Journal of Hydrology 519 (2014) 512–522

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Estimation procedures for the GEV distribution for the minima

Jose A. Raynal-Villasenor *, M. Elena Raynal-Gutierrez

Department of Civil and Environmental Engineering, Universidad de las Americas Puebla, Ex-Hacienda de Santa Catarina Martir S/N, 72810 Cholula, Puebla, Mexico

ARTICLE INFO

Article history:
Received 25 February 2014
Received in revised form 1 July 2014
Accepted 21 July 2014
Available online 1 August 2014
This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Attilio Castellarin, Associate Editor

Keywords:
Distribution sampling experiments
Computer analysis
Estimation
Distribution functions
Parameters
Statistics

SUMMARY

The biased and unbiased moments (MOM1 and MOM2), maximum likelihood (ML), sextiles (SEX1 and SEX2) and probability weighted moments (PWM) methods for the estimation the parameters and quantiles of the General Extreme Value (GEV) Distribution for the minima were analyzed and compared by using data generation techniques of the type of distribution sampling experiments. Considering bias, variance and mean square error criteria of estimates of parameters and quantiles, it is concluded that in general for the values of the shape parameter considered: −0.1, −0.3, and −0.5 and 0.1, 0.3 and 0.5, the sample sizes analyzed: $9 \leqslant N \leqslant 99$ and non-exceedance probabilities: $0.01 \leqslant \Pi(x) \leqslant 0.10$, the ML method performed better than the other five. However, for sample sizes bigger than 49, most of the methods, with the exception of SEX1, produced similar results. As a general conclusion of the study reported here, it can be stated that the ML method resulted to be better to the other five when estimating the parameters and quantiles of the GEV distribution for the minima, for the cases analyzed in this study.

1. Introduction

One of the major problems in hydrologic frequency analysis is the selection of an appropriate probability distribution function to describe the distribution of hydrologic events. After all, there are many non-negative functions that integrate to one, which is one of the necessary conditions of being a probability distribution function. In practice, it is often assumed that the correct distribution function is one member of a parametric family of distribution functions. For instance, a parametric family of wide application in flood and low flow frequency analyzes is the General Extreme Value (GEV) distribution. For any given parametric family, the form of the probability density function is known, except for the unknown parameters. Once a particular parametric family is assumed, the unknown parameters are then estimated from actual data. In the study reported herein, the GEV distribution for the minima (GEVM) is the particular parametric family selected for further analysis towards its application in low flow frequency analysis.

The GEVM distribution, is defined in the following section, and has not been studied to some extent in both the statistical and hydrologic literature. It is quite flexible since it has three parameters and depending of the value of the shape parameter β, it become the Extreme Value type I (EVI) distribution for the minima (EVIM) when $\beta = 0$ and type II (EVII) distribution for the minima (EVIIM) when $\beta < 0$ and type III (EVIII) distribution for the minima (EVIIIM) when $\beta > 0$. In particular, the EVI distribution for the maxima has been most widely studied so far in flood frequency analysis, and it is known in the field of hydrology as the Gumbel's distribution. The literature abounds with alternative estimation techniques with regard to the EVI distribution for the maxima, namely: moments, maximum likelihood, least squares, mode and interquartile range, probability weighted moments, and best linear combination of order statistics. In addition, variations of such methods are also identified as the method of regression and Gumbel's method (Lowery and Nash, 1970). As such, a number of studies comparing among alternative estimating techniques for the parameters and quantiles of the GEV for the maxima have been made such those by Lowery and Nash (1970), Maciunas Landwher et al. (1979), Hosking et al. (1985), Raynal and Salas (1986), Lu and Stedinger (1992a,b), Martins and Stedinger (2000), and Raynal-Villasenor (2012b). They have compared the methods of moments (MOM), maximum likelihood (ML), sextiles (SEX), and probability weighted moments (PWM).

Less plentiful is the literature describing and comparing among estimation methods for the EVIIM and EVIIIM distributions or in general for the GEVM distribution. Most studies are related to the development and application of estimation procedures.

* Corresponding author.
E-mail addresses: josea.raynal@udlap.mx (J.A. Raynal-Villasenor), malenaraynal@gmail.com (M.E. Raynal-Gutierrez).

http://dx.doi.org/10.1016/j.jhydrol.2014.07.045
0022-1694/

Raynal-Villasenor (1995, 2012a, 2013a,b) has used the GEVM distribution with much better results that those produced by the extreme value type III distribution for the minima in several Mexican rivers. Zaidman et al. (2003) used the GEV, the generalized Pareto, generalized Logistic and Pearson type III distributions to characterize low flows in British rivers and they found that the GEV distribution was the model with more applicability to different conditions. Hewa et al. (2007) explored the application of LH-moments to low flow frequency analysis using the GEV distribution. Wang et al. (2011) applied several distribution functions to study the 41 Californian streams to assess the impact of dams on the flow regime and maximum/minimum flow probability distribution, they found that the GEV distribution gave the best result in the sites where dams have a major impact, this was with regard to the seven-day low flow. Yurekli et al. (2012) analyzed 17 rainfall gauging stations in the Cekerek watershed in Turkey to perform seasonal regional drought analysis based on the standardized precipitation index (SPI) method, one of the candidates regional distributions having the minimum Z (DIST) for k-reference periods was the GEV distribution. Wang et al. (2013) developed a generalized extreme value (GEV) distribution analysis approach, namely, a GEV tree approach that allows for both stationary and non-stationary cases. Changes in 20 year return values were estimated from the most suitable GEV distribution chosen from a GEV tree. Twenty year return values of extreme low minimum temperature are found to have warmed strongly over the century in most parts of the continent. Teimouri et al. (2013) made a comprehensive comparison of the following methods for the Weibull distribution: the method of maximum likelihood estimation (MLE), the method of logarithmic moments, the percentile method, the method of moments and the method of L-moments. Rusticucci and Tencer (2008), fitted a GEV distribution to extreme temperature indexes and return values are calculated for the period 1956–2003 for Argentina. Melchers (2008) showed that the Frechet extreme value distribution is more appropriate than Gumbel to represent the maximum pit depth in steel corrosion. Jaruskova and Rencova (2008) applied a GEV distribution to test statistics for detecting a change in a location parameter of annual maxima and minima temperature series at five European cities.

The estimation techniques that will be used in this study had been obtained by Jenkinson (1955, 1969) who developed and applied the maximum likelihood (ML) and sextiles (SEX) methods for fitting the EVIII distribution to data of annual floods, while NERC (1975) reviewed and applied Jenkinson's procedures, in addition to the biased and unbiased moments (MOM1 and MOM2) method, for estimating the parameters of the GEV distribution for the maxima. Likewise, Clarke (1973) applied Jenkinson's methods for flood frequency analysis. A related study on the properties of the three types of extreme value distributions was reported by Ochoa et al. (1980). The research reported herein is an attempt to bring information by considering the method of sextiles, with the approaches devised by Jenkinson (1969), SEX1, and Clarke (1973), SEX2, and the method of probability weighted moments proposed by Raynal-Villasenor (1987), PWM.

To serve the purpose of quantify the differences among the previous methods of estimation of parameters, three different measures, e.g. variance, bias and mean squared error, will be used to assess the methods of MOM1, MOM2, ML, SEX1, SEX2, and PWM to be used and statistically compared towards its application in low flow frequency analysis. Then following this introduction there is a section where the GEVM distribution is defined and the estimation procedures are described in detail in Appendix A. Then follows a section describing the experimental study that has been carried out and the results are presented and discussed. The next section is that of application of the methods of estimation of parameters for the GEVM to actual low flow data. The paper ends with a section that briefly summarizes and describes the conclusions reached in the study.

2. Methods

The GEVM distribution, exceedance probability, $Pr(X > x)$, is (Raynal-Villasenor, 2012a):

$$\Pi(x) = \exp\left\{-[1-\beta(\omega-x)/\alpha]^{\frac{1}{\beta}}\right\} \quad (1)$$

where ω, α and β are the location, scale and shape parameters, respectively. $\Pi(x)$ is the probability distribution function of the random variable x and for the case of low flow frequency analysis is equal to the exceedance probability, $Pr(X > x)$. The scale parameter must meet the condition that $\alpha > 0$. The GEVM distribution is a family of distributions, each member is defined by the value of the shape parameter, β, as it can be seen in Fig. 1.

The domain of variable x in the GEVM distribution is as follows:

(1) For $\beta < 0$:

$$-\infty < x \leqslant \omega - \alpha/\beta \quad (2)$$

It can be clear from Eq. (2), that the fact of having negative values on the shape parameter, an upper bound is set to the GEVM distribution denoted by $(\omega - \alpha/\beta)$ and by observing Fig. 1 it can be seen that most of the values of this distributions reach a zero value for Gumbel's reduced variates bigger than 2, which value correspond to a small return period. This is also true for the value zero of such parameter, the so-called Gumbel's distribution for the minima.

(2) For $\beta > 0$:

$$\omega - \alpha/\beta \leqslant x < \infty \quad (3)$$

The probability density function for the GEVM distribution is, Raynal-Villasenor (2012a):

$$\pi(x) = \frac{1}{\alpha}\exp\left\{-[1-\beta(\omega-x)/\alpha]^{1/\beta}\right\}[1-\beta(\omega-x)/\alpha]^{1/\beta-1} \quad (4)$$

where $\pi(x)$ is the probability density distribution of random variable x.

The detailed procedures for each of the estimation techniques for the parameters of the GEVM distribution, namely biased and unbiased moments (MOM1 and MOM2), maximum likelihood

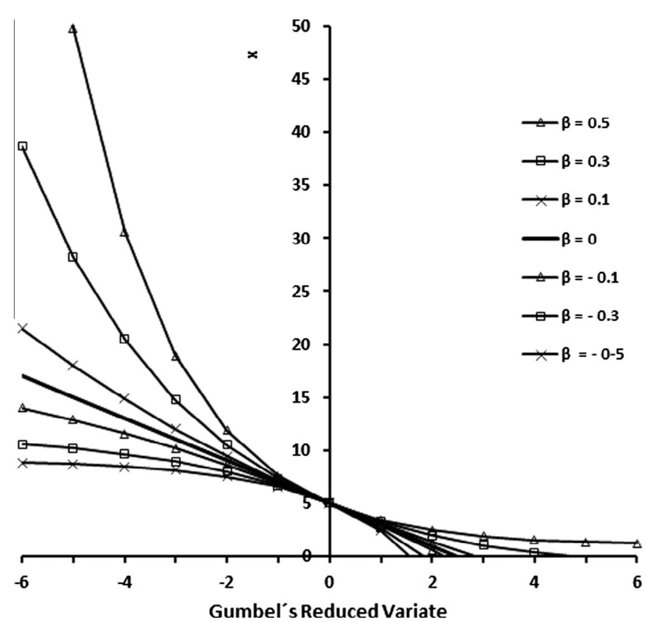

Fig. 1. The general extreme value family of distributions for the minima.