Figure D1.1, Figure D1.2 and Figure D1.3 with streamflow gauges provided in Table D1.2 and mapped Figure D1.4, Figure D1.5, and Figure D1.6 in Appendix D1. For ease of reference, the three homogenous regions identified for the IFM and IRM are

referred to as the climatic zones used by Gericke (2015) within which they are situated, *i.e.* ESC, NI and SWC.

Growth curves were derived within each homogenous region by using two gauge density reduction methods, *i.e.* a systematic removal and a random removal of gauges and design events estimated using the GEV PD. The systematic removal of gauges involved removing the closest gauges from a PoI and thereafter removing the furthest gauges within each homogenous region, and the random removal of gauge method involved removing gauges at random around a PoI. Details of the systematic and removal of gauges are provided in Section 7.2.1 and Section 7.3.1 respectively. The GEV PD was chosen to provide a conservative approach due to the degree of impact as detailed in Chapter 4, and due its widespread use as highlighted in Section 4.1.1. For each gauge reduction method, the median growth curves from each site was selected for each scenario of the gauged network to estimate the 2-, 5-, 10-, 20-, 50-, 100- and 200-year design event values. As an example, the index rainfall growth curve for the NI region is shown in Figure 7.3, in which R_T represents the T-year design rainfall estimate and R_2 represents the 2-year design rainfall estimate. Relationships between R_2 and MAP, and Q_2 and catchment area were developed to estimate R_2 or Q_2 at the PoI. Q_2 represents the 2-year design flood estimate. MAP was obtained from the Lynch (2004) database and catchment areas were obtained from the Nathanael (2015) study. An example of the R_2 vs MAP relationship for Region NI is shown in Figure 7.4. R_2 or Q_2 and the appropriate growth curve factor (R_T/R_2 or Q_T/Q_2) is then used to estimate the desired R_T or Q_T at the PoI using the growth curve using Equation 7.1 assuming that the PoI is ungauged. The index rainfall and flood approach were applied to a systematic and random gauge removal method, which are detailed in Section 7.2 and Section 7.3 respectively.

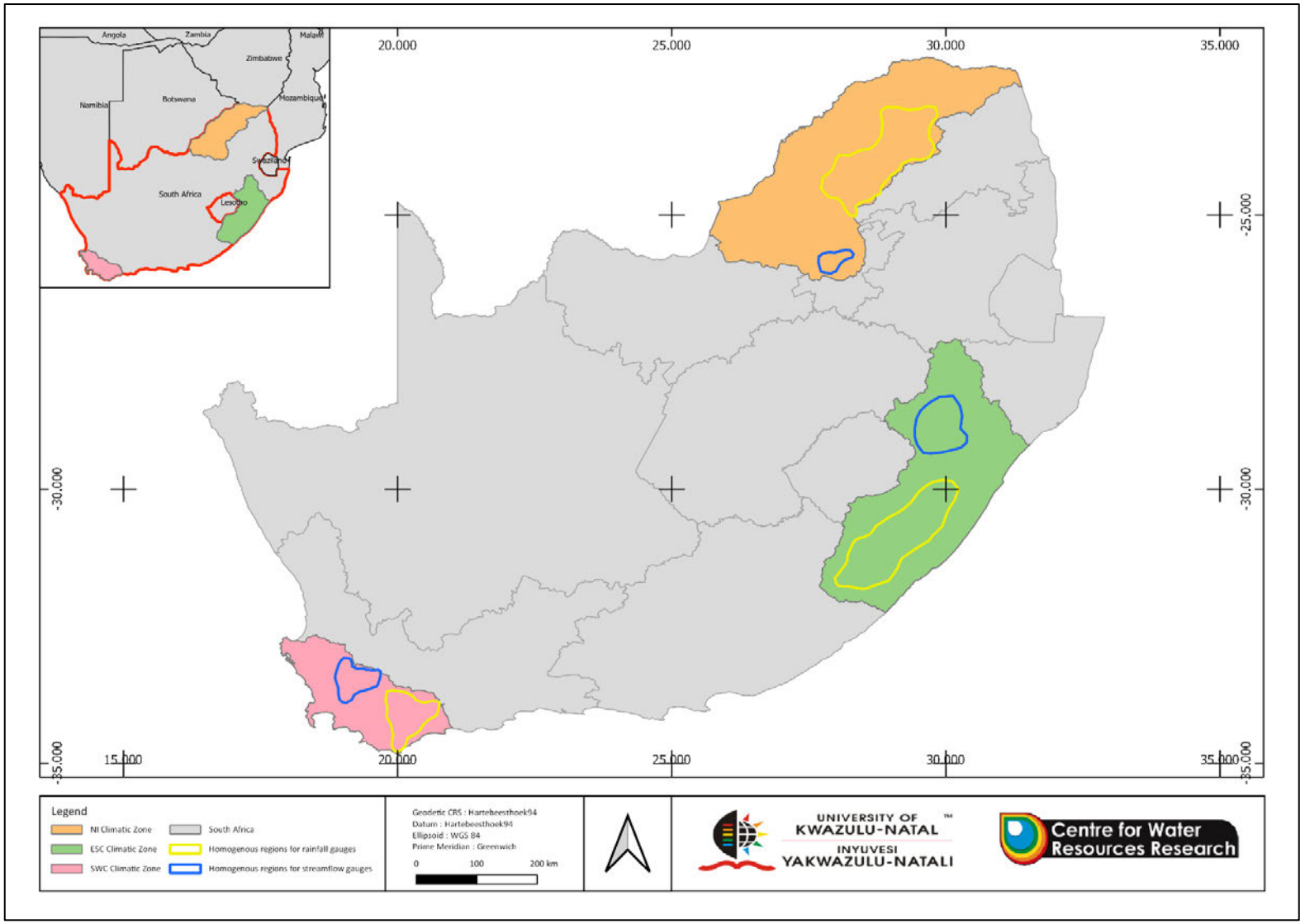

Figure 7.2 Homogenous regions used in the IFM and IRM

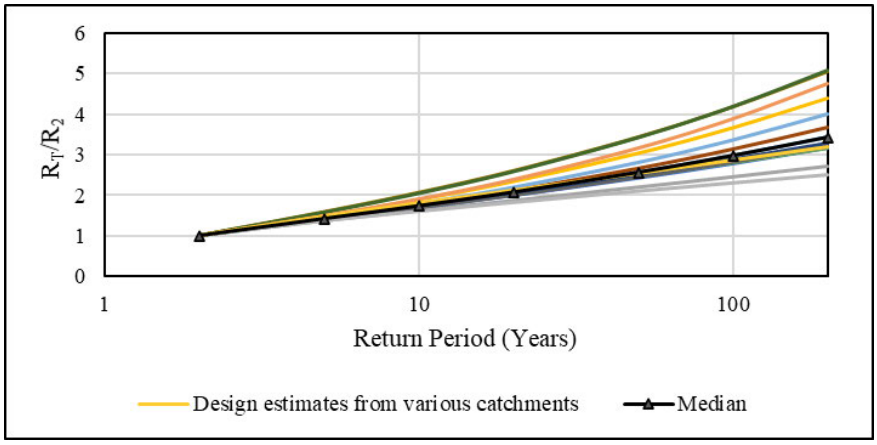

Figure 7.3 Growth curve for design rainfall in Region NI

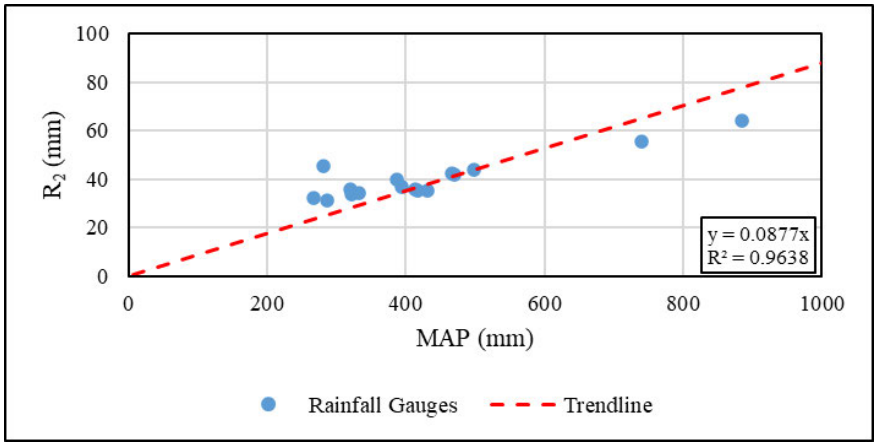

Figure 7.4 R_2 vs MAP for the NI homogenous region

$$P(T,N) = P_2 . q_R(T,N) \quad (7.1)$$

where:

P = Design rainfall or flood estimate (mm or $m^3.s^{-1}$),

N = Percentage of gauged network (100%, 75% or 50%),

P_2 = 2-year design rainfall or flood event (mm or $m^3.s^{-1}$), and

q_R = Index rainfall or flood growth curve.

7.2 Systematic Removal of Gauges

This analysis aimed to evaluate the impact of a reduced gauged network on DRE and DFE and the subsequent impact of using neighbouring gauges with varying proximity from a pre-selected PoI. The method, results and summary of results on the application of a systematic gauge removal method are provided in Section 7.2.1, Section 7.2.2, Section 7.2.3 and Section 7.2.4 respectively.

7.2.1 Methodology

The gauged networks were reduced to 75% and 50% of the total available network to follow historic gauged reduction trends in SA as detailed in Section 7.1. Four scenarios (*Closest.75%, Closest.50%, Furthest.75%* and *Furthest.50%*) were created to represent this systematic removal of gauges as explained in Table 7.1.

Table 7.1 Scenarios of reduced gauge density using a systematic removal of gauges

Scenario ID	Explanation
Closest.75%	Closest 75% of the total available gauged network around the PoI.
Closest.50%	Closest 50% of the total available gauged network around the PoI.
Furthest.75%	Furthest 75% of the total available gauged network around the PoI.
Furthest.50%	Furthest 50% of the total available gauged network around the PoI.

The index approach explained in Section 7.1 was then applied using each systematic gauge reduction scenario to estimate the design rainfall and floods events at the PoI. Comparison were then drawn between rainfall and flood events estimated from each of the four scenarios, and from the total available network. The total available rainfall and flow networks within each homogenous region defined in Section 7.1 were mapped on QGIS 10.3.1 and provide in Appendix D1. Thereafter, the PoI within each homogenous region rainfall network was chosen based on centrality within the homogenous region and having a MAP similar to the median MAP from all gauges within the homogenous region. Similarly, the PoI within each homogenous region streamflow network was chosen based on centrality within the homogenous region and having a catchment area similar to the median catchment area from all gauges within the homogenous region.

It was assumed that the total available gauged network resulted in the most accurate and representative design events compared to events estimated from a reduced gauged network. It was assumed that no outliers were present in the initial AMS data and that all observed rainfall and streamflow data fitted the GEV PD. Stationarity within each observed rainfall and streamflow dataset has also been assumed. It is acknowledged that: (a) factors such as rainfall type, elevation, seasonality of precipitation and land use have not been accounted for within each homogenous region, (b) the homogeneity of the growth curves within each region has not

been tested, and (c) any uncertainties in the observed data have not been taken into consideration in this analysis.

MRD, *MARD*, *PBIAS* and *NSE* were computed and used to evaluate the impact of network density on DRE and DFE. *RD* as described in Section 4.1.1 was calculated using Equation 7.2. *MRD* were then calculated as the arithmetic mean of the *RD* values across all return periods and *MARD* was calculated as the arithmetic mean of the absolute *RD* across all return periods. *NSE* and *PBIAS* were calculated using Equation 7.3 and Equation 7.4 respectively.

$$RD_{den.} = \frac{[E_{rand.,T,i} - E_{100,T,i}]}{E_{100,T,i}} \times 100 \quad (7.2)$$

where:

$RD_{den.}$ = RD calculated between design rainfall or flood events estimated using a scenario of reduced network density and entire network density (%),

$E_{rand.}$ = Design rainfall or flood estimate from *Random.75%* or *Random.50%* (mm or $m^3.s^{-1}$),

i = Set of gauges (1, 2, 3…100), and

E_{100} = Rainfall or flood estimated using 100% of gauged network (mm or $m^3.s^{-1}$).

$$NSE = \left(1 - \frac{\sum_{T=1}^{7}(E_{100,T,i} - E_{rand,T,i})^2}{\sum_{T=1}^{7}(E_{100,T,i} - \bar{E})^2}\right) \quad (7.3)$$

where:

NSE = Statistic quantifying the fit of estimated design events against the 1:1 line, and

$\bar{E}$ = Avg. rainfall or flood estimated by the entire record length (mm or $m^3.s^{-1}$).

$$PBIAS = \sum_{T=1}^{7}\left[\frac{(E_{100,T,i} - E_{rand,T,i}) \times 100}{E_{100,T,i}}\right] \quad (7.4)$$

The following two assessments were performed to evaluate the impact a systematic removal of gauges:

(a) Detailed assessment: Statistics (*MRD, MARD, NSE* and *PBIAS*) were computed per scenario *Closest.75%, Closest.50%, Furthest.75%* and *Furthest.50%* for each homogenous region.

(b) Summative assessment: Computed statistics were averaged ($Avg._s$) across scenarios with 75%, 50%, closest and furthest gauged network for each homogenous region and thereafter averaged across all catchments. Statistics from *Closest.75%* and *Furthest.75%* were averaged and compared to averaged statistics from *Closest.50%* and

Furthest.50% to provide an overall density impact when using 75% and 50% of the gauged network. Statistics from *Closest.75%* and *Closest.50%* were averaged and compared to averaged *Furthest.75%* and *Furthest.50%* to provide an overall impact for proximity of gauges.

Sections 7.2.2 and Section 7.2.3 presents results for the impact of a systematically reduced network density on DRE and DFE respectively.

7.2.2 Design rainfall

Design rainfall results for the detailed assessment are provided in Section 7.2.2.1 with results for the summative assessment provided in Section 7.2.2.2.

7.2.2.1 Detailed assessment

Design rainfall estimated by the *Closest.75%* and *Closest.50%* scenarios are consistently under-estimated across all regions indicated by negative *MRD* and *PBIAS* values, whereas estimates by the *Furthest.75%* and *Furthest.50%* scenarios are consistently over-estimated as indicated by positive *MRD* and *PBIAS* as shown in Figure 7.5. Design rainfall events by the *Closest.50%* scenario are the most under-estimated compared to all other scenarios across all homogenous regions, and events from the *Furthest.50%* scenario are the most over-estimated as indicated by largest positive *MRD* and *PBIAS* values shown in Figure 7.5.

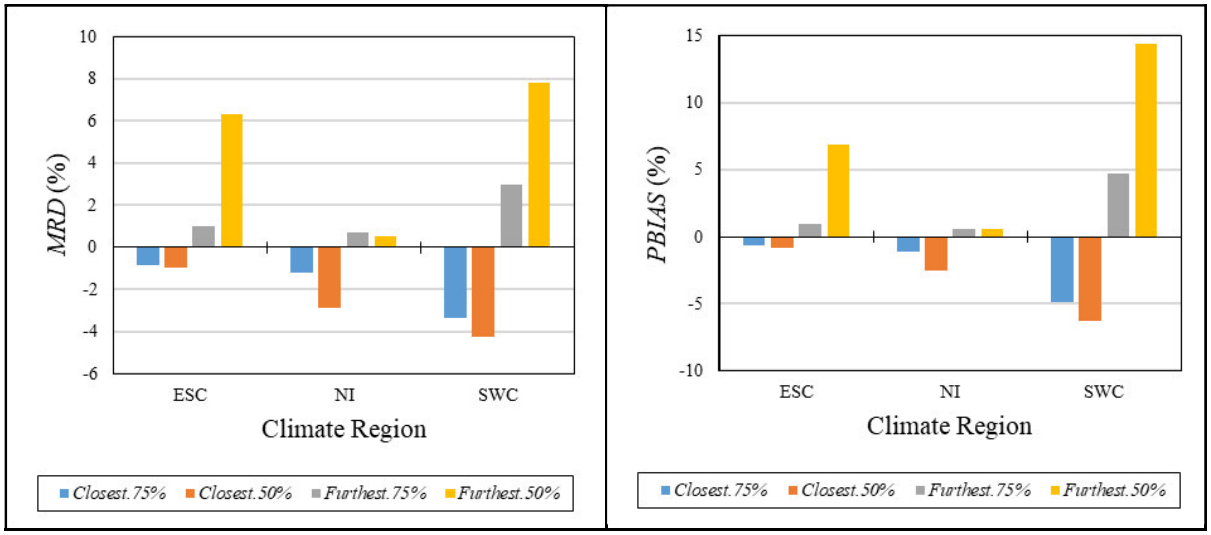

Figure 7.5 *MRD* and *PBIAS* for design rainfall events estimated using scenarios (*Closest.75%*, *Closest.50%*, *Furthest.75%* and *Furthest.50%*) of a systematically reduced gauge network

Design rainfall events are impacted by up to 9% as indicated by *MARD* values as shown in Figure 7.6. Events are most impacted by the *Furthest.50%* scenario in the ESC and SWC regions and by the *Closest.50%* scenario in the NI regions as indicated by the largest *MARD* and smallest *NSE* values as shown in Figure 7.6. Furthermore, acceptable events are estimated by all scenarios across all regions as classified by *PBIAS* values (< ±15%) and *NSE* values (> 0.75) as shown in Figure 7.5 and Figure 7.6 respectively.

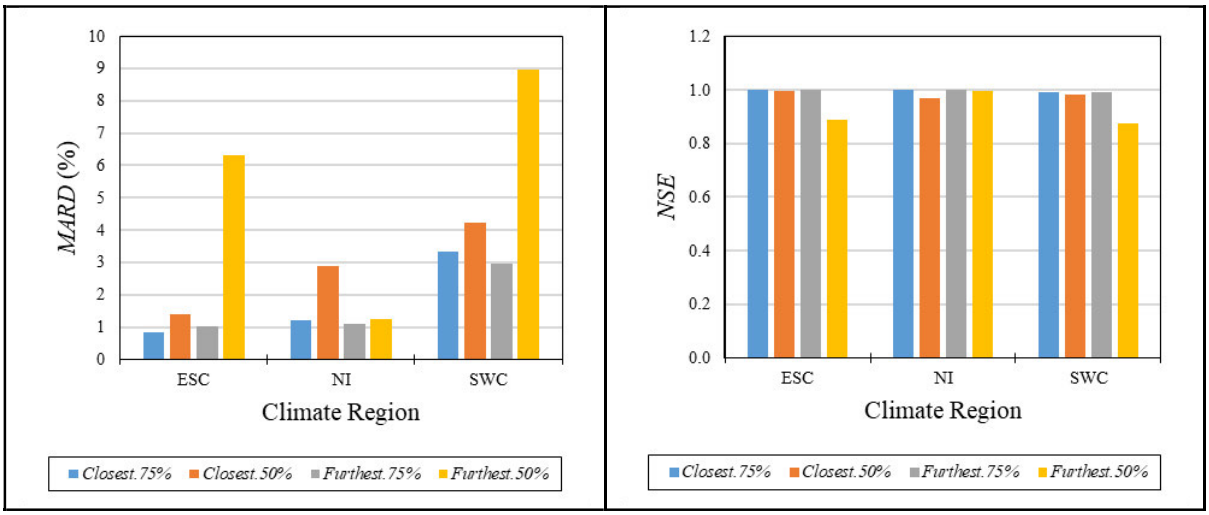

Figure 7.6 *MARD* and *NSE* for design rainfall events estimated using scenarios (*Closest.75%*, *Closest.50%*, *Furthest.75%* and *Furthest.50%*) of a systematically reduced gauge network

7.2.2.2 Summative assessment

On average, scenarios using *75%* of rainfall gauges and the *closest* rainfall gauges are under-estimated as indicated by negative *Avg.$_S$ MRD* and *Avg.$_S$ PBIAS* values as shown Figure 7.7, and the converse is true for scenarios using *50%* and the *furthest* rainfall gauges

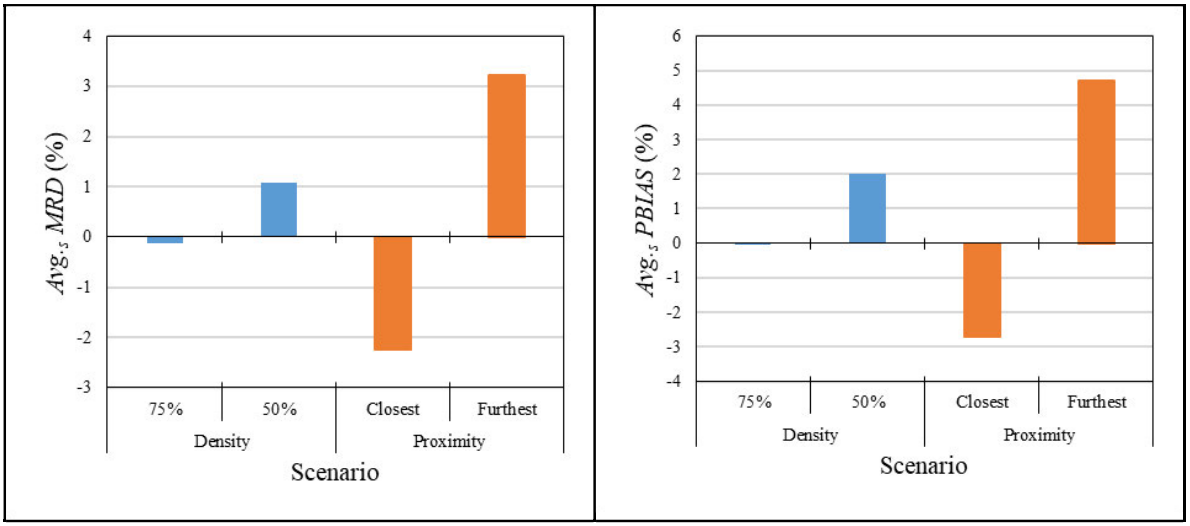

Figure 7.7 *Avg.s MRD* and *Avg.s PBIAS* for design rainfall events estimated using scenarios (*Closest.75%*, *Closest.50%*, *Furthest.75%* and *Furthest.50%*) of a systematically reduced gauge network

Design rainfall events are impacted by up to 4.2% from scenarios using *50%* of the gauged network which are greater than impacts from scenarios using *75%* of the gauged network *i.e.* impacted by up to 2% as indicated by larger $Avg._s$ *MARD* values and smaller $Avg._s$ *NSE* values as shown in Figure 7.8, and by larger $Avg._s$ *MRD* and $Avg._s$ *PBIAS* values, as shown in Figure 7.7. Design rainfall events are impacted by up to 4% from scenarios using the *furthest* gauged network which are greater than impacts from scenarios using the *closest* gauged network *i.e.* impacted by up to 2.5%.

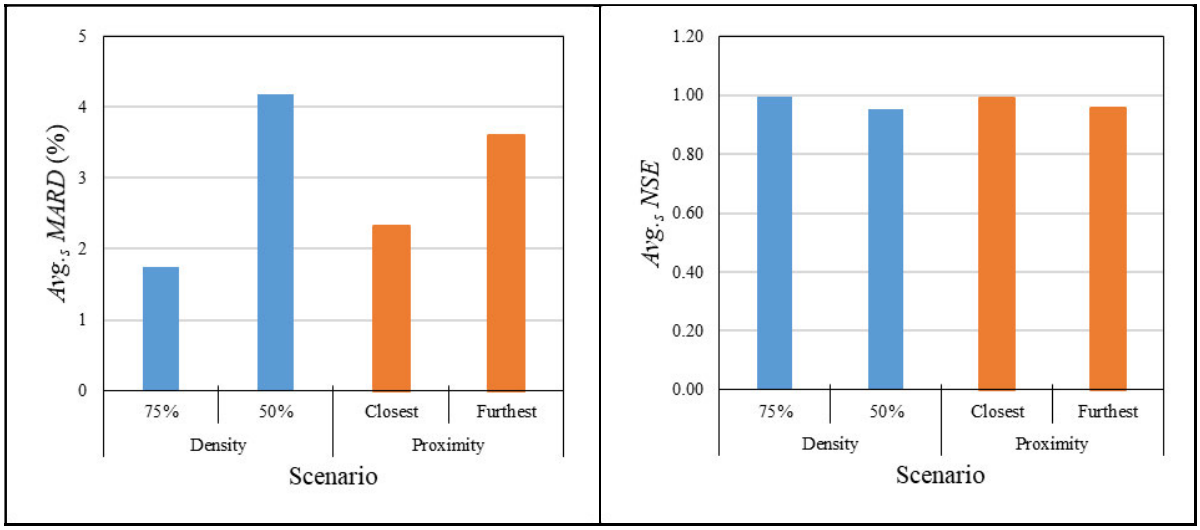

Figure 7.8 *Avg.s MARD* and *Avg.s NSE* for design rainfall events estimated using scenarios (*Closest.75%*, *Closest.50%*, *Furthest.75%* and *Furthest.50%*) of a systematically reduced gauge network

7.2.3 Design floods

Design flood results for the summative assessment are provided in this section and results for the detailed assessment provided in Figure D2.1 in Appendix D2. Design flood events are generally under-estimated from scenarios using *75%* of streamflow gauges and the *furthest* streamflow gauges as indicated by positive *Avg.s MRD* and *Avg.s PBIAS* values, and over-estimated from scenarios using *50%* of streamflow gauges and the *closest* streamflow gauging network as indicated by negative *Avg.s MRD* and *Avg.s PBIAS* values as shown in Figure 7.9.

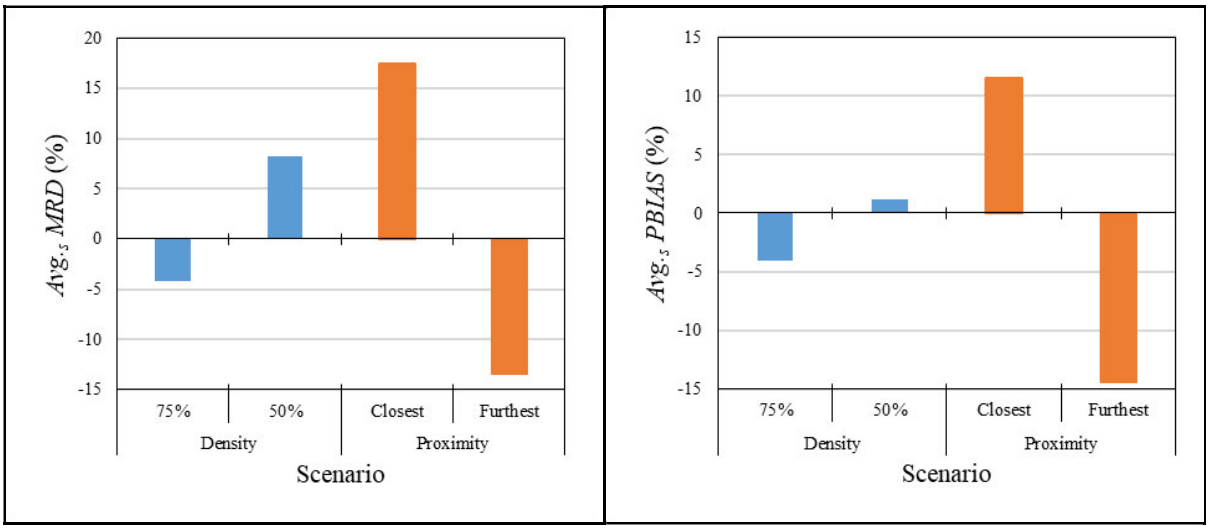

Figure 7.9 *Avg.s MRD* and *Avg.s PBIAS* for design flood events estimated using scenarios (*Closest.75%*, *Closest.50%*, *Furthest.75%* and *Furthest.50%*) of a systematically reduced gauge network

Design floods events are impacted by up to 28% from scenarios using *50%* of the gauged network which are greater than impacts from scenarios using *75%* of the gauged network *i.e.* impacted by up to 17%, and impacted by up to 24% from scenarios using the *closest* gauged network which are slightly greater than impacts from scenarios using the *furthest* gauged network *i.e.* impacted by up to 22% as indicated by larger *Avg.s MARD* values and smaller *Avg.s NSE* values as shown in Figure 7.10, and by larger *Avg.s MRD* values as shown in Figure 7.9.

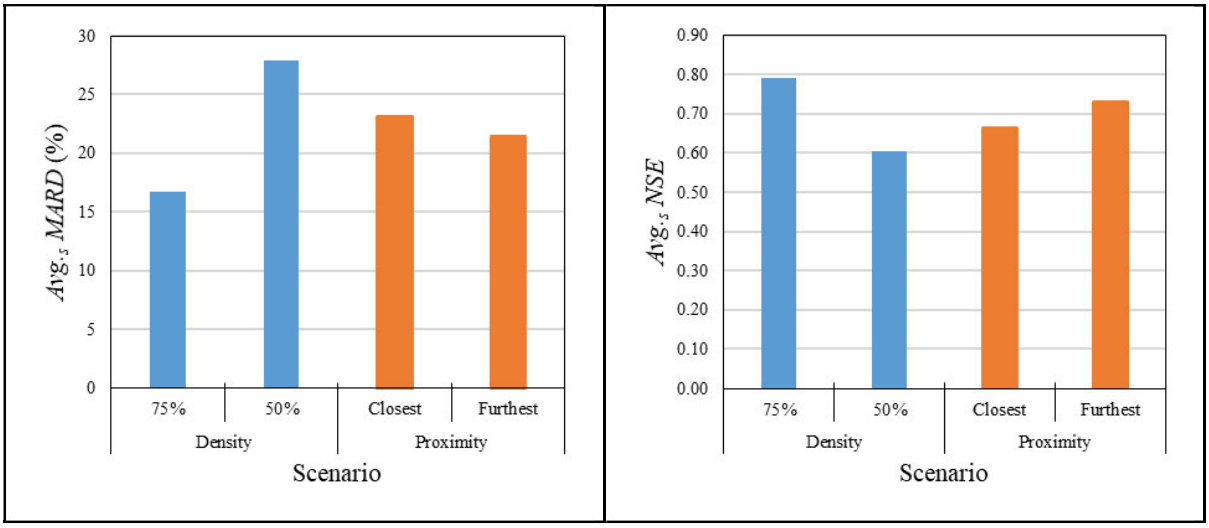

Figure 7.10 *Avg.s MARD* and *Avg.s NSE* for design flood events estimated using scenarios (*Closest.75%*, *Closest.50%*, *Furthest.75%* and *Furthest.50%*) of a systematically reduced gauge network

7.2.4 Summary of results

The following results are summarised for design rainfall and flood events estimated from the systematic reduction of gauged networks:

(a) Design rainfall events are impacted by up to 4.2% from scenarios using *50%* of the gauged network which are greater than impacts from scenarios using *75%* of the gauged network *i.e.* impacted by up to 2%.
(b) Design floods events are impacted by up to 28% from scenarios using *50%* of the gauged network which are greater than impacts from scenarios using *75%* of the gauged network *i.e.* impacted by up to 17%.
(c) Design rainfall events are impacted by up to 4% from scenarios using the *furthest* gauged network which are greater than impacts from scenarios using the *closest* gauged network *i.e.* impacted by up to 2.5%.
(d) Design floods events are impacted by up to 24% from scenarios using the *closest* gauged network which are slightly greater than impacts from scenarios using the *furthest* gauged network *i.e.* impacted by up to 22%.

The greater impact in design events estimated by the *75%* gauged network scenario compared to the *50%* gauged network scenario may be a result of a reduced regionalisation efficiency (Lebecherel *et al.*, 2016) as mentioned in Section 2.1.2. The greater impact in design rainfall estimates by the *furthest* gauged network compared to the *closest* gauged network may be a result of the principle of Tobler's first law of geography (Tobler, 1970) as mentioned in Section

2.1.2. The greater impact in design flood estimates by the *closest* gauged network compared to the *furthest* gauged network is contrary to the trend on design rainfall estimates and contrary to findings of the HDes method applied by (Lebecherel *et al.*, 2016). This contrary trend may be as a result of factors such as rainfall type, seasonality of precipitation, catchment elevation, geology, soil, topography and land use not being accounted for as mentioned in Section 7.2.1, and an insufficient number of study sites used in this study. . Design floods are more impacted by a reduced gauged network than design rainfall estimated which may be attributed to: (a) the use of more gauges in the rainfall network than the streamflow network, which may affect the regionalisation efficiency, (b) any uncertainties or inaccuracies in the delineated homogenous regions, and (c) the period of available records and stationarity as detailed in Section 6.1.4.

7.3 Random Removal of Gauges

This section focuses on evaluating the impact of a reduced rainfall and streamflow gauge network on DRE and DFE by the random removal of gauges.

7.3.1 Methodology

A PoI required for the IRM and IFM within each of the three homogenous regions as detailed in Section 7.1 were chosen as the gauge having the median MAP or catchment area. Thereafter, gauges were removed at random such that 75% of the network remained for analysis. This process of random removal of gauges was repeated 100 times to create 100 networks with each having 75% of the total available gauged network. Similarly, 100 networks with each having 50% of the total available gauged network were also generated. Two scenarios of a randomly reduced gauged network were created *i.e.* containing 75% and 50% of the total gauged network area and are referred to as *Random.75%* and *Random.50%*, respectively. The IRM and IFM were then applied 100 times to each of the 100 networks having 75% or 50% of gauge network to estimate the 2-, 5-, 10-, 20-, 50-, 100- and 200-year estimates at a PoI. Comparisons were then drawn between the estimated design events at the PoI when using the total available network and for each reduced network from *Random.75%* and *Random.50%*. *MRD*, as described in Section 7.2.1, was computed for each of the 100 networks of gauges and an *Avg. MRD* was then calculated as the average across all 100 *MRD* values. Similarly, *Avg. MARD*, *Avg. PBIAS* and *Avg. NSE* were computed. The assumptions made in the systematic gauge removal method as detailed in Section 7.2.1 are also made in the method of random gauge removal.

Section 7.3.2 and Section 7.3.3 presents results for the impact of a randomly reduced network density on estimated design rainfall and flood events respectively.

7.3.2 Design rainfall

Negative *Avg. MRD* and *Avg. PBIAS* values were computed for design rainfall events from *Random.75%* and *Random.50%* across all homogenous regions indicating a general under-estimation as shown in Figure 7.11.

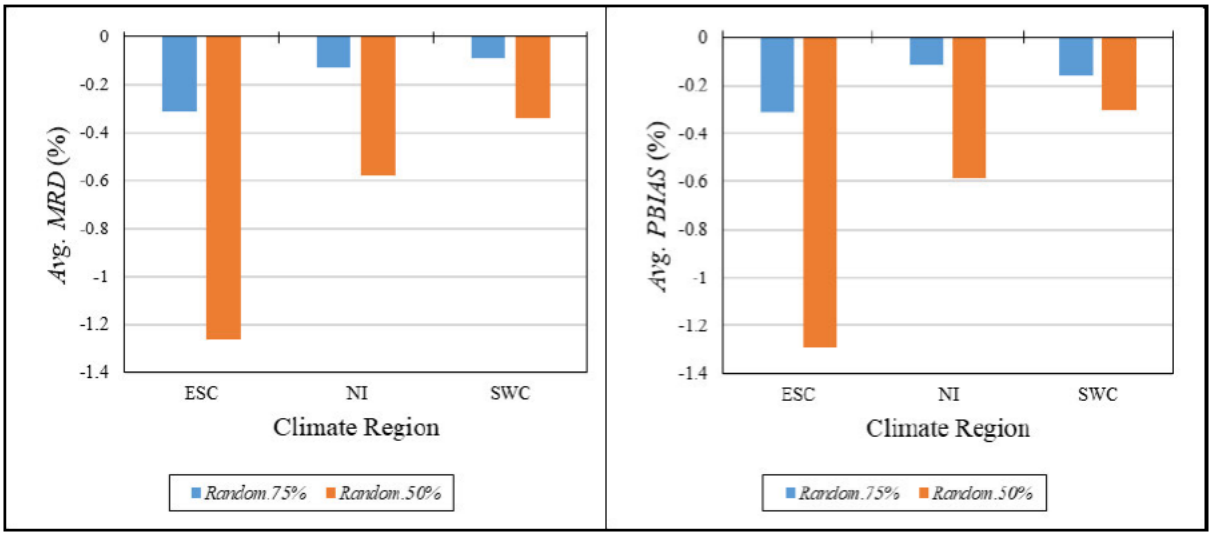

Figure 7.11 *Avg. MRD* and *Avg. PBIAS* of design rainfall events computed using a 75% and 50% randomly reduced gauge network

Design rainfall is impacted by up to 4.5% when estimated from the *Random.50%* scenario and by up to 2.2% when estimated from the *Random.75%* scenario across all regions as indicated by *Avg. MARD* values as shown in Figure 7.12, and this greater impact is also indicated by larger *Avg. MRD* and *Avg. PBIAS* as shown in Figure 7.11. It is further observed that acceptable design rainfall events are estimated from the *Random.75%* and *Random.50%* scenarios across all homogenous regions as classified by *Avg. PBIAS* values (< 10%) and *Avg. NSE* values (> 0.75) as shown in Figure 7.11 and Figure 7.12, respectively.

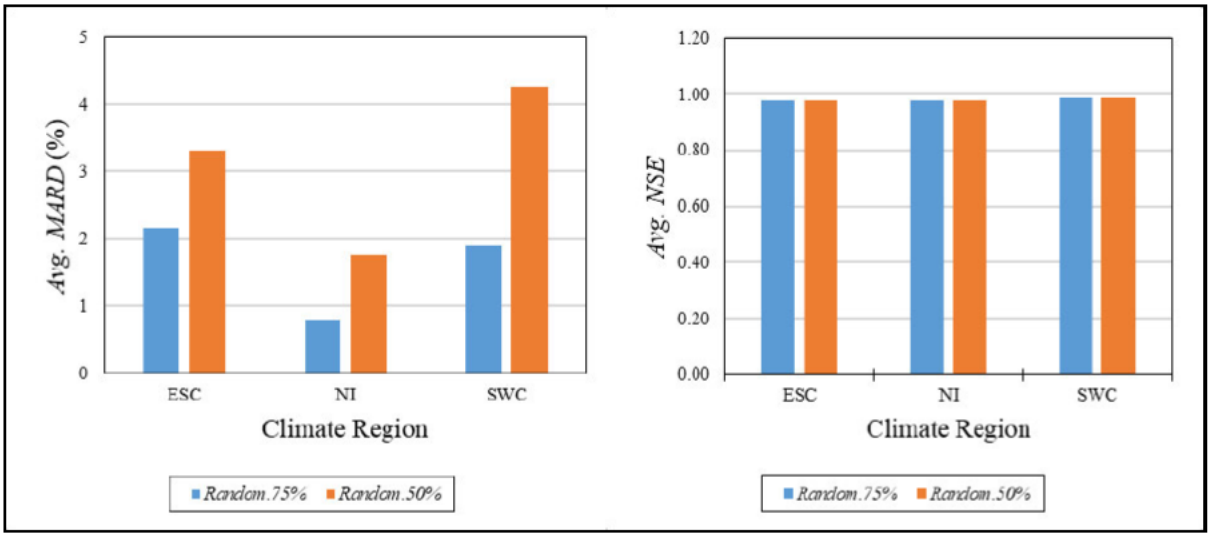

Figure 7.12 *Avg. MARD* and *Avg. NSE* of design rainfall events computed using a 75% and 50% randomly reduced gauge network

7.3.3 Design floods

Design flood events are over-estimated from both the *Random.75%* and *Random.50%* scenarios across all climate regions as indicated by positive *Avg. MRD* and *Avg. PBIAS* values as shown in Figure 7.13.

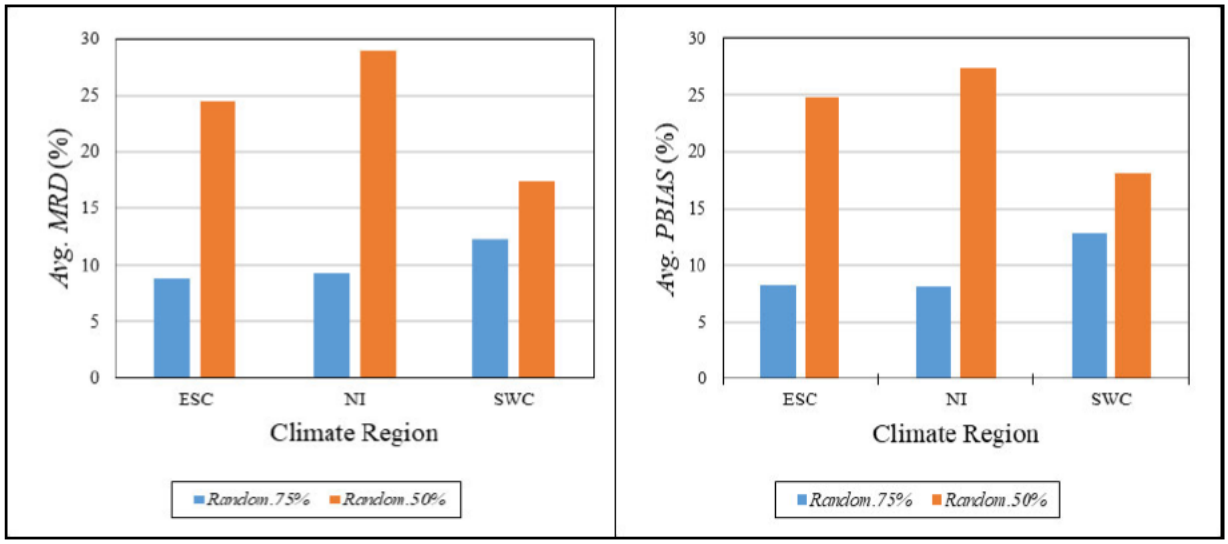

Figure 7.13 *Avg. MRD* and *Avg. PBIAS* of design flood events computed using a 75% and 50% randomly reduced gauge network

Design flood events are impacted by up to 60% from the *Random.50%* scenario which are greater than impacts from the *Random.75%* scenario *i.e.* impacted by up to 30% across all regions as indicated by *Avg. MARD* and smaller *Avg. NSE* values evident in Figure 7.14, and

larger *Avg. MRD* and *Avg. PBIAS* values in Figure 7.13. Design floods are classified as satisfactory from the *Random.50%* scenario and acceptable from the *Random.75%* scenario by *Avg. NSE* values, *i.e.* NSE < 0.5 and NSE > 0.75 respectively, as shown in Figure 7.14.

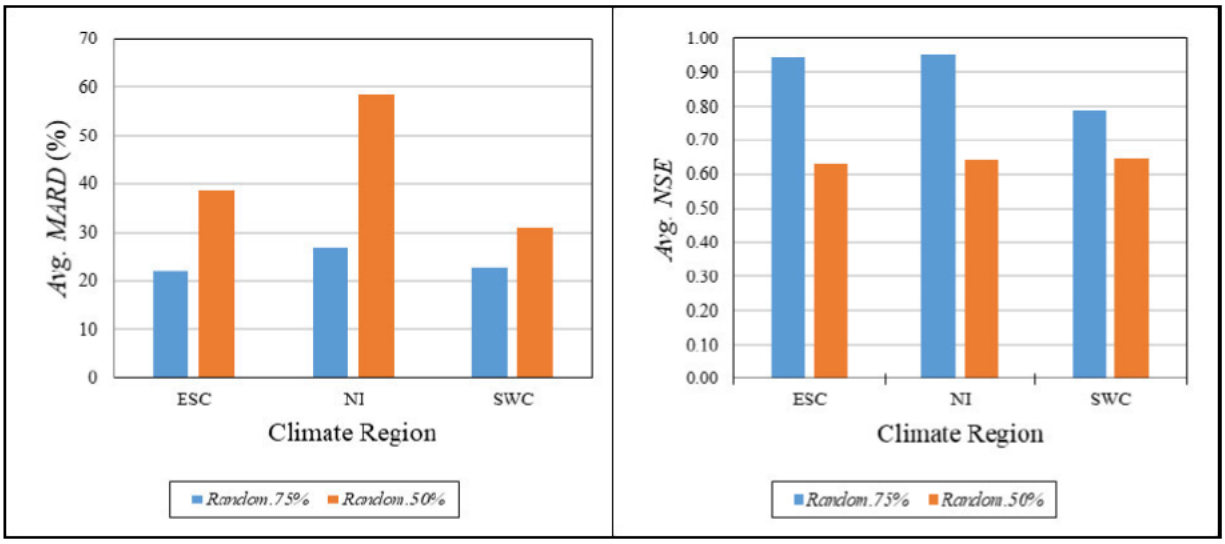

Figure 7.14 *Avg. MARD* and *Avg. NSE* of design flood events computed using a 75% and 50% randomly reduced gauge network

7.3.4 Summary of results

The following results are summarised for design rainfall and flood events estimated from the random reduction of gauged networks:

(a) Design rainfall events are impacted by up to 4.5% from the *Random.50%* scenario which are greater than impacts from the *Random.75%* scenario *i.e.* impacted by up to 2.2% across all regions.
(b) Design flood events are impacted by up to 60% from the *Random.50%* scenario which are greater than impacts from the *Random.75%* scenario *i.e.* impacted by up to 30% across all regions.

The greater impact in design events estimated by the 75% gauged network scenario compared to the 50% gauged network scenario may be a result of factors as detailed in Section 7.2.4.

7.4 Chapter Summary and Conclusion

Two methods were applied to evaluate the impact of a reduced network density on design rainfall and flood events *i.e.,* a systematic and random removal of gauges. The first analysis evaluated the difference between design rainfall and flood events estimated with 75% and 50%

of the gauged network which was systematically reduced. A total of four systematic gauge removal scenarios were created (*Closest.75%, Closest.50%, Furthest.75% and Furthest.50%*) which represent the closest and furthest 75% and 50% of the gauged network around a PoI.

For the first analysis, *i.e.,* a systematic removal of gauges design rainfall events are impacted by up to 4.2% from scenarios using *50%* of the gauged network and by up to 2% from scenarios using *75%* of the gauged network. Design floods events are impacted by up to 28% from scenarios using *50%* of the gauged network and by up to 17% from scenarios using *75%* of the gauged network. Design rainfall events are impacted by up to 4% from scenarios using the *furthest* gauged network which are greater than impacts from scenarios using the *closest* gauged network *i.e*., impacted by up to 2.5%. Design floods events are impacted by up to 24% from scenarios using the *closest* gauged network which are slightly greater than impacts from scenarios using the *furthest* gauged network *i.e*. impacted by up to 22%. A discussion of these findings is provided in Section 7.2.4.

For the second analysis, *i.e.,* a random removal of gauges, differences between design rainfall and floods estimated with the total available network and with 75% and 50% of the gauged network were assessed. A total of 100 networks each having 75% of the gauged network (*Random.75%*) and 50% of the gauged network (*Random.50%*) were generated and used to estimate the design rainfall and flood events at a PoI. Results highlight that design rainfall events are impacted by up to 4.5% from the *Random.50%* scenario which are greater than impacts from the *Random.75%* scenario *i.e.,* impacted by up to 2.2% across all regions. Design flood events are impacted by up to 60% from the *Random.50%* scenario which are greater than impacts from the *Random.75%* scenario *i.e.,* impacted by up to 30% across all regions.

The results indicate that in practice and research the use of a denser monitoring network and rainfall gauges in a closer proximity to a PoI will improve both the accuracy and confidence in DRE and DFE at ungauged sites, thereby increasing the accuracy of infrastructural design which limits the risk to the safety of lives, and severe economic, environmental, and social consequences. Design floods estimated from scenarios using the *closest* gauged network are impacted slightly greater compared to floods estimated from scenarios using the *furthest* gauged network which is contrary to the trend observed for design rainfall estimates and contrary to findings of the HDes method applied by (Lebecherel *et al.,* 2016). It is recommended that this study be expanded to other regions in SA to verify and have more confidence in the findings.

8. DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

The following research question were addressed in this study:

(a) What is the impact of Low Outlier (LO) and High Outlier (HO) events on DRE and DFE and should outliers be excluded from DRE and DFE in SA?

To address this question, the impacts of LO and HO events on DRE and DFE in SA were assessed [Aim (a)]. Undertaking a detailed review of relevant literature [Objective (a)] and calculating the impact of design rainfall and flood estimates to the introduction of LO and HO events using observed data and synthetically generated data series [Objective (b) and Objective (c)] addressed Aim (a).

Objective (a) was met by conducting a detailed literature review which included the importance of data screening and the influence of LO and HO events on DRE and DFE. In summary, data screening is necessary to ensure reliable design flood estimations and is regular practice in many European countries, USA, Australia and UK (Robson and Reed, 1999; Madsen, 2013; Ball *et al.*, 2016; England Jr *et al.*, 2019). Guidelines for data screening and quality control for DFE includes the detection and treatment of outliers (England Jr *et al.*, 2019). Outlier events are data points which significantly depart from the trend of the remaining dataset (Lamontagne *et al.*, 2013; Lamontagne *et al.*, 2016; England Jr *et al.*, 2019). LO events affect, *inter alia*, sample statistics which results in biased parameter estimates, influences flood frequency analysis and the estimation of rare flood events (Lamontagne *et al.*, 2013; Asikoglu, 2017; England Jr *et al.*, 2019). HO observations have been found to directly influence FFA and cause over-estimation of design events (England Jr *et al.*, 2019). There are no prescribed guidelines for data screening and quality control of rainfall or streamflow data for regular practice in South Africa apart for the use of the standardised Z-Score approach to detect outliers (Van der Spuy and Rademeyer, 2018) which, however, has fundamental shortcomings (Iglewicz and Hoaglin, 1993). Further details are provided in Chapter 2.

Objective (b) and Objective (c) were met by calculating and substituting LOs and HOs into observed and synthetically generated rainfall and streamflow AMS of six rainfall and streamflow gauges selected in three different climatic regions in SA. The actual PD of each observed dataset is not known, hence the analysis using observed data may be

biased for or against a particular PD. Synthetic datasets were generated to improve confidence of the analysis by creating AMS datasets from a defined PD. RFA and FFA were performed on observed data and synthetically generated data series per gauge using the GEV, GPA, Kappa, LN and LP3 PDs to estimate the 2-, 5-, 10-, 20-, 50-, 100- and 200-year return period events with and without synthetic outliers. Comparisons between estimated design rainfall and floods with and without substituted outliers were then undertaken. Details on the methodology, assumptions and limitations are provided in Chapter 4.

Results from the analysis of observed datasets showed that design events estimated by the LN PD are the most impacted by LOs i.e., by up to 22% for design rainfall and 45% for design floods across catchments, and the least impacted by HOs i.e. by up to 6% for design rainfall and 22% for design floods across catchments. Design events estimated by the GEV and GPA PDs are the least impacted by LOs i.e. by up to 3% for design rainfall and 2% for design floods across catchments, and the most impacted by HOs i.e. by up to 16% for design rainfall and 46% for design floods across catchments. Regarding the analysis of synthetically generated data series, design rainfall events estimated by the GEV and GPA PDs are impacted by up to 2% in the presence of LOs across catchments and up to 1% for design flood events. Design rainfall and flood events estimated by the GEV and GPA PDs in the presence of HOs are impacted by up to 12% and 13% respectively across catchments. The difference in results between using observed and synthetically generated datasets may be a result of the incorrect PD being applied on the observed datasets and subsequently the generation and substitution of biased synthetic LOs and HOs.

In both practice and research, the results highlight the impact of outlier events and indicate that outlier events must not be ignored in DRE and DFE. In practice, the presence of LOs generally results in an under-estimation of design rainfall and floods thereby reducing the accuracy of infrastructural design which increases the risk of failure and increases risk to the safety of lives and severe economic, environmental, and social consequences. The presence of HOs generally results in an over-estimation of design rainfall and floods thereby resulting in an over-design of infrastructure which provides a conservative approach, however, the economic viability of the design may

be questioned. The LN and GEV PD are most impacted by LOs and HOs respectively therefore special care should be taken in their application.

It is recommended from this study that: (a) outlier events must not be ignored in DRE and DFE, (b) LOs be excluded and HOs should not be excluded from DRE and DFE in SA. Judgement from the analyst is ultimately required on whether to include or exclude HOs from further analysis, as also recommended by (England Jr *et al.*, 2019), after such events have been verified against events from neighbouring stations, (c) special caution be taken when applying the Kappa, LN and GEV PDs for DRE and DFE in SA due to the impact of outliers when using these PDs, and (d) this study be expanded to other regions in SA to have more confidence in the findings and thereafter be used in a South African national guideline.

(b) What is the performance of outlier detection methods in detecting LO and HO events under South African conditions and should outlier detection be regular practice in DRE and DFE in SA?

To address this question, the performance of outlier detection methods in detecting outlier observations were assessed [Aim (b)]. Undertaking a detailed review of relevant literature [Objective (a)], and applying and assessing the performance of the BP, MZS and MGBT methods in detecting outlier observations [Objective (d)] addressed Aim (b).

Objective (a) was met by conducting a detailed review of relevant literature on the BP, MZS and MGBT as contained in Section 2.2. In summary, there are numerous outlier detection methods available and outlier detection is included in many international guidelines. The performance of various methods under South African conditions needs to be investigated. The MZS, BP and the MGBT methods were applied in this study.

Objective (d) was met by applying the BP, MZS and MGBT methods to detect substituted outliers in observed and synthetically generated rainfall and streamflow datasets. Outlier detection methods were applied to observed data scenarios *i.e. Obs.L1* to *Obs.H3* for each driver rainfall and streamflow gauge and also to each of the 100 synthetically generated rainfall and streamflow data series per scenario *i.e. Syn.L1* to *Syn.H3* and per gauge as detailed in Section 4.1.1 and Section 4.2.1 respectively. The BP and MZS were applied to LO and HO scenarios whereas the MGBT was only

applied to LO scenarios as it is designed to only detect LOs. Details on the methodology, assumptions and limitations are provided in Chapter 5.

From observed data, the MGBT outperforms the BP and MZS in detecting LOs with a *RD* of Avg. detection, as detailed in Section 5.1.1, of up to -6% and -30% in observed rainfall and streamflow data, respectively. The MZS outperforms the BP method in detecting HOs with a *RD* of Avg. detection of up to 50% and 100% in observed rainfall and streamflow data respectively. From synthetically generated data series, the MZS outperforms the BP and MGBT in detecting LOs in rainfall datasets of up to -30%. The BP outperforms the MGBT and MZS in detecting LOs in streamflow datasets of up to -15%. The MZS outperforms the BP method in detecting HOs with an *RD* of Avg. detection of by up to 20% and -3% in synthetically generated rainfall and streamflow data respectively.

The performance of different outlier detection methods as presented in this study is aimed to inform the future application of such methods in both practice and research which will lead to a more accurate DRE and DFE. It is recommended from this study that the MGBT be used to detect LOs and the MZS be used to detect HOs in both rainfall and streamflow data in SA.

(c) What is the impact of declining data availability, *i.e.* rainfall and streamflow record lengths and monitoring network density, on DRE and DFE in SA?

To address this question, the impacts of reduced data availability on DRE and DFE in SA were assessed [Aim (c)]. A detailed review of relevant literature [Objective (a)], an evaluation of the impact of both a reduced rainfall and streamflow record length and of different periods of record on design rainfall and flood estimation [Objective (e)], and an evaluation of the impact of a reduced rainfall and streamflow gauged network density on design rainfall and flood estimation by means of a random and systematic reduction of gauges [Objective (f)] were undertaken to address Aim (c).

Objective (a) was met by a detailed review of literature which included information about the current rainfall and streamflow gauging networks in South Africa and the importance of record length, gauge density and proximity on design rainfall and flood estimation. In summary, there is a decline of hydrological monitoring in SA (Pitman, 2011; Pegram *et al.*, 2016) , which affects the availability of gauge network density and

record length and severely impacts on water resource management. An increased rainfall gauge density has been shown in other studies to improve, *inter alia*, simulated streamflow estimates, areal estimates of rainfall, MAP estimates and reduced under-estimation of cumulative rainfall (Krajewski *et al.*, 2003; St-Hilaire *et al.*, 2003; Bárdossy and Das, 2008; Xu *et al.*, 2013). A dense streamflow gauge network is important for, *inter alia*, information transfer from gauged to ungauged catchments (Hrachowitz *et al.*, 2013). Pool *et al.* (2019) further highlighted the importance of spatial proximity of gauges. Boughton (2007) concluded that rainfall record length is influential on estimated rainfall characteristics and on the performance of rainfall-runoff modelling. Using streamflow datasets with shorter records can result in an over-estimation of simulated streamflow and introduced more errors in rainfall-runoff model estimates (Boughton, 2007), however, these datasets are valuable for decision making in ungauged catchments (Pool *et al.*, 2019). Further details are provided in Chapter 2.

Objective (e) was met by reducing the length of the initial observed rainfall and streamflow AMS record of six rainfall and streamflow gauges within South Africa to 75% and 50% by using a moving window approach. Three windows each representing a chronological period of time for both 75% and 50% of the total length of the AMS were chosen. A total of six record length scenarios were created. Each record length scenario was then used to estimate design rainfall and floods, and these were compared to design events estimated using the entire record length. This comparison evaluated the impact of periods of records, *i.e.*, windows within the total record length each representing a different chronological period of time, on DRE and DFE. Details on the methodology, assumptions and limitations are provided in Chapter 6. Regarding the impact of period of record, results show that estimated design rainfall events are impacted by up to 8% whereas design floods are impacted up 95% when using different equal length periods of record.

Regarding the impact of reduced record lengths, impacts across scenarios using 75% of the record (*f.75, m.75, l.75*) and across scenarios using 50% of the record (*f.50, m.50 or l.*50) were averaged to evaluate the overall impact of 75% and 50% of reduced record length on DRE and DFE. On average, design rainfall events are impacted by up to 4% and design floods impacted by up to 24%. There is a negligible difference in impact between design rainfall estimated using the 75% and 50% of AMS records scenario.

Design floods are impacted more when using the 50% of AMS record scenario compared to using 75% of the AMS record scenario. For the impact of period of record and the impact of reduced record length, significant differences between the impacts of design rainfall and floods may be attributed to the period of available rainfall (on average between 1880 and 2000) and streamflow (on average between 1954 and 2013) data *i.e.*, rainfall data covers a longer time period and possibly different climate regimes compared to streamflow data. The identification of climate regimes is beyond the scope of this study. The difference in impacts may also be associated to the assumption of stationarity as detailed in Section 2.1.2. Assessing the stationarity of rainfall and streamflow data is not in the scope of this study.

Objective (f) was met by reducing the total available rainfall and streamflow gauged network density within three homogenous regions in South Africa to 75% and 50% using both random and systematic gauge reduction methods. For both gauge reduction methods, design rainfall and flood events at a PoI were estimated using an index rainfall and flood approach with the reduced network density scenarios and these were compared to design events estimated using the total available gauged network density to evaluate the impact of reduced network density. Details on the methodology, assumptions and limitations are provided in Chapter 7.

Results from the systematic gauge removal highlight that design rainfall events are impacted by up to 4.2% from scenarios using 50% of the gauged network which are greater than impacts from scenarios using 75% of the gauged network *i.e.*, impacted by up to 2%. Design floods events are impacted by up to 28% from scenarios using *50%* of the gauged network which are greater than impacts from scenarios using *75%* of the gauged network *i.e.*, impacted by up to 17%. Design rainfall events are impacted by up to 4% from scenarios using the *furthest* gauged network which are greater than impacts from scenarios using the *closest* gauged network *i.e.*, impacted by up to 2.5%. Design floods events are impacted by up to 24% from scenarios using the *closest* gauged network which are slightly greater than impacts from scenarios using the *furthest* gauged network *i.e.*, impacted by up to 22%. The greater impact on design events estimated by the 50% gauged network scenario compared to the 75% gauged network scenario may be a result of a reduced regionalisation efficiency and principles from Tobler's law of geography as detailed in Section 7.2.4.

Results highlight that design rainfall events are impacted by up to 4.5% from the *Random.50%* scenario which are greater than impacts from the *Random.75%* scenario *i.e.,* impacted by up to 2.2% across all regions. Design flood events are impacted by up to 60% from the *Random.50%* scenario which are greater than impacts from the *Random.75%* scenario *i.e.,* impacted by up to 30% across all regions.

In general, results computed from a reduced gauge network indicate in practice and research that the use of a denser monitoring network and gauges in a closer proximity to a PoI will improve both the accuracy and confidence in DRE and DFE at ungauged sites thereby increasing the accuracy of infrastructural design which limits the risk to the safety of lives, and severe economic, environmental, and social consequences.

It is recommended that this study be expanded to other regions in SA to have more confidence in the findings, and thereafter be used in a South African national guideline. It is also recommended from this study that additional national resources be directed towards maintaining and improving the hydrological monitoring networks in SA ensure the availability of long-term and high-quality hydrological data which will increase the accuracy of DRE and DFE.

Future recommendations for assessing the impacts of outlier events and data availability on design rainfall and flood estimations in SA include: (a) test the performance of each PD for each rainfall and streamflow record to reduce any inherent error in the design rainfall or flood estimate, (b) test the assumption of stationarity with each rainfall and streamflow record, and (c) identify possible climatic regimes in each rainfall and streamflow record to inform the methodology applied to address the impact of reduced record lengths on DRE and DFE.

9. REFERENCES

Alexander, W. 2002. Statistical analysis of extreme floods. *Journal of the South African Institution of Civil Engineering* 44 (1): 20-25.

Archibald, J, Buchanan, B, Fuka, D, Georgakakos, C, Lyon, S and Walter, M. 2014. A simple regionally parameterized model for predicting nopoint source areas in the northeastern US. *Journal of Hydrology: Regional Studies* 1 (1): 74-91.

Asikoglu, O. 2017. Outlier Detection in Extreme Value Series. *Journal of Multidisciplinary Engineering Science and Technology* 4 (5): 7314-7318.

Bailey, A and Pitman, W. 2016. *Water Resources of South Africa, 2012 Study (WR2012): WR2012 Study User's Guide*. WRC Report No. TT 684/16. Water Research Commission, Pretoria, South Africa.

Ball, J, Babister, M, Nathan, R, Weeks, W, Weinmann, P, Retallick, M and Testoni, I. 2016. *Australian Rainfall and Runoff: A Guide to Flood Estimation*. Commonwealth of Australia, Murdoch, Australia.

Bárdossy, A and Das, T. 2008. Influence of rainfall observation network on model calibration and application. *Hydrology and Earth System Sciences* 12 (1): 77-89.

Bezak, N, Brilly, M and Šraj, M. 2015. Flood frequency analyses, statistical trends and seasonality analyses of discharge data: a case study of the Litija station on the Sava River. *Journal of Flood Risk Management* 9 (2): 154-168.

Boughton, W. 2007. Effect of data length on rainfall–runoff modelling. *Environmental Modelling & Software* 22 (3): 406-413.

Calitz, J and Smithers, J. 2020. *Development and assessment of a probabilistic rational method for design flood estimation in South Africa*. WRC Report No. K5/2748. Water Research Commission, Pretoria, South Africa.

CGaTA. 2009. *National Disaster Management Centre Annual Report 2008-2009*. Cooperative Governance and Traditional Affairs, Pretoria, South Africa.

Chebana, F and Ouarda, T. 2009. Index flood–based multivariate regional frequency analysis. *Water Resources Research* 45 (10): 1-15.

Chen, H and Ramachandra, R. 2002. Testing Hydrologic Time Series for Stationarity. *Journal of Hydrological Engineering* 7 (1): 129-136.

Cohn, T, England, J, Berenbrock, C, Mason, R, Stedinger, J and Lamontagne, J. 2013. A generalized Grubbs-Beck test statistic for detecting multiple potentially influential low outliers in flood series. *Water Resources Research* 49 (8): 5047-5058.

Costa, J and Jarrett, R. 2008. *An evaluation of selected extraordinary floods in the United States reported by the US Geological Survey and implications for future advancement of flood science*. Report No. 2008-5164. Report No. 2008-5164. United States Geological Survey Reston, Virginia.

Dalrymple, T. 1960. *Manual of hydrology: Part 3 Flood flow techniques. Flood frequency Analyses*. United States Department of the Interior, Washington DC, USA.

Dent, J. 1994. The parlous state of hydrometry overseas. *Circulation* 44 (1): 14-15.

Dent, M, Lynch, S and Schulze, R. 1987. *Mapping mean annual and other rainfall statistics over southern Africa*. WRC Report No. 109/1/89. Water Research Commission, Pretoria, South Africa.

DWS. 2011. Hydrological Services - Surface Water (Data, Dams, Floods and Flows). [Internet]. Department of Water and Sanitation. Available from: http://www.dwa.gov.za/Hydrology/Verified/hymain.aspx/ [Accessed: 20 September 2018].

Dyler, T and Tyson, P. 1977. Evaluating above and below normal rainfall periods over South Africa. *Journal of Applied Meteorology* 16 (1): 145-147.

England Jr, JF, Cohn, TA, Faber, BA, Stedinger, JR, Thomas Jr, WO, Veilleux, AG, Kiang, JE and Mason Jr, RR. 2019. *Guidelines for determining flood flow frequency—Bulletin 17C*. United States Geological Survey, Reston, Virginia.

Farquharson, F, Meigh, J and Sutcliffe, J. 1992. Regional flood frequency analysis in arid and semi-arid areas. *Journal of Hydrology* 138 (3): 487-501.

Gericke, O. 2015. Estimation of Catchment Response Time in Medium to Large Catchments in South Africa. Unpublished thesis, School of Engineering, University of KwaZulu-Natal Pietermaritzburg, South Africa.

Gilleland, E. 2018. extRemes: Extreme Value Analysis. V2.1. Available from: https://cran.r-project.org/web/packages/extRemes/index.html [Accessed: 10 May 2019].

Görgens, A. 2007. *Joint peak-volume (JPV) design flood hydrographs for South Africa*. WRC Report No. 1420/3/07. WRC Report No. 1420/3/07. Water Research Commission, Pretoria, South Africa.

Grubbs, F and Beck, G. 1972. Extension of sample sizes and percentage points for significance tests of outlying observations. *Technometrics* 14 (4): 847-854.

Haile, A. 2011. Regional Flood Frequency Analysis on Southern Africa. Unpublished thesis, Department of Geosciences, Unisersity of Oslo, Oslo, Norway.

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

Hosking, J and Wallis, J. 1997. *Regional frequency analysis—An approach based on L-moments*. Cambridge University Press, Cambridge, UK.

Hrachowitz, M, Savenije, H, Blöschl, G, McDonnell, J, Sivapalan, M, Pomeroy, J, Arheimer, B, Blume, T, Clark, M, Ehret, U, Fenicia, F, Freer, J, Gelfan, A, Gupta, H, Hughes, D, Hut, R, Montanari, A, Pande, S, Tetzlaff, D, Troch, P, Uhlenbrook, S, Wagener, T, Winsemius, H, Woods, R, Zehe, E and Cudennec, C. 2013. A decade of Predictions in Ungauged Basins (PUB)-A Review. *Hydrological Sciences Journal* 58 (6): 1198-1255.

HRU. 1972. *Design flood determination in South Africa*. HRU Report No. 1/72. Hydrological Research Institute, Johannesburg, RSA.

Hughes, D. 2008. Modelling semi-arid and arid hydrology and water resources: The southern Africa experience. *Hydrological modelling in arid and semi-arid areas* (1): 29-40.

Iglewicz, B and Hoaglin, D. 1993. *How to detect and handle outliers*. ASQC Qualty Press, California, USA.

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.

Knisel, W, Renard, K and Lane, L.1979. Hydrologic data collection: How long is long enough? *Speciality Conference on Irrigation and Drainage in the Nineteen Eighties*, American Society or Civil Engineers, Albuquerque, New Mexico.

Kondragunta, C.2001. An outlier detection technique to quality control rain gauge measurements. *Spring Meeting*, American Geophysical Union, Silver Spring, USA.

Koutsoyiannis, D, Montanari, M, Lins, H and Cohn, T. 2009. Climate, hydrology and freshwater: Towards an interactive incorporation of hydrological experience into climate research – discussion of "The implications of projected climate change for freshwater resources and their management". *Hydrological sciences journal* 54 (2): 394-405.

Kovacs, Z. 1988. *Regional maximum flood peaks in Southern Africa*. Technical Report TR137. Technical Report No. TR137. Department of Water Affairs, Directorate of Hydrology, Pretoria, South Africa.

Krajewski, WF, Ciach, GJ and Habib, E. 2003. An analysis of small-scale rainfall variability in different climatic regimes. *Hydrological sciences journal* 48 (2): 151-162.

Krasovskaia, I, Gottschalk, L, Sælthun, NR and Berg, H. 2001. Perception of the risk of flooding: the case of the 1995 flood in Norway. *Hydrological Sciences Journal* 46 (6): 855-868.

Lamontagne, J, Stedinger, J, Yu, X, Whealton, C and Xu, Z. 2016. Robust flood frequency analysis: Performance of EMA with multiple Grubbs-Beck outlier tests. *Water Resources Research* 52 (4): 3068-3084.

Lamontagne, JR, Stedinger, JR, Cohn, TA and Barth, NA.2013. Robust national flood frequency guidelines: What is an outlier? *World Environmental and Water Resources Congress 2013: Showcasing the Future*, 2454-2466. American Society of Civil Engineers, Cincinnati, Ohio.

Lebecherel, L, Andréassian, V and Perrin, C. 2016. On evaluating the robustness of spatial-proximity-based regionalization methods. *Journal of Hydrology* 539 (1): 196-203.

Lim, K, Engel, B, Tang, Z, Muthukrishnan, S, Choi, J and Kim, K. 2006. Effects of calibration on L-THIA GIS runoff and pollutant estimation. *J Environ Manag* 78 (1): 35-43.

Liu, J, Doan, C, Liong, S, Sanders, R, Dao, A and Fewtrell, T. 2015. Regional frequency analysis of extreme rainfall events in Jakarta. *Journal of Natural Hazards* 75 (1): 1075-1104.

Liu, X, Liu, Z, Zhang, Y and Jiang, B. 2017. The effects of floods on the incidence of bacillary dysentery in Baise (Guangxi Province, China) from 2004 to 2012. *International Journal of Environmental Research and Public Health* 14 (2): 179-190.

Lorenz, C and Kunstmann, H. 2012. The hydrological cycle in three state-of-the-art reanalyses: Intercomparison and performance analysis. *Journal of Hydrometeorology* 13 (5): 1397-1420.

Lynch, S. 2004. *Development of a Raster Database of Annual, Monthly and Daily Rainfall for Southern Africa*. WRC Report No. 1156/1/04. WRC Report No. 1156/1/03. Water Research Commission, Pretoria, South Africa.

Madsen, H. 2013. *Floodfreq COST Action ES0901 - European procedures for flood frequency estimation : a review of applied methods in Europe for flood-frequency analysis in a changing environment : WG4 - Flood frequency estimation methods and environmental change*. Centre for Ecology & Hydrology, Bailrigg, England.

Middleton, B and Bailey, A. 2008. *Water Resources of South Africa 2005*. WRC Report Nos. KTT 380 to 382/08. Water Research Commission Pretoria, South Africa.

Montanari, A and Koutsoyiannis, D. 2014. Modeling and mitigating natural hazards: Stationarity is immortal! *Water Resources Research* 50 (12): 9748-9756.

Moriasi, D, Arnold, J, Liew, MV, Bingner, R, Harmel, R and Veith, T. 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans ASABE* 50 (3): 885–900.

Muller, C, Chapman, L, Johnston, S, Kidd, C, Illingworth, S, Foody, G, Overeem, A and Leigh, R. 2015. Crowdsourcing for climate and atmospheric sciences: current status and future potential. *International Journal of Climatology* 35 (11): 3185-3203.

Nash, J and Sutcliffe, J. 1970. River flow forecasting through conceptual models. Part I: A Discussion of Principles. *Journal of Hydrology* 10 (1): 282–290.

Nathanael, J. 2015. Assessing the Performance of Regional Flood Frequency Analysis Methods in South Africa. Unpublished Unpublished thesis, Centre for Water Resources Research University of KwaZulu-Natal, Pietermaritzburg, South Africa.

Parajuli, P, Mankin, K and Barnes, P. 2009. Source specific fecal bacteria modeling using soil and water assessment tool model. *Bioresource Technology* 100 (2): 953–963.

Paretti, N, Kennedy, J and Cohn, T. 2014. *Evaluation of the Expected Moments Algorithm and a Multiple Low-Outlier Test for Flood Frequency Analysis at Streamgaging Stations in Arizona*. Report No. 5026. United States Geological Survey, Reston, Virginia.

Pegram, G. 1997. Patching rainfall data using regression methods. Grouping, patching and outlier detection. *Journal of hydrology* 198 (3): 319-334.

Pegram, G, Sinclair, S and Bárdossy, A. 2016. *New Methods of Infilling Southern African Raingauge Records Enhanced by Annual, Monthly and Daily Precipitation Estimates Tagged with Uncertainty*. WRC Report No. 2241/1/15. WRC Report No. 2241/1/15. Water Research Commission, Pretoria, South Africa.

Pilgrim, D. 2001. Australian rainfall and runoff : a guide to flood estimation. *Institution of Engineers*, Barton, Australia.

Pitman, W. 2011. Overview of water resource assessment in South Africa: Current state and future challenges. *Water SA* 37 (5): 659-664.

Plavšić, J, Mihailović, V and Blagojević, B.2014. Assessment of methods for outlier detection and treatment in flood frequency analysis. *Mediterranean Meeting on "Monitoring, modelling and early warning of extreme events triggered by heavy rainfall "*, 181-190. University of Calabria, Cosenza, Italy.

Pool, S, Viviroli, D and Seibert, J. 2019. Value of a limited number of discharge observations for improving regionalization: A large-sample study across the United States. *Water Resources Research* 55 (1): 363–377.

Rahman, A, Rahman, A, Zaman, M, Haddad, K, Ahsan, A and Imteaz, M. 2013. A study on selection of probability distributions for at-site flood frequency analysis in Australia. *Natural Hazards* 69 (1): 1803–1813.

Rahman, A, Haddad, K and Rahman, A. 2014. Identification of outliers in flood frequency analysis: Comparison of original and multiple Grubbs-Beck Test. *World Academy of Science, Engineering and Technology* 8 (1): 732-740.

Robson, A and Reed, D. 1999. *Flood estimation handbook: Statistical procedures for flood frequency estimation*. Institute of Hydrology Wallingford, UK.

Rojas-Serna, C, Lebecherel, L, Perrin, C, Andréassian, V and Oudin, L. 2016. How should a rainfall-runoff model be parameterized in an almost ungauged catchment? A methodology tested on 609 catchments. *Water Resources Research* 52 (1): 4765–4784.

Rowinski, P, Strupczewski, W and Singh, V. 2002. A note on the applicability of log-Gumbel and log-logistic probability distributions in hydrological analyses. *Hydrological Sciences Journal* 47 (1): 107–122.

Rulfova, Z, Buishand, A, Roth, M and Kysely, J. 2016. Two-component generalized extreme value distribution for precipitation frequency analysis. *Journal of Hydrology* 534 (1): 659–668.

SANRAL. 2013. *Drainage Manual*. South African National Roads Agency, Pretoria, South Africa.

Santhi, C, Arnold, J, Williams, J, Dugas, W, RSrinivasan and Hauck, L. 2001. Validation of the SWAT model on a large river basin with point and nonpoint sources. *American Water Resources Association* 37 (5): 1169–1188.

Sene, K and Farquharson, F. 1998. Sampling errors for water resources design: the need for improved hydrometry in developing countries. *Water Resources Management* 12 (2): 121-138.

Singh, J, Knapp, H, Arnald, J and Demissie, M. 2004. Hydrologicmodeling of the Iroquois River watershed using HSPF and SWAT. *American Water Resources Association* 41 (2): 343–360.

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

Smithers, J and Schulze, R. 2000a. *Long duration design rainfall estimates for South Africa*. WRC Report No. 811/1/00. WRC Report No. 811/1/00. Water Research Commission, Pretoria, South Africa.

Smithers, J and Schulze, R. 2000b. *Development and evaluation of techniques for estimating short duration design rainfall in South Africa*. WRC Report No. 681/1/00. WRC Report No. 681/1/00. Water Research Commision Pretoria, South Africa.

Smithers, J. 2012. Methods for design flood estimation in South Africa. *Water SA* 38 (4): 633-646.

Smithers, J, Gorgens, A, Gericke, O, Jonker, V and Roberts, P. 2014. *The initiation of a National Flood Studies Programme for South Africa*. SANCOLD, Pretoria, South Africa.

Šraj, M, Viglione, A, Parajka, J and Blöschl, G. 2016. The influence of non-stationarity in extreme hydrological events on flood frequency estimation. *Journal of Hydrology and Hydromechanics* 64 (4): 426-437.

St-Hilaire, A, Ouarda, T, Lachance, M, Bobée, B, Gaudet, J and Gignac, C. 2003. Assessment of the impact of meteorological network density on the estimation of basin precipitation and runoff: a case study. *Hydrological Processes* 17 (18): 3561-3580.

Stedinger, J, Vogel, R and Foufoula-Georgiou, E. 1993. *Frequency Analysis of Extreme Events. In: Maidment DR (ed.) Handbook of Hydrology.* McGraw-Hill, New York, USA.

Stephenson, AG. 2018. evd: Functions for Extreme Value Distributions. V2.3.3. Available from: https://cran.r-project.org/web/packages/evd/index.html [Accessed: 10 May 2019].

Stewart, B.2015. Measuring what we manage–the importance of hydrological data to water resources management. *Hydrological Sciences and Water Security: Past, Present and Future. Proceedings of the 11th Kovacs Colloquium*, 80-85. International Association of Hydrological Sciences, Paris, France.

Sun, X, Mein, R, Keenan, T and Elliott, J. 2000. Flood estimation using radar and raingauge data. *Journal of Hydrology* 239 (1): 4-18.

Sunilkumar, K, Narayana Rao, T and Satheeshkumar, S. 2016. Assessment of small-scale variability of rainfall and multi-satellite precipitation estimates using measurements from a dense rain gauge network in Southeast India. *Hydrology and Earth System Sciences* 20 (5): 1719-1735.

Tobler, W. 1970. A computer movie simulating urban growth in the Detroit Region. *Economic Geography* 11 (77): 620-622.

USGS. 2006. Benefits of the USGS Stream Gauging Program – Users and uses of US streamflow data. [Internet]. United States Geological Survey. Available from: http://water.usgs.gov/osw/pubs/nhwc_report.pdf [Accessed: 13 September 2029].

Van Bladeren, D, Zawada, P and Mahlangu, D. 2007. *Statistical based regional flood frequency estimation study for South Africa using systematic, historical and palaeoflood data*. WRC Report No 1260/1/07. WRC Report No 1260/1/07. Water Research Commission, Pretoria, South Africa.

Van der Spuy, D and Rademeyer, P. 2018. *Flood frequency estimation methods as applied in the Department of Water Affairs*. Department of Water and Sanitation, Pretoria, South Africa.

Van Liew, M, Arnold, J and Garbrecht, J. 2003. Hydrologic simulation on agricultural watersheds: choosing between two models. *Transactions of the American Society of Association Executives* 46 (6): 1539–1551.

Van Vuuren, S, Van Dijk, M and Coetzee, G. 2013. *Status review and requirements of overhauling flood determination methods in South Africa*. 1431204196. WRC Project No. K8/994 (TT 563/13). Water Research Commission, Pretoria, South Africa.

Viviroli, D and Seibert, J. 2015. Can a regionalized model parameterisation be improved with a limited number of runoff measurements? *Journal of Hydrology* 529 (1): 49-61.

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

Whitacre, J, Pham, T and Sarker, R.2006. Use of statistical outlier detection method in adaptive evolutionary algorithms. *Proceedings of the 8th annual conference on Genetic and evolutionary computation*, 1345-1352. University of New South Wales, Washington, USA.

Wolodzko, T. 2018. extraDistr: Additional Univariate and Multivariate Distributions. V1.8.10. Available from: https://cran.r-project.org/web/packages/extraDistr/index.html [Accessed: 10 May 2019].

Xu, H, Xu, C, Chen, H, Zhang, Z and Li, L. 2013. Assessing the influence of rain gauge density and distribution on hydrological model performance in a humid region of China. *Journal of Hydrology* 505 (1): 1-12.

Zalina, M, Desa, M, Nguyen, V and Kassim, A. 2002. Selecting a probability distribution for extreme rainfall series in Malaysia. *Water Science and Technology* 45 (1): 63-72.

Zhang, X, Harvey, K, Hogg, W and Yuzyk, T. 2001. Trends in Canadian streamflow. *Water Resources Research* 37 (4): 987-998.

10. APPENDIX A: INVENTORY OF SELECTED RAINFALL AND STREAMFLOW GAUGES

This appendix contains an inventory of the driver rainfall and streamflow gauges used in this study, shown in Table A.1 and Table A.2 respectively. The reliable record lengths in Table A.1 and Table A.2 refer to the unedited observed record *i.e.* no patching or infilling as detailed in Chapter 3).

Table A.1 Driver rainfall gauges and associated attributes

Station Number	Streamflow Gauge ID	Climate Zone Gericke (2015)	Latitude (^{O})	Longitude (^{O})	Start Date	End Date	Total Record Length (Years)	Reliable Record Length (Years)	MAP (mm)	Altitude (m)
0239097	U2H013	ESC	-29.62	30.07	09/1882	10/2001	113	46	1007	1540
0268640	V2H004	ESC	-29.17	29.87	09/1882	08/2001	107	82	877	1520
0476031	A2H012	NI	-26.01	28.05	05/1886	08/2000	107	49	649	1375
0589670	A6H011	NI	-24.67	28.37	02/1903	08/2000	96	82	598	1234
0042227	G1H008	SWC	-33.29	19.14	01/1850	12/2001	150	99	473	165
0025414	H7H004	SWC	-33.91	20.73	01/1878	12/2001	121	74	281	375

Table A.2 Driver streamflow gauges and associated attributes

Streamflow Gauge ID	Climate Zone Gericke (2015)	HRU (1972) Zone	Kovacs (1988) K-Region	Latitude (^{O})	longitude (^{O})	Start date	End date	Record length (Years)	Actual Record Length (Years)	Area (km^2)
U2H013	ESC	5	5	-29.51	30.09	08/1960	09/2013	54	46	295.70
V2H004	ESC	9	5	-29.36	29.88	09/1972	11/2013	42	40	269.13
A2H012	NI	8	5	-25.81	27.91	10/1922	12/2013	92	54	2579.65
A6H011	NI	8	4.6	-24.76	28.34	11/1966	08/2013	48	40	73.66
G1H008	SWC	2	5	-33.31	19.07	05/1954	11/2013	60	43	396.07
H7H004	SWC	2	5	-33.91	20.71	05/1951	10/2013	63	42	25.60

11. APPENDIX B: IMPACT OF OUTLIER EVENTS ON DESIGN RAINFALL AND FLOOD ESTIMATION

This appendix contains results for the detailed assessment (*cf.* Section 4.1.1) for all study catchments. Appendix B1 and Appendix B2 in Section 11.1 and Section 11.2 respectively contains results for observed data and synthetically generated data series respectively.

11.1 Appendix B1: Observed Data

Design rainfall results are presented in Section 11.1.1 followed by design floods in Section 11.1.2.

11.1.1 Design rainfall

Figure B1.1 to Figure B1.5 provides the calculated *MRD*, *MARD*, *PBIAS* and *NSE* for estimated design rainfall events of Rainfall Gauge 0476031, 0589670, 0025414, 0239097 and 0268640 respectively.

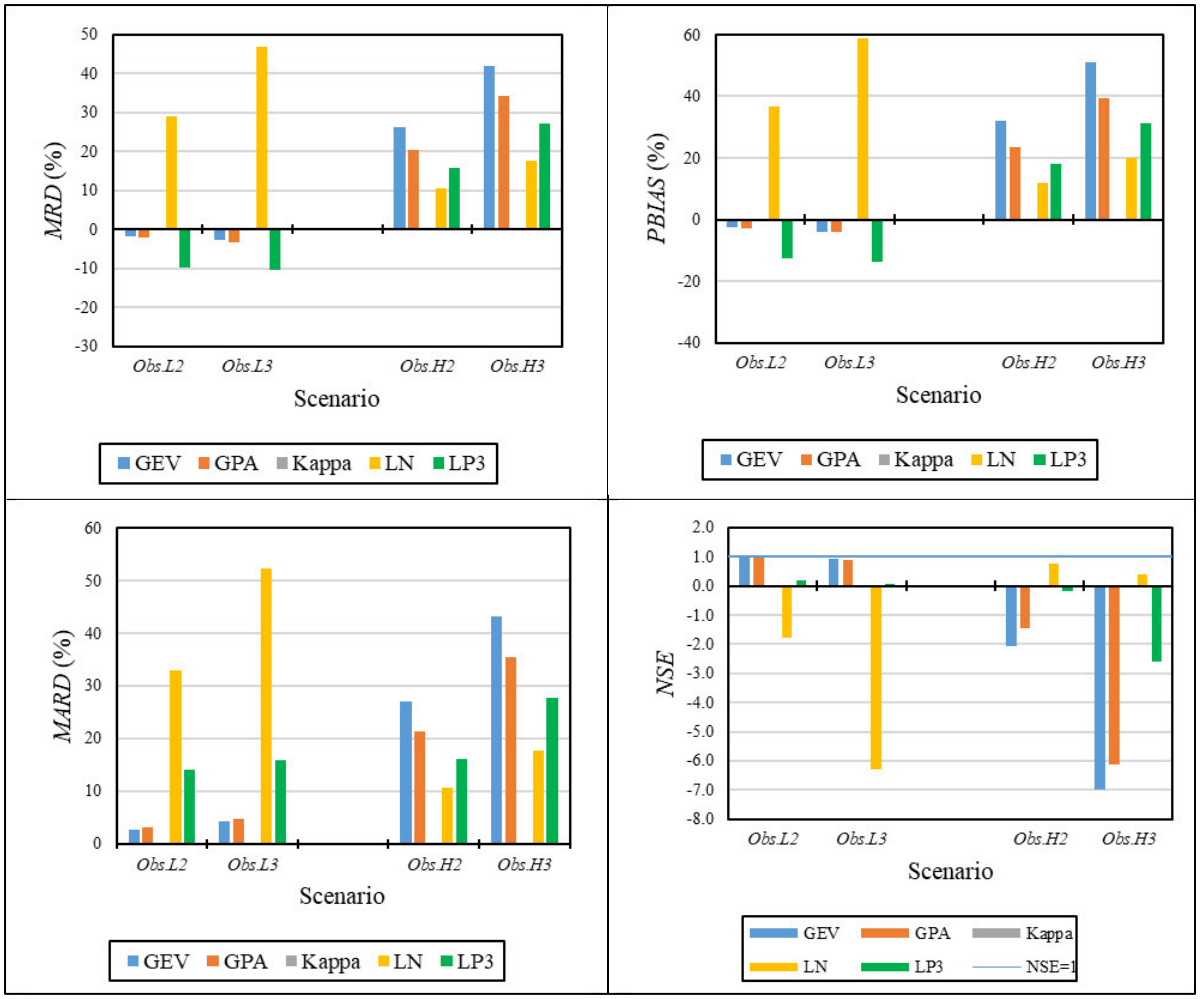

Figure B1.1 Impact of LO (*Obs.L2* and *Obs.L3*) and HO (*Obs.H2* and *Obs.H3*) scenarios on design rainfall estimated by different PDs from observed data for Rainfall Gauge 0476031

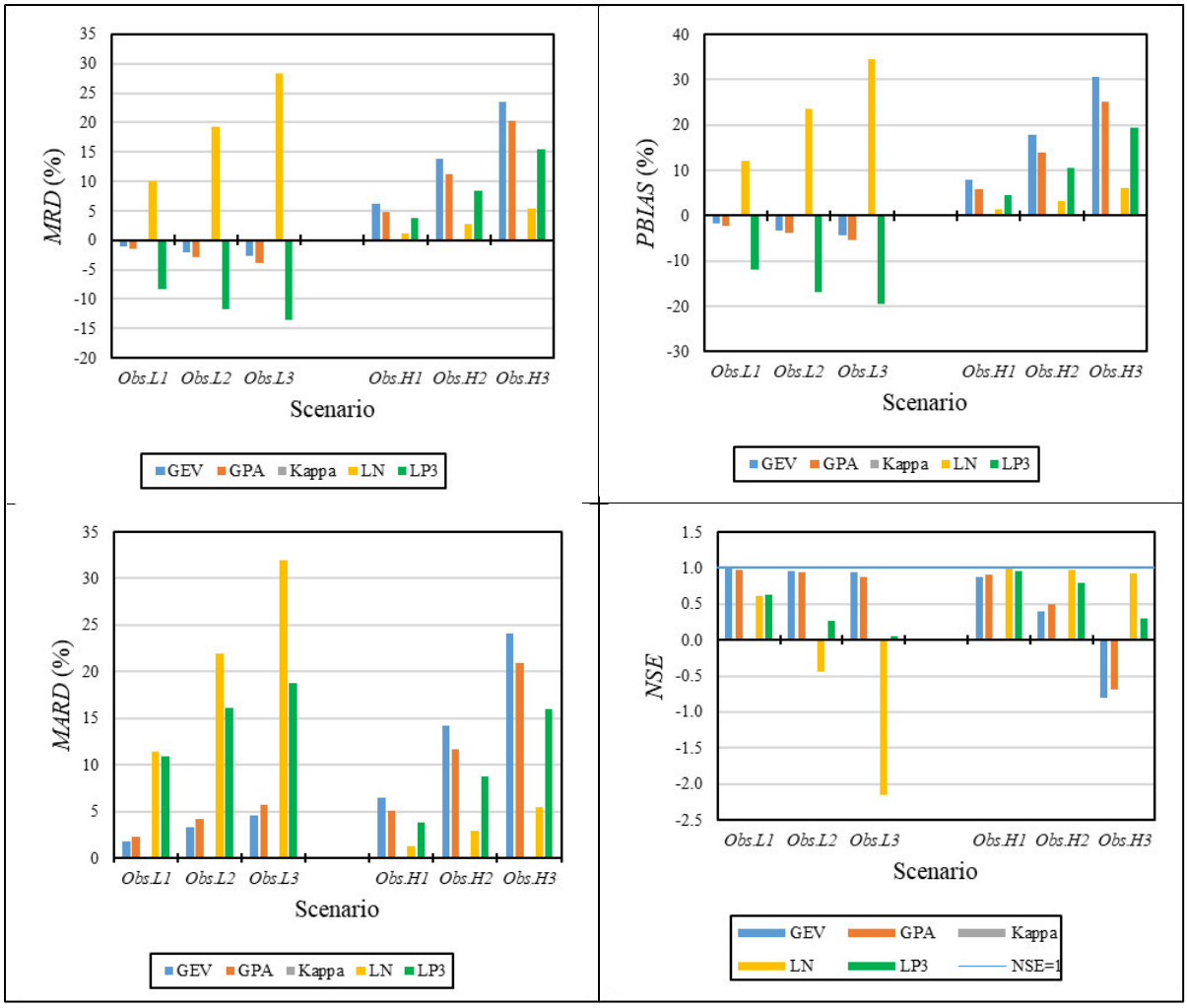

Figure B1.2 Impact of LO (*Obs.L1*, *Obs.L2* and *Obs.L3*) and HO (*Obs.H1*, *Obs.H2* and *Obs.H3*) scenarios on design rainfall estimated by different PDs from observed data for Rainfall Gauge 0589670

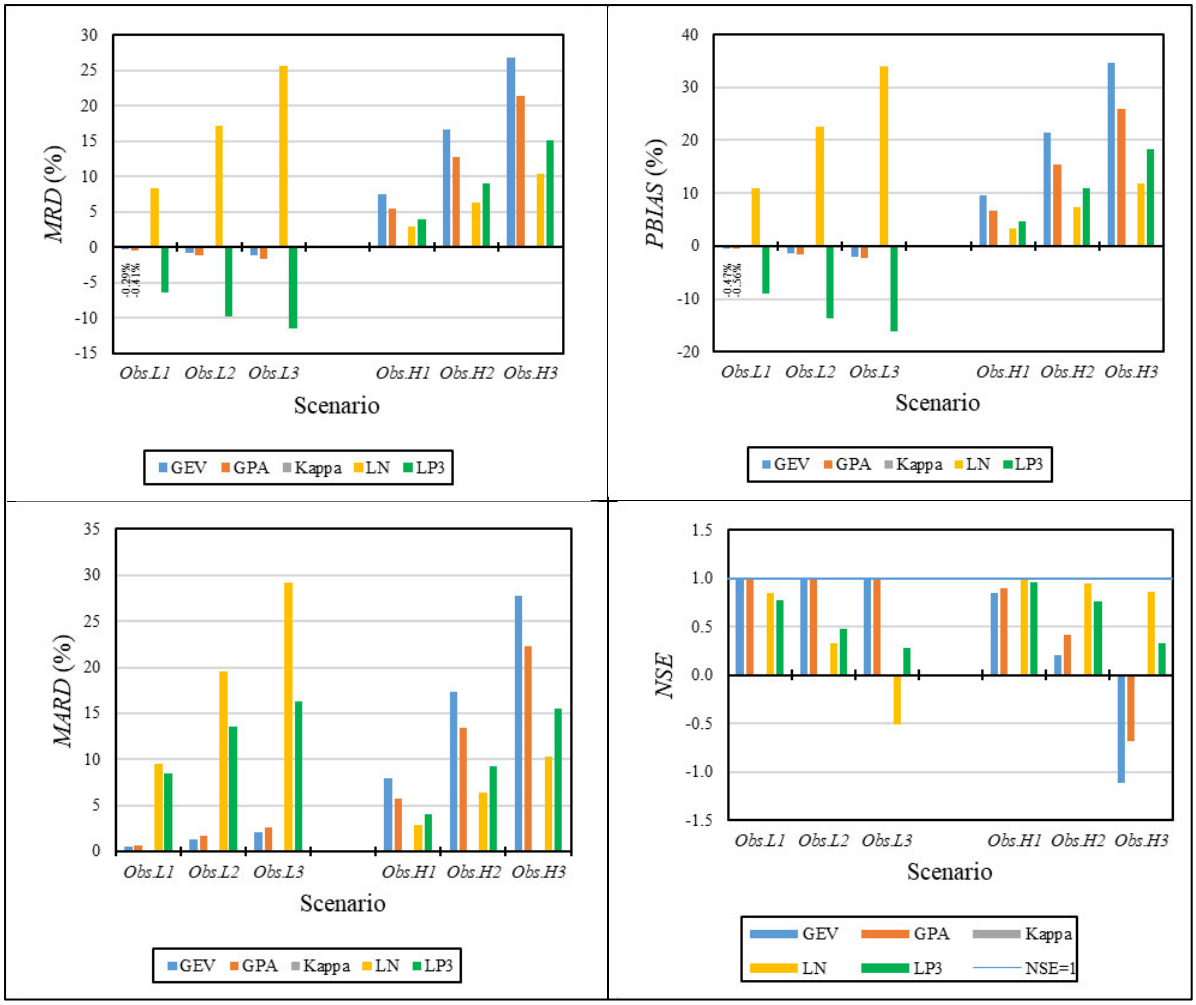

Figure B1.3 Impact of LO (*Obs.L1*, *Obs.L2* and *Obs.L3*) and HO (*Obs.H1*, *Obs.H2* and *Obs.H3*) scenarios on design rainfall estimated by different PDs from observed data for Rainfall Gauge 0025414

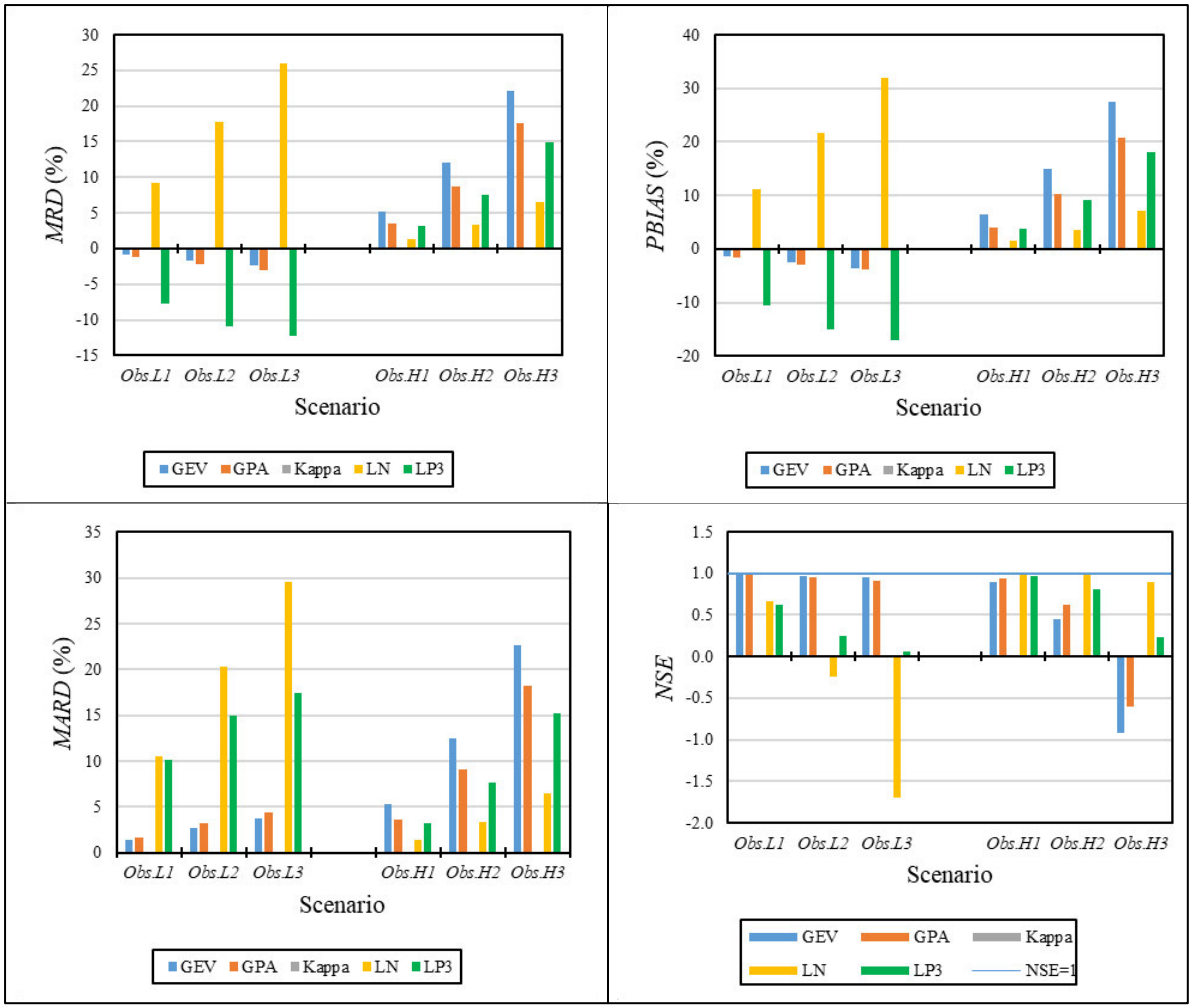

Figure B1.4 Impact of LO (*Obs.L1*, *Obs.L2* and *Obs.L3*) and HO (*Obs.H1*, *Obs.H2* and *Obs.H3*) scenarios on design rainfall estimated by different PDs from observed data for Rainfall Gauge 0239097

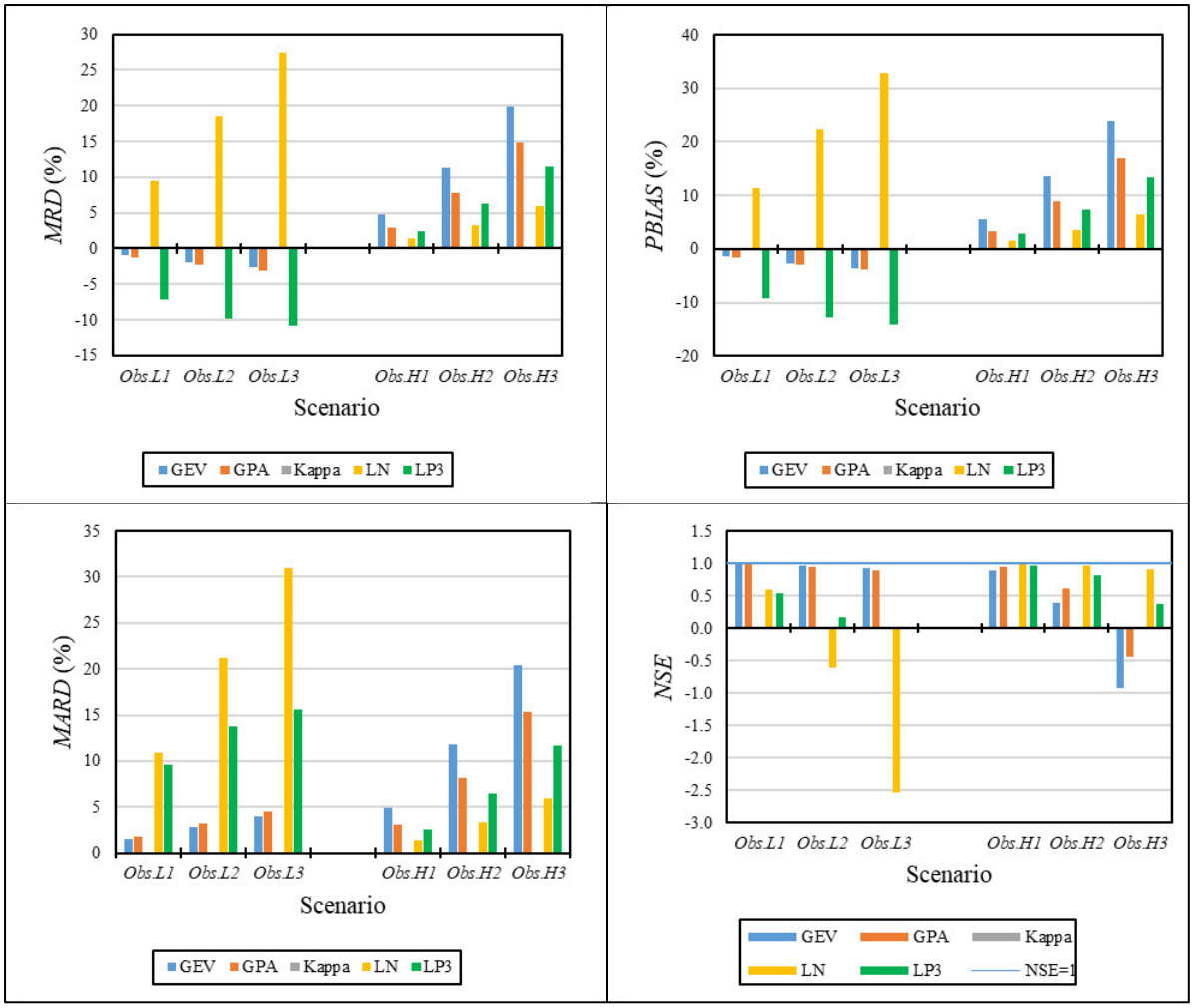

Figure B1.5 Impact of LO (*Obs.L1*, *Obs.L2* and *Obs.L3*) and HO (*Obs.H1*, *Obs.H2* and *Obs.H3*) scenarios on design rainfall estimated by different PDs from observed data for Rainfall Gauge 0268640

11.1.2 Design floods

Figure B1.6 to Figure B1.11 provides the calculated *MRD*, *MARD*, *PBIAS* and *NSE* for estimated design flood events in Catchment A2H012, A6H011, H7H004, G1H008, U2H013 and V2H004 respectively.

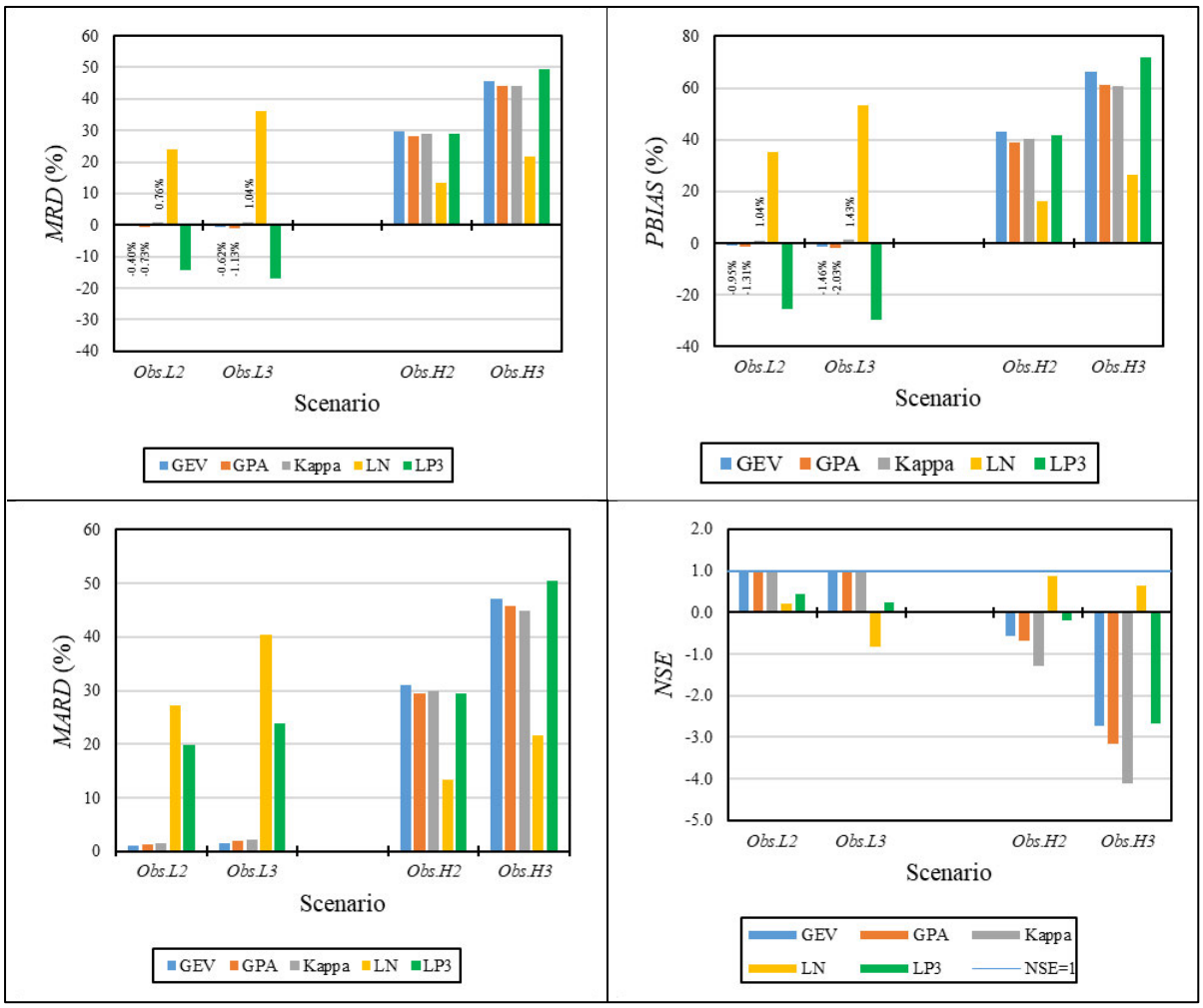

Figure B1.6 Impact of LO (*Obs.L2* and *Obs.L3*) and HO (*Obs.H2* and *Obs.H3*) scenarios on design floods estimated by different PDs from observed data in Catchment A2H012

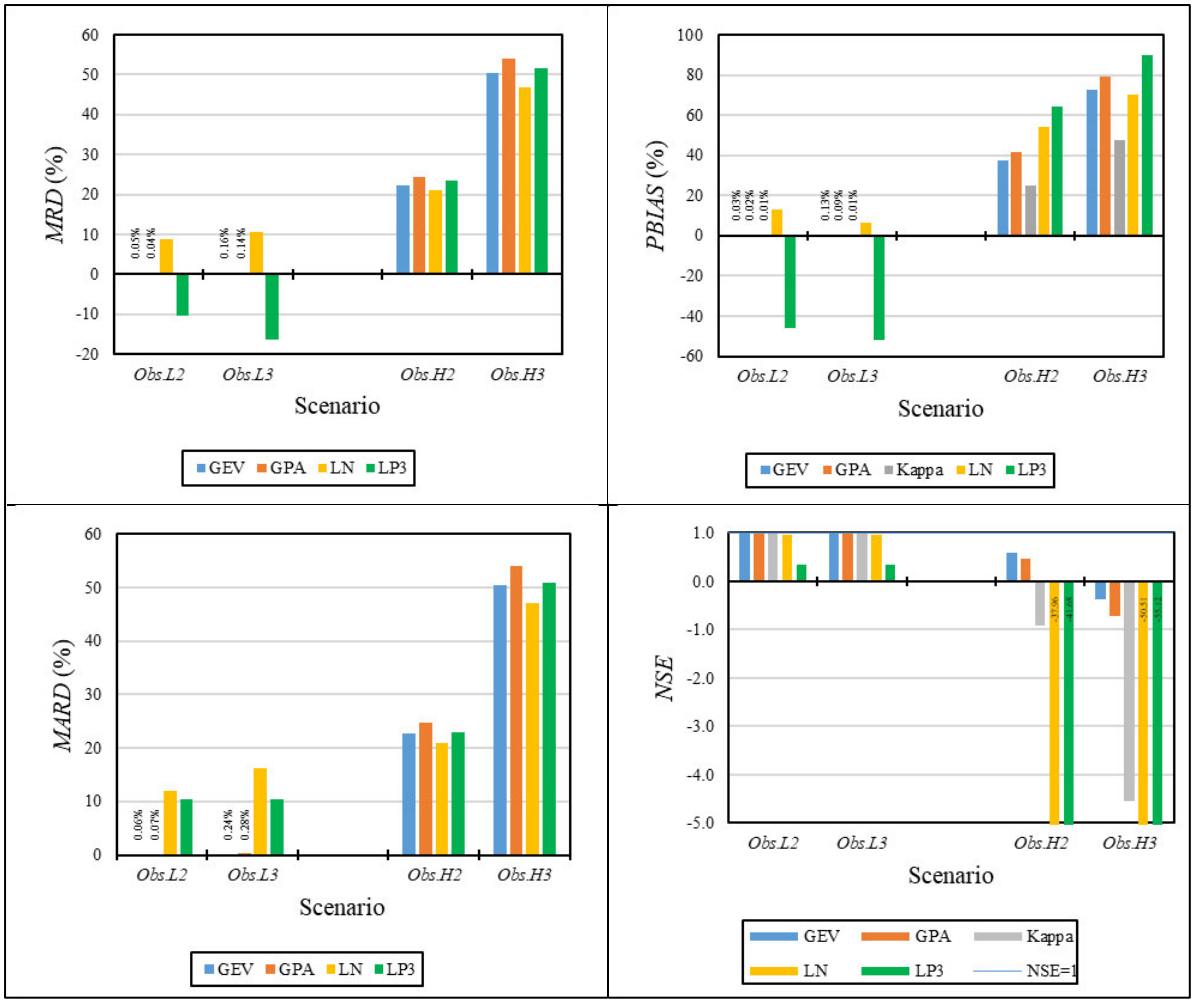

Figure B1.7 Impact of LO (*Obs.L2* and *Obs.L3*) and HO (*Obs.H2* and *Obs.H3*) scenarios on design floods estimated by different PDs from observed data in Catchment A6H011

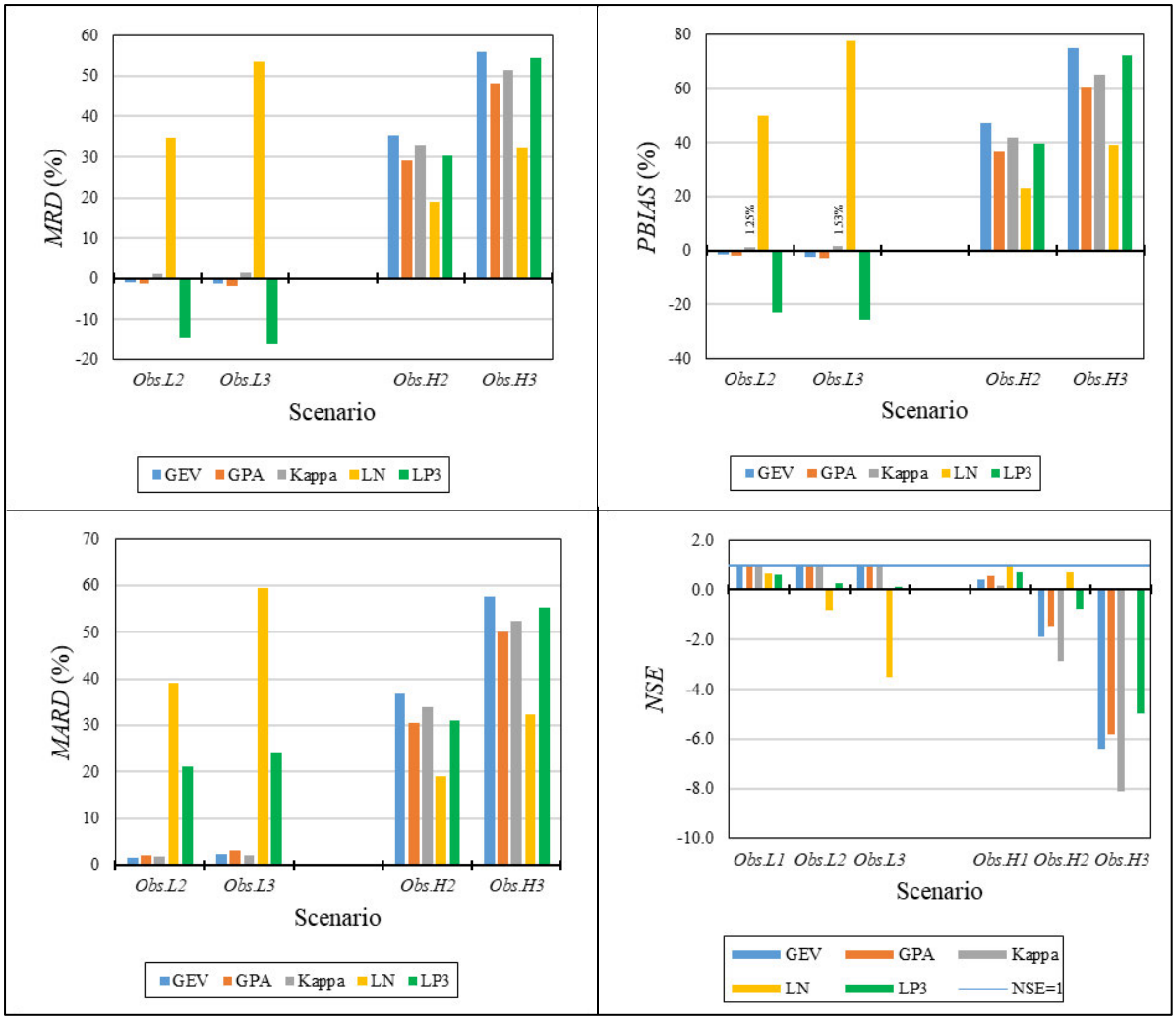

Figure B1.8 Impact of LO (*Obs.L2* and *Obs.L3*) and HO (*Obs.H2* and *Obs.H3*) scenarios on design floods estimated by different PDs from observed data in Catchment G1H008

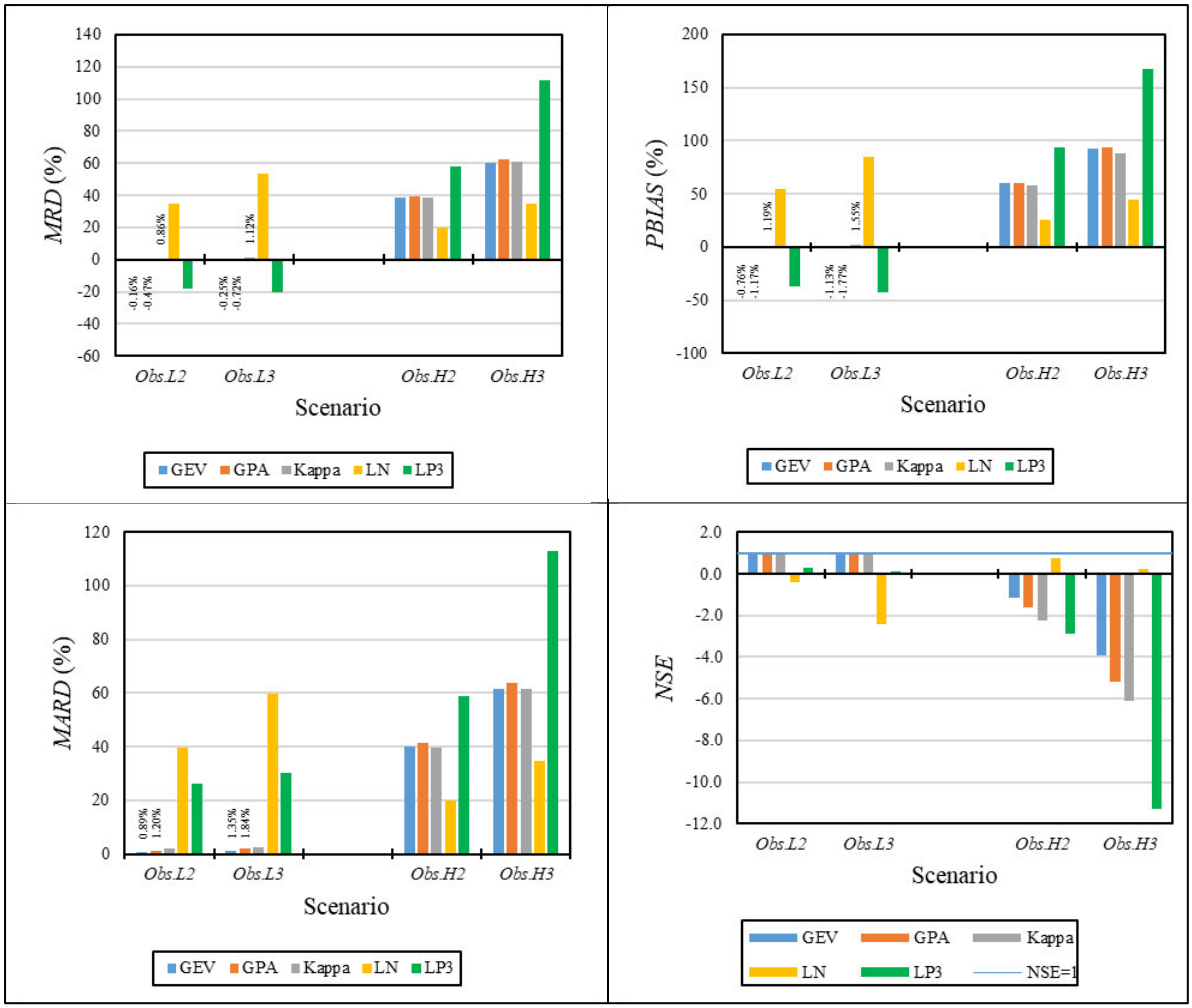

Figure B1.9 Impact of LO (*Obs.L2* and *Obs.L3*) and HO (*Obs.H2* and *Obs.H3*) scenarios on design floods estimated by different PDs from observed data in Catchment H7H004

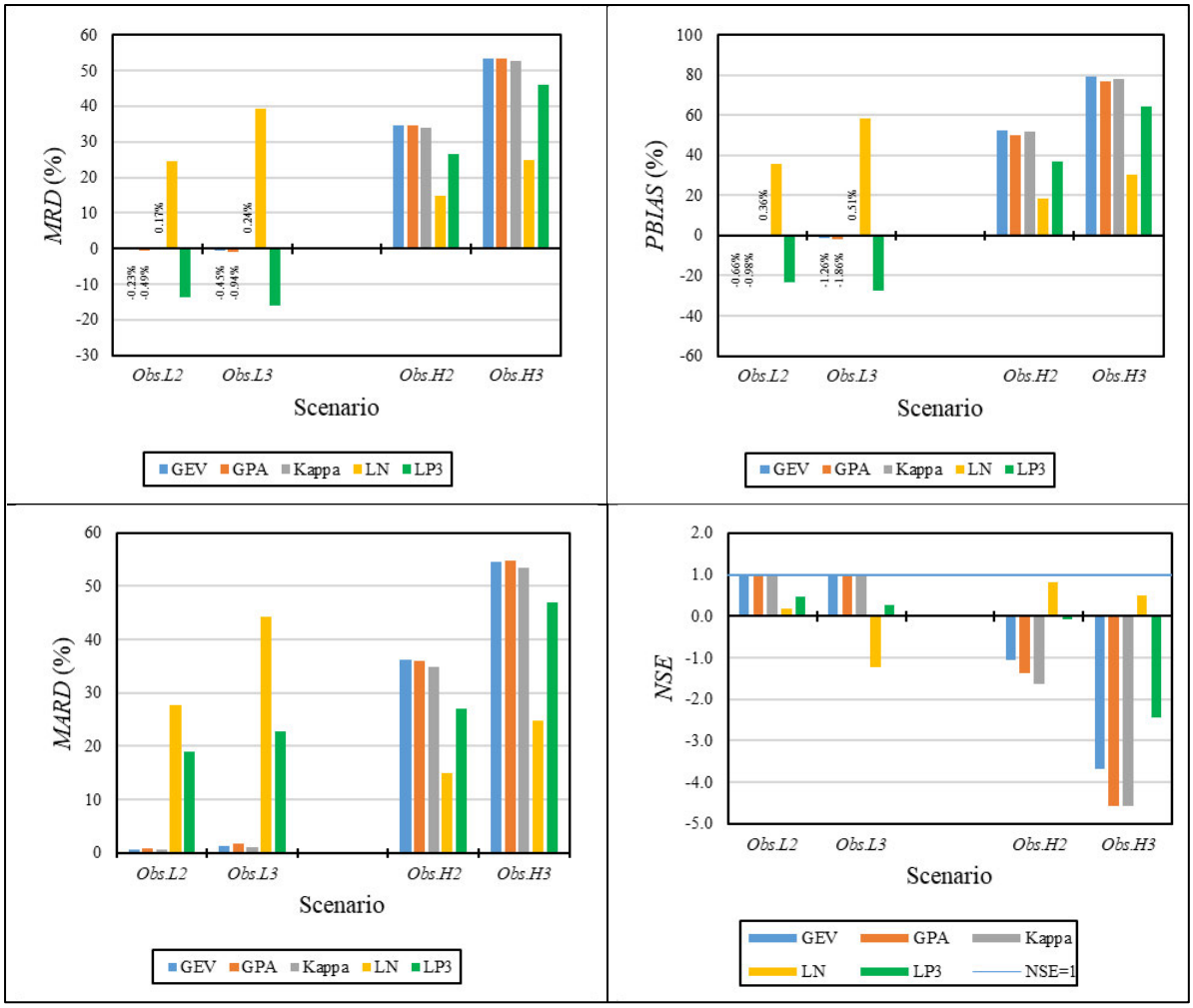

Figure B1.10 Impact of LO (*Obs.L2* and *Obs.L3*) and HO (*Obs.H2* and *Obs.H3*) scenarios on design floods estimated by different PDs from observed data in Catchment U2H013

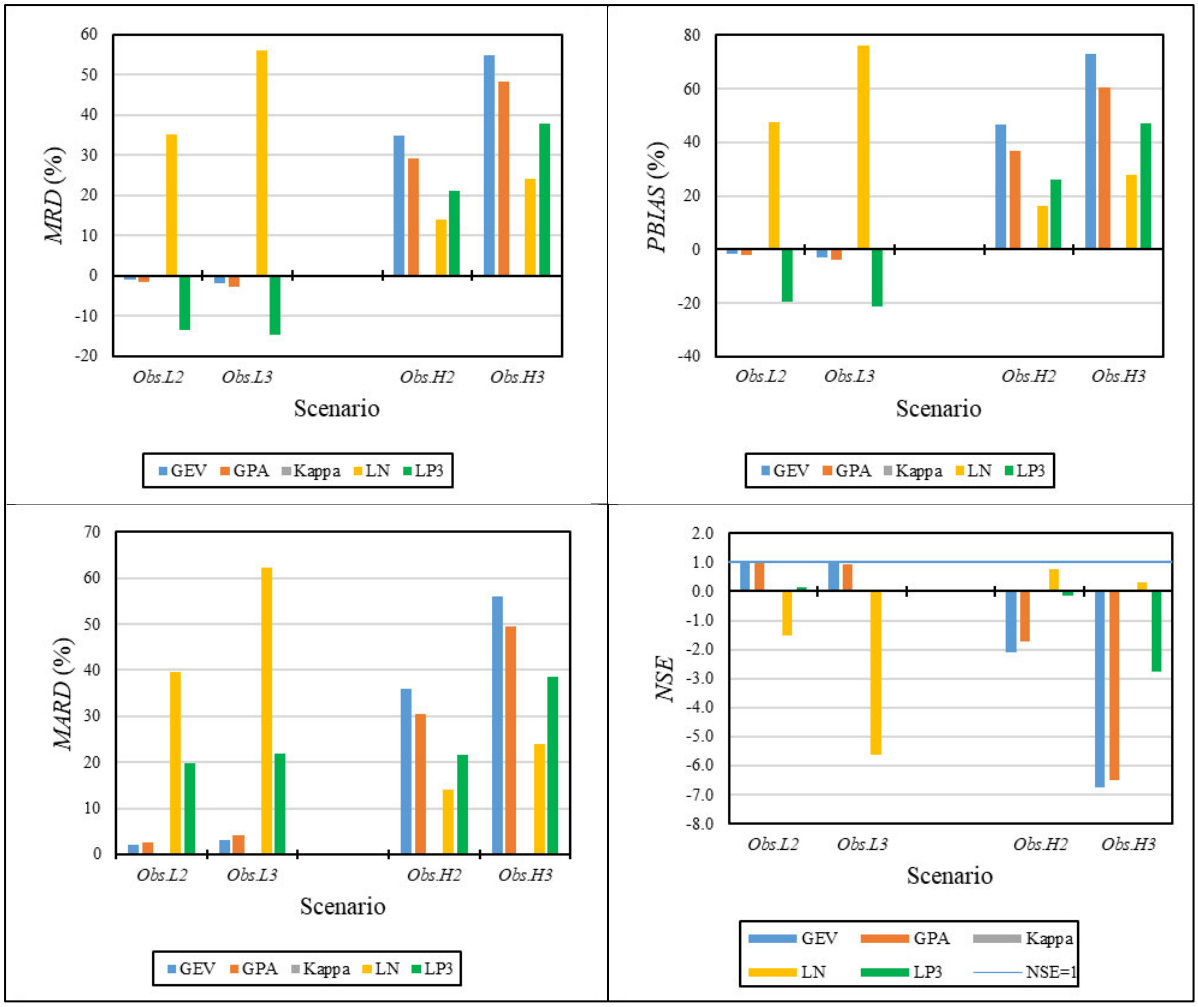

Figure B1.11 Impact of LO (*Obs.L2* and *Obs.L3*) and HO (*Obs.H2* and *Obs.H3*) scenarios on design floods estimated by different PDs from observed data in Catchment V2H004

11.2 Appendix B2: Synthetically Generated Data Series

Results for design rainfall estimates are first presented in Section 11.2.1 followed by results for design flood estimates in Section 11.2.2.

11.2.1 Design rainfall

Figure B2.1 to Figure B2.6 provides the calculated $Avg_{d..}$ *MRD*, $Avg_{d..}$ *MARD*, $Avg_{d..}$ *PBIAS* and $Avg_{d..}$ *NSE* for estimated design rainfall events of Rainfall Gauge 0476031, 0589670, 0025414, 0239097 and 0268640 respectively.

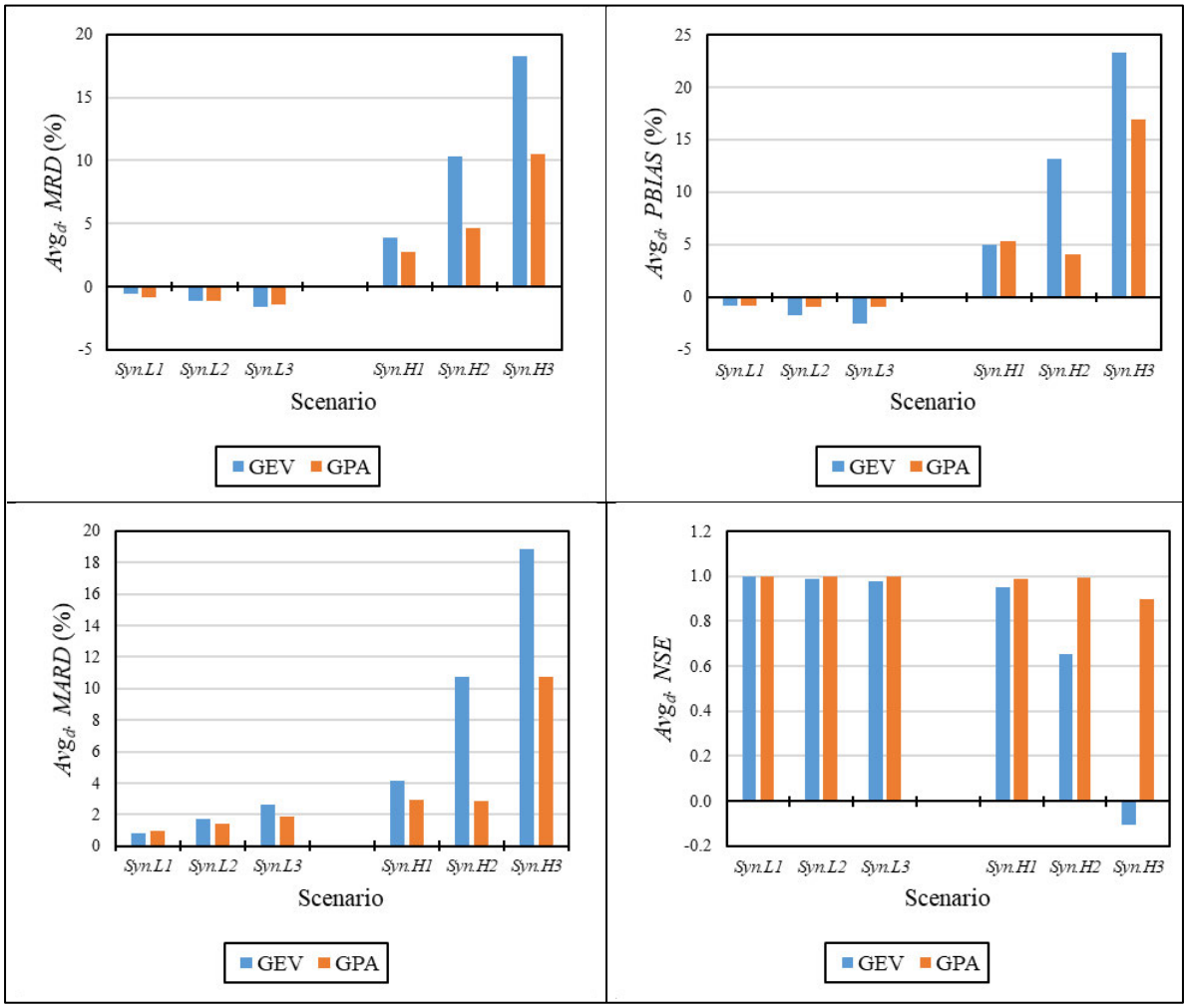

Figure B2.1 Impact of LO (*Obs.L1*, *Obs.L2* and *Obs.L3*) and HO (*Obs.H1*, *Obs.H2* and *Obs.H3*) scenarios on design rainfall estimated by different PDs from synthetically generated data series for Rainfall Gauge 0476031

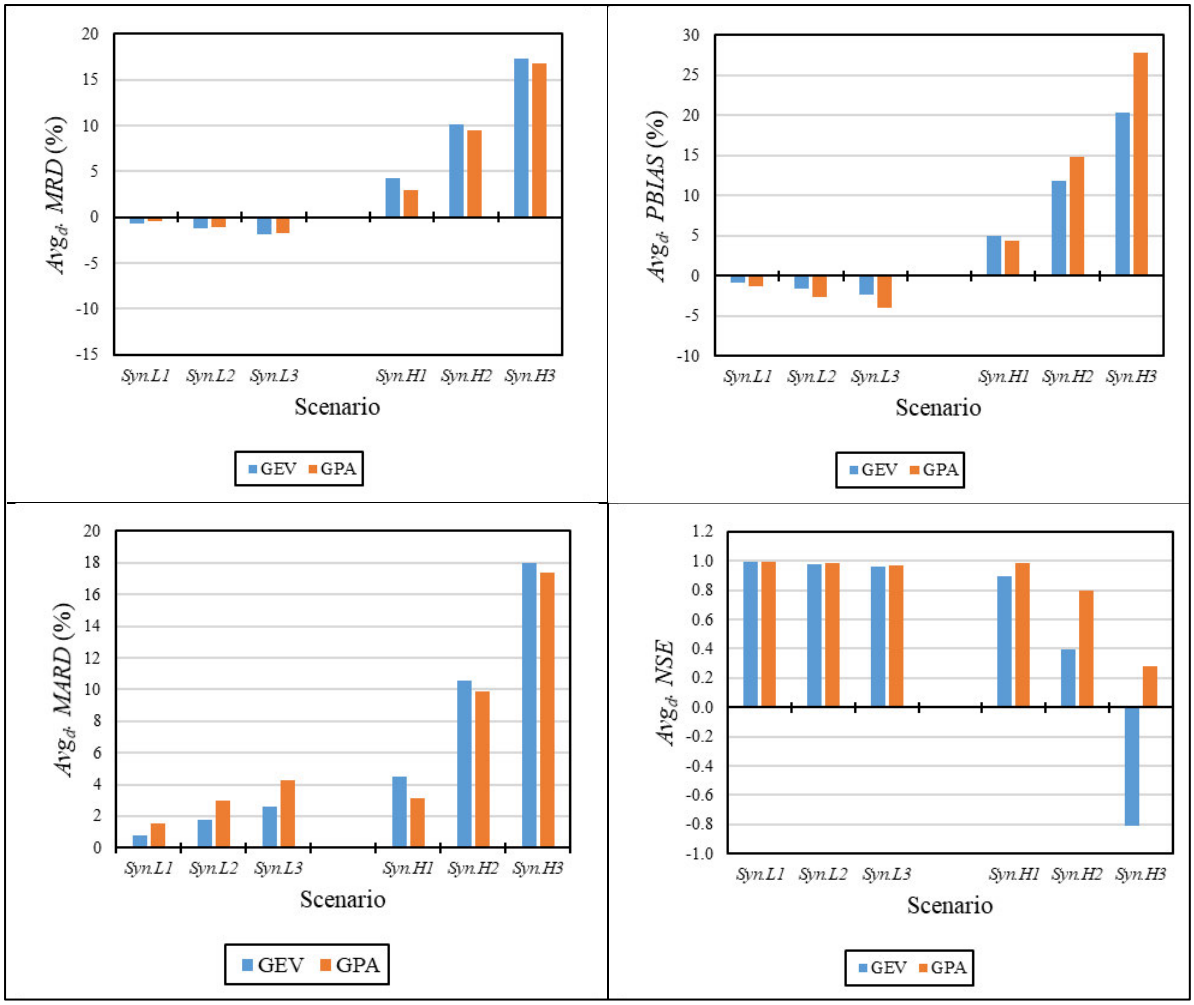

Figure B2.2 Impact LO (*Obs.L1*, *Obs.L2* and *Obs.L3*) and HO (*Obs.H1*, *Obs.H2* and *Obs.H3*) scenarios on design rainfall estimated by different PDs from synthetically generated data series for Rainfall Gauge 0589670

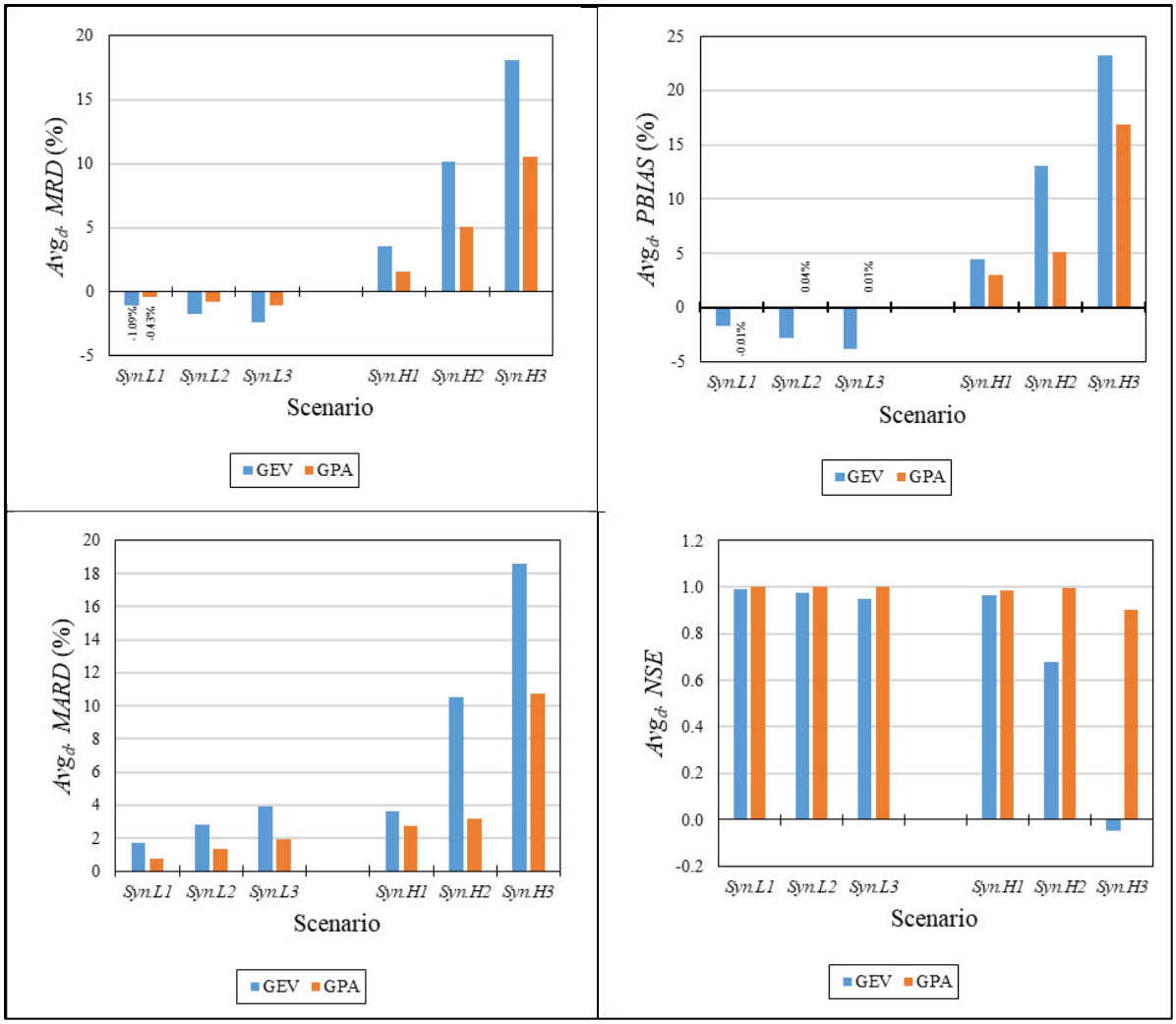

Figure B2. 3 Impact of LO (*Obs.L1*, *Obs.L2* and *Obs.L3*) and HO (*Obs.H1*, *Obs.H2* and *Obs.H3*) scenarios on design rainfall estimated by different PDs from synthetically generated data for Rainfall Gauge 0042227

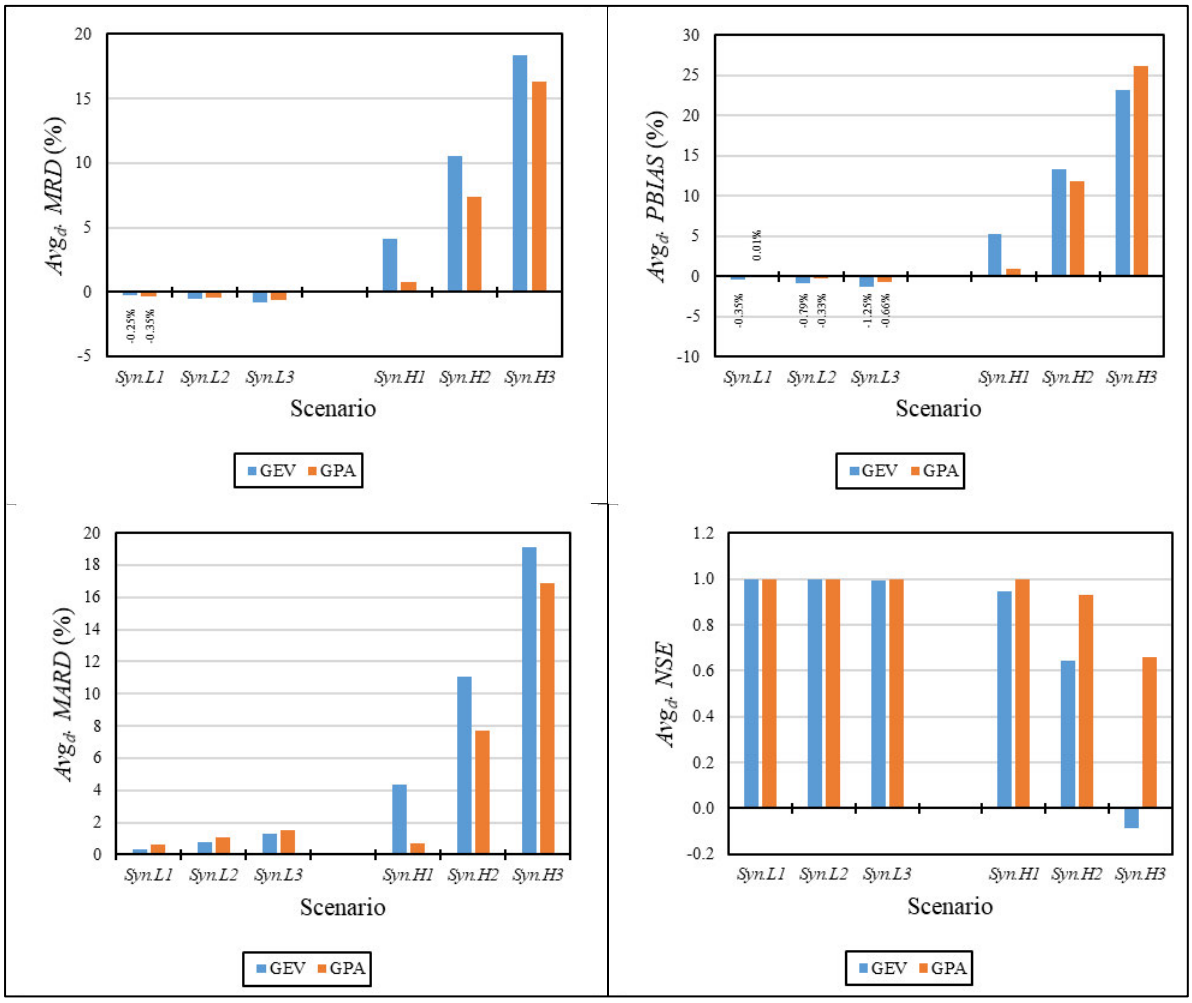

Figure B2.4 Impact of LO (*Obs.L1*, *Obs.L2* and *Obs.L3*) and HO (*Obs.H1*, *Obs.H2* and *Obs.H3*) scenarios on design rainfall estimated by different PDs from synthetically generated data for Rainfall Gauge 0025414

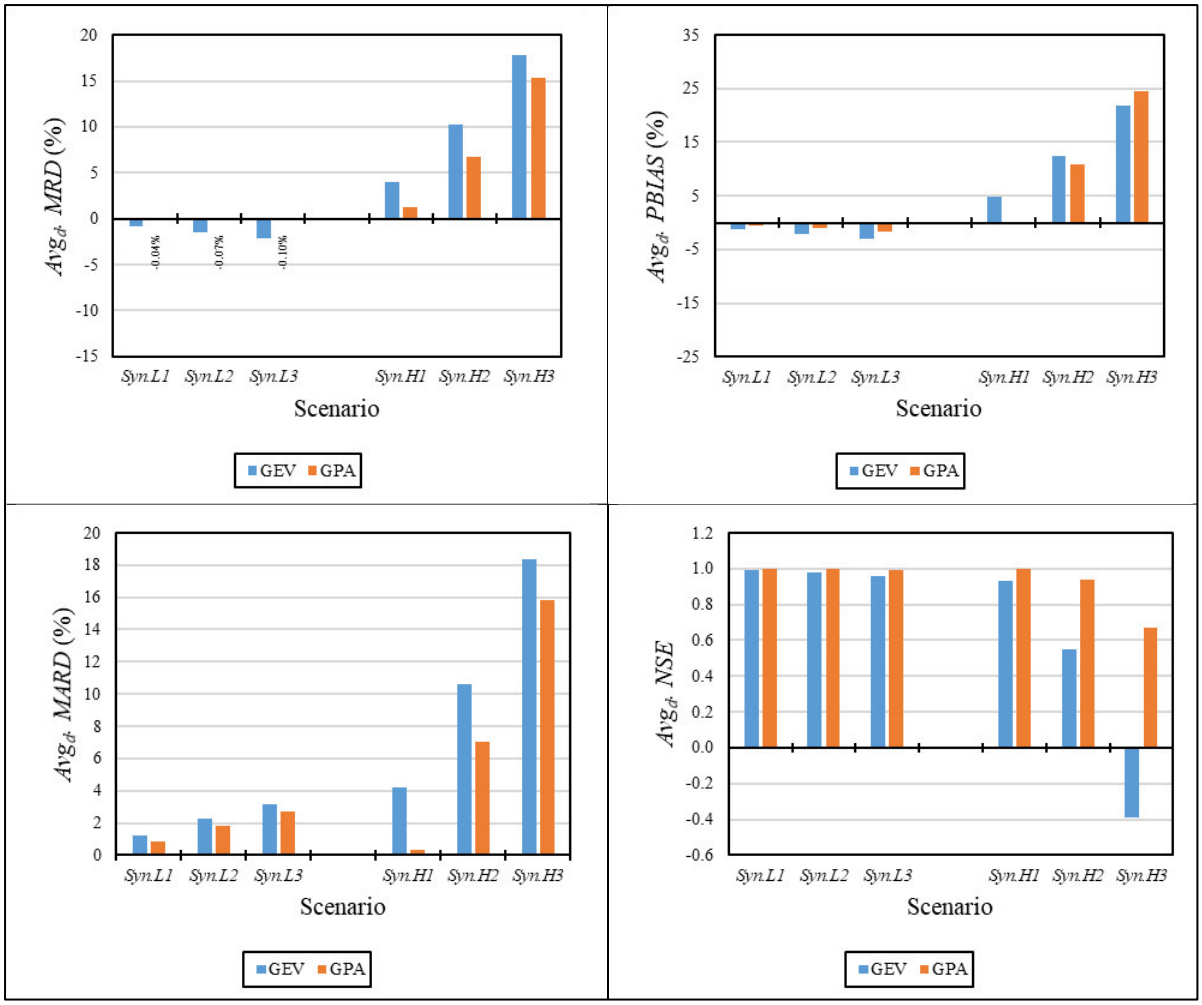

Figure B2.5 Impact of LO (*Obs.L1*, *Obs.L2* and *Obs.L3*) and HO (*Obs.H1*, *Obs.H2* and *Obs.H3*) scenarios on design rainfall estimated by different PDs from synthetically generated data series for Rainfall Gauge 0239097

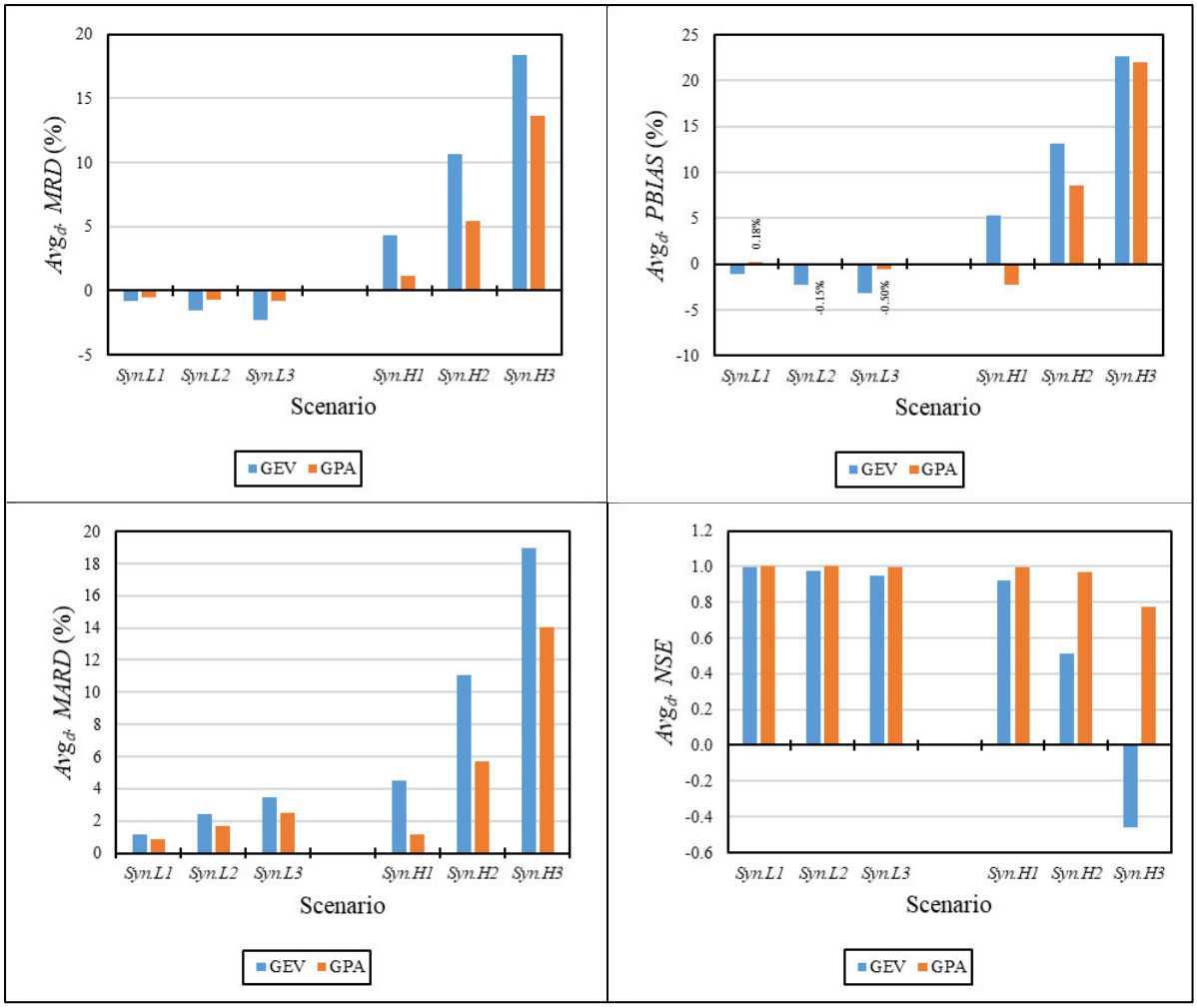

Figure B2.6 Impact LO (*Obs.L1*, *Obs.L2* and *Obs.L3*) and HO (*Obs.H1*, *Obs.H2* and *Obs.H3*) scenarios on design rainfall estimated by different PDs from synthetically generated data series for Rainfall Gauge 0268640

11.2.2 Design floods

Figure B2.7 to Figure B2.12 provides the calculated $Avg_{d..}$ *MRD*, $Avg_{d..}$ *MARD*, $Avg_{d..}$ *PBIAS* and $Avg_{d..}$ *NSE* for estimated design flood events in Catchment A2H012, A6H011, H7H004, G1H008, U2H013 and V2H004 respectively.

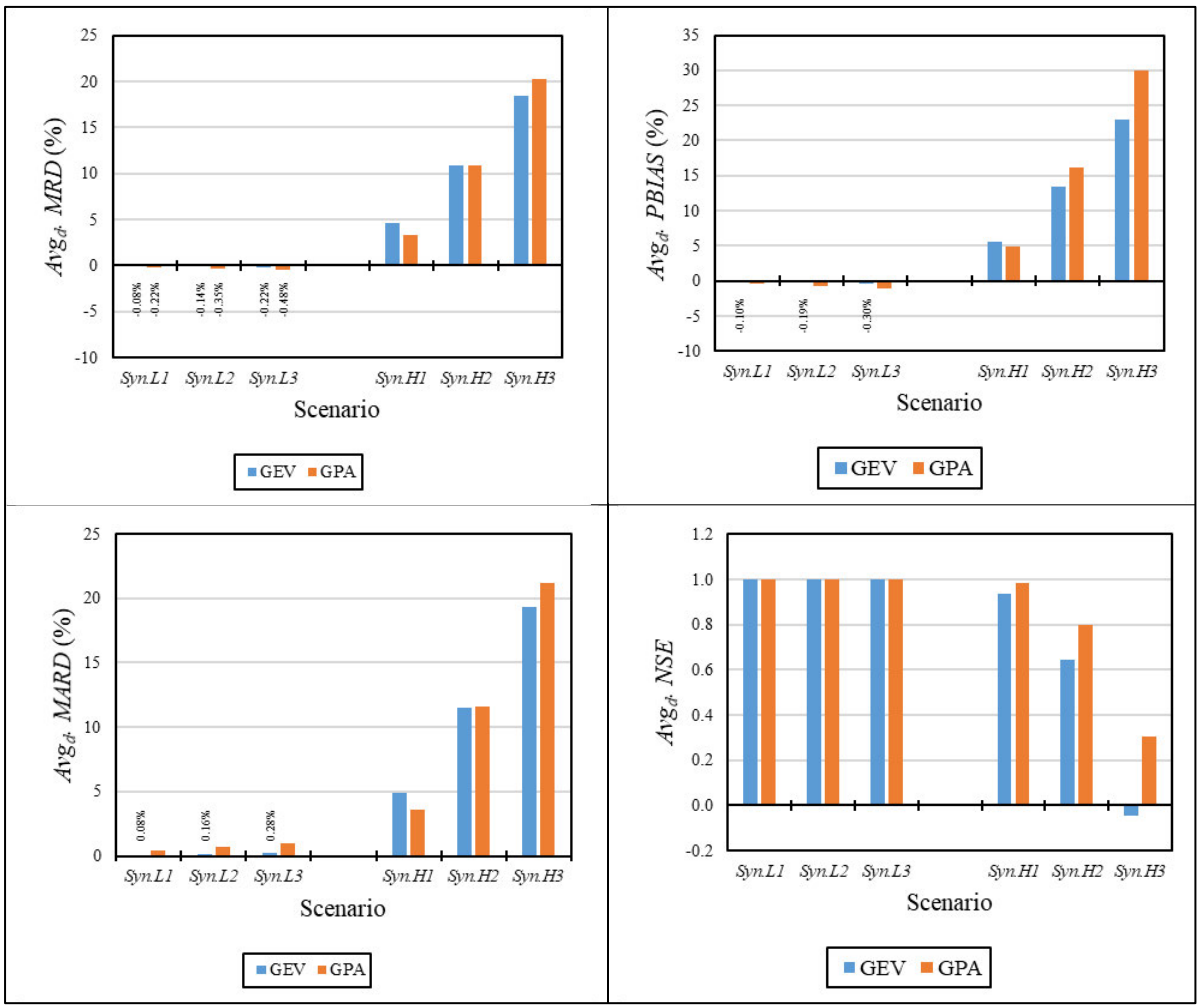

Figure B2.7 Impact of LO (*Obs.L1*, *Obs.L2* and *Obs.L3*) and HO (*Obs.H1*, *Obs.H2* and *Obs.H3*) scenarios on design floods estimated by different PDs from synthetically generated data series in Catchment A2H012

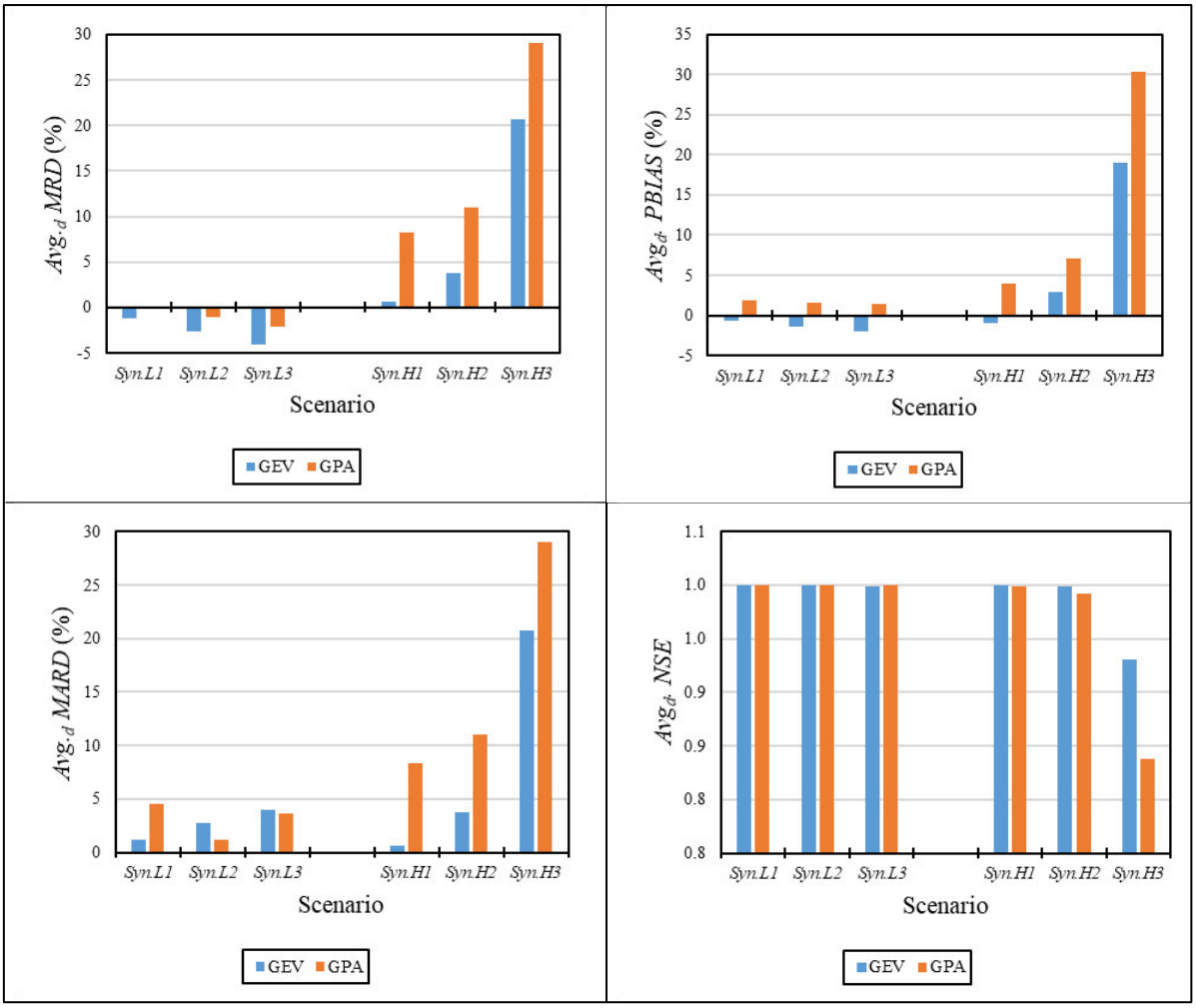

Figure B2.8 Impact of LO (*Obs.L1*, *Obs.L2* and *Obs.L3*) and HO (*Obs.H1*, *Obs.H2* and *Obs.H3*) scenarios on design floods estimated by different PDs from synthetically generated data series in Catchment A6H011

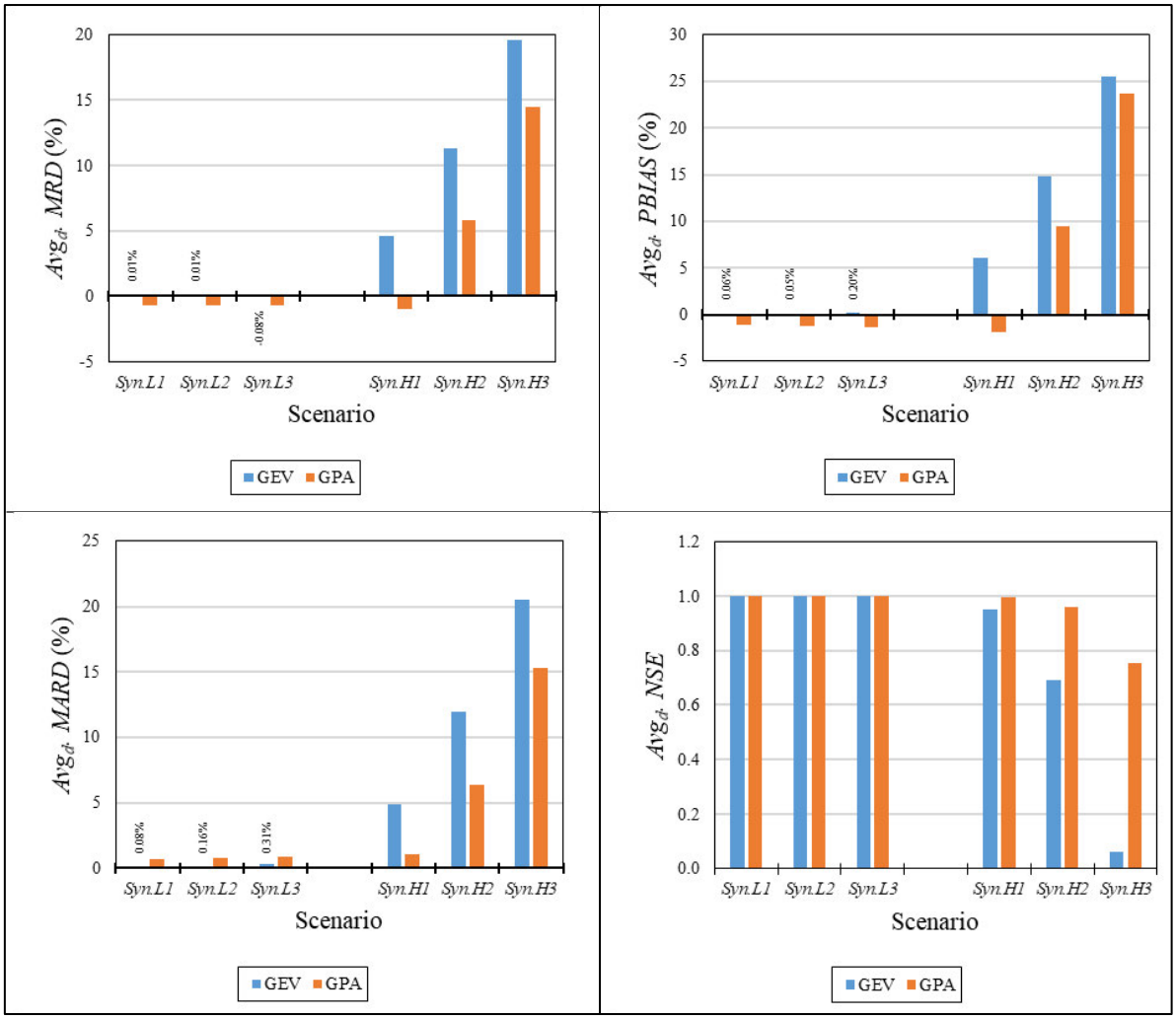

Figure B2.9 Impact of LO (*Obs.L1*, *Obs.L2* and *Obs.L3*) and HO (*Obs.H1*, *Obs.H2* and *Obs.H3*) scenarios on design floods estimated by different PDs from synthetically generated data series in Catchment G1H008

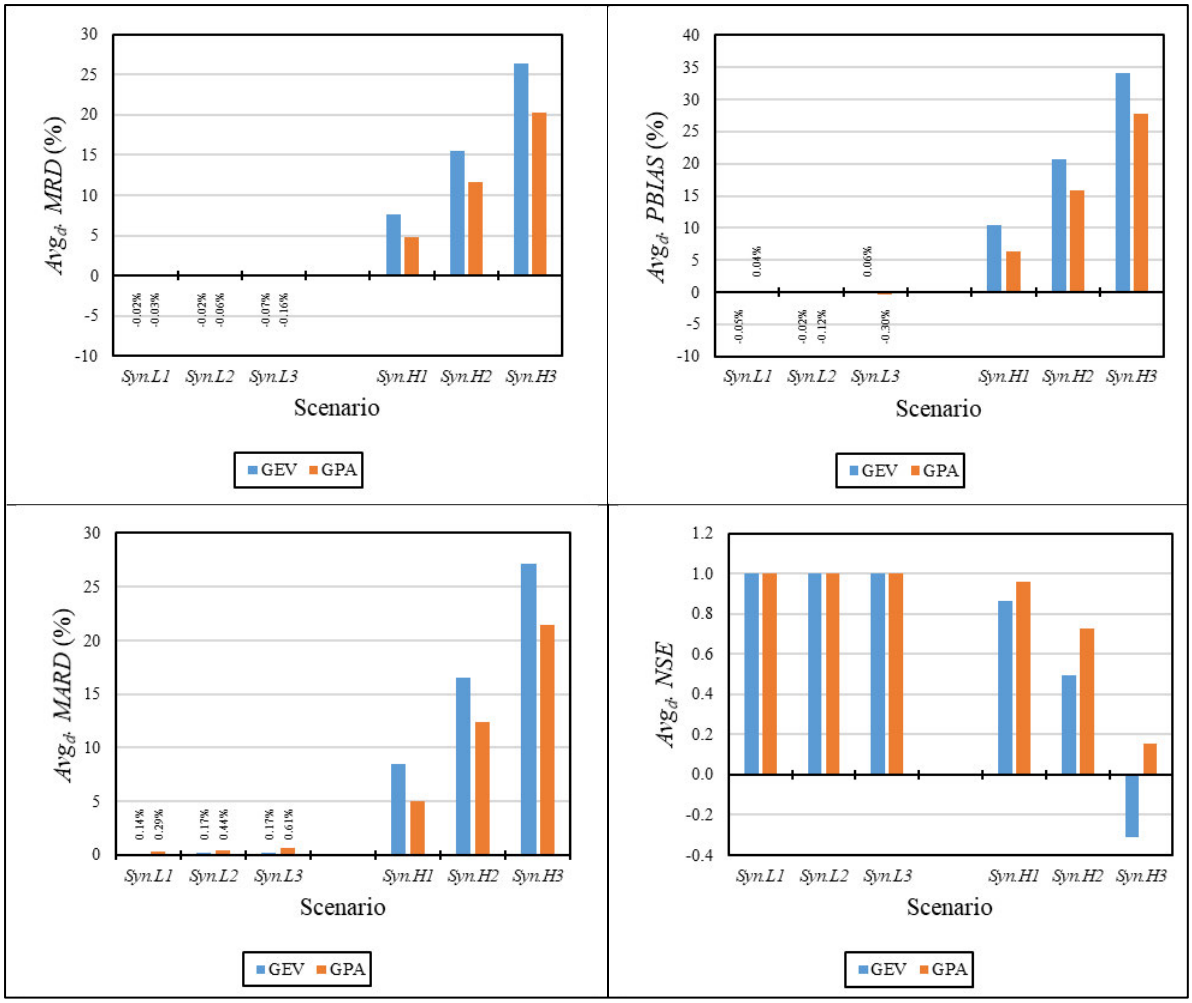

Figure B2.10 Impact of LO (*Obs.L1*, *Obs.L2* and *Obs.L3*) and HO (*Obs.H1*, *Obs.H2* and *Obs.H3*) scenarios on design floods estimated by different PDs from synthetically generated data series in Catchment H7H004

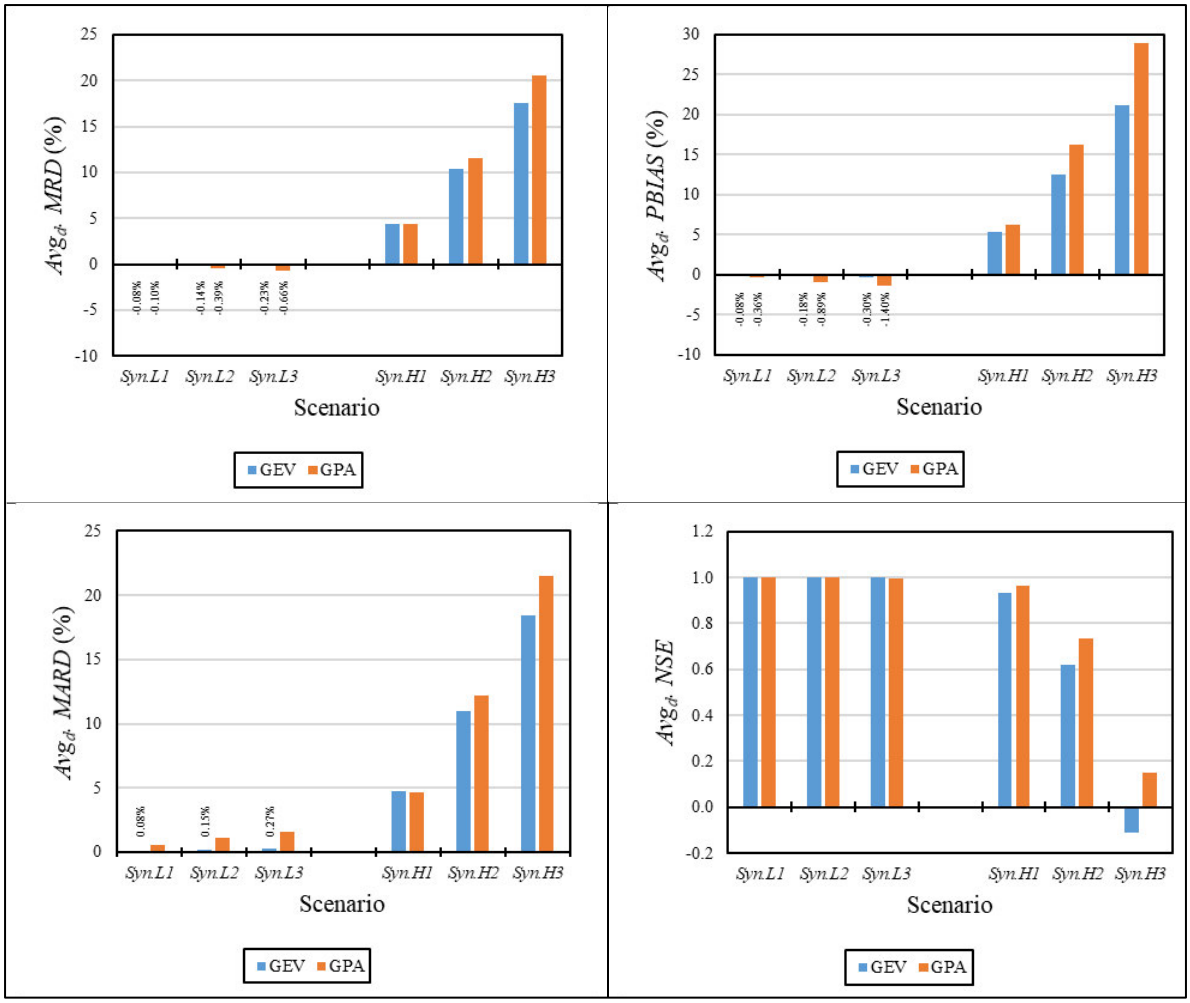

Figure B2.11 Impact of LO (*Obs.L1*, *Obs.L2* and *Obs.L3*) and HO (*Obs.H1*, *Obs.H2* and *Obs.H3*) scenarios on design floods estimated by different PDs from synthetically generated data series in Catchment U2H013

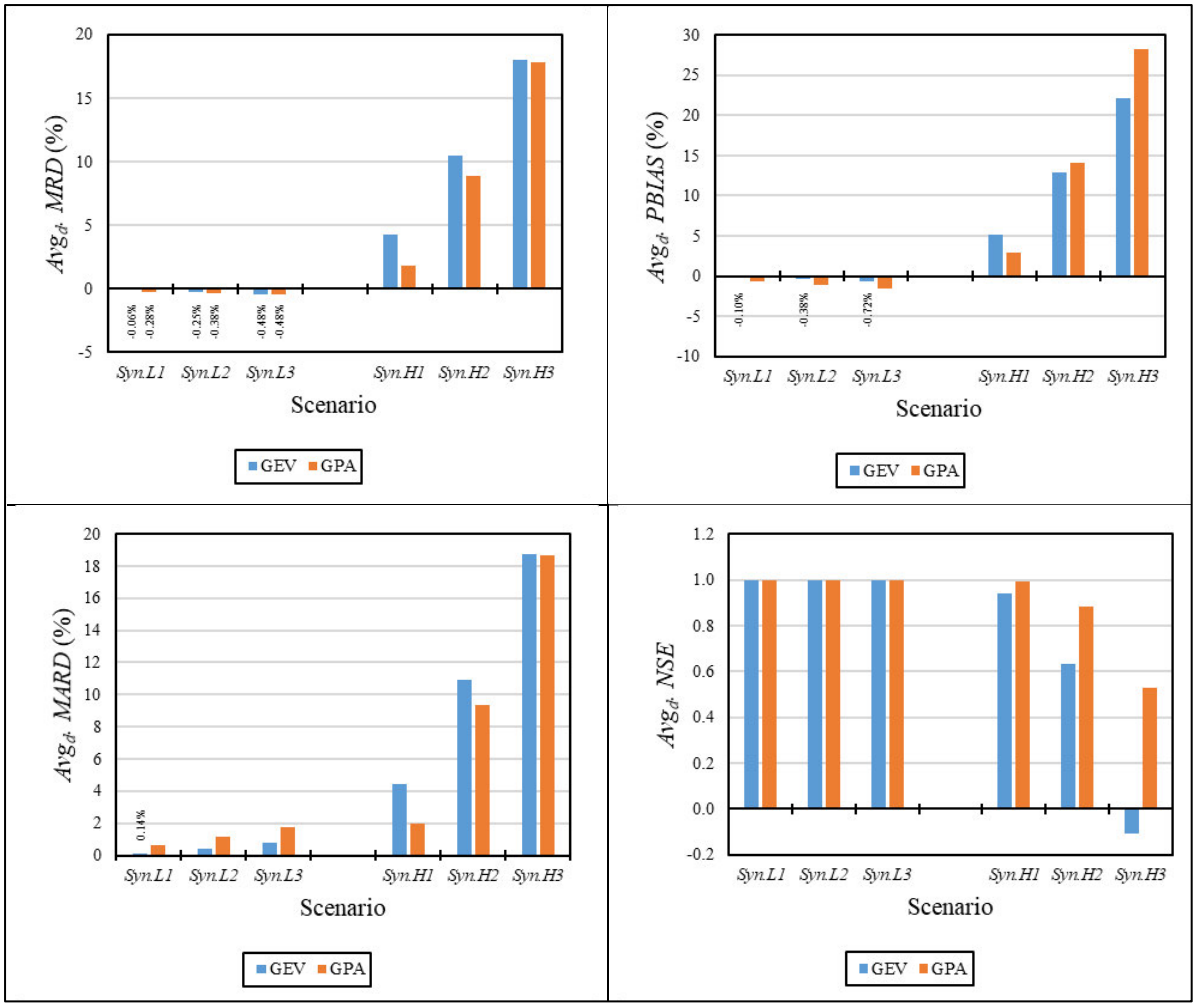

Figure B2.12 Impact of LO (*Obs.L1*, *Obs.L2* and *Obs.L3*) and HO (*Obs.H1*, *Obs.H2* and *Obs.H3*) scenarios on design floods estimated by different PDs from synthetically generated data series in Catchment V2H004

12. APPENDIX C: PERFORMANCE OF OUTLIER DETECTION METHODS

This appendix provides results on the percentage of substituted and detected outliers using the BP, MZS and MGBT method on observed streamflow data and synthetically generated rainfall and streamflow data series with LO and HO scenarios, as shown in Figure C.1 to Figure C.9.

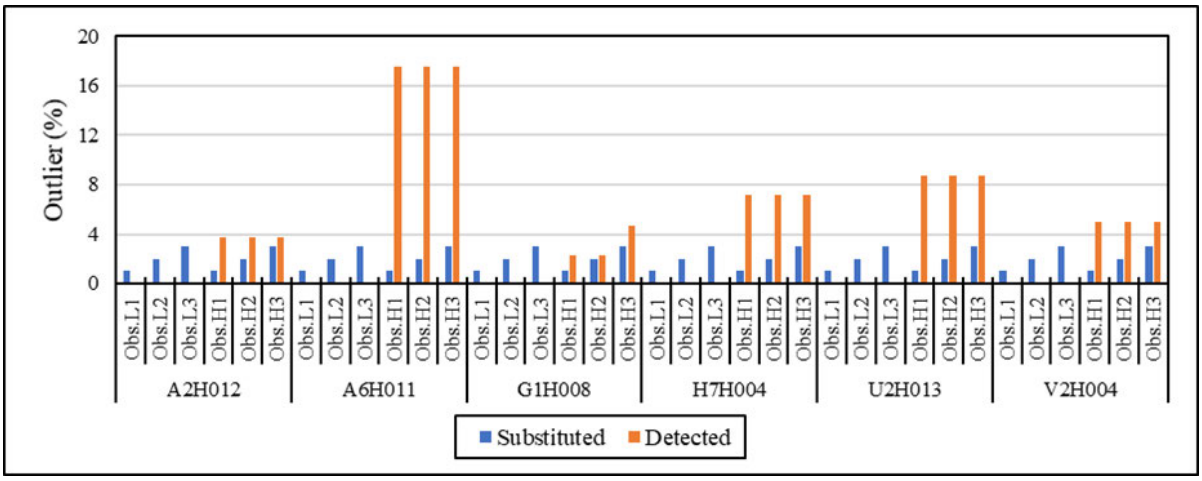

Figure C.1 Percentage of substituted and detected outliers using the BP method on observed streamflow data with LO (*Obs.L1*, *Obs.L2* and *Obs.L3*) and HO (*Obs.H1*, *Obs.H2* and *Obs.H3*) scenarios

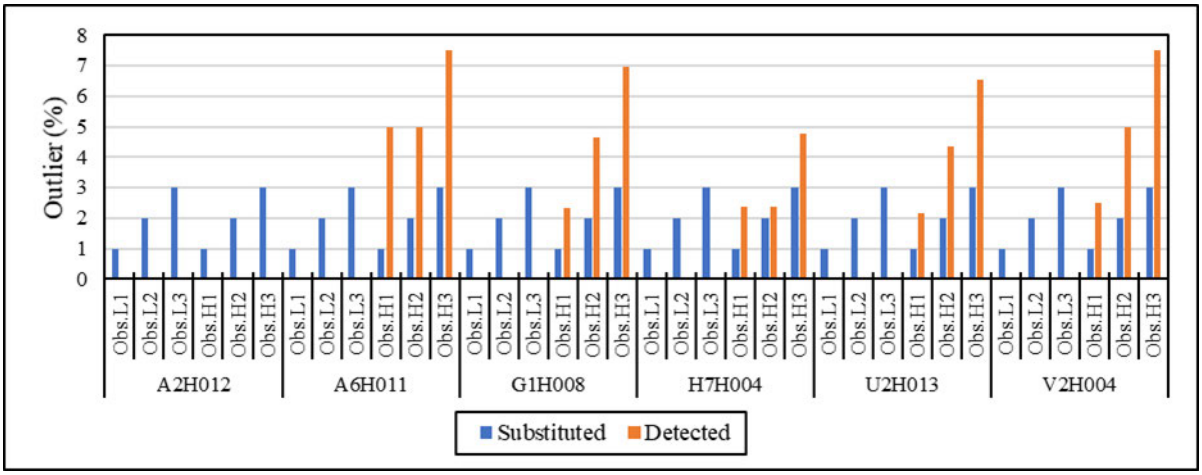

Figure C.2 Percentage of substituted and detected outliers using the MZS method on observed streamflow data with LO (*Obs.L1*, *Obs.L2* and *Obs.L3*) and HO (*Obs.H1*, *Obs.H2* and *Obs.H3*) scenarios

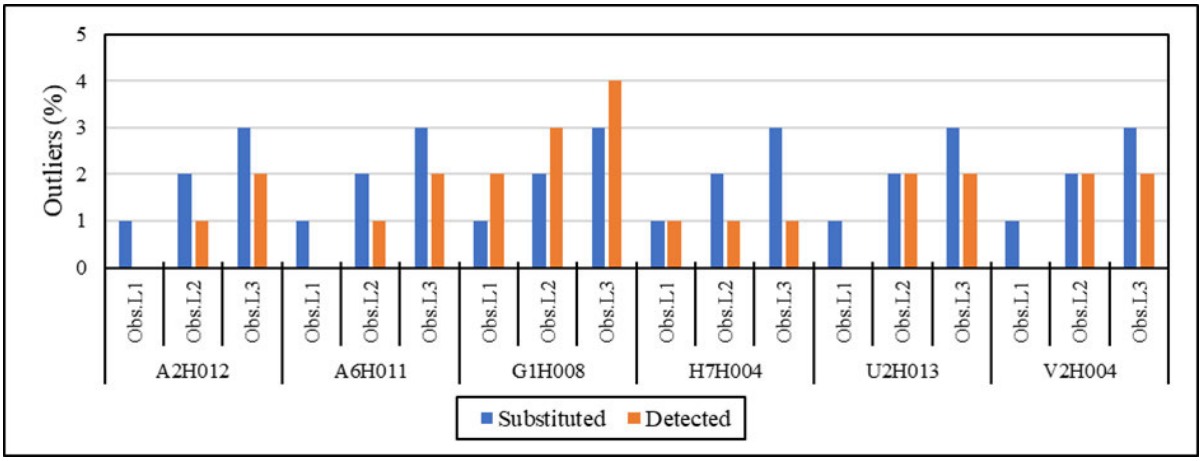

Figure C.3 Percentage of substituted and detected outliers using the MGBT method on observed streamflow data with LO (*Obs.L1*, *Obs.L2* and *L3*) scenarios

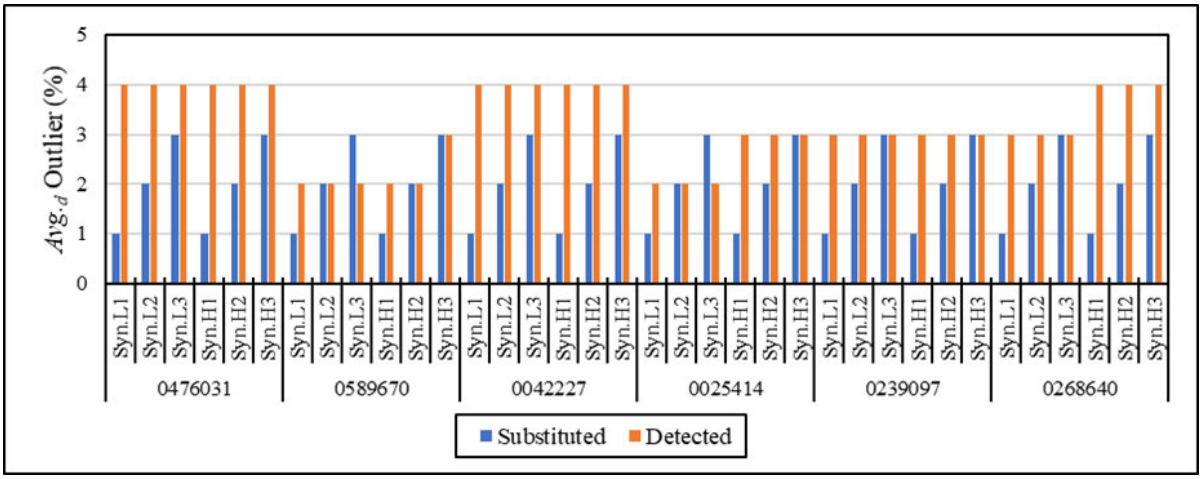

Figure C.4 Percentage of substituted and detected outliers using the BP method on synthetically generated rainfall data series with LO (*Syn.L1*, *Syn.L2* and *Syn.L3*) and HO (*Syn.H1*, *Syn.H2* and *Syn.H3*) scenarios

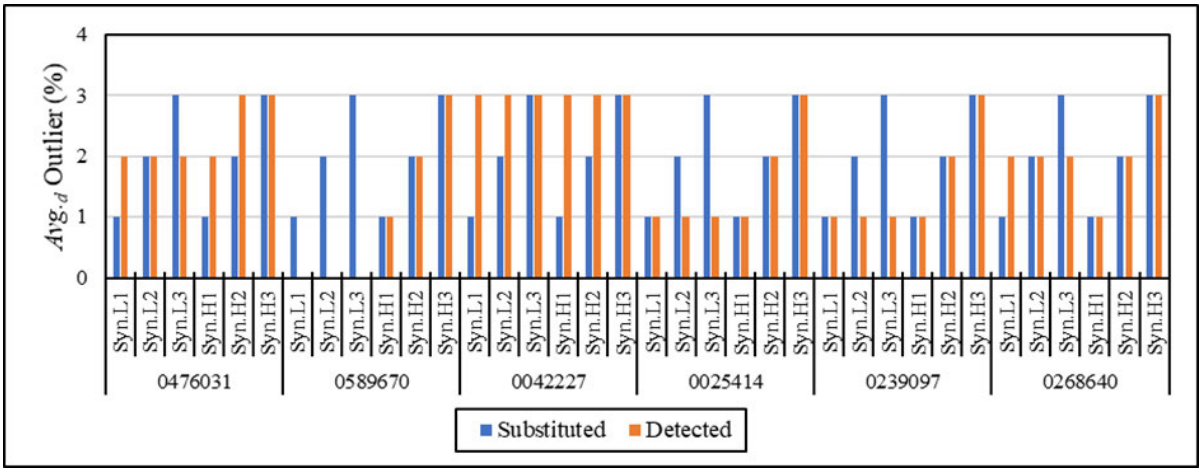

Figure C.5 Percentage of substituted and detected outliers using the MZS method synthetically generated rainfall data series with LO (*Syn.L1*, *Syn.L2* and *Syn.L3*) and HO (*Syn.H1*, *Syn.H2* and *Syn.H3*) scenarios

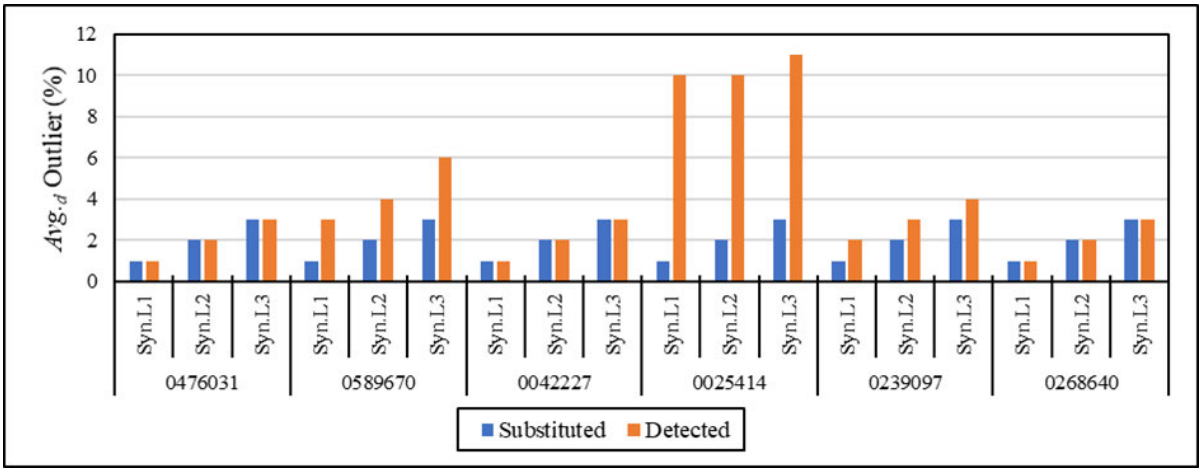

Figure C.6 Percentage of substituted and detected outliers using the MGBT method on synthetically generated rainfall data series with LO (*Syn.L1*, *Syn.L2* and *Syn.L3*) scenarios

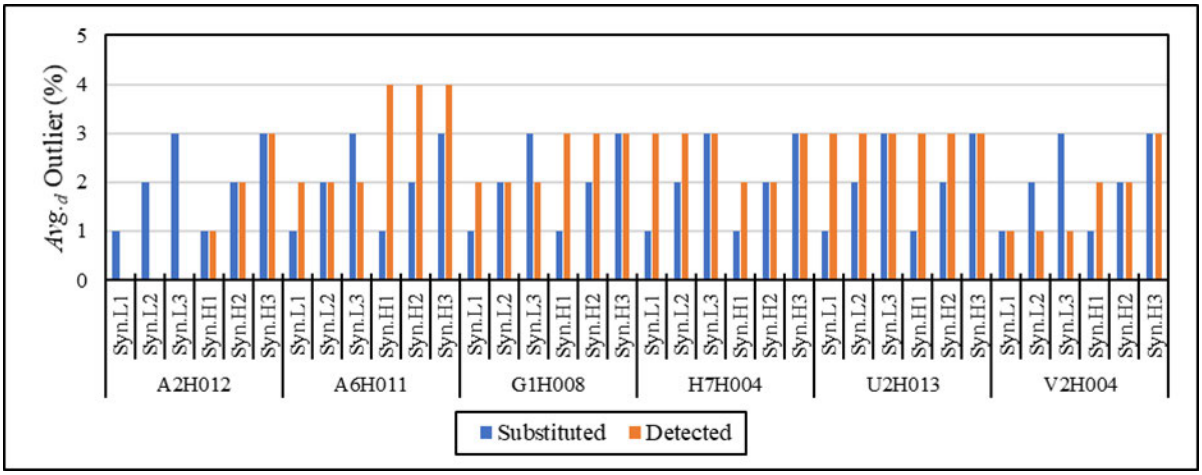

Figure C.7 Percentage of substituted and detected outliers using the BP method on synthetically generated streamflow data series with LO (*Syn.L1*, *Syn.L2* and *Syn.L3*) and HO (*Syn.H1*, *Syn.H2* and *Syn.H3)* scenarios

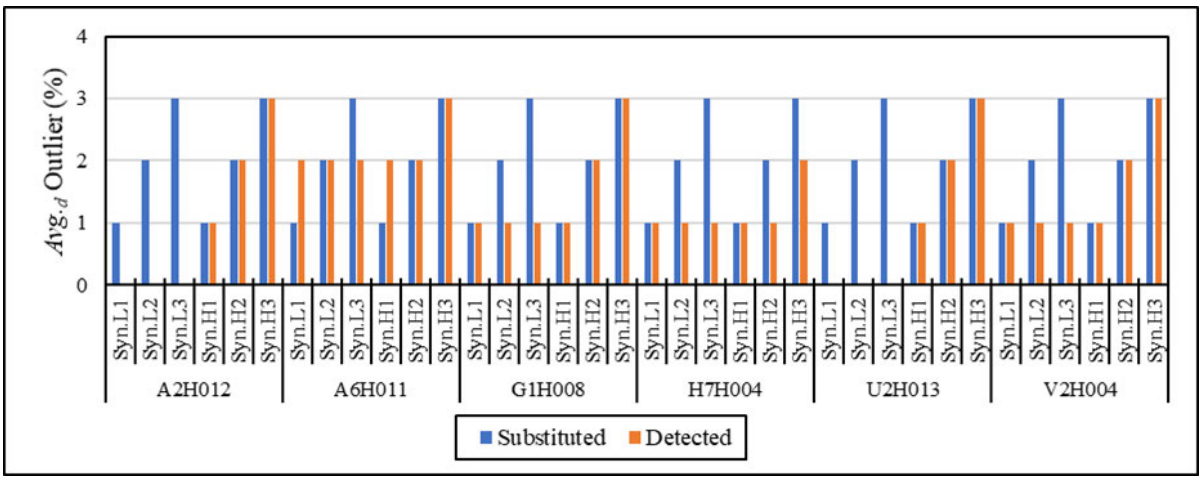

Figure C.8 Percentage of substituted and detected outliers using the MZS method synthetically generated streamflow data series with LO (*Syn.L1*, *Syn.L2* and *Syn.L3*) and HO (*Syn.H1*, *Syn.H2* and *Syn.H3)* scenarios

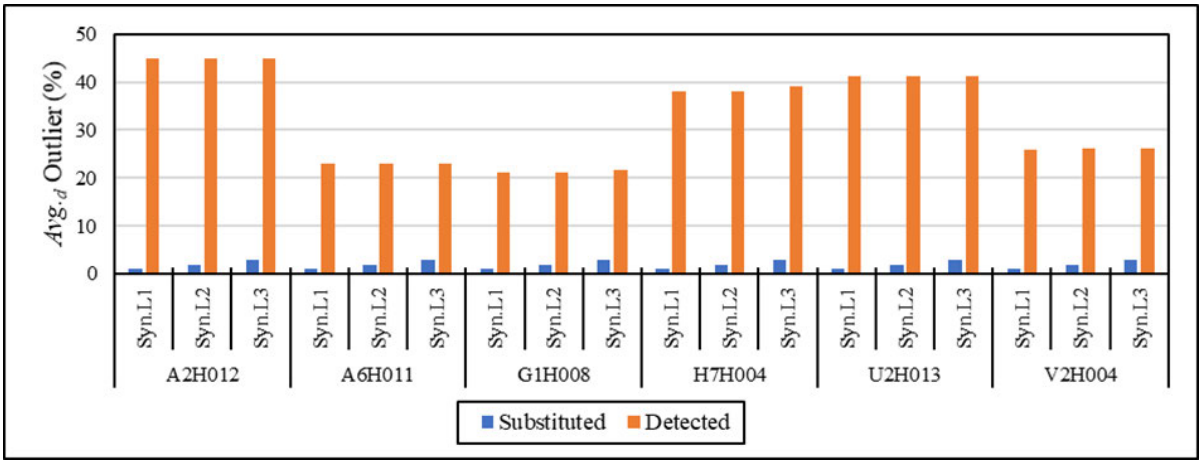

Figure C.9 Percentage of substituted and detected outliers using the MGBT method on synthetically generated streamflow data series with LO (*Syn.L1*, *Syn.H2* and *Syn.L3*) scenarios

13. APPENDIX D: IMPACT OF REDUCED NETWORK DENSITY ON DESIGN RAINFALL AND FLOOD ESTIMATION

This appendix contains attribute information and results regarding the impact of a reduced gauged network *cf.* Chapter 7.

13.1 Appendix D1: Gauge Inventory for the Impact of Reduced Network Density on Design Rainfall And Flood Estimations

Appendix D1 contains an inventory of the rainfall and streamflow gauges used to evaluate the impact of reduced record length on DRE and DFE. Table D1.1 and Figure D1.1, Figure D1.2 and Figure D1.3 detailing rainfall gauges with Table D1.2 and Figure D1.4, Figure D1.5 and Figure D1.6 detailing streamflow gauges with associated attributes.

Table D1.1 Rainfall gauges and associated attributes used in the analyse of a reduced gauge network on design rainfall estimation

Station Number	Cluster	Gericke (2015) Climate Zone	Latitude (^{0})	Longitude (^{0})	Start Date	End Date	Total Record Length (Years)	Reliable Record Length (Years)	MAP (mm)	Altitude (m)
0151604	74	ESC	-31.07	28.35	1884/03	2001/12	115	99	746	1250
0151623	74	ESC	-31.40	28.34	1880/01	2000/07	116	52	637	1338
0152190	74	ESC	-31.17	28.62	1888/04	2001/08	111	74	1091	1378
0152259	74	ESC	-31.36	28.60	1884/03	2001/12	116	69	913	1036
0152475	74	ESC	-31.42	28.77	1888/04	2001/12	110	77	1211	1200
0152792	74	ESC	-31.20	28.92	1888/04	2000/07	109	71	834	991
0153631	74	ESC	-31.02	29.36	1882/09	2001/12	107	74	1168	1039
0179353	74	ESC	-30.88	28.68	1884/03	2001/07	112	68	981	1298
0179713	74	ESC	-30.89	28.89	1888/04	2001/12	111	78	1130	1311
0179790	74	ESC	-30.67	28.95	1890/03	2000/08	109	68	897	1463

Station Number	Cluster	Gericke (2015) Climate Zone	Latitude (^{0})	Longitude (^{0})	Start Date	End Date	Total Record Length (Years)	Reliable Record Length (Years)	MAP (mm)	Altitude (m)
0179864	74	ESC	-30.90	28.99	1888/04	2000/07	110	65	845	1134
0180030	74	ESC	-31.00	29.02	1888/04	2001/12	111	74	754	1154
0180123	74	ESC	-30.56	29.08	1882/09	2000/08	110	70	982	1481
0180439	74	ESC	-30.82	29.26	1882/09	2001/12	112	79	921	1118
0180537	74	ESC	-30.96	29.30	1882/09	2001/12	111	80	646	1129
0590361	28	NI	-24.52	28.72	1903/09	2000/08	96	89	608	1100
0590486	28	NI	-24.60	28.79	1903/09	2000/08	96	64	604	1082
0632274	28	NI	-24.06	28.17	1903/02	2000/08	96	62	653	1463
0632726	28	NI	-24.11	28.42	1903/09	2000/08	96	48	625	1320
0633463	28	NI	-24.22	28.77	1903/09	2000/08	96	56	591	1130
0633796	28	NI	-24.27	28.96	1903/09	2000/08	96	65	604	1082
0633881	28	NI	-24.18	29.01	1903/09	2000/08	96	82	621	1094
0634131	28	NI	-24.18	29.07	1903/09	2000/08	96	77	584	1219
0676237	28	NI	-23.95	28.63	1903/09	2000/08	96	56	514	988
0676705	28	NI	-23.75	28.91	1903/09	2000/08	96	50	489	1082
0002885	1	SWC	-34.75	20.00	1875/02	2001/12	124	87	471	12
0003032	1	SWC	-34.53	20.04	1875/02	2000/07	122	109	467	91

Station Number	Cluster	Gericke (2015) Climate Zone	Latitude (^{0})	Longitude (^{0})	Start Date	End Date	Total Record Length (Years)	Reliable Record Length (Years)	MAP (mm)	Altitude (m)
0003192	1	SWC	-34.71	20.10	1875/02	2001/12	122	48	413	12
0007698	1	SWC	-34.13	19.90	1875/02	2001/09	123	52	431	295
0007699	1	SWC	-34.16	19.89	1875/02	2001/12	124	51	418	162
0008136	1	SWC	-34.27	20.08	1875/02	2000/07	122	49	402	244
0008367	1	SWC	-34.12	20.23	1875/02	2001/12	122	67	333	168
0008470	1	SWC	-34.32	20.30	1875/02	2001/12	124	67	387	168
0008751	1	SWC	-34.00	20.44	1883/05	2001/12	115	68	884	250
0008782	1	SWC	-34.03	20.45	1883/05	2001/12	115	108	740	143
0023619	1	SWC	-33.81	19.88	1877/10	2000/07	121	60	287	183
0023674	1	SWC	-33.76	19.88	1877/12	2000/07	121	61	499	375
0023678	1	SWC	-33.80	19.88	1877/10	2000/07	121	104	322	183
0024146	1	SWC	-33.94	20.10	1877/10	2000/07	121	56	267	120
0024197	1	SWC	-33.79	20.13	1877/10	2001/12	122	86	321	223
0025414	1	SWC	-33.91	20.73	1878/01	2001/12	121	74	281	375

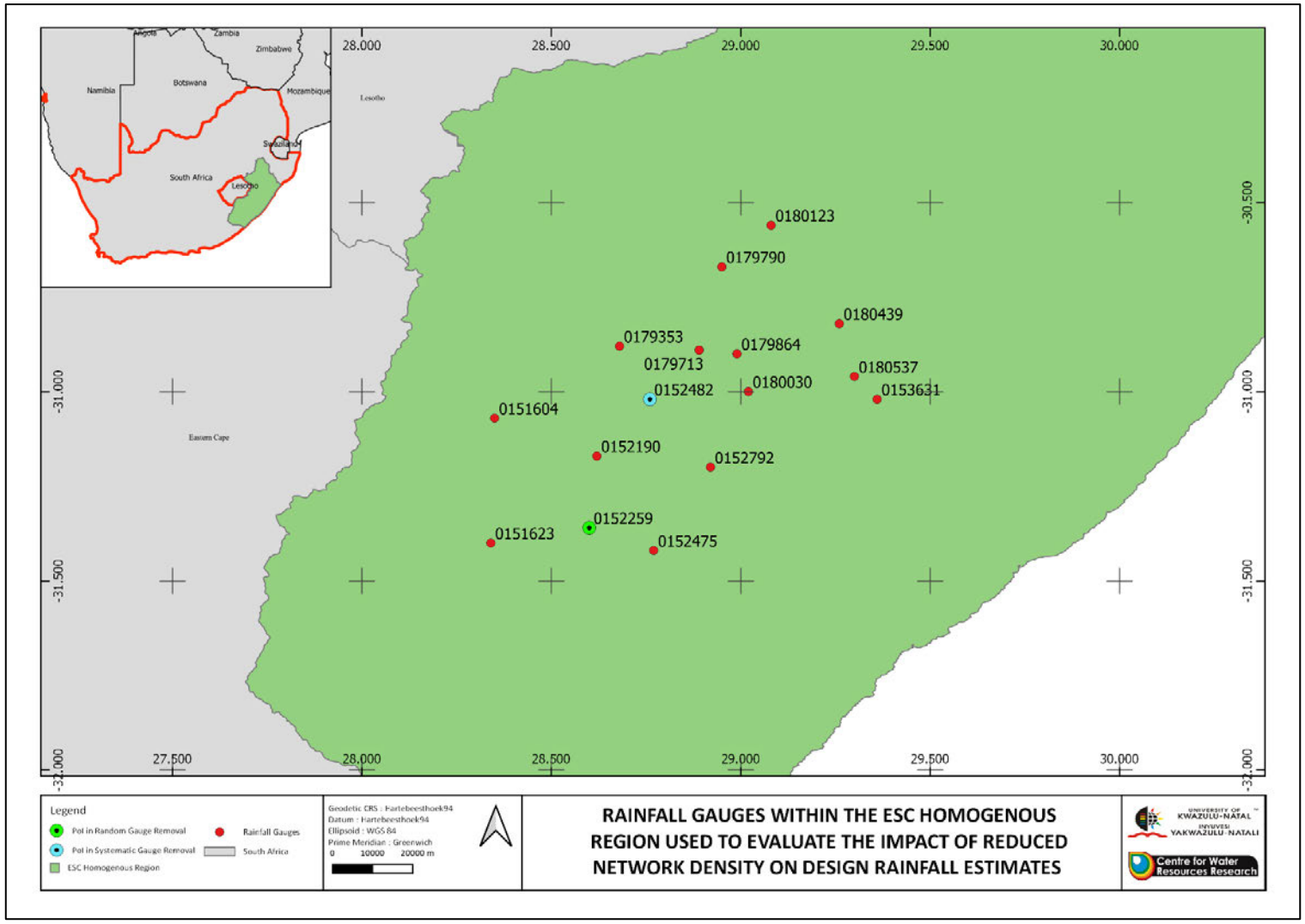

Figure D1.1 Rainfall gauges within the ESC homogenous region used to evaluate the impact of reduced network density on design rainfall estimates

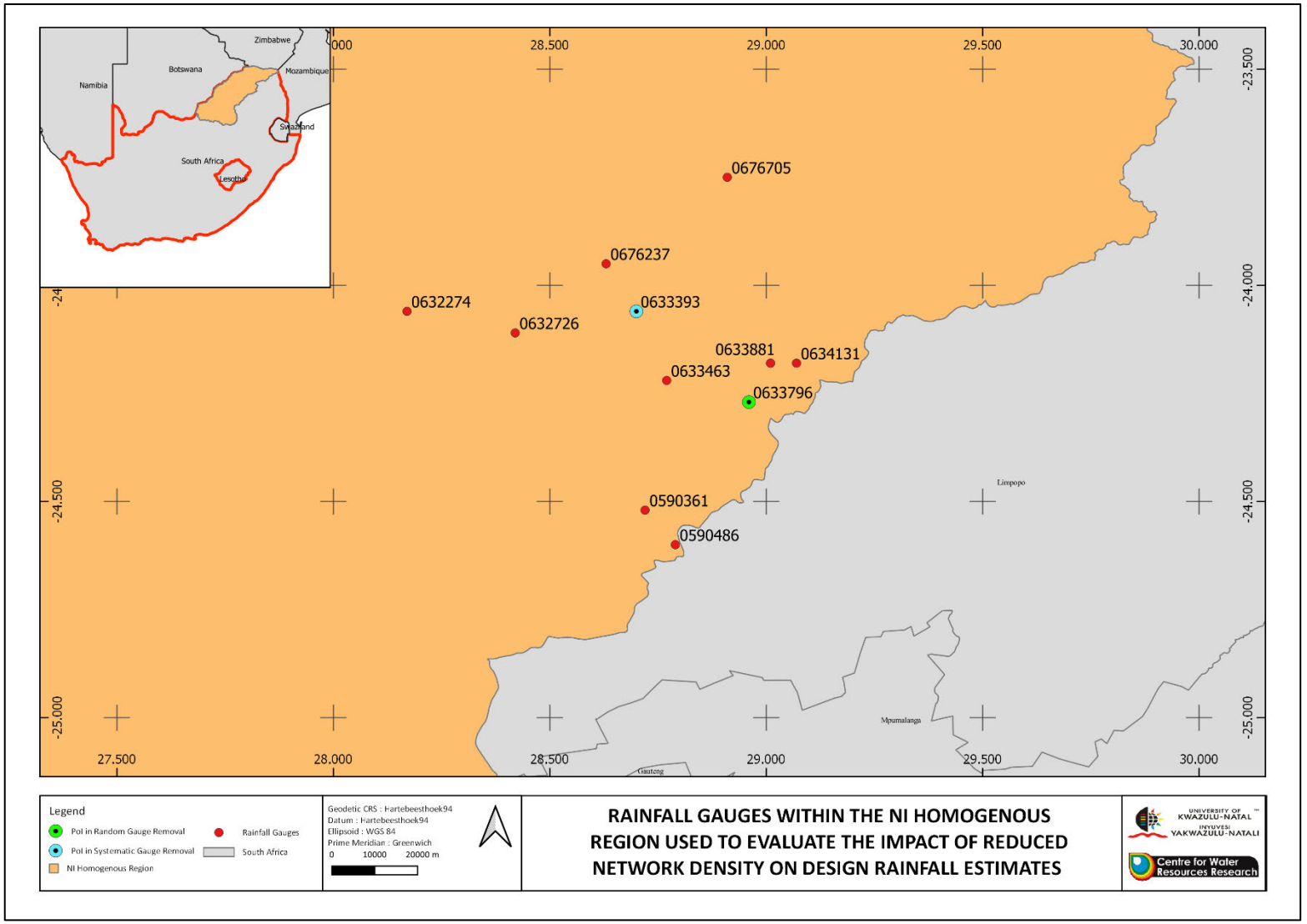

Figure D1.2 Rainfall gauges within the NI homogenous region used to evaluate the impact of reduced network density on design rainfall estimates

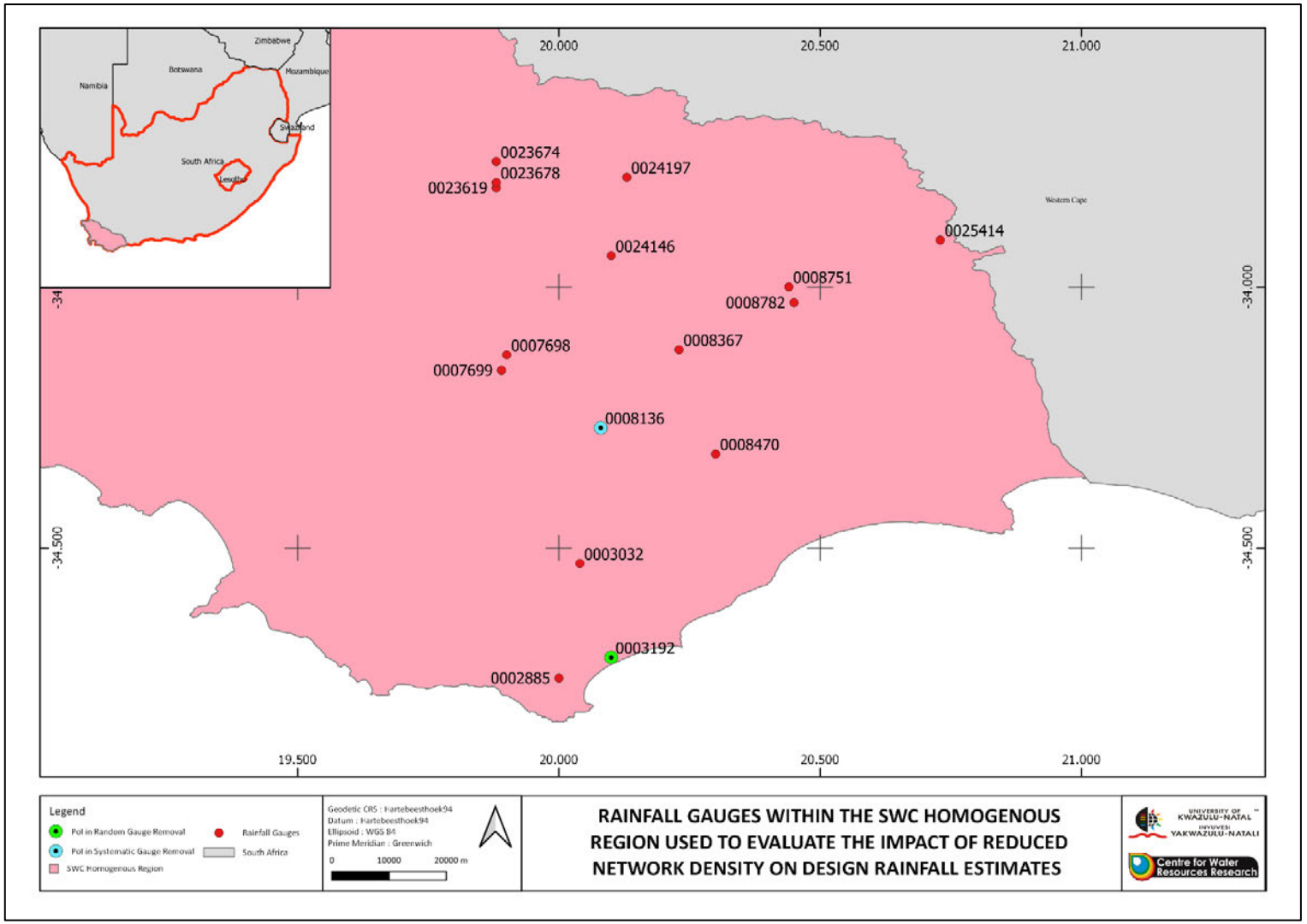

Figure D1.3 Rainfall gauges within the SWC homogenous region used to evaluate the impact of reduced network density on design rainfall estimates

Table D1.2 Streamflow gauges and associated attributes used in the analyse of a reduced gauge network on design flood estimation

Gauge ID	Gericke (2015) Climate Zone	HRU (1972) Zone	Kovacs (1988) K-Region	Latitude (^{0})	longitude (^{0})	Start date	End date	Record length (Years)	Actual Record Length (Years)	Area (km^2)
V1H009	ESC	9	5	-28.89	29.77	1954/01/15	2013/11/04	61	44	196.69
V1H010	ESC	9	5	-28.82	29.55	1964/11/26	2014/01/06	51	36	786.65
V1H038	ESC	9	5	-28.56	29.75	1971/10/19	2013/11/05	43	29	1660.1
V2H004	ESC	9	5	-29.07	30.25	1960/05/01	2013/11/14	54	39	1555.77
V3H007	ESC	9	5	-29.24	29.79	1972/07/28	2013/11/12	42	45	115.33
V6H004	ESC	9	5	-28.4	30.01	1954/01/01	2014/01/08	61	41	663.92
V7H012	ESC	9	5	-29.01	29.88	1962/11/17	2013/11/12	52	30	199.89
V7H016	ESC	9	5	-29.19	29.63	1972/10/23	2013/11/07	43	28	122.12
V7H017	ESC	9	5	-29.19	29.64	1972/10/23	2013/11/12	42	28	282.24
A2H012	NI	8	5	-25.81	27.91	1922/10/01	2013/12/11	92	57	2579.65
A2H013	NI	8	5	-25.78	27.76	1922/10/01	2013/12/11	93	58	1164.43
A2H023	NI	8	5	-25.95	27.96	1965/10/23	2013/12/09	50	32	689.85
A2H044	NI	8	5	-25.9	27.93	1971/07/18	2013/12/11	43	37	764.04
A2H045	NI	8	5	-25.89	27.91	1972/05/25	2013/12/11	42	29	663.72
A2H049	NI	8	5	-25.98	27.84	1972/07/04	2013/12/09	43	34	372.83

Gauge ID	Gericke (2015) Climate Zone	HRU (1972) Zone	Kovacs (1988) K-Region	Latitude (^{0})	longitude (^{0})	Start date	End date	Record length (Years)	Actual Record Length (Years)	Area (km^2)
A2H050	NI	8	5	-25.99	27.84	1973/04/06	2013/12/09	41	34	152.63
A2H063	NI	8	5	-25.7	28.19	1984/05/10	2013/12/10	30	24	33.28
G1H010	SWC	2	5	-33.39	19.16	1964/05/05	2013/11/07	50	42	10.4
G1H011	SWC	2	5	-33.38	19.15	1964/04/29	2013/11/07	50	38	18.29
G1H012	SWC	2	5	-33.35	19.1	1964/04/20	1996/06/04	33	21	34.43
G1H015	SWC	2	5	-33.82	19.06	1964/06/06	1988/07/18	25	19	1.8
G1H017	SWC	2	5	-33.83	19.03	1964/06/06	1988/07/19	25	20	1.76
G1H018	SWC	2	5	-33.82	19.05	1964/06/06	2013/08/27	50	25	3.49
G1H029	SWC	2	5	-33.16	19.05	1972/11/30	2013/11/07	42	29	35.66
G1H040	SWC	2	5	-33.36	18.96	1979/08/16	2013/11/06	35	32	37.62
H1H013	SWC	2	5	-33.36	19.3	1965/02/24	2013/11/04	49	29	62.63
H2H005	SWC	2	5	-33.46	19.62	1969/09/26	2013/11/12	45	38	14.81
H2H008	SWC	2	5	-33.33	19.64	1982/06/29	2013/11/04	32	25	9.69
J1H016	SWC	2	5	-33.29	19.73	1974/06/24	2013/11/04	40	26	30.88

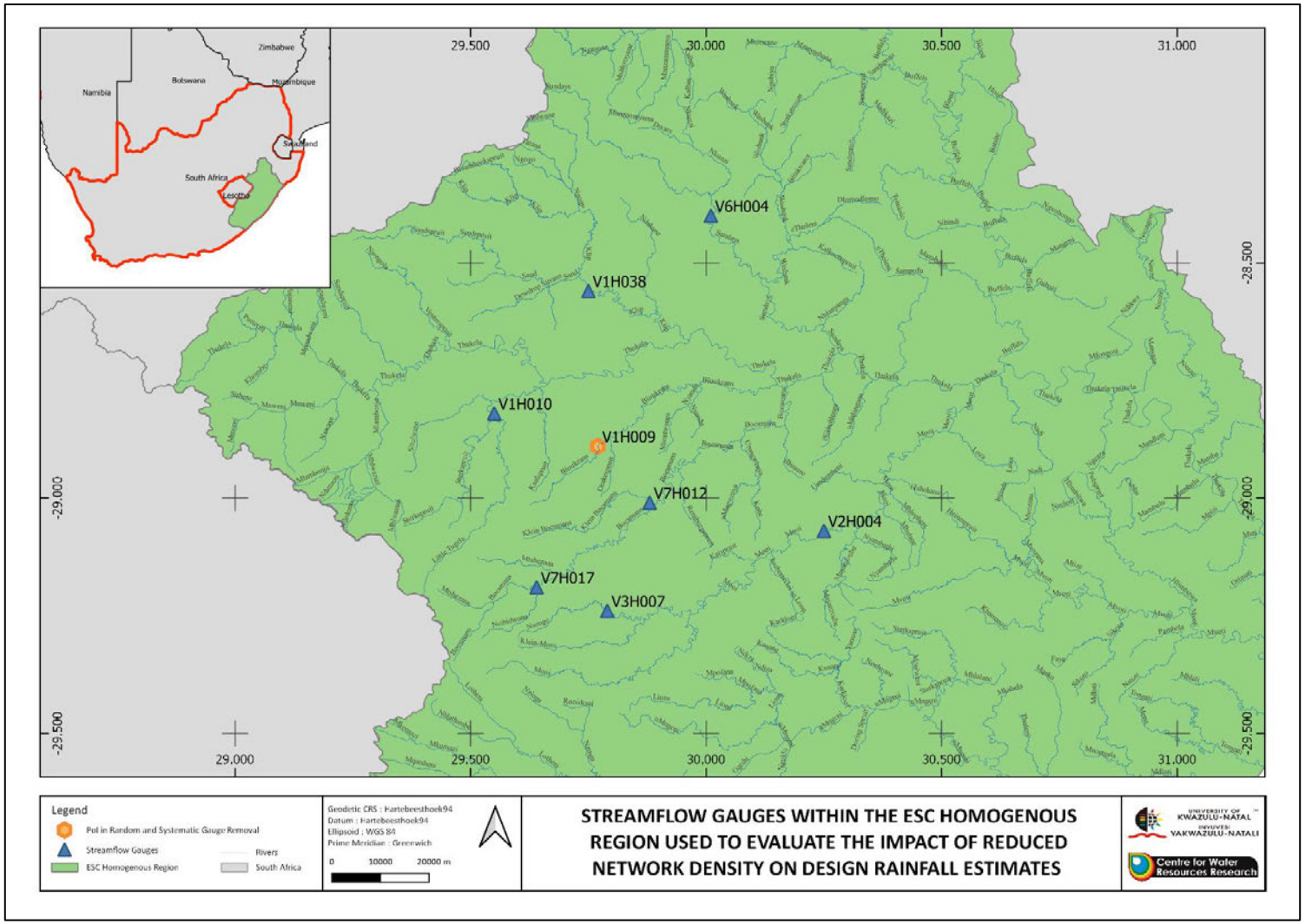

Figure D1.4 Selected streamflow gauges within the ESC homogenous region used to evaluate the impact of reduced network density on design flood estimates

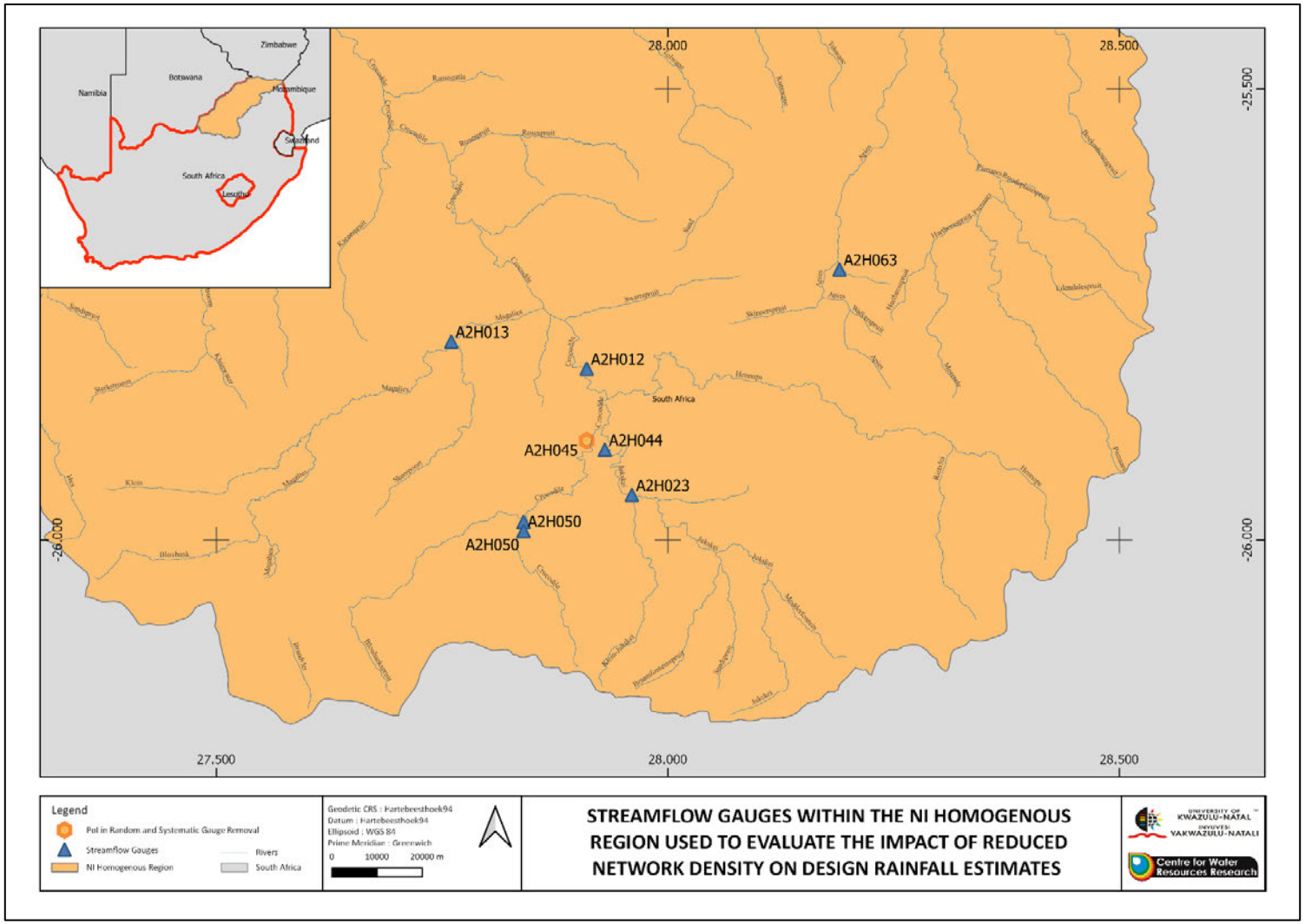

Figure D1.5 Selected streamflow gauges within the NI homogenous region used to evaluate the impact of reduced network density on design flood estimates

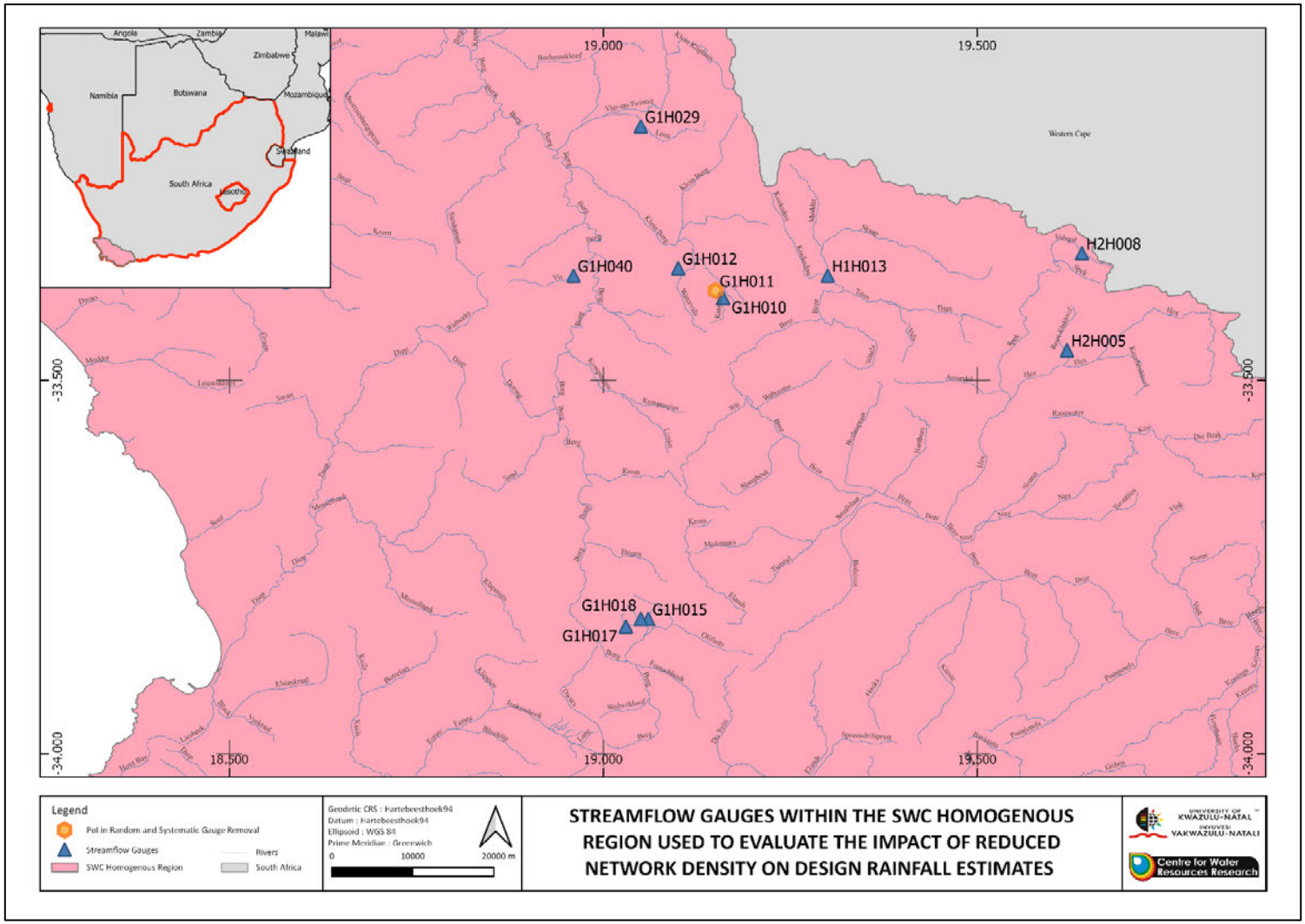

Figure D1.6 Selected streamflow gauges within the SWC homogenous region used to evaluate the impact of reduced network density on design flood estimates

13.2 Appendix D2: Impact of Reduced Network Density on Design Flood Estimation Using Systematically Generated Data Series

Figure D2.1 provides results *of MRD, MARD, PBIAS* and *NSE* for design flood events estimated using scenarios (*Closest.75%, Closest.50%, Furthest.75%* and *Furthest.50%)* of a systematically reduced gauge network.

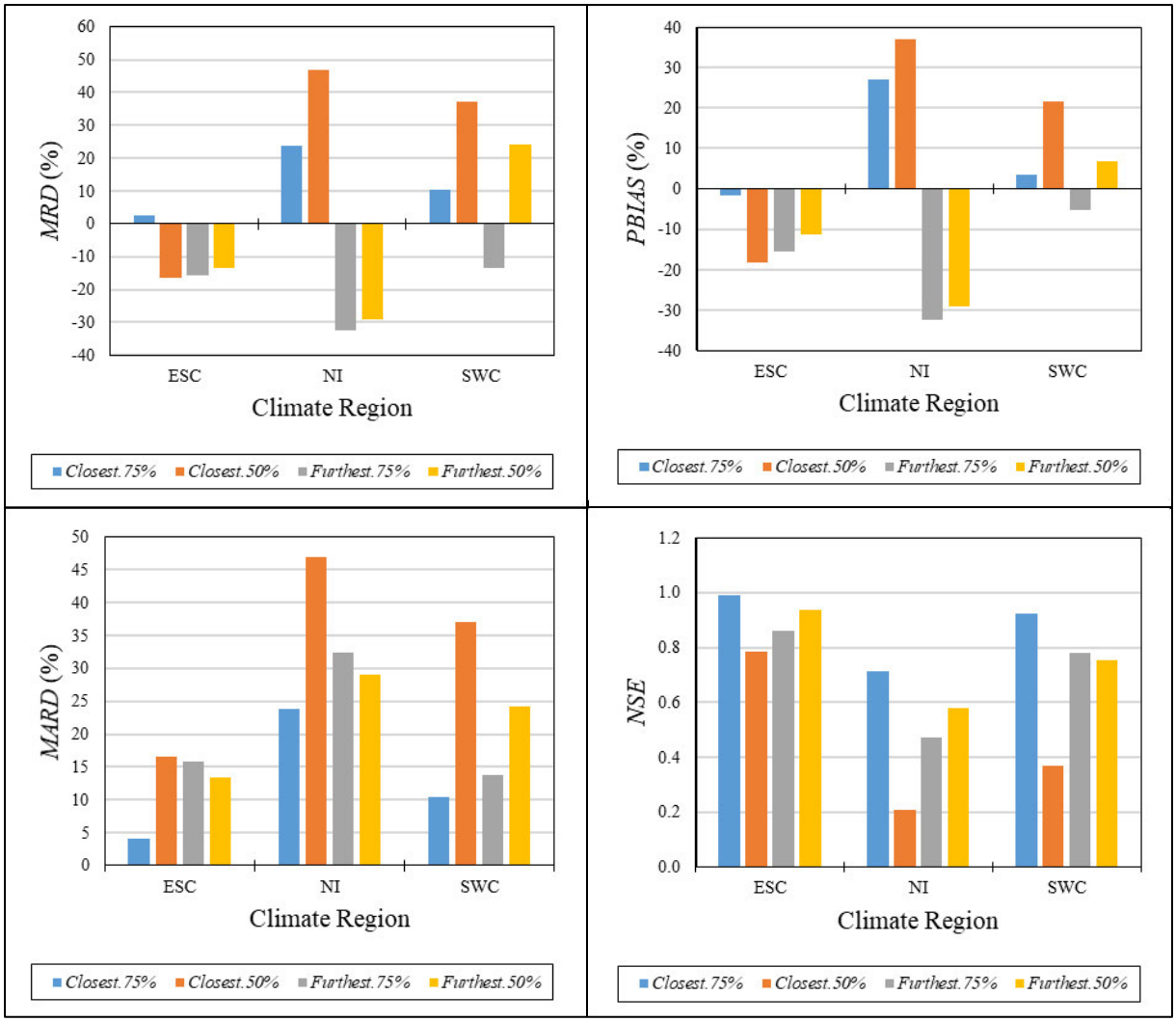

Figure D2.1 *MRD, PBIAS, MARD and NSE* for design flood events estimated using scenarios (*Closest.75%, Closest.50%, Furthest.75% and Furthest.50%*) of a systematically reduced gauge network