Water SA 50(1) 1–19 / Jan 2024
https://doi.org/10.17159/wsa/2024.v50.i1.4022

Research paper

Improved flood quantile estimation for South Africa

D van der Spuy[1] and JA du Plessis[1]

[1]Department of Civil Engineering, Stellenbosch University, P/Bag X1, Matieland 7602, South Africa

The performance of the most frequently used flood frequency probability distributions in South Africa (Log-Normal, Log Pearson3 and Generalised Extreme Value) were reviewed and all tend to perform poorly when lower exceedance probability frequency events are estimated, especially where outliers are present in the dataset. This can be attributed to the challenge when analysing very limited 'samples' of annual flood peak populations, which are an unknown. At present outliers are inadequately 'managed' by attempting to 'normalise' the flood peak dataset, which conceals the significance of the observed data. Thus, to adequately consider the outliers, this study was undertaken with the aim to improve the current statistical approach by developing a more stable and consistent methodology to estimate flood quantiles. The approach followed in the development of the new methodology, called IPZA, might be considered as unconventional, given that a multiple regression approach was used to accommodate the strongly skewed data, which are often associated with annual flood peak series. The main advantages of IPZA are 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, thereby eliminating the need to adjust any moments. The performance of IPZA exceeded initial expectations. The results are more consistent and, by taking outliers into account, appear to be more sensible than existing probability distributions. It is recommended that IPZA should be used as a valuable addition to the existing set of decision-making tools for hydrologists/engineers performing flood frequency analyses.

CORRESPONDENCE
D van der Spuy

EMAIL
vds.danie@gmail.com

DATES
Received: 16 September 2022
Accepted: 5 December 2023

KEYWORDS
flood
quantile
frequency
probability
exceedance
distribution
practical

COPYRIGHT
© The Author(s)
Published under a Creative Commons Attribution 4.0 International Licence
(CC BY 4.0)

INTRODUCTION

Floods rank among the topmost devasting natural disasters. The 1931 Yangtze (Yangzi) River flood, in Southern China, is indisputably still considered the worst natural disaster of the 20th century. The flood inundated 180 000 km^2 and an estimated 3.7 million people lost their lives (Courtney, 2018) – however, one should 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 flood in the Buffels River in 1981, which inundated the town of Laingsburg, caused 90 people to lose their lives and caused damage in excess of 10 million ZAR (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, causing widespread damage to road infrastructure and property of more than 100 million ZAR (1984 values) and over 200 reported deaths in South Africa alone (Kovács et al., 1985). The main cause of 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 has occurred in SA. Damage to property, agriculture (including several hundred farm dams that were breached), communications and infrastructure amounted to approximately 400 million ZAR, 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 of approximately 1.5 billion ZAR (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 use to perform flood frequency analyses (FFAs). 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 an entire population of AMS of flood peak records, to test the AMS sample records against, does not exist. From very 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 engineers and hydrologists, to increase the confidence in FFA.

Existing probability distributions commonly applied in South Africa

The probability distributions most used in SA for FFA are the Log-Normal (LN), Log-Pearson Type III (LP3) and the Generalised Extreme Value (GEV) distributions (Alexander, 1990; SANRAL, 2013; Van der Spuy and Du Plessis, 2022a). Both the LP3 and GEV distributions have provided good results (DWS, 1993–2021) and the practice is to apply both and use the one that seems to fit the observed data best. Van der Spuy and Du Plessis (2022a) illustrated that all of the distributions lack consistency, which is mainly caused by the inability to deal properly with outliers in an AMS of flood peak data. They showed that, although the GEV proved to be insensitive to low outliers, all distributions are susceptible to high outliers. Ball et al. (2019) consider outliers as observations that are inconsistent with the general trend of the rest of the data. Frost (2019) stated that, while there is no strict statistical rule or mathematical definition to identify outliers, guidelines exist through which possible outliers can be identified. The Zscore (one such guideline) also provides a measure to compare relative probabilities associated with relative magnitudes of the data. 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 be outliers (Brownlee, 2018; Frost, 2019). Van der Spuy and Du Plessis (2022a) expressed the need for a probability distribution which will provide more consistent flood frequency estimations. They also proposed that the plotting position (PP) approach should be improved, especially concerning outliers. Van der Spuy and Du Plessis (2022b) developed an alternative PP technique, the Z-set PP. The Z-set PP offers an improved and more realistic representation of outliers, with a further advantage of being less susceptible to different record lengths than the existing PP techniques. The Z-set PP was considered as an essential tool to the flood frequency analyst to assist in more consistent choices of appropriate probability distributions.

Purpose of study

The purpose of this study was to develop a more consistent methodology to estimate flood quantiles (discharges with specified probabilities of exceedance) that would not be as sensitive to outliers as the most frequently used flood frequency probability distributions in South Africa. Consequently, the Z-set PP formed an integral part in the study.

The sensitivity to outliers of the developed methodology will be evaluated against that of the LP3 and GEV.

METHODOLOGY

Statistics were identified, computed from the AMS data at selected flow gauging sites (referred to as sites henceforth), which could be considered as potential parameters of the 'underlying probability distribution', in an FFA. The study was performed in three phases to find the most appropriate statistics that would be used as parameters for a new flood quantile methodology.

Phase 1: Evaluation of statistics as potential independent variables

Multiple regression analyses (MRAs) were used to determine which of the identified potential statistics have little or no impact on the development of a new flood quantile methodology. For the MRAs the following variables were identified (detailed in Table 1):

- As independent variables – the identified potential statistics were considered
- As dependent variables – flood peak values, referred to as 'observed' flood peaks (Q_{obs}) at identified annual exceedance probabilities (AEPs), were used

Values for Q_{obs} were estimated for the identified AEPs by interpolation and extrapolation of the flood frequency relationship between the AEPs (applying the Z-set PP) and the observed AMS flood peaks (estimated from observed (recorded) stage levels).

In this phase of the study MRAs were conducted for every identified AEP to estimate a flood peak (Q_{est}) that could be compared graphically, corroborated by the coefficient of determination (R^2), to the corresponding Q_{obs}, for every flow site.

The identified variables, which did not contribute to improve the results in any way, were eliminated.

Phase 2: Identify parameters for flood quantile model

Using an MRA-approach, various clusters (henceforth used in referring to combinations of the remaining independent variables) were assessed to estimate the dependent variables (Q_{obs}).

Due to the parameters used in most existing probability distributions, it was anticipated that a cluster 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 study. Consequently, Table 2 depicts examples of clusters that were considered and assessed.

For every cluster at every site, MRAs were conducted at all selected AEPs to estimate values of Q_{est}. Instead of just graphically comparing Q_{est} with the corresponding Q_{obs}, key performance indicators (KPIs) were used to rank the relative performance of the different clusters. All these KPIs perform differently, mainly

Table 1. Initial statistic variable list considered for the MRAs

Independent variables	Description	Comments
Q_{max}	Maximum peak of the AMS (m^3/s)	* Indicates excluding largest flood peak (Q_{max}) from AMS dataset *⁰ Indicates excluding both largest flood peak (Q_{max}) and smallest flood peak (Q_{min}) from AMS dataset
Q_{min}	Minimum peak of the AMS (m^3/s)	
Q_{ave}, Q_{ave}*	Average(s) of the AMS (m^3/s)	
Q_{mdn}, Q_{mdn}*	Median(s) of the AMS (m^3/s)	
Q_{gmn}	Geometric mean of the AMS (m^3/s)	
SD, SD*, SD*^{0}	Standard deviation(s) of the AMS (m^3/s)	
g, *g**	Skewness statistics of the AMS	
kurt, kurt*	Kurtosis statistics of the AMS	
COV, COV*	Coefficient of variation of the AMS	
Z_{max} and Z_{min}	*Z*-scores of Q_{max} and Q_{min}, respectively	
Dependent variable	**Description**	**Comments**
Q_{obs}	Flood peak(s) corresponding to selected AEPs, estimated through interpolation/ extrapolation of observed AMS flood peaks, applying the Z-set PP technique	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

Table 2. Example of clusters considered with different measures of central tendency

Use of average values	Use of median values	Use of geometric mean values
Q_{ave}, SD	Q_{mdn}, SD	Q_{gmn}, SD
Q_{ave}, SD, g	Q_{mdn}, SD, g	Q_{gmn}, SD, g
etc.	etc.	etc.

related to the magnitude of the values involved – therefore, two KPIs were chosen, namely, the mean absolute percentage error (MAPE) and the mean absolute error (MAE) depicted in Eq. 1 and Eq. 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 were considered as relative assessments of performance, rather than absolute, for the purpose of this study.

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

$$\text{MAPE} = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{Qi_{\text{est}} - Qi_{\text{obs}}}{Qi_{\text{obs}}}\right| \quad (1)$$

and

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|Qi_{\text{est}} - Qi_{\text{obs}}\right| \quad (2)$$

where:

MAPE = mean absolute percentage error (%)

MAE = mean absolute error (m^3/s)

n = sample size = number of AEP values used for AEP≤ 50% and AEP range > 50% ranges

Qi_{obs} = observed flood peak related to an AEP at observation 'i' (m^3/s)

Qi_{est} = estimated flood peak related to an AEP at observation 'i' (m^3/s)

From these results the cluster that provided the best fit, on average, to the observed AMS flood peaks was used to develop a flood quantile model.

Phase 3: Flood quantile model

Wheeler (2022) 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 the AMS of flood peaks, Wheeler explained that it is a "mathematical fact of life" that no skewed probability model exists to fit these data. Wheeler (2009) also stated that if data are transformed to make them appear to be 'more normal', it is likely to "end up with a beautiful, but completely incorrect, analysis". As a practical alternative, the MRA results of the chosen cluster were used to find an empirical solution for the flood quantile model. Consequently, the independent variables were used as the parameters for the flood quantile model, while the regression coefficients, associated with the AEPs, correspond to the frequency factors. However, given that the wide range of AEPs did not provide sensible independent variables, an associated standardised variable was considered. 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). Both are considered to be suitable, but the standardised variate W_P was chosen, merely because of the simplicity of its equation.

$$W_P = -\ln(-\ln(1-\text{AEP})) \quad (3)$$

To estimate continuous frequency factors, at different AEPs used in the study, polynomial functions were used. For each statistical parameter the regression coefficients, determined for every AEP, were considered to be the dependent variables (y), while the associated standard variate W_P was considered to be the independent variable (x).

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

$$K_P = a_m(W_P)^m + a_{m-1}(W_P)^{m-1} + \ldots + a_2(W_P)^2 + a_1(W_P)^1 + a_0 \quad (4)$$

where:

K_P = frequency factor

M = denotes the degree of the polynomial (determined by optimum correlation coefficient)

a_m = coefficient of the m^{th} degree term

The general form of the flood quantile model can be written as:

$$Q_{AEP} = K_{AEP-S1} \cdot S1 + K_{AEP-S2} \cdot S2 + K_{AEP-S3} \cdot S3 + \ldots + K_{AEP-SN} \cdot SN \quad (5)$$

where:

Q_{AEP} = flood peak associated with an AEP (m^3/s)

Si = i^{th} statistical parameter, for i from 1 to N

K_{AEP-Si} = frequency factor, related to AEP for Si

N = total number of statistics considered

DATA

Similar to the study by Van der Spuy and Du Plessis (2022b), only AMS flood peak data were considered in this study. It has been reported that the partial duration series (PDS) flood peak data approach does not really improve results, especially for an ARI (annual recurrence interval) higher than 10 years (Mkhandi et al., 2005; Karim et al., 2017). Srikanthan (2014) also concluded that, since the AMS results in the smallest bias in most cases, the use of AMS in FFA is preferred to PDS.

Data sources

Van der Spuy and Du Plessis (2022b) used 18 sites for their study. A further 23 sites were added to these sites for this study, ensuring more comprehensive coverage across South Africa. The objective was to find suitable sites, with verified reliable data, in areas and drainage regions not covered previously. There was an attempt to keep the record lengths at 60 years and above; however, 2 sites had record lengths of 58 and 57 years, respectively, and 1 site a record length of only 39 years. These sites were included to provide some data in areas where appropriate sites are very limited.

The location of the selected 41 flow sites is depicted in Fig. 1. Relevant metadata for the flow sites, which include 37 dam and

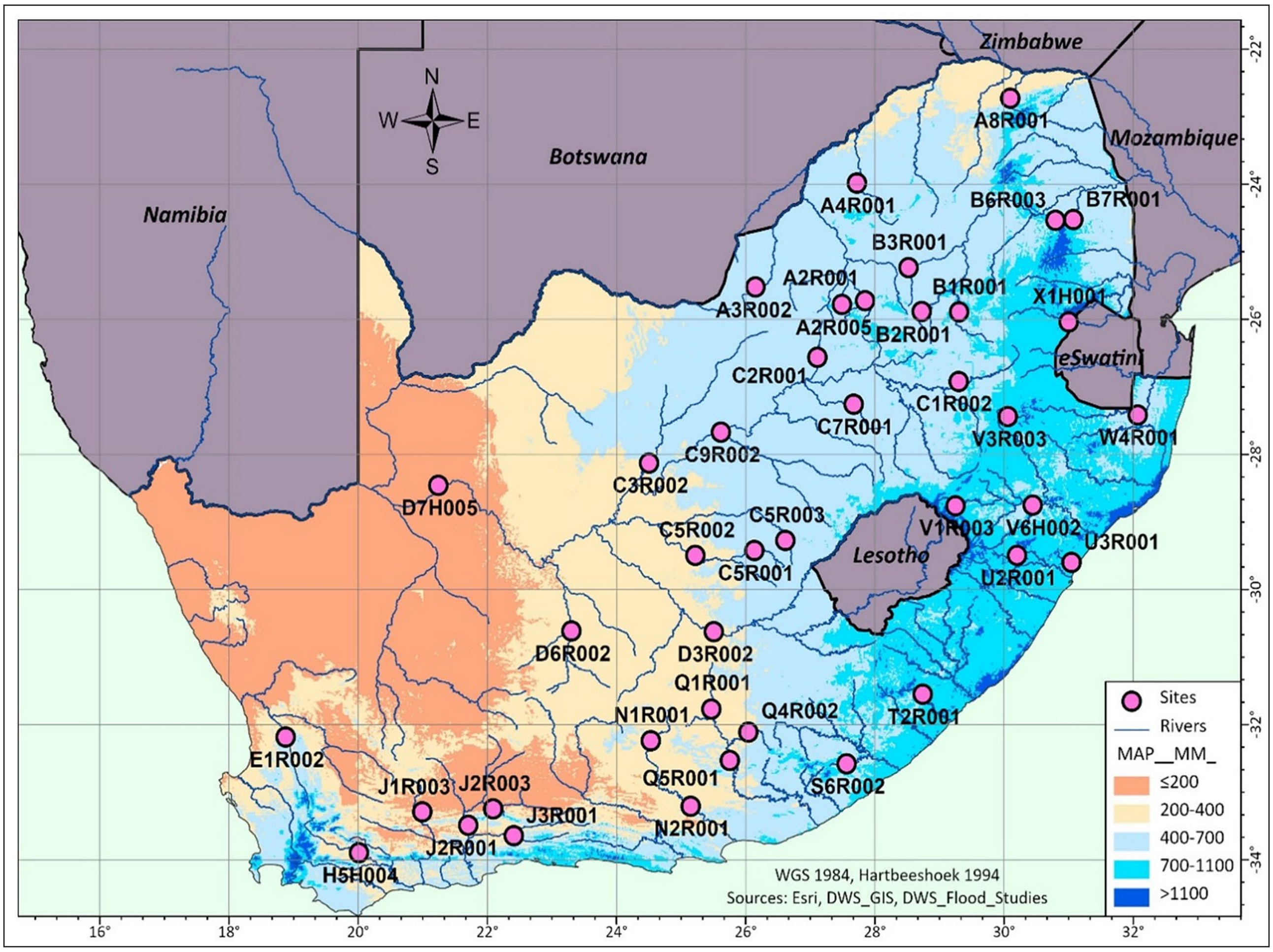

Figure 1. Distribution of sites

4 gauging weir sites with a combined database of 3 615 years of usable AMS flood peaks, are provided in Table A1 (Appendix). The sites amply cover the vast diversity of meteorological and hydrological conditions experienced across SA.

Extracted data used in study

Selected statistics computed from the AMS at each site, utilised as independent variables in the MRA approach, are presented in Table A2 (Appendix). The most appropriate statistics, from these selected statistics, constitute the input into a probability model to produce estimated flood peak values (Q_{est}) at corresponding AEPs.

Selected flood frequency data (Q_{obs} vs selected AEPs), applied as dependent variables in the MRA, are presented in Table A3 (Appendix).

RESULTS

The results are presented below in sections corresponding to the phases defined in the methodology. The acronym of the proposed flood quantile model is identified in advance, since it appears on figures presented as part of the results and is defined as the IPZA (Improved Probability-model, South Africa).

Phase 1: Evaluation of statistics as potential independent variables

MRA clusters, containing numerous combinations of the initially selected statistics (Q_{max}, Q_{min}, Q_{ave}, Q_{ave}*, Q_{mdn}, Q_{mdn}*, Q_{gmn}, SD, SD*, SD^{*0}, *g*, *g**, COV, COV*, COV^{*0}, Z_{max} and Z_{min}), were performed and assessed. As could be expected, variables such as Q_{ave}, Q_{mdn}, Q_{gmn} and SD were considered as potentially significant. Unexpectedly, and contrary to existing practice and beliefs, the initial evaluations suggested that the effect of including skewness (*g*) might be statistically insignificant. It was thus hypothesised that omitting *g* would either have no effect or potentially improve the results. Since the SD relates to the slope of the relative distribution of the data points (in this case a scatterplot of flood peaks against their corresponding Zscores), and is thus affected by outliers, it was furthermore hypothesised that replacing *g* with SD* might improve the outcome. In addition, SD*, SD^{*0} and SD** were also considered as being potentially significant. The example in Fig. 2 (from an MRA at a chosen AEP) illustrates what led to the suggestion that replacing *g* with SD* might improve results.

Several clusters were evaluated graphically, corroborated by the coefficient of determination (R^2), to determine whether they could contribute to improve the results, or not (also illustrated by Fig. 2). After much experimentation and evaluation, the following statistics/parameters were excluded (Q_{max}, Q_{min}, Z_{max}, Z_{min}, Q_{ave}*, Q_{mdn}*, *g**, kurt, kurt*, COV, COV*) and the following retained (Q_{ave}, Q_{mdn}, Q_{gmn}, SD, SD*, *g*) for more detailed analyses.

Phase 2: Identify parameters for flood quantile model

To find the best grouping of parameters, different clusters 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, clusters were added where *g* was replaced by SD*. The clusters considered and the selected statistics are depicted in Table 3.

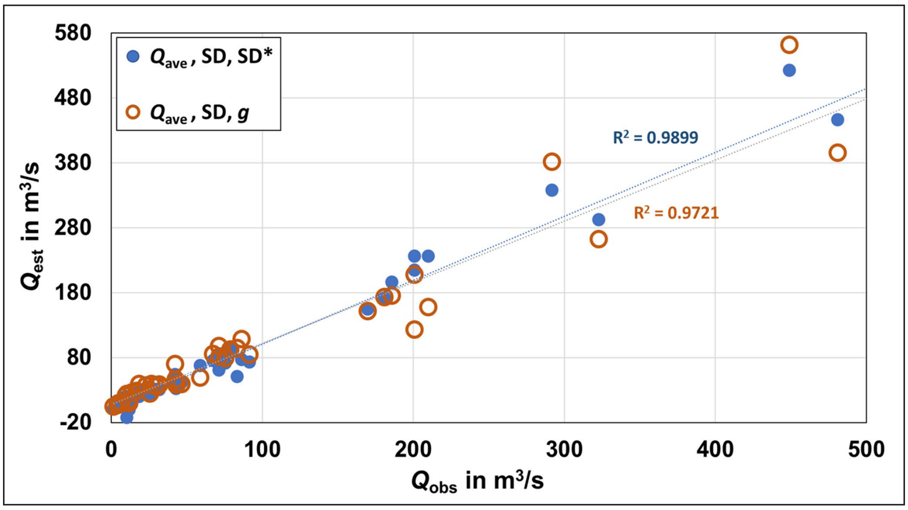

Figure 2. MRA results at AEP = 80%: inclusion of *g* vs SD*

Table 3. Clusters considered with different measures of tendency

Average (Q_{ave})		Median (Q_{mdn})		Geometric mean (Q_{gmn})	
Acronym	Statistics	Acronym	Statistics	Acronym	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*

Table 4. Ranking results

AEP range	Q_{ave} clusters			Q_{mdn} clusters			Q_{gmn} clusters		
	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					
≤ 50%	3	5	1	7	9	4	8	6	2
> 50%	8	9	7	5	6	4	2	3	1
				MAE ranking					
≤ 50%	5	4	1	8	6	3	9	7	2
> 50%	8	9	3	4	7	5	2	6	1

MRAs were conducted at all stations for all selected AEPs for every cluster (a total of 198 MRAs). The summarised outcome of the ranking results of the MAPE and MAE KPIs, based on the average value for an AEP range at all stations, is presented in Table 4 – the MAPE results for the AEP ≤ 50% range are provided in Table A4 (Appendix).

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 clusters with the *g* statistic generally performed the worst – the exception being Gmn1 for AEPs ≤ 50%.
- Regardless of the measure of central tendency used, the clusters with the SD* statistic performed the best overall – the exception being Mdn0 for AEPs > 50% (only with MAE).
- The poor performance of the Q_{mdn} clusters was disappointing since the median (as with the geometric mean) is generally considered to be closer to the estimated 50% AMS flood peak (≡ to a 2-year ARI).
- For AEPs ≤ 50% both MAPE and MAE indicated that Cluster Ave2 (Q_{ave} SD and SD*) performed the best – with Cluster Gmn2 being 2nd best.
- For AEPs > 50% (ARI < 2 years) both MAPE and MAE indicated that Cluster Gmn2 (Q_{gmn} SD and SD*) performed the best.

To illustrate some of the above findings, visual comparisons of clusters are presented, indicating which cluster is categorised as noticeably better or only slightly better, than the other. For each category the sites, which indicate the most noticeable difference between the clusters, were selected to illustrate the differences and are depicted below, as follows:

- The suggestion, that the g-statistic is not a statistically significant variable, is visually illustrated in Fig. 3 (Ave0 vs Ave1), supported by the estimated significance levels, at selected AEPs in Table 5.

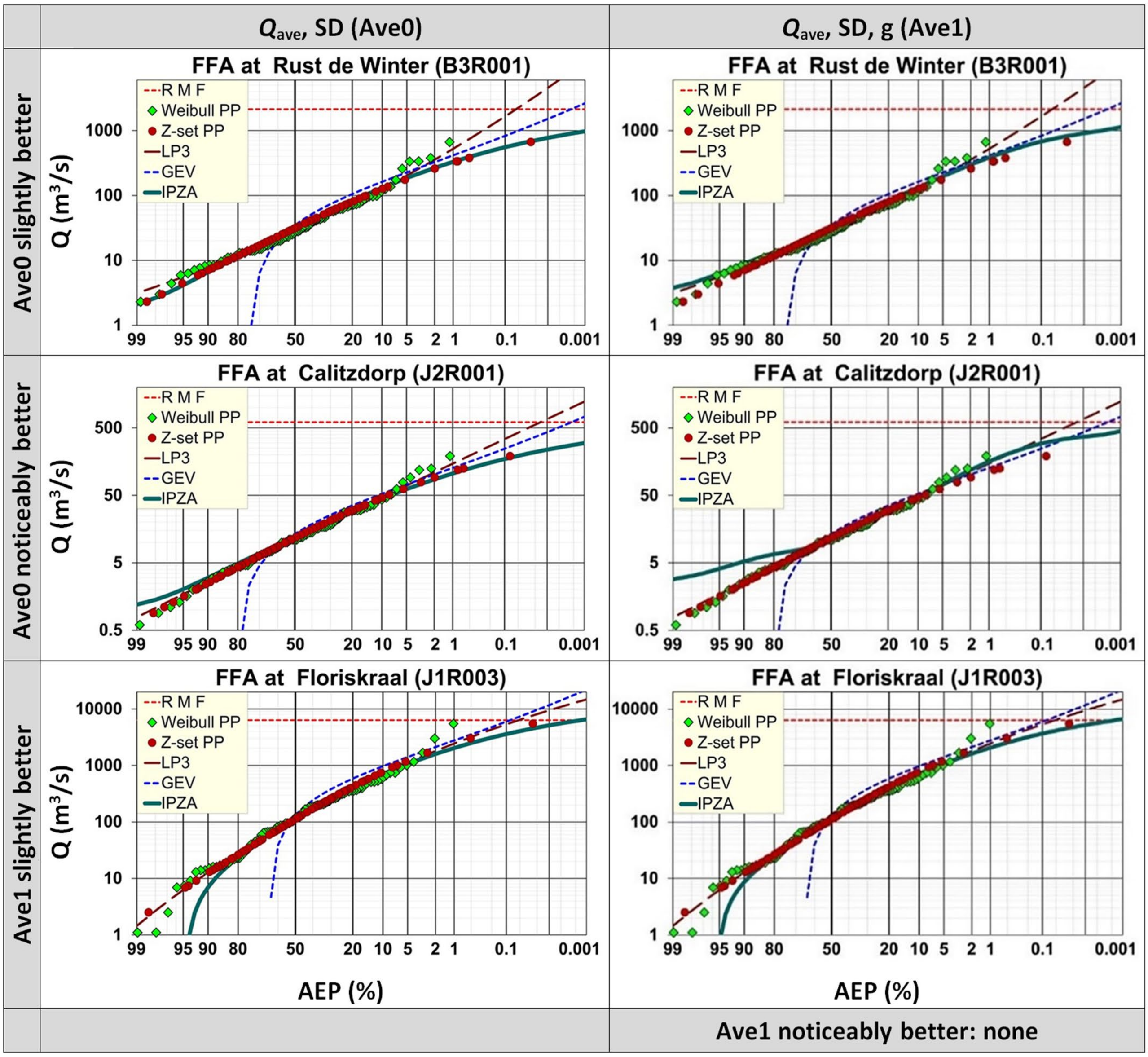

Figure 3. Compare Clusters Ave0 and Ave1 to illustrate non-performance of g

Table 5. Levels of significance for Cluster Ave1

AEP (%)	R^2	Levels of significance				Shading indicating evidence of significance (after Ganesh and Cave, 2018)
		F-value	*P*-value			
			Q_{ave}	SD	*g*	
90	0.9598	4.640 x 10^{-26}	4.667 x 10^{-19}	4.037 x 10^{-08}	0.6172	
70	0.9851	4.673 x 10^{-34}	4.501 x 10^{-25}	6.055 x 10^{-10}	0.8068	*P*-value < 0.001 Very strong evidence
50	0.9963	3.108 x 10^{-45}	1.821 x 10^{-34}	1.948 x 10^{-13}	0.8031	
20	0.9995	7.675 x 10^{-61}	2.515 x 10^{-46}	2.691 x 10^{-04}	0.5708	*P*-value < 0.01 Strong evidence
10	0.9973	8.345 x 10^{-48}	3.775 x 10^{-30}	5.875 x 10^{-05}	0.8985	
5	0.9962	5.518 x 10^{-45}	6.755 x 10^{-25}	4.938 x 10^{-09}	0.6267	*P*-value < 0.05 Moderate evidence
2	0.9943	9.466 x 10^{-42}	1.211 x 10^{-18}	1.349 x 10^{-11}	0.4455	
1	0.9941	1.834 x 10^{-41}	2.357 x 10^{-16}	1.376 x 10^{-13}	0.3663	*P*-value < 0.1 Weak evidence
0.5	0.9940	2.606 x 10^{-41}	3.863 x 10^{-14}	2.987 x 10^{-15}	0.1972	
0.2	0.9940	2.435 x 10^{-41}	8.557 x 10^{-12}	5.563 x 10^{-17}	0.1166	*P*-value ≥ 0.1 Insufficient evidence
0.1	0.9944	6.902 x 10^{-42}	1.281 x 10^{-10}	1.814 x 10^{-18}	0.0989	
0.01	0.9963	3.810 x 10^{-45}	2.374 x 10^{-08}	8.437 x 10^{-24}	0.1067	

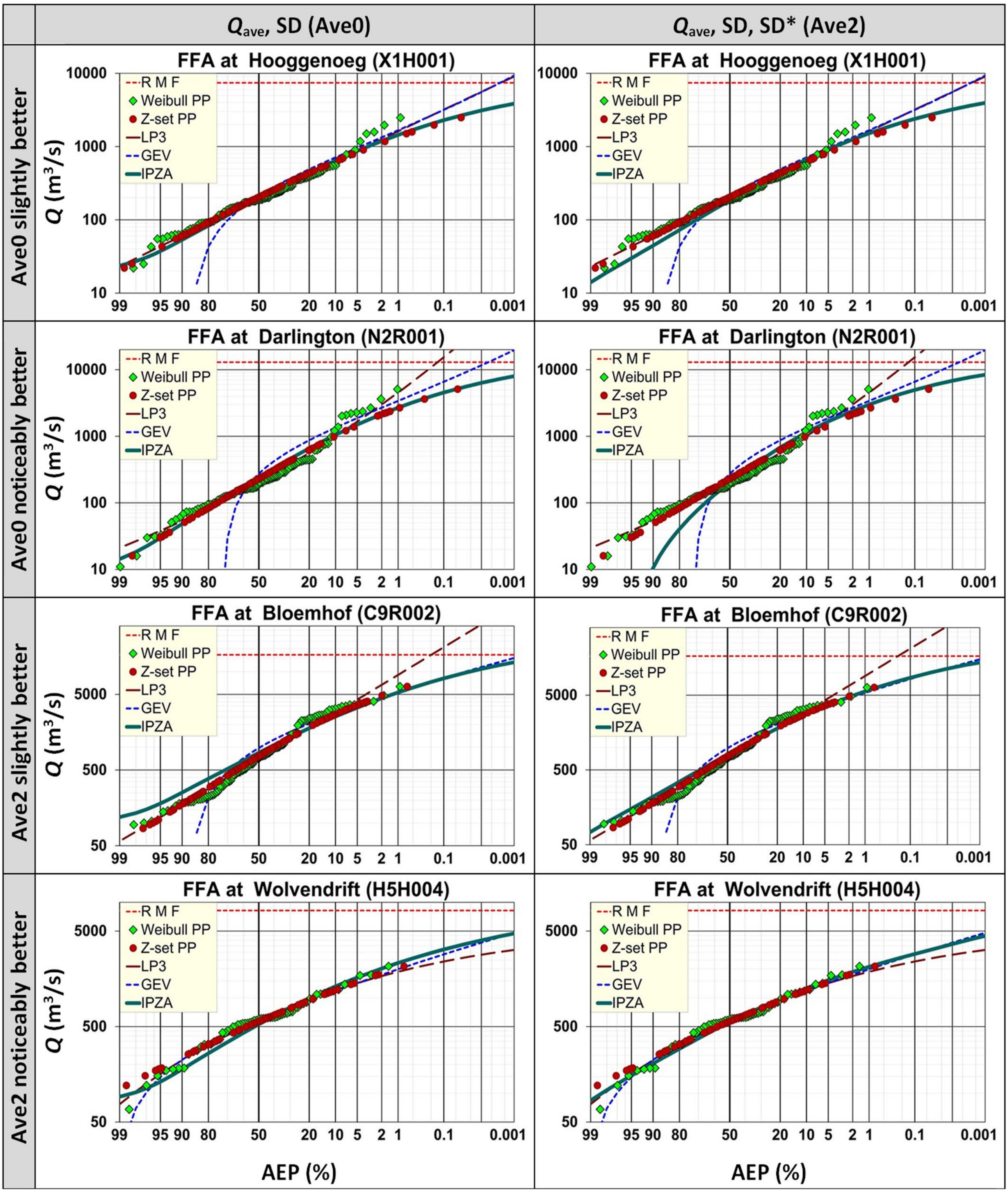

Figure 4. Compare Clusters Ave0 and Ave2 to illustrate effect of introducing SD*

- The effect of introducing the SD* statistic is illustrated in Fig. 4 (Ave0 vs Ave2).
- The effect of replacing Q_{ave} with Q_{gmn} is illustrated in Fig. 5 (Ave2 vs Gmn2).
- In all ensuing figures, flood quantile plots are fitted using method of moments.
- The estimated RMF (regional maximum flood; after Kovács, 1988) is also depicted on the relevant figures.

Combining the results of Ave2 (for AEPs ≤ 50%) and Gmn2 (for AEPs > 50%) was considered; however, this created a discontinuous transition in the region where the AEP = 50%.

To confirm that Cluster Ave2 might not be further improved, by inadvertently excluding more pertinent statistics, it was compared with Cluster Ave3 (Q_{ave}, SD, SD^{*0}) and Cluster Ave4 (Q_{ave}, SD, SD**) and the results are presented in Table 6, where the MAE was used in the AEP range > 50% and MAPE was used in the AEP range ≤ 50%. SD^{*0} 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.

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 cluster.

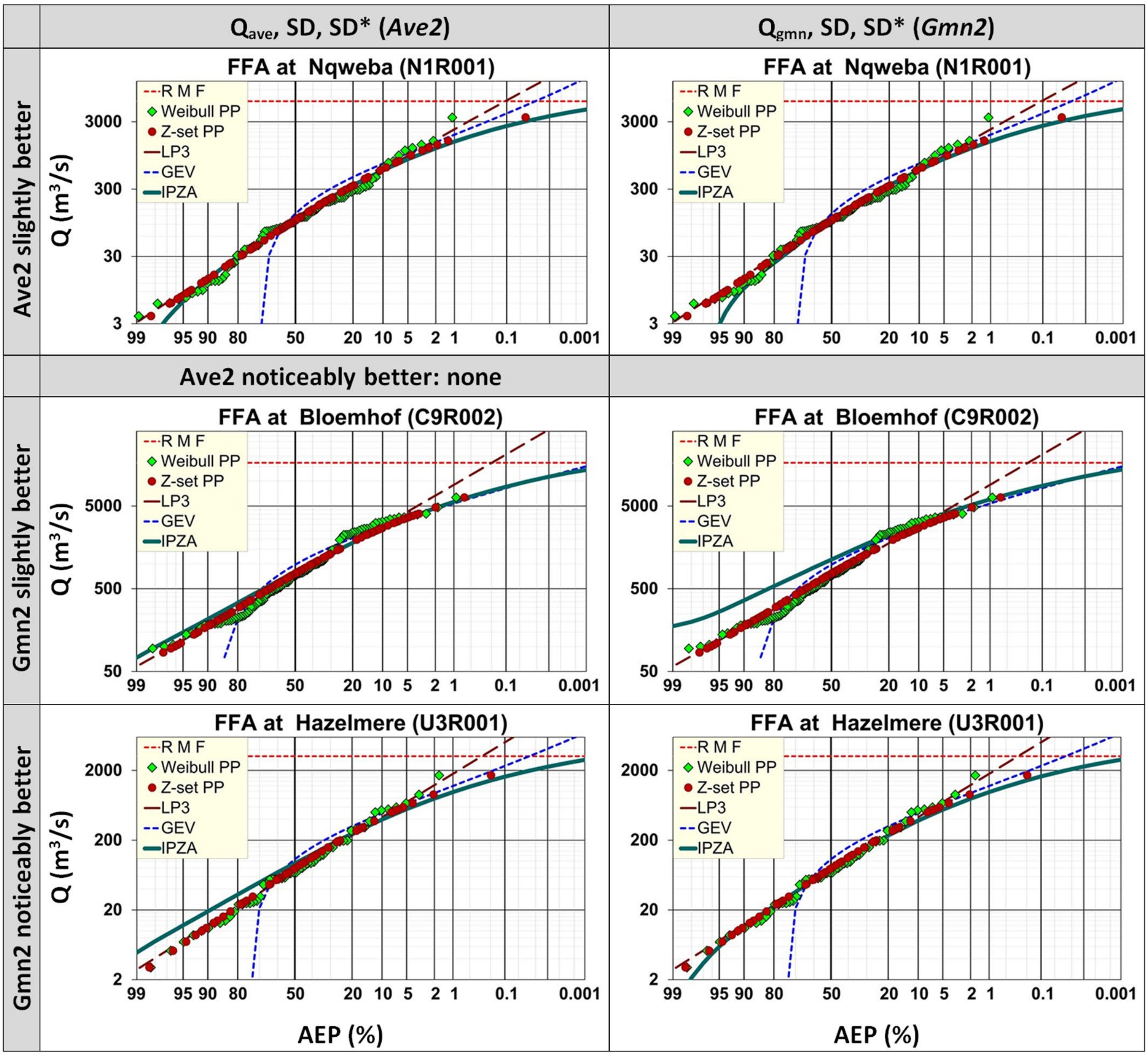

Figure 5. Compare Clusters Ave2 and Gmn2 to illustrate effect if Q_{gmn} replaces Q_{ave}

Table 6. Ranking results of various clusters of Q_{ave}

Site	MAE used for AEP range > 50%						MAPE used for AEP range ≤ 50%					
	Clusters						Clusters					
	Ave2	Ave3	Ave4	Ave2	Ave3	Ave4	Ave2	Ave3	Ave4	Ave2	Ave3	Ave4
	Q_{ave}, including			Ranking			Q_{ave}, including			Ranking		
	SD, SD*	SD, SD^{*0}	SD, SD**	1	2	3	SD, SD*	SD, SD^{*0}	SD, SD**	1	2	3
				Σ						Σ		
				45	89	112				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

Table 6 Continued. Ranking results of various clusters of Q_{ave}

Site	MAE used for AEP range > 50%						MAPE used for AEP range ≤ 50%					
	Clusters						Clusters					
	Ave2	Ave3	Ave4	Ave2	Ave3	Ave4	Ave2	Ave3	Ave4	Ave2	Ave3	Ave4
	Q_{ave}, including			Ranking			Q_{ave}, including			Ranking		
	SD, SD*	SD, SD*⁰	SD, SD**	1	2	3	SD, SD*	SD, SD*⁰	SD, SD**	1	2	3
				Σ						Σ		
				45	89	112				78	80	88
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

Phase 3: Flood quantile model

Estimated regression coefficients, for the independent variables used in the MRAs of Cluster Ave2, are provided in Table 7 and Table 8, for every chosen AEP with the associated standardised variate W_P.

To limit the size of the tables presented, only the most common AEPs ≤ 50% are included. A comparison of Q_{obs} with Q_{est} results, from MRA for Cluster Ave2, is provided in Table 9, with illustrative scatterplots in Fig. 6.

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

- Relevant statistics for B1R001: Q_{ave} = 280 m^3/s, SD = 384 m^3/s and SD* = 317 m^3/s (Table A2)
- At AEP = 1%: the respective regression coefficients are 1.1443, 1.0642 and 2.5224 (Table 8)
- Therefore, $Q_{est} = 1.1443 \cdot 280 + 1.0642 \cdot 384 + 2.5224 \cdot 317 =$ 1 528.66 m^3/s

Table 7. Estimated MRA regression coefficients at corresponding AEPs (99.5%–20%)

AEP(%)	99.5	99	95	90	80	70	60	**50**	40	30	20
W_P	–1.6674	–1.5272	–1.0972	–0.834	–0.4759	–0.1856	0.0874	**0.3665**	0.6717	1.0309	1.4999
MRA regression coefficients											
Q_{ave}	0.1756	0.2225	0.3889	0.5021	0.6779	0.8308	0.9688	**1.1036**	1.2365	1.3645	1.4769
SD	–0.0132	–0.0155	–0.0276	–0.0361	–0.0567	–0.0798	–0.0977	**–0.1240**	–0.1327	–0.1548	–0.1294
SD*	–0.1221	–0.1540	–0.2509	–0.3035	–0.3603	–0.3807	–0.3757	**–0.3301**	–0.2659	–0.1039	0.1416

Table 8. Estimated MRA regression coefficients at corresponding AEPs (10%–0.001%)

AEP(%)	10	5	2	1	0.5	0.2	0.1	0.03	0.01	0.003	0.001
W_P	2.2504	2.9702	3.9019	4.6001	5.2958	6.2136	6.9073	8.1116	9.2103	10.4143	11.5129
MRA regression coefficients											
Q_{ave}	1.5089	1.4806	1.2683	1.1443	0.9282	0.6254	0.4426	0.2181	0.0025	-0.0513	0.0838
SD	0.0275	0.1991	0.6210	1.0642	1.6532	2.4266	3.0775	4.3237	5.3624	6.6915	8.0279
SD*	0.7250	1.2545	2.0693	2.5224	2.8720	3.3928	3.6379	3.7631	4.0032	3.7216	3.0235

Table 9. MRA estimates; Q_{est} vs Q_{obs} for Cluster 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

**The value in the shaded cell is illustrated below in estimating the 1% flood peak at B1R001*

Table 9 Continued. MRA estimates; Q_{est} vs Q_{obs} for Cluster 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}
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
R^2				0.9988		0.9995		0.9987		0.9985		0.9979		0.9979		0.9975	

**The value in the shaded cell, is illustrated below in estimating the 1% flood peak at B1R001*

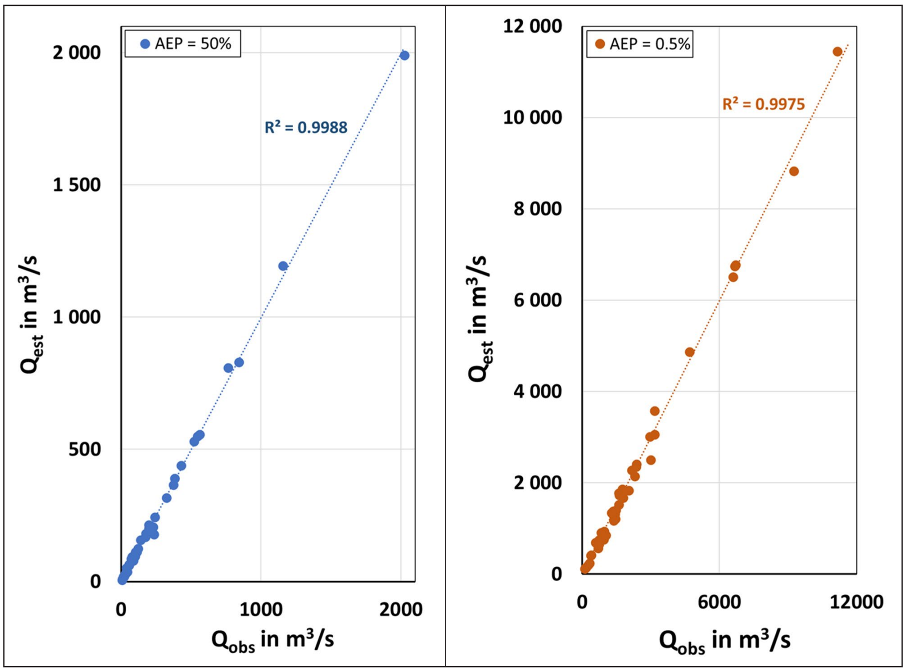

Figure 6. Scatterplots illustrating Q_{est} vs Q_{obs} at AEP = 50% and AEP = 0.5%

The regression coefficients become the frequency factors and a polynomial function (Eq. 4) was fitted to establish a continuous relationship between the frequency factors and W_P to estimate intermediate values. The graphs, showing the relationships between W_P and the estimated frequency factors K_{P-Q}, K_{P-SD} and K_{P-SD*}, respectively associated with Q_{ave}, SD and SD*, are provided in Fig. 7.

From the selected Ave2 cluster, the parameters for the flood quantile model are Q_{ave}, SD and SD*. Equation 5 can thus be rephrased as the general form of the IPZA flood quantile equation:

$$Q_{AEP} = K_{P-Q} \cdot Q_{ave} + K_{P-SD} \cdot SD + K_{P-SD*} \cdot SD^* \qquad (6)$$

where: K_{P-Q}, K_{P-SD} and K_{P-SD*} are the frequency factors for their respective parameters.

The estimated frequency factors for selected AEP values are provided in Table 10.

Assessment of IPZA

In this section 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 or flood quantile model tends to overestimate or underestimate. Applicable criteria are not readily available and appear to depend on the focus of the study. Haddad and Rahman (2012),

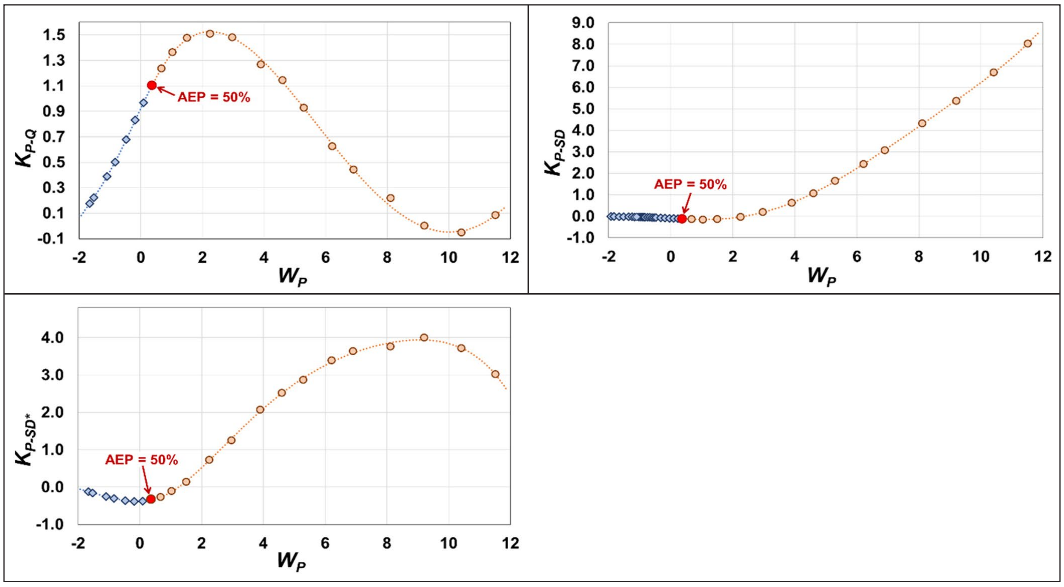

Figure 7. Graphs showing the relationship between W_P and the frequency factors

Table 10. Frequency factors for IPZA

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

Table 11. Criteria to establish the tendency of IPZA to under- or overestimate

Limit	Criteria (all Q values 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	$(Q_{est}/Q_{obs} - 1) > 20\%$

Table 12. Number of sites compliant with set criteria

Criteria	AEP > 50%	AEP ≤ 50%
	Number of sites	
Underestimation unacceptable	6	0
Underestimation acceptable	3	3
Estimation good	19	37
Overestimation acceptable	9	1
Overestimation unacceptable	4	0

considering the focus 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 completely unacceptable for this study and the suggested criteria in Table 11 were utilised to determine whether IPZA tends to underestimate or overestimate, and whether it was considered acceptable or not.

The criteria were applied to all sites at every AEP and the average outcome was calculated for every site separately for the AEPs > 50% and the AEPs ≤ 50%. The number of sites compliant with the criteria is depicted in Table 12.

In the design flood range IPZA performs very well using the defined criteria. For AEP > 50% the performance is slightly lower, but it still outperforms the GEV in this range (refer to Fig. 3 to Fig. 5, as well as Fig. 8).

Outliers

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 study sites with the most distinctive outliers are depicted in Fig. 8 as an illustration.

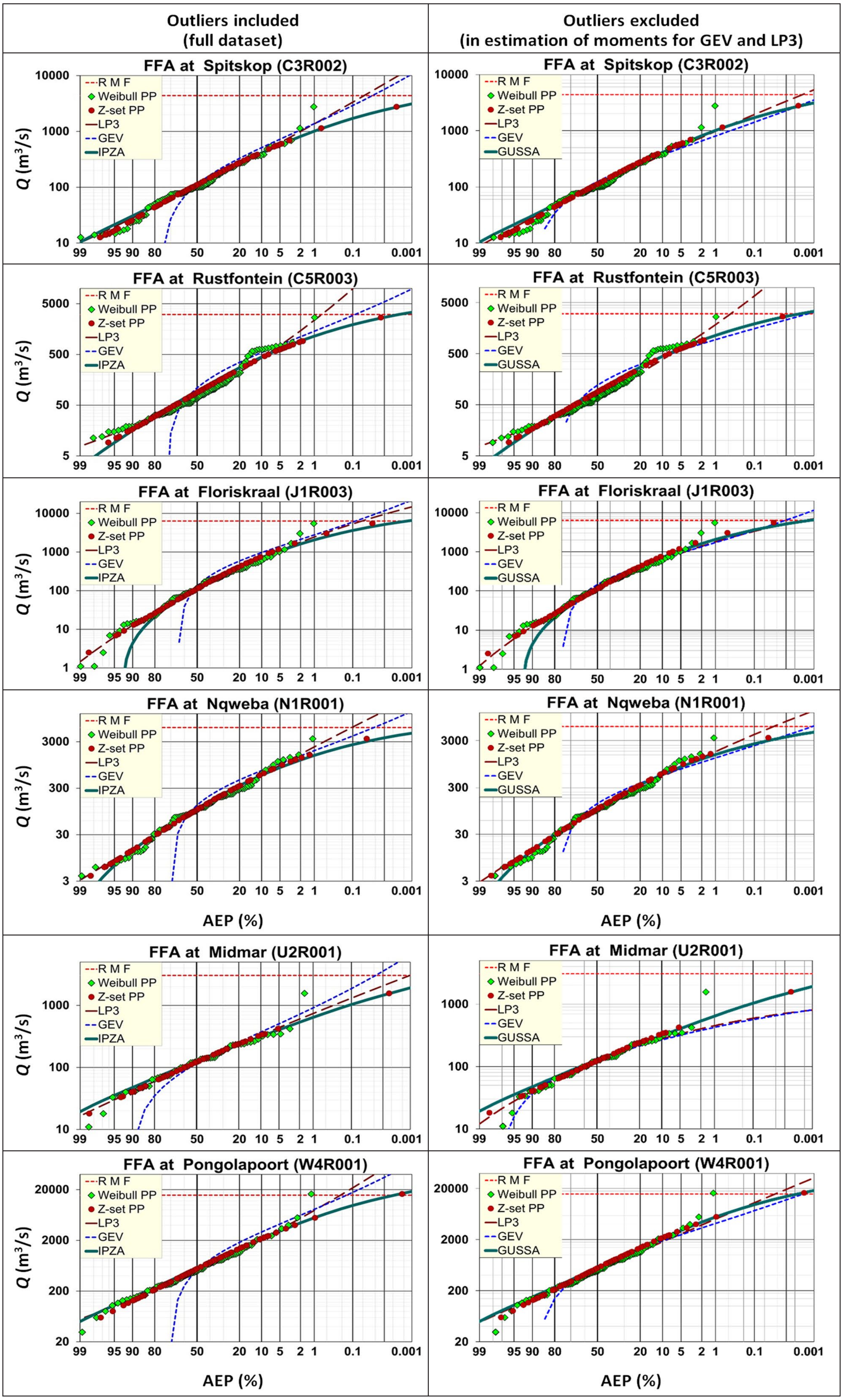

Figure 8. Illustration of impact of outliers on IPZA

It is important to note in Fig. 8 that:

- The PPs represent the full AMS for both columns (outlier still part of the AMS)
- The outlier was removed in the estimation of the moments for GEV and LP3 to assess if it would fit the rest of the data better (IPZA was developed to accommodate outliers – there is no need to remove it)

From Fig. 8 it is evident that the GEV and LP3 probability distributions, with the outlier excluded in the estimation of moments, plot much closer to IPZA, especially in the case of the GEV – the exception being the Midmar Dam, where PILFs (potential influential low flows) can severely influence the LP3.

DISCUSSION

The primary objective of this study was to propose a more consistent flood quantile methodology. The Z-set PP established the foundation for the development of the proposed IPZA flood quantile model. Although only 41 sites were used in the study, due to record length and data accuracy constraints, the coverage of sites across South Africa and the diversity of meteorological and hydrological conditions were well covered.

The study leading to the culmination of the IPZA flood quantile model, has unveiled some noteworthy observations:

- Wheeler (2022) 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 simple to use.
- Wheeler (2022) also explained that no skewed probability model exists to fit strongly skewed data. To use the MRA approach to estimate parameters for a flood quantile model 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 was proved correct. This might seem to be a surprising result, but it could have been expected, considering the statement by Wheeler (2011) 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 be the reason why some distributions often do not perform well, considering that the sample *g* is most probably not representative of the skewness of the underlying distribution.
- The hypothesis that the addition of SD* might improve results, was also proved correct. It appears that SD* might compensate for the instances where SD is affected by extremely high flood peaks, resulting in an inaccurate slope for the relative distribution of data points. Where no outlier is present SD ≈ SD*, with no effect on results, as illustrated in Fig. 4 at Wolvendfrift and Bloemhof Dam.
- Wide-ranging speculation about whether the median statistic should be used instead of the average, was challenged by the outcome of this study, demonstrating the average value to be the best option (even the geometric mean proved to be a more viable option than the median).
- The IPZA seems to be virtually unaffected by outliers. Consequently, the problem to decide whether to ignore a suspected outlier, or not, should no longer be a concern.
- The IPZA 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.

CONCLUSION

The Z-set PP, which adjusted the PP of outliers, provided the basis for finding a corresponding solution for a more consistent flood quantile methodology. Whether the IPZA can be considered as bounded, or not, is inconclusive at this stage. However, from the results presented it can be concluded that the IPZA is probably more likely to have an upper bound than the GEV or LP3.

The IPZA seems to be much more stable and consistent in the estimation of flood quantiles, compared to the GEV and the LP3, as depicted on all the included FFA results.

Further research to improve the IPZA should 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 study sites did cover diverse meteorological and hydrological conditions, it should be tested with datasets from outside South Africa to further evaluate its performance.

It is recommended that the IPZA be used as a valuable addition 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 FFA method 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) is used to determine a 'best fit', a visual check to verify, or even change, the outcome is strongly recommended.

REFERENCES

ALEXANDER WJR (1990) *Flood Hydrology for Southern Africa*. South African National Committee on Large Dams, 1990.

BALL J, BABISTER M, NATHAN R, WEEKS W, WEINMANN E, RETALLICK M and TESTONI I (eds) (2019) *Australian Rainfall and Runoff: A Guide to Flood Estimation*. Commonwealth of Australia (Geoscience Australia). URL: https://arr.ga.gov.au/

BROWNLEE J (2018) How to use statistics to identify outliers in data. Machine Learning Mastery. URL: https://machinelearningmastery.com/how-to-use-statistics-to-identify-outliers-in-data/ (Accessed 8 May 2019).

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

DYSON LL and VAN HEERDEN J (2001) The heavy rainfall and floods over the northeastern interior of South Africa. *S. Afr. J. Sci.* **97** (3) 80–86.

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

FROST J (2019) 5 ways to find outliers in your data. Statistics by Jim. URL: https://statisticsbyjim.com/basics/outliers/ (Accessed 8 May 2019).

GANESH S and CAVE V (2018) P-values, P-values everywhere! *N. Zeal. Vet. J.* **66** (2) 55–56. https://doi.org/10.1080/00480169.2018.1415604

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. *J. Hydrol.* **430–431** 142–161. https://doi.org/10.1016/j.jhydrol.2012.02.012

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. https://doi.org/10.3390/w9070481

KOVÁCS Z (1988) Regional maximum flood peaks in Southern Africa. Technical Report 137. Department of Water Affairs, South Africa.

KOVÁCS ZP, DU PLESSIS DB, BRACHER PR, DUNN P and MALLORY GCL (1985) Documentation of the 1984 Domoina Floods. Technical Report 122. Department of Water Affairs, South Africa.

MKHANDI S, OPERE AO and WILLEMS P (2005) Comparison between annual maximum and peaks over threshold models for flood frequency prediction. In: *Proceedings of the International Conference on UNESCO FRIEND/Nile Project: Towards a better Cooperation*, 12–15 November 2005, Sharm-El-Sheikh, Egypt.

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

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.

ROBERTS CPR and ALEXANDER WJR (1982) Lessons learnt from the 1981 Laingsburg Flood. *Die Siviele Ingenieur (Civil Engineering)* **24** (1) 17–27.

VAN BLADEREN D and BURGER CE (1989) Documentation of the September 1987 Natal Floods. Technical Report 139. Department of Water Affairs, South Africa.

VAN DER SPUY D and DU PLESSIS JA (2022a) Flood frequency analysis – Part 1: Review of the statistical approach in South Africa. *Water SA* **48** (2) 110–119. https://doi.org/10.17159/wsa/2022.v48.i2.3848.1

VAN DER SPUY D and DU PLESSIS JA (2022b) Flood frequency analysis – Part 2: Development of a modified plotting position. *Water SA* **48** (2) 120–130. https://doi.org/10.17159/wsa/2022.v48.i2.3848.2

WHEELER DJ (2009) Transforming the data can be fatal to your analysis. Quality Digest. URL: https://www.qualitydigest.com/inside/quality-insider-column/transforming-data-can-be-fatal-your-analysis.html (Accessed 20 December 2022).

WHEELER DJ (2011) Problems with skewness and kurtosis, Part Two: What do the shape parameters do? Quality Digest. URL: https://www.qualitydigest.com/inside/quality-insider-article/problems-skewness-and-kurtosis-part-two.html (Accessed 4 August 2018).

WHEELER DJ (2022) *How Can a Control Chart Work Without a Distribution?* Quality Digest. URL: https://www.qualitydigest.com/inside/six-sigma-column/how-can-control-chart-work-without-distribution-042522.html? (Accessed 3 May 2022).

APPENDIX

Table A1. Sites megadata

Site			Augment record with data from site[1]	Catchment area (km²)	AMS flood peaks (m^3/s)		Record length (years)	
Number	Name	River			Lowest	Highest	Total	Usable
A2R001	Hartebeespoort	Crocodile	A2H012/013	4 014	13	1 670	97	97
A2R005	Buffelspoort	Sterkstroom	-	119	0.9	398	84	84
A3R002	Klein Maricopoort	Klein Marico	A3H001	1 157	0.6	506	112	110
A4R001	Mokolo	Mokolo	A4H005	4 319	1.2	1 291	57	57
A8R001	Nzhelele	Nzhelele	A8H001	830	2.0	2 795	86	74
B1R001	Witbank	Olifants	B1H001	3 579	3.5	2 565	114	112
B2R001	Bronkhorstspruit	Bronkhorstspruit	B2H001	1 244	2.2	995	114	114
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
C1R002	Grootdraai	Vaal (upper)	C1H001	7 982	42	2 275	114	114
C2R001	Boskop	Mooi	C2H001	3 297	2.1	112	114	114
C3R002	Spitskop	Harts	C3H007	26 730	13	2 760	96	96
C5R001	Tierpoort	Tierpoort	-	922	2.1	1 670	97	93
C5R002	Kalkfontein	Riet	C5H001	10 260	3.2	9 800	106	106
C5R003	Rustfontein	Modder	C5H003	937	9.2	2 670	101	101
C7R001	Koppies	Renoster	-	2 142	23	1 480	99	99
C9R002	Bloemhof	Vaal (middle)	C9H006	108 360	85	6 340	110	108
D3R002	Gariep	Orange (upper)	D3H002	70 655	106	11 460	115	114
D6R002	Smartt Syndicate	Ongers	-	13 340	1.0	1 190	92	71
D7H005	Upington	Orange (lower)	D7H003	361 512	130	8 315	87	87
E1R002	Clanwilliam	Olifants	-	2 025	30	1 385	84	84
H5H004	Wolvendrift	Bree	H5H002	6 713	68	2 137	65	65
J1R003	Floriskraal	Buffels	J1H004	4 024	1.1	5 475	98	96
J2R001	Calitzdorp	Nels	-	37	0.6	191	99	89
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
N2R001	Darlington	Sondags	N2H002	16 820	11	5 090	96	96
Q1R001	Grassridge	Great Brak	-	4 326	13	805	94	94
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
V6H002	Tugela Ferry	Tugela	-	12 862	38	2 438	92	92
W4R001	Pongolapoort	Phongolo	W4H002	7 800	31	16 450	71	71
X1H001	Hooggenoeg	Komati	-	5 503	8.3	2 481	110	110

[1]*Flow gauging weirs close to an existing dam was used to augment the inflow record at the dam.*

Table A2. Selected statistics considered in the MRA

Site	Full record[1]							Record, excluding some data				
	Q_{min}	Q_{max}	Q_{ave}	Q_{mdn}	Q_{gmn}	SD	g	Q_{ave}*	Q_{mdn}*	SD*	SD*^{0}	g*
A2R001	13.0	1 670	278	181	164	301	2.133	263	180	267	267	1.749
A2R005	0.90	398	33.9	22.0	18.2	49.8	5.126	29.5	22.0	29.6	29.7	1.982
A3R002	0.60	506	37.4	19.5	15.4	63.3	4.696	33.1	18.0	44.7	44.8	2.916
A4R001	1.17	1 291	181	58.0	68.7	281	2.509	161	53.5	240	241	2.495
A8R001	2.00	2 795	227	89.5	98.8	409	4.380	192	88.0	278	279	3.334
B1R001	3.50	2 565	280	124	129	384	3.032	259	120	317	318	2.155
B2R001	2.20	995	144	79.0	74.9	188	2.690	137	78.0	171	171	2.637
B3R001	2.30	665	58.5	27.0	29.9	96.5	4.084	51.3	26.0	70.7	70.9	3.134
B6R003	19.0	1 615	283	229	181	282	2.254	264	224	233	232	1.475
B7R001	0.20	915	89.4	38.0	35.4	137	3.795	77.0	38.0	92.6	92.8	1.829
C1R002	42.0	2 275	494	367	357	421	1.849	478	366	388	387	1.634
C2R001	2.10	112	23.2	18.0	15.2	22.0	1.755	22.5	18.0	20.4	20.4	1.584
C3R002	12.6	2 760	192	99.5	107	318	6.008	165	99.0	178	178	2.614
C5R001	2.10	1 670	101	34.0	36.9	226	4.807	84.0	33.5	156	157	3.718
C5R002	3.20	9 800	547	210	221	1 223	5.642	459	207	824	827	5.076
C5R003	9.20	2 670	193	67.0	84.5	336	4.567	168	66.5	226	227	1.865
C7R001	23.0	1 480	251	170	164	264	2.413	239	167	234	234	2.144
C9R002	85.0	6 340	1 216	719	710	1 251	1.507	1 168	712	1 154	1 154	1.226
D3R002	106	11 460	2 614	1 976	1 941	2 094	1.769	2 536	1 967	1 928	1 923	1.532
D6R002	1.00	1 190	122	61.0	50.2	195	3.870	106	59.5	147	148	3.704
D7H005	130	8 315	1 730	1 294	1 070	1 690	1.762	1 653	1 285	1 540	1 540	1.566
E1R002	30.0	1 385	408	307	314	309	1.459	397	304	291	290	1.415
H5H004	67.9	2 137	661	595	545	409	1.367	638	588	367	342	1.050
J1R003	1.10	5 475	298	118	102	665	5.934	243	115	398	399	4.438
J2R001	0.60	191	20.4	11.0	11.0	29.1	3.573	18.5	11.0	22.8	22.9	2.878
J2R003	0.20	196	17.5	5.80	5.98	32.9	3.557	15.5	5.65	26.8	26.9	3.358
J3R001	1.60	2 755	184	42.0	43.8	428	3.946	160	41.0	347	348	3.683
N1R001	3.90	3 470	243	106	95.4	455	4.760	207	101	301	302	2.785
N2R001	11.0	5 090	467	193	220	796	3.523	418	192	641	645	2.980
Q1R001	13.0	805	123	80.5	85.0	139	3.037	116	80.0	121	121	2.924
Q4R002	4.80	802	142	93.5	93.3	147	2.514	135	89.0	130	130	2.345
Q5R001	9.00	3 888	452	236	229	639	2.986	417	226	538	539	2.513
S6R002	1.34	881	108	50.4	51.5	146	3.007	95.7	49.8	110	110	1.840
T2R001	8.40	920	285	235	177	237	0.883	269	225	215	214	0.721
U2R001	11.0	1 565	166	127	118	202	5.565	143	125	93.8	93.0	0.915
U3R001	3.00	1 700	174	69.0	72.3	279	3.521	147	69.0	192	192	2.119
V1R003	41.0	2 140	566	415	407	437	1.258	546	410	401	399	0.997
V3R003	13.4	382	102	88.0	81.3	73.9	1.793	97.9	85.5	66.3	65.9	1.586
V6H002	38.0	2 438	949	897	800	492	0.761	933	896	469	462	0.634
W4R001	31.0	16 450	1 003	488	528	1 931	6.470	820	485	929	931	2.676
X1H001	8.30	2 481	302	188	196	371	3.564	282	188	308	308	3.167

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

**indicates the largest value (Q_{max}) is omitted from the dataset to determine statistic*
**^{0}indicates that both the largest value (Q_{max}) and the smallest value (Q_{min}) are omitted from the dataset to determine statistics*

Table A3. Flood frequency data considered in the MRA approach

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														
A2R001	12	26	39	68	177	422	630	841	1 137	1 364	1 591	1 886	2 102	2 757	3 272
A2R005	0.93	3.2	3.7	6.8	22	51	81	112	160	223	262	309	342	465	621
A3R002	0.47	1.2	2.3	4.5	17	54	95	140	208	262	317	390	444	610	736
A4R001	2.0	5.5	10	19	73	261	443	649	953	1 204	1 438	1 750	1 987	2 791	3 629
A8R001	3.8	11	17	32	103	319	541	742	1 173	1 454	1 795	2 221	2 542	3 450	4 205
B1R001	5.2	13	22	42	141	418	683	969	1 384	1 620	1 793	2 338	2 749	3 571	4 285
B2R001	3.7	9.5	15	27	81	215	334	470	662	809	956	1 149	1 290	1 717	2 039
B3R001	1.9	4.3	7.0	12	30	79	129	181	258	332	395	486	550	770	1 026
B6R003	16	32	49	79	198	427	630	820	1 089	1 299	1 474	1 792	1 964	2 611	3 058
B7R001	0.72	1.5	4.7	9.8	39	138	236	350	466	558	693	936	1 035	1 331	1 587
C1R002	48	90	123	181	373	721	983	1 260	1 614	1 888	2 162	2 523	2 795	3 676	4 503
C2R001	1.5	3.0	4.3	6.9	16	35	49	65	87	101	118	140	157	217	281
C3R002	8.6	17	25	44	112	273	421	587	824	1 027	1 449	1 735	1 936	2 531	3 160
C5R001	1.4	3.4	5.9	12	34	133	237	357	534	680	834	1 116	1 283	1 828	2 489
C5R002	7.1	21	38	72	236	738	1 301	1 973	2 985	3 821	4 700	5 907	6 841	9 951	12 829
C5R003	4.5	11	17	30	89	256	432	628	897	1 120	1 354	1 674	1 923	2 768	3 610
C7R001	18	34	49	76	173	368	527	694	928	1 105	1 282	1 512	1 680	2 188	2 579
C9R002	57	114	178	292	764	1827	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
D6R002	1.2	3.7	7.0	14	56	185	313	404	678	863	1 043	1 254	1 365	1 775	2 113
D7H005	89	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
E1R002	38	90	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	1194	1 519	1 815	2 051	2 352	2 642	2 903	3 787	4 702
J1R003	1.7	6.6	13	26	112	421	783	1 141	1 753	2 360	2 991	3 760	4 344	6 026	7 288
J2R001	0.66	1.6	2.5	4.3	11	30	46	65	91	113	138	165	184	248	310
J2R003	0.14	0.39	0.78	1.7	6.0	23	44	66	98	129	159	195	227	344	474
J3R001	0.82	2.5	4.9	11	45	203	406	665	1 068	1 409	1 756	2 255	2 582	3 670	4 532
N1R001	2.8	7.9	14	30	104	342	597	898	1 342	1 685	2 037	2 522	2 906	4 024	4 888
N2R001	12	30	48	84	229	618	1 005	1 465	2 129	2 634	3 156	3 896	4 455	6 218	7 743
Q1R001	11	22	29	43	87	174	244	317	418	500	579	689	768	1 065	1 381
Q4R002	7.7	19	28	43	98	212	301	397	531	642	736	874	978	1 319	1 639
Q5R001	12	29	48	86	241	646	1 033	1 441	2 076	2 518	2 962	3 604	4 115	5 486	6 763
S6R002	1.8	5.3	9.0	17	55	162	266	376	511	598	730	935	1 029	1 308	1 657
T2R001	11	27	41	72	199	454	652	852	1 083	1 235	1 374	1 628	1 775	2 236	2 666
U2R001	15	30	40	59	122	239	335	495	694	827	948	1 111	1 237	1 530	1 936
U3R001	2.1	4.8	12	23	78	251	426	614	911	1 143	1 378	1 691	1 875	2 414	2 881
V1R003	52	94	132	201	429	848	1 126	1 447	1 875	2 176	2 367	2 758	3 122	4 007	4 867
V3R003	15	26	36	47	85	147	188	237	290	334	377	436	481	635	794
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
W4R001	36	83	129	210	546	1 376	2 150	3 049	4 386	5 500	6 675	8 287	9 536	13 760	17 997
X1H001	20	41	60	92	202	431	628	837	1 129	1 358	1 591	1 904	2 141	2 914	3 630

Table A4. MAPE ranking for AEP range ≤ 50%

Site	Clusters									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*	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