

# Uncertainty in calibrating generalised Pareto distribution to rainfall extremes in Lake Victoria basin

Charles Onyutha and Patrick Willems

## ABSTRACT

Uncertainty in the calibration of the generalised Pareto distribution (GPD) to rainfall extremes is assessed based on observed and large number of global climate model rainfall time series for nine locations in the Lake Victoria basin (LVB) in Eastern Africa. The class of the GPD suitable for capturing the tail behaviour of the distribution and extreme quantiles is investigated. The best parameter estimation method is selected following comparison of the method of moments, maximum likelihood, L-moments, and weighted linear regression in quantile plots (WLR) to quantify uncertainty in the extreme intensity quantiles by employing the Jackknife method and nonparametric percentile bootstrapping in a combined way. The normal tailed GPD was found suitable. Although the performance of each parameter estimation method was acceptable in a number of evaluation criteria, generally the WLR technique appears to be more robust than others. The difference between upper and lower limits of the 95% confidence intervals expressed as a percentage of the empirical 10-year rainfall intensity quantile ranges from 9.25 up to 59.66%. The assessed uncertainty will be useful in support of risk based planning, design and operation of water engineering and water management applications related to floods in the LVB.

**Key words** | extreme value analysis, generalised Pareto distribution, Lake Victoria basin, parameter estimation, rainfall extremes, uncertainty analysis

**Charles Onyutha** (corresponding author)

**Patrick Willems**

Hydraulics Laboratory,  
KU Leuven,  
Kasteelpark Arenberg 40,  
Leuven B-3001,  
Belgium

E-mail: [charles.onyutha@bwk.kuleuven.be](mailto:charles.onyutha@bwk.kuleuven.be)

**Patrick Willems**

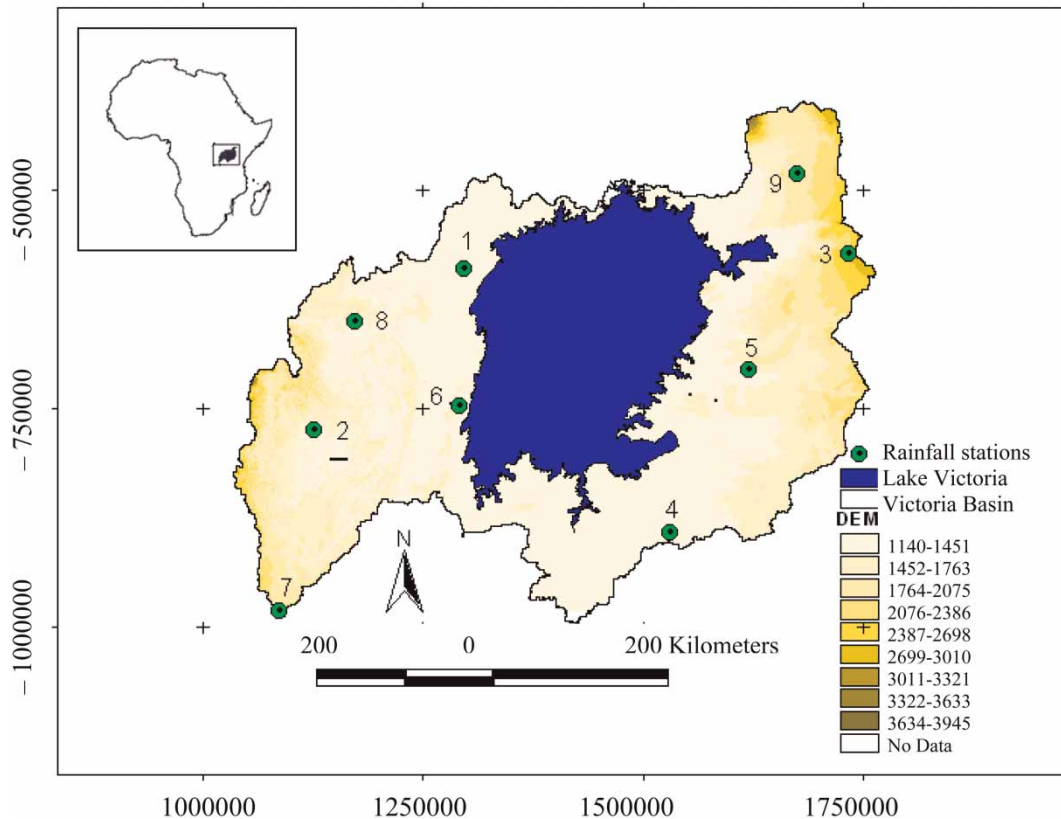
Department of Hydrology and Hydraulic  
Engineering,  
Vrije Universiteit Brussels,  
Pleinlaan 2, 1050 Elsene,  
Brussel,  
Belgium

## INTRODUCTION

The Lake Victoria basin (LVB), which has a total catchment area of about 184,000 km<sup>2</sup>, stretches 355 km in the east-west direction (31°37' E to 34°53' E) and 412 km in the north-south direction (00°30' N to 3°12' S). It comprises of Lake Victoria which is the largest freshwater body in Africa and the second largest in the world. Lake Victoria, which has a shoreline of 4,828 km and a surface area of 68,800 km<sup>2</sup>, is situated at an altitude of 1,134 m above sea level. It has a relatively small drainage basin which is slightly less than three times the lake's surface in area (Figure 1). According to Awange & Ong'ang'a (2006), the climate of the LVB is generally tropical humid, with temperatures ranging from 15 °C in the highlands, to 28 °C in the semi-arid areas. The mean annual rainfall in the LVB varies from about 886 to 2600 mm. According to Anyah *et al.* (2006), the general climate of the LVB ranges from a modified

equatorial type with substantial rainfall occurring throughout the year, especially over the lake and its vicinity, to a semi-arid type characterised by intermittent droughts over some areas located even within short distances from the lake shore. Albeit Lake Victoria is geopolitically shared by Kenya, Uganda and Tanzania in the proportion of 6%, 43% and 51% respectively (Awange & Ong'ang'a 2006), the LVB stretches in five countries including Burundi, Kenya, Rwanda, Tanzania and Uganda.

On a yearly basis, extreme rainfall and associated flooding events inflict severe damage to public life and property in many parts of the world. According to the United Nations Environmental Programme (UNEP) (2006), flooding is one of the most devastating natural hazards in the LVB. Areas close to Lake Victoria and generally low-lying parts of the LVB are characterised by episodes of floods, for instance,



**Figure 1** | Locations of the meteorological stations (see Table 1 for details); the longitudes and latitudes are in kilometres.

downstream of River Nzoia and River Nyando, around Budalang'i and Kano plains (Gichere *et al.* 2013). These frequent flooding events in the LVB are due to conversion of wetlands, deforestation of the catchments and destruction of river embankments (UNEP 2006). The flood damage in the LVB is exacerbated by the already existing crisis of an untamed poverty associated with high population growth of its inhabitants. Importantly, for development of the LVB through proper management of its water resources under variable climatic conditions and anthropogenic influences, there is an inevitable need for an accurate descriptive study of hydrometeorological extremes and their recurrence rates or return period  $T$  based on long-term time series of observations of rainfall intensities, discharges or water levels, etc. However, data limitation of the historical time series is a common setback to such a study. This creates uncertainty in the estimation of rainfall intensity quantiles for planning, design, operation and management of water resources. The quantification and reduction of predictive

uncertainty is a crucial issue in statistical modelling of a hydrological variable that allows judgment of the degree of confidence in estimated results. It needs to be taken into account in decision making related to risk-based water engineering and management. Some of the ways for dealing with uncertainty in quantile estimates include regionalisation and the at-site approach (Yang *et al.* 2010). Regionalisation attempts in the LVB were made by Kizza *et al.* (2013) on the parameters of a conceptual water balance model to estimate catchment inflows, Nyeko-Ogiramoi *et al.* (2012) on an elusive search for regional flood frequency estimates, and Onyutha & Willems (2014) on empirical-statistical means of estimating flood quantiles from physiographic characteristics. Uncertainty analysis of flood quantile estimates at a regional level was also carried out by Mkhandi *et al.* (1996) for only Tanzania that shares the southern part of the LVB. However, the accuracy of quantiles obtained for ungauged watersheds based on a regional analysis depends on the density of the stations,

and above all the quality of the data used. This may limit the validity of the regionalisation procedures in real life applications. This means that it is important to first deal with the at-site uncertainties, as done in this study, before the data can be used for regionalisation procedures for ungauged catchments. Such at-site uncertainty studies in the LVB were carried out on rainfall-runoff modelling by Kizza *et al.* (2011) and on sampling peak flow extremes by Onyutha & Willems (2013).

One of the commonly used methods for assessing errors in statistical models is the Jackknife approach of Tukey (1958). Bootstrapping was introduced by Efron (1979) to further strengthen the Jackknife method for estimating bias or standard error. The use of the Jackknife method can be found in a number of studies (see e.g. Miller 1974; Hinkley 1977; Wu 1986; Beersma & Buishand 1999; Friedl & Stampfer 2002). Other statistical approaches of uncertainty assessment can be found elaborated in Montanari (2011).

Hydrometeorological extreme value analysis is commonly based on partial-duration-series (PDS) or peak-over-threshold (POT) extremes extracted from the available series of historical flow and hydro-climatic observations. Following the extreme value theory by Pickands (1975), generalised Pareto distribution (GPD) can be calibrated to these extremes (see e.g. Madsen *et al.* 1997a, b; Kjeldsen *et al.* 2000; Skaugen 2002; Li *et al.* 2013; Willems *et al.* 2007). In frequency analysis, the reliability of the theoretical quantiles estimated from the extreme value distribution (EVD) obviously depends on the accuracy of the GPD calibration. This is, however, challenged by the proper choice of the type of GPD (normal, heavy or light) and by the selection of the best parameter estimation method. Different parameter estimation methods for GPD are known including, among others, the method of moments (MOM), L-moments (Lmom), maximum likelihood (ML), and the weighted linear regression in quantile plots (WLR).

The objectives of this study are to: (1) determine the most suitable GPD class; (2) find out the best performing parameter estimation method; and (3) assess, based on (1) and (2), the uncertainties in estimating extreme rainfall intensity quantiles for the LVB. Sampling uncertainty of quantile estimates is addressed by applying the delete-one Jackknife method and nonparametric percentile bootstrapping in a combined way.

## DATA SERIES AND CONSIDERATION

The EVD analysis in this study is based on large number (1,143) of daily rainfall intensity series obtained from the following:

- (1) The database of the FRIEND/Nile (River Nile basin Flow Regimes from International, Experimental and Network Data) project (United Nations Educational Scientific & Cultural Organization (UNESCO) 2014). Daily observed rainfall data (for the period 1961–2000) of nine selected meteorological stations distributed in the LVB were used. According to Kizza *et al.* (2009), the 1960s represent a significant upward jump in the LVB rainfall. Variability, trend and step change analysis can give information on the impact of man's activities (such as deforestation, urbanisation, agricultural practices, etc.) on the hydrology of a catchment. Thus the rainfall series starting from 1961 onwards were selected to get an insight of the impact of the step change on sampling uncertainty in rainfall extremes in the LVB. Furthermore, the period 1961–1990 is the baseline for climate studies as recommended by the World Meteorological Organisation. The 30-year averaging period is recommended by the Intergovernmental Panel on Climate Change (IPCC) (2001) as well for a common definition of a baseline climate.
- (2) Global climate model (GCM) data for the selected stations in (1). The daily rainfall time series of a large number of GCMs obtained from the Program for Climate Model Diagnosis and Inter-comparison (PCMDI) database were used in this study. A total of 1,134 daily rainfall series were obtained from the archived daily data at PCMDI (last accessed online at [http://www2-pcmdi.llnl.gov/esg\\_data\\_portal](http://www2-pcmdi.llnl.gov/esg_data_portal) on 8 June 2011). To select the best parameter estimation method to be used for estimating uncertainty in quantiles at the selected stations, large series data are required. The data from the GCMs were used as the readily available synthetic rainfall series over the LVB in order to compliment the station data and to minimise, on the study findings, the effect of the data quality problems at the selected stations. GCM data are also commonly applied on the basis of climate change impact studies, which is one of the important

applications of the extreme quantile estimation dealt with in this paper.

For the rainfall data in (1), a number of potential data problems, for instance, missing values, data entry errors, outliers, etc., were solved by careful inspection. The data were screened and comparison between stations was made using the statistical metrics standard deviation, coefficient of variation ( $C_v$ ), skewness ( $S_k$ ) and kurtosis ( $K_u$ ). Missing data were filled-in using the inverse distance weighted (IDW) interpolation technique of Shepard (1968). The missing rainfall intensity ( $P_X$ ) at station  $X$  for a certain period using rainfall values ( $P_j$ ) at  $g$  neighbouring stations for the same period is given by

$$P_X = \sum_{j=1}^g P_j \times d_j^{-r} \left( \sum_{j=1}^g d_j^{-r} \right)^{-1} \quad (1)$$

where  $d_j$  is the distance between station  $X$  (the one with the missing record) and that in the neighborhood being used for interpolation, and  $r$  is the power parameter (an arbitrary positive real number).

The IDW approach has been applied to climate data such as rainfall in a number of studies including (Legates & Willmott 1990; Frei & Schär 1998; Rudolf & Rubel 2005; Ahrens 2006; Ly *et al.* 2011). It is important to note that the reliability of the IDW interpolation lies in the value of  $r$ . A small value of  $r$  tends to approximate the interpolated record to averages of the neighbouring stations used, while a larger value of  $r$  down-weights

points further away and assigns larger weights to the nearest points (Lu & Wong 2008). According to Dirks *et al.* (1998), interpolation errors appear to be minimised by taking  $r=2$  for daily or monthly data, 3 for hourly and 1 for annual data. However following Goovaerts (2000) and Lloyd (2005),  $r$  is usually set to 2. Hence  $r=2$  was adopted in this study.

Table 1 shows for all the stations the coordinates, station name, station ID, long-term mean,  $C_v$ ,  $S_k$ ,  $K_u$  and maximum intensity ( $M_{ax}$ ) of the 1961–2000 daily rainfall series. The  $C_v$  ranges from 1.4 (station 2) to 2.7 (station 8) which shows that there is considerable variability on a day to day basis in the LVB. The  $M_{ax}$  varies from 79.7 mm/day (station 3) to 129.5 mm/day (station 4).

## METHODOLOGY

### Selection of independent extremes

For frequency analysis, the data are required to be independent and identically distributed. This is normally achievable through extraction from full series, either annual maxima series (AMS) or PDS and/or POT extreme events. In AMS, the time slice is chosen to be one hydrological year. Although extreme events obtained from the AMS model show strong independence, the limitation of their number tends to lower the accuracy of extreme value analysis. The PDS or POT model ensures that the most severe events are selected, independent of the time of occurrence. Hence, it intuitively provides a more consistent definition

**Table 1** | Overview of selected stations

No.	Station ID	Station name	Lon. [deg.]	Lat. [deg.]	Mean [mm/day]	$C_v$ [-]	$S_k$ [-]	$K_u$ [-]	$M_{ax}$ [mm/day]	Country
1	9,031,026	Kamenyamigo	31.67	- 0.30	2.7	2.6	4.1	21.9	90.0	Uganda
2	70,009	Kigali Aero Nyabarongo	30.13	- 1.97	6.6	1.4	2.9	12.0	91.9	Rwanda
3	9,035,002	Londiani Forest	35.60	- 0.15	3.2	2.2	3.3	14.4	79.7	Kenya
4	9,333,005	Maswa Hydromet	33.77	- 3.17	2.5	3.1	5.4	42.3	129.5	Tanzania
5	9,134,008	Nyabassi	34.57	- 1.35	3.9	2.3	3.7	19.7	120.0	Tanzania
6	9,131,001	Rubya Mission	31.62	- 1.72	3.8	2.3	3.8	19.3	124.8	Tanzania
7	10,161	Ruvyironza	29.77	- 3.82	3.6	2.0	3.2	13.7	90.0	Burundi
8	9,030,012	Rwoho Forest	30.55	- 0.85	2.6	2.7	4.4	29.4	120.6	Uganda
9	8,935,076	Turbo Forest	35.02	0.67	3.6	2.2	3.4	16.1	82.5	Kenya

Lon.: Longitude; Lat.: Latitude; deg.: Degree.

of the extreme value region than the AMS approach. Comparison of the relative performance of the AMS and PDS/POT approaches for rainfall series can be found in Madsen *et al.* (1997a), and Madsen *et al.* (1997b). According to Madsen *et al.* (1997a), PDS is more efficient than the AMS in estimation of flood quantiles. Madsen *et al.* (1997b) concluded that in general, one should use PDS with exponentially distributed magnitudes for close to zero shape parameter. It was shown that PDS yields more precise quantile estimates than AMS (Yevjevich & Taesombut 1979; Martins & Stedinger 2001). An overview of the recent advances in PDS/POT modelling of extreme hydrological events can be found in the paper by Rosbjerg & Madsen (2004). These authors summarise important extensions of the PDS/POT method since the mid-1990s. They show the PDS model to be competitive with that of the AMS model and highly efficient for estimation of quantiles. Therefore, given the above support for PDS, it is adopted for the LVB.

In this study, independent rainfall intensity extremes were selected using independence criteria based on threshold values for the time between two successive independent rainfall intensity peaks, the ratio of the minimum rainfall intensity between the two peaks over the peak value, and the peak height (Willems 2009). The independence criteria are merged from a number of concepts. Firstly, subsequent rainfall intensity peaks can be considered largely independent if the inter-event time exceeds the recession time and if the lowest inter-event intensity drops below a specific low intensity level. Secondly, in line with statistical extreme value analysis, the POTs are selected as the highest events in periods where the hydro-climatic variable up crosses a selected threshold. The threshold component of the method suggested by Willems (2009) to extract independent extreme events makes use of the paper by Lang *et al.* (1999) which reviews tests and methods useful for modelling the process of POT values, the choice of the threshold level, the verification of the independence of the values and the stationarity of the process.

**GPD and calculation of the quantile for a given return period**

Whereas the extreme events from AMS follow the generalised extreme value distribution of Jenkinson (1955), it has

been shown in some asymptotic sense that for independent rainfall intensity extremes as used in this study, the conditional distribution of these extremes follows the GPD of Pickands (1975). This is confirmed in Figure 2 in a visual way using the L-moment diagram (Hosking & Wallis 1997) where the GPD passes through the average of the scatter points of the plot. Other distributions as well are used for comparison including the Gamma or Pearson III distribution (PE3), generalised normal distribution, and generalised logistic distribution (Hosking 1991).

Considering  $x_t$ ,  $\alpha$  and  $\gamma$  as threshold (analogous to location,  $\xi$ ), scale and shape parameters respectively, the cumulative distribution function  $G(x)$  of the GPD is given by

$$G(x) = 1 - \left\{ 1 + \gamma \frac{(x - x_t)}{\alpha} \right\}^{-\frac{1}{\gamma}}, \text{ for } \gamma \neq 0 \tag{2}$$

$$G(x) = 1 - \exp\left\{ -\left(\frac{x - x_t}{\alpha}\right) \right\}, \text{ for } \gamma = 0 \tag{3}$$

This distribution is valid for values of  $x$  above the threshold  $x_t$ . Based on the shape parameter  $\gamma$  or ( $k = -\gamma$ ), also called the extreme value index, three GPD classes can be identified including normal ( $\gamma = k = 0$ ), heavy ( $\gamma > 0$  or  $k < 0$ ) and light ( $\gamma < 0$  or  $k > 0$ ) tails. In the case  $\gamma < 0$ , the GPD has a final right-end point, whereas in the cases  $\gamma = 0$  or  $\gamma > 0$  the upper GPD tail goes up to infinite values. The GPD for  $\gamma = 0$  equals the exponential distribution.

Considering  $t$  as the number of observed rainfall intensity events above  $x_t$ ,  $w$  as the sample size of the POT events, and  $R_T$  as the rainfall intensity corresponding to a

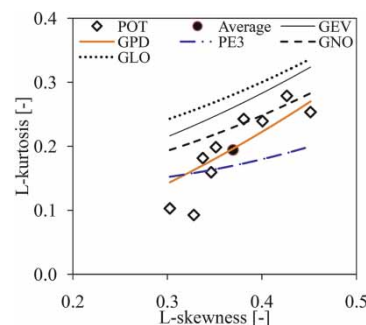


Figure 2 | L-moment ratio diagram for daily observed rainfall POT intensities.

given  $T$  or  $T$ -year quantile, the relationship between the  $R_T$  and  $T$  based on calibrated GPD is given by (Rosbjerg 1985)

$$T[\text{years}] = \left(\frac{n}{f}\right) \frac{1}{1 - G(R_T)} \quad (4)$$

where  $n$  is the data record length in years, and  $f = t$  (for WLR method) and  $f = w$  (for MOM, Lmom and ML methods).

Considering  $i$  as the rank of the POT events ( $i = 1$  for the highest), the empirical exceedance probability or  $T$  is given by

$$T[\text{years}] = \frac{n}{i} \quad (5)$$

### Estimation of GPD parameters

Commonly used parameter estimation methods for calibrating the GPD to extreme hydrometeorological events include the MOM, Lmom and ML methods. The ML method (Fisher 1922) holds that the best value of a parameter of a probability distribution should be that value which maximises the likelihood or joint probability of occurrence of the observed sample. The ML method is mainly used because of its desirable mathematical and optimality properties. However, ML estimates can be heavily biased for small samples. The mathematics involved in its usage is also often non-trivial. The MOM approach (Pearson 1902a, b) follows the rule that the estimators of the (conventional) population moments should be equal to the sample moments. The main advantage of the MOM approach is its undeniable simplicity when moments are available. However, it lacks the desirable optimality properties like those of the ML method. The Lmom are certain linear combinations of probability weighted moments (PWMs) that have simple interpretations as measures of the location, dispersion and shape of the data sample (such as is the case for conventional moments). The main advantage of Lmom is that they suffer less from the effect of sample variability compared with the conventional product moment estimators (Hosking 1990; Stedinger *et al.* 1993). The MOM, ML and Lmom estimations of GPD parameters were reviewed by Hosking & Wallis (1987). These

authors showed that ML estimators do not exist for shape parameter less than  $-1$ . They furthermore proposed PWMs to estimate the GPD parameters to improve the asymptotic variance of those based on the MOM approach.

The comparison of these parameter estimation methods for calibrating GPDs was undertaken in a number of studies for different regions of the world (see e.g. Hosking 1990; Castillo & Hadi 1997; Öztekin 2005; Deidda & Puliga 2009). For continuous samples, Deidda & Puliga (2009) found that the performance of the ML was better than those of other methods. However, they remarked that the determination of efficient estimators for rounded-off samples is still an open problem. Öztekin (2005) found for peak flows of 50 different rivers, most of them in Turkey, that the MOM approach was superior over other parameter estimation methods. According to Hosking (1990), L-moment estimators when compared with those of ML estimators usually show the Lmom method to be reasonably efficient. Castillo & Hadi (1997) remarked that the most serious problem with the MOM and PWM methods is that they can produce nonsensical estimates (estimates inconsistent with the observed data). The above summary shows that the suitability of parameter estimation methods may vary from region to region. It is important to note that apart from MOM, Lmom and ML, there were also other methods introduced for calibrating GPDs, for instance the elemental percentile method (EPM) of Castillo & Hadi (1997), and the principle of maximum entropy (POME) proposed by Singh & Guo (1995). The POME approach was formulated by Jaynes (1957a, 1957b, 1982). Castillo & Hadi (1997) showed that EPM performs well compared to existing methods. By comparing with the existing methods, the parameter estimates yielded by the POME were shown by Singh & Guo (1995) to be either superior or comparable for high skewness. The WLR, which is based on the quantile incremental properties of the GPDs (Beirlant *et al.* 1996), is another method for calibrating GPDs to extreme events. It has been used in a number of studies including Onyutha & Willems (2014), Csörgö *et al.* (1985), Kratz & Resnick (1996), Schultze & Steinebach (1996), and Willems *et al.* (2007). The main advantage of WLR is the selection in a visual way in quantile plots, of the most optimal threshold, the extreme value index and the corresponding distribution class of the EVD.

### Normal versus heavy tail

When the aim of the GPD calibration is to extrapolate to extreme quantile estimates (e.g. for water engineering design or planning applications), the shape of the GPD's tail is of primary importance. Major attention in this respect has to go to the sign of the extreme value index  $\gamma$ . According to [Cai et al. \(2013\)](#), empirical literature has documented that random variables investigated in meteorology and environmental science exhibit generally values of  $\gamma$  around zero (normal tail GPD). In analysis of hydrometeorological variables, the tendency not to make assumptions or fix the distribution class as either normal, heavy-tailed or with finite endpoint is necessitated by estimators of  $\gamma$  without having prior knowledge of its sign. This presents the possibility of systematic over/under-estimation in the tail of the EVD. However, not to lose the said generality so soon in this study,  $\gamma$  is computed using MOM, Lmom, ML and WLR methods and the statistical significance of its difference from zero is investigated. This was done based on the calculation of 95% confidence interval (CI) on the  $\gamma$  estimate. This evaluation of the sign of parameter  $\gamma$  was twofold; firstly for each of the rainfall series individually and secondly for all the rainfall series collectively. The idea here is that should  $\gamma$  have a value that is not significantly different from zero across the study area, the exponential distribution as a special case of the GPD with  $\gamma = 0$  can be adopted for the statistical modelling of the rainfall extremes.

The MOM, Lmom, ML, and WLR methods provide estimators for the different GPD parameters, including the shape parameter  $\gamma$  as follows:

- for MOM approach, as reviewed by [Hosking & Wallis \(1987\)](#)

$$Z = \frac{2(1 + \gamma)(1 - 2\gamma)^{0.5}}{1 - 3\gamma} \tag{6}$$

where Z is the sample skewness;

- for Lmom method ([Hosking 1990](#))

$$\gamma = \frac{3\tau_3 - 1}{1 + \tau_3} \tag{7}$$

where  $\tau_3$  is the L-skewness.

For the ML method, no estimator can be obtained for the location  $\xi$  because the likelihood function is unbounded with respect to  $\xi$  as shown by [Singh & Guo \(1995\)](#). These authors applied the constraint  $\xi = x_1$  which is the lowest POT sample value, since for this constraint the likelihood function is maximum. For that constraint, the ML estimator of  $\gamma$  (or  $k = -\gamma$ ) can be computed together with the GPD scale parameter  $\alpha$  using ([Singh & Guo 1995](#))

$$\sum_{i=1}^w \frac{\alpha^{-1} \times (x_i - \xi)}{1 + \alpha^{-1} \times \gamma(x_i - \xi)} = \frac{w}{1 + \gamma} \tag{8}$$

$$\sum_{i=1}^w \ln \left\{ 1 + \frac{\gamma(x_i - \xi)}{\alpha} \right\} = w\gamma \tag{9}$$

For the WLR method, the  $\gamma$  value can be computed as the inverse slope of the empirical POT values in the Pareto quantile plot when  $\gamma$  is positive ([Beirlant et al. 1996](#); [Willems et al. 2007](#)). In this plot the logarithmically transformed empirical POT values are plotted versus minus exceedance probability, i.e.  $-\ln(1 - G(x))$  in abscissa, and  $\ln(x)$  in ordinate. The slope of the estimate for  $\gamma > 0$  can be obtained by least square weighted linear regression using Hill weights ([Hill 1975](#)):

$$\gamma_t = \frac{1}{t - 1} \left\{ \sum_{i=1}^{t-1} \ln(x_i) \right\} - \ln(x_t) \tag{10}$$

[Beirlant et al. \(1996\)](#) developed another approach valid for any sign of  $\gamma$  based on the so-called generalised quantile plot

$$\gamma_t = \frac{\sum_{i=1}^{t-1} \left\{ -\ln\left(\frac{i}{t}\right) \right\}^{-1} \ln\left(\frac{i}{t}\right) \{ \ln(x_t) - \ln(x_i) \}}{\sum_{i=1}^{t-1} \left\{ -\ln\left(\frac{i}{t}\right) \right\}^{-1} \ln^2\left(\frac{i}{t}\right)} \tag{11}$$

To compute the GPD parameters for  $\gamma = 0$ , hence the exponential distribution parameters  $\alpha$  and  $\xi$ , Equations (12)–(15) apply:

- for MOM approach

$$\alpha = \sigma \tag{12}$$

$$\xi = \mu - \alpha \tag{13}$$

- for the Lmom

$$\alpha = 2\lambda_2 \tag{14}$$

$$\xi = \lambda_1 - \alpha \tag{15}$$

where  $\lambda_1$  is the first L-moment,  $\lambda_2$  the second L-moment,  $\mu$  the sample mean, and  $\sigma$  the standard deviation.

For the ML method as explained above, the  $\xi$  was constrained to  $x_1$ . For a known  $\xi$ , the estimate of  $\alpha$  using the ML method is identical to that of Lmom or MOM. Therefore, in this study,  $\alpha$  of the ML method was determined using Equation (12).

For the WLR method, the estimate of  $\alpha$  is obtained by a least squares weighted linear regression in the exponential plot, defined as the plot similar to the Pareto quantile plot but with the  $\ln(x)$  taken as  $x$  in ordinate. This regression estimate constrained to  $t$  observations is given by (Csörgő *et al.* 1985; Beirlant *et al.* 1996)

$$\alpha_t = \frac{1}{t-1} \left( \sum_{i=1}^{t-1} x_i \right) - x_t \tag{16}$$

Equation (16) is again based on the weighting factors proposed by Hill (1975). The optimal threshold  $x_t$  can be defined as the POT value above which the mean squared error (MSE) of the regression is minimal. For a selected  $t$

value, the MSE of the weighted linear regression in the exponential quantile plot is

$$MSE_t = \frac{1}{t-1} \left\{ \sum_{i=1}^{t-1} \left\{ -\ln\left(\frac{i}{t}\right) \right\}^{-1} \left( x_i - x_t - \alpha_t \ln\left(\frac{i}{t}\right) \right)^2 \right\} \tag{17}$$

For the exponential distribution of Equation (3), Equation (4) transfers to a linear relationship between  $R_T$  and  $\log(T)$  as in Equation (18)

$$R_T = \alpha(\log(T) - \log(T_0)) + R_{T_0} \tag{18}$$

where  $R_{T_0}$  is the rainfall intensity at return period  $T_0$ , and  $T_0$  is equal or higher than the empirical return period  $n/t$  of the threshold  $x_t$ .

Figure 3(a) shows an example of such a calibrated normal tailed GPD as a linear regression line in the exponential quantile plot. An example of the calculated slope and MSE for selected  $t$  of the same station is shown in Figure 3(b). Because of high fluctuations in the slope of the quantile plot for high thresholds (see Figure 3(b)), stemming from randomness in the available POT values, the slope estimates for these high thresholds have high statistical uncertainty. Instead for very low thresholds, the slope estimates might result in pronounced bias because according to Pickands (1975) the slope converges asymptotically for higher thresholds. The selection of optimal threshold values  $x_t$  above which the EVD is calibrated is, in this study, ensured to be at a point above which the MSE of

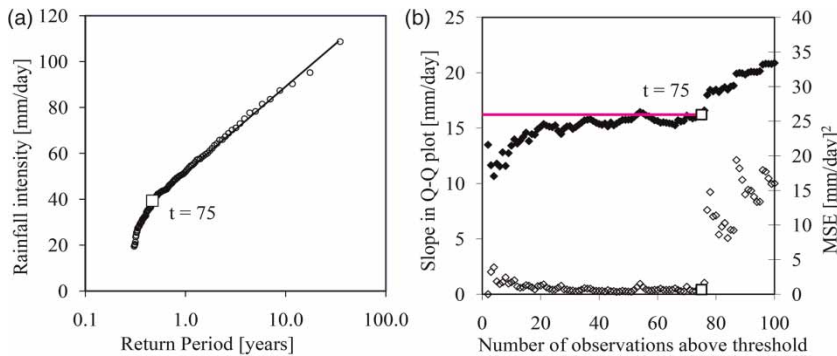


Figure 3 | (a) Observations (O) and calibrated exponential distribution above selected optimal threshold (□) in the exponential quantile plot for daily observed rainfall intensities at station 2. (b) Slope (◆) and MSE (◇) for selected  $t$ .

the linear regression is minimal, i.e., within nearly horizontal sections in the plot of the slope versus the number of observations above the threshold. In cases where it is possible for the MSE to reach local minima at different threshold orders within the range where the optimal threshold is situated, the EVD is calibrated for the threshold at each of the local minima for visual aid of selecting the most suitable one. For the example of Figure 3, the optimal threshold was determined as the POT value with threshold rank  $t = 75$  (the 75th highest rainfall intensity value). A linear tail behaviour in the exponential quantile plot can be observed towards the higher rainfall intensity values.

This procedure for selection of the type of distribution and optimal  $x_t$  based on Beirlant *et al.* (1996) can be found in a systematic methodology by Willems *et al.* (2007).

## Evaluation of calibrated EVD

### Goodness-of-fit statistics

The goodness-of-fit between the empirical and theoretical quantiles was evaluated both statistically and graphically. This was for all the parameter estimation methods done over the same quantile region, i.e. for POT events above  $x_t$ . Statistically, the Kolmogorov–Smirnov (*K-S*) (Kolmogorov 1933; Smirnov 1936) test (Equation (19)) at 5% level of significance was conducted on the null hypothesis that the data follow the GPD. The *K-S* statistic  $K_{\text{stat}}$  was compared with the critical values ( $K_{\text{CV}}$ ) of the said test. If  $K_{\text{stat}}$  was found to be less than  $K_{\text{CV}}$ , the null hypothesis was accepted; otherwise it was rejected. In this study, the values of  $K_{\text{stat}}$  for which the null hypothesis was accepted were also used to judge on the performance of the parameter estimation methods in the calibration of the EVD to empirical quantiles. The lower the value of  $K_{\text{stat}}$  for which the null hypothesis is accepted, the higher the acceptability of the parameter estimation method used to calibrate the EVD.

For all the parameter estimation methods,  $K_{\text{stat}}$  is computed as

$$K_{\text{stat}} = \max_{i=1}^t \left\{ G(R_{e,i}) - G(R_{g,i}) \right\} \quad (19)$$

where  $G(R_{g,i})$  and  $G(R_{e,i})$  are the cumulative distributions of theoretical (GPD based) and empirical quantiles, respectively.

It is worth noting that the *K-S* test has well known limitations. *K-S* only applies to continuous distributions and tends to be more sensitive near the centre of the distribution than in the tails. Therefore, this test was complimented with the computation and comparison of other statistics such as the standard error of the estimate *SEE* [mm/day], also referred to as the root MSE; the *bias* [%], which represents the mean error; and the probability plot correlation coefficient (*PPCC*) [-]. The *PPCC* was originally developed for normality testing (Filliben 1975; Looney & Gulledge 1985) and modified by Vogel (1986). It is a measure of correlation between the theoretical ( $R_{g,i}$ ) and empirical ( $R_{e,i}$ ) quantiles

$$SEE \text{ [mm/day]} = \left( \frac{1}{t} \sum_{i=1}^t (R_{e,i} - R_{g,i})^2 \right)^{0.5} \quad (20)$$

$$Bias \text{ [%]} = \frac{1}{t} \sum_{i=1}^t \left( \frac{R_{e,i} - R_{g,i}}{R_{g,i}} \times 100 \right) \quad (21)$$

$$PPCC \text{ [-]} = \frac{\sum_{i=1}^t (R_{e,i} - \bar{R}_{e,i})(R_{g,i} - \bar{R}_{g,i})}{\left( \sum_{i=1}^t (R_{e,i} - \bar{R}_{e,i})^2 \sum_{i=1}^t (R_{g,i} - \bar{R}_{g,i})^2 \right)^{0.5}} \quad (22)$$

where  $\bar{R}_{g,i}$  and  $\bar{R}_{e,i}$  are the mean values of  $R_{g,i}$  and  $R_{e,i}$  respectively.

Graphically the statistics  $K_{\text{stat}}$ , *SEE*, *bias*, and *PPCC* for the EVD calibrated using one parameter estimation method was plotted against those of others to visualise the differences. Performance of the parameter estimation methods was furthermore compared through the percentage of times: (1) the null hypothesis in the *K-S* test was accepted; (2) the lowest values of  $K_{\text{stat}}$ , *SEE*, and *bias* were realised; and (3) the highest value of *PPCC* was obtained.

### Tail behaviour

The calibrated GPDs moreover were evaluated graphically in quantile plots, i.e. exponential quantile plot for  $\gamma = 0$ , for checking the shape and accuracy of the slope of the

GPD. The key parameter defining the slope in the exponential quantile plot  $\alpha$ , obtained using the various parameter estimation methods was graphically compared as well. This was done by plotting values of  $\alpha$  obtained by one parameter estimation method against those of others.

### Error analysis in the quantile estimation

The distribution of the *bias* in the extreme quantiles and 95% CI on the *bias* were computed for the various parameter estimation methods assuming that they are random and follow a normal distribution. These were in turn used to determine if the values of *bias* were statistically significant or not. We also constructed and compared 95% CIs on the rainfall intensity quantile itself based on the EVD parameters calculated by the different parameter estimation methods. By combining the delete-one Jackknife method with nonparametric percentile bootstrapping, the 95% CI was constructed as follows:

- (1) From the original full series of POT values of size  $w$ ,  $i$ th event is deleted or left out to obtain  $(w - 1)$  data points.
- (2) Using the  $(w - 1)$  data points, parameters of the EVD are estimated.
- (3) Steps 1 and 2 are repeated  $w$  times to obtain  $w$  sets of estimated parameters.
- (4) Values of each EVD parameter are ranked from the highest to the lowest.
- (5) The limits (upper, lower) of 95% CI for the theoretical quantiles are constructed using the  $([0.025w]$ th,  $[0.975w]$ th) sets of estimated EVD parameter values.

An example of EVDs calibrated to daily empirical quantiles using parameters estimated as described earlier is shown in Figures 4(a). Figure 4(b) shows, for the WLR method, the CIs on the theoretical quantiles.

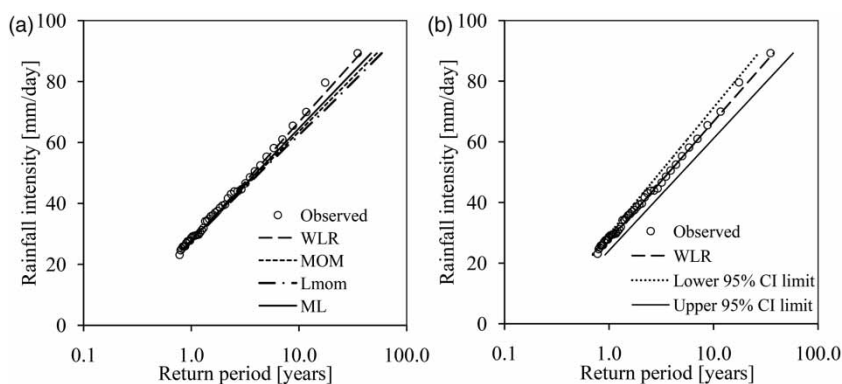
After defining CIs for the entire range of empirical return periods, 10-year rainfall quantile ( $R_{10}$ ) was considered for comparison of CIs between the various parameter estimation methods. A return period of 10 years was selected because it is commonly applied in planning, design and operation of water resource projects. The difference (Diff) was computed between the upper (*UL*) and lower (*LL*) limits of the 95% CI expressed as a percentage of  $R_{10}$

$$\text{Diff} [\%] = \frac{UL - LL}{R_{10}} \times 100 \quad (23)$$

This Diff is indicative of the influence of the parameter estimation uncertainty on the rainfall quantile estimate. It is also worth noting that it can be possible for an empirical quantile to be located outside the CI constructed on the calibrated EVD. A small (large) Diff but large (small) *bias* indicates rigidity (flexibility) of the region over which parameters for the calibrated EVD were estimated. The best parameter estimation method would have both its Diff and *bias* small.

### Uncertainty in empirical extreme quantiles

After selecting the best parameter estimation method using all the datasets (combined observed time series and those



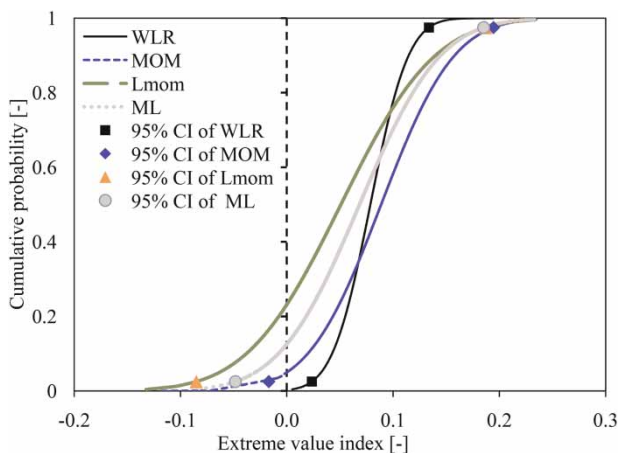
**Figure 4** | (a) Exponential distribution calibrated following the different methods. (b) CIs constructed on the WLR-based theoretical 7 quantiles. The daily observed rainfall intensities are at station 1.

from the GCMs), uncertainty in the quantiles estimated only using observed data (excluding the GCMs, in this case) at the different selected stations is assessed. This is done graphically by comparing empirical  $T$  versus EVD based  $T$ , and  $K_{\text{stat}}$  versus  $PPCC$ ; but also by checking the distribution of the residuals on the modelled  $T$ , and comparing the CIs on the quantiles. The difference between the GPD parameters at the different stations is also investigated.

## RESULTS AND DISCUSSION

### Normal versus heavy tail

Figure 5 shows collective distribution of  $\gamma$  computed from all the series used in this study. It can be seen that the entire distribution of  $\gamma$  for WLR falls to the right of  $\gamma = 0$ . This follows from Equation (11), which is valid for  $\gamma > 0$  only. In other words, WLR was here used to check how close the  $\gamma$  distribution values are to those by other calibration methods, and from zero. The differences between zero and the centre of the distribution of  $\gamma$  obtained by WLR, MOM, Lmom and ML methods are 0.072, 0.089, 0.051 and 0.068 respectively. It is also noted from Figure 5 that for MOM, Lmom and ML methods, a higher fraction of the  $\gamma$  distribution values falls on the positive than the negative side. This implies that the possibility of the rainfall series exhibiting light tail for the GPD is far lower than that for heavy or normal tails.



**Figure 5** | Cumulative probability distribution of the extreme value index  $\gamma$  obtained from all the rainfall series (both station-based observed data and those from the GCMs) for the various parameter estimation methods.

This was also expected, given that the light tail case does not occur that often in hydrometeorology. Because rainfall does not show upper limits, normal and heavy tail distributions are most often found (see e.g. Buishand 1989; Harremoës & Mikkelsen 1995; Willems *et al.* 2007).

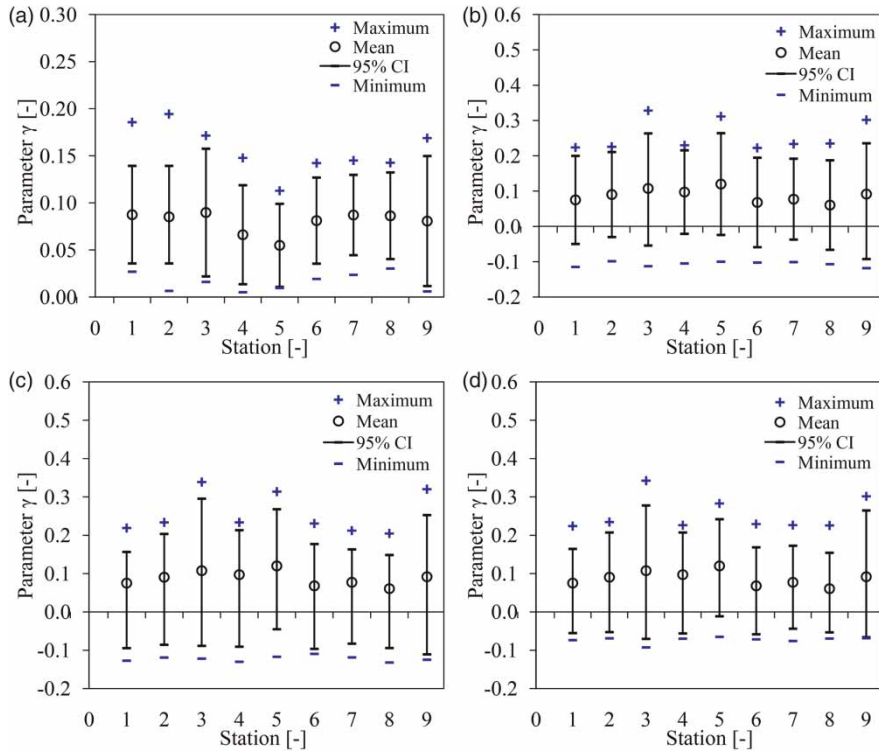
Figure 6 shows the variation in CI for the estimated  $\gamma$  at selected rainfall stations using the different parameter estimation methods. No significant difference is seen between the CIs of the various parameter estimation methods.

It is shown from Figures 5 and 6 that except for WLR, which was conditioned for  $\gamma > 0$ , the  $\gamma = 0$  case falls within the 95% CI. This means that the GPD tails do not deviate much from the normal case ( $\gamma = 0$ ). It implies that the GPD can as well be obtainable by approximating  $\gamma$  to zero as also adopted for this study for the EVD results discussed next.

### Evaluation of calibrated EVDs

#### Goodness-of-fit statistics

Figure 7 shows the differences between the parameter estimation methods in terms of *bias* and *SEE* in extreme daily rainfall quantiles. There are more points about the origin of the plots than for higher values of the *bias*. In the positive *bias* region in the plots, there are systematic differences between the *bias* obtained by Lmom and MOM or ML methods. However, for the same region in the plots, the *bias* on the WLR method is about the origin (or nearly zero in value). The question is why the *bias* of WLR is smaller than that of other methods. The best approach to such influences would be to analyse the theoretical versus empirical distributions that control the different regions of the plot. Investigation of the tail behaviour in the quantile plots (e.g. Figure 8) shows that the MOM, ML and Lmom methods, although they lead to parameter estimates that give rise to good overall distribution match, may show biased tail behaviour (systematically increasing bias towards the higher extremes). The advantage of the WLR parameter estimation can also be seen in Figures 7(g)–7(i) by its lower values of *SEE*. This reward for WLR can be ascribed to the focus of the method on the EVD tail behaviour by means of the quantile plot, including selection of the most optimal threshold to calibrate the EVD. However, for 21.87% of the cases, the

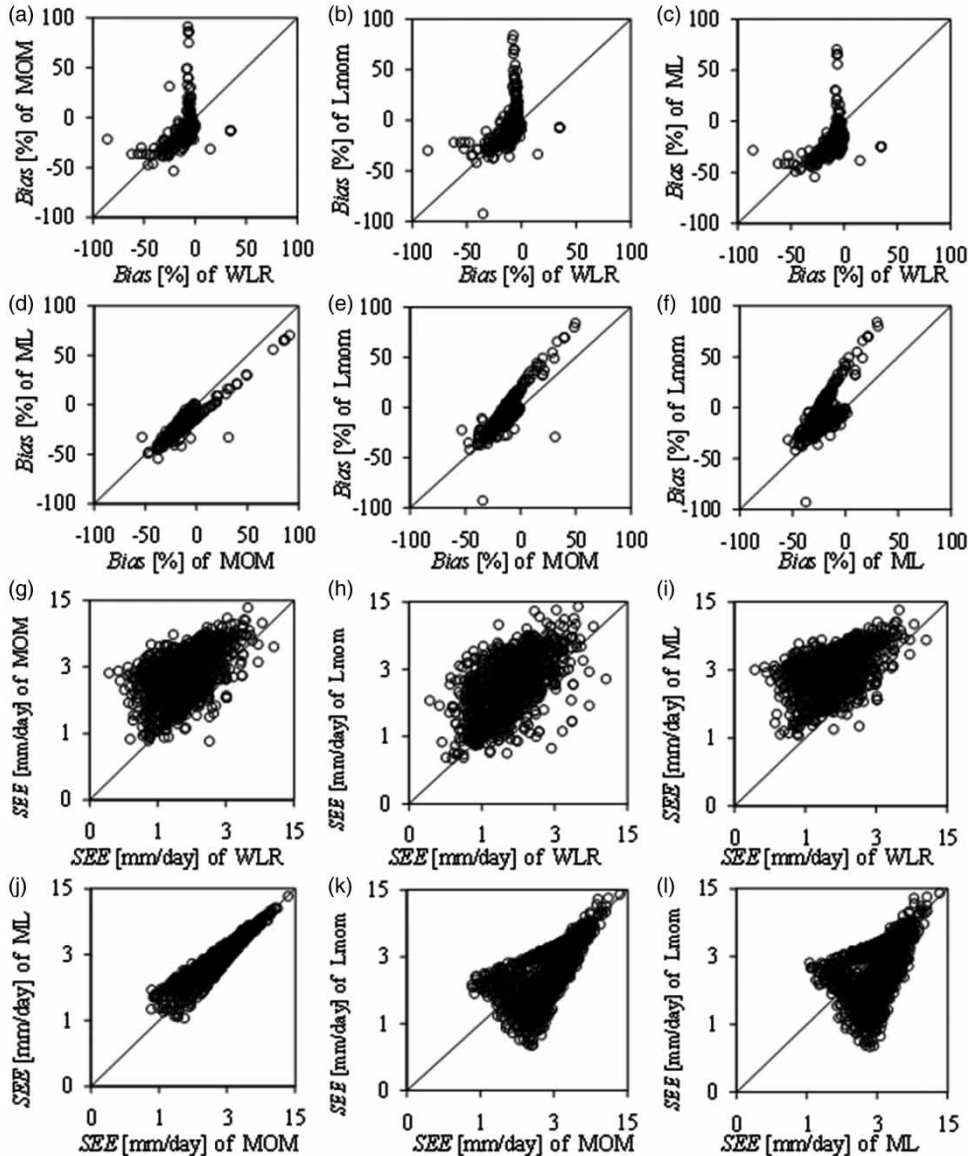


**Figure 6** | Maximum, mean, minimum and 95% CI of the  $\gamma$  distribution obtained from all the rainfall series (both station-based observed data and those from the GCMs) for different stations calculated using: (a) WLR, (b) MOM, (c) Lmom, and (d) ML.

value of *SEE* obtained for the WLR method was higher than those of the other methods. From Figures 7(j)–7(l), it is shown that the differences in proportion of the scatter points of *SEE* above and below the bisector line are not significant. This generally indicates the closeness of the EVD parameters estimated by the MOM, ML and Lmom methods.

Figure 9 shows the probability density of the *SEE* and *bias* for the different parameter estimation methods. Figure 9(a) shows, as already concluded from Figure 7, that the values of *SEE* obtained by WLR are smaller than those of other methods. Figure 9(b) shows the distribution of *bias* in a pattern similar to that of *SEE* in Figure 9(a). It can be noted that the zero value is within these CIs for WLR, MOM and Lmom methods; thus their biases are not significant. However, the *bias* for ML is significant. Despite the acceptability of the *bias* in the EVD calibrated using the MOM and Lmom, their uncertainties are higher than those of ML and WLR. These uncertainties are reflected in the difference between the upper and lower limits of the CIs on *bias* as evidenced from Figure 9(b).

The *PPCC* and  $K_{\text{stat}}$  results are shown in Figure 10. It is generally shown that values of *PPCC* are closer to each other in the lower region of the scatter plots. In the opposite manner, the values of  $K_{\text{stat}}$  are closer between the different parameter estimation methods in the higher region of the plots. This follows, for some samples, the slight deviation of  $\gamma$  from being zero. Although it was seen before that the GPDs calibrated to the rainfall extremes of the LVB can be approximated by the exponential distribution,  $\gamma$  values of some samples deviate from zero. As this deviation increases (though not significantly), the *PPCC* reduces and  $K_{\text{stat}}$  increases. The closeness to each other of lower *PPCC* but higher  $K_{\text{stat}}$  values for the different parameter estimation methods is due to more or less similar deviations of the calibrated exponential distributions in capturing empirical quantiles whose actual  $\gamma$  values deviate from zero. On average the values of  $K_{\text{stat}}$  for which the null hypothesis is accepted for MOM, ML, Lmom and WLR methods are 0.145, 0.146, 0.151 and 0.125. These low values of  $K_{\text{stat}}$  indicate acceptability of the exponential distribution as a special



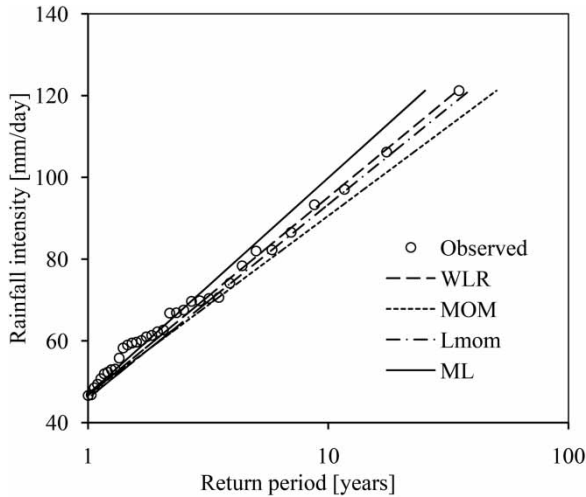
**Figure 7** | Comparison of bias in (a) to (f) and SEE in (g) to (l) obtained for extreme daily rainfall quantiles (from combined series of both station-based observed data and those from the GCMs) by the various parameter estimation methods. All axes of charts (g) to (l) are to log base 10 for clarity.

case of the GPD calibrated to rainfall extremes in the LVB. It also shows the acceptability of the different parameter estimation methods in calibrating the exponential distribution to these extremes.

### Tail behaviour

Figure 11 shows the differences in the values of parameter  $\alpha$  obtained using the various parameter estimation methods. It

is shown that the values of parameter  $\alpha$  obtained by MOM, Lmom and ML are closer to each other than with those of WLR. This can be explained by the difference in the region of POT events for which the parameter  $\alpha$  is calculated. The WLR method constrains the computation of parameter  $\alpha$  to POT events above  $x_t$ , whereas the computation of parameter  $\alpha$  by MOM, Lmom and ML is unconstrained. With increase in parameter  $\alpha$ , some systematic deviations (higher  $\alpha$ ) between the values estimated by



**Figure 8** | Systematic deviations of calibrated exponential distribution from empirical quantiles of station 4.

Lmom and those by MOM or ML are seen in Figures 11(d) and 11(e). This is due to difference in sensitivity of the mentioned methods to variation or dispersion from the mean of the independent POT events. For the MOM, Lmom, ML methods versus WLR, no systematic differences in  $\alpha$  estimates are noted.

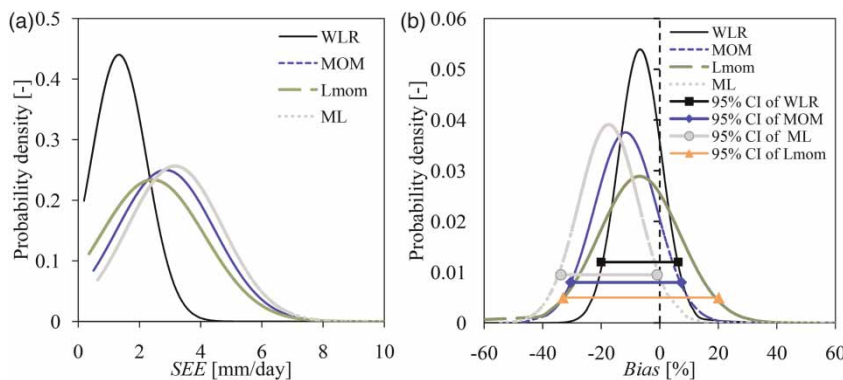
**CIs on estimated quantiles**

Figure 12 shows comparative scatter plots for Diff. No clear systematic differences between the methods are seen. Using CIs constructed on EVDs for which the null hypothesis of the  $K$ -S test is accepted, the overall values (minimum, average, maximum) of Diff for WLR, MOM,

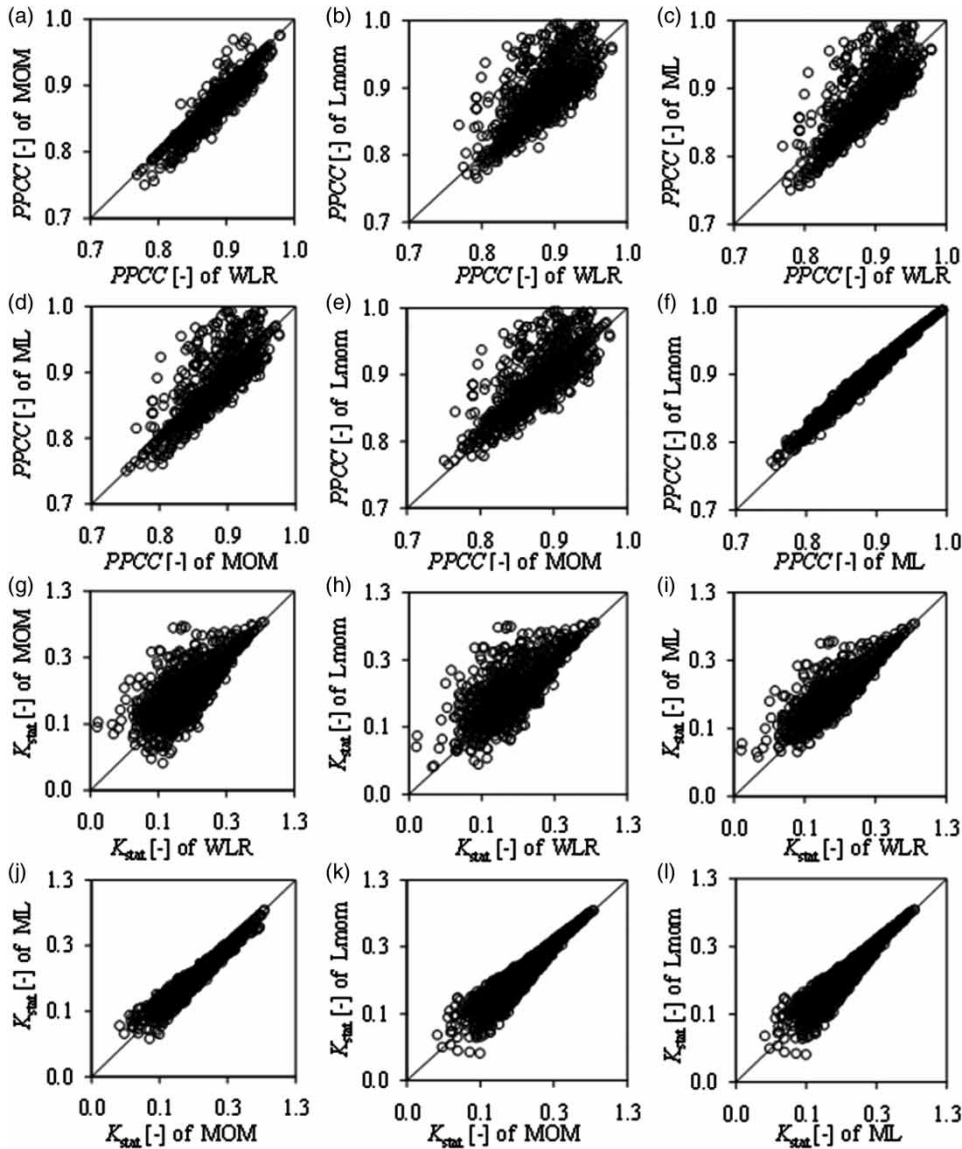
Lmom and ML are (0.20%, 40.24%, 79.75%), (0.02%, 43.10%, 94.45%), (0.05%, 35.79%, 95.27%) and (0.47%, 47.20%, 92.63%) respectively.

**Summary on evaluation of the calibrated EVDs**

Figure 13 shows a summary of the performance of the various EVD parameter calibration methods. It is shown that there is no significant difference between the methods in the number of times the null hypothesis of the  $K$ -S test on the calibrated EVD is accepted. The  $K$ -S null hypothesis is rejected in 19.01, 24.39, 26.91 and 24.37% of the cases for WLR, MOM, Lmom and ML methods respectively. The percentages of the cases in which  $K_{stat}$  ( $PPCC$ ) is the lowest (highest) for WLR, MOM, Lmom and ML are 50.70% (58.10%), 17.40% (1.99%), 13.98% (37.92%) and 17.92% (1.99%) respectively. With respect to Diff, the Lmom approach presents the smallest value followed by WLR, MOM and finally ML. Based on these results, it is suggested that the most to the least robust method for GPD parameter estimation to rainfall extremes in the LVB with respect to  $SEE$ ,  $K_{stat}$ ,  $PPCC$  and  $bias$  are WLR, Lmom, MOM and ML. These results are consistent with those of Mkhandi *et al.* (1996) in a study on the southern part of the LVB that showed that Lmom performs better than MOM and ML for flood quantile estimation. However, instead of the conventional calibration of EVDs to the entire POT series, it is expected that better estimations of quantiles using Lmom, MOM and ML methods might be obtained with focus of the methods on the EVD tail behaviour or the



**Figure 9** | Probability distribution of the  $SEE$  in (a) and  $bias$  in (b) on estimated extreme daily rainfall quantiles (from combined series of both station-based observed data and those from the GCMs) for the various parameter estimation methods.



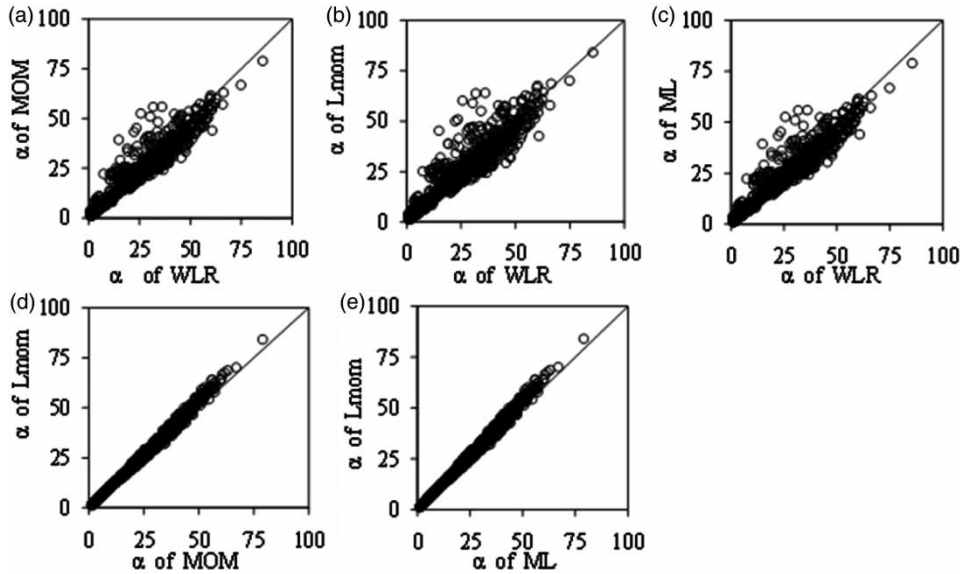
**Figure 10** | Comparison of  $PPCC$  in (a) to (f), and  $K_{stat}$  in (g) to (l) obtained for extreme daily rainfall quantiles (from combined series of both station-based observed data and those from the GCMs) by the various parameter estimation methods. All axes of charts (g) to (l) are to log base 10 for clarity.

region over which GPD (Equation (2) or (3)) is valid. This can be achieved when the most optimal thresholds as defined earlier are first selected in a quantile plot, as is the case for the WLR method.

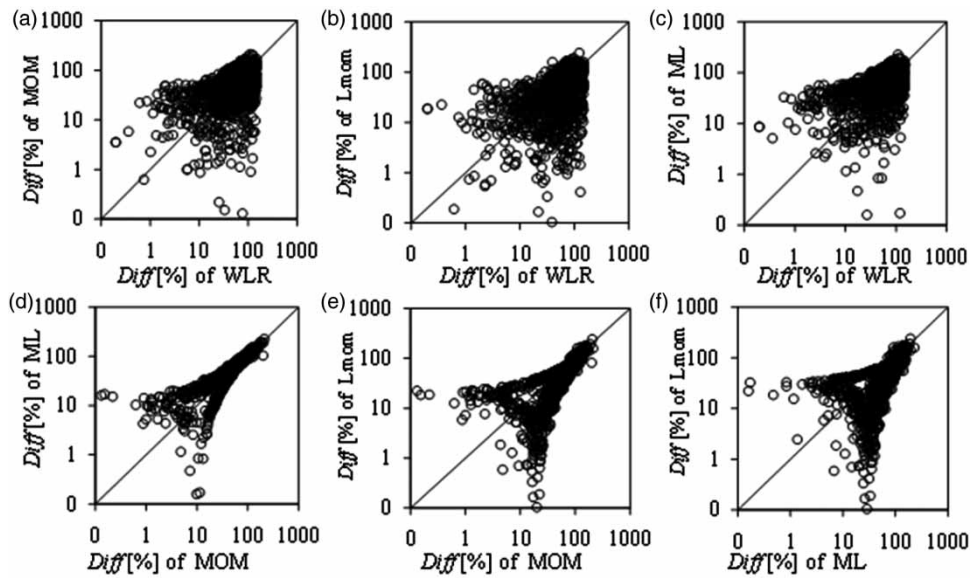
### Uncertainty in empirical extreme quantiles

Figure 14 shows, for station-based observed rainfall intensity quantiles, the difference in the GPD parameters across the study area. It is shown in Figure 14(a) that stations 3, 4

and 9 (found on the eastern half) of the LVB are characterised by higher values of  $x_t$  than those of 1, 2, and 6–8 (on the western half). This is because of the higher rainfall extreme intensities in the eastern half than in the western part of the study area. Figure 14(b) shows parameter  $\alpha$  of the different stations. No strong difference in the values of  $\alpha$  is seen among stations 1–2 and 4–8. Conversely, stations 3 and 9 are characterised by high values of  $\alpha$ . This might have to do with spatial difference in variability as is explained hereafter.



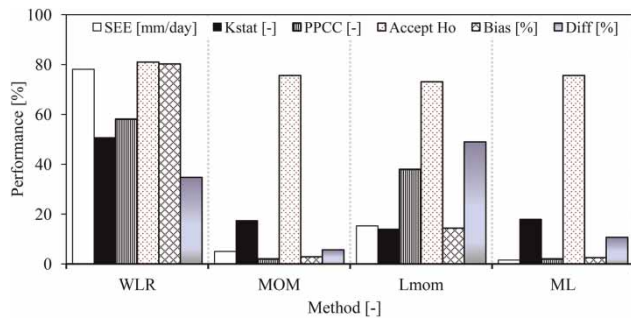
**Figure 11** | Comparison of parameter  $\alpha$  obtained by the different parameter estimation methods using combined series of both station-based data and those from the GCMs.



**Figure 12** | Comparison of Diff obtained by the various parameter estimation methods using combined series of both station-based observed data and those from the GCMs.

Figure 15 shows, for the WLR method, graphical representations of the uncertainty in modelled  $T_s$  and  $T$ -year events. Figure 15(a) shows comparison between the EVD based and empirical  $T$  values. The random deviations are due to data limitations. The systematic deviations of the data points from the bisector shows statistical modelling

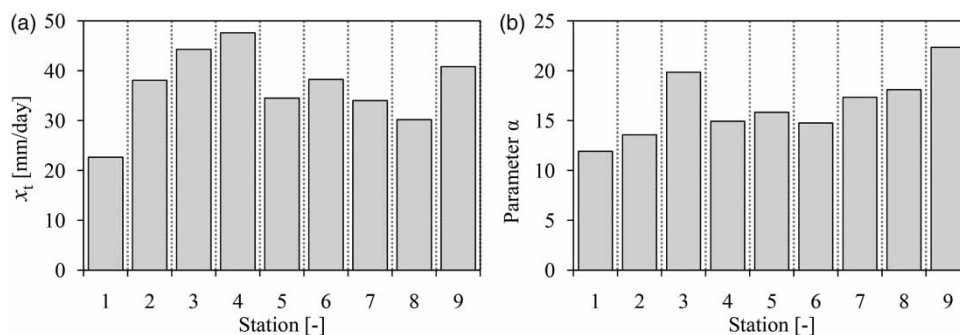
uncertainty on  $T$  and are proportional to the EVD parameter  $\alpha$ . Figure 15(b) shows distribution of residuals on modelled  $T$  for station 7. It can be seen that the distribution of these residuals follow well the assumed Gaussian distribution; this was found for modelled  $T$  at all the selected stations and confirmed by  $t$ -tests. Figure 15(c) plots the  $K_{\text{stat}}$  versus



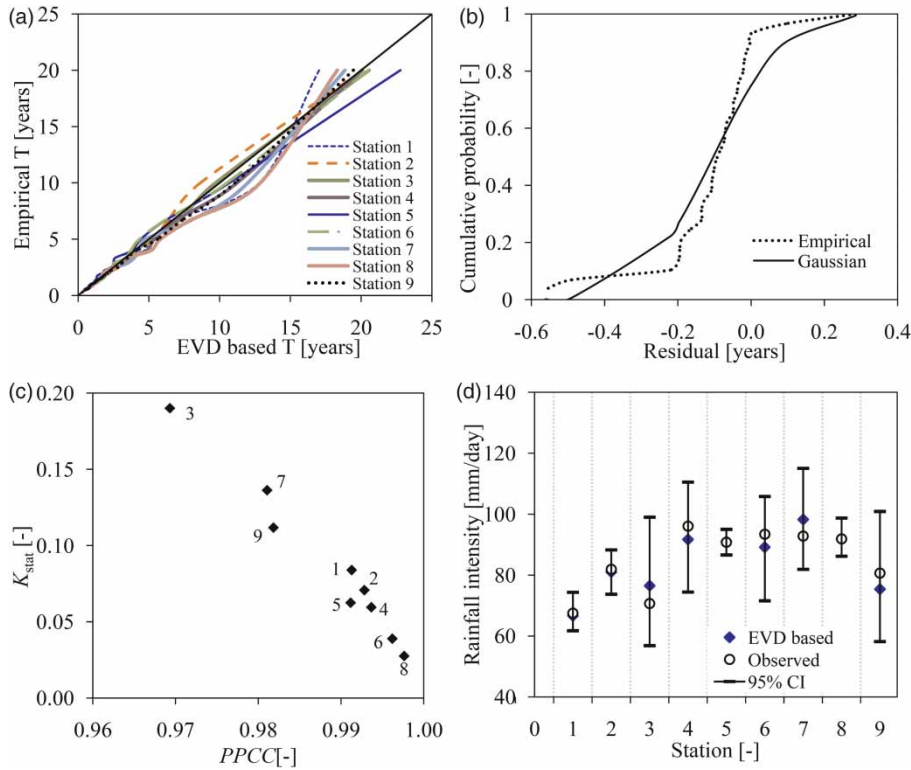
**Figure 13** | Performance of the different parameter estimation methods evaluated by the percentage of times the null hypothesis of the  $K$ - $S$  test is accepted or has the lowest values of  $SEE$ ,  $bias$ ,  $K_{stat}$  [-],  $Diff$  or the highest value of  $PPCC$ .

$PPCC$ . It is shown that  $PPCC$  values for all the stations are close to unity. It can also be noted that both statistics lead to consistent conclusions: the higher the  $K_{stat}$  the lower the  $PPCC$ . Figure 15(c) confirms the acceptability of the EVD calibrations to rainfall extremes at all the stations. Figure 15(d) shows 10-year rainfall intensity quantiles and their estimation uncertainty. The width (difference between upper and lower limits) of the 95% CIs expressed as a percentage of the empirical 10-year rainfall intensity quantile ranges from 9.25% (at station 5) up to 59.66% (at station 3). These uncertainties indicate the need for their consideration in decision making related to risk based water resources engineering and management. The shifts in the CI between different stations indicate spatial differences across the study area. The stations at which wider CIs are obtained (stations 3 and 9 in Figure 15(d)) present in general higher temporal variability in rainfall intensity as shown in Figure 16. The rainfall extreme intensities in the Nile basin indeed have strong spatial variation as seen from the regional extreme value analysis by Nyeko-Ogiramo *et al.* (2012).

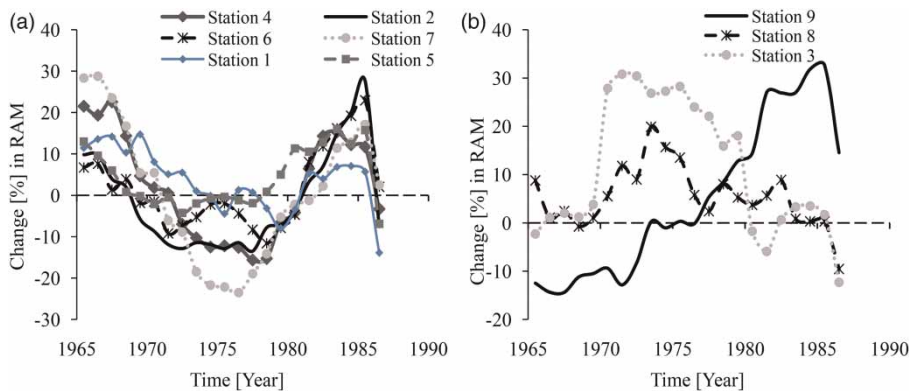
Figure 16 shows the temporal variability computed using series obtained by extracting the highest rainfall intensity in each year (RAM) of all the selected stations in Table 1. The RAM is deemed to be representative of the POT events. The oscillation pattern in Figure 16 was calculated and the significance tested using the quantile anomaly indicator on which the details can be found fully discussed in Ntegeka & Willems (2008), and Willems (2013). From Figure 16(a), it is shown that oscillation highs (OHs) and lows (OLs) of stations 1–2 and 4–7 are more or less similar, and hence showing the closeness in the CI for the estimated quantiles. The OHs in RAM of the mentioned stations occurred in the mid-1960s and the early 1980s. According to Kizza *et al.* (2009), there was a significant upward jump in the rainfall of the LVB in the 1960s. The authors also remarked that the amount of rainfall received over the LVB in the 1960s, early 1980s as well as the late 1990s was above average. It is shown from Figure 16(a) that the OL was in the period 1970–1980; this was significant at station 7. The OH of the early 1980s was significant at station 2. These results are consistent with the findings on the variability in rainfall extremes in the LVB by Nyeko-Ogiramo *et al.* (2013) and Mbungu *et al.* (2012). Whereas extreme rainfall in the LVB is influenced by changes that occur both in the Pacific and Atlantic Oceans (Nyeko-Ogiramo *et al.* 2013), the drivers for the variability in the annual or seasonal rainfall might be those that influence variations in the ‘short rains’ which generally occur from October to December (Kizza *et al.* 2009). The variability in the annual or seasonal rainfall in the equatorial region is ascribed to the El-Niño Southern



**Figure 14** | GPD parameters, (a)  $x_t$ , and (b)  $\alpha$ , estimated by the WLR method using observed rainfall intensity at the different stations.



**Figure 15** | (a) Empirical  $T$  versus EVD-based  $T$ , (b) distribution of residuals on  $T$  at station 3, (c)  $K_{sat}$  versus PPCC goodness-of-fit statistics on calibrated EVD; the labels on the data points refer to the stations of the study, (d) 10-year rainfall intensity quantiles and CIs.



**Figure 16** | Temporal variability in the RAM of LVB for the period 1960–1990 using a time slice of 10 years.

Oscillation (Nicholson & Entekhabi 1986; Ropelewski & Halpert 1987; Ogallo 1988; Camberlin 1995; Nicholson 1996; Nicholson & Kim 1997; Indeje *et al.* 2000; Philipps & McIntyre 2000; Schreck & Semazzi 2004). It is importantly seen in Figure 16(b) that there are three stations

whose variability patterns in RAM are different from those shown in Figure 16(a). Such differences may be due to the quality of the data being too low to obtain a clear variability pattern in rainfall extremes at some stations of the LVB.

## CONCLUSION

This paper assessed the most suitable class of the GPD for extreme rainfall intensity in the LVB. Investigation of the tails of the GPDs did not show any clear evidence against the extreme value index being zero. This finding agrees with the results of Onyutha & Willems (2013) which suggested that the exponential distribution as the normal tailed GPD is the most suited for hydrological extremes in most of the catchments of the LVB. However, more studies need to be carried out using longer and up-to-date time series of hydro-meteorological variables at many other stations within the study area before making a generalisation.

The exponential distribution was calibrated to rainfall extremes using a number of parameter estimation methods including the MOM, Lmom, ML and WLR based on weighted regression in quantile plots. The performance of each parameter estimation method was found acceptable based on the goodness-of-fit statistics, analysis of the tail behaviour in quantile plots, and quantification of the uncertainty in extreme quantiles. However, generally for the selected rainfall series of the LVB, the WLR technique was found to be more robust than others. The WLR method was followed by the methods Lmom, MOM and finally ML. Next to analysing longer records and more stations, it would be worth testing more parameter estimation methods such as EPM (Castillo & Hadi 1997), and POME (Singh & Guo 1995). It is also proposed that for Lmom, MOM and ML methods, the optimal threshold (i.e. the POT value above which the MSE on the calibrated EVD is minimal) should be selected as for the WLR method so as to minimise possible bias in capturing the tails of the extreme distributions.

Finally, using the obtained best parameter estimation method, uncertainty in the extreme intensity quantiles across the LVB was quantified using the Jackknife method and nonparametric percentile bootstrapping in a combined way. The width (difference between upper and lower limits) of the 95% CIs expressed as a percentage of the empirical 10-year rainfall intensity quantile ranges from 9.25% up to 59.66%. The assessed uncertainty in estimation of rainfall extreme quantiles will be useful for water resources managers of the LVB in accounting for this uncertainty in their planning, design and operation of water resource projects.

## ACKNOWLEDGEMENTS

The authors would like to thank the two anonymous reviewers for their insightful comments and suggestions to enhance the quality of the paper. The research was linked to the FRIEND/NILE project of UNESCO and the Flanders in Trust Fund, and was financially supported by an IRO PhD scholarship of KU Leuven.

## REFERENCES

- Ahrens, B. 2006 Distance in spatial interpolation of daily rain gauge data. *Hydrol. Earth Syst. Sci.* **10**, 197–208.
- Anyah, R. O., Semazzi, F. H. M. & Xie, L. 2006 Simulated physical mechanisms associated with climate variability over Lake Victoria basin in East Africa. *Mon. Weather Rev.* **134**, 3588–3609.
- Awange, J. L. & Ong'ang'a, O. 2006 *Lake Victoria: Ecology, Resources, Environment*. Springer Berlin, Germany.
- Beersma, J. J. & Buishand, T. A. 1999 A simple test for equality of monthly variances in climate time series. *J. Clim.* **12** (6), 1770–1779.
- Beirlant, J., Teugels, J. L. & Vynckier, P. 1996 *Practical analysis of extreme values*. Leuven University Press, Leuven, Belgium, 137 pp.
- Buishand, T. A. 1989 Statistics of extremes in climatology. *Stat. Neerl.* **43** (1), 1–30.
- Cai, J.-J., de Haan, L. & Zhou, C. 2013 Bias correction in extreme value statistics with index around zero. *Extremes* **16** (2), 173–201.
- Camberlin, P. 1995 June-September rainfall in North-Eastern Africa and atmospheric signals over the tropics: A zonal perspective. *Int. J. Climatol.* **15**, 773–783.
- Castillo, E. & Hadi, A. S. 1997 Fitting the generalized Pareto distribution to data. *J. Am. Stat. Assoc.* **92** (440), 1609–1620.
- Csörgö, S., Deheuvels, P. & Mason, D. 1985 Kernel estimators of the tail index of a distribution. *Ann. Stat.* **13** (3), 1050–1077.
- Deidda, R. & Puliga, M. 2009 Performances of some parameter estimators of the generalized Pareto distribution over rounded-off samples. *Phys. Chem. Earth* **34**, 626–634.
- Dirks, K. N., Hay, J. E., Stow, C. D. & Harris, D. 1998 High-resolution of rainfall on Norfolk Island, Part II: Interpolation of rainfall data. *J. Hydrol.* **208**, 187–193.
- Efron, B. 1979 Bootstrap methods: Another look at Jackknife. *Ann. Stat.* **7** (1), 1–26.
- Filliben, J. J. 1975 The probability plot correlation coefficient test for normality. *Technometrics* **17** (1), 111–117.
- Fisher, R. A. 1922 The mathematical foundations of theoretical statistics. *Phil. Trans. Roy. Soc. A* **222**, 309–368.

- Frei, C. & Schär, C. 1998 A precipitation climatology of the Alps from high-resolution rain-gauge observations. *Int. J. Climatol.* **18**, 873–900.
- Friedl, H. & Stampfer, E. 2002 Jackknife resampling. In: *Encyclopedia of Environmetrics*, vol. 2 (A. El-Shaarawi & W. Piegorisch, eds). Wiley, Chichester, pp. 1089–1098.
- Gichere, S. K., Olado, G., Anyona, D. N., Matano, A. S., Dida, G. O., Abuom, P. O., Amayi, A. J. & Ofulla, A. V. O. 2013 Effects of drought and floods on crop and animal losses and socio-economic status of households in the Lake Victoria basin of Kenya. *J. Emerg. Trends Econ. Manage. Sci.* **4** (1), 31–41.
- Goovaerts, P. 2000 Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *J. Hydrol.* **228**, 113–129.
- Harremoës, P. & Mikkelsen, P. S. 1995 Properties of extreme point rainfall I: Results from a rain gauge system in Denmark. *Atmos. Res.* **37**, 277–286.
- Hill, B. M. 1975 A simple and general approach to inference about the tail of a distribution. *Ann. Stat.* **3** (5), 1163–1174.
- Hinkley, D. V. 1977 Jackknifing in unbalanced situations. *Technometrics* **19** (3), 285–292.
- Hosking, J. R. M. 1990 Lmoments: Analysis and estimation of distribution using linear combinations of order statistics. *J. R. Stat. Soc. B* **51**, 105–124.
- Hosking, J. R. M. 1991 Approximations for use in constructing L-moment ratio diagrams. Research Report RC 16635, IBM Research Division, Yorktown Heights, New York, USA.
- Hosking, J. R. M. & Wallis, J. R. 1987 Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics* **29** (3), 339–349.
- Hosking, J. R. M. & Wallis, J. R. 1997 *Regional Frequency Analysis: An Approach Based on L-Moments*. Cambridge University Press, Cambridge, UK.
- Indeje, M., Semazzi, H. F. M. & Ogallo, L. J. 2000 ENSO signals in East African rainfall seasons. *Int. J. Climatol.* **20**, 19–46.
- IPCC 2001 *Climate Change: The Scientific Basis*. Cambridge University Press, Cambridge, UK and New York, NY, USA, 996 pp.
- Jaynes, E. T. 1957a Information theory and statistical mechanics, I. *Phys. Rev.* **106**, 620–630.
- Jaynes, E. T. 1957b Information and statistical mechanics, I. *Phys. Rev.* **108**, 171–190.
- Jaynes, E. T. 1982 On the rationale of maximum entropy methods. *P. IEEE* **70**, 939–952.
- Jenkinson, A. F. 1955 The frequency distribution of the annual maximum (or minimum) of meteorological elements. *Q. J. Roy. Meteor. Soc.* **81**, 158–171.
- Kizza, M., Rodhe, A., Xu, C.-Y., Ntale, H. K. & Halldin, S. 2009 Temporal rainfall variability in the Lake Victoria basin in East Africa during the twentieth century. *Theor. Appl. Climatol.* **98**, 119–135.
- Kizza, M., Rodhe, A., Xu, C.-Y. & Ntale, H. K. 2011 Modelling catchment inflows into Lake Victoria: Uncertainties in rainfall-runoff modelling for the Nzoia River. *Hydrol. Sci. J.* **56** (7), 1210–1226.
- Kizza, M., Guerrero, J.-L., Rodhe, A., Xu, C.-Y. & Ntale, H. K. 2013 Modelling catchment inflows into Lake Victoria: Regionalisation of the parameters of a conceptual water balance model. *Hydrol. Res.* **44** (5), 789–808.
- Kjeldsen, T. R., Lundorf, A. & Rosbjerg, D. 2000 Use of a two-component exponential distribution in partial duration modeling of hydrological droughts in Zimbabwean rivers. *Hydrol. Sci. J.* **45** (2), 285–298.
- Kolmogorov, A. N. 1933 *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Springer, Berlin.
- Kratz, M. & Resnick, S. I. 1996 The qq-estimator and heavy tails. *Commun. Stat. Stoch. Model.* **12** (4), 699–724.
- Lang, M., Ouarda, T. B. M. J. & Bobée, B. 1999 Towards operational guidelines for over-threshold modeling. *J. Hydrol.* **225**, 103–117.
- Legates, D. R. & Willmott, C. J. 1990 Mean seasonal and spatial variability in global surface air temperature. *Theor. Appl. Climatol.* **41**, 11–21.
- Li, Z., Brissette, F. & Chen, J. 2013 Finding the most appropriate precipitation probability distribution for stochastic weather generation and hydrological modeling in Nordic watersheds. *Hydrol. Process.* **27** (25), 3718–3729.
- Lloyd, C. D. 2005 Assessing the effect of integrating elevation data into the estimation of monthly precipitation in Great Britain. *J. Hydrol.* **308**, 128–150.
- Looney, S. W. & Gullidge, T. R. J. 1985 Use of the correlation coefficient with normal probability plots. *Am. Stat.* **39** (1), 75–79.
- Lu, G. Y. & Wong, D. W. 2008 An adaptive inverse-distance weighting spatial interpolation technique. *Comput. Geosci.* **34**, 1044–1055.
- Ly, S., Charles, C. & Degré, A. 2011 Geostatistical interpolation of daily rainfall at catchment scale: The use of several variogram models in the Ourthe and Ambleve catchments, Belgium. *Hydrol. Earth Syst. Sci.* **15**, 2259–2274.
- Madsen, H., Rasmussen, P. & Rosbjerg, D. 1997a Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events, 1. At-site modeling. *Water Resour. Res.* **33**, 747–757.
- Madsen, H., Pearson, P. & Rosbjerg, D. 1997b Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events, 2. Regional modeling. *Water Resour. Res.* **33**, 759–769.
- Martins, E. S. & Stedinger, J. R. 2001 Generalized maximum likelihood Pareto-Poisson estimators for partial duration series. *Water Resour. Res.* **37** (10), 2551–2557.
- Mbungu, W., Ntegeka, V., Kahimba, F. C., Taye, M. & Willems, P. 2012 Temporal and spatial variations in hydro-climatic extremes in the Lake Victoria basin. *Phys. Chem. Earth* **50–52**, 24–33.
- Miller, R. G. 1974 The jackknife – a review. *Biometrika* **61** (1), 1–15.
- Mkhandi, S. H., Kachroo, R. K. & Guo, S. L. 1996 Uncertainty analysis of flood quantile estimates with reference to Tanzania. *J. Hydrol.* **185**, 317–333.
- Montanari, A. 2011 Uncertainty of hydrological predictions. In: *Treatise on Water Science* (P. Wilderer, ed.). Academic Press, Oxford, UK, pp. 459–478.
- Nicholson, S. E. 1996 A review of climate dynamics and climate variability in Eastern Africa. In: *The Limnology, Climatology*

- and *Paleoclimatology of the East African Lakes* (T.C. Johnson & E.O. Odada, eds). Gordon and Breach, Amsterdam, The Netherlands, pp. 25–56.
- Nicholson, S. E. & Entekhabi, D. 1986 The quasi-periodic behavior of rainfall variability in Africa and its relationship to Southern Oscillation. *J. Clim. Appl. Meteorol.* **26**, 561–578.
- Nicholson, S. E. & Kim, J. 1997 The relationship of the El-Niño Southern Oscillation to African rainfall. *Int. J. Climatol.* **17**, 117–135.
- Ntegeka, V. & Willems, P. 2008 Trends and multidecadal oscillations in rainfall extremes, based on a more than 100-year time series of 10 min rainfall intensities at Uccle, Belgium. *Water Resour. Res.* **44** (7), W07402.
- Nyeko-Ogiramoi, P., Willems, P., Mutua, F. & Moges, S. A. 2012 An elusive search for regional flood frequency estimates in the River Nile basin. *Hydrol. Earth Syst. Sci.* **16**, 3149–3163.
- Nyeko-Ogiramoi, P., Willems, P. & Gaddi-Ngirane, K. 2013 Trend and variability in observed hydrometeorological extremes in the Lake Victoria basin. *J. Hydrol.* **489**, 56–73.
- Ogallo, L. J. 1988 Relationships between seasonal rainfall in East Africa and the Southern Oscillation. *J. Climatol.* **8**, 31–43.
- Onyutha, C. & Willems, P. 2013 Uncertainties in flow-duration-frequency relationships of high and low flow extremes in Lake Victoria basin. *Water* **5** (4), 1561–1579.
- Onyutha, C. & Willems, P. 2014 Empirical statistical characterisation and regionalisation of amplitude-duration-frequency curves for extreme peak flows in the Lake Victoria basin. *Hydrol. Sci. J.* (in press).
- Öztekin, T. 2005 Comparison of parameter estimation methods for the three-parameter generalized Pareto distribution. *Turk. J. Agric. For.* **29**, 419–428.
- Pearson, K. 1902a On the systematic fitting of curves to the observations and measurements. *Biometrika* **1**, 265–305.
- Pearson, K. 1902b On the systematic fitting of curves to the observations and measurements, part II. *Biometrika* **2**, 1–23.
- Phillipps, J. & McIntyre, B. 2000 ENSO and interannual rainfall variability in Uganda: Implications for agricultural management. *Int. J. Climatol.* **20**, 171–182.
- Pickands, J. 1975 Statistical inference using extreme order statistics. *Ann. Stat.* **3** (1), 119–131.
- Ropelewski, C. F. & Halpert, M. S. 1987 Global and regional scale precipitation patterns associated with El Niño/Southern Oscillation. *Mon. Weather Rev.* **115**, 1606–1626.
- Rosbjerg, D. 1985 Estimation in partial duration series with independent and dependent peak values. *J. Hydrol.* **76**, 183–195.
- Rosbjerg, D. & Madsen, H. 2004 Advanced approaches in PDS/POT modeling of extreme hydrological events. In: *Hydrology: Science and Practice for the 21st Century, Volume 1, Proceedings of the British Hydrological Society International Conference* (B. Webb, N. Arnell, C. Onof, N. MacIntyre, R. Gurney & C. Kirby, eds). Imperial College, London, UK, pp. 217–220.
- Rudolf, B. & Rubel, F. 2005 Global precipitation. In: *Observed Global Climate: Numerical Data and Functional Relationships in Science and Technology – New Series, Group 5, Geophysics, 6(A)* (M. Hantel & Landolt-Börnstein, eds). Springer, Berlin, Germany, pp. 11.1–11.53.
- Schreck, C. J. & Semazzi, F. H. M. 2004 Variability of the recent climate of Eastern Africa. *Int. J. Climatol.* **24**, 681–701.
- Schultze, J. & Steinebach, J. 1996 On least squares estimates of an exponential tail coefficient. *Stat. Decis.* **14** (4), 353–372.
- Shepard, D. 1968 A two-dimensional interpolation function for irregularly spaced data. Proceedings of the 23rd National Conference. Harvard College-Cambridge, Massachusetts, USA, pp. 517–523.
- Singh, V. P. & Guo, H. 1995 Parameter estimation for 3-parameter generalized Pareto distribution by the principle of maximum entropy (POME). *Hydrol. Sci. J.* **40** (2), 165–181.
- Skaugen, T. A. 2002 Spatial disaggregation procedure for precipitation. *Hydrol. Sci. J.* **47** (6), 943–956.
- Smirnov, N. V. 1936 Sur la distribution de  $w_2$  (criterium de M. R. von Mises). *C. R. Acad. Sci.* **202**, 449–452.
- Stedinger, J. R., Vogel, R. M. & Foufoula-Georgiou, E. 1993 Frequency analysis of extreme events. In: *Handbook of Hydrology* (D. R. Maidment, ed.). McGraw-Hill, New York, USA, pp. 18.1–18.66.
- Tukey, J. 1958 Bias and confidence in not quite large samples (abstract). *Ann. Math. Stat.* **29**, 614.
- UNEP 2006 *Environment for Development: An Ecosystems Assessment of Lake Victoria Basin*, UNEP/PASS (E. O. Odada, D. O. Olago & W. Ochola, eds). UNEP/ Pan African START Secretariat (PASS), Nairobi, Kenya, 200 pp.
- UNESCO 2014 The FRIEND/Nile project. Retrieved January 02, 2014, from <http://www.unesco.org/new/en/cairo/natural-sciences/hydrology-programme/friendnile/>.
- Vogel, R. M. 1986 The probability plot correlation coefficient test for the normal, lognormal, and Gumbel distributional hypotheses. *Water Resour. Res.* **22**, 587–590.
- Willems, P. 2009 A time series tool to support the multi-criteria performance evaluation of rainfall-runoff models. *Environ. Model. Softw.* **24**, 311–321.
- Willems, P. 2013 Multidecadal oscillatory behaviour of rainfall extremes in Europe. *Clim. Chang.* **120**, 931–944.
- Willems, P., Guillou, A. & Beirlant, J. 2007 Bias correction in hydrologic GPD based extreme value analysis by means of a slowly varying function. *J. Hydrol.* **338**, 221–236.
- Wu, C. F. J. 1986 Jackknife, Bootstrap and other resampling methods in regression analysis. *Ann. Stat.* **14**, 1343–1350.
- Yang, T., Shao, Q., Hao, Z.-C., Chen, X., Zhang, Z., Xu, C.-Y. & Sun, L. 2010 Regional frequency analysis and spatio-temporal pattern characterization of rainfall extremes in the Pearl River Basin, China. *J. Hydrol.* **380**, 386–405.
- Yevjevich, V. & Taesombut, V. 1979 Information on flood peaks in daily flow series. In: *Input for Risk Analysis in Water Systems* (E. McBean, K. Hipel & T. Unny, eds). Water Resources Publications, Highland Ranch, Colorado, USA, pp. 171–192.