Stellenbosch
UNIVERSITY
IYUNIVESITHI
UNIVERSITEIT

Improved Statistical Flood Frequency Estimation Approach for South Africa

by

Danie van der Spuy

Dissertation presented for the degree of Doctor of Philosophy
in the Faculty of Engineering
at Stellenbosch University

Supervisor: Prof. Jakobus Andries du Plessis

December 2023

Declaration

By submitting this dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third-party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: December 2023

Abstract

The performance of the most frequently used flood frequency probability distributions in South Africa, using the current statistical approach, were reviewed and all tend to perform poorly when lower probability frequency events are estimated, especially where outliers are present in the dataset. This can, essentially, be attributed to the challenge to analyse very limited 'samples' of annual flood peak populations, which are an unknown. As a result, outliers were inadequately 'managed' by attempting to 'normalise' the flood peak dataset, which concealed the significance of the observed data. Thus, by considering the outliers properly, the research was undertaken with the aim to improve the current statistical approach and to develop a more stable and consistent probability distribution. The re-evaluation of the plotting position was considered as a precondition to the research.

The plotting position concept was initially applied more than a century ago. Since then, many alternative plotting position techniques have been developed that, practically, do not differ much from each other. The ineffective management of outliers is regarded as the main reason for the inconsistency in flood frequency results. Exploiting a more pragmatic approach, managing outliers by using Z-scores, an alternative plotting position technique was developed, referred to as the 'Z-set' plotting position. The main advantages of the Z-set plotting position technique, distinguishing it from the other plotting position techniques, is its consistency, the more sensible inclusion of outliers and a more homogeneous appearance. The Z-set plotting position technique also proved to be less susceptible to different record lengths than the existing plotting position techniques, which may encourage more consistent choices of appropriate probability distributions.

The methodology followed in the development of the proposed new probability distribution, named IPZA, might be considered as unconventional, but the multiple regression approach was used to accommodate the strongly skewed data, often associated with annual flood peak series. The main advantages of the IPZA distribution, is its consistency, the simplicity of application (only one set of frequency factors for every parameter, regardless of the skewness), the integrated handling of outliers and the use of conventional method of moments, eliminating the need to adjust any moments.

The performance of the Z-set plotting position and the IPZA probability distribution exceeded initial expectations. The results of both are more consistent than and, by taking outliers into account, appear to be more sensible than existing plotting positions and probability distributions.

It is concluded that the Z-set plotting technique and the IPZA probability distribution should be used as valuable additions to the existing set of decision-making tools for hydrologists/engineers performing flood frequency analyses. Evidently, this does not exclude the use of existing probability distributions. It is sound practice to use more than one probability distribution to assist in making a more informed scientific decision.

Samevatting

Die meer algemene waarskynlikheidsverdelings, wat tans in Suid-Afrika in die statistiese vloedfrekwensie benadering gebruik word, is geëvalueer en daar is gevind dat dit nie goed vaar in die beraming van laer waarskynlikheids-vloedpieke nie, veral waar uitskieters in die datastel voorkom. Dit kan toegeskryf word aan die uitdaging om vloedfrekwensie-ontledings te doen met 'n steekproef van die populasie van jaarlikse vloedpieke, wat 'n onbekende is. Gevolglik word uitskieters nie effektief hanteer nie, deurdat daar gepoog word om die datastel te normaliseer wat meebring dat die ware betekenis van die data verberg word. Derhalwe, met behoorlike inagneming van uitskieters, is die navorsing onderneem met die doel om die huidige statistiese benadering te verbeter en om 'n meer stabiele en konsekwente waarskynlikheidsverdeling te ontwikkel. Die her-evaluering van die plotposisie-tegniek was gestel as 'n voorvereiste tot die navorsing.

Die plotposisie-tegniek is aanvanklik meer as 'n eeu gelede ontwikkel. Sedertdien is verskeie alternatiewe plotposisies ontwikkel wat, uit 'n praktiese oogpunt, nie veel van mekaar verskil nie. Oneffektiewe hantering van uitskieters word beskou as die hoofrede vir die inkonsekwente resultate in vloed waarskynlikheids-ontledings. 'n Meer praktiese benadering is ondersoek, waar uitskieters hanteer is met behulp van Z-waardes, wat gelei het tot die ontwikkeling van 'n alternatiewe plotposisie-tegniek, bekend as die 'Z-set' plotposisie. Die ooglopende voordele van die Z-set plotposisie-tegniek, bo die ander plotposisie-tegnieke, is groter konsekwentheid, meer sinvolle hantering van uitskieters en 'n meer homogene voorkoms. Die Z-set plotposisie-tegniek het ook getoon dat dit meer bestendig reageer, as die bestaande plotposisies, op verskillend rekordlengtes, wat meer konsekwente keuses van waarskynlikheidsverdelings kan meebring.

Die metode wat gebruik was in die ontwikkeling van die voorgestelde waarskynlikheidsverdeling, genaamd 'IPZA', mag beskou word as onkonvensioneel, maar die meervoudige regressie-ontleding model is gebruik om uiters skeef verdeelde data, wat dikwels in jaarlikse vloedpiekdata voorkom, te akkommodeer. Die konsekwentheid, eenvoud van toepassing (slegs een stel frekwensie-faktore vir elke parameter, ongeag die skeefheid), die ingeslote hantering van uitskieters en die gebruik van die konvensionele metode van momente, wat die transformasie van momente onnodig maak, is die grootste voordele van die IPZA verdeling.

Die Z-set plotposisie en die IPZA waarskynlikheidsverdeling het alle aanvanklike verwagtinge oorskry. Beide se resultate is aansienlik meer konsekwent en, deur die uitskieters in ag te neem, blyk dit ook meer sinvol te wees as dié van bestaande plotposisies en waarskynlikheidsverdelings.

Die IPZA waarskynlikheidsverdeling, asook die Z-set plotposisie-tegniek, behoort gebruik te word in vloedfrekwensie-ontledings deur hidroloë/-ingenieurs, as waardevolle toevoegings tot bestaande besluitnemings-hulpmiddels. Uiteraard sluit dit nie ander verdelings uit nie. Dit is steeds gesonde praktyk om meer as een waarskynlikheidsverdeling te gebruik om 'n beter ingeligte wetenskaplike besluit te kan neem.

Acknowledgements

In addition to my sincere recognition of Prof JA du Plessis, for his guidance, support, supervision, encouragement and patience throughout the research, I wish to express my gratitude to:

- My ex-colleagues, from the Flood Studies component of the Department of Water Affairs and Sanitation, Archie Thobejane, Charles Linström, Ernest Oakes, Dumisani Shezi, Jeremy Naidoo Jermaine Nathanael, Tankiso Skosana and, especially, Pieter Rademeyer who was my indispensable right hand at Flood Studies. The enormous amount of flood frequency analyses that we performed, deliberated- and (sometimes) heavily debated on, led to an enhanced insight into the challenges associated with flood frequency analyses. We learnt from one another, and this is an outcome (hopefully not the last) from the knowledge and experience we all gained. I also highly value your mental support throughout this research;
- Other ex-colleagues/friends that expressed words of encouragement, often just at the right time;
- My children and my extended family for their encouragement throughout this research;
- My wife, Adri, for her patience (sometimes, understandably, a tad impatience), understanding and support during the lengthy period of this research.

Gratitude, above all, to the grace of my Heavenly Father in allowing me to give back to others the knowledge He allowed me to gain.

Table of Contents

List of Figures

List of Tables

List of acronyms, initialisms, abbreviations and symbols

$\varphi(z)$	-	Probability Density Function of the standard normal distribution
$\Phi(z)$	-	Cumulative Distribution Function of the standard normal distribution
7ym	-	7-year-multiple(s)
AEP(s)	-	Annual Exceedance Probability(ies)
AFP	-	Annual Flood Peak
AMS	-	Annual Maximum Series (of flood peak data)
ARI(s)	-	Annual Recurrence Interval(s)
ARR	-	Australian Rainfall and Runoff
AVE, $\overline{X}$	-	Average
CA	-	Catchment Area
CDF	-	Cumulative distribution function
COST	-	European Cooperation in Science and Technology
COV	-	Coefficient of variation
DWA	-	Department of Water Affairs
DWS	-	Department of Water and Sanitation
EV1	-	Gumbel (or Extreme Value 1) distribution
EV4	-	Four Parameter Extreme Value distribution
EXP	-	Exponential distribution
F-value	-	The probability of a statistical significance between model and the dependent variable
FEH	-	Flood Estimation Handbook (United Kingdom)
FFA(s)	-	Flood Frequency Analysis(es)
F(Z)	-	standard normal cumulative distribution function
g	-	Skewness statistic
GEV	-	Generalised Extreme Value distribution
GEV_{MM}	-	Generalised Extreme Value distribution (using method of moments)
GEV_{PWM}	-	Generalised Extreme Value distribution (using probability weighted moments)

GLO	-	Generalised Logistic distribution
GPA	-	Generalised Pareto distribution
H0	-	Null Hypothesis
$H0_F$	-	Null Hypothesis that no statistical significance exists between model and the dependent variable
$H0_P$	-	Null Hypothesis that no statistical significance exists between an independent variable and the dependent variable
HRU	-	Hydrological Research Unit (of the University of the Witwatersrand)
HYDSTRA	-	DWS hydrological database
IPZA	-	Improved Probability-analysis, South Africa
KAP3	-	3-Parameter Kappa distribution
$K_{P\text{-}SD}$	-	Frequency factor for standard deviation
$K_{P\text{-}SD^*}$	-	Frequency factor for standard deviation, omitting largest flood peak from dataset
$K_{P\text{-}Q}$	-	Frequency factor for the average flood peak
kurt	-	kurtosis statistic
LM	-	L-moment
LN	-	Log-Normal distribution
LN2	-	Two-parameter Log-Normal distribution
LN3	-	Three-parameter Log-Normal distribution
LN4	-	Four-parameter Log-Normal distribution
LP3	-	Log-Pearson Type III distribution
MAE	-	Mean Absolute Error
MAPE	-	Mean Absolute Percentage Error
MGBT	-	Multiple Grubbs-Beck Test
MLVA	-	Mean Logarithm Value Approach
MM	-	Method of Moments
MOZ	-	Mozambique
MRA	-	Multiple Regression Analysis
NERC	-	National Environment Research Council (re Flood Studies Report)

n	-	Sample size
P-value	-	The probability of a statistical significance between an independent variable and the dependent variable
PE	-	Probability of exceedance
P_i	-	Plotting probability of the i^{th} order statistic
P_P	-	Probability of exceedance corresponding to Peak over Threshold approach
PDS	-	Partial Duration Series (of flood peak data)
PE3	-	Pearson type III distribution
PDF(s)	-	Probability Density Function(s)
PILFs	-	Potentially Influential Low Flows
PMF	-	Probable Maximum Flood
POT	-	Peaks over Threshold approach
POTλ	-	Peaks over Threshold approach, producing on average λ flood peak(s) per year
PP(s)	-	Plotting Position(s)
Q_{AEP}	-	Flood peak associated with an Annual Exceedance Probability
Q_{ave}	-	Average flood peak of an Annual Maximum Series
Q_{est}	-	Estimated flood peak
Q_{gmn}	-	Geometric mean flood peak of an Annual Maximum Series
Q_{max}	-	Maximum flood peak in an Annual Maximum Series
Q_{mdn}	-	Median flood peak in an Annual Maximum Series
Q_{min}	-	Minimum flood peak in an Annual Maximum Series
Q_{obs}	-	Observed flood peak
PWM	-	Probability Weighted Moment
RFFA	-	Regional Flood Frequency Analysis
RL1	-	Record length result from first approach applied
$RL1_P$	-	Proposed minimum length from first approach applied
RL2	-	Record length result from second approach applied
$RL2_P$	-	Proposed minimum record length from second approach applied

RL_R	-	Recommended record length
RMF	-	Regional Maximum Flood
R-W	-	Rolling Window
R-W*m*	-	*m* year (record length) Rolling Window sub-dataset
SA	-	South Africa
SANRAL	-	South African National Road Agency
SCS	-	Soil Conservation Service
SD	-	Standard deviation of an Annual Maximum Series
SD*	-	Standard deviation of dataset with highest value omitted
SD**	-	Standard deviation of dataset with 2 highest values omitted
SD*^o	-	Standard deviation of dataset with lowest and highest values omitted
SDF	-	Standard Design Flood
SFFA	-	Site-Specific (At-Site) Flood Frequency Analysis
T	-	Return period – also Recurrence interval (alternative for annual recurrence interval in figures/tables)
TCEV	-	Two-Component Extreme Value distribution
TEV	-	Transformed Extreme Value distribution
UK	-	United Kingdom
USA	-	United States of America
USACE	-	U.S. Army Corps of Engineers
WMO	-	World Meteorological Organisation
W_P	-	Standardised variate linked to the Gumbel distribution
W_T	-	Standardised variate linked to the log-normal distribution
$\overline{X}$	-	sample mean
Z-score	-	distance from the mean in terms of standard deviation multiple

1 Introduction

This chapter commences with a brief background to the research. It is followed by the problem definition and subsequent objectives of this research. The chapter concludes with a brief overview of the dissertation structure.

1.1 Background

Floods rank among the topmost devasting natural disasters. The 1931 Yangtze (Yangzi) River floods, in Southern China, is indisputably still considered the worst natural disaster of the 20^{th} century; the flood inundated 180 000 km^2; and an estimated 3.7 million people lost their lives (Courtney, 2018) – be aware that figures, relating to damages and deaths, are estimates and may vary from reference to reference. In South Africa (SA) the floods of 1981, 1984, 1987, 1988 and more recently 2000 immediately come to mind. The floods in the Buffels River in 1981, which inundated the town of Laingsburg and caused 90 people to lose their lives and caused damage in excess of R10 million (1981 values), was caused by what is popularly referred to as a Black South-Easter weather situation (Roberts and Alexander, 1982). Although Domoina in 1984 is commonly referred to as a tropical cyclone, it only officially reached the status of a severe tropical storm. Domoina caused widespread flooding in Northern KwaZulu-Natal (KZN) in SA, Mozambique (MOZ) and eSwatini (ESW), causing widespread damage to road infrastructure and property of more than R100 million (1984 values) and over 200 reported deaths, in South Africa alone (Kovács, Du Plessis, Bracher, Dunn, and Mallory, 1985). The main cause for the 1987 floods in the southern part of KZN was an intense off-shore cut-off low pressure system. The damage was calculated to have been amongst the most devasting that occurred in SA. Damage to property, agriculture (that included several hundred farm dams that breached), communications and infrastructure amounted to about R400 million (1987 values), with 388 reported deaths and 150 000 people left homeless (Van Bladeren and Burger, 1989). During two periods in February 2000 tropical weather systems, which included the tropical cyclone Eline making landfall at the MOZ coast in the Beira area, resulted in extreme rainfall leading to devastating floods. The excessive financial impact (approximately R1.5 billion (2000 values) to infrastructure and water services in SA), as well as social impact (at least 600 people in MOZ lost their lives and hundreds of thousands of people were displaced) was most probably the worst experienced in living memory in Southern Africa (Dyson and Van Heerden, 2001).

These extreme events occur in relatively short annual maximum series (AMS) flood peak record lengths, which engineers and hydrologists have to use to perform a flood frequency analysis (FFA) with. These analyses are used to determine design floods for dams and bridges, for example, as well as to determine flood lines along rivers, optimal time frames for construction, inundation of vulnerable crops, etc.

In flood hydrology, the benefit of having a proper population of an AMS of flood peak records, to test the AMS sample records against, does not exist. From these short AMS samples, hydrologists have to estimate the underlying probability distribution of the AMS population. It is thus of extreme importance to constantly strive to improve the tools, available to hydrologists, to continuously increase the confidence in FFA.

1.1.1 History of flood model development in South Africa

Different groupings of FFA approaches, which associate a probability of exceedance, typically expressed as Annual Exceedance Probability (AEP), with a flood peak, are reported in literature. AEP can also be expressed in terms of return period (T), or annual recurrence interval (ARI).

In SA three classifications, commonly referred to as empirical-, deterministic- and statistical methods (Alexander, 1990; Pegram and Parak, 2004; Smithers, 2012; Van der Spuy and Rademeyer, 2018), are used to estimate the AEP of peak flows in an FFA.

All these FFA approaches associate flood peak magnitudes with annual exceedance probabilities (AEPs). The main difference lies in the way these flood peaks and associated AEPs are obtained:

- In statistical FFA methods flood peak frequencies are estimated through statistical analyses of observed annual flood peak data. Statistical methods refer to that part of statistics, indicated as inferential statistics in which probability theory is applied to draw conclusions from data.
- In deterministic FFA methods, storm rainfall is used in conceptual rainfall-runoff models. Flood peak frequencies are estimated by assuming that the statistical properties of a flood are analogous to that of the storm rainfall causing it; and
- In empirical FFA methods (based, per definition, on observation or experience rather than theory or pure logic) graphical or numerical relationships are applied to translate catchment characteristics into flood peak frequencies.

The methods under the last two approaches are developed by applying statistical methods to observed flood peak data.

SA's climatic conditions essentially range from Mediterranean in the southwestern corner of SA, changing to semi-arid and arid regions, to temperate in the interior plateau, to subtropical in the northeast. This very wide range of climatic conditions, consequently, causes very different hydrological responses, ranging from high flow to virtually no-flow conditions.

The most comprehensive study on Flood Hydrology in SA, up to date, was done by the Hydrological Research Unit of the University of the Witwatersrand (HRU), at the request of the South African Institution of Civil Engineers. Their first report, HRU 4/69 (Midgley, Pullen and Pitman, 1969) was published in 1969. Some of the methods proposed in this report, as well as in some of the follow-up reports (Pitman and Midgley, 1971; Midgley, 1972 and Bauer and Midgley, 1974), were developed a few years before the first report was published – for example an empirical method developed by Pitman and Midgley (1967). The methodologies that evolved from these studies belong to the deterministic-, as well as the empirical approaches and were developed primarily to be used in the large ungauged parts of SA. The deterministic and empirical methods proposed in SANRAL's Drainage Manual (SANRAL, 2013), as well as in Alexander's *Flood Hydrology for Southern Africa* (1990) handbook, are mostly a repeat of the methodologies developed by the HRU during the late 1960s.

In addition to the approaches, presented in *Flood Hydrology for Southern Africa* (Alexander, 1990), Alexander (2002 and 2003) introduced a new empirical method called the Standard Design Flood (SDF). The method generally overestimates flood peak frequencies and, in several cases, even underestimates (Gericke and Du Plessis, 2012), causing some mistrust in the methodology. The latest edition of the *Drainage Manual* (SANRAL, 2013) also included the deterministic Soil Conservation Service (SCS) technique for small catchments, adapted for South African conditions (SCS-SA) by Schmidt and Schulze (1987).

Both Alexander (1990) and SANRAL (2013) include the FFA for the most commonly used probability distributions in SA, namely the Log-Normal- (LN), Log-Pearson Type III- (LP3) and the General Extreme Value (GEV) distributions.

1.1.2 Current situation in South Africa

The need for developing regional methods, for the estimation of design flood characteristics in ungauged basins, was highlighted by Mimikou, Niadas, Hadjissava and Kouvopoulos (1993). Smithers (2012) also remarked that the HRU studies (noted in 1.1.1 above) indicated that the most frequent need for design flood estimation exist for small catchments (smaller than 15 km^2). For practical and monetary reasons, most of these catchments are ungauged and observed flow and flood data do not exist and, consequently, deterministic and empirical methods must be applied.

The deterministic- and empirical methodologies that evolved from the HRU studies are still currently being used in SA. The results produced by these approaches are not consistent when compared to statistical analyses of AMS flood peaks (Naidoo, 2020). The need, to update most of these methods (Van der Spuy, Rademeyer and Linström, 2004; Smithers, Görgens, Gericke, Jonker and Roberts, 2014), is long overdue and critical to enable an impartial assessment of the results obtained from applying the various FFA methods used in SA. Currently, more than 50 additional years of data are available that can be utilised to improve the methods developed by the HRU.

The South African hydrological science and engineering community, under the umbrella of the National Flood Studies Programme (NFSP), is currently engaged in updating/replacing, primarily, appropriate deterministic and empirical flood frequency estimation methods in SA (Smithers, *et al.*, 2014). Numerous statistical flood frequency analyses will be carried out in this research, on an estimated 300 to 500 AMS of flood peaks from various flow gauging sites. To realistically accomplish this, a sound statistical FFA will be essential.

1.2 Problem Statement

Observed streamflow signifies the catchment-response to a rainfall event. Consequently, a statistical analysis of streamflow data is still considered to be the most accurate 'modelling' of catchment response.

The South African flood hydrology community, unfortunately, mostly still avoid the statistical approach. Van Vuuren, Van Dijk and Coetzee (2012) found that only 17% of all practitioners will consider the use of statistical analyses and that only 23% of those will consider the GEV as one of the possible probability distributions. This was confirmed by Du Plessis (2014), with figures of 11% and 22%, respectively. It is thus evident that deterministic and empirical models, rather than a statistical analysis on available AMS flood peak data, are still the preferred approach of most practitioners to determine flood frequencies. The main reason for this appears to be a lack of confidence in the statistical methodologies. The deterministic and empirical methodologies are also perceived to be 'easier' to apply. This perception, unfortunately, mainly stems from the avoidance of the statistical approach by practitioners, which might explain their lack in confidence to deal properly with the uncertainties attached to the statistical approach. These uncertainties are addressed in Chapter 2 and mainly include, but are not limited to, issues like (1) appropriate record length (Hattingh, Mostert, Muir and Basson, 2011; Van der Spuy, 2018); (2) plotting positions with special emphasis on plotting of 'outliers' (United States Army Corps of Engineers; USACE, 1993 and World Meteorological Organisation; WMO, 2009); (3) applicability of current probability distributions and/or combinations thereof (NERC, 1975; Vogel, McMahon and Chiew, 1993); (4) single site or regional analysis (Alexander, 1990; Haile, 2011; Smithers, 2012).

The status quo, of avoiding the use of statistical methods, will remain if a clear and consistent statistical approach is not formulated. It is also evident that, to develop proper deterministic- and empirical methodologies during the NFSP, a need exists for an improved consistent statistical approach. The literature review (Chapter 2) confirmed that, in general, the existing statistical methodologies, in analysing streamflow data, appears to be ineffective in consistently predicting reliable lower AEPs of AMS flood peak data (Van der Spuy, 2018).

Van der Spuy (2018), supported by DWS (1993-2021), suggested that the inconsistent way in which outliers are dealt with is the source of most concerns with statistical FFAs. Subsequently, the inherent differences in the various probability distributions are intensified by not acknowledging outliers for what they are, by minimising their impact through approaches attempting to 'normalise' datasets. Normalising of datasets is accomplished, for example, by using the logarithms of the flood peak data or/and by applying adapted moments like probability weighted moments (PWM) or L-moments (LM). In the analysis of data, the nature of the data should determine the statistic (Wheeler 2022b) and not the other way round. By tampering with the data and/or the moments (statistics) the context of the original data is lost.

Instead of repeatedly trying to address the symptoms, the root cause of the problems, namely outliers, needs to be addressed. If outliers are addressed more sensibly it is foreseen that the plotting position (PP) and, consequently, the probability distribution will have to be addressed too. If the root cause of the concerns will not be addressed people will continue to "*do wrong things more precisely*" (Galloway, 2010).

1.3 Objectives considered for an Improved Statistical Approach

The inclusive objective of the research was the development of an improved statistical approach that any practitioner will be able to use with confidence to determine more accurate and sensible AEPs (or ARIs) from observed AMS flood peak data.

The inclusive objective comprises of a primary objective, supported by specific secondary objectives, associated with statistical analyses of streamflow data. To effectively address the primary objective, namely an improved probability distribution, a few secondary objectives were considered and addressed.

1.3.1 Single site or regional analysis

Single site/site-specific/at-Site Flood Frequency Analysis (SFFA), as opposed to a regional Flood Frequency Analysis (RFFA) in the statistical analysis of streamflow data, is a much-debated topic. RFFA is an essential part of developing deterministic- and empirical methods and should be investigated appropriately, in future. However, it was not deemed essential in the context of this research, which concentrated on SFFA and, thus, RFFA was not investigated further.

1.3.2 Plotting positions

Plotting positions (PPs) serve as visual guides to assist in choosing the most appropriate probability distribution. Currently outliers are not considered appropriately, in determining PPs (to be highlighted in Chapter 2, Literature review, Sections 2.2.2, 2.3.5 and 2.5.1).

Furthermore, cases where AMS flood peak values of the same order of magnitude might occur several times, in a specific record (observed mostly in drier parts), may cause the visual appearance of a plotting position (PP) to be confusing, rather than useful.

Hence, a more suitable PP ought to be developed.

1.3.3 Record length

From Van der Spuy (2018), DWS (1993-2021) and Hattingh *et al*. (2011) it is evident that the GEV distribution tends to 'stabilise' in the estimation of the lower AEPs with record lengths exceeding 30-40 years. This observation provided a point of departure to thoroughly investigate record lengths that would become a criterion in the choice of flow sites, for the development of the improved probability distribution and approach.

1.3.4 An Improved Probability Distribution

In developing a new statistical FFA procedure, an improved probability distribution would be the ultimate objective. Since none of the existing probability distributions, used in South Africa, demonstrates a trend to support an upper limit (bound) for flood peak records, it would be a bonus if the envisioned new probability distribution would suggest that an upper bound do indeed exist.

1.3.5 Summary of objectives

The main objective of the research was to develop a new statistical approach, which will preferably include a proper bounded probability distribution and will be more consistent and reliable in its performance. It might also help to change the existing negative perceptions relating to statistical analyses of AMS flood peak data.

In the process, preceding the main objective, conventional views on certain aspects of flood hydrology, like PPs (specifically in relation to outliers) and acceptable record lengths, were also challenged, as secondary objectives.

It is thus hypothesised that:

- a minimum record length does exist, beyond which FFA results will not vary much;
- a new PP can be developed, which will consider outliers in a more sensible way;
- it is not necessary to normalise datasets in developing a new probability distribution and that the method of moments (MM) should thus be adequate and
- a new probability distribution can be developed, where the FFA results will not be affected by outliers, with an upper limit (bounded) for flood peaks.

This new approach will further contribute to the flood discipline in:

- providing an improved statistical FFA tool, with which deterministic and empirical methodologies can be improved/developed for ungauged catchments, with an increased level of confidence; and
- opening numerous opportunities for possible further research and studies.

It is anticipated that an improved PP technique might act as a 'guiding principle', in the research, in developing an improved probability distribution. The importance of the PP should never be underestimated. It remains the only tangible tool that enables the hydrologist/engineer to verify that the fitted probability distribution is consistent with the data.

1.4 Chapter overviews

In Chapter 1 (the introduction) the research problem was defined with the subsequent research objectives.

The first part of Chapter 2, the literature review, focused on various aspects of a flood frequency analysis that can influence the eventual choice of the most appropriate probability distribution. These aspects include, for instance, AMS versus Peaks over Threshold (POT), record lengths, PPs and outliers. The PP techniques were also critically reviewed to identify concerns and problems with the existing approaches, to aid the research of an improved PP technique. The second part of the literature review focused mainly on a critical appraisal of the performance of the three most widely used probability distributions in SA, especially in relation to their sensitivity to outliers and their likely performance with very long record lengths. The third part is a short review on concerns encountered with present FFA approaches. The chapter concluded by deliberating on the literature review.

Chapter 3 (data acquisition) concentrated on the selection of flow sites, according to certain set criteria, to provide suitable data for the relevant research actions. Some of the literature, from Chapter 2 (Literature study) were referenced in this chapter too, to justify decisions that were made.

Chapters 4 to 6 relate to the three objectives and in each one the relevant methodology, research results and discussion were addressed. In Chapters 4 and 5 references were also used to justify decisions/assumptions that were made. Where deemed necessary, additional references were used, not necessarily relevant to the literature study per se, to justify decisions/assumptions, as well as to substantiate certain results.

In Chapter 4 record lengths were researched to determine minimum required record lengths, to ensure that research results can be used with a higher degree of confidence. The research explored two alternative routes to find a possible solution. The results of both alternatives were taken into account to recommend a minimum record length. Since this research will not affect the final results and will only assist in selecting additional sites for the primary research, the outcome is summarised in the main text of the dissertation, but the bulk of the research is presented in Appendix 4.

In Chapter 5 a practical approach, rather than a pure theoretical one, was followed to research a more consistent PP technique. The PP technique results were also assessed against an existing PP technique to confirm if concerns were addressed satisfactorily and whether there was any significant improvement, compared to the existing PP techniques. The outcome of this research was considered important since it was intended to use the PP technique as a benchmark for the primary research.

In Chapter 6 the research focused on an improved FFA approach that concentrated on the development of a bounded probability distribution.

Finally, Chapter 7 summarises and reflects on the findings of the research and makes recommendations for possible further research.

2 Literature review

In this chapter the literature review is discussed. Firstly, specific aspects of an FFA that would influence the choice of a probability distribution to yield sensible results are discussed, followed by a review of the performance of the three commonly used probability distributions in SA. A short review on concerns encountered with present FFA approaches is presented before the chapter concludes with a reflection on existing PP techniques followed by general observations and conclusions from the literature review.

2.1 Introduction

Many studies, encountered in doing the literature review, yielded similar outcomes, due to similar research approaches and similar applications. Only the papers, considered to convey a specific issue most appropriately, are referenced.

Most studies are very limited concerning general application, since most of these studies concentrated in their research only on a specific site of interest. This is particularly true in most cases where aspects, like the comparison between the applications of different probability distributions, moments, etc. were investigated.

There are, however, a few publications that address the topic of flood hydrology from a much wider perspective, like the Australian Rainfall and Runoff (ARR) Flood Estimation Guide (Ball, Babister, Nathan, Weeks, Weinmann, Retallick and Testoni, 2019), NERC (1975) which, in the United Kingdom (UK), which was replaced by the new national guidelines (Institute of Hydrology, 1999) that is referred to as the Flood Estimation Handbook (FEH), Bulletin 17C (England, Cohn, Faber, Stedinger, Thomas, Veilleux, Kiang, and Mason, 2018) in the United States of America (USA), Alexander (1990), USACE (1994) and WMO (2009).

Lengthy debates in literature exists about various aspects of an FFA which primarily includes, but is not limited to, statistical moments, probability distributions and PPs.

This research will concentrate on the flood frequency challenge problem from a more practical point of view. The literature review was conducted accordingly.

2.2 Aspects of an FFA influencing the choice of a probability distribution

The most common aspects that can have an influence on FFA are considered below. It includes reflecting on: (1) the choice between the use of AMS or POT, (2) the transformation of data, (3) the RFFA is a viable option to SFFA, (4) an upper bound probability distribution to flood peak data, (5) record lengths and (6) the different plotting position techniques.

2.2.1 AMS, PDS or POT series

Three approaches can be applied to compile a flood peak dataset, from a hydrological flood peak record, namely:

- the AMS consisting of the maximum peak flow, from each hydrological year of record;
- the POT series consists of all well-defined flood peaks, from the total hydrological record, above a specified threshold value;

  Note: From various references (e.g. Mohssen, 2009; Guru, 2016; Karim, Hasan and Marvanek, 2017) the acronym POTλ, where λ is known as the arrival rate and is the average number of peaks per record year, can be used – e.g. if the 200 highest flood peaks are extracted from a 50-year record, it will be a POT4 series (on average 4 floods per record year)

- the PDS (Partial duration series) that contains the highest n flood peaks from a hydrological record length of n years (hence, the POT1 case).

Each approach attempts to represent the flood peak aspects of the entire series of flow hydrographs by a simple series of flood peak values. For ease of reference POT and PDS will simply be referred to as POT.

Mkhandi, Opere and Willems (2005) carried out a study in the Equatorial Lake Victoria sub-basin, Africa, to compare the AMS- and POT approaches. They concluded that fitted probability distributions to the AMS and POT extreme events have indicated that both gave similar or comparable predictions of flood magnitudes, at lower AEP frequencies (ARIs higher than 10 years). Karim *et al*. (2017) did a similar study in Australia and reached a similar conclusion, in stating that both methods produce comparable magnitudes for large floods, but extensive differences were observed for small floods in the one to five years average recurrence interval.

Bezak, Brilly and Šraj (2014) performed a study in Slovenia and concluded that even though the POT method is a little more complex than the AMS series method, the statistical and graphical results showed that the POT method presented better fits to data. Karim *et al*. (2017) indicates that choosing the 'right' threshold flood might require an iteration process and recognised that although some historical floods might be excluded in the AMS series, conversely, the POT series may produce some unrealistic floods depending on the selection of the threshold flood-value in the analysis. Mohssen (2009) also pointed out that the value chosen for the threshold is considered as one of the main concerns associated with POT series.

Mohssen (2009) highlighted another major concern, namely the transformation of exceedance probabilities, obtained in the POT domain, to the AMS domain. He remarked that the commonly used formula (Equation 2-1) assumes that the arrival rate follows a Poisson distribution, which might not be true for many flood events.

$$P_A = 1 - e^{(-\lambda P_P)}$$ **Equation 2-1**

Where

P_A – AEP, corresponding to AMS

P_P – Exceedance probability, corresponding to POT

λ – arrival rate = N/n (with n = record length, N = number of chosen flood peaks)

Consequently, Mohssen (2009) derived a formula (Equation 2-2) from the Binomial distribution for any value of λ, where $P_P < 1/\lambda$.

$$P_A = \lambda P_P (1 - P_P)^{\lambda - 1}$$ **Equation 2-2**

Although a study by Schlögl and Laaha (2017) concentrated on rainfall and temperature, in considering extreme weather exposure identification for road networks, some of their conclusions are also relevant to flood frequency analyses. For instance, they found that the AMS/GEV outperformed the POT/GPA for precipitation and temperature differences, while for temperature minima and maxima the POT/GPA appeared to be better suited. With regards to AEPs they commented that the POT/GPA approach tends to underestimate AEPs as compared to the AMS/GEV approach in the higher AEP range, whereas the opposite was true for the lower AEP range (extreme events). They attributed this behaviour to two factors, namely sampling uncertainty and threshold selection.

Comparing AMS and POT in a frequency analysis, Srikanthan (2014) remarked that most literature, supporting the use of the POT, only considered the variance of the root mean square error of the quantile estimates and not the bias. He concluded that, since the AMS gave the smallest bias in most cases of his research, the use of AMS in frequency analysis is preferred to POT.

2.2.2 Transformation of data and moments

Transformation of data and/or moments primarily consider the use of log-transformed data and/or the use of different moments, than MM, such as PWM, or related LM. The rationale behind the transformation of data seems to be an attempt to make data sets appear more normally distributed. In doing so important statistical indicators, like outliers are ignored and not properly addressed. Cunnane (1985) confirms that proper consideration of possible outliers is avoided by affirming that low weight is in effect given to the maximum sample value when parameters are estimated by maximum likelihood (ML) or PWMs.

Hosking, Wallis and Wood (1985) argued that the application of PWMs outperforms the other applications in many cases and will usually be the preferred approach. Vogel *et al*. (1993) used LM diagrams as a goodness-of-fit (GOF) evaluation and commented that although the GEV procedures seemed to perform well for all regions considered, the LM- diagrams did not always favour the GEV procedure. Gunasekara and Cunnane (1992), cited by Vogel *et al*. (1993), also confirmed the above findings. Van der Spuy (2018) reported that the GEV_{PWM} did not improve the results of the GEV_{MM} (GEV, using MM), as is generally expected and assumed.

Cunnane (1985) queried the necessity for transformation of data by questioning if log distributions are indeed necessary, and substantiated it by citing the conclusion of Landwehr, Matalas and Wallis (1978) that log-transformation of data has no advantage and might even be considered as counter-productive. Wheeler (2011a&b) published two very insightful articles on the skewness (g) and kurtosis (kurt) shape statistics and advised that there is no need to transform the data to change the shape of the histogram, when the data are not homogeneous, but that consideration should rather be given to the impact of a lack of homogeneity, in the context of the original observations – for instance, in FFA, including a 1 000-year flood peak in a 50-year AMS record will indicate a lack of homogeneity and, in context of the relative short record length, it will have an impact on the estimated lower AEP flood peaks.

This above sound practical advice, unfortunately, is not generally applied, which leads to the subconscious ignoring of the fact that lower AEP flood peaks may occur in a relative short record. This results in practitioners still using various practices to transform observed data into 'more acceptable' data sets that will fit one of the currently used probability distributions (unbounded).

These views will be considered very carefully in this research, because of its direct relevance to outliers.

2.2.3 Single site vs regional analysis

FFA is used for establishing a relationship between flood magnitude and frequency of occurrence (return period) and for estimating the design-flood at a given location of interest (Castellarin *et al*., 2012). The approach can be implemented locally (SFFA) or regionally (RFFA). Support for the regional approach is expressed by Smithers (2012), who claims that the advantages of a regional approach for frequency analyses are evident from many studies (e.g., Potter, 1987; Stedinger, Vogel and Foufoula-Georgiou, 1993; Hosking and Wallis, 1997; Cordery and Pilgrim, 2000).

Haile (2011) intended to develop regional flood frequency curves, to be used in ungauged catchments, to improve the design and economic appraisal of civil engineering structures in Southern Africa. Haile (2011) identified nine regions, of which five were in SA, for his study. Haile (2011) considered seven theoretical probability distributions to evaluate which will be suitable to represent the average flood frequency distribution of regional data. Applying PWMs and LMs in evaluating the GPA, GEV, Gumbel (EV1), Exponential (EXP), LN3, Pearson type III (PE3) and GLO, it was concluded that the GPA, GEV, LN3 and PE3 emerged as underlying regional distributions and that none of the three remaining probability distributions should be considered for regionalisation in Southern Africa.

Due to a lack of sufficient streamflow-gauging sites, regional curves could only be developed and verified for five homogeneous regions, as identified by Haile (2011), of which four are in SA. The regional curves for one of these regions are depicted in Figure 2-1.

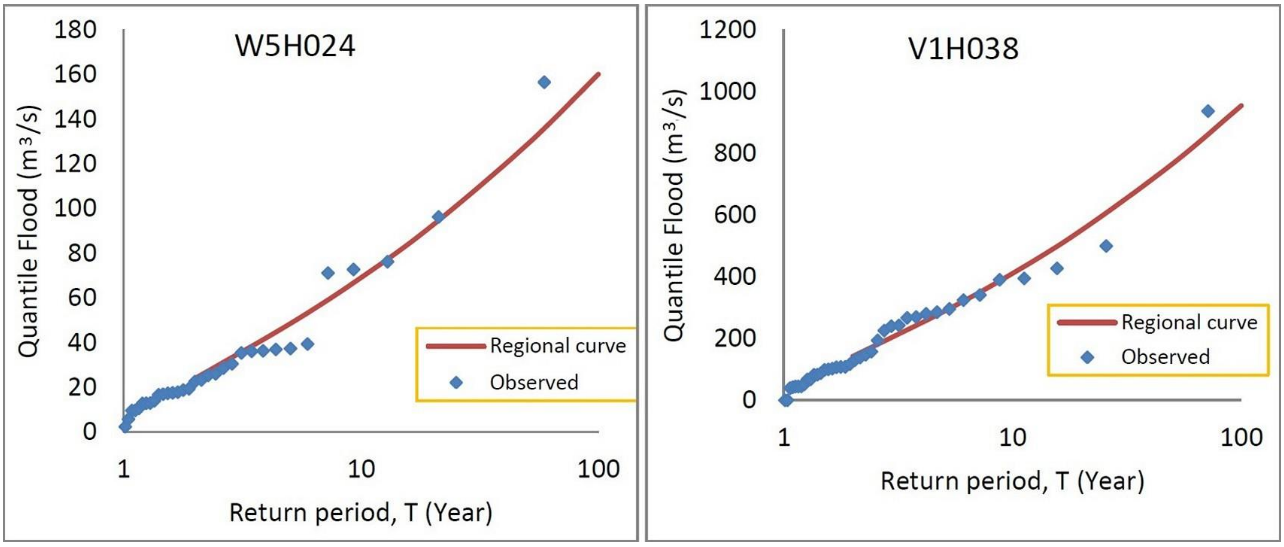


Figure 2-1: Regional curves for Region ZA_R3, (after Haile, 2011)

From Figure 2-1 it can be concluded that outliers were clearly not adequately addressed, producing regional curves (LN3 chosen for ZA_R3), suggesting an evident absence of an upper bound for the AMS flood peaks – if the outliers had not been considered as part of the samples, the regional curves might have changed considerably.

Although Alexander (1990) does discuss and promote the concept of a regional analysis, the practitioner is cautioned in the use thereof, by emphasising the need to consider the basic assumptions in this procedure, which are:

- The region within which the stations are located must be hydrologically homogeneous.
- There should be no spatial correlation between the stations used in an FFA, in a region.

Alexander (1990) also alerted the practitioner to the fact that most of the severe floods in southern Africa are caused by widespread storms which are likely to cover most of the region. Consequently, there will nearly always be some degree of correlation between the records from the stations within the region. Alexander (2000) reiterates his concern that in regional analyses there is a probability of floods occurring simultaneously at two or more sites within a large region. Faber (2010) confirms this concern by observing that the difficulty to put together a collection of gauged sites, that are independent of one another, is one of the major challenges of the regionalisation techniques. In addition, Faber (2010) commented that the large flood events tend to span multiple sites, or even an entire region, thereby causing cross-correlation between the records and consequently reducing the effective size of the data set. This is the case in most extreme events that occurred in SA, which raises some concerns regarding regional analyses, considering the South African conditions.

Alexander (1990) also highlighted, regarding his second basic assumption above, that analytical methods for determining the grouping of station within hydrologically homogeneous regions have indeed been developed overseas, but that these methods required a denser network of stations than what is available over most of SA, thereby indicating the need for the development of a unique approach for SA. Castellarin *et al*. (2012) suggested the use of the regional approach when available data record lengths are short, as compared to the AEP of interest, or for predicting the flooding potential at locations where no observed data are available.

Consequently, a regional FFA approach, for South African conditions, is deemed appropriate for the development of deterministic and empirical methods, but it is not considered realistic in a statistical FFA approach.

2.2.4 An upper bound to AMS flood peak data

The most popular probability distribution functions, used in FFA, typically have two to four parameters with the common feature of having no upper bound. This is particularly true of log-transformed AMS flood peak data portraying a positive skewness coefficient. Information about upper bounded (and/or lower bounded) probability distributions applicable to FFA are not commonly available.

Botero and Francés (2010) interrogate the practice of accepting unbounded probability distributions by stating that the estimated annual maximum flood peaks in the lower AEP range increase without any limit as the AEPs decrease, in the case of unbounded probability distributions. They argued that, considering the specific characteristics of a catchment of interest, like catchment area and geomorphologic characteristics, the obvious question to ask is whether it is possible that a flood peak, with no restriction on its magnitude, can occur. Their own unequivocal response was: *"The straight answer is no, this is not possible"*.

In their study Botero and Francés (2010) defined relative short series, recorded systematically at a flow-gauging station, as 'Systematic' information; and historic and palaeoflood information as 'Non-Systematic' information. A site on the Jucar river in Spain with a Systematic AMS of fifty-six years and four Non-Systematic observed floods, dating back to between 1778 and 1864, were used as a case study. Botero and Francés (2010) chose the Four Parameter Extreme Value (EV4) distribution and included an estimated Probable Maximum Flood (PMF) value, to generate a flood peak record of 450 years using Monte Carlo simulations.

The EV4-, the Four Parameter Log-Normal- (LN4) and the Transformed Extreme Value (TEV) distribution were applied to the data as "upper-bounded" distributions. They used the EV4-, GEV- and the Two-Component Extreme Value (TCEV) distribution, in their robustness analysis.

Botero and Francés (2010) cited the GPA and GEV as commonly used distribution functions in hydrology, which have an upper bound. The GPA is not generally used in SA in FFA and the validity of it being an upper bound distribution is thus unknown. The GEV indeed appears to be an upper bound distribution, if the AMS flood peaks are plotted on a log-probability scale – when plotted on a normal-probability scale, it is obvious that the GEV is not a true upper-bound distribution.

However, their study was again a case where a very site-specific research was performed and should be considered being primarily relevant to Mediterranean rivers only, as acknowledged by Botero and Francés (2010) by concluding that the EV4 distribution function appears to be more robust in a typical Mediterranean river.

The statement by Botero and Francés (2010), that extreme FFA should have an upper-bound, presents the potential for research into a true upper-bound distribution. Their inference that the TCEV should not be used for estimating very low AEP flood peaks and that the GEV should only be used for high AEPs, in an upper bounded population, should be noted in future studies. Bardsley (2016) intuitively recognises that an upper bound should exist and his solution suggests the introduction of a sufficiently high upper truncation point to the flood distribution, with a subsequent modified ARI-scale. For example, he suggested that if a 10 000-year ARI is considered as a truncation point, the ARI-scale must be adjusted to indicate the 10 000-year ARI as infinity (∞). However, the choice of a truncation point is extremely subjective, and a bounded probability distribution might provide a more acceptable solution, which concurs with the primary objective of this study.

2.2.5 Record length

Studies on record lengths are very limited. Bulletin 17C (England *et al*., 2018) recommends a minimum record length of 10 years but proceed in cautioning practitioners against the use of short records that may not be representative of long-term conditions.

Both Tasker (1983) and McCuen and Galloway (2010) used Monte Carlo simulations to determine relative errors for the upper levels of confidence for various AEP AMS flood peaks. Both only considered the LP3 distribution for their studies. McCuen and Galloway (2010) compared their findings with the confidence intervals presented in Bulletin 17B (IACWD {Interagency Advisory Committee on Water Data}, 1982). The findings of both Tasker (1983) and McCuen and Galloway (2010) focused more on confidence in predictions with certain record lengths, rather than evaluating record lengths, per se. Bulletin 17C (England *et al*., 2018) replaced Bulletin 17B (IACWD, 1982) in 2018, but the recommendations did not change much in principle. England *et al.* (2018) stated, as a matter-of-fact, that extended records in time is needed to achieve a more representative sample.

Hattingh *et al*. (2011) concluded, from an investigation done in Namibia, using the GEV distribution, that more than 40 years of data are needed to produce more consistent forecasts in the lower AEP range. Van der Spuy (2018) demonstrated the findings of numerous statistical analyses, which were carried out, by presenting two typical examples with more than 100 years of streamflow data. The purpose was to highlight the minimum number of years, of streamflow data, needed for the probability distributions to produce consistent predictions in the lower AEP range. It was illustrated that the GEV distribution produced consistent results after 20 years and 30 years, respectively.

In Figure 2-2 the results at one of these sites are depicted. The GEV, LN and LP3 distributions, using MM, as well as the GEV_{PWM} were fitted to the AMS. Four record lengths were considered, namely, the observed record length of 105 years, as well as the first 20 years-, the first 40 years- and the first 70 years of the 105-year long record. The shaded areas indicate the range of results obtained, for the different record lengths considered, for just the GEV and LP3, since they produced the best range of fits of the applied distributions. The Regional Maximum Flood (RMF), as an indication of the maximum expected flood for this site, is also shown (Kovács, 1988).

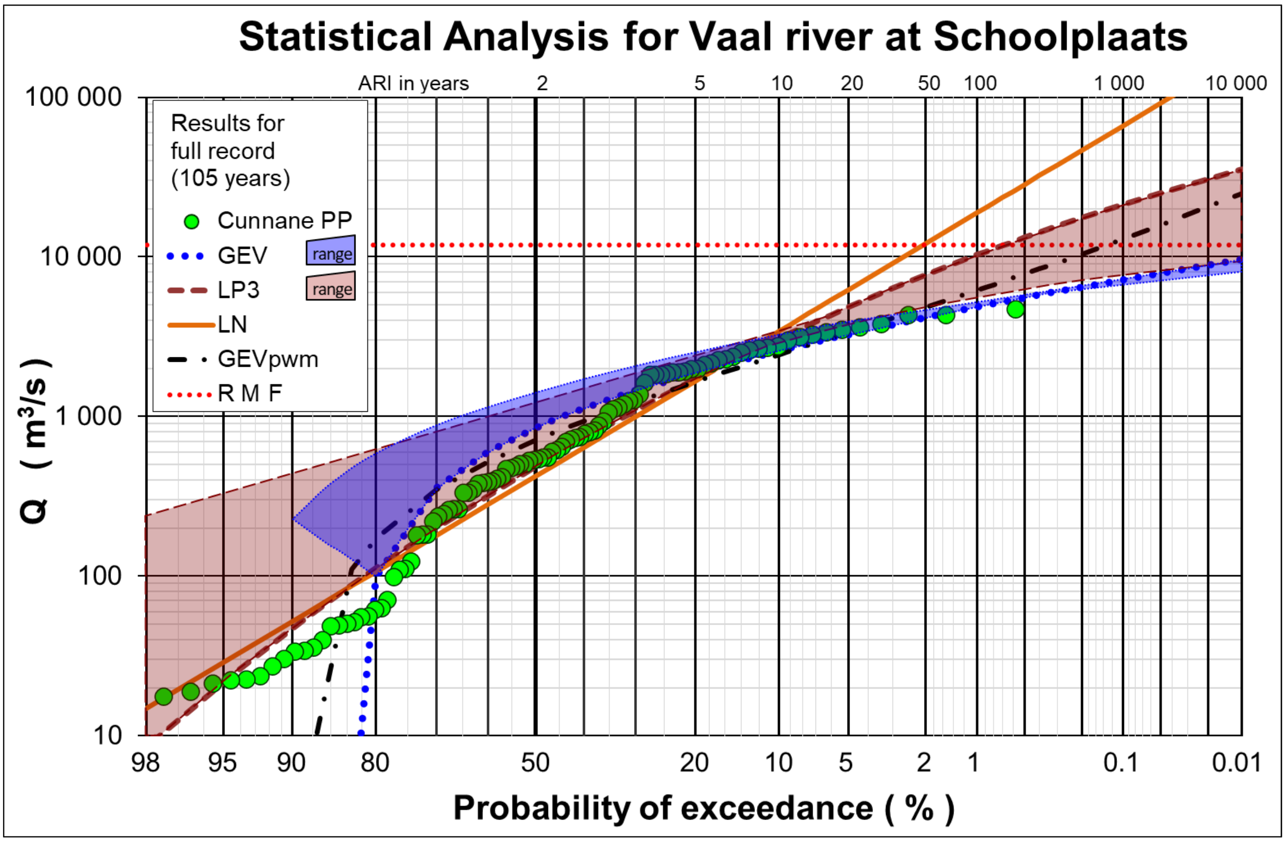


Figure 2-2: Typical example of difference in consistency between the GEV and LP3 distributions (after Van der Spuy, 2018)

Regardless of using 20 years, 40 years, 70 years or 105 years of annual peak flow data, the GEV produced results in the lower AEP range (AEP < 5% or ARI > 20 years) consistently in an exceptionally narrow band (Figure 2-2).

At both representative flow sites, presented by Van der Spuy (2018), the GEV demonstrated a higher consistency in predicting lower AEPs than the LP3, irrespective of record length, and both performed better than the LN and GEV_{PWM}.

Hattingh *et al*. (2011) and Van der Spuy (2018) approached the problem independently and in different ways, but obtained very similar findings, admittedly, using only a limited number of flow sites. These preliminary findings encourage further investigation into the matter of determining adequate record lengths for statistical FFA.

2.2.6 Plotting position

The first mention of a plotting order-ranked data technique, or PP in short, can be found in Hazen (1913). Since then, numerous other PP techniques were introduced as 'alternatives', most probably attempting to cater for outliers that invariably may appear in shorter data record lengths.

PP refers to the probability value assigned to each of the data to be plotted.
Most of the existing PP techniques available in literature can be expressed as

$$P_i = \frac{i+a}{n+1+2a}$$ **Equation 2-3**

where P_i is the 'plotting probability' of the i^{th} order statistic; n is the sample size and a is claimed to be an unbiased plotting parameter, determined 'to fit' different probability distributions.

Table 2-1 provides a list of the most widely used PP techniques (Cunnane, 1978; Adamowski, 1981), indicating the respective value of parameter 'a' in each case.

Table 2-1: Plotting Parameters for most common PP techniques [(1)]

PP technique	*a*	PP technique	*a*
Weibull (1939)	0	Blom (1958)	-0.375
Adamowski (1981)	-0.25	Cunane (1978)	-0.40
Beard (1943)	-0.31	Gringorten (1963)	-0.44
Tukey (1962)	-0.3333	Hazen (1913), Foster (1936)	-0.50

(1) The more recent, so-called unbiased, PP techniques still use ranking of the data as its basis and does not differ much, visually, from the 'older' PP techniques (as illustrated by Yu and Huang 2001; Kim *et al*., 2012)

The PP technique is summarised as follows:

- Arrange the given data-series – in this case the AMS – in descending order.
- Assign an order number to each of the data (termed as ranking of the data), starting at the highest flood peak with i = 1 to i = n for the lowest flood peak.

  The above ranking order is preferred, since it relates directly to an AEP, which directly relates to risk. If the flood peak data are sorted in ascending order (noted in some references) the probability value, assigned to a flood peak data point, indicates probability of non-exceedance.
- Apply Equation 2-3 to assign a probability value P_i to every flood peak. P_i indicates the probability that the corresponding flood peak, Q_i, will be exceeded.

The PP technique is still commonly used (SANRAL, 2013) to assess the performance of the different probability distributions, which is used to estimate flood peak frequencies of AMS flood peak data. However, the PP are probably one of the most misunderstood and misused elements of a statistical analysis. The U.S. Army Corps of Engineers (USACE, 1994) claims that if (flood peak) data are truly drawn from the distribution of a log-normal parent population, log-transformed data will plot as a straight line on a log-probability grid. Hence, it might be misinterpreted by practitioners to suggest that if observed data do not fall on a straight line, the parent population cannot be a log-normal distribution, which is not necessarily correct. PPs are merely rough estimates of AEPs, based on ranking of observed annual maximum events. It will thus, most probably, not reflect the true AEP of observed events and it can be inferred that the PP approach will result in (1) different PPs for two or more events with the same magnitude and (2) a gross overestimation of the AEPs of high outliers.

The description by the WMO (2009) provides a fair explanation of the PP by identifying it as a means to provide a visual display of the data, which also serves to verify that the fitted probability distribution is consistent with the data. It is thus clearly meant as a visual check and a quote by Watt (Posner, s.a.) seems fitting here: "*Do not put your faith in what statistics say, until you have carefully considered what they do not say".* All the PPs, according to the WMO (2009, citing Hirsch and Stedinger, 1987) are only basic estimates of the relative range of AEPs linked to the largest events.

Although PPs are the only visual GOF check to judge the fitted probability distribution(s) against the AMS data, it is an important factor that can influence the analyst to make the wrong choice in selecting a "best fit" probability distribution. For instance, Pegram and Parak (2004) suggested that the RMF had an ARI of approximately 200 years, based on the Weibull PP. Pegram and Parak (2004) then used AMS records, from the three largest RMF-regions at the time in SA, in their study to substantiate their claim. By their own admission they did not exclude excessively large flood peaks from the relative short records. Thus, they did not consider that although observed outliers (including the RMF) will be part of the population, they are not part of the relatively small AMS sample hydrologists must use to estimate what the underlying population should look like. By including these outliers in their FFAs and PPs, lower AEP flood peaks were grossly overestimated, leading to their conclusion that: *"…the return period of the RMF is approximately 200 years…"*. In contrast, Van der Spuy (2018) illustrated that the ARI of the RMF seems to vary between 10 000 years and 100 000 years.

The WMO (2009) also stated that PPs should only be regarded as rough estimates of the relative range of AEPs linked with the largest events. Occasionally, hydrologists fail to recognize that the PP primarily offers a rough visual aid, rather than an absolute measure of probability, against which probability distributions can be compared. This perception is clearly demonstrated by Makkonen (2006) who stated that the PP is considered a standard technique to estimate the probability of extreme weather events.

Scientists seem to be unable to reach consensus about which of the many PP techniques should be considered as the "mathematically correct" one. This PP-controversy already spans more than a century and the scientific interpretation thereof differs from one researcher to another. Makkonen (2006) claimed that, without repeating the extensive and controversial discussions about PP formulas, many of the discussions in papers lacked a theoretical basis and that, consequently, a rather fatalistic attitude, towards selecting a proper formula, emerged. He supported his conclusion, concerning this apathetic viewpoint, by providing three examples: Langbein (1960), cited in Makkonen (2006), compared the choice of a PP technique to "… *like taking a stand on a political question*"; Benson (1962), cited in Makkonen (2006), claimed that the selection of a PP technique

"… *cannot be made by comparing the principles on which each is based."*; and Jordaan (2005), cited in Makkonen (2006) commented on the PP techniques that "… *there appear to be almost as many opinions as there are statisticians.*"

Makkonen (2006) also cited Gumbel (1958) and Castillo (1988) to argue that order ranking and PP techniques have been rigorously analysed mathematically and, consequently, that the theoretical foundations are well known in principle. Makkonen (2006) pondered on the long and controversial history of the PP formulas and the many different types of probability papers to be responsible for the lack of transformation of the mathematical theory into a correct and generally accepted practice.

Horton, Folland and Parker (2001) applied the empirical Jenkinson's method (Equation 2-4) to estimate cumulative probabilities in their research to investigate changes in the incidence of extremes in temperatures. Folland and Anderson (2002) cited Beard (1943) as the source of this PP technique, which is widely referred to as the Jenkinson's method, probably ever since Jenkinson used it extensively in the Flood Studies report (NERC, 1975a,1975b).

$$P = (i - 0.31)/(n + 0.38) \qquad \textbf{Equation 2-4}$$

where P is the 'plotting probability' of the i^{th} order statistic and n is the sample size.

Folland and Anderson (2002), after testing the Jenkinson's method against four other widely used PP techniques, concluded that the Jenkinson's ranking method is likely to be satisfactory for consistently ranking time series of most climatological data, when changes of moderate extremes in the form of percentiles are calculated. Folland and Anderson (2002) also recommended that the Weibull PP should not be used, since it produced estimates smaller than the estimates from the other methods with higher i-values (where data are ranked in ascending order).

Makkonen (2006) – citing Gumbel (1958), Cook (1982, 1985), and Cook, Harris and Whiting (2003) in support – argued that the Jenkinson's method is based on a view that a natural estimate for the PP is the median of its probability density distribution. Subsequently, he claimed that the conclusion reached by Folland and Anderson (2002) on the Weibull formula is strange, since the Weibull PP is generally used and may be considered as an essential part of the standard Gumbel extreme value method.

Cook (2011), drew attention to three claims from Makkonen (2006), namely; (1) that the Weibull estimator $P = i/(n + 1)$ should be used exclusively to derive all statistical properties, including the annual recurrence interval (ARI) and PP; (2) that the improvements, since 1939, in extreme value analysis methods were invalid; and (3) that weather-related building codes and regulations should be updated, by re-estimating previous risk-evaluations. Cook (2011) responded by pointing out that this should have provoked an immediate and urgent response but that the only published response, in the Journal of Applied Meteorology and Climatology, has been some comments by De Haan (2007). Cook (2011) seems to agree with De Haan (2007) on most of his comments and found the issue raised, about the purpose of extreme value analysis not being discussed, as particularly relevant. De Haan (2007) did point out that he only intended to cast some doubt on the claims made in the paper, not meaning to criticise the paper itself. He indicated that the field has seen revolutionary changes since the 1980s and that much research has been done on statistical methods for extremes. He also referred interested readers to the books by Embrechts *et al.* (1997), Coles (2001), and Beirlant *et al.* (2004). De Haan (2007) concluded that the PP is irrelevant to modern extreme value statistics.

Makkonen (2008) subsequently stated that this scientific article is intended to mark the end of the century-long controversial discussion on the PP. He insisted that the Weibull PP should still be used, given that it is unique and independent of the parent distribution. He explained his viewpoint by stating that the core purpose of the extreme value analysis is to estimate the cumulative distribution function (CDF) by order-ranked sample data, so that $Pi = i/(n+1)$, not be considered as an estimate, with Pi the 'plotting probability' of the i^{th} order statistic in n observations. He concluded that all the other PP techniques are based on inappropriate assumptions and should be discarded.

Makkonen's claims were challenged by several subsequent papers (Mehdi and Mehdi, 2011; Cook, 2012; Kim, Shin, Joo and Heo, 2012; Yahaya, Nor, Jali, Ramli, Ahmad and Ul-Saufie, 2012a; Yahaya, Yee, Ramli and Ahmad, 2012b; Fuglem, Parr and Jordaan, 2013), who maintained that other PP techniques, developed by renowned statisticians, were just as valid.

Makkonen, Pajari and Tikanmäki (2013) and Makkonen and Pajari (2014) continue this dispute where they again concluded with similar remarks, as in previous papers. For example, Makkonen *et al*. (2013) claimed that their results indicated that random simulations, similar to those of Fuglem *et al.* (2013), and also presented by others (e.g. Harris, 2001; Mehdi and Mehdi, 2011; Cook, 2012; Kim *et al.*, 2012 and Yahaya *et al.*, 2012a, 2012b), are misleading and their conclusions regarding the plotting methods are invalid.

Makkonen *et al*. (2013) again persisted in their conclusion that their theoretical analysis and random simulations, indicated that the Weibull formula is the true rank probability and should be used as the unique PP irrespectively of the underlying probability distribution.

From the above the following can be concluded:

- Scientists, in the PP debate, were only considering the issue from a theoretical perspective.
- The hydrologists seem to ignore the obvious reality that the analysts indeed only have a very small sample of the underlying population of AMS flood peaks to analyse. Consequently, the PP-issue seems to demand a more holistic approach, rather than a pure theoretical one.
- All existing PP techniques, although useful to some extent, can only provide basic estimates of probabilities of extreme events.

It was illustrated by Van der Spuy (2018) that the differences between the existing PP techniques are so insignificant that there is, for practical purposes, no real difference between them (Figure 2-3). The Weibull and Hazen PPs represent the two extremes. The Adamowski PP plots approximately halfway between these extreme PPs. The rest of the PP techniques that include the familiar Cunnane, Gringorten and Blom PPs, all fit in between the Adamowski and the Hazen PPs.

Noticeable differences exist only at two to four data points at both the lower- and higher ends of the probability scale. It is postulated that using any PP technique will result in the same choice of a probability distribution.

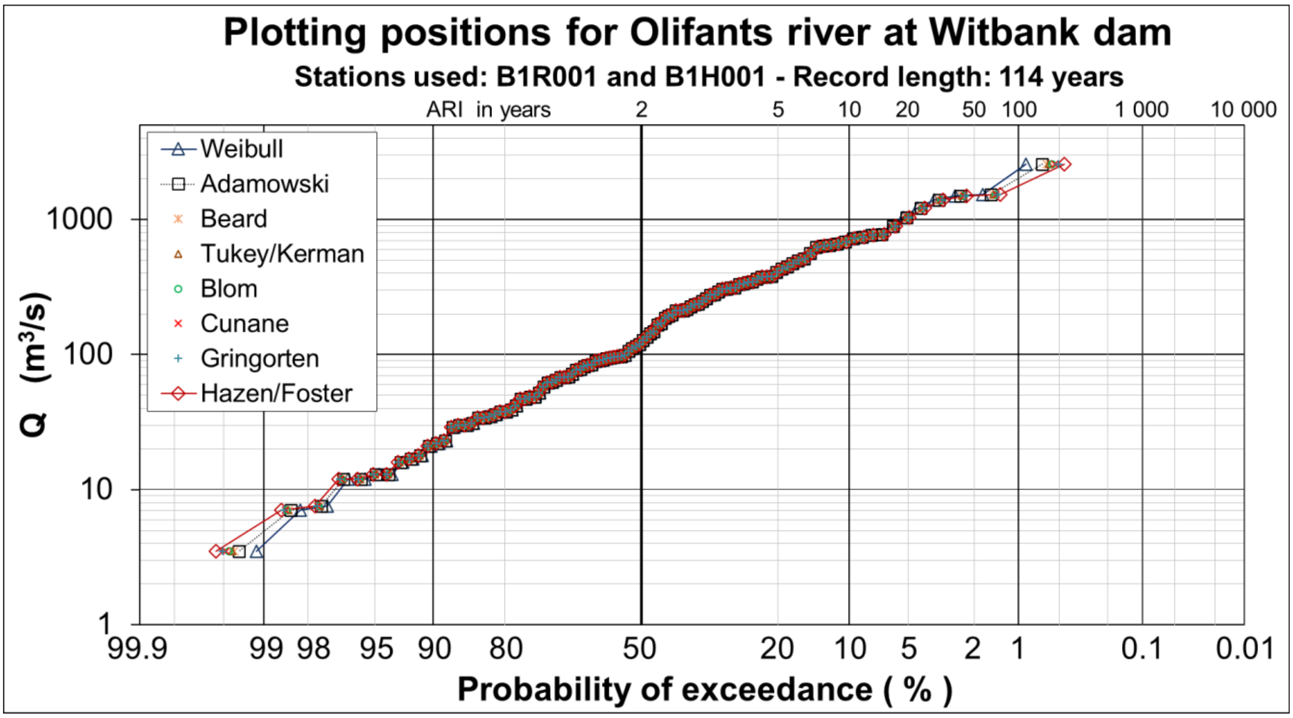


Figure 2-3: Illustrate 'differences' between existing PPs (after Van der Spuy, 2018)

It would seem as if the PP techniques, developed after the Hazen PP (1913) and Weibull PP (1939), attempted to accommodate outliers in a very simplistic way. Unfortunately, if the data-record contains a significant high outlier, the assigned PP (AEP) would most probably be too high.

2.3 Performance of commonly used probability distributions

The following discussion primarily considers the most commonly used methodologies of choice in SA, supplemented by some perspective from international literature. Key aspects will also be deliberated, to illustrate general differences in approaches and to draw attention to some misconceptions that may lead to erroneous application and consequent flawed deductions from fitted probability distributions.

2.3.1 International preferences

It is important to acknowledge that there is still, in flood hydrology, a disagreement on which probability distribution provides the 'best' fit to observed AMS flood peak data. Despite numerous studies, three probability distributions seem to have withstood the test of time, namely the LN, LP3 and the GEV, which contains three extreme value continuous probability distributions, with some variations in application (for example, applying different moments; using 2 or 3 parameters; etc.).

In general, the LP3 is the preferred distribution in the USA (Faber, 2010; England *et al*., 2018), although the use of more probability distributions is encouraged if a reasonable fit to the data is not provided by the LP3 (USACE, 1993). In research, to identify the preferred probability distribution for at-site FFA in Canada, Zhang, Stadnyk and Burn (2019) concluded that the GEV is better than the other considered probability distributions.

In Australia the Generalized Pareto (GPA), GEV, Three-parameter Log-Normal (LN3), LP3 and the two-parameter Log-Normal (LN2), are all considered as acceptable alternative distributions (Vogel *et al*., 1993). Rahman, Rahman, Zaman, Haddad, Ahsan and Imteaz (2013) confirmed that the LP3, the GEV and GPA distributions have been identified as the top three best-fit flood distributions, in Australia.

In the UK the GEV distribution was originally preferred according to NERC (1975). Robson and Reed (1999) as well as Asikoglu (2018), citing Salinas, Castellarin, Viglione, Kohnová and Kjeldsen (2014), indicated that the Generalised Logistic distribution (GLO) replaced the GEV as the preferred distribution in the UK. Castellarin, Kohnova, Gaál, Fleig, Salinas, Toumazis, Kjeldsen, and Macdonald (2012) produced a comprehensive review on the European Cooperation in Science and Technology (COST) Action ES0901 European procedure for flood frequency estimation, initiated in 2010, and concluded that the GEV, LN3 and GLO are generally the most suitable distributions for the various nations included in the dataset. Castellarin *et al.* (2012) also stated that, because the database included annual maximum series, the GPA distribution was found to be generally inappropriate, as should have been expected, since the GPA has been shown to be mostly suitable for representing frequencies of partial duration series.

However, in SA the LN, LP3 and GEV, either with method of moments (MM) or probability weighted moments (PWM) are proposed (Alexander, 1990). Both LP3 and GEV distributions provided good results and the practice is to apply both and choose the one that visually seems to fit the data best (Van der Spuy and Rademeyer, 2018).

Unfortunately, from experience gained in more than thirty years of performing numerous flood frequency analyses, as well as observations made by researchers around the world (Wallis and Wood, 1985; Gunasekara and Cunnane, 1992; Mutua, 1994; Galloway, 2010; Lettenmaier, 2010), it has to be admitted that the current probability distributions do not perform very well in estimating AMS flood peaks towards the lower AEP range; as will be disclosed in the discussions that follow.

AMS flood peak frequencies are generally overestimated and many times grossly so, if the observed log-transformed data record yields a positive skewness coefficient.

2.3.2 Considering the LN distribution

The normal distribution was first developed by De Moivre in 1753 (Van der Spuy and Rademeyer, 2018). The distribution is widely used in meteorology and hydrology, as well as in other civil engineering applications, such as measurement errors (survey).

Van der Spuy and Rademeyer (2018) commented that Hazen (1913) is credited with having observed that, while hydrological data are usually strongly skewed, the logarithms of the data have a near symmetrical distribution. The LN distribution, which is a normal distribution fitted to the logarithms of the observed values, is thus also used in analysing AMS flood peak data.

This distribution is symmetrical about the mean and is therefore only suitable for data where the skewness coefficient of the log-transformed data is equal to, or close to zero. The LN distribution is a special case of the LP3 distribution, where g equals (or forced to be) zero.

2.3.3 Considering the LP3 distribution

In Pakistan the results of an FFA on the river Swat indicated that the LP3 and GEV ranked as the top two distributions at all four sites along the river, with LP3 the top distribution at two of the four sites (Farooq, Shafique and Khattak, 2018). In a Kenya study it was observed that the Wakeby (named after a lake in Australia) and the LN3 distributions were considered the best models for FFA in Kenya (Mutua, 1994). However, due to the complexity of the Wakeby distribution, it was suggested that the LN3 be considered as the best model for flood frequency analysis in Kenya. Conversely, Mutua (1994) stated that the three-parameter log Pearson distribution was found to be one of the worst fitting distributions, despite its popularity within the country.

Wallis and Wood (1985), also cited by Alexander (1990), indicated that the flood quintile estimates obtained from the LP3 distribution are significantly worse than those obtained from the GEV and Wakeby. They further commented that the U.S. Water Resources Council (USWRC) Bulletin 17B guideline is in need of re-evaluation. Twenty-five years later Lettenmaier (2010) remarked that he personally believes that it is time to *"give Bulletin 17B a decent burial"* and to move on. 17B was eventually replaced by Bulletin 17C (England *et al.*, 2018), eight years later. Galloway (2010) simultaneously stressed that the replacement of Bulletin 17B should be considered as urgent and voiced the need for scientists to proceed forward with good science and to look differently at solutions and added: *"Don't help people do wrong things more precisely"*.

Van der Spuy (2018) and Van der Spuy and Rademeyer (2018) indicated that the LP3 distribution is not only sensitive to high outliers but can be seriously affected by low outliers, as illustrated in Section 2.3.5.

2.3.4 Considering the GEV distribution

The GEV is an extreme value distribution which seems to be largely less affected by high outliers, when compared to the LP3 distribution, and is virtually immune to lower outliers. The downside is that the GEV does not normally produce reliable estimates for the higher AEPs ≥ 50% (ARI ≤ 1 in 2-year). It will, however, produce (generally) more reliable estimates at the lower AEPs; i.e. AEPs ≤ 20% (ARIs ≥ 5 years).

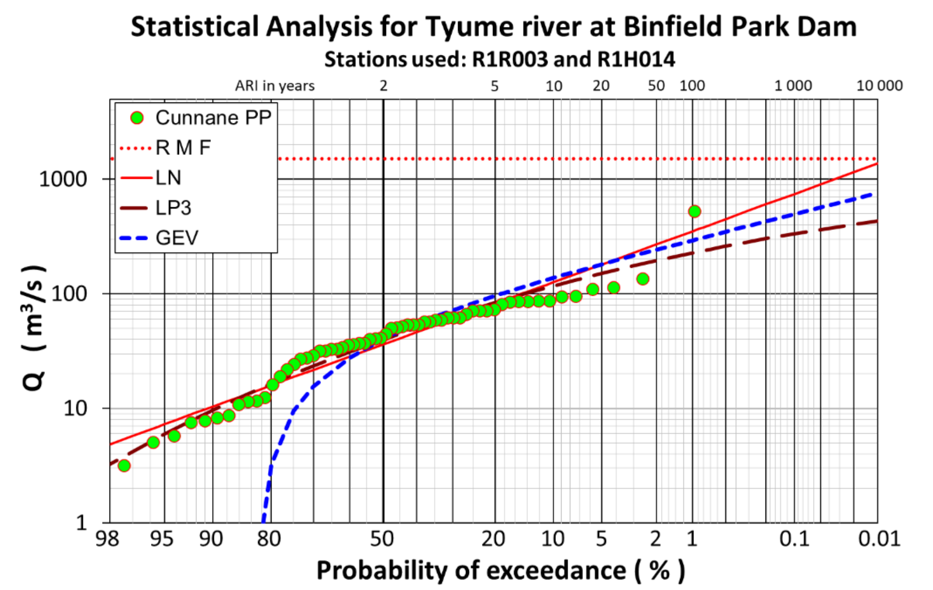
Vogel *et al*. (1993) found that the GPA and GEV approaches are preferred in Australia outside the winter-dominated rainfall regimes and that they are also considered to provide probably the best description of flood flows across the entire Australia.

Van der Spuy and Rademeyer (2018) also indicated that the GEV generally outperforms the other probability distributions used in SA, in FFA (see also Figure 2-2 and the accompanying discussion in Section 2.2.5)

2.3.5 Sensitivity of probability distributions to outliers

Van der Spuy (2018) and Van der Spuy and Rademeyer (2018) suggested practical guidelines to deal with various aspects, encountered in a statistical analysis. The importance of dealing sensibly with outliers in a data record, was one of the highlighted aspects. FFAs performed at various flow sites (DWS, 1993-2021) have shown that outliers will affect how probability distributions fit the observed data. MM was used in all the analyses, since PWMs and LMs tend to adjust the weight of data points, thereby inherently considering an outlier as part of the sample. Two of these analyses were chosen to illustrate the impact of outliers on various probability distributions, mainly concerning the lower AEP-range that relates to the design flood range and associated risks (see Figure 2-4 and Figure 2-5).

In Figure 2-4, the sensitivity of the probability distributions towards high outliers is illustrated.



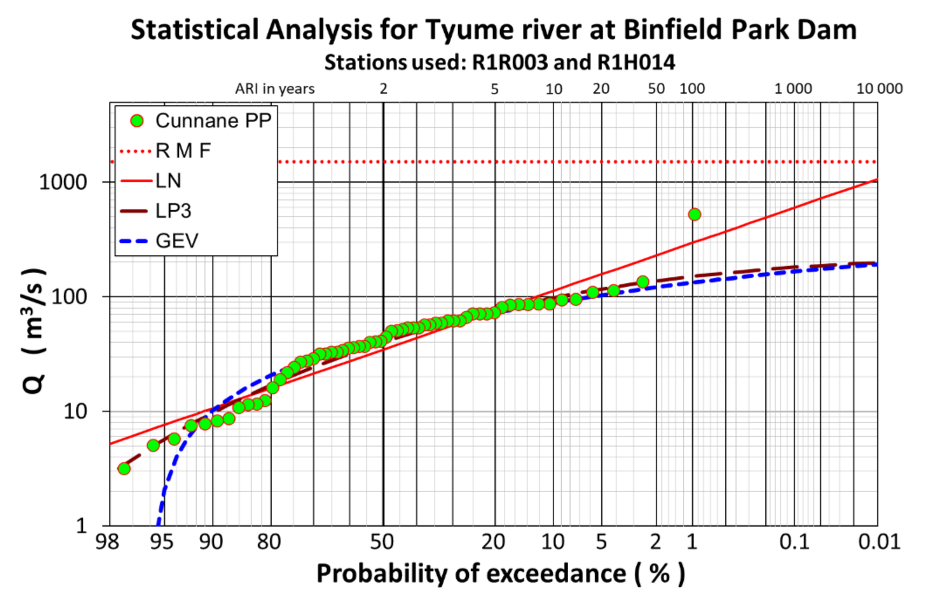
Record length 62 years

No points considered as outliers – used all data points in FFA

The estimated 0.01% AEP (ARI = 10 000 year) AMS flood peak:

GEV – 758 m³/s
LP3 – 430 m³/s
LN – 1 370 m³/s

(Highest AMS flood peak in record: 525 m³/s)



Regard flow higher than 200 m³/s as high outlier (thus, 61 years of record used in FFA)*

The estimated 0.01% AEP (ARI = 10 000 year) AMS flood peak:

GEV – 191 m³/s
LP3 – 198 m³/s
LN – 1 067 m³/s

Figure 2-4: Sensitivity of GEV and LP3 distributions, towards high outliers

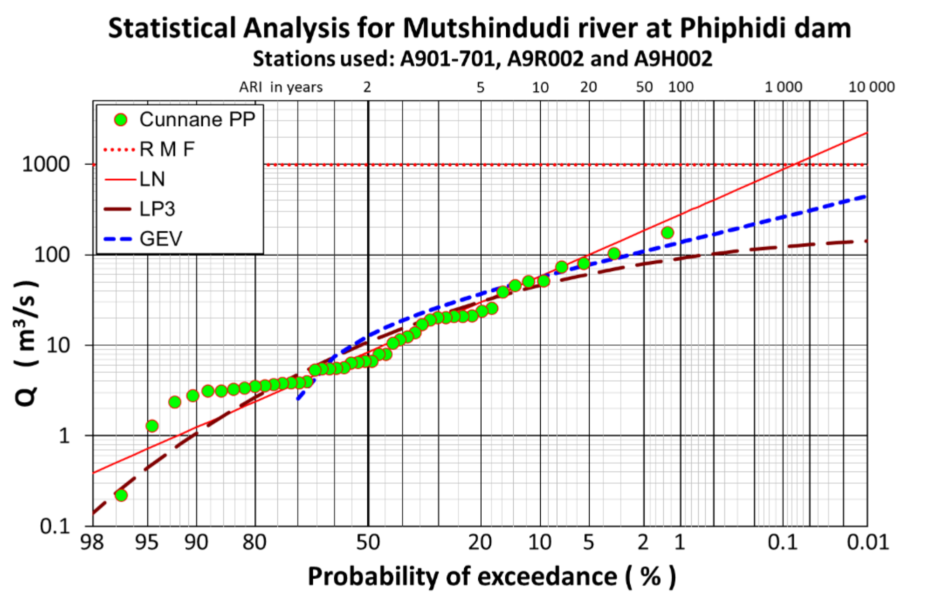
***** In the event where outliers are excluded, the data-points are still displayed in the PP-plots since it remains proper observations. Hence, the outliers are only excluded in estimating the moments for the FFA.

The inclusion of the high outlier, or not, seems to have a slightly bigger effect on the GEV than on the LP3 distribution. The effect on the LN distribution proved to be much less.

However, what is more interesting is that the exclusion of the outlier seems to have an adverse effect on both the GEV and the LP3 – to the extent that both distributions seem to deem this flood peak not as part of the fitted distribution (no associated frequency).

Even when no data point is considered an outlier, the LP3 is not performing well – estimating the AEP, of the maximum flood peak in the record, as 0.0005% (a 1-in-200 000 event!). The most appropriate solution for this site would seem to be not to exclude the outlier and to use a combination of the GEV, for the lower AEPs, and the LP3, for the higher AEPs. (Gericke and Du Plessis, 2012)

In Figure 2-5 the sensitivity of the probability distributions to lower outliers/flows is illustrated.



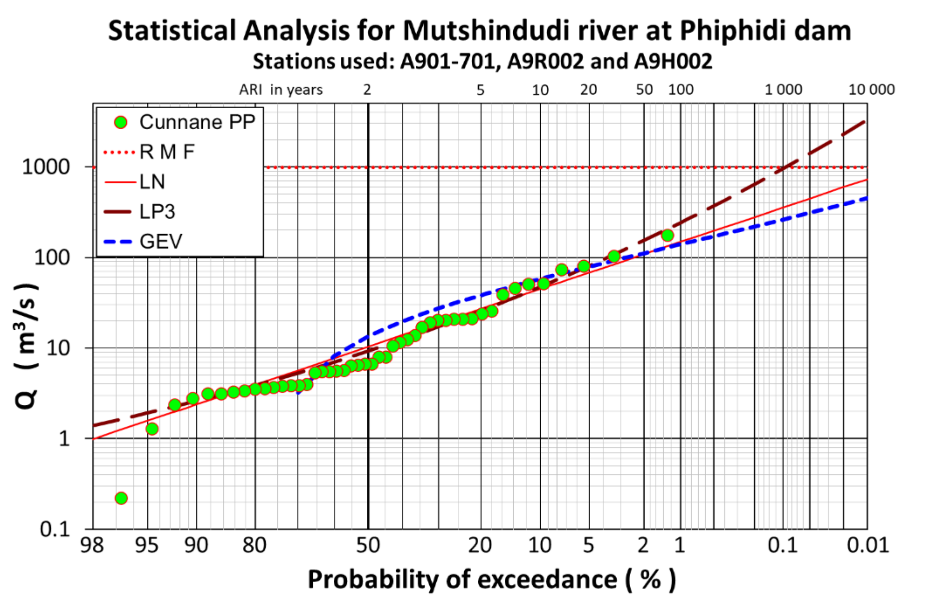
Record length: 51 years
Missing data: 3 years

No points considered as outliers – used all valid data points (48 years)

The estimated 0.01% AEP (ARI = 10 000 year) AMS flood peak:

GEV	–	449 m³/s
LP3	–	142 m³/s
LN	–	2 253 m³/s

Note:
lowest peak not shown on graphs (< 0.1 m³/s)



Regard flows less than 1 m³/s as low outliers (thus 46 years of record used)

The estimated 0.01% AEP (ARI = 10 000 year) AMS flood peak:

GEV	–	451 m³/s
LP3	–	3 316 m³/s
LN	–	732 m³/s

Figure 2-5: Sensitivity of GEV and LP3 distributions, towards low outliers

In this case the inclusion of the low outlier, or not, seem to have virtually no effect on the GEV, but the LN and, particularly, the LP3 are adversely affected. It was further observed in informal studies and numerous FFAs (DWS, 1993-2021) that the way in which the lower peaks are distributed can also affects the LP3 and LN, but the effect on the GEV is still negligible – see Figure 2-2 where this is also illustrated.

It is concluded that the GEV appears to be markedly less impacted by low outliers in estimating lower AEP flood peaks (higher ARI flood peaks) than the other two distributions, in all relevant observed cases (DWS, 1993-2021). Conversely, all the probability distributions are affected by high outliers. If high outliers are present, either the GEV or the LP3 might be impacted more than the other, but there is no clear indication of the impelling cause(s) and it is clear that more research is needed.

2.3.6 Performance of probability distributions with longer records

How the different probability distributions might perform with very long records will only be truthfully answered in the distant future, as truly long records become available.

However, to get some insight, a 100+ year observed flow record was used to stochastically generate a 10 000-year flow record (independently, using Monte Carlo simulation techniques) and the AMS was extracted from the generated record. The results, as an example, are presented in Figure 2-6 (Van der Spuy, 2018).

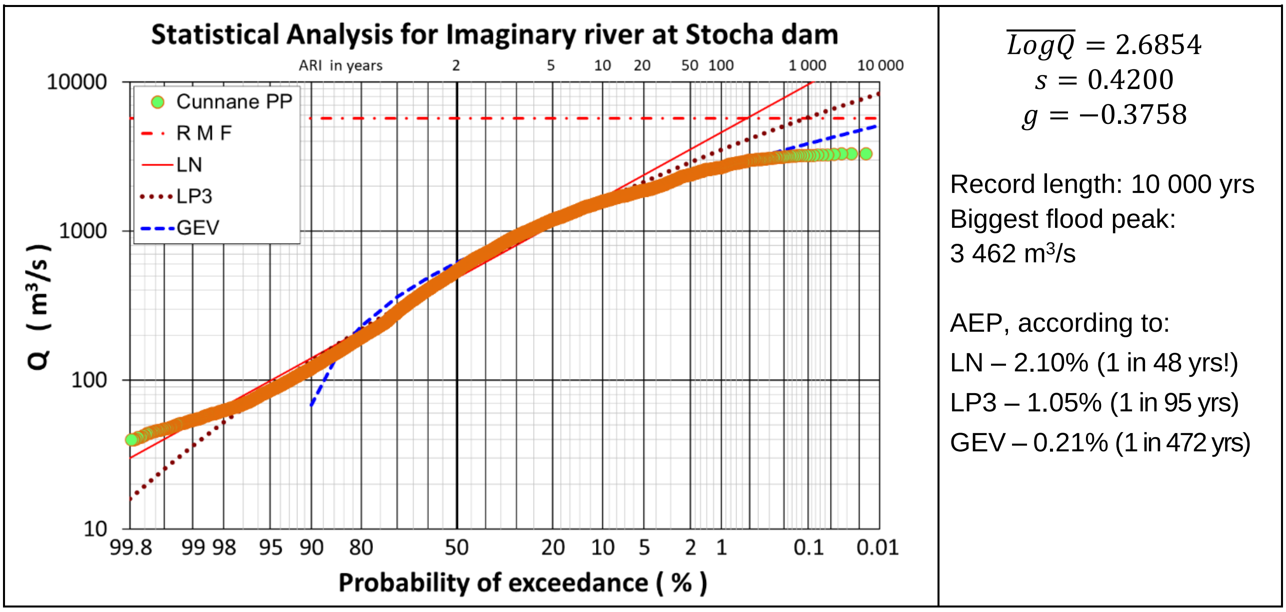
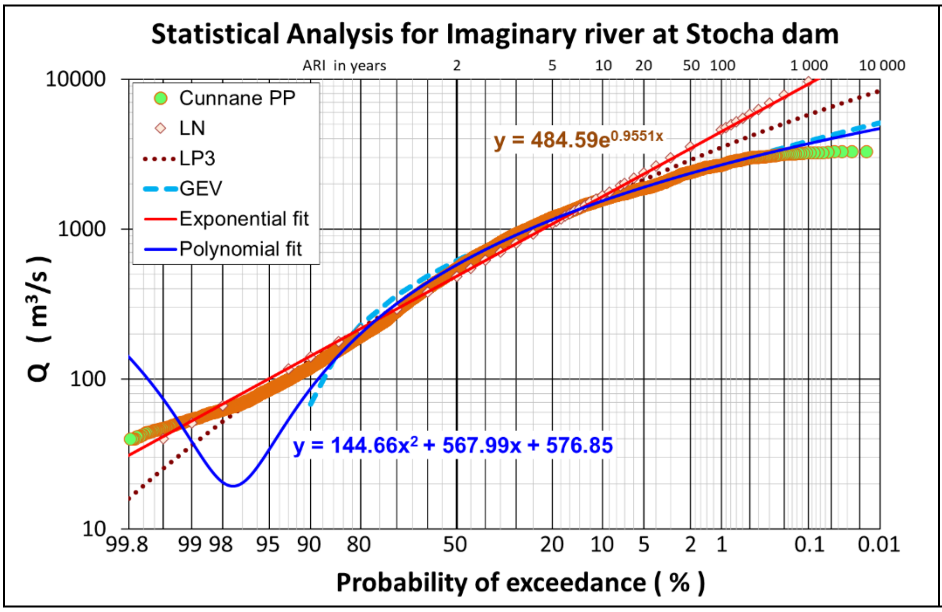


Figure 2-6: Probability distributions' performance, with a 10 000-year generated flow record

The GEV again outperformed the other two distributions, in estimating lower AEP flood peaks, by a large margin (as expected), but it nonetheless 'only' assigned an AEP of 0.21% (ARI = 472 years) to the highest flood peak in the 10 000-year record. This case-study supported the hypothesis that the probability distributions are more likely to overestimate estimated flood peaks at the lower AEPs, but the degree of overestimation with such a long record (albeit generated) was not expected.

These findings led to the question of how would these probability distributions compare with simple mathematical trendline equations, fitted to the data? (It might seem contrary to what was said previously about PPs and outliers in Section 1.3.2 and Section 2.3.5, but the mentioned arguments are not applicable to a 10 000-year record, with no obvious outliers).

Mathematical trendlines were fitted to the data in the example portrayed in Figure 2-6, to mimic the LN and the GEV distributions. The results are depicted in Figure 2-7.



Not surprisingly, the LN is almost identical to an exponential fit.

The GEV closely resembles a 2nd degree polynomial fit.

The LP3 could not be mimiced by a single simple mathematical equation.

Figure 2-7: Comparing the probability distributions with simple curves fitted to the data

Concerning present global practices, the observations above trigger thoughts like:

- If the LN already mimics an exponential mathematical trend, should the LP3 even be considered if the log-transformed data indicate a positive skewness?
- It may, practically, be more sensible to use combinations of probability distributions; for example:
  - Fitted probability distributions can be 'linked':

    For example, in Figure 2-4, with the high outlier included, accept the LP3 distribution for AEPs higher than 50% and the GEV distribution for AEPs equal to or lower than 50%.
  - Consequently, probability distributions can be 'combined' by means of weighted averages:

    Both Van der Spuy (2018) and Gericke and Du Plessis (2012) concluded that an option to use combinations of multiple probability distributions can add value to finding a realistic flood peak, especially in the lower AEP domain. Gericke and Du Plessis (2012) used the Mean Logarithm Value Approach (MLVA) to illustrate the success thereof to overcome limitations, like short record lengths, associated with the single site approach.

    For example, in Figure 2-4, with the outlier excluded, accept the LP3 for AEPs higher than 20% and the MLVA of the LN and GEV for AEPs equal to or lower than 20%.

Although the above approaches will, visually, produce more pleasing fits to the plotted data, it is appreciated that it will probably receive criticism and serious opposition from statisticians, but who can honestly claim that the best probability distribution, for a particular data set, is NOT the average of two (or more) known probability distributions if, by omitting one or two outliers from a data set, the forecasted 10 000 year flood peak can change (using the same probability distribution) by more than 2 200% (difference for LP3 in Figure 2-5)?

2.4 Concerns associated with existing trends in FFA

2.4.1 Background

Present trends in FFA raise some concerns, as indicated in Section 1.2, that poses the question if hydrologists are not just doing the wrong things more precisely? (Galloway, 2010). A recent study by Calitz (2020), reflecting some of the recent trends in FFA, is used to illustrate some of these concerns, which includes outliers, normalisation of data and goodness of fit (GOF) tests.

Although Calitz (2020) considered regionalised approaches in SA, he also researched the *'Identification of a distribution suitable for at-site flood frequency analysis in South Africa'*. Calitz (2020) considered the LP3, LN, GEV, PE3, GPA and 3-Parameter Kappa (KAP3) probability distributions for comparison.

Since the way outliers are perceived seems to be the underlying cause of most concerns related to present practices, an introductory discussion on the topic will be presented in Section 2.4.2. In Section 2.4.3 the assumptions and results of Calitz (2020) will be considered. Section 2.4.4 will elaborate on the related concerns with existing practices.

2.4.2 Outliers

Frost (s.a.) stated that, while there is no strict statistical rule or mathematical definition to identify outliers, guidelines exist through which possible outliers can be identified. He emphasised that a sound knowledge of the subject-area and understanding of the data collection process is crucial to accurately identify outliers. He described five methods, including: (1) sorting data, (2) graphing data, (3) using Z-scores, (4) using the interquartile range and (5) hypothesis tests, to identify outliers in datasets, noting the advantages and disadvantages of each. Frost (s.a.) indicated that the biggest disadvantage of the Z-score approach is that a high outlier in the dataset inflates the mean and standard deviation. However, the Z-score provides a measure to compare relative probabilities, associated with relative magnitudes of the data. It is however recognised that, if low- and high outliers are present, it will most probably have a bigger effect on the standard deviation than on the mean. Since nearly 100% of the data will be within 3 standard deviations of the mean, data with a Z-score higher than 3, or lower than -3, can considered to be outliers (Brownlee, 2018; Frost, s.a.).

Ball, *et al.* (2019) consider outliers as observations that are inconsistent with the general trend of the rest of the data. They suggest censoring of outliers, especially Potentially Influential Low Flows (PILFs). In statistics, censoring implies that the value of an observation is only partially known. While Ball, *et al.* (2019) illustrate removal of identified PILFs in their examples, it seems that high outliers are not excluded. England, *et al.* (2019) also suggested the same approach with more emphasis on PILFs and, additionally, dedicated a full Appendix on PILFs. This most probable emanates from the sensitivity of the LP3, the preferred probability distribution, to PILFs. Both Ball, *et al.* (2019) and England, *et al.* (2019) suggest the Multiple Grubbs-Beck Test (MGBT) to identify PILFs in an AMS.

Reed (1999), citing Barnett and Lewis (1984), proposes four possible treatments for outliers, of which an erroneous observation is the only reason considered for rejecting the outlier, since he considers it as "*bad practice*" to ignore outliers in FFA. Reed (1999) also questions the possibility of an extraordinary flood occurring within the gauging period and shifts the onus to the analyst to 'prove' that the flood peak exceeds anything experienced in a period several times longer than the gauged period (which in almost all cases amounts to an impossibility!). This view is deemed to be very naive, since there is statistically approximately a 22% probability that a flood peak with an ARI of 4 times the record length can occur in a record – which is equivalent to the probability that a 5-year flood peak can occur in any given year. The view by Reed (1999) that FFA procedures were 'designed' to be robust against outliers is also flawed (see Section 2.4.4) and it results in a flood peak, with potentially a very low AEP, to be considered as a flood peak with an AEP corresponding to the record length.

2.4.3 Assumptions and results re Calitz (2020)

Calitz (2020) claims that several authors recommended the LN, LP3 and the GEV for FFA in SA, but that neither provided guidance for which method is the most suitable. However, Van der Spuy and Rademeyer (2018) clearly stated that the GEV seems to be less sensitive to outliers than the LP3, but advised to use both to provide the analyst with a wider choice, since all probability distributions have limitations. Van der Spuy (2018) also indicated that the GEV_{PWM} did not improve the results of the GEV_{MM}.

Calitz (2020) used logarithms of data (LN and LP3) and adopted the LM for estimating moments, by claiming that it is more robust where outliers are present in the dataset, thereby 'normalising' the data (directly or indirectly), which is a concern to be addressed in Section 2.4.4.

Calitz (2020) used the LM-diagram, with the LM-kurtosis representing the y-axis and the LM-skewness the x-axis, as the GOF test. The LM-diagram as a GOF test raises some serious questions, which will also be addressed in Section 2.4.4. Calitz (2020) used the GOF test to eliminate the GEV, thereby ignoring the observation by Vogel *et al.* (1993) that although the GEV procedures seemed to perform well for all regions considered, the LM-diagrams did not always favour the GEV procedure.

2.4.4 Concerns related to present practices

Outliers and normalising data

Many authoritative publications (e.g. Reed, 1999; Robson and Reed, 1999; England, *et al.*, 2019, Ball, *et al.*, 2019) consider PWMs and LMs as robust against outliers in a dataset. This belief suggests a fundamental distortion of the underlying truth. By assigning a lower weighting to the larger values (Cunnune, 1985) the data appears to be more 'normally distributed', which conceals the signal (implication) contained by the outlier in the original data. Hence, while LM might 'appear' to be 'more robust' if outliers are present, in reality it disguises the truth about outliers, by assuming normality of data, which is exactly what Wheeler (2009a, 2009b and 2009c) cautioned against (see next paragraph). Transforming either the data or the moments alter the results, which is considered unacceptable since it no longer reflects the significance of the observed data.

With regards to normality of data, interesting viewpoints are raised by Breyfogle (2009a, 2009b) and Wheeler (2009a, 2009b and 2009c). If data are routinely examined for normality, Wheeler (2009a) suggests that it is symptomatic of what he defines as 'leptokurtophobia' (leptokurtosis refer to probability models with heavier tails than a normal distribution). Although Breyfogle (2009a) and Wheeler (2009a) do disagree on certain aspects relating to behaviour charts (a time-series analogue to an AMS), they do agree that results should be considered in the context of the original data and that to transform data to achieve certain statistical properties is unacceptable and misleading. Transformation is only acceptable if it makes sense both in terms of the original data and the objectives of the analysis. Wheeler (2009b, 2022b) stated that data is not generated by a probability model but by a process or system, subject to change and upsets over time – which is a true and accurate observation applicable to flood hydrology. He continues by stating that if data are not homogeneous, a single probability model will never be suitable, and if the data are transformed to make it appear 'more normal' one can "*end up with a beautiful, but completely incorrect, analysis*". Wheeler (2009a; 2009c, 2010) explains that transformation unavoidably distorts and obscures the signals contained in the original data and change the results of the analysis, which explains why outliers 'disappear' in present FFA approaches. In support, Kline (2016) warns that applying a transformation causes a sacrifice of the original meaningful observation and one should consider if the expectation of normality is reasonable, before applying a normalising transformation. Wheeler (2022a) also expressed his concern that many believe the first step in data analysis is to check for normality. If the data are strongly skewed, as is often the case with AMS of flood peaks, Wheeler emphasised that it is a "*mathematical fact of life*" that no skewed probability model exists to fit these data. Wheeler (2009b) concluded by stating that if being taught inaccurate ideas, it is hard to change and he also reminded the reader of a caution from Boorstin (s.a.): "*The greatest obstacle to discovery is not ignorance, but rather the illusion of knowledge*".

<u>LM-diagram as a GOF test</u>

The statistical meaning of skewness and kurtosis is well documented. However, the statistical meaning of the transformed moments is not clear, since it does not reflect the perspective of the original data. Wheeler (2011b) also expressed his concerns on the use of skewness- and kurtosis-statistics. He explained that the mean (location) and SD (dispersion) will always be estimated with greater precision than the skewness and kurtosis. For example, respectively 6x and 24x the data is needed to estimate the skewness and kurtosis with the same precision as the mean and SD estimates. A query to Wheeler (2023) evoked the following remark from him "*I have not come across the nomenclature of probability weighted moments and L moments, but fitting the moments of any probability model is going to end up being an exercise in fitting the noise. As you have discovered,* *once you have characterized the location and dispersion of a data set, you gain little by moving on to higher moments.*"

It is, thus, concerning that the transformed version of the two least accurate moments of the first four moments, be used in a GOF test. It might also explain the observation by Vogel *et al.* (1993) that the GEV procedures seemed to perform well in Australia, even though the LM-diagrams do not always favour the GEV procedure.

2.5 Literature review deliberations and conclusions

Due to the inconclusive outcome of the PP-debate, as well as what has been observed about the PP techniques, it was considered fitting to deliberate in more detail on the PP concerns, before concluding on the observations from the literature review in its entirety.

2.5.1 Reflecting on existing Plotting Position techniques

An improved PP technique is seen as a key benchmark in researching an enhanced probability distribution. Certain gaps and concerns were observed in the existing PP techniques. It was considered essential to identify these concerns through a critical review, to address them properly in this research. It was anticipated that the existing PP techniques might form part of the development of an improved PP technique. Hence, the existing PP techniques were considered as an integral part to the research and will be discussed as an overview for the research methodology in Chapter 5.

In Figure 2-8 the AMS flood peaks at Woodstock Dam (V1R003) were used to highlight concerns associated with the current PP techniques, namely:

- Given that PPs are meant to provide a visual trend, to assist in the choice of the most appropriate probability distribution, it is evident from Figure 2-8 that any of the PP techniques will yield the same outcome in the choice of a probability distribution, which supports the hypothesis under Section 2.2.6, based on a presentation by Van der Spuy (2018). Despite the century long controversy, around the different PP techniques, it would seem as if the differences between the existing PP techniques are, for practical purposes, insignificant.
- The 'ranked' PP (i.e. probability) assigned to an outlier (indicated by the red rectangle on Figure 2-8) is most probably incorrect and can result in the analyst not choosing the most appropriate theoretical probability distribution.
- Due to the ranking process, values in the dataset having the same, or very similar, magnitudes will be assigned different PPs (i.e. probabilities). This may distort the visual appearance of the PP to such an extent that it may complicate the choice of the most applicable probability distribution (indicated by red ellipses on Figure 2-8).

An additional concern is that estimated PP probabilities are used as a benchmark against which "goodness-of-fit" tests are performed to determine which probability distribution is "a better fit" to the data – it is quite possible that it might indicate the "worst fit", especially if outliers are present in the record.

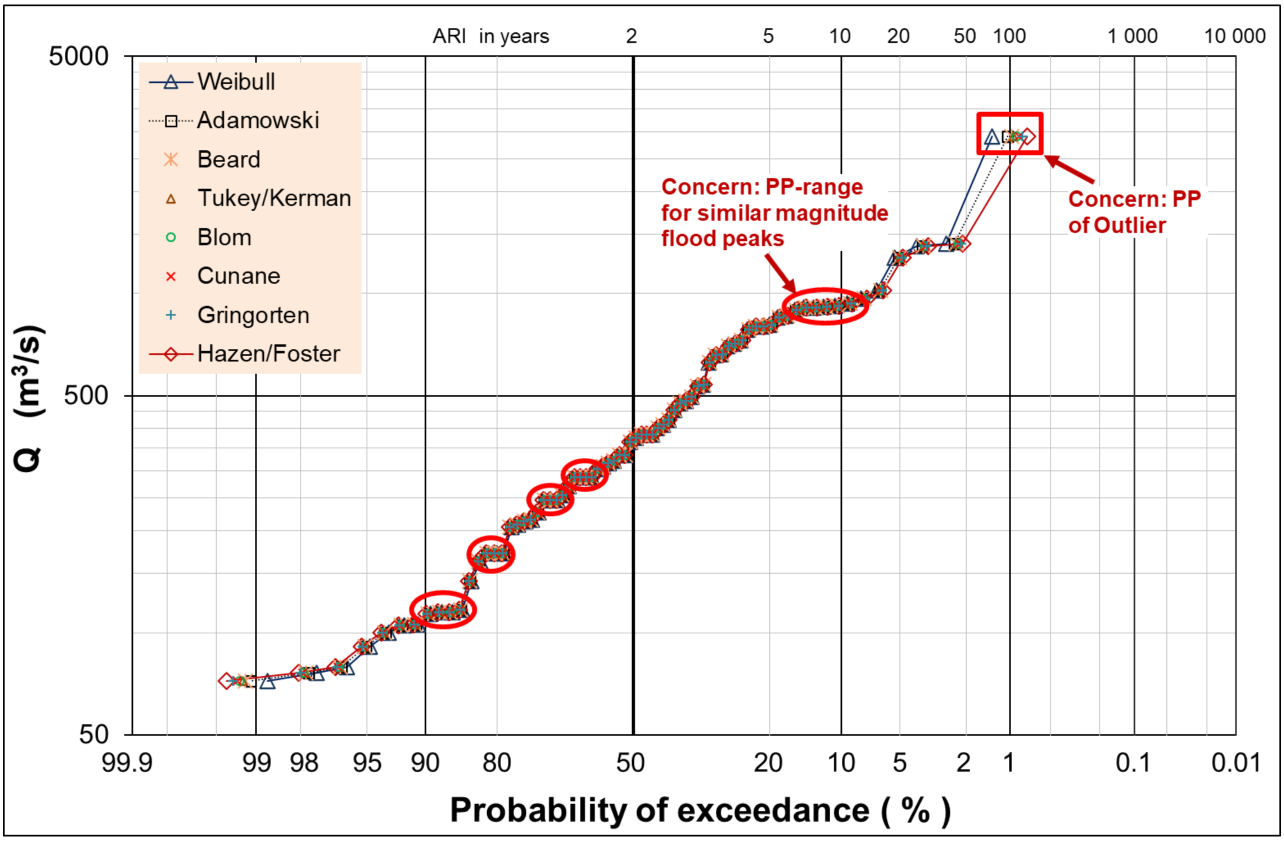


Figure 2-8: Concerns with current PPs, considering Woodstock Dam AMS

In support of the observations from Figure 2-8, the first 7 ranked positions for various PPs are compared in Table 2-2 (74 valid flood peaks in record). From ranked position 4 there is already, virtually, no difference between the AEPs and associated ARIs (in brackets) for the different PPs.

See *Appendix 2-A* for full table.

Table 2-2: AEPs for various PPs, for record length of 74 years (show first 7 rank positions)

Rank	AEPs (with ARI in years in brackets), according to:							
	Weibull	Adamowski	Beard	Tukey	Blom	Cunane	Gringorten	Hazen
1	0.0133 (75)	0.0101 (99)	0.0093 (108)	0.009 (111)	0.0084 (119)	0.0081 (124)	0.0076 (132)	0.0068 (148)
2	0.027 (38)	0.023 (43)	0.023 (44)	0.022 (45)	0.022 (46)	0.022 (46)	0.021 (48)	0.020 (49)
3	0.040 (25)	0.037 (27)	0.036 (28)	0.036 (28)	0.035 (28)	0.035 (29)	0.035 (29)	0.034 (30)
4	0.053 (19)	0.050 (20)	0.050 (20)	0.049 (20)	0.049 (20)	0.049 (21)	0.048 (21)	0.047 (21)
5	0.067 (15)	0.064 (16)	0.063 (16)	0.063 (16)	0.062 (16)	0.062 (16)	0.062 (16)	0.061 (16)
6	0.080 (13)	0.077 (13)	0.076 (13)	0.076 (13)	0.076 (13)	0.075 (13)	0.075 (13)	0.074 (13)
7	0.093 (11)	0.091 (11)	0.090 (11)	0.090 (11)	0.089 (11)	0.089 (11)	0.089 (11)	0.088 (11)
… …	… …	… …	… …	… …	… …	… …	… …	… …

It is evident that the current PP techniques can and should be improved, to provide more credible estimations of probabilities of extreme flood events.

2.5.2 Literature review observations and conclusions

Observations from the literature review indicates the following concise conclusions:

- The GEV proved to be superior to the LP3 and LN in the lower AEP range. In the higher AEP range the LP3 and LN produce more acceptable results than the GEV.
- Contradictory opinions exist on the use of AMS and POT in FFA.
- It would appear as if the transformation of data, that is abundantly present in existing FFA approaches, might not be the best solution to analyse flood peak data statistically.
- Due to the nature of the larger flood events in SA the gauged flow sites are mostly not statistically independent of one another. Consequently, the RFFA is not considered realistic in a statistical analysis approach.
- An investigation into appropriate record lengths, to produce more consistent estimates of lower AEPs, might prove useful to identify flow sites with suitable record lengths, for research application.
- The existing attempts to solve the PP-dilemma seem to be not very successful. Given the background of economics, the comments from De Haan (2007) might be reasonably valid, but the practice of ignoring the PP, regrettably, is also noted in some flood hydrology studies (e.g. Xiong Guo, Chen, Yin and Liu, 2018; Langat, Kumar and Koech, 2019; Ul Hassen, Hayat, and Noreen, 2019; Zhang et al., 2019). "Regrettably", because in FFA no population of AMS flood peaks exist, against which the observed data records can be assessed, while the PP provides the only visual GOF test against which the fitted distribution can be checked. Since the existing probability distributions also proved to be inadequate under certain conditions, a more realistic PP technique might prove to be a useful tool in FFA – in sync with the development of an improved statistical FFA approach.
- A distinct need exists for an investigation into an improved statistical FFA approach, preferably a proper bounded probability distribution, without having to normalise data/moments.

3 Data acquisition

3.1 Introduction

For the investigations into a new PP technique and an appropriate record length, primary flow sites were selected providing they have long verified records and are representative of the high variability in flow data experienced across SA, which is typical of a semi-arid region. Additional flow sites, referred to as secondary flow sites, were included in the research for a bounded probability distribution. The flow data used had been continuously subjected to a verification and validation process, to ensure that the data were reliable enough to be used with a reasonable degree of confidence.

Continuous water level data are recorded at a network of flow sites, consisting of around 210 reservoir- and 550 active river flow sites, by the Department of Water and Sanitation (DWS). These data and related information are freely available from the DWS hydrological database, HYDSTRA, which was also the source of water level- and streamflow data and -information for this research.

3.1.1 Flow sites in South Africa

Flow sites in rivers may consist of a natural control section (less common in SA) or an artificial control section (gauging weir – more common in SA), where a continuous record of stage (water level) is obtained. Stage can be translated to a discharge (flow) through a stage-discharge relationship, referred to as a stage-discharge rating. Gauging weirs are primarily designed to record low to medium flows (water resource and water quality purposes) – higher water levels will be recorded but the stage-discharge rating at a weir generally extends only to the weir structure limit. Gauging weir sites were also not selected with higher floods in mind and, therefore, not all sites lend themselves to accurate high flow calibration. More complex direct- (e.g. streamflow gauging) and indirect (e.g. stepped backwater) hydraulic approaches can be and are used, to extend stage-discharge ratings at suitable gauging weir sites.

Peak inflow data at dams were also considered, as a source of high flow data. Inflow hydrographs, at all dams with suitable information, were estimated through reservoir routing techniques, based on the continuity principle. Because of the large changes in storage volumes with relatively small changes in level, the accuracy of lower flows might be considered suspect. However, the methodology proved to be acceptably accurate with higher flows and, specifically, flood peaks. (DWS, 2006-2021)

The term 'flow site' is to be used for both weir- and dam sites, henceforth, for ease of reference.

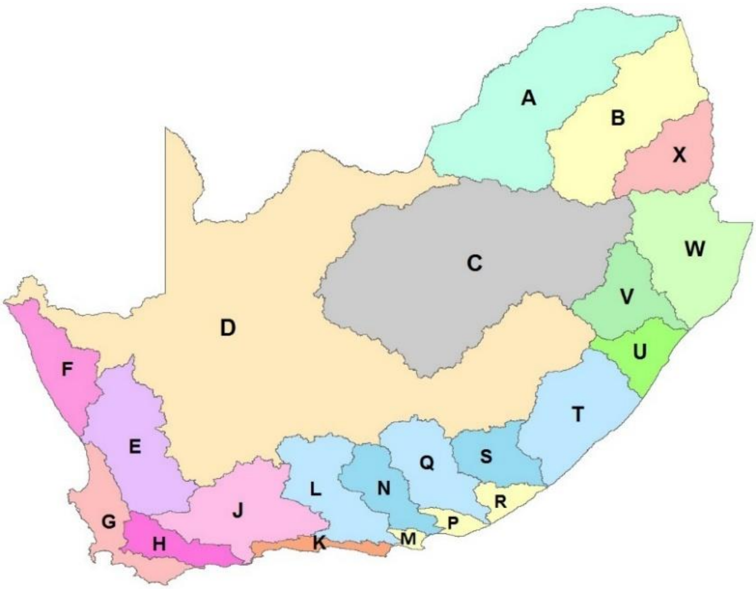
3.1.2 Flow sites identification

The flow sites in SA contains some 'intelligence' in the adopted numbering system. The country is divided into 22 primary drainage regions (Figure 3-1), which in turn is sub-divided into secondary-, tertiary- and quaternary drainage areas.

The allocated number of a flow site consists of 3 parts:

- the first two digits represents the secondary drainage region in which the flow site is situated
- the next digit indicates whether the flow site is a dam/reservoir (R) or a river flow site (H)
- the last 3 digits represent the sequential number of establishing the flow site

For instance: D7H005 will be the 5th (005) river flow site (H), established in secondary region D7.



Primary drainage regions in SA

For example, drainage regions D and C drain the Orange River and its largest tributary, the Vaal River (C).

The Orange River system drains 100% of Lesotho and 47% of SA.

(It also explains the 40% of primary stations coming from C and D)

Figure 3-1: Primary drainage regions of South Africa

3.2 Flood peak data considered

AMS of flood peak data were considered in this research. According to Mkhandi *et al*. (2005) and Karim *et al*. (2017) the POT/PDS flood peak approach does not really improve results, for ARIs higher than 10 years.

Srikanthan (2014) also concluded that, since the AMS gave the smallest bias in most cases of his research, the use of AMS in frequency analysis is preferred to POT/PDS. The largest part of SA can also be considered as semi-arid to arid. Consequently, it is a rarity to experience more than one sizeable independent flood peak, in any given year, across SA.

Primary flow sites, with long reliable records, were chosen for the initial research on the plotting position and record length. Secondary flow sites were chosen, primarily to augment the database for the research on an improved probability distribution, simultaneously covering areas not covered by the primary sites. Secondary flow sites were also used to validate results of the initial research.

3.2.1 Primary flow sites

Primary flow sites, having reliable and tested data, were selected for *Plotting Position*- and *Record Length* research purposes, consistent with the following criteria:

- **Representative spread across the country;** in choosing representative high flow sites, an attempt was made to cover as big an area, as possible, across the whole country.
- **Ensure ample flow diversity;** to avoid generating a database containing similar records, it was ensured that the choice of flow sites covered the following:
  - drier and wetter areas
  - a range from smaller to larger catchment areas (CA)
  - a variety of peak flow magnitudes – ranging from low to high
- **Reasonably long, verified flow records;** adhering to the following arbitrary criteria:
  - where possible, sites with minimum record lengths of 90 years
  - shorter lengths – preferably, of not less than 80 years – were considered to enable improved compliance to the previous criteria

Unfortunately, very few gauging weirs have long enough records AND suitable stage-discharge ratings to adhere to the criteria set above.

Thus, AMS peak inflow data at dams had to be considered. Few dams have sufficiently long records, themselves. Fortunately, most dams were built on, or close to, sites where there already existed a flow gauging weir for some time. In such a case, the AMS flood peak record from the flow gauging weir(s) was combined with the estimated inflow AMS flood peak record at the dam, to provide an extended AMS flood peak record at the dam.

Accordingly, the following 15 primary flow sites were identified (for all research purposes):

- 12 dams/reservoirs, spread across the country, with good reliable data
- 3 river gauging weirs, to improve areal coverage

The 15 flow sites have a combined database of 1556 years of AMS flood peaks, and the relevant metadata of the flow sites are depicted in Figure 3-2.

The observed AMS flood peaks, at all fifteen identified reservoir- and weir flow sites, are presented in Appendix 3-A and Appendix 3-B, respectively.

The metadata in Table 3-1 confirm that the set criteria were mostly met:

- Acceptable spread across the country, although the Western Cape (WC), Eastern Cape (EC), KZN and Limpopo provinces were underrepresented.
- Catchment Areas ranges from 37 km^2 to 361 512 km^2
- Flood peaks range from 0.6 m^3/s to 11 460 m^3/s
- Only 2 of the 15 flow sites have usable record lengths slightly shorter than 90 years.

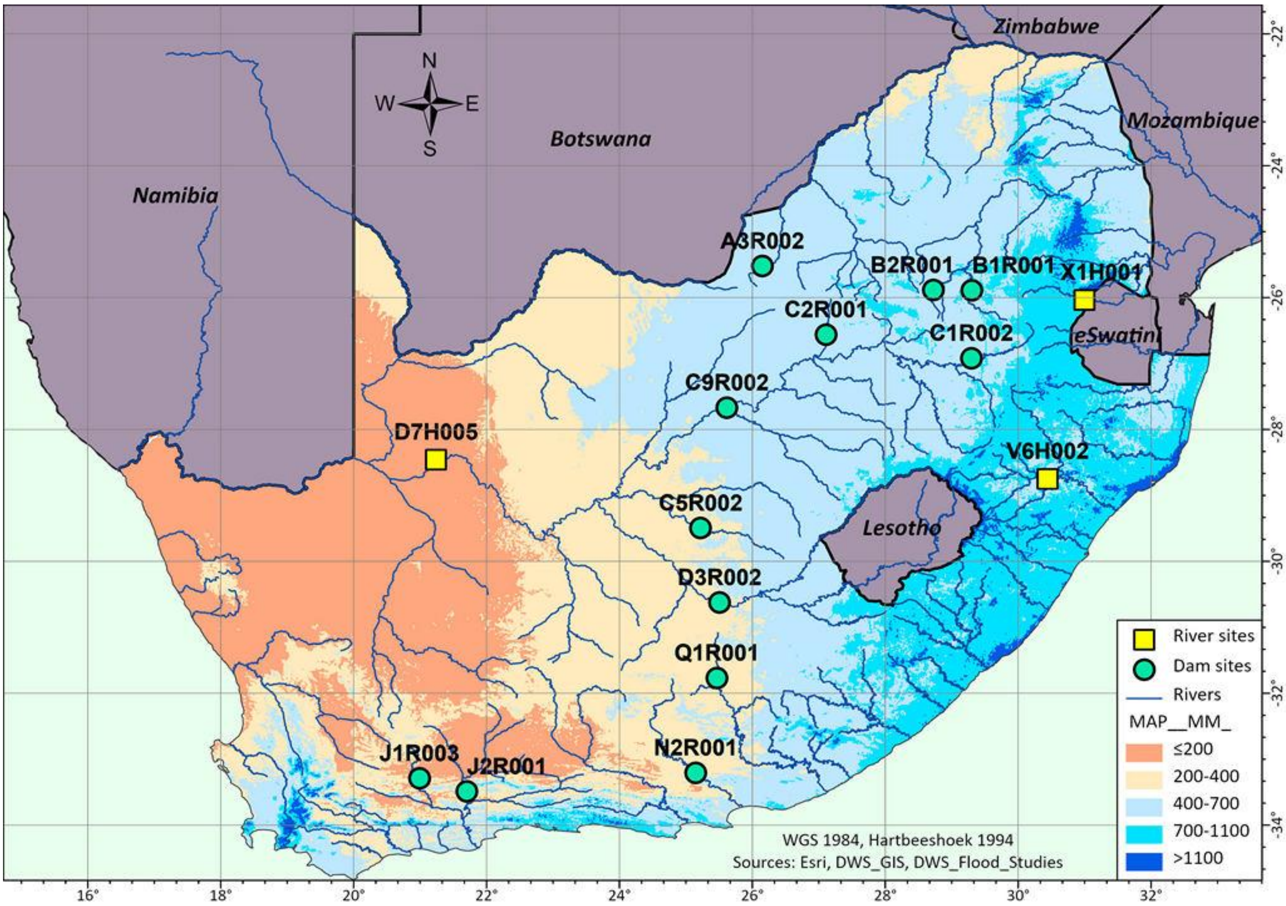


Figure 3-2: Distribution of primary flow sites

Table 3-1: Primary flow sites

Flow Site			Augment record with data from flow site: (1)	CA (km^2)	AMS Flood Peaks (m^3/s)		Record length (years)	
Number	**Dam/Site name**	**River**			**Lowest**	**Highest**	**Total**	**Usable**
Dam/Reservoir flow sites								
A3R002	Klein Maricopoort	Klein Marico	A3H001	1 157	0.6	506	112	110
B1R001	Witbank	Olifants	B1H001	3 579	3.5	2 565	114	112
B2R001	Bronkhorstspruit	Bronkhorstspruit	B2H001	1 244	2.2	995	114	114
C1R002	Grootdraai	Vaal (upper)	C1H001	7 982	42	2 275	114	114
C2R001	Boskop	Mooi	C2H001	3 480	2.1	112	114	114
C5R002	Kalkfontein	Riet	C5H001	10 265	3.2	9 800	106	106
C9R002	Bloemhof	Vaal (middle)	C9H006	108 360	85	6 340	110	108
D3R002	Gariep	Orange (upper)	D3H002	70 665	106	11 460	115	114
J1R003	Floriskraal	Buffels	J1H004	4 024	1.1	5 475	98	96
J2R001	Calitzdorp	Nels	-	37	0.6	191	99	89
N2R001	Darlington	Sondags	N2H002	16 820	11	5 090	96	96
Q1R001	Grassridge	Great Brak	-	4 326	13	805	94	94
River flow sites								
D7H005	Upington	Orange (lower)	D7H003	361 512	130	8 315	87	87
V6H002	Tugela Ferry	Tugela	-	12 862	38	2 438	92	92
X1H001	Hooggenoeg	Komati	-	5 503	8.3	2 481	110	110

(1) Flow gauging weirs close to an existing dam used to augment the inflow record at the dam. These gauging weirs, typically, existed prior to the construction of the dam. After construction of the dam these weirs either disappeared (submerged by dam) or remained to record spills and/or flow releases from dams (downstream) – in a very few cases the weirs are situated upstream of the dams and were retained to verify the inflows into the dams.

3.2.2 Secondary flow sites

The secondary sites were identified to cover areas, not covered by the primary flow sites and to:

- augment the primary sites for the research on developing an alternative probability distribution
- provide sites for validation of the results of the PP- and record length research

The choice of secondary sites was influenced by the outcome of the record length research (Chapter 4). The following secondary flow sites were identified:

- 26 reservoir sites (depicted in Table 3-2 and Figure 3-3), also spread across the country, not bounded by the criteria for the primary sites, but attempting to find suitable sites in areas and drainage regions not covered by the primary sites.

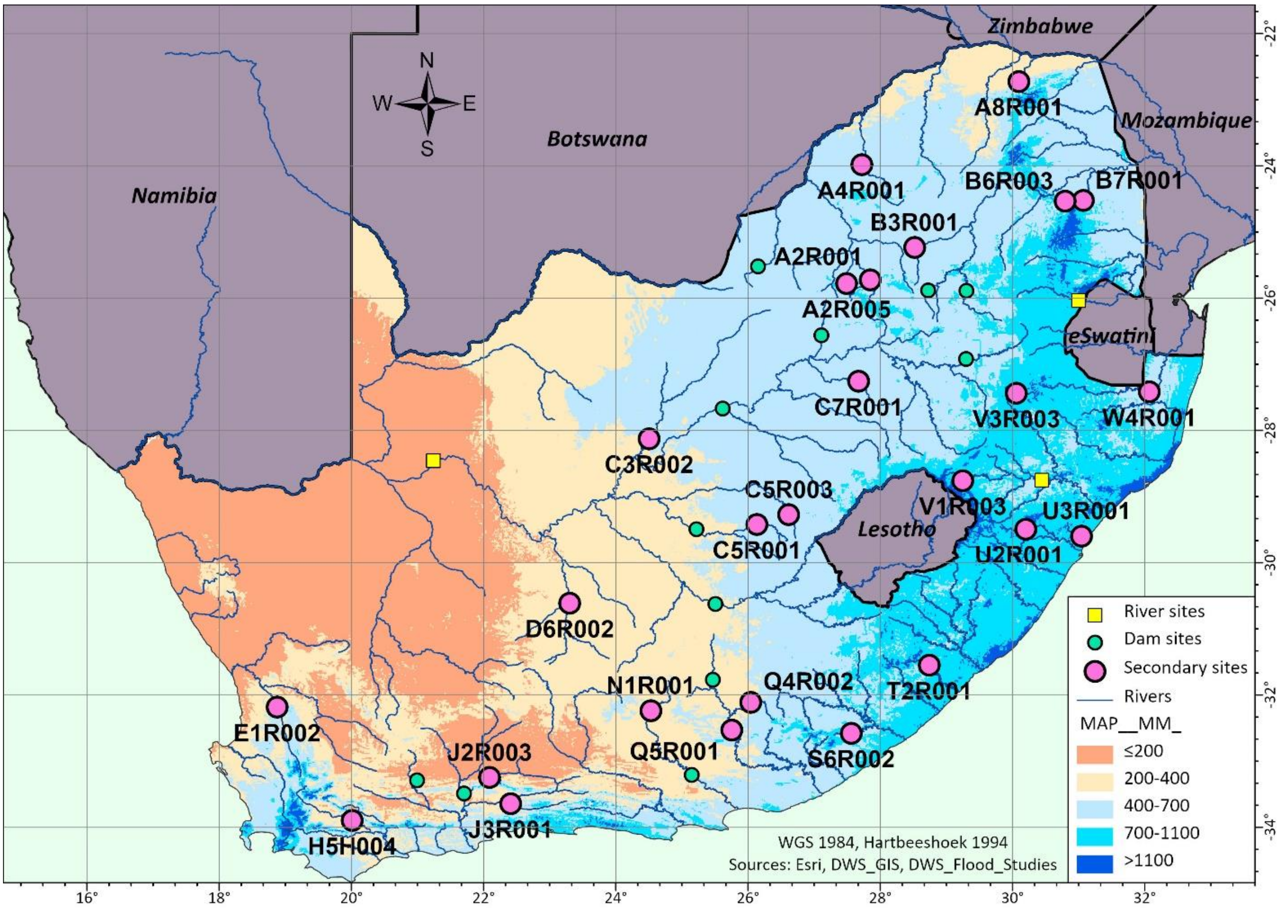


Figure 3-3: Secondary flow sites

It was anticipated to keep the record lengths 60 years and above, but 3 sites had record lengths shorter than 60 years – two sites with slightly shorter record lengths of 58 and 57 years, respectively, and one site with a record length of only 39 years. These sites were included to provide some data in areas where appropriate flow sites are limited. The observed AMS flood peaks, at these sites, are presented in Appendix 3-C and Appendix 3-D.

Table 3-2: Secondary flow sites, additional for research on bounded probability distribution

Flow Site			Augment record with data from flow site: (1)	CA (km²)	AMS Flood Peaks (m³/s)		Record length (years)	
Number	Dam/Site name	River			Lowest	Highest	Total	Usable
Flow sites								
A2R001	Hartebeespoort	Crocodile	A2H012/013	4 014	13	1 670	97	97
A2R005	Buffelspoort	Sterkstroom	-	119	0.9	398	84	84
A4R001	Mokolo	Mokolo	A4H005	4 319	1.2	1 291	57	57
A8R001	Nzhelele	Nzhelele	A8H001	830	2.0	2 795	86	74
B3R001	Rust de Winter	Elands	-	1 133	2.3	665	86	86
B6R003	Blyderivierspoort	Blyde	B6H004/5	2 169	19	1 615	69	69
B7R001	Klaserie	Klaserie	B7H004	164	0.2	915	69	68
C3R002	Spitskop	Harts	C3H007	26 730	13	2 760	96	96
C5R001	Tierpoort	Tierpoort	-	922	2.1	1 670	97	93
C5R003	Rustfontein	Modder	C5H003	937	9.2	2 670	101	101
C7R001	Koppies	Renoster	-	2 142	23	1 480	99	99
D6R002	Smartt Syndicate	Ongers	-	13 340	1.0	1 190	92	71
E1R002	Clanwilliam	Olifants	-	2 025	30	1 385	84	84
H5H004	Wolvendrift	Bree	H5H002	6 713	68	2 137	65	65
J2R003	Oukloof	Cordiers	-	154	0.2	196	88	87
J3R001	Kammanassie	Kammanassie	J3H001	1 525	1.6	2 755	106	105
N1R001	Nqweba	Sundays	-	3667	3.9	3 470	91	91
Q4R002	Kommandodrift	Tarka	Q4R001	3 623	4.8	802	90	90
Q5R001	Elandsdrift	GreatFish	Q7H001	16 854	9.0	3 888	113	99
S6R002	Wriggleswade	Kubisi	S6H002	447	1.3	881	68	66
T2R001	Mtata	Mtata	-	888	8.4	920	39	39
U2R001	Midmar	Mgeni	U2H001	928	11	1 565	63	63
U3R001	Hazelmere	Mdloti	U3H002	376	3.0	1 700	58	58
V1R003	Woodstock	Tugela	V1H002/026	1 149	41	2 140	84	75
V3R003	Zaaihoek	Slang	V3H005	620	13	382	71	71
W4R001	Pongolapoort	Phongolo	W4H002	7 800	31	16 450	71	71

(1) Flow gauging weirs close to an existing dam used to augment the inflow record at the dam.

3.3 Data considered for developing a new distribution

Numerous statistics were identified, from the AMS at each flow site, as potential parameters for a probability distribution. A condensed version of the identified statistics is presented as Table 3-3 (extended table in Appendix 3-E). Selected statistics, from these identified statistics, will constitute the input into a probability model to produce estimated flood peak values (Q_{est}) at corresponding AEPs.

Thus, to perform the research into the development of an improved probability distribution, known flood peak values, referred to as 'observed' flood peaks (Q_{obs}), at various AEPs are required. These values will be estimated from the results of the research on the plotting position technique and will, therefore, be presented in the chapter dealing with the research on a new probability distribution (Chapter 6).

Table 3-3: Statistics from flow site to be considered in distribution research

Site	n (used)	Full record [1]								Record, excluding Q_{max} [2]			
		Q_{min}	Q_{max}	Q_{mdn}	Q_{ave}	SD	g	kurt	Q_{gmn}	Q_{mdn}*	Q_{ave}*	SD*	g*
A3R002	110	0.60	506	19.5	37.4	63.3	4.696	28.939	15.4	18.0	33.1	44.7	2.916
B1R001	112	3.50	2 565	124	280	384	3.032	12.368	129	120	259	317	2.155
B2R001	114	2.20	995	79.0	144	188	2.690	8.125	74.9	78.0	137	171	2.637
C1R002	114	42.0	2 275	367	494	421	1.849	4.196	357	366	478	388	1.634
C2R001	114	2.10	112	18.0	23.2	22.0	1.755	3.403	15.2	18.0	22.5	20.4	1.584
C5R002	106	3.20	9 800	210	547	1 223	5.642	36.533	221	207	459	824	5.076
C9R002	108	85.0	6 340	719	1 216	1 251	1.507	2.075	710	712	1 168	1 154	1.226
D3R002	114	106	11 460	1 976	2 614	2 094	1.769	3.911	1 941	1 967	2 536	1 928	1.532
D7H005	87	130	8 315	1 294	1 730	1 690	1.762	3.504	1 070	1 285	1 653	1 540	1.566
J1R003	96	1.10	5 475	118	298	665	5.934	41.598	102	115	243	398	4.438
J2R001	89	0.60	191	11.0	20.4	29.1	3.573	15.546	11.0	11.0	18.5	22.8	2.878
N2R001	96	11.0	5 090	193	467	796	3.523	14.379	220	192	418	641	2.980
Q1R001	94	13.0	805	80.5	123	139	3.037	10.341	85.0	80.0	116	121	2.924
V6H002	92	38.0	2 438	897	949	492	0.761	0.665	800	896	933	469	0.634
X1H001	110	8.30	2 481	188	302	371	3.564	15.155	196	188	282	308	3.167
A2R001	97	13.0	1 670	181	278	301	2.133	5.468	164	180	263	267	1.749
A2R005	84	0.90	398	22.0	33.9	49.8	5.126	34.648	18.2	22.0	29.5	29.6	1.982
A4R001	57	1.17	1 291	58.0	181	281	2.509	6.254	68.7	53.5	161	240	2.495
A8R001	74	2.00	2 795	89.5	227	409	4.380	23.051	98.8	88.0	192	278	3.334
B3R001	86	2.30	665	27.0	58.5	96.5	4.084	20.109	29.9	26.0	51.3	70.7	3.134
B6R003	69	19.0	1 615	229	283	282	2.254	6.915	181	224	264	233	1.475
B7R001	68	0.20	915	38.0	89.4	137	3.795	19.482	35.4	38.0	77.0	92.6	1.829
C3R002	96	12.6	2 760	99.5	192	318	6.008	45.414	107	99.0	165	178	2.614
C5R001	93	2.10	1 670	34.0	101	226	4.807	27.223	36.9	33.5	84.0	156	3.718
C5R003	101	9.20	2 670	67.0	193	336	4.567	29.383	84.5	66.5	168	226	1.865
C7R001	99	23.0	1 480	170	251	264	2.413	6.764	164	167	239	234	2.144
D6R002	71	1.00	1 190	61.0	122	195	3.870	17.691	50.2	59.5	106	147	3.704
E1R002	84	30.0	1 385	307	408	309	1.459	1.758	314	304	397	291	1.415
H5H004	65	67.9	2 137	595	661	409	1.367	2.344	545	588	638	367	1.050
J2R003	87	0.20	196	5.80	17.5	32.9	3.557	14.049	5.98	5.65	15.5	26.8	3.358
J3R001	105	1.60	2 755	42.0	184	428	3.946	17.425	43.8	41.0	160	347	3.683
N1R001	91	3.90	3 470	106	243	455	4.760	28.956	95.4	101	207	301	2.785
Q4R002	90	4.80	802	93.5	142	147	2.514	7.866	93.3	89.0	135	130	2.345
Q5R001	99	9.00	3 888	236	452	639	2.986	10.487	229	226	417	538	2.513
S6R002	66	1.34	881	50.4	108	146	3.007	11.992	51.5	49.8	95.7	110	1.840
T2R001	39	8.40	920	235	285	237	0.883	0.139	177	225	269	215	0.721
U2R001	63	11.0	1 565	127	166	202	5.565	38.024	118	125	143	93.8	0.915
U3R001	58	3.00	1 700	69.0	174	279	3.521	15.780	72.3	69.0	147	192	2.119
V1R003	79	41.0	2 140	415	566	437	1.258	410	407	410	546	401	0.997
V3R003	71	13.4	382	88.0	102	73.9	1.793	85.5	81.3	85.5	97.9	66.3	1.586
W4R001	85	31.0	16 450	488	1 003	1 931	6.470	485	528	485	820	929	2.676

(1) Statistics include: Q_{max} - maximum flood peak; Q_{min} - minimum flood peak; Q_{mdn} – median flood peak; Q_{ave} - average flood peak; Q_{gmn} - geometric mean flood peak; SD - standard deviation (Q and SD in m^3/s)

(2) * indicates that largest value (Q_{max}) is omitted from dataset to determine statistic

The term 'geometric mean' is not commonly used in statistics, and it is defined by:

$$\mathrm{Q}_{gmn}(Q_1, \dots, Q_n) = \sqrt[n]{Q_1 * Q_2 * Q_3 * \dots * Q_n}$$ **Equation 3-1**

which, incidentally, is identical to the antilogarithm of the arithmetic mean of the logarithms of the data.

The amount of flow sites for the secondary and primary research, respectively 15 and 41, were mainly restricted by the length of the data-sets and by the verified quality of the data. The number of flow sites might be considered as relatively insufficient, at first glance. However, the wide spread of the flow sites across the country, which resulted into a significant diversity in flood peak magnitudes, as well as a variety in extents of flood peaks in data-sets, adequately compensate for the lack of more flow sites.

4 Record length

The research on suitable record lengths, summarised in this chapter, guided the choice of research sites to be included in the research to address the main objective, namely the development of an improved probability distribution. The research also required a vast number of statistical FFAs that resulted in numerous tables. Nevertheless, the research produced interesting findings and, consequently, a summary of this secondary objective will be presented in this chapter, while the bulk of the research is presented in Appendix 4-A.

4.1 Record length methodology

The purpose of this part of the research, was directed to establish a minimum data record length (RL), that could be used with a reasonable degree of confidence, to identify additional appropriate sites for conducting research on the primary objective. From Section 2.2.5, it was clear that related studies on this subject focussed more on the expected confidence in predictions, with different record lengths. The limited studies by Hattingh, *et al*. (2011) and Van der Spuy (2018), suggested minimum record lengths of between 30 and 40 years, which needed to be verified.

4.1.1 The research approach

An overlapping rolling window (R-W) technique was used to find an RL that, if extracted from any part of the existing record length, would practically give similar FFA results. R-W sizes in periods of 7-year multiples (7ym) were considered. The R-W concept is illustrated in Figure 4-1; with a dataset of record length n and an R-W of length m years (R-Wm), the dataset is partitioned into sub-datasets:

Number of sub-datasets = $n - m + 1$ **Equation 4-1**

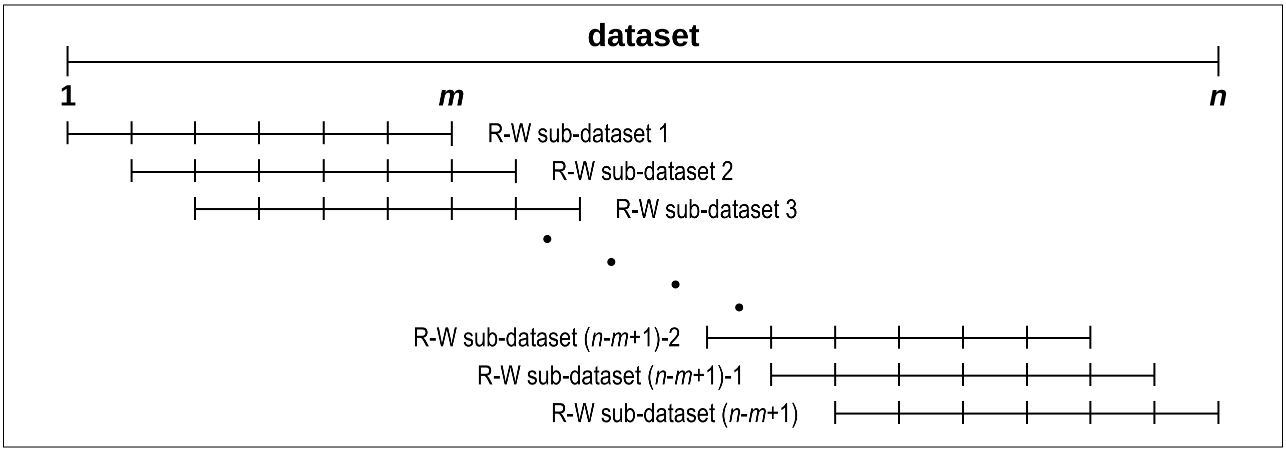


Figure 4-1: R-W concept, illustrating partitioned sub-datasets

Two approaches used set criteria (provided under the results in Section 4.2) to estimate RL1 and RL2, respectively. In the first approach variances in the FFA results of the sub-datasets of every 7ym R-W were considered for every AEP. In the second approach the variances in the averaged FFA results of all the 7ym R-Ws were considered at every AEP. The results were compared for the GEV and LP3 distributions. The methodology is explained in detail in Appendix 4-A1.

4.2 Record length analysis

4.2.1 *Prologue*

Selected primary sites, D3R002 and N2R001, are used to illustrate the results of the record length research, for the GEV an LP3 distributions, which are summarised in the following paragraphs.

The comprehensive results are provided in Appendix 4-A2.

4.2.2 *Extracted results at the two selected primary sites to estimate RL1*

The set criteria for RL1 were (see Section 4-A1.3 in Appendix 4):

- COV1 ≤ 0.25, for AEPs from 50% to 0.01% (2 ≤ ARI ≤ 10 000 year), on condition that
- COV1 ≤ 0.2, for AEPs from 10% to 0.5% (10 ≤ ARI ≤ 200 year)

<u>The results from flow site D3R002 are provided as follows:</u>

- It is shown in Table 4-1 that the set criteria, for RL1 were met for the GEV from R-W35 and for the LP3 from R-W42, which are visually depicted in Figure 4-2.
- Therefore, RL1 = 35 years for the GEV and RL1 = 42 years for the LP3.

Table 4-1: Selected 7ym R-W results for flow site D3R002

7ym R-W (yrs)	Stats	AEP (%)											
		50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
		GEV distribution: Flood peaks in m^3/s											
28	min1	1 355	2 828	3 676	4 545	5 521	6 079	6 604	7 251	7 707	8 093	8 532	8 835
	max1	3 655	5 884	7 423	8 907	10 821	12 313	13 796	15 989	18 018	20 799	25 101	29 197
	AVE1	2 476	4 265	5 499	6 730	8 401	9 721	11 102	13 042	14 608	16 271	18 638	20 574
	SD1	667	896	1 035	1 187	1 458	1 754	2 150	2 853	3 535	4 357	5 678	6 870
	COV1	0.27	0.21	0.19	0.18	0.17	0.18	0.19	0.22	0.24	0.27	0.30	0.33
35	min1	1 481	3 143	4 342	5 209	6 336	7 102	7 792	8 646	9 261	9 853	10 600	11 139
	max1	3 512	5 626	7 046	8 417	10 236	11 736	13 273	15 386	17 375	19 602	23 639	27 169
	AVE1	2 495	4 298	5 551	6 804	8 510	9 859	11 271	13 249	14 842	16 530	18 922	20 869
	SD1	599	801	903	992	1 134	1 298	1 545	2 035	2 547	3 189	4 255	5 239
	COV1	0.24	0.19	0.16	0.15	0.13	0.13	0.14	0.15	0.17	0.19	0.22	0.25
		LP3 distribution: Flood peaks in m^3/s											
35	min1	1 333	2 871	4 231	5 564	6 817	7 641	8 425	9 405	10 107	10 777	11 617	12 220
	max1	3 357	5 588	7 239	8 939	11 725	13 986	16 376	19 995	23 091	26 788	33 751	39 904
	AVE1	2 330	4 212	5 651	7 151	9 249	10 935	12 714	15 218	17 235	19 364	22 362	24 777
	SD1	566	832	984	1 125	1 347	1 584	1 918	2 547	3 185	3 975	5 271	6 456
	COV1	0.24	0.20	0.17	0.16	0.15	0.14	0.15	0.17	0.18	0.21	0.24	0.26
42	min1	1 448	2 938	4 152	5 460	7 002	8 135	9 253	10 706	11 783	12 838	14 199	15 201
	max1	3 166	5 428	6 963	8 528	11 210	13 530	16 039	19 663	22 646	25 967	30 781	34 750
	AVE1	2 360	4 275	5 725	7 220	9 286	10 923	12 626	14 984	16 849	18 787	21 460	23 569
	SD1	492	723	861	995	1 200	1 404	1 668	2 126	2 565	3 086	3 907	4 630
	COV1	0.21	0.17	0.15	0.14	0.13	0.13	0.13	0.14	0.15	0.16	0.18	0.20

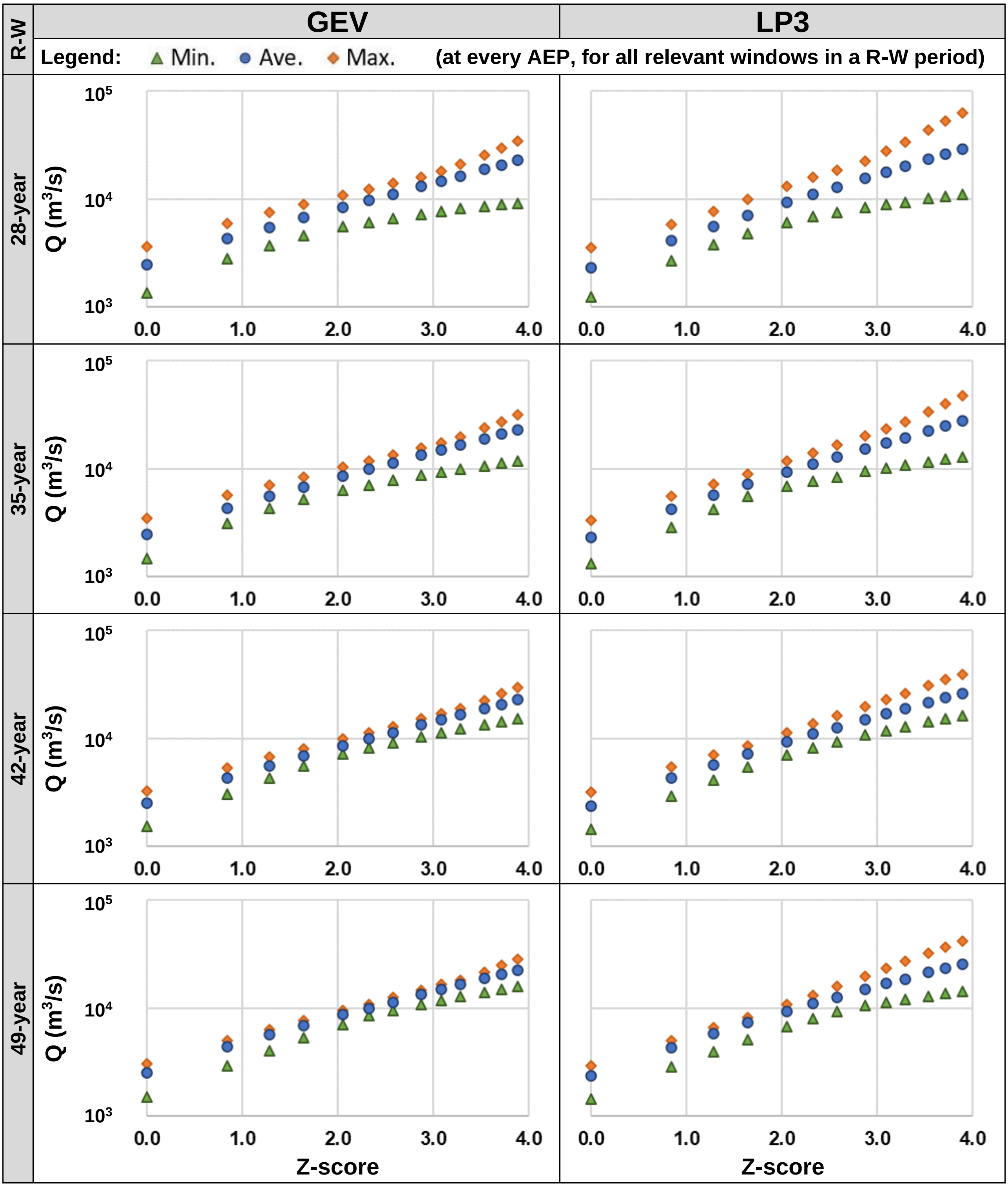


Figure 4-2: Results of selected 7ym R-Ws, for flow site D3R002

<u>**The results from flow site N2R001 are provided as follows:**</u>

- It is shown in Table 4-2 that the set criteria for RL1 were met for the GEV from R-W42 and for the LP3 from R-W77, which are visually depicted in Figure 4-3.
- Therefore, RL1 = 42 years for the GEV and RL1 = 77 years for the LP3.

Table 4-2: Selected 7ym R-W results for flow site N2R001

7ym R-W (yrs)	Stats	AEP (%)											
		50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
		GEV distribution: Flood peaks in m³/s											
35	min1	207	375	489	600	746	858	971	1 123	1 240	1 358	1 518	1 641
	max1	423	1 167	1 807	2 503	3 542	4 438	5 443	6 967	8 288	9 828	12 228	14 338
	AVE1	312	947	1 441	1 979	2 784	3 479	4 262	5 451	6 487	7 657	9 441	10 996
	SD1	45	230	386	561	832	1 074	1 353	1 793	2 188	2 646	3 367	4 016
	COV1	0.14	0.24	0.27	0.28	0.30	0.31	0.32	0.33	0.34	0.35	0.36	0.37
42	min1	236	545	783	1 042	1 425	1 754	2 123	2 680	3 162	3 704	4 521	5 227
	max1	384	1 179	1 808	2 502	3 538	4 431	5 440	6 975	8 310	9 815	12 105	14 097
	AVE1	306	973	1 497	2 073	2 942	3 700	4 559	5 877	7 034	8 351	10 377	12 158
	SD1	38	157	257	367	537	688	864	1 143	1 397	1 697	2 177	2 619
	COV1	0.12	0.16	0.17	0.18	0.18	0.19	0.19	0.19	0.20	0.20	0.21	0.22
		LP3 distribution: Flood peaks in m³/s											
70	min1	180	506	906	1 501	2 716	4 092	5 883	9 221	12 722	17 326	25 632	34 095
	max1	234	700	1 290	2 183	4 058	6 372	9 785	16 751	24 713	35 997	58 214	82 844
	AVE1	209	606	1 098	1 831	3 327	5 022	7 390	11 953	16 891	23 568	36 013	49 112
	SD1	13	51	112	220	483	829	1 375	2 571	4 019	6 162	10 575	15 651
	COV1	0.06	0.08	0.10	0.12	0.15	0.17	0.19	0.22	0.24	0.26	0.29	0.32
77	min1	189	540	974	1 617	2 818	4 106	5 832	9 003	12 278	16 529	24 081	31 664
	max1	227	666	1 206	2 006	3 680	5 609	8 330	13 623	19 395	27 246	41 969	57 536
	AVE1	212	604	1 081	1 783	3 194	4 770	6 946	11 077	15 483	21 363	32 154	43 334
	SD1	12	46	94	172	346	561	881	1 543	2 306	3 394	5 537	7 908
	COV1	0.06	0.08	0.09	0.10	0.11	0.12	0.13	0.14	0.15	0.16	0.17	0.18

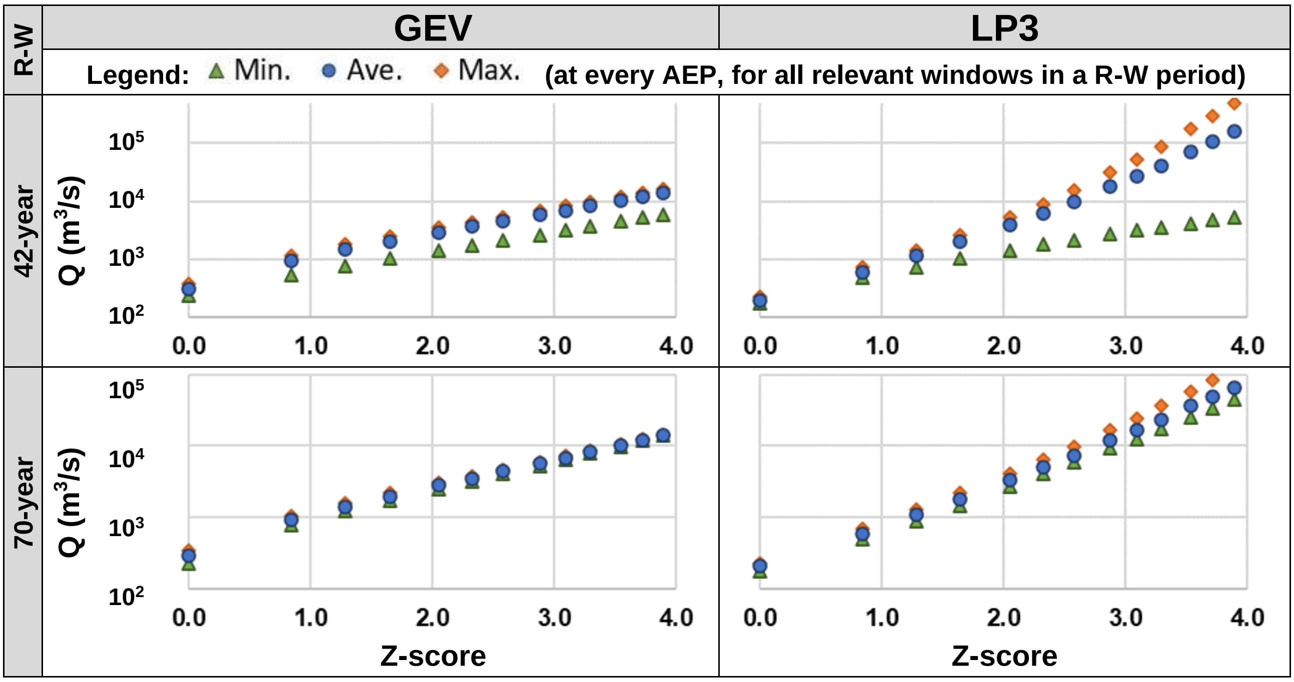


Figure 4-3: Results of selected 7ym R-Ws, for flow site N2R001

4.2.3 Extracted results at the two selected primary sites to illustrate RL2 estimation

The set criteria for RL2 were (see Section 4-A1.4 in Appendix 4):

- COV2 ≤ 0.05, for AEPs from 50% to 0.5% (2 ≤ ARI ≤ 200 year)
- COV2 ≤ 0.1, for AEPs from 0.2% to 0.01% (500 ≤ ARI ≤ 10 000 year)

<u>The results from flow site D3R002 are provided as follows:</u>

- The average results of all 7ym R-Ws, also visually displayed in Figure 4-4, are depicted in Table 4-3 (GEV) and Table 4-4 (LP3).
- If sub-datasets, shorter than R-W14 for the GEV and shorter than R-W21 for the LP3, are excluded from the dataset, the set COV criteria are met for RL2.
- Therefore, RL2 = 14 years for the GEV and RL2 = 21 years for the LP3.

Table 4-3: Combined results of 7ym R-W averages, for the GEV, at D3R002

7ym R-Ws (years)	**AEP (%)**											
	50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
	Average flood peaks (GEV: AVE1) per R-W (m^3/s)											
14	2 430	4 108	5 222	6 303	7 731	8 832	9 962	11 516	12 746	14 032	15 831	17 277
21	2 454	4 209	5 404	6 583	8 172	9 419	10 719	12 536	13 999	15 551	17 757	19 560
28	2 476	4 265	5 499	6 730	8 401	9 721	11 102	13 042	14 608	16 271	18 638	20 574
35	2 495	4 298	5 551	6 804	8 510	9 859	11 271	13 249	14 842	16 530	18 922	20 869
42	2 524	4 345	5 610	6 873	8 586	9 935	11 339	13 292	14 853	16 494	18 797	20 651
49	2 541	4 373	5 645	6 911	8 625	9 969	11 363	13 291	14 823	16 422	18 650	20 429
56	2 554	4 397	5 675	6 947	8 663	10 005	11 390	13 299	14 806	16 370	18 532	20 244
63	2 544	4 406	5 698	6 985	8 723	10 081	11 483	13 412	14 932	16 508	18 680	20 395
70	2 518	4 392	5 697	6 998	8 757	10 134	11 557	13 517	15 064	16 668	18 881	20 628
77	2 483	4 357	5 668	6 980	8 762	10 163	11 616	13 625	15 217	16 875	19 170	20 990
84	2 448	4 336	5 665	7 001	8 825	10 266	11 769	13 857	15 520	17 259	19 680	21 610
91	2 396	4 269	5 595	6 933	8 768	10 225	11 748	13 876	15 577	17 362	19 859	21 857
98	2 340	4 189	5 501	6 829	8 657	10 111	11 637	13 774	15 488	17 291	19 821	21 851
105	2 277	4 093	5 385	6 698	8 509	9 954	11 474	13 608	15 325	17 136	19 683	21 733
112	2 202	3 982	5 257	6 557	8 359	9 805	11 331	13 485	15 225	17 068	19 674	21 780
Descriptive statistics with sub-datasets shorter than R-W14 excluded												
min2	2 277	4 093	5 222	6 303	7 731	8 832	9 962	11 516	12 746	14 032	15 831	17 277
max2	2 554	4 406	5 698	7 001	8 825	10 266	11 769	13 876	15 577	17 362	19 859	21 857
AVE2	2 463	4 288	5 558	6 827	8 549	9 905	11 316	13 278	14 843	16 483	18 779	20 619
SD2	81	104	142	197	292	379	476	617	734	859	1 034	1 176
COV2	0.03	0.02	0.03	0.03	0.03	0.04	0.04	0.05	0.05	0.05	0.06	0.06

Table 4-4: Combined results of 7ym R-W averages, for the LP3, at D3R002

7ym R-Ws (years)	AEP (%)											
	50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
	Average flood peaks (LP3: AVE1) per Rolling-window (m^3/s)											
21	2 274	4 100	5 511	7 005	9 150	10 928	12 862	15 693	18 070	20 679	24 533	27 796
28	2 302	4 163	5 599	7 109	9 248	10 990	12 851	15 510	17 684	20 012	23 344	26 074
35	2 330	4 212	5 651	7 151	9 249	10 935	12 714	15 218	17 235	19 364	22 362	24 777
42	2 360	4 275	5 725	7 220	9 286	10 923	12 626	14 984	16 849	18 787	21 460	23 569
49	2 375	4 312	5 772	7 268	9 323	10 940	12 613	14 914	16 724	18 593	21 157	23 169
56	2 387	4 345	5 812	7 310	9 353	10 951	12 598	14 848	16 608	18 417	20 885	22 810
63	2 372	4 348	5 835	7 354	9 427	11 048	12 714	14 989	16 762	18 581	21 055	22 978
70	2 338	4 322	5 824	7 366	9 479	11 138	12 851	15 198	17 036	18 929	21 514	23 533
77	2 308	4 288	5 786	7 322	9 425	11 072	12 771	15 094	16 910	18 777	21 325	23 311
84	2 274	4 265	5 777	7 328	9 452	11 116	12 831	15 174	17 005	18 886	21 450	23 447
91	2 229	4 203	5 699	7 231	9 319	10 950	12 626	14 909	16 687	18 508	20 985	22 909
98	2 192	4 141	5 593	7 057	9 019	10 525	12 047	14 085	15 644	17 218	19 319	20 924
105	2 151	4 067	5 465	6 846	8 653	10 004	11 339	13 076	14 367	15 637	17 281	18 498
112	2 064	3 922	5 317	6 729	8 628	10 086	11 561	13 529	15 030	16 537	18 536	20 052
Descriptive statistics with sub-datasets shorter than R-W21 excluded												
min2	2 151	4 067	5 465	6 846	8 653	10 004	11 339	13 076	14 367	15 637	17 281	18 498
max2	2 387	4 348	5 835	7 366	9 479	11 138	12 862	15 693	18 070	20 679	24 533	27 796
AVE2	2 299	4 234	5 696	7 197	9 260	10 886	12 573	14 899	16 737	18 645	21 282	23 369
SD2	74	94	123	157	223	305	427	665	911	1 220	1 739	2 225
COV2	0.03	0.02	0.02	0.02	0.02	0.03	0.03	0.04	0.05	0.07	0.08	0.10

<u>The results from flow site N2R001 are provided as follows:</u>

- The average results of all 7ym R-Ws, also visually displayed in Figure 4-4, are depicted in Table 4-5 (GEV) and Table 4-6 (LP3).
- If sub-datasets, shorter than R-W28 for the GEV and shorter than R-W70 for the LP3, are excluded from the dataset, the set COV criteria are met for RL2.
- Therefore, RL2 = 28 years for the GEV and RL2 = 70 years for the LP3.

Table 4-5: Combined results of 7ym R-W averages, for the GEV, at N2R001

7ym R-Ws (years)	AEP (%)											
	50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
	Average flood peaks (GEV: AVE1) per Rolling-window (m³/s)											
28	318	924	1 390	1 893	2 635	3 269	3 976	5 040	5 958	6 986	8 538	9 879
35	312	947	1 441	1 979	2 784	3 479	4 262	5 451	6 487	7 657	9 441	10 996
42	306	973	1 497	2 073	2 942	3 700	4 559	5 877	7 034	8 351	10 377	12 158
49	304	992	1 537	2 139	3 053	3 855	4 769	6 180	7 426	8 852	11 057	13 009
56	302	966	1 493	2 075	2 958	3 733	4 617	5 982	7 187	8 567	10 702	12 591
63	298	941	1 452	2 017	2 877	3 632	4 493	5 825	7 003	8 352	10 441	12 292
70	294	919	1 416	1 968	2 807	3 546	4 390	5 696	6 853	8 179	10 236	12 059
77	292	908	1 399	1 943	2 771	3 500	4 334	5 624	6 766	8 076	10 107	11 908
84	289	899	1 385	1 924	2 746	3 470	4 298	5 580	6 716	8 019	10 040	11 833
91	289	902	1 389	1 929	2 751	3 473	4 297	5 571	6 697	7 987	9 983	11 749
Descriptive statistics with sub-datasets shorter than R-W28 excluded												
min2	289	899	1 385	1 893	2 635	3 269	3 976	5 040	5 958	6 986	8 538	9 879
max2	318	992	1 537	2 139	3 053	3 855	4 769	6 180	7 426	8 852	11 057	13 009
AVE2	300	937	1 440	1 994	2 833	3 566	4 399	5 682	6 813	8 102	10 092	11 847
SD2	10	32	54	80	124	167	222	314	403	514	698	873
COV2	0.03	0.03	0.04	0.04	0.04	0.05	0.05	0.06	0.06	0.06	0.07	0.07

Table 4-6: Combined results of 7ym R-W averages, for the LP3, at N2R001

7ym R-Ws (years)	AEP (%)											
	50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
	Average flood peaks (LP3: AVE1) per Rolling-window (m³/s)											
70	209	606	1 098	1 831	3 327	5 022	7 390	11 953	16 891	23 568	36 013	49 112
77	212	604	1 081	1 783	3 194	4 770	6 946	11 077	15 483	21 363	32 154	43 334
84	214	597	1 060	1 734	3 084	4 583	6 646	10 546	14 691	20 207	30 296	40 717
91	212	590	1 052	1 734	3 119	4 682	6 861	11 053	15 581	21 696	33 076	45 030
Descriptive statistics with sub-datasets shorter than R-W70 excluded												
min2	209	590	1 052	1 734	3 084	4 583	6 646	10 546	14 691	20 207	30 296	40 717
max2	214	606	1 098	1 831	3 327	5 022	7 390	11 953	16 891	23 568	36 013	49 112
AVE2	212	599	1 073	1 770	3 181	4 764	6 961	11 157	15 661	21 709	32 885	44 548
SD2	2	7	21	46	108	188	313	584	911	1 394	2 385	3 522
COV2	0.01	0.01	0.02	0.03	0.03	0.04	0.04	0.05	0.06	0.06	0.07	0.08

Note: Analogous results, for all primary flow sites, can be found in Appendix 4-L, with the resulting RL2 values summarised in Table 4-7 (see 4.2.4).

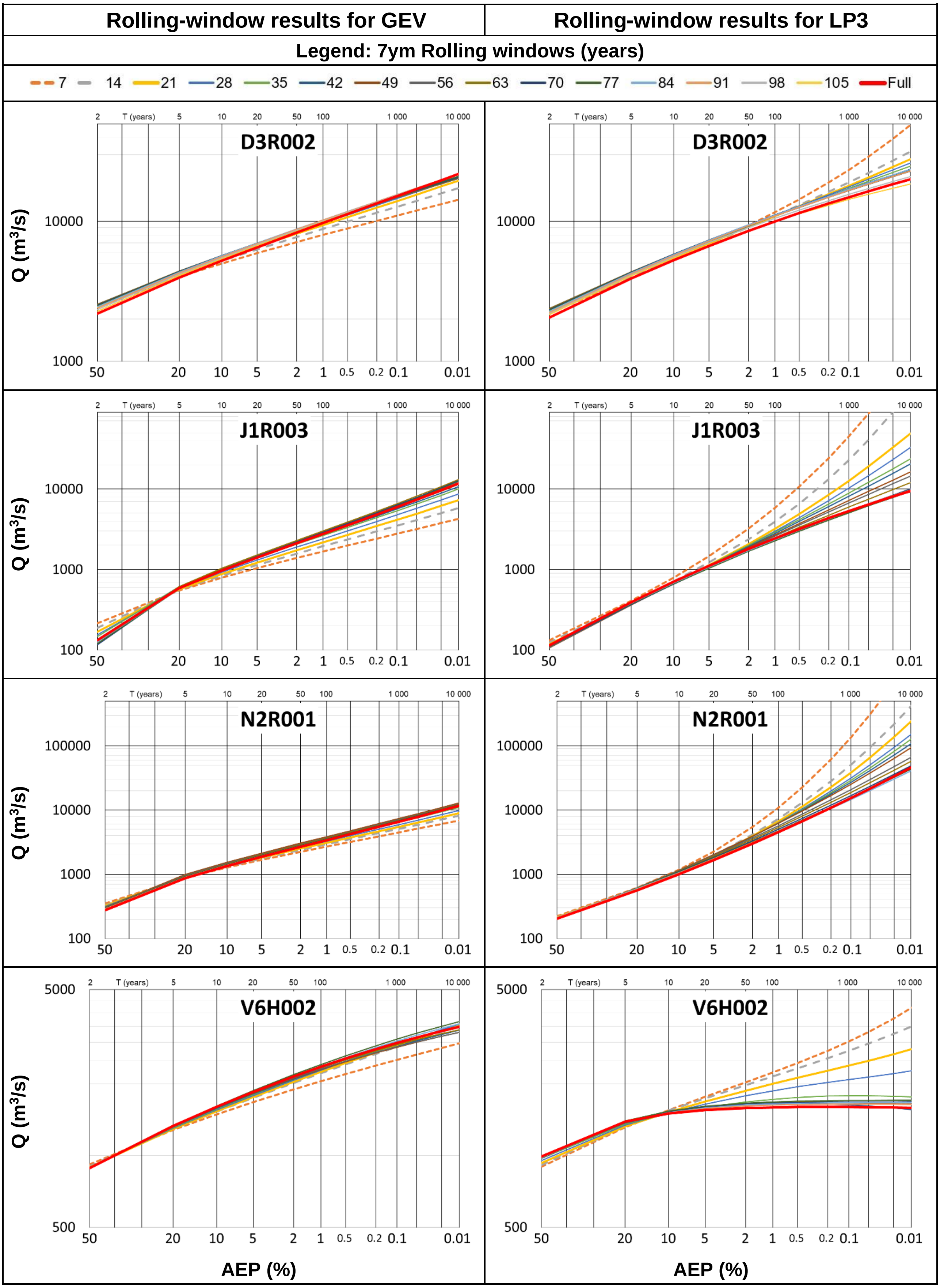


Figure 4-4: A visual comparison of all the 7ym R-W average results at selected flow sites.

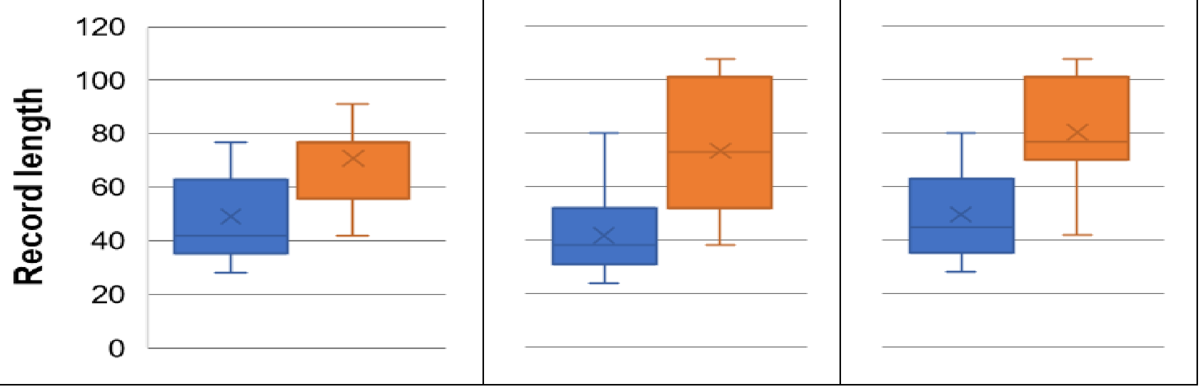
4.2.4 Summarised Record length results for all primary sites

The minimum record length results of all the primary flow sites, with the illustrated sites (D3R002 and N2R001) highlighted, were assessed and are summarised in Table 4-7, as follows:

- $RL1_P$, and $RL2_P$, the proposed values for RL1 and RL2 respectively, are shown for every flow site, for both the GEV and LP3 distributions.
- The recommended record length (RL_R), which represents the maximum of $RL1_P$ and $RL2_P$, is depicted in the last 2 columns, respectively, for the GEV and LP3 distributions.
- Percentiles for all the results are estimated and graphically shown as box plots in the last row of the table.

Table 4-7: Recommended record lengths

Flow site	$RL1_P$		$RL2_P$		RL_R	
	GEV	LP3	GEV	LP3	GEV	LP3
A3R002	77	77	80	108	80	108
B1R001	42	77	38	101	42	101
B2R001	49	77	31	45	49	77
C1R002	63	77	31	52	63	77
C2R001	70	77	59	73	70	77
C5R002	63	91	52	101	63	101
C9R002	35	70	31	59	35	70
D3R002	35	42	31	38	35	42
J1R003	63	91	52	80	63	91
J2R001	49	56	38	38	49	56
N2R001	42	77	45	87	45	87
Q1R001	42	49	45	73	45	73
D7H005	35	70	31	87	35	87
V6H002	28	42	24	52	28	52
X1H001	42	91	38	108	42	108
min	28	42	24	38	28	42
25%	35	56	31	52	35	70
50%	42	77	38	73	45	77
75%	63	77	52	101	63	101
max	77	91	80	108	80	108



4.3 Record length discussion

One of the secondary objectives, to determine the minimum record length that can be considered as adequate to provide reasonable results, was researched using two research procedures, both involving R-Ws. Literature on comparable research could not be found and, consequently, the results of this research could not be assessed against similar research. Although some literature (Hattingh, *et al.*, 2011; Van der Spuy, 2018) did propose minimum record lengths between 30 and 40 years, it was based on FFA studies on individual sites.

Although the number of sites used in this part of the research was relatively small, the focus was to use sites with long record lengths. Furthermore, a representative spread across the country, covering dry (low rainfall) and wet (higher rainfall) parts, as well as small and large catchment areas, were considered as being more important than simply the number of sites used. It did, however, limit the conclusions on an unexpected observation that did not affect the objective of this part of the research but might be worthwhile to research further. The observation, evident on Figure 4-4, Figure 4-2, and Figure 4-3, relates to the so-called confidence intervals. Currently, outliers have a large impact on the estimation of confidence intervals. It seems as if the R-W concept might provide an alternative way of estimating confidence intervals, considering the increasingly relatively narrow intervals as the record lengths increase, between the maximum and minimum results of the 7ym R-Ws, as depicted on the above-mentioned figures; especially in the case of the GEV.

As might have been expected, both explored research avenues confirmed that the longer the record length, the more reliable the results. From Table 4-7 it was concluded that sites, with record lengths of 63 years (75 percentile with only 2 results exceeding that), or longer, should be used for reasonably acceptable results in research. In exceptional cases, for instance where there were clearly gaps in the coverage across South Africa, flow sites with shorter record lengths may be considered for research; preferably not shorter than 35 years.

5 Plotting position

In this chapter the research is aimed at producing a more consistent PP technique that will also reflect outliers more realistically. The outcome of this research will also serve as a benchmark for the primary research of developing a new probability distribution.

Due to the vast amount of generated data/information only pertinent concise results, at selected flow sites are provided in the document itself. Where deemed necessary, supplementary corresponding tables and figures are presented in the appendices. (Comprehensive results are available, on request, in electronic format).

5.1 Plotting position methodology

In this section the research approach followed, to deal with the PP dispute, is deliberated. An alternative PP technique was envisaged, which will not replicate shortcomings associated with existing PP techniques. The main shortcomings of existing PP techniques relate to the plotting of similar magnitude flood peaks at different probabilities and not adequately addressing PPs of outliers (see Section 2.5.1).

5.1.1 Prologue

This research is not about partaking in the many lengthy theoretical mathematical debates as to which PP technique is superior to the others. These theoretical arguments may very well be valid for a 'population' of flood peaks. The reality, however, is that available flood peak data records cannot necessarily be considered as representative samples of the 'population' of AMS flood peaks. The analyst is, consequently, left with the challenge to attempt a good-as-possible analysis with relatively very short samples (data records) of the population.

Klemeš (1987) is rather cynical in his view on the science of hydrology and FFA, but he also made a very valid point when stating that FFA was not motivated by science, but the need to make decisions demanded by society.

The reality is that an FFA should never be done without prudently considering all available data and information available – keeping in mind the relevant advice from Wheeler (2011b): "*This limitation on what we can obtain from a collection of data is inherent in the statistics themselves, and must be respected in our analysis of the data… … …interpret your data in their context*."

Considering the above, it was recognised that the PP-issue demands a more pragmatic approach, rather than a purely theoretical one.

5.1.2 Alternative approaches considered for more pragmatic PP technique

A more realistic PP was envisaged to address the concerns highlighted in Section 2.5.1. Various alternative methodologies were considered in the development of a more practical PP technique, but it was obvious that introducing variables, like catchment- and rainfall characteristics would not be of any benefit to solve the issues depicted in Figure 2-8. Neither would the inclusion of basic statistical principles like record lengths, steepness of flood peak distribution (or the 'probability-slope' between lowest and highest peaks) or magnitude of flood peaks, be able to solve these issues.

The only plausible option that remained, seemingly, was to reinvestigate potential statistics options. Wheeler (2011b) cautioned against merely applying statistics in a probability model. He stated that the dataset should be homogeneous before statistical parameters, intended for a probability model, can be estimated sensibly. He considered it of paramount importance that one should identify suspect data and examine it for evidence of lack of homogeneity.

Homogeneity is an indication whether sample data sets are similar and representative of the population. The application of homogeneity tests in most other fields, like economic studies, demography and population studies, etc., usually have the advantage of working with a well-established worldwide population. Unfortunately, in FFA the luxury of having a population does not (yet) exist. Furthermore, populations will most probably tend to be more site- or area-specific, rather than global. A practical suggestion would be to (visually) inspect PPs for a noticeable trend, mimicking a probable distribution, which possibly will be an indication that the data can be assumed to be homogeneous.

In the field of FFA issues like low- and high outliers, in relative short data records, can render a reasonable homogeneous dataset to appear completely non-homogeneous. Using statistics like skewness and kurtosis, that depict the tails of a probability distribution, was thus considered as a possible resort to deal with the outlier-issue.

Wheeler (2011b) expressed his concerns on the use of skewness- and kurtosis-statistics (Section 2.4), but also, coincidentally, provided a guideline of how to consider outliers in a data set, by stating that the limitation on what can be obtained from a dataset is inherent in the statistics themselves, and that it must be respected in the analysis of the data. Consequently, it was considered to approach the PP problem from a different perspective, in studying the possibility that outliers themselves may provide a possible solution.

Thus, a technique had to be found that could provide a quantifiable value, comparable to the existing PP-values, which could also incorporate the relative probabilities, associated with the relative magnitudes of the AMS flood peaks. The only outlier-identification technique that could achieve this was the Z-score statistic (see also Section 2.4), which was selected to be used in this research.

5.1.3 Research approach

Various scenarios, that involved multiple linear regression analysis (MRA) of combinations of the various Z-scores, were investigated to develop an improved PP technique. The Z-scores considered, include the *Z*-score of the original (untransformed) AMS flood peaks (Z_Q), the *Z*-score of the log-transformed data (Z_{logQ}) and the associated *Z*-score from one of the existing PP techniques. At the outset the Z-score for the Weibull PP ($Z_{Weibull}$) was selected, since the equation is simple, it is widely used and, amongst most of the common PP techniques, it is most conservative towards risk. It is hypothesised that any of the other common PP techniques could have been used (recognising, of course, that different PPs will produce different coefficients) and the soundness of this assumption will be established under the results (in Section 5.2.4).

To synchronise the Z-scores determined from the AMS, with that of the Weibull PP, the same preferred ranking order as proposed in Section 2.2.6 was used and the Z-scores for the Weibull PP probabilities were estimated by using the standard normal distribution – also referred to as the Z-distribution.

The PDF for the standard normal random variable, z, is given by:

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{z^2}{2}\right)} \quad \text{for } -\infty < z < \infty \qquad \textbf{Equation 5-1}$$

and the cumulative distribution function (CDF) of the standard normal distribution, by:

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2}\, dx \quad \text{for } -\infty < z < \infty \qquad \textbf{Equation 5-2}$$

The variables, that were considered for the MRAs, are defined in Table 5-1.

Table 5-1: Variables considered for the MRAs

Independent variable		Description	Comment
1	$Z_{Q_i} = \frac{Q_i - \overline{Q}}{S_Q}$	*Z*-score of the original AMS flood peak data	where *i* indicates the *i*th order statistic, in the ranked dataset
2	$Z_{logQ_i} = \frac{\log Q_i - \overline{logQ}}{S_{logQ}}$	*Z*-score of the Log-transformed AMS flood peak data	
3	$Z_{i.Weibull}$	*Z*-score of the Weibull PP (AEP) (= 0 – the inverse of Φ(Z) of the Weibull PP)	
Dependent variable		**Description**	
	$Z_{i.PP}$	*Z*-scores of the corresponding AEP, according to the fitted probability distribution(s) – see also Table 5-2	

Three MRA scenarios were investigated, namely:

- MRA1, where independent variables 1 and 3 were considered
- MRA2, where independent variables 1 and 2 were considered
- MRA3, where independent variables 1, 2 and 3 were considered

Each MRA scenario contained 3 sequential regression attempts (referred to as a sequent), related to a progressively improved dependent variable, which is defined in Table 5-2.

Table 5-2: Estimation of dependent variables for sequential MRAs (Sequents)

Sequent	Dependent variable	Description
1	$Z_{i.PP_1}$	Z-scores of the corresponding AEPs of the AMS – obtained from a single probability distribution [(1)] that visually fit the Weibull PPs the best. An exception is made for the GEV, since it offers a poor fit at the higher AEPs. If the GEV is considered for the lower AEPs, it is combined with another probability distribution for the higher AEPs.
2	$Z_{i.PP_2}$	Z-scores of the corresponding AEPs of the AMS, obtained from a single or combination of probability distributions [(2)], that fit the PPs of Sequent 1 the best.
3	$Z_{i.PP_3}$	Z-scores of the corresponding AEPs of the AMS, obtained from a single or combination of probability distributions [(2)] and the regression fit from Sequent 2, that fit the PPs of Sequent 2 the best.

(1) The LN, LP3 and GEV distributions were considered.
(2) Using the MLVA from Gericke and Du Plessis (2012)

The combinations of independent- and dependent variables for the MRA-scenarios are depicted in Table 5-3.

The flow-diagram in Figure 5-1, depicts the 3 sequential attempts of scenario MRA3 – the flow-diagrams for MRA1 and MRA2 will be similar, but with different independent variables, as designated in Table 5-3.

Table 5-3: Combinations of independent- and dependent variables for different MRAs

MRAs	Description	Dependent variable	Independent variables		
MRA1.1	MRA1, Sequent 1	$Z_{i.PP_1}$	Z_{Q_i}		$Z_{i.Weibull}$
MRA1.2	MRA1, Sequent 2	$Z_{i.PP_2}$			
MRA1.3	MRA1, Sequent 3	$Z_{i.PP_3}$			
MRA2.1	MRA2, Sequent 1	$Z_{i.PP_1}$	Z_{Q_i}	Z_{logQ_i}	
MRA2.2	MRA2, Sequent 2	$Z_{i.PP_2}$			
MRA2.3	MRA2, Sequent 3	$Z_{i.PP_3}$			
MRA3.1	MRA3, Sequent 1	$Z_{i.PP_1}$	Z_{Q_i}	Z_{logQ_i}	$Z_{i.Weibull}$
MRA3.2	MRA3, Sequent 2	$Z_{i.PP_2}$			
MRA3.3	MRA3, Sequent 3	$Z_{i.PP_3}$			

The 3 MRA scenarios, each with 3 sequents, resulted in a total of 9 MRAs being carried out. The general equation, resulting from an MRA, was depicted as follows:

$$Ze_{i.PP_y} = C_{1_{xy}} * Z_{i.Weibull} + C_{2_{xy}} * Z_{Q_i} + C_{3_{xy}} * Z_{logQ_i} + C_{4_{xy}} \qquad \textbf{Equation 5-3}$$

Where x signifies a specific MRA-scenario and y the sequent in a specific MRA-scenario, with $Ze_{i.PP_y}$ the estimated dependent variable from the relevant MRA to be compared to $Z_{i.PPy}$.

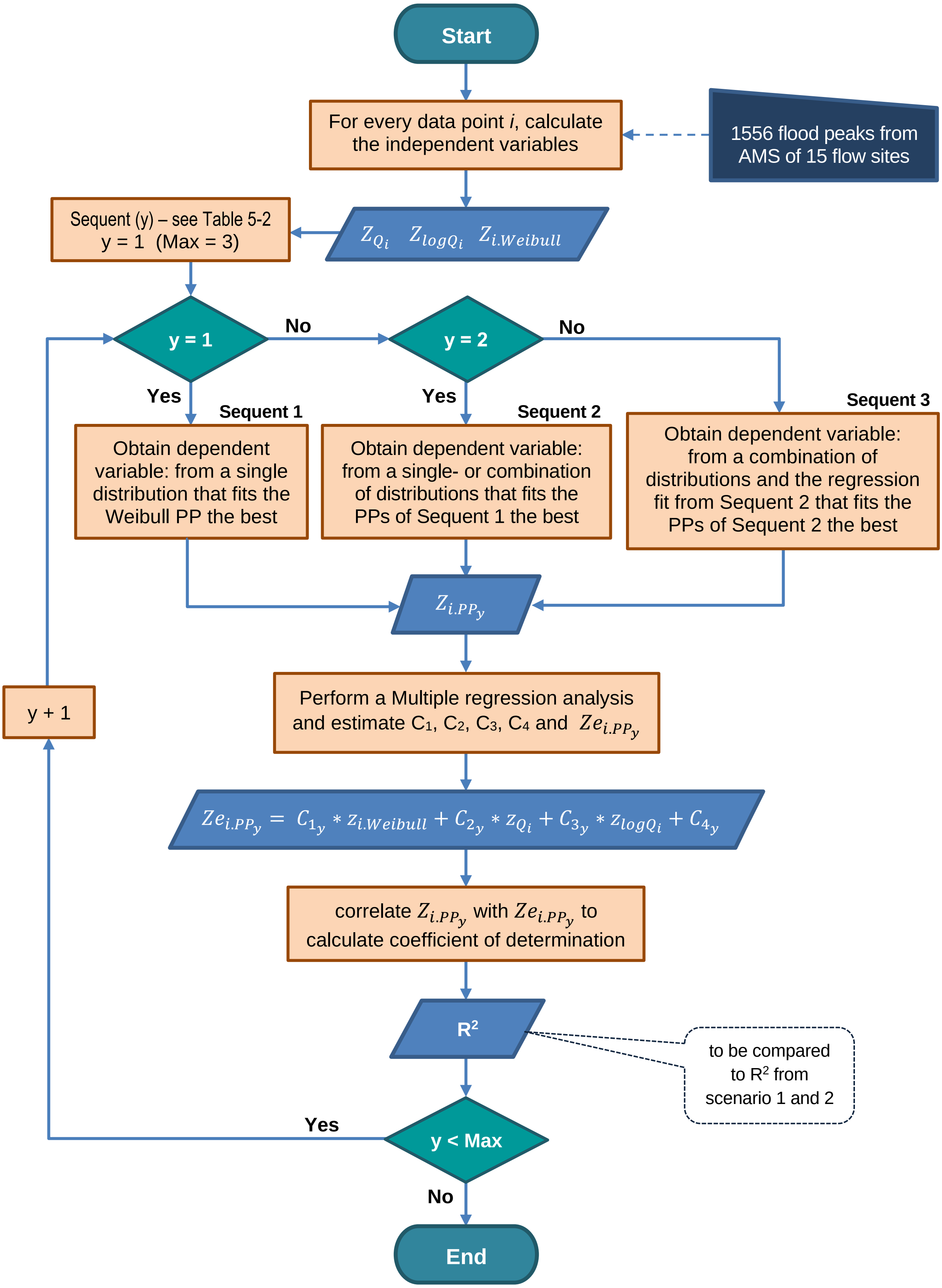


Figure 5-1: Flow chart illustrating the PP research methodology, using scenario MRA3

5.1.4 Assessment norms for the proposed PP technique

The coefficient of determination (R^2) was applied to compare the outcomes of the different scenarios with each other, supported by a statistical significance test, to identify the scenario that produced the best results.

The MRA models use more than one independent variable in a scenario. To perform a statistical significance test, two null hypotheses (H0) were thus specified:

- $H0_F$ suggesting there is no statistical significance (correlation) between the scenario and the dependant variable, with the F-value providing the probability of the hypothesis test, and
- $H0_P$, suggesting a particular independent variable displays statistical insignificance towards the dependent variable, with the P-value providing the probability of the hypothesis test.

Thus, if the F-value or P-value is less than the significance level, which is generally 0.05, evidence exists to reject the null hypothesis. Ganesh and Cave (2018) suggests the following guidelines in terms of evidence to reject the null hypothesis:

- P-value < 0.001 indicate very strong evidence of significance
- P-value < 0.01 indicate strong evidence of significance
- P-value < 0.05 indicate moderate evidence of significance
- P-value < 0.1 indicate weak evidence of significance
- P-value ≥ 0.1 indicate insufficient evidence of significance

The intended research, relating to the primary objective of an improved probability distribution, relied on an improved PP technique, to provide data intended for the primary objective research. The results emanating from the proposed PP technique were weighed up against that of the existing Weibull PP technique, to assess if the concerns expressed earlier (Section 2.5.1) were sufficiently addressed and whether the proposed PP technique proved to be an improvement, compared to the existing PP techniques.

5.2 Plotting position analysis

In the first part of this section the proposed Z-scores are appraised, which is followed by the results of the ensuing MRAs. The research used different combinations of the following Z-scores: Z_Q, Z_{logQ} and $Z_{Weibull}$. The combinations of the independent variables, considered for the 3 MRA scenarios, are repeated here for ease of reference:

- MRA1: independent variables Z_Q and $Z_{Weibull}$
- MRA2: independent variables Z_Q and Z_{logQ}
- MRA3, independent variables Z_Q, Z_{logQ} and $Z_{Weibull}$

The hypothesis that any of the common PPs, other than the Weibull (Section 5.1.3), could also have been used was also explored and the conclusion is presented at the end of this section.

5.2.1 Z-score appraisal

In Figure 5-2 the *Z*-scores for the AMS and log transformed AMS are depicted, with the *Z*-scores of the Weibull PP, for all the primary sites.

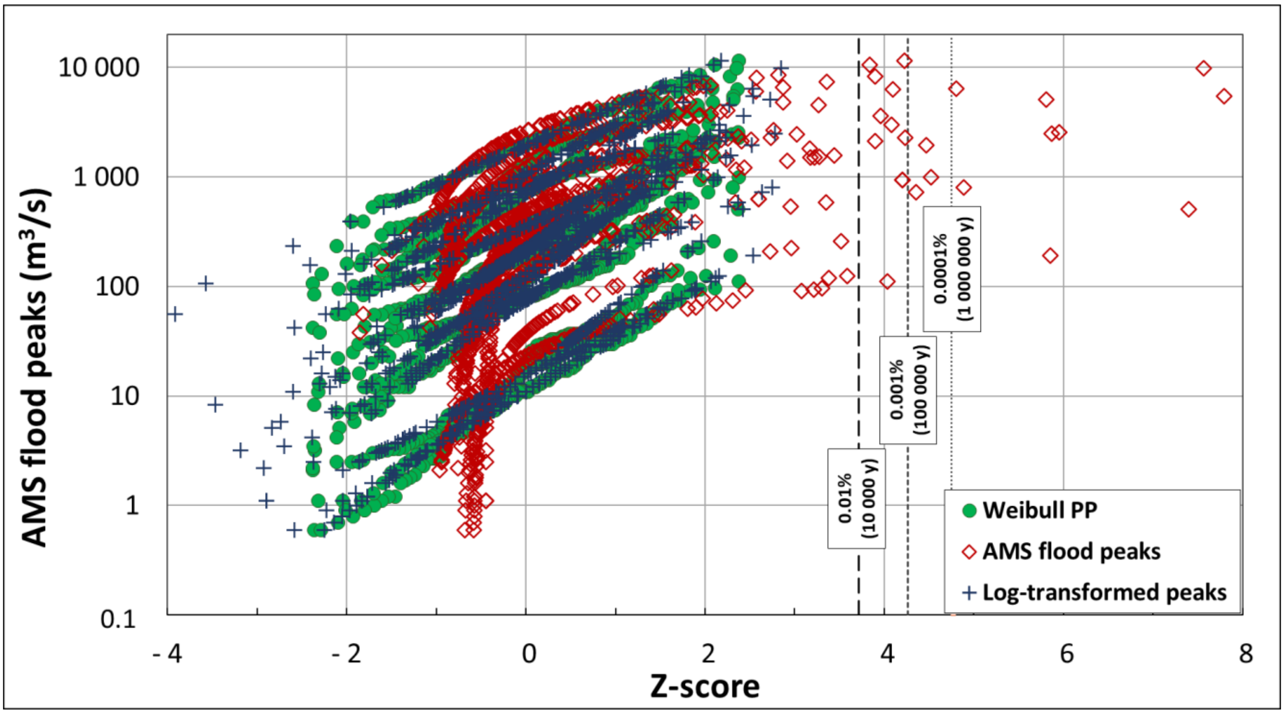


Figure 5-2: Various Z-scores for the AMS at all primary flow sites

The Z_Q yields AEPs that, in comparison to existing distributions, were considered excessively low. Several flood peaks had Z-scores higher than 3.719 (AEP < 0.01%; ARI > 10 000 years); and about 50% of these even have AEPs of less than 0.0001% (ARI > 1 000 000 years).

Hence, Z_{logQ} were also considered, although it tends to be very similar to the *Z*-scores of the existing PPs in that outliers are not adequately addressed. The caution expressed by Wheeler (2011b), against the practice of transformation of data that can hide the fact that data are not sufficiently homogeneous and, consequently, prevent the accurate treatment of problematic data, was taken into account. The only reason for considering the Z_{logQ} was to determine if it could be used to curtail the excessively low Z_Q values.

5.2.2 MRA results

Two MRA approaches were followed in determining coefficients C_1, C_2, C_3 and C_4 (see Equation 5-3, Section 5.1.3), namely (1) a single MRA considering all 1556 flood peaks as one record (total record approach) and (2) individual MRAs for every flow site and then taking the averages of the coefficients of each analysis (average approach).

Initially the average approach, mostly, resulted in the most favourable coefficient of determination, whilst the total record approach yielded the most favourable coefficient of determination when the process was repeated with adjusted Z_{PP} values (as described in Section 5.1.3 and illustrated in Table 5-2 and Figure 5-1, three iterative sequential MRAs were used, where the results of a sequence was used to improve Z_{PP} for the next sequential MRA). The results, relating to the inclusion of the Weibull PPs, are depicted in Table 5-4. In addition, all the MRA results are depicted graphically in Figure 5-3.

Both MRA2.3 and MRA3.3 produced very good results. However, it is hypothesised that any existing PP (in this case Weibull) will produce acceptable results as the AMS flood peak datasets get increasingly larger, approaching a sensible population-size. All the PPs will hence yield similar results, which suggests that the Weibull PP (or any other PP) can be used on its own, if the sample data record is sufficiently large.

Table 5-4: MRA, including Weibull PP: Parameters and coefficients of determination

$$Z_{i.PP_y} = C_{1_{xy}} * Z_{i.Weibull} + C_{2_{xy}} * Z_{Q_i} + C_{3_{xy}} * Z_{logQ_i} + C_{4_{xy}}$$

Multiple Regression Analysis			Coefficients				
			C_1	C_2	C_3	C_4	R^2
MRA1	MRA1.1	average	1.0084	0.0492	-	0.0050	0.9820
	MRA1.2	average	0.9587	0.1229	-	0.0011	0.9854
	MRA1.3	total	0.9001	0.1745	-	-	0.9828
MRA2	MRA2.1	average	-	0.0725	0.9556	0.0050	0.9938
	MRA2.2	average	-	0.1513	0.9015	0.0011	0.9949
	MRA2.3	average	-	0.1629	0.8886	-	0.9994
MRA3	MRA3.1	total	0.2691	0.0554	0.7135	0.0056	0.9955
	MRA3.2	total	0.2769	0.1272	0.6601	0.0019	0.9967
	MRA3.3	total	0.0902	0.1564	0.8083	-	0.9996

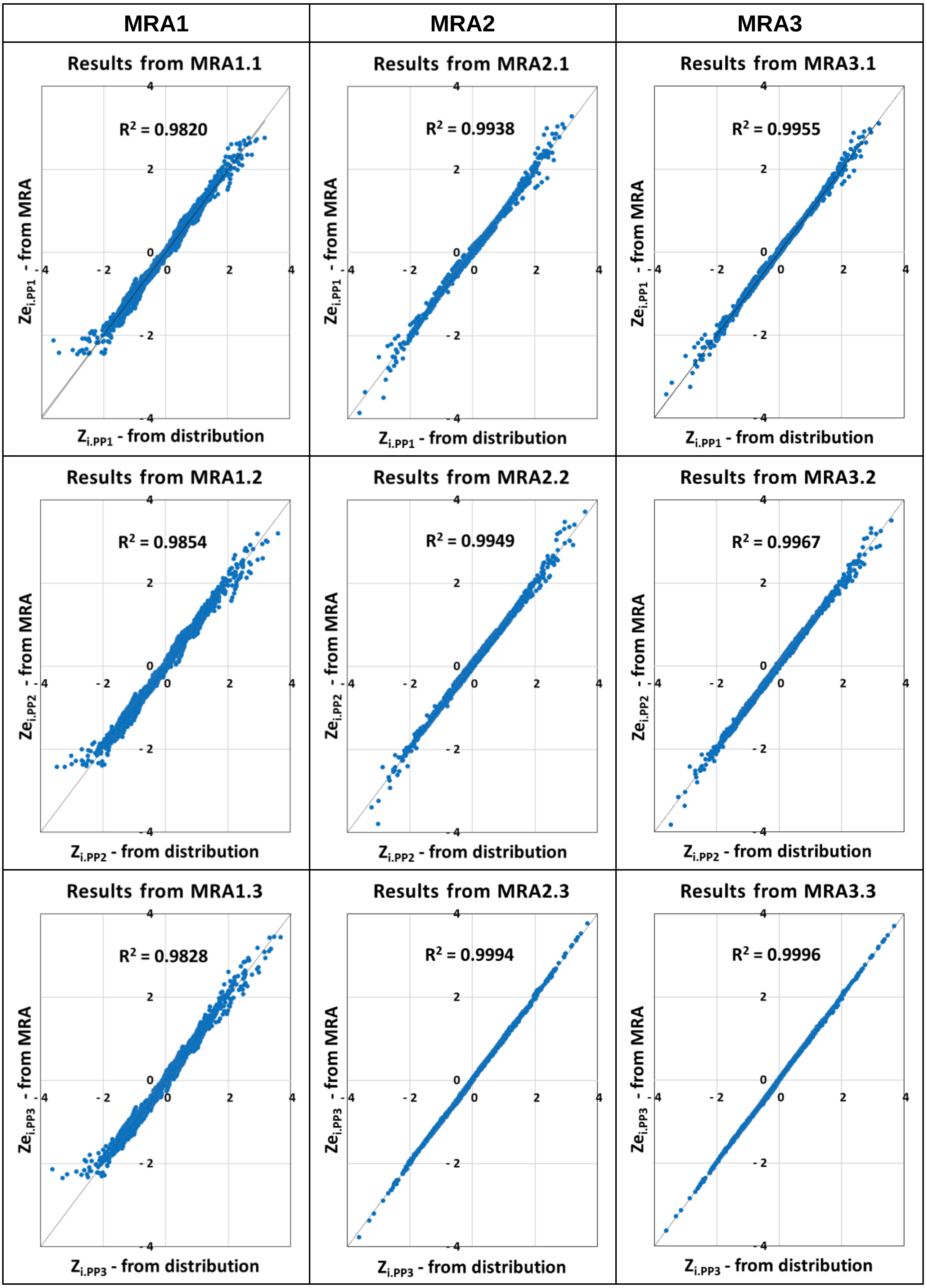


Figure 5-3: Comparing results from the Multiple Regression Analyses

A test for significance on MRA3.3 resulted in an F-value of 0.000 and P-values of 0.000, 0.000 and $2.24*10^{-130}$, respectively for Z_Q, Z_{logQ} and $Z_{Weibull}$ indicating very strong evidence of significance.

To illustrate the MRA-results the final (3rd) sequent, using the AMS of Woodstock Dam (one of the secondary flow sites), is demonstrated in Table 5-5.

The proposed PP technique, containing a set of Z-scores, is named the Z-set PP, for ease of reference.

Consequently, the proposed Z-set PP from the MRA3.3-results, containing the Weibull PP and yielding a slightly larger coefficient of determination of 0.9996, is proposed – the ranked Z-score is thus depicted by Equation 5-4.

$$Z_{i.Z\text{-}set} = 0.0902 * Z_{i.Weibull} + 0.1564 * Z_{Q_i} + 0.8083 * Z_{logQ_i}$$ **Equation 5-4**

The Z-set PPs (AEPs) can now be determined from the corresponding $Z_{i.Z\text{-}set}$ value by subtracting the related CDF of the standard normal distribution value from 1.

$$P_{i.Z\text{-}set} = 1 - \Phi(Z_{i.Z\text{-}set})$$ **Equation 5-5**

[in Excel: $P_{i.Z\text{-}set} = 1 - NORM.S.DIST(Z_{i.Z\text{-}set})$]

Table 5-5: Woodstock Dam (V1R003), as an illustration to calculate the Z-set PP

Statistics			$Q_{ave} = 498\ m^3/s$		$S = 436\ m^3/s$		$logQ_{ave} = 2.5605$	$S_{log} = 0.3542$	
rank	AMS flood peak data		$P_{Weibull}$ (Eq. 2-3)	Independent variables (Z-scores):			Dependent variable from:		P_{Z-set} (Eq. 5-5)
				Weibull PP	AMS flood peaks		Distributions	MRA	
i	Q	logQ		$Z_{Weibull}$ (Table 5-1)	Z_Q (Table 5-1)	Z_{logQ} (Table 5-1)	Z_{Distr} (Table 5-2)	$Z_{Z\text{-}set}$ (Eq. 5-4)	
1	2915	3.4646	0.0133	2.21636	5.54256	2.55231	2.68558	3.13002	0.0009
2	1400	3.1461	0.0267	1.93221	2.06825	1.65318	2.54353	1.83422	0.0333
3	1380	3.1399	0.0400	1.75069	2.02239	1.63554	2.26433	1.79640	0.0362
4	1275	3.1055	0.0533	1.61336	1.78159	1.53852	1.87618	1.66792	0.0477
5	1020	3.0086	0.0667	1.50109	1.19681	1.26496	1.74291	1.34519	0.0893
6	967	2.9854	0.0800	1.40507	1.07526	1.19954	1.69123	1.26464	0.1030
7	935	2.9708	0.0933	1.32050	1.00188	1.15828	1.66693	1.21218	0.1127
8	920	2.9638	0.1067	1.24445	0.96748	1.13845	1.65141	1.18390	0.1182
9	915	2.9614	0.1200	1.17499	0.95601	1.13177	1.62033	1.17043	0.1209
10	910	2.9590	0.1333	1.11077	0.94455	1.12506	1.55112	1.15741	0.1236
11	910	2.9590	0.1467	1.05084	0.94455	1.12506	1.36567	1.15200	0.1247
12	895	2.9518	0.1600	0.99446	0.91015	1.10468	1.35663	1.12506	0.1303
13	860	2.9345	0.1733	0.94107	0.82988	1.05577	1.29002	1.06815	0.1427
14	853	2.9309	0.1867	0.89025	0.81383	1.04575	1.25179	1.05296	0.1462
15	810	2.9085	0.2000	0.84162	0.71522	0.98234	1.25735	0.98188	0.1631
16	801	2.9036	0.2133	0.79491	0.69458	0.96864	1.13048	0.96337	0.1677
17	799	2.9025	0.2267	0.74987	0.69000	0.96558	1.13594	0.95610	0.1695
18	784	2.8943	0.2400	0.70630	0.65560	0.94234	1.11740	0.92801	0.1767

Statistics			$Q_{ave} = 498\ m^3/s$		$S = 436\ m^3/s$		$logQ_{ave} = 2.5605$	$S_{log} = 0.3542$	
rank	AMS flood peak data		$P_{Weibull}$ (Eq. 2-3)	Independent variables (Z-scores):			Dependent variable from:		P_{Z-set} (Eq. 5-5)
				Weibull PP	AMS flood peaks		Distributions	MRA	
i	Q	logQ		$Z_{Weibull}$ (Table 5-1)	Z_Q (Table 5-1)	Z_{logQ} (Table 5-1)	Z_{Distr} (Table 5-2)	Z_{Z-set} (Eq. 5-4)	
19	727	2.8615	0.2533	0.66404	0.52488	0.84980	1.05464	0.82895	0.2036
20	713	2.8531	0.2667	0.62293	0.49277	0.82596	1.02879	0.80095	0.2116
21	700	2.8451	0.2800	0.58284	0.46296	0.80340	1.01840	0.77443	0.2193
22	664	2.8222	0.2933	0.54367	0.38040	0.73868	1.00573	0.70566	0.2402
23	658	2.8182	0.3067	0.50532	0.36664	0.72755	0.91150	0.69105	0.2448
24	625	2.7959	0.3200	0.46770	0.29097	0.66447	0.86165	0.62483	0.2660
25	538	2.7308	0.3333	0.43073	0.09145	0.48070	0.77928	0.44175	0.3293
26	533	2.7267	0.3467	0.39434	0.07999	0.46926	0.74296	0.42742	0.3345
27	495	2.6946	0.3600	0.35846	-0.00716	0.37858	0.72614	0.33725	0.3680
28	483	2.6839	0.3733	0.32304	-0.03468	0.34849	0.70933	0.30543	0.3800
29	475	2.6767	0.3867	0.28802	-0.05302	0.32802	0.67712	0.28285	0.3886
30	453	2.6561	0.4000	0.25335	-0.10348	0.26988	0.64263	0.22483	0.4111
31	424	2.6274	0.4133	0.21898	-0.16998	0.18877	0.60936	0.14577	0.4421
32	411	2.6138	0.4267	0.18487	-0.19979	0.15059	0.58103	0.10717	0.4573
33	402	2.6042	0.4400	0.15097	-0.22043	0.12345	0.56717	0.07894	0.4685
34	385	2.5855	0.4533	0.11724	-0.25942	0.07047	0.55145	0.02698	0.4892
35	385	2.5855	0.4667	0.08365	-0.25942	0.07047	0.53944	0.02394	0.4904
36	385	2.5855	0.4800	0.05015	-0.25942	0.07047	0.52912	0.02092	0.4917
37	376	2.5752	0.4933	0.01671	-0.28006	0.04147	0.50869	-0.00877	0.5035
38	366	2.5635	0.5067	-0.01671	-0.30299	0.00843	0.49007	-0.04209	0.5168
39	335	2.5250	0.5200	-0.05015	-0.37408	-0.10007	0.46940	-0.14393	0.5572
40	335	2.5250	0.5333	-0.08365	-0.37408	-0.10007	0.42718	-0.14695	0.5584
41	321	2.5065	0.5467	-0.11724	-0.40619	-0.15241	0.38863	-0.19731	0.5782
42	317	2.5011	0.5600	-0.15097	-0.41536	-0.16778	0.37388	-0.21421	0.5848
43	310	2.4914	0.5733	-0.18487	-0.43141	-0.19516	0.32358	-0.24191	0.5956
44	305	2.4843	0.5867	-0.21898	-0.44288	-0.21509	0.29516	-0.26290	0.6037
45	288	2.4594	0.6000	-0.25335	-0.48187	-0.28540	0.26882	-0.32893	0.6289
46	288	2.4594	0.6133	-0.28802	-0.48187	-0.28540	0.26882	-0.33206	0.6301
47	288	2.4594	0.6267	-0.32304	-0.48187	-0.28540	0.20918	-0.33523	0.6313
48	288	2.4594	0.6400	-0.35846	-0.48187	-0.28540	0.18966	-0.33843	0.6325
49	270	2.4314	0.6533	-0.39434	-0.52314	-0.36453	0.16702	-0.41208	0.6599
50	254	2.4048	0.6667	-0.43073	-0.55984	-0.43942	0.14689	-0.48164	0.6850
51	246	2.3909	0.6800	-0.46770	-0.57818	-0.47865	0.13525	-0.51956	0.6983
52	246	2.3909	0.6933	-0.50532	-0.57818	-0.47865	0.11765	-0.52295	0.6995
53	246	2.3909	0.7067	-0.54367	-0.57818	-0.47865	0.07279	-0.52642	0.7007
54	226	2.3541	0.7200	-0.58284	-0.62405	-0.58261	0.04513	-0.62116	0.7328
55	215	2.3324	0.7333	-0.62293	-0.64927	-0.64378	0.03265	-0.67817	0.7512
56	213	2.3284	0.7467	-0.66404	-0.65386	-0.65524	-0.00548	-0.69186	0.7555
57	208	2.3181	0.7600	-0.70630	-0.66533	-0.68436	-0.02804	-0.72101	0.7645

Statistics			$Q_{ave} = 498\ m^3/s$		$S = 436\ m^3/s$		$logQ_{ave} = 2.5605$	$S_{log} = 0.3542$	
rank	AMS flood peak data		$P_{Weibull}$ (Eq. 2-3)	Independent variables (Z-scores):			Dependent variable from:		P_{Z-set} (Eq. 5-5)
				Weibull PP	AMS flood peaks		Distributions	MRA	
i	Q	logQ		$Z_{Weibull}$ (Table 5-1)	Z_Q (Table 5-1)	Z_{logQ} (Table 5-1)	Z_{Distr} (Table 5-2)	$Z_{Z\text{-}set}$ (Eq. 5-4)	
58	205	2.3118	0.7733	-0.74987	-0.67221	-0.70217	-0.03129	-0.74041	0.7705
59	172	2.2355	0.7867	-0.79491	-0.74788	-0.91735	-0.03456	-0.93024	0.8239
60	172	2.2355	0.8000	-0.84162	-0.74788	-0.91735	-0.07775	-0.93446	0.8250
61	172	2.2355	0.8133	-0.89025	-0.74788	-0.91735	-0.12248	-0.93885	0.8261
62	162	2.2095	0.8267	-0.94107	-0.77082	-0.99079	-0.12944	-1.00639	0.8429
63	142	2.1523	0.8400	-0.99446	-0.81668	-1.15233	-0.13292	-1.14896	0.8747
64	117	2.0682	0.8533	-1.05084	-0.87401	-1.38974	-0.16106	-1.35492	0.9123
65	115	2.0607	0.8667	-1.11077	-0.87860	-1.41088	-0.23059	-1.37813	0.9159
66	115	2.0607	0.8800	-1.17499	-0.87860	-1.41088	-0.24957	-1.38393	0.9168
67	114	2.0569	0.8933	-1.24445	-0.88089	-1.42159	-0.26474	-1.39922	0.9191
68	106	2.0253	0.9067	-1.32050	-0.89924	-1.51079	-0.31116	-1.48105	0.9307
69	105	2.0212	0.9200	-1.40507	-0.90153	-1.52241	-0.31907	-1.49844	0.9330
70	100	2.0000	0.9333	-1.50109	-0.91300	-1.58223	-0.33507	-1.55725	0.9403
71	91	1.9590	0.9467	-1.61336	-0.93364	-1.69785	-0.37606	-1.66408	0.9520
72	79	1.8976	0.9600	-1.75069	-0.96116	-1.87121	-0.38443	-1.82092	0.9657
73	76	1.8808	0.9733	-1.93221	-0.96804	-1.91868	-0.43029	-1.87675	0.9697
74	72	1.8573	0.9867	-2.21636	-0.97721	-1.98496	-0.43882	-1.95742	0.9749

5.2.3 Assessing the results from the Z-set PP technique

Impact on outliers and similar magnitude flood peaks

The AMS flood peaks of Woodstock Dam (a secondary site), used to illustrate concerns about the current PPs in Figure 2-8, is shown in Figure 5-4 where the Z-set PP are depicted against existing PPs in use (only Weibull and Hazen, being the 2 extremes, are shown).

A similar illustration, at one of the primary flow sites with a longer record, is shown in Figure 5-5.

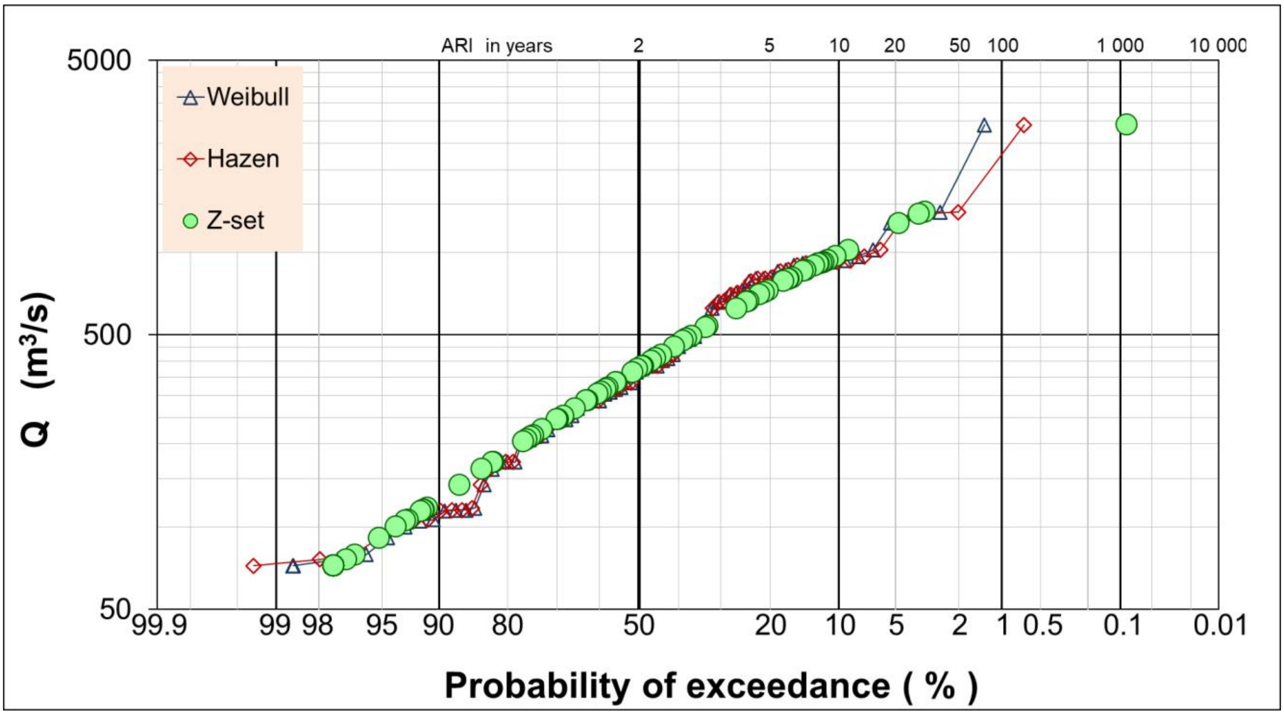


Figure 5-4: Comparison of the Z-set PP technique with existing techniques, at Woodstock Dam (secondary site)

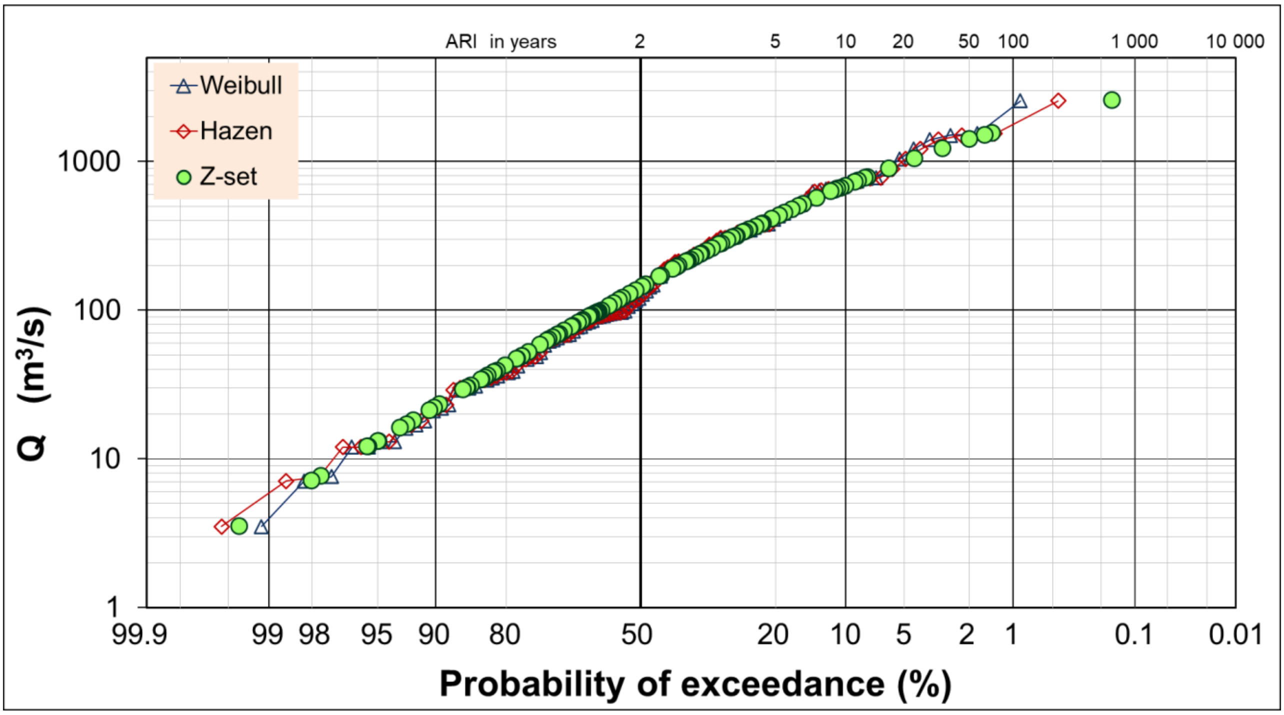


Figure 5-5: Comparison of the Z-set PP technique with existing techniques, at Witbank Dam (primary site)

Figure 5-4 and Figure 5-5 illustrates that the concerns, raised about the outliers and flood peaks of similar magnitude, depicted in Figure 2-8 and accompanying discussion, have been reasonably addressed without adversely affecting the rest of the PPs.

Although the results indicate an improvement, the observation from Figure 2-6 should be kept in mind – suggesting that, since the existing probability distributions were used in the compilation of the Z-set PP, the PPs might still be conservative towards risk.

Impact of Record length on Z-set PP

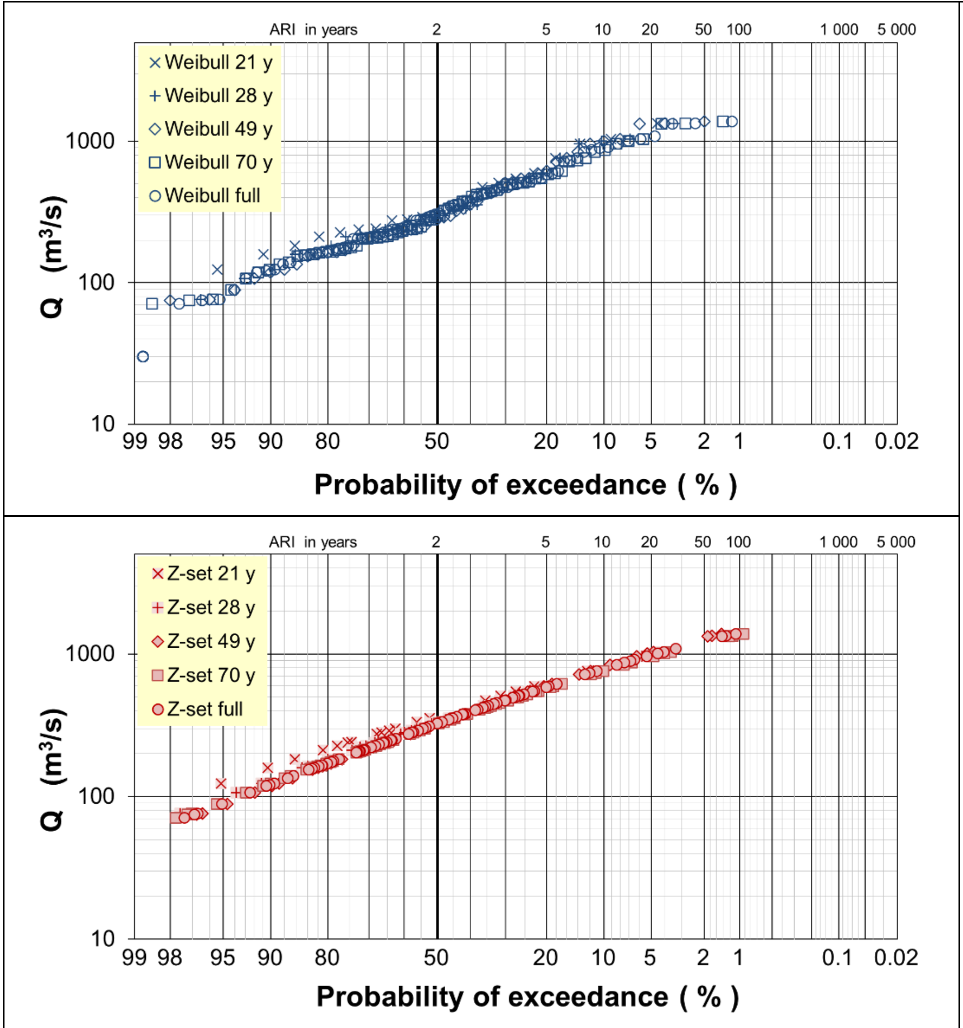
To illustrates the consequential effect of different record lengths on the Z-set PP it was compared to the Weibull PP. The effect on existing outliers was still evident, but it illustrates the (unforeseen) added benefit of using the Z-set PP.

Four case studies were considered to illustrate the benefit, namely:

- an AMS where no obvious outlier was present (Figure 5-6)
- an AMS where an outlier seemed to be present rather early in the record, but turned out to be just one of several higher flows as the record increased (Figure 5-7)
- an AMS where the highest peak, not necessarily an outlier, appeared rather late in the record (Figure 5-8)
- an AMS with a very high outlier occurring later in the record, staying a high outlier for the full record (Figure 5-9)

The 7ym concept used in the record length research (see Section 4.1.1), was applied to choose relevant record lengths. To avoid excessive cluttering on the figures, only relevant record lengths, where a change in the appearance of the PPs can be observed, are shown.

The AMS at Clanwilliam Dam (a secondary flow site), with homogeneously distributed flood peak data, is used for illustration purposes of the first scenario. Figure 5-6 illustrates the relationship.



Olifants River at Clanwilliam Dam.

Site chosen where the flood peaks are mainly caused by the same type of rainfall; in this case frontal rainfall.

No obvious outliers present in any record length.

CA = 2 025 km^2

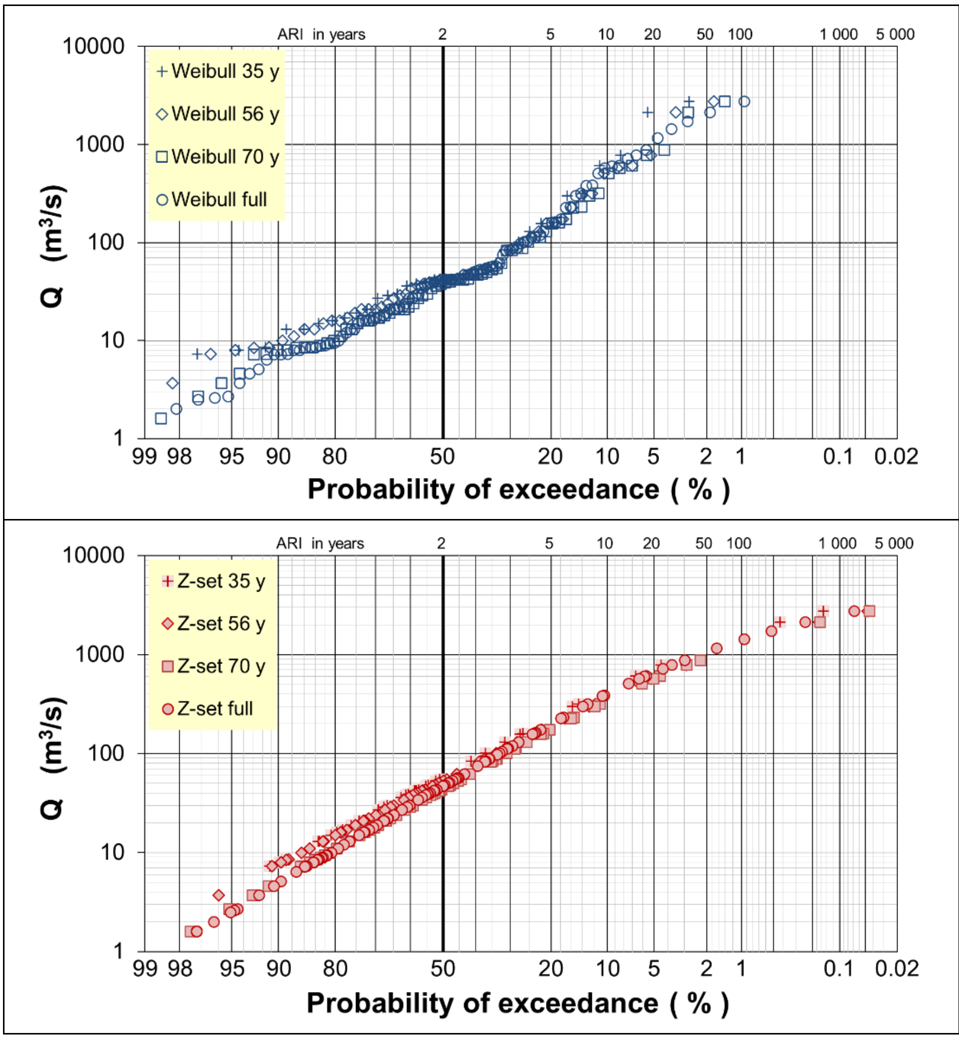
Record length: 84 y

Figure 5-6: The impact of AMS record lengths on the Z-set PP and Weibull PP (considering homogeneously distributed flood peak data)

Observations from Figure 5-6:

- Weibull PP: Slightly inferior to Z-set. With no outliers present in the AMS, there is virtually very little difference between the two PPs. However, the Z-set PP has a smoother appearance and remains effectively the same, regardless of record length.
- Z-set PP: Effectively no difference in the PPs, regardless of record length, from 28 years onwards. A slight deviation observed at higher AEPs (> 50%) for 21-year record.
- AEPs range visually very similar for Z-set and Weibull.
- The data appear to be remarkably stationary and homogeneous.

The AMS at Kammanassie Dam (also a secondary flow site) is used to illustrate (Figure 5-7) the effect on a record length, with an outlier relatively early in the record.



Kammanassie River at Kammanassie Dam.

Site chosen where the highest flood peak on record occurred within the first 21 years.

CA = 1 525 km^2

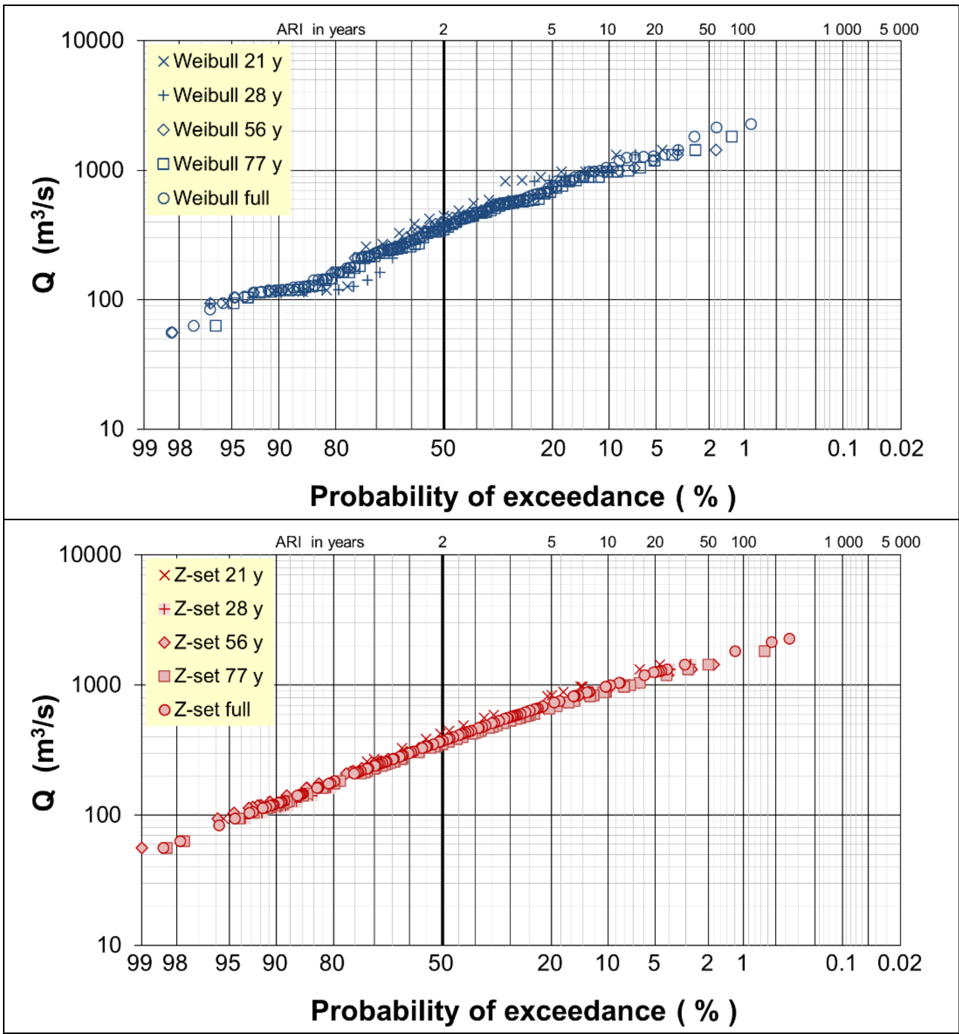
Record length: 106 y

Figure 5-7: The impact of AMS record lengths on the Z-set PP and Weibull PP (considering highest flood peak in early part of record)

Observations from Figure 5-7:

- Weibull PP: Appears more disorderly than the Z-set PP.
- Z-set PP: Effectively, little difference in PPs at AEPs ≤ 50%. Due to several low flows added after 56 years of record, 2 distinct groupings can be observed in the higher AEP range (> 50%) – PPs for 35 to 56 years are grouped and PPs for 70 to 106 years are grouped;
- The AEP range is not the same for Z-set and Weibull, due to failure of the Weibull PP to deal effectively with outliers and higher flows.
- The data appear to be relatively stationary and homogeneous.

In Figure 5-8 the results, at a site with the highest peak occurring later in the record, are depicted. The AMS at Grootdraai Dam, a primary flow site, is used.



Vaal River at Grootdraai Dam.

Site chosen where the highest flood peak on record occurred after 75 years.

CA = 7 982 km^2

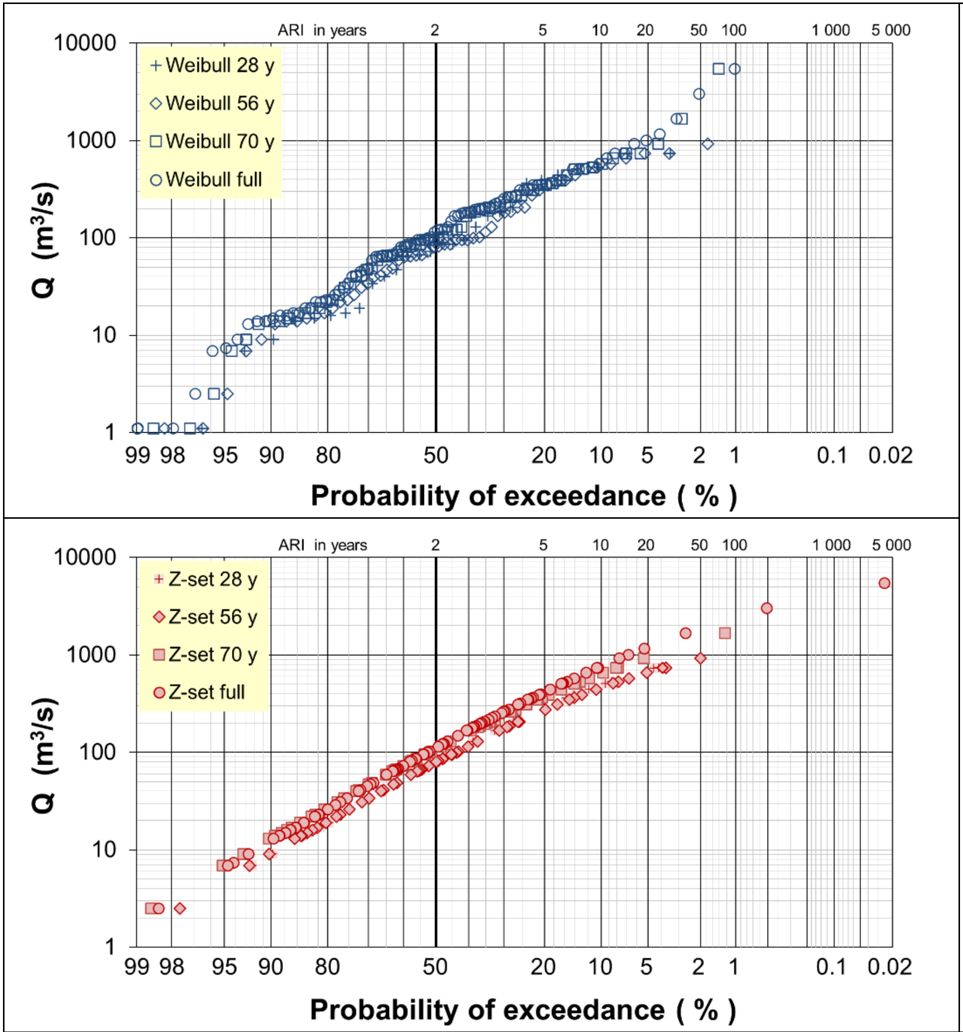
Record length: 114 y

Figure 5-8: The impact of AMS record lengths on the Z-set PP and Weibull PP (considering highest flood peak in later part of record)

Observations from Figure 5-8:

- Weibull PP: somewhat inferior to Z-set; slightly more scattered, particularly at lower AEPs, with different record lengths.
- Z-set PP: Effectively no difference in PPs, irrespective of record length.
- Z-set PP has a slightly wider AEP range than the Weibull PP.
- The data appear to be relatively stationary and homogeneous.

The effect that a very high outlier can have on a record length, is depicted in Figure 5-9. The AMS at Floriskraal Dam, one of the primary flow sites, is used as an example.



Buffels River at Floriskraal Dam.

Site chosen with a very high outlier occurring in 1981 (year 61).

CA = 4 024 km^2

Record length: 98 y

Figure 5-9: The impact of AMS record lengths on the Z-set PP and Weibull PP (considering a very high outlier)

Observations from Figure 5-9:

- Weibull PP: Two groups of PPs exist – one for 28 to 56-year record lengths and one for 70 to 98-year record lengths. However, it is less visible in the Weibull PP since the PPs are much more scattered than in the Z-set PP.
- Z-set PP: The two PP groups are clearly distinguishable. The split is caused by the high outlier in 1981 – in year 61 of the 98-year AMS.
- A large difference in AEP range between Weibull and Z-set, for record lengths ≥ 70 y. It is due to the inability of existing PPs to make any provision for outliers.
- The data appear to be reasonably homogeneous (apart from the split caused by the high outlier).

Figure 5-6 to Figure 5-9 illustrate that the Z-set PPs are reasonably similar for varying record lengths. This appears not to be the case if a relative high outlier occurs somewhere in the record (in the example in Figure 5-9, the outlier emanates from the devastating 'Laingsburg 1981' flood event).

It is interesting to note that the ARI associated with the 1981-event (Z-set PP), of approximately 4 000 years, is consistent with the dating of other palaeoflood evidence in the J-drainage region, of about 3 000 years ago (Van Bladeren, Zawada and Mahlangu, 2007). Zawada (1994) observed that, while palaeoflood evidence exists in the region for other rivers, no palaeoflood evidence could be obtained in that part of the Buffels River. He concluded that the 1981-flood most probably scoured any palaeoflood evidence and that no evidence exists for a flood event larger than the 1981-event in the Buffels River.

Boxplots were used to further illustrate the benefit of using the Z-set PP compared to the Weibull PP. To explain how the boxplots were generated, the FFA on Clanwilliam Dam is used as an example (see Table 5-6).

- An FFA was performed, using the complete AMS record, and the most suitable probability distribution was chosen that best fits the AMS data.
- Seven commonly used AEPs (50, 20, 10, 5, 2, 1 and 0.5%) were chosen with their corresponding Z-scores and flood peaks.
- Using various record lengths (28, 49, 70, and 84 years in this example) Z-scores, for these matching flood peaks were determined, using both the Weibull- and Z-set PPs
- The record lengths were chosen in the same way as described in Section 4.1.1 (also see Figure 5-6).
- For the first set of boxplots the Z-score 'variance' (Var_Z) is used, as the average of the squared deviation from the expected Z-score of every applicable AEP, using different record lengths. Thus, Var_Z for every AEP was determined, for various record lengths – for example, from Table 5-6, for an AEP of 10% the Var_Z for the Z-set PP is given by:

  $Var_Z = [(1.284 - 1.282)^2 + (1.257 - 1.282)^2 + (1.361 - 1.282)^2 + (1.308 - 1.282)^2]/4 = \mathbf{0.00189}$

  A similar procedure was used to calculate the variance, using the Weibull PP.

  The boxplot of the Var_Z -values, for both the Z-set and Weibull PPs, across the AEP-range is depicted in Figure 5-10.

- For the second set of boxplots the range of Z-scores ($Z_{max} - Z_{min}$), obtained from the different record lengths, associated with every AEP, were determined. For the same example:

  $Z_{max} - Z_{min} = 1.361 - 1.257 = \mathbf{0.104}$

  The boxplot of the '$Z_{max} - Z_{min}$'-values, for both the Z-set and Weibull PPs, across the AEP-range is depicted in Figure 5-11.

Table 5-6: Data for boxplots at Clanwilliam Dam (E1R002) – full record length 84 years

Results from FFA distribution			AMS record lengths used (years)								Data for Boxplots			
			28		49		70		84					
			Z-scores from Weibull and Z-set PPs								*Var*$_Z$		Range (Max-Min)	
AEP (%)	Z-score	Q (m³/s)	Wei bull	Z-set	Wei bull	Z-set	Wei bull	Z-set	Wei bull	Z-set	Weibull	Z-set	Wei bull	Z-set
50	0.000	332	0.195	-0.011	0.189	0.036	0.079	0.040	0.066	0.027	0.0210	0.0009	0.129	0.051
20	0.842	602	0.825	0.850	0.797	0.842	0.923	0.917	0.900	0.880	0.0031	0.0018	0.126	0.075
10	1.282	802	0.977	1.284	1.040	1.257	1.179	1.361	1.104	1.308	0.0483	**0.00189**	0.202	**0.104**
5	1.645	986	1.176	1.621	1.228	1.574	1.424	1.704	1.430	1.652	0.1221	0.0023	0.255	0.131
2	2.054	1233	2.426	2.095	1.508	1.951	1.682	2.106	1.756	2.050	0.1659	0.0038	0.918	0.155
1	2.326	1426	2.045	2.277	2.353	2.263	2.255	2.409	2.540	2.371	0.0327	0.0038	0.496	0.146
0.5	2.576	1624	2.273	2.501	3.701	2.603	2.528	2.658	3.781	2.713	0.7033	0.0080	1.508	0.212

Appendix 5-A contains the statistics of the remaining stations used in Figure 5-10 and Figure 5-11.

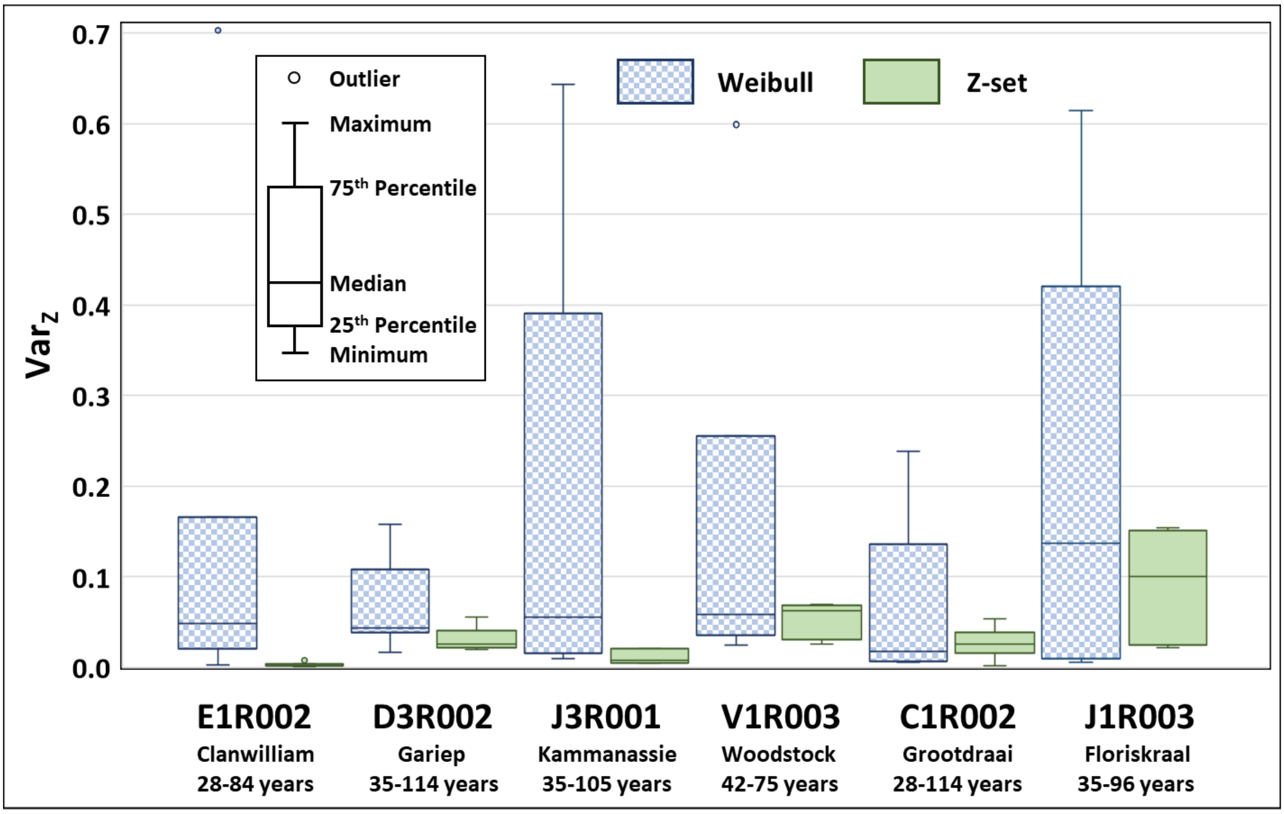


Figure 5-10: Weibull and Z-set PPs: Variance of PPs considering an array of record lengths

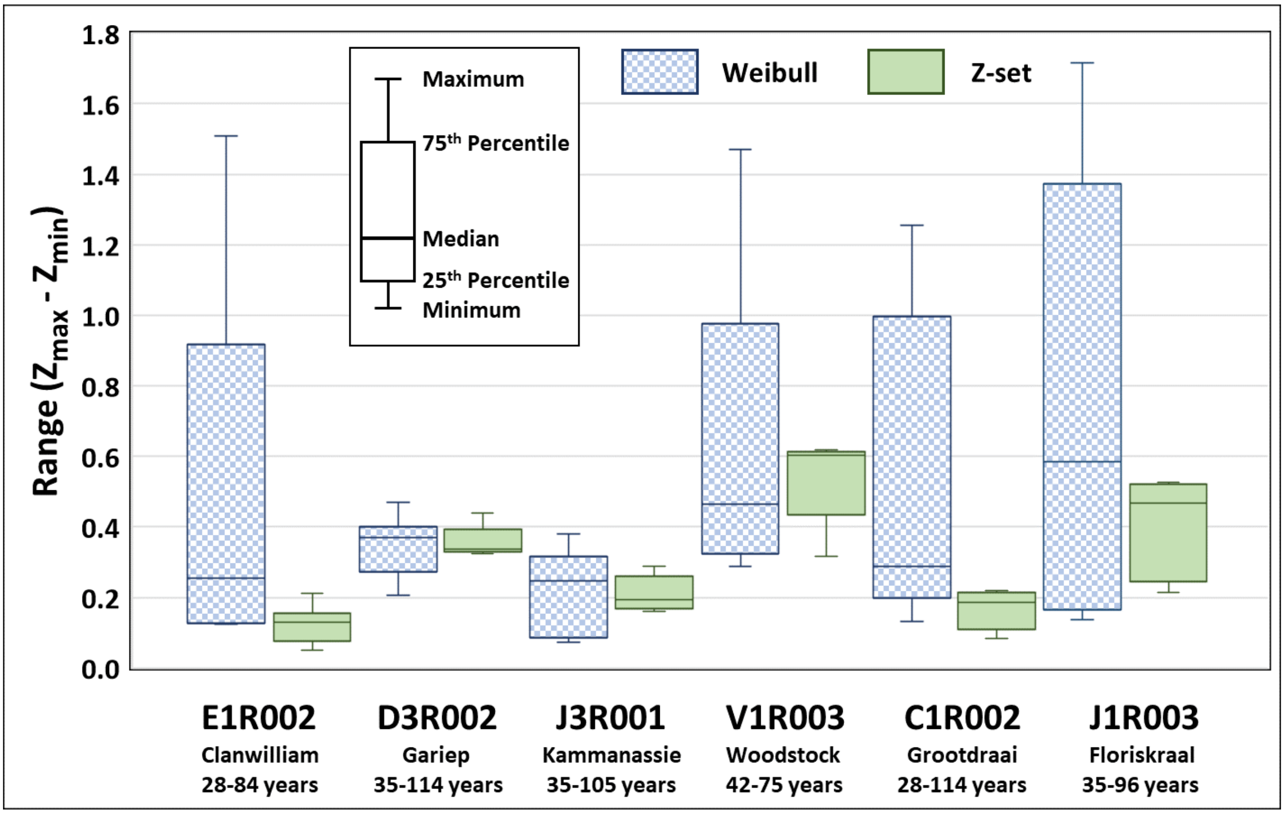


Figure 5-11: Weibull and Z-set PPs: Range of PPs considering an array of record lengths

The advantage of using the Z-set PP is evident from the sets of two boxplots that reveals a higher degree of consistency in the Z-set PPs, regardless of record length. This is especially true for sites with record lengths longer than 35 to 40 years.

5.2.4 Demonstrate confirmation of PP-hypothesis

It was hypothesised (in Section 5.1.3) that any of the existing PPs can be used as an independent variable in performing an iterative MRA to estimate the Z-set PP. The hypothesis was tested in this research, using the two extreme existing PP techniques, the Weibull PP and the Hazen PP. The results of using the Hazen PP are presented in Table 5-7 (analogue to Table 5-4, where the Hazen PP was used)

Table 5-7: MRA, including Hazen PP: Parameters and coefficients of determination

$Z_{i.PP_y} = C_{1_{xy}} * Z_{i.Hazen} + C_{2_{xy}} * Z_{Q_i} + C_{3_{xy}} * Z_{logQ_i} + C_{4_{xy}}$								
Multiple Regression Analysis			Coefficients					
			C_1	C_2	C_3	C_4	R^2
MRA1	MRA1.1	average	0.9931	0.0286	-	0.0050	0.9823
	MRA1.2	average	0.9414	0.1060	-	0.0011	0.9855
	MRA1.3	total	0.8830	0.1615	-	-	0.9833
MRA2	MRA2.1	average	-	0.0727	0.9554	0.0050	0.9938
	MRA2.2	average	-	0.1516	0.9013	0.0011	0.9949
	MRA2.3	average	-	0.1644	0.8887	-	0.9994
MRA3	MRA3.1	total	0.2598	0.0518	0.7168	0.0056	0.9955
	MRA3.2	total	0.2646	0.1237	0.6660	0.0019	0.9966
	MRA3.3	total	0.0880	0.1571	0.8082	-	0.9996

Thus, in using the Hazen PP, equation 5-4 merely changes to:

$$Z_{i.Z\text{-}set} = 0.0880 * Z_{i.Hazen} + 0.1571 * Z_{Q_i} + 0.8082 * Z_{logQ_i} \qquad \textbf{Equation 5-6}$$

To illustrate the effect that a choice of an existing PP technique might have on the Z-set PP, the AMS at two selected sites were used; Woodstock Dam, with relatively larger flood peaks (Figure 5-12), and Maricopoort Dam with relatively lower flood peaks (Figure 5-13).

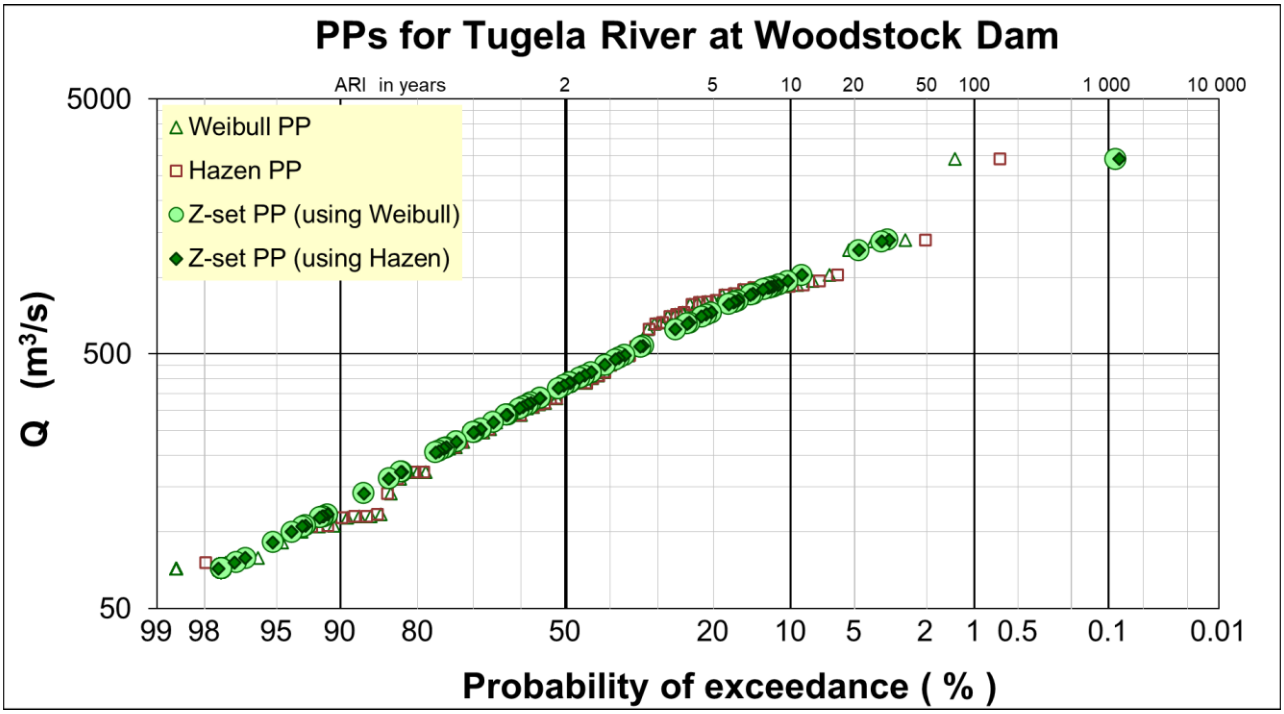


Figure 5-12: Comparison of the Z-set PP, using either Weibull or Hazen PPs at V1R003

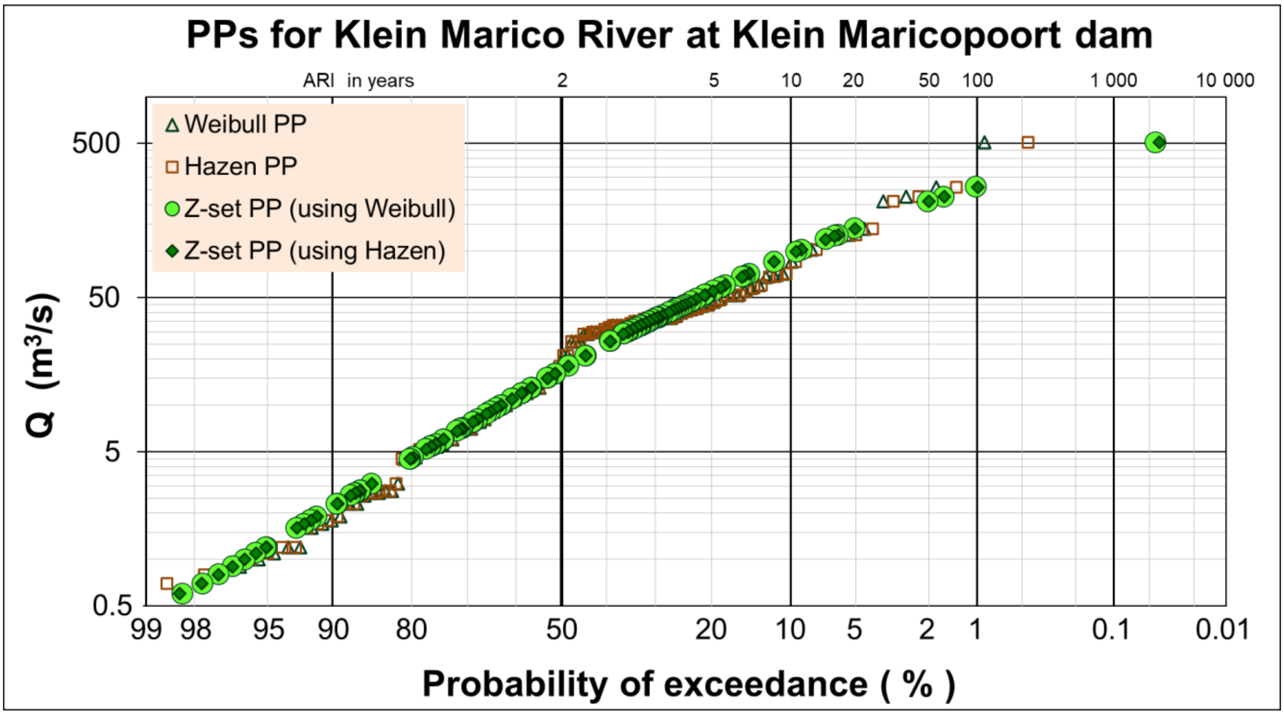


Figure 5-13: Comparison of the Z-set PP, using either Weibull or Hazen PPs at A3R002

Thus, the hypothesis was confirmed that the choice of an existing PP technique had, virtually, no effect on the Z-set PP technique.

5.3 Plotting Position discussion

The second secondary objective was to propose a more sensible PP, against which probability distributions can be compared, to ensure the selection of the most appropriate probability distribution.

Primary sites were used in this part of the research and although the number of sites was relatively small, the focus was to use sites with long record lengths. A representative spread across the country, covering dry (low rainfall) and wet (higher rainfall) parts, as well as small and large catchment areas, were considered as being more important than simply the number of sites used.

Considering the enormous differences in flood peak magnitudes, as well as the noticeable differences in the slope of the trends of the PPs between the sites (see Figure 5-2), the extremely high coefficient of determination is an indication that the outcome of this research can be considered as particularly satisfactory.

The proposed approach, that include some statistical parameters, produced promising results, leading to a new proposed PP (Z-set PP) with the following observed characteristics:

- The general trend of the Z-set PP technique does not differ much from that of the existing PPs.
- By appreciating the relative magnitudes of the flood peaks in the PP, through the inclusion of the Z-scores of the data, notable differences can be observed where the Z-set PP:
    - eliminates the allocating of noticeably different PPs (probabilities), to flood peaks of similar magnitude;
    - visually, appears much smoother than the jagged appearance of the existing PPs, mimicking the shape of a probability distribution;
    - offers an improved and more realistic representation of outliers.
- The Z-set PP proved to be less susceptible to different record lengths than the existing PP techniques, depending on the relative magnitude of an outlier; hence, it may encourage more consistent choices of appropriate probability distributions.
- It was also noted that the Z-set PP, from record lengths of 35 years and above, seem to follow a similar PP trend – depending on the size of outliers.

The practice to normalise datasets does not do justice to the appearance of an outlier(s) in a dataset. Supplemental to the ranking technique, the relative magnitudes of the flood peaks in a dataset were also considered, in estimating the Z-set PP. Consequently, the Z-set PP technique assigns a more sensible PP to an outlier, by acknowledging the relative magnitude of the outlier.

A preliminary supplementary observation might suggest that sites require longer minimum record lengths in areas with more than one rainfall-causing system (e.g., V1R003; 42 years) than sites in areas with primarily one rainfall-causing system (e.g., E1R002; 28 years).

6 Probability distribution

In this chapter an improved probability distribution is researched. A comprehensive methodology to pursue, applying the improved Z-set PP technique, was not clear from the start and the methodology evolved as the research progressed. The Z-set PP served as basis for the development of an improved probability distribution.

6.1 Probability distribution methodology

This section is divided into 3 parts; in the first part a screening of the preliminary selection of statistics was conducted to eliminate statistics that have an insignificant effect on the development of a probability distribution. The second part covers the more detailed research and comparison of various scenarios of the more significant statistics, which resulted in the best combination of statistics that conclusively produced the best probability distribution model. In the third part the scenario, considered to be the most appropriate combination of statistics, was used to estimate the various parameters for a 'new' probability distribution and the appropriate regression coefficients.

6.1.1 Part 1: Screening of preliminary potential statistics

Table 3-3, and the associated table in Appendix 3-E, covers the principal identified statistics, which were initially considered as potential parameters for a probability distribution. MRAs were utilised to screen the identified potential statistics, defined in Table 6-1 for clarity, to establish which would have a negligible- or no impact on the development of a new 'underlying' probability distribution.

Table 6-1: Complete initial statistic variable list considered for the MRAs

<table>
<tr><th>Independent variables</th><th>Description</th><th>Comments</th></tr>
<tr><td>Q_{max}</td><td>Maximum peak of the AMS (m^3/s)</td><td rowspan="10">* Indicates excluding largest flood peak (Q_{max}) from AMS dataset

*° Indicates excluding both largest flood peak (Q_{max}) and smallest flood peak (Q_{min}) from AMS dataset.

** Indicates excluding largest 2 flood peaks from AMS dataset</td></tr>
<tr><td>Q_{min}</td><td>Minimum peak of the AMS (m^3/s)</td></tr>
<tr><td>Q_{gmn}</td><td>Geometric mean of the AMS (m^3/s)</td></tr>
<tr><td>Q_{mdn}, Q_{mdn}*, Q_{mdn}*°</td><td>Median(s) of the AMS (m^3/s)</td></tr>
<tr><td>Q_{ave}, Q_{ave}*, Q_{ave}*°, Q_{ave}**</td><td>Average(s) of the AMS (m^3/s)</td></tr>
<tr><td>SD, SD*, SD*°, SD**</td><td>Standard deviation(s) of the AMS (m^3/s)</td></tr>
<tr><td>g, g*, g*°</td><td>Skewness statistics of the AMS</td></tr>
<tr><td>kurt, kurt*</td><td>Kurtosis statistics of the AMS</td></tr>
<tr><td>COV, COV*, COV*°</td><td>Coefficient of variation of the AMS</td></tr>
<tr><td>Z_{max} and Z_{min}</td><td>Z-scores of Q_{max} and Q_{min}</td></tr>
<tr><th>Dependent variable</th><th>Description</th><th>Comments</th></tr>
<tr><td>Q_{obs}
(see also Table 6-2)</td><td>The observed flood peak(s), corresponding to selected AEPs, estimated from observed AMS flood peaks, applying the Z-set PP technique</td><td>Selected AEPs (in %):
99.5, 99, 95, 90, 80, 70, 60, 50, 40, 30, 20, 10, 5, 2, 1, 0.5, 0.2, 0.1, 0.03, 0.01, 0.003, 0.001.</td></tr>
</table>

The Q_{obs} values, at identified AEPs for all the flow sites, were considered as dependent variables. Values for Q_{obs} were estimated, for the identified AEPs, by interpolation and extrapolation of the flood frequency relationship between AEP (from the Z-set PP) and the AMS flood peaks. The concise flood frequency data (Q_{obs} vs selected AEPs) are presented in Table 6-2 (full table in Appendix 6-A).

Table 6-2: Flood frequency data (Q_{obs} vs AEP) to be considered as dependent variables

Site	**AEP (%)**														
	99	**95**	**90**	**80**	**50**	**20**	**10**	**5**	**2**	**1**	**0.5**	**0.2**	**0.1**	**0.01**	**0.001**
	Associated Q_{obs} (m^3/s) – from Z-set PP														
A3R002	0.472	1.20	2.27	4.51	16.5	54.3	95.4	140	208	262	317	390	444	610	736
B1R001	5.18	13.0	22.2	42.3	141	418	683	969	1 384	1 620	1 793	2 338	2 749	3 571	4 285
B2R001	3.68	9.46	15.3	26.8	80.9	215	334	470	662	809	956	1 149	1 290	1 717	2 039
C1R002	48.2	90.4	123	181	373	721	983	1 260	1 614	1 888	2 162	2 523	2 795	3 676	4 503
C2R001	1.46	3.01	4.29	6.91	16.1	35.0	49.2	64.7	86.6	101	118	140	157	217	281
C5R002	7.06	20.8	37.9	71.5	236	738	1 301	1 973	2 985	3 821	4 700	5 907	6 841	9 951	12 829
C9R002	56.7	114	178	292	764	1 827	2 703	3 611	4 789	5 749	6 591	7 694	8 416	11 115	13 359
D3R002	268	505	689	1 006	2 024	3 847	5 209	6 575	8 377	9 852	11 146	12 941	14 301	18 710	22 868
D7H005	88.6	191	278	449	1 157	2 596	3 810	5 001	6 742	7 984	9 247	10 871	12 055	15 578	18 238
J1R003	1.66	6.59	12.6	25.9	112	421	783	1 141	1 753	2 360	2 991	3 760	4 344	6 026	7 288
J2R001	0.662	1.57	2.49	4.29	11.4	29.8	45.8	64.6	91.4	113	138	165	184	248	310
N2R001	11.9	30.2	48.2	83.5	229	618	1 005	1 465	2 129	2 634	3 156	3 896	4 455	6 218	7 743
Q1R001	11.1	21.7	29.1	43.1	87.1	174	244	317	418	500	579	689	768	1 065	1 381
V6H002	160	272	353	481	841	1 360	1 719	2 064	2 507	2 835	3 159	3 582	3 898	4 927	5 925
X1H001	20.0	41.1	60.0	91.5	202	431	628	837	1 129	1 358	1 591	1 904	2 141	2 914	3 630
A2R001	11.6	25.8	38.9	67.5	177	422	630	841	1 137	1 364	1 591	1 886	2 102	2 757	3 272
A2R005	0.933	3.20	3.71	6.81	22.0	51.3	80.5	112	160	223	262	309	342	465	621
A4R001	2.03	5.52	10.1	18.7	73.4	261	443	649	953	1 204	1 438	1 750	1 987	2 791	3 629
A8R001	3.82	10.5	17.1	32.1	103	319	541	742	1 173	1 454	1 795	2 221	2 542	3 450	4 205
B3R001	1.86	4.33	6.97	11.6	29.7	79.3	129	181	258	332	395	486	550	770	1 026
B6R003	15.6	31.9	48.7	78.8	198	427	630	820	1 089	1 299	1 474	1 792	1 964	2 611	3 058
B7R001	0.721	1.54	4.71	9.76	38.7	138	236	350	466	558	693	936	1 035	1 331	1 587
C3R002	8.55	16.6	25.1	43.8	112	273	421	587	824	1 027	1 449	1 735	1 936	2 531	3 160
C5R001	1.35	3.38	5.94	11.9	33.8	133	237	357	534	680	834	1 116	1 283	1 828	2 489
C5R003	4.49	10.8	16.9	29.9	89.0	256	432	628	897	1 120	1 354	1 674	1 923	2 768	3 610
C7R001	18.1	33.6	48.5	75.8	173	368	527	694	928	1 105	1 282	1 512	1 680	2 188	2 579
D6R002	1.16	3.70	7.04	13.8	56.2	185	313	404	678	863	1 043	1 254	1 365	1 775	2 113
E1R002	38.4	89.9	121	170	325	591	785	978	1 232	1 425	1 619	1 876	2 073	2 732	3 403
H5H004	111	182	235	323	563	972	1 194	1 519	1 815	2 051	2 352	2 642	2 903	3 787	4 702
J2R003	0.142	0.393	0.781	1.65	6.00	23.3	43.7	65.7	98.3	129	159	195	227	344	474
J3R001	0.815	2.53	4.87	10.5	45.0	203	406	665	1 068	1 409	1 756	2 255	2 582	3 670	4 532
N1R001	2.79	7.89	13.7	29.7	104	342	597	898	1 342	1 685	2 037	2 522	2 906	4 024	4 888
Q4R002	7.72	18.6	27.5	42.7	98.1	212	301	397	531	642	736	874	978	1 319	1 639
Q5R001	11.6	29.2	48.4	86.3	241	646	1 033	1 441	2 076	2 518	2 962	3 604	4 115	5 486	6 763
S6R002	1.81	5.27	9.01	16.9	55.1	162	266	376	511	598	730	935	1 029	1 308	1 657
T2R001	11.4	26.9	41.4	71.7	199	454	652	852	1 083	1 235	1 374	1 628	1 775	2 236	2 666
U2R001	14.6	29.9	40.0	59.0	122	239	335	495	694	827	948	1 111	1 237	1 530	1 936
U3R001	2.14	4.76	11.5	23.3	77.7	251	426	614	911	1 143	1 378	1 691	1 875	2 414	2 881
V1R003	52.3	93.8	132	201	429	848	1 126	1 447	1 875	2 176	2 367	2 758	3 122	4 007	4 867
V3R003	15.4	26.4	35.5	46.5	84.5	147	188	237	290	334	377	436	481	635	794
W4R001	36.2	83.2	129	210	546	1 376	2 150	3 049	4 386	5 500	6 675	8 287	9 536	13 760	17 997

Thus, MRAs were conducted for every identified AEP to estimate a flood peak (Q_{est}) that could be compared to the corresponding Q_{obs}, for every site.

Numerous combinations of the identified statistics were explored as potential independent variables in the MRAs, in the screening process. The identified variables, which were found to be insignificant in improving the results in any way, were eliminated. The flow-diagram in Figure 6-1 offers a visual depiction of the applied methodology.

6.1.2 Part 2: Finding the best combination of significant statistics

With a vast array of scenarios still possible, the MRA-approach was also used in the pursuing research. With the least significant statistics eliminated, the independent statistic variables that were left (shown as an example in Table 6-3) were further researched, while the dependent variables (Q_{obs}) stayed the same.

Table 6-3: Example of most significant statistic independent variables

Remaining independent variables	Comment
Q_{gmn}, Q_{mdn}, Q_{ave}, g, SD, … …	Descriptions and comments identical to that in Table 6-1

Due to the structure of most existing probability distributions, it was anticipated that a scenario with one of the measures of central tendency as integral part, accompanied by the SD and g, would result in the most likely solution to the research. Consequently, for example, the following scenarios were considered and compared:

Table 6-4: Example of scenarios considered with different measures of central tendency

Scenarios considered, with shown measure of tendency:		
Average (Q_{ave})	**Median (Q_{mdn})**	**Geometric mean (Q_{gmn})**
Q_{ave}, SD	Q_{mdn}, SD	Q_{gmn}, SD
Q_{ave}, SD, g	Q_{mdn}, SD, g	Q_{gmn}, SD, g
etc…	etc…	etc…

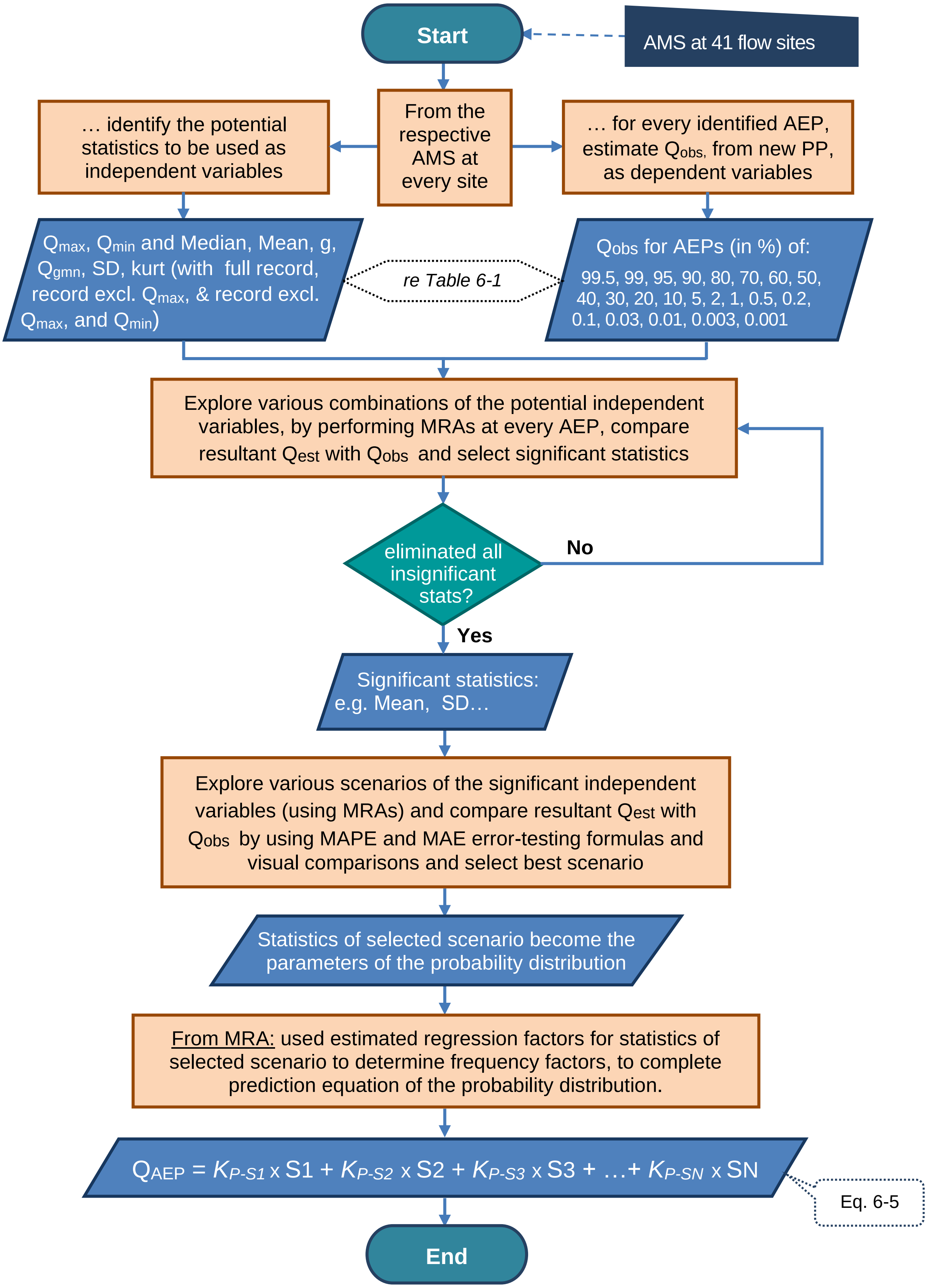


Figure 6-1: Flow chart. illustrating the probability distribution methodology

For every scenario and every site, MRAs were conducted at all selected AEPs to estimate values of Q_{est}. Instead of just comparing Q_{est} with the corresponding Q_{obs}, as in the screening process, so-called 'error testing models' were used to rank the performance of the different scenarios against each other. All these models have advantages and disadvantages, mainly related to the magnitude of the values involved – therefore, two formulas were chosen, namely the mean absolute percentage error (MAPE) and the mean absolute error (MAE), depicted in Equation 6-1 and Equation 6-2 respectively. The MAE is likely to suggest relatively large 'errors' at lower AEPs (large flood peaks), while the percentage difference might be very small; on the other hand, MAPE might suggest large percentage 'errors' at higher AEPs, while the absolute errors (in terms of flood peaks) could be insignificant in relation to flood hydrology. Therefore, these tests should be considered as relative, rather than absolute, for the purpose of this research.

In FFA the 'design flood' range is of more practical value than the AEP range > 50%. Thus, both MAPE and MAE were considered separately for AEP range ≤ 50% and AEP range > 50%.

$$\text{MAPE} = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{Qi_{est} - Qi_{obs}}{Qi_{obs}}\right| \qquad \textbf{Equation 6-1}$$

and

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|Qi_{est} - Qi_{obs}\right| \qquad \textbf{Equation 6-2}$$

where

MAPE - mean absolute percentage error

MAE - mean absolute error

n - sample size = number of observations considered in the applicable AEP range

Qi - flood peak related to an AEP at observation '*i*'

From the results the scenario that provided, on average, the best fit to the observed AMS flood peaks will be used to develop a new probability distribution.

6.1.3 Part 3: Probability distribution parameters and regression coefficients

Kline (2016) indicated that most researchers consider data with absolute values for skewness and kurtosis larger than 3.0 and 8.0, respectively, as problematic (in this research: g > 3 and kurt > 8 at about 56% of the sites and at a further 9.7% of the sites the kurt > 8; refer to Table 3-3). Thus, as a practical alternative, also considering the statement of Wheeler (2022a) that no skewed probability model exists that will fit these data, the MRA results of the chosen scenario was used to find an empirical solution for the probability distribution. Consequently, the independent variables turn out to be the parameters for the probability distribution, while the regression coefficients, associated with the AEPs, became the frequency factors.

However, the wide range of AEPs did not provide sensible independent variables and associated standardised variables were considered, instead. Two standardised variables were considered, namely the standardised variate used for the log-normal distribution (W_T) and the standardised variate for the EV1 distribution (W_P). The standardised variate W_T is equal to the Z-score of the standard normal distribution, which is defined as:

W_T = 0 – the inverse of Φ(Z) of the standard normal distribution (Φ(Z) from Equation 5-2)

The standardised variate W_P is defined as:

$$W_P = -\ln(-\ln(1-AEP))$$ **Equation 6-3**

To estimate regression coefficients at AEPs between those used in the research, polynomial functions were used. For each statistical parameter, the regression coefficients determined at every AEP, were considered to be the dependent variables (y) while the associated standard variates (e.g., W_P) were considered to be the independent variables (x).

The general form of the polynomial function (using W_P, for example) is:

$$K_P = a_n(W_P)^n + a_{n-1}(W_P)^{n-1} + \ldots + a_2(W_P)^2 + a_1(W_P)^1 + a_0$$ **Equation 6-4**

where

K_P - frequency factor

n - denotes the degree of the polynomial

a_n (etc.) - coefficient of the n^{th} degree term (etc.)

The general form of the prediction equation of the probability distribution can be written as:

$$Q_{AEP} = K_{P\text{-}S1} \times S1 + K_{P\text{-}S2} \times S2 + K_{P\text{-}S3} \times S3 + \ldots + K_{P\text{-}SN} \times SN$$ **Equation 6-5**

where

Q_{AEP} - Flood peak associated with an AEP

K_P - frequency factor related to an AEP (used *P* for conciseness)

S1 (etc.) - 1^{st} Statistical parameter (etc.)

$K_{P\text{-}S1}$ (etc.) - frequency factor for S1 (etc.)

N - Total number of statistics considered

This concludes the methodologies followed for the three research objectives. The analyses and their respective results will be presented in the same sequence in the next chapter.

6.2 Probability distribution analysis

In this section the results are presented corresponding to the sequence in the 'Methodology' chapter (Chapter 4). In part 1 the least significant potential statistics are identified and eliminated. In part 2 various scenarios were researched where the remaining statistics were combined with different measures of tendency. Part 3 covers the choice of the best scenario and the resultant parameters of the proposed probability distribution, while part 4 assess the performance of the newly defined probability distribution. The acronym of the proposed probability distribution is identified in advance, since it appears on numerous figures that will be presented as part of the results; it is defined as the Improved Probability-analysis, South Africa (IPZA) probability distribution.

6.2.1 Part 1: Screening of preliminary potential statistics

An MRA approach was used to screen the potential statistics (as listed in Table 6-1) to eliminate statistics that proved to be insignificant in improving the performance of the MRA prediction model. The screening of various scenarios of potential statistics (independent variables) was performed by simply comparing Q_{est} with the Q_{obs}, visually.

Numerous preliminary scenarios of the initial identified statistics (Q_{max}, Q_{min}, Q_{ave}, Q_{ave}*, Q_{ave}*º, Q_{mdn}, Q_{mdn}*, Q_{mdn}*º, Q_{gmn}, g, g*, g*º, SD, SD*, SD*º, SD**, kurt, kurt*, COV, COV*, COV*º, Z_{max} and Z_{min}) were researched and, to avoid cramming too much immaterial screening analyses results into the chapter, it is deemed sufficient to illustrate the important observations. As could be expected, variables like Q_{ave}, Q_{mdn}, Q_{gmn} and SD were considered as being significant. Unexpectedly, contrary to existing beliefs, the screening part suggested that the effect of including g might be trivial or even detrimental, rather than beneficial. It was hypothesised that to omit g would either have no effect or improve the results. Since the general slope of a probability distribution is related to the SD, and thus affected by outliers, it was furthermore hypothesised that replacing g with SD* might improve the outcome. Incidentally, SD*, SD*º and SD** were also considered as not being insignificant. The example in Figure 6-2, resulting from MRAs at one of the chosen AEPs, suggests that replacing g with SD* might improve results.

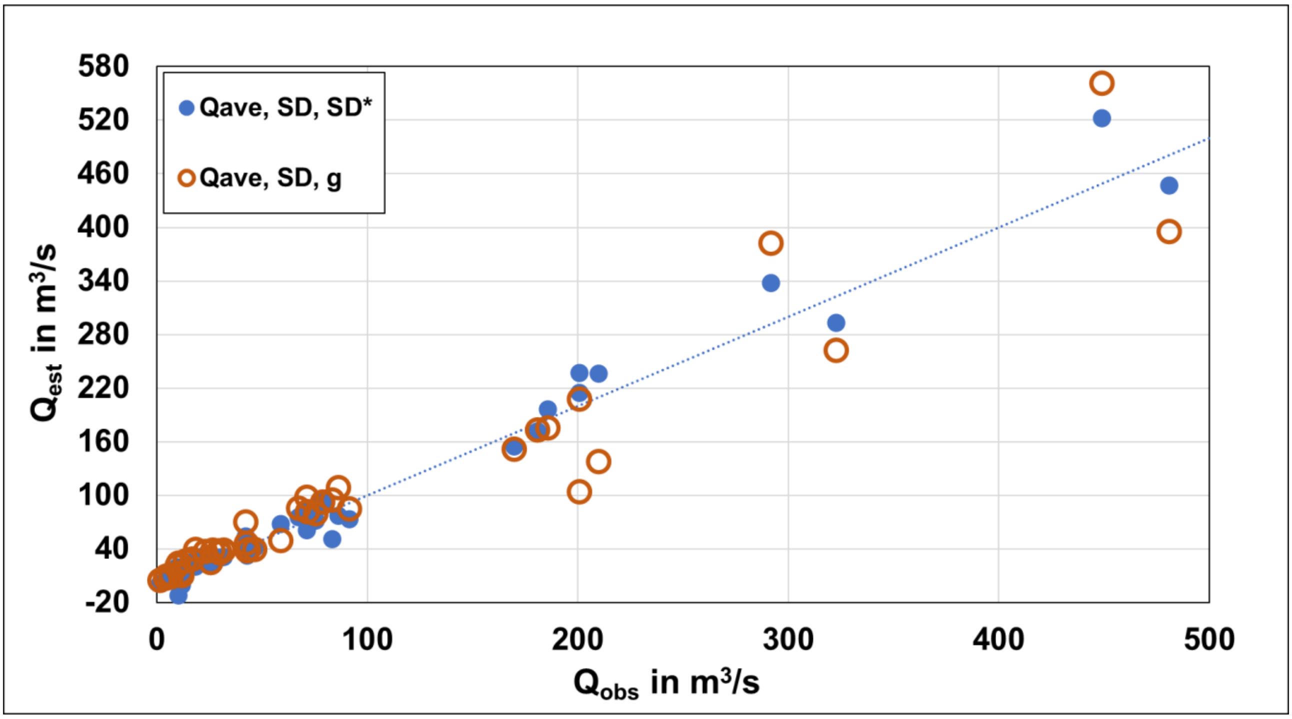


Figure 6-2: Comparing MRA results, relating to the inclusion of g vs the inclusion of SD*

The rest of the identified potential statistics were eliminated by the screening process; two illustrations are provided to indicate why these statistics were considered as insignificant. The first illustration, provided in Figure 6-3, involved the Q_{max} statistic, which typically improved the results at some flow sites but either had a negligible effect or worsened the results at most of the flow sites, across the chosen AEP range.

The second illustration, provided in Figure 6-4, involved the Z_{max} statistic, which typically had a negligible effect at most flow sites, but worsened the results at some of the flow sites, across the chosen AEP range.

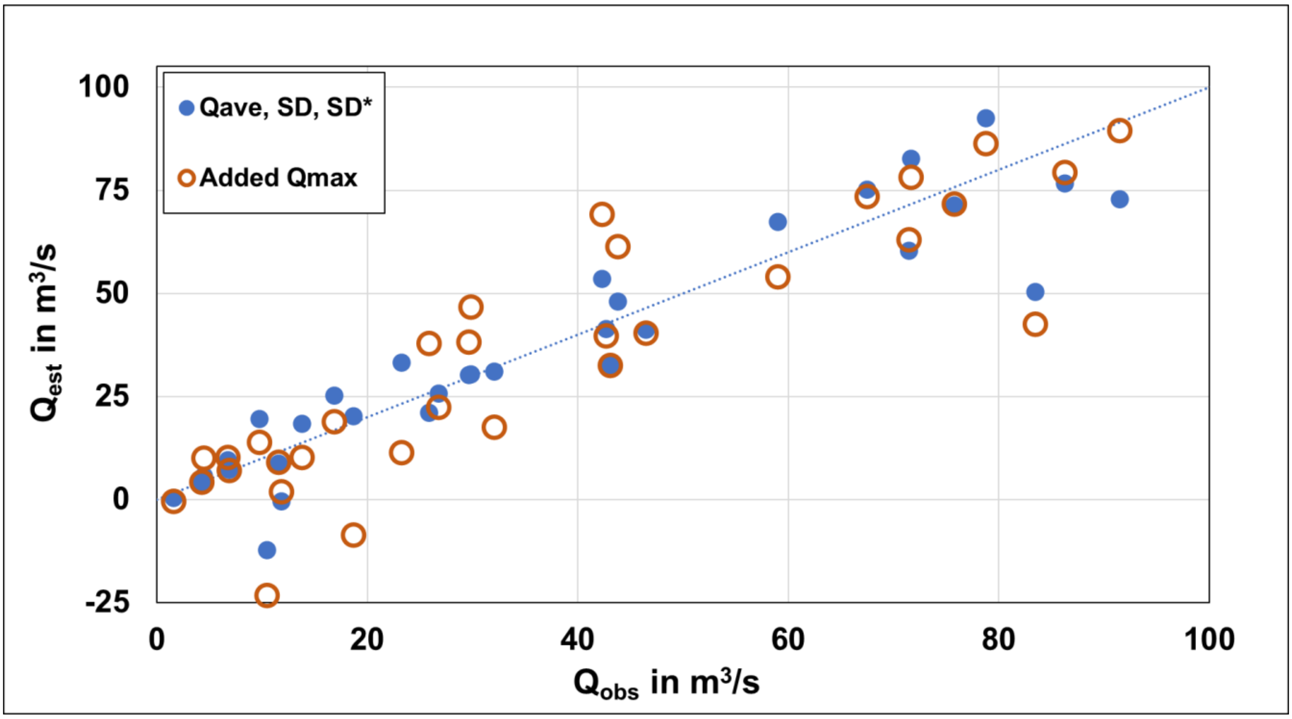


Figure 6-3: Screening illustration, where adding a parameter (Q_{max}) mostly worsened results

Similar impacts were observed when adding Q_{min}.

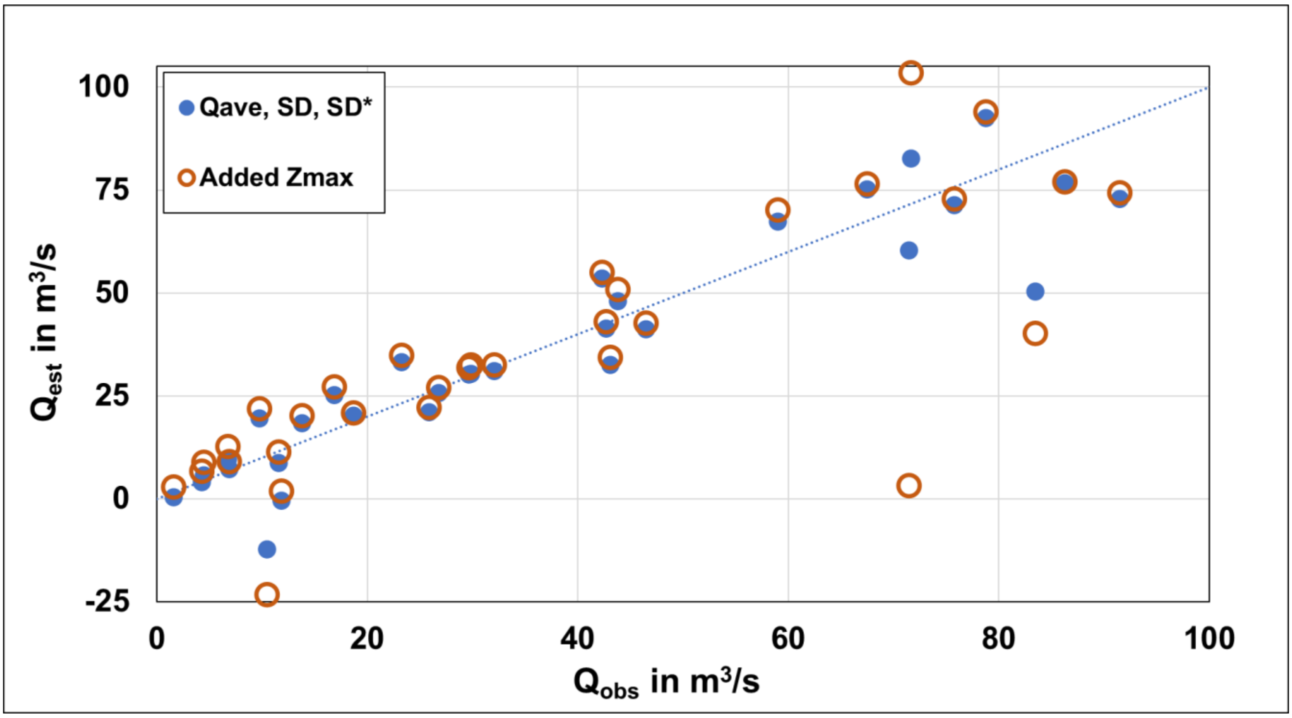


Figure 6-4: Screening illustration, where adding a parameter (Z_{max}) did not affect results

Similar impacts were observed when adding Z_{min}, COV, COV*, COV*°, kurt and kurt*

Similarly, the effect of replacing of Q_{ave} with Q_{ave}* or Q_{ave}*° had no or little effect on the outcome of the MRAs (the same is true of Q_{mdn}, Q_{gmn} and g).

6.2.2 Part 2: Finding the best combination of significant statistics

To find the best combination, different scenarios were used with one of the measures of central tendency as the integral statistic, supplemented by SD and g.

To test the hypothesis that replacing g with SD* may improve results, scenarios were added where g was replaced by SD*.

For ease of reference Table 6-4 is duplicated as Table 6-5, adding the extra scenarios and their abbreviations.

Table 6-5: Scenarios considered with different measures of central tendency

Scenarios considered with shown measure of central tendency:					
Average (Q_{ave})		Median (Q_{mdn})		Geometric mean (Q_{gmn})	
Scenario	**Statistics**	**Scenario**	**Statistics**	**Scenario**	**Statistics**
ave0	Q_{ave}, SD	mdn0	Q_{mdn}, SD	gmn0	Q_{gmn}, SD
ave1	Q_{ave}, SD, g	mdn1	Q_{mdn}, SD, g	gmn1	Q_{gmn}, SD, g
ave2	Q_{ave}, SD, SD*	mdn2	Q_{mdn}, SD, SD*	gmn2	Q_{gmn}, SD, SD*

MRAs were conducted at all selected AEPs for every scenario (a total of 198 MRAs). Owing to the vast amount of data and resulting information, only the essential results of the ultimate scenario, utilise to develop the new probability distribution, are presented in Section 6.2.3.

However, to comprehend the motivation for the choice of the final scenario, it is essential to present comparative performances of the different scenarios. The ranking results of the MAPE and MAE error testing models, at all the research flow sites, is presented in Appendix 5.3. The MAPE rankings are presented in Appendix 6-B, for AEPs ≤ 50% and in Appendix 6-C for AEPs > 50%. The MAE rankings are presented in Appendix 6-D, for AEPs ≤ 50% and in Appendix 6-E for AEPs > 50%.The summarised outcome of the rankings is presented in Table 6-6.

Table 6-6: Ranking results

AEP range	Q_{ave} scenarios			Q_{mdn} scenarios			Q_{gmn} scenarios		
	ave0	**ave1**	**ave2**	**mdn0**	**mdn1**	**mdn2**	**gmn0**	**gmn1**	**gmn2**
	Q_{ave}, SD	Q_{ave}, SD, g	Q_{ave}, SD, SD*	Q_{mdn}, SD	Q_{mdn}, SD, g	Q_{mdn}, SD, SD*	Q_{gmn}, SD	Q_{gmn}, SD, g	Q_{gmn}, SD, SD*
	MAPE Ranking (Appendix 6-B&C)								
≤ 50%	**3**	**5**	**1**	**7**	**9**	**4**	**8**	**6**	**2**
> 50%	**8**	**9**	**7**	**5**	**6**	**4**	**2**	**3**	**1**
	MAE Ranking (Appendix 6-D&E)								
≤ 50%	**5**	**4**	**1**	**8**	**6**	**3**	**9**	**7**	**2**
> 50%	**8**	**9**	**3**	**4**	**7**	**5**	**2**	**6**	**1**

Based on the ranking results, bearing in mind that the MAPE is considered to be more relevant for AEPs ≤ 50% and the MAE for AEPs > 50%, the following was noted:

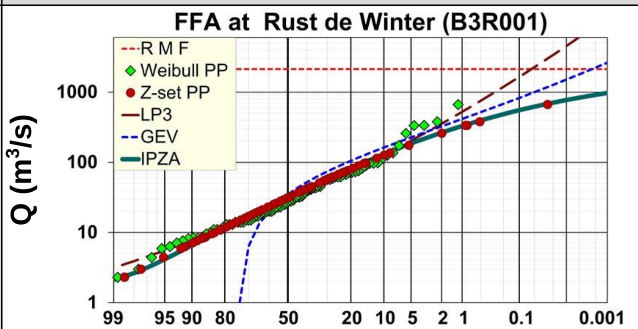
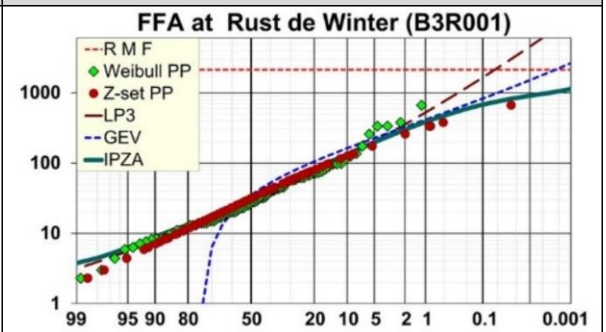
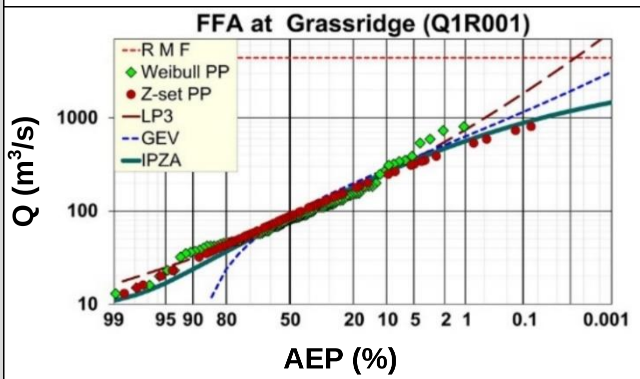
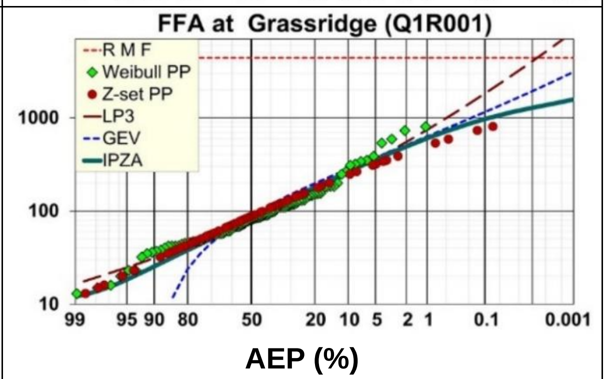
- As hypothesised, regardless of the measure of tendency used, the scenarios with the g statistic performed generally the worst – the exception being gmn1 for AEPs ≤ 50%.

- Regardless of the measure of tendency used, the scenarios with the SD* statistic performed overall the best – the exception being mdn0 for AEPs > 50% (only with MAE).
- The performance of the Q_{mdn} scenarios was unexpectedly disappointing since the median (as with the geometric mean) is generally considered to be closer to the estimated 50% AMS flood peak (ARI = 2 year).
- In the design flood AEP range (AEPs ≤ 50%) both MAPE and MAE indicated that the ave2 scenario (Q_{ave} SD and SD*) performed the best – with the gmn2 scenario 2nd best.
- For AEPs > 50% (ARI < 2 years) both MAPE and MAE indicated that the gmn2 scenario (Q_{gmn} SD and SD*) performed the best.

To illustrate some of the above findings, visual comparisons of scenarios are provided in the Appendix 6-F to Appendix 6-H, at all sites. Concise tables, indicating which scenario is categorised as noticeably better or only slightly better (a maximum of 2 examples, if available, per scenario), are depicted below (with the Appendix from which it was extracted in brackets), as follows:

- The suggestion, that the g-statistic does not add any value to the results, is visually illustrated in Table 6-7 (Appendix 6-F), supported by the estimated significance levels, at all AEPs for scenario ave1, shown in Table 6-8.
- The effect of introducing the SD* statistic is illustrated in Table 6-9 (Appendix 6-G), with the matching significance levels, for scenario ave2, shown in Table 6-10.
- The effect of replacing Q_{ave} with Q_{gmn} is illustrated in
- **Table 6-11** (Appendix 6-H).

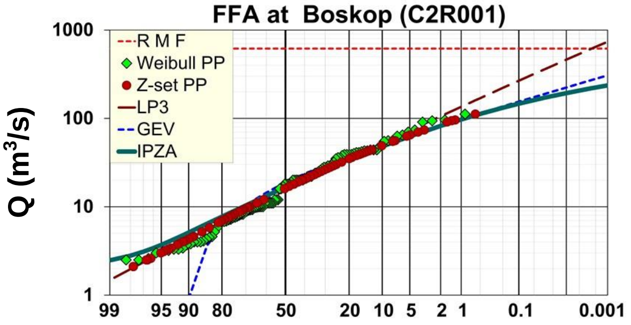
Table 6-7: Compare scenarios ave0 and ave1 to illustrate non-performance of g

<table>
<tr><th></th><th>Qave, SD (ave0)</th><th>Qave, SD, g (ave1)</th></tr>
<tr><td rowspan="2">ave0 slightly better</td><td>
</td><td>
</td></tr>
<tr><td>
</td><td>
</td></tr>
</table>

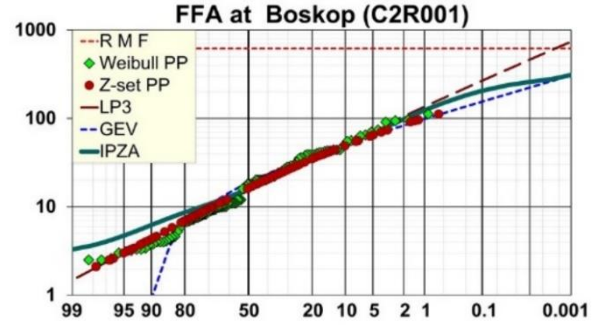
	Qave, SD (ave0)	Qave, SD, g (ave1)

ave0 noticeably better

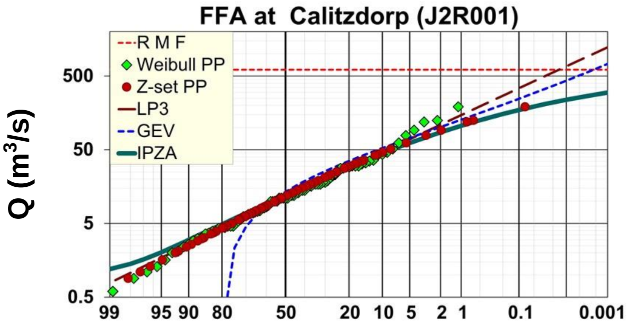
Qave, SD (ave0)



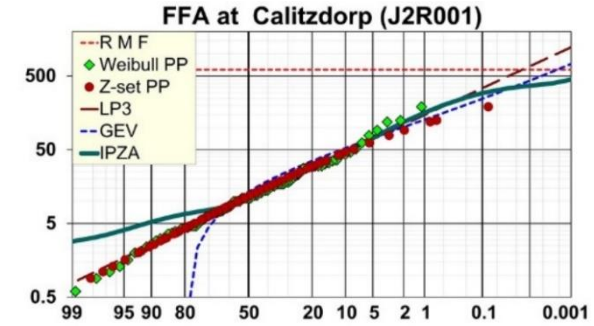
Qave, SD, g (ave1)



Qave, SD (ave0)

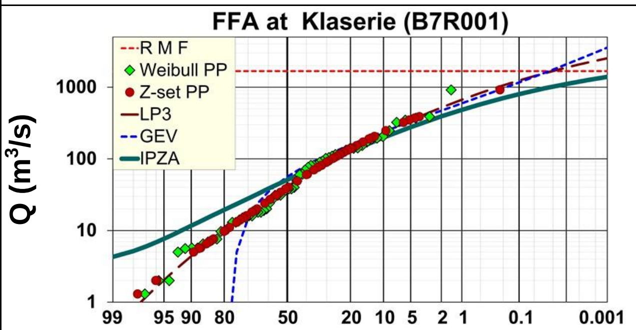


Qave, SD, g (ave1)

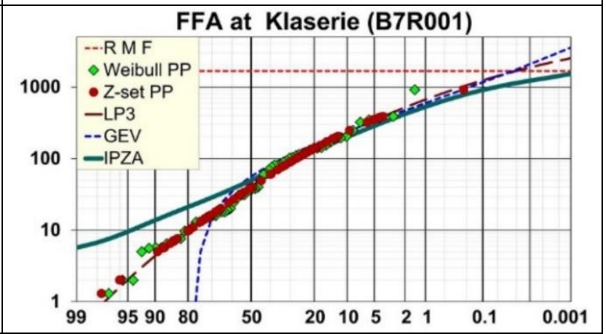


ave1 slightly better

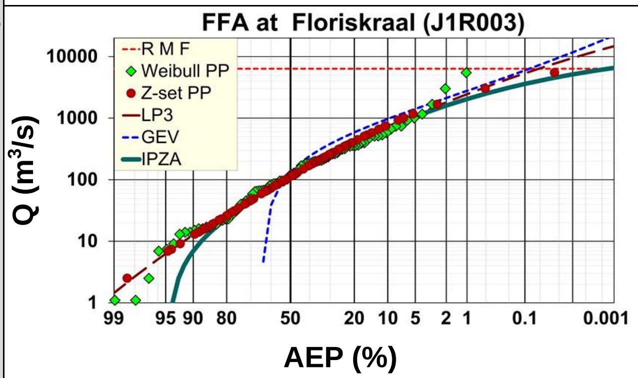
Qave, SD (ave0)



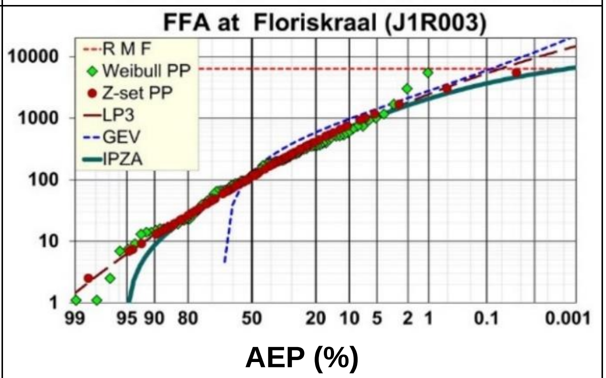
Qave, SD, g (ave1)



Qave, SD (ave0)



Qave, SD, g (ave1)

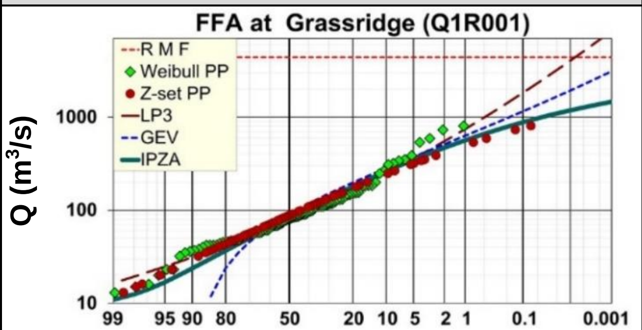
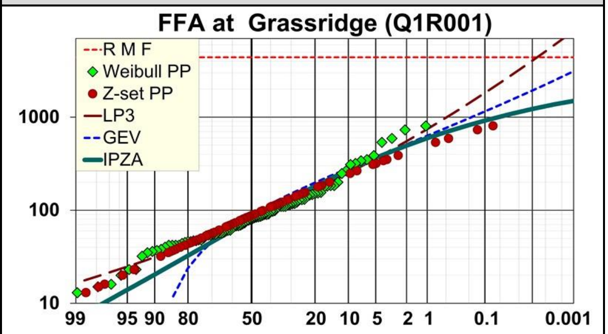
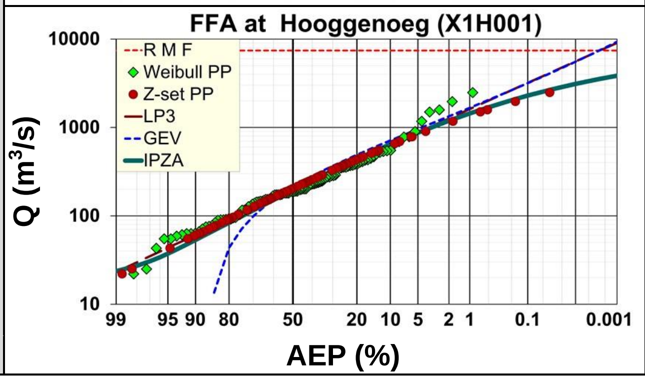
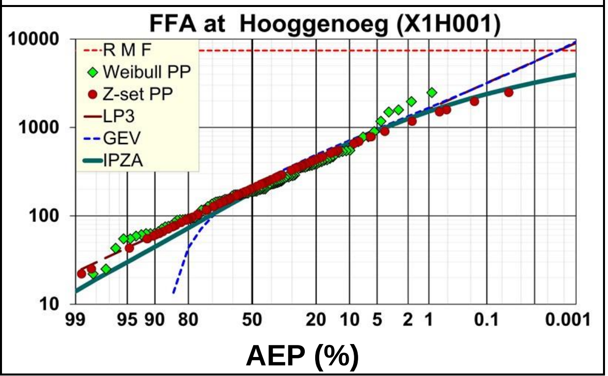


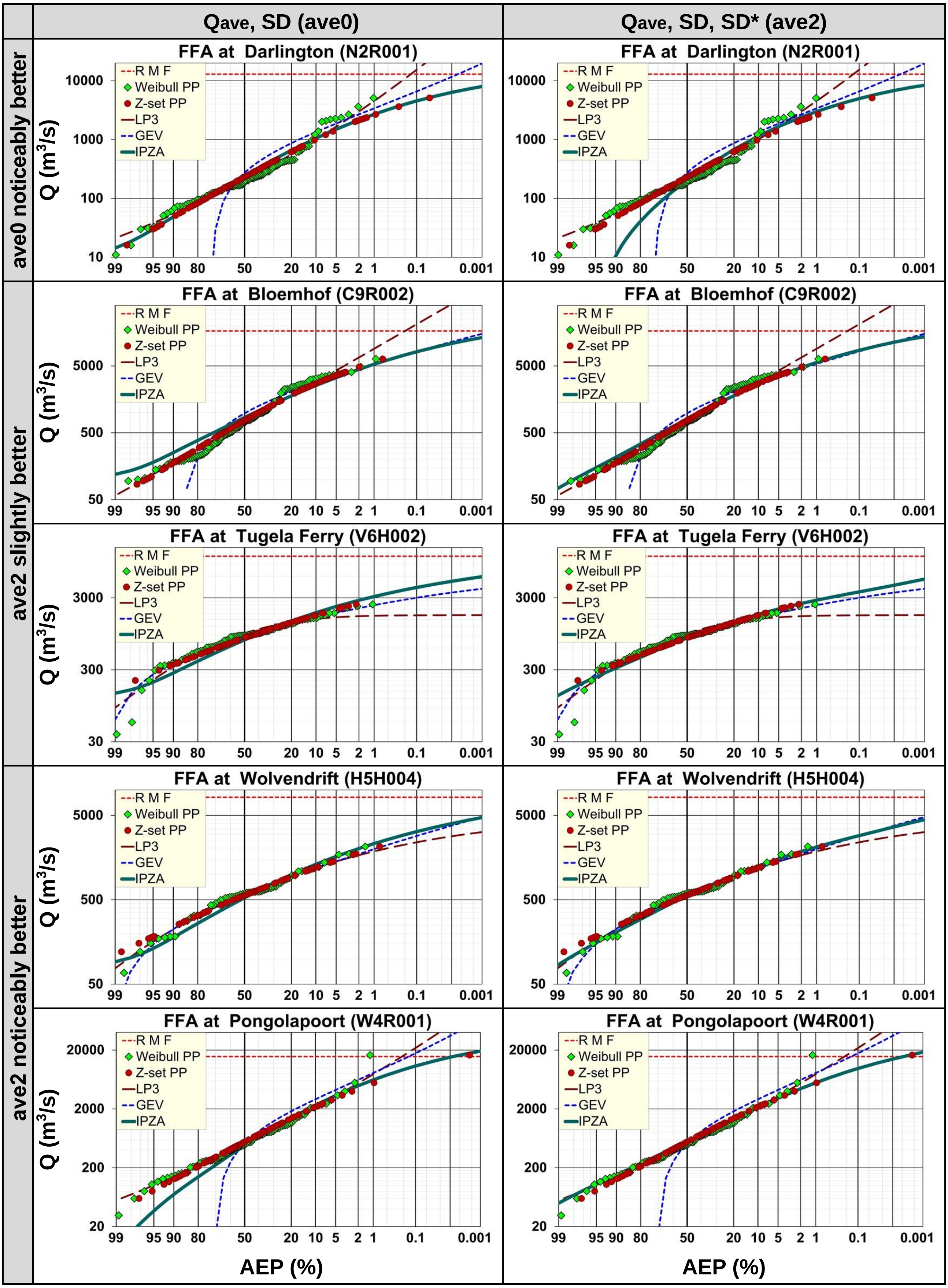
ave1 noticeably better:
None

Table 6-8: Levels of significance for scenario ave1

AEP (%)	R^2	Levels of Significance				Shading indicating evidence of significance
		F-value	P-value			
			Q_{ave}	SD	g	
0.995	0.9080	2.042E-19	1.012E-14	4.381E-07	0.4766	
99	0.9229	7.776E-21	9.078E-16	1.524E-07	0.4624	Very strong evidence
95	0.9466	8.716E-24	1.272E-17	6.450E-08	0.5452	
90	0.9598	4.640E-26	4.667E-19	4.037E-08	0.6172	Strong evidence
80	0.9748	8.253E-30	9.197E-22	7.234E-09	0.6745	
70	0.9851	4.673E-34	4.501E-25	6.055E-10	0.8068	Moderate evidence
60	0.9918	8.056E-39	6.871E-29	4.196E-11	0.9981	
50	0.9963	3.108E-45	1.821E-34	1.948E-13	0.8031	Weak evidence
40	0.9986	8.220E-53	4.327E-41	1.030E-15	0.5573	
30	0.9997	1.967E-64	9.802E-52	1.626E-19	0.1492	Insufficient evidence
20	0.9995	7.675E-61	2.515E-46	2.691E-04	0.5708	
10	0.9973	8.345E-48	3.775E-30	5.875E-05	0.8985	
5	0.9962	5.518E-45	6.755E-25	4.938E-09	0.6267	
2	0.9943	9.466E-42	1.211E-18	1.349E-11	0.4455	
1	0.9941	1.834E-41	2.357E-16	1.376E-13	0.3663	
0.5	0.9940	2.606E-41	3.863E-14	2.987E-15	0.1972	
0.2	0.9940	2.435E-41	8.557E-12	5.563E-17	0.1166	
0.1	0.9944	6.902E-42	1.281E-10	1.814E-18	0.0989	
0.03	0.9953	3.018E-43	4.636E-09	4.447E-21	0.0816	
0.01	0.9963	3.810E-45	2.374E-08	8.437E-24	0.1067	
0.003	0.9973	7.250E-48	4.588E-08	3.523E-27	0.0729	
0.001	0.9981	1.149E-50	3.761E-08	2.009E-30	0.0276	

Table 6-9: Compare scenarios ave0 and ave2 to illustrate effect of introducing SD*

<table>
<tr><th></th><th>Qave, SD (ave0)</th><th>Qave, SD, SD* (ave2)</th></tr>
<tr><td rowspan="2">ave0 slightly better</td><td>
</td><td>
</td></tr>
<tr><td>
</td><td>
</td></tr>
</table>

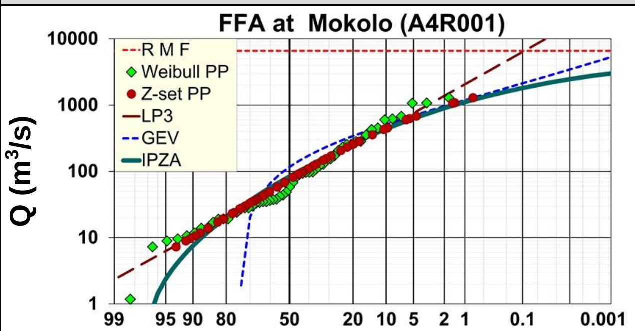
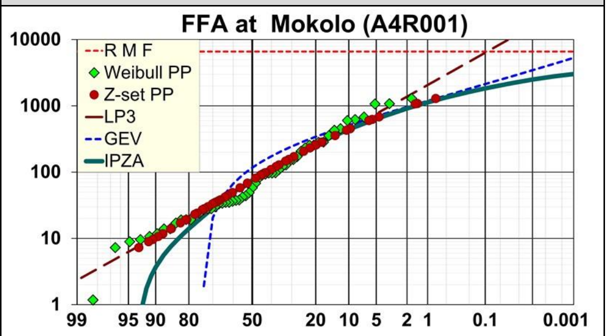
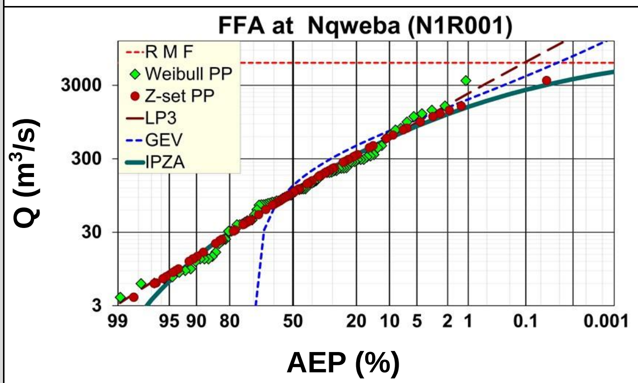
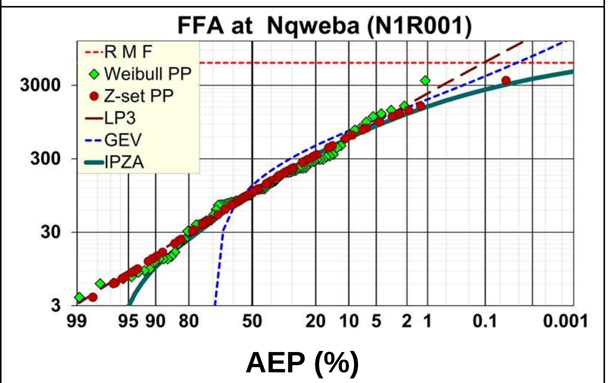


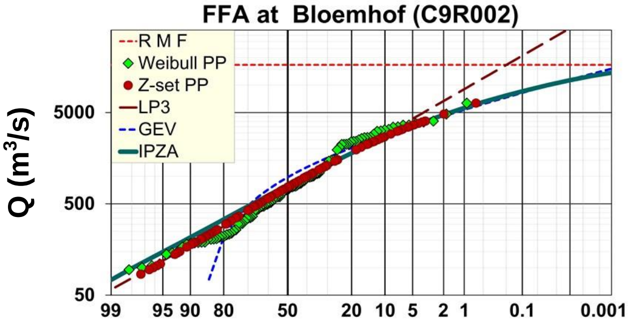
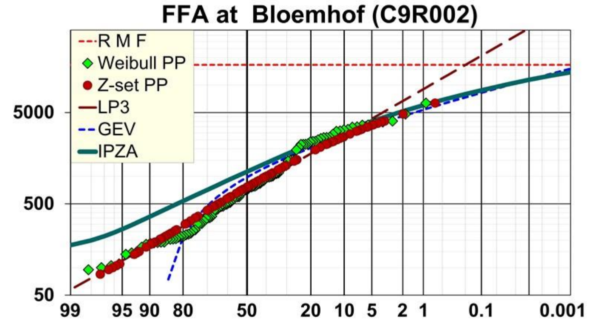
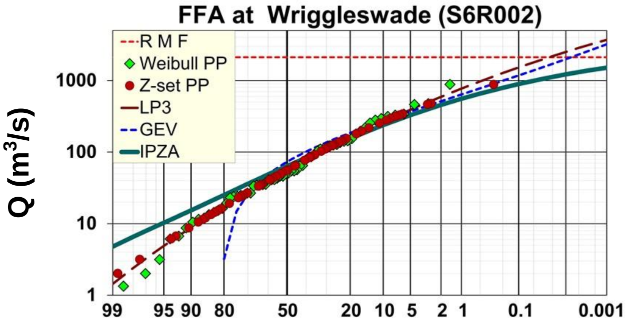
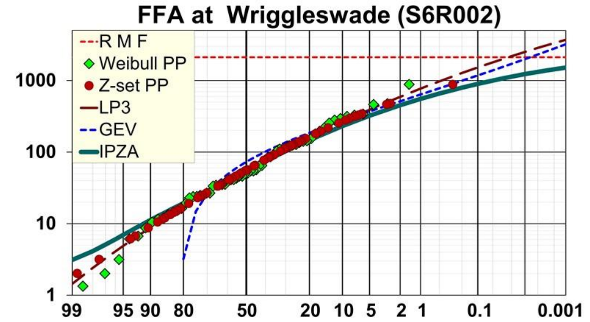
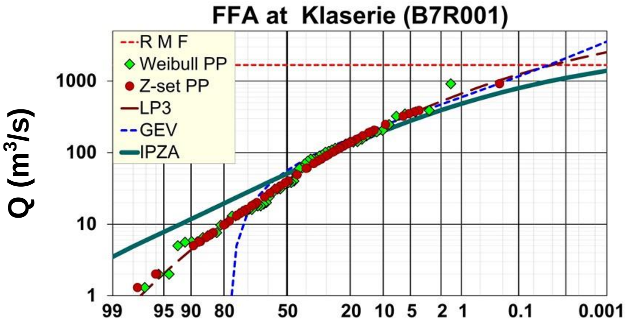
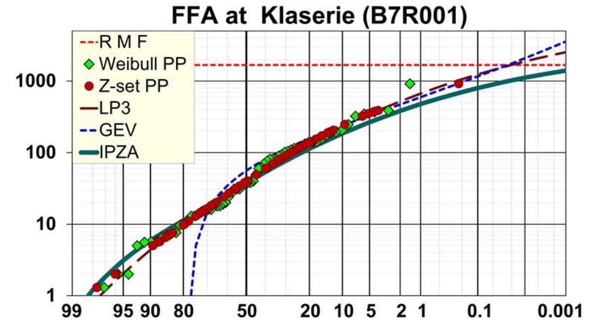
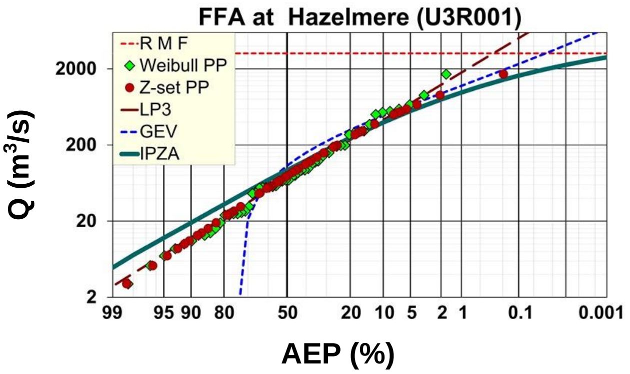
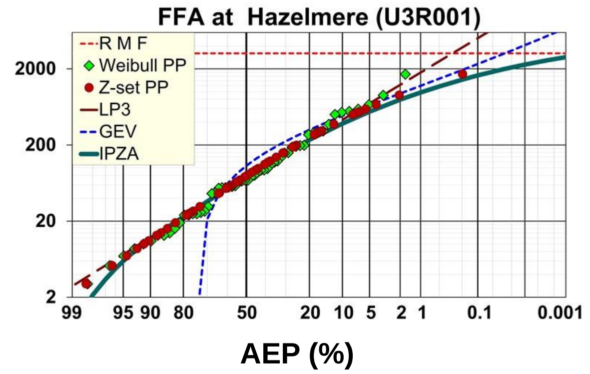
Qave, SD (ave0)
Qave, SD, SD* (ave2)
ave0 noticeably better
ave2 slightly better
ave2 noticeably better
FFA at Darlington (N2R001)
FFA at Bloemhof (C9R002)
FFA at Tugela Ferry (V6H002)
FFA at Wolvendrift (H5H004)
FFA at Pongolapoort (W4R001)
R M F
Weibull PP
Z-set PP
LP3
GEV
IPZA
Q (m³/s)
99 95 90 80 50 20 10 5 2 1 0.1 0.001
AEP (%)

Table 6-10: Levels of significance for scenario ave2

AEP (%)	R^2	Levels of Significance				Shading indicating evidence of significance
		F-value	P-value			
			Qave	SD	SD*	
0.995	0.9436	2.389E-23	2.540E-16	0.3070	1.399E-05	
99	0.9566	1.907E-25	6.040E-18	0.2688	2.621E-06	Very strong evidence
95	0.9716	7.239E-29	5.926E-20	0.1810	9.154E-07	
90	0.9793	2.040E-31	2.040E-21	0.1356	4.976E-07	Strong evidence
80	0.9876	1.698E-35	5.838E-24	0.0449	2.326E-07	
70	0.9932	2.341E-40	2.170E-27	0.0053	5.612E-08	Moderate evidence
60	0.9967	3.888E-46	9.132E-32	0.0002	5.009E-09	
50	0.9987	5.997E-54	3.787E-38	7.648E-08	1.696E-10	Weak evidence
40	0.9996	4.649E-63	9.420E-46	3.225E-12	3.760E-12	
30	0.9998	4.640E-67	1.261E-47	3.364E-14	0.0002	Insufficient evidence
20	0.9995	7.362E-62	1.816E-39	5.990E-06	0.0243	
10	0.9986	4.134E-53	2.214E-25	0.7256	6.368E-07	
5	0.9984	6.747E-52	4.795E-20	0.0129	1.171E-08	
2	0.9978	2.762E-49	7.303E-12	1.475E-05	2.196E-09	
1	0.9977	3.291E-49	1.244E-08	2.568E-08	1.173E-09	
0.5	0.9973	9.798E-48	3.222E-05	2.391E-10	1.816E-08	
0.2	0.9972	1.612E-47	0.0129	1.388E-12	2.173E-08	
0.1	0.9973	1.037E-47	0.0925	2.520E-14	4.477E-08	
0.03	0.9974	4.166E-48	0.4264	5.019E-17	3.921E-07	
0.01	0.9982	4.292E-51	0.8770	5.587E-21	3.577E-08	
0.003	0.9986	4.718E-53	0.9665	1.944E-24	1.732E-07	
0.001	0.9987	1.647E-53	0.6227	2.626E-26	2.243E-05	

Table 6-11: Compare scenarios ave2 and gmn2 to illustrate effect if Q_{gmn} replaces Q_{ave}

	Q_{ave}, SD, SD* (ave2)	Q_{gmn}, SD, SD* (gmn2)
ave2 slightly better		
		

	Qave, SD, SD* (ave2)	Qgmn, SD, SD* (gmn2)
	ave2 noticeably better: None	
gmn2 slightly better		
		
gmn2 noticeably better		
		

Consideration was given to combine the results of ave2 (for AEPs ≤ 50%) and gmn2 (for AEPs > 50%). However, this created an uneven transition in the region where the AEP = 50%.

To appreciate the real difference between ave2 and gmn2, the difference between the 2 scenarios was evaluated according to the rankings (Appendix 6-B to Appendix 6-E), as well as visually (Appendix 6-H), for every flow site. Only the AEP > 50% range was considered, and the outcome is presented in Table 6-12, indicating the individual sites where differences were identified.

Table 6-12: Outcome of comparing ave2 with gmn2 at individual flow sites for AEP ≥ 50%

	ave2 the better fit		gmn2 the better fit			
	noticeably	slightly	slightly		noticeably	
Sites where differences were identified	None	A4R001 A8R001 C5R003 J1R003 N1R001	C9R002 D6R002 D7H005 S6R002 T2R001	C5R001* C5R002* N2R001* Q1R001*	B1R001 B7R001 U3R001	X1H001*
Total number	**0**	**5**	**5**	**4**	**3**	**1**

* Indicates that the better fit at AEP ≥ 50% becomes irrelevant, since the ave2 fit at AEP ≤ 50% is better

Essentially, both ave2 and gmn2 provided slightly better fits, at 5 sites each, and at only 3 sites was gmn2 found to provide a noticeably better fit than ave2 to the Q_{obs}. The ave2 ranked 1st in both the MAPE and MAE comparisons in the design flood range, and 3rd in the AEP > 50% range (re MAE).

To confirm that scenario ave2 might not further be improved, by inadvertently excluding more pertinent statistics, it was compared with scenario ave3 (Q_{ave}, SD, SD*°) and scenario ave4 (Q_{ave}, SD, SD**) and the results are presented in Table 6-13, where the MAE was used in the AEP range > 50% and MAPE was used in the AEP range ≤ 50%.

SD*° was calculated by excluding the highest and lowest flood peaks from the AMS dataset and SD** was calculated by excluding the highest 2 flood peaks from the AMS dataset.

Table 6-13: Ranking results of various scenarios of Q_{ave}

Site	MAE used for AEP range > 50%							MAPE used for AEP range ≤ 50%						
	Scenarios							Scenarios						
	ave2	ave3	ave4		ave2	ave3	ave4	ave2	ave3	ave4		ave2	ave3	ave4
	Qave, including				Ranking			Qave, including				Ranking		
					1	2	3					1	2	3
	SD, SD*	SD, SD*°	SD, SD**	Σ	45	89	112	SD, SD*	SD, SD*°	SD, SD**	Σ	78	80	88
A3R002	1.0	2.2	2.3		1	2	3	11%	11%	12%		2	1	3
B1R001	9.0	15.4	14.4		1	3	2	6.0%	5.9%	5.3%		3	2	1
B2R001	1.5	5.4	8.7		1	2	3	1.5%	1.6%	2.3%		1	2	3
C1R002	8.4	14.9	18.0		1	2	3	3.5%	3.5%	3.1%		2	3	1
C2R001	0.3	0.6	0.7		1	2	3	3.5%	3.4%	4.0%		2	1	3
C5R002	53.1	78.4	49.1		2	3	1	5.8%	5.9%	3.4%		2	3	1
C9R002	35.0	49.6	48.9		1	3	2	2.0%	2.1%	2.1%		1	3	2
D3R002	35.7	65.4	76.6		1	2	3	1.7%	1.7%	1.4%		2	3	1
D7H005	48.5	61.5	78.1		1	2	3	2.7%	2.7%	3.2%		2	1	3
J1R003	5.0	17.9	24.7		1	2	3	9.4%	9.5%	15%		1	2	3
J2R001	0.2	0.8	1.1		1	2	3	1.4%	7.0%	7.9%		1	2	3
N2R001	31.8	49.4	57.0		1	2	3	11%	11%	9.8%		2	3	1
Q1R001	7.7	10.3	9.6		1	3	2	13%	13%	11%		2	3	1
V6H002	27.7	29.4	28.2		1	3	2	4.2%	4.3%	4.0%		2	3	1
X1H001	14.7	21.6	21.4		1	3	2	9.0%	9.0%	7.4%		3	2	1
A2R001	5.8	9.1	11.3		1	2	3	2.8%	2.7%	3.2%		2	1	3
A2R005	1.9	2.1	2.3		1	2	3	14%	14%	14%		2	3	1
A4R001	5.2	10.9	13.4		1	2	3	8.1%	8.1%	8.9%		2	1	3
A8R001	2.2	9.9	14.4		1	2	3	4.0%	3.9%	6.6%		2	1	3
B3R001	1.7	3.6	4.7		1	2	3	4.6%	4.7%	5.0%		1	2	3
B6R003	9.8	10.7	13.3		1	2	3	4.6%	4.5%	4.6%		2	1	3
B7R001	8.0	9.1	8.8		1	3	2	16%	16%	15%		3	2	1
C3R002	3.1	7.6	10.1		1	2	3	4.7%	4.7%	5.2%		2	1	3
C5R001	8.6	13.4	16.7		1	2	3	6.1%	6.2%	7.0%		1	2	3
C5R003	2.5	8.6	16.3		1	2	3	3.9%	3.7%	4.6%		2	1	3
C7R001	3.9	8.6	11.2		1	2	3	4.9%	4.9%	4.6%		2	3	1
D6R002	3.0	6.7	13.7		1	2	3	8.8%	8.6%	16%		2	1	3
E1R002	11.5	12.3	15.1		1	2	3	4.0%	4.0%	3.8%		2	3	1
H5H004	24.0	20.7	20.6		3	2	1	2.2%	2.8%	3.7%		1	2	3
J2R003	0.9	1.6	1.9		1	2	3	9.2%	9.1%	10.0%		2	1	3
J3R001	25.0	35.0	38.6		1	2	3	11%	12%	9.3%		2	3	1
N1R001	3.1	11.8	16.5		1	2	3	7.3%	7.3%	6.2%		3	2	1
Q4R002	1.6	3.4	4.0		1	2	3	1.7%	1.6%	2.2%		2	1	3
Q5R001	5.0	18.0	31.4		1	2	3	1.2%	1.3%	2.0%		1	2	3
S6R002	6.5	7.8	8.0		1	2	3	9.0%	8.9%	8.5%		3	2	1
T2R001	22.4	20.4	22.5		2	1	3	7.1%	7.1%	7.6%		1	2	3
U2R001	6.2	6.4	7.5		1	2	3	10%	10%	8.8%		2	3	1
U3R001	9.3	13.0	14.6		1	2	3	9.8%	9.6%	9.8%		2	1	3
V1R003	11.4	12.9	22.4		1	2	3	1.6%	1.5%	2.2%		2	1	3
V3R003	4.8	5.6	5.0		1	3	2	4.9%	4.9%	3.2%		2	3	1
W4R001	22.2	45.9	53.3		1	2	3	2.6%	2.3%	3.8%		2	1	3

The final choice was primarily guided by the relevance of the results in the design flood range, as well as avoiding unnecessarily complicated solutions. Accordingly, ave2 was selected as the best scenario.

6.2.3 Part 3: Probability distribution parameters and frequency factors

Consequently, from the chosen scenario ave2, the parameters for the probability distribution are Q_{ave}, SD and SD*. Equation **6-5** can be rephrased as the general form of the prediction equation for the IPZA probability distribution:

$$Q_{AEP} = K_{P\text{-}Q} \cdot Q_{ave} + K_{P\text{-}SD} \cdot SD + K_{P\text{-}SD^*} \cdot SD^* \qquad \textbf{Equation 6-6}$$

where

$K_{P\text{-}Q}$, $K_{P\text{-}SD}$ and $K_{P\text{-}SD^*}$ are the frequency factors, associated with the corresponding parameters.

Estimated regression coefficients, for the independent variables used in the MRAs of scenario ave2, are provided in Table 6-14, for every chosen AEP with associated standardised variates W_T and W_P, respectively used in the LN and EV1 probability distributions.

Table 6-14: Estimated MRA regression coefficients at corresponding AEPs

	Standardised variates		**Independent variables (parameters)**		
	LN	**GEV (EV1)**	**Q_{ave}**	**SD**	**SD***
AEP (%)	**W_T**	**W_P**	**MRA regression coefficients**		
99.5	- 2.5758	-1.6674	0.1756	- 0.0132	-0.1221
99	- 2.3263	-1.5272	0.2225	- 0.0155	-0.1540
95	- 1.6449	-1.0972	0.3889	- 0.0276	-0.2509
90	- 1.2816	-0.8340	0.5021	- 0.0361	-0.3035
80	- 0.8416	-0.4759	0.6779	- 0.0567	-0.3603
70	- 0.5244	-0.1856	0.8308	- 0.0798	-0.3807
60	- 0.2533	0.0874	0.9688	- 0.0977	-0.3757
50	**-**	**0.3665**	**1.1036**	**- 0.1240**	**-0.3301**
40	0.2533	0.6717	1.2365	- 0.1327	-0.2659
30	0.5244	1.0309	1.3645	- 0.1548	-0.1039
20	0.8416	1.4999	1.4769	- 0.1294	0.1416
10	1.2816	2.2504	1.5089	- 0.0275	0.7250
5	1.6449	2.9702	1.4806	0.1991	1.2545
2	2.0537	3.9019	1.2683	0.6210	2.0693
1	2.3263	4.6001	1.1443	1.0642	2.5224
0.5	2.5758	5.2958	0.9282	1.6532	2.8720
0.2	2.8782	6.2136	0.6254	2.4266	3.3928
0.1	3.0902	6.9073	0.4426	3.0775	3.6379
0.03	3.4316	8.1116	0.2181	4.3237	3.7631
0.01	3.7190	9.2103	0.0025	5.3624	4.0032
0.003	4.0128	10.4143	-0.0513	6.6915	3.7216
0.001	4.2649	11.5129	0.0838	8.0279	3.0235

The comparison of Q_{obs} with Q_{est} results, from the MRA for scenario ave2, are provided in Appendix 6-I and a concise version of it is presented in Table 6-15.

Table 6-15: MRA estimates; Q_{est} compared to Q_{obs} for scenario ave2

Site	Moments			AEP (%) 50		20		10		5		2		1*		0.5	
	Q and SD values in m³/s																
	Q_{ave}	SD	SD*	Q_{obs}	Q_{est}	Q_{obs}	Q_{est}	Q_{obs}	Q_{est}	Q_{obs}	Q_{est}	Q_{obs}	Q_{est}	Q_{obs}	Q_{est}	Q_{obs}	Q_{est}
A3R002	37.4	63.3	44.7	16.5	18.7	54.3	53.4	95.4	87.1	140	124	208	179	262	223	317	268
B1R001	280	384	317	141	156	418	408	683	642	969	889	1 384	1 250	1 620	1 529	1 793	1 806
B2R001	144	188	171	80.9	79.5	215	213	334	336	470	465	662	653	809	796	956	935
C1R002	494	421	388	373	365	721	730	983	1 015	1 260	1 301	1 614	1 690	1 888	1 992	2 162	2 268
C2R001	23.2	22.0	20.4	16.1	16.2	35.0	34.4	49.2	49.3	64.7	64.4	86.6	85.4	101	102	118	117
C5R002	547	1 223	824	236	180	738	766	1 301	1 389	1 973	2 087	2 985	3 158	3 821	4 006	4 700	4 896
C9R002	1 216	1 251	1 154	764	806	1 827	1 797	2 703	2 636	3 611	3 496	4 789	4 706	5 749	5 633	6 591	6 510
D3R002	2 614	2 094	1 928	2 024	1 990	3 847	3 863	5 209	5 285	6 575	6 707	8 377	8 606	9 852	10 083	11 146	11 425
D7H005	1 730	1 690	1 540	1 157	1 191	2 596	2 554	3 810	3 680	5 001	4 830	6 742	6 430	7 984	7 662	9 247	8 822
J1R003	298	665	398	112	115	421	410	783	720	1 141	1 072	1 753	1 613	2 360	2 051	2 991	2 517
J2R001	20.4	29.1	22.8	11.4	11.4	29.8	29.6	45.8	46.6	64.6	64.7	91.4	91.2	113	112	138	133
N2R001	467	796	641	229	205	618	677	1 005	1 147	1 465	1 653	2 129	2 412	2 634	2 998	3 156	3 590
Q1R001	123	139	121	87.1	79.2	174	181	244	270	317	362	418	493	500	594	579	691
V6H002	949	492	469	841	831	1 360	1 404	1 719	1 759	2 064	2 092	2 507	2 481	2 835	2 794	3 159	3 043
X1H001	302	371	308	202	186	431	442	628	669	837	907	1 129	1 251	1 358	1 518	1 591	1 779
A2R001	278	301	267	177	181	422	409	630	604	841	806	1 137	1 092	1 364	1 312	1 591	1 523
A2R005	33.9	49.8	29.6	22.0	21.5	51.3	47.9	80.5	71.3	112	97.4	160	135	223	167	262	199
A4R001	181	281	240	73.4	85.3	261	265	443	439	649	625	953	901	1 204	1 112	1 438	1 322
A8R001	227	409	278	103	109	319	322	541	533	742	766	1 173	1 117	1 454	1 396	1 795	1 685
B3R001	58.5	96.5	70.7	29.7	29.2	79.3	83.9	129	137	181	194	258	280	332	348	395	417
B6R003	283	282	233	198	201	427	415	630	589	820	768	1 089	1 016	1 299	1 212	1 474	1 398
B7R001	89.4	137	92.6	38.7	51.1	138	127	236	198	350	276	466	390	558	482	693	575
C3R002	192	318	178	112	114	273	267	421	409	587	570	824	809	1 027	1 007	1 449	1 215
C5R001	101	226	156	33.8	32.0	133	142	237	260	357	391	534	592	680	750	834	916
C5R003	193	336	226	89.0	96.6	256	273	432	446	628	636	897	921	1 120	1 148	1 354	1 383
C7R001	251	264	234	173	168	368	370	527	541	694	718	928	966	1 105	1 158	1 282	1 341
D6R002	122	195	147	56.2	61.4	185	175	313	285	404	403	678	580	863	718	1 043	858
E1R002	408	309	291	325	316	591	604	785	819	978	1 031	1 232	1 312	1 425	1 530	1 619	1 725
H5H004	661	409	367	563	557	972	975	1 194	1 252	1 519	1 520	1 815	1 852	2 051	2 117	2 352	2 344
J2R003	17.5	32.9	26.8	6.00	6.44	23.3	25.4	43.7	45.0	65.7	66.1	98.3	98.1	129	123	159	148
J3R001	184	428	347	45.0	35.9	203	266	406	518	665	793	1 068	1 217	1 409	1 541	1 756	1 875
N1R001	243	455	301	104	112	342	342	597	572	898	827	1 342	1 213	1 685	1 521	2 037	1 841
Q4R002	142	147	130	98.1	95.9	212	209	301	305	397	403	531	541	642	647	736	749
Q5R001	452	639	538	241	243	646	662	1 033	1 055	1 441	1 472	2 076	2 084	2 518	2 555	2 962	3 022
S6R002	108	146	110	55.1	64.4	162	156	266	238	376	326	511	454	598	555	730	656
T2R001	285	237	215	199	215	454	421	652	580	852	740	1 083	954	1 235	1 121	1 374	1 274
U2R001	166	202	93.8	122	127	239	232	335	313	495	403	694	530	827	641	948	757
U3R001	174	279	192	77.7	94.4	251	248	426	394	614	554	911	790	1 143	979	1 378	1 173
V1R003	566	437	401	429	438	848	837	1 126	1 133	1 447	1 428	1 875	1 819	2 176	2 124	2 367	2 399
V3R003	102	73.9	66.3	84.5	81.4	147	150	188	200	237	249	290	312	334	362	377	407
W4R001	1 003	1 931	929	546	561	1 376	1 364	2 150	2 135	3 049	3 036	4 386	4 395	5 500	5 547	6 675	6 793

* The estimation of the value in the yellow highlighted cell, is illustrated below

To illustrate the estimation of the 1% flood peak at B1R001:

- Relevant statistics for B1R001: Q_{ave} = 280 m³/s, SD = 384 m³/s and SD* = 317 m³/s (Table 3-3)
- At AEP = 1%: the regression coefficients, respectively, are 1.1443, 1.0642 and 2.5224 (Table 6-14)
- Therefore, Q_{est} = 1.1443 * 280 + 1.0642 * 384 + 2.5224 * 317 = 1528.66 m³/s (Equation 6-6)

The regression coefficients become the frequency factors and were fitted against the standardised variates to enable the estimate of intermediate values. Both standardised variates (W_T and W_P) were considered. Since W_T = 0 at AEP = 50%, it occasionally caused some discontinuity problems if, when experimenting with the equation, the equation is divided by the standardised variates. However, considering the format of the eventual equation, it is not a problem and both standardised variates produced fairly accurate fits against the regression coefficients for the respective variables (Q_{ave}, SD and SD*). The W_P was chosen, merely because of the simplicity of its equation. The fitted graphs, showing the relationships between W_P and the estimated frequency factors $K_{P\text{-}Q}$, $K_{P\text{-}SD}$ and $K_{P\text{-}SD*}$, respectively associated with Q_{ave}, SD and SD*, are provided in Appendix 6-J and, as an example, the graph for $K_{P\text{-}SD}$ is presented in Figure 6-5.

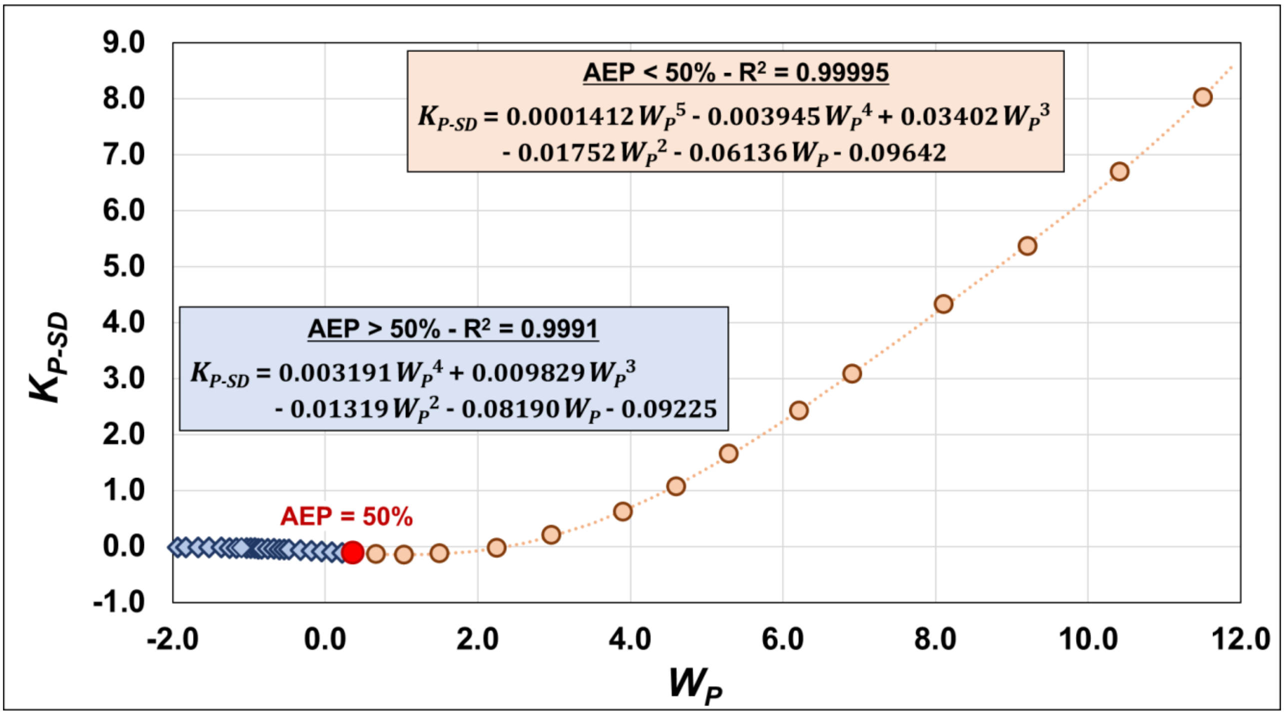


Figure 6-5: Relationship between W_P and estimated frequency factor for SD ($K_{P\text{-}SD}$)

A comprehensive table, providing frequency factors for AEP-values between 99.5% and 0.001%, is accessible in Appendix 6-K, with a concise version thereof presented as Table 6-16.

Table 6-16: Frequency factors for IPZA

ARI (y)	2	5	10	20	50	100	200	**500**	**1 000**	2 000	5 000	10 000
AEP (%)	**50**	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
W_P	0.3665	1.4999	2.2504	2.9702	3.9019	4.6001	5.2958	6.2136	6.9073	7.6007	8.5171	9.2103
Frequency factors												
K_{P-Q}	1.1035	1.4673	1.5258	1.4791	1.3099	1.1296	0.9249	0.6444	0.4429	0.2641	0.0803	-0.0082
K_{P-SD}	-0.1216	-0.1320	-0.0286	0.1838	0.6317	1.0865	1.6253	2.4345	3.0952	3.7787	4.6980	5.4022
K_{P-SD*}	-0.3379	0.1553	0.7155	1.3020	2.0310	2.5124	2.9205	3.3465	3.5892	3.7695	3.9131	3.9379

6.2.4 Part 4: Supplementary assessment of the IPZA probability distribution

In this part a brief assessment is done on the tendency of IPZA to under- or overestimate, and on the performance of IPZA in relation to outliers.

Under- or overestimation

It is beneficial to identify whether a particular probability distribution tends to overestimate or underestimate. Analogue to Section 4.1.1, applicable criteria are not readily available and appears to depend on the results of the research/study. Haddad and Rahman (2012), considering the results for their study, suggested that a Q_{est}/Q_{obs} ratio value of between 0.5 and 2 may be considered as acceptable. However, an underestimation of 50% and overestimation of 100% is considered as completely unacceptable for this research and the suggested criteria in Table 6-17 were utilised to determine whether IPZA tend to underestimate or overestimate, and whether it was considered acceptable or not.

Table 6-17: Criteria to establish the tendency of IPZA to under- or overestimate

Limit	Criteria (all Qs in m³/s)		
Underestimation unacceptable	$Q_{est} - Q_{obs} < -6$	or	$(Q_{est}/Q_{obs} - 1) < -20\%$
Underestimation acceptable	$-6 \leq Q_{est} - Q_{obs} < -3$	or	$-20\% \leq (Q_{est}/Q_{obs} - 1) < -10\%$
Estimation good	$-3 \leq Q_{est} - Q_{obs} \leq 3$	or	$-10\% \leq (Q_{est}/Q_{obs} - 1) \leq 10\%$
Overestimation acceptable	$3 < Q_{est} - Q_{obs} \leq 6$	or	$10\% < (Q_{est}/Q_{obs} - 1) \leq 20\%$
Overestimation unacceptable	$Q_{est} - Q_{obs} > 6$	or	$(Q_{est}/Q_{obs} - 1) > 20\%$

The criteria were applied to all sites at every AEP and the average outcome was calculated for every site for the AEPs > 50%, as well as separately for the AEPs ≤ 50%.

Since g does not feature in IPZA, the results were plotted in Figure 6-6 and Figure 6-7 against g, to confirm whether the under- or overestimation cannot perchance be related to g.

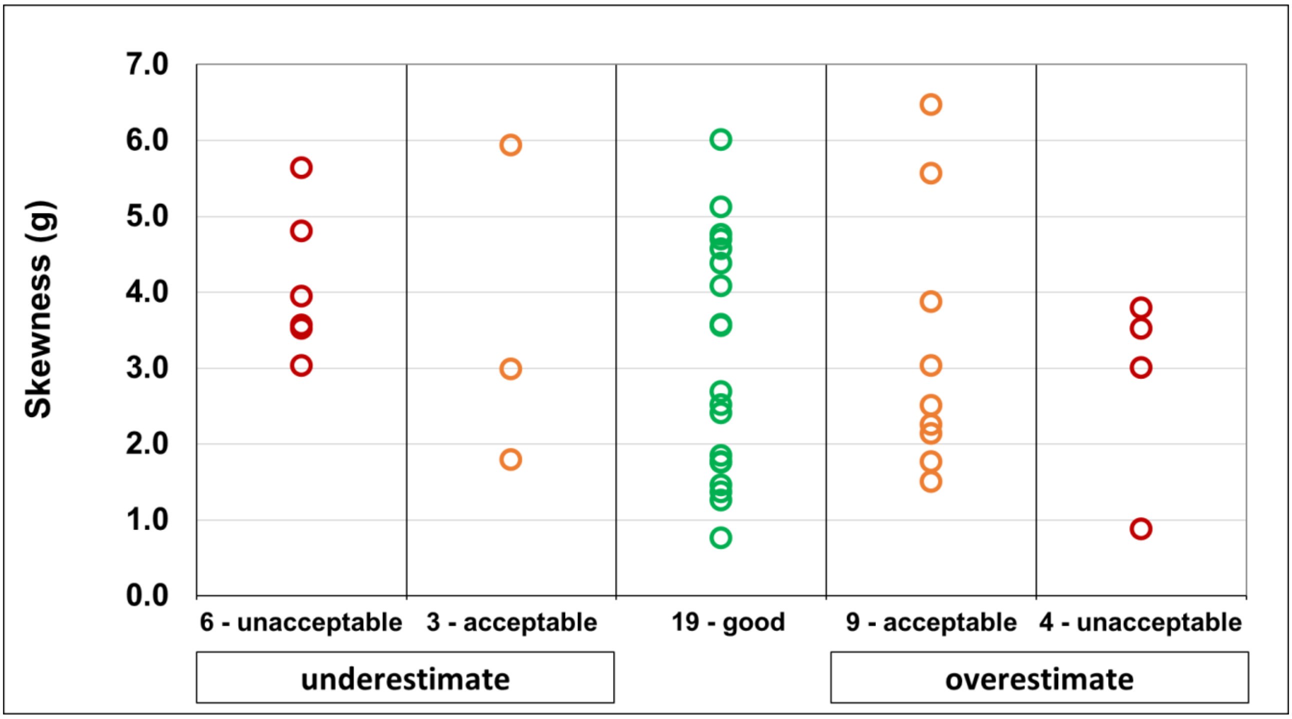


Figure 6-6: Under- or overestimation of Q at each site, by IPZA, for AEP > 50%

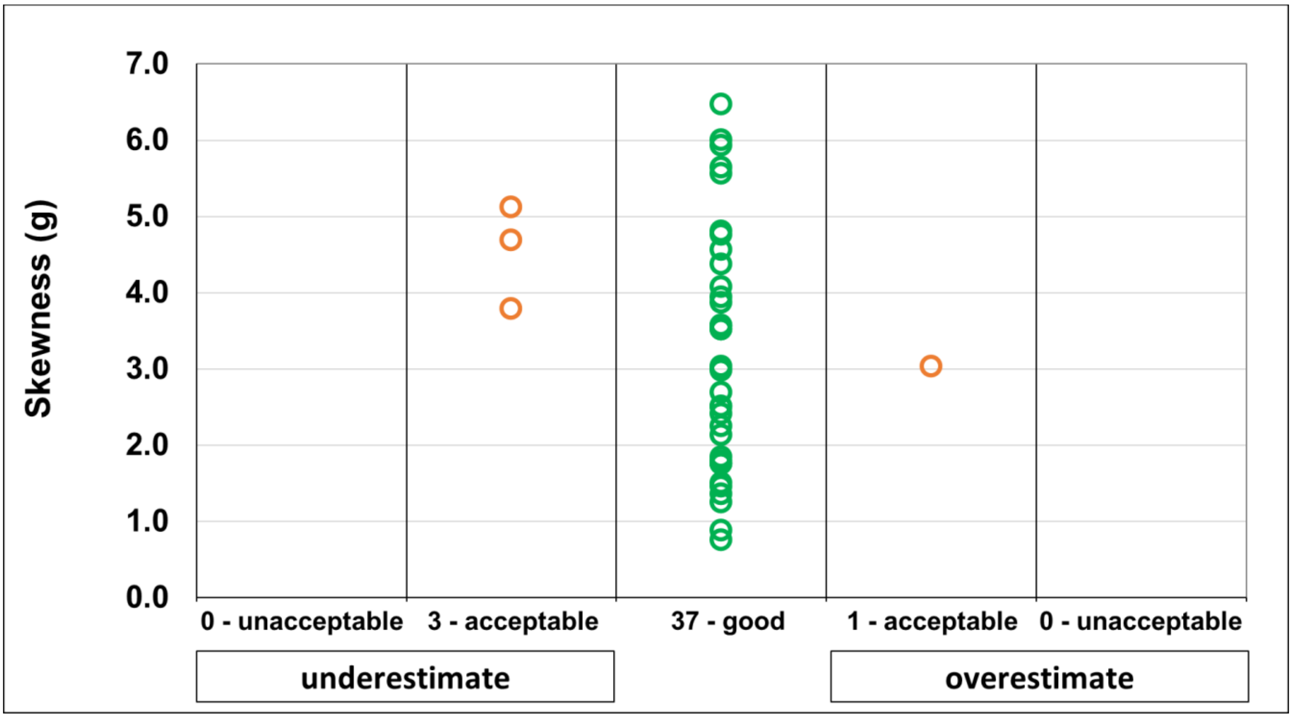


Figure 6-7: Under- or overestimation of Q at each site, by IPZA**,** for AEP ≤ 50**%**

In the design flood range (Figure 6-7) IPZA performs very well. IPZA slightly overestimates at only 1 site and slightly underestimates at 3 sites. For AEP > 50% (see Figure 6-6) the performance is slightly worse, but better than expected – for example, it outperforms the GEV in this range, regarding Q_{est} in relation to Q_{obs} (refer to

Table 6-11 and Appendix 6-H).

It is also clear from the figures above that the degree of under- or overestimation is not related to g.

<u>Performance of IPZA in relation to outliers</u>

The performance of IPZA, in relation to how it is affected by outliers, will be pivotal in whether its performance is an improvement to existing probability distributions, used in South Africa, or not. The sites with the most distinctive outliers were selected from Appendix 6-H, to illustrate the effect of an outlier on the probability distributions.

In Figure 6-8 the selected sites are depicted with the highest flood peak (outlier) excluded in the calculation of moments for the GEV and LP3 probability distributions.

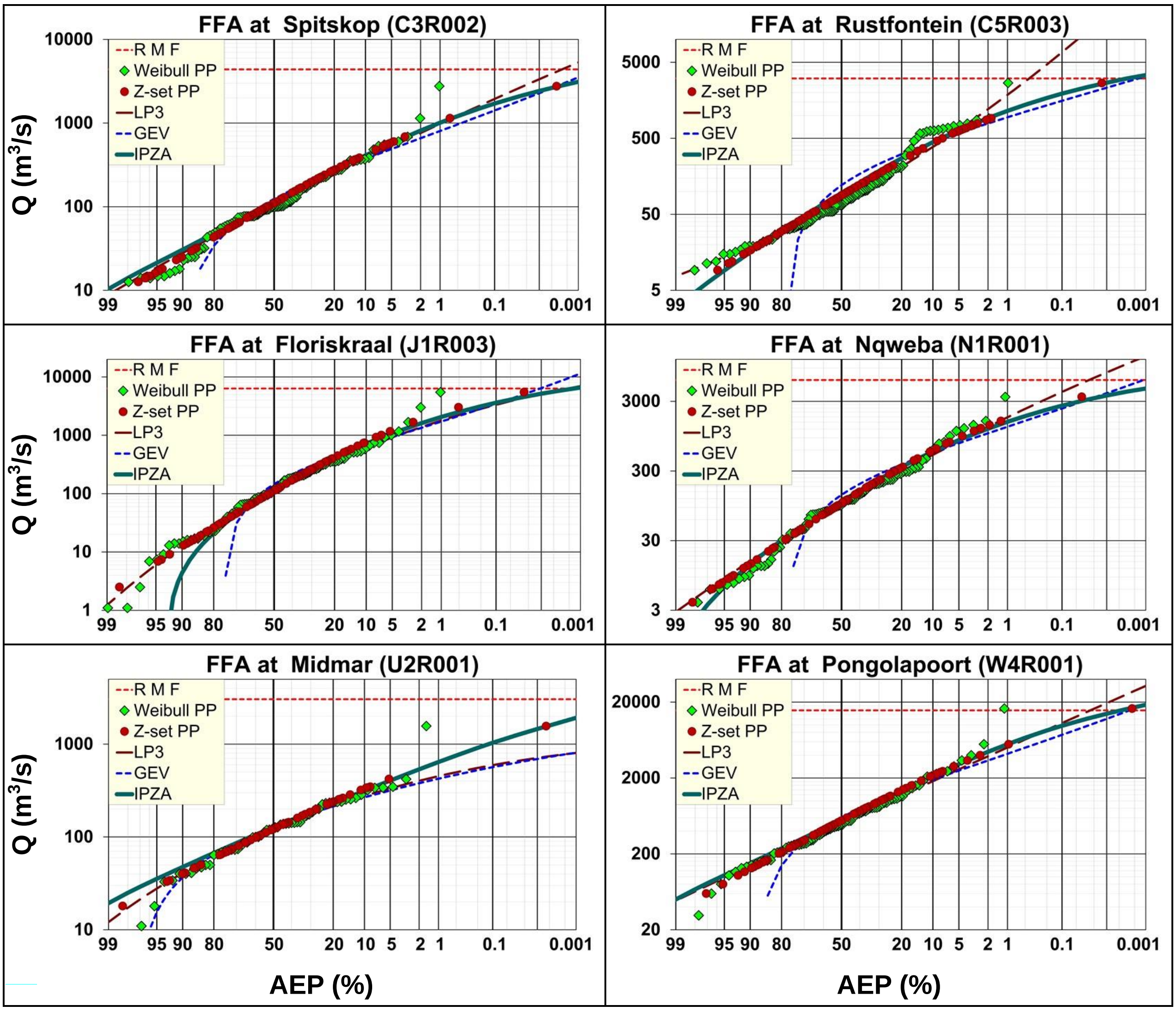


Figure 6-8: FFA results if highest flood peak is excluded in calculation of moments

At one of the selected sites (U2R001) the IPZA results are distinctly higher than that of the GEV and the LP3. At the other 5 sites the FFA results of the IPZA and the GEV are relatively close (and even with the LP3 at 3 sites), with the IPZA results slightly higher than that of the GEV, in the design flood range.

In Figure 6-9 the selected sites are depicted with no flood peaks excluded in the calculation of moments.

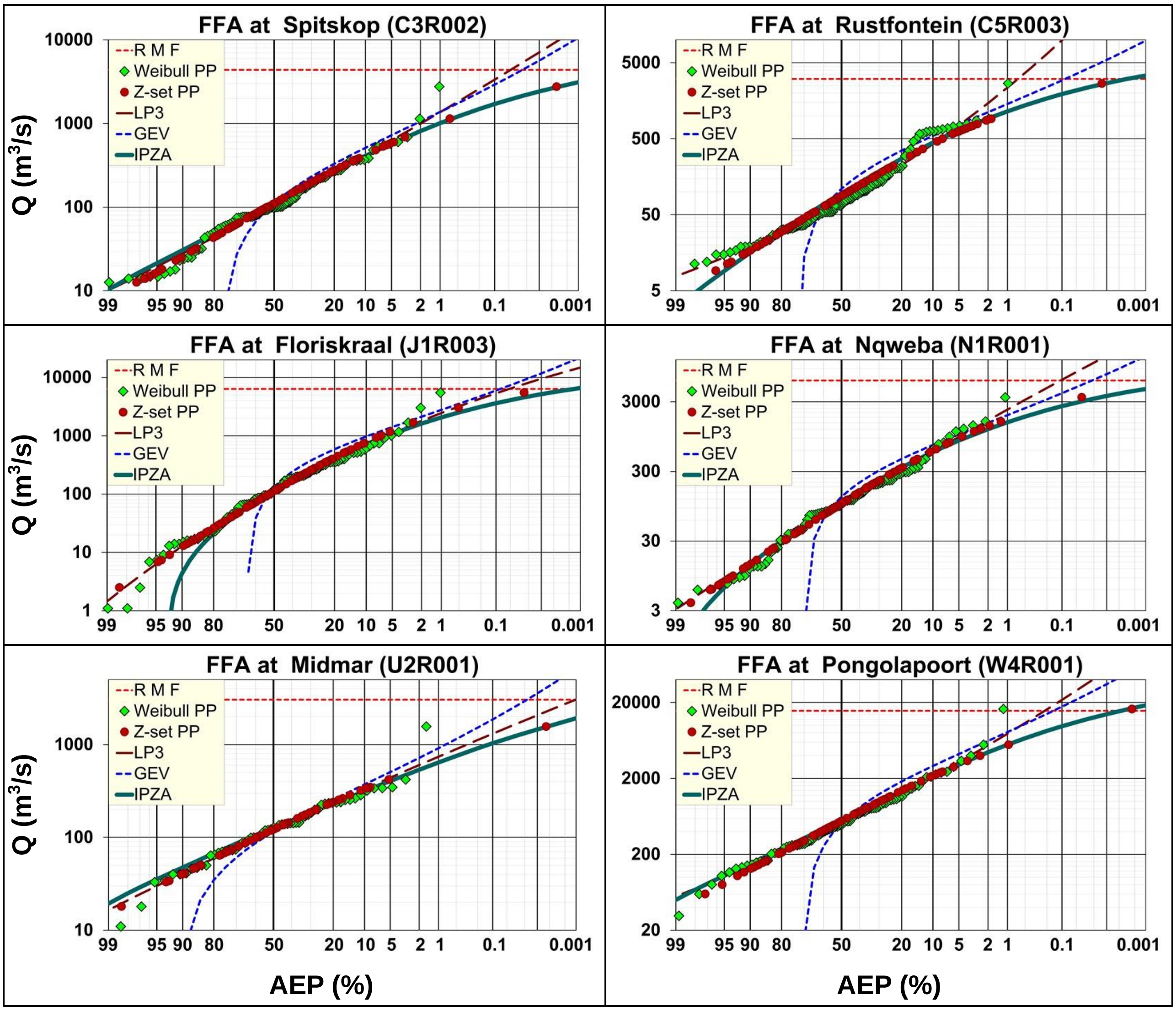


Figure 6-9: FFA results with no flood peaks excluded in calculation of moments

While the exclusion or inclusion of the outlier clearly affects the FFA results in the case of the GEV and LP3, the outlier is sufficiently dealt with in the IPZA probability distribution with the inclusion of SD and SD*.

6.3 Probability Distribution discussion

The primary objective of this research was to propose a more consistent, preferably bounded, probability distribution. The proposed Z-set PP established the foundation for the development of the proposed IPZA probability distribution.

Although the research used only 41 flow sites, due to record length and data accuracy constraints, the general coverage of sites across South Africa and the diversity of meteorological and hydrological conditions was well covered.

The research leading to the culmination of the IPZA probability distribution, has unveiled some noteworthy observations:

- Wheeler (2022a) expressed his concern that many believe the first step in data analysis is to check for normality. With IPZA there is no need to transform data or moments to make the distribution 'more normal', which makes it straightforward to use.
- Wheeler (2022a) also explained that no skewed probability model exists to fit strongly skewed data. The decision to consider the MRA approach to estimate parameters for a prediction equation of the underlying probability distribution might perhaps be considered as unconventional, but the results proved to be exceptionally good.
- The hypothesis that to omit g would either have no effect or improve the results (in Section 6.2.1) was proved correct in Section 6.2.2 and confirmed in Section 6.2.4. This might seem to be a surprising result, but it could have been expected, considering the statement by Wheeler (2011b) that to estimate g with a similar degree of confidence than the location statistics (Q_{ave} and SD), 6 times the record length is required. This might very well be the reason why some distributions often do not perform well, considering that the sample g might not be representative of the skewness of the underlying distribution.
- The hypothesis that the addition of SD* might improve results, was also proved correct in Section 6.2.2. It appears that SD* might compensate for the cases where SD is affected by extremely high flood peaks, resulting in an inaccurate slope for the probability distribution.
- Wide-ranging speculation that the median statistic should be used instead of the average, was challenged by the outcome of this research. The geometric mean appears to be a more viable option than the median.
- The IPZA probability distribution is extremely consistent in the design flood range (AEP ≤ 50%; ARI ≥ 2 years) in estimating flood peak frequencies from AMS. According to the set criteria, the estimated flood peaks are considered as 'good' at 37 of the 41 sites – and 'acceptable' at the remaining sites.
- The IPZA distribution seems to be unaffected by outliers. Consequently, the problem to decide whether to ignore a suspected outlier, or not, is no longer a concern.
- Due to the IPZA not being affected by outliers it generally estimates lower flood peaks at the lower AEPs, than existing probability distributions that do not exclude the outliers when estimating moments (refer to Figure 6-8 and Figure 6-9).
- It should also be noted (Table 6-9 and associated Appendix 6-G) that the RMF, which is indicated on all figures, is approximated by IPZA at an AEP < 0.001% (23 sites), at an AEP ≈ 0.001% (11 sites) and at an AEP ≈ 0.01% (7 sites).
- Overall, considering the CDFs portrayed in Table 6-9 and associated Appendix 6-G, the IPZA estimates flood peak frequencies from AMS data better than the GEV, for AEP > 50% (ARI < 2 years). Inconsistencies observed in this AEP-range may be either due to the level

pool routing technique applied at dams to estimate inflows (which is less accurate for low flows), or it may be attributed to calibration problems of stage-discharge relationships at spillways of dams (erroneous outflows generate erroneous inflow estimates).

Although the latter comment addresses specific concern at dam sites, especially for low flows, it also raises a concern at weir sites, specifically for high flows. Erroneous estimates at either side of the AMS flood peak spectrum may cause a lack of homogeneity of the AMS flood peak data, which can affect the performance of IPZA, or any other probability distribution. Incidentally, some of the observed unsatisfactory fittings, for AEP > 50%, can be related to known low flow calibration problems (e.g., C5R001, C5R002, J3R001, N2R001, T2R001).

7 Recommendations and Conclusions

This chapter presents the combined conclusions and recommendations, pertaining to record length, PP and the development of an improved probability distribution.

7.1 Recommendations

Although the secondary objectives of this research were to find a minimum acceptable record length and to improve the PP, the results encouraged further research. The primary objective, namely the development of an appropriate probability distribution, also exposed possibilities for further research.

Suggested further research regarding record length:

- It is hypothesised that the R-W approach can produce improved results, by excluding extreme outliers, or by using a robust probability distribution like IPZA.
- The GEV, generally, produced more consistent estimates from shorter 7ym R-Ws, than the LP3. This research can be extended to include other probability distributions, such as GPA, GLO and especially IPZA, for example.
- Minimum lengths of less than 50 years were recommended, at 10 of the 15 sites (considering the GEV results in Table 4-7). At the other sites minimum lengths of up to 80 years were recommended. The reason for this, e.g. meteorological systems, presents an opportunity for further research.
- The results of both R-W procedures suggest that the R-W might present a different option for estimating confidence intervals, which should be explored.

Suggested further research regarding PP:

- The Z-set PP and the procedures applied in the record length research may be combined to reinvestigate/improve the Z-set PP.
- Investigate whether the combination of these approaches might not, effectively, shorten the AMS lengths that can be used with acceptable degrees of confidence.
- The preliminary observation, that sites with more than one rainfall-causing system might require longer minimum record lengths than sites in areas with primarily one rainfall-causing system, should be investigated further.

Further research to improve the IPZA distribution include:

- Investigate the less satisfactory estimates for AEPs > 50%.
- The performance of IPZA should be evaluated with shorter AMS record lengths, at sites with good, verified data.
- Although the research sites did cover diverse meteorological and hydrological conditions, it should be tested with datasets from outside South Africa to further evaluate its performance.

7.2 Conclusions

The absence of similar studies relating to the record length research, resulted in the development of a R-W technique that proved to exceed all expectations. On average (50 percentile) record lengths shorter than 50 years (even as low as 28- or 35 years) were found to be adequate in meeting the set criteria – partially supporting the preliminary observations by Hatting *et al*. (2011) and Van der Spuy (2018). It also provided a more holistic picture across SA indicating where a need for longer records exists.

For the purpose of research, the suggested 63 years (75 percentile), as a minimum record length, is considered to be ample in proving the hypothesis in Section 1.3.5.

Finding a more sensible PP technique was considered as being crucial to the outcome of the primary objective, of finding an improved probability distribution. Since existing PPs use a simple ranking technique, it does not differentiate between high and low flood peaks and, therefore, also cannot accommodate outliers sensibly.

A more practical approach, still considering statistical properties, was adopted that led to the Z-set PP technique. In addition to addressing the raised concerns, related to the existing PP techniques, the proposed Z-set PP technique also displayed a similar PP trend, for record lengths longer than 35 to 42 years. It also supports the initial claims by Hatting *et al*. (2011) and Van der Spuy (2018) and compliments the findings on the record lengths research (see Section 7.1).

The Z-set equation is repeated, for ease of reference:

$$Z_{i.Z\text{-}set} = 0.0902\ Z_{i.Weibull} + 0.1564\ Z_{Q_i} + 0.8083\ Z_{logQ_i}$$

Consequently, the Z-set PPs (AEPs) can be determined by subtracting the related CDF of the standard normal distribution value from 1:

$$P_{i.Z\text{-}set} = 1 - \Phi(Z_{i.Z\text{-}set})$$

The Z-set PP proved that, if AMS data are viewed from a more holistic perspective, the challenges related to outliers can be managed sensibly. Thus, the Z-set PP provided the basis for the research to find a corresponding solution for a more consistent probability distribution.

In terms of risk the IPZA returns similar results at lower AEPs (design flood range) than the other distributions considered, provided that high outliers are excluded in the calculation of moments. However, if outliers are not excluded, the IPZA seems to be much more stable and consistent in estimating lower AEPs, compared to the GEV and the LP3.

For ease of reference, the prediction equation for the IPZA probability distribution and the related frequency factors from Table 6-16 are repeated below:

$$Q_{AEP} = K_{P\text{-}Q} \cdot Q_{ave} + K_{P\text{-}SD} \cdot SD + K_{P\text{-}SD^*} \cdot SD^*$$

AEP(%)	50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
$K_{P\text{-}Q}$	1.1035	1.4673	1.5258	1.4791	1.3099	1.1296	0.9249	0.6444	0.4429	0.2641	0.0803	-0.0082
$K_{P\text{-}SD}$	-0.1216	-0.1320	-0.0286	0.1838	0.6317	1.0865	1.6253	2.4345	3.0952	3.7787	4.6980	5.4022
$K_{P\text{-}SD*}$	-0.3379	0.1553	0.7155	1.3020	2.0310	2.5124	2.9205	3.3465	3.5892	3.7695	3.9131	3.9379

Regarding the hypotheses in Section 1.3.5 it can be concluded that:

- A minimum record length was established.
- A new PP was developed that considered outliers more sensibly.
- It was not necessary to normalise the AMS datasets and the MM performed very well, as expected.
- A new probability distribution was developed that proved to be not affected by outliers. However, due to the MRA approach that was followed, it is inconclusive whether the IPZA is (or will be) a bounded distribution, or not. From the figures, presented in this research, it can be stated that the IPZA is probably more likely to have an upper bound than the GEV or LP3.

It was shown in 6.2.4 that, if the outliers are ignored, the LP3 and especially the GEV compare rather well with IPZA in the lower AEP range. However, with the outliers included it was not the case. It might be considered as a limitation that the methodology was developed with minimum record lengths of 60 years and IPZA should be tested with shorter record lengths.

It is concluded that the proposed Z-set PP and the proposed IPZA probability distribution be used as valuable additions to the existing set of decision-making tools for flood hydrologists/engineers performing FFA.

The above conclusion specifically does not exclude the use of other probability distributions. It is sound practice to use more than one probability distribution and to choose the one that fits the observed AMS flood peaks the best, according to the hydrologist's own sound scientific judgement. Therefore, regardless of which approach (or even software) was used to determine a 'best fit', a visual check to verify, or even change, the outcome is strongly recommended.

References

Adamowski, K. 1981. Plotting Formula for Flood Frequency. *Water Resources Association* 17(2):197-202. doi:10.1111/j.1752-1688.1981.tb03922.x

Alexander, W.J.R. 1990. *Flood Hydrology for Southern Africa*. South African National Committee on Large Dams, 1990.

Alexander, W.J.R. 2000. *Flood Risk Reduction Measures*. Department of Civil Engineering, University of Pretoria, South Africa.

Alexander, W.J.R. 2002. The Standard Design Flood: technical paper. *Journal of the South African Institution of Civil Engineering* 44(1):26-30.

Alexander, W.J.R. 2003. *The Standard Design Flood - a new design philosophy*. Department of Civil Engineering, University of Pretoria, South Africa.

Alexander, W.J.R. 2005a. Linkages between solar activity and climatic responses. *Energy & Environment* 16(2):239-253. doi:10.1260/0958305053749462

Alexander, W.J.R. 2005b. Development of a multi-year climate prediction model. *Water SA* 31(2):209-218. doi:10.4314/wsa.v31i2.5204

Asikoglu, O. L. 2018. Parent Flood Frequency Distribution of Turkish Rivers. *Polish Journal of Environmental Studies* 27(2):529-539. doi:10.15244/pjoes/75963

Ball, J., Babister, M., Nathan, R., Weeks, W., Weinmann, E., Retallick, M. and Testoni, I. (ed.). 2019. *Australian Rainfall and Runoff: A Guide to Flood Estimation*. © Commonwealth of Australia (Geoscience Australia).

Bardsley, E. 2016. Note on a modified return period scale for upper-truncated unbounded flood distributions. *Journal of Hydrology* 544:452-455. doi:10.1016/j.jhydrol.2016.11.050

Barnett, V. and Lewis, T. 1984. *Outliers in statistical data* (2nd ed.). John Wiley.

Bauer, S.W. and Midgley, D.C. 1974. *A simple procedure for synthesising Direct Runoff Hydrograph. HRU* Report 1/74, Department of Civil Engineering, University of the Witwatersrand, Johannesburg.

Beard, L.R. 1943. Statistical Analysis in Hydrology. *Transactions of the American Society of Civil Engineers* 108(1):1110-1121. doi:10.1061/TACEAT.0005568

Beirlant, J., Goegebeur, Y., Teugels, J. and Segers, J. 2004. *Statistics of Extremes*. John Wiley and Sons 490 pp. doi:10.1002/0470012382

Benson, M.A. 1962. Plotting Positions and Economics of Engineering Planning. *Journal of the Hydraulics Division* 88(6):57-71. doi:10.1061/JYCEAJ.0000817

Bezak, N., Brilly, M. and Šraj, M. 2014. Comparison between the peaks-over-threshold method and the annual maximum method for flood frequency analysis. *Hydrological Sciences Journal* 59(5):959-977. doi:10.1080/02626667.2013.831174

Bobbitt, Z. 2021. *What is Considered a Good Coefficient of Variation?* Statistics, Simplified, Statology. [Online] Available: https://www.statology.org/what-is-a-good-coefficient-of-variation/#:~:text=In%20most%20fields%2C%20lower%20values,less%20variability%20around%20the%20mean. [2021, October 18]

Boorstin, D.J. s.a.. Brainy Quote. [Online] Available: https://www.brainyquote.com/authors/daniel-j-boorstin-quotes [2022, December 20]

Botero, B.A. and Francés, F. 2010. Estimation of high return period flood quantiles using additional non-systematic information with upper bounded statistical models. *Hydrology and Earth System Sciences* 14:2617–2628. doi:10.5194/hess-14-2617-2010

Breyfogle, F. 2009a. Non-normal data: *To Transform or Not to Transform.* Quality Digest. [Online] Available: https://www.qualitydigest.com/inside/quality-insider-column/individuals-control-chart-and-data-normality.html [2022, December 20]

Breyfogle, F. 2009b. *NOT Transforming the Data Can Be Fatal to Your Analysis.* Quality Digest. [Online] Available: https://www.qualitydigest.com/inside/six-sigma-column/not-transforming-data-can-be-fatal-your-analysis-091609.html [2022, December 20]

Brownlee, J. 2018. *How to Use Statistics to Identify Outliers in Data*. Machine Learning Mastery. [Online] Available: https://machinelearningmastery.com/how-to-use-statistics-to-identify-outliers-in-data/ [2019, May 8]

Calitz, J.P. 2020. Development and Assessment of Regionalised Approaches to Design Flood Estimation in South Africa. Unpublished doctoral dissertation. University of KwaZulu-Natal, College of Agriculture Engineering and Science, Pietermaritzburg, South Africa.

Castellarin, A., Kohnova, S., Gaál, L., Fleig, A., Salinas, J.L., Toumazis, A., Kjeldsen, T.R. and Macdonald, N. 2012. *Review of applied-statistical methods for flood-frequency analysis in Europe.* NERC/Centre for Ecology & Hydrology, 122pp. (ESSEM COST Action ES0901).

Castillo, E. 1988. *Extreme Value Theory in Engineering*. Academic Press 389 pp.

Coles, S. 2001. *An Introduction to Statistical Modeling of Extreme Values*. Springer 228 pp. doi:10.1007/978-1-4471-3675-0

Cook, N. 2011. Comments on "Plotting Positions in Extreme Value Analysis". *Journal of Applied Meteorology and Climatology* 50(1):255-266. doi:10.1175/2010JAMC2316.1

Cook, N.J. 1982. Towards better estimation of extreme winds. *Journal of Wind Engineering and Industrial Aerodynamics* 9(3):295-323. doi:10.1016/0167-6105(82)90021-6

Cook, N.J. 1985. *The Designer's Guide to Wind Loading on Building Structures. Part I: Background, Damage Survey, Wind Data and Structural Classification.* Butterworths. 371 pp.

Cook, N.J. 2012. Rebuttal of "Problems in the extreme value analysis". *Structural Safety* 34(1):418-423. doi:10.1016/j.strusafe.2011.08.002

Cook, N.J., Harris, R.I. and Whiting, R. 2003. Extreme wind speeds in mixed climates revisited. *Journal of Wind Engineering and Industrial Aerodynamics* 91(3):403-422. doi:10.1016/S0167-6105(02)00397-5

Cordery, I. and Pilgrim, D.H. 2000. *The State of the Art of Flood Prediction*. In: Parker DJ (ed.) Floods. Volume II. Routledge, London. 185–197.

Courtney, C. 2018. The Nature of Disaster in China: The 1931 Yangzi River Flood (Studies in Environment and History). Cambridge: Cambridge University Press. doi:10.1017/9781108278362

Cunnane, C. 1978. Unbiased Plotting Positions - a review. *Journal of Hydrology* 37(3-4):205-222. doi:10.1016/0022-1694(78)90017-3

Cunnane, C. 1985. Factors affecting choice of distribution for flood series. *Hydrological Sciences Journal* 30(1):25-36. doi:10.1080/02626668509490969

De Haan, L. 2007. Comments on "Plotting Positions in Extreme Value Analysis". *Journal of Applied Meteorology and Climatology* 46(3):396. doi:10.1175/JAM2471.1

Du Plessis, J.A. 2014. Unpublished results of a frequency analysis survey. Short Course on Flood Hydrology, August 2014, University of Stellenbosch.

Dyson, L.L. and Van Heerden, J. 2001. The heavy rainfall and floods over the northeastern interior of South Africa. *South African Journal of Science* 97(3):80-86.

DWS. 1993–2021. Flood frequency analyses intended for dam safety evaluation (numerous reports). Department of Water and Sanitation, Pretoria.

DWS. 2006–2021. Extension of rating curves at flow gauging weirs to accommodate high flows (numerous reports). Department of Water and Sanitation, Pretoria.

Embrechts, P., Klüppelberg, C. and Mikosch, T. 1997. *Modelling Extremal Events for Insurance and Finance*. Springer. 648 pp. doi:10.1007/978-3-642-33483-2

England, J.F. (Jr.), Cohn, T.A., Faber, B.A., Stedinger, J.R., Thomas, W.O. (Jr.), Veilleux, A.G., Kiang, J.E. and Mason, R.R. (Jr.) 2019. *Guidelines for determining flood flow frequency - Bulletin 17C (ver. 1.1, May 2019)*. U.S. Geological Survey Techniques and Methods, book 4, chap. B5 148 pp. doi:10.3133/tm4B5

Faber, B. 2010. Current Methods for Flood Frequency Analysis. In: *Workshop on Non-Stationarity, Hydrologic Frequency Analysis and Water Management*, January 13-15, 2010, Boulder, Colorado. 268-273.

Farooq, M., Shafique, M. and Khattak, M.S. 2018. Flood frequency analysis of river Swat using Log Pearson type 3, Generalized Extreme Value, Normal, and Gumbel Max distribution methods. *Arabian Journal of Geosciences* 11(9):216. doi:10.1007/s12517-018-3553-z

Folland, C. and Anderson, C. 2002. Estimating Changing Extremes Using Empirical Ranking Methods. *Journal of Climate* 15:2954–2960. doi:10.1175/1520-0442(2002)015<2954:ECEUER>2.0.CO;2

FORMPLUS BLOG. s.a.. *Coefficient of Variation: Definition, Formula, Interpretation, Examples & FAQs*. [Online] Available: https://www.formpl.us/blog/coefficient-variation [2022, July 18]

Frost, J. s.a. *5 Ways to Find Outliers in Your Data*. Statistics by Jim. [Online] Available: https://statisticsbyjim.com/basics/outliers/ [2019, May 8]

Fuglem, M., Parr, G. and Jordaan, I.J. 2013. Plotting positions for fitting distributions and extreme value analysis. Canadian *Journal of Civil Engineering* 40(2):130-139. doi:10.1139/cjce-2012-0427

Galloway, G.E. 2010. If Stationarity is Dead, What Do We Do Now? In: *Workshop on Non-Stationarity, Hydrologic Frequency Analysis and Water Management*, January 13-15, 2010, Boulder, Colorado. 274-280.

Ganesh, S. and Cave, V. 2018. P-values, P-values everywhere! *New Zealand Veterinary Journal* 66(2):55-56. doi:10.1080/00480169.2018.1415604

Gericke, O.J. and Du Plessis, J.A. 2012. Evaluation of the standard design flood method in selected basins in South Africa. *Journal of the South African Institution of Civil Engineering* 54(2):2-14.

Gumbel, E.J. 1958. *Statistics of Extremes*. Columbia University Press. 375 pp.

Gunasekara, T.A.G. and Cunnane, C. 1992. Split sampling technique for selecting a flood frequency analysis procedure. *Journal of Hydrology* 130(1):189-200. doi:10.1016/0022-1694(92)90110-H

Guru, N. 2016. Flood Frequency Analysis of Partial Duration Series using Soft Computing Techniques for Mahanadi River Basin in India. Unpublished doctoral dissertation. National Institute of Technology, Rourkela-769008, India, January-2016.

Haddad, K. and Rahman, A. 2012. Regional flood frequency analysis in eastern Australia: Bayesian GLS regression-based methods within fixed region and ROI framework – Quantile Regression vs Parameter Regression Technique. *Journal of Hydrology* 430-431:142-161. doi:10.1016/j.jhydrol.2012.02.012

Haile, A.T. 2011. Regional Flood Frequency Analysis in Southern Africa. Unpublished master's Thesis. University of Oslo, Norway.

Harris, R.I. 2001. The accuracy of design values predicted from extreme value analysis. *Journal of Wind Engineering and Industrial Aerodynamics* 89(2):153-164. doi:10.1016/S0167-6105(00)00060-X

Hattingh, L., Mostert, A., Muir, C. and Basson, T. 2011. The impact of hydrology on the adequacy of existing dams to safety standards – the Namibian experience. In: *Proceedings of the SANCOLD conference: Management and Design of Dams in Africa*, 8-10 November 2011, Johannesburg, South Africa.

Hazen, A. 1913. Storage to be Provided in Impounding Reservoirs for Municipal Water Supply. In: *Proceedings of the American Society of Civil Engineers* 39(9):1943-2044.

Hirsch, R.M. and Stedinger, J.R. 1987. Plotting positions for historical floods and their precision. *Water Resources Research* 23(4):715-727. doi:10.1029/WR023i004p00715

Horton, E.B., Folland, C.K. and Parker, D.E. 2001. The changing incidence of extremes in worldwide and central England temperatures to the end of the twentieth century. *Climatic Change* 50:267-295. doi:10.1023/A:1010603629772

Hosking, J.R.M. and Wallis, J.R. 1997. *Regional Frequency Analysis: An Approach Based on L-Moments*. Cambridge University Press, Cambridge. 224 pp.

Hosking, J.R.M., Wallis, J.R. and Wood, E.F. 1985. Estimation of the Generalized Extreme-Value Distribution by the Method of Probability-Weighted Moments. *Technometrics* 27(3):251-261. doi:10.2307/1269706

IACWD (Interagency Advisory Committee on Water Data). 1982. *Guidelines for determining flood flow frequency, Bulletin 17B*. Technical report, Interagency Advisory Committee on Water Data, Hydrology Sub-committee.

Institute of Hydrology. 1999. *Flood Estimation Handbook (five volumes)*. Centre for Ecology & Hydrology, Wallingford, Oxfordshire, UK.

Jordaan, I. 2005. *Decisions under Uncertainty: Probabilistic Analysis for Engineering Decisions*. Cambridge University Press 688 pp. doi:10.1017/CBO9780511804861

Karim, F., Hasan, M. and Marvanek, S. 2017. Evaluating Annual Maximum and Partial Duration Series for Estimating Frequency of Small Magnitude Floods. *Water* 2017 9(7):article 481. doi:10.3390/w9070481

Kim, S., Shin, H., Joo, K. and Heo, J-H. 2012. Development of plotting position for the general extreme value distribution. *Journal of Hydrology* 475:259–269. doi:10.1016/j.jhydrol.2012.09.055

Klemeš, V. 1987. *Hydrological and Engineering Relevance of Flood Frequency Analysis*. In: Singh V.P. (eds) Hydrologic Frequency Modeling. Springer, Dordrecht. doi:10.1007/978-94-009-3953-0_1

Kline, R.B. 2016. *Principles and practice of structural equation modeling* (4th ed.). Guilford Press. 534 pp.

Kovács, Z. 1988. *Regional Maximum Flood Peaks in Southern Africa*. Technical Report 137. Department of Water Affairs, South Africa.

Kovács, Z.P., Du Plessis, D.B., Bracher, P.R., Dunn, P., and Mallory, G.C.L. 1985. *Documentation of the 1984 Domoina Floods*. Technical Report 122. Department of Water Affairs, South Africa.

Landwehr, J.M., Matalas, N.C. and Wallis, J.R. 1978. Some Comparisons of Flood Statistics in Real and Log Space. *Water Resources Research* 14(5):902-920. doi:10.1029/WR014i005p00902

Langat, P.K., Kumar, L. and Koech, R. 2019. Identification of the Most Suitable Probability Distribution Models for Maximum, Minimum, and Mean Streamflow. *Water* 11(4):734. doi:10.3390/w11040734

Langbein, W.B. 1960. Plotting positions in frequency analysis. In: *Flood-frequency analyses, Manual of Hydrology: Part 3. Flood Flow Techniques*, Geological Survey Water-Supply, Paper 1543-A 48-51. doi:10.3133/wsp1543A

Lettenmaier, D. 2010. Workshop Summary I. In: Wor*kshop on Non-Stationarity, Hydrologic Frequency Analysis and Water Management*, January 13-15, 2010, Boulder, Colorado. 268-273.

Makkonen, L. 2006. Plotting Positions in Extreme Value Analysis. *Journal of Applied Meteorology and Climatology* 45:334-340. doi:10.1175/JAM2349.1

Makkonen, L. 2008. Bringing Closure to the Plotting Position Controversy. *Communications in Statistics - Theory and Methods* 37(3):460-467. doi:10.1080/03610920701653094

Makkonen, L. and Pajari, M. 2014. Defining Sample Quantiles by the True Rank Probability. *Journal of Probability and Statistics* 2014(4):6 pp. doi:10.1155/2014/326579

Makkonen, L., Pajari, M. and Tikanmäki, M. 2013. Discussion on "Plotting positions for fitting distributions and extreme value analysis". *Canadian Journal of Civil Engineering* 40(9):927-929. doi:10.1139/cjce-2013-0227

McCuen, R.H. and Galloway, K.E. 2010. Record Length Requirements for Annual Maximum Flood Series. *Journal of Hydraulic Engineering* 15(9):704-707. doi:10.1061/(ASCE)HE.1943-5584.0000223

Mehdi, F. and Mehdi, J. 2011. Determination of Plotting Position Formula for the Normal, Log-Normal, Pearson(III), Log-Pearson(III) and Gumble Distributional Hypotheses using the Probability Plot Correlation Coefficient Test. *World Applied Sciences Journal* 15(8):1181-1185.

Midgley, D.C. 1972. *Design Flood Determination in South Africa*. HRU report no 1/72, Department of Civil Engineering, University of the Witwatersrand, Johannesburg.

Midgley, D.C., Pullen, R.A. and Pitman, W.V. 1969. *Design Flood Determination in South Africa*. HRU report no 4/69, Department of Civil Engineering, University of the Witwatersrand, Johannesburg.

Mimikou, M., Niadas, P., Hadjissava, P. and Kouvopoulos, Y. 1993. Regional prediction of extreme design storm and flood characteristics. *Water Power & Dam Construction*, December 1993.

Mkhandi, S., Opere, A.O. and Willems, P. 2005. *Comparison between annual maximum and peaks over threshold models for flood frequency prediction*. Department of Meteorology, University of Nairobi, Kenya.

Mohssen, M. 2009. Partial duration series in the annual domain. In: *Proceedings of the 18th World IMACS and MODSIM International Congress*, 13-17 July 2009 (pp. 2694-2700), Cairns, Australia.

Mutua, F.M. 1994. The use of the Akaike Information Criterion in the identification of an optimum flood frequency model. *Hydrological Sciences Journal* 39(3):235-244.

Naidoo, J. 2020. An Assessment of the Performance of Deterministic and Empirical Design Flood Estimation Methods in South Africa. Unpublished master's Thesis. University of KwaZulu-Natal, South Africa.

NERC (Natural Environment Research Council). 1975a. *Hydrological Studies*, Vol. I, Flood Studies Report. Natural Environment Research Council, London, UK.

NERC (Natural Environment Research Council). 1975b. *Meteorological Studies*, Vol. II, Flood Studies Report. Natural Environment Research Council, London, UK.

Pegram, G. and Parak, M. 2004. A review of the regional maximum flood and rational formula using geomorphological information and observed floods. *Water SA* 30(3):377-392.

Pitman, W.V. and Midgley, D.C. 1967. Flood Studies in South Africa: Frequency Analysis of Peak Discharges. *Die Siviele Ingenieur (Civil Engineering)* 9(8):193-200.

Pitman, W.V. and Midgley, D.C. 1971. *Amendments to Design Flood Manual HRU 4/69*. HRU Report 1/71, Department of Civil Engineering, University of the Witwatersrand, Johannesburg.

Posner, M.A. s.a. Michael A. Posner's Page of Statistics Quotes. [Online] Available: http://www.posnersmith.net/quotes_s.html [2015, October 12]

Potter, K.W. 1987. Research on flood frequency analysis: 1983–1986. *Reviews of Geophysics* 25(2):113-118. doi:10.1029/RG025i002p00113

Rahman, A.S., Rahman, A., Zaman, M.A., 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(3):1803-1813. doi:10.1007/s11069-013-0775-y

Reed, D.W. 1999. *Overview*. Volume 1 of the Flood Estimation Handbook. Centre for Ecology & Hydrology, Wallingford, Oxfordshire, UK.

Roberts, C.P.R. and Alexander, W.J.R. 1982. Lessons learnt from the 1981 Laingsburg Flood. *Die Siviele Ingenieur (Civil Engineering)* 24(1):17-27.

Robson, A.J. and Reed, D.W. 1999. *Statistical procedures for flood frequency estimation*. Volume 3 of the Flood Estimation Handbook. Centre for Ecology & Hydrology, Wallingford, Oxfordshire, UK.

Salinas, J.L., Castellarin, A., Viglione, A., Kohnová, S. and Kjeldsen, T.R. 2014. Regional parent flood frequency distributions in Europe – Part 1: Is the GEV model suitable as a pan-European parent? *Hydrology and Earth System Sciences* 18(11):4381-4389. doi:10.5194/hess-18-4381-2014

SANRAL (South African National Road Agency). 2013. *Drainage Manual*, 6th edition. South African National Roads Agency Ltd, Pretoria.

Schlögl, M. and Laaha, G. 2017. Extreme weather exposure identification for road networks - a comparative assessment of statistical methods. *Natural Hazards and Earth Systems Sciences* 17(4):515-531. doi:10.5194/nhess-17-515-2017

Schmidt, E.J. and Schulze, R.E. 1987. *Flood Volume and Peak Discharge from Small Catchments in Southern Africa based on the SCS Technique*. ACRU Report No. 24 (WRC-TT 31/8), Department of Agricultural Engineering, University of Natal, Pietermaritzburg, South Africa. 164 pp.

Smithers, J.C. 2012. Methods for design flood estimation in South Africa. *Water SA* 38(4):633-646. doi:10.4314/wsa.v38i4.19

Smithers, J.C., Görgens, A., Gericke, J., Jonker, V. and Roberts, C.P.R. 2014. The Initiation of a National Flood Studies Programme for South Africa. In: *Proceedings of the SANCOLD National Conference*, 5–7 November 2014, Johannesburg.

Srikanthan, S. 2014. A Comparison of Annual Maximum and Partial Duration Series in Frequency Analysis. In: *Proceedings of the Hydrology and Water Resources Symposium*, HWRS 2014 374-381.

Stedinger, J.R., Vogel, R.M. and Foufoula-Georgiou, E. 1993. Frequency Analysis of Extreme Events. In: *Handbook of Hydrology*. 18. McGraw-Hill, New York, USA.

Tasker, G.D. 1983. Effective record length for the T-year event. *Journal of Hydrology* 64(1-4):39-47. doi:10.1016/0022-1694(83)90059-8

Ul Hassen, M., Hayat, O. and Noreen, Z. 2019. Selecting the best probability distribution for at-site flood frequency analysis; a study of Torne River. *SN Applied Sciences* 1:1629. doi:10.1007/s42452-019-1584-z

USACE (U.S. Army Corps of Engineers). 1993. *Engineering and Design: Hydrologic Frequency Analysis*. Department of the Army, USACE EM1110-2-1415.

USACE (U.S. Army Corps of Engineers). 1994. *Engineering and Design: Flood-Runoff Analysis*. Department of the Army, USACE EM1110-2-1417.

Van Bladeren, D. and Burger, C.E. 1989. *Documentation of the September 1987 Natal Floods*. Technical Report 139. Department of Water Affairs, South Africa.

Van Bladeren, D., Zawada, P.K. and Mahlangu, D. 2007. Statistical Based Regional Flood *Frequency Estimation Study for South Africa Using Systematic, Historical and Palaeoflood Data, Pilot Study – Catchment Management Area 15*. WRC Report No 1260/1/07, Water Research Commission, Pretoria. ISBN 078-1-77005-537-7

Van der Spuy, D. 2018. *A practical approach to Probability Analyses* [PowerPoint presentation]. Flood Hydrology Course, August 2018, University of Stellenbosch, South Africa.

Van der Spuy, D. and Rademeyer, P.F. 2018. *Flood Frequency Estimation Methods applied in the Flood Studies Component, Department of Water and Sanitation*. In: Proceedings of Flood Hydrology Course, August 2018, University of Stellenbosch, South Africa.

Van der Spuy, D., Rademeyer, P.F. and Linström, C.R. 2004. *Flood Frequency Estimation Methods as applied in the Department of Water Affairs and Forestry*. In: Proceedings of Flood Hydrology Course, April 2004, University of Stellenbosch, South Africa.

Van Vuuren, S.J., Van Dijk, M. and Coetzee, G.L. 2012. *Status review and requirements of overhauling Flood Determination Methods in South Africa*. WRC Report No. TT563/13, Water Research Commission, Pretoria. 91 pp.

Vogel, R.M., McMahon, T.A. and Chiew, F.H.S. 1993. Floodflow frequency model selection in Australia. *Journal of Hydrology* 146:421-449. doi:10.1016/0022-1694(93)90288-K

Wallis, J.R. and Wood, E.F. 1985. Relative Accuracy of Log Pearson III Procedures. *Journal of Hydraulic Engineering* 111(7):1043-1056. doi:10.1061/(ASCE)0733-9429(1985)111:7(1043)

Wheeler, D.J. 2009a. *Do You Have Leptokurtophobia?* Quality Digest. [Online] Available: https://www.qualitydigest.com/inside/quality-insider-column/do-you-have-leptokurtophobia.html [2022, December 20]

Wheeler, D.J. 2009b. *Transforming the Data can be Fatal to your Analysis.* Quality Digest. [Online] Available: https://www.qualitydigest.com/inside/quality-insider-column/transforming-data-can-be-fatal-your-analysis.html [2022, December 20]

Wheeler, D.J. 2009c. *Avoiding Statistical Jabberwocky.* Quality Digest. [Online] Available: https://www.qualitydigest.com/inside/six-sigma-column/avoiding-statistical-jabberwocky-100709.html [2022, December 20]

Wheeler, D.J. 2010. *Are You Sure We Don't Need Normally Distributed Data?* More about the misuses of probability theory. Quality Digest, Manuscript No. 220. [Online] Available: https://www.qualitydigest.com/inside/quality-insider-column/are-you-sure-we-don-t-need-normally-distributed-data.html [2022, July 14]

Wheeler, D.J. 2011a. *Problems with Skewness and Kurtosis, Part One: What do the shape parameters do?* Quality Digest. [Online] Available: https://www.qualitydigest.com/inside/quality-insider-article/problems-skewness-and-kurtosis-part-one.html [2018, August 4]

Wheeler, D.J. 2011b. *Problems with Skewness and Kurtosis, Part Two: What do the shape parameters do?* Quality Digest. [Online] Available: https://www.qualitydigest.com/inside/quality-insider-article/problems-skewness-and-kurtosis-part-two.html [2018, August 4]

Wheeler, D.J. 2022a. *How Can a Control Chart Work Without a Distribution?* Quality Digest. [Online] Available: https://www.qualitydigest.com/inside/six-sigma-column/how-can-control-chart-work-without-distribution-042522.html? [2022, May 3]

Wheeler, D.J. 2022b*. The Secret of Data Analysis.* Quality Digest. [Online] Available: https://www.qualitydigest.com/inside/statistics-column/secret-data-analysis-120522.html [2022, December 7]

Wheeler, D.J. 2023. Probability weighted moments and 'Normality' of data, e-mail to D van der Spuy, 4 January. Available e-mail: djwheeler@spcpress.com.

WMO (World Meteorological Organization). 2009. *Guide to Hydrological Practices, Volume II: Management of Water Resources and Applications of Hydrological Practices*, Sixth Edition. WMO-No. 168, 2009.

Xiong, F., Guo, S., Chen, L., Yin, J. and Liu, P. 2018. Flood Frequency Analysis Using Halphen Distribution and Maximum Entropy. *Journal of Hydrologic Engineering* 23(5):04018012. doi:10.1061/(ASCE)HE.1943-5584.0001637

Yahaya, A.S., Nor, N.M., Jali, N.R.M., Ramli, N.A., Ahmad, F. and Ul-Saufie, A.Z. 2012a. Determination of the Probability Plotting Position for Type I Extreme Value Distribution. *Journal of Applied Sciences* 12(14):1501-1506. doi:10.3923/jas.2012.1501.1506

Yahaya, A.S., Yee, C.S., Ramli, N.A. and Ahmad, F. 2012b. Determination of the Best Probability Plotting Position for Predicting Parameters of the Weibull Distribution. *International Journal of Applied Science and Technology* 2(3):106-111.

Yu, G.H. and Huang, C.C. 2001. A Distribution Free Plotting Position. Stochastic Environmental *Research and Risk Assessment* 15:462-476. doi:10.1007/s004770100083

Zawada, P.K. 1994. Palaeoflood hydrology of the Buffels River, Laingsburg, South Africa: was the 1981 flood the largest? *South African Journal of Geology* 97(1):21-32. doi:10.10520/AJA10120750_807

Zhang, Z., Stadnyk, T.A. and Burn, D.H. 2019. Identification of a preferred statistical distribution for at-site flood frequency analysis in Canada. Canadian Water Resources *Journal / Revue canadienne des ressources hydriques* 45(1):43-58. doi:10.1080/07011784.2019.1691942

APPENDICES

Please note that, for ease of reference, Appendices are numbered to concur with Chapter numbers. Tables and Figures will thus be identified by Appendix numbers and not by captions.

For example: Appendix 3-A will be the first Appendix of Chapter 3.

APPENDIX 2: Literature study

APPENDIX 2-A: Similarity between PP techniques

Compare AEPs for various PP techniques to illustrate the similarity between them. The AMS record of 74 years, at Woodstock dam (V1R003), is used – Note: the ranking, in relation to record length, determine the probability; not the AMS values.

Q (m^3/s)	Rank	AEPs (with ARI in brackets), according to:							
		Weibull	Adamowski	Beard	Tukey	Blom	Cunane	Gringorten	Hazen
2 915	1	0.0133 (75)	0.0101 (99)	0.0093 (108)	0.009 (111)	0.0084 (119)	0.0081 (124)	0.0076 (132)	0.0068 (148)
1 400	2	0.027 (38)	0.023 (43)	0.023 (44)	0.022 (45)	0.022 (46)	0.022 (46)	0.021 (48)	0.020 (49)
1 380	3	0.040 (25)	0.037 (27)	0.036 (28)	0.036 (28)	0.035 (28)	0.035 (29)	0.035 (29)	0.034 (30)
1 275	4	0.053 (19)	0.050 (20)	0.050 (20)	0.049 (20)	0.049 (20)	0.049 (21)	0.048 (21)	0.047 (21)
1 020	5	0.067 (15)	0.064 (16)	0.063 (16)	0.063 (16)	0.062 (16)	0.062 (16)	0.062 (16)	0.061 (16)
967	6	0.080 (13)	0.077 (13)	0.076 (13)	0.076 (13)	0.076 (13)	0.075 (13)	0.075 (13)	0.074 (13)
935	7	0.093 (11)	0.091 (11)	0.090 (11)	0.090 (11)	0.089 (11)	0.089 (11)	0.089 (11)	0.088 (11)
920	8	0.107 (9.4)	0.104 (9.6)	0.103 (9.7)	0.103 (9.7)	0.103 (9.7)	0.102 (9.8)	0.102 (9.8)	0.101 (9.9)
915	9	0.120 (8.3)	0.117 (8.5)	0.117 (8.6)	0.117 (8.6)	0.116 (8.6)	0.116 (8.6)	0.115 (8.7)	0.115 (8.7)
910	10	0.133 (7.5)	0.131 (7.6)	0.130 (7.7)	0.130 (7.7)	0.130 (7.7)	0.129 (7.7)	0.129 (7.8)	0.128 (7.8)
910	11	0.147 (6.8)	0.144 (6.9)	0.144 (7.0)	0.143 (7.0)	0.143 (7.0)	0.143 (7.0)	0.142 (7.0)	0.142 (7.0)
895	12	0.160 (6.3)	0.158 (6.3)	0.157 (6.4)	0.157 (6.4)	0.157 (6.4)	0.156 (6.4)	0.156 (6.4)	0.155 (6.4)
860	13	0.173 (5.8)	0.171 (5.8)	0.171 (5.9)	0.170 (5.9)	0.170 (5.9)	0.170 (5.9)	0.169 (5.9)	0.169 (5.9)
853	14	0.187 (5.4)	0.185 (5.4)	0.184 (5.4)	0.184 (5.4)	0.184 (5.4)	0.183 (5.5)	0.183 (5.5)	0.182 (5.5)
810	15	0.200 (5.0)	0.198 (5.1)	0.197 (5.1)	0.197 (5.1)	0.197 (5.1)	0.197 (5.1)	0.196 (5.1)	0.196 (5.1)
801	16	0.213 (4.7)	0.211 (4.7)	0.211 (4.7)	0.211 (4.7)	0.210 (4.8)	0.210 (4.8)	0.210 (4.8)	0.209 (4.8)
799	17	0.227 (4.4)	0.225 (4.4)	0.224 (4.5)	0.224 (4.5)	0.224 (4.5)	0.224 (4.5)	0.223 (4.5)	0.223 (4.5)
784	18	0.240 (4.2)	0.238 (4.2)	0.238 (4.2)	0.238 (4.2)	0.237 (4.2)	0.237 (4.2)	0.237 (4.2)	0.236 (4.2)
727	19	0.253 (3.9)	0.252 (4.0)	0.251 (4.0)	0.251 (4.0)	0.251 (4.0)	0.251 (4.0)	0.250 (4.0)	0.250 (4.0)
713	20	0.267 (3.8)	0.265 (3.8)	0.265 (3.8)	0.265 (3.8)	0.264 (3.8)	0.264 (3.8)	0.264 (3.8)	0.264 (3.8)
700	21	0.28 (3.6)	0.279 (3.6)	0.278 (3.6)	0.278 (3.6)	0.278 (3.6)	0.278 (3.6)	0.277 (3.6)	0.277 (3.6)
664	22	0.293 (3.4)	0.292 (3.4)	0.292 (3.4)	0.291 (3.4)	0.291 (3.4)	0.291 (3.4)	0.291 (3.4)	0.291 (3.4)
658	23	0.307 (3.3)	0.305 (3.3)	0.305 (3.3)	0.305 (3.3)	0.305 (3.3)	0.305 (3.3)	0.304 (3.3)	0.304 (3.3)
625	24	0.320 (3.1)	0.319 (3.1)	0.318 (3.1)	0.318 (3.1)	0.318 (3.1)	0.318 (3.1)	0.318 (3.1)	0.318 (3.1)
538	25	0.333 (3.0)	0.332 (3.0)	0.332 (3.0)	0.332 (3.0)	0.332 (3.0)	0.332 (3.0)	0.331 (3.0)	0.331 (3.0)
533	26	0.347 (2.9)	0.346 (2.9)	0.345 (2.9)	0.345 (2.9)	0.345 (2.9)	0.345 (2.9)	0.345 (2.9)	0.345 (2.9)
495	27	0.360 (2.8)	0.359 (2.8)	0.359 (2.8)	0.359 (2.8)	0.359 (2.8)	0.358 (2.8)	0.358 (2.8)	0.358 (2.8)
483	28	0.373 (2.7)	0.372 (2.7)	0.372 (2.7)	0.372 (2.7)	0.372 (2.7)	0.372 (2.7)	0.372 (2.7)	0.372 (2.7)
475	29	0.387 (2.6)	0.386 (2.6)	0.386 (2.6)	0.386 (2.6)	0.386 (2.6)	0.385 (2.6)	0.385 (2.6)	0.385 (2.6)
453	30	0.400 (2.5)	0.399 (2.5)	0.399 (2.5)	0.399 (2.5)	0.399 (2.5)	0.399 (2.5)	0.399 (2.5)	0.399 (2.5)
424	31	0.413 (2.4)	0.413 (2.4)	0.413 (2.4)	0.413 (2.4)	0.412 (2.4)	0.412 (2.4)	0.412 (2.4)	0.412 (2.4)
411	32	0.427 (2.3)	0.426 (2.3)	0.426 (2.3)	0.426 (2.3)	0.426 (2.3)	0.426 (2.3)	0.426 (2.3)	0.426 (2.3)
402	33	0.440 (2.3)	0.440 (2.3)	0.439 (2.3)	0.439 (2.3)	0.439 (2.3)	0.439 (2.3)	0.439 (2.3)	0.439 (2.3)
385	34	0.453 (2.2)	0.453 (2.2)	0.453 (2.2)	0.453 (2.2)	0.453 (2.2)	0.453 (2.2)	0.453 (2.2)	0.453 (2.2)
385	35	0.467 (2.1)	0.466 (2.1)	0.466 (2.1)	0.466 (2.1)	0.466 (2.1)	0.466 (2.1)	0.466 (2.1)	0.466 (2.1)
385	36	0.480 (2.1)	0.480 (2.1)	0.480 (2.1)	0.480 (2.1)	0.480 (2.1)	0.480 (2.1)	0.480 (2.1)	0.480 (2.1)

Q (m³/s)	Rank	AEPs (with ARI in brackets), according to:							
		Weibull	Adamowski	Beard	Tukey	Blom	Cunane	Gringorten	Hazen
376	37	0.493 (2.0)	0.493 (2.0)	0.493 (2.0)	0.493 (2.0)	0.493 (2.0)	0.493 (2.0)	0.493 (2.0)	0.493 (2.0)
366	38	0.507 (2.0)	0.507 (2.0)	0.507 (2.0)	0.507 (2.0)	0.507 (2.0)	0.507 (2.0)	0.507 (2.0)	0.507 (2.0)
335	39	0.520 (1.9)	0.520 (1.9)	0.520 (1.9)	0.520 (1.9)	0.520 (1.9)	0.520 (1.9)	0.520 (1.9)	0.520 (1.9)
335	40	0.533 (1.9)	0.534 (1.9)	0.534 (1.9)	0.534 (1.9)	0.534 (1.9)	0.534 (1.9)	0.534 (1.9)	0.534 (1.9)
321	41	0.547 (1.8)	0.547 (1.8)	0.547 (1.8)	0.547 (1.8)	0.547 (1.8)	0.547 (1.8)	0.547 (1.8)	0.547 (1.8)
317	42	0.560 (1.8)	0.560 (1.8)	0.561 (1.8)	0.561 (1.8)	0.561 (1.8)	0.561 (1.8)	0.561 (1.8)	0.561 (1.8)
310	43	0.573 (1.7)	0.574 (1.7)	0.574 (1.7)	0.574 (1.7)	0.574 (1.7)	0.574 (1.7)	0.574 (1.7)	0.574 (1.7)
305	44	0.587 (1.7)	0.587 (1.7)	0.587 (1.7)	0.587 (1.7)	0.588 (1.7)	0.588 (1.7)	0.588 (1.7)	0.588 (1.7)
288	45	0.600 (1.7)	0.601 (1.7)	0.601 (1.7)	0.601 (1.7)	0.601 (1.7)	0.601 (1.7)	0.601 (1.7)	0.601 (1.7)
288	46	0.613 (1.6)	0.614 (1.6)	0.614 (1.6)	0.614 (1.6)	0.614 (1.6)	0.615 (1.6)	0.615 (1.6)	0.615 (1.6)
288	47	0.627 (1.6)	0.628 (1.6)	0.628 (1.6)	0.628 (1.6)	0.628 (1.6)	0.628 (1.6)	0.628 (1.6)	0.628 (1.6)
288	48	0.640 (1.6)	0.641 (1.6)	0.641 (1.6)	0.641 (1.6)	0.641 (1.6)	0.642 (1.6)	0.642 (1.6)	0.642 (1.6)
270	49	0.653 (1.5)	0.654 (1.5)	0.655 (1.5)	0.655 (1.5)	0.655 (1.5)	0.655 (1.5)	0.655 (1.5)	0.655 (1.5)
254	50	0.667 (1.5)	0.668 (1.5)	0.668 (1.5)	0.668 (1.5)	0.668 (1.5)	0.668 (1.5)	0.669 (1.5)	0.669 (1.5)
246	51	0.680 (1.5)	0.681 (1.5)	0.682 (1.5)	0.682 (1.5)	0.682 (1.5)	0.682 (1.5)	0.682 (1.5)	0.682 (1.5)
246	52	0.693 (1.4)	0.695 (1.4)	0.695 (1.4)	0.695 (1.4)	0.695 (1.4)	0.695 (1.4)	0.696 (1.4)	0.696 (1.4)
246	53	0.707 (1.4)	0.708 (1.4)	0.708 (1.4)	0.709 (1.4)	0.709 (1.4)	0.709 (1.4)	0.709 (1.4)	0.709 (1.4)
226	54	0.720 (1.4)	0.721 (1.4)	0.722 (1.4)	0.722 (1.4)	0.722 (1.4)	0.722 (1.4)	0.723 (1.4)	0.723 (1.4)
215	55	0.733 (1.4)	0.735 (1.4)	0.735 (1.4)	0.735 (1.4)	0.736 (1.4)	0.736 (1.4)	0.736 (1.4)	0.736 (1.4)
213	56	0.747 (1.3)	0.748 (1.3)	0.749 (1.3)	0.749 (1.3)	0.749 (1.3)	0.749 (1.3)	0.750 (1.3)	0.750 (1.3)
208	57	0.760 (1.3)	0.762 (1.3)	0.762 (1.3)	0.762 (1.3)	0.763 (1.3)	0.763 (1.3)	0.763 (1.3)	0.764 (1.3)
205	58	0.773 (1.3)	0.775 (1.3)	0.776 (1.3)	0.776 (1.3)	0.776 (1.3)	0.776 (1.3)	0.777 (1.3)	0.777 (1.3)
172	59	0.787 (1.3)	0.789 (1.3)	0.789 (1.3)	0.789 (1.3)	0.790 (1.3)	0.790 (1.3)	0.790 (1.3)	0.791 (1.3)
172	60	0.800 (1.3)	0.802 (1.2)	0.803 (1.2)	0.803 (1.2)	0.803 (1.2)	0.803 (1.2)	0.804 (1.2)	0.804 (1.2)
172	61	0.813 (1.2)	0.815 (1.2)	0.816 (1.2)	0.816 (1.2)	0.816 (1.2)	0.817 (1.2)	0.817 (1.2)	0.818 (1.2)
162	62	0.827 (1.2)	0.829 (1.2)	0.829 (1.2)	0.830 (1.2)	0.830 (1.2)	0.830 (1.2)	0.831 (1.2)	0.831 (1.2)
142	63	0.840 (1.2)	0.842 (1.2)	0.843 (1.2)	0.843 (1.2)	0.843 (1.2)	0.844 (1.2)	0.844 (1.2)	0.845 (1.2)
117	64	0.853 (1.2)	0.856 (1.2)	0.856 (1.2)	0.857 (1.2)	0.857 (1.2)	0.857 (1.2)	0.858 (1.2)	0.858 (1.2)
115	65	0.867 (1.2)	0.869 (1.2)	0.870 (1.1)	0.870 (1.1)	0.870 (1.1)	0.871 (1.1)	0.871 (1.1)	0.872 (1.1)
115	66	0.880 (1.1)	0.883 (1.1)	0.883 (1.1)	0.883 (1.1)	0.884 (1.1)	0.884 (1.1)	0.885 (1.1)	0.885 (1.1)
114	67	0.893 (1.1)	0.896 (1.1)	0.897 (1.1)	0.897 (1.1)	0.897 (1.1)	0.898 (1.1)	0.898 (1.1)	0.899 (1.1)
106	68	0.907 (1.1)	0.909 (1.1)	0.910 (1.1)	0.910 (1.1)	0.911 (1.1)	0.911 (1.1)	0.911 (1.1)	0.912 (1.1)
105	69	0.920 (1.09)	0.923 (1.08)	0.924 (1.08)	0.924 (1.08)	0.924 (1.08)	0.925 (1.08)	0.925 (1.08)	0.926 (1.08)
100	70	0.933 (1.07)	0.936 (1.07)	0.937 (1.07)	0.937 (1.07)	0.938 (1.07)	0.938 (1.07)	0.938 (1.07)	0.939 (1.06)
91	71	0.947 (1.06)	0.950 (1.05)	0.950 (1.05)	0.951 (1.05)	0.951 (1.05)	0.951 (1.05)	0.952 (1.05)	0.953 (1.05)
79	72	0.960 (1.04)	0.963 (1.04)	0.964 (1.04)	0.964 (1.04)	0.965 (1.04)	0.965 (1.04)	0.965 (1.04)	0.966 (1.03)
76	73	0.973 (1.03)	0.977 (1.02)	0.977 (1.02)	0.978 (1.02)	0.978 (1.02)	0.978 (1.02)	0.979 (1.02)	0.980 (1.02)
72	74	0.987 (1.01)	0.990 (1.01)	0.991 (1.01)	0.991 (1.01)	0.992 (1.01)	0.992 (1.01)	0.992 (1.01)	0.993 (1.01)

APPENDIX 3: Data

APPENDIX 3-A: AMS at primary reservoir flow sites

AMS of flood peaks at identified primary reservoir flow sites, intended for all research

<u>Note:</u> The hydrological year (Hydro year), also referred to as a water year, is globally defined by the calendar year in which it ends, so the 2018 Hydro year (South Africa) started on 1 October 2017 and ended on 30 September 2018.

Hydro year	Reservoirs (Hydro flow site numbers)											
	A3R002	B1R001	B2R001	C1R002	C2R001	C5R002	C9R002	D3R002	J1R003	J2R001	N2R001	Q1R001
1904								1 235				
1905		16	2.2	127	3.3			751				
1906		12	20	270	7.8			997				
1907	37	498	107	889	12			997				
1908	52	63	31	489	2.1			789				
1909	127	891	633	974	11		3 633	2 665				
1910	8.8	378	139	824	6.9		2 235	1 369				
1911	37	378	24	1 437	20		2 247	663				
1912	13	21	4.2	259	3.0		1 055	598				
1913	38	34	29	114	10	166	877	1 235				
1914	44	13	12	116	9.1	142	701	106				
1915	260	227	584	386	74	515	3 485	2 089				
1916	5.4	68	45	446	7.1	255	2 695	789				
1917	35	49	12	328	9.7	395	467	1 817				
1918	42	625	417	974	21	83	4 807	1 510				
1919	26	305	77	424	36	297	3 231	235				
1920	37	29	32	94	8.7	1 751	725	5 194		10		
1921	55	363	292	589	20	130	1 229	655	391	8.8		
1922	2.8	188	12	838	11	311	2 271	1 161	168	21		
1923	21	449	135	1 320	22	1 580	2 417	4 094	86	32	125	
1924	125	238	178	119	6.8	503	803	2 161	34	1.3	253	
1925	52	96	162	558	22	1 427	3 150	10 650	741	17	218	23
1926	15	3.5	58	104	23	251	461	534	1.1	0.9	101	145
1927	58	106	24	210	4.6	336	685	2 345	65	3.2	128	147
1928	71	135	165	162	5.2	213	652	888	6.9	28	2 260	46
1929	85	68	126	366	8.2	270	981	888	365	7.3	305	68
1930	60	409	321	584	20	1 061	1 468	2 665	9.1	8.4	770	131
1931	37	77	8.1	120	4.1	284	-99	3 424	15	2.9	166	80
1932	41	49	56	142	10	572	229	910	19	78	3 620	50
1933	68	91	21	56	11	331	233	395	67	-99	275	590
1934	506	218	326	658	16	772	-99	6 930	512	-99	202	350
1935	37	38	82	248	18	447	253	2 161	445	-99	223	387
1936	69	642	935	574	7.2	42	1 951	613	129	-99	448	118
1937	140	514	633	578	20	919	3 116	3 424	187	-99	161	23
1938	34	128	56	439	24	5.8	239	1 817	736	-99	443	110
1939	37	473	288	998	21	186	2 411	1 510	532	-99	343	37
1940	1.8	345	110	509	4.0	197	712	1 185	73	-99	202	151

Hydro year	Reservoirs (Hydro flow site numbers)											
	A3R002	B1R001	B2R001	C1R002	C2R001	C5R002	C9R002	D3R002	J1R003	J2R001	N2R001	Q1R001
1941	16	683	455	560	11	500	1 071	1 112	206	-99	153	122
1942	30	90	18	218	10	178	301	1 984	86	-99	2 670	46
1943	225	248	292	304	8.9	1 450	1 091	2 969	40	4.3	94	181
1944	119	1 537	935	1 045	65	161	3 651	4 094	47	11	356	97
1945	1.1	294	67	230	9.3	25	566	1 883	95	1.1	82	318
1946	210	346	383	399	20	207	422	1 510	16	7.4	438	15
1947	1.2	236	70	257	9.2	58	437	647	14	11	161	184
1948	0.7	97	108	336	12	1 865	596	3 102	17	3.6	1 210	69
1949	2.3	211	63	119	3.4	98	208	613	183	8.0	166	129
1950	5.2	275	17	468	9.8	1 270	566	1 753	1.1	13	433	57
1951	9.6	39	142	215	2.5	61	191	2 270	95	42	152	535
1952	2.3	-99	12	339	20	185	189	1 817	658	5.7	621	83
1953	13	-99	12	549	7.6	85	985	869	207	11	77	42
1954	1.9	35	16	162	3.1	324	222	2 270	82	14	2 355	340
1955	102	727	80	1 194	19	812	1 517	5 044	41	12	245	310
1956	1.7	1 500	131	684	4.0	268	471	3 424	-99	2.4	73	36
1957	7.8	167	270	733	18	168	1 288	5 794	350	4.4	110	48
1958	31	148	51	468	2.5	442	3 456	4 097	66	3.7	166	53
1959	10	97	440	431	2.6	70	110	3 803	311	2.1	113	88
1960	1.2	42	9.9	175	3.7	85	304	1 802	26	2.0	100	85
1961	47	47	60	478	3.2	138	782	1 967	100	10	2 200	57
1962	12	95	114	337	18	424	325	4 029	80	4.6	171	67
1963	0.8	30	18	228	6.6	402	353	3 124	102	3.9	240	117
1964	3.1	38	21	240	5.8	72	206	3 231	64	16	149	90
1965	1.6	22	87	531	4.3	117	957	4 092	14	1.6	73	42
1966	29	7.1	57	42	39	695	764	5 411	2.5	28	60	16
1967	49	84	106	756	30	335	2 664	8 007	577	6.3	244	110
1968	2.6	18	26	146	6.8	75	85	1 300	115	3.9	157	730
1969	7.1	12	26	129	3.3	742	150	1 815	49	18	68	20
1970	2.8	47	129	272	3.8	169	168	1 802	13	5.6	192	42
1971	29	65	207	125	20	172	425	-99	923	17	5 090	248
1972	12	78	398	888	12	1 555	840	6 530	22	4.6	30	86
1973	11	23	8.5	63	4.5	12	100	625	31	0.6	81	47
1974	5.6	213	15	352	8.2	6 425	1 510	6 760	59	5.0	2 015	44
1975	12	780	218	1 827	16	317	4 030	2 280	23	8.1	33	805
1976	30	331	44	602	112	1 030	2 375	8 520	276	6.6	194	13
1977	8.1	431	21	881	18	132	2 900	6 610	95	17	2 081	110
1978	21	314	324	182	91	413	2 655	2 290	265	4.4	11	115
1979	4.5	13	7.0	163	7.5	183	260	4 630	19	4.6	615	40
1980	32	1 215	48	252	11	63	1 010	2 680	123	46	56	130
1981	43	1 040	76	245	29	125	900	4 795	5 475	191	975	47
1982	33	92	69	143	56	5.1	140	1 825	1 675	22	230	184
1983	6.0	111	37	106	43	14	105	2 130	217	12	1 380	61
1984	6.0	1 400	67	282	29	293	215	1 455	200	11	160	58
1985	99	17	76	144	26	73	350	1 530	199	13	361	42

Hydro year	Reservoirs (Hydro flow site numbers)											
	A3R002	B1R001	B2R001	C1R002	C2R001	C5R002	C9R002	D3R002	J1R003	J2R001	N2R001	Q1R001
1986	26	7.6	105	414	36	207	205	2 315	310	30	422	110
1987	33	650	250	126	55	317	220	4 280	196	36	86	97
1988	30	2 565	112	651	40	9 800	6 340	11 460	347	6.8	51	55
1989	2.7	116	78	420	44	420	2 080	5 240	509	15	73	178
1990	4.5	305	175	390	28	110	350	1 035	261	29	657	78
1991	5.4	144	112	177	17	1 445	490	2 480	121	3.1	31	32
1992	1.0	52	74	121	28	347	95	5 220	66	119	233	45
1993	9.1	72	86	145	4.0	99	830	855	-99	51	218	57
1994	6.8	62	72	286	32	893	365	3 395	23	12	332	155
1995	16	34	26	300	24	78	145	395	16	9.9	760	90
1996	45	770	995	2 135	70	135	3 275	3 285	358	30	152	35
1997	35	198	92	665	39	62	2 600	1 850	255	25	725	73
1998	52	382	188	618	63	109	1 180	3 200	149	7.3	16	48
1999	15	194	81	367	94	3.2	600	1 195	123	14	136	54
2000	33	740	323	1 045	96	101	2 240	1 915	234	62	172	77
2001	2.7	313	59	568	43	148	765	1 595	45	18	415	153
2002	11	85	40	84	41	321	780	4 055	86	19	284	81
2003	7.8	69	89	212	56	86	180	1 395	317	29	287	71
2004	10	58	113	242	39	123	520	1 085	1 005	92	166	55
2005	18	99	72	365	49	41	185	820	179	23	381	265
2006	33	280	82	2 275	11	330	1 490	2 945	226	43	406	113
2007	4.6	260	35	635	25	281	300	3 855	22	11	286	84
2008	36	566	108	821	8.0	141	510	1 465	1 165	125	136	66
2009	7.0	213	166	1 285	27	199	590	2 475	508	33	95	61
2010	0.9	170	81	1 250	44	379	3 935	3 680	29	9.9	132	141
2011	5.4	335	116	825	41	4 550	3 800	5 285	397	35	222	68
2012	-99	31	58	215	22	57	205	1 420	68	14	127	54
2013	-99	30	31	515	44	75	190	885	169	17	88	38
2014	40	663	84	1 275	42	150	1 315	2 254	3 010	30	454	61
2015	7.1	36	264	309	38	281	495	1 236	17	16	122	108
2016	13	82	40	288	9.9	28	187	692	358	6.9	95	99
2017	26	120	60	1 265	35	175	2 520	2 672	7.4	2.6	117	91
2018	0.6	94	189	374	39	366	995	2 372	182	6.2	36	200

Note: -99 indicates missing or suspect data not considered

APPENDIX 3-B: AMS at primary weir flow sites

AMS of flood peaks at identified primary weir flow sites, intended for all research

Hydro year	Flow site no.			Hydro year	Flow site no.			Hydro year	Flow site no.		
	D7H005	V6H002	X1H001		D7H005	V6H002	X1H001		D7H005	V6H002	X1H001
1910			202	1947	679	858	92	1984	323	498	691
1911			349	1948	3 851	1 083	175	1985	284	705	425
1912			150	1949	340	623	95	1986	363	534	104
1913			143	1950	1 852	1 191	181	1987	389	2 134	229
1914			139	1951	1 749	890	146	1988	7 403	38	173
1915			275	1952	1 355	962	181	1989	2 849	1 425	188
1916			146	1953	1 079	962	780	1990	385	673	373
1917			235	1954	1 627	1 353	544	1991	929	566	355
1918			1 498	1955	4 463	2 187	784	1992	493	427	55
1919			275	1956	2 240	1 036	386	1993	130	411	25
1920			202	1957	4 451	1 563	173	1994	285	935	22
1921			96	1958	4 684	1 579	247	1995	226	368	188
1922			8.3	1959	2 085	1 173	446	1996	2 369	1 174	509
1923			655	1960	882	872	215	1997	2 498	901	90
1924			117	1961	2 417	898	1 174	1998	1 275	612	92
1925			410	1962	2 066	1 018	382	1999	484	345	59
1926			61	1963	2 226	1 245	464	2000	1 912	1 097	900
1927			90	1964	1 452	1 517	227	2001	727	526	188
1928		640	190	1965	2 558	1 102	285	2002	1 825	548	172
1929		1 035	129	1966	3 981	1 173	153	2003	767	213	43
1930		1 838	198	1967	6 072	1 318	285	2004	217	766	63
1931		854	173	1968	784	348	551	2005	162	691	89
1932		883	190	1969	1 294	56	256	2006	2 000	773	144
1933	564	377	76	1970	1 140	766	202	2007	724	459	63
1934	6 583	973	181	1971	265	892	91	2008	176	849	78
1935	2 260	1 467	181	1972	2 085	1 549	243	2009	1 024	919	369
1936	1 236	1 035	168	1973	399	652	155	2010	3 135	611	535
1937	2 622	1 252	523	1974	8 315	950	1 961	2011	4 802	1 306	354
1938	1 789	550	103	1975	3 054	1 828	270	2012	219	896	116
1939	2 971	1 483	2 481	1976	4 699	1 710	1 580	2013	337	941	153
1940	2 280	946	403	1977	1 476	998	55	2014	1 048	376	364
1941	1 479	1 238	226	1978	1 401	156	431	2015	250	505	66
1942	2 102	1 514	84	1979	345	892	71	2016	236	617	75
1943	2 307	2 297	326	1980	242	449	238	2017	571	1 687	136
1944	3 513	2 438	345	1981	365	865	197	2018	599	1 319	290
1945	1 606	918	127	1982	376	628	116	2019	325	295	231
1946	993	754	157	1983	237	341	63				

APPENDIX 3-C: AMS at secondary flow sites (group 1)

AMS of flood peaks at secondary flow sites (group1) intended for research, pertaining to a new bounded probability distribution – some of these sites can also be used as test flow sites to assess the results of the record length- and/or plotting position research.

Hydro year	Flow sites (Hydro numbers)												
	A2R001	A2R005	A4R001	A8R001	B3R001	B6R003	B7R001	C3R002	C5R001	C5R003	C7R001	D6R002	E1R002
1919										22			
1920										686			
1921										54	102		
1922										128	662		
1923	224									668	152	30	
1924	37							115	25	98	270	279	
1925	273							101	850	739	901	121	
1926	237							101	3.5	37	63	-99	
1927	38							153	95	366	65	17	
1928	224							182	23	66	66	410	
1929	100							573	72	613	469	64	
1930	56							43	17	54	209	186	
1931	55							686	33	137	40	320	
1932	29							122	66	88	85	296	
1933	69			372				195	14	33	47	126	
1934	444			183	76			146	91	628	613	308	
1935	79			11	15			221	16	184	973	142	545
1936	123	60		446	114			130	7.8	23	96	33	508
1937	198	60		1 683	378			271	26	628	1 250	156	276
1938	256	11		141	11			227	2.2	44	247	208	239
1939	138	90		330	25			55	17	11	373	98	183
1940	150	26		163	28			108	56	54	132	9.8	124
1941	70	85		668	43			77	30	723	252	104	593
1942	37	8.1		572	15			77	28	33	343	53	962
1943	498	36		181	55			86	62	647	185	111	299
1944	1 670	94		82	62			1 142	86	137	238	193	242
1945	203	33		572	14			24	25	335	634	58	1 345
1946	174	16		278	71			97	60	128	53	21	761
1947	73	13		-99	8.4			217	6.6	20	107	12	227
1948	82	15		128	21			205	751	915	416	123	291
1949	70	8.5		-99	11			169	2.1	49	135	-99	159
1950	268	398		-99	23			77	70	781	216	-99	334
1951	83	29		11	12	56	5.7	269	19	36	30	-99	355
1952	20	7.9		21	8.4	39	13	114	4.9	297	180	-99	212
1953	67	25		30	10	366	81	77	14	54	1 480	-99	281
1954	24	5.7		176	20	109	31	77	17	55	170	-99	1 035
1955	777	34		-99	45	30	109	357	395	205	437	-99	473
1956	162	17		-99	665	126	247	321	167	580	23	-99	221
1957	222	10		-99	29	84	18	100	45	49	356	-99	1 010
1958	233	61		-99	7.1	366	171	15	18	29	220	-99	163
1959	79	6.9		-99	7.6	239	123	63	24	54	73	-99	376
1960	38	10		-99	11	652	142	357	91	22	89	-99	107
1961	22	11		-99	15	308	-99	225	30	40	303	-99	432
1962	13	5.3		-99	19	111	16	46	18	23	227	-99	325
1963	26	4.0	36	16	13	108	18	239	67	52	150	-99	1 335
1964	81	0.9	12	19	14	70	16	79	50	66	114	-99	213
1965	44	40	19	56	51	291	31	357	6.9	27	398	9.8	89

Hydro year	Flow sites (Hydro numbers)												
	A2R001	A2R005	A4R001	A8R001	B3R001	B6R003	B7R001	C3R002	C5R001	C5R003	C7R001	D6R002	E1R002
1966	50	30	1	88	28	282	31	184	647	205	123	47	207
1967	805	32	452	45	174	498	61	482	19	102	374	95	1 385
1968	27	3.2	11	2.0	38	34	10	162	13	104	207	18	208
1969	17	5.3	14	7.0	40	240	138	77	34	101	73	419	76
1970	46	1.8	37	12	98	47	5.0	65	26	34	101	94	135
1971	347	23	257	91	25	343	104	83	35	19	153	96	249
1972	442	13	206	70	57	719	388	531	92	160	200	63	75
1973	104	7.9	19	95	4.4	168	6.9	49	5.0	19	224	101	425
1974	125	65	81	70	16	252	101	553	60	220	181	1 000	520
1975	370	32	599	70	335	75	24	260	9.9	12	309	120	154
1976	709	39	1 291	81	62	357	370	600	177	165	140	273	550
1977	984	50	623	353	67	255	140	110	96	67	111	22	719
1978	824	37	232	333	72	435	76	100	57	35	108	12	171
1979	45	54	7	108	9.6	229	15	18	7.6	9.2	156	57	232
1980	283	22	428	129	136	296	6.5	25	17	16	97	61	166
1981	702	41	128	176	73	309	40	102	300	130	218	91	841
1982	120	1.5	146	60	31	32	13	14	28	27	52	10	119
1983	49	0.9	19	29	126	29	7.6	31	39	17	123	-99	619
1984	125	10	27	1 280	20	50	81	16	60	15	106	-99	871
1985	179	39	28	593	3.0	486	60	17	134	15	176	35	509
1986	149	1.6	34	88	17	174	14	23	66	96	103	138	380
1987	244	6.7	35	101	6.3	230	18	99	133	91	175	12	205
1988	365	39	58	303	14	494	347	2 760	1 670	2 670	847	1 190	140
1989	281	10	35	62	13	203	15	196	-99	201	302	43	175
1990	193	7.7	49	103	32	102	70	14	-99	66	135	7.7	492
1991	89	24	90	85	19	380	20	300	48	82	464	1.3	586
1992	115	3.2	38	28	5.9	19	0.2	15	74	155	25	1.7	445
1993	101	10	30	74	2.3	1 018	205	49	3.1	32	92	1.0	238
1994	138	11	109	190	18	331	38	74	71	500	242	78	452
1995	360	22	33	47	23	218	93	90	3.3	48	34	21	177
1996	940	107	1 070	175	336	795	115	45	136	462	515	27	734
1997	742	114	98	77	82	660	11	32	7.3	80	423	20	420
1998	181	13	24	77	62	267	132	92	19	118	47	2.1	204
1999	202	31	42	140	96	130	153	29	2.7	32	164	-99	347
2000	1 360	72	680	2 795	89	970	89	92	99	91	304	30	166
2001	215	13	97	40	14	184	5.6	93	60	111	77	286	408
2002	170	8.8	121	97	33	86	19	167	141	870	247	-99	498
2003	192	8.4	14	21	59	70	9.7	60	37	66	25	46	157
2004	202	31	171	128	37	424	49	384	33	52	26	148	71
2005	264	42	9	77	26	63	2.0	129	11	36	363	3.6	475
2006	452	39	283	74	97	865	116	365	153	200	280	77	363
2007	139	9.3	28	30	8	77	2.0	61	25	40	50	6.8	910
2008	309	9.1	355	50	259	131	27	80	117	31	81	23	1 090
2009	221	62	282	15	45	387	195	99	111	32	256	54	895
2010	474	37	254	417	67	109	34	370	65	80	1 050	97	248
2011	1025	166	97	216	43	230	81	278	760	582	328	27	283
2012	252	13	155	46	20	1 615	915	56	53	19	86	22	304
2013	367	7.8	93	767	13	259	38	13	27	19	63	152	403
2014	668	51	1 080	356	57	202	322	75	25	75	84	10	377
2015	141	36	68	33	17	42	189	75	13	36	59		121
2016	537	40	17	230	23	53	1.3	25	-99	42	35		310
2017	565	55	44	50	41	456	38	275	10	77	625		30
2018	655	19	10		32	39	5.7	58	16	119	127		255
2019	515	1.8	23		9	184	7.4	65	54	72	270		

APPENDIX 3-D: AMS at secondary flow sites (group 2)

AMS of flood peaks at secondary flow sites (group 2) intended for additional research sites pertaining to a new bounded probability distribution – some of these sites can also be used as test flow sites to assess the results of the record length- and/or plotting position research.

Hydro year	Flow sites (Hydro numbers												
	H5H004	J2R003	J3R001	N1R001	Q4R002	Q5R001	S6R002	T2R001	U2R001	U3R001	V1R003	V3R003	W4R001
1906						534							
1907						534							
1908						197							
1909						570							
1910						1 305							
1911						2 669							
1912			42			139							
1913			130			1 305							
1914			48			607							
1915			157			400							
1916			2 755			607							
1917			84			903							
1918			782			2 531							
1919			13			107							
1920			101			948							
1921			300			1 360							
1922			47			570							
1923			43			607							
1924			36			123							
1925			318	43		499							
1926			53	12	31	91							
1927			55	1 370	61	135							
1928			19	950	128	204							
1929			15	193	276	113							
1930			42	220	331	154							
1931		8.2	8.5	1 125	56	125							
1932		133	2 130	1 230	784	268					115		
1933		19	27	172	260	152					-99		
1934		0.5	8.0	327	274	1 003					172		
1935		5.5	608	207	112	303					335		
1936		0.3	30	80	62	117					246		
1937		3.2	42	294	67	180					288		
1938		1.4	13	291	41	236					115		
1939		15	38	226	107	606					910		
1940		2.9	16	274	118	693					246		
1941		1.3	17	93	204	322					-99		
1942		67	39	107	114	322					-99		
1943		3.0	7.3	77	192	121					288		

Hydro year	Flow sites (Hydro numbers											
	H5H004	J2R003	J3R001	N1R001	Q4R002	Q5R001	S6R002	T2R001	U2R001	U3R001	V1R003	V3R003	W4R001
1944		2.2	-99	451	152	2 145					385		546
1945		5.8	160	90	81	136					-99		203
1946		1.2	21	114	125	322					246		300
1947		0.2	29	38	101	277					288		-99
1948		4.4	16	740	87	195	464				91	35	390
1949		4.6	50	83	4.8	13	2				-99	41	-99
1950		2.1	112	423	802	2 349	41				72	28	353
1951		38	510	269	265	119	43				-99	33	273
1952		1.2	40	205	82	91	34				-99	34	264
1953		5.4	88	71	65	55	50				-99	132	240
1954		20	572	72	322	1 340	341				-99	35	250
1955	973	0.3	34	150	74	9.0	-99				920	149	1 069
1956	559	2.4	21	14	33	38	-99				288	72	2 259
1957	1 087	3.4	21	80	221	570	16		142		335	51	426
1958	496	0.7	10	13	53	47	15		50		385	77	2 451
1959	541	2.3	22	9.5	88	22	27		160		411	35	104
1960	281	23	8.4	106	29	69	25		18		385	118	142
1961	577	6.6	3.7	1 565	105	193	24		41	80	172	114	116
1962	635	2.5	227	39	84	293	36		73	7.0	142	70	369
1963	780	2.6	62	317	160	384	23		46	16	226	134	5 570
1964	257	4.2	173	24	47	31	33		40	46	117	37	213
1965	153	0.6	11	38	29	22	6.7		50	11	172	150	1 835
1966	685	10	24	31	54	144	19		40	3.0	801	39	1 407
1967	791	19	83	75	157	236	25		65	27	205	138	543
1968	625	4.9	47	43	19	59	46		338	55	76	102	305
1969	325	0.9	8.0	96	100	135	6.1		93	93	321	59	-99
1970	173	-99	7.2	779	598	91	10		138	24	100	55	-99
1971	431	34	233	341	405	1 128	144		285	138	658	33	163
1972	183	1.0	4.6	21	102	107	128		127	87	625	127	812
1973	688	3.6	8.9	113	209	91	14		200	98	254	33	2 100
1974	1 145	7.7	17	3 470	449	3 888	65		320	158	799	50	482
1975	431	15	8.5	41	38	-99	14		127	157	853	106	1 040
1976	899	5.3	9.4	194	381	-99	316		175	305	910	78	1 010
1977	1 420	93	88	137	118	96	55		340	503	664	138	255
1978	327	19	2.7	6.0	40	74	115		140	535	727	170	355
1979	626	8.0	18	13	98	120	195	36	144	118	915	19	351
1980	347	0.2	1.6	3.9	20	159	3.2	8.4	104	196	366	36	166
1981	2 137	59	880	197	290	91	64	40	71	47	533	78	268
1982	179	8.6	113	9.0	51	90	1.3	29	74	14	114	52	80
1983	714	2.6	385	293	186	480	8.7	22	263	26	79	13	31
1984	1 382	14	104	52	17	244	11	56	64	285	317	241	16 450
1985	455	5.0	2.6	117	115	1 615	54	365	347	552	784	79	1 600
1986	843	6.7	27	268	241	655	880	92	100	69	860	31	403
1987	581	21	39	23	48	90	27	164	1 565	1 700	935	89	1 055

Hydro year	Flow sites (Hydro numbers												
	H5H004	J2R003	J3R001	N1R001	Q4R002	Q5R001	S6R002	T2R001	U2R001	U3R001	V1R003	V3R003	W4R001
1988	306	12	6.4	223	80	259	34	71	231	677	967	98	465
1989	595	1.5	5.1	172	42	360	57	398	87	374	700	88	730
1990	667	20	75	181	40	421	110	146	169	900	538	80	3 405
1991	1 220	0.3	40	8.5	142	57	12	68	253	94	810	151	1 470
1992	973	7.0	51	6.1	22	284	41	64	225	123	2 915	81	417
1993	1 744	6.8	600	224	68	35	24	28	70	57	483	149	462
1994	1 178	6.9	163	144	183	-99	136	210	240	198	453	92	670
1995	121	4.7	16	44	16	-99	103	336	68	76	162	44	545
1996	1 710	3.6	56	13	192	-99	84	428	255	273	895	345	2 090
1997	868	196	1 435	116	56	-99	255	424	144	191	208	124	815
1998	525	7.6	42	7.0	75	-99	77	798	99	64	310	304	910
1999	272	45	22	139	16	-99	46	214	185	56	1 020	62	670
2000	324	12	9.3	7.4	150	-99	180	435	420	592	1 400	168	1 130
2001	654	1.0	12	31	58	-99	154	657	142	69	1 380	113	770
2002	537	26	98	138	85	-99	329	647	139	298	495	88	1 585
2003	1 114	9.0	719	16	30	-99	51	127	80	10	215	96	700
2004	184	4.0	16	61	252	-99	280	473	100	19	402	120	680
2005	522	1.8	8.8	195	106	93	295	920	136	31	475	94	290
2006	709	155	1 160	630	276	635	86	614	122	187	424	223	161
2007	622	8.7	7.2	77	78	255	141	325	47	55	105	46	945
2008	610	28	1 730	95	56	216	92	409	119	54	305	83	275
2009	1 092	8.3	42	74	89	248	41	239	198	58	713	92	450
2010	358	7.8	2.5	116	71	260	35	342	34	25	376	242	1 305
2011	497	15	57	567	261	616	476	381	33	13	1 275	139	597
2012	618	45	380	70	31	82	125	305	73	25	270	65	802
2013	608	22	43	32	24	86	218	665	236	111	213	96	488
2014	605	45	53	71	224	276	112	136	120	13	106	382	445
2015	541	3.1	119	87	118	84	123	140	234	8.7	825	94	135
2016	368	3.5	57			155		235	11	5.2		54	60
2017	68	85	2.0			400		82	238	24		70	130
2018	550	2.3							87	67		243	
2019	465								112				

APPENDIX 3-E: Probability distribution research – flow site statistics

Info from flow site to be considered in probability distribution research.

Site	n (used)	Full record							logtransformed data		
		Qmin	Qmax	Qmdn	Qave	SD	g	kurt	Qave	SD	g
A3R002	110	0.60	506	19.5	37.4	63.3	4.696	28.939	1.1870	0.6272	- 0.2609
B1R001	112	3.50	2 565	124	280	384	3.032	12.368	2.1113	0.5814	- 0.2000
B2R001	114	2.20	995	79.0	144	188	2.690	8.125	1.8742	0.5241	- 0.2069
C1R002	114	42.0	2 275	367	494	421	1.849	4.196	2.5525	0.3603	- 0.0879
C2R001	114	2.10	112	18.0	23.2	22.0	1.755	3.403	1.1805	0.4206	- 0.0891
C5R002	106	3.20	9 800	210	547	1 223	5.642	36.533	2.3448	0.5784	- 0.2679
C9R002	108	85.0	6 340	719	1 216	1 251	1.507	2.075	2.8514	0.4723	0.0154
D3R002	114	106	11 460	1 976	2 614	2 094	1.769	3.911	3.2880	0.3534	- 0.4183
D7H005	87	130	8 315	1 294	1 730	1 690	1.762	3.504	3.0293	0.4544	- 0.1385
J1R003	96	1.10	5 475	118	298	665	5.934	41.598	2.0094	0.6807	- 0.4054
J2R001	89	0.60	191	11.0	20.4	29.1	3.573	15.546	1.0413	0.4889	- 0.0276
N2R001	96	11.0	5 090	193	467	796	3.523	14.379	2.3425	0.5002	0.4152
Q1R001	94	13.0	805	80.5	123	139	3.037	10.341	1.9297	0.3546	0.4590
V6H002	92	38.0	2 438	897	949	492	0.761	0.665	2.9031	0.2952	- 1.7737
X1H001	110	8.30	2 481	188	302	371	3.564	15.155	2.2927	0.3960	- 0.0222
A2R001	97	13.0	1 670	181	278	301	2.133	5.468	2.2140	0.4723	- 0.1591
A2R005	84	0.90	398	22.0	33.9	49.8	5.126	34.648	1.2595	0.5156	- 0.3623
A4R001	57	1.17	1 291	58.0	181	281	2.509	6.254	1.8369	0.6333	0.0081
A8R001	74	2.00	2 795	89.5	227	409	4.380	23.051	1.9947	0.5676	- 0.0871
B3R001	86	2.30	665	27.0	58.5	96.5	4.084	20.109	1.4752	0.4774	0.3890
B6R003	69	19.0	1 615	229	283	282	2.254	6.915	2.2574	0.4378	- 0.2135
B7R001	68	0.20	915	38.0	89.4	137	3.795	19.482	1.5493	0.6734	- 0.5807
C3R002	96	12.6	2 760	99.5	192	318	6.008	45.414	2.0280	0.4585	0.1200
C5R001	93	2.10	1 670	34.0	101	226	4.807	27.223	1.5672	0.5846	0.2626
C5R003	101	9.20	2 670	67.0	193	336	4.567	29.383	1.9266	0.5304	0.5590
C7R001	99	23.0	1 480	170	251	264	2.413	6.764	2.2156	0.4061	0.0277
D6R002	71	1.00	1 190	61.0	122	195	3.870	17.691	1.7007	0.6470	- 0.5122
E1R002	84	30.0	1 385	307	408	309	1.459	1.758	2.4964	0.3268	- 0.2233
H5H004	65	67.9	2 137	595	661	409	1.367	2.344	2.7364	0.2880	- 0.6244
J2R003	87	0.20	196	5.80	17.5	32.9	3.557	14.049	0.7764	0.6592	- 0.0257
J3R001	105	1.60	2 755	42.0	184	428	3.946	17.425	1.6415	0.7085	0.4581
N1R001	91	3.90	3 470	106	243	455	4.760	28.956	1.9797	0.6149	- 0.0991
Q4R002	90	4.80	802	93.5	142	147	2.514	7.866	1.9698	0.4129	- 0.1880
Q5R001	99	9.00	3 888	236	452	639	2.986	10.487	2.3589	0.5166	- 0.0062
S6R002	66	1.34	881	50.4	108	146	3.007	11.992	1.7114	0.5761	- 0.3836
T2R001	39	8.40	920	235	285	237	0.883	0.139	2.2488	0.4966	- 0.7360
U2R001	63	11.0	1 565	127	166	202	5.565	38.024	2.0706	0.3541	- 0.0946
U3R001	58	3.00	1 700	69.0	174	279	3.521	15.780	1.8589	0.6030	- 0.0200
V1R003	79	41.0	2 140	415	566	437	1.258	410	2.6100	0.3824	- 0.4312
V3R003	71	13.4	382	88.0	102	73.9	1.793	85.5	1.9099	0.2977	- 0.0709
W4R001	85	31.0	16 450	488	1 003	1 931	6.470	485	2.7230	0.4612	0.2834

Site	n (used)	Data record											
		Full			Excluding Qmax [1]					Excluding Qmax and Qmin [2]			
		Qmin	Qmax	Qgmn	Qmdn*	Qave*	SD*	g*	kurt*	Qmdn*°	Qave*°	SD*°	g*°
A3R002	110	0.60	506	19.5	18.0	33.1	44.7	2.916	10.184	19.5	33.4	44.8	2.908
B1R001	112	3.50	2 565	124	120	259	317	2.155	5.039	124	261	318	2.147
B2R001	114	2.20	995	79.0	78.0	137	171	2.637	8.224	79.0	138	171	2.631
C1R002	114	42.0	2 275	367	366	478	388	1.634	3.211	367	482	387	1.637
C2R001	114	2.10	112	18.0	18.0	22.5	20.4	1.584	2.612	18.0	22.6	20.4	1.580
C5R002	106	3.20	9 800	210	207	459	824	5.076	31.367	210	463	827	5.060
C9R002	108	85.0	6 340	719	712	1168	1154	1.226	0.446	719	1 178	1 154	1.217
D3R002	114	106	11 460	1 976	1 967	2 536	1928	1.532	2.804	1 976	2 558	1 923	1.542
D7H005	87	130	8 315	1 294	1 285	1 653	1540	1.566	2.678	1 294	1 671	1 540	1.561
J1R003	96	1.10	5 475	118	115	243	398	4.438	26.058	118	246	399	4.424
J2R001	89	0.60	191	11.0	11.0	18.5	22.8	2.878	9.631	11.0	18.9	22.9	2.864
N2R001	96	11.0	5 090	193	192	418	641	2.980	9.176	194	427	645	2.952
Q1R001	94	13.0	805	80.5	80.0	116	121	2.924	10.055	80.5	117	121	2.923
V6H002	92	38.0	2 438	897	896	933	469	0.634	0.397	897	943	462	0.688
X1H001	110	8.30	2 481	188	188	282	308	3.167	12.165	188	284	308	3.169
A2R001	97	13.0	1 670	181	180	263	267	1.749	3.079	181	266	267	1.743
A2R005	84	0.90	398	22.0	22.0	29.5	29.6	1.982	5.273	22.0	29.9	29.7	1.980
A4R001	57	1.17	1 291	58.0	53.5	161	240	2.495	6.452	58.0	164	241	2.473
A8R001	74	2.00	2 795	89.5	88.0	192	278	3.334	13.657	89.5	195	279	3.321
B3R001	86	2.30	665	27.0	26.0	51.3	70.7	3.134	10.583	27.0	51.9	70.9	3.122
B6R003	69	19.0	1 615	229	224	264	233	1.475	2.031	229	268	232	1.472
B7R001	68	0.20	915	38.0	38.0	77.0	92.6	1.829	3.143	38.0	78.2	92.8	1.815
C3R002	96	12.6	2 760	99.5	99.0	165	178	2.614	9.632	99.5	166	178	2.611
C5R001	93	2.10	1 670	34.0	33.5	84.0	156	3.718	13.963	34.0	84.9	157	3.699
C5R003	101	9.20	2 670	67.0	66.5	168	226	1.865	2.280	67.0	170	227	1.852
C7R001	99	23.0	1 480	170	167	239	234	2.144	5.179	170	241	234	2.141
D6R002	71	1.00	1 190	61.0	59.5	106	147	3.704	19.269	61.0	108	148	3.694
E1R002	84	30.0	1 385	307	304	397	291	1.415	1.699	307	401	290	1.428
H5H004	65	67.9	2 137	595	588	638	367	1.050	1.144	581	620	342	0.916
J2R003	87	0.20	196	5.80	5.65	15.5	26.8	3.358	12.699	5.80	15.6	26.9	3.340
J3R001	105	1.60	2 755	42.0	41.0	160	347	3.683	15.027	42.0	161	348	3.664
N1R001	91	3.90	3 470	106	101	207	301	2.785	8.066	106	209	302	2.772
Q4R002	90	4.80	802	93.5	89.0	135	130	2.345	7.598	93.5	136	130	2.350
Q5R001	99	9.00	3 888	236	226	417	538	2.513	6.633	236	422	539	2.504
S6R002	66	1.34	881	50.4	49.8	95.7	110	1.840	3.105	50.4	97.1	110	1.828
T2R001	39	8.40	920	235	225	269	215	0.721	-0.348	235	276	214	0.707
U2R001	63	11.0	1 565	127	125	143	93.8	0.915	0.276	127	145	93.0	0.932
U3R001	58	3.00	1 700	69.0	69.0	147	192	2.119	4.421	69.0	150	192	2.101
V1R003	79	41.0	2 140	415	410	546	401	0.997	0.973	415	553	399	1.000
V3R003	71	13.4	382	88.0	85.5	97.9	66.3	1.586	3.018	88.0	99.1	65.9	1.608
W4R001	85	31.0	16 450	488	485	820	929	2.676	9.057	488	829	931	2.672

[1] * Denotes largest value (Q_{max}) omitted from dataset

[2] *° Denotes largest value (Q_{max}), as well as smallest value (Q_{min}), omitted from dataset

APPENDIX 4: Record length

Contents: Appendix 4-A

List of figures: Appendix 4-A

List of tables: Appendix 4-A

APPENDIX 4-A Record length research

Appendix 4-A covers the research to decide on appropriate record lengths for research. However, only pertinent concise results, at selected flow sites, are provided due to the vast amount of generated data/information. Where deemed necessary, supplementary corresponding tables and figures are presented in the follow-up appendices. (Comprehensive results are available, on request, in electronic format).

4-A1 Record length methodology

It was foreseen that a larger number of flow sites would be required to deal with the primary objective, than the number of flow sites used to address the secondary objectives.

The purpose of this part of the research, therefore, was directed to establish a minimum data record length, that could be used with a reasonable degree of confidence even at lower AEPs, to identify additional appropriate sites for conducting research on the primary objective.

4-A1.1 Prologue

From Section 2.2.5, it was clear that related studies on this subject focussed more on the expected confidence in predictions, with different record lengths.

Bulletin 17C (England *et al*, 2018) recommends a minimum record length of 10 years but, from practical experience, it was considered to be way too short. The limited studies by Hattingh, *et al*. (2011) and Van der Spuy (2018), suggested minimum record lengths of between 30 and 40 years, which seemed more plausible. However, the studies were limited to site-specific investigations and the results needed to be verified, conducting more widespread research.

4-A1.2 The research approach

The objective was to develop a technique for finding a record length that, if extracted from any part of the existing record length, would practically give similar FFA results within specified criteria. An overlapping rolling window (R-W; R-Ws for plural) approach was considered. To select a R-W size the observations made by Alexander (2005a and 2005b) were used as a guideline. He concluded that the length of wet and dry sequences, in SA, typically varied between 6 to 8 years (on average 7 years). He also investigated a linkage with solar activity and concluded that a 21-year periodicity (a multiple of 7) is evident.

Consequently, R-W sizes in periods of 7-year multiples (7ym) were used, since it was expected that it would increase the probability of having almost equal wetter and drier periods. The R-W concept is illustrated in Figure 4-A.1; with a dataset of record length n and an R-W of length m years (R-Wm), the dataset is partitioned into the following number of sub-datasets:

Number of sub-datasets = $n - m + 1$ **Equation 4-A.1**

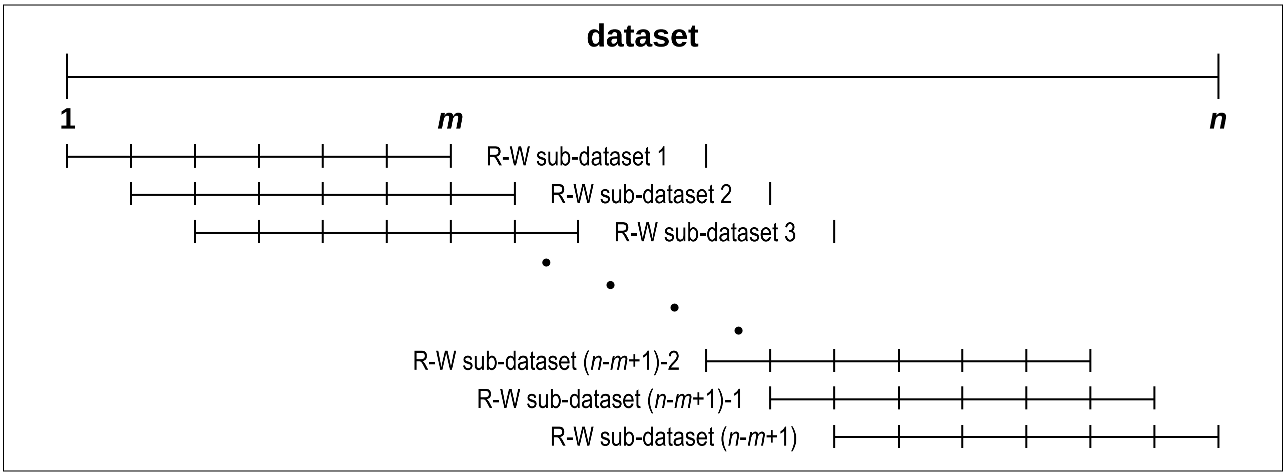


Figure 4-A.1: R-W concept, illustrating partitioned sub-datasets

Applying an R-W approach, on all 15 primary sites, thus required a large number of analyses (12 787) to be performed. A simple C++ program was written (Appendix 4-N), producing output to be used as input to Excel-spreadsheets. To ensure consistency, in applying the methodology, data records were considered as follows:

- no data points were considered/treated as outliers
- gaps in records were considered as missing data and were not filled

To properly illustrate the methodology an arbitrary AMS flow record of 40 years, presented in Table 4-A.1 was used. The methodology is explained through concise descriptions, supplemented by a schematic flowchart in Figure 4-A.2 and tables portraying selected sections of the results. It is important to note that the R-W technique is not applied to resample the data (as is the case, for example, with bootstrapping) but rather to evaluate and compare the FFA results of the probability distributions GEV and LP3. It is thus important to maintain the order in which the annual flood peaks occurred in nature, to establish at which R-W record length dry and wet spells will no longer have a noticeable impact on the results.

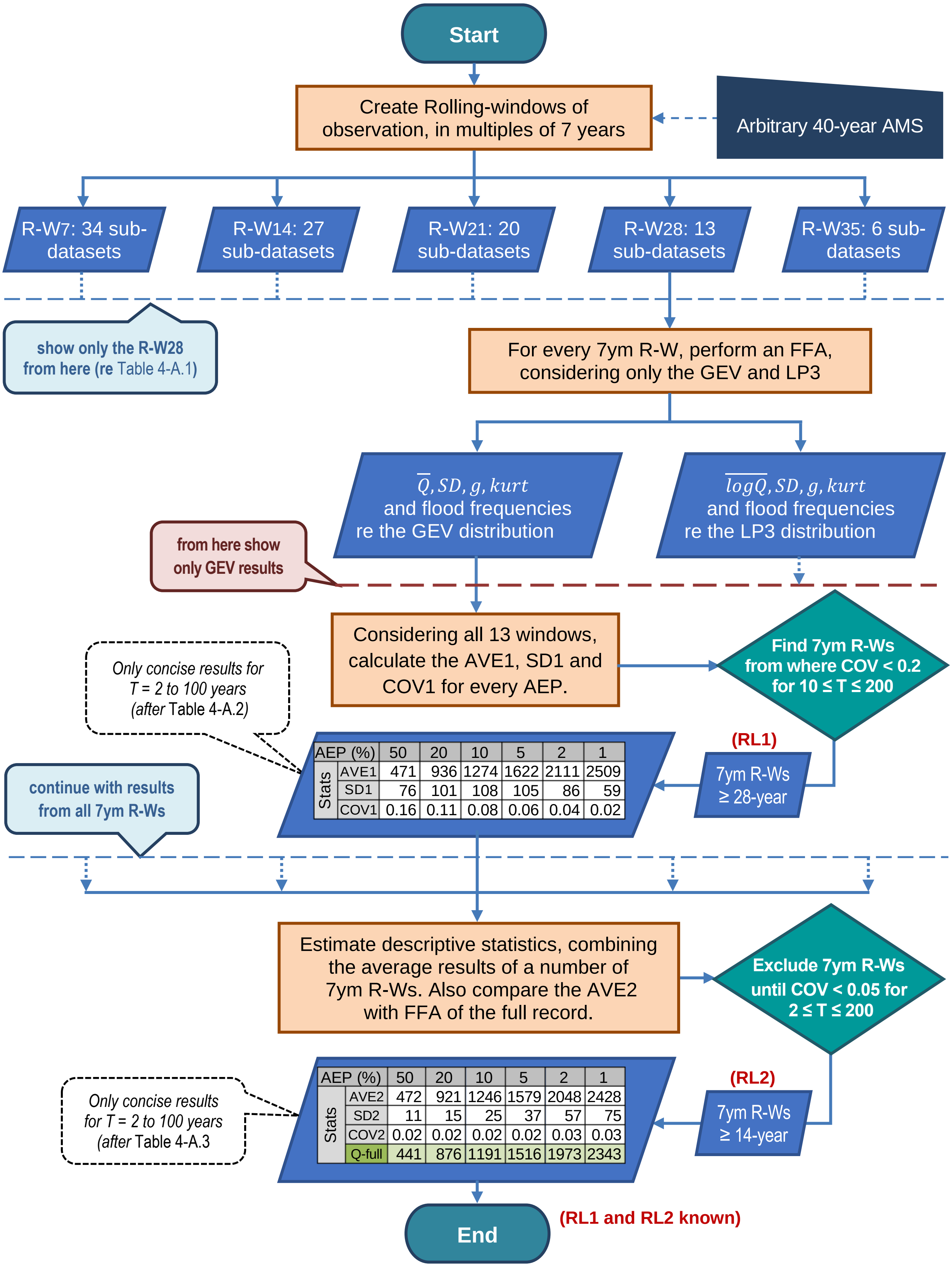
Preparing the data for the R-W analysis

- R-W sub-datasets were extracted, in multiples of 7-year periods (7-, 14-, 21 years and so forth)
- For every R-W sub-dataset, descriptive stats were estimated to perform a statistical analysis.

Table 4-A.1: Arbitrary 40-year record, Illustrating the R-W concept, using R-W28

40-year AMS		28-year rolling windows*												
		1	2	3	4	5	6	7	8	9	10	11	12	13
Hydro Year	AMS flood peaks (m^3/s)													
1978	163	163												
1979	252	252	252											
1980	245	245	245	245										
1981	143	143	143	143	143									
1982	106	106	106	106	106	106								
1983	282	282	282	282	282	282	282							
1984	144	144	144	144	144	144	144	144						
1985	414	414	414	414	414	414	414	414	414					
1986	126	126	126	126	126	126	126	126	126	126				
1987	651	651	651	651	651	651	651	651	651	651	651			
1988	420	420	420	420	420	420	420	420	420	420	420	420		
1989	390	390	390	390	390	390	390	390	390	390	390	390	390	
1990	177	177	177	177	177	177	177	177	177	177	177	177	177	177
1991	121	121	121	121	121	121	121	121	121	121	121	121	121	121
1992	145	145	145	145	145	145	145	145	145	145	145	145	145	145
1993	286	286	286	286	286	286	286	286	286	286	286	286	286	286
1994	300	300	300	300	300	300	300	300	300	300	300	300	300	300
1995	2 135	2 135	2 135	2 135	2 135	2 135	2 135	2 135	2 135	2 135	2 135	2 135	2 135	2 135
1996	665	665	665	665	665	665	665	665	665	665	665	665	665	665
1997	618	618	618	618	618	618	618	618	618	618	618	618	618	618
1998	367	367	367	367	367	367	367	367	367	367	367	367	367	367
1999	1 045	1 045	1 045	1 045	1 045	1 045	1 045	1 045	1 045	1 045	1 045	1 045	1 045	1 045
2000	568	568	568	568	568	568	568	568	568	568	568	568	568	568
2001	84	84	84	84	84	84	84	84	84	84	84	84	84	84
2002	212	212	212	212	212	212	212	212	212	212	212	212	212	212
2003	242	242	242	242	242	242	242	242	242	242	242	242	242	242
2004	365	365	365	365	365	365	365	365	365	365	365	365	365	365
2005	2 275	2 275	2 275	2 275	2 275	2 275	2 275	2 275	2 275	2 275	2 275	2 275	2 275	2 275
2006	635		635	635	635	635	635	635	635	635	635	635	635	635
2007	821			821	821	821	821	821	821	821	821	821	821	821
2008	1 285				1 285	1 285	1 285	1 285	1 285	1 285	1 285	1 285	1 285	1 285
2009	1 250					1 250	1 250	1 250	1 250	1 250	1 250	1 250	1 250	1 250
2010	825						825	825	825	825	825	825	825	825
2011	215							215	215	215	215	215	215	215
2012	515								515	515	515	515	515	515
2013	1 275									1 275	1 275	1 275	1 275	1 275
2014	309										309	309	309	309
2015	288											288	288	288
2016	1 265												1 265	1 265
2017	374													374
Stats	$\bar{Q}$	462.2	479.0	499.4	536.5	576.0	601.7	599.3	612.6	643.3	649.9	636.9	667.1	666.5
	S	538.5	536.2	538.1	555.5	565.7	559.9	561.4	554.6	566.9	561.8	565.9	576.4	576.7
	g	2.6736	2.6153	2.4795	2.1539	1.9062	1.8452	1.8380	1.8513	1.6232	1.6534	1.6809	1.4756	1.4758

* The R-W28 was chosen as illustration, since 13 windows was the optimum to reasonably fit a one-page table and, although Figure 4-A.2 indicates 5 possible 7ym R-W sub-datasets, Table 4-A.1, Figure 4-A.2 and Table 4-A.2 only show the results of R-W28, for ease of presentation



	AEP (%)	50	20	10	5	2	1
Stats	AVE1	471	936	1274	1622	2111	2509
	SD1	76	101	108	105	86	59
	COV1	0.16	0.11	0.08	0.06	0.04	0.02

	AEP (%)	50	20	10	5	2	1
Stats	AVE2	472	921	1246	1579	2048	2428
	SD2	11	15	25	37	57	75
	COV2	0.02	0.02	0.02	0.02	0.03	0.03
	Q-full	441	876	1191	1516	1973	2343

Figure 4-A.2: Flow chart illustrating the record length methodology

Performing FFAs and obtain relevant statistics for each AEP; Table 4-A.2

The GEV and LP3 distributions were used in the statistical analysis for every R-W sub-dataset.

These results were considered differently by two adopted procedures, explained in the subsequent paragraphs, to provide estimates of appropriate record lengths for research. The COV was used as evaluation criteria to estimate appropriate record lengths. Bobbitt (2021) stated that no explicit criteria exist to declare a COV as 'good', since it will be dictated by the situation. FORMPLUS BLOG (s.a.) considers, generally, that a COV between 20 – 30% is acceptable. Hence, the criteria for the two procedures were dictated by the observed results of each procedure.

4-A1.3 First procedure to estimate a minimum record length (RL1)

The variances in the FFA results of the sub-datasets for every 7ym R-W were considered at every AEP – resulting in an average (AVE1), standard deviation (SD1) and coefficient of variation (COV1) at every AEP. COV1 is thus an indication of the expected variability in the FFA results, if any R-W sub-dataset would have been selected, and should comply to the following criteria:

- COV1 ≤ 0.25, for AEPs from 50% to 0.01% (2 ≤ ARI ≤ 10 000 year), on condition that;
- COV1 ≤ 0.2, for AEPs from 10% to 0.5% (10 ≤ ARI ≤ 200 year); i.e., design flood range.

The shortest R-W, to comply to this criteria, was referred to as RL1, and in this illustration the above criteria were met at R-W28, considering the GEV distribution.

Thus RL1 = 28 years, re the GEV distribution (incidentally, RL1 = 35 years re the LP3).

Table 4-A.2: Estimated flood frequencies for R-W28, after the GEV

R-W28	AEP (%)											
	50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
	GEV distribution: Flood peaks (m^3/s)											
1	338	763	1 089	1 441	1 960	2 404	2 897	3 637	4 272	4 981	6 047	6 963
2	356	781	1 106	1 455	1 970	2 408	2 894	3 621	4 243	4 937	5 975	6 865
3	377	808	1 134	1 483	1 993	2 424	2 899	3 605	4 205	4 870	5 858	6 699
4	415	870	1 208	1 562	2 070	2 489	2 945	3 608	4 161	4 763	5 641	6 373
5	457	930	1 273	1 627	2 125	2 530	2 962	3 580	4 086	4 628	5 404	6 040
6	486	955	1 295	1 643	2 131	2 525	2 945	3 541	4 027	4 546	5 285	5 887
7	483	954	1 295	1 644	2 132	2 527	2 947	3 543	4 029	4 546	5 284	5 885
8	498	962	1 299	1 644	2 128	2 519	2 935	3 528	4 011	4 526	5 261	5 860
9	532	1 016	1 357	1 702	2 175	2 549	2 940	3 486	3 922	4 380	5 018	5 529
10	538	1 017	1 356	1 699	2 170	2 544	2 937	3 486	3 926	4 388	5 036	5 555
11	524	1 005	1 346	1 693	2 171	2 551	2 951	3 511	3 961	4 436	5 102	5 637
12	558	1 056	1 401	1 745	2 209	2 571	2 944	3 457	3 861	4 279	4 854	5 308
13	558	1 056	1 401	1 745	2 209	2 571	2 945	3 458	3 862	4 281	4 856	5 310
	Selected Descriptive Statistics											
min	338	763	1 089	1 441	1 960	2 404	2 894	3 457	3 861	4 279	4 854	5 308
max	558	1 056	1 401	1 745	2 209	2 571	2 962	3 637	4 272	4 981	6 047	6 963
AVE1	471	936	1 274	1 622	2 111	2 509	2 934	3 543	4 044	4 582	5 355	5 993
SD1	76	101	108	105	86	59	22	62	141	240	408	567
COV1	0.16	0.11	0.08	0.06	0.04	0.02	0.01	0.02	0.03	0.05	0.08	0.09

* In Table 4-A.2 only the results of the GEV for R-W28 are presented as illustration; corresponding to the concise table embedded in Figure 4-A.2.

4-A1.4 Second procedure to estimate a minimum record length (RL2)

The variances in the averaged FFA results (AVE1) of all the 7ym R-Ws were considered at every AEP – resulting in an average (AVE2), standard deviation (SD2) and coefficient of variation (COV2) at every AEP. COV2 is thus an indication of the expected variability in the averaged FFA results, if the average of any R-W (7, 14, 21, …) would have been selected:

- The results of the shorter 7ym R-Ws were progressively excluded, starting with R-W7, to determine a second set of average descriptive stats, until the following chosen criteria (used tighter criteria, since only the averages of all the R-Ws were used) were met:
  - COV2 ≤ 0.05, for AEPs from 50% to 0.5% (2 ≤ ARI ≤ 200 year)
  - COV2 ≤ 0.1, for AEPs from 0.2% to 0.01% (500 ≤ ARI ≤ 10 000 year).
- The first 7ym R-W, not to be excluded, was referred to as RL2.

In Table 4-A.3, considering **all** the 7ym R-Ws, it is shown that the above criteria were not met. However, by excluding the results of R-W7, the criteria were met.

Thus RL2 = 14 years, re the GEV distribution (incidentally, RL2 = 21 years, re the LP3)

AVE2 was also compared to the FFA results of the full record; in this case it was noted that the combined R-W procedure differed by less than 5% from the results of the full record FFA, for AEPs equal or less than 10%.

Table 4-A.3: Combining the average results all the 7ym R-Ws, after the GEV

R-Ws lengths (years)	AEP (%)											
	50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
	Average flood peaks (GEV: AVE1) per Rolling-window (m^3/s)											
7	500	881	1 135	1 384	1 716	1 975	2 243	2615	2912	3225	3665	4021
14	480	913	1 222	1 538	1 979	2 336	2 715	3257	3701	4178	4865	5432
21	479	932	1 259	1 597	2 073	2 462	2 879	3480	3977	4515	5297	5947
28	471	936	1 274	1 622	2 111	2 509	2 934	3543	4044	4582	5355	5993
35	457	904	1 227	1 560	2 027	2 405	2 808	3383	3852	4354	5072	5658
Descriptive statistics from all 7ym Rolling-windows averages (m^3/s)												
AVE2	478	913	1 224	1 540	1 981	2 337	2 716	3 255	3 697	4 171	4 851	5 410
SD2	16	22	54	93	156	213	277	374	458	551	691	809
COV2	0.033	0.024	0.044	0.060	0.079	0.091	0.102	0.115	0.124	0.132	0.142	0.150
Descriptive statistics, with R-W7 sub-dataset excluded and compared with results from full FFA												
min	457	904	1 222	1 538	1 979	2 336	2 715	3 257	3 701	4 178	4 865	5 432
max	480	936	1 274	1 622	2 111	2 509	2 934	3 543	4 044	4 582	5 355	5 993
AVE2	472	921	1 246	1 579	2 048	2 428	2 834	3 416	3 893	4 407	5 147	5 758
SD2	11	15	25	37	57	75	95	125	151	180	225	263
COV2	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.04	0.04	0.04	0.04	0.05
Full FFA	441	876	1 191	1 516	1 973	2 343	2 738	3 303	3 764	4 258	4 964	5 542
Δ%*	7.0%	5.2%	4.6%	4.2%	3.8%	3.6%	3.5%	3.4%	3.4%	3.5%	3.7%	3.9%

* The difference between the AVE2 and the Full FFA

4-A1.5 Recommended minimum record length

The estimated minimum record lengths, RL1 and RL2, were considered to estimate proposed minimum record lengths from the two procedures, respectively ($RL1_P$ and $RL2_P$). These proposed minimum record lengths were then considered to confirm a recommended minimum record length (RL_R).

The following was considered from the first procedure:

- The result, RL1, was considered as the proposed minimum record length.
- Thus, the proposed minimum record length: $RL1_P$ = RL1 years.

The following was considered from the second procedure:

- The result, RL2, was considered as the lowest 7ym R-W to be considered, in following this procedure to determine a record length.
- However, to get sensible average results it was assumed that at least three 7ym R-Ws should be available (for estimating AVE2), with the largest 7ym R-W having at least four R-W sub-datasets (for estimating AVE1).
- Thus, the proposed minimum record length: $RL2_P$ = RL2 + 17 years (RL2+7+7+3 years).

  To clarify the last two bullets, consider the following example:

  - Assume a dataset with a record length of 42 years and RL2 was found to be 28 years.
  - R-W28, R-W35 and R-W42 are available; thus, the three 7ym R-Ws criterium was met;
  - However, R-W42 only has one sub-dataset, namely 42 years; thus, criterium not met; to meet the criterium the record length should have been 45 years (3 years longer);
    from Equation 4-A.1: Number of sub-datasets = 45 – 42 + 1 = 4, which implies that a sensible average result can also be obtained for the largest 7ym R-W
    Hence $RL2_P$ = 28 + 7 + 7 + 3 = 45 years.

The recommended minimum record length, considering the proposed minimum record lengths from both procedures, was taken as the maximum of $RL1_P$ and $RL2_P$.

thus

$$RL_R = \max(RL1_P, RL2_P) \qquad \textbf{Equation 4-A.2}$$

4-A2 Record length analysis

4-A2.1 Prologue

It was considered impractical to present the substantial results of all the R-Ws, at all 15 primary sites, in tables or figures in this section of Appendix 4-A. Where deemed necessary, supplementary corresponding tables and figures are presented in the follow-up appendices. (Comprehensive results are available, on request, in electronic format).

Appendix 4-B to Appendix 4-K demonstrate the results of this research linked to RL1 at two selected primary sites, D3R002 and N2R001. Appendix 4-L and Appendix 4-M illustrate the results related to RL2.

The two selected primary sites, supplemented by summarised versions of the appendices for convenience, were used to present the results in estimating RL1 and RL2, respectively in Section 4-A2.2 and Section 4-A2.3, for significant R-Ws.

The summarised results of these two sites, in the estimation of RL1 and RL2, is reflected in Section 4-A2.4.

4-A2.2 Extracted results at the two selected primary sites in estimating RL1

As indicated in Section 4-A1.3 the set criteria for RL1 were:

- COV1 ≤ 0.25, for AEPs from 50% to 0.01% (2 ≤ ARI ≤ 10 000 year), on condition that
- COV1 ≤ 0.2, for AEPs from 10% to 0.5% (10 ≤ ARI ≤ 200 year)

The results from flow site D3R002 are provided as follows:

- The FFA results for the sub-datasets for R-W35 (GEV) and R-W42 (LP3), respectively presented in Appendix 4-B and Appendix 4-C, provide the source of the descriptive statistics in Appendices 4-D and 4-E; graphically presented in Appendix 4-F.
- Table 4-A.4 is a condensed version of Appendices 4-B and 4-C, which include the cov for the sub-datasets to illustrate the difference with COV1.
- Table 4-A.5 depicts certain descriptive statistics results of four intentionally selected 7ym R-Ws, from Appendices 4-D, 4-E and 4-F, for ease of reference.
- It is clear from Table 4-A.5 that the set criteria, for RL1 were met for the GEV from R-W35 and for the LP3 from R-W42, which are visually depicted in Figure 4-A.3.
- Therefore, RL1 = 35 years for the GEV and RL1 = 42 years for the LP3.

Table 4-A.4: Condensed FFA results for flow site D3R002

R-W35 results, considering GEV distribution																
Sub set	**FFA moments**				**Annual exceedance probability (%)**											
	ave	**sdev**	**cov**	**skew**	**50**	**20**	**10**	**5**	**2**	**1**	**0.5**	**0.2**	**0.1**	**0.05**	**0.02**	**0.01**
	Flows in m³/s				**Flood peaks in m³/s**											
1	1 964	2 104	1.071	2.641	1 481	3 143	4 419	5 792	7 815	9 538	11 454	14 323	16 783	19 525	23 639	27 169
2	1 972	2 101	1.065	2.639	1 489	3 150	4 424	5 795	7 817	9 537	11 450	14 314	16 770	19 507	23 613	27 136
3	1 985	2 095	1.055	2.648	1 503	3 158	4 428	5 796	7 812	9 530	11 440	14 302	16 757	19 494	23 601	27 127
4	1 988	2 094	1.053	2.649	1 506	3 160	4 429	5 796	7 812	9 528	11 438	14 299	16 753	19 489	23 596	27 122
5	2 016	2 087	1.035	2.637	1 537	3 186	4 451	5 813	7 819	9 528	11 426	14 269	16 707	19 423	23 498	26 993
6	2 078	2 081	1.001	2.571	1 603	3 256	4 519	5 874	7 864	9 552	11 423	14 216	16 602	19 254	23 218	26 607
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
75	2 804	2 076	0.740	2.250	2 345	4 033	5 295	6 627	8 547	10 146	11 890	14 446	16 590	18 935	22 378	25 268
76	2 736	2 054	0.751	2.408	2 274	3 924	5 171	6 499	8 432	10 058	11 845	14 490	16 730	19 200	22 860	25 962
77	2 694	2 069	0.768	2.406	2 229	3 892	5 149	6 487	8 434	10 071	11 871	14 534	16 789	19 276	22 960	26 082
78	2 577	2 063	0.801	2.547	2 107	3 748	5 000	6 342	8 310	9 978	11 825	14 577	16 926	19 534	23 428	26 753
79	2 601	2 059	0.792	2.527	2 133	3 773	5 023	6 361	8 322	9 981	11 817	14 550	16 880	19 465	23 319	26 607
80	2 608	2 058	0.789	2.521	2 140	3 780	5 029	6 366	8 325	9 982	11 814	14 542	16 867	19 444	23 288	26 565
Average descriptive statistics (flood peaks in m³/s)																
min	1 964	1 525	0.776	**min1**	1 481	3 143	4 342	5 209	6 336	7 102	7 792	8 646	9 261	9 853	10 600	11 139
max	3 913	2 579	0.659	**max1**	3 512	5 626	7 046	8 417	10 236	11 736	13 273	15 386	17 375	19 602	23 639	27 169
AVE	2 889	2 117	0.733	**AVE1**	2 495	4 298	5 551	6 804	8 510	9 859	11 271	13 249	14 842	16 530	18 922	20 869
				SD1	599	801	903	992	1 134	1 298	1 545	2 035	2 547	3 189	4 255	5 239
				COV1	0.24	0.19	0.16	0.15	0.13	0.13	0.14	0.15	0.17	0.19	0.22	0.25
R-W42 results, considering LP3 distribution																
Sub set	**FFA moments**				**Annual exceedance probability (%)**											
	ave	**sdev**	**cov**	**skew**	**50**	**20**	**10**	**5**	**2**	**1**	**0.5**	**0.2**	**0.1**	**0.05**	**0.02**	**0.01**
	(log-transformed)				**Flood peaks in m³/s**											
1	3.142	0.382	0.122	-0.295	1 448	2 938	4 152	5 460	7 337	8 867	10 492	12 782	14 620	16 547	19 229	21 358
2	3.144	0.382	0.122	-0.312	1 458	2 953	4 163	5 460	7 309	8 808	10 392	12 612	14 384	16 233	18 794	20 816
3	3.142	0.384	0.122	-0.304	1 452	2 949	4 167	5 476	7 350	8 874	10 489	12 759	14 577	16 478	19 119	21 209
4	3.154	0.387	0.123	-0.373	1 507	3 052	4 282	5 579	7 395	8 840	10 344	12 413	14 036	15 705	17 975	19 738
5	3.149	0.390	0.124	-0.342	1 484	3 035	4 291	5 631	7 533	9 066	10 678	12 922	14 702	16 550	19 091	21 084
6	3.157	0.388	0.123	-0.406	1 526	3 086	4 315	5 599	7 376	8 776	10 218	12 182	13 707	15 261	17 354	18 963
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
68	3.384	0.325	0.096	-0.170	2 474	4 576	6 236	8 005	10 534	12 602	14 809	17 948	20 495	23 195	27 003	30 070
69	3.373	0.318	0.094	-0.136	2 401	4 395	5 972	7 658	10 078	12 066	14 199	17 248	19 736	22 386	26 146	29 192
70	3.380	0.308	0.091	-0.073	2 421	4 373	5 927	7 601	10 029	12 046	14 229	17 388	19 995	22 800	26 827	30 126
71	3.357	0.311	0.093	-0.044	2 285	4 159	5 671	7 314	9 724	11 747	13 954	17 178	19 864	22 775	26 991	30 475
72	3.358	0.311	0.093	-0.060	2 299	4 178	5 688	7 322	9 709	11 702	13 870	17 023	19 638	22 463	26 537	29 891
73	3.345	0.298	0.089	-0.110	2 242	3 955	5 283	6 687	8 684	10 313	12 051	14 525	16 535	18 671	21 693	24 136
Average descriptive statistics (flood peaks in m³/s)																
min	3.142	0.283	0.090	**min1**	1 448	2 938	4 152	5 460	7 002	8 135	9 253	10 706	11 783	12 838	14 199	15 201
max	3.474	0.390	0.112	**max1**	3 166	5 428	6 963	8 528	11 210	13 530	16 039	19 663	22 646	25 967	30 781	34 750
AVE	3.347	0.326	0.097	**AVE1**	2 360	4 275	5 725	7 220	9 286	10 923	12 626	14 984	16 849	18 787	21 460	23 569
				SD1	492	723	861	995	1 200	1 404	1 668	2 126	2 565	3 086	3 907	4 630
				COV1	0.21	0.17	0.15	0.14	0.13	0.13	0.13	0.14	0.15	0.16	0.18	0.20

Table 4-A.5: Selected 7ym R-W results for flow site D3R002

7ym R-W (yrs)	Stats	AEP (%)											
		50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
		GEV distribution: Flood peaks in m³/s											
28	min1	1 355	2 828	3 676	4 545	5 521	6 079	6 604	7 251	7 707	8 093	8 532	8 835
	max1	3 655	5 884	7 423	8 907	10 821	12 313	13 796	15 989	18 018	20 799	25 101	29 197
	AVE1	2 476	4 265	5 499	6 730	8 401	9 721	11 102	13 042	14 608	16 271	18 638	20 574
	SD1	667	896	1 035	1 187	1 458	1 754	2 150	2 853	3 535	4 357	5 678	6 870
	COV1	0.27	0.21	0.19	0.18	0.17	0.18	0.19	0.22	0.24	0.27	0.30	0.33
35	min1	1 481	3 143	4 342	5 209	6 336	7 102	7 792	8 646	9 261	9 853	10 600	11 139
	max1	3 512	5 626	7 046	8 417	10 236	11 736	13 273	15 386	17 375	19 602	23 639	27 169
	AVE1	2 495	4 298	5 551	6 804	8 510	9 859	11 271	13 249	14 842	16 530	18 922	20 869
	SD1	599	801	903	992	1 134	1 298	1 545	2 035	2 547	3 189	4 255	5 239
	COV1	0.24	0.19	0.16	0.15	0.13	0.13	0.14	0.15	0.17	0.19	0.22	0.25
42	min1	1 536	3 082	4 270	5 550	7 185	8 146	9 092	10 326	11 254	12 178	13 393	14 307
	max1	3 264	5 337	6 717	8 060	9 806	11 262	12 839	15 052	16 802	18 809	22 319	25 654
	AVE1	2 524	4 345	5 610	6 873	8 586	9 935	11 339	13 292	14 853	16 494	18 797	20 651
	SD1	525	714	802	865	937	1 010	1 129	1 406	1 732	2 173	2 941	3 672
	COV1	0.21	0.16	0.14	0.13	0.11	0.10	0.10	0.11	0.12	0.13	0.16	0.18
49	min1	1 507	2 950	4 066	5 276	7 070	8 506	9 493	10 787	11 759	12 725	13 994	14 948
	max1	3 042	5 023	6 373	7 695	9 469	10 880	12 443	14 627	16 360	18 166	21 504	24 792
	AVE1	2 541	4 373	5 645	6 911	8 625	9 969	11 363	13 291	14 823	16 422	18 650	20 429
	SD1	445	612	691	745	792	825	877	1 019	1 213	1 498	2 029	2 553
	COV1	0.18	0.14	0.12	0.11	0.09	0.08	0.08	0.08	0.08	0.09	0.11	0.12
		LP3 distribution: Flood peaks in m³/s											
28	min1	1 238	2 705	3 764	4 761	6 116	6 839	7 471	8 261	8 827	9 369	10 050	10 542
	max1	3 516	5 750	7 739	9 938	13 078	15 684	18 485	22 491	27 583	33 691	43 372	52 098
	AVE1	2 302	4 163	5 599	7 109	9 248	10 990	12 851	15 510	17 684	20 012	23 344	26 074
	SD1	637	933	1 122	1 335	1 733	2 170	2 757	3 798	4 808	6 033	8 020	9 835
	COV1	0.28	0.22	0.20	0.19	0.19	0.20	0.21	0.24	0.27	0.30	0.34	0.38
35	min1	1 333	2 871	4 231	5 564	6 817	7 641	8 425	9 405	10 107	10 777	11 617	12 220
	max1	3 357	5 588	7 239	8 939	11 725	13 986	16 376	19 995	23 091	26 788	33 751	39 904
	AVE1	2 330	4 212	5 651	7 151	9 249	10 935	12 714	15 218	17 235	19 364	22 362	24 777
	SD1	566	832	984	1 125	1 347	1 584	1 918	2 547	3 185	3 975	5 271	6 456
	COV1	0.24	0.20	0.17	0.16	0.15	0.14	0.15	0.17	0.18	0.21	0.24	0.26
42	min1	1 448	2 938	4 152	5 460	7 002	8 135	9 253	10 706	11 783	12 838	14 199	15 201
	max1	3 166	5 428	6 963	8 528	11 210	13 530	16 039	19 663	22 646	25 967	30 781	34 750
	AVE1	2 360	4 275	5 725	7 220	9 286	10 923	12 626	14 984	16 849	18 787	21 460	23 569
	SD1	492	723	861	995	1 200	1 404	1 668	2 126	2 565	3 086	3 907	4 630
	COV1	0.21	0.17	0.15	0.14	0.13	0.13	0.13	0.14	0.15	0.16	0.18	0.20
49	min1	1 444	2 847	3 969	5 143	6 748	7 997	9 286	10 512	11 299	12 031	12 920	13 536
	max1	2 921	5 039	6 587	8 176	10 762	13 168	15 835	19 794	23 143	26 819	32 221	36 746
	AVE1	2 375	4 312	5 772	7 268	9 323	10 940	12 613	14 914	16 724	18 593	21 157	23 169
	SD1	412	606	735	873	1 097	1 320	1 601	2 079	2 528	3 056	3 883	4 610
	COV1	0.17	0.14	0.13	0.12	0.12	0.12	0.13	0.14	0.15	0.16	0.18	0.20

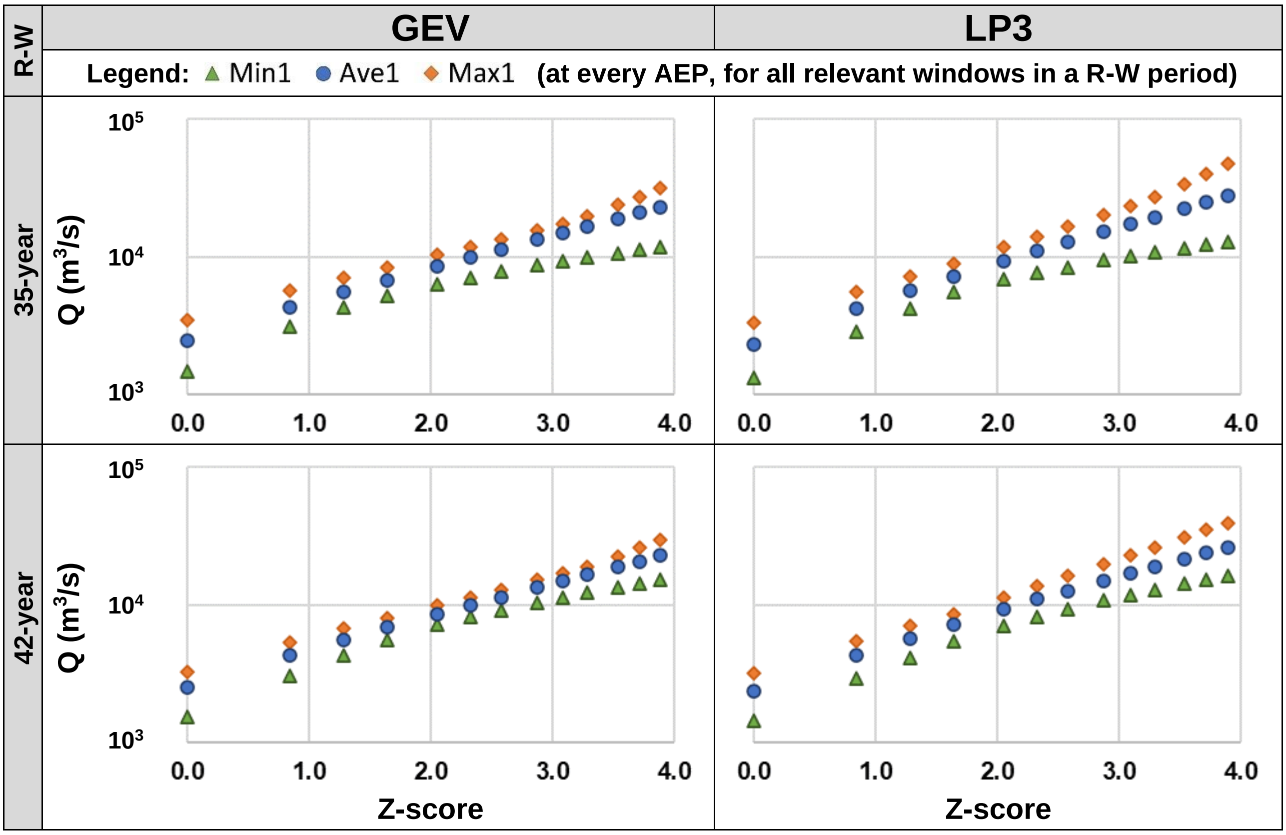


Figure 4-A.3: Results of selected 7ym R-Ws, for flow site D3R002

<u>The results from flow site N2R001 are provided as follows:</u>

- The FFA results for the <u>sub-datasets</u> for R-W42 (GEV) and R-W77 (LP3), respectively presented in Appendix 4-G and Appendix 4-H, provide the source of the descriptive statistics in Appendices 4-I and 4-J; and graphically presented in Appendix 4-K.
- Table 4-A.6 is a condensed version of Appendices 4-G and 4-H, which include the cov for the sub-datasets to illustrate the difference with COV1.
- Table 4-A.7 depicts certain descriptive statistics results of six intentionally selected 7ym R-Ws, from Appendices 4-I, 4-J and 4-K, for ease of reference.
- It is clear from Table 4-A.7 that the set criteria for RL1 were met for the GEV from R-W42 and for the LP3 from R-W77, which are visually depicted in Figure 4-A.4
- Therefore, RL1 = 42 years for the GEV and RL1 = 77 years for the LP3.

Table 4-A.6: Condensed FFA results for flow site N2R001

R-W42 results, considering GEV distribution																
Sub set	**FFA moments**				**Annual exceedance probability (%)**											
	ave	**sdev**	**cov**	**skew**	**50**	**20**	**10**	**5**	**2**	**1**	**0.5**	**0.2**	**0.1**	**0.05**	**0.02**	**0.01**
	Flow in m³/s				**Flood peaks in m³/s**											
1	542	822	1.517	2.445	357	1 015	1 515	2 047	2 824	3 478	4 199	5 268	6 175	7 177	8 666	9 930
2	541	823	1.521	2.442	355	1 015	1 514	2 047	2 824	3 479	4 200	5 269	6 176	7 178	8 666	9 929
3	537	825	1.536	2.437	350	1 012	1 513	2 047	2 826	3 481	4 204	5 274	6 182	7 184	8 672	9 936
4	537	825	1.536	2.437	351	1 012	1 513	2 047	2 826	3 481	4 203	5 273	6 181	7 183	8 671	9 934
5	538	824	1.532	2.439	353	1 013	1 513	2 047	2 825	3 480	4 202	5 272	6 179	7 182	8 669	9 933
6	537	825	1.536	2.436	351	1 012	1 513	2 047	2 826	3 481	4 203	5 273	6 181	7 183	8 670	9 933
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
50	365	473	1.296	2.538	257	633	921	1 229	1 680	2 062	2 485	3 115	3 652	4 249	5 139	5 898
51	375	471	1.256	2.525	268	642	928	1 234	1 682	2 061	2 480	3 105	3 637	4 227	5 108	5 859
52	376	470	1.250	2.530	269	643	928	1 234	1 681	2 060	2 480	3 104	3 636	4 227	5 108	5 860
53	330	394	1.194	2.811	238	546	784	1 043	1 427	1 756	2 125	2 682	3 163	3 704	4 521	5 227
54	332	392	1.181	2.830	240	547	784	1 042	1 425	1 754	2 123	2 680	3 162	3 704	4 523	5 233
55	328	395	1.204	2.806	236	545	783	1 042	1 427	1 757	2 126	2 684	3 165	3 706	4 523	5 230
Average descriptive statistics																
min	328	392	1.195	**min1**	236	545	783	1 042	1 425	1 754	2 123	2 680	3 162	3 704	4 521	5 227
max	618	1 058	1.712	**max1**	384	1 179	1 808	2 502	3 538	4 431	5 440	6 975	8 310	9 815	12 105	14 097
AVE	513	873	1.702	**AVE1**	306	973	1 497	2 073	2 942	3 700	4 559	5 877	7 034	8 351	10 377	12 158
				SD1	38	157	257	367	537	688	864	1 143	1 397	1 697	2 177	2 619
				COV1	0.12	0.16	0.17	0.18	0.18	0.19	0.19	0.19	0.20	0.20	0.21	0.22
R-W77 results, considering LP3 distribution																
Sub set	**FFA moments**				**Annual exceedance probability (%)**											
	ave	**sdev**	**cov**	**skew**	**50**	**20**	**10**	**5**	**2**	**1**	**0.5**	**0.2**	**0.1**	**0.05**	**0.02**	**0.01**
	(log-transformed)				**Flood peaks in m³/s**											
1	2.367	0.538	0.227	0.332	218	645	1 184	1 997	3 680	5 609	8 330	13 623	19 395	27 246	41 969	57 536
2	2.369	0.538	0.227	0.324	219	647	1 185	1 995	3 665	5 573	8 255	13 455	19 107	26 772	41 097	56 194
3	2.372	0.538	0.227	0.308	221	653	1 194	2 004	3 666	5 554	8 197	13 292	18 800	26 235	40 050	54 529
4	2.373	0.538	0.227	0.300	222	656	1 197	2 006	3 660	5 533	8 148	13 172	18 586	25 874	39 371	53 471
5	2.379	0.537	0.226	0.273	226	664	1 205	2 003	3 618	5 426	7 925	12 672	17 730	24 472	36 816	49 570
6	2.381	0.536	0.225	0.267	227	666	1 206	2 001	3 605	5 396	7 866	12 543	17 513	24 123	36 190	48 625
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
15	2.329	0.508	0.218	0.331	200	557	988	1 618	2 880	4 285	6 220	9 889	13 795	19 005	28 555	38 442
16	2.335	0.509	0.218	0.297	204	568	1 002	1 631	2 878	4 250	6 124	9 633	13 329	18 209	27 051	36 101
17	2.328	0.508	0.218	0.337	199	556	988	1 620	2 891	4 310	6 270	9 996	13 976	19 295	29 074	39 226
18	2.320	0.509	0.219	0.374	194	546	977	1 617	2 920	4 396	6 458	10 436	14 743	20 570	31 437	42 876
19	2.317	0.510	0.220	0.389	192	542	974	1 618	2 938	4 441	6 555	10 656	15 123	21 200	32 604	44 681
20	2.309	0.517	0.224	0.379	189	540	976	1 629	2 974	4 511	6 676	10 889	15 487	21 752	33 529	46 018
Average descriptive statistics																
min	2.309	0.506	0.219	**min1**	189	540	974	1 617	2 818	4 106	5 832	9 003	12 278	16 529	24 081	31 664
max	2.381	0.538	0.226	**max1**	227	666	1 206	2 006	3 680	5 609	8 330	13 623	19 395	27 246	41 969	57 536
AVE	2.351	0.520	0.221	**AVE1**	212	604	1 081	1 783	3 194	4 770	6 946	11 077	15 483	21 363	32 154	43 334
				SD1	12	46	94	172	346	561	881	1 543	2 306	3 394	5 537	7 908
				COV1	0.06	0.08	0.09	0.10	0.11	0.12	0.13	0.14	0.15	0.16	0.17	0.18

Table 4-A.7: Selected 7ym R-W results for flow site N2R001

7ym R-W (yrs)	Stats	AEP (%)											
		50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
		GEV distribution: Flood peaks in m³/s											
35	min1	207	375	489	600	746	858	971	1 123	1 240	1 358	1 518	1 641
	max1	423	1 167	1 807	2 503	3 542	4 438	5 443	6 967	8 288	9 828	12 228	14 338
	AVE1	312	947	1 441	1 979	2 784	3 479	4 262	5 451	6 487	7 657	9 441	10 996
	SD1	45	230	386	561	832	1 074	1 353	1 793	2 188	2 646	3 367	4 016
	COV1	0.14	0.24	0.27	0.28	0.30	0.31	0.32	0.33	0.34	0.35	0.36	0.37
42	min1	236	545	783	1 042	1 425	1 754	2 123	2 680	3 162	3 704	4 521	5 227
	max1	384	1 179	1 808	2 502	3 538	4 431	5 440	6 975	8 310	9 815	12 105	14 097
	AVE1	306	973	1 497	2 073	2 942	3 700	4 559	5 877	7 034	8 351	10 377	12 158
	SD1	38	157	257	367	537	688	864	1 143	1 397	1 697	2 177	2 619
	COV1	0.12	0.16	0.17	0.18	0.18	0.19	0.19	0.19	0.20	0.20	0.21	0.22
49	min1	233	816	1 293	1 834	2 679	3 441	4 330	5 739	7 016	8 511	10 887	12 733
	max1	377	1 192	1 819	2 496	3 498	4 353	5 307	6 740	7 980	9 415	11 597	13 492
	AVE1	304	992	1 537	2 139	3 053	3 855	4 769	6 180	7 426	8 852	11 057	13 009
	SD1	47	124	174	220	273	305	328	338	326	294	224	191
	COV1	0.15	0.12	0.11	0.10	0.09	0.08	0.07	0.05	0.04	0.03	0.02	0.01
		LP3 distribution: Flood peaks in m³/s											
70	min1	180	506	906	1 501	2 716	4 092	5 883	9 221	12 722	17 326	25 632	34 095
	max1	234	700	1 290	2 183	4 058	6 372	9 785	16 751	24 713	35 997	58 214	82 844
	AVE1	209	606	1 098	1 831	3 327	5 022	7 390	11 953	16 891	23 568	36 013	49 112
	SD1	13	51	112	220	483	829	1 375	2 571	4 019	6 162	10 575	15 651
	COV1	0.06	0.08	0.10	0.12	0.15	0.17	0.19	0.22	0.24	0.26	0.29	0.32
77	min1	189	540	974	1 617	2 818	4 106	5 832	9 003	12 278	16 529	24 081	31 664
	max1	227	666	1 206	2 006	3 680	5 609	8 330	13 623	19 395	27 246	41 969	57 536
	AVE1	212	604	1 081	1 783	3 194	4 770	6 946	11 077	15 483	21 363	32 154	43 334
	SD1	12	46	94	172	346	561	881	1 543	2 306	3 394	5 537	7 908
	COV1	0.06	0.08	0.09	0.10	0.11	0.12	0.13	0.14	0.15	0.16	0.17	0.18
84	min1	195	533	937	1 516	2 642	3 873	5 544	8 659	11 925	16 224	23 990	31 918
	max1	226	638	1 134	1 857	3 331	4 978	7 254	11 582	16 201	22 371	33 700	45 442
	AVE1	214	597	1 060	1 734	3 084	4 583	6 646	10 546	14 691	20 207	30 296	40 717
	SD1	11	40	78	137	262	405	609	1 006	1 442	2 038	3 161	4 356
	COV1	0.05	0.07	0.07	0.08	0.08	0.09	0.09	0.10	0.10	0.10	0.10	0.11

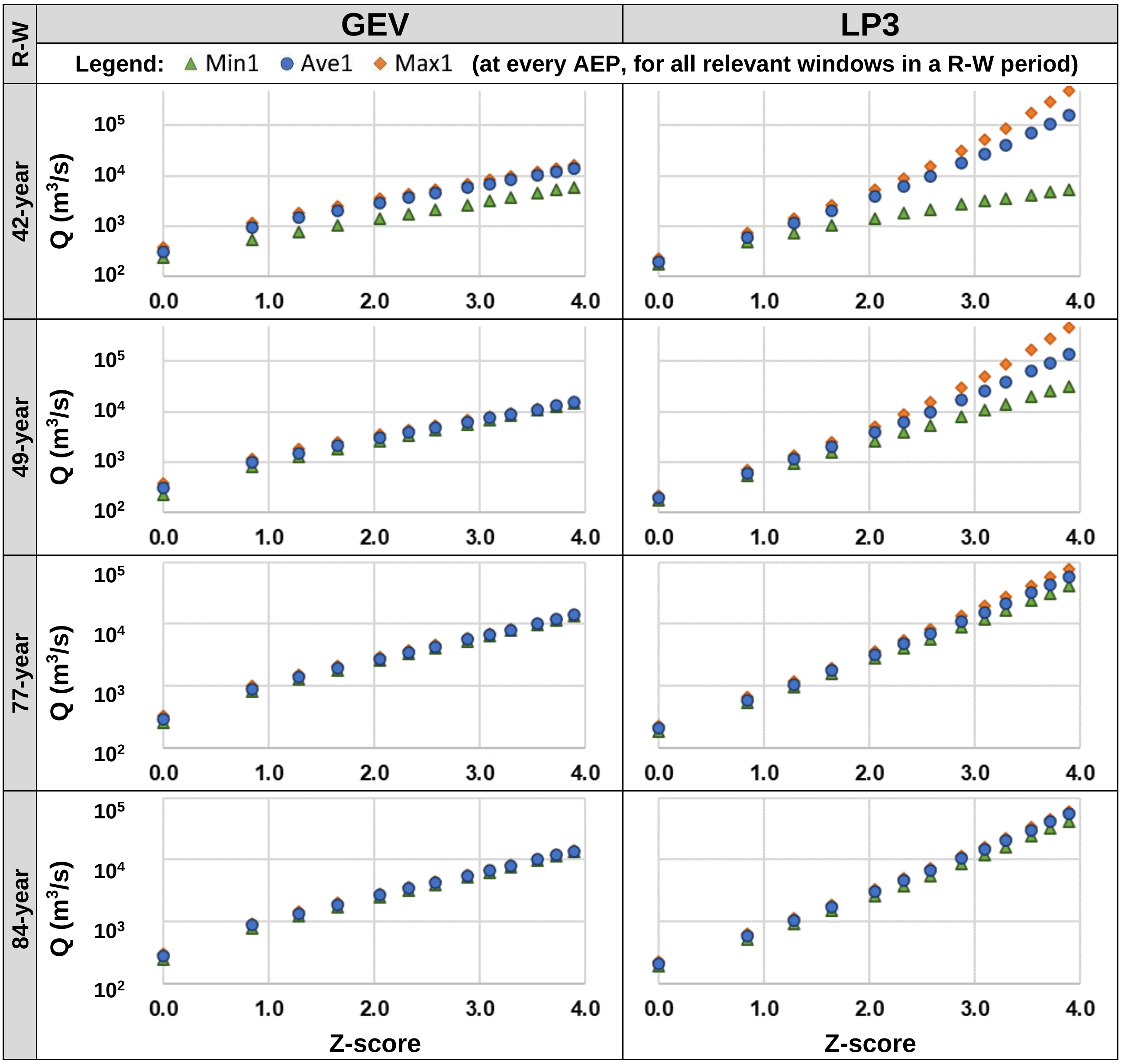


Figure 4-A.4: Results of selected 7ym R-Ws, for flow site N2R001

4-A2.3 Extracted results at the two selected primary sites in estimating RL2

As indicated in Section 4-A1.4 the set criteria for RL2 were:

- COV2 ≤ 0.05, for AEPs from 50% to 0.5% (2 ≤ ARI ≤ 200 year)
- COV2 ≤ 0.1, for AEPs from 0.2% to 0.01% (500 ≤ ARI ≤ 10 000 year)

<u>**The results from flow site D3R002 are provided as follows:**</u>

- The average results of all 7ym R-Ws are depicted in Table 4-A.8 (GEV) and Table 4-A.9 (LP3), and visually displayed in Figure 4-A.5.
- If sub-datasets, shorter than R-W14 for the GEV and shorter than R-W21 for the LP3, are excluded from the dataset, the set COV criteria are met for RL2.
- Therefore, RL2 = 14 years for the GEV and RL2 = 21 years for the LP3.

Table 4-A.8: Combined results of 7ym R-W averages, for the GEV, at D3R002

7ym R-Ws (years)	AEP (%)											
	50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
	Average flood peaks (GEV: AVE1) per R-W (m^3/s)											
7	2 442	4 007	4 999	5 931	7 123	8 014	8 908	10 104	11 028	11 974	13 267	14 283
14	2 430	4 108	5 222	6 303	7 731	8 832	9 962	11 516	12 746	14 032	15 831	17 277
21	2 454	4 209	5 404	6 583	8 172	9 419	10 719	12 536	13 999	15 551	17 757	19 560
28	2 476	4 265	5 499	6 730	8 401	9 721	11 102	13 042	14 608	16 271	18 638	20 574
35	2 495	4 298	5 551	6 804	8 510	9 859	11 271	13 249	14 842	16 530	18 922	20 869
42	2 524	4 345	5 610	6 873	8 586	9 935	11 339	13 292	14 853	16 494	18 797	20 651
49	2 541	4 373	5 645	6 911	8 625	9 969	11 363	13 291	14 823	16 422	18 650	20 429
56	2 554	4 397	5 675	6 947	8 663	10 005	11 390	13 299	14 806	16 370	18 532	20 244
63	2 544	4 406	5 698	6 985	8 723	10 081	11 483	13 412	14 932	16 508	18 680	20 395
70	2 518	4 392	5 697	6 998	8 757	10 134	11 557	13 517	15 064	16 668	18 881	20 628
77	2 483	4 357	5 668	6 980	8 762	10 163	11 616	13 625	15 217	16 875	19 170	20 990
84	2 448	4 336	5 665	7 001	8 825	10 266	11 769	13 857	15 520	17 259	19 680	21 610
91	2 396	4 269	5 595	6 933	8 768	10 225	11 748	13 876	15 577	17 362	19 859	21 857
98	2 340	4 189	5 501	6 829	8 657	10 111	11 637	13 774	15 488	17 291	19 821	21 851
105	2 277	4 093	5 385	6 698	8 509	9 954	11 474	13 608	15 325	17 136	19 683	21 733
112	2 202	3 982	5 257	6 557	8 359	9 805	11 331	13 485	15 225	17 068	19 674	21 780
Descriptive statistics from all 7ym Rolling-windows averages (m^3/s)												
AVE2	2 461	4 270	5 521	6 767	8 454	9 779	11 156	13 067	14 589	16 183	18 411	20 197
SD2	78	124	199	299	464	610	773	1 013	1 213	1 428	1 737	1 990
COV2	0.03	0.03	0.04	0.04	0.05	0.06	0.07	0.08	0.08	0.09	0.09	0.10
Descriptive statistics with sub-datasets shorter than R-W14 excluded												
min2	2 277	4 093	5 222	6 303	7 731	8 832	9 962	11 516	12 746	14 032	15 831	17 277
max2	2 554	4 406	5 698	7 001	8 825	10 266	11 769	13 876	15 577	17 362	19 859	21 857
AVE2	2 463	4 288	5 558	6 827	8 549	9 905	11 316	13 278	14 843	16 483	18 779	20 619
SD2	81	104	142	197	292	379	476	617	734	859	1 034	1 176
COV2	0.03	0.02	0.03	0.03	0.03	0.04	0.04	0.05	0.05	0.05	0.06	0.06
Full FFA	2 188	3 954	5 221	6 516	8 313	9 758	11 285	13 444	15 191	17 045	19 669	21 795
Δ%	12.6%	8.5%	6.5%	4.8%	2.8%	1.5%	0.3%	1.2%	2.3%	3.3%	4.5%	5.4%

Table 4-A.9: Combined results of 7ym R-W averages, for the LP3, at D3R002

7ym R-Ws (years)	AEP (%)											
	50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
	Average flood peaks (LP3: AVE1) per Rolling-window (m^3/s)											
7	2 203	3 923	5 322	6 896	9 369	11 645	14 391	18 992	23 464	29 092	38 969	48 983
14	2 229	4 002	5 390	6 886	9 088	10 968	13 073	16 273	19 072	22 261	27 194	31 572
21	2 274	4 100	5 511	7 005	9 150	10 928	12 862	15 693	18 070	20 679	24 533	27 796
28	2 302	4 163	5 599	7 109	9 248	10 990	12 851	15 510	17 684	20 012	23 344	26 074
35	2 330	4 212	5 651	7 151	9 249	10 935	12 714	15 218	17 235	19 364	22 362	24 777
42	2 360	4 275	5 725	7 220	9 286	10 923	12 626	14 984	16 849	18 787	21 460	23 569
49	2 375	4 312	5 772	7 268	9 323	10 940	12 613	14 914	16 724	18 593	21 157	23 169
56	2 387	4 345	5 812	7 310	9 353	10 951	12 598	14 848	16 608	18 417	20 885	22 810
63	2 372	4 348	5 835	7 354	9 427	11 048	12 714	14 989	16 762	18 581	21 055	22 978
70	2 338	4 322	5 824	7 366	9 479	11 138	12 851	15 198	17 036	18 929	21 514	23 533
77	2 308	4 288	5 786	7 322	9 425	11 072	12 771	15 094	16 910	18 777	21 325	23 311
84	2 274	4 265	5 777	7 328	9 452	11 116	12 831	15 174	17 005	18 886	21 450	23 447
91	2 229	4 203	5 699	7 231	9 319	10 950	12 626	14 909	16 687	18 508	20 985	22 909
98	2 192	4 141	5 593	7 057	9 019	10 525	12 047	14 085	15 644	17 218	19 319	20 924
105	2 151	4 067	5 465	6 846	8 653	10 004	11 339	13 076	14 367	15 637	17 281	18 498
112	2 064	3 922	5 317	6 729	8 628	10 086	11 561	13 529	15 030	16 537	18 536	20 052
Descriptive statistics from all 7ym Rolling-windows averages (m^3/s)												
AVE2	2 288	4 198	5 651	7 157	9 256	10 942	12 727	15 264	17 341	19 583	22 855	25 623
SD2	74	130	165	181	213	343	621	1 252	1 985	3 011	4 978	7 104
COV2	0.03	0.03	0.03	0.03	0.02	0.03	0.05	0.08	0.11	0.15	0.22	0.28
Descriptive statistics with sub-datasets shorter than R-W21 excluded												
min2	2 151	4 067	5 465	6 846	8 653	10 004	11 339	13 076	14 367	15 637	17 281	18 498
max2	2 387	4 348	5 835	7 366	9 479	11 138	12 862	15 693	18 070	20 679	24 533	27 796
AVE2	2 299	4 234	5 696	7 197	9 260	10 886	12 573	14 899	16 737	18 645	21 282	23 369
SD2	74	94	123	157	223	305	427	665	911	1 220	1 739	2 225
COV2	0.03	0.02	0.02	0.02	0.02	0.03	0.03	0.04	0.05	0.07	0.08	0.10
Full FFA	2 054	3 894	5 275	6 674	8 558	10 007	11 474	13 434	14 931	16 436	18 435	19 951
Δ%	11.9%	8.7%	8.0%	7.8%	8.2%	8.8%	9.6%	10.9%	12.1%	13.4%	15.4%	17.1%

The results from flow site N2R001 are provided as follows:

- The average results of all 7ym R-Ws are depicted in Table 4-A.10 (GEV) and Table 4-A.11 (LP3), and visually displayed in Figure 4-A.5.
- If sub-datasets, shorter than R-W28 for the GEV and shorter than R-W70 for the LP3, are excluded from the dataset, the set COV criteria are met for RL2.
- Therefore, RL2 = 28 years for the GEV and RL2 = 70 years for the LP3.

Table 4-A.10: Combined results of 7ym R-W averages, for the GEV, at N2R001

7ym R-Ws (years)	AEP (%)											
	50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
	Average flood peaks (GEV: AVE1) per Rolling-window (m^3/s)											
7	353	890	1 277	1 677	2 240	2 700	3 195	3 911	4 505	5 149	6 088	6 872
14	332	898	1 319	1 763	2 404	2 940	3 526	4 392	5 123	5 930	7 126	8 142
21	325	911	1 354	1 826	2 516	3 097	3 739	4 695	5 510	6 414	7 765	8 919
28	318	924	1 390	1 893	2 635	3 269	3 976	5 040	5 958	6 986	8 538	9 879
35	312	947	1 441	1 979	2 784	3 479	4 262	5 451	6 487	7 657	9 441	10 996
42	306	973	1 497	2 073	2 942	3 700	4 559	5 877	7 034	8 351	10 377	12 158
49	304	992	1 537	2 139	3 053	3 855	4 769	6 180	7 426	8 852	11 057	13 009
56	302	966	1 493	2 075	2 958	3 733	4 617	5 982	7 187	8 567	10 702	12 591
63	298	941	1 452	2 017	2 877	3 632	4 493	5 825	7 003	8 352	10 441	12 292
70	294	919	1 416	1 968	2 807	3 546	4 390	5 696	6 853	8 179	10 236	12 059
77	292	908	1 399	1 943	2 771	3 500	4 334	5 624	6 766	8 076	10 107	11 908
84	289	899	1 385	1 924	2 746	3 470	4 298	5 580	6 716	8 019	10 040	11 833
91	289	902	1 389	1 929	2 751	3 473	4 297	5 571	6 697	7 987	9 983	11 749
	Descriptive statistics from all 7ym Rolling-windows averages (m^3/s)											
AVE2	309	928	1 412	1 939	2 730	3 415	4 189	5 371	6 405	7 578	9 377	10 954
SD2	19	33	73	129	230	331	458	671	875	1 122	1 527	1 905
COV2	0.06	0.04	0.05	0.07	0.08	0.10	0.11	0.12	0.14	0.15	0.16	0.17
	Descriptive statistics with sub-datasets shorter than R-W28 excluded											
min2	289	899	1 385	1 893	2 635	3 269	3 976	5 040	5 958	6 986	8 538	9 879
max2	318	992	1 537	2 139	3 053	3 855	4 769	6 180	7 426	8 852	11 057	13 009
AVE2	300	937	1 440	1 994	2 833	3 566	4 399	5 682	6 813	8 102	10 092	11 847
SD2	10	32	54	80	124	167	222	314	403	514	698	873
COV2	0.03	0.03	0.04	0.04	0.04	0.05	0.05	0.06	0.06	0.06	0.07	0.07
Full FFA	275	871	1 347	1 876	2 683	3 394	4 208	5 471	6 591	7 877	9 874	11 647
Δ%	9.3%	7.6%	6.9%	6.3%	5.6%	5.1%	4.5%	3.9%	3.4%	2.9%	2.2%	1.7%

Table 4-A.11: Combined results of 7ym R-W averages, for the LP3, at N2R001

7ym R-Ws (years)	AEP (%)											
	50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
	Average flood peaks (LP3: AVE1) per Rolling-window (m^3/s)											
7	224	612	1 180	2 252	5 419	10 862	22 451	61 470	136 277	310 438	957 048	2 300 094
14	218	589	1 095	1 959	4 151	7 324	13 005	28 205	51 357	94 700	216 806	411 403
21	213	590	1 090	1 920	3 924	6 666	11 286	22 671	38 603	66 142	136 404	238 271
28	208	592	1 101	1 928	3 857	6 380	10 430	19 734	31 763	50 926	94 657	150 954
35	204	598	1 121	1 962	3 877	6 316	10 131	18 610	29 216	45 613	81 677	126 470
42	200	606	1 143	1 998	3 904	6 270	9 883	17 670	27 125	41 357	71 687	108 237
49	200	614	1 161	2 024	3 914	6 213	9 646	16 838	25 321	37 757	63 429	93 444
56	203	616	1 147	1 964	3 694	5 721	8 636	14 447	20 944	29 983	47 410	66 359
63	206	610	1 121	1 896	3 510	5 371	8 014	13 201	18 913	26 756	41 639	57 575
70	209	606	1 098	1 831	3 327	5 022	7 390	11 953	16 891	23 568	36 013	49 112
77	212	604	1 081	1 783	3 194	4 770	6 946	11 077	15 483	21 363	32 154	43 334
84	214	597	1 060	1 734	3 084	4 583	6 646	10 546	14 691	20 207	30 296	40 717
91	212	590	1 052	1 734	3 119	4 682	6 861	11 053	15 581	21 696	33 076	45 030
Descriptive statistics from all 7ym Rolling-windows averages (m^3/s)												
AVE2	209	602	1 112	1 922	3 767	6 168	10 102	19 806	34 013	60 808	141 715	287 000
SD2	7	10	39	138	608	1 646	4 173	13 546	32 506	77 975	250 597	613 771
COV2	0.03	0.02	0.03	0.07	0.16	0.27	0.41	0.68	0.96	1.28	1.77	2.14
Descriptive statistics with sub-datasets shorter than R-W70 excluded												
min2	209	590	1 052	1 734	3 084	4 583	6 646	10 546	14 691	20 207	30 296	40 717
max2	214	606	1 098	1 831	3 327	5 022	7 390	11 953	16 891	23 568	36 013	49 112
AVE2	212	599	1 073	1 770	3 181	4 764	6 961	11 157	15 661	21 709	32 885	44 548
SD2	2	7	21	46	108	188	313	584	911	1 394	2 385	3 522
COV2	0.01	0.01	0.02	0.03	0.03	0.04	0.04	0.05	0.06	0.06	0.07	0.08
Full FFA	203	563	1 004	1 659	3 002	4 530	6 678	10 850	15 401	21 599	33 255	45 626
Δ%	4.1%	6.5%	6.9%	6.7%	6.0%	5.2%	4.2%	2.8%	1.7%	0.5%	1.1%	2.4%

Note: Analogous results, for all primary flow sites, can be found in Appendix 4-L, with the resulting RL2 values summarised in Table 4-A.12 (see Section 4-A.2.4).

The two selected sites D3R002 and N2R001, shown in Figure 4-A.5, illustrate the difference in results, that can be expected, between the GEV and LP3 distributions. Flow site N2R001 illustrates a case where the results of the GEV-distribution differed significantly from the results of the LP3-distribution. Flow site D3R002 illustrates a case where the results, relating to record lengths required to perform a reasonable consistent FFA, of the GEV-distribution were very much comparable to the results of the LP3-distribution. The other two sites displayed in Figure 4-A.5 were selected to illustrate the least satisfactory spread of results for the research sites, respectively for the GEV-distribution (J1R003) and for the LP3-distribution (V6H002).

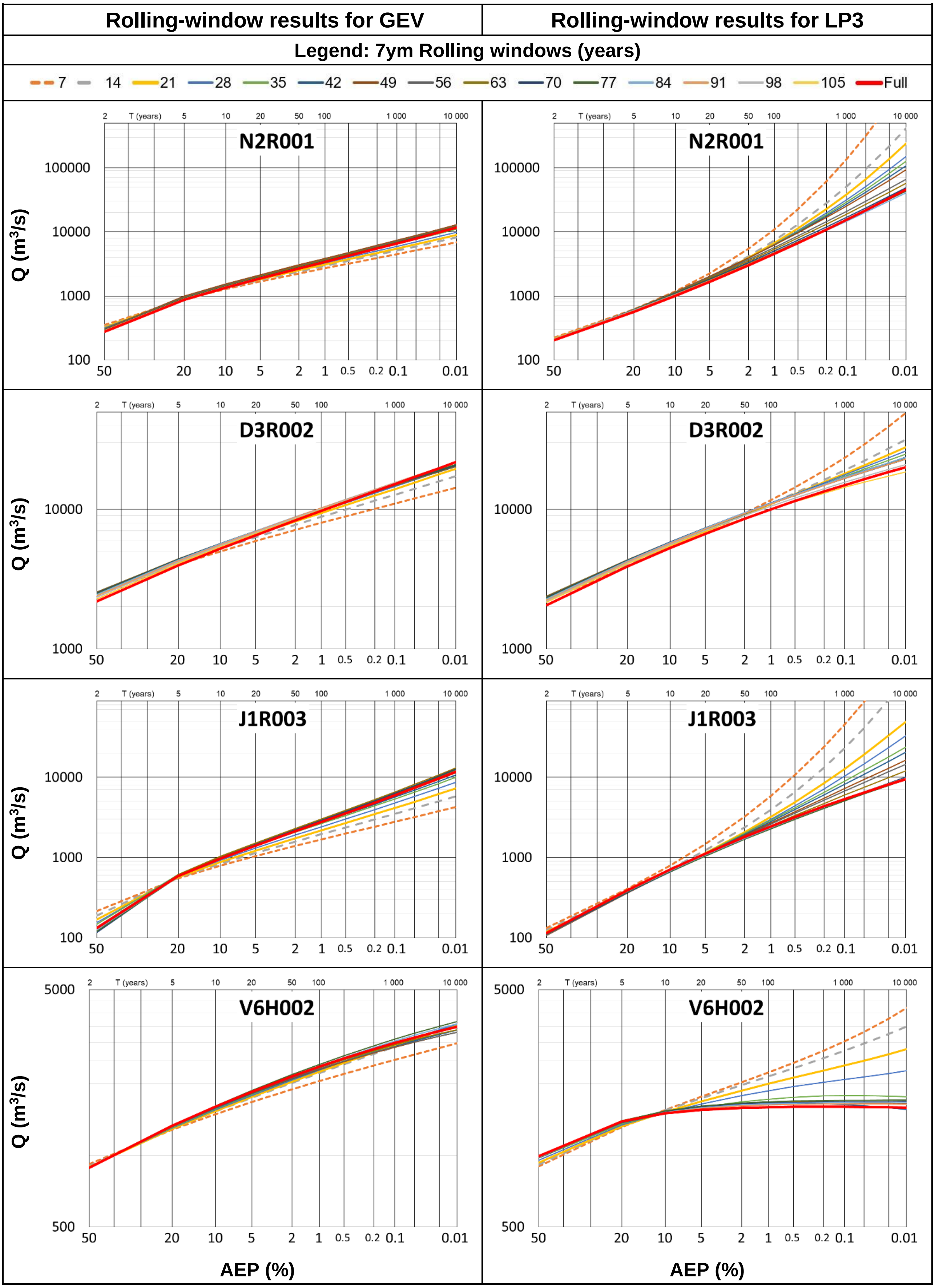


Figure 4-A.5: A visual comparison of all the 7ym R-W average results at selected flow sites.

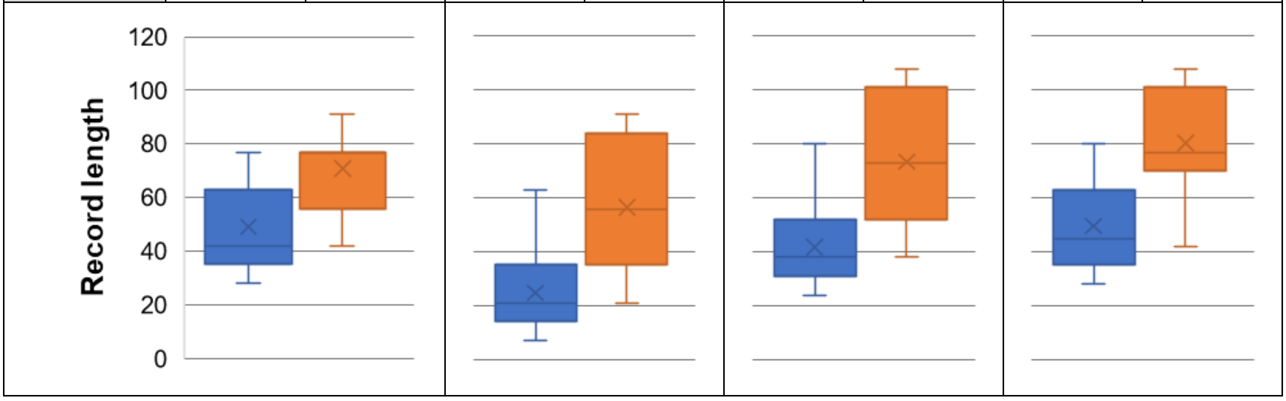
4-A2.4 Summarised Record length results for all primary sites

The minimum record length results of all the primary flow sites, with the illustrated sites (D3R002 and N2R001) highlighted, were assessed and are summarised in Table 4-A.12, as follows:

- $RL1_P$, RL2 and $RL2_P$ are shown for every flow site, for both the GEV and LP3 distributions.
- RL_R is depicted in the last 2 columns, respectively, for the GEV and LP3 distributions.
- Percentiles for all the results are estimated and graphically shown as box plots in the last row of the table.

Table 4-A.12: Recommended record lengths

Flow site	$RL1_P$		RL2		$RL2_P$		RL_R	
	GEV	LP3	GEV	LP3	GEV	LP3	GEV	LP3
A3R002	77	77	63	91	80	108	80	108
B1R001	42	77	21	84	38	101	42	101
B2R001	49	77	14	28	31	45	49	77
C1R002	63	77	14	35	31	52	63	77
C2R001	70	77	42	56	59	73	70	77
C5R002	63	91	35	84	52	101	63	101
C9R002	35	70	14	42	31	59	35	70
D3R002	35	42	14	21	31	38	35	42
J1R003	63	91	35	63	52	80	63	91
J2R001	49	56	21	21	38	38	49	56
N2R001	42	77	28	70	45	87	45	87
Q1R001	42	49	28	56	45	73	45	73
D7H005	35	70	14	70	31	87	35	87
V6H002	28	42	7	35	24	52	28	52
X1H001	42	91	21	91	38	108	42	108
Percentiles and box plots								
min	28	42	7	21	24	38	28	42
25%	35	56	14	35	31	52	35	70
50%	42	77	21	56	38	73	45	77
75%	63	77	35	84	52	101	63	101
max	77	91	63	91	80	108	80	108



with (from Section 4-A1.5):

- $RL1_P$ = RL1 (proposed record length from first research procedure)
- $RL2_P$ = RL2 + 17 (proposed record length from second research procedure)
- RL_R = max($RL1_P$, $RL2_P$) (recommended record length, considering both research procedures)

4-A3 Record length discussion, Recommendations and Conclusion

To retain continuity and avoid unnecessary duplication, the Record Length discussion is presented in the main text of the dissertation as Section 4.3 (p.51).

Recommendations and Conclusions are dealt with in Chapter 7 (p.103-105)

APPENDIX 4-B: Selected R-W35 results at D3R002 re GEV

The R-W35 FFA results, from where the GEV distribution met the set criteria for RL1, at flow site D3R002.

Sub set	FFA moments				Annual exceedance probability (%)											
	ave	sdev	skew	kurt	50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
	Flow in m^3/s				Flood peaks in m^3/s											
1	1 964	2 104	2.641	8.356	1 481	3 143	4 419	5 792	7 815	9 538	11 454	14 323	16 783	19 525	23 639	27 169
2	1 972	2 101	2.639	8.362	1 489	3 150	4 424	5 795	7 817	9 537	11 450	14 314	16 770	19 507	23 613	27 136
3	1 985	2 095	2.648	8.425	1 503	3 158	4 428	5 796	7 812	9 530	11 440	14 302	16 757	19 494	23 601	27 127
4	1 988	2 094	2.649	8.440	1 506	3 160	4 429	5 796	7 812	9 528	11 438	14 299	16 753	19 489	23 596	27 122
5	2 016	2 087	2.637	8.437	1 537	3 186	4 451	5 813	7 819	9 528	11 426	14 269	16 707	19 423	23 498	26 993
6	2 078	2 081	2.571	8.199	1 603	3 256	4 519	5 874	7 864	9 552	11 423	14 216	16 602	19 254	23 218	26 607
7	2 119	2 107	2.443	7.462	1 644	3 332	4 612	5 977	7 967	9 644	11 491	14 229	16 553	19 120	22 931	26 167
8	2 134	2 103	2.435	7.458	1 659	3 346	4 624	5 986	7 971	9 643	11 484	14 211	16 525	19 079	22 871	26 089
9	2 158	2 091	2.452	7.582	1 686	3 360	4 630	5 985	7 962	9 628	11 464	14 188	16 501	19 056	22 853	26 078
10	2 160	2 090	2.455	7.596	1 687	3 361	4 630	5 984	7 960	9 626	11 463	14 187	16 501	19 058	22 857	26 085
11	2 213	2 089	2.381	7.327	1 744	3 426	4 696	6 045	8 006	9 652	11 460	14 130	16 388	18 875	22 554	25 668
12	2 227	2 076	2.423	7.501	1 760	3 426	4 687	6 030	7 987	9 634	11 447	14 131	16 405	18 916	22 639	25 798
13	2 218	2 077	2.432	7.522	1 749	3 415	4 677	6 022	7 982	9 633	11 450	14 143	16 426	18 947	22 688	25 862
14	2 260	2 063	2.431	7.602	1 795	3 449	4 702	6 037	7 983	9 621	11 425	14 097	16 362	18 864	22 575	25 724
15	2 260	2 063	2.431	7.602	1 795	3 449	4 702	6 037	7 983	9 621	11 425	14 097	16 362	18 864	22 575	25 724
16	2 242	2 072	2.417	7.502	1 775	3 439	4 698	6 038	7 991	9 633	11 440	14 114	16 380	18 881	22 588	25 731
17	2 300	2 043	2.462	7.788	1 838	3 472	4 713	6 037	7 970	9 601	11 399	14 067	16 333	18 839	22 565	25 731
18	2 296	2 036	2.478	7.917	1 834	3 462	4 699	6 020	7 950	9 580	11 378	14 049	16 320	18 832	22 571	25 751
19	2 375	2 025	2.420	7.734	1 919	3 544	4 774	6 084	7 992	9 597	11 363	13 978	16 195	18 640	22 266	25 341
20	2 507	2 093	2.130	5.970	2 052	3 769	5 041	6 375	8 280	9 855	11 561	14 041	16 105	18 348	21 614	24 334
21	2 507	2 093	2.130	5.969	2 052	3 769	5 042	6 375	8 281	9 855	11 561	14 041	16 105	18 348	21 613	24 333
22	2 554	2 104	2.035	5.585	2 103	3 842	5 120	6 451	8 340	9 891	11 562	13 973	15 968	18 122	21 237	23 815
23	2 301	1 565	1.139	1.182	2 044	3 427	4 343	5 221	6 357	7 209	8 057	9 176	10 022	10 868	11 985	12 830
24	2 342	1 536	1.167	1.324	2 086	3 441	4 342	5 209	6 336	7 183	8 030	9 152	10 002	10 856	11 988	12 847
25	2 390	1 562	1.051	0.917	2 146	3 534	4 439	5 298	6 396	7 208	8 009	9 053	9 832	10 602	11 608	12 360
26	2 454	1 544	0.987	0.886	2 222	3 599	4 487	5 323	6 381	7 157	7 915	8 895	9 619	10 329	11 248	11 929
27	2 521	1 525	0.923	0.868	2 302	3 666	4 536	5 347	6 364	7 102	7 818	8 732	9 402	10 054	10 889	11 501
28	2 562	1 548	0.832	0.562	2 354	3 744	4 616	5 417	6 407	7 115	7 792	8 646	9 261	9 853	10 600	11 139
29	2 619	1 615	0.786	0.161	2 410	3 863	4 766	5 591	6 601	7 318	7 999	8 852	9 462	10 044	10 774	11 296
30	2 822	1 826	1.016	0.861	2 542	4 167	5 222	6 218	7 484	8 416	9 331	10 517	11 398	12 266	13 393	14 231
31	2 847	1 797	1.079	0.967	2 562	4 156	5 201	6 196	7 472	8 421	9 359	10 586	11 506	12 419	13 616	14 513
32	2 701	1 658	1.182	1.713	2 423	3 884	4 858	5 797	7 019	7 940	8 862	10 086	11 016	11 950	13 191	14 135
33	2 691	1 662	1.188	1.694	2 411	3 876	4 853	5 795	7 024	7 950	8 878	10 110	11 047	11 989	13 242	14 196
34	2 860	1 744	1.105	1.032	2 579	4 123	5 140	6 111	7 362	8 295	9 221	10 437	11 352	12 263	13 461	14 363
35	2 780	1 781	1.118	0.981	2 491	4 067	5 107	6 101	7 385	8 344	9 297	10 551	11 496	12 438	13 679	14 616
36	2 921	1 895	0.984	0.357	2 637	4 326	5 416	6 441	7 737	8 687	9 615	10 813	11 699	12 567	13 690	14 520
37	2 943	1 882	0.978	0.391	2 662	4 341	5 423	6 439	7 724	8 664	9 583	10 767	11 642	12 499	13 606	14 425
38	3 153	2 079	1.025	0.447	2 833	4 682	5 884	7 020	8 466	9 533	10 581	11 942	12 954	13 952	15 249	16 215
39	3 310	2 127	0.883	0.002	3 013	4 919	6 126	7 244	8 637	9 642	10 611	11 841	12 736	13 602	14 705	15 508
40	3 319	2 122	0.880	0.013	3 023	4 924	6 128	7 243	8 631	9 632	10 596	11 821	12 711	13 573	14 668	15 466
41	3 366	2 133	0.804	-0.131	3 087	5 004	6 199	7 294	8 639	9 597	10 510	11 655	12 477	13 263	14 252	14 962

Sub set	FFA moments				Annual exceedance probability (%)											
	ave	sdev	skew	kurt	50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
	Flow in m^3/s				Flood peaks in m^3/s											
42	3 326	2 132	0.863	-0.055	3 032	4 944	6 151	7 265	8 648	9 642	10 598	11 809	12 686	13 534	14 608	15 388
43	3 409	2 131	0.760	-0.192	3 141	5 058	6 244	7 323	8 637	9 566	10 446	11 540	12 321	13 063	13 987	14 647
44	3 418	2 123	0.764	-0.174	3 150	5 060	6 243	7 319	8 631	9 559	10 439	11 534	12 316	13 059	13 986	14 648
45	3 460	2 081	0.813	-0.094	3 185	5 055	6 223	7 294	8 613	9 553	10 450	11 577	12 388	13 164	14 141	14 845
46	3 413	2 107	0.827	-0.145	3 132	5 024	6 210	7 300	8 644	9 605	10 523	11 680	12 514	13 314	14 324	15 052
47	3 439	2 077	0.877	-0.101	3 150	5 012	6 190	7 280	8 636	9 613	10 555	11 749	12 618	13 457	14 524	15 301
48	3 456	2 066	0.879	-0.074	3 168	5 019	6 191	7 276	8 625	9 598	10 536	11 725	12 591	13 427	14 491	15 265
49	3 513	2 060	0.805	-0.146	3 243	5 094	6 249	7 307	8 607	9 533	10 415	11 523	12 318	13 079	14 035	14 723
50	3 788	2 437	1.218	1.569	3 373	5 515	6 951	8 341	10 159	11 536	12 920	14 765	16 174	17 595	19 493	20 942
51	3 913	2 394	1.181	1.552	3 512	5 622	7 028	8 384	10 149	11 479	12 810	14 577	15 919	17 267	19 059	20 422
52	3 878	2 428	1.136	1.423	3 480	5 626	7 046	8 408	10 170	11 490	12 804	14 538	15 847	17 156	18 885	20 192
53	3 805	2 431	1.219	1.565	3 390	5 527	6 959	8 346	10 161	11 535	12 917	14 760	16 167	17 586	19 482	20 931
54	3 856	2 441	1.143	1.374	3 455	5 615	7 045	8 417	10 193	11 523	12 849	14 598	15 920	17 242	18 989	20 310
55	3 715	2 469	1.216	1.524	3 294	5 465	6 919	8 327	10 168	11 562	12 963	14 830	16 256	17 693	19 612	21 078
56	3 695	2 468	1.243	1.572	3 269	5 436	6 893	8 308	10 164	11 575	12 996	14 898	16 354	17 826	19 799	21 311
57	3 598	2 530	1.207	1.412	3 168	5 394	6 884	8 324	10 205	11 628	13 056	14 956	16 406	17 866	19 813	21 298
58	3 640	2 512	1.191	1.452	3 217	5 429	6 905	8 330	10 188	11 590	12 995	14 862	16 282	17 710	19 611	21 058
59	3 637	2 514	1.190	1.443	3 213	5 428	6 905	8 331	10 190	11 593	12 998	14 865	16 285	17 713	19 613	21 060
60	3 613	2 514	1.220	1.493	3 184	5 394	6 875	8 310	10 187	11 609	13 039	14 945	16 401	17 869	19 830	21 330
61	3 558	2 546	1.217	1.401	3 124	5 362	6 862	8 314	10 214	11 652	13 097	15 023	16 494	17 977	19 957	21 470
62	3 520	2 561	1.234	1.381	3 080	5 329	6 840	8 305	10 227	11 685	13 153	15 114	16 615	18 132	20 161	21 715
63	3 449	2 579	1.283	1.415	2 996	5 253	6 781	8 271	10 236	11 736	13 255	15 296	16 868	18 464	20 613	22 268
64	3 410	2 559	1.348	1.644	2 949	5 178	6 700	8 196	10 184	11 714	13 273	15 386	17 024	18 700	20 973	22 738
65	3 221	2 451	1.556	2.646	2 748	4 850	6 324	7 802	9 814	11 398	13 044	15 324	17 134	19 021	21 640	23 719
66	3 215	2 456	1.549	2.620	2 742	4 850	6 327	7 806	9 819	11 402	13 046	15 322	17 127	19 008	21 616	23 686
67	3 187	2 479	1.521	2.515	2 713	4 845	6 333	7 821	9 838	11 420	13 060	15 323	17 113	18 973	21 545	23 581
68	3 219	2 468	1.506	2.527	2 750	4 875	6 355	7 833	9 833	11 400	13 020	15 254	17 018	18 848	21 375	23 371
69	3 143	2 403	1.651	3.285	2 667	4 712	6 162	7 628	9 643	11 243	12 921	15 266	17 145	19 119	21 885	24 102
70	3 167	2 381	1.691	3.416	2 690	4 711	6 149	7 609	9 622	11 228	12 915	15 284	17 189	19 196	22 019	24 290
71	3 044	2 299	1.927	4.649	2 561	4 477	5 872	7 315	9 346	10 999	12 766	15 298	17 375	19 602	22 797	25 419
72	3 084	2 298	1.877	4.521	2 605	4 528	5 921	7 357	9 369	11 001	12 740	15 220	17 247	19 413	22 508	25 038
73	2 992	2 132	2.006	6.185	2 537	4 303	5 598	6 944	8 850	10 412	12 090	14 509	16 504	18 656	21 761	24 325
74	2 844	2 052	2.295	8.134	2 388	4 050	5 297	6 617	8 524	10 117	11 859	14 418	16 571	18 932	22 408	25 334
75	2 804	2 076	2.250	7.820	2 345	4 033	5 295	6 627	8 547	10 146	11 890	14 446	16 590	18 935	22 378	25 268
76	2 736	2 054	2.408	8.649	2 274	3 924	5 171	6 499	8 432	10 058	11 845	14 490	16 730	19 200	22 860	25 962
77	2 694	2 069	2.406	8.516	2 229	3 892	5 149	6 487	8 434	10 071	11 871	14 534	16 789	19 276	22 960	26 082
78	2 577	2 063	2.547	9.282	2 107	3 748	5 000	6 342	8 310	9 978	11 825	14 577	16 926	19 534	23 428	26 753
79	2 601	2 059	2.527	9.248	2 133	3 773	5 023	6 361	8 322	9 981	11 817	14 550	16 880	19 465	23 319	26 607
80	2 608	2 058	2.521	9.239	2 140	3 780	5 029	6 366	8 325	9 982	11 814	14 542	16 867	19 444	23 288	26 565
Selected average descriptive statistics																
min1	1 964	1 525	0.760	-0.192	1 481	3 143	4 342	5 209	6 336	7 102	7 792	8 646	9 261	9 853	10 600	11 139
max1	3 913	2 579	2.649	9.282	3 512	5 626	7 046	8 417	10 236	11 736	13 273	15 386	17 375	19 602	23 639	27 169
AVE1	2 889	2 117	1.624	3.688	2 495	4 298	5 551	6 804	8 510	9 859	11 271	13 249	14 842	16 530	18 922	20 869
SD1					599	801	903	992	1 134	1 298	1 545	2 035	2 547	3 189	4 255	5 239
COV1					0.24	0.19	0.16	0.15	0.13	0.13	0.14	0.15	0.17	0.19	0.22	0.25

APPENDIX 4-C: Selected R-W42 results at D3R002 re LP3

The R-W42 FFA results, from where the LP3 distribution met the set criteria for RL1, at flow site D3R002.

Sub set	FFA moments				Annual exceedance probability (%)											
	ave	sdev	skew	kurt	50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
	(log-transformed)				Flood peaks in m³/s											
1	3.142	4 152	-0.295	1.042	1 448	2 938	4 152	5 460	7 337	8 867	10 492	12 782	14 620	16 547	19 229	21 358
2	3.144	0.382	-0.312	1.051	1 458	2 953	4 163	5 460	7 309	8 808	10 392	12 612	14 384	16 233	18 794	20 816
3	3.142	0.384	-0.304	1.000	1 452	2 949	4 167	5 476	7 350	8 874	10 489	12 759	14 577	16 478	19 119	21 209
4	3.154	0.387	-0.373	0.931	1 507	3 052	4 282	5 579	7 395	8 840	10 344	12 413	14 036	15 705	17 975	19 738
5	3.149	0.390	-0.342	0.787	1 484	3 035	4 291	5 631	7 533	9 066	10 678	12 922	14 702	16 550	19 091	21 084
6	3.157	0.388	-0.406	0.892	1 526	3 086	4 315	5 599	7 376	8 776	10 218	12 182	13 707	15 261	17 354	18 963
7	3.156	0.387	-0.401	0.924	1 519	3 068	4 289	5 566	7 335	8 730	10 170	12 133	13 659	15 215	17 315	18 931
8	3.159	0.388	-0.424	0.924	1 534	3 092	4 310	5 574	7 312	8 673	10 067	11 954	13 411	14 888	16 866	18 379
9	3.161	0.385	-0.441	1.015	1 548	3 100	4 302	5 541	7 233	8 546	9 885	11 685	13 064	14 456	16 308	17 717
10	3.175	0.382	-0.538	1.234	1 619	3 179	4 337	5 489	7 002	8 135	9 253	10 706	11 783	12 838	14 199	15 201
11	3.190	0.390	-0.559	1.043	1 682	3 342	4 576	5 803	7 410	8 609	9 790	11 316	12 443	13 542	14 954	15 987
12	3.226	0.347	-0.128	0.009	1 711	3 311	4 632	6 082	8 221	10 019	11 981	14 840	17 213	19 776	23 469	26 502
13	3.236	0.357	-0.124	-0.169	1 752	3 456	4 881	6 461	8 813	10 805	12 993	16 202	18 881	21 789	26 001	29 478
14	3.253	0.357	-0.222	-0.143	1 847	3 608	5 034	6 571	8 788	10 610	12 561	15 339	17 593	19 981	23 345	26 049
15	3.261	0.361	-0.263	-0.222	1 891	3 701	5 153	6 703	8 915	10 714	12 621	15 308	17 465	19 728	22 884	25 394
16	3.263	0.361	-0.279	-0.208	1 904	3 717	5 162	6 698	8 877	10 640	12 500	15 107	17 190	19 367	22 387	24 777
17	3.285	0.332	-0.070	-0.500	1 944	3 674	5 100	6 668	8 991	10 956	13 113	16 280	18 930	21 812	26 000	29 469
18	3.282	0.329	-0.081	-0.457	1 935	3 632	5 019	6 536	8 772	10 652	12 708	15 711	18 212	20 922	24 843	28 078
19	3.298	0.322	-0.160	-0.314	2 027	3 728	5 070	6 500	8 544	10 217	12 004	14 547	16 612	18 804	21 898	24 394
20	3.309	0.321	-0.247	-0.268	2 099	3 824	5 145	6 517	8 427	9 948	11 537	13 743	15 491	17 308	19 813	21 787
21	3.309	0.321	-0.247	-0.268	2 099	3 824	5 145	6 517	8 426	9 948	11 537	13 743	15 491	17 307	19 812	21 786
22	3.318	0.328	-0.268	-0.392	2 153	3 961	5 348	6 788	8 789	10 380	12 038	14 332	16 145	18 024	20 606	22 634
23	3.315	0.322	-0.375	-0.544	2 165	3 896	5 164	6 436	8 136	9 438	10 755	12 515	13 861	15 216	17 022	18 398
24	3.325	0.310	-0.345	-0.428	2 200	3 885	5 114	6 345	7 992	9 257	10 538	12 257	13 575	14 909	16 692	18 056
25	3.322	0.310	-0.319	-0.443	2 180	3 861	5 098	6 347	8 032	9 336	10 667	12 467	13 858	15 274	17 182	18 651
26	3.329	0.305	-0.365	-0.267	2 227	3 887	5 081	6 266	7 835	9 028	10 229	11 827	13 043	14 265	15 889	17 122
27	3.350	0.308	-0.417	-0.176	2 351	4 103	5 344	6 561	8 148	9 338	10 521	12 073	13 238	14 395	15 912	17 049
28	3.335	0.319	-0.360	-0.422	2 259	4 050	5 363	6 683	8 451	9 810	11 186	13 033	14 448	15 879	17 789	19 249
29	3.342	0.327	-0.320	-0.505	2 287	4 176	5 595	7 046	9 027	10 576	12 167	14 335	16 021	17 746	20 083	21 891
30	3.351	0.321	-0.387	-0.320	2 355	4 225	5 588	6 948	8 757	10 138	11 528	13 379	14 788	16 203	18 081	19 506
31	3.383	0.310	-0.217	-0.519	2 480	4 437	5 929	7 478	9 635	11 357	13 159	15 668	17 663	19 742	22 620	24 897
32	3.383	0.310	-0.224	-0.518	2 479	4 427	5 906	7 438	9 566	11 260	13 030	15 487	17 436	19 464	22 266	24 477
33	3.383	0.310	-0.230	-0.513	2 484	4 433	5 910	7 437	9 553	11 235	12 989	15 420	17 346	19 346	22 105	24 278
34	3.404	0.298	-0.253	-0.442	2 611	4 547	5 981	7 440	9 430	10 989	12 596	14 797	16 519	18 292	20 710	22 597
35	3.402	0.297	-0.230	-0.426	2 589	4 513	5 947	7 414	9 429	11 016	12 662	14 929	16 714	18 559	21 092	23 079
36	3.412	0.299	-0.307	-0.450	2 673	4 645	6 081	7 523	9 460	10 954	12 476	14 529	16 113	17 725	19 894	21 564
37	3.414	0.298	-0.323	-0.401	2 690	4 655	6 077	7 497	9 392	10 847	12 321	14 300	15 821	17 360	19 423	21 004
38	3.420	0.293	-0.364	-0.242	2 738	4 682	6 060	7 415	9 196	10 542	11 889	13 674	15 026	16 381	18 174	19 531
39	3.423	0.290	-0.369	-0.158	2 756	4 684	6 042	7 374	9 116	10 429	11 740	13 471	14 779	16 087	17 814	19 119
40	3.420	0.292	-0.346	-0.232	2 733	4 669	6 048	7 409	9 207	10 573	11 945	13 771	15 161	16 559	18 418	19 831
41	3.417	0.292	-0.318	-0.245	2 708	4 639	6 027	7 409	9 248	10 656	12 081	13 992	15 458	16 941	18 927	20 448

Sub set	FFA moments				Annual exceedance probability (%)											
	ave	sdev	skew	kurt	50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
	(log-transformed)				Flood peaks in m^3/s											
42	3.418	0.292	-0.320	-0.254	2 712	4 647	6 039	7 423	9 266	10 677	12 104	14 018	15 485	16 970	18 957	20 479
43	3.436	0.308	-0.231	-0.235	2 807	4 991	6 641	8 345	10 702	12 572	14 521	17 219	19 353	21 567	24 619	27 022
44	3.449	0.308	-0.327	-0.191	2 924	5 153	6 784	8 424	10 626	12 324	14 051	16 378	18 171	19 990	22 433	24 310
45	3.454	0.299	-0.238	-0.261	2 924	5 113	6 745	8 415	10 705	12 508	14 376	16 946	18 967	21 055	23 918	26 161
46	3.452	0.299	-0.213	-0.272	2 900	5 084	6 725	8 415	10 751	12 603	14 532	17 205	19 322	21 520	24 552	26 942
47	3.474	0.283	-0.146	-0.353	3 026	5 173	6 786	8 452	10 770	12 622	14 565	17 282	19 452	21 726	24 894	27 416
48	3.467	0.293	-0.202	-0.437	2 995	5 193	6 838	8 528	10 860	12 707	14 631	17 295	19 405	21 597	24 621	27 007
49	3.471	0.292	-0.246	-0.409	3 038	5 244	6 871	8 521	10 769	12 527	14 338	16 816	18 756	20 750	23 472	25 596
50	3.455	0.321	-0.543	0.140	3 046	5 367	6 963	8 482	10 400	11 790	13 131	14 832	16 067	17 259	18 771	19 868
51	3.469	0.310	-0.621	0.509	3 166	5 428	6 918	8 292	9 968	11 145	12 251	13 615	14 578	15 486	16 607	17 400
52	3.467	0.311	-0.599	0.443	3 144	5 413	6 923	8 326	10 053	11 275	12 432	13 868	14 890	15 859	17 065	17 923
53	3.462	0.309	-0.575	0.498	3 100	5 331	6 826	8 224	9 955	11 189	12 364	13 834	14 886	15 891	17 149	18 051
54	3.451	0.315	-0.484	0.230	2 994	5 252	6 827	8 348	10 298	11 734	13 139	14 951	16 287	17 595	19 280	20 522
55	3.440	0.312	-0.412	0.259	2 890	5 082	6 647	8 187	10 206	11 725	13 240	15 233	16 734	18 228	20 192	21 668
56	3.430	0.313	-0.325	0.175	2 797	4 971	6 573	8 190	10 370	12 057	13 777	16 101	17 895	19 719	22 175	24 065
57	3.430	0.313	-0.329	0.164	2 802	4 983	6 588	8 206	10 387	12 072	13 790	16 108	17 896	19 713	22 157	24 035
58	3.428	0.315	-0.311	0.084	2 780	4 970	6 596	8 247	10 488	12 231	14 017	16 442	18 324	20 246	22 845	24 854
59	3.422	0.320	-0.281	-0.088	2 733	4 947	6 620	8 340	10 706	12 569	14 498	17 147	19 224	21 364	24 288	26 570
60	3.405	0.328	-0.202	-0.287	2 607	4 830	6 573	8 417	11 033	13 154	15 403	18 576	21 131	23 820	27 584	30 592
61	3.405	0.328	-0.197	-0.284	2 602	4 821	6 564	8 411	11 034	13 165	15 427	18 622	21 198	23 913	27 719	30 765
62	3.406	0.329	-0.210	-0.305	2 617	4 851	6 600	8 447	11 063	13 181	15 423	18 580	21 117	23 784	27 511	30 484
63	3.396	0.329	-0.123	-0.341	2 526	4 726	6 500	8 419	11 210	13 530	16 039	19 663	22 646	25 848	30 427	34 165
64	3.388	0.325	-0.079	-0.228	2 466	4 593	6 324	8 211	10 984	13 311	15 851	19 554	22 634	25 967	30 781	34 750
65	3.380	0.316	-0.113	-0.085	2 430	4 435	6 028	7 738	10 204	12 241	14 434	17 586	20 171	22 936	26 878	30 087
66	3.394	0.317	-0.208	-0.092	2 541	4 611	6 207	7 879	10 226	12 113	14 100	16 882	19 108	21 438	24 681	27 259
67	3.392	0.319	-0.189	-0.150	2 521	4 595	6 209	7 910	10 314	12 260	14 320	17 224	19 559	22 017	25 457	28 207
68	3.384	0.325	-0.170	-0.319	2 474	4 576	6 236	8 005	10 534	12 602	14 809	17 948	20 495	23 195	27 003	30 070
69	3.373	0.318	-0.136	-0.171	2 401	4 395	5 972	7 658	10 078	12 066	14 199	17 248	19 736	22 386	26 146	29 192
70	3.380	0.308	-0.073	-0.079	2 421	4 373	5 927	7 601	10 029	12 046	14 229	17 388	19 995	22 800	26 827	30 126
71	3.357	0.311	-0.044	-0.104	2 285	4 159	5 671	7 314	9 724	11 747	13 954	17 178	19 864	22 775	26 991	30 475
72	3.358	0.311	-0.060	-0.110	2 299	4 178	5 688	7 322	9 709	11 702	13 870	17 023	19 638	22 463	26 537	29 891
73	3.345	0.298	-0.110	0.070	2 242	3 955	5 283	6 687	8 684	10 313	12 051	14 525	16 535	18 671	21 693	24 136
Average descriptive statistics																
min1	3.142	0.283	-0.621	-0.544	1 448	2 938	4 152	5 460	7 002	8 135	9 253	10 706	11 783	12 838	14 199	15 201
max1	3.474	0.390	-0.044	1.234	3 166	5 428	6 963	8 528	11 210	13 530	16 039	19 663	22 646	25 967	30 781	34 750
AVE1	3.347	0.326	-0.286	-0.025	2 360	4 275	5 725	7 220	9 286	10 923	12 626	14 984	16 849	18 787	21 460	23 569
SD1					492	723	861	995	1 200	1 404	1 668	2 126	2 565	3 086	3 907	4 630
COV1					0.21	0.17	0.15	0.14	0.13	0.13	0.13	0.14	0.15	0.16	0.18	0.20

APPENDIX 4-D: 7ym R-W results at D3R002, for GEV

A comparison of all the 7ym R-W average results at D3R002, applying the GEV distribution. D3R002 was chosen as an example, to illustrate a case where the results of the GEV and LP3 were very much comparable – see also Appendix 4-F for a graphic comparison.

7ym R-W (yrs)	Stats	AEP (%)											
		50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
		GEV distribution: Flood peaks in m³/s											
7	min1	910	1 492	1 840	2 148	2 497	2 615	2 708	2 800	2 853	2 894	2 936	2 960
	max1	4 954	7 471	9 034	11 145	13 965	16 469	19 431	23 940	27 735	31 897	38 027	43 189
	AVE1	2 442	4 007	4 999	5 931	7 123	8 014	8 908	10 104	11 028	11 974	13 267	14 283
	SD1	956	1 550	2 006	2 501	3 238	3 871	4 574	5 619	6 503	7 474	8 899	10 092
	COV1	0.39	0.39	0.40	0.42	0.45	0.48	0.51	0.56	0.59	0.62	0.67	0.71
14	min1	1 051	1 651	2 031	2 385	2 829	3 136	3 416	3 751	3 986	4 207	4 478	4 668
	max1	4 238	6 611	8 065	9 778	12 049	13 919	15 869	18 841	21 693	24 799	29 395	33 545
	AVE1	2 430	4 108	5 222	6 303	7 731	8 832	9 962	11 516	12 746	14 032	15 831	17 277
	SD1	797	1 207	1 532	1 908	2 516	3 075	3 731	4 758	5 664	6 689	8 239	9 572
	COV1	0.33	0.29	0.29	0.30	0.33	0.35	0.37	0.41	0.44	0.48	0.52	0.55
21	min1	1 225	2 264	3 016	3 788	4 867	5 602	6 178	6 733	7 077	7 390	7 763	8 016
	max1	3 851	6 135	7 845	9 470	11 550	13 093	14 615	16 784	19 426	22 356	27 381	31 750
	AVE1	2 454	4 209	5 404	6 583	8 172	9 419	10 719	12 536	13 999	15 551	17 757	19 560
	SD1	726	1 005	1 209	1 452	1 881	2 315	2 860	3 766	4 606	5 591	7 135	8 504
	COV1	0.30	0.24	0.22	0.22	0.23	0.25	0.27	0.30	0.33	0.36	0.40	0.43
28	min1	1 355	2 828	3 676	4 545	5 521	6 079	6 604	7 251	7 707	8 093	8 532	8 835
	max1	3 655	5 884	7 423	8 907	10 821	12 313	13 796	15 989	18 018	20 799	25 101	29 197
	AVE1	2 476	4 265	5 499	6 730	8 401	9 721	11 102	13 042	14 608	16 271	18 638	20 574
	SD1	667	896	1 035	1 187	1 458	1 754	2 150	2 853	3 535	4 357	5 678	6 870
	COV1	0.27	0.21	0.19	0.18	0.17	0.18	0.19	0.22	0.24	0.27	0.30	0.33
35	min1	1 481	3 143	4 342	5 209	6 336	7 102	7 792	8 646	9 261	9 853	10 600	11 139
	max1	3 512	5 626	7 046	8 417	10 236	11 736	13 273	15 386	17 375	19 602	23 639	27 169
	AVE1	2 495	4 298	5 551	6 804	8 510	9 859	11 271	13 249	14 842	16 530	18 922	20 869
	SD1	599	801	903	992	1 134	1 298	1 545	2 035	2 547	3 189	4 255	5 239
	COV1	0.24	0.19	0.16	0.15	0.13	0.13	0.14	0.15	0.17	0.19	0.22	0.25
42	min1	1 536	3 082	4 270	5 550	7 185	8 146	9 092	10 326	11 254	12 178	13 393	14 307
	max1	3 264	5 337	6 717	8 060	9 806	11 262	12 839	15 052	16 802	18 809	22 319	25 654
	AVE1	2 524	4 345	5 610	6 873	8 586	9 935	11 339	13 292	14 853	16 494	18 797	20 651
	SD1	525	714	802	865	937	1 010	1 129	1 406	1 732	2 173	2 941	3 672
	COV1	0.21	0.16	0.14	0.13	0.11	0.10	0.10	0.11	0.12	0.13	0.16	0.18
49	min1	1 507	2 950	4 066	5 276	7 070	8 506	9 493	10 787	11 759	12 725	13 994	14 948
	max1	3 042	5 023	6 373	7 695	9 469	10 880	12 443	14 627	16 360	18 166	21 504	24 792
	AVE1	2 541	4 373	5 645	6 911	8 625	9 969	11 363	13 291	14 823	16 422	18 650	20 429
	SD1	445	612	691	745	792	825	877	1 019	1 213	1 498	2 029	2 553
	COV1	0.18	0.14	0.12	0.11	0.09	0.08	0.08	0.08	0.08	0.09	0.11	0.12

7ym R-W (yrs)	Stats	AEP (%)											
		50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
		GEV distribution: Flood peaks in m³/s											
56	min1	1 737	3 279	4 427	5 636	7 368	8 800	10 099	11 411	12 389	13 355	14 613	15 552
	max1	2 968	4 888	6 194	7 472	9 261	10 644	12 105	14 107	15 677	17 338	19 726	22 283
	AVE1	2 554	4 397	5 675	6 947	8 663	10 005	11 390	13 299	14 806	16 370	18 532	20 244
	SD1	342	469	534	583	631	663	701	781	884	1 036	1 327	1 620
	COV1	0.13	0.11	0.09	0.08	0.07	0.07	0.06	0.06	0.06	0.06	0.07	0.08
63	min1	1 896	3 449	4 580	5 749	7 395	8 735	10 169	12 221	13 905	15 711	17 704	19 119
	max1	2 791	4 728	6 178	7 632	9 610	11 166	12 782	15 020	16 795	18 645	21 211	23 247
	AVE1	2 544	4 406	5 698	6 985	8 723	10 081	11 483	13 412	14 932	16 508	18 680	20 395
	SD1	249	323	362	390	411	416	413	405	406	427	507	618
	COV1	0.10	0.07	0.06	0.06	0.05	0.04	0.04	0.03	0.03	0.03	0.03	0.03
70	min1	2 035	3 772	5 007	6 259	7 982	9 350	10 785	12 796	14 402	15 842	17 804	19 333
	max1	2 680	4 687	6 104	7 522	9 436	10 931	12 505	14 699	16 446	18 272	20 809	22 826
	AVE1	2 518	4 392	5 697	6 998	8 757	10 134	11 557	13 517	15 064	16 668	18 881	20 628
	SD1	166	205	243	286	349	402	461	548	624	710	840	953
	COV1	0.07	0.05	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.05
77	min1	2 229	4 083	5 369	6 644	8 352	9 686	11 054	12 916	14 376	15 885	17 955	19 581
	max1	2 628	4 581	5 974	7 362	9 237	10 720	12 313	14 561	16 356	18 238	20 866	22 967
	AVE1	2 483	4 357	5 668	6 980	8 762	10 163	11 616	13 625	15 217	16 875	19 170	20 990
	SD1	114	157	211	273	364	439	520	636	732	836	985	1 109
	COV1	0.05	0.04	0.04	0.04	0.04	0.04	0.04	0.05	0.05	0.05	0.05	0.05
84	min1	2 251	4 188	5 453	6 694	8 385	9 717	11 094	12 967	14 447	15 985	18 110	19 791
	max1	2 524	4 441	5 797	7 158	9 029	10 522	12 118	14 425	16 284	18 249	21 020	23 254
	AVE1	2 448	4 336	5 665	7 001	8 825	10 266	11 769	13 857	15 520	17 259	19 680	21 610
	SD1	73	73	116	174	265	344	431	563	675	798	981	1 136
	COV1	0.03	0.02	0.02	0.02	0.03	0.03	0.04	0.04	0.04	0.05	0.05	0.05
91	min1	2 259	4 133	5 353	6 574	8 242	9 561	10 936	12 849	14 373	15 960	18 165	19 921
	max1	2 527	4 404	5 722	7 046	8 852	10 328	11 909	14 122	15 896	17 761	20 375	22 473
	AVE1	2 396	4 269	5 595	6 933	8 768	10 225	11 748	13 876	15 577	17 362	19 859	21 857
	SD1	72	67	84	113	161	204	252	324	384	451	549	631
	COV1	0.03	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.03
98	min1	2 249	4 108	5 432	6 776	8 633	10 097	11 597	13 696	15 366	17 112	19 550	21 499
	max1	2 434	4 260	5 558	6 877	8 685	10 121	11 682	13 876	15 643	17 510	20 139	22 257
	AVE1	2 340	4 189	5 501	6 829	8 657	10 111	11 637	13 774	15 488	17 291	19 821	21 851
	SD1	69	63	54	42	21	7	30	73	113	160	233	297
	COV1	0.03	0.01	0.01	0.01	0.002	0.001	0.003	0.01	0.01	0.01	0.01	0.01
105	min1	2 208	4 017	5 312	6 634	8 467	9 937	11 454	13 574	15 278	17 076	19 604	21 636
	max1	2 317	4 121	5 406	6 714	8 523	9 965	11 493	13 687	15 461	17 340	19 998	22 149
	AVE1	2 277	4 093	5 385	6 698	8 509	9 954	11 474	13 608	15 325	17 136	19 683	21 733
	SD1	32	34	33	29	20	10	10	33	57	86	134	177
	COV1	0.014	0.008	0.006	0.004	0.002	0.001	0.001	0.002	0.004	0.005	0.007	0.008
112	min1	2 187	3 972	5 250	6 553	8 356	9 798	11 320	13 468	15 204	17 042	19 641	21 743
	max1	2 217	3 992	5 263	6 559	8 361	9 811	11 342	13 504	15 251	17 101	19 717	21 834
	AVE1	2 202	3 982	5 257	6 557	8 359	9 805	11 331	13 485	15 225	17 068	19 674	21 780
	SD1	15	10	7	3	3	7	11	18	24	30	39	47
	COV1	0.007	0.003	0.001	0.001	0.000	0.001	0.001	0.001	0.002	0.002	0.002	0.002

APPENDIX 4-E: 7ym R-W results at D3R002, for LP3

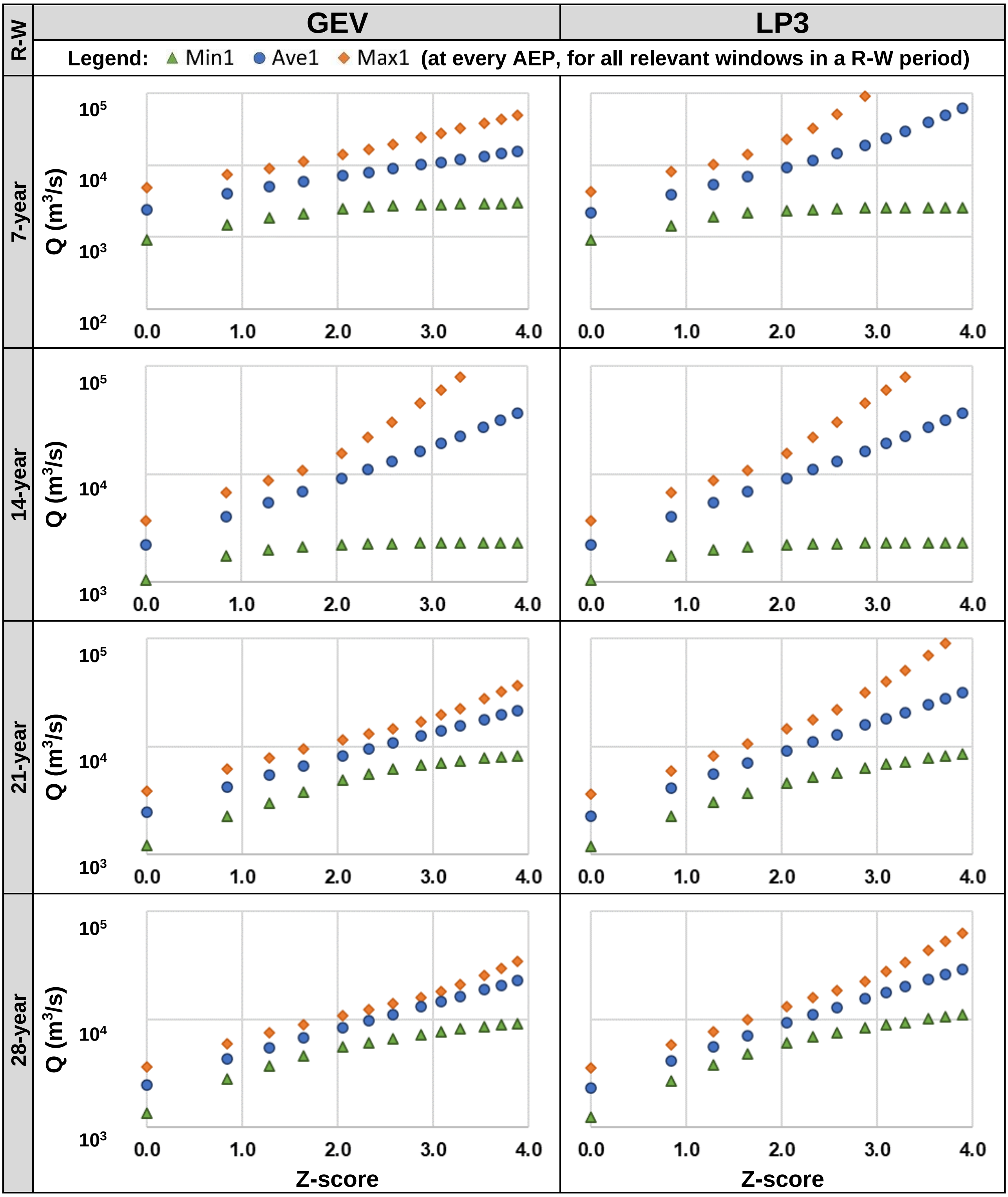
A comparison of all the 7ym R-W average results at D3R002, applying the LP3 distribution. D3R002 was chosen as an example, to illustrate a case where the results of the GEV and LP3 were very much comparable – see also Appendix 4-F for a graphic comparison.

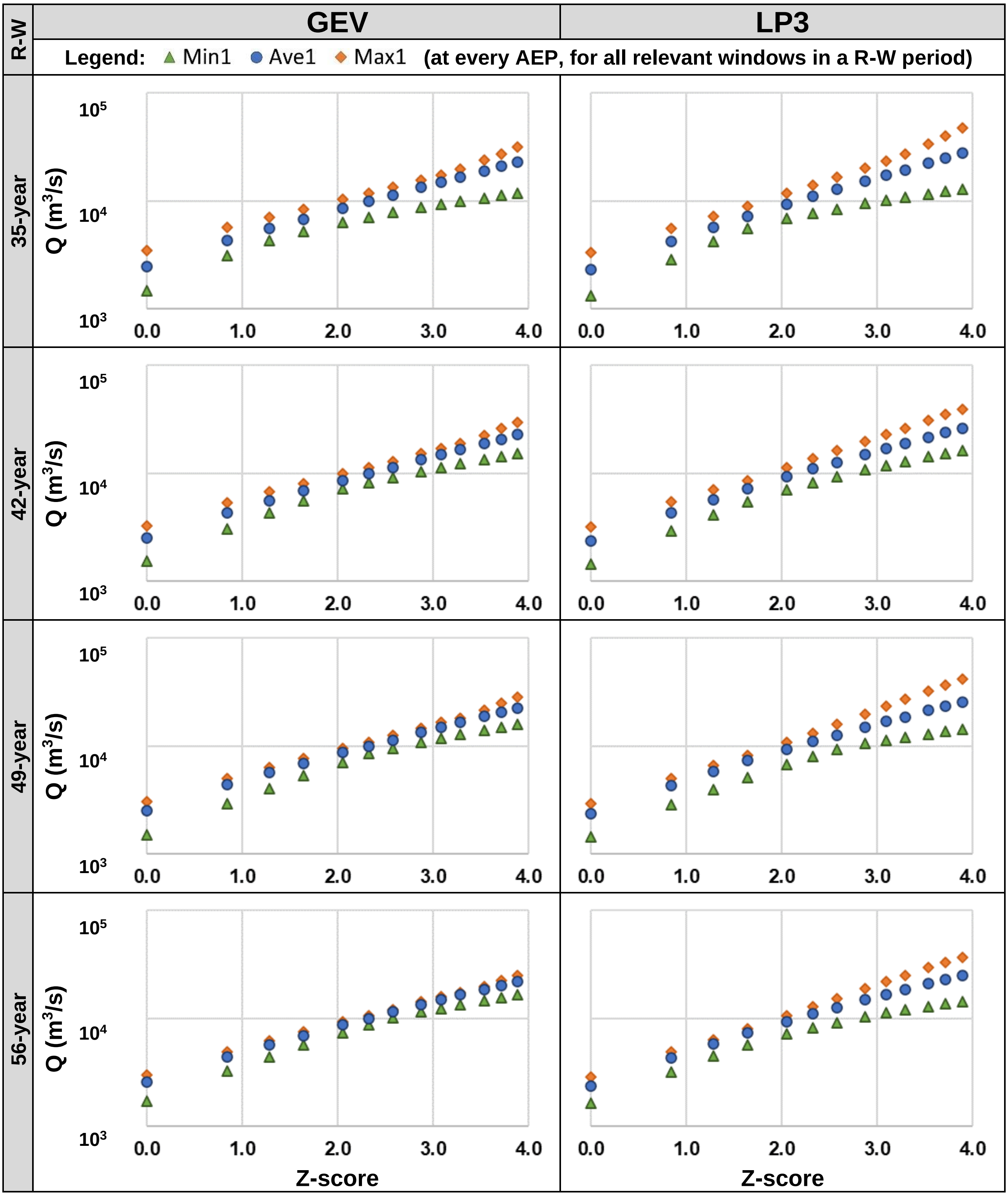
7ym R-W (yrs)	Stats	AEP (%)											
		50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
		LP3 distribution: Flood peaks in m³/s											
7	min1	922	1 465	1 955	2 175	2 360	2 447	2 506	2 555	2 578	2 593	2 605	2 610
	max1	4 347	8 021	10 349	14 052	22 518	32 174	50 028	89 178	137 633	211 904	373 643	572 517
	AVE1	2 203	3 923	5 322	6 896	9 369	11 645	14 391	18 992	23 464	29 092	38 969	48 983
	SD1	889	1 519	2 200	3 206	5 293	7 699	11 131	17 977	25 730	36 772	58 981	84 501
	COV1	0.40	0.39	0.41	0.46	0.56	0.66	0.77	0.95	1.10	1.26	1.51	1.73
14	min1	1 048	1 747	1 980	2 119	2 221	2 263	2 289	2 308	2 316	2 322	2 326	2 329
	max1	3 742	6 719	8 682	10 795	15 548	21 883	30 173	45 023	60 045	79 237	112 772	145 970
	AVE1	2 229	4 002	5 390	6 886	9 088	10 968	13 073	16 273	19 072	22 261	27 194	31 572
	SD1	747	1 183	1 608	2 213	3 413	4 715	6 439	9 520	12 600	16 474	23 101	29 516
	COV1	0.33	0.30	0.30	0.32	0.38	0.43	0.49	0.59	0.66	0.74	0.85	0.93
21	min1	1 201	2 290	3 035	3 729	4 576	5 165	5 713	6 377	6 838	7 264	7 779	8 134
	max1	3 647	5 983	8 087	10 634	14 464	17 751	22 054	31 060	39 617	50 256	69 605	88 387
	AVE1	2 274	4 100	5 511	7 005	9 150	10 928	12 862	15 693	18 070	20 679	24 533	27 796
	SD1	698	1 035	1 284	1 629	2 352	3 165	4 246	6 145	7 993	10 254	13 995	17 498
	COV1	0.31	0.25	0.23	0.23	0.26	0.29	0.33	0.39	0.44	0.50	0.57	0.63
28	min1	1 238	2 705	3 764	4 761	6 116	6 839	7 471	8 261	8 827	9 369	10 050	10 542
	max1	3 516	5 750	7 739	9 938	13 078	15 684	18 485	22 491	27 583	33 691	43 372	52 098
	AVE1	2 302	4 163	5 599	7 109	9 248	10 990	12 851	15 510	17 684	20 012	23 344	26 074
	SD1	637	933	1 122	1 335	1 733	2 170	2 757	3 798	4 808	6 033	8 020	9 835
	COV1	0.28	0.22	0.20	0.19	0.19	0.20	0.21	0.24	0.27	0.30	0.34	0.38
35	min1	1 333	2 871	4 231	5 564	6 817	7 641	8 425	9 405	10 107	10 777	11 617	12 220
	max1	3 357	5 588	7 239	8 939	11 725	13 986	16 376	19 995	23 091	26 788	33 751	39 904
	AVE1	2 330	4 212	5 651	7 151	9 249	10 935	12 714	15 218	17 235	19 364	22 362	24 777
	SD1	566	832	984	1 125	1 347	1 584	1 918	2 547	3 185	3 975	5 271	6 456
	COV1	0.24	0.20	0.17	0.16	0.15	0.14	0.15	0.17	0.18	0.21	0.24	0.26
42	min1	1 448	2 938	4 152	5 460	7 002	8 135	9 253	10 706	11 783	12 838	14 199	15 201
	max1	3 166	5 428	6 963	8 528	11 210	13 530	16 039	19 663	22 646	25 967	30 781	34 750
	AVE1	2 360	4 275	5 725	7 220	9 286	10 923	12 626	14 984	16 849	18 787	21 460	23 569
	SD1	492	723	861	995	1 200	1 404	1 668	2 126	2 565	3 086	3 907	4 630
	COV1	0.21	0.17	0.15	0.14	0.13	0.13	0.13	0.14	0.15	0.16	0.18	0.20
49	min1	1 444	2 847	3 969	5 143	6 748	7 997	9 286	10 512	11 299	12 031	12 920	13 536
	max1	2 921	5 039	6 587	8 176	10 762	13 168	15 835	19 794	23 143	26 819	32 221	36 746
	AVE1	2 375	4 312	5 772	7 268	9 323	10 940	12 613	14 914	16 724	18 593	21 157	23 169
	SD1	412	606	735	873	1 097	1 320	1 601	2 079	2 528	3 056	3 883	4 610
	COV1	0.17	0.14	0.13	0.12	0.12	0.12	0.13	0.14	0.15	0.16	0.18	0.20

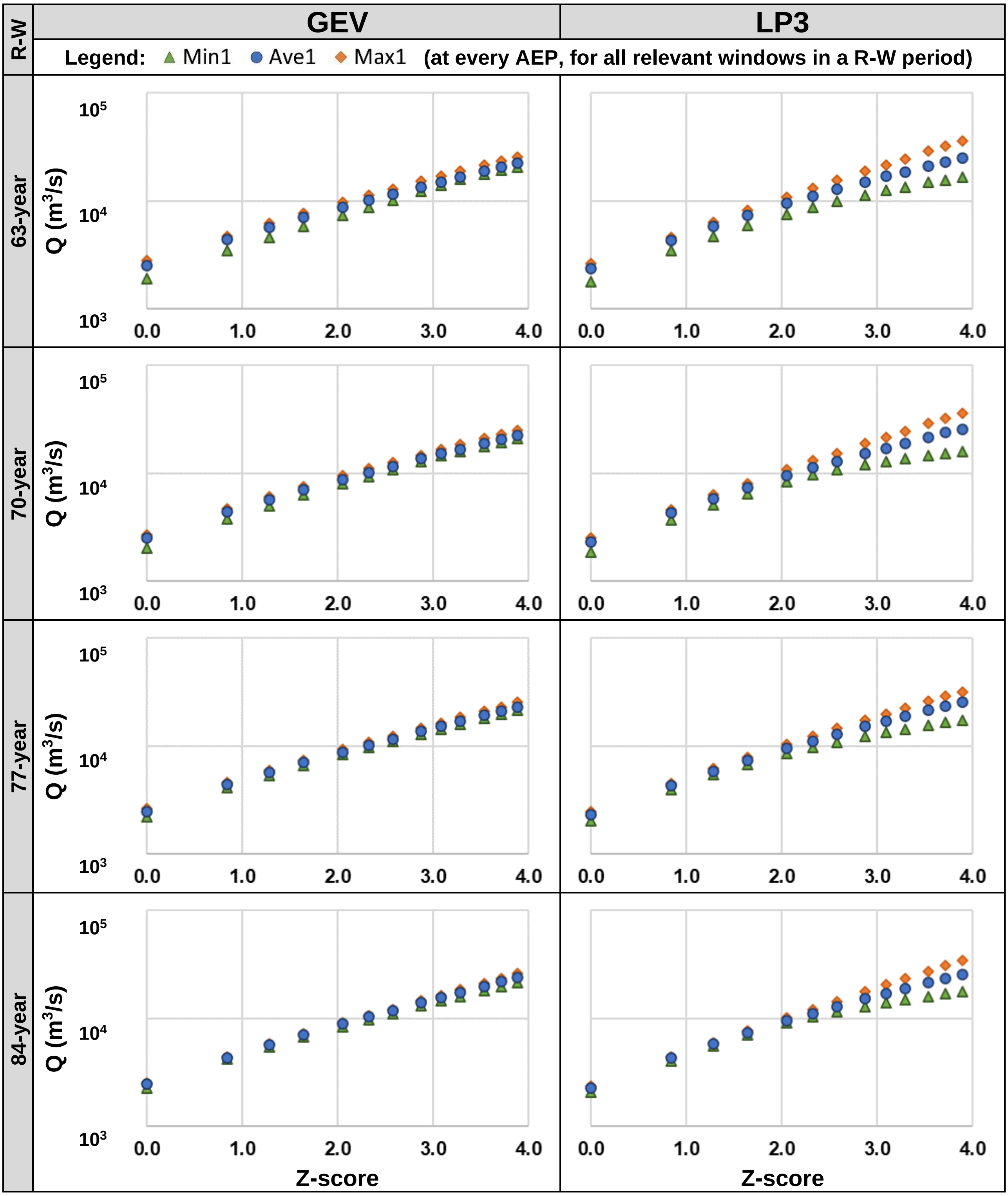
7ym R-W (yrs)	Stats	AEP (%)											
		50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
		LP3 distribution: Flood peaks in m^3/s											
56	min1	1 634	3 220	4 451	5 684	7 239	8 231	9 172	10 339	11 165	11 946	12 909	13 589
	max1	2 841	4 902	6 358	7 923	10 605	12 860	15 303	18 830	21 734	24 848	29 297	32 924
	AVE1	2 387	4 345	5 812	7 310	9 353	10 951	12 598	14 848	16 608	18 417	20 885	22 810
	SD1	318	443	550	692	959	1 230	1 564	2 104	2 589	3 139	3 969	4 675
	COV1	0.13	0.10	0.09	0.09	0.10	0.11	0.12	0.14	0.16	0.17	0.19	0.20
63	min1	1 802	3 464	4 683	5 890	7 466	8 643	9 804	11 308	12 423	13 482	14 728	15 627
	max1	2 658	4 636	6 279	8 130	10 803	13 009	15 381	18 778	21 554	24 513	28 713	32 115
	AVE1	2 372	4 348	5 835	7 354	9 427	11 048	12 714	14 989	16 762	18 581	21 055	22 978
	SD1	237	301	375	484	695	909	1 173	1 599	1 981	2 416	3 073	3 632
	COV1	0.10	0.07	0.06	0.07	0.07	0.08	0.09	0.11	0.12	0.13	0.15	0.16
70	min1	1 863	3 680	5 081	6 520	8 431	9 726	10 752	11 996	12 858	13 656	14 620	15 285
	max1	2 552	4 604	6 278	8 068	10 734	12 936	15 305	18 704	21 483	24 447	28 710	32 213
	AVE1	2 338	4 322	5 824	7 366	9 479	11 138	12 851	15 198	17 036	18 929	21 514	23 533
	SD1	170	196	254	357	574	803	1 092	1 566	1 997	2 491	3 243	3 889
	COV1	0.07	0.05	0.04	0.05	0.06	0.07	0.08	0.10	0.12	0.13	0.15	0.17
77	min1	2 025	3 994	5 484	6 811	8 490	9 702	10 865	12 327	13 377	14 380	15 635	16 533
	max1	2 474	4 523	6 174	7 870	10 343	12 342	14 455	17 422	19 799	22 291	25 760	28 519
	AVE1	2 308	4 288	5 786	7 322	9 425	11 072	12 771	15 094	16 910	18 777	21 325	23 311
	SD1	113	148	224	331	530	736	998	1 432	1 831	2 290	2 992	3 595
	COV1	0.05	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.11	0.12	0.14	0.15
84	min1	2 070	4 077	5 578	7 053	9 028	10 251	11 405	12 829	13 833	14 778	15 938	16 753
	max1	2 372	4 444	5 943	7 599	10 047	12 131	14 383	17 633	20 305	23 169	27 262	30 599
	AVE1	2 274	4 265	5 777	7 328	9 452	11 116	12 831	15 174	17 005	18 886	21 450	23 447
	SD1	76	86	125	181	323	501	744	1 167	1 567	2 037	2 766	3 401
	COV1	0.03	0.02	0.02	0.02	0.03	0.05	0.06	0.08	0.09	0.11	0.13	0.15
91	min1	2 063	4 051	5 453	6 907	8 795	10 053	11 262	12 788	13 888	14 941	16 264	17 213
	max1	2 336	4 310	5 849	7 470	9 771	11 755	13 895	16 970	19 492	22 189	26 032	29 159
	AVE1	2 229	4 203	5 699	7 231	9 319	10 950	12 626	14 909	16 687	18 508	20 985	22 909
	SD1	68	68	87	144	321	534	816	1 294	1 739	2 257	3 055	3 747
	COV1	0.03	0.02	0.02	0.02	0.03	0.05	0.06	0.09	0.10	0.12	0.15	0.16
98	min1	2 091	4 042	5 502	6 929	8 574	9 731	10 830	12 197	13 170	14 091	15 234	16 044
	max1	2 298	4 279	5 678	7 259	9 568	11 465	13 497	16 400	18 765	21 282	24 848	27 731
	AVE1	2 192	4 141	5 593	7 057	9 019	10 525	12 047	14 085	15 644	17 218	19 319	20 924
	SD1	69	72	67	130	331	553	833	1 291	1 705	2 178	2 891	3 499
	COV1	0.03	0.02	0.01	0.02	0.04	0.05	0.07	0.09	0.11	0.13	0.15	0.17
105	min1	2 066	3 952	5 369	6 778	8 436	9 645	10 815	12 301	13 381	14 422	15 740	16 693
	max1	2 219	4 126	5 489	6 891	8 734	10 215	11 710	13 703	15 219	16 740	18 753	20 276
	AVE1	2 151	4 067	5 465	6 846	8 653	10 004	11 339	13 076	14 367	15 637	17 281	18 498
	SD1	43	52	41	37	96	169	259	401	523	658	854	1 014
	COV1	0.020	0.013	0.007	0.005	0.011	0.017	0.023	0.031	0.036	0.042	0.049	0.055
112	min1	2 040	3 896	5 305	6 704	8 538	9 931	11 327	13 171	14 563	15 949	17 771	19 138
	max1	2 090	3 948	5 325	6 743	8 696	10 210	11 753	13 830	15 427	17 042	19 202	20 851
	AVE1	2 064	3 922	5 317	6 729	8 628	10 086	11 561	13 529	15 030	16 537	18 536	20 052
	SD1	25	26	11	22	81	142	216	333	436	551	721	862
	COV1	0.012	0.007	0.002	0.003	0.009	0.014	0.019	0.025	0.029	0.033	0.039	0.043

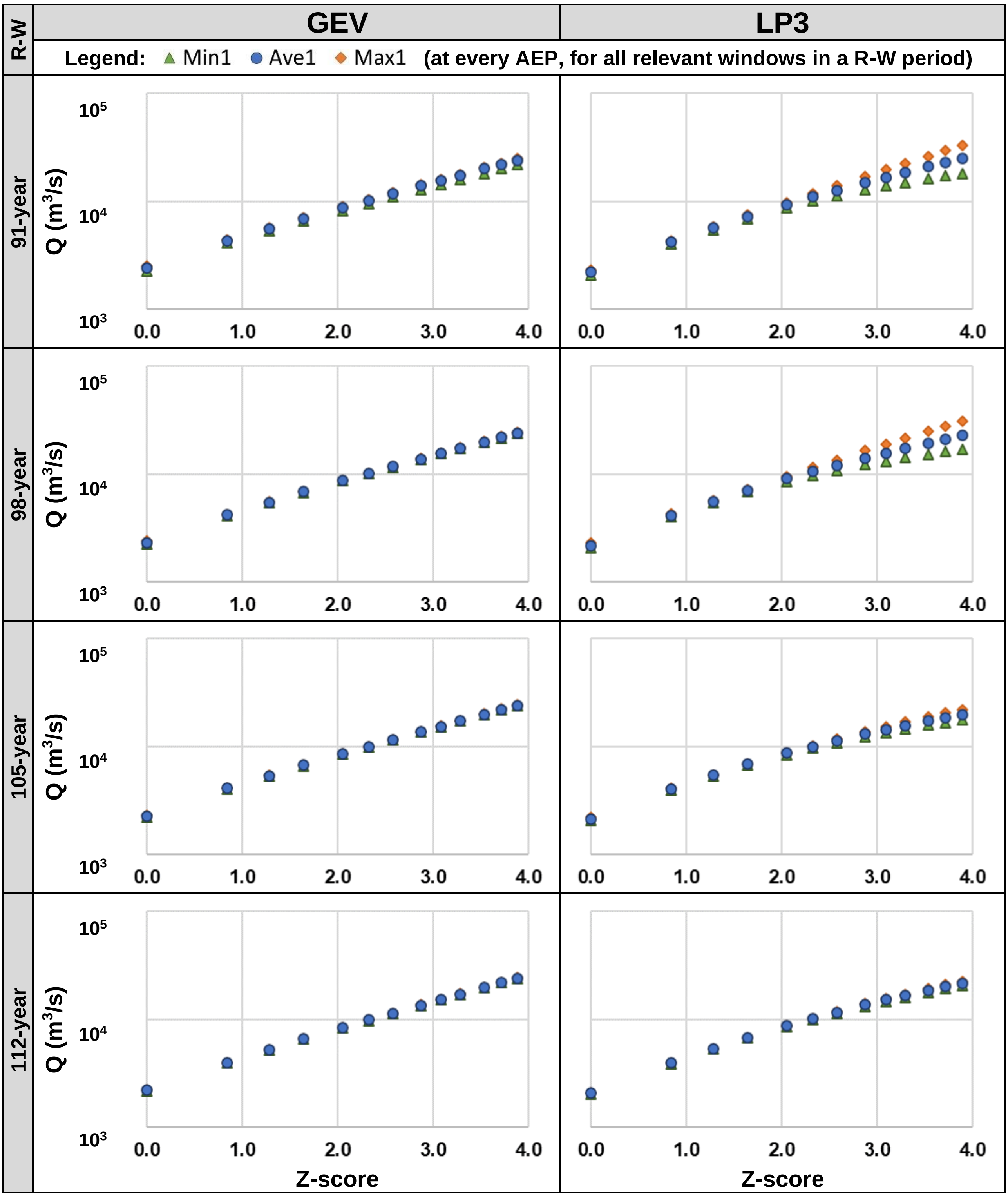
APPENDIX 4-F: All 7ym R-W results at D3R002

A visual comparison of all the 7ym R-W average results at D3R002, which was chosen as an example, to illustrate a case where the results of the GEV and LP3 were very much comparable.




R-W
GEV
LP3
Legend: Min1 Ave1 Max1 (at every AEP, for all relevant windows in a R-W period)
35-year
42-year
49-year
56-year
Q (m3/s)
10^5
10^4
10^3
0.0
1.0
2.0
3.0
4.0
Z-score


R-W
GEV
LP3
Legend: Min1 Ave1 Max1 (at every AEP, for all relevant windows in a R-W period)
63-year
70-year
77-year
84-year
Q (m³/s)
10⁵
10⁴
10³
0.0
1.0
2.0
3.0
4.0
Z-score


R-W
GEV
LP3
Legend: Min1 Ave1 Max1 (at every AEP, for all relevant windows in a R-W period)
91-year
98-year
105-year
112-year
Q (m3/s)
10⁵
10⁴
10³
0.0
1.0
2.0
3.0
4.0
Z-score

APPENDIX 4-G: Selected R-W42 results at N2R001 re GEV

The R-W42 FFA results, from where the GEV distribution met the set criteria for RL1, at flow site N2R001.

Sub set	FFA moments				Annual exceedance probability (%)											
	ave	sdev	skew	kurt	50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
	Flow in m³/s				Flood peaks in m³/s											
1	542	822	2.445	5.310	357	1 015	1 515	2 047	2 824	3 478	4 199	5 268	6 175	7 177	8 666	9 930
2	541	823	2.442	5.299	355	1 015	1 514	2 047	2 824	3 479	4 200	5 269	6 176	7 178	8 666	9 929
3	537	825	2.437	5.272	350	1 012	1 513	2 047	2 826	3 481	4 204	5 274	6 182	7 184	8 672	9 936
4	537	825	2.437	5.273	351	1 012	1 513	2 047	2 826	3 481	4 203	5 273	6 181	7 183	8 671	9 934
5	538	824	2.439	5.283	353	1 013	1 513	2 047	2 825	3 480	4 202	5 272	6 179	7 182	8 669	9 933
6	537	825	2.436	5.271	351	1 012	1 513	2 047	2 826	3 481	4 203	5 273	6 181	7 183	8 670	9 933
7	488	780	2.802	7.483	306	916	1 388	1 899	2 659	3 311	4 041	5 142	6 093	7 160	8 773	10 168
8	602	1 054	2.926	8.773	355	1 172	1 808	2 502	3 538	4 431	5 435	6 960	8 284	9 776	12 044	14 015
9	584	1 057	2.947	8.840	336	1 155	1 793	2 489	3 529	4 428	5 438	6 973	8 308	9 813	12 103	14 095
10	582	1 058	2.945	8.825	334	1 153	1 792	2 489	3 530	4 429	5 440	6 975	8 310	9 815	12 105	14 097
11	544	971	3.231	11.741	312	1 052	1 635	2 278	3 250	4 097	5 060	6 539	7 838	9 318	11 593	13 594
12	538	973	3.223	11.683	306	1 048	1 633	2 277	3 250	4 100	5 064	6 545	7 845	9 326	11 602	13 603
13	538	974	3.223	11.681	306	1 048	1 633	2 277	3 250	4 100	5 064	6 545	7 845	9 326	11 602	13 603
14	582	1 001	2.923	9.702	348	1 124	1 728	2 386	3 370	4 218	5 171	6 618	7 874	9 289	11 441	13 311
15	572	1 004	2.919	9.648	336	1 116	1 722	2 383	3 370	4 220	5 177	6 627	7 887	9 306	11 463	13 337
16	583	1 002	2.905	9.616	348	1 126	1 732	2 391	3 375	4 222	5 174	6 618	7 871	9 283	11 426	13 287
17	573	1 005	2.903	9.578	338	1 119	1 726	2 387	3 374	4 224	5 179	6 627	7 883	9 298	11 447	13 313
18	588	1 007	2.848	9.328	353	1 138	1 747	2 408	3 392	4 238	5 186	6 620	7 862	9 257	11 371	13 202
19	589	1 006	2.848	9.332	354	1 139	1 747	2 408	3 392	4 237	5 185	6 619	7 861	9 256	11 370	13 201
20	618	1 011	2.731	8.750	384	1 179	1 791	2 453	3 433	4 271	5 206	6 612	7 824	9 179	11 221	12 981
21	559	960	3.137	11.774	331	1 066	1 643	2 277	3 233	4 064	5 005	6 445	7 706	9 137	11 331	13 254
22	565	958	3.141	11.823	338	1 071	1 647	2 280	3 233	4 062	5 001	6 439	7 698	9 127	11 319	13 239
23	566	957	3.139	11.819	339	1 072	1 648	2 281	3 234	4 063	5 001	6 438	7 696	9 124	11 314	13 232
24	567	957	3.139	11.820	339	1 072	1 648	2 281	3 234	4 063	5 001	6 438	7 696	9 124	11 314	13 232
25	557	960	3.135	11.760	330	1 065	1 643	2 277	3 233	4 064	5 006	6 446	7 707	9 139	11 333	13 255
26	555	961	3.130	11.729	327	1 064	1 642	2 277	3 234	4 066	5 007	6 448	7 709	9 141	11 335	13 257
27	542	956	3.219	12.259	314	1 043	1 617	2 250	3 205	4 039	4 986	6 438	7 715	9 167	11 400	13 362
28	539	957	3.212	12.210	311	1 041	1 616	2 249	3 206	4 041	4 988	6 442	7 718	9 171	11 403	13 365
29	534	958	3.216	12.214	306	1 036	1 612	2 246	3 205	4 040	4 989	6 445	7 724	9 180	11 417	13 383
30	536	958	3.218	12.229	308	1 037	1 613	2 247	3 204	4 040	4 988	6 444	7 722	9 177	11 414	13 380
31	529	958	3.236	12.304	300	1 030	1 606	2 240	3 199	4 036	4 986	6 447	7 730	9 192	11 440	13 417
32	545	956	3.207	12.195	317	1 047	1 621	2 254	3 209	4 042	4 988	6 438	7 712	9 161	11 387	13 343
33	493	914	3.652	15.717	271	951	1 497	2 104	3 035	3 859	4 806	6 278	7 590	9 100	11 455	13 553
34	504	914	3.615	15.530	283	964	1 510	2 117	3 047	3 869	4 811	6 278	7 582	9 083	11 420	13 499
35	503	915	3.611	15.498	282	964	1 509	2 117	3 048	3 870	4 813	6 279	7 584	9 084	11 421	13 500
36	503	914	3.612	15.508	282	964	1 510	2 117	3 047	3 869	4 812	6 279	7 583	9 083	11 420	13 499
37	503	914	3.612	15.510	282	964	1 510	2 117	3 047	3 869	4 812	6 279	7 583	9 083	11 420	13 499
38	511	912	3.613	15.545	290	970	1 515	2 121	3 049	3 869	4 810	6 274	7 575	9 072	11 404	13 479
39	515	911	3.618	15.598	295	974	1 517	2 122	3 049	3 868	4 808	6 270	7 570	9 066	11 396	13 470
40	469	871	4.110	19.770	256	890	1 406	1 986	2 887	3 693	4 628	6 101	7 428	8 972	11 407	13 601
41	469	871	4.110	19.768	256	890	1 406	1 986	2 887	3 693	4 628	6 101	7 428	8 972	11 407	13 601

Sub set	FFA moments				Annual exceedance probability (%)											
	ave	sdev	skew	kurt	50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
	Flow in m^3/s				Flood peaks in m^3/s											
42	473	871	4.107	19.765	260	894	1 408	1 989	2 888	3 694	4 628	6 100	7 426	8 968	11 400	13 592
43	479	869	4.107	19.788	266	899	1 413	1 992	2 891	3 695	4 628	6 097	7 420	8 959	11 387	13 575
44	484	867	4.116	19.872	272	903	1 416	1 994	2 891	3 693	4 625	6 092	7 414	8 952	11 379	13 566
45	486	867	4.123	19.922	274	904	1 416	1 994	2 890	3 692	4 623	6 090	7 412	8 950	11 377	13 564
46	482	868	4.115	19.857	270	901	1 415	1 993	2 890	3 693	4 625	6 093	7 416	8 955	11 382	13 570
47	482	868	4.114	19.844	269	901	1 414	1 993	2 890	3 694	4 626	6 094	7 417	8 956	11 383	13 571
48	485	867	4.123	19.922	273	904	1 416	1 993	2 890	3 692	4 623	6 090	7 412	8 951	11 378	13 566
49	484	867	4.120	19.892	271	902	1 415	1 993	2 890	3 693	4 624	6 092	7 414	8 953	11 380	13 568
50	365	473	2.538	6.607	257	633	921	1 229	1 680	2 062	2 485	3 115	3 652	4 249	5 139	5 898
51	375	471	2.525	6.593	268	642	928	1 234	1 682	2 061	2 480	3 105	3 637	4 227	5 108	5 859
52	376	470	2.530	6.617	269	643	928	1 234	1 681	2 060	2 480	3 104	3 636	4 227	5 108	5 860
53	330	394	2.811	9.616	238	546	784	1 043	1 427	1 756	2 125	2 682	3 163	3 704	4 521	5 227
54	332	392	2.830	9.721	240	547	784	1 042	1 425	1 754	2 123	2 680	3 162	3 704	4 523	5 233
55	328	395	2.806	9.572	236	545	783	1 042	1 427	1 757	2 126	2 684	3 165	3 706	4 523	5 230
Average descriptive statistics																
min1	328	392	2.436	5.271	236	545	783	1 042	1 425	1 754	2 123	2 680	3 162	3 704	4 521	5 227
max1	618	1 058	4.123	19.922	384	1 179	1 808	2 502	3 538	4 431	5 440	6 975	8 310	9 815	12 105	14 097
AVE1	513	873	3.212	12.125	306	973	1 497	2 073	2 942	3 700	4 559	5 877	7 034	8 351	10 377	12 158
SD1					38	157	257	367	537	688	864	1 143	1 397	1 697	2 177	2 619
COV1					0.12	0.16	0.17	0.18	0.18	0.19	0.19	0.19	0.20	0.20	0.21	0.22

APPENDIX 4-H: Selected R-W77 results at N2R001 re LP3

The R-W77 FFA results, from where the LP3 distribution met the set criteria for RL1, at flow site N2R001.

Sub set	FFA moments				Annual exceedance probability (%)											
	ave	sdev	skew	kurt	50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
	(log-transformed)				Flood peaks in m³/s											
1	2.367	0.538	0.332	0.225	218	645	1 184	1 997	3 680	5 609	8 330	13 623	19 395	27 246	41 969	57 536
2	2.369	0.538	0.324	0.237	219	647	1 185	1 995	3 665	5 573	8 255	13 455	19 107	26 772	41 097	56 194
3	2.372	0.538	0.308	0.212	221	653	1 194	2 004	3 666	5 554	8 197	13 292	18 800	26 235	40 050	54 529
4	2.373	0.538	0.300	0.207	222	656	1 197	2 006	3 660	5 533	8 148	13 172	18 586	25 874	39 371	53 471
5	2.379	0.537	0.273	0.230	226	664	1 205	2 003	3 618	5 426	7 925	12 672	17 730	24 472	36 816	49 570
6	2.381	0.536	0.267	0.241	227	666	1 206	2 001	3 605	5 396	7 866	12 543	17 513	24 123	36 190	48 625
7	2.371	0.525	0.258	0.392	223	637	1 138	1 865	3 308	4 899	7 068	11 126	15 388	21 002	31 137	41 470
8	2.372	0.525	0.249	0.375	224	641	1 143	1 870	3 311	4 894	7 047	11 061	15 267	20 790	30 731	40 837
9	2.367	0.522	0.274	0.461	220	628	1 120	1 837	3 265	4 844	7 003	11 057	15 333	20 982	31 224	41 709
10	2.366	0.522	0.279	0.453	219	626	1 119	1 838	3 274	4 864	7 044	11 145	15 480	21 219	31 645	42 342
11	2.345	0.506	0.249	0.545	211	580	1 012	1 626	2 818	4 106	5 832	9 003	12 278	16 529	24 081	31 664
12	2.341	0.506	0.272	0.540	208	574	1 006	1 625	2 838	4 159	5 944	9 252	12 698	17 205	25 284	33 465
13	2.341	0.506	0.269	0.539	208	575	1 007	1 626	2 836	4 152	5 930	9 220	12 643	17 117	25 124	33 224
14	2.338	0.507	0.286	0.526	206	571	1 003	1 627	2 854	4 198	6 023	9 424	12 986	17 667	26 103	34 692
15	2.329	0.508	0.331	0.533	200	557	988	1 618	2 880	4 285	6 220	9 889	13 795	19 005	28 555	38 442
16	2.335	0.509	0.297	0.487	204	568	1 002	1 631	2 878	4 250	6 124	9 633	13 329	18 209	27 051	36 101
17	2.328	0.508	0.337	0.519	199	556	988	1 620	2 891	4 310	6 270	9 996	13 976	19 295	29 074	39 226
18	2.320	0.509	0.374	0.516	194	546	977	1 617	2 920	4 396	6 458	10 436	14 743	20 570	31 437	42 876
19	2.317	0.510	0.389	0.503	192	542	974	1 618	2 938	4 441	6 555	10 656	15 123	21 200	32 604	44 681
20	2.309	0.517	0.379	0.398	189	540	976	1 629	2 974	4 511	6 676	10 889	15 487	21 752	33 529	46 018
Average descriptive statistics																
min1	2.309	0.506	0.249	0.207	189	540	974	1 617	2 818	4 106	5 832	9 003	12 278	16 529	24 081	31 664
max1	2.381	0.538	0.389	0.545	227	666	1 206	2 006	3 680	5 609	8 330	13 623	19 395	27 246	41 969	57 536
AVE1	2.351	0.520	0.302	0.407	212	604	1 081	1 783	3 194	4 770	6 946	11 077	15 483	21 363	32 154	43 334
SD1					12	46	94	172	346	561	881	1 543	2 306	3 394	5 537	7 908
COV1					0.06	0.08	0.09	0.10	0.11	0.12	0.13	0.14	0.15	0.16	0.17	0.18

APPENDIX 4-I: 7ym R-W results at N2R001, for GEV

A comparison of all the 7ym R-W average results at N2R001, applying the GEV distribution. N2R001 was chosen as an example, to illustrate a case where the results of the GEV and LP3 were quite different – see also Appendix 4-K for a graphic comparison.

7ym R-W (yrs)	Stats	AEP (%)											
		50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
		GEV distribution: Flood peaks in m^3/s											
7	min1	118	193	229	259	294	309	320	331	337	342	347	350
	max1	999	2 621	3 755	4 890	6 430	7 740	9 304	11 863	14 057	16 501	20 168	23 314
	AVE1	353	890	1 277	1 677	2 240	2 700	3 195	3 911	4 505	5 149	6 088	6 872
	SD1	171	566	875	1 205	1 685	2 089	2 534	3 194	3 755	4 377	5 304	6 095
	COV1	0.49	0.64	0.69	0.72	0.75	0.77	0.79	0.82	0.83	0.85	0.87	0.89
14	min1	146	292	369	432	488	521	549	580	599	616	635	647
	max1	616	1 755	2 608	3 511	4 815	5 909	7 184	9 106	10 885	12 895	16 005	18 723
	AVE1	332	898	1 319	1 763	2 404	2 940	3 526	4 392	5 123	5 930	7 126	8 142
	SD1	103	414	670	951	1 373	1 738	2 149	2 772	3 313	3 922	4 846	5 649
	COV1	0.31	0.46	0.51	0.54	0.57	0.59	0.61	0.63	0.65	0.66	0.68	0.69
21	min1	183	300	372	436	513	567	618	680	724	766	817	853
	max1	484	1 393	2 141	2 946	4 154	5 184	6 330	8 048	9 592	11 378	14 122	16 563
	AVE1	325	911	1 354	1 826	2 516	3 097	3 739	4 695	5 510	6 414	7 765	8 919
	SD1	69	342	569	823	1 209	1 548	1 934	2 528	3 050	3 645	4 557	5 358
	COV1	0.21	0.38	0.42	0.45	0.48	0.50	0.52	0.54	0.55	0.57	0.59	0.60
28	min1	201	362	474	585	735	852	973	1 139	1 252	1 361	1 506	1 616
	max1	466	1 273	1 963	2 707	3 807	4 746	5 793	7 390	8 766	10 388	12 950	15 248
	AVE1	318	924	1 390	1 893	2 635	3 269	3 976	5 040	5 958	6 986	8 538	9 879
	SD1	55	286	482	702	1 041	1 343	1 690	2 231	2 713	3 268	4 131	4 899
	COV1	0.17	0.31	0.35	0.37	0.40	0.41	0.43	0.44	0.46	0.47	0.48	0.50
35	min1	207	375	489	600	746	858	971	1 123	1 240	1 358	1 518	1 641
	max1	423	1 167	1 807	2 503	3 542	4 438	5 443	6 967	8 288	9 828	12 228	14 338
	AVE1	312	947	1 441	1 979	2 784	3 479	4 262	5 451	6 487	7 657	9 441	10 996
	SD1	45	230	386	561	832	1 074	1 353	1 793	2 188	2 646	3 367	4 016
	COV1	0.14	0.24	0.27	0.28	0.30	0.31	0.32	0.33	0.34	0.35	0.36	0.37
42	min1	236	545	783	1 042	1 425	1 754	2 123	2 680	3 162	3 704	4 521	5 227
	max1	384	1 179	1 808	2 502	3 538	4 431	5 440	6 975	8 310	9 815	12 105	14 097
	AVE1	306	973	1 497	2 073	2 942	3 700	4 559	5 877	7 034	8 351	10 377	12 158
	SD1	38	157	257	367	537	688	864	1 143	1 397	1 697	2 177	2 619
	COV1	0.12	0.16	0.17	0.18	0.18	0.19	0.19	0.19	0.20	0.20	0.21	0.22
49	min1	233	816	1 293	1 834	2 679	3 441	4 330	5 739	7 016	8 511	10 887	12 733
	max1	377	1 192	1 819	2 496	3 498	4 353	5 307	6 740	7 980	9 415	11 597	13 492
	AVE1	304	992	1 537	2 139	3 053	3 855	4 769	6 180	7 426	8 852	11 057	13 009
	SD1	47	124	174	220	273	305	328	338	326	294	224	191
	COV1	0.15	0.12	0.11	0.10	0.09	0.08	0.07	0.05	0.04	0.03	0.02	0.01

7ym R-W (yrs)	Stats	AEP (%)											
		50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
		GEV distribution: Flood peaks in m³/s											
56	min1	211	754	1 201	1 711	2 512	3 237	4 088	5 445	6 683	8 139	10 465	12 415
	max1	405	1 199	1 806	2 458	3 418	4 234	5 139	6 519	7 717	9 058	11 079	12 823
	AVE1	302	966	1 493	2 075	2 958	3 733	4 617	5 982	7 187	8 567	10 702	12 591
	SD1	50	126	174	219	269	300	320	328	313	277	189	115
	COV1	0.17	0.13	0.12	0.11	0.09	0.08	0.07	0.05	0.04	0.03	0.02	0.01
63	min1	215	761	1 208	1 715	2 507	3 221	4 055	5 378	6 579	7 985	10 220	12 160
	max1	371	1 117	1 694	2 319	3 246	4 041	4 933	6 279	7 443	8 770	10 838	12 642
	AVE1	298	941	1 452	2 017	2 877	3 632	4 493	5 825	7 003	8 352	10 441	12 292
	SD1	44	109	152	191	237	267	288	302	296	273	212	149
	COV1	0.15	0.12	0.10	0.09	0.08	0.07	0.06	0.05	0.04	0.03	0.02	0.01
70	min1	233	786	1 236	1 743	2 531	3 237	4 057	5 350	6 517	7 866	9 980	11 883
	max1	355	1 060	1 610	2 209	3 104	3 877	4 748	6 080	7 238	8 547	10 584	12 392
	AVE1	294	919	1 416	1 968	2 807	3 546	4 390	5 696	6 853	8 179	10 236	12 059
	SD1	35	93	132	168	212	242	266	286	290	280	240	187
	COV1	0.12	0.10	0.09	0.09	0.08	0.07	0.06	0.05	0.04	0.03	0.02	0.02
77	min1	258	833	1 295	1 812	2 607	3 311	4 121	5 387	6 518	7 825	9 871	11 701
	max1	329	999	1 526	2 103	2 972	3 728	4 584	5 893	7 037	8 336	10 368	12 188
	AVE1	292	908	1 399	1 943	2 771	3 500	4 334	5 624	6 766	8 076	10 107	11 908
	SD1	28	73	105	134	171	197	219	241	249	249	232	203
	COV1	0.09	0.08	0.07	0.07	0.06	0.06	0.05	0.04	0.04	0.03	0.02	0.02
84	min1	251	798	1 240	1 737	2 504	3 188	3 978	5 220	6 335	7 628	9 660	11 486
	max1	312	950	1 455	2 012	2 855	3 593	4 432	5 721	6 856	8 150	10 180	12 009
	AVE1	289	899	1 385	1 924	2 746	3 470	4 298	5 580	6 716	8 019	10 040	11 833
	SD1	23	61	87	112	144	166	186	207	218	222	215	201
	COV1	0.08	0.07	0.06	0.06	0.05	0.05	0.04	0.04	0.03	0.03	0.02	0.02
91	min1	288	901	1 388	1 928	2 750	3 472	4 296	5 569	6 694	7 983	9 979	11 745
	max1	292	904	1 391	1 930	2 752	3 474	4 298	5 572	6 698	7 989	9 987	11 756
	AVE1	289	902	1 389	1 929	2 751	3 473	4 297	5 571	6 697	7 987	9 983	11 749
	SD1	1.5	1.1	0.9	0.7	0.6	0.7	0.9	1.2	1.5	2.0	2.8	3.7
	COV1	0.0052	0.0013	0.0007	0.0004	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0003	0.0003

APPENDIX 4-J: 7ym R-W results at N2R001, for LP3

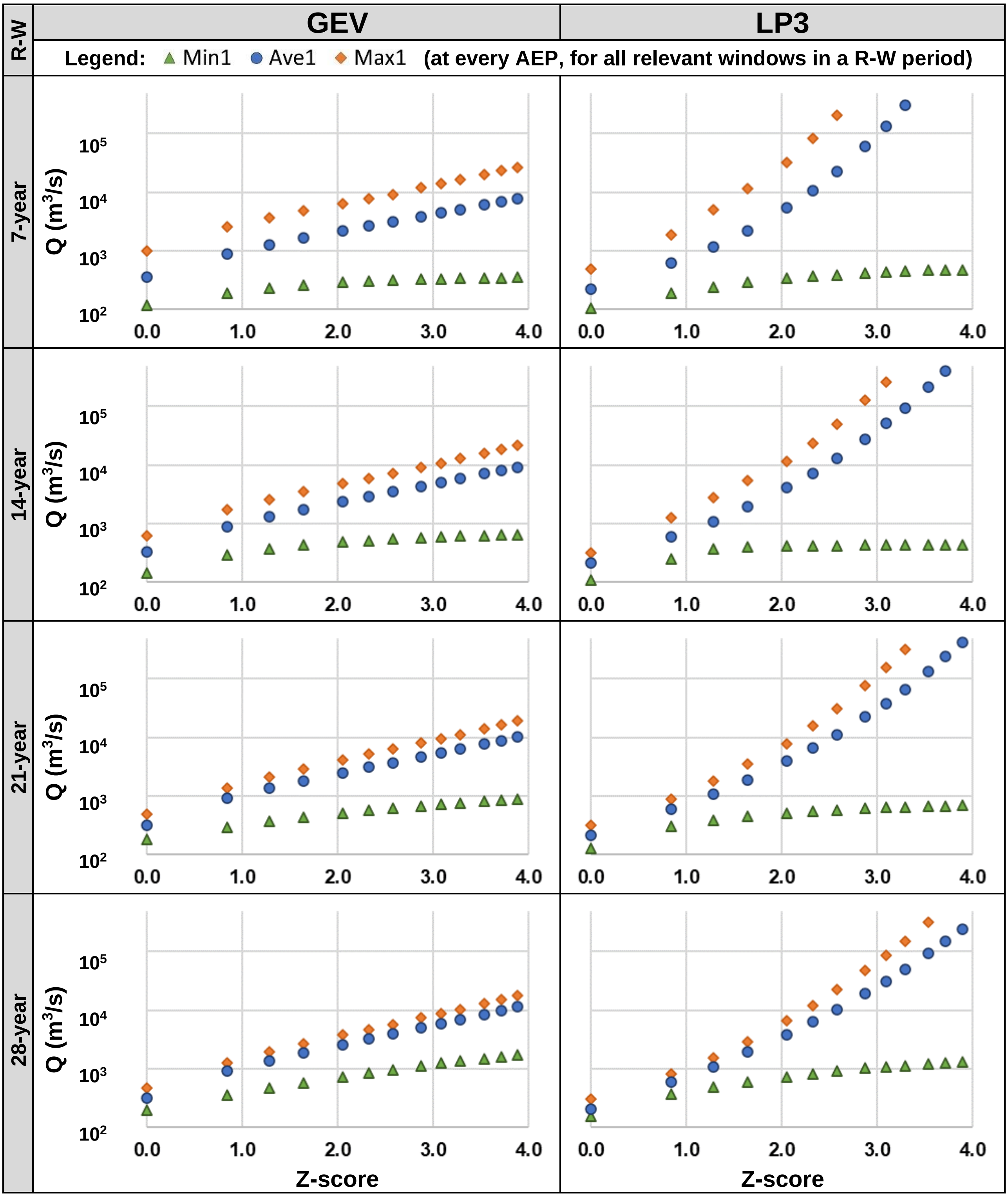
A comparison of all the 7ym R-W average results at N2R001, applying the LP3 distribution. N2R001 was chosen as an example, to illustrate a case where the results of the GEV and LP3 were quite different – see also Appendix 4-K for a graphic comparison.

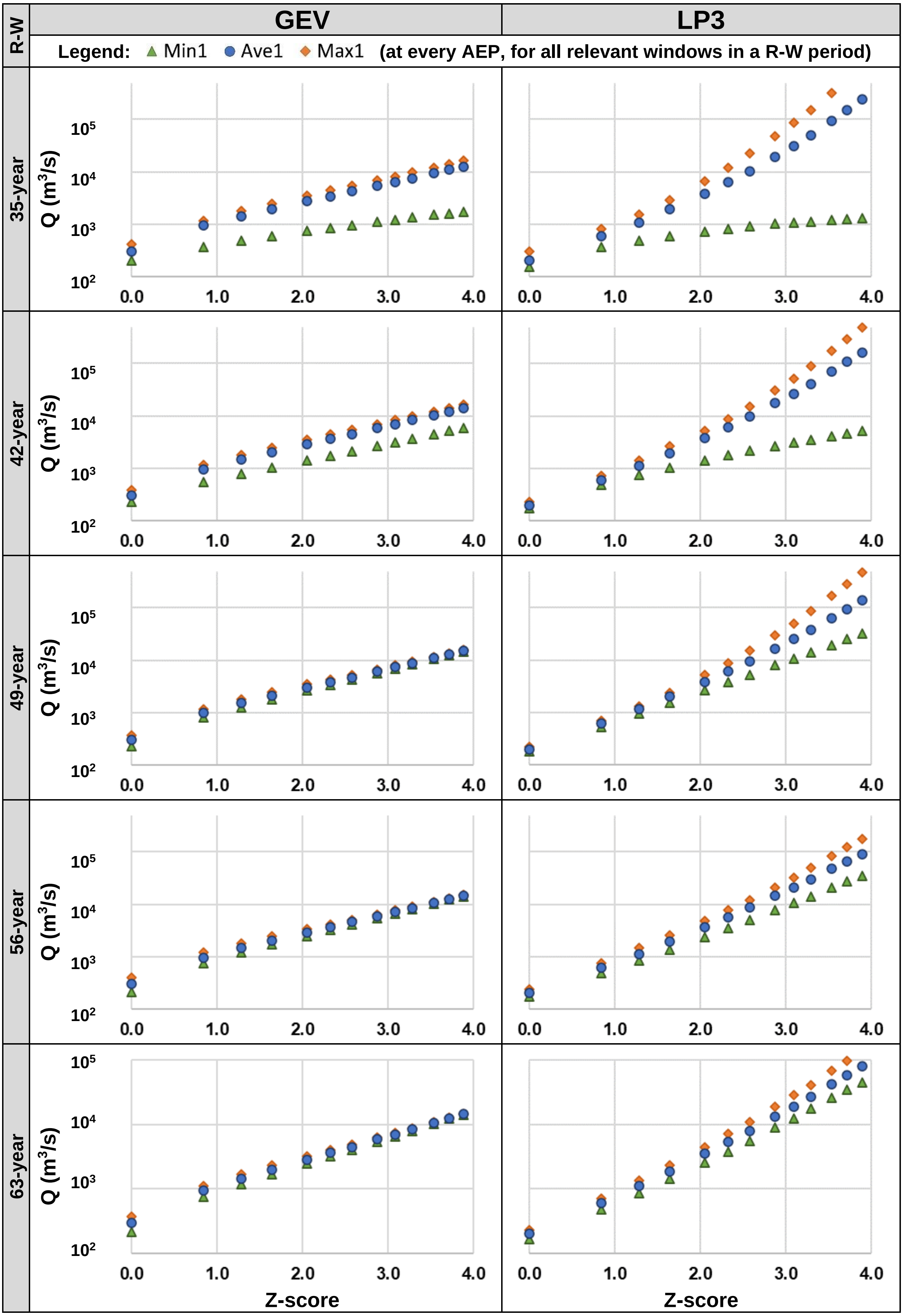
7ym R-W (yrs)	**Stats**	**AEP (%)**											
		50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
		LP3 distribution: Flood peaks in m³/s											
7	min1	105	190	242	294	341	368	391	418	436	452	469	469
	max1	485	1 891	5 107	11 796	31 839	82 029	204 985	663 482	1 628 620	4 669 930	18 774 600	53 725 500
	AVE1	224	612	1 180	2 252	5 419	10 862	22 451	61 470	136 277	310 438	957 048	2 300 094
	SD1	85	335	854	2 079	6 559	15 672	37 916	125 050	314 903	807 469	2 878 406	7 662 820
	COV1	0.38	0.55	0.72	0.92	1.21	1.44	1.69	2.03	2.31	2.60	3.01	3.33
14	min1	110	247	368	407	421	425	428	429	430	431	431	431
	max1	320	1 277	2 841	5 487	11 789	23 811	49 250	127 394	262 208	589 223	1 712 700	3 829 630
	AVE1	218	589	1 095	1 959	4 151	7 324	13 005	28 205	51 357	94 700	216 806	411 403
	SD1	54	207	532	1 201	3 197	6 509	13 181	33 690	69 073	142 651	375 674	786 066
	COV1	0.25	0.35	0.49	0.61	0.77	0.89	1.01	1.19	1.34	1.51	1.73	1.91
21	min1	127	309	385	447	511	550	582	616	637	654	672	682
	max1	318	885	1 849	3 571	7 938	15 850	31 770	78 078	157 284	325 632	846 233	1 734 470
	AVE1	213	590	1 090	1 920	3 924	6 666	11 286	22 671	38 603	66 142	136 404	238 271
	SD1	40	137	381	886	2 321	4 554	8 782	20 871	40 420	78 792	191 816	377 094
	COV1	0.19	0.23	0.35	0.46	0.59	0.68	0.78	0.92	1.05	1.19	1.41	1.58
28	min1	154	369	490	603	740	834	922	1 028	1 094	1 151	1 217	1 262
	max1	309	810	1 575	2 928	6 677	12 165	22 331	48 639	86 356	151 714	315 200	543 096
	AVE1	208	592	1 101	1 928	3 857	6 380	10 430	19 734	31 763	50 926	94 657	150 954
	SD1	35	107	300	701	1 810	3 439	6 293	13 503	23 668	41 116	84 545	145 109
	COV1	0.17	0.18	0.27	0.36	0.47	0.54	0.60	0.68	0.75	0.81	0.89	0.96
35	min1	163	372	510	646	825	959	1 090	1 261	1 386	1 509	1 667	1 782
	max1	264	720	1 409	2 594	5 566	10 109	18 002	37 716	65 069	111 682	225 747	381 433
	AVE1	204	598	1 121	1 962	3 877	6 316	10 131	18 610	29 216	45 613	81 677	126 470
	SD1	25	78	228	544	1 417	2 686	4 871	10 254	17 638	29 996	59 817	100 203
	COV1	0.12	0.13	0.20	0.28	0.37	0.43	0.48	0.55	0.60	0.66	0.73	0.79
42	min1	175	492	752	1 035	1 454	1 802	2 176	2 708	3 137	3 587	4 213	4 710
	max1	231	725	1 450	2 642	5 358	8 755	15 234	31 141	52 928	89 292	176 532	293 720
	AVE1	200	606	1 143	1 998	3 904	6 270	9 883	17 670	27 125	41 357	71 687	108 237
	SD1	14	45	144	361	989	1 927	3 566	7 624	13 172	22 374	44 216	73 224
	COV1	0.07	0.07	0.13	0.18	0.25	0.31	0.36	0.43	0.49	0.54	0.62	0.68
49	min1	180	537	961	1 565	2 680	3 839	5 349	8 025	10 692	14 045	19 778	25 320
	max1	221	690	1 345	2 435	5 179	8 950	15 181	29 861	50 584	85 450	169 320	282 301
	AVE1	200	614	1 161	2 024	3 914	6 213	9 646	16 838	25 321	37 757	63 429	93 444
	SD1	13	46	123	290	788	1 558	2 935	6 428	11 284	19 419	38 864	64 762
	COV1	0.06	0.08	0.11	0.14	0.20	0.25	0.30	0.38	0.45	0.51	0.61	0.69

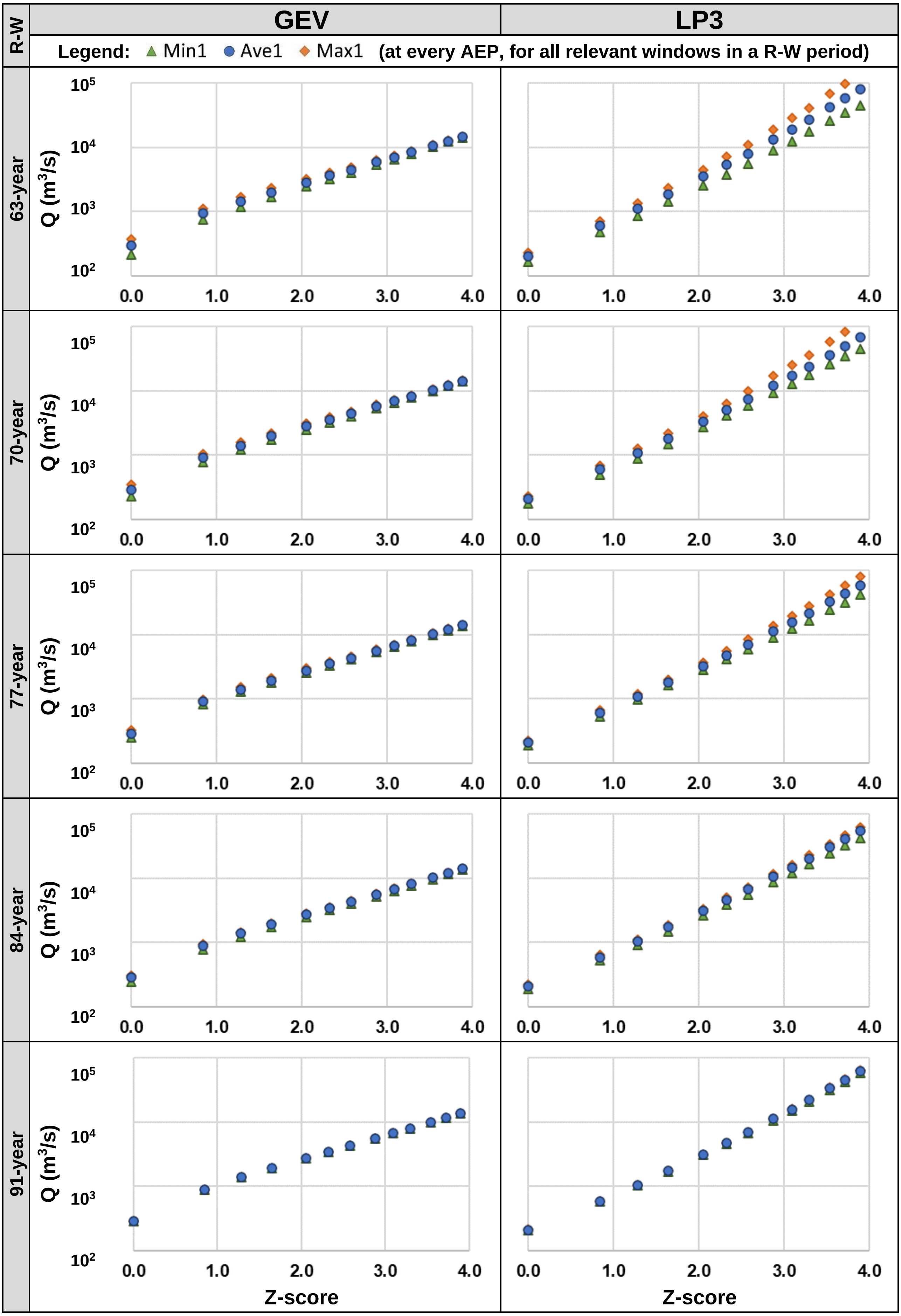
7ym R-W (yrs)	Stats	AEP (%)											
		50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
		LP3 distribution: Flood peaks in m³/s											
56	min1	175	492	870	1 400	2 429	3 541	5 033	7 773	10 603	14 277	20 802	27 352
	max1	245	770	1 464	2 550	4 884	7 680	11 986	21 345	32 590	49 150	83 257	122 745
	AVE1	203	616	1 147	1 964	3 694	5 721	8 636	14 447	20 944	29 983	47 410	66 359
	SD1	17	66	150	304	695	1 226	2 081	4 002	6 383	9 981	17 566	26 495
	COV1	0.08	0.11	0.13	0.15	0.19	0.21	0.24	0.28	0.30	0.33	0.37	0.40
63	min1	170	479	861	1 423	2 547	3 811	5 566	8 926	12 538	17 398	26 188	34 390
	max1	236	714	1 347	2 332	4 461	7 075	10 938	18 875	28 001	41 000	66 732	95 400
	AVE1	206	610	1 121	1 896	3 510	5 371	8 014	13 201	18 913	26 756	41 639	57 575
	SD1	16	61	133	262	582	1 009	1 688	3 193	5 032	7 777	13 481	20 099
	COV1	0.08	0.10	0.12	0.14	0.17	0.19	0.21	0.24	0.27	0.29	0.32	0.35
70	min1	180	506	906	1 501	2 716	4 092	5 883	9 221	12 722	17 326	25 632	34 095
	max1	234	700	1 290	2 183	4 058	6 372	9 785	16 751	24 713	35 997	58 214	82 844
	AVE1	209	606	1 098	1 831	3 327	5 022	7 390	11 953	16 891	23 568	36 013	49 112
	SD1	13	51	112	220	483	829	1 375	2 571	4 019	6 162	10 575	15 651
	COV1	0.06	0.08	0.10	0.12	0.15	0.17	0.19	0.22	0.24	0.26	0.29	0.32
77	min1	189	540	974	1 617	2 818	4 106	5 832	9 003	12 278	16 529	24 081	31 664
	max1	227	666	1 206	2 006	3 680	5 609	8 330	13 623	19 395	27 246	41 969	57 536
	AVE1	212	604	1 081	1 783	3 194	4 770	6 946	11 077	15 483	21 363	32 154	43 334
	SD1	12	46	94	172	346	561	881	1 543	2 306	3 394	5 537	7 908
	COV1	0.06	0.08	0.09	0.10	0.11	0.12	0.13	0.14	0.15	0.16	0.17	0.18
84	min1	195	533	937	1 516	2 642	3 873	5 544	8 659	11 925	16 224	23 990	31 918
	max1	226	638	1 134	1 857	3 331	4 978	7 254	11 582	16 201	22 371	33 700	45 442
	AVE1	214	597	1 060	1 734	3 084	4 583	6 646	10 546	14 691	20 207	30 296	40 717
	SD1	11	40	78	137	262	405	609	1 006	1 442	2 038	3 161	4 356
	COV1	0.05	0.07	0.07	0.08	0.08	0.09	0.09	0.10	0.10	0.10	0.10	0.11
91	min1	209	587	1 045	1 723	3 088	4 604	6 700	10 692	14 964	20 686	31 231	42 206
	max1	216	597	1 058	1 744	3 150	4 741	6 963	11 242	15 870	22 171	33 947	46 370
	AVE1	212	590	1 052	1 734	3 119	4 682	6 861	11 053	15 581	21 696	33 076	45 030
	SD1	2.4	3.9	4.3	6.8	23	50	98	211	355	575	1 041	1 587
	COV1	0.011	0.007	0.004	0.004	0.007	0.011	0.014	0.019	0.023	0.026	0.031	0.035

APPENDIX 4-K: All 7ym R-W results at N2R001

A visual comparison of all the 7ym R-W average results at N2R001, which was chosen as an example, to illustrate a case where the results of the GEV and LP3 were quite different.




R-W
GEV
LP3
Legend: Min1 Ave1 Max1 (at every AEP, for all relevant windows in a R-W period)
35-year
42-year
49-year
56-year
63-year
Q (m³/s)
10²
10³
10⁴
10⁵
0.0
1.0
2.0
3.0
4.0
Z-score


R-W
GEV
LP3
Legend: Min1 Ave1 Max1 (at every AEP, for all relevant windows in a R-W period)
63-year
70-year
77-year
84-year
91-year
Q (m3/s)
10^5
10^4
10^3
10^2
0.0
1.0
2.0
3.0
4.0
Z-score

APPENDIX 4-L: Combined 7ym R-W average statistics; primary sites

Average statistics results (AVE2) at all primary sites, for 7ym R-Ws that demonstrate adherence to the set criteria for RL2 – all 7ym R-W results portrayed in Appendix 4-M.

Site	GEV / LP3	RL2	Stats	AEP (%)											
				50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
				Flood peaks in m^3/s											
A3R002	GEV	63	Min2	21	64	98	137	197	250	311	406	492	591	747	886
			Max2	22	70	109	154	225	289	365	485	594	722	928	1 115
			AVE2	21	67	105	148	214	274	344	456	557	675	863	1 034
			SD2	0.34	1.9	3.5	5.6	9.4	13	18	26	35	45	62	79
			COV2	0.02	0.03	0.03	0.04	0.04	0.05	0.05	0.06	0.06	0.07	0.07	0.08
			Full FFA	22	67	104	146	212	272	343	455	557	677	869	1 044
			Δ%	2.0%	0.8%	1.0%	1.0%	0.9%	0.7%	0.5%	0.2%	0.1%	0.3%	0.7%	1.0%
	LP3	91	Min2	16	52	93	149	249	347	468	667	851	1 068	1 409	1 713
			Max2	16	52	95	154	261	368	502	726	936	1 186	1 586	1 948
			AVE2	16	52	94	151	254	357	483	693	889	1 121	1 488	1 819
			SD2	0.15	0.15	0.94	2.5	6.2	11	17	30	43	60	90	119
			COV2	0.01	0.003	0.01	0.02	0.02	0.03	0.04	0.04	0.05	0.05	0.06	0.07
			Full FFA	16	53	94	148	243	335	445	624	785	970	1 257	1 507
			Δ%	2.1%	0.9%	0.6%	2.2%	4.6%	6.5%	8.5%	11.2%	13.3%	15.5%	18.4%	20.7%
B1R001	GEV	21	Min2	193	495	730	976	1 310	1 585	1 885	2 322	2 688	3 089	3 679	4 176
			Max2	215	540	793	1 069	1 479	1 832	2 232	2 840	3 367	3 960	4 861	5 644
			AVE2	202	523	769	1 034	1 425	1 758	2 129	2 685	3 164	3 698	4 503	5 196
			SD2	6.4	13	19	28	46	64	88	129	170	219	302	379
			COV2	0.03	0.02	0.03	0.03	0.03	0.04	0.04	0.05	0.05	0.06	0.07	0.07
			Full FFA	189	485	716	970	1 349	1 678	2 050	2 616	3 109	3 668	4 520	5 264
			Δ%	7.0%	7.9%	7.4%	6.7%	5.6%	4.7%	3.9%	2.7%	1.8%	0.8%	0.4%	1.3%
	LP3	84	Min2	140	414	709	1 090	1 743	2 361	3 099	4 275	5 331	6 543	8 404	10 024
			Max2	142	427	746	1 170	1 917	2 647	3 538	4 999	6 348	7 933	10 434	12 671
			AVE2	141	421	726	1 123	1 810	2 467	3 258	4 530	5 683	7 016	9 085	10 902
			SD2	0.80	5.5	16	35	79	130	201	333	470	644	945	1 235
			COV2	0.01	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.11
			Full FFA	135	403	696	1 080	1 746	2 386	3 158	4 405	5 538	6 852	8 894	10 691
			Δ%	4.3%	4.4%	4.2%	4.0%	3.7%	3.4%	3.2%	2.8%	2.6%	2.4%	2.1%	2.0%
B2R001	GEV	14	Min2	102	249	357	462	609	729	857	1 042	1 195	1 360	1 601	1 802
			Max2	114	261	379	507	696	856	1 034	1 303	1 534	1 791	2 179	2 512
			AVE2	106	256	369	490	667	817	983	1 230	1 442	1 677	2 030	2 333
			SD2	3.6	4.1	6.6	12	22	34	47	71	93	119	161	198
			COV2	0.03	0.02	0.02	0.02	0.03	0.04	0.05	0.06	0.06	0.07	0.08	0.08
			Full FFA	101	249	363	486	668	823	995	1 255	1 478	1 727	2 101	2 423
			Δ%	5.1%	2.6%	1.6%	0.8%	0.1%	0.7%	1.3%	2.0%	2.4%	2.9%	3.4%	3.7%
	LP3	28	Min2	76	206	346	522	825	1 116	1 469	2 044	2 572	3 193	4 173	5 050
			Max2	81	216	364	559	910	1 263	1 708	2 473	3 217	4 141	5 708	7 224
			AVE2	78	211	354	542	874	1 202	1 610	2 293	2 940	3 719	4 991	6 167
			SD2	1.8	3.6	6.0	11	25	43	71	128	193	281	449	629
			COV2	0.02	0.02	0.02	0.02	0.03	0.04	0.04	0.06	0.07	0.08	0.09	0.10
			Full FFA	78	209	341	506	779	1 031	1 325	1 784	2 190	2 649	3 343	3 940
			Δ%	0.4%	1.1%	3.7%	7.0%	12.2%	16.7%	21.5%	28.5%	34.3%	40.4%	49.3%	56.5%

Site	GEV / LP3	RL2	Stats	AEP (%)											
				50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
				Flood peaks in m^3/s											
C1R002	GEV	14	Min2	366	679	902	1 131	1 451	1 705	1 973	2 327	2 604	2 897	3 309	3 642
			Max2	406	744	999	1 263	1 637	1 943	2 270	2 742	3 129	3 547	4 148	4 643
			AVE2	377	698	929	1 166	1 499	1 769	2 057	2 469	2 807	3 170	3 694	4 125
			SD2	12	22	29	39	57	76	99	137	171	211	272	324
			COV2	0.03	0.03	0.03	0.03	0.04	0.04	0.05	0.06	0.06	0.07	0.07	0.08
			Full FFA	406	760	1 015	1 277	1 644	1 941	2 257	2 707	3 073	3 464	4 021	4 476
			Δ%	7.3%	8.1%	8.5%	8.7%	8.8%	8.9%	8.9%	8.8%	8.7%	8.5%	8.1%	7.8%
	LP3	35	Min2	332	648	908	1 196	1 626	1 994	2 398	2 996	3 501	4 057	4 875	5 563
			Max2	352	704	1 005	1 345	1 861	2 306	2 803	3 544	4 174	4 878	6 189	7 387
			AVE2	337	659	930	1 234	1 692	2 089	2 533	3 201	3 776	4 418	5 381	6 206
			SD2	5.9	18	31	47	74	99	129	180	232	303	436	578
			COV2	0.02	0.03	0.03	0.04	0.04	0.05	0.05	0.06	0.06	0.07	0.08	0.09
			Full FFA	361	720	1 025	1 368	1 885	2 329	2 823	3 556	4 176	4 856	5 852	6 684
			Δ%	6.8%	8.4%	9.2%	9.8%	10.2%	10.3%	10.3%	10.0%	9.6%	9.0%	8.1%	7.1%
C2R001	GEV	42	Min2	17	34	46	58	75	89	104	125	143	163	191	215
			Max2	19	38	52	66	85	101	118	142	161	184	216	244
			AVE2	18	35	48	62	80	96	112	136	155	176	207	232
			SD2	0.84	1.7	2.4	3.1	4.0	4.7	5.3	6.2	6.9	7.6	8.7	9.7
			COV2	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.04	0.04	0.04	0.04
			Full FFA	19	37	51	64	83	98	114	137	155	174	202	224
			Δ%	6.3%	5.3%	4.6%	4.0%	3.2%	2.5%	1.7%	0.6%	0.3%	1.2%	2.7%	3.8%
	LP3	56	Min2	14	31	47	66	96	124	158	211	260	317	408	491
			Max2	15	35	52	74	109	141	178	236	289	349	449	538
			AVE2	14	33	50	70	104	134	171	228	280	340	434	517
			SD2	0.70	1.5	2.2	3.1	4.4	5.6	6.8	8.6	10	12	14	18
			COV2	0.05	0.05	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.03	0.03	0.03
			Full FFA	15	34	52	73	106	135	169	222	267	319	396	462
			Δ%	6.1%	5.4%	4.5%	3.5%	2.0%	0.6%	0.8%	3.0%	4.7%	6.7%	9.5%	11.9%
C5R002	GEV	35	Min2	238	1 088	1 798	2 616	3 871	4 940	6 171	8 099	9 826	11 829	14 978	17 811
			Max2	290	1 182	1 967	2 869	4 303	5 616	7 170	9 674	11 981	14 718	19 136	23 208
			AVE2	255	1 149	1 887	2 729	4 054	5 259	6 675	8 943	11 020	13 471	17 404	21 012
			SD2	19	29	46	75	141	215	316	502	693	939	1 370	1 796
			COV2	0.07	0.02	0.02	0.03	0.03	0.04	0.05	0.06	0.06	0.07	0.08	0.09
			Full FFA	239	1 082	1 788	2 603	3 900	5 090	6 503	8 787	10 897	13 407	17 468	21 223
			Δ%	6.9%	6.2%	5.5%	4.8%	4.0%	3.3%	2.6%	1.8%	1.1%	0.5%	0.4%	1.0%
	LP3	84	Min2	225	681	1 173	1 811	2 907	3 949	5 193	7 179	8 952	10 944	13 983	16 612
			Max2	231	703	1 217	1 894	3 066	4 187	5 535	7 705	9 675	11 958	15 511	18 647
			AVE2	228	694	1 202	1 861	2 992	4 065	5 347	7 393	9 231	11 343	14 592	17 424
			SD2	2.5	9.6	21	37	71	107	157	249	345	470	688	904
			COV2	0.01	0.01	0.02	0.02	0.02	0.03	0.03	0.03	0.04	0.04	0.05	0.05
			Full FFA	228	678	1 165	1 794	2 871	3 892	5 107	7 042	8 775	10 760	13 802	16 442
			Δ%	0.2%	2.3%	3.2%	3.7%	4.2%	4.5%	4.7%	5.0%	5.2%	5.4%	5.7%	6.0%

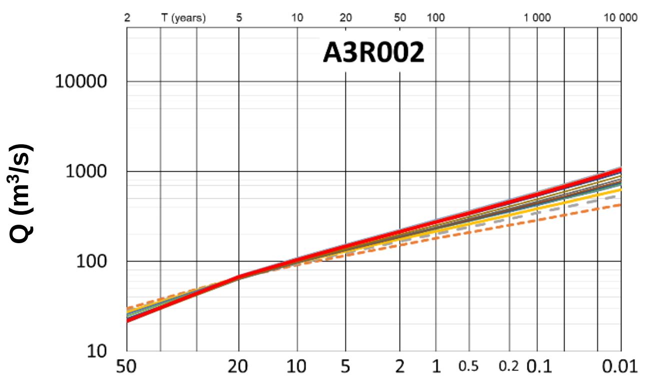
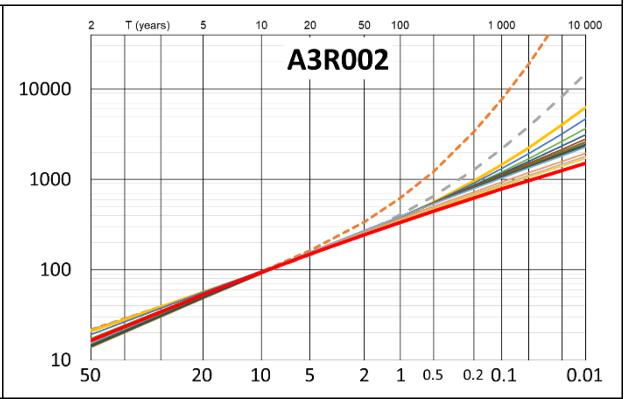
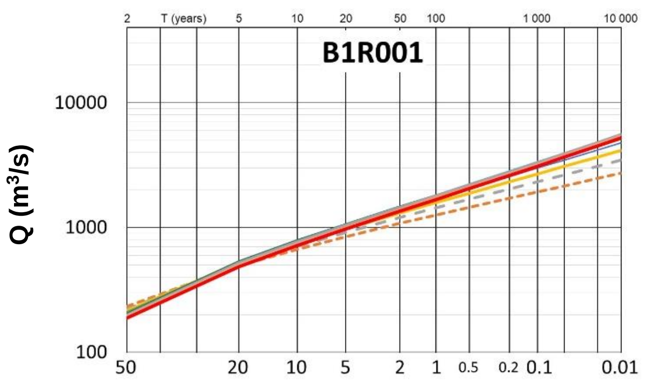
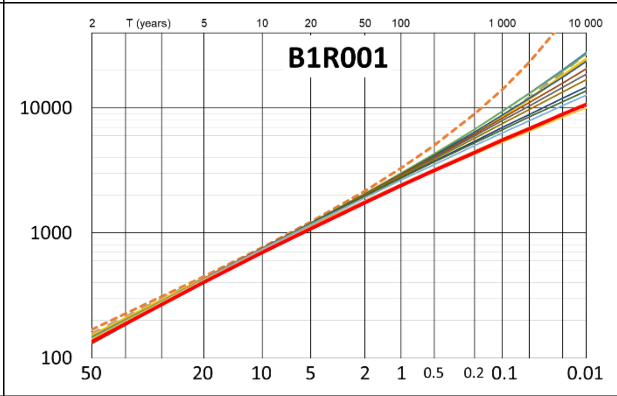
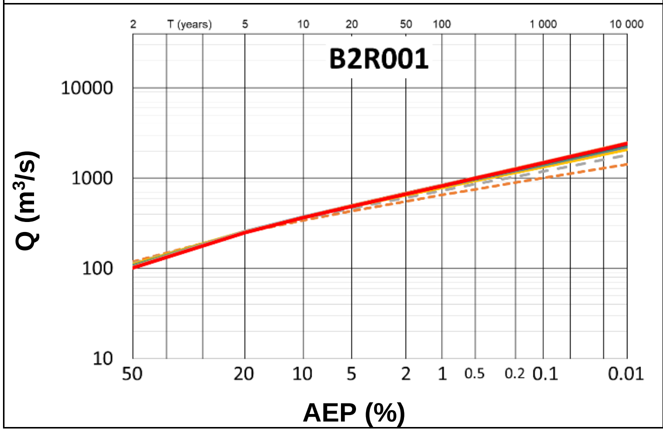
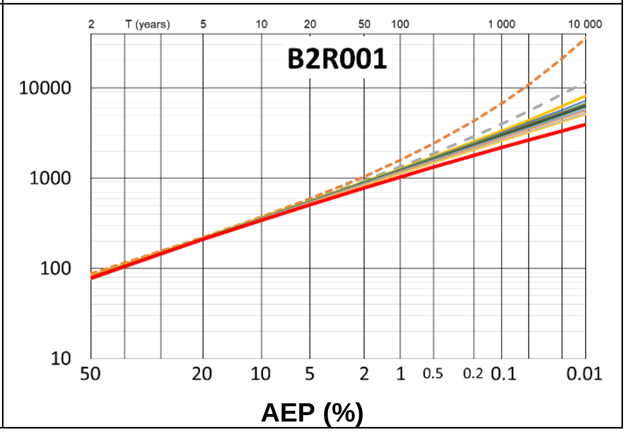
Site	GEV / LP3	RL2	Stats	AEP (%)											
				50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
				Flood peaks in m^3/s											
C9R002	GEV	14	Min2	825	1 826	2 544	3 278	4 289	5 040	5 827	6 931	7 822	8 766	10 108	11 203
			Max2	947	2 013	2 762	3 515	4 542	5 352	6 217	7 502	8 564	9 699	11 319	12 642
			AVE2	869	1 888	2 614	3 352	4 373	5 190	6 053	7 272	8 259	9 307	10 794	12 002
			SD2	41	57	65	73	85	100	120	158	195	240	312	377
			COV2	0.05	0.03	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.03
			Full FFA	978	2 055	2 806	3 556	4 570	5 365	6 187	7 321	8 216	9 145	10 427	11 441
			Δ%	11.2%	8.1%	6.8%	5.7%	4.3%	3.3%	2.2%	0.7%	0.5%	1.8%	3.5%	4.9%
	LP3	42	Min2	590	1 502	2 496	3 839	6 307	8 732	11 685	16 693	21 287	26 754	35 606	43 735
			Max2	687	1 720	2 793	4 180	6 599	9 097	12 536	18 651	24 792	32 590	46 119	59 426
			AVE2	621	1 572	2 593	3 951	6 406	8 894	12 063	17 556	22 938	29 626	40 942	51 804
			SD2	34	76	103	116	103	106	238	648	1 180	1 974	3 582	5 374
			COV2	0.06	0.05	0.04	0.03	0.02	0.01	0.02	0.04	0.05	0.07	0.09	0.10
			Full FFA	708	1 772	2 867	4 268	6 686	9 024	11 877	16 576	20 950	26 141	34 455	42 008
			Δ%	12.4%	11.3%	9.5%	7.4%	4.2%	1.4%	1.6%	5.9%	9.5%	13.3%	18.8%	23.3%
D3R002	GEV	14	Min2	2 277	4 093	5 222	6 303	7 731	8 832	9 962	11 516	12 746	14 032	15 831	17 277
			Max2	2 554	4 406	5 698	7 001	8 825	10 266	11 769	13 876	15 577	17 362	19 859	21 857
			AVE2	2 463	4 288	5 558	6 827	8 549	9 905	11 316	13 278	14 843	16 483	18 779	20 619
			SD2	81	104	142	197	292	379	476	617	734	859	1 034	1 176
			COV2	0.03	0.02	0.03	0.03	0.03	0.04	0.04	0.05	0.05	0.05	0.06	0.06
			Full FFA	2 188	3 954	5 221	6 516	8 313	9 758	11 285	13 444	15 191	17 045	19 669	21 795
			Δ%	12.6%	8.5%	6.5%	4.8%	2.8%	1.5%	0.3%	1.2%	2.3%	3.3%	4.5%	5.4%
	LP3	21	Min2	2 151	4 067	5 465	6 846	8 653	10 004	11 339	13 076	14 367	15 637	17 281	18 498
			Max2	2 387	4 348	5 835	7 366	9 479	11 138	12 862	15 693	18 070	20 679	24 533	27 796
			AVE2	2 299	4 234	5 696	7 197	9 260	10 886	12 573	14 899	16 737	18 645	21 282	23 369
			SD2	74	94	123	157	223	305	427	665	911	1 220	1 739	2 225
			COV2	0.03	0.02	0.02	0.02	0.02	0.03	0.03	0.04	0.05	0.07	0.08	0.10
			Full FFA	2 054	3 894	5 275	6 674	8 558	10 007	11 474	13 434	14 931	16 436	18 435	19 951
			Δ%	11.9%	8.7%	8.0%	7.8%	8.2%	8.8%	9.6%	10.9%	12.1%	13.4%	15.4%	17.1%
D7H005	GEV	14	Min2	1 419	2 825	3 751	4 676	5 936	6 935	7 986	9 468	10 672	11 955	13 794	15 305
			Max2	1 484	2 989	4 099	5 234	6 813	8 086	9 445	11 380	12 960	14 649	17 063	19 041
			AVE2	1 451	2 916	3 959	5 019	6 487	7 664	8 908	10 668	12 095	13 612	15 770	17 527
			SD2	20	61	116	181	279	363	455	589	700	819	990	1 130
			COV2	0.01	0.02	0.03	0.04	0.04	0.05	0.05	0.06	0.06	0.06	0.06	0.06
			Full FFA	1 408	2 859	3 892	4 943	6 393	7 551	8 770	10 483	11 862	13 317	15 366	17 016
			Δ%	3.0%	2.0%	1.7%	1.5%	1.5%	1.5%	1.6%	1.8%	2.0%	2.2%	2.6%	3.0%
	LP3	70	Min2	1 132	2 669	4 094	5 769	8 393	10 707	13 319	17 254	20 613	24 319	29 772	34 337
			Max2	1 148	2 693	4 131	5 856	8 609	11 087	13 942	18 353	22 221	26 595	33 231	38 964
			AVE2	1 140	2 682	4 115	5 807	8 475	10 847	13 545	17 655	21 204	25 163	31 073	36 093
			SD2	8.2	12	19	44	117	209	344	607	884	1 246	1 882	2 507
			COV2	0.01	0.005	0.005	0.01	0.01	0.02	0.03	0.03	0.04	0.05	0.06	0.07
			Full FFA	1 111	2 658	4 107	5 822	8 524	10 918	13 632	17 736	21 253	25 143	30 885	35 705
			Δ%	2.6%	0.9%	0.2%	0.3%	0.6%	0.7%	0.6%	0.5%	0.2%	0.1%	0.6%	1.1%

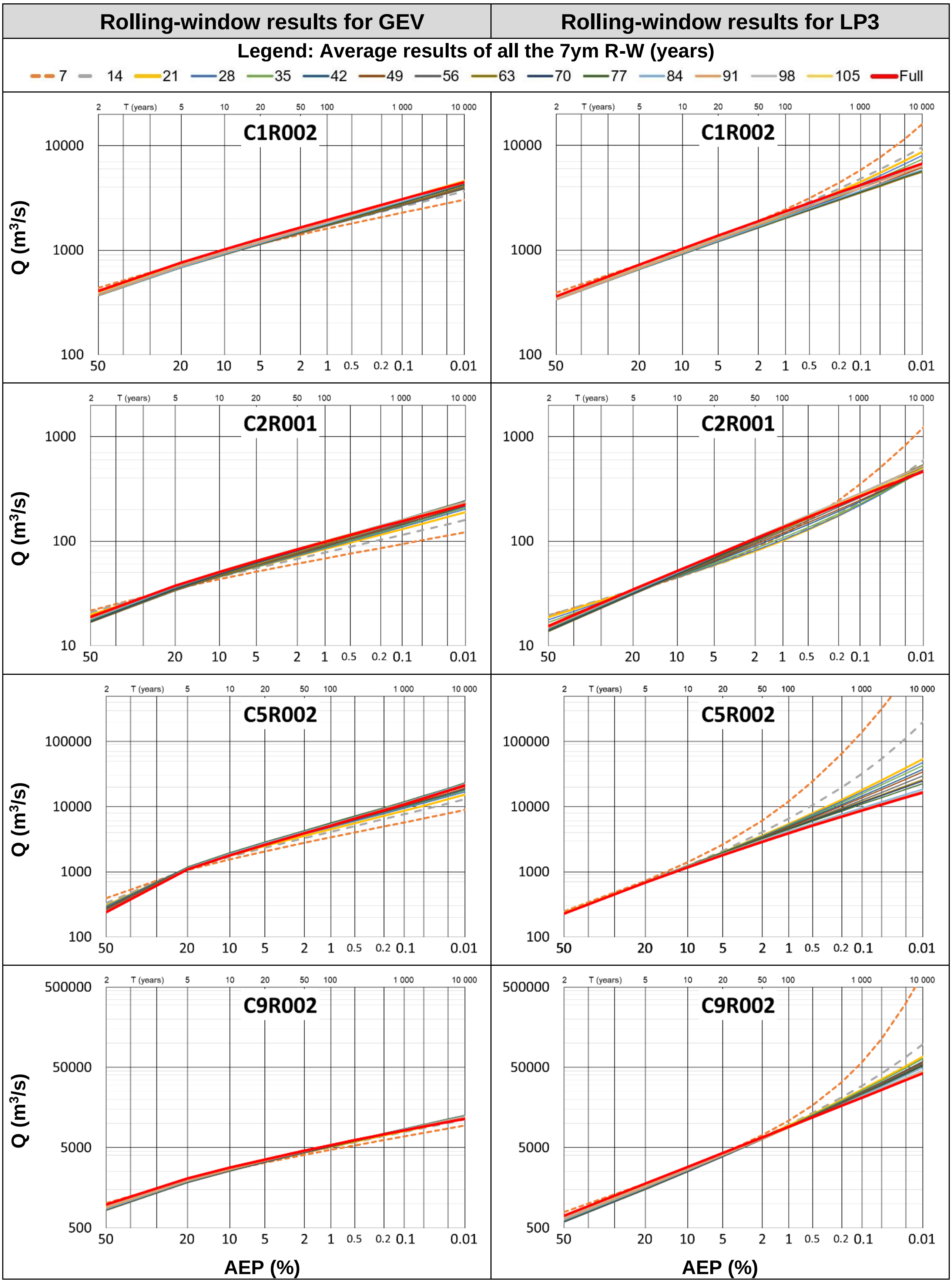
Site	GEV / LP3	RL2	Stats	AEP (%)											
				50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
				Flood peaks in m^3/s											
J1R003	GEV	35	Min2	116	571	949	1 388	2 059	2 645	3 326	4 403	5 378	6 518	8 328	9 972
			Max2	147	614	1 032	1 518	2 298	3 018	3 878	5 277	6 578	8 135	10 673	13 035
			AVE2	127	597	990	1 442	2 162	2 823	3 607	4 876	6 051	7 450	9 719	11 823
			SD2	10	17	28	45	78	116	168	267	370	505	746	989
			COV2	0.08	0.03	0.03	0.03	0.04	0.04	0.05	0.05	0.06	0.07	0.08	0.08
			Full FFA	132	586	968	1 409	2 115	2 766	3 541	4 798	5 964	7 355	9 616	11 714
			Δ%	4.2%	1.9%	2.3%	2.3%	2.2%	2.1%	1.9%	1.6%	1.5%	1.3%	1.1%	0.9%
	LP3	63	Min2	110	368	660	1 037	1 678	2 279	2 988	4 099	5 080	6 187	7 856	9 281
			Max2	114	389	701	1 110	1 807	2 490	3 339	4 724	5 996	7 485	9 829	11 923
			AVE2	112	376	675	1 067	1 743	2 384	3 146	4 351	5 423	6 645	8 503	10 107
			SD2	1.7	8.2	17	29	55	87	136	235	346	495	768	1 045
			COV2	0.02	0.02	0.02	0.03	0.03	0.04	0.04	0.05	0.06	0.07	0.09	0.10
			Full FFA	114	390	703	1 110	1 800	2 442	3 190	4 344	5 343	6 451	8 085	9 448
			Δ%	1.2%	3.6%	4.0%	3.9%	3.2%	2.4%	1.4%	0.2%	1.5%	3.0%	5.2%	7.0%
J2R001	GEV	21	Min2	13	35	51	68	93	114	138	173	204	239	291	337
			Max2	15	37	55	76	107	135	168	219	265	318	401	476
			AVE2	13	36	53	73	103	129	159	205	246	293	366	431
			SD2	0.67	0.60	1.4	2.7	4.9	7.2	10	15	20	26	35	45
			COV2	0.05	0.02	0.03	0.04	0.05	0.06	0.06	0.07	0.08	0.09	0.10	0.10
			Full FFA	13	35	53	72	102	128	158	204	245	293	367	432
			Δ%	0.6%	1.6%	1.6%	1.4%	1.1%	0.9%	0.7%	0.4%	0.2%	0.0%	0.2%	0.4%
	LP3	21	Min2	11	27	44	67	106	145	196	281	354	441	579	703
			Max2	12	29	48	74	118	162	217	312	406	522	721	937
			AVE2	11	28	46	69	111	152	205	295	383	492	676	854
			SD2	0.56	0.76	1.5	2.6	4.6	6.3	8.3	12	16	23	40	63
			COV2	0.05	0.03	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.05	0.06	0.07
			Full FFA	11	28	46	69	109	147	194	270	341	425	557	677
			Δ%	1.2%	1.3%	1.2%	0.5%	1.3%	3.2%	5.4%	9.0%	12.2%	15.8%	21.3%	26.1%
N2R001	GEV	28	Min2	289	899	1 385	1 893	2 635	3 269	3 976	5 040	5 958	6 986	8 538	9 879
			Max2	318	992	1 537	2 139	3 053	3 855	4 769	6 180	7 426	8 852	11 057	13 009
			AVE2	300	937	1 440	1 994	2 833	3 566	4 399	5 682	6 813	8 102	10 092	11 847
			SD2	10	32	54	80	124	167	222	314	403	514	698	873
			COV2	0.03	0.03	0.04	0.04	0.04	0.05	0.05	0.06	0.06	0.06	0.07	0.07
			Full FFA	275	871	1 347	1 876	2 683	3 394	4 208	5 471	6 591	7 877	9 874	11 647
			Δ%	9.3%	7.6%	6.9%	6.3%	5.6%	5.1%	4.5%	3.9%	3.4%	2.9%	2.2%	1.7%
	LP3	70	Min2	209	590	1 052	1 734	3 084	4 583	6 646	10 546	14 691	20 207	30 296	40 717
			Max2	214	606	1 098	1 831	3 327	5 022	7 390	11 953	16 891	23 568	36 013	49 112
			AVE2	212	599	1 073	1 770	3 181	4 764	6 961	11 157	15 661	21 709	32 885	44 548
			SD2	2.1	7.1	21	46	108	188	313	584	911	1 394	2 385	3 522
			COV2	0.01	0.01	0.02	0.03	0.03	0.04	0.04	0.05	0.06	0.06	0.07	0.08
			Full FFA	203	563	1 004	1 659	3 002	4 530	6 678	10 850	15 401	21 599	33 255	45 626
			Δ%	4.1%	6.5%	6.9%	6.7%	6.0%	5.2%	4.2%	2.8%	1.7%	0.5%	1.1%	2.4%

Site	GEV / LP3	RL2	Stats	AEP (%)											
				50	20	10	5	2	1	0.5	0.2	0.1	0.05	0.02	0.01
				Flood peaks in m^3/s											
Q1R001	GEV	28	Min2	89	199	284	377	507	616	736	915	1 068	1 238	1 492	1 711
			Max2	94	212	308	412	568	703	855	1 087	1 290	1 520	1 872	2 180
			AVE2	91	205	295	392	537	662	802	1 016	1 202	1 411	1 732	2 011
			SD2	1.4	4.1	7.5	12	19	27	36	51	65	83	112	138
			COV2	0.02	0.02	0.03	0.03	0.04	0.04	0.04	0.05	0.05	0.06	0.06	0.07
			Full FFA	91	198	282	374	512	631	766	972	1 151	1 354	1 664	1 935
			Δ%	0.1%	3.8%	4.5%	4.8%	4.8%	4.8%	4.7%	4.5%	4.4%	4.2%	4.1%	3.9%
	LP3	56	Min2	76	164	250	362	561	762	1 019	1 469	1 915	2 476	3 440	4 381
			Max2	79	167	259	383	613	852	1 164	1 723	2 292	3 023	4 313	5 604
			AVE2	78	165	254	372	586	804	1 086	1 584	2 084	2 719	3 823	4 914
			SD2	1.4	1.0	2.9	7.5	19	32	53	95	142	208	335	473
			COV2	0.02	0.01	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10
			Full FFA	80	165	250	359	552	745	988	1 409	1 821	2 335	3 206	4 047
			Δ%	2.6%	0.1%	1.7%	3.6%	6.1%	7.9%	9.8%	12.4%	14.4%	16.5%	19.2%	21.4%
V6H002	GEV	7	Min2	887	1 284	1 494	1 679	1 901	2 056	2 204	2 392	2 529	2 663	2 838	2 970
			Max2	920	1 343	1 621	1 878	2 194	2 418	2 631	2 898	3 090	3 273	3 502	3 667
			AVE2	895	1 321	1 581	1 815	2 101	2 303	2 495	2 735	2 909	3 076	3 288	3 442
			SD2	8.7	20	39	56	77	92	105	121	133	146	164	180
			COV2	0.01	0.02	0.02	0.03	0.04	0.04	0.04	0.04	0.05	0.05	0.05	0.05
			Full FFA	886	1 334	1 609	1 856	2 155	2 365	2 562	2 805	2 977	3 139	3 340	3 482
			Δ%	0.9%	1.0%	1.7%	2.2%	2.5%	2.6%	2.6%	2.5%	2.3%	2.0%	1.6%	1.2%
	LP3	35	Min2	974	1 359	1 501	1 572	1 619	1 633	1 640	1 645	1 637	1 618	1 587	1 561
			Max2	989	1 402	1 539	1 616	1 682	1 726	1 757	1 780	1 786	1 786	1 778	1 769
			AVE2	980	1 387	1 523	1 596	1 647	1 667	1 678	1 684	1 685	1 683	1 677	1 672
			SD2	5.1	13	14	14	20	27	33	40	43	46	51	56
			COV2	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.03	0.03	0.03	0.03
			Full FFA	992	1 390	1 504	1 557	1 587	1 597	1 602	1 604	1 603	1 601	1 597	1 592
			Δ%	1.2%	0.2%	1.2%	2.5%	3.7%	4.4%	4.8%	5.0%	5.1%	5.1%	5.0%	5.0%
X1H001	GEV	21	Min2	215	500	727	978	1 336	1 637	1 974	2 482	2 921	3 414	4 163	4 814
			Max2	244	551	793	1 059	1 458	1 805	2 198	2 817	3 369	4 001	4 979	5 845
			AVE2	232	532	768	1 027	1 418	1 759	2 145	2 738	3 259	3 852	4 764	5 567
			SD2	9.9	17	23	29	40	52	68	96	124	160	219	275
			COV2	0.04	0.03	0.03	0.03	0.03	0.03	0.03	0.04	0.04	0.04	0.05	0.05
			Full FFA	212	490	712	958	1 336	1 668	2 050	2 642	3 168	3 772	4 712	5 547
			Δ%	9.3%	8.6%	7.9%	7.1%	6.2%	5.4%	4.6%	3.6%	2.9%	2.1%	1.1%	0.4%
	LP3	91	Min2	197	430	644	899	1 308	1 677	2 105	2 772	3 360	4 029	5 050	5 937
			Max2	199	438	663	936	1 383	1 799	2 293	3 084	3 806	4 650	5 985	7 189
			AVE2	198	434	654	917	1 341	1 728	2 180	2 891	3 525	4 254	5 380	6 372
			SD2	0.97	4.0	9.1	18	39	64	99	169	244	344	525	708
			COV2	0.005	0.01	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.10	0.11
			Full FFA	197	423	630	874	1 263	1 613	2 016	2 642	3 191	3 814	4 762	5 581
			Δ%	0.6%	2.6%	3.8%	4.8%	6.2%	7.1%	8.1%	9.4%	10.4%	11.5%	13.0%	14.2%

APPENDIX 4-M: Comparison of all 7ym R-W for primary flow sites

A visual comparison of the average results (AVE2), for all the 7ym R-Ws at all primary flow sites.

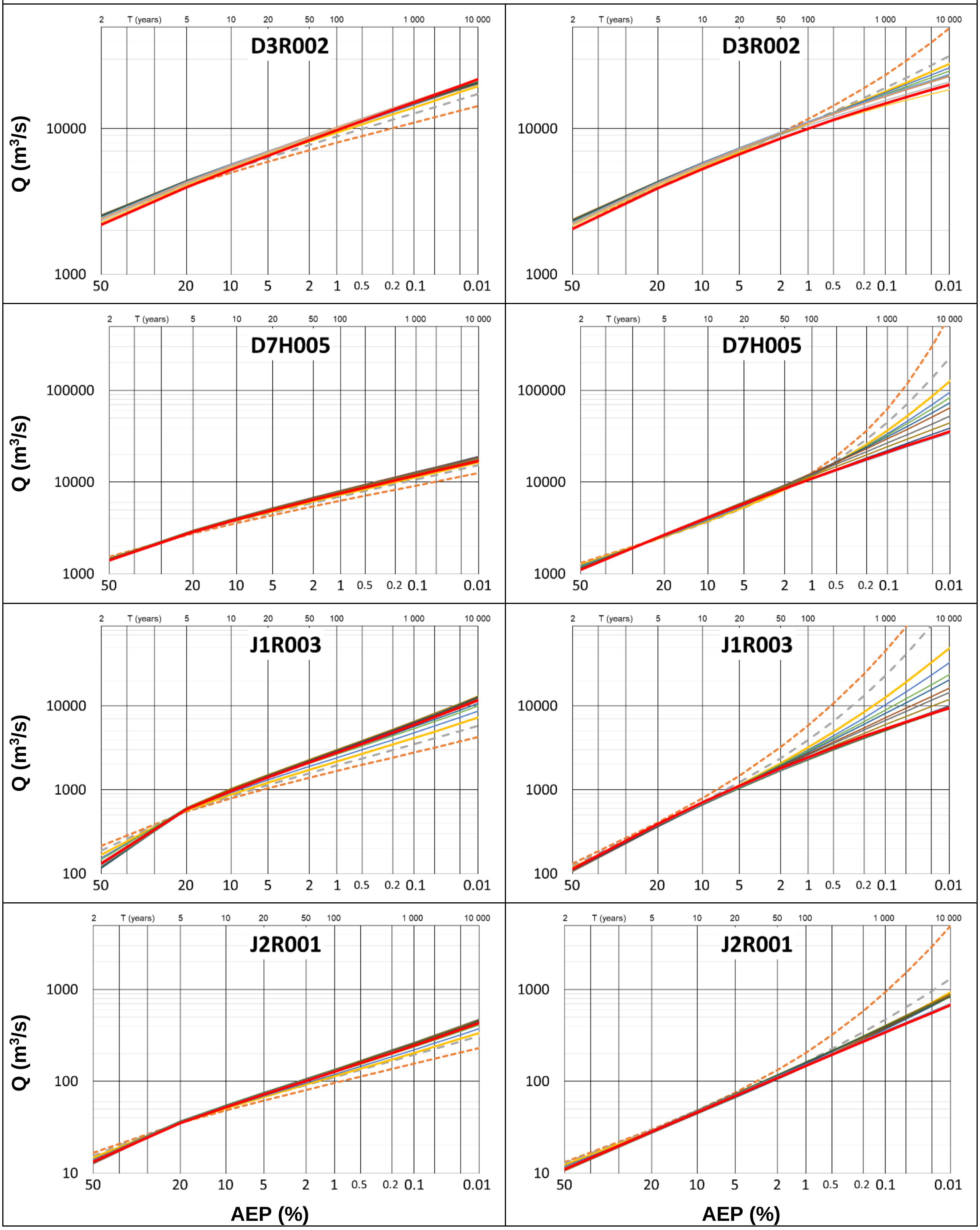
Rolling-window results for GEV	Rolling-window results for LP3
Legend: Average results of all the 7ym R-W (years) 7 14 21 28 35 42 49 56 63 70 77 84 91 98 105 Full	
	
	
	

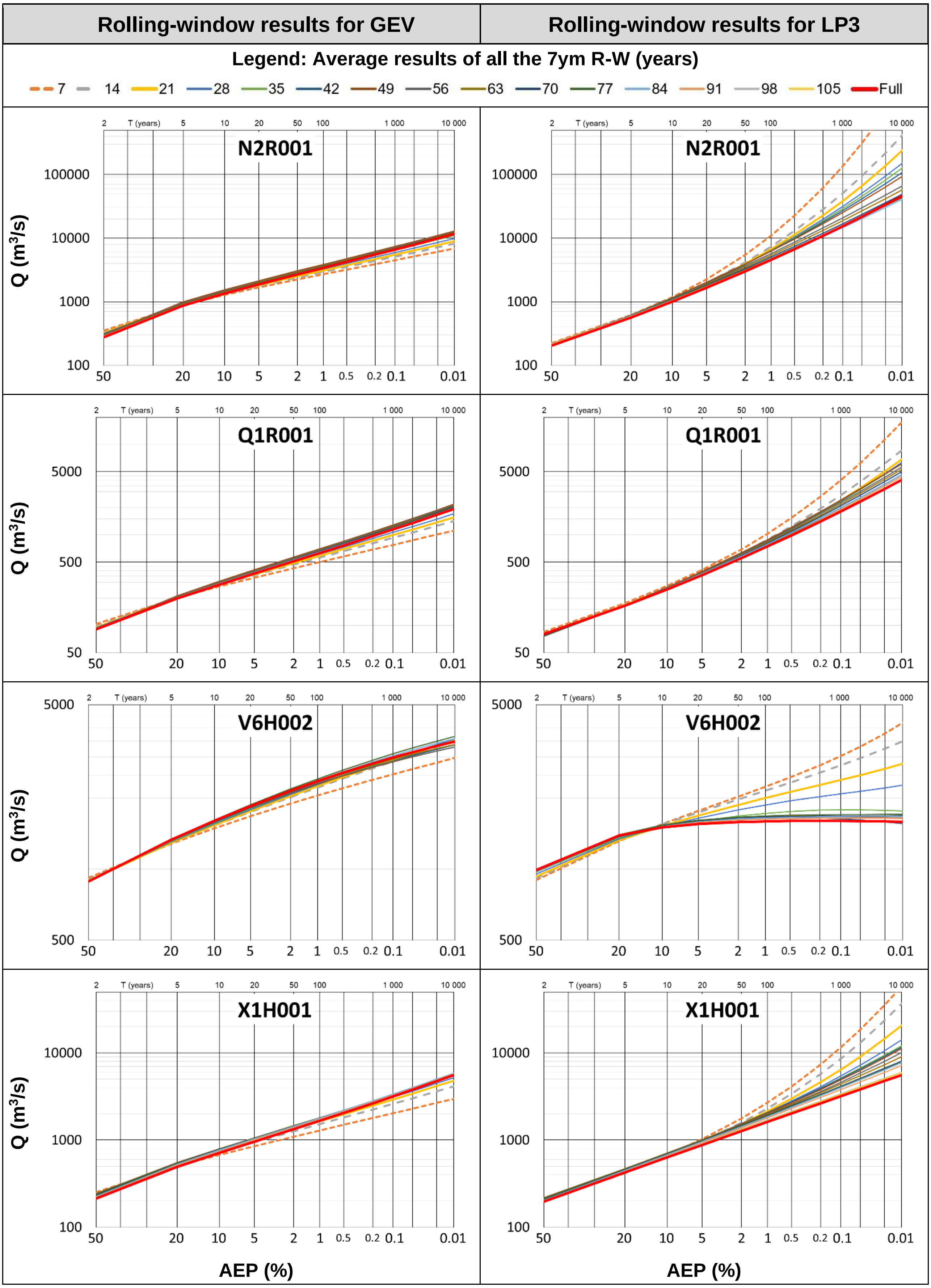

Rolling-window results for GEV
Rolling-window results for LP3
Legend: Average results of all the 7ym R-W (years)
7 14 21 28 35 42 49 56 63 70 77 84 91 98 105 Full
C1R002
C2R001
C5R002
C9R002
T (years)
Q (m³/s)
AEP (%)

Rolling-window results for GEV	Rolling-window results for LP3

Legend: Average results of all the 7ym R-W (years)

7, 14, 21, 28, 35, 42, 49, 56, 63, 70, 77, 84, 91, 98, 105, Full





Rolling-window results for GEV
Rolling-window results for LP3
Legend: Average results of all the 7ym R-W (years)
7
14
21
28
35
42
49
56
63
70
77
84
91
98
105
Full
N2R001
Q1R001
V6H002
X1H001
Q (m³/s)
T (years)
AEP (%)

APPENDIX 4-N: Program code regarding record length research

C++ program code, producing output to be used as input for Excel-spreadsheet, for research on record lengths.

```
#include <iostream>
#include <stdio.h>      /* printf */
#include <math.h>       /* log */
#include <fstream>
#include <iomanip>
#include <string>
#include <sstream>
#include <cstring>
#include <cstdlib>
#include <vector>
#include <dirent.h>

using namespace std;

int main()
{
   int count1 = 1, count2 = 1, start, finish, nyear, nyearbase;
   double ave  = 0, std = 0, skew = 0, kurt = 0, avelog = 0, stdlog = 0, skewlog = 0, kurtlog = 0;
   double Qlp3[15], Qgev[15], P[15], Wp0, WpLP, Wp114, WpGEV, k, Ey, vary;
   string name, LNup;

   ifstream file;
   string fileName;

// cout << "" << endl
// << " Enter input file name:   ";
// cin >> fileName;
// fileName += ".csv";
   fileName = "afp.csv";
   string temp;
   int count = 0;

   file.open(fileName.c_str());

   while(!file.eof())
   {
         file >> temp;
      // cout << temp << endl;
         count++;
   }
   file.close();

   string stringData[count-1];
   double data[count-1];
   int size = count-1;

   file.open(fileName.c_str());
   for(int i = 0; i < size; i++)
   {
               file >> stringData[i];
               data[i] = atof(stringData[i].c_str());
               //cout << setprecision(10) << data[i] << endl;
               //cout << stringData[i] << endl;
   }

   cout << "" << endl
   << " It is a " << size << " year record  -  How many 'moving' years do you want to use?   ";
   cin >> nyearbase;

   nyear = nyearbase;

```

```
// cout << "" << endl
// << " Do you choose to consider the LN as the 'upper limit' of the LP3? yes/no:   ";
// cin >> LNup;

   LNup = "no";

   ofstream Excelfile;
   Excelfile.open("Output.csv");
   Excelfile << "Record length is " << size << " years" <<endl
   <<

        //string Outputfile;

        //cout << "Enter file name: ";
        //cin >> Outputfile;

        //Outputfile += ".csv";

   while (nyear <= size)
   {
        start = 0;
        finish = nyear;

        Excelfile << "Results for " << nyear << " year intervals" <<endl;

        while (finish <= size)
        {
             //************************** AVERAGE ***************************

             for(int i = start; i < finish; i++)
             {
                  ave += data[i];
             }
             ave = ave/nyear;

             //******************** STANDARD DEVIATION ***********************

             for(int i = start; i < finish; i++)
             {
                  std += pow((data[i]-ave),2);
             }
             std = pow(std/(nyear-1),0.5);
             //std = pow(std,0.5);

             //************************* SKEWNESS *****************************

             for(int i = start; i < finish; i++)
             {
                  skew += pow((data[i]-ave)/std,3);
             }
             skew = skew*nyear/((nyear-1)*(nyear-2));

             //************************* KURTOSIS *****************************

             for(int i = start; i < finish; i++)
             {
                  kurt += pow((data[i]-ave)/std,4);
             }
             kurt = kurt*nyear*(nyear+1)/((nyear-1)*(nyear-2)*(nyear-3)) - 3*pow((nyear-1),2)/((nyear-2)*(nyear-

             //************************** LOG AVERAGE ***************************

             for(int i = start; i < finish; i++)
             {
                  avelog += log10(data[i]);
             }
             avelog = avelog/nyear;
```

```
//******************** LOG STANDARD DEVIATION ***********************

for(int i = start; i < finish; i++)
{
	stdlog += pow((log10(data[i])-avelog),2);
}
stdlog = stdlog/(nyear-1);
stdlog = pow(stdlog,0.5);

//************************ LOG SKEWNESS ****************************

for(int i = start; i < finish; i++)
{
	skewlog += pow((log10(data[i])-avelog)/stdlog,3);
}
skewlog = skewlog*nyear/((nyear-1)*(nyear-2));

//************************ LOG KURTOSIS ****************************

for(int i = start; i < finish; i++)
{
	kurtlog += pow((log10(data[i])-avelog)/stdlog,4);
}
kurtlog = kurtlog*nyear*(nyear+1)/((nyear-1)*(nyear-2)*(nyear-3)) - 3*pow((nyear-1),2)/((nyear-

//*************************************** end stat parameters

for(int i = 0; i < 15; i++)
{
	P[i] = (0.1*pow(count1,2) - 0.6*count1+1)/count2;

	//******************************* The GEV distribution ****************************

	if (skew <= 1.060)
	{
		k = -0.00193687*pow(skew,5) - 0.00573999*pow(skew,4) + 0.01650298*pow(skew,3) +
	}
	else
	{
		k = -0.00010487*pow(skew,5) + 0.00240956*pow(skew,4) - 0.02313297*pow(skew,3) +
	}
		Ey = tgamma(1+k);
		vary = tgamma(1+2*k) - pow(Ey,2);
		Wp114 = -log(-log(1-P[i]));
		WpGEV = (1-exp(-k*Wp114))/k;

	if (skew < 1.1396)
	{
		Qgev[i] = ave + (std*(k*WpGEV + Ey - 1)/(pow(vary,0.5)));
	}
	else if (skew > 1.144)
	{
		Qgev[i] = ave - (std*(k*WpGEV + Ey - 1)/(pow(vary,0.5)));
	}
	else
	{
		Qgev[i] = ave + std*(0.78071*Wp114 - 0.45064);
	}

	//****************** The LP3 distribution; if g > 0, set g = 0 *******************

	if (skewlog > 0.000 && LNup == "yes" )
	{
	skewlog = 0.000;
	}
	else
	{
		skewlog = skewlog;
```

```
                }
                if (P[i] == 0.5)
                {
                        Wp0 = 0.000;
                }
                else
                {
                        Wp0 = 3.2654469*pow(pow(-log(P[i]),0.1502),3) - 5.9262054*pow(pow(-
                }

                WpLP = Wp0 + (skewlog/6)*(-0.000889*pow(Wp0,5) + 0.006974*pow(Wp0,4) -
                Qlp3[i] = pow(10,avelog + stdlog*WpLP);

                if (i == 2 || i == 5 || i == 8 || i == 11)
                {
                        count1 = 1;
                        count2*= 10;
                }
                else
                {
                        count1+= 1;
                }
        }

        //******************************** Write results to Output file **********************************************

        Excelfile << "Q_LP3," << avelog << "," << stdlog << "," << skewlog << "," << kurtlog << "," << Qlp3[0]
        << "Q_GEV," << ave << "," << std << "," << skew << "," << kurt << "," << Qgev[0] << "," << Qgev[1]

        start+= 1;
        finish+= 1;
        count1 = 1;
        count2 = 1;
        ave = 0;
        std = 0;
        skew = 0;
        kurt = 0;
        avelog = 0;
        stdlog = 0;
        skewlog = 0;
        kurtlog = 0;
    }

    //nyear = nyear + nyearbase;

    if (nyear + nyearbase < size || nyear == size)
    if (nyear + nyearbase < size || nyear == size)
    {
        nyear = nyear + nyearbase;
    }
    else
    {
        nyear = size;
    }

}

Excelfile.close();

return 0;
}
```

APPENDIX 5: Plotting position

APPENDIX 5-A: Statistics for Boxplots, comparing PP techniques

Boxplots were used to further illustrate the benefit of using the Z-set PP. The estimated statistics for the boxplots, depicted in Figure 5-10and Figure 5-11, are presented below. The statistics and the associated boxplots, for all the individual stations, are provided electronically.

Clanwilliam dam

Results from FFA distribution			AMS record length (years)								Data for Boxplots			
			28		49		70		84					
			Z-scores from Weibull and Z-set PPs								Var_Z		Max-Min	
AEP (%)	Z-score	Q (m^3/s)	Weibull	Z-set	Weibull	Z-set	Weibull	Z-set	Weibull	Z-set	Weibull	Z-set	Weibull	Z-set
50	0.000	332	0.195	-0.011	0.189	0.036	0.079	0.040	0.066	0.027	0.0210	0.0009	0.129	0.051
20	0.842	602	0.825	0.850	0.797	0.842	0.923	0.917	0.900	0.880	0.0031	0.0018	0.126	0.075
10	1.282	802	0.977	1.284	1.040	1.257	1.179	1.361	1.104	1.308	0.0483	0.0019	0.202	0.104
5	1.645	986	1.176	1.621	1.228	1.574	1.424	1.704	1.430	1.652	0.1221	0.0023	0.255	0.131
2	2.054	1233	2.426	2.095	1.508	1.951	1.682	2.106	1.756	2.050	0.1659	0.0038	0.918	0.155
1	2.326	1426	2.045	2.277	2.353	2.263	2.255	2.409	2.540	2.371	0.0327	0.0038	0.496	0.146
0.5	2.576	1624	2.273	2.501	3.701	2.603	2.528	2.658	3.781	2.713	0.7033	0.0080	1.508	0.212

Gariep dam

Results from FFA distribution			AMS record length (years)										Data for Boxplots			
			35		56		77		98		114					
			Z-scores from Weibull and Z-set PPs										Var_Z		Max-Min	
AEP (%)	Z-score	Q (m^3/s)	Weibull	Z-set	Weibull	Z-set	Weibull	Z-set	Weibull	Z-set	Weibull	Z-set	Weibull	Z-set	Weibull	Z-set
50	0.000	2054	0.422	0.440	0.275	0.285	0.054	0.039	0.030	0.002	0.026	0.017	0.161	0.157	0.396	0.438
20	0.842	3894	1.183	1.201	1.054	1.109	0.712	0.825	0.712	0.807	0.788	0.858	0.173	0.139	0.471	0.394
10	1.282	5227	1.388	1.573	1.483	1.536	1.145	1.233	1.188	1.226	1.227	1.290	0.118	0.132	0.338	0.347
5	1.645	6533	1.550	1.881	1.748	1.879	1.345	1.553	1.402	1.552	1.482	1.632	0.181	0.134	0.402	0.329
2	2.054	8340	1.732	2.254	1.939	2.290	1.887	1.960	1.832	1.953	1.898	2.047	0.196	0.128	0.207	0.337
1	2.326	9785	1.851	2.506	2.049	2.565	2.124	2.238	1.984	2.227	2.047	2.331	0.315	0.122	0.273	0.338
0.5	2.576	11306	1.959	2.734	2.149	2.813	2.307	2.489	2.272	2.496	2.329	2.610	0.373	0.119	0.370	0.325

Kammanassie dam

Results from FFA distribution			AMS record length (years)								Data for Boxplots			
			35		56		70		105					
			Z-scores from Weibull and Z-set PPs								Var_Z		Max-Min	
AEP (%)	Z-score	Q (m^3/s)	Weibull	Z-set	Weibull	Z-set	Weibull	Z-set	Weibull	Z-set	Weibull	Z-set	Weibull	Z-set
50	0.000	46	0.197	-0.146	0.182	-0.090	0.254	0.031	0.181	-0.008	0.203	0.069	0.073	0.177
20	0.842	188	0.889	0.721	0.955	0.827	0.976	0.883	0.930	0.809	0.096	0.052	0.087	0.162
10	1.282	377	1.121	1.173	1.194	1.308	1.240	1.341	1.115	1.238	0.114	0.060	0.126	0.168
5	1.645	664	1.277	1.569	1.525	1.754	1.512	1.763	1.406	1.650	0.215	0.077	0.248	0.194
2	2.054	1110	1.456	2.022	1.687	2.255	1.773	2.250	1.659	2.091	0.410	0.117	0.317	0.233
1	2.326	1438	1.511	2.261	1.736	2.522	1.827	2.522	1.779	2.348	0.613	0.119	0.316	0.261
0.5	2.576	1798	1.558	2.468	1.778	2.752	1.873	2.757	1.938	2.607	0.789	0.124	0.380	0.289

Woodstock dam

Results from FFA distribution			AMS record length (years)								Data for Boxplots			
			35		56		63		74					
			Z-scores from Weibull and Z-set PPs								***Var*z**		**Max-Min**	
AEP (%)	Z-score	Q (m^3/s)	Weibull	Z-set	Weibull	Z-set	Weibull	Z-set	Weibull	Z-set	Weibull	Z-set	Weibull	Z-set
50	0.000	384	0.450	0.307	0.086	0.077	0.014	-0.009	0.107	0.023	0.164	0.104	0.436	0.316
20	0.842	698	0.855	1.164	0.578	0.797	0.531	0.731	0.581	0.771	0.212	0.137	0.324	0.433
10	1.282	936	2.790	1.789	1.625	1.263	1.321	1.169	1.323	1.214	0.483	0.177	1.470	0.620
5	1.645	1187	1.850	2.078	1.866	1.657	1.605	1.517	1.577	1.564	0.133	0.164	0.289	0.561
2	2.054	1537	2.365	2.533	1.935	2.079	1.900	1.931	1.969	2.000	0.167	0.170	0.465	0.602
1	2.326	1818	2.698	2.827	1.981	2.353	1.967	2.219	2.034	2.296	0.343	0.166	0.732	0.608
0.5	2.576	2113	2.997	3.091	2.021	2.598	2.026	2.478	2.092	2.562	0.502	0.162	0.976	0.613

Grootdraai dam

Results from FFA distribution			AMS record length (years)								Data for Boxplots			
			28		56		77		114					
			Z-scores from Weibull and Z-set PPs								***Var*z**		**Max-Min**	
AEP (%)	Z-score	Q (m^3/s)	Weibull	Z-set	Weibull	Z-set	Weibull	Z-set	Weibull	Z-set	Weibull	Z-set	Weibull	Z-set
50	0.000	361	-0.054	0.008	-0.111	-0.052	0.021	0.031	-0.014	-0.039	0.050	0.033	0.133	0.084
20	0.842	720	0.659	0.870	0.850	0.922	0.857	0.950	0.803	0.840	0.061	0.055	0.198	0.110
10	1.282	1015	1.292	1.370	1.399	1.469	1.461	1.476	1.246	1.325	0.086	0.128	0.215	0.151
5	1.645	1277	1.460	1.709	1.748	1.865	1.724	1.845	1.566	1.675	0.112	0.128	0.289	0.190
2	2.054	1644	2.351	2.166	2.127	2.304	2.180	2.285	1.886	2.091	0.166	0.158	0.465	0.213
1	2.326	1941	3.006	2.475	2.366	2.591	2.463	2.571	2.008	2.371	0.294	0.176	0.998	0.219
0.5	2.576	2257	3.600	2.755	2.583	2.851	2.721	2.832	2.345	2.664	0.352	0.200	1.255	0.187

Floriskraal dam

Results from FFA distribution			AMS record length (years)								Data for Boxplots			
			35		56		70		96					
			Z-scores from Weibull and Z-set PPs								***Var*z**		**Max-Min**	
AEP (%)	Z-score	Q (m^3/s)	Weibull	Z-set	Weibull	Z-set	Weibull	Z-set	Weibull	Z-set	Weibull	Z-set	Weibull	Z-set
50	0.000	112	0.250	0.194	0.374	0.235	0.082	0.089	-0.018	0.020	0.181	0.135	0.392	0.215
20	0.842	367	0.771	0.992	0.938	1.077	0.907	0.931	0.897	0.831	0.072	0.121	0.166	0.246
10	1.282	638	1.359	1.506	1.446	1.601	1.355	1.419	1.310	1.269	0.086	0.173	0.137	0.332
5	1.645	982	2.135	1.995	2.191	2.128	1.744	1.888	1.606	1.660	0.294	0.273	0.586	0.468
2	2.054	1381	2.774	2.385	2.651	2.546	1.849	2.323	1.798	2.024	0.445	0.280	0.976	0.521
1	2.326	1790	3.261	2.682	3.001	2.863	1.929	2.653	1.887	2.336	0.612	0.307	1.374	0.527
0.5	2.576	1624	2.273	2.501	3.701	2.603	2.528	2.658	3.781	2.713	0.7033	0.0080	1.508	0.212

APPENDIX 6: Probability distribution

APPENDIX 6-A: Probability distribution research – flood frequency data

Flood frequency data (Q vs AEP) to be considered in distribution research.

Site	AEP (%)											
	99.5	**99**	**95**	**90**	**80**	**70**	**60**	**50**	**40**	**30**	**20**	**10**
	Associated flood peak (m^3/s) – from Z-set PP											
A3R002	0.400	0.472	1.20	2.27	4.51	7.46	11.4	16.5	25.8	36.7	54.3	95.4
B1R001	3.13	5.18	13.0	22.2	42.3	67.8	97.3	141	199	285	418	683
B2R001	2.57	3.68	9.46	15.3	26.8	41.3	59.6	80.9	109	150	215	334
C1R002	38.0	48.2	90.4	123	181	240	302	373	462	568	721	983
C2R001	1.19	1.46	3.01	4.29	6.91	9.51	12.1	16.1	20.5	26.3	35.0	49.2
C5R002	4.89	7.06	20.8	37.9	71.5	112	165	236	331	477	738	1 301
C9R002	43.9	56.7	114	178	292	423	571	764	995	1 325	1 827	2 703
D3R002	208	268	505	689	1 006	1 319	1 651	2 024	2 469	3 051	3 847	5 209
D7H005	67.2	88.6	191	278	449	647	880	1 157	1 493	1 965	2 596	3 810
J1R003	1.11	1.66	6.59	12.6	25.9	45.9	73.0	112	172	261	421	783
J2R001	0.464	0.662	1.57	2.49	4.29	6.29	8.69	11.4	15.5	21.1	29.8	45.8
N2R001	10.4	11.9	30.2	48.2	83.5	124	168	229	310	428	618	1 005
Q1R001	5.58	11.1	21.7	29.1	43.1	56.5	70.7	87.1	108	134	174	244
V6H002	132	160	272	353	481	600	710	841	971	1 138	1 360	1 719
X1H001	15.5	20.0	41.1	60.0	91.5	125	162	202	256	330	431	628
A2R001	8.62	11.6	25.8	38.9	67.5	97.6	134	177	228	302	422	630
A2R005	0.674	0.933	3.20	3.71	6.81	10.0	14.0	22.0	26.4	36.5	51.3	80.5
A4R001	1.11	2.03	5.52	10.1	18.7	29.0	46.8	73.4	112	169	261	443
A8R001	2.59	3.82	10.5	17.1	32.1	50.5	73.9	103	146	211	319	541
B3R001	1.36	1.86	4.33	6.97	11.6	16.0	21.7	29.7	40.8	56.8	79.3	129
B6R003	11.7	15.6	31.9	48.7	78.8	111	148	198	250	321	427	630
B7R001	0.462	0.721	1.54	4.71	9.76	16.3	26.0	38.7	59.8	89.4	138	236
C3R002	6.68	8.55	16.6	25.1	43.8	62.7	84.6	112	147	197	273	421
C5R001	0.961	1.35	3.38	5.94	11.9	18.3	27.2	33.8	52.5	82.5	133	237
C5R003	3.27	4.49	10.8	16.9	29.9	44.8	64.3	89.0	123	172	256	432
C7R001	14.6	18.1	33.6	48.5	75.8	104	135	173	219	278	368	527
D6R002	0.800	1.16	3.70	7.04	13.8	24.3	37.3	56.2	83.2	121	185	313
E1R002	21.0	38.4	89.9	121	170	217	268	325	392	476	591	785
H5H004	93.3	111	182	235	323	400	479	563	660	778	972	1 194
J2R003	0.097	0.142	0.393	0.781	1.65	2.77	4.35	6.00	9.55	15.0	23.3	43.7
J3R001	0.540	0.815	2.53	4.87	10.5	18.6	30.3	45.0	72.8	116	203	406
N1R001	1.88	2.79	7.89	13.7	29.7	45.9	70.6	104	150	221	342	597
Q4R002	5.52	7.72	18.6	27.5	42.7	58.9	77.1	98.1	123	159	212	301
Q5R001	8.36	11.6	29.2	48.4	86.3	126	176	241	324	456	646	1 033
S6R002	1.26	1.81	5.27	9.01	16.9	26.8	39.2	55.1	78.8	112	162	266
T2R001	7.75	11.4	26.9	41.4	71.7	107	147	199	259	339	454	652
U2R001	11.5	14.6	29.9	40.0	59.0	77.6	98.5	122	150	189	239	335
U3R001	1.06	2.14	4.76	11.5	23.3	36.2	53.5	77.7	113	163	251	426
V1R003	34.0	52.3	93.8	132	201	271	341	429	535	667	848	1 126
V3R003	12.7	15.4	26.4	35.5	46.5	58.3	70.4	84.5	99.2	118	147	188
W4R001	25.3	36.2	83.2	129	210	301	414	546	730	977	1 376	2 150

Site	AEP (%)									
	5	2	1	0.5	0.2	0.1	0.03	0.01	0.003	0.001
	Associated flood peak (m^3/s) – from Z-set PP									
A3R002	140	208	262	317	390	444	535	610	682	736
B1R001	969	1 384	1 620	1 793	2 338	2 749	3 195	3 571	3 955	4 285
B2R001	470	662	809	956	1 149	1 290	1 524	1 717	1 901	2 039
C1R002	1 260	1 614	1 888	2 162	2 523	2 795	3 261	3 676	4 117	4 503
C2R001	64.7	86.6	101	118	140	157	188	217	250	281
C5R002	1 973	2 985	3 821	4 700	5 907	6 841	8 476	9 951	11 512	12 829
C9R002	3 611	4 789	5 749	6 591	7 694	8 416	9 579	11 115	12 321	13 359
D3R002	6 575	8 377	9 852	11 146	12 941	14 301	16 631	18 710	20 922	22 868
D7H005	5 001	6 742	7 984	9 247	10 871	12 055	13 987	15 578	17 094	18 238
J1R003	1 141	1 753	2 360	2 991	3 760	4 344	5 338	6 026	6 705	7 288
J2R001	64.6	91.4	113	138	165	184	218	248	281	310
N2R001	1 465	2 129	2 634	3 156	3 896	4 455	5 397	6 218	7 055	7 743
Q1R001	317	418	500	579	689	768	921	1 065	1 228	1 381
V6H002	2 064	2 507	2 835	3 159	3 582	3 898	4 440	4 927	5 453	5 925
X1H001	837	1 129	1 358	1 591	1 904	2 141	2 549	2 914	3 298	3 630
A2R001	841	1 137	1 364	1 591	1 886	2 102	2 459	2 757	3 047	3 272
A2R005	112	160	223	262	309	342	395	465	545	621
A4R001	649	953	1 204	1 438	1 750	1 987	2 404	2 791	3 225	3 629
A8R001	742	1 173	1 454	1 795	2 221	2 542	3 101	3 450	3 859	4 205
B3R001	181	258	332	395	486	550	654	770	902	1 026
B6R003	820	1 089	1 299	1 474	1 792	1 964	2 275	2 611	2 883	3 058
B7R001	350	466	558	693	936	1 035	1 196	1 331	1 469	1 587
C3R002	587	824	1 027	1 449	1 735	1 936	2 259	2 531	2 828	3 160
C5R001	357	534	680	834	1 116	1 283	1 531	1 828	2 168	2 489
C5R003	628	897	1 120	1 354	1 674	1 923	2 363	2 768	3 211	3 610
C7R001	694	928	1 105	1 282	1 512	1 680	1 958	2 188	2 410	2 579
D6R002	404	678	863	1 043	1 254	1 365	1 597	1 775	1 957	2 113
E1R002	978	1 232	1 425	1 619	1 876	2 073	2 416	2 732	3 081	3 403
H5H004	1 519	1 815	2 051	2 352	2 642	2 903	3 361	3 787	4 262	4 702
J2R003	65.7	98.3	129	159	195	227	286	344	410	474
J3R001	665	1 068	1 409	1 756	2 255	2 582	3 153	3 670	4 140	4 532
N1R001	898	1 342	1 685	2 037	2 522	2 906	3 569	4 024	4 489	4 888
Q4R002	397	531	642	736	874	978	1 158	1 319	1 490	1 639
Q5R001	1 441	2 076	2 518	2 962	3 604	4 115	4 866	5 486	6 255	6 763
S6R002	376	511	598	730	935	1 029	1 181	1 308	1 439	1 657
T2R001	852	1 083	1 235	1 374	1 628	1 775	2 021	2 236	2 464	2 666
U2R001	495	694	827	948	1 111	1 237	1 416	1 530	1 706	1 936
U3R001	614	911	1 143	1 378	1 691	1 875	2 167	2 414	2 665	2 881
V1R003	1 447	1 875	2 176	2 367	2 758	3 122	3 589	4 007	4 459	4 867
V3R003	237	290	334	377	436	481	561	635	717	794
W4R001	3 049	4 386	5 500	6 675	8 287	9 536	11 736	13 760	15 979	17 997

APPENDIX 6-B: MAPE Ranking for AEP range ≤ 50%

Ranking of the different investigated scenarios, related to the MAPE for the AEP range ≤ 50%.

Site	Scenarios										Qave (ave)			Qmdn (mdn)			Qgmn (gmn)		
	ave0	ave1	ave2	mdn0	mdn1	mdn2	gmn0	gmn1	gmn2		0	1	2	0	1	2	0	1	2
	Qave, including			Qmdn, including			Qgmn, including				Ranking								
	SD	SD, g	SD, SD*	SD	SD, g	SD, SD*	SD	SD, g	SD, SD*		3	5	1	7	9	4	8	6	2
										Σ	186	203	142	237	240	191	238	230	178
A3R002	14%	8.5%	11%	14%	8.8%	13%	15%	11%	11%		7	1	4	8	2	6	9	5	3
B1R001	11%	10%	6.0%	15%	15%	8.5%	14%	13%	5.3%		5	4	2	9	8	3	7	6	1
B2R001	8.8%	6.0%	1.5%	10.0%	7.5%	1.3%	11%	9.2%	1.5%		6	4	3	8	5	1	9	7	2
C1R002	3.0%	3.7%	3.5%	2.6%	3.3%	3.5%	2.9%	3.3%	4.8%		3	8	7	1	4	6	2	5	9
C2R001	5.8%	14%	3.5%	6.1%	14%	5.6%	7.5%	11%	3.1%		4	8	2	5	9	3	6	7	1
C5R002	3.2%	3.9%	5.8%	4.3%	4.5%	6.5%	4.0%	4.2%	8.2%		1	2	7	5	6	8	3	4	9
C9R002	4.8%	5.2%	2.0%	8.6%	9.1%	4.0%	7.4%	7.8%	2.1%		4	5	1	8	9	3	6	7	2
D3R002	2.4%	2.3%	1.7%	2.1%	2.0%	1.9%	2.6%	2.5%	2.9%		6	5	1	4	3	2	8	7	9
D7H005	5.2%	5.6%	2.7%	5.1%	5.3%	3.4%	8.1%	8.4%	2.6%		5	7	2	4	6	3	8	9	1
J1R003	10%	9.2%	9.4%	12%	11%	12%	9.3%	8.9%	8.7%		6	3	5	8	7	9	4	2	1
J2R001	10%	23%	1.4%	11%	21%	6.8%	10%	18%	7.6%		5	9	1	6	8	2	4	7	3
N2R001	3.7%	3.6%	11%	1.9%	1.8%	10%	3.1%	3.0%	15%		6	5	8	2	1	7	4	3	9
Q1R001	9.0%	15%	13%	8.6%	14%	13%	12%	15%	16%		2	7	5	1	6	4	3	8	9
V6H002	6.1%	6.9%	4.2%	10%	11%	4.7%	7.2%	7.6%	4.3%		4	5	1	8	9	3	6	7	2
X1H001	5.0%	6.8%	9.0%	4.8%	6.6%	8.7%	6.7%	7.7%	12%		2	5	8	1	3	7	4	6	9
A2R001	6.5%	5.5%	2.8%	7.1%	6.1%	3.0%	8.6%	8.1%	2.1%		6	4	2	7	5	3	9	8	1
A2R005	11%	18%	14%	9.6%	16%	14%	12%	15%	16%		2	9	5	1	7	4	3	6	8
A4R001	17%	16%	8.1%	22%	21%	11%	18%	18%	6.1%		5	4	2	9	8	3	7	6	1
A8R001	6.1%	4.6%	4.0%	6.8%	5.4%	4.4%	4.9%	4.1%	3.7%		8	5	2	9	7	4	6	3	1
B3R001	2.5%	12%	4.6%	1.6%	11%	4.4%	3.3%	9.1%	8.1%		2	9	5	1	8	4	3	7	6
B6R003	4.1%	2.8%	4.6%	4.5%	3.7%	5.6%	6.4%	5.6%	4.9%		3	1	5	4	2	7	9	8	6
B7R001	17%	10%	16%	18%	13%	17%	19%	16%	17%		5	1	4	8	2	7	9	3	6
C3R002	4.4%	5.8%	4.7%	4.2%	5.6%	6.3%	5.1%	5.4%	4.2%		3	8	4	2	7	9	5	6	1
C5R001	5.7%	6.6%	6.1%	6.8%	7.2%	9.9%	6.6%	6.1%	12%		1	4	3	6	7	8	5	2	9
C5R003	4.0%	4.6%	3.9%	6.8%	5.2%	4.0%	2.0%	1.9%	3.5%		6	7	4	9	8	5	2	1	3
C7R001	1.7%	3.4%	4.9%	2.0%	3.4%	4.7%	1.6%	2.8%	6.4%		2	5	8	3	6	7	1	4	9
D6R002	13%	9.2%	8.8%	13%	9.6%	9.5%	16%	14%	8.8%		7	3	2	6	5	4	9	8	1
E1R002	5.2%	6.2%	4.0%	4.3%	5.2%	4.2%	6.1%	6.7%	5.6%		4	8	1	3	5	2	7	9	6
H5H004	6.9%	7.9%	2.2%	11%	12%	3.3%	8.8%	9.4%	2.8%		4	5	1	8	9	3	6	7	2
J2R003	19%	15%	9.2%	17%	14%	10%	17%	14%	11%		9	6	1	8	4	2	7	5	3
J3R001	16%	16%	11%	12%	11%	16%	9.2%	8.4%	19%		8	6	4	5	3	7	2	1	9
N1R001	9.6%	8.1%	7.3%	9.4%	7.9%	8.2%	9.1%	8.4%	6.1%		9	4	2	8	3	5	7	6	1
Q4R002	2.2%	2.0%	1.7%	4.4%	2.9%	1.7%	2.6%	1.7%	2.6%		5	4	2	9	8	3	6	1	7
Q5R001	6.1%	6.1%	1.2%	6.5%	6.4%	1.4%	6.3%	6.4%	3.5%		5	4	1	9	7	2	6	8	3
S6R002	11%	6.5%	9.0%	14%	11%	11%	13%	11%	8.9%		4	1	3	9	5	7	8	6	2
T2R001	6.5%	6.0%	7.1%	5.7%	5.1%	6.9%	12%	12%	9.0%		4	3	6	2	1	5	9	8	7
U2R001	5.7%	11%	10%	7.5%	14%	9.8%	6.3%	10%	11%		1	7	6	3	9	4	2	5	8
U3R001	11%	9.0%	9.8%	13%	11%	10%	12%	11%	8.3%		7	2	3	9	6	4	8	5	1
V1R003	1.1%	1.3%	1.6%	2.9%	2.7%	3.2%	1.4%	1.3%	1.4%		1	3	6	8	7	9	5	2	4
V3R003	8.8%	15%	4.9%	12%	18%	6.6%	11%	15%	6.5%		4	8	1	6	9	3	5	7	2
W4R001	7.1%	6.4%	2.6%	8.3%	7.9%	3.1%	11%	10%	2.3%		5	4	2	7	6	3	9	8	1

APPENDIX 6-C: MAPE Ranking for AEP range 50% to 95%

Ranking of the different investigated scenarios, re the MAPE for AEP range 50% to 95%

Site	Scenarios										Q_{ave} (ave)			Q_{mdn} (mdn)			Q_{gmn} (gmn)		
	ave0	ave1	ave2	mdn0	mdn1	mdn2	gmn0	gmn1	gmn2		0	1	2	0	1	2	0	1	2
	Q_{ave}, including			Q_{mdn}, including			Q_{gmn}, including				Ranking								
	SD	SD, g	SD, SD*	SD	SD, g	SD, SD*	SD	SD, g	SD, SD*		8	9	7	5	6	4	2	3	1
										Σ	242	271	232	204	226	202	147	187	134
A3R002	53%	118%	23%	64%	160%	53%	7.0%	139%	8.1%		5	7	3	6	9	4	1	8	2
B1R001	64%	67%	25%	14%	17%	13%	21%	24%	2.2%		8	9	7	3	4	2	5	6	1
B2R001	38%	42%	6.6%	20%	26%	2.7%	16%	24%	3.1%		8	9	3	5	7	1	4	6	2
C1R002	4.2%	4.0%	5.2%	5.0%	4.7%	5.9%	1.1%	1.1%	1.4%		5	4	8	7	6	9	2	1	3
C2R001	13%	25%	4.9%	22%	40%	19%	8.5%	34%	4.4%		4	7	2	6	9	5	3	8	1
C5R002	44%	45%	101%	26%	27%	45%	32%	35%	47%		5	7	9	1	2	6	3	4	8
C9R002	31%	30%	15%	10%	9.3%	3.2%	14%	13%	6.0%		9	8	7	4	3	1	6	5	2
D3R002	5.6%	5.7%	4.1%	6.7%	6.8%	6.6%	1.3%	1.4%	1.0%		5	6	4	8	9	7	2	3	1
D7H005	23%	22%	14%	30%	29%	26%	13%	12%	7.5%		6	5	4	9	8	7	3	2	1
J1R003	29%	22%	37%	24%	22%	27%	51%	41%	50%		5	2	6	3	1	4	9	7	8
J2R001	14%	59%	5.7%	4.8%	72%	4.8%	6.9%	99%	1.6%		6	7	4	3	8	2	5	9	1
N2R001	3.5%	3.5%	50%	20%	21%	46%	7.3%	8.3%	26%		1	2	9	5	6	8	3	4	7
Q1R001	13%	11%	22%	15%	11%	20%	5.5%	1.5%	8.0%		6	4	9	7	5	8	2	1	3
V6H002	16%	16%	6.4%	7.5%	7.2%	3.6%	8.7%	8.2%	5.9%		9	8	3	5	4	1	7	6	2
X1H001	7.8%	6.7%	19%	9.9%	8.3%	15%	2.6%	2.4%	4.5%		5	4	9	7	6	8	2	1	3
A2R001	26%	27%	12%	20%	22%	13%	14%	16%	7.1%		8	9	2	6	7	3	4	5	1
A2R005	12%	50%	33%	29%	84%	43%	7.6%	84%	16%		2	7	5	4	9	6	1	8	3
A4R001	100%	104%	25%	11%	6.6%	72%	12%	19%	37%		8	9	5	2	1	7	3	4	6
A8R001	14%	18%	8.5%	13%	6.6%	23%	5.3%	3.9%	12%		7	8	4	6	3	9	2	1	5
B3R001	5.1%	17%	21%	12%	15%	22%	6.6%	35%	11%		1	6	7	4	5	8	2	9	3
B6R003	16%	17%	17%	35%	36%	37%	12%	13%	11%		4	5	6	7	8	9	2	3	1
B7R001	132%	162%	134%	59%	103%	50%	37%	95%	26%		7	9	8	4	6	3	2	5	1
C3R002	14%	8.3%	12%	11%	6.0%	5.9%	3.5%	9.1%	9.2%		9	4	8	7	3	2	1	5	6
C5R001	40%	20%	112%	41%	12%	72%	47%	9.3%	71%		4	3	9	5	2	8	6	1	7
C5R003	12%	17%	7.2%	33%	25%	43%	8.5%	3.0%	13%		4	6	2	8	7	9	3	1	5
C7R001	3.8%	4.9%	5.0%	2.2%	3.7%	2.7%	4.2%	6.2%	0.9%		5	7	8	2	4	3	6	9	1
D6R002	73%	88%	24%	63%	83%	39%	17%	44%	8.7%		7	9	3	6	8	4	2	5	1
E1R002	10%	10.0%	8.3%	13%	13%	13%	4.2%	3.7%	3.8%		6	5	4	9	7	8	3	1	2
H5H004	17%	17%	8.1%	8.9%	8.4%	5.0%	8.6%	8.2%	5.9%		9	8	3	7	5	1	6	4	2
J2R003	59%	207%	100%	16%	210%	95%	25%	276%	92%		3	7	6	1	8	5	2	9	4
J3R001	66%	58%	409%	113%	99%	299%	156%	140%	312%		2	1	9	4	3	7	6	5	8
N1R001	25%	30%	7.8%	21%	29%	14%	11%	4.6%	20%		7	9	2	6	8	4	3	1	5
Q4R002	5.1%	7.5%	2.7%	3.0%	5.6%	5.5%	5.1%	10%	1.9%		4	8	2	3	7	6	5	9	1
Q5R001	29%	29%	8.9%	12%	13%	7.4%	9.4%	9.2%	6.9%		8	9	3	6	7	2	5	4	1
S6R002	60%	69%	48%	18%	31%	7.3%	22%	39%	14%		8	9	7	3	5	1	4	6	2
T2R001	40%	40%	42%	54%	55%	56%	24%	24%	21%		4	5	6	7	8	9	2	3	1
U2R001	20%	17%	12%	4.7%	3.2%	14%	5.6%	2.6%	6.8%		9	8	6	3	2	7	4	1	5
U3R001	70%	77%	58%	11%	22%	4.1%	15%	28%	5.6%		8	9	7	3	5	1	4	6	2
V1R003	3.8%	4.0%	7.2%	2.9%	3.3%	3.2%	4.4%	4.5%	4.5%		4	5	9	1	3	2	6	7	8
V3R003	17%	16%	12%	9.1%	6.4%	6.4%	8.6%	5.2%	6.9%		9	8	7	6	2	3	5	1	4
W4R001	30%	31%	18%	15%	16%	11%	8.6%	11%	13%		8	9	7	5	6	2	1	3	4

APPENDIX 6-D: MAE Ranking for AEP range ≤ 50%

Ranking of the different investigated scenarios, related to the MAE for the AEP range ≤ 50%.

Site	Scenarios										Qave (ave)			Qmdn (mdn)			Qgmn (ave)		
	ave0	ave1	ave2	mdn0	mdn1	mdn2	gmn0	gmn1	gmn2		0	1	2	0	1	2	0	1	2
	Qave, including			Qmdn, including			Qgmn, including				Ranking								
	SD	SD, g	SD, SD*	SD	SD, g	SD, SD*	SD	SD, g	SD, SD*		5	4	1	8	6	3	9	7	2
										Σ	210	207	159	237	212	179	250	218	173
A3R002	49	33	42	48	31	42	53	20	41		8	3	6	7	2	5	9	1	4
B1R001	211	185	120	271	248	139	254	238	109		5	4	2	9	7	3	8	6	1
B2R001	82	47	12	95	61	14	96	72	8		7	4	2	8	5	3	9	6	1
C1R002	66	86	75	55	76	77	64	78	101		3	8	4	1	5	6	2	7	9
C2R001	10	21	7	9	21	8	12	13	7		5	9	2	4	8	3	6	7	1
C5R002	200	228	144	175	185	174	160	174	206		7	9	1	5	6	3	2	4	8
C9R002	210	249	120	412	451	172	371	407	152		4	5	1	8	9	3	6	7	2
D3R002	261	241	153	191	185	168	262	251	274		7	5	1	4	3	2	8	6	9
D7H005	455	503	239	396	423	213	647	691	197		6	7	3	4	5	2	8	9	1
J1R003	393	355	379	360	318	378	370	343	357		9	3	8	5	1	7	6	2	4
J2R001	28	36	2	29	33	23	28	26	23		5	9	1	7	8	3	6	4	2
N2R001	78	71	333	43	39	324	86	79	402		4	3	8	2	1	7	6	5	9
Q1R001	59	106	83	57	101	84	70	103	97		2	9	4	1	7	5	3	8	6
V6H002	165	196	169	273	317	152	183	209	163		3	6	4	8	9	1	5	7	2
X1H001	94	134	156	90	129	156	115	143	187		2	5	7	1	4	8	3	6	9
A2R001	97	76	42	109	88	43	125	111	29		6	4	2	7	5	3	9	8	1
A2R005	35	56	45	31	54	44	35	36	45		3	9	6	1	8	5	2	4	7
A4R001	259	237	153	313	295	170	291	278	138		5	4	2	9	8	3	7	6	1
A8R001	109	69	76	125	85	82	104	75	63		8	2	4	9	6	5	7	3	1
B3R001	8	63	17	8	58	16	11	49	23		2	9	5	1	8	4	3	7	6
B6R003	65	40	73	38	23	64	86	68	67		5	3	8	2	1	4	9	7	6
B7R001	116	58	117	130	77	123	132	92	118		4	1	5	8	2	7	9	3	6
C3R002	49	78	77	50	81	81	54	79	69		1	6	5	2	9	8	3	7	4
C5R001	48	45	45	49	45	49	46	41	59		6	3	2	7	4	8	5	1	9
C5R003	45	42	42	74	36	37	41	33	42		8	6	7	9	2	3	4	1	5
C7R001	36	65	76	34	61	77	34	55	93		3	6	7	2	5	8	1	4	9
D6R002	129	75	92	130	79	93	144	107	86		7	1	4	8	2	5	9	6	3
E1R002	82	103	59	61	82	58	90	106	79		5	8	2	3	6	1	7	9	4
H5H004	152	181	62	218	253	66	180	204	61		4	6	2	8	9	3	5	7	1
J2R003	38	25	25	40	21	26	39	15	24		7	4	5	9	2	6	8	1	3
J3R001	163	138	66	167	141	75	159	141	104		8	4	1	9	5	2	7	6	3
N1R001	208	165	173	203	159	175	210	181	158		8	3	4	7	2	5	9	6	1
Q4R002	22	17	11	32	14	10	24	10	15		7	6	3	9	4	2	8	1	5
Q5R001	162	159	27	194	188	26	189	190	68		5	4	2	9	6	1	7	8	3
S6R002	75	32	64	95	55	72	90	60	62		7	1	5	9	2	6	8	3	4
T2R001	77	68	86	57	45	76	131	126	92		5	3	6	2	1	4	9	8	7
U2R001	53	124	89	67	151	83	62	124	91		1	8	5	3	9	4	2	7	6
U3R001	121	85	108	149	112	119	139	112	103		7	1	3	9	5	6	8	4	2
V1R003	23	31	40	45	35	51	29	23	31		2	4	7	8	6	9	3	1	5
V3R003	34	66	18	44	75	22	40	63	22		4	8	1	6	9	3	5	7	2
W4R001	441	370	226	554	508	228	654	608	220		5	4	2	7	6	3	9	8	1

APPENDIX 6-E: MAE Ranking for AEP range 50% to 95%

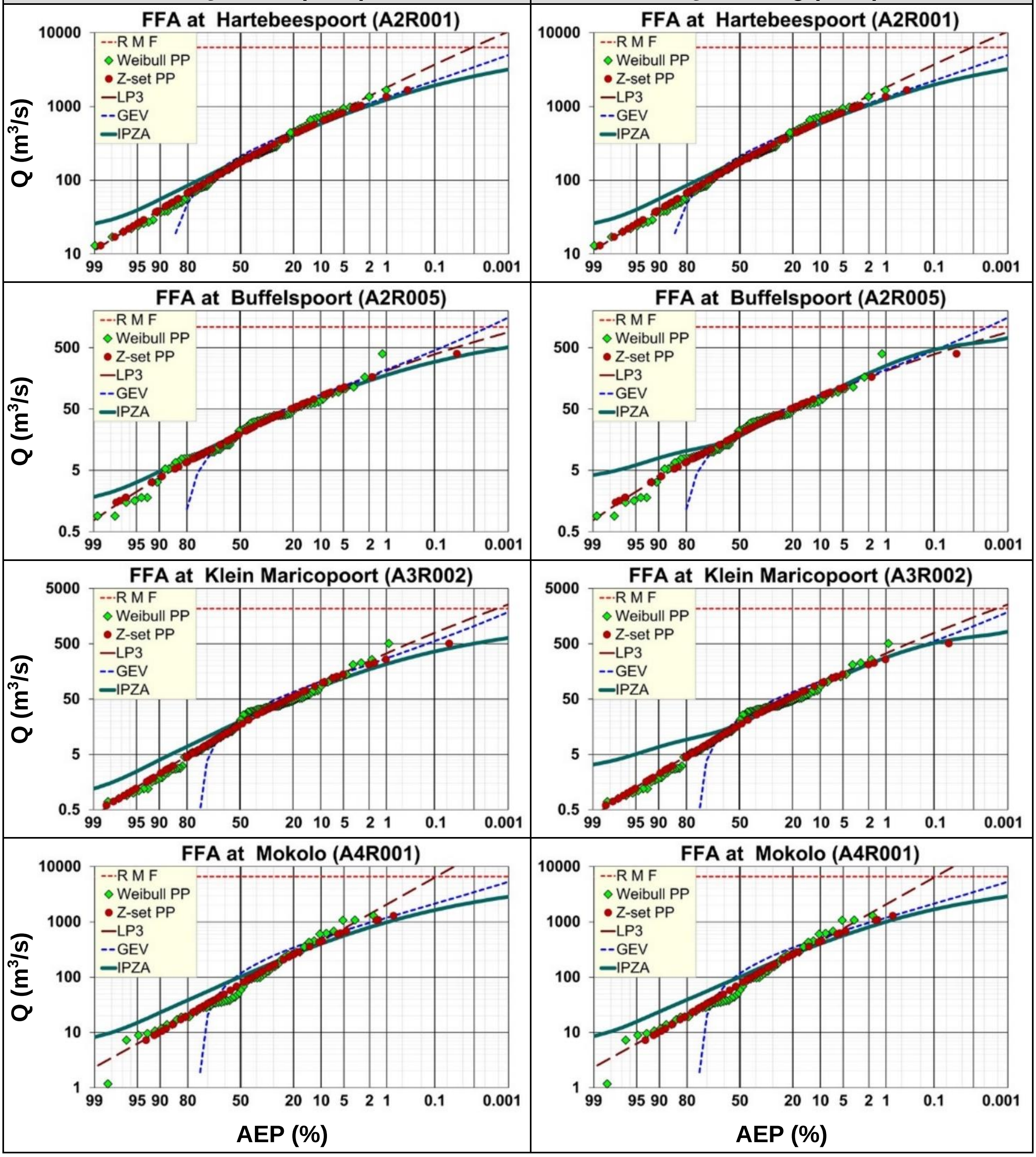
Ranking of the different investigated scenarios, re the MAE for AEP range 50% to 95%

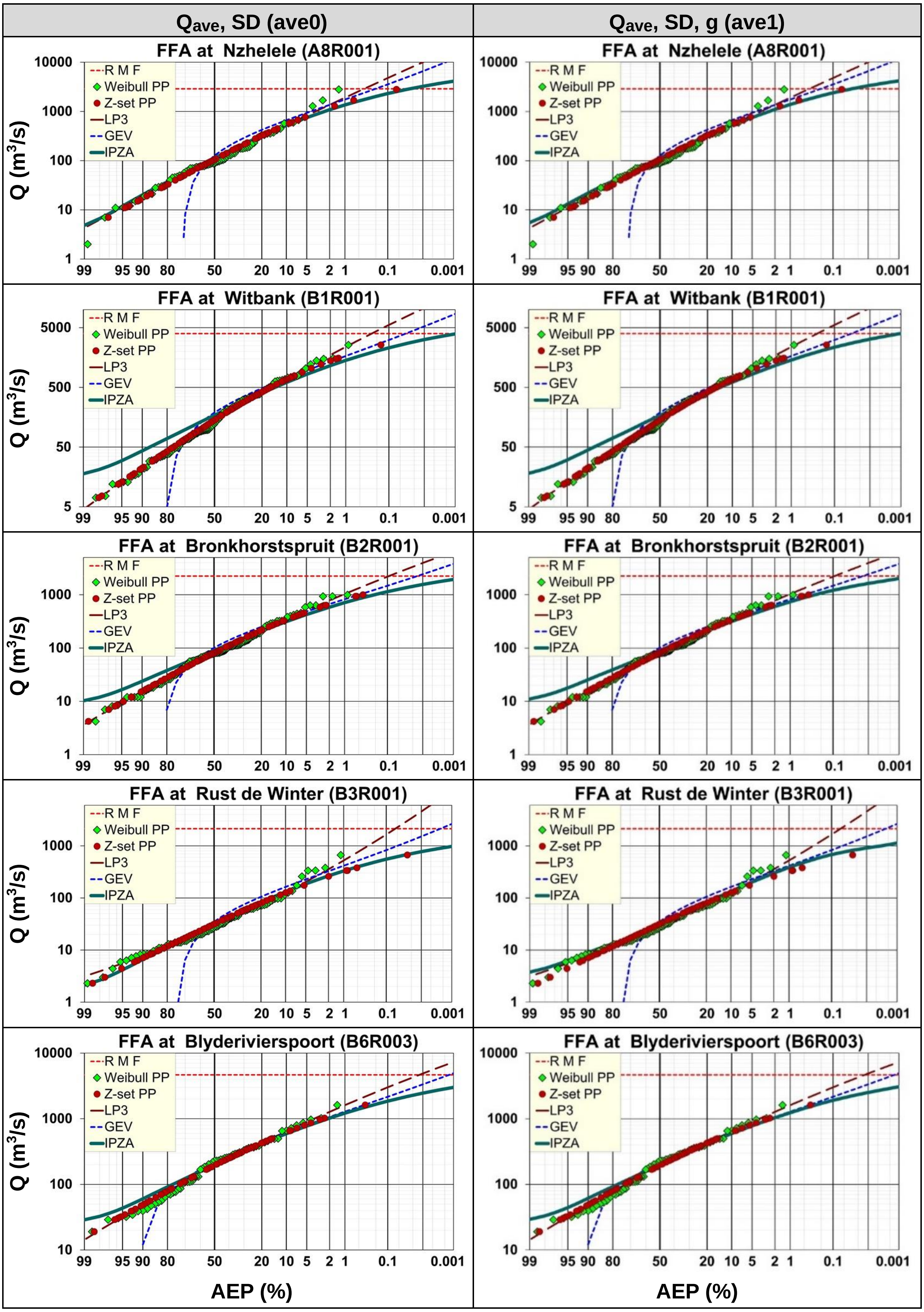
Site	Scenarios										Qave (ave)			Qmdn (mdn)			Qgmn (gmn)		
	ave0	ave1	ave2	mdn0	mdn1	mdn2	gmn0	gmn1	gmn2		0	1	2	0	1	2	0	1	2
	Qave, including			Qmdn, including			Qgmn, including				Ranking								
	SD	SD, g	SD, SD*	SD	SD, g	SD, SD*	SD	SD, g	SD, SD*		8	9	3	4	7	5	2	6	1
										Σ	248	272	194	203	247	209	149	213	110
A3R002	2.0	3.3	0.97	2.5	4.8	2.0	0.35	3.9	0.32		4	7	3	6	9	5	2	8	1
B1R001	23	23	9.0	4.5	5.7	5.7	7.5	8.6	1.2		8	9	7	2	3	4	5	6	1
B2R001	8.9	9.4	1.5	4.2	5.4	0.67	4.1	5.7	0.53		8	9	3	5	6	2	4	7	1
C1R002	8.9	8.8	8.4	10	11	10	7.6	8.1	5.0		6	5	4	7	9	8	2	3	1
C2R001	0.88	1.4	0.26	1.5	2.5	1.3	0.72	2.1	0.36		4	6	1	7	9	5	3	8	2
C5R002	24	25	53	12	13	23	12	13	23		7	8	9	2	3	6	1	4	5
C9R002	80	79	35	27	27	13	41	38	16		9	8	5	3	4	1	7	6	2
D3R002	60	60	36	69	76	66	44	47	29		5	6	2	8	9	7	3	4	1
D7H005	95	93	49	127	129	111	62	58	34		6	5	2	8	9	7	4	3	1
J1R003	5.2	4.3	5.0	9.6	9.2	11	6.5	4.8	8.8		4	1	3	8	7	9	5	2	6
J2R001	0.59	1.7	0.19	0.21	2.3	0.18	0.38	3.2	0.10		6	7	3	4	8	2	5	9	1
N2R001	4.9	4.4	32	13	13	31	6.6	7.2	16		2	1	9	6	5	8	3	4	7
Q1R001	4.7	4.5	7.7	5.5	5.2	6.6	3.2	2.8	2.7		5	4	9	7	6	8	3	2	1
V6H002	59	58	28	26	28	20	33	32	24		9	8	4	3	5	1	7	6	2
X1H001	6.9	6.8	15	8.4	8.6	11	5.3	5.7	3.9		5	4	9	6	7	8	2	3	1
A2R001	15	15	5.8	12	13	7.2	9.2	10	4.1		8	9	2	6	7	3	4	5	1
A2R005	1.0	3.0	1.9	1.9	4.8	2.6	1.0	5.0	1.3		2	7	4	5	8	6	1	9	3
A4R001	17	18	5.2	2.2	1.7	10	2.9	3.8	4.7		8	9	6	2	1	7	3	4	5
A8R001	4.9	5.3	2.2	3.1	1.8	6.5	2.2	2.7	2.9		7	8	2	6	1	9	3	4	5
B3R001	0.74	1.5	1.7	0.85	1.6	2.0	1.2	3.5	1.1		1	5	7	2	6	8	4	9	3
B6R003	11	11	9.8	26	27	26	10	11	8.0		4	6	2	7	9	8	3	5	1
B7R001	8.2	9.1	8.0	3.0	4.6	2.4	1.8	4.4	1.1		8	9	7	4	6	3	2	5	1
C3R002	5.3	4.3	3.1	4.1	3.5	2.1	2.4	4.2	2.9		9	8	4	6	5	1	2	7	3
C5R001	3.6	2.2	8.6	3.7	1.5	6.4	3.5	1.8	6.1		5	3	9	6	1	8	4	2	7
C5R003	4.3	5.0	2.5	8.6	7.0	11	2.1	2.1	2.8		5	6	3	8	7	9	2	1	4
C7R001	3.5	4.0	3.9	2.5	3.9	2.2	4.9	6.2	1.6		4	7	5	3	6	2	8	9	1
D6R002	8.4	9.1	3.0	7.1	8.7	4.9	2.3	4.6	1.2		7	9	3	6	8	5	2	4	1
E1R002	17	17	11	22	22	21	13	13	10		6	5	2	8	9	7	3	4	1
H5H004	41	41	24	20	19	12	20	20	14		9	8	7	4	3	1	5	6	2
J2R003	0.95	1.9	0.92	0.26	1.8	1.0	0.22	2.6	0.96		4	8	3	2	7	6	1	9	5
J3R001	8.1	7.7	25	7.8	7.0	19	8.9	7.6	20		5	3	9	4	1	7	6	2	8
N1R001	6.7	7.1	3.1	6.1	7.5	6.2	2.3	2.4	4.2		7	8	3	5	9	6	1	2	4
Q4R002	3.0	3.6	1.6	2.8	3.8	3.0	3.6	5.3	1.7		4	7	1	3	8	5	6	9	2
Q5R001	22	22	5.0	9.3	9.9	5.2	8.8	9.3	4.2		9	8	2	5	7	3	4	6	1
S6R002	8.6	9.2	6.5	2.1	3.7	0.82	3.2	5.1	1.6		8	9	7	3	5	1	4	6	2
T2R001	24	24	22	34	35	34	15	15	12		5	6	4	7	9	8	2	3	1
U2R001	9.6	9.0	6.2	3.0	3.8	8.6	3.8	4.0	5.0		9	8	6	1	3	7	2	4	5
U3R001	12	12	9.3	0.99	2.5	1.7	2.3	4.0	0.85		8	9	7	2	5	3	4	6	1
V1R003	11	11	11	12	13	11	14	15	12		1	2	4	5	7	3	8	9	6
V3R003	6.0	5.6	4.8	2.8	2.5	2.1	3.0	2.5	2.2		9	8	7	5	3	1	6	4	2
W4R001	53	55	22	24	25	11	18	21	17		8	9	5	6	7	1	3	4	2

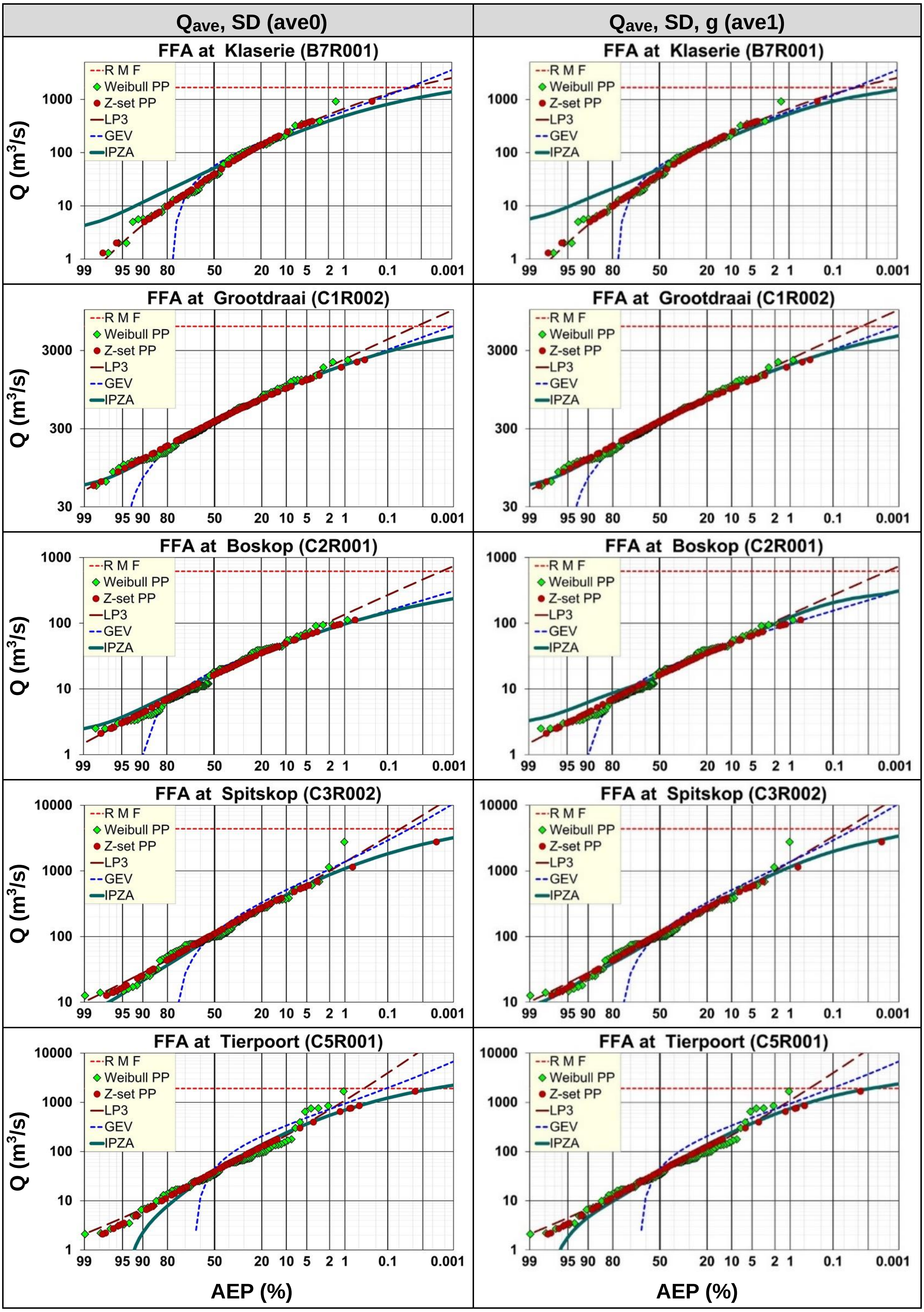
APPENDIX 6-F: Visual comparison of ave0 with ave1

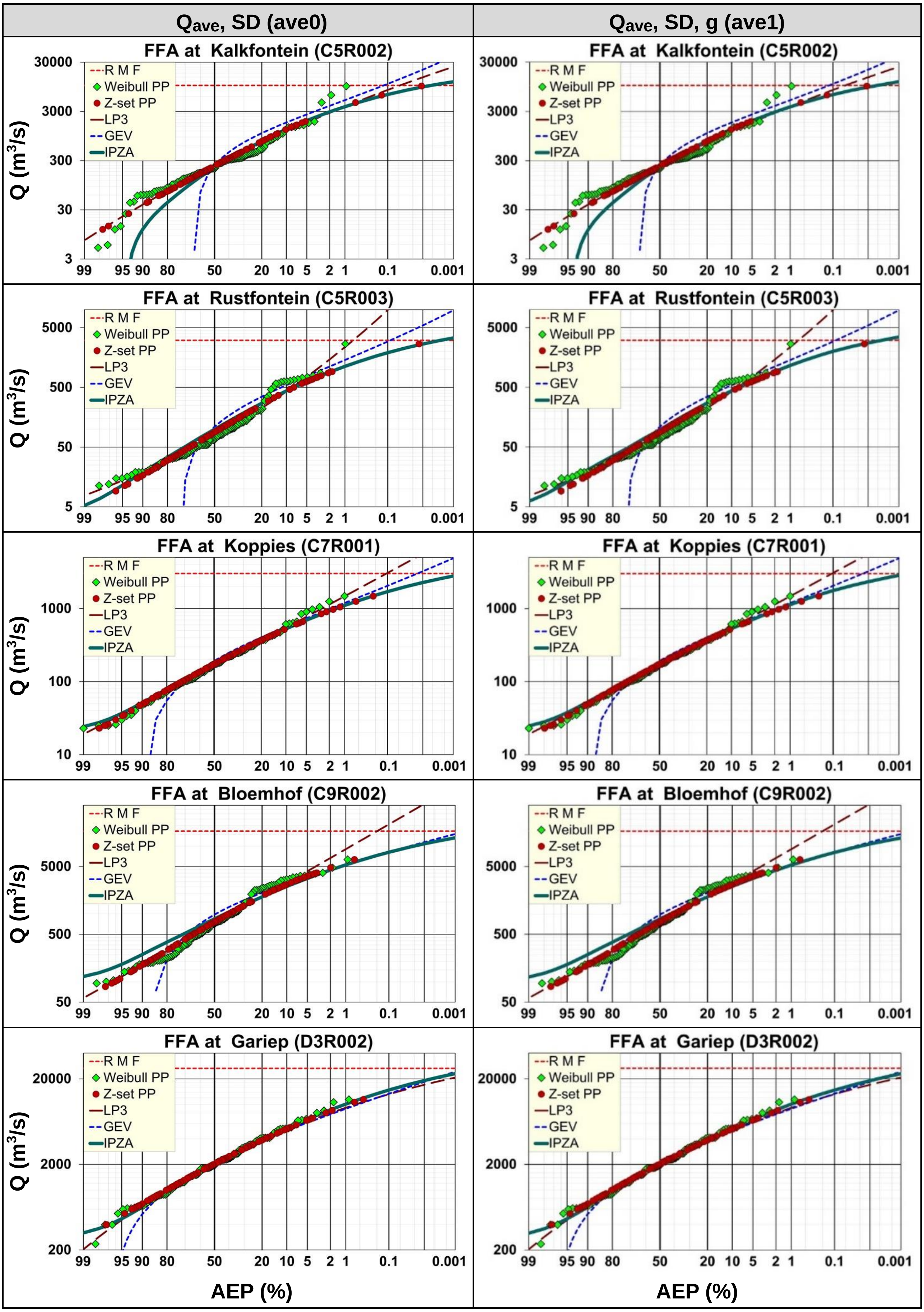
Visually compare the ave0-scenario, which included the Q_{ave} and the SD, against the ave1-scenario, which included the Q_{ave}, SD and g.

Q_{ave}, SD (ave0)	Q_{ave}, SD, g (ave1)

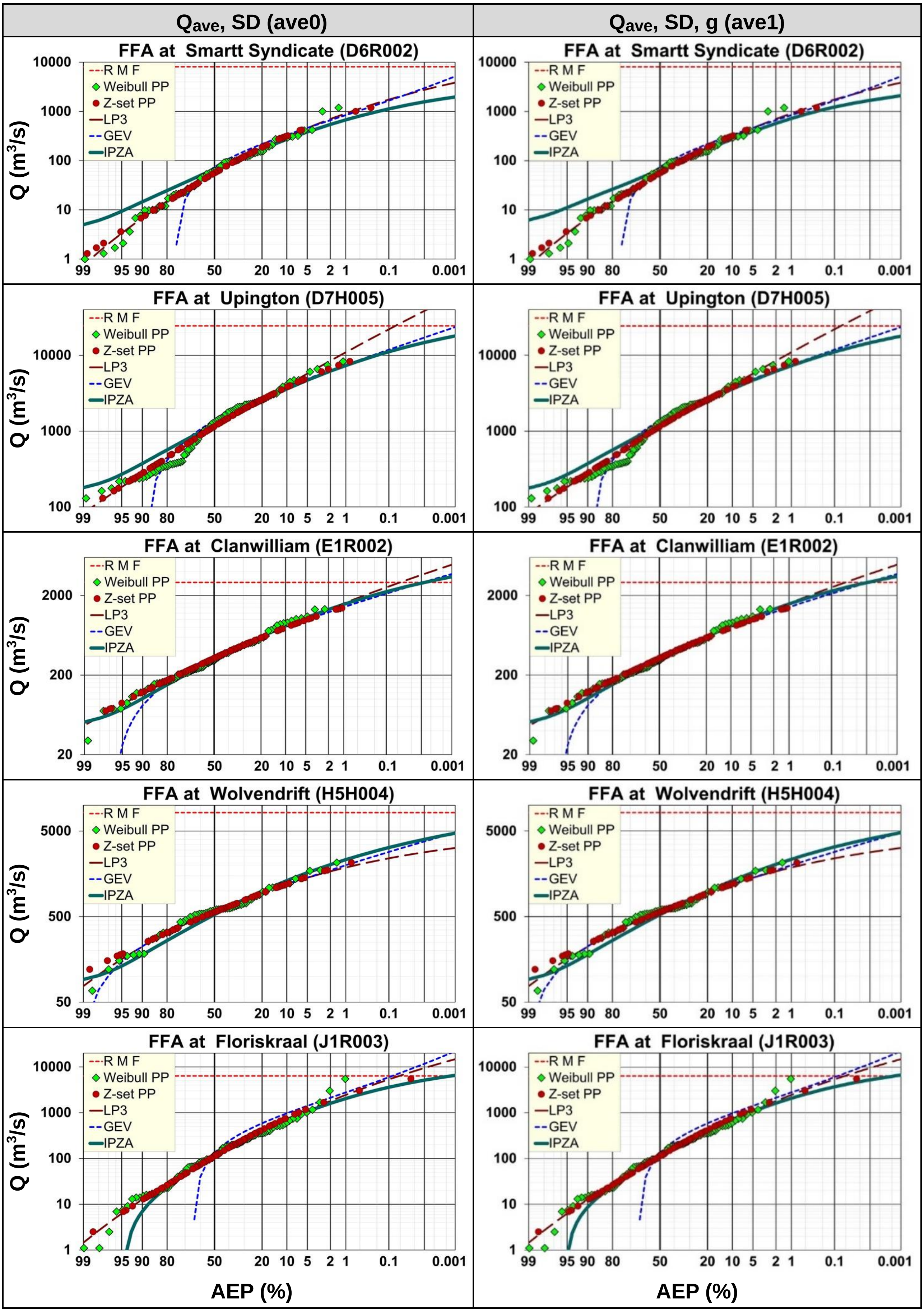



Qave, SD (ave0)
Qave, SD, g (ave1)
FFA at Nzhelele (A8R001)
FFA at Witbank (B1R001)
FFA at Bronkhorstspruit (B2R001)
FFA at Rust de Winter (B3R001)
FFA at Blyderivierspoort (B6R003)
R M F
Weibull PP
Z-set PP
LP3
GEV
IPZA
Q (m3/s)
AEP (%)
99 95 90 80 50 20 10 5 2 1 0.1 0.001
1 10 100 1000 10000
5 50 500 5000


Qave, SD (ave0)
Qave, SD, g (ave1)
FFA at Klaserie (B7R001)
FFA at Grootdraai (C1R002)
FFA at Boskop (C2R001)
FFA at Spitskop (C3R002)
FFA at Tierpoort (C5R001)
R M F
Weibull PP
Z-set PP
LP3
GEV
IPZA
Q (m³/s)
AEP (%)
99 95 90 80 50 20 10 5 2 1 0.1 0.001

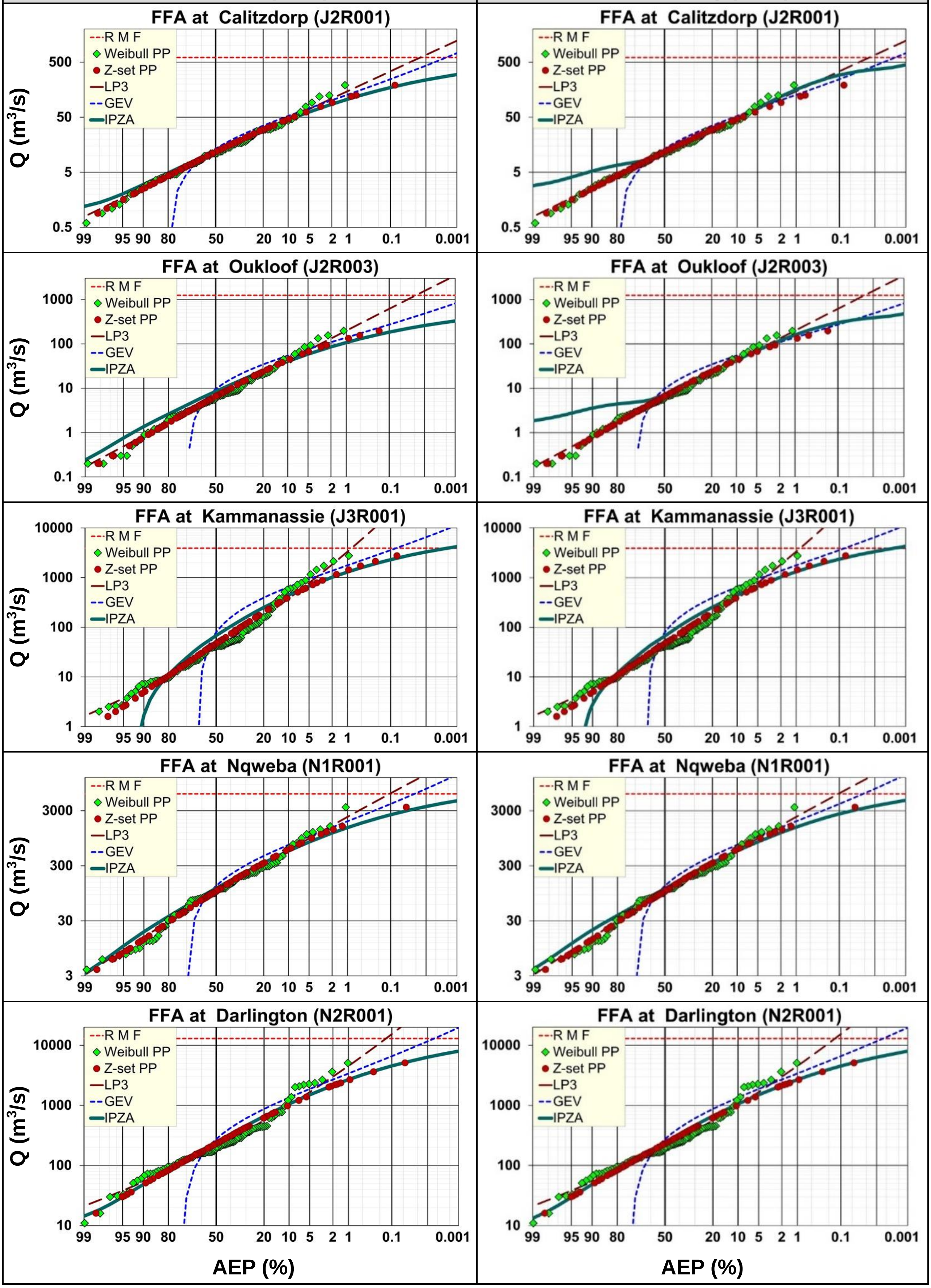


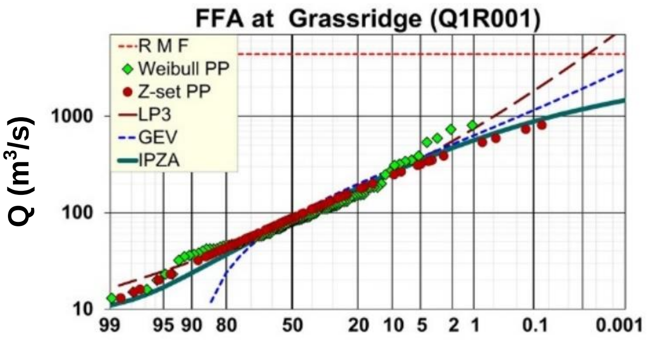
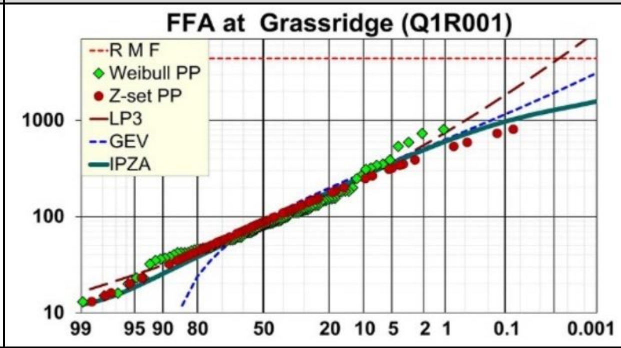
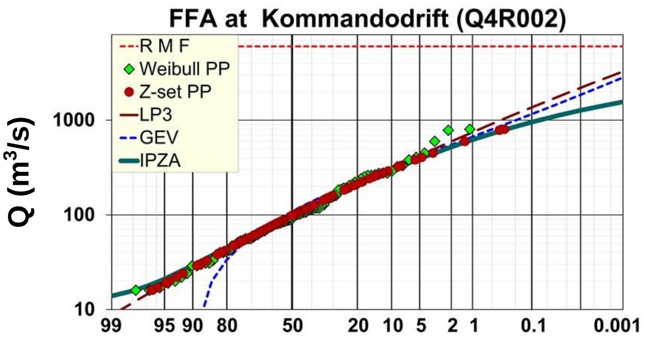
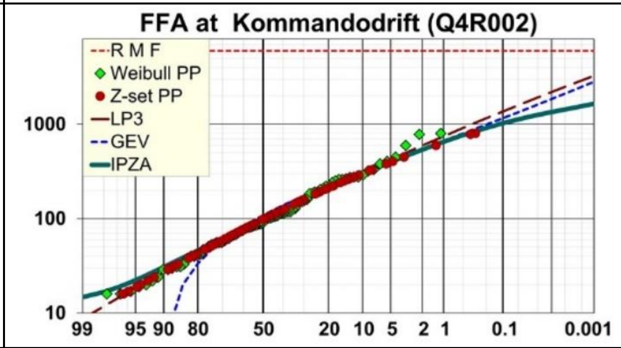
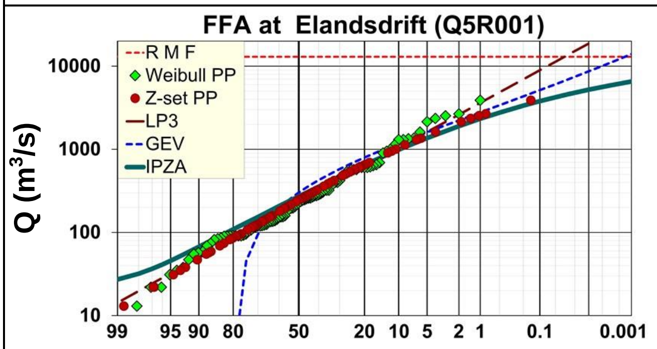
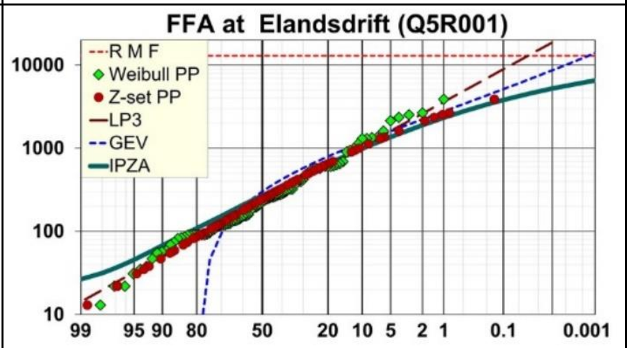
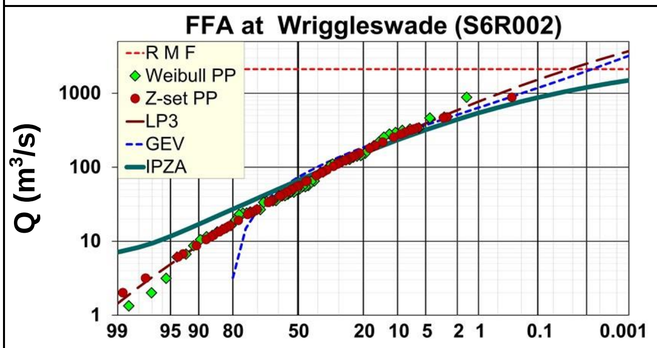
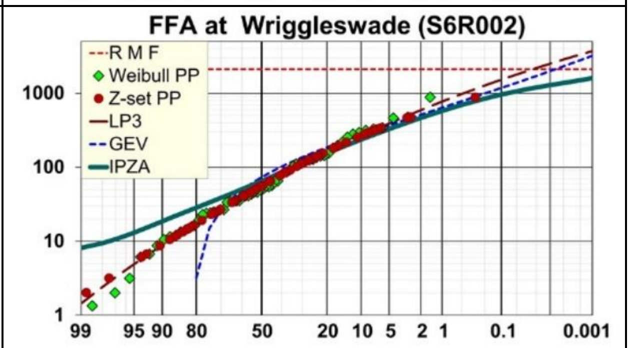
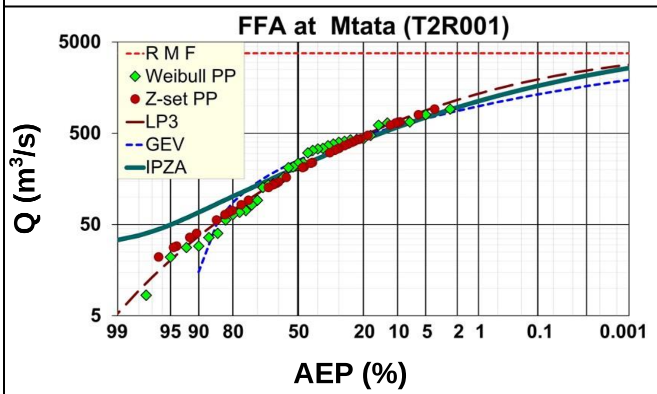
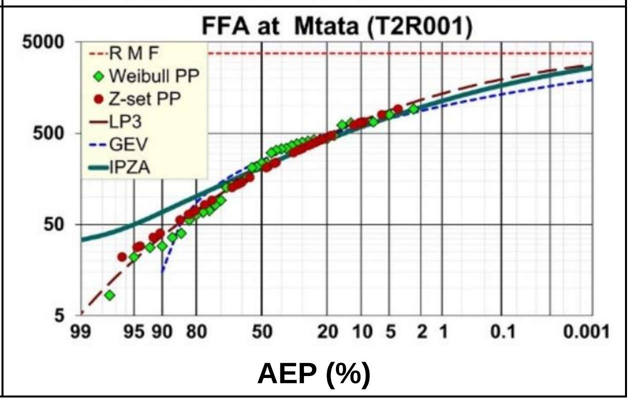
Q_{ave}, SD (ave0)
Q_{ave}, SD, g (ave1)
FFA at Kalkfontein (C5R002)
FFA at Rustfontein (C5R003)
FFA at Koppies (C7R001)
FFA at Bloemhof (C9R002)
FFA at Gariep (D3R002)
R M F
Weibull PP
Z-set PP
LP3
GEV
IPZA
Q (m³/s)
AEP (%)
99 95 90 80 50 20 10 5 2 1 0.1 0.001

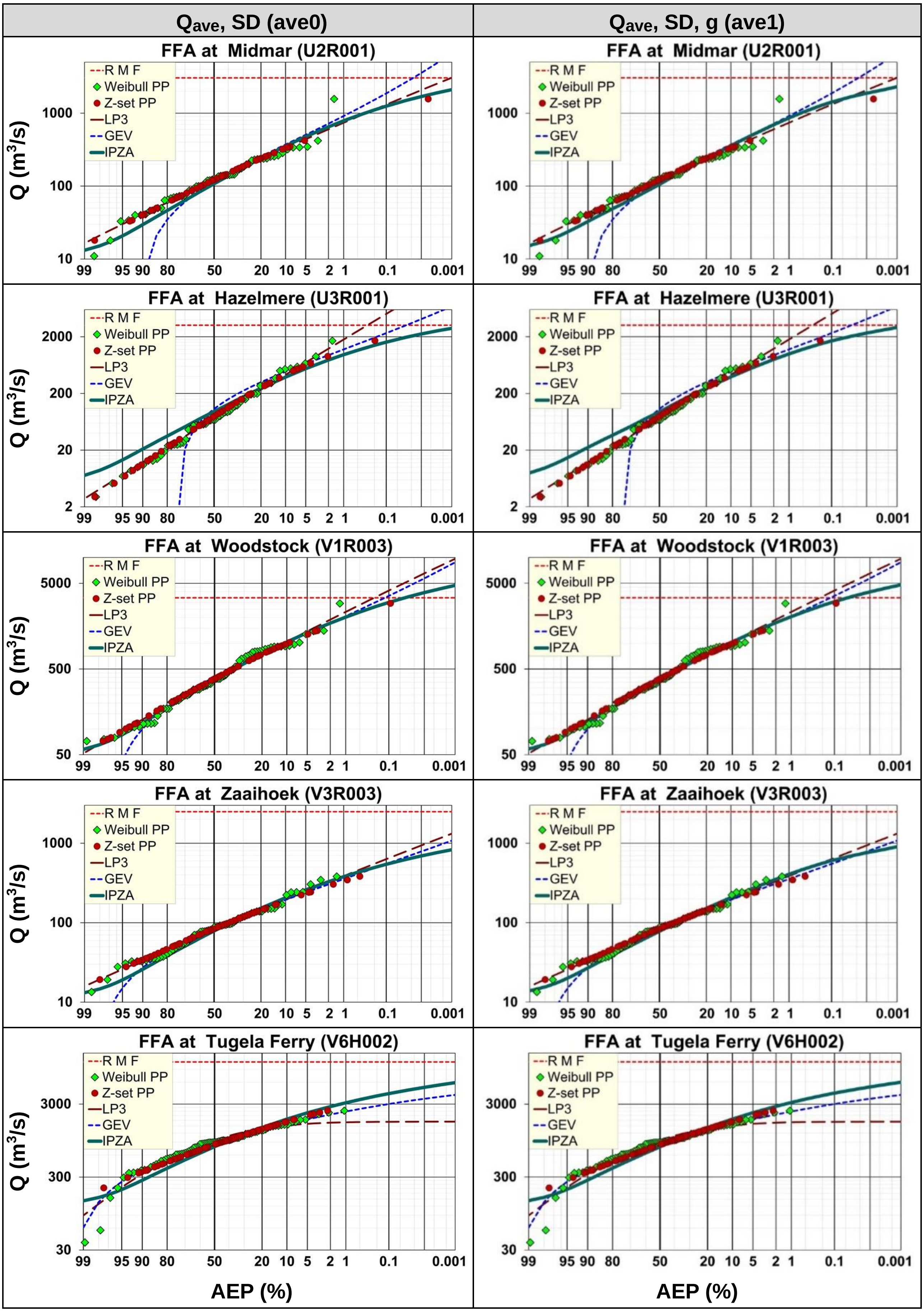


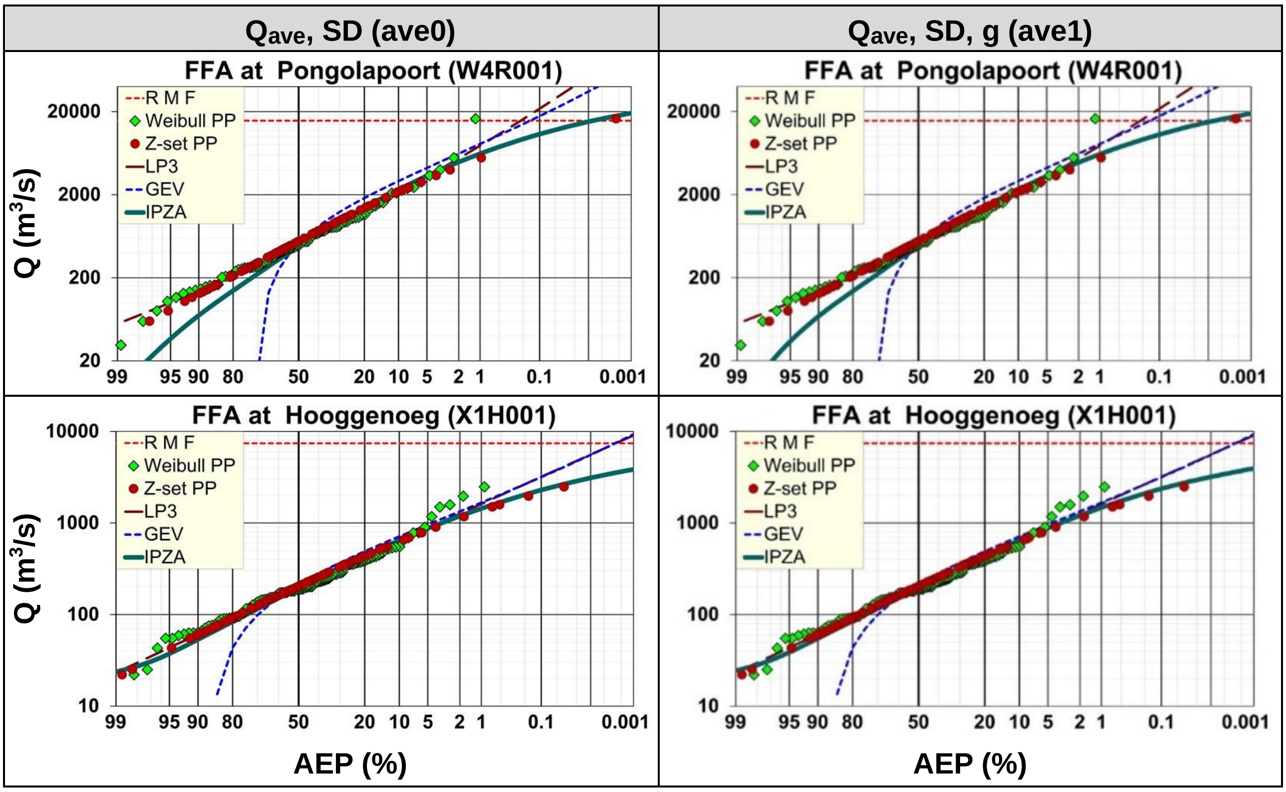
Qave, SD (ave0)
Qave, SD, g (ave1)
FFA at Smartt Syndicate (D6R002)
FFA at Upington (D7H005)
FFA at Clanwilliam (E1R002)
FFA at Wolvendrift (H5H004)
FFA at Floriskraal (J1R003)
R M F
Weibull PP
Z-set PP
LP3
GEV
IPZA
Q (m³/s)
AEP (%)
99
95
90
80
50
20
10
5
2
1
0.1
0.001

Q_{ave}, SD (ave0)	Q_{ave}, SD, g (ave1)



Q_{ave}, SD (ave0)	Q_{ave}, SD, g (ave1)
	
	
	
	
	

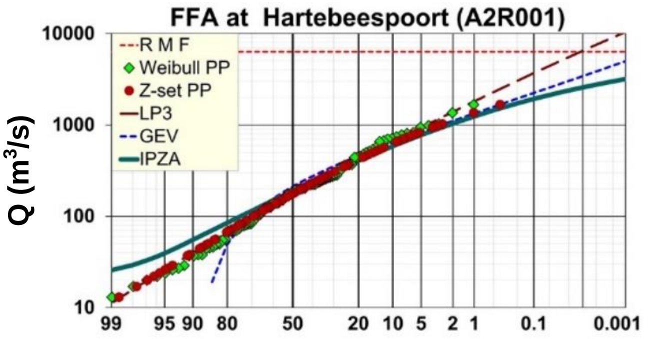

Qave, SD (ave0)
Qave, SD, g (ave1)
FFA at Midmar (U2R001)
FFA at Hazelmere (U3R001)
FFA at Woodstock (V1R003)
FFA at Zaaihoek (V3R003)
FFA at Tugela Ferry (V6H002)
R M F
Weibull PP
Z-set PP
LP3
GEV
IPZA
Q (m3/s)
99 95 90 80 50 20 10 5 2 1 0.1 0.001
AEP (%)

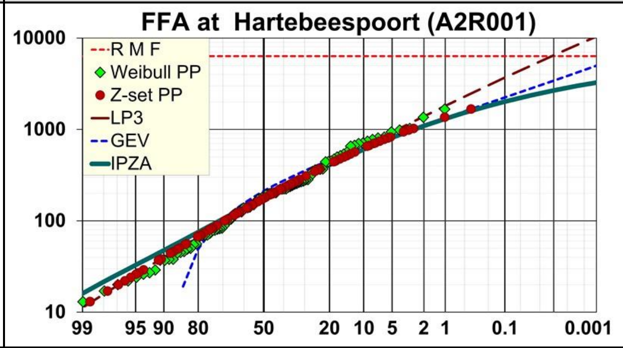

Qave, SD (ave0)
Qave, SD, g (ave1)
FFA at Pongolapoort (W4R001)
FFA at Hooggenoeg (X1H001)
R M F
Weibull PP
Z-set PP
LP3
GEV
IPZA
Q (m3/s)
AEP (%)
99 95 90 80 50 20 10 5 2 1 0.1 0.001
20000
2000
200
20
10000
1000
100
10

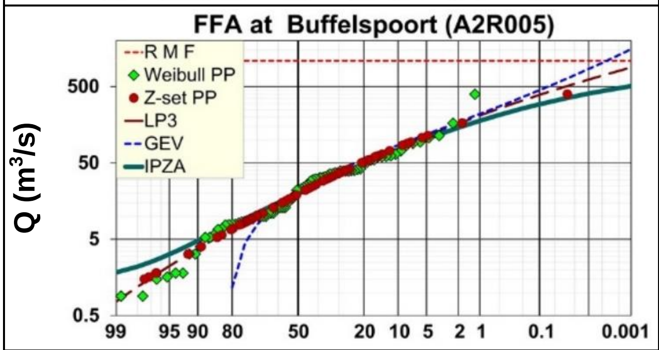
APPENDIX 6-G: Visual comparison of ave0 with ave2

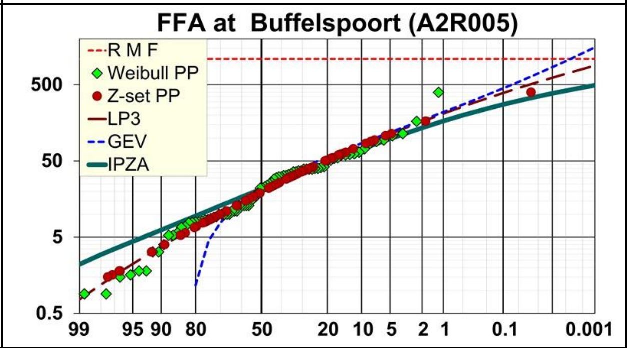
Visually compare the ave0-scenario, which included the Q_{ave} and the SD, against the ave2-scenario, which included the Q_{ave}, SD and SD*.

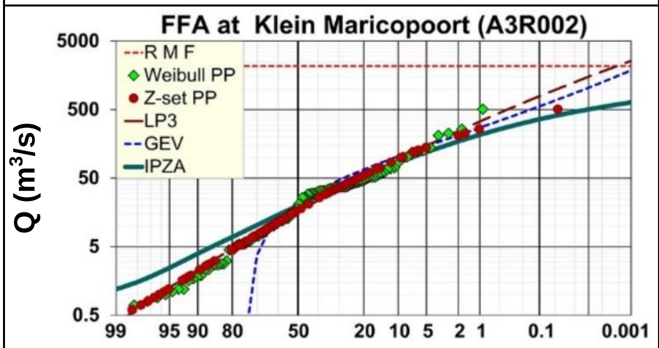
Q_{ave}, SD (ave0) | **Q_{ave}, SD, SD* (ave2)**

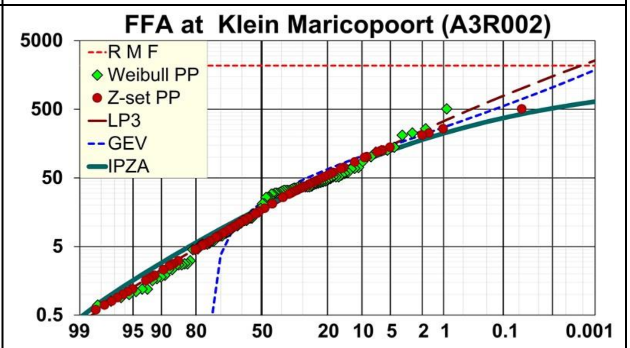


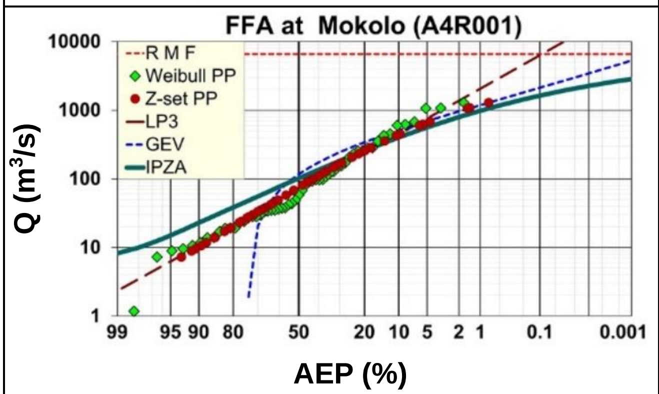


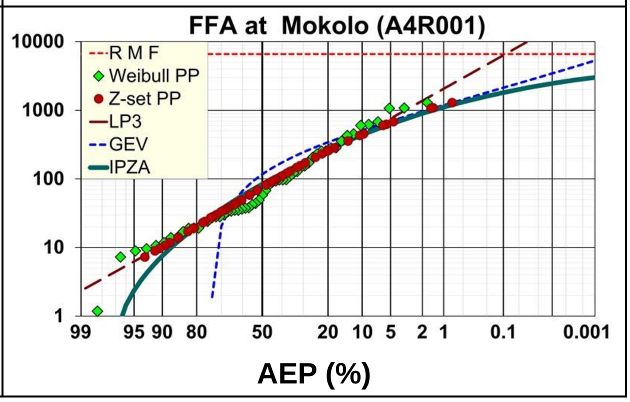


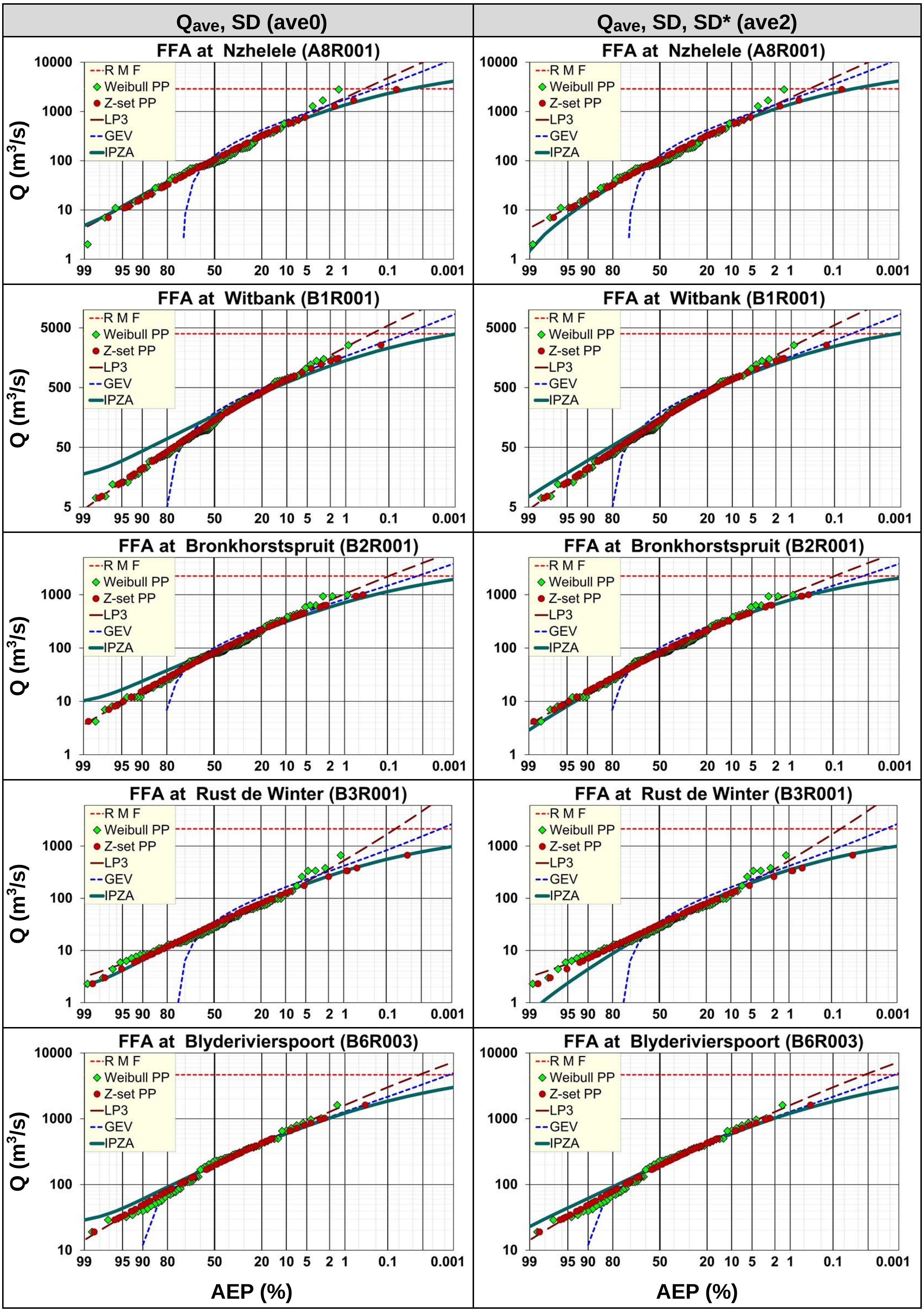




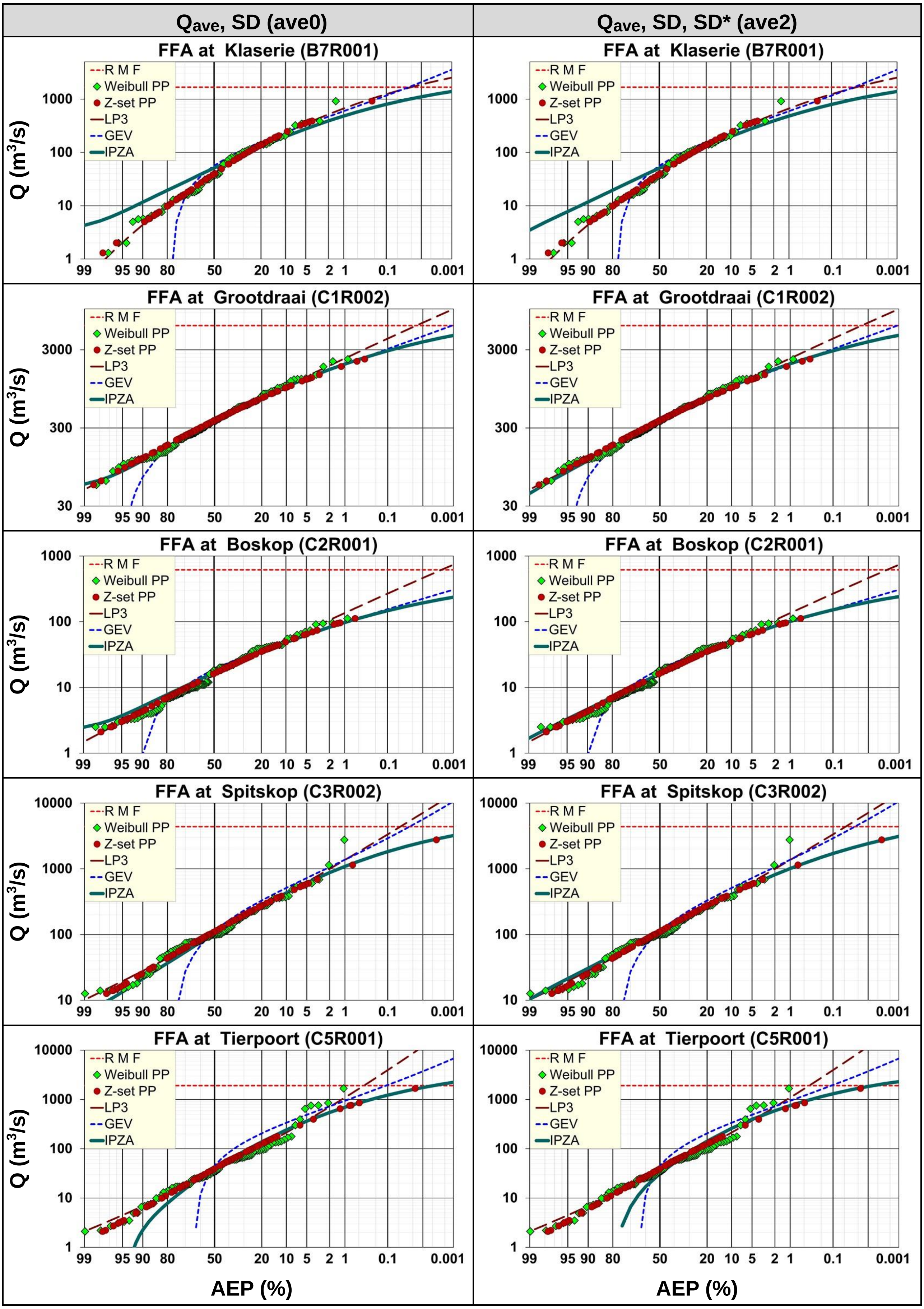






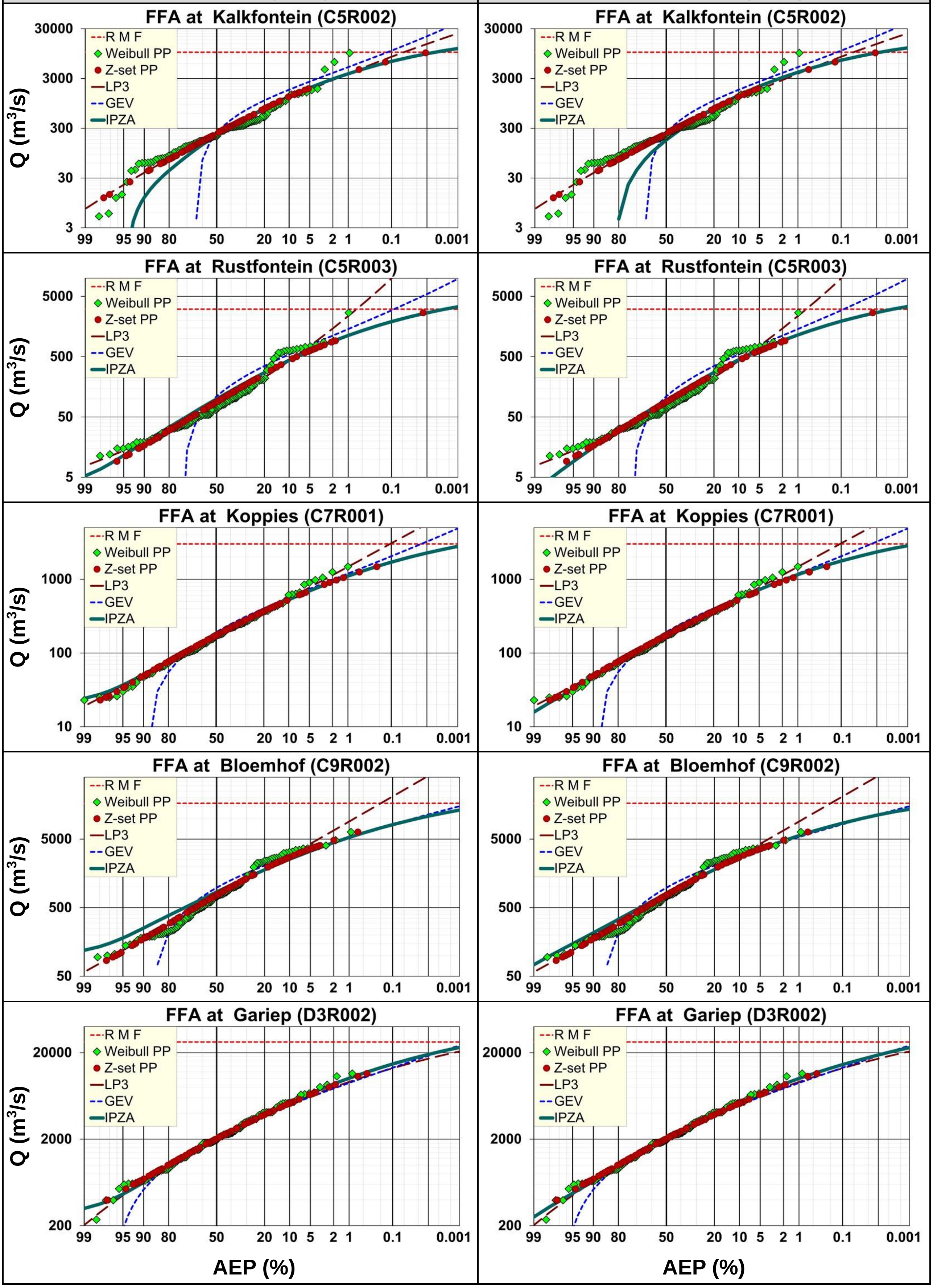


Qave, SD (ave0)
Qave, SD, SD* (ave2)
FFA at Nzhelele (A8R001)
FFA at Witbank (B1R001)
FFA at Bronkhorstspruit (B2R001)
FFA at Rust de Winter (B3R001)
FFA at Blyderivierspoort (B6R003)
R M F
Weibull PP
Z-set PP
LP3
GEV
IPZA
Q (m³/s)
99 95 90 80 50 20 10 5 2 1 0.1 0.001
AEP (%)

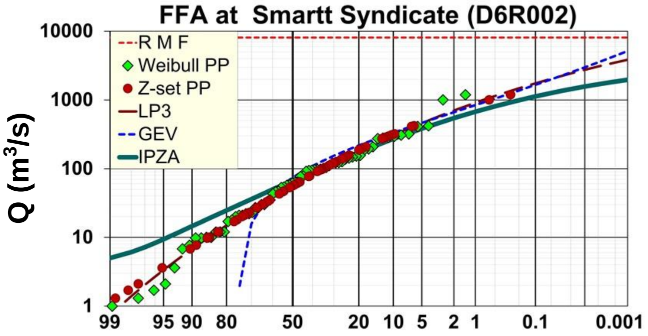


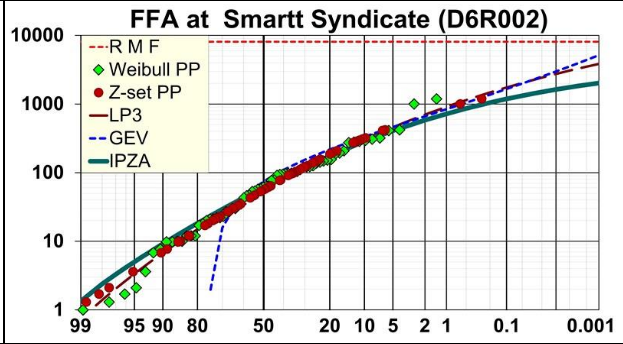
Qave, SD (ave0)
Qave, SD, SD* (ave2)
FFA at Klaserie (B7R001)
FFA at Grootdraai (C1R002)
FFA at Boskop (C2R001)
FFA at Spitskop (C3R002)
FFA at Tierpoort (C5R001)
R M F
Weibull PP
Z-set PP
LP3
GEV
IPZA
Q (m3/s)
99 95 90 80 50 20 10 5 2 1 0.1 0.001
AEP (%)

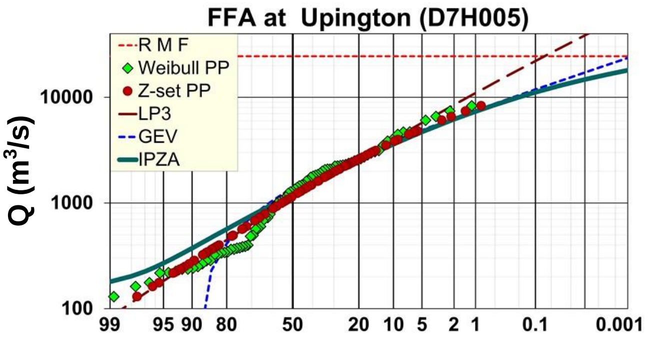
Q_{ave}, SD (ave0)	Q_{ave}, SD, SD* (ave2)

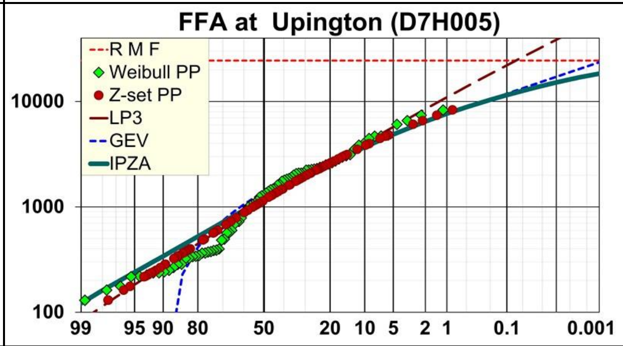


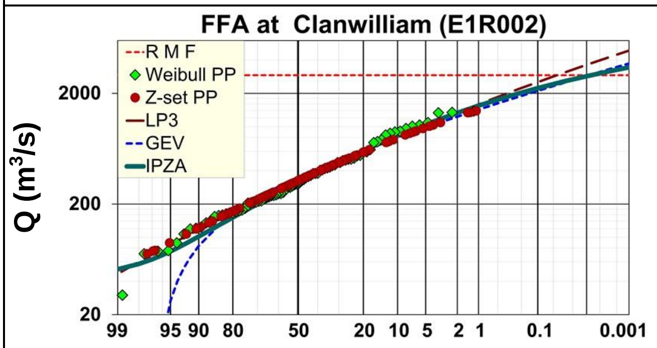
Q_{ave}, SD (ave0)	Q_{ave}, SD, SD* (ave2)

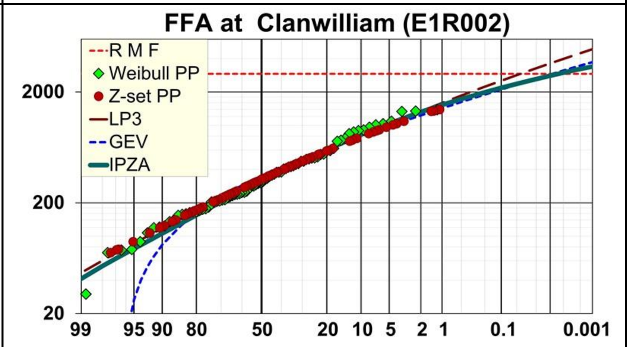


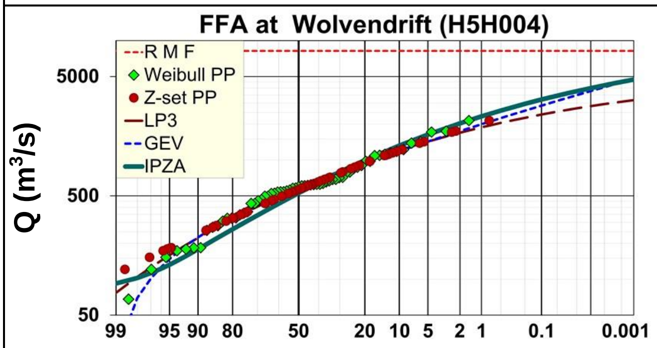


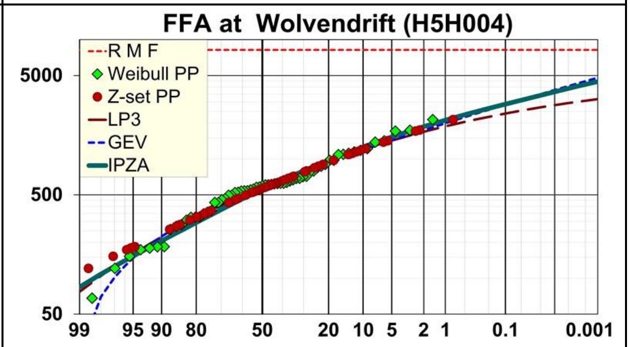


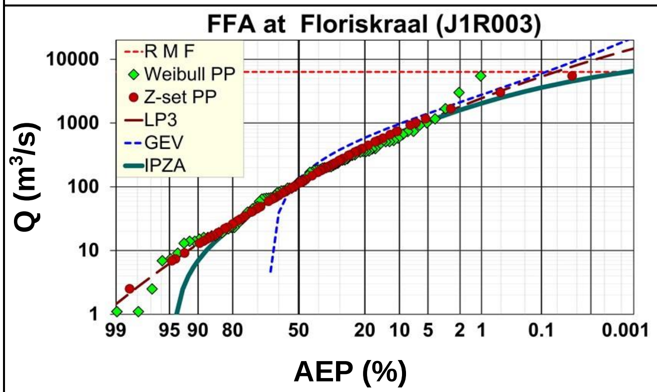


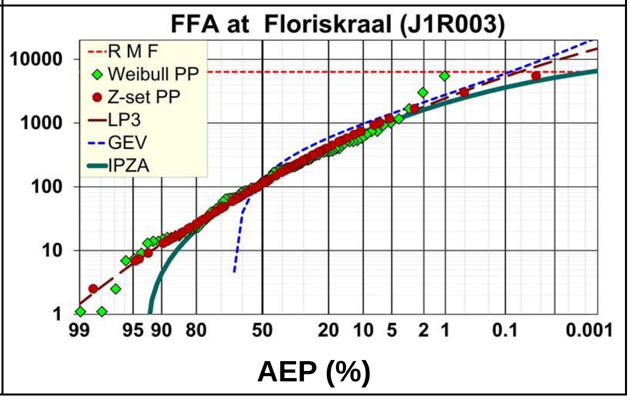


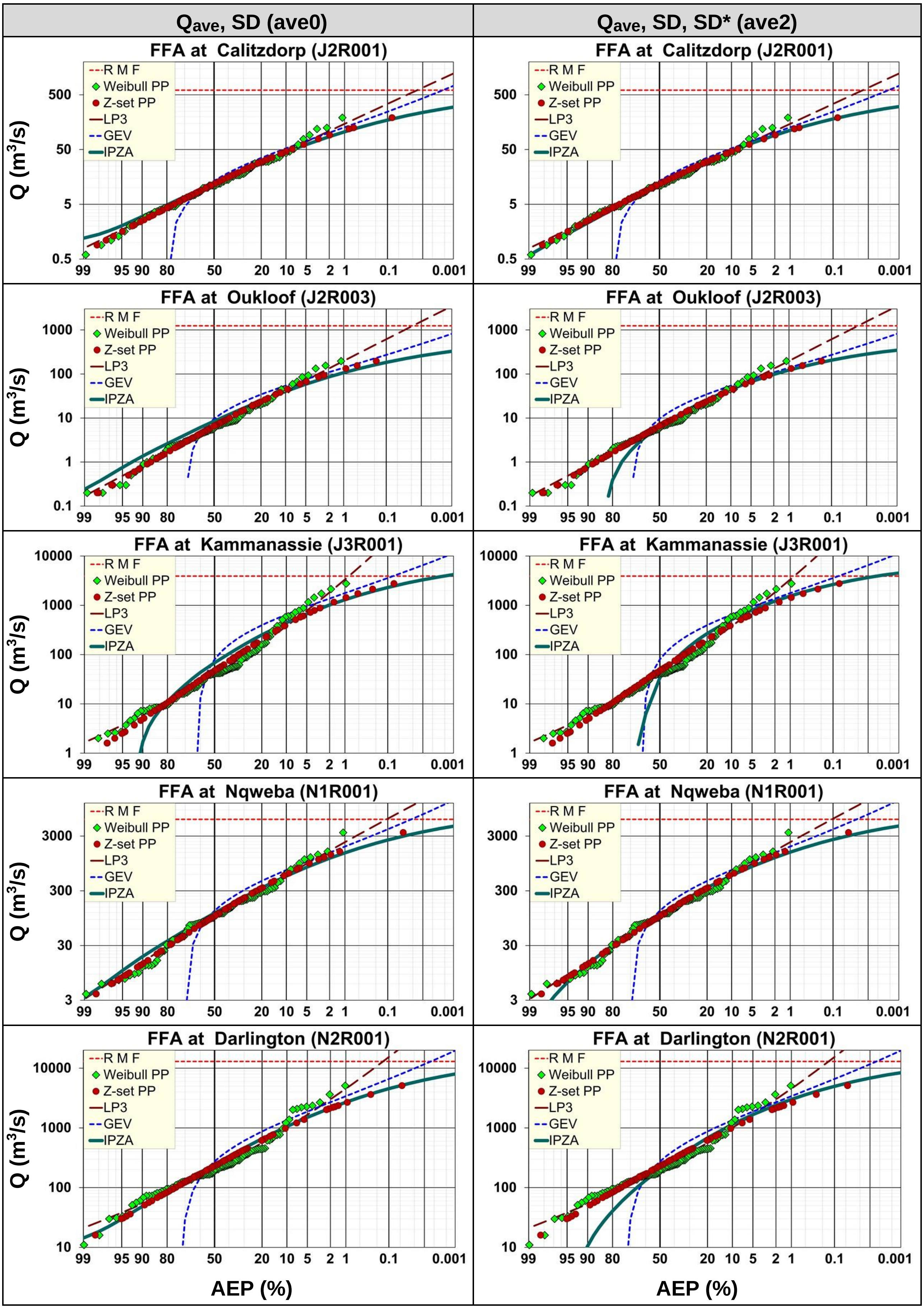


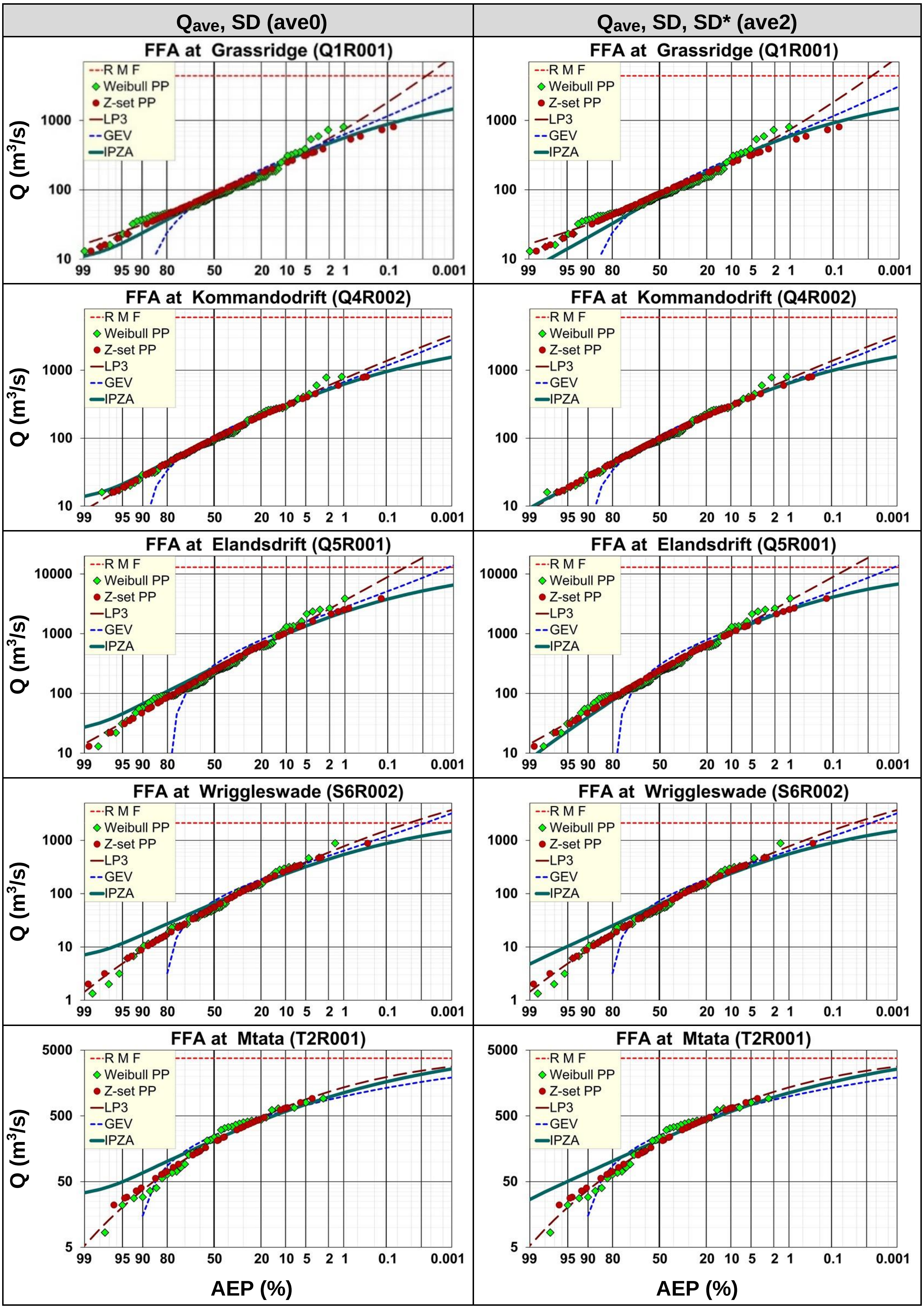


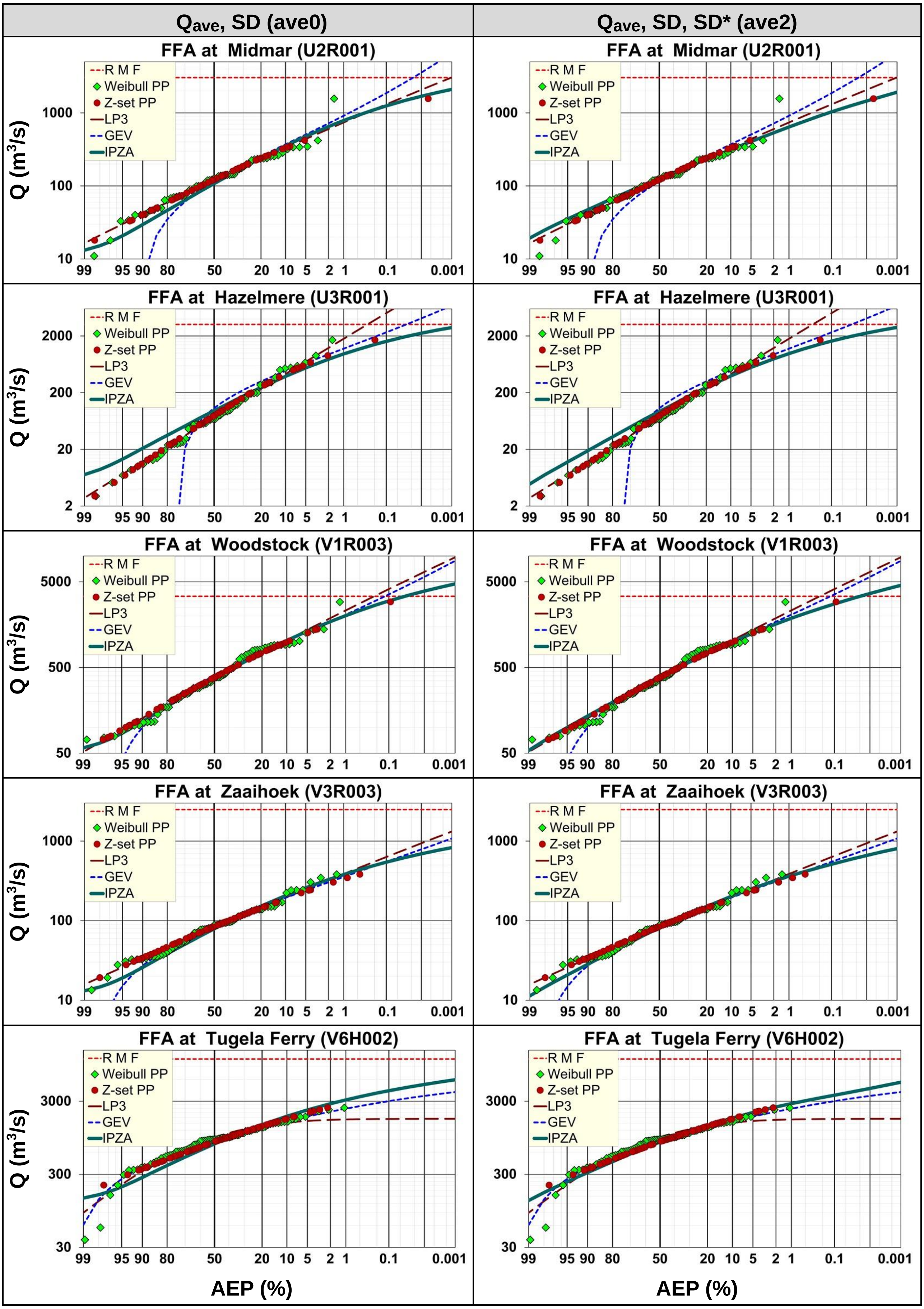


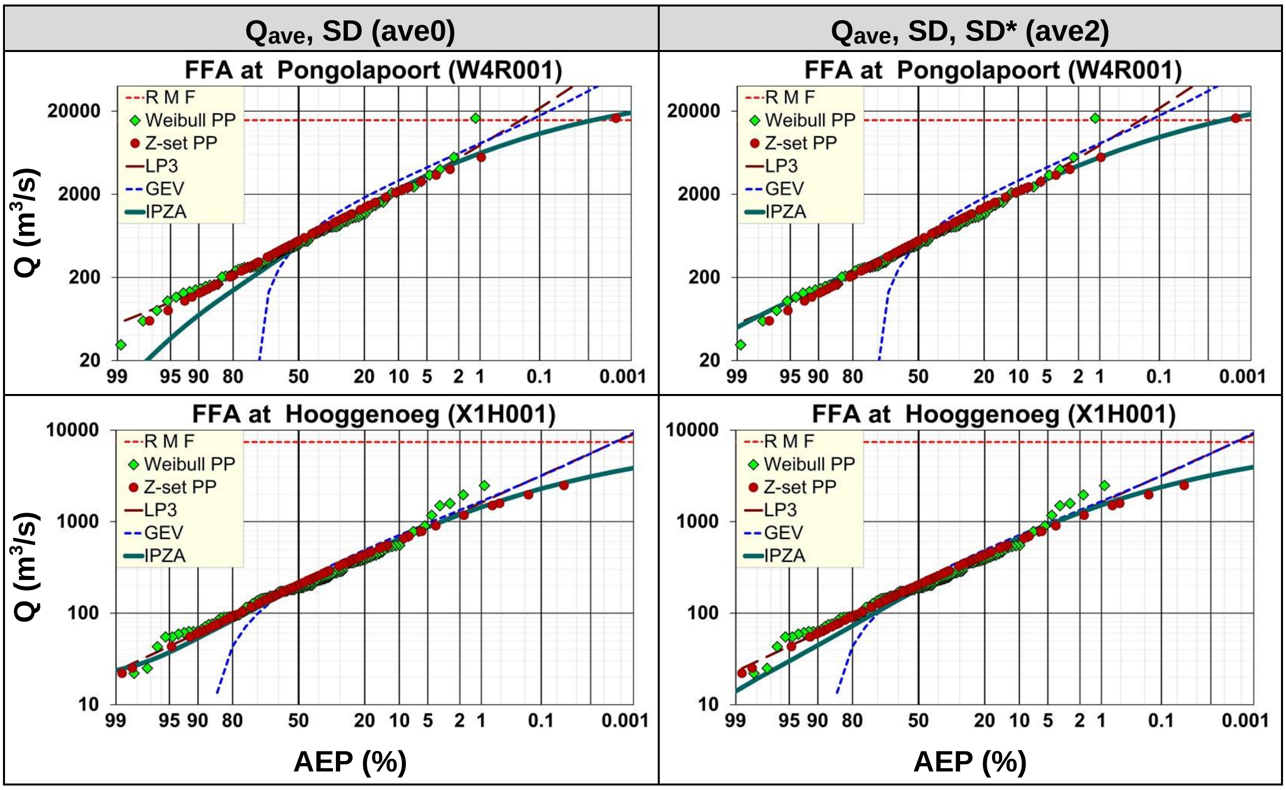





Qave, SD (ave0)
Qave, SD, SD* (ave2)
FFA at Calitzdorp (J2R001)
FFA at Oukloof (J2R003)
FFA at Kammanassie (J3R001)
FFA at Nqweba (N1R001)
FFA at Darlington (N2R001)
R M F
Weibull PP
Z-set PP
LP3
GEV
IPZA
Q (m3/s)
99 95 90 80 50 20 10 5 2 1 0.1 0.001
AEP (%)


Qave, SD (ave0)
Qave, SD, SD* (ave2)
FFA at Grassridge (Q1R001)
FFA at Kommandodrift (Q4R002)
FFA at Elandsdrift (Q5R001)
FFA at Wriggleswade (S6R002)
FFA at Mtata (T2R001)
R M F
Weibull PP
Z-set PP
LP3
GEV
IPZA
Q (m3/s)
99 95 90 80 50 20 10 5 2 1 0.1 0.001
AEP (%)

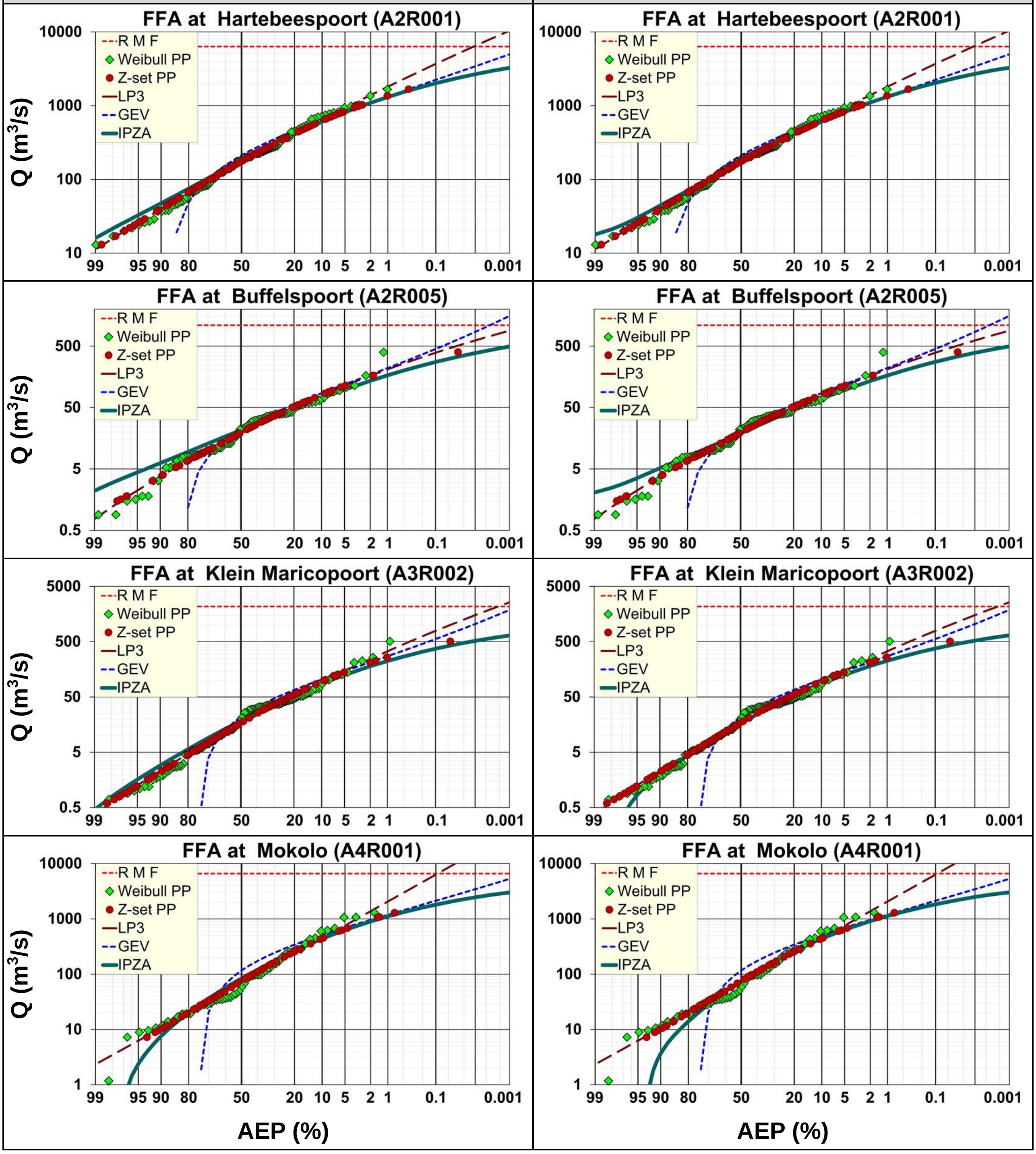

Qave, SD (ave0)
Qave, SD, SD* (ave2)
FFA at Midmar (U2R001)
FFA at Hazelmere (U3R001)
FFA at Woodstock (V1R003)
FFA at Zaaihoek (V3R003)
FFA at Tugela Ferry (V6H002)
R M F
Weibull PP
Z-set PP
LP3
GEV
IPZA
Q (m3/s)
AEP (%)
99 95 90 80 50 20 10 5 2 1 0.1 0.001


Qave, SD (ave0)
Qave, SD, SD* (ave2)
FFA at Pongolapoort (W4R001)
FFA at Hooggenoeg (X1H001)
R M F
Weibull PP
Z-set PP
LP3
GEV
IPZA
Q (m3/s)
20000
2000
200
20
10000
1000
100
10
99
95 90 80
50
20 10 5 2 1
0.1
0.001
AEP (%)

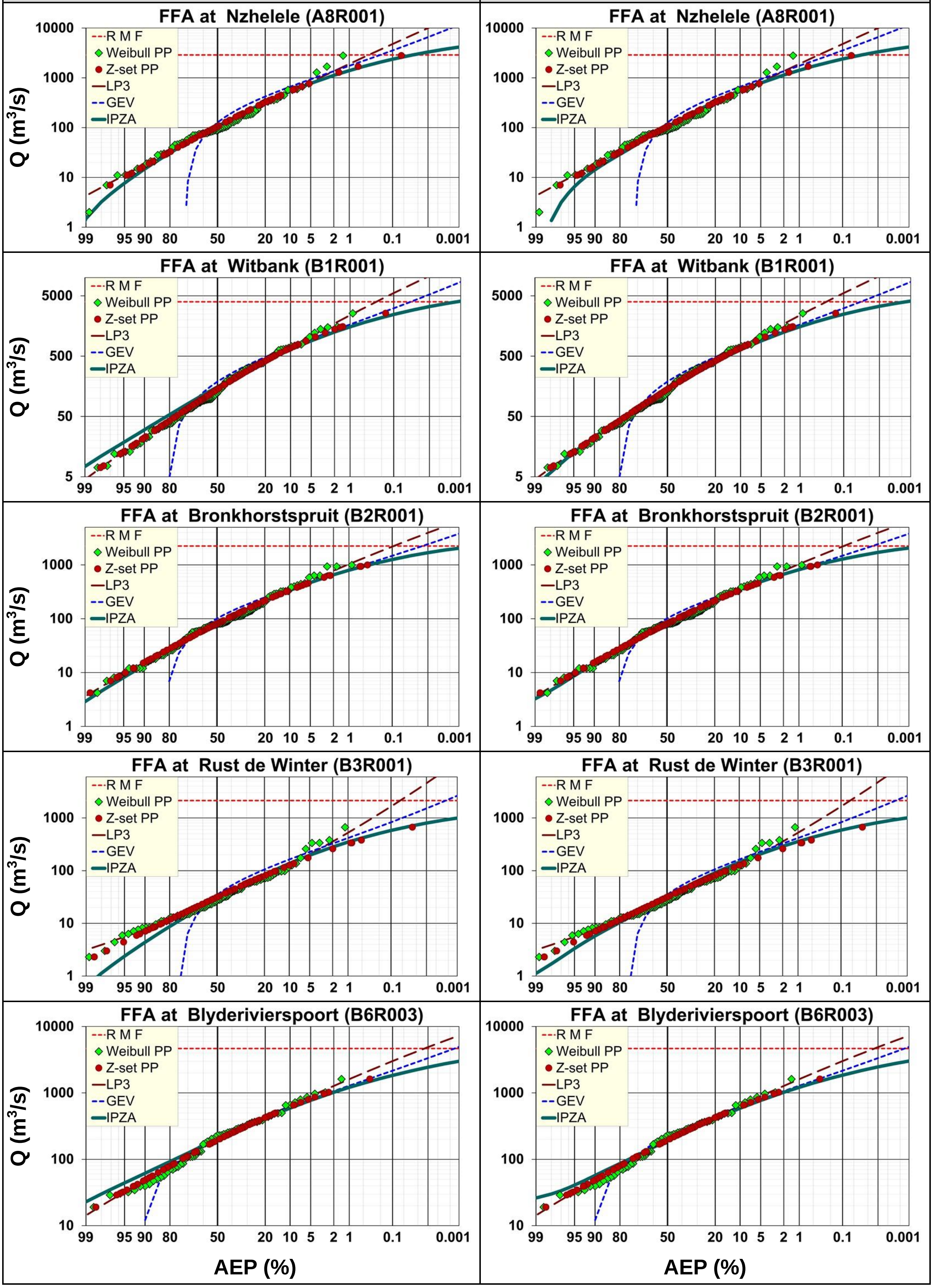
APPENDIX 6-H: Visual comparison of ave2 with gmn2

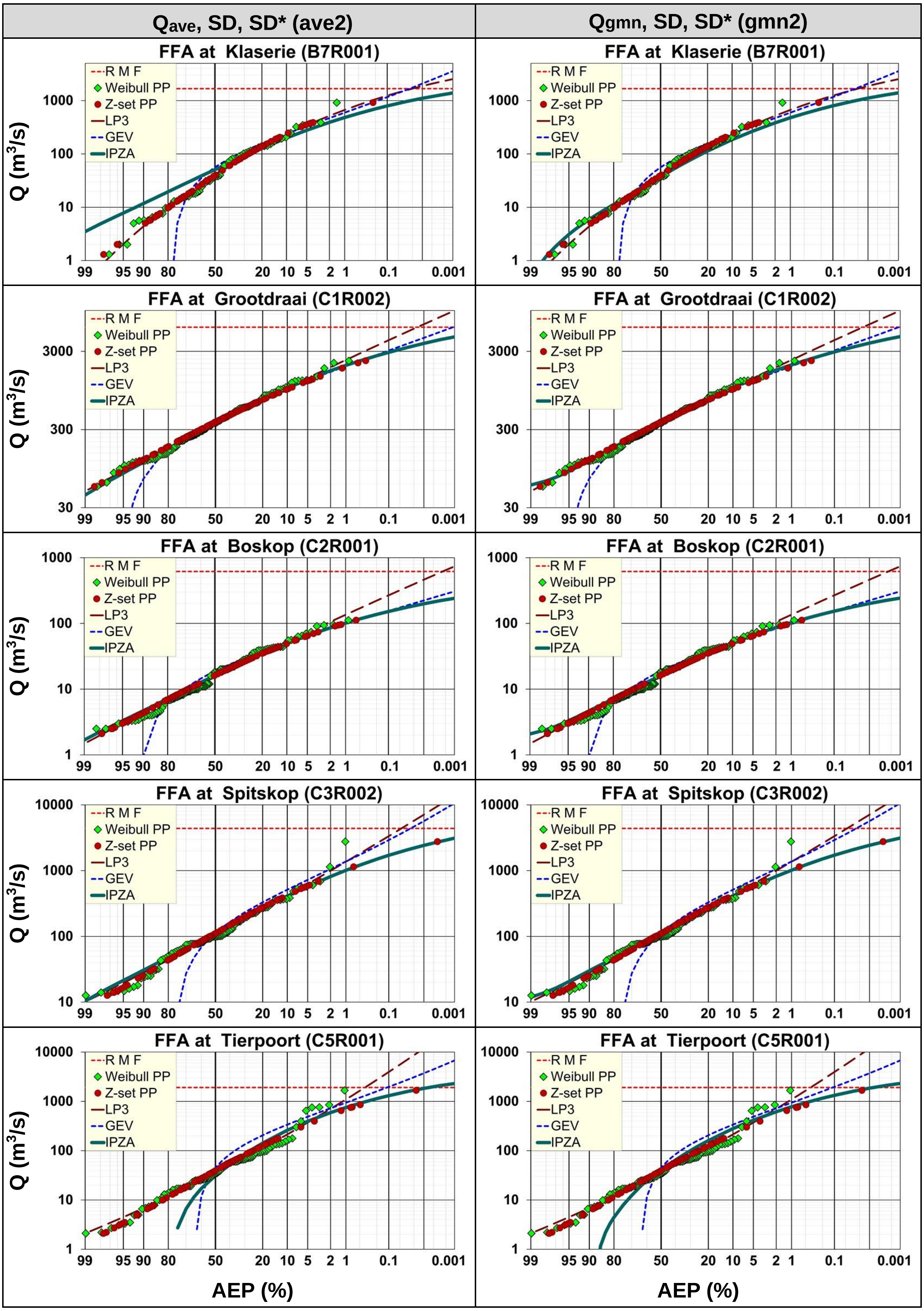
Visually compare the ave2 scenario, which included the Q_{ave}, SD and SD*, against the gmn2 scenario, which included the Q_{gmn}, SD and SD*.

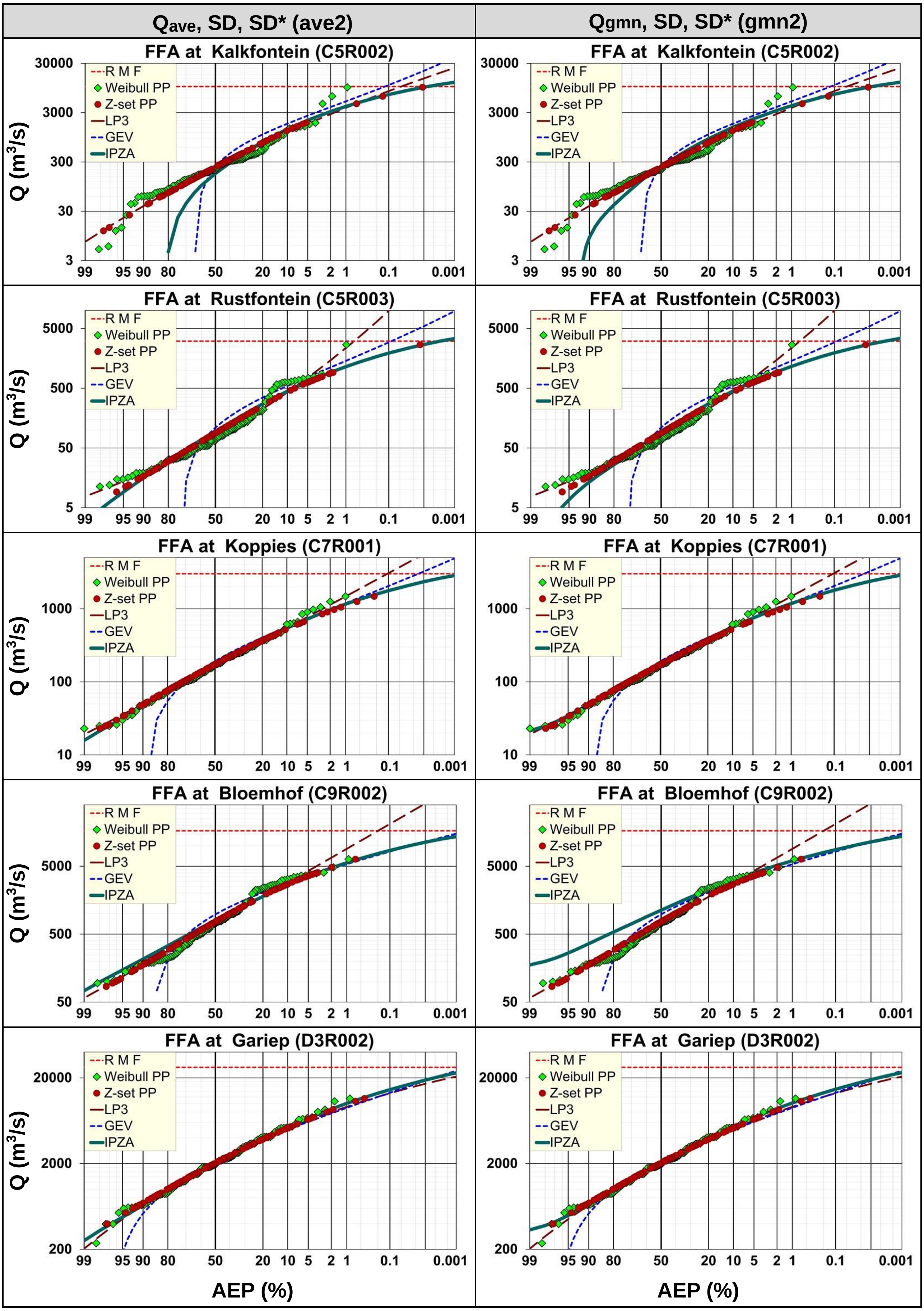
Q_{ave}, SD, SD* (ave2)	Q_{gmn}, SD, SD* (gmn2)

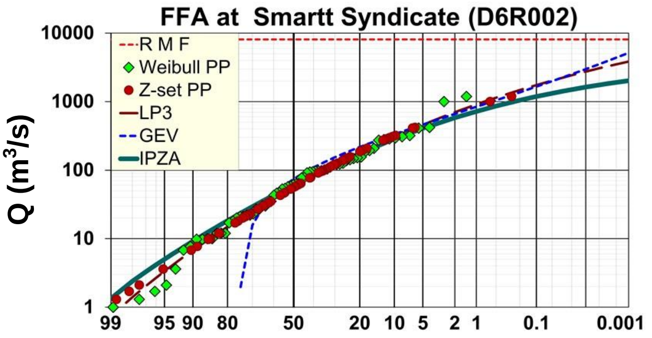
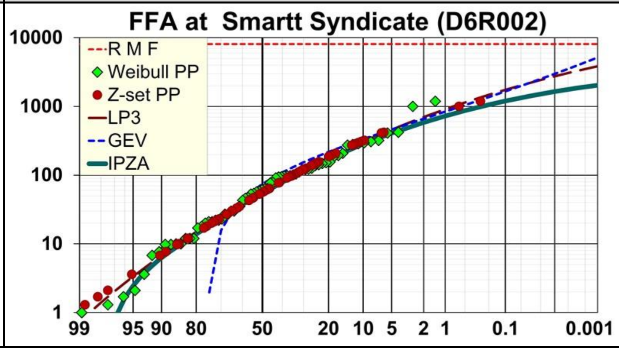
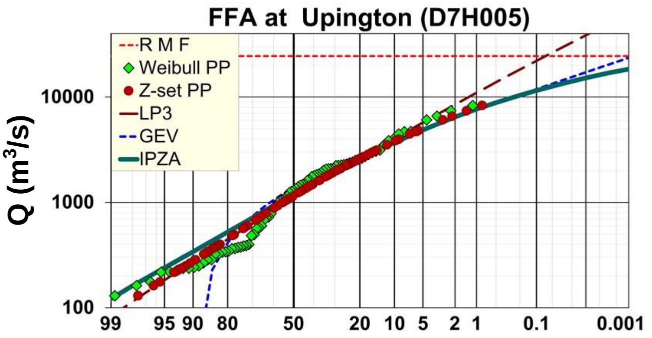
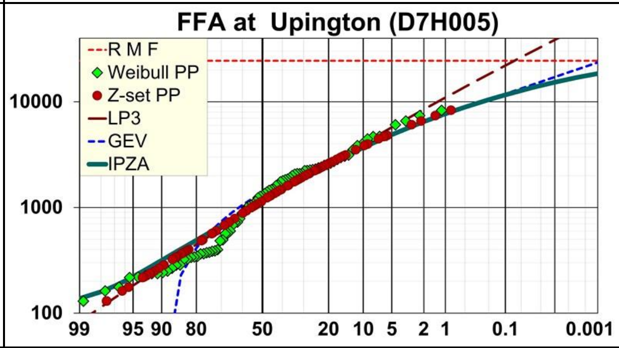
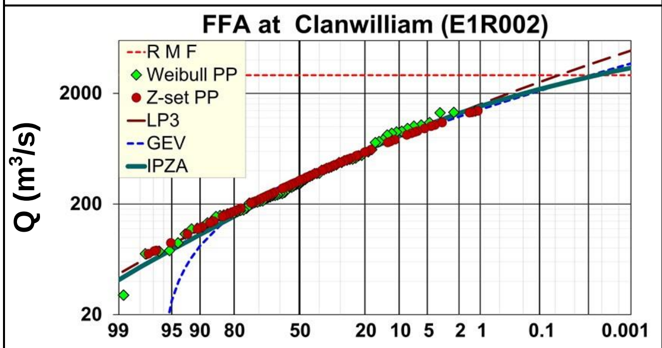
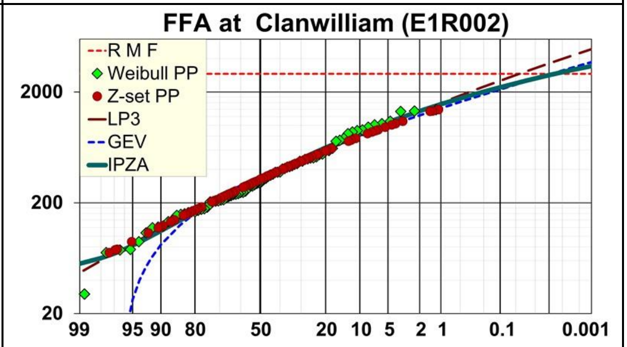
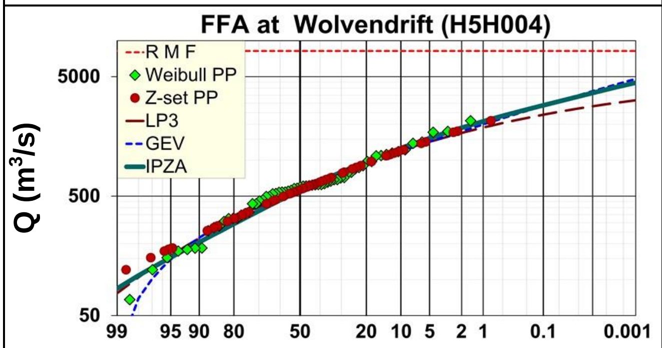
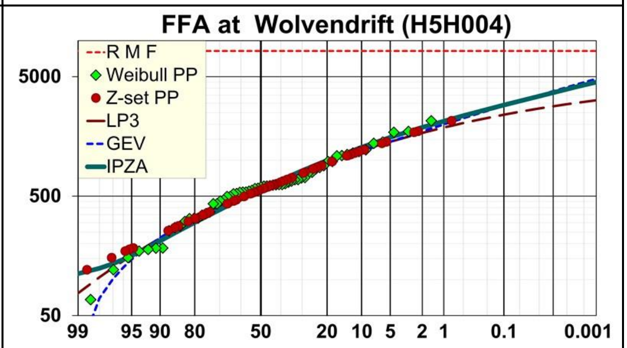
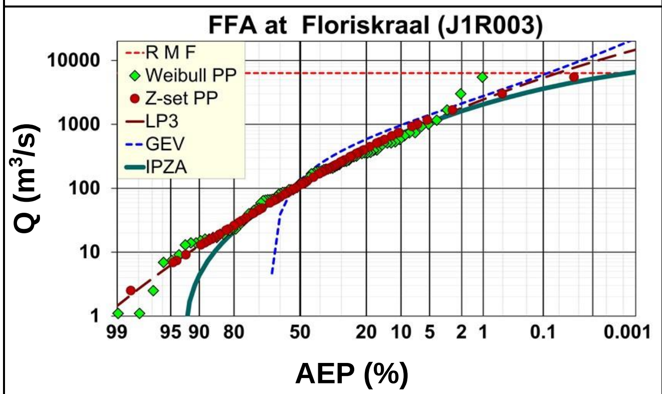
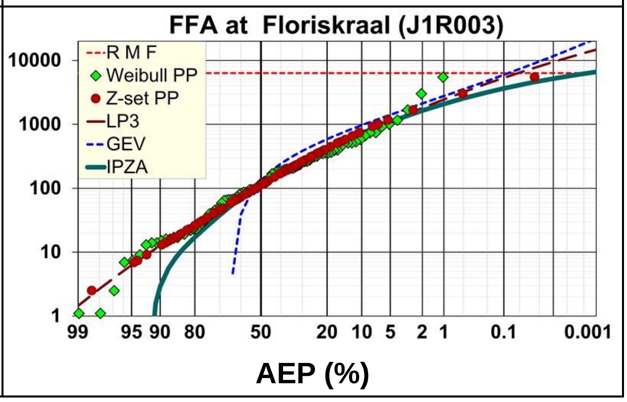


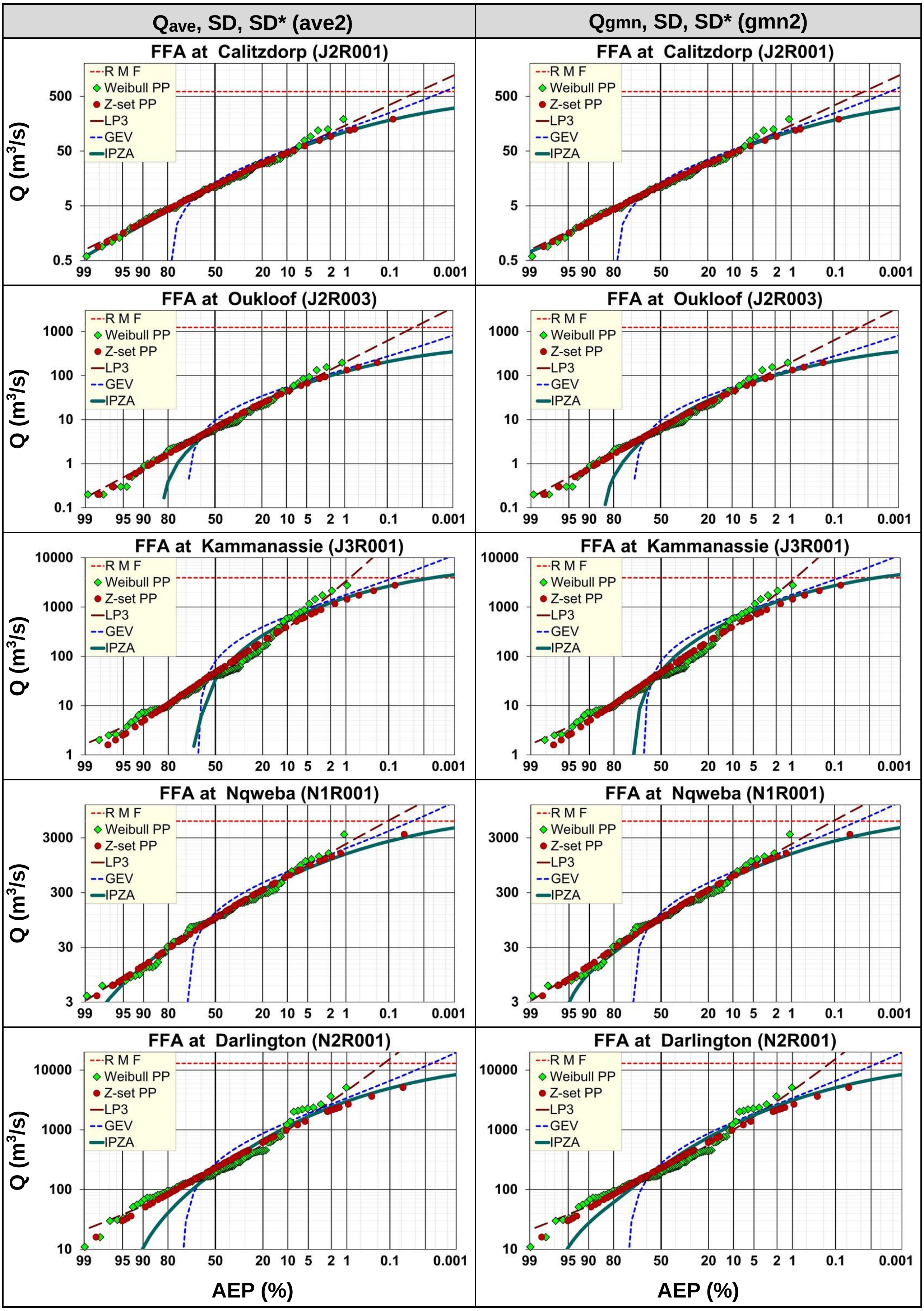
Q_{ave}, SD, SD* (ave2)	Q_{gmn}, SD, SD* (gmn2)



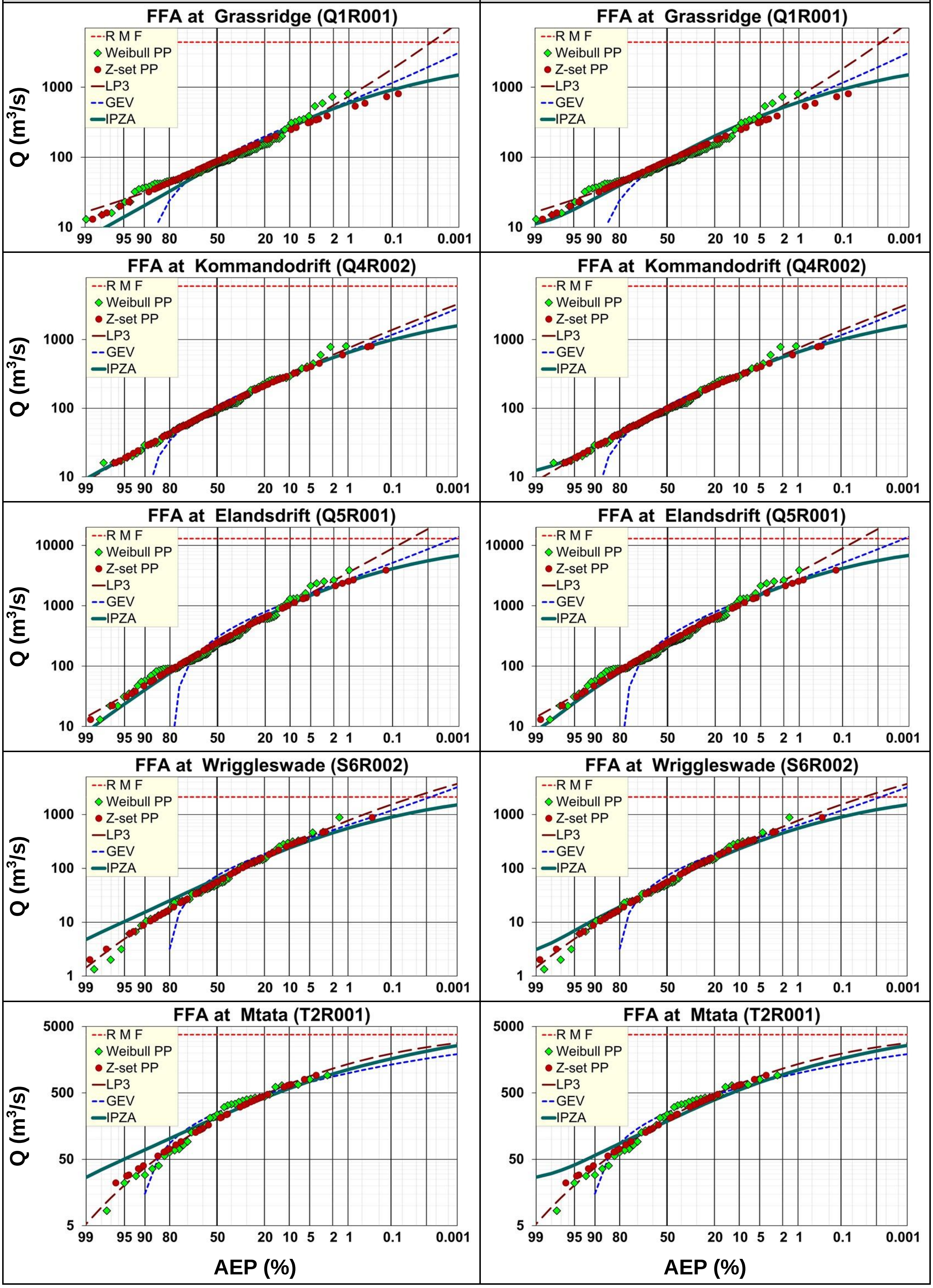

Qave, SD, SD* (ave2)
Qgmn, SD, SD* (gmn2)
FFA at Klaserie (B7R001)
FFA at Grootdraai (C1R002)
FFA at Boskop (C2R001)
FFA at Spitskop (C3R002)
FFA at Tierpoort (C5R001)
R M F
Weibull PP
Z-set PP
LP3
GEV
IPZA
Q (m3/s)
99 95 90 80 50 20 10 5 2 1 0.1 0.001
AEP (%)

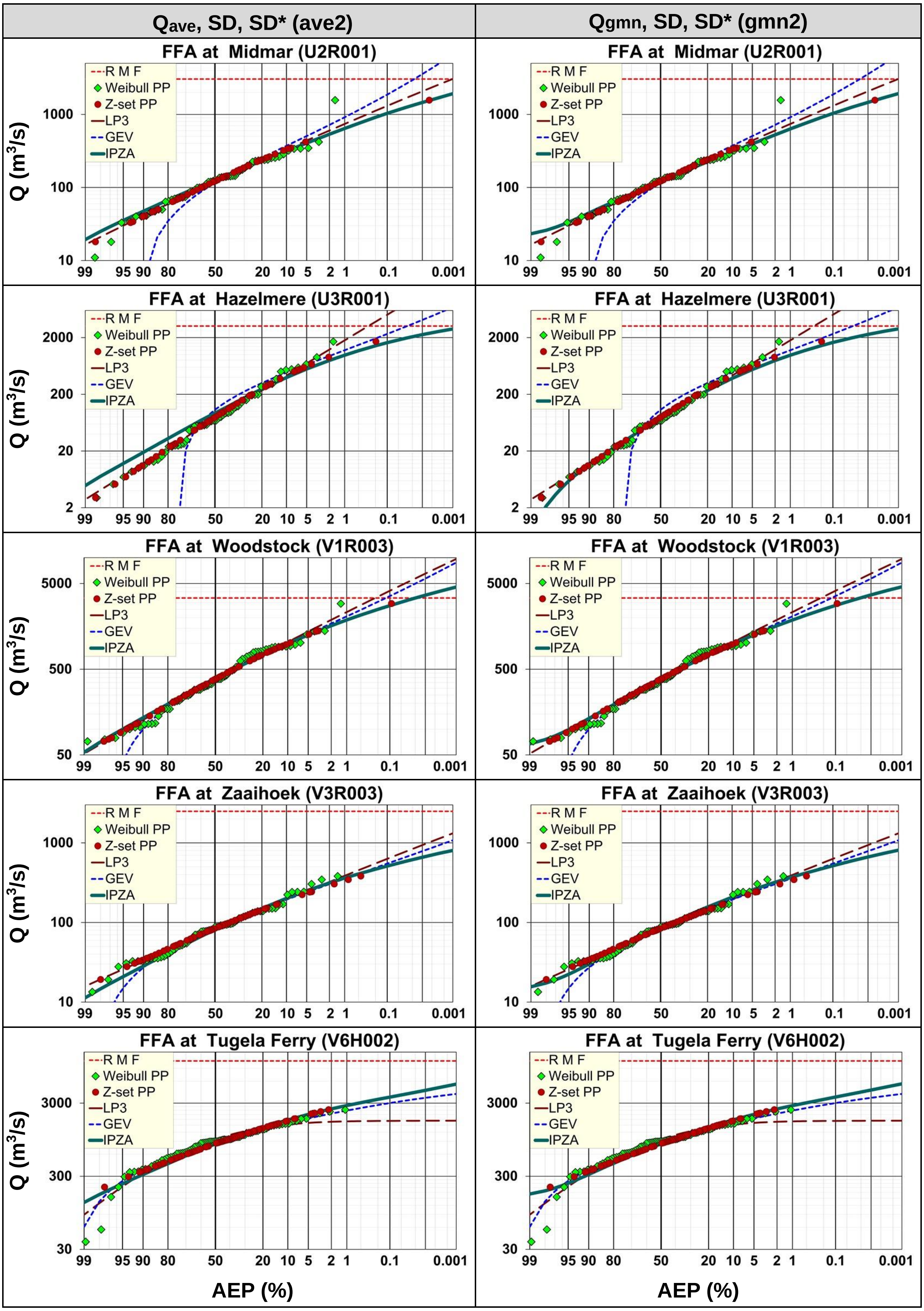

Qave, SD, SD* (ave2)
Qgmn, SD, SD* (gmn2)
FFA at Kalkfontein (C5R002)
FFA at Rustfontein (C5R003)
FFA at Koppies (C7R001)
FFA at Bloemhof (C9R002)
FFA at Gariep (D3R002)
R M F
Weibull PP
Z-set PP
LP3
GEV
IPZA
Q (m³/s)
AEP (%)
99
95
90
80
50
20
10
5
2
1
0.1
0.001
30000
3000
300
30
3
5000
500
50
5
1000
100
10
20000
2000
200

Qave, SD, SD* (ave2)	Qgmn, SD, SD* (gmn2)
	
	
	
	
	

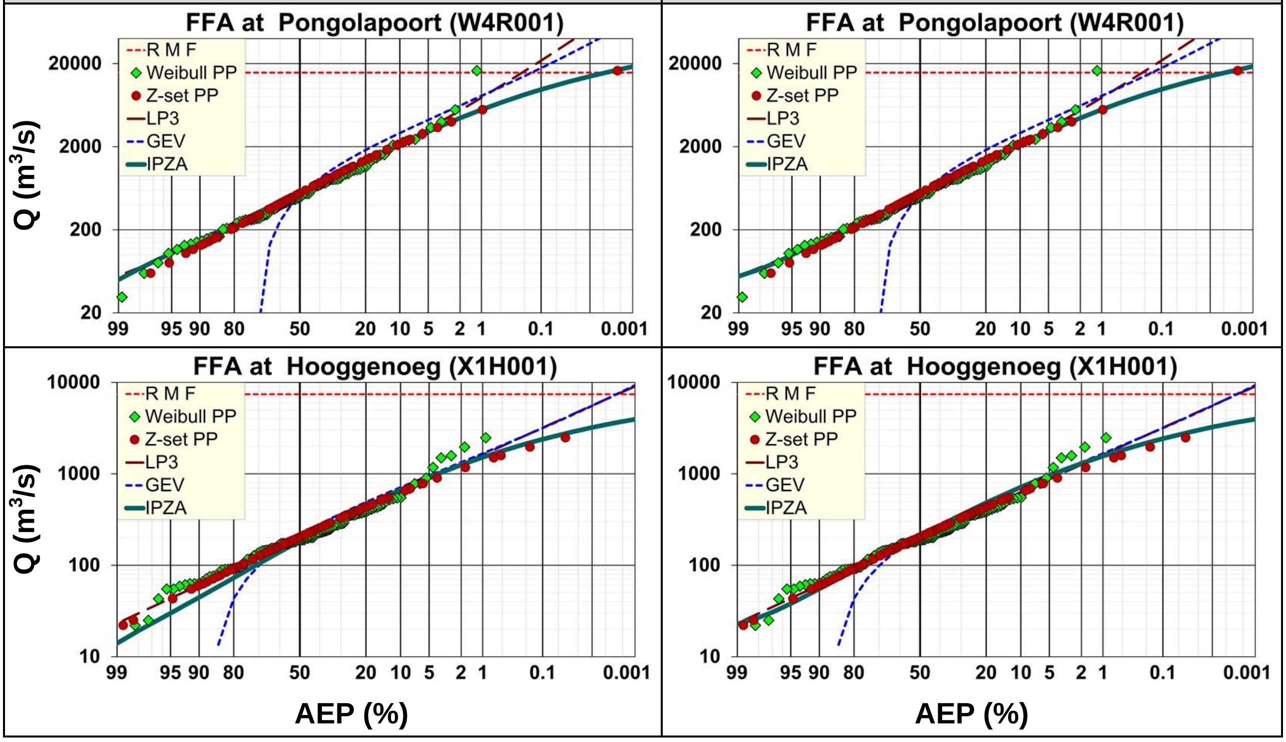

Qave, SD, SD* (ave2)
Qgmn, SD, SD* (gmn2)
FFA at Calitzdorp (J2R001)
FFA at Oukloof (J2R003)
FFA at Kammanassie (J3R001)
FFA at Nqweba (N1R001)
FFA at Darlington (N2R001)
R M F
Weibull PP
Z-set PP
LP3
GEV
IPZA
Q (m³/s)
AEP (%)
99 95 90 80 50 20 10 5 2 1 0.1 0.001

Qave, SD, SD* (ave2)	Qgmn, SD, SD* (gmn2)




Qave, SD, SD* (ave2)
Qgmn, SD, SD* (gmn2)
FFA at Midmar (U2R001)
FFA at Hazelmere (U3R001)
FFA at Woodstock (V1R003)
FFA at Zaaihoek (V3R003)
FFA at Tugela Ferry (V6H002)
R M F
Weibull PP
Z-set PP
LP3
GEV
IPZA
Q (m³/s)
99 95 90 80 50 20 10 5 2 1 0.1 0.001
AEP (%)

Qave, SD, SD* (ave2)	Qgmn, SD, SD* (gmn2)



APPENDIX 6-I: MRA results for selected scenario ave2

Selected scenario ave2: MRA results, comparing Q_{est} against Q_{obs}, for all selected AEPs.

Site	Moments			AEP (%)											
				99.5		99		95		90		80		70	
	Q and SD values in m^3/s														
	Q_{ave}	SD	SD*	Q_{obs}	Q_{est}	Q_{obs}	Q_{est}	Q_{obs}	Q_{est}	Q_{obs}	Q_{est}	Q_{obs}	Q_{est}	Q_{obs}	Q_{est}
A3R002	37.4	63.3	44.7	0.400	0.271	0.472	0.456	1.20	1.59	2.27	2.93	4.51	5.66	7.46	9.00
B1R001	280	384	317	3.13	5.27	5.18	7.39	13.0	18.6	22.2	30.2	42.3	53.5	67.8	80.9
B2R001	144	188	171	2.57	1.98	3.68	2.88	9.46	8.07	15.3	13.8	26.8	25.6	41.3	39.8
C1R002	494	421	388	38.0	33.8	48.2	43.6	90.4	83.1	123	115	181	171	240	229
C2R001	23.2	22.0	20.4	1.19	1.29	1.46	1.68	3.01	3.30	4.29	4.67	6.91	7.14	9.51	9.77
C5R002	547	1 223	824	4.89	-20.8	7.06	-24.2	20.8	-27.9	37.9	-19.8	71.5	4.33	112	42.8
C9R002	1 216	1 251	1 154	43.9	56.1	56.7	73.5	114	149	178	215	292	338	423	471
D3R002	2 614	2 094	1 928	208	196	268	252	505	475	689	652	1 006	959	1 319	1 271
D7H005	1 730	1 690	1 540	67.2	93.4	88.6	122	191	240	278	340	449	522	647	716
J1R003	298	665	398	1.11	-5.04	1.66	-5.26	6.59	-2.23	12.6	4.90	25.9	21.0	45.9	43.1
J2R001	20.4	29.1	22.8	0.464	0.417	0.662	0.582	1.57	1.42	2.49	2.29	4.29	3.98	6.29	5.97
N2R001	467	796	641	10.4	-6.85	11.9	-7.20	30.2	-1.23	48.2	11.1	83.5	40.3	124	80.2
Q1R001	123	139	121	5.58	5.12	11.1	6.75	21.7	13.9	29.1	20.4	43.1	32.4	56.5	45.6
V6H002	949	492	469	132	103	160	131	272	238	353	316	481	446	600	571
X1H001	302	371	308	15.5	10.5	20.0	14.0	41.1	29.9	60.0	44.7	91.5	72.7	125	104
A2R001	278	301	267	8.62	12.2	11.6	16.0	25.8	32.8	38.9	47.6	67.5	75.0	97.6	105
A2R005	33.9	49.8	29.6	0.674	1.68	0.933	2.21	3.20	4.38	3.71	6.24	6.81	9.50	10.00	12.9
A4R001	181	281	240	1.11	-1.30	2.03	-1.12	5.52	2.30	10.1	7.72	18.7	20.1	29.0	36.3
A8R001	227	409	278	2.59	0.617	3.82	1.50	10.5	7.51	17.1	15.2	32.1	30.9	50.5	50.6
B3R001	58.5	96.5	70.7	1.36	0.358	1.86	0.627	4.33	2.34	6.97	4.42	11.6	8.70	16.0	14.0
B6R003	283	282	233	11.7	17.6	15.6	22.9	31.9	44.1	48.7	61.5	78.8	92.3	111	124
B7R001	89.4	137	92.6	0.462	2.57	0.721	3.50	1.54	7.74	4.71	11.8	9.76	19.4	16.3	28.0
C3R002	192	318	178	6.68	7.77	8.55	10.4	16.6	21.2	25.1	30.9	43.8	48.0	62.7	66.3
C5R001	101	226	156	0.961	-4.31	1.35	-5.06	3.38	-6.09	5.94	-4.79	11.9	-0.543	18.3	6.50
C5R003	193	336	226	3.27	1.81	4.49	2.88	10.8	9.01	16.9	16.1	29.9	30.2	44.8	47.3
C7R001	251	264	234	14.6	12.1	18.1	15.8	33.6	31.8	48.5	45.7	75.8	71.2	104	98.8
D6R002	122	195	147	0.800	0.782	1.16	1.34	3.70	4.95	7.04	9.30	13.8	18.3	24.3	29.4
E1R002	408	309	291	21.0	32.1	38.4	41.3	89.9	77.3	121	106	170	155	217	204
H5H004	661	409	367	93.3	65.8	111	84.1	182	154	235	206	323	292	400	376
J2R003	17.5	32.9	26.8	0.097	-0.625	0.142	-0.731	0.393	-0.802	0.781	-0.506	1.65	0.379	2.77	1.75
J3R001	184	428	347	0.540	-15.7	0.815	-19.0	2.53	-27.1	4.87	-28.2	10.5	-24.3	18.6	-13.1
N1R001	243	455	301	1.88	-0.184	2.79	0.559	7.89	6.28	13.7	14.0	29.7	30.2	45.9	50.6
Q4R002	142	147	130	5.52	7.16	7.72	9.35	18.6	18.7	27.5	26.7	42.7	41.3	58.9	57.0
Q5R001	452	639	538	8.36	5.29	11.6	7.91	29.2	23.4	48.4	40.8	86.3	76.6	126	120
S6R002	108	146	110	1.26	3.55	1.81	4.76	5.27	10.3	9.01	15.4	16.9	25.1	26.8	35.9
T2R001	285	237	215	7.75	20.7	11.4	26.7	26.9	50.5	41.4	69.4	71.7	103	107	136
U2R001	166	202	93.8	11.5	15.0	14.6	19.3	29.9	35.4	40.0	47.5	59.0	67.2	77.6	86.0
U3R001	174	279	192	1.06	3.49	2.14	4.91	4.76	12.0	11.5	19.2	23.3	33.2	36.2	49.5
V1R003	566	437	401	34.0	44.7	52.3	57.5	93.8	108	132	147	201	215	271	283
V3R003	102	73.9	66.3	12.7	8.83	15.4	11.3	26.4	21.0	35.5	28.4	46.5	41.0	58.3	53.5
W4R001	1 003	1 931	929	25.3	37.2	36.2	50.2	83.2	104	129	152	210	236	301	326

Site	Moments			AEP (%)											
				60		50		40		30		20		10	
	Q and SD values in m^3/s														
	Q_{ave}	SD	SD*	Q_{obs}	Q_{est}	Q_{obs}	Q_{est}	Q_{obs}	Q_{est}	Q_{obs}	Q_{est}	Q_{obs}	Q_{est}	Q_{obs}	Q_{est}
A3R002	37.4	63.3	44.7	11.4	13.2	16.5	18.7	25.8	26.0	36.7	36.6	54.3	53.4	95.4	87.1
B1R001	280	384	317	97.3	114	141	156	199	211	285	289	418	408	683	642
B2R001	144	188	171	59.6	57.2	80.9	79.5	109	108	150	150	215	213	334	336
C1R002	494	421	388	302	291	373	365	462	452	568	568	721	730	983	1 015
C2R001	23.2	22.0	20.4	12.1	12.7	16.1	16.2	20.5	20.4	26.3	26.2	35.0	34.4	49.2	49.3
C5R002	547	1 223	824	165	100	236	180	331	295	477	471	738	766	1 301	1 389
C9R002	1 216	1 251	1 154	571	622	764	806	995	1 031	1 325	1 345	1 827	1 797	2 703	2 636
D3R002	2 614	2 094	1 928	1 651	1 604	2 024	1 990	2 469	2 442	3 051	3 043	3 847	3 863	5 209	5 285
D7H005	1 730	1 690	1 540	880	932	1 157	1 191	1 493	1 505	1 965	1 939	2 596	2 554	3 810	3 680
J1R003	298	665	398	73.0	74.3	112	115	172	174	261	262	421	410	783	720
J2R001	20.4	29.1	22.8	8.69	8.38	11.4	11.4	15.5	15.3	21.1	21.0	29.8	29.6	45.8	46.6
N2R001	467	796	641	168	134	229	205	310	301	428	447	618	677	1 005	1 147
Q1R001	123	139	121	70.7	60.7	87.1	79.2	108	102	134	134	174	181	244	270
V6H002	949	492	469	710	695	841	831	971	983	1 138	1 170	1 360	1 404	1 719	1 759
X1H001	302	371	308	162	141	202	186	256	242	330	323	431	442	628	669
A2R001	278	301	267	134	139	177	181	228	232	302	305	422	409	630	604
A2R005	33.9	49.8	29.6	14.0	16.9	22.0	21.5	26.4	27.5	36.5	35.5	51.3	47.9	80.5	71.3
A4R001	181	281	240	46.8	57.4	73.4	85.3	112	122	169	178	261	265	443	439
A8R001	227	409	278	73.9	76.0	103	109	146	153	211	218	319	322	541	533
B3R001	58.5	96.5	70.7	21.7	20.7	29.7	29.2	40.8	40.7	56.8	57.5	79.3	83.9	129	137
B6R003	283	282	233	148	160	198	201	250	251	321	319	427	415	630	589
B7R001	89.4	137	92.6	26.0	38.4	38.7	51.1	59.8	67.7	89.4	91.1	138	127	236	198
C3R002	192	318	178	84.6	87.9	112	114	147	148	197	194	273	267	421	409
C5R001	101	226	156	27.2	17.2	33.8	32.0	52.5	53.5	82.5	86.7	133	142	237	260
C5R003	193	336	226	64.3	69.1	89.0	96.6	123	134	172	188	256	273	432	446
C7R001	251	264	234	135	130	173	168	219	214	278	278	368	370	527	541
D6R002	122	195	147	37.3	43.4	56.2	61.4	83.2	85.3	121	120	185	175	313	285
E1R002	408	309	291	268	256	325	316	392	387	476	479	591	604	785	819
H5H004	661	409	367	479	462	563	557	660	665	778	800	972	975	1 194	1 252
J2R003	17.5	32.9	26.8	4.35	3.72	6.00	6.44	9.55	10.2	15.0	16.1	23.3	25.4	43.7	45.0
J3R001	184	428	347	30.3	6.44	45.0	35.9	72.8	78.9	116	149	203	266	406	518
N1R001	243	455	301	70.6	77.5	104	112	150	160	221	229	342	342	597	572
Q4R002	142	147	130	77.1	74.6	98.1	95.9	123	122	159	158	212	209	301	305
Q5R001	452	639	538	176	174	241	243	324	332	456	463	646	662	1 033	1 055
S6R002	108	146	110	39.2	48.7	55.1	64.4	78.8	84.5	112	113	162	156	266	238
T2R001	285	237	215	147	172	199	215	259	264	339	330	454	421	652	580
U2R001	166	202	93.8	98.5	106	122	127	150	153	189	185	239	232	335	313
U3R001	174	279	192	53.5	69.4	77.7	94.4	113	127	163	175	251	248	426	394
V1R003	566	437	401	341	355	429	438	535	536	667	663	848	837	1 126	1 133
V3R003	102	73.9	66.3	70.4	66.6	84.5	81.4	99.2	98.6	118	121	147	150	188	200
W4R001	1 003	1 931	929	414	434	546	561	730	737	977	974	1 376	1 364	2 150	2 135

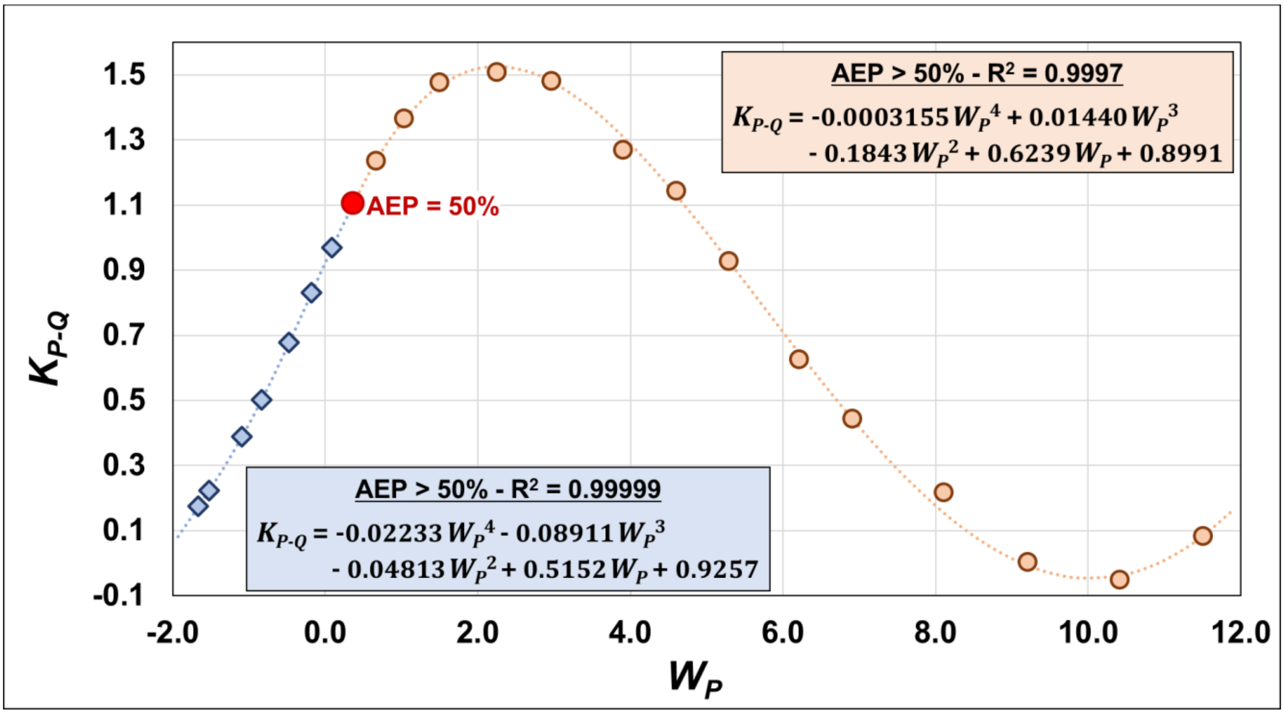
Site	Moments			AEP (%)									
				5		2		1		0.5		0.2	
	Q and SD values in m^3/s												
	Q_{ave}	SD	SD*	Q_{obs}	Q_{est}	Q_{obs}	Q_{est}	Q_{obs}	Q_{est}	Q_{obs}	Q_{est}	Q_{obs}	Q_{est}
A3R002	37.4	63.3	44.7	140	124	208	179	262	223	317	268	390	329
B1R001	280	384	317	969	889	1 384	1 250	1 620	1 529	1 793	1 806	2 338	2 184
B2R001	144	188	171	470	465	662	653	809	796	956	935	1 149	1 126
C1R002	494	421	388	1 260	1 301	1 614	1 690	1 888	1 992	2 162	2 268	2 523	2 647
C2R001	23.2	22.0	20.4	64.7	64.4	86.6	85.4	101	102	118	117	140	137
C5R002	547	1 223	824	1 973	2 087	2 985	3 158	3 821	4 006	4 700	4 896	5 907	6 106
C9R002	1 216	1 251	1 154	3 611	3 496	4 789	4 706	5 749	5 633	6 591	6 510	7 694	7 711
D3R002	2 614	2 094	1 928	6 575	6 707	8 377	8 606	9 852	10 083	11 146	11 425	12 941	13 257
D7H005	1 730	1 690	1 540	5 001	4 830	6 742	6 430	7 984	7 662	9 247	8 822	10 871	10 407
J1R003	298	665	398	1 141	1 072	1 753	1 613	2 360	2 051	2 991	2 517	3 760	3 148
J2R001	20.4	29.1	22.8	64.6	64.7	91.4	91.2	113	112	138	133	165	161
N2R001	467	796	641	1 465	1 653	2 129	2 412	2 634	2 998	3 156	3 590	3 896	4 398
Q1R001	123	139	121	317	362	418	493	500	594	579	691	689	824
V6H002	949	492	469	2 064	2 092	2 507	2 481	2 835	2 794	3 159	3 043	3 582	3 381
X1H001	302	371	308	837	907	1 129	1 251	1 358	1 518	1 591	1 779	1 904	2 135
A2R001	278	301	267	841	806	1 137	1 092	1 364	1 312	1 591	1 523	1 886	1 811
A2R005	33.9	49.8	29.6	112	97.4	160	135	223	167	262	199	309	243
A4R001	181	281	240	649	625	953	901	1 204	1 112	1 438	1 322	1 750	1 610
A8R001	227	409	278	742	766	1 173	1 117	1 454	1 396	1 795	1 685	2 221	2 077
B3R001	58.5	96.5	70.7	181	194	258	280	332	348	395	417	486	511
B6R003	283	282	233	820	768	1 089	1 016	1 299	1 212	1 474	1 398	1 792	1 652
B7R001	89.4	137	92.6	350	276	466	390	558	482	693	575	936	703
C3R002	192	318	178	587	570	824	809	1 027	1 007	1 449	1 215	1 735	1 496
C5R001	101	226	156	357	391	534	592	680	750	834	916	1 116	1 142
C5R003	193	336	226	628	636	897	921	1 120	1 148	1 354	1 383	1 674	1 702
C7R001	251	264	234	694	718	928	966	1 105	1 158	1 282	1 341	1 512	1 590
D6R002	122	195	147	404	403	678	580	863	718	1 043	858	1 254	1 048
E1R002	408	309	291	978	1 031	1 232	1 312	1 425	1 530	1 619	1 725	1 876	1 992
H5H004	661	409	367	1 519	1 520	1 815	1 852	2 051	2 117	2 352	2 344	2 642	2 651
J2R003	17.5	32.9	26.8	65.7	66.1	98.3	98.1	129	123	159	148	195	182
J3R001	184	428	347	665	793	1 068	1 217	1 409	1 541	1 756	1 875	2 255	2 331
N1R001	243	455	301	898	827	1 342	1 213	1 685	1 521	2 037	1 841	2 522	2 276
Q4R002	142	147	130	397	403	531	541	642	647	736	749	874	887
Q5R001	452	639	538	1 441	1 472	2 076	2 084	2 518	2 555	2 962	3 022	3 604	3 659
S6R002	108	146	110	376	326	511	454	598	555	730	656	935	793
T2R001	285	237	215	852	740	1 083	954	1 235	1 121	1 374	1 274	1 628	1 483
U2R001	166	202	93.8	495	403	694	530	827	641	948	757	1 111	912
U3R001	174	279	192	614	554	911	790	1 143	979	1 378	1 173	1 691	1 435
V1R003	566	437	401	1 447	1 428	1 875	1 819	2 176	2 124	2 367	2 399	2 758	2 774
V3R003	102	73.9	66.3	237	249	290	312	334	362	377	407	436	468
W4R001	1 003	1 931	929	3 049	3 036	4 386	4 395	5 500	5 547	6 675	6 793	8 287	8 466

Site	Moments			AEP (%)									
				0.1		0.03		0.01		0.003		0.001	
	Q and SD values in m³/s												
	Q_{ave}	SD	SD*	Q_{obs}	Q_{est}	Q_{obs}	Q_{est}	Q_{obs}	Q_{est}	Q_{obs}	Q_{est}	Q_{obs}	Q_{est}
A3R002	37.4	63.3	44.7	444	374	535	450	610	519	682	588	736	647
B1R001	280	384	317	2 749	2 460	3 195	2 916	3 571	3 330	3 955	3 736	4 285	4 065
B2R001	144	188	171	1 290	1 264	1 524	1 487	1 717	1 693	1 901	1 887	2 039	2 038
C1R002	494	421	388	2 795	2 926	3 261	3 388	3 676	3 812	4 117	4 237	4 503	4 595
C2R001	23.2	22.0	20.4	157	152	188	177	217	200	250	222	281	240
C5R002	547	1 223	824	6 841	7 004	8 476	8 509	9 951	9 859	11 512	11 223	12 829	12 357
C9R002	1 216	1 251	1 154	8 416	8 586	9 579	10 017	11 115	11 331	12 321	12 604	13 359	13 636
D3R002	2 614	2 094	1 928	14 301	14 614	16 631	16 878	18 710	18 951	20 922	21 050	22 868	22 856
D7H005	1 730	1 690	1 540	12 055	11 568	13 987	13 478	15 578	15 230	17 094	16 948	18 238	18 365
J1R003	298	665	398	4 344	3 624	5 338	4 435	6 026	5 156	6 705	5 912	7 288	6 562
J2R001	20.4	29.1	22.8	184	182	218	216	248	248	281	279	310	305
N2R001	467	796	641	4 455	4 988	5 397	5 956	6 218	6 836	7 055	7 688	7 743	8 368
Q1R001	123	139	121	768	922	921	1 083	1 065	1 230	1 228	1 375	1 381	1 494
V6H002	949	492	469	3 898	3 643	4 440	4 102	4 927	4 521	5 453	4 993	5 925	5 451
X1H001	302	371	308	2 141	2 397	2 549	2 831	2 914	3 225	3 298	3 616	3 630	3 938
A2R001	278	301	267	2 102	2 022	2 459	2 369	2 757	2 686	3 047	2 997	3 272	3 251
A2R005	33.9	49.8	29.6	342	276	395	334	465	386	545	442	621	493
A4R001	181	281	240	1 987	1 819	2 404	2 158	2 791	2 469	3 225	2 765	3 629	2 998
A8R001	227	409	278	2 542	2 370	3 101	2 864	3 450	3 307	3 859	3 760	4 205	4 144
B3R001	58.5	96.5	70.7	550	580	654	696	770	801	902	906	1 026	993
B6R003	283	282	233	1 964	1 841	2 275	2 159	2 611	2 447	2 883	2 742	3 058	2 995
B7R001	89.4	137	92.6	1 035	798	1 196	960	1 331	1 106	1 469	1 257	1 587	1 388
C3R002	192	318	178	1 936	1 711	2 259	2 087	2 531	2 419	2 828	2 782	3 160	3 110
C5R001	101	226	156	1 283	1 309	1 531	1 588	1 828	1 838	2 168	2 089	2 489	2 296
C5R003	193	336	226	1 923	1 941	2 363	2 344	2 768	2 705	3 211	3 077	3 610	3 394
C7R001	251	264	234	1 680	1 773	1 958	2 075	2 188	2 351	2 410	2 622	2 579	2 846
D6R002	122	195	147	1 365	1 189	1 597	1 423	1 775	1 634	1 957	1 845	2 113	2 019
E1R002	408	309	291	2 073	2 189	2 416	2 519	2 732	2 821	3 081	3 128	3 403	3 392
H5H004	661	409	367	2 903	2 887	3 361	3 294	3 787	3 664	4 262	4 069	4 702	4 449
J2R003	17.5	32.9	26.8	227	206	286	247	344	284	410	319	474	347
J3R001	184	428	347	2 582	2 661	3 153	3 196	3 670	3 685	4 140	4 146	4 532	4 501
N1R001	243	455	301	2 906	2 601	3 569	3 151	4 024	3 643	4 489	4 150	4 888	4 580
Q4R002	142	147	130	978	989	1 158	1 157	1 319	1 310	1 490	1 461	1 639	1 586
Q5R001	452	639	538	4 115	4 124	4 866	4 886	5 486	5 582	6 255	6 255	6 763	6 794
S6R002	108	146	110	1 029	895	1 181	1 066	1 308	1 221	1 439	1 378	1 657	1 510
T2R001	285	237	215	1 775	1 637	2 021	1 895	2 236	2 131	2 464	2 369	2 666	2 574
U2R001	166	202	93.8	1 237	1 036	1 416	1 262	1 530	1 458	1 706	1 691	1 936	1 918
U3R001	174	279	192	1 875	1 632	2 167	1 964	2 414	2 262	2 665	2 569	2 881	2 831
V1R003	566	437	401	3 122	3 053	3 589	3 520	4 007	3 948	4 459	4 385	4 867	4 765
V3R003	102	73.9	66.3	481	514	561	591	635	662	717	736	794	802
W4R001	1 003	1 931	929	9 536	9 767	11 736	12 064	13 760	14 077	15 979	16 327	17 997	18 394

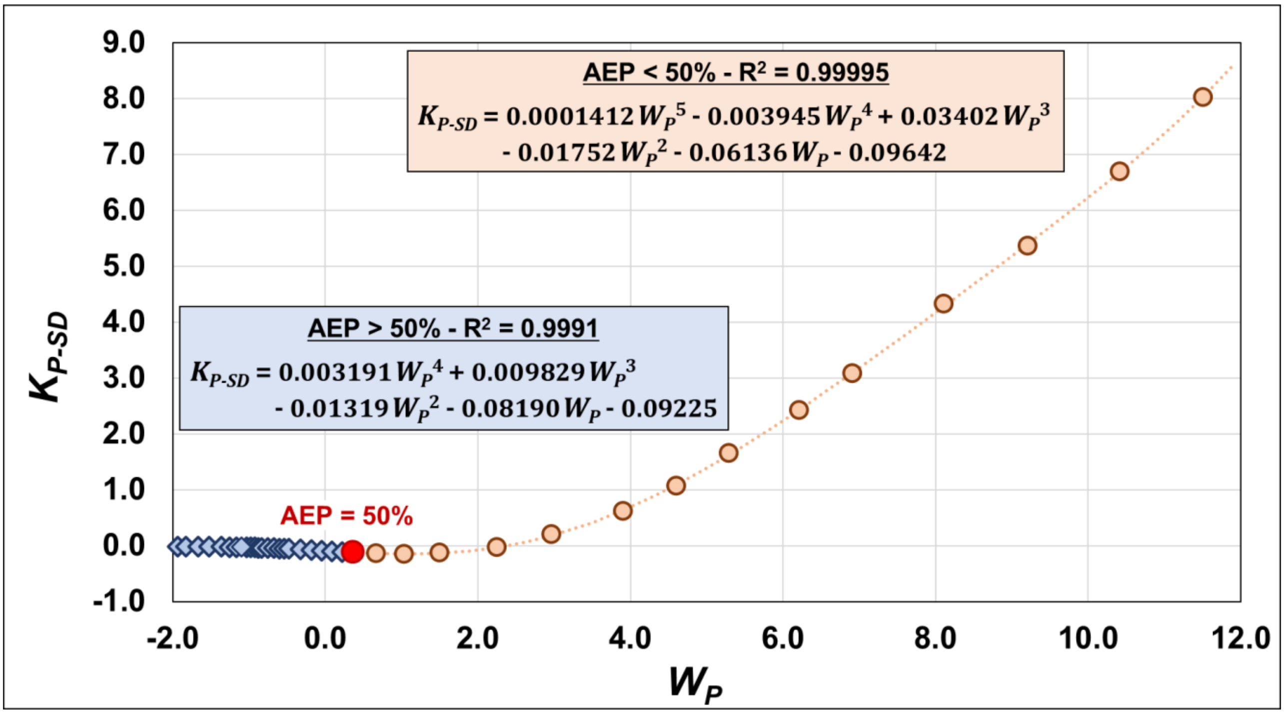
APPENDIX 6-J: Probability distribution parameters

Present the estimated relationship between the standardised variate W_P and the three frequency factors, for selected scenario ave2.

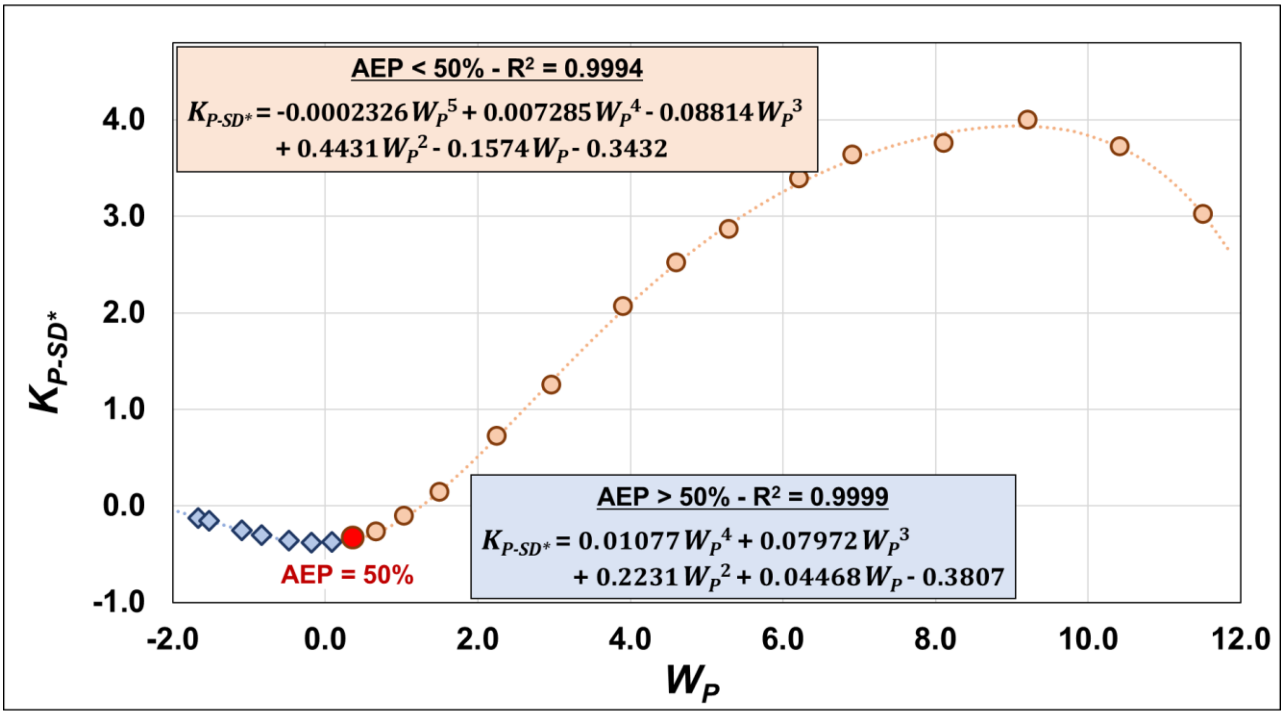
Relationship between W_P and estimated frequency factor for Q_{ave} (K_{P-Q})



Relationship between W_P and estimated frequency factor for SD (K_{P-SD})



Relationship between W_P and estimated frequency factor for SD* ($K_{P\text{-}SD*}$)



APPENDIX 6-K: Frequency factors for IPZA probability distribution

Estimated frequency factors, corresponding to the AEP and associated standardised variable W_P.

AEP ≥ 50%

AEP (%)	W_P	K_{P-Q}	K_{P-SD}	K_{P-SD*}
99.9	-1.9326	0.0820	-0.0097	-0.0590
99.8	-1.8269	0.1184	-0.0110	-0.0838
99.5	-1.6674	0.1733	-0.0133	-0.1212
99	-1.5272	0.2226	-0.0156	-0.1540
98	-1.3641	0.2822	-0.0190	-0.1916
97	-1.2546	0.3242	-0.0218	-0.2163
96	-1.1690	0.3583	-0.0243	-0.2353
95	-1.0972	0.3878	-0.0266	-0.2508
94	-1.0344	0.4143	-0.0289	-0.2641
93	-0.9780	0.4387	-0.0310	-0.2757
92	-0.9265	0.4615	-0.0332	-0.2860
91	-0.8788	0.4829	-0.0352	-0.2954
90	-0.8340	0.5034	-0.0373	-0.3038
88	-0.7515	0.5420	-0.0413	-0.3187
86	-0.6761	0.5783	-0.0453	-0.3313
84	-0.6057	0.6128	-0.0492	-0.3422
82	-0.5393	0.6459	-0.0532	-0.3515
80	-0.4759	0.6781	-0.0572	-0.3595
75	-0.3266	0.7551	-0.0672	-0.3741
70	-0.1856	0.8290	-0.0776	-0.3818
65	-0.0486	0.9005	-0.0883	-0.3824
60	0.0874	0.9703	-0.0995	-0.3750
55	0.2250	1.0381	-0.1112	-0.3584
50	**0.3665**	**1.1035**	**-0.1216**	**-0.3379**

AEP ≤ 50%

AEP (%)	W_P	K_{P-Q}	K_{P-SD}	K_{P-SD*}
50	**0.3665**	**1.1035**	**-0.1216**	**-0.3379**
45	0.5144	1.1732	-0.1283	-0.3184
40	0.6717	1.2393	-0.1360	-0.2743
35	0.8422	1.3023	-0.1421	-0.2106
30	1.0309	1.3618	-0.1453	-0.1232
25	1.2459	1.4174	-0.1434	-0.0051
20	1.4999	1.4673	-0.1320	0.1553
14	1.8916	1.5132	-0.0920	0.4356
10	2.2504	1.5258	-0.0286	0.7155
7	2.6232	1.5125	0.0669	1.0180
5	2.9702	1.4791	0.1838	1.3020
3	3.4914	1.3968	0.4107	1.7192
2	3.9019	1.3099	0.6317	2.0310
1.4	4.2617	1.2212	0.8542	2.2875
1.0	4.6001	1.1296	1.0865	2.5124
0.7	4.9583	1.0263	1.3544	2.7318
0.5	5.2958	0.9249	1.6253	2.9205
0.3	5.8076	0.7681	2.0652	3.1735
0.2	6.2136	0.6444	2.4345	3.3465
0.14	6.5706	0.5386	2.7707	3.4795
0.10	6.9073	0.4429	3.0952	3.5892
0.07	7.2641	0.3473	3.4450	3.6896
0.05	7.6007	0.2641	3.7787	3.7695
0.03	8.1116	0.1531	4.2897	3.8635
0.02	8.5171	0.0803	4.6980	3.9131
0.014	8.8738	0.0287	5.0590	3.9363
0.010	9.2103	-0.0082	5.4022	3.9379
0.007	9.5670	-0.0343	5.7705	3.9142
0.005	9.9035	-0.0460	6.1247	3.8633
0.003	10.4143	-0.0385	6.6826	3.7196
0.002	10.8198	-0.0102	7.1520	3.5333
0.0014	11.1764	0.0314	7.5930	3.3026
0.0010	11.5129	0.0851	8.0410	3.0152